

**Report Title**

**Updated Bioinformatics Evaluation of the Transfer DNA in MON 87705 Utilizing the  
AD\_2010, TOX\_2010, and PRT\_2010 Databases**

**Authors**

**Haidi Tu  
Andre Silvanovich, Ph.D.**

**Report Completed On**

**March 15, 2010**

**Sponsor and Performing Laboratory**

**Monsanto Company  
Regulatory Product Characterization Center  
800 North Lindbergh Blvd.  
St. Louis, MO 63167**

**Laboratory Project ID**

**Study Number: RAR-10-070**

## Table of Contents

	<b>Page</b>
Report Title .....	1
Table of Contents .....	2
1.0 Summary .....	4
2.0 Sequence Database Preparation .....	5
3.0 Sequence Database Searches .....	5
4.0 Significance of the Alignments.....	6
5.0 Results and Conclusions .....	6
6.0 References.....	8

## Figures

Figure 1. Frame 1 of the MON 87705 T-DNA Sequence Translation .....	9
Figure 2. Frame 2 of the MON 87705 T-DNA Sequence Translation .....	10
Figure 3. Frame 3 of the MON 87705 T-DNA Sequence Translation .....	11
Figure 4. Frame 4 of the MON 87705 T-DNA Sequence Translation .....	12
Figure 5. Frame 5 of the MON 87705 T-DNA Sequence Translation .....	13
Figure 6. Frame 6 of the MON 87705 T-DNA Sequence Translation .....	14

## Tables

Table 1. Summary of the top alignments for the FASTA search of the allergen sequence database (AD_2010) using sequences derived from the inserted DNA in MON 87705.....	15
Table 2. Summary of the top alignments for the FASTA search of the toxin sequence database (TOX_2010) using sequences derived from the inserted DNA in MON 87705. ....	15
Table 3. Summary of the top alignments for the FASTA search of the toxin sequence database (PRT_2010) using sequences derived from the inserted DNA in MON 87705. ....	16
Appendix 1. Bioinformatic analysis of polypeptide frame1 .....	17
Appendix 2. Bioinformatic analysis of polypeptide frame2 .....	30
Appendix 3. Bioinformatic analysis of polypeptide frame3 .....	48
Appendix 4. Bioinformatic analysis of polypeptide frame4 .....	123

Appendix 5. Bioinformatic analysis of polypeptide frame5 .....	150
Appendix 6. Bioinformatic analysis of polypeptide frame6 .....	163

## 1.0 Summary

Monsanto Company has developed biotechnology-derived soybean MON 87705 with an improved fatty acid profile that results in enhanced nutritional characteristics. MON 87705 was developed to selectively suppress two key enzymes, FATB and FAD2, involved in the soybean seed fatty acid biosynthetic pathway. As a result, MON 87705 soybean oil contains lower levels of saturated fatty acids (16:0 palmitic acid and 18:0 stearic acid) and higher levels of monounsaturated 18:1 oleic acid, with an associated decrease in polyunsaturated 18:2 linoleic acid levels relative to commodity soybean oil.

MON 87705 also contains the 5-enolpyruvylshikimate-3-phosphate synthase gene derived from *Agrobacterium* sp. strain CP4 (*cp4 epsps*). Expression of the gene product (CP4 EPSPS) renders the plant tolerant to glyphosate, which is the active ingredient in the Roundup<sup>®</sup> family of agricultural herbicides. Glyphosate binds to the endogenous plant EPSPS enzyme and blocks the biosynthesis of shikimate-3-phosphate, thereby depriving plants of aromatic amino acids (Haslam, 1993; Steinrücken et al., 1984). The CP4 EPSPS protein is structurally similar and functionally identical to endogenous plant EPSPS enzymes but has a much reduced affinity for glyphosate relative to endogenous plant EPSPS (Padgett et al., 1996). Introduction of the *cp4 epsps* gene into MON 87705 allows for the production of aromatic amino acids and other metabolites even in the presence of glyphosate (Padgett et al., 1996).

Previous bioinformatic evaluation of the transfer DNA in MON 87705 concluded that the putative peptides encoded by translation of reading frames 1 through 6 of the T-DNA in MON 87705 were not similar to known allergens, toxins, or other proteins that may adversely affect human or animal health (Tu and Silvanovich, 2009). Periodically, the databases used to evaluate proteins are updated. Since the most recent report was completed, the allergen (AD\_2010), toxin (TOX\_2010) and protein (PRT\_2010) databases have been revised and updated (Tu and Silvanovich, 2010). In order to determine if potential opening reading frames (ORFs) from the T-DNA share significant sequence similarity to new sequences contained in the updated allergen, toxin, and protein databases, a bioinformatic search was made between the six query sequences and sequences contained in AD\_2010, TOX\_2010, and PRT\_2010 databases, respectively. Bioinformatic comparisons were performed using both a FASTA software program and an eight amino acid sliding window search program against the allergen database, and FASTA searches against the toxin and protein databases.

This report describes the bioinformatic evaluation of the inserted sequence contained in MON 87705. No evidence exists to indicate that the insert in MON 87705 results in the production of any protein that shares relevant similarities between known allergens, toxins or other biologically active proteins. As a result, in the event that an unexpected translation product was derived from reading frames 1-6, these putative polypeptides are not expected to possess functional cross-reactivity with known allergenic proteins or be toxic or display adverse biological activity.

---

<sup>®</sup> Roundup is a registered trademark of Monsanto Technology, LLC.

## 2.0 Sequence Database Preparation

The allergen, gliadin, and glutenin sequence database (AD\_2010) was obtained from FARRP (2010)<sup>1</sup> and was used as provided. The AD\_2010 database contains 1,471 sequences. A complete description of the AD\_2010 database can be found in Tu and Silvanovich (2010).

GenBank protein database, release 175.0 (December 15, 2009), was downloaded from NCBI and formatted for use in these bioinformatic analyses. It is referred to herein as the PRT\_2010 database and contains 17,815,538 sequences. A complete description of the PRT\_2010 database can be found in Tu and Silvanovich (2010).

The toxin database is a subset of sequences derived from the PRT\_2010 database that was selected using a keyword search and filtered to remove likely non-toxin proteins. It is referred to herein as the TOX\_2010 database and contains 8,448 sequences. A complete description of the TOX\_2010 database can be found in Tu and Silvanovich (2010).

## 3.0 Sequence Database Searches

Sequences corresponding to frames 1 to 6 (Figures 1-6) of the complete T-DNA in MON 87705 used in this analysis were the same as those used in the previously conducted bioinformatics analysis described in Tu and Silvanovich (2009).

FASTA analyses using the AD\_2010, TOX\_2010, and PRT\_2010 databases were performed on a virtual machine loaded with a SUSE LINUX version 10 operating system and FASTA version 3.4t 26 (July 7, 2006).

The structural similarity of translated T-DNA sequence in MON 87705 to sequences in the AD\_2010, TOX\_2010, and PRT\_2010 databases were assessed using the FASTA algorithm (Lipman and Pearson, 1985; Pearson and Lipman, 1988). FASTA comparisons are initiated by aligning the first match of a specific wordsize. The alignment is then extended based on the chosen scoring matrix. Specific FASTA comparison parameters used in this study included a wordsize (*k-tuple*) of two, a gap creation penalty of 10, a gap extension penalty of two, and an expectation threshold (*E-score*) of one. FASTA comparisons were performed using the BLOSUM50 scoring matrix (Henikoff and Henikoff, 1992). The BLOSUM matrix series (Henikoff and Henikoff, 1992) was derived from a set of aligned, ungapped regions from protein families, called the BLOCKS database. Sequences from each block were clustered based on the percent of identical residues in the alignments (Henikoff and Henikoff, 1996). The BLOSUM50 matrix will identify blocks of conserved residues that are at least 50% identical. BLOSUM50 works well for identifying sequence similarities that include gaps, and thus recognizes distant evolutionary relationships (Pearson, 2000).

In addition to the FASTA comparisons of each query sequence to allergens (to assess overall structural similarity), an eight amino acid sliding window search was performed. An algorithm was developed to identify whether or not a linearly contiguous match of eight amino acids existed between the query sequence and sequences within the allergen database (AD\_2010).

---

<sup>1</sup> located at <http://www.allergenonline.com>

This program compares the query sequence to each protein sequence in the allergen database using a sliding-window of eight amino acids; that is, with a seven amino acid overlap relative to the preceding window.

#### 4.0 Significance of the Alignments

An *E*-score of  $1e-5$  ( $1 \times 10^{-5}$ ) was set as an initial high cut-off value for alignment significance for identifying potential allergens and toxins. Although all alignments were inspected visually, any aligned sequence that yielded an *E*-score less than or equal to  $1e-5$  was analyzed further to determine if such an alignment represented relevant sequence homology.

#### 5.0 Results and Conclusions

No alignment met or exceeded the Codex (2003) FASTA alignment threshold for potential allergenicity of 35% identity over 80 amino acids when using the AD\_2010 database to run a FASTA search. Furthermore, no FASTA alignment displayed an *E*-score of less than or equal to  $1e-5$  and no eight amino acid matches were identified (Table 1, Appendices 1-6).

The results of the search comparisons showed that no relevant alignments were observed against proteins in the TOX\_2010 database using any of the six query sequences. No FASTA alignment displayed an *E*-score of less than or equal to  $1e-5$  (Table 2, Appendices 1-6).

When used as queries for a FASTA search the PRT\_2010 database, frames 1 to 6 yielded alignments that displayed *E*-scores less than or equal to  $1e-5$ . Two classes of alignments with frame 1 yielding an *E*-score less than  $1e-5$  were observed. One class, that yielded the lowest numerical *E*-score, was with an ORF derived from the Figwort Mosaic Virus that aligned with sequence translated from the promoter for CP4 EPSPS. The second class of alignments was with soy acyl-ACP thioesterase. Neither of these classes of alignment is unexpected because the promoter for CP4 EPSPS is partially derived from Figwort Mosaic Virus and soy acyl-ACP thioesterase is the product of the FATB gene and is an intended suppression target. While these alignments reflect conserved structure, there is no indication that they reflect the potential for adverse biological activity. Inspection of the frame 2, 4, and 6 alignments revealed that they were punctuated with numerous stop codons in the query sequence and required numerous gaps to optimize the alignment. As a result, it is unlikely these alignments reflect conserved structure. Numerous alignments with frame 3 yielded an *E*-score less than or equal to  $1e-5$ . Of these, the top alignment was with a patent sequences described as being “5-enolpyruvylshikimate-3-phosphate synthase”, CP4 EPSPS. While the top alignment positively identifies CP4 EPSPS, there is no indication that it or any other alignment with frame 3 reflects the potential for adverse biological activity. For those alignments with frame 5 yielding an *E*-score less than  $1e-5$ , all were an identical 100 amino acid overlap displaying 82% identity with an acyl-ACP thioesterase or a presumptive acyl-ACP thioesterase. This alignment is not unexpected because the T-DNA contains segments of the FATB gene sequence on the forward and reverse complement strands. While these alignments reflect conserved structure, there is no indication that they reflect the potential for adverse biological activity.

Together, these data demonstrate that the putative peptides encoded by translation of reading frames 1 through 6 of the T-DNA in MON 87705 do not have the potential to share structurally or immunologically relevant similarities with known allergens, toxins or biologically active proteins associated with adverse effects for human or animal health. These results further support the findings of Tu and Silvanovich (2009) which concludes that the T-DNA in MON 87705 soybean does not contain protein sequences that are likely to adversely affect human or animal health.

## 6.0 References

- Codex Alimentarius 2003. Guideline for the conduct of food safety assessment of foods derived from recombinant-DNA plants. CAC/GL 45-2003.
- FARRP (Food Allergy Research and Resource Program Database) 2010.  
www.allergenonline.com. University of Nebraska.
- Haslam, E. 1993. Shikimic Acid: Metabolism and Metabolites. John Wiley and Sons, Chichester, England.
- Henikoff, S., and Henikoff, J. G. 1992. Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci USA 89:10915-10919.
- Henikoff, J.G. and Henikoff, S. 1996. Blocks database and its applications. Methods Enzymol 266:88-105.
- Lipman D.J., and Pearson W.R. 1985. Rapid and sensitive protein similarity searches. Science Mar 227:1435-1441.
- Padgett, S. R., Re, D. B., Barry, G. F., Eichholtz, D. E., Delannay, X., Fuchs, R. L., Kishore, G. M., and Fraley, R. T. 1996. New Weed Control Opportunities: Development of Soybeans with a Roundup Ready Gene. CRC Handbook. 4:53-84.
- Pearson, W. R., and Lipman, D. J. 1988. Improved tools for biological sequence comparison. Proc Natl Acad Sci USA 85:2444-2448.
- Pearson, W.R. 2000. Flexible sequence similarity searching with the FASTA3 program package. Methods Mol Biol 132:185-219.
- Steinrücken, H. and Amrhein, N. 1984. 5-Enolpyruvylshikimate-3-Phosphate Synthase of *Klebsiella Pneumoniae*. Eur.J.Biochem. 143:351-357.
- Tu, H. and Silvanovich, A. 2009. Bioinformatics Evaluation of the Transfer DNA Insert in MON 87705. Monsanto Technical Report MSL0021928, St. Louis, MO.
- Tu, H. and Silvanovich, A. 2010. The Assembly of Databases Used for FASTA, BLAST and Sliding Window Searches in 2010. Monsanto Technical Report MSL0022498, St. Louis, MO.



FMSGKSTWIS	NEYDGQYGEK	ERVITNFFSI	QKCRCPQRY	KMKVHFDKTT	50
NYDPSYL.AK	AINKLF.FGN	LYFDVSTFTS	KWGLR.ETSR	SMRPSSGS.F	100
SVQSLNKVRV	QSLQTISQKL	QEINEESSIK	VNYCSSTCIM	VSKFQKKTST	150
EDLKLVGIFE	SNLVNIEQLA	CGDQTKKEWC	RIVRRTYQKH	LCLYCKDKAD	200
SSSTSGEQNN	VEKSCPDSPL	TNAYDERSDD	HKRISLSSGF	SSIPDWVQST	250
RYEPYHFIQI	GIAKTKKELP	SSKVCKEEFD	IKLDIGSFSL	EGGCSWNGTH	300
MVVIINHFC	HEILRLIYTL	LVHYFIWCLN	K.YKWLLIIC	IH.DIKYLL.	350
RRVSYRLMVH	NQ.NFWETNM	YNHLE.P.L	KGVIK.CL.H	VVLNIHKTQP	400
TQEVLSIHG.	TGA.W.FKE.	YYFMLNPNG	FGFNAINKTT	LVADSIYRSY	450
.P.PLHTFLR	YL.GKR.RTL	RSFFFFFFVC.	SLILFR.LL	DSMLFVD.SF	500
.KF.SV.V.	IL.SIVNINVSLL	LNDNLFMKLF	SHSDRSCFEI	LICSIDCWL	550
DLYTSELVDL	RI.DVCRFEL	.LGNSGVA.Q	IQCSLFLRVI	LGLFALYLLK	600
LPI.LNLARL	AQMIEHHNFC	GRNRFDSDSG	FLL.LFCVKD	DFHNGYICLL	650
FLLIQYLIVL	FFADLLTRDL	PWRKLAESAM	VCRTHLLSPI	SRNPVNANLP	700
YRFL.RRSSI	HELIRFRRRG	D.RRVG.R.L	ALSFVLLRSC	LLFPRRACFT	750
VQAAVQQLLV	SPLVFLEPSV	FQVTSLSPTG	PSCLEVSLAV	KLVSPVFWKV	800
KMLSTLVRLC	KLWVPESVRK	VILGSLMVLV	TVDSLRLRL	SISVTLQLVA	850
V.LWVLLVFT	ISIALSLVTL	LSLSVQWVVC	.THFAKWVCR	.SLKTVIVFQ	900
LPCVDQRLQR	QSPTGYLWLP	LK.SPLFCLL	VSTPQVSPLL	SSQS.LVTTL	950
KRCFKVLVLT	LPLRLMLTVC	VPSVLKVVS	SPVK.LMFQV	IHPLLLSHWL	1000
LPCLFQVPTS	PSLTF..TQP	VLVSS.LCRK	WVPTSK.STH	VLLVEKTWLT	1050
CVFVLLL.RV	LLFQKTVLLL	.STSIQFSL	QLHSLKVLPL	.TVWKNSVLR	1100
KATVFLLSQT	VSSSTVLIAM	KVRLSSCVV	VLTVRVSVTL	LEQLSLPTSI	1150
TVSL.ASSLW	VSFLKTLLLL	MMLL.SLLAS	QSSWI.WLVL	ELRSNSPTLR	1200
LLDELKNSSS	VPDPKILGSS	S.SFRSYHRF	RQRSSSMHQ	FHCAHTRILL	1250
SLSIMALGKL	FFLYHLLCL.	FTVFFIRFSL	SNCMEMDGE	ELMNDMVLLF	1300
ILKLILFVFS	LICCVLNLKL	.EICKHFVLS	KNVSNRGL..	PKLI.GVKHL	1350
.LYHYAYSLG	NKYIFRPRKA	ANVTEYKYVL	LCFRHL.TFL	YVIFQNPCQI	1400
LIIAL.L.LY	SWICS.V.KY	FLMHFMTCQL	IDNMHQSTCS	HSKRPNPGAY	1450
MARSGGRGTV	DSRGSPAKTF	NTYYLRKKYV	IIYLYFNIYS	YVFFKNLLYI	1500
DQLKYFYIYT	YFAFLSIFLR	FLAYLIMTIL	.SIIILTWT	ILLEPYEVS	1550
AFDYVDSVLI	QASICRLLIN	LVTVRTNQLL	ILPPS.LILV	VSNATTLTSL	1600
LDHKKKPRNK	RRQNTMRVSF	A.QCLSS.NS	NKNAITHSGH	HLSTS.SGSP	1650
RQEKKTGPQK	PCTTTRTHKG	VNRAAQNIHQ	LNPS.AHTFV	VSNPTSNSYS	1700
LPPPHFCLFQ	HPSNCPMPRG	QMSMHVNKTY	DYKYLQSRPR	FSSSRTGYRA	1750
RA.G.IKLCL	HLGIFHVGFI	IFVEEKKLPK	LNIAFIYPFS	FLDFLKA.V.	1800
APS..LQYQH	FSLINNWL.	CRSRGRSQHF	VLVTFCFMVM	ISRLDVLPEK	1850
RP.KFIMVAT	AATSSFFPVT	SPSPDSGGAG	SKLGGGPANL	GGLKSKSASS	1900
GGLKAKAQAP	SKINGTTVVT	SKESFKHDDD	LPSPPPRTFI	NQSCRFKLSV	1950
FENGFMMSGKS	TWISKGRSSS	CLKLSLDVTT	VVPLIFEGAC	AFAFKPPEDA	2000
DLDFSPPRFA	GPPPSLLPAP	PESGEGEVTG	KNDEVAAVAT	IMNF.GRFSG	2050
RTSSLEIMTI	KQKVTSTKCC	DRPLLRYHSQ	LFINREKC.Y	CSY.LGAYT.	2100
AFRKSRENEK	YINA.FNFGS	FFSSTNMMNP	T.NIPRCRHN	LIYP.DA.QK	2150
SASHLMQ.LC	NVSFICLFRP	CFSRILWADV	SYGTFLLES	RYFTIILYSL	2200
RHYYYFFTF.	GHSFFVFEG	YCLFCLEYVW	LDLFDSYILY	KSNYVIKKRS	2250
KNLQI.GSNP	ELHANQFTTH	TSLTLN.RDW	LVNSCFSV.F	VEFVHYKGEF	2300
CRYPHWRPR	VPYIYSDRGA	PNREVSHLSP	HLDVNVDTSK	.RFPN.NNLF	2350
IAFAYKYDGS	.FVVLKCTF	IL..RCGHLH	F.IEKKLVT	LSFSPY.PSY	2400
SLLIHVDFPD	MKPFTIE				2417

**Figure 1. Frame 1 of the MON 87705 T-DNA Sequence Translation**

SCPGNLHGSA	MSMMVNMEKK	KE.LPIFFQF	KNVDVRSVII	K.KYILIKRQ	50
ITIRRIYRRK	Q.TNYSNSEI	FISTCLHSRP	NGGLDEKLHD	RCGRRVEANS	100
QSKASTRSGY	RVSKPLAKSY	RRSMKNLQSK	.TTVPAHASW	SVSFRKRHPP	150
KT.S.WASLK	VILSTSSSWL	VGTRQKRNGA	ELLGAPTKSI	FAFIAKIKQI	200
PLVQVGNKIT	WKRAVLTAHS	LMRMTNAVTT	TKELA.AQDL	AAFQIGFNQQ	250
GTSHITLFLK	VSPKPRRNSH	PQRFVRKNSI	SSLISEVSLI	REVARGMGHI	300
WLL..TISIV	MRF.G.YILY	LFIIILFGV.I	NDINGS..SA	FIEISNIYSR	350
EECHID.WST	INEIFGRRTC	ITICLNNLN.	KV.LNDVCNM	.Y.TFIKHNQ	400
PKRY.VFTAK	QGHNGNLKND	IILC.TLTLV	SDSTL.IKPL	SLIPFIVLI	450
DPSRYTLFCD	ISEVSVNVPL	DRSFSFSSAD	RCSYFDDCW	IRCSLLIDRS	500
ENSDLLFRFY	RLLISTFHCF	.TIIYS.NYF	PILIDLVLRF	.FVRLIVGWW	550
IYIRVNLLIC	VFKMYVDLNC	DWVILE.HNK	SSVPFF.G.F	SDCLLYIS.N	600
CRFD.I.LA.	LR..STTIFV	VEIGLTPIAA	FYYDCFVLKM	IFIMVIYVYC	650
FY.FNI.LFF	FLQIC.PEIIY	HGAS.QNLQW	CAEPISYLQS	LEIQSTQISL	700
IGFSEDAAAS	TSLSDFVVVG	IEEEWDDVNW	L.ASSS.GHV	FCFHGVHASR	750
CKQPSSNCS.	VLWSFWNRPY	SR.QVYLPQV	LHVWRSR.R.	NSYHRSFGR.	800
RCYQHW.GYA	SYGCQNP.GR	.YLDH.WCW.	RWTPCS.GSS	RFR.RCNWLP	850
FDYGSCWCLR	FR.HFWH.RF	SH.ASNGSCV	EPTSRNGCAG	EV.RR.SSSS	900
YLAWTKDSNA	NHLQGTYGFR	SSEVRCSACW	SQHPRYHHCY	RANHDS.PH.	950
KDASRFWC.P	YR.D.C.RCA	YHPS.RSW.A	HRSSD.CSR.	SILYCFPIGC	1000
CLACSRFRRH	HP.RFDEPNP	YWSHLDSAGN	GCRHRSQDPT	SCWRRRG.L	1050
ACSFFYFEGC	YCSRRPCSFY	DRRVNSRCC	SCIR.RCYRY	ERFGRTPC.G	1100
KRPSFCCRKR	SQAQRC.LR.	R.DFSRRAWS	S.R.GSR.RF	WSSCRYPPRS	1150
PYRYELPRYG	SRF.KPCYC.	.CYYDRY.LP	RVHGF DGWSW	S.DRTL RH.G	1200
CLMSSRIRAR	YRILRS.DPL	ARAFVRIIGF	DNVRQVQCIS	FIAHTPESY.	1250
V.VLWHWENC	FSCTICCACN	LLCFLFGFRY	RTVKWKWMEK	S..MIWSFCS	1300
FSN.YYLFFL	LFVVC.I.NY	KRYANILF.V	KMCQIVASND	RS.YEE.NTC	1350
SCTIMLIH.A	TNIFSDLEKL	QMLLNTSMSS	CVLDIYELSF	M.FSRILVRF	1400
.SLLYNYSY	HGFVVEYENI	F.CIL.LAN.	LTTICINRPAA	TRSGPTRGPI	1450
WGPAAAVRS	TLEDPRQKHL	IRII.EKNM.	.YIYILISIL	MYFLKIYYIL	1500
IN.NIFISTL	ILHFYQFSCV	FWHI...LFF	NNQSLFLHGT	YCWNHMKCPL	1550
HLTMWIVF.S	RPPFAAY.LI	W.QSVLISYL	SFLHHN.SW.	SRMPQH.LVS	1600
WIIRKSQGTK	EDKTQ.EYPL	HSNV.VHKIQ	TKTQSHTVDI	TYPLADQDRR	1650
VKKKKLDPKS	HAQQHVLTKV	SIEQPKTFTN	STHHEPHTLL	FLTQPQTRIL	1700
FRHLIFVYFN	TRQTACHPVA	KCPCMLTRPM	TINICNLGPG	FHHQEPGTEL	1750
EPRGKLNCAC	ISGYFMWGSS	YLLRKRNSRN	.IMHLYILFH	F.IS.RLRCR	1800
HLASSYNIST	SLY..TIGCN	AAVEDDHNIS	CWLLFVLWS.	FQD.TFYRRS	1850
DLRNSLWWQQ	LLLHHFSLLL	HPRTLVEQA	ANLVVGLQTL	ED.NPNLRL	1900
VA.RQRRKPL	RKLMEPQLLH	LKKASSMMMI	YLRLPELLS	TSPAGLNYQC	1950
LKMASCPCNL	HGSAKVDHHH	A.SFL.M.QL	WFH.FSKGLA	PLPSSHQKTQ	2000
IWILVLQGLQ	AHQVCCLLH	QSPARVK.QG	KMMK.QLLPP	..ISKVASPV	2050
ERLVLKS.P.	NKK.PARNVV	IVLYCGITAN	CLSIERSADI	VATS.VPTPK	2100
PSGNLEMKKD	I.MHNSISGV	SFPQOI..TP	HEISR DAGTI	.FTPRTNRKR	2150
VPHI.CNSSV	MFHSFAYFGL	VFLVFYGLMS	HMGLFY.RAY	VTLPLYCIL.	2200
DIIIIFLPFE	DTLFLYLKEF	IVYFVWNMFG	WIYSIHIYYI	KVIMLLRNVV	2250
RTYKYKDRI	NFMQINLQPT	QV.H.INVIG	..IHVSLFNL	LNLYIIRANS	2300
ADIIHTGGRG	SHIYIAIAAR	QIVKFLI.AP	IWT.M.TRRN	KDFRIRIICL	2350
LLSPINTTDR	NLSFYQNVLS	FYNNAADIYI	FELKKNW.LL	FLFLHIDHHT	2400
HC.SM.ISRT	.SHLQL				2416

**Figure 2. Frame 2 of the MON 87705 T-DNA Sequence Translation**

HVREIYMDQQ	.V.WSIWRKR	KSNYQFFFN	KM.MSAALL	NESTF..NDK	50
LRSVVFIGES	NKQIILIRKS	LFRRVYIHVQ	MGA.MRNFTI	DAAVEWKLIL	100
SPKPQQGGT	ESPNH.PKAT	GDQ.RIFNQS	KLLFQHMHHG	Q.VSEKDIHR	150
RLKVSGLH.K	.SCQHRAAGL	WGPDKKGMVQ	NC.AHLPKAS	LPLLQR.SRF	200
L.YKMGTK.R	GKELS.QPTH	.CV.RTQ.RP	QKN.LELRI	QHSRLGSINK	250
VRAISLYSNW	YRQNQEGTPI	LKGL.GRIRY	QA.YRKFLS	GRLLVEWDTY	300
GCYNKPFFLS	.DFEVNIYFT	CSLFYLVFE	MI.MALDNLH	SLRYQIFTLE	350
KSVI.IDGPQ	SMKFLGDEHV	.PFA.ITLIK	RCD.MMFVTC	STKHS.NTTN	400
PRGIEYSRLN	RGIMVI.RMI	LFYVKP.HWF	RIQRYK.NHS	RC.FHLSFLL	450
TLAATHFSAI	SLR.ALTP	IVLFLFRLLI	VAHIISMIVG	FDALC.LIVL	500
KILICCLDFI	DC.YQRFTAS	KR.FIHETIF	PF.SILF.DF	NLFD.LLVGG	550
SIYE.TC.FA	YLRCMSI.IV	IG.FWSSITN	PVFPFSKGN	RIVCFISLEI	600
ADLIEFSSLS	SDDRAPQFLW	.KSV.LR.RL	FTMIVLC.R	FS.WLYMSTV	650
FIDSIFDCSF	FCRFVDQRST	MAQVSRICNG	VQNPSLISNL	SKSSQRKSPL	700
SVSLKTQQHP	RAYPISSSWG	LKKSGMTLIG	SELRPLKVMS	SVSTACMLHG	750
ASSRPATARK	SSGLSGTVRI	PGDKSISHRS	FMFGGLASGE	TRITGLLEGE	800
DVINTGKAMQ	AMGARIRKEG	DTWIIDGVGN	GGLLAPEAPL	DFGNAATGCR	850
LTMGLVGVYD	FDSTFIGDAS	LTKRPMGRVL	NPLREMGVQV	KSEDGDRLPV	900
TLRGPKTPTP	ITYRVPMAA	QVKSALLAG	LNTPGITTVI	EPIMTRDTE	950
KMLQGFGANL	TVETDADGVR	TIRLEGRGKL	TGQVIDVPGD	PSSTAFPLVA	1000
ALLVPGSDVT	ILNVLMNPTR	TGLILTLQEM	GADIEVINPR	LAGGEDVADL	1050
RVRSSTLKGV	TVPEDRAPSM	IDEYPILAVA	AAFAEGATVM	NGLEELRVKE	1100
SDRLSAVANG	LKLNGVDCDE	GETSLVVRGR	PDGKGLGNAS	GAAVATHLDH	1150
RIAMSFLVMG	LVSENPVTVD	DATMIATSFP	EFMDLMAGLG	AKIELSDTKA	1200
A..AQEFELG	TGS.DLRIL	LELSFVSSVS	TTFVKFNASV	SLRTHQNPT	1250
FEYYGIGKTV	FLVPFVVLVI	YCVFYSVFAI	EL.NGNGWRR	VNE.YGPFVH	1300
SQINIICFFS	YLLCVEFEII	RDMQTFCFE	KCVKSWPLMT	EVNMRSKTLV	1350
VVPLCLFTRQ	QIYFQT.KSC	KCY.IQVCPL	VF.TFMNFPL	CNFPELSDS	1400
NHCFIIVIL	MDL.LSMKIF	FNAFYDLPID	.QHASIDLQP	LEAAQPGGLY	1450
GPVRRPRYGR	L.RIPGKNI	YVLFKKKICN	NIFIF.YLFL	CIF.KSIIY	1500
STKIFLYLHL	FCIFINFLAF	FGIFNNDYSL	IINHYSYMH	IVGTI.SVHC	1550
I.LCG.CFDP	GLHLPLIN.F	GNSPY.SVTY	PSSIIINLGS	LECHNTD.SL	1600
GS.EKAKEQK	KTKHNESILC	IAMSKFIKFK	QKRNTQWTS	LIH.LIRIAA	1650
SRKKNWTPKA	MHNNTYSQRC	QSSSPKHSPT	QPIMSPHIC	F.PNLKLVFS	1700
SATSFLFIST	PVKLHATPWP	NVHAC.QDL	L.ISAISAQV	FIIKNRVPSS	1750
SLGVN.IVPA	SRDISCGVHH	IC.GKETPEI	ELCIYISFFI	SRFPEGLGVG	1800
T.LVATISAL	LSIDKQLAVM	PQ.RTITTFR	AGYFLFYGHD	FKTRRSTGEA	1850
TLEIHYGGNS	CYFIIFPCYF	TLAGLWWSRQ	QTWWWACKPW	RTKIQICVFW	1900
WLEGKGASPF	EN.WNHSCYI	.RKLQA...S	TFASPQNFYQ	PVLQV.TISV	1950
.KWLHVREIY	MDQQR.IIIM	LEAFFRCNNC	GSINFRRGLR	LCLQATRRRR	2000
FGF.SSKVCR	PTTKFAACST	RVRRG.SNRE	K...SSSCCHH	NEFLRSLLR	2050
NV.S.NHDHK	TKSNQHEML	SSSTAALQPI	VYQ.REVLIL	.LLARCLHLS	2100
LQEI.K.KRI	YKCIIQFREF	LFLNKYDEPH	MKYPEMQAQF	NLPLGRVTKE	2150
CLTFDAIAL	CFIHLLISAL	FFSYSMG.CL	IWDFSTREPT	LLYHYIVFFE	2200
TLLFFFYLLR	TLFFCI.RNL	LFILFGICLV	GFIRFIYII	K.LCY.ET..	2250
ELTNIRIESR	TSCKSIYNPH	KFNIKLT.LV	SKFMFLCLIC	.ICTL.GRIL	2300
QISITLAAAG	PIYI.RSRA	KS.SFSSKPP	FGRECRHVEI	KISELE.FVY	2350
CFRL.IRRIV	ICRFIKMYFH	FIITLRTSTF	LN.KKIGNYS	FFFSILTIIIL	2400
IADPCRFPGH	EAIYN.				2416

**Figure 3. Frame 3 of the MON 87705 T-DNA Sequence Translation**

FNCKWLHVRE	IYMDQQ.V.W	SIWRKRKSNY	QFFFN SKM.M	SAALL.NEST	50
F..NDKLR SV	VFIGESNKQI	ILIRKSLFRR	VYIHVQMG A.	MRNFTIWR AA	100
IAIYIWDPRP	PV.WISAEFA	LIMYKF NKL N	RET.IY.PIT	LI.C.TCVGC	150
KLICMKFGIR	SLYL.VLTTF	LNNIITFI.Y	I.IE.IQPNI	FQTK.TINSF	200
KYKKRVSSKG	KKIIIMSQRI	QYNGKVT.AL	..KSPI.DIS	P.NTRKTRPK	250
.ANE.NITEL	LHQM.GTLLL	RVLGVN.IVP	ASRDISCGVH	HIC.GKETPE	300
IELCIYISFF	ISRFPEGLGV	GT.LVATISA	LLSIDKQLAV	MPQ.RTITTF	350
RAGYFLFYGH	DFKTRRSTGE	ATLEIHYGGN	SCYFIIFPCY	FTLAGLWWSR	400
QQTWWWACKP	WRTKIQICVF	WWLEGKGASP	FEN.WNHSCY	I.RKLQA...	450
STFADPCRFP	GHEAIFKH..	FKPAGLV DKS	SGGRRR.III	MLEAFFRCNN	500
CGSINFRRGL	RLCLQATRRR	RFGF.SSKVC	RPTTKFAACS	TRVRRG.SNR	550
EK..SSSCCH	HNEFLRSLLR	.NV.S.NHDH	KTKSNQHEML	.SSSTAALQP	600
IVYQ.REVLI	L.LLARCLHL	SLQEI.K.KR	IYKCIIQFRE	FLFLNKYDEP	650
HMKYPEMQAQ	FNLPLGSSSV	PGS...KPGP	RLQIFIVIGL	VNMHGHLATG	700
WHAV.RVLK.	TKMRWRKRIR	V.GWVRNNKC	VGS.WVELVN	VLGCSIDTFV	750
STCCCAWLLG	SSFFFLTRRS	.SASG.VMST	VCDCVFV.IL	.T.TLLCKGY	800
SHCVLSSFPV	WLFLMIQETS	QCCGIRDYQD	.L.WRKDK.L	ISTDCYQIN.	850
.AANGGLDQN	TIHIVKCNGH	FIWFQQYVPC	KNND.LLKNS	HY.ICQKTQE	900
N..KCKISVD	IKIF.LINI.	.IFKKYIRID	IKI.IYYYIF	FS.IIRIKCF	950
CRGSSRDVRT	AAAGPGHIGP	RVGPLRVAAG	RLMHVVNQLA	SHKMH.KIFS	1000
YSTTNP.V.L	.L.SND.NLT	RILENYIKES	S.MSKTQEDI	LVFSNIC SFS	1050
RSENI FVA..	ISIMVQLQVF	YSSY.LRSLE	ATI.HIFTQN	KMFAYLL.FQ	1100
IQHTTNKRKN	K.Y.FENEQK	DHIIH.LFSI	HFHFTVR.RK	PNKKH SKLQA	1150
QQMVQEQKFS	QCHNTQTQ.D	SGVCAMKLMH	.T.RTL SKPM	IRTKALARGS	1200
.DLRIRYRAR	ILELIKQP.C	RRVRS.LQDQ	PSNP.TLGS.	.RS..HHQQ.	1250
QGFQKRDP.R	GSS.RYGDRG	G.RQLLQKRY	RD PYRQDDHA	RREKSHLHRN	1300
QHR.A.DRLR	QQKDGRFP.H	GVLPNRS.R.	HLQRMQLQQR	ELDTRRS.KE	1350
HGLLEQ.HPS	K.KNEHASQP	RLLHQQDVG.	SLRCRHPFPA	ESR.DQYGLG	1400
SSKR.GW.RR	NLEQARQQPM	GKQ.RMDHLE	HQSLDR.AYH	DLQDGWY AHR	1450
QHQSQR.G.H	QNLEASFQCG	HES.LAR.QW	.YLG C.DQQA	EQRTSLERKP	1500
.VPCR.LALE	SLVHAR.LED	DHRLQTS PAH	PFREVGSTHD	PLDA..EK RH	1550
Q.KCYRNRKH	QQDP.SNGNQ	LQRYRNREEP	QE QGVHRYQH	HQ.SKYHLPY	1600
GFWHP.LA.P	YQC..HLHLP	KDR.YEFHR.	RDLQT.RTCG	R.TCHLEYGR	1650
FQKDQRTYEQ	LLDGCLHREA	CTPWKQKT.P	.EDEAQSQLT	SSHSSSIPTT	1700
TKSDKLVDA A	ASSEKPIREI	CVDWISRDWR	.EMGSAH HCR	FC.LAPW.IS	1750
GQQICKKKNN	QILNQ.KQ.T	YITIMKIIFN	TKQS..KAAI	GVKPISTTKI	1800
VVLYHLS.AS	.IQSNRQFQE	I.SKQSENYP	.KKGTLDLLC	YSRITQSQFK	1850
STYILNTQIN	KFTRI.IHQP	TINRTN.NLK	TRSIRMGK.F	HE.IIV.KQ.	1900
NVDINNR.NL	NNRSEFSERS	INKEHRIQQS	SK.YEQRSAD	EKEKERSKGT	1950
LTLTSEISQK	SV.RLGSIRT	INGISNESGF	IYSVESETNV	RV.HKIIISFF	2000
KLPLCPCLAV	NTQYLLGWLC	FMNV.YYMLQ	TSFNHTF.LR	LFKQMVIVHR	2050
LPKISLIVDH	QSI.HSSLE.	IFDISMNADY	QEPFISFIQT	PNKIMNK.SI	2100
Y.PQNLMTME	MVYYNNH MCP	IPRATSLKRE	TS DIKLDIEF	FLTNL.GWEF	2150
LLGFGDTNLN	KVIWL VPC.L	NPIWNAAKS.	AQANSFVVVT	AFVIRISEWA	2200
VRTALFHVIL	FPTCTRGICF	IFAIKAKMLL	VGAPNNSAPF	LFCLVPTS QL	2250
LDVDKITFKD	AH.L.VFGGC	LFLKLT DHDA	CAGTVVYFD.	RFFIDLL.LL	2300
ANGLETLYPD	LVEALD.ELA	STRRPHRS.S	FSSKPPFGRE	CRHVEIKISE	2350
LE.FVYCFRL	.IRRIVICRF	IKMYFHFIIT	LRTSTFLN.K	KIGNYSFFF S	2400
ILTIILIADP	CRFP GHE				2417

**Figure 4. Frame 4 of the MON 87705 T-DNA Sequence Translation**

SIVNGFMMSGK	STWISNEYDG	QYGEKERVIT	NFFSIQKCRC	PQRYYKMKVH	50
FDKTTNYDPS	YL.AKAINKL	F.FGNLYFDV	STFTSKWGLR	.ETSRFGAPR	100
SLYIYGTRGR	QCDGYLQNSP	L.CTNSTN.T	EKHEFTNQSR	.FNVKLVWVW	150
N.FA.SSGFD	PYICKFLLRF	LIT.LLLYNI	YESNKSNTY	SKQNKQ.IPS	200
NTKKECPQKV	KK...CLKEY	NIMVK.RRLS	SRKVPYETSA	HRIREKQGRN	250
KQMNETLQSY	CIKCEALFCY	AS.G.IKLCL	HLGIFHVGF	IFVEEKKLPK	300
LNAYFIYPFS	FLDFLKA.V.	APS..LQYQH	FSLLINNWL.	CRSRGRSQHF	350
VLVTFCFMVM	ISRLDVLPEK	RP.KFIMVAT	AATSSFFPVT	SPSPDSGGAG	400
SKLGGGPANL	GGLKSKSASS	GGLKAKAQAP	SKINGTTVVT	SKESFKHDDD	450
LPLLIHVDFP	DMKPFSTNDS	LNLDWLIVK	LGGGEGRSSS	CLKLSLDVTT	500
VVPLIFEGAC	AFAFKPPEDA	DLDFSPRFA	GPPPSLLPAP	PESGEGETVG	550
KNDEVAAVAT	IMNF.GRFSG	RTSSLEIMTI	KQKVTSTKCC	DRPLLRHYSQ	600
LFINREKC.Y	CSY.LGAYT.	AFRKSNEKG	YINA.FNFGS	FFSSTNMMNP	650
T.NIPRCRHN	LIYP.ARARY	PVLDDENLGR	DCRYL.S.VL	LTCMDIWPRG	700
GMQFDGC.NK	QK.GGGREYE	FEVGLETTNV	WAHDGLSW.M	FWAARLTPL.	750
VRVVVHGFVG	PVFFS.RGDP	DQLVDK.CPL	CVIAFLFEFY	ELRHICYAKDT	800
LIVFCLLLFL	GFFL.SKRLV	SVVAFETTKI	NYDGGRISN.	LVRTVTKLIN	850
KRQMEAWIKT	LST.SNAMDT	SYGSNNMYHV	RIMIDY.RIV	IIKYAKKRKK	900
IDKNAK.V.I	.KYFS.SIYN	RFLKNT.E.I	LKYKIITYF	FLK.YVLNVF	950
AGDPLESTVP	RPPDRAI.AP	GLGRFEWLQV	D.CMLSINWQ	VIKCIKKYFH	1000
TQLQIHEYNY	NYKAMIRI.Q	GFWKIT.RKV	HKCLKHKRTY	LYSVTFAAFL	1050
GLKIYLLPSE	.A.WYNYKCF	TPHINFGH.R	PRFDTFLLKT	KCLHISYNFK	1100
FNTQQIREKT	NNINLRMNKR	TISFINSSPS	ISISQFDSN	RIKNTVNYKH	1150
NKWKYKNSFP	NAIILKLSRI	LVCAQ.N.CI	ELDERCRNR.	YERKL.LEDP	1200
KILGSGTELE	FLSSSSSLSV	GEFDLSSKTS	HQIHELWEAS	SDHSSIINSN	1250
RVFRNETHNE	EAHSDTVIEV	GSDSCRSRVT	ETLTVRTTTH	DERSLTFIAI	1300
NTVELETVCD	SRKTVAFLNT	EFFQTVHNGS	TFSECSCNSE	NWILVDHRRS	1350
TVFWNSNTLQ	SRRTNTQVSH	VFSTSKTWVD	HFDVGTFLQ	SQDETSTGWV	1400
HQNVKDGVDG	TWNKQGSNW	ESSRGWITWN	INHLTGELTT	TFKTDGHTV	1450
SISLNGKVS	KTLKHLFSVV	TSHDWLDNSG	DTWGVETSKQ	NSGLHLSGSH	1500
RYPVGDWRS	LWSTQGNWKT	ITVFRHLHT	HFAKWVQHTT	HWTLSERSVT	1550
NESAIEIVNT	NKTHSQTATS	CSVTEIERSL	RSKESTVTNT	INDPSITFLT	1600
DSGTHSLHSL	TSVDNIFTFQ	KTGDTSTFTAS	ETSKHEGPVG	DRLVTWNTDG	1650
SRKTRGLTSS	CWTAACVTKH	ARRGNRRHDL	KRTKLRAN.R	HPTLLQSPRR	1700
RNRISWMLL	RLQRNR.GRF	ALTGFREIGD	KRWVLHTIAD	SANLRHGRSL	1750
VNKSAKKRTI	KY.INKNSRH	I.PL.KSSLT	QNNHSHKPLS	ESNRFLPQKL	1800
WCSII.AKRA	KFNQIGNFKR	YKANNPRITL	RKREHWICYA	TPELPNHNSN	1850
RHTS.IRKST	SSLVYRSTNQ	QSIEQIKISK	QDRSEWENSF	MNKLSFRSSE	1900
TLILTIDKI.	TTDQNFQNDQ	STKSIESNNH	RNNMSNDQQT	KKKKNDLVR	1950
.RLPQRYRRK	VCSG.GQ.ER	.MESATRVVL	FIALNPKPML	GFNIK.YHSL	2000
NYHYAPV.P.	ILNTSWGVCV	L.MFSTTCYK	HHLITPFN.G	YSSKWLYMFV	2050
SQKFH.LWTI	NLYDTLL.SK	YLISQ.MQII	KSHLYHLFKH	QIK..TSKVY	2100
INLKIS.QWK	WFIITTICVP	FHEQPPSREK	LPISSLISNS	SLQTFEDGSS	2150
FLVLAIP.I	K.YGSYLV.D	TQSGMLLNPE	LKLILLWSSL	RSSYALVSGL	2200
SGQLFSTLFC	SPLVLEESAL	SLQ.RQRCFW	.VRLTILHHS	FFVWSPQASC	2250
SMLTRLLSKM	PTNFKSSVDV	FF.NLLTMMH	VLEQ.FTLIE	DSSLISCSFW	2300
LMVWRLCTLT	LLRLWTEN.L	PLDGRIDREV	SHLSPHLDVN	VDTSK.RFPN	2350
.NNLFIAFAY	KYDGS.FVVL	SKCTFIL..R	CGHLHF.IEK	KLVITLSFSP	2400
Y.PSYSLLIH	VDFPDM				2416

**Figure 5. Frame 5 of the MON 87705 T-DNA Sequence Translation**

QL.MASCPGN	LHGSAMSMV	NMEKKKE.LP	IFFQFKNVDV	RSVVIK.KYI	50
LIKROITIRR	IYRRKQ.TNY	SNSEIFISTC	LHSRPNGLD	EKLHDLARRD	100
RYIYMGPAAA	SVMDICRIRP	YNVQIQQIKQ	RNMNLLTNHV	NLMLNLCGL.	150
IDLHEVRDSI	LIFVSSYYVS	..HNYFYIIY	MNRINPTKHI	PNKINNKFLO	200
IQKKSVLKR.	KNNNNVSKNT	I.W.SNVGSL	VEKSHMRHQP	IEYEKNKAEI	250
SK.MKHYRAI	ASNVRSFVT	RPRGKLNAC	ISGYFMWGSS	YLLRKRNSRN	300
.IMHLYILFH	F.IS.RLRCR	HLASSYNIST	SLY..TIGCN	AAVEDDHNIS	350
CWLLFVLWS.	FQD.TFYRRS	DLRNSLWWQQ	LLLHHFSLLL	HPRRTLVEQA	400
ANLVVGLQTL	ED.NPNLRL	VA.RQRRKPL	RKLMEPQLLH	LKKASSMMMI	450
YLC.SM.ISR	T.SHFQTLIV	.TCRTG..KF	WGEAKVDHHH	A.SFL.M.QL	500
WFH.FSKGLA	PLPSSHQKTQ	IWILVLQGLQ	AHHQVCCLLH	QSPARVK.QG	550
KMMK.QLLPP	..ISKVASPV	ERLVLS.P.	NKK.PARNVV	IVLYCGITAN	600
CLSIERSADI	VATS.VTPPK	PSGNLEMKKD	I.MHNSISGV	SFPQQI..TP	650
HEISRDAGTI	.FTPRLELGT	RFLMMKTWAE	IADIYSHRSC	.HAWTFGHGV	700
ACSLTGVEIN	KNEVAEENTS	LRLG.KQQMC	GLMMG.VGEC	FGLLD.HLCE	750
YVLLCMAFGV	QFFFLDAAIL	IS.WISDVHC	V.LRFCLNFM	NLDIAMQRIL	800
SLCFVFFCSL	AFSYDPRD.S	VLWHSRLPRL	IMMEEG.VTD	.YGLLPN.LI	850
SGKWRPGSKH	YPHSQMQWTL	HMVPTICTM.	E..LIIKE.S	LLNMPKNARK	900
LIKMQNKCRY	KNILVDQYII	DF.KIHKNRY	.NINILLHIF	FLNNTY.MFL	950
PGIL.SRPYR	GRRTGYPYRP	GWAASSGCRS	IDACCQSIGK	S.NALKNIFI	1000
LNYSKMSITI	IIKQ.LESDK	DSGKLHKGKF	INV.NTRGHT	CIQ.HLQLF.	1050
V.KYICCLVN	KHNGTTTSVL	LLILTSVIRG	HDLTHFYSKQ	NVCISLIISN	1100
STHNK.EKKQ	IILI.E.TKG	PYHSLTLLHP	FPFHSSIAKT	E.KTQ.ITST	1150
TNGTRKTVFP	MP.YSNSVGF	WCVRNETDAL	NLTNVVETDD	TNESSS.RIL	1200
RS.DPVPSN	S.AHQAAALS	ESSILAPRPA	IKSMNSGKLV	AIIVASSTVT	1250
GFSETRPITR	KLIAIR.SRW	VATAAPEALP	RPLPSGRPRT	TREVSPSSQS	1300
TPLSLRPFAT	AERRSLSLTR	SSSKPFITVA	PSANAAATAR	IGYSSIIEGA	1350
RSSGTVTPFK	VEERTRKSAT	SSPPARRGLI	TSMSAPISCR	VKMRPVVRVGF	1400
IKTLRMVTSE	PGTSKAATNG	KAVEDGSPGT	SIT.PVSLPR	PSRRMVRTPS	1450
ASVSTVRMLAP	KP.SIFSWS	RVMIGSITVV	IPGVLRPASR	TADFT.AEAI	1500
GTL.VIGGV	FGPRKVTGRR	SPSSDFTCTP	ISRSGFNTRP	IGRLVREASP	1550
MKVLSSK.TP	TRPIVKRQPV	AALPKSRGAS	GARSPPLPTP	SMIQVSPSLR	1600
ILAPIACIAL	PVLITSSPSK	RPVIRVSPLA	RPPNMKDLWE	IDLSPGIRTV	1650
PERPEDLRAV	AGRLAP.SM	HAVETEDMTL	RGRSSEPINV	IPLFFNPHDD	1700
EIG.ARGCCC	VFRETDKGD	R.LDFERLEI	RDGFCTPLQI	LLTCAMVDLW	1750
STNLQKKEQS	NIESIKTVDI	YNHYENHL.H	KTIIIVKSRYR	SQTDIFYHKNC	1800
GALSSELSEL	NSIKSAISR	IKQTIRELPL	EKGNTGFVML	LQNYPITIQI	1850
DIHLKYANQQ	VHSYIDPPTN	NQSNKLKSN	KIDQNGKIVS	.INYLEAVK	1900
R.Y.QSIKSK	QQIRIFRTIN	QQRASNPTII	EII.ATISRR	KRKRTI.GYV	1950
NAYLRDIAEK	CVAARVNKND	KWNQQREWFY	L.R.IRNQC.	GLT.NNIIL.	2000
ITIMPLFSRE	YSIPLGLVVF	YECLVLHVTN	II.SHLLIKV	IQANGYTCSS	2050
PKNFIDCGPS	IYMTLFSRVN	I.YLNECRLS	RAIYIIYSNT	K.NNEQVKYI	2100
LTSKSHDNGN	GLL.QPYVSH	STSNLPQERN	FRYQA.YRIL	PYKPLRMGVP	2150
SWFWRYQFE.	SDMARTLLIE	PNLECC.ILS	SS.FFCGRHC	VRHTH..VGC	2200
QDSSFPRYFV	PHLY.RNLLY	LCNKGKDAFG	RCA.QFCTIP	FLSGPHKPAA	2250
RC.QDYFQRC	PLTSLRWMS	FSETY.P.CM	CWNSSL.LK	ILH.SPVAFG	2300
.WFGDSVP.P	C.GFGLRISF	HSTAASIVKF	LI.APIWT.M	.TRRNKDFRI	2350
RIICLLLSPI	NTTDRNLSFY	QNVLSFYNNA	ADIIYIFELKK	NW.LLFLFLH	2400
IDHHTHC.SM	.ISRT.				2416

**Figure 6. Frame 6 of the MON 87705 T-DNA Sequence Translation**

**Table 1.** Summary of the top alignments for the FASTA search of the allergen sequence database (AD\_2010) using sequences derived from the inserted DNA in MON 87705.

Appendix	Polypeptide	AD_2010 Sequence Database						
		Sliding Window	FASTA search					
		Hits	# Hits	GI #	Description	% Identity	aa Overlap	E score
1	frame1	No	2	76782247	hydrophobic seed protein pr (134 aa)	32.787	61	0.067
2	frame2	No	-	-	-	-	-	-
3	frame3	No	-	-	-	-	-	-
4	frame4	No	17	73912496	omega-5 gliadin [Triticum (439 aa)	17.370	403	0.00067
5	frame5	No	1	76782247	hydrophobic seed protein pr (134 aa)	30.645	62	0.26
6	frame6	No	3	66849793	allergen Asp F7 [Aspergillu (270 aa)	21.101	218	0.42

**Table 2.** Summary of the top alignments for the FASTA search of the toxin sequence database (TOX\_2010) using sequences derived from the inserted DNA in MON 87705.

Appendix	Polypeptide	TOX_2010 Sequence Database					
		FASTA search					
		# Hits	GI #	Description	% Identity	aa Overlap	E score
1	frame1	1	158113520	Zeta toxin family protein (383 aa)	33.333	54	0.43
2	frame2	3	62693	precrotoxin B (AA -16 to 122) (138 aa)	28.829	111	0.064
3	frame3	-	-	-	-	-	-
4	frame4	1	28194389	myotoxic phospholipase A2 [ (138 aa)	24.793	121	0.91
5	frame5	1	6224908	enterotoxin [Bacillus cereus (431 aa)	26.708	161	0.59
6	frame6	1	197625399	structural toxin protein R (2003 aa)	25.989	177	0.086

**Table 3.** Summary of the top alignments for the FASTA search of the protein sequence database (PRT\_2010) using sequences derived from the inserted DNA in MON 87705.

Appendix	Polypeptide	PRT_2010 Sequence Database					
		FASTA search					
		# Hits	GI #	Description	% Identity	aa Overlap	<i>E</i> score
1	frame1	107 <sup>1</sup>	58814	unnamed protein product [Figw (512 aa)	100.000	135	2.4e-47
2	frame2	748 <sup>1</sup>	136192920	hypothetical protein GOS_8 (396 aa)	26.905	420	3.2e-10
3	frame3	6026 <sup>1</sup>	62318479	5-enol-pyruvylshikimate-3- (527 aa)	88.701	531	1.2e-184
4	frame4	83 <sup>1</sup>	193900736	GH12332 [Drosophila grimsh (1317 aa)	16.446	529	6.5e-06
5	frame5	76 <sup>1</sup>	90192131	acyl-ACP thioesterase [Glyc (416 aa)	82.000	100	7.2e-22
6	frame6	113 <sup>1</sup>	228385049	Sequence 21516 from patent (331 aa)	40.909	242	1.2e-17

---

<sup>1</sup> Only the top 50 alignments are shown in the appendices.



Appendix 1. Bioinformatic analysis of polypeptide frame1

```
>frame1
FMSGKSTWISNEYDGGYGEKERVITNFFSIQKCRCPQRYYKMKVHFDKTINYDPSYL.AKAINKLF.FGNLYFDVSTFTSKWG
LR.ETSRSMRPSGGS.FSVQSLNKVRVQSLQTIQKQLQEQINEESSIKVNYCSTCIMVSKFQKKTSTEDLKLVGIFESNLVNI
EQLACGDQTKKEWCRIVRRTYQKHLCLYCKDKADSSSTSGEQNNVEKSCPDSPLTNAYDERSDDHKRISLSSGFSIPDWQVS
TRYEPYHFTQIGIAKTKKELPSSKVCKEEFDIKLDIGSFSLEGGSWNGTHMVVIIINHFCHEILRLIYTLVLVHYFIWCLNK.
YKWLIIICIH.DIKYLL.RRVSYRLMVHNQ.NFWETNMYNHLE.P.LKGVIK.CL.HVVNLNIHKTQPTQEVLSIHG.TGA.W
.FKE.YYFMLNPNGIFGFNAINKTTLVADSIYRSY.P.PLHTFLRYL.GKR.RTLRSFFFFFVC.SLLILFR.LLDSMLFVD.
SF.KF.SVV.ILSVININVSLLLNDNLFMKLFSDRSCEILICSIDCWLVDLYTSELVDLRI.DVCRFEL.LGNSGVA.QI
QCSLFLRVILGLFALYLLKLPI.LNLARLAQMIEHHNFCGRNRFDSDSGFLL.LFCVKDDFHNGYICLLFLLIQYLIVLFFAD
LLTRDLPWRKLAESAMVCRTHLLSPISRNPVNANLPYRFL.RRSSIHLEIRFRRRGD.RRVG.R.LALSFVLLRSCLLFPRRA
CFTVQAQVQLLVSPLVFLEPSVFQVTSLSPTGSPCLEVSLAVKLVSPVFWKVKMLSTLVRLCKLWVPESVRKVILGSLMLV
TVDSLLRLLSISVTLQLVAV.LWVLVFTISIALSLVTLSSLVQWVVC.THFAKWVCR.SLKTIVVFQQLPCVDQRLQRQSP
TGYLWLPK.SPLFCLLVSTPQVSPLLSSQS.LVTTLKRCFKVLVLTPLRLMLTVCVPSVLKVVVSSPVK.LMFQVIHPLLL
SHWLLPCLEQVPTSPSLTF..TQPVLVSS.LCRKWVPTSK.STHVLLVEKTMWLTCTVFVLLL.RVLLFQKTVLLL.STSIQFSL
LQLHSLKVLPL.TVWKNSVLRKATVFLLSQTVSSSTVLIAMKVRLSSCVVLTVRVSVTLLEQLSLPTSITVSL.ASSLWVS
FKTLLLLMMLL.SLLASQSSWI.WLVLELRNSPTLRLLDELKNSSVDPDPKILGSSS.SFRSYHRFRQRSSSMHQFHCAH
TRILLSLSIMALGKLFLLYHLLCL.FTVFFIRFSLSNCEMEMDGEELMNDMVLLFILKLILFVFSLICCVLNLKL.EICKHFV
LSKNVSNRGL..PKLI.GVKHL.LYHYAYSLGNKYIFRPRKAANVTEYKYVLLCFRHL.TFLYVIFQNPQCILIIAL.L.LYS
WCSS.V.KYFLMHFMTCLQIDNMHQSTCSHSCRPNPGAYMARSGRGTVDSRGSPAKTFNTYYLRKKYVYIYLYFNIYSYVFF
KNLLYIDQLKYFYIYTYFAFLSIFLRFLLAYLIMTL..SIIILTWYILLEPEVSTIAFDYVDSVLIQASICRLLINLVTVRTN
QLLILPPS.LILVVSNAATTLTSLLDHKKKPRNKRQNTMRVSFA.QCLSS.NSNKNAITHSGHHLSTS.SGSPRQEKKTGPQK
PCTTTTRTHKGVNRAQNTIHLNPS.AHTFVVSNPTSNYSYSLPPPHFCLFQHPSCMCMPPRGQMSMHVNKTYDYKYLQSRPRFSS
SRGTYRARA.G.IKCLHLGIFHVGFIIFVEEKKLPKLNAYFIYPFSFLDFLKA.V.APS..LQYQHFSLLINNW.L.CRSRGR
SQHFVLVTFCFMVMISRDLVLEKRP.KFIMVATAATSSFFPVTSPSPDGGAGSKLGGGPANLGGKSKSASSGGLKAKAQA
PSKINGTTVVTSKESFKHDDDLSPPPRTFINQSCRFLSVFENGFMGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAF
AFKPPEDADLDFSPPRFAGPPSLLPAPPESGEVETGKNDEVAAVATIMNF.GRFSGRTSSEIMTIKQKVTSTKCCDRPLL
RHYSQLFINREKC.YCSY.LGAYT.AFRKSRNEKGYINA.FNFGSFFSSTNMMNPT.NIPCRHNLIIYP.DA.QKSASHLMQ.
LCNVSFICLFRPCFSRILWADVSYGTFLLESRLRYFTIILYSLRHYYFFTF.GHSFFVFEGIYCLFCLFYVWLDLFDYIYK
SNYVIKKRSKNLQI.GSNPELHANQFTTHTSLTLN.RDWLVNSCFSV.FVEPVHYKEFCRVPVSHWRPRVPYIYSDRGAPNRE
VSHLSPHLDVNVDTSK.RFPN.NNLFIAFAYKYDGS.FVVLKCTFIL..RCGLH.F.IEKKLVITLSFSPY.PSYSLLIHVD
FPDMKPFTEE
```

Sliding 8 amino acid window search  
Database searched = AD\_2010  
Query = frame1

Start time: Wed Feb 3 22:04:48 GMT 2010 Finish time: Wed Feb 3 22:04:48 GMT 2010

No 8 amino acid matches exist between frame1 and the AD\_2010 database

# fasta34 frame1.pep /genedata/1/db/AD\_2010 -Q -E 1 -O frame1.pep\_ad.fasta  
FASTA searches a protein or DNA sequence data bank version 3.4t26 July 7, 2006  
Please cite:  
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

frame1, 2417 aa  
vs /genedata/1/db/AD\_2010 library

	opt	E()
< 20	2	0:==
22	0	0:
24	0	0:
26	0	0:
28	1	0:==
30	8	2:*==
32	9	8:==*
34	32	21:=====*

one = represents 3 library sequences

36	59	44:=====*
38	74	72:=====*
40	99	101:=====*
42	89	123:=====*
44	98	136:=====*
46	145	138:=====*
48	120	132:=====*
50	150	121:=====*
52	125	106:=====*
54	91	91:=====*
56	64	76:=====*
58	69	62:=====*
60	41	50:=====*
62	73	40:=====*
64	22	32:=====*
66	21	25:=====*
68	19	20:=====*
70	9	16:=====*
72	11	12:=====*
74	7	10:=====*
76	6	7:=====*
78	2	6:=====*
80	3	4:=====*
82	4	3:=====*
84	1	3:=====*
86	7	2:=====*
88	0	2:=====*
90	2	1:=====*
92	1	1:=====*
94	2	1:=====*
96	2	1:=====*
98	1	0:=====*
100	0	0:=====*
102	0	0:=====*
104	1	0:=====*
106	0	0:=====*
108	0	0:=====*
110	0	0:=====*
112	0	0:=====*
114	0	0:=====*
116	0	0:=====*
118	0	0:=====*
>120	1	0:=====*

inset = represents 1 library sequences

331323 residues in 1471 sequences  
Expectation\_n fit: rho(ln(x))= 3.60570.00375; mu= 22.4960 0.193  
mean\_var=57.895415.216, 0's: 2 Z-trim: 3 B-trim: 8 in 1/42  
Lambda= 0.168559  
Kolmogorov-Smirnov statistic: 0.0289 (N=28) at 48  
  
FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2  
join: 44, opt: 32, open/ext: -10/-2, width: 16  
The best scores are: opt bits E(1471)  
gi|76782247|gb|ABA54897.1| hydrophobic seed protei ( 134) 96 32.7 0.067  
gi|2833325|sp|Q25641.1|CRPI\_PERAM RecName: Full=A1 ( 685) 87 31.5 0.83  
  
>>gi|76782247|gb|ABA54897.1| hydrophobic seed protein pr (134 aa)  
initn: 97 initl: 91 opt: 96 Z-score: 123.4 bits: 32.7 E(): 0.067  
Smith-Waterman score: 96; 32.787% identity (54.098% similar) in 61 aa overlap  
(1964-2024:3-61)

1940 1950 1960 1970 1980 1990

```
frame1 PPPRTFINQSCRFKLSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTTVPLIFEGACAFA
      ::  .:  : :...  .  .  .
gi|767      MGSKVVASVALLSINIL-FISMVSSSSHYDP
              10          20          30

      2000      2010      2020      2030      2040      2050
frame1 FKPPEDADLDFSPPRFAGPPPSLLPAPPESGEGETVGKNDEVAATIMNF.GRFSGRTS
      ::  .  .:..  : :...  : : :
gi|767 PPPCYVPAPLTTPPPSLSPPPSLSP-PPPSGPPSCPDLSVCLNILDGSPADCCALIADLV
              40          50          60          70          80          90

      2060      2070      2080      2090      2100      2110
frame1 SLEIMTIKQKVSTKCCDRPLLRRHYSQLFINREK.YCSY.LGAYT.AFRKSRNEKGYIN
gi|767 DLEASVCLCIQLRVLGIVNLDLNLQLILNACGPSYSPNATCPRT
              100          110          120          130

>>gi|2833325|sp|Q25641.1|CRPI_PERAM RecName: Full=Allerg (685 aa)
  initn: 56 init1: 56 opt: 87 Z-score: 103.8 bits: 31.5 E(): 0.83
Smith-Waterman score: 87; 26.263% identity (56.566% similar) in 99 aa overlap
(2178-2273:213-310)

      2150      2160      2170      2180      2190      2200
frame1 .QKSASHLMQ.LCNVSFICLFRPCFSRILWADVSYGTFLLESRLRYFTIILYSLRHHYYFF
              .:..  .:  : :...  : :  .  .  .
gi|283 KNAGSGEHHSHVIPVNFTLRTQDHLLAYFTSDVNLNAFNTY-YRYYPSWYNTTLYGHNI
              190          200          210          220          230          240

      2210      2220      2230      2240      2250      2260
frame1 TF.GHSFF-VFEGYCLFCLEYVWLDLFDYSIILYKSNIYIKRKNLQI.GSN--PELHA
      :...  .:  .  .:  .  .:  .  .:  .  .:  .  .:  .  .:  .  .:  .  .
gi|283 DRERGEQFYTYKQIYARYFLERLSNDLPDVYPFYYSKPVKSAYNPNLRYHNGEEMPVRPS
              250          260          270          280          290          300

      2270      2280      2290      2300      2310      2320
frame1 NQFTTHTSLTLN.RDWLVNSCFV.FVEFVHYKGFCRYPSHWRPRVPYIYSDRGAPNRE
      :...  .:  .
gi|283 NMYVTNFDLYIADIKNYEKRVEDAIDFGYAFDEHMKPHSLYHDVHGMEYLADMIENMD
              310          320          330          340          350          360

2417 residues in 1 query sequences
331323 residues in 1471 library sequences
Scomplib [34t26]
start: Wed Feb 3 22:04:47 2010 done: Wed Feb 3 22:04:47 2010
Total Scan time: 0.100 Total Display time: 0.020

Function used was FASTA [version 3.4t26 July 7, 2006]

# fasta34 frame1.pep /genedata/1/db/TOX_2010 -Q -E 1 -O frame1.pep_tx.fasta
FASTA searches a protein or DNA sequence data bank version 3.4t26 July 7, 2006
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

frame1, 2417 aa
vs /genedata/1/db/TOX_2010 library

      opt      E()
< 20      60      0:====
```

```
22      0      0:      one = represents 17 library sequences
24      0      0:
26      0      0:
28      1      2:*
30      6      12:*
32      23      45:==*
34      113      122:====*
36      186      250:====*
38      399      414:====*
40      889      577:====*
42      636      706:====*
44      607      779:====*
46      1019      793:====*
48      813      759:====*
50      533      693:====*
52      559      609:====*
54      474      520:====*
56      359      435:====*
58      347      357:====*
60      428      289:====*
62      127      232:====*
64      168      184:====*
66      130      146:====*
68      118      115:====*
70      94      90:====*
72      97      70:====*
74      76      55:====*
76      62      43:====*
78      31      33:==*
80      14      26:==*
82      11      20:==*
84      11      16:*
86      4      12:*
88      4      9:*
90      3      7:*
92      10      6:*
94      3      4:*
96      8      3:*
98      5      3:*
100      1      2:*
102      0      2:*
104      0      1:*
106      1      1:*
108      5      1:*
110      5      1:*
112      2      0:=
114      0      0:
116      0      0:
118      0      0:
>120      1      0:=
2069351 residues in 8448 sequences
  Expectation_n fit: rho(ln(x))= 3.57820.000588; mu= 22.2556 0.030
  mean_var=54.324412.024, 0's: 60 Z-trim: 61 B-trim: 133 in 1/61
  Lambda= 0.174011
  Kolmogorov-Smirnov statistic: 0.0278 (N=29) at 48

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 44, opt: 32, open/ext: -10/-2, width: 16
The best scores are:
gi|158113520|gb|ABW15717.1| Zeta toxin family prot ( 383) 97 34.1 0.43
```

>>gi|158113520|gb|ABW15717.1| Zeta toxin family protein (383 aa)  
initn: 49 init1: 49 opt: 97 Z-score: 122.5 bits: 34.1 E(): 0.43  
Smith-Waterman score: 97; 33.333% identity (59.259% similar) in 54 aa overlap  
(1983-2031:330-383)

```

      1960      1970      1980      1990      2000
frame1 NGFMGSKSTWISKGRSSCLKLSLDVTTVVPLIFEGACAFKPPEDADLDIFS---PPR
      : : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . :
gi|158 ETSTKAGRSRSPIPASGTDTAATPDAAASPADPQITPAAPSHTTQSPAQASFGSGVKPSPPR
      300      310      320      330      340      350

      2010      2020      2030      2040      2050      2060
frame1 F-AGPPPSLLPAPPESGEVITGKNDEVAAVATIMNF.GRFSGRTSLSLEIMTIKQKVTST
      . : . : . : . : . : . : . : . : . : . : . : . : . : . : . : . :
gi|158 RGSGLPPPGTTPPGPGQIGQGPAKGR
      360      370      380
```

2417 residues in 1 query sequences  
2069351 residues in 8448 library sequences  
Scomplib [34t26]  
start: Wed Feb 3 22:04:48 2010 done: Wed Feb 3 22:04:48 2010  
Total Scan time: 0.610 Total Display time: 0.010

Function used was FASTA [version 3.4t26 July 7, 2006]

# fasta34 frame1.pep /genedata/1/db/PRT\_2010 -Q -E 1 -O frame1.pep\_prt.fasta  
FASTA searches a protein or DNA sequence data bank version 3.4t26 July 7, 2006  
Please cite:  
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

frame1, 2417 aa  
vs /genedata/1/db/PRT\_2010 library

```

      opt      E()
< 20 279198      0:=====
22 102      0:===== one = represents 26111 library sequences
24 226      17:*
26 570      374:*
28 2238      4038:*
30 11944      24533:*
32 59499      94862:===*
34 193895      257255:===== *
36 442020      528341:===== *
38 783292      873150:===== *
40 1105434      1217968:===== *
42 1354950      1488818:===== *
44 1516386      1642304:===== *
46 1566653      1672728:===== *
48 1542343      1601444:===== *
50 1429997      1461324:===== *
52 1281649      1284748:===== *
54 1125766      1097400:===== *
56 974486      916665:===== *
58 792708      752565:===== *
60 658401      609622:===== *
62 523004      488736:===== *
64 418360      388689:===== *
66 335973      307208:===== *
```

```

68 289181      241643:===== *
70 229872      189366:===== *
72 181669      147971:===== *
74 148018      115368:===== *
76 122907      89794:===== *
78 94144      69795:===== *
80 72307      54195:===== *
82 57224      41458:===== *
84 45701      32840:===== *
86 34674      25410:===== *
88 28518      19661:===== *
90 22718      15212:===== *
92 18421      11771:===== *
94 14554      9107:===== *
96 11193      7047:===== *
98 8192      5452:===== *
100 6494      4219:===== *
102 5111      3264:===== *
104 3988      2526:===== *
106 3435      1954:===== *
108 2468      1512:===== *
110 2045      1170:===== *
112 1665      905:===== *
114 1481      700:===== *
116 1151      542:===== *
118 949      419:===== *
>120 4051      324:===== *
4761287459 residues in 17815538 sequences
statistics sampled from 60000 to 17807527 sequences
Expectation_n fit: rho(ln(x))= 5.08390.000188; mu= 15.6481 0.010
mean_var=76.429815.305, 0's: 954 Z-trim: 958 B-trim: 61 in 1/64
Lambda= 0.146704
```

Kolmogorov-Smirnov statistic: 0.0491 (N=29) at 52

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2  
join: 44, opt: 32, open/ext: -10/-2, width: 16

The best scores are:

	opt	bits	E(17815538)
gi 58814 emb CAA29528.1  unnamed protein product [ ( 512)	889	199.2	2.4e-47
gi 90192131 gb ABD91726.1  acyl-ACP thioesterase [ ( 416)	568	131.2	5.8e-27
gi 133754634 gb ABO38556.1  acyl ACP-thioesterase ( 413)	394	94.4	7e-16
gi 133754638 gb ABO38558.1  acyl ACP-thioesterase ( 413)	394	94.4	7e-16
gi 133754636 gb ABO38557.1  acyl ACP-thioesterase ( 413)	394	94.4	7e-16
gi 133754632 gb ABO38555.1  acyl ACP-thioesterase ( 413)	394	94.4	7e-16
gi 133754630 gb ABO38554.1  acyl ACP-thioesterase ( 413)	394	94.4	7e-16
gi 325369 gb AAA43816.1  VI [Figwort mosaic virus] ( 538)	382	91.9	5e-15
gi 156900676 gb ABU96744.1  chloroplast acyl-ACP t ( 418)	347	84.4	7e-13
gi 248570279 gb ACT09366.1  stearyl acyl-ACP-thio ( 418)	347	84.4	7e-13
gi 229410505 gb ACQ57188.1  acyl acyl-carrier-prot ( 436)	345	84.0	9.7e-13
gi 229410506 gb ACQ63293.1  acyl acyl-carrier-prot ( 434)	330	80.8	8.7e-12
gi 229410507 gb ACQ57190.1  acyl acyl-carrier-prot ( 434)	329	80.6	1e-11
gi 229358082 gb ACQ57189.1  acyl acyl-carrier-prot ( 420)	327	80.2	1.3e-11
gi 83778888 gb ABC47311.1  FATB [Populus tomentosa ( 421)	324	79.6	2e-11
gi 217072522 gb ACJ84621.1  unknown [Medicago trun ( 181)	314	77.1	4.7e-11
gi 3410682 gb AAC31104.1 I90742 Sequence 23 from p ( 417)	313	77.2	1e-10
gi 147858246 emb CAN81819.1  hypothetical protein ( 400)	310	76.6	1.5e-10
gi 3410692 gb AAC31114.1 I90752 Sequence 33 from p ( 416)	310	76.6	1.6e-10
gi 4529965 emb CAA02764.1  unnamed protein product ( 177)	301	74.4	3.1e-10
gi 81362039 gb ABB71581.1  acyl-acyl carrier prote ( 414)	296	73.6	1.2e-09
gi 118489518 gb ABK96561.1  unknown [Populus trich ( 418)	295	73.4	1.4e-09
gi 222855220 gb EEE92767.1  predicted protein [Pop ( 418)	295	73.4	1.4e-09
gi 495227 emb CAA54060.1  acyl-ACP thioesterase [C ( 414)	286	71.5	5.3e-09
gi 2298907 emb CAA02760.1  unnamed protein product ( 414)	286	71.5	5.3e-09

gi|2298917|emb|CAA02765.1| unnamed protein product ( 414) 286 71.5 5.3e-09  
gi|12057063|emb|CAC19933.1| thioesterase type B [C ( 414) 286 71.5 5.3e-09  
gi|60392546|gb|AA19379.1| acyl-ACP thioesterase [ ( 430) 286 71.5 5.5e-09  
gi|60392552|gb|AA19382.1| acyl-ACP thioesterase [ ( 430) 286 71.5 5.5e-09  
gi|60392556|gb|AA19384.1| acyl-ACP thioesterase [ ( 430) 286 71.5 5.5e-09  
gi|60392558|gb|AA19385.1| acyl-ACP thioesterase [ ( 430) 286 71.5 5.5e-09  
gi|60392548|gb|AA19380.1| acyl-ACP thioesterase [ ( 430) 286 71.5 5.5e-09  
gi|2687425|gb|AAB88824.1| FatB thioesterase [Helia ( 430) 286 71.5 5.5e-09  
gi|60392562|gb|AA19377.1| acyl-ACP thioesterase [ ( 430) 286 71.5 5.5e-09  
gi|33325193|gb|AAQ08202.1| acyl-ACP thioesterase F ( 430) 286 71.5 5.5e-09  
gi|60392560|gb|AA19386.1| acyl-ACP thioesterase [ ( 430) 286 71.5 5.5e-09  
gi|60392542|gb|AA19377.1| acyl-ACP thioesterase [ ( 430) 286 71.5 5.5e-09  
gi|60392550|gb|AA19381.1| acyl-ACP thioesterase [ ( 430) 286 71.5 5.5e-09  
gi|60392544|gb|AA19378.1| acyl-ACP thioesterase [ ( 430) 286 71.5 5.5e-09  
gi|60392554|gb|AA19383.1| acyl-ACP thioesterase [ ( 430) 286 71.5 5.5e-09  
gi|18073131|emb|CAC80370.1| Acyl-ACP thioesterase ( 430) 277 69.6 2.1e-08  
gi|6048397|gb|AAF02215.1|AF076535\_1 palmitoyl-acyl ( 413) 276 69.4 2.3e-08  
gi|157417724|gb|ABV54795.1| acyl-ACP thioesterase ( 419) 276 69.4 2.3e-08  
gi|223545508|gb|EEF47013.1| palmitoyl-acyl carrier ( 419) 276 69.4 2.3e-08  
gi|3410687|gb|AAC31109.1|I90747 Sequence 28 from p ( 415) 268 67.7 7.5e-08  
gi|758701|gb|AAC48990.1| 16:0-ACP thioesterase pre ( 415) 268 67.7 7.5e-08  
gi|4104242|gb|AAD01982.1| palmitoyl-acyl carrier p ( 410) 259 65.8 2.8e-07  
gi|229358077|gb|ACQ57187.1| acyl acyl-carrier-prot ( 423) 259 65.8 2.8e-07  
gi|18073133|emb|CAC80371.1| Acyl-ACP thioesterase ( 430) 258 65.6 3.3e-07  
gi|3859830|gb|AAC72882.1| thioesterase FatB1-1 [Cu ( 414) 254 64.7 5.8e-07  
gi|222866396|gb|EEF03527.1| predicted protein [Pop ( 409) 250 63.9 1e-06  
gi|223271795|emb|CAX36780.1| unnamed protein produ ( 166) 238 61.0 3.1e-06  
gi|804948|emb|CAA85388.1| acyl-(acyl carrier prote ( 412) 239 61.6 5.2e-06  
gi|634003|emb|CAA85387.1| acyl-(acyl carrier prote ( 412) 238 61.3 6.1e-06  
gi|6664317|gb|AAF22899.1|AC006932\_16 T27G7.19 [Ara ( 412) 238 61.3 6.1e-06  
gi|15451198|gb|AAK96870.1| putative oleoyl-[acyl-c ( 412) 238 61.3 6.1e-06  
gi|17978739|gb|AAL47363.1| putative oleoyl-[acyl-c ( 412) 238 61.3 6.1e-06  
gi|16648915|gb|AAL24309.1| putative oleoyl-[acyl-c ( 412) 238 61.3 6.1e-06  
gi|30725684|gb|AAP37864.1| At1g08510 [Arabidopsis ( 412) 238 61.3 6.1e-06  
gi|223271141|emb|CAX36774.1| unnamed protein produ ( 168) 219 57.0 5.1e-05  
gi|223271117|emb|CAX36756.1| unnamed protein produ ( 168) 219 57.0 5.1e-05  
gi|223271145|emb|CAX36777.1| unnamed protein produ ( 169) 214 55.9 0.00011  
gi|223271121|emb|CAX36759.1| unnamed protein produ ( 169) 214 55.9 0.00011  
gi|223271109|emb|CAX36750.1| unnamed protein produ ( 169) 214 55.9 0.00011  
gi|193290692|gb|ACF17654.1| putative acyl-ACP thio ( 421) 218 57.1 0.00012  
gi|223271105|emb|CAX36747.1| unnamed protein produ ( 168) 212 55.5 0.00014  
gi|223271129|emb|CAX36765.1| unnamed protein produ ( 168) 212 55.5 0.00014  
gi|12057065|emb|CAC19934.1| fatty acyl-(ACP) thioe ( 412) 216 56.7 0.00015  
gi|2298919|emb|CAA02766.1| unnamed protein product ( 412) 216 56.7 0.00015  
gi|222855219|gb|EEE92766.1| predicted protein [Pop ( 407) 214 56.3 0.0002  
gi|6434144|emb|CAB60830.1| acyl-(ACP) thioesterase ( 419) 214 56.3 0.00021  
gi|2298913|emb|CAA02763.1| unnamed protein product ( 419) 214 56.3 0.00021  
gi|1930081|gb|AAB51525.1| acyl-ACP thioesterase [G ( 406) 213 56.0 0.00024  
gi|223271133|emb|CAX36768.1| unnamed protein produ ( 169) 208 54.7 0.00025  
gi|10050769|gb|AAE27676.1| Sequence 3 from patent ( 411) 208 55.0 0.00049  
gi|1215718|gb|AAC49179.1| thioesterase [Cuphea pal ( 411) 208 55.0 0.00049  
gi|223271101|emb|CAX36744.1| unnamed protein produ ( 168) 201 53.2 0.00071  
gi|3410686|gb|AAC31108.1|I90746 Sequence 27 from p ( 415) 205 54.4 0.00077  
gi|1292906|gb|AAC49269.1| FatB2 [Cuphea hookeriana ( 415) 205 54.4 0.00077  
gi|2459529|gb|AAB71729.1| acyl-ACP thioesterase [M ( 429) 203 53.9 0.0011  
gi|223271137|emb|CAX36771.1| unnamed protein produ ( 166) 197 52.3 0.0013  
gi|223271113|emb|CAX36753.1| unnamed protein produ ( 166) 197 52.3 0.0013  
gi|111610624|gb|ABH11710.1| palmitoyl-ACP thioeste ( 414) 200 53.3 0.0016  
gi|90018255|gb|ABD83939.1| palmitoyl-ACP thioester ( 417) 199 53.1 0.0019  
gi|223271125|emb|CAX36762.1| unnamed protein produ ( 167) 193 51.5 0.0023  
gi|1336008|gb|AAC49784.1| acyl-ACP thioesterase [C ( 408) 194 52.0 0.0038  
gi|32351197|gb|AAP75616.1| inclusion body matrix p ( 504) 189 51.0 0.0094

gi|2298909|emb|CAA02761.1| unnamed protein product ( 375) 186 50.3 0.012  
gi|241933353|gb|EES06498.1| hypothetical protein S ( 416) 185 50.1 0.015  
gi|81361963|gb|ABB71580.1| seed-specific acyl-acyl ( 417) 183 49.7 0.02  
gi|81361906|gb|ABB71579.1| seed-specific acyl-acyl ( 417) 183 49.7 0.02  
gi|215768010|dbj|BAH00239.1| unnamed protein produ ( 427) 181 49.3 0.027  
gi|113594816|dbj|BAF18690.1| Os06g0143400 [Oryza s ( 427) 181 49.3 0.027  
gi|5803272|dbj|BAA83582.1| putative acyl-ACP thioe ( 427) 181 49.3 0.027  
gi|125554012|gb|EAY99617.1| hypothetical protein O ( 427) 181 49.3 0.027  
gi|125596021|gb|EAE35801.1| hypothetical protein O ( 427) 181 49.3 0.027  
gi|215704277|dbj|BAG93117.1| unnamed protein produ ( 427) 181 49.3 0.027  
gi|30962820|gb|AAD42220.2|AF147879\_1 palmitoyl-acy ( 417) 176 48.2 0.055  
gi|112455672|gb|ABI18986.1| palmitoyl-ACP thioeste ( 411) 174 47.8 0.073  
gi|10050770|gb|AAE27677.1| Sequence 4 from patent ( 411) 171 47.2 0.11  
gi|1215720|gb|AAC49180.1| thioesterase [Cuphea pal ( 411) 171 47.2 0.11  
gi|241914680|gb|EER87824.1| hypothetical protein S ( 428) 171 47.2 0.12  
gi|270240896|emb|CB128125.1| unnamed protein produ ( 408) 164 45.7 0.31  
gi|211967563|gb|EEB02759.1| hypothetical protein T ( 542) 163 45.6 0.45  
gi|223947635|gb|ACN27901.1| unknown [Zea mays] ( 426) 161 45.1 0.5  
gi|195635773|gb|ACG37355.1| myristoyl-acyl carrier ( 426) 161 45.1 0.5  
gi|141446876|gb|ABO87585.1| inclusion body matrix ( 112) 153 42.9 0.6

>>gi|58814|emb|CAA29528.1| unnamed protein product [Figw (512 aa)  
initn: 889 initl: 889 opt: 889 Z-score: 1012.7 bits: 199.2 E(): 2.4e-47  
Smith-Waterman score: 889; 100.000% identity (100.000% similar) in 135 aa overlap  
(100-234:375-509)

	70	80	90	100	110	120
frame1	NLYFDVSTFTSKWGLR.ETSRMRPSSGS.FSVQSLNKVRVQSLQTISQKLQINEESSI					
		.....				
gi 588	STRYEPYHFIFIQIGIAKTKKELPSSKVKCEEFSVQSLNKVRVQSLQTISQKLQINEESSI					
	350	360	370	380	390	400
	130	140	150	160	170	180
frame1	KVNYCSSTCIMVSKFKQKTTSTEDLKLGVGFESNLVNIQIACGDQTKKEWCRIVRRTYQK					
	.....					
gi 588	KVNYCSSTCIMVSKFKQKTTSTEDLKLGVGFESNLVNIQIACGDQTKKEWCRIVRRTYQK					
	410	420	430	440	450	460
	190	200	210	220	230	240
frame1	HLCLYCKDKADSSSTSGEQNNVEKSCPDSPLTNAYDERSDDHKRISLSSGFSIPDWVQS					
	.....					
gi 588	HLCLYCKDKADSSSTSGEQNNVEKSCPDSPLTNAYDERSDDHKRIPSI					
	470	480	490	500	510	
	250	260	270	280	290	300
frame1	TRYEPYHFIFIQIGIAKTKKELPSSKVKCEEFDIKLDIGSFSLEGGSWNGTHMVVIINHFFH					

>>gi|90192131|gb|ABD91726.1| acyl-ACP thioesterase [Glyc (416 aa)  
initn: 594 initl: 560 opt: 568 Z-score: 646.7 bits: 131.2 E(): 5.8e-27  
Smith-Waterman score: 568; 81.651% identity (92.661% similar) in 109 aa overlap  
(1857-1963:1-109)

	1830	1840	1850	1860	1870	1880
frame1	SQHFLVLTFCFMMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG					
	.....					
gi 901				MVATAATSSFFPVTSPSPDSGGAGSKLGGG		
				10	20	30
	1890	1900	1910	1920	1930	1940
frame1	PANLGLGKSKSASSGGGLKAKAQAQPSKINGTTVVTSKESFKHDDLPSPPPRTFINQSCRF					
	.....					

Monsanto Company

Regulatory Product Characterization Center  
RAR-10-070





```

      2000      2010      2020      2030      2040      2050
frame1 EDADLDFSPPRFAGPPPSLLPAPPESGEGEVTGKNDEVA AAVATIMNF.GRFSGR TSSLEI

gi|217 GADRTASINTIMNHLQETALNHVKTAGLLGNFGFS
      150      160      170      180

>>gi|3410682|gb|AAC311104.1|I90742 Sequence 23 from paten (417 aa)
  initn: 345 initl: 96 opt: 313 Z-score: 355.0 bits: 77.2 E(): 1e-10
Smith-Waterman score: 313; 46.099% identity (71.631% similar) in 141 aa overlap
(1857-1986:1-138)

      1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
      ..... : : : : :
gi|341 MASTAATAACFFPVSSPSDSVAKTKNIGS-
      10      20

      1890      1900      1910      1920      1930      1940
frame1 PANLGGGLKSKSASSGGL--KAKAQAPSKINGTTVVTSK--ESFKHDDDLPSPPPTFINQ
      ..... : : : : :
gi|341 -ASLGGMKAQS--SSGGLQVKASAQAPSKINGTSVGLTKPSESLKNEDEMPSSHPTFINQ
      30      40      50      60      70      80

      1950      1960      1970      1980      1990
frame1 SCRFBK--LSVFENGFMMSGKSTWIS---KGRSSSCLKLSLDVTTVVP--LIFEGACAFAFK
      .. : : : : :
gi|341 LPDWSMLLAIAITTFILAAEKQWMMMLDWKPRRSDMLIDPFGIGRIVQDGLIFRQNFIRSIRSY
      90      100      110      120      130      140

      2000      2010      2020      2030      2040      2050
frame1 PPEDADLDFSPPRFAGPPPSLLPAPPESGEGEVTGKNDEVA AAVATIMNF.GRFSGR TSSSL

gi|341 EIGADRTASIEITLMNHLQETALNHVKSAGLLGDGFGSTPGMKCKNLIWVVTRMQVVVDRY
      150      160      170      180      190      200

>>gi|147858246|emb|CAN81819.1| hypothetical protein [Vit (400 aa)
  initn: 248 initl: 107 opt: 310 Z-score: 351.9 bits: 76.6 E(): 1.5e-10
Smith-Waterman score: 310; 46.479% identity (67.606% similar) in 142 aa overlap
(1857-1986:1-142)

      1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
      ..... : : : : :
gi|147 MVATAATSAFFAVASPSDDPAKPKPGX
      10      20      30

      1890      1900      1910      1920      1930      1940
frame1 GPANLGGGLKSKSASSGGL--KAKAQAPSKINGTTV--VTSKESFKHDDDLPSPPPTFIN
      : : : : :
gi|147 GSAILRGIKSRNAPSGSLQVKANAQAPPKINGTTVGYTSSAEGVKIEDDMSSPPPTFIN
      40      50      60      70      80      90

      1950      1960      1970      1980      1990
frame1 QSCRFBK--LSVFENGFMMSGKSTWIS---KGRSSSCLKLSLDVTTVVP--LIFEGACAFAF
      : .. : : : :
gi|147 QLPDWSMLLAIAITTFILAAEKQWMMMLDWKPRRSDMLIDPFGLGKIVQDGLVFRQNFIRS
      100      110      120      130      140      150

      2000      2010      2020      2030      2040      2050
frame1 KPPEDADLDFSPPRFAGPPPSLLPAPPESGEGEVTGKNDEVA AAVATIMNF.GRFSGR TSS

```

```

gi|147 YEIGADRTASIEITLMNHLQETALNHVRTAGLLGDGFGSTPEMSIRNLIWVVTRMQVVVDR
      160      170      180      190      200      210

>>gi|3410692|gb|AAC31114.1|I90752 Sequence 33 from paten (416 aa)
  initn: 167 initl: 167 opt: 310 Z-score: 351.6 bits: 76.6 E(): 1.6e-10
Smith-Waterman score: 310; 47.518% identity (71.631% similar) in 141 aa overlap
(1857-1986:1-137)

      1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
      ..... : : : : :
gi|341 MASTAATAAFFPVSS--STDSVAKPKNIGS-
      10      20

      1890      1900      1910      1920      1930      1940
frame1 PANLGGGLKSKSASSGGL--KAKAQAPSKINGTTVVTSK--ESFKHDDDLPSPPPTFINQ
      ..... : : : : :
gi|341 -AGLGGGLKSKS--SSGRLQVKATAQAPSKINGTSVGLTKPVEGLKNEDDMPSPPPTFINQ
      30      40      50      60      70      80

      1950      1960      1970      1980      1990
frame1 SCRFBK--LSVFENGFMMSGKSTWIS---KGRSSSCLKLSLDVTTVVP--LIFEGACAFAFK
      .. : : : : :
gi|341 LPDWSMLLAIAITTFILAAEKQWMMMLDWKPRRSDMLIDPFGIGRIVQDGLIFRQNFIRSIRSY
      90      100      110      120      130      140

      2000      2010      2020      2030      2040      2050
frame1 PPEDADLDFSPPRFAGPPPSLLPAPPESGEGEVTGKNDEVA AAVATIMNF.GRFSGR TSSSL

gi|341 EIGADRTASIEITLMNHLQETALNHVKTAGLLGDGFGATPEMCKKNLIWVVTRMQVVVDRY
      150      160      170      180      190      200

>>gi|4529965|emb|CAA02764.1| unnamed protein product [Cu (177 aa)
  initn: 161 initl: 161 opt: 301 Z-score: 346.3 bits: 74.4 E(): 3.1e-10
Smith-Waterman score: 301; 47.321% identity (76.786% similar) in 112 aa overlap
(1857-1963:1-110)

      1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
      ..... : : : : :
gi|452 MVATAASSAFFPV--PSADTSSRPKGLNG
      10      20

      1890      1900      1910      1920      1930      1940
frame1 PANLGGGLKSKSASSGGL--KAKAQAPSKINGTTVVTSKESFKHDDDLPS-PPPTFINQS
      ..... : : : : :
gi|452 PSSFSPLKPKSIPNGGLQVKASASAPPKINGSSVGLKSGGLKTHDDAPSAPPPTFINQL
      30      40      50      60      70      80

      1950      1960      1970      1980      1990      2000
frame1 CRFBK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFAFKPPEDAD
      .. : : : : :
gi|452 PDWSMLLAIAITTFILAAEKQWMMMLDWKPKRLDMLIDPFGLGRIVQDGLVFRQNFIRSIRSYE
      90      100      110      120      130      140

>>gi|81362039|gb|ABB71581.1| acyl-acyl carrier protein t (414 aa)
  initn: 244 initl: 171 opt: 296 Z-score: 335.6 bits: 73.6 E(): 1.2e-09
Smith-Waterman score: 296; 47.321% identity (76.786% similar) in 112 aa overlap
(1857-1963:1-110)

      1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG

```



```

gi|813
          ::::::::::: ::::... :::::
          MVATAASSAFFPV--PSPDTSSRPKGLGNG
          10      20

1890      1900      1910      1920      1930      1940
frame1 PANLGGGLKSKSASSGGL--KAKAQAPSKINGTTVVTSKESFKHDDDLPS-PPPTFTINQS
      .... :: : ..::: :::::~: ::::: .. ::: .. :~: :::::~:
gi|813 SSSLSPKPKLMANGGLQVKANASAPPKINGSSVGLKSGSLKTQEDTPSAPPPTFTINQL
      30      40      50      60      70      80

          1950      1960      1970      1980      1990      2000
frame1 CRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
      .. ..:: . :::::~: .
gi|813 PDWSMLLAAITTVFLAAEKQWMMMLDWKPKRPDMLVDPFGLGRIVQDGLVFRQNFSIRSIE
      90      100      110      120      130      140

>>gi|118489518|gb|ABK96561.1| unknown [Populus trichocar (418 aa)
      initn: 308 init1: 132 opt: 295 Z-score: 334.4 bits: 73.4 E(): 1.4e-09
      Smith-Waterman score: 295; 48.673% identity (76.991% similar) in 113 aa overlap
      (1857-1963:1-111)

          1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSSGAGSKLGGG
      :::::::::::~: :::::~: .. .. ..:::
gi|118 MVATAAASSFFFPVPSPSGDA--KASKFGSV
          10      20

1890      1900      1910      1920      1930      1940
frame1 PANLGGGLKSKSASSGGLKAKA--QAPSKINGTTV--VTSKESFKHDDDLPSPPPTFTINQ
      :::::::::::~: :::::~: :::::~: : ..:~: :::::~: : ..:~: :::::~:
gi|118 SASLGGIKTKSASSGALQVNTNGQAPPKINGPPVGLAASVETLKNEDVSSPAPPTFTINQ
      30      40      50      60      70      80

          1950      1960      1970      1980      1990      2000
frame1 SCRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDA
      .. ..:: . :::::~: .
gi|118 LPDWSMLLAAITTMFLAAEKQWMMMLDWKPKRPDMLIDPFGIGRIVQDGLVFRQNFSIRSIE
      90      100      110      120      130      140

>>gi|222855220|gb|EEE92767.1| predicted protein [Populus (418 aa)
      initn: 313 init1: 132 opt: 295 Z-score: 334.4 bits: 73.4 E(): 1.4e-09
      Smith-Waterman score: 295; 48.673% identity (76.991% similar) in 113 aa overlap
      (1857-1963:1-111)

          1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSSGAGSKLGGG
      :::::::::::~: :::::~: .. .. ..:::
gi|222 MVAAAAASSFFFPVPSPSGDA--KASKFGSV
          10      20

1890      1900      1910      1920      1930      1940
frame1 PANLGGGLKSKSASSGGLKAK--AQAPSKINGTTV--VTSKESFKHDDDLPSPPPTFTINQ
      :::::::::::~: :::::~: :::::~: : ..:~: :::::~: : ..:~: :::::~:
gi|222 SASLGGIKTKSASSGALQVNTNAQAPPKINGPPVGLTASVETLKNEDVSSPAPPTFTINQ
      30      40      50      60      70      80

          1950      1960      1970      1980      1990      2000
frame1 SCRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDA
      .. ..:: . :::::~: .
gi|222 LPDWSMLLAAITTMFLAAEKQWMMMLDWKPKRPDMLIDPFGIGRIVQDGLVFRQNFSIRSIE
      90      100      110      120      130      140

```

```

>>gi|495227|emb|CAA54060.1| acyl-ACP thioesterase [Cuphe (414 aa)
      initn: 237 init1: 165 opt: 286 Z-score: 324.2 bits: 71.5 E(): 5.3e-09
      Smith-Waterman score: 286; 45.536% identity (77.679% similar) in 112 aa overlap
      (1857-1963:1-110)

          1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSSGAGSKLGGG
      :::::::::::~: :::::~: .. .. ..:::
gi|495 MVATAASSAFFPL--PSPDTSSRPKGLGNG
          10      20

1890      1900      1910      1920      1930      1940
frame1 PANLGGGLKSKSASSGGLKAKAQ--PSKINGTTVVTSKESFKHDDDLPS-PPPTFTINQS
      .... :: : ..:~: :::::~: : ..:~: .. :::~: : ..:~: :::::~:
gi|495 SSSLSPKPKFVANAGLVKVKASASAPPKINGSSVGLKSGSLKTQEDTPSVPPPTFTINQL
      30      40      50      60      70      80

          1950      1960      1970      1980      1990      2000
frame1 CRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
      .. ..:: . :::::~: .
gi|495 PDWSMLLAAITTVFLAAEKQWMMMLDWKPKRPDMLVDPFGLGSIVQGGLVFRQNFSIRSIE
      90      100      110      120      130      140

>>gi|2298907|emb|CAA02760.1| unnamed protein product [Cu (414 aa)
      initn: 237 init1: 165 opt: 286 Z-score: 324.2 bits: 71.5 E(): 5.3e-09
      Smith-Waterman score: 286; 45.536% identity (77.679% similar) in 112 aa overlap
      (1857-1963:1-110)

          1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSSGAGSKLGGG
      :::::::::::~: :::::~: .. .. ..:::
gi|229 MVATAASSAFFPL--PSPDTSSRPKGLGNG
          10      20

1890      1900      1910      1920      1930      1940
frame1 PANLGGGLKSKSASSGGLKAKAQ--PSKINGTTVVTSKESFKHDDDLPS-PPPTFTINQS
      .... :: : ..:~: :::::~: : ..:~: .. :::~: : ..:~: :::::~:
gi|229 SSSLSPKPKFVANAGLVKVKASASAPPKINGSSVGLKSGSLKTQEDTPSVPPPTFTINQL
      30      40      50      60      70      80

          1950      1960      1970      1980      1990      2000
frame1 CRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
      .. ..:: . :::::~: .
gi|229 PDWSMLLAAITTVFLAAEKQWMMMLDWKPKRPDMLVDPFGLGSIVQGGLVFRQNFSIRSIE
      90      100      110      120      130      140

>>gi|2298917|emb|CAA02765.1| unnamed protein product [Cu (414 aa)
      initn: 237 init1: 165 opt: 286 Z-score: 324.2 bits: 71.5 E(): 5.3e-09
      Smith-Waterman score: 286; 45.536% identity (77.679% similar) in 112 aa overlap
      (1857-1963:1-110)

          1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSSGAGSKLGGG
      :::::::::::~: :::::~: .. .. ..:::
gi|229 MVATAASSAFFPL--PSPDTSSRPKGLGNG
          10      20

1890      1900      1910      1920      1930      1940
frame1 PANLGGGLKSKSASSGGLKAKAQ--PSKINGTTVVTSKESFKHDDDLPS-PPPTFTINQS
      .... :: : ..:~: :::::~: : ..:~: .. :::~: : ..:~: :::::~:
gi|229 SSSLSPKPKFVANAGLVKVKASASAPPKINGSSVGLKSGSLKTQEDTPSVPPPTFTINQL
      30      40      50      60      70      80

```

```

1950      1960      1970      1980      1990      2000
frame1 CRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
..  .... .  .... .
gi|229 PDWSMLLAAITTVFLAAEKQWMMMLDWKPKRPDMLVDPFGLGSIVQGGLVFRQNFSIRSIE
90      100      110      120      130      140

>>gi|12057063|emb|CAC19933.1| thioesterase type B [Cuphe (414 aa)
  initn: 237 initl: 165 opt: 286  Z-score: 324.2  bits: 71.5 E(): 5.3e-09
Smith-Waterman score: 286;  45.536% identity (77.679% similar) in 112 aa overlap
(1857-1963:1-110)

1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
.....:.....  ....:
gi|120          MVATAASSAFFPL--PSPDTSSRPGKLGNG
                        10      20

1890      1900      1910      1920      1930      1940
frame1 PANLGGKSKSASSGGLKAKAQA--PSKINGTTVVTSKESFKHDDDLPS-PPPRTFINQS
....  :::::::::::  : :::::  ..  ::  ::  ::::::::::
gi|120 SSSLSPKPKFVANAGLKVKASASAPPKINGSSVGLKSGSLKTQEDTPSVPPPRTFINQL
30      40      50      60      70      80

1950      1960      1970      1980      1990      2000
frame1 CRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
..  .... .  .... .
gi|120 PDWSMLLAAITTVFLAAEKQWMMMLDWKPKRPDMLVDPFGLGSIVQGGLVFRQNFSIRSIE
90      100      110      120      130      140

>>gi|60392546|gb|AA19379.1| acyl-ACP thioesterase [Heli (430 aa)
  initn: 264 initl: 106 opt: 286  Z-score: 324.0  bits: 71.5 E(): 5.5e-09
Smith-Waterman score: 286;  44.737% identity (75.439% similar) in 114 aa overlap
(1857-1963:1-113)

1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
::  :::::::::::  ::  ::::
gi|603          MVAMSATASLFPVSSPKPHSGAKTSDKLGG
                        10      20      30

1890      1900      1910      1920      1930      1940
frame1 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLPSPPPRTFIN
.....:.....:.....  ..  ::  ::  ::::::::::
gi|603 EPGSAVVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDDYSTSPAPRTFIN
40      50      60      70      80

1950      1960      1970      1980      1990
frame1 QSCRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPED
:  ..  .... .  ....:
gi|603 QLPDWSMLLAAITTIFLAAEKQWMMLEWKTMRPDIADMDPFGLGRIVQDGLVFRQNFSI
90      100      110      120      130      140

>>gi|60392552|gb|AA19382.1| acyl-ACP thioesterase [Heli (430 aa)
  initn: 264 initl: 106 opt: 286  Z-score: 324.0  bits: 71.5 E(): 5.5e-09
Smith-Waterman score: 286;  44.737% identity (75.439% similar) in 114 aa overlap
(1857-1963:1-113)

1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
::  :::::::::::  ::  ::::
gi|603          MVAMSATASLFPVSSPKPHSGAKTSDKLGG
                        10      20      30

1890      1900      1910      1920      1930      1940
frame1 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLPSPPPRTFIN
.....:.....:.....  ..  ::  ::  ::::::::::
gi|603 EPGSAVVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDDYSTSPAPRTFIN
40      50      60      70      80

1950      1960      1970      1980      1990
frame1 QSCRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPED
:  ..  .... .  ....:
gi|603 QLPDWSMLLAAITTIFLAAEKQWMMLEWKTMRPDIADMDPFGLGRIVQDGLVFRQNFSI
90      100      110      120      130      140

>>gi|60392548|gb|AA19380.1| acyl-ACP thioesterase [Heli (430 aa)
  initn: 264 initl: 106 opt: 286  Z-score: 324.0  bits: 71.5 E(): 5.5e-09
```

```

10      20      30
1890      1900      1910      1920      1930      1940
frame1 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLPSPPPRTFIN
.....:.....:.....  ..  ::  ::  ::::::::::
gi|603 EPGSAVVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDDYSTSPAPRTFIN
40      50      60      70      80

1950      1960      1970      1980      1990
frame1 QSCRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPED
:  ..  .... .  ....:
gi|603 QLPDWSMLLAAITTIFLAAEKQWMMLEWKTMRPDIADMDPFGLGRIVQDGLVFRQNFSI
90      100      110      120      130      140

>>gi|60392556|gb|AA19384.1| acyl-ACP thioesterase [Heli (430 aa)
  initn: 264 initl: 106 opt: 286  Z-score: 324.0  bits: 71.5 E(): 5.5e-09
Smith-Waterman score: 286;  44.737% identity (75.439% similar) in 114 aa overlap
(1857-1963:1-113)

1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
::  :::::::::::  ::  ::::
gi|603          MVAMSATASLFPVSSPKPHSGAKTSDKLGG
                        10      20      30

1890      1900      1910      1920      1930      1940
frame1 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLPSPPPRTFIN
.....:.....:.....  ..  ::  ::  ::::::::::
gi|603 EPGSAVVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDDYSTSPAPRTFIN
40      50      60      70      80

1950      1960      1970      1980      1990
frame1 QSCRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPED
:  ..  .... .  ....:
gi|603 QLPDWSMLLAAITTIFLAAEKQWMMLEWKTMRPDIADMDPFGLGRIVQDGLVFRQNFSI
90      100      110      120      130      140

>>gi|60392558|gb|AA19385.1| acyl-ACP thioesterase [Heli (430 aa)
  initn: 264 initl: 106 opt: 286  Z-score: 324.0  bits: 71.5 E(): 5.5e-09
Smith-Waterman score: 286;  44.737% identity (75.439% similar) in 114 aa overlap
(1857-1963:1-113)

1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
::  :::::::::::  ::  ::::
gi|603          MVAMSATASLFPVSSPKPHSGAKTSDKLGG
                        10      20      30

1890      1900      1910      1920      1930      1940
frame1 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLPSPPPRTFIN
.....:.....:.....  ..  ::  ::  ::::::::::
gi|603 EPGSAVVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDDYSTSPAPRTFIN
40      50      60      70      80

1950      1960      1970      1980      1990
frame1 QSCRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPED
:  ..  .... .  ....:
gi|603 QLPDWSMLLAAITTIFLAAEKQWMMLEWKTMRPDIADMDPFGLGRIVQDGLVFRQNFSI
90      100      110      120      130      140

>>gi|60392548|gb|AA19380.1| acyl-ACP thioesterase [Heli (430 aa)
  initn: 264 initl: 106 opt: 286  Z-score: 324.0  bits: 71.5 E(): 5.5e-09
```

Smith-Waterman score: 286; 44.737% identity (75.439% similar) in 114 aa overlap (1857-1963:1-113)

```

      1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|603              MVAMSATASLFPVSSPKPHSGAKTSDKLGG
                        10      20      30
```

```

      1890      1900      1910      1920      1930      1940
frame1 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLPSPPRTFFIN
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|603 EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDDYSTSPAPRTFFIN
      40      50      60      70      80
```

```

      1950      1960      1970      1980      1990
frame1 QSCRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPED
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|603 QLPDWSMLLAAITTIFLAAEKQWMMLEWTKRPMIADMDPFGGLGRIVQDGLVFRQNFSI
      90      100      110      120      130      140
```

>>gi|2687425|gb|AAB88824.1| FatB thioesterase [Helianthu (430 aa)  
initn: 264 initl: 106 opt: 286 Z-score: 324.0 bits: 71.5 E(): 5.5e-09  
Smith-Waterman score: 286; 44.737% identity (75.439% similar) in 114 aa overlap (1857-1963:1-113)

```

      1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|268              MVAMSATASLFPVSSPKPHSGAKTSDKLGG
                        10      20      30
```

```

      1890      1900      1910      1920      1930      1940
frame1 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLPSPPRTFFIN
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|268 EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDDYSTSPAPRTFFIN
      40      50      60      70      80
```

```

      1950      1960      1970      1980      1990
frame1 QSCRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPED
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : ~
gi|268 QLPDWSMLLAAITTIFLAAEKQWMMLEWTKRPMIADMDPFGGLGRIVQDGLVFRQNFSI
      90      100      110      120      130      140
```

>>gi|60392562|gb|AAX19387.1| acyl-ACP thioesterase [Heli (430 aa)  
initn: 264 initl: 106 opt: 286 Z-score: 324.0 bits: 71.5 E(): 5.5e-09  
Smith-Waterman score: 286; 44.737% identity (75.439% similar) in 114 aa overlap (1857-1963:1-113)

```

      1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : ~
gi|603              MVAMSATASLFPVSSPKPHSGAKTSDKLGG
                        10      20      30
```

```

      1890      1900      1910      1920      1930      1940
frame1 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLPSPPRTFFIN
      : : : : : : : : : : : : : : : : : : : : : : : : : : ~
gi|603 EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDDYSTSPAPRTFFIN
      40      50      60      70      80
```

1950 1960 1970 1980 1990

```

frame1 QSCRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPED
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : ~
gi|603 QLPDWSMLLAAITTIFLAAEKQWMMLEWTKRPMIADMDPFGGLGRIVQDGLVFRQNFSI
      90      100      110      120      130      140
```

>>gi|33325193|gb|AAQ08202.1| acyl-ACP thioesterase FATB (430 aa)  
initn: 264 initl: 106 opt: 286 Z-score: 324.0 bits: 71.5 E(): 5.5e-09  
Smith-Waterman score: 286; 44.737% identity (75.439% similar) in 114 aa overlap (1857-1963:1-113)

```

      1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
      : : : : : : : : : : : : : : : : : : : : : : : : : : ~
gi|333              MVAMSATASLFPVSSPKPHSGAKTSDKLGG
                        10      20      30
```

```

      1890      1900      1910      1920      1930      1940
frame1 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLPSPPRTFFIN
      : : : : ~
gi|333 EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDDYSTSPAPRTFFIN
      40      50      60      70      80
```

```

      1950      1960      1970      1980      1990
frame1 QSCRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPED
      : : : : ~
gi|333 QLPDWSMLLAAITTIFLAAEKQWMMLEWTKRPMIADMDPFGGLGRIVQDGLVFRQNFSI
      90      100      110      120      130      140
```

>>gi|60392560|gb|AAX19386.1| acyl-ACP thioesterase [Heli (430 aa)  
initn: 264 initl: 106 opt: 286 Z-score: 324.0 bits: 71.5 E(): 5.5e-09  
Smith-Waterman score: 286; 44.737% identity (75.439% similar) in 114 aa overlap (1857-1963:1-113)

```

      1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
      : : : : ~
gi|603              MVAMSATASLFPVSSPKPHSGAKTSDKLGG
                        10      20      30
```

```

      1890      1900      1910      1920      1930      1940
frame1 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLPSPPRTFFIN
      : : : : ~
gi|603 EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDDYSTSPAPRTFFIN
      40      50      60      70      80
```

```

      1950      1960      1970      1980      1990
frame1 QSCRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPED
      : : : : ~
gi|603 QLPDWSMLLAAITTIFLAAEKQWMMLEWTKRPMIADMDPFGGLGRIVQDGLVFRQNFSI
      90      100      110      120      130      140
```

>>gi|60392542|gb|AAX19377.1| acyl-ACP thioesterase [Heli (430 aa)  
initn: 264 initl: 106 opt: 286 Z-score: 324.0 bits: 71.5 E(): 5.5e-09  
Smith-Waterman score: 286; 44.737% identity (75.439% similar) in 114 aa overlap (1857-1963:1-113)

```

      1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
      : : : : ~
gi|603              MVAMSATASLFPVSSPKPHSGAKTSDKLGG
                        10      20      30
```

```

      1890      1900      1910      1920      1930      1940
frame1 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLSPPPRTFIN
      :.... :..... :..... :..... :..... :.....
gi|603 EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDDYSTSPAPRTFIN
      40      50      60      70      80

      1950      1960      1970      1980      1990
frame1 QSCRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPED
      : .. :... :..... :.
gi|603 QLPDWSMLLAAITTTIFLAAEKQWMMLEWTKRPMIADMDPFGLGRIVQDGLVFRQNFSI
      90      100      110      120      130      140

>>gi|60392550|gb|AAx19381.1| acyl-ACP thioesterase [Heli (430 aa)
      initn: 264 init1: 106 opt: 286 Z-score: 324.0 bits: 71.5 E(): 5.5e-09
Smith-Waterman score: 286; 44.737% identity (75.439% similar) in 114 aa overlap
(1857-1963:1-113)

      1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSSGGAGS-KLGG
      :... :..... :... :.....
gi|603 MVAMSATASLFPVSSPKPHSGAKTSDKLGG
      10      20      30

      1890      1900      1910      1920      1930      1940
frame1 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLSPPPRTFIN
      :.... :..... :..... :..... :..... :.....
gi|603 EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDDYSTSPAPRTFIN
      40      50      60      70      80

      1950      1960      1970      1980      1990
frame1 QSCRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPED
      : .. :... :..... :.
gi|603 QLPDWSMLLAAITTTIFLAAEKQWMMLEWTKRPMIADMDPFGLGRIVQDGLVFRQNFSI
      90      100      110      120      130      140

>>gi|60392544|gb|AAx19378.1| acyl-ACP thioesterase [Heli (430 aa)
      initn: 264 init1: 106 opt: 286 Z-score: 324.0 bits: 71.5 E(): 5.5e-09
Smith-Waterman score: 286; 44.737% identity (75.439% similar) in 114 aa overlap
(1857-1963:1-113)

      1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSSGGAGS-KLGG
      :... :..... :... :.....
gi|603 MVAMSATASLFPVSSPKPHSGAKTSDKLGG
      10      20      30

      1890      1900      1910      1920      1930      1940
frame1 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLSPPPRTFIN
      :.... :..... :..... :..... :..... :.....
gi|603 EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDDYSTSPAPRTFIN
      40      50      60      70      80

      1950      1960      1970      1980      1990
frame1 QSCRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPED
      : .. :... :..... :.
gi|603 QLPDWSMLLAAITTTIFLAAEKQWMMLEWTKRPMIADMDPFGLGRIVQDGLVFRQNFSI
      90      100      110      120      130      140

>>gi|60392554|gb|AAx19383.1| acyl-ACP thioesterase [Heli (430 aa)
      initn: 264 init1: 106 opt: 286 Z-score: 324.0 bits: 71.5 E(): 5.5e-09
Smith-Waterman score: 286; 44.737% identity (75.439% similar) in 114 aa overlap
(1857-1963:1-113)

```

```

      1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSSGGAGS-KLGG
      :... :..... :... :.....
gi|603 MVAMSATASLFPVSSPKPHSGAKTSDKLGG
      10      20      30

      1890      1900      1910      1920      1930      1940
frame1 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLSPPPRTFIN
      :.... :..... :..... :..... :..... :.....
gi|603 EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDDYSTSPAPRTFIN
      40      50      60      70      80

      1950      1960      1970      1980      1990
frame1 QSCRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPED
      : .. :... :..... :.
gi|603 QLPDWSMLLAAITTTIFLAAEKQWMMLEWTKRPMIADMDPFGLGRIVQDGLVFRQNFSI
      90      100      110      120      130      140

>>gi|18073131|emb|CAC80370.1| Acyl-ACP thioesterase [Hel (430 aa)
      initn: 257 init1: 106 opt: 277 Z-score: 313.7 bits: 69.6 E(): 2.1e-08
Smith-Waterman score: 277; 43.860% identity (73.684% similar) in 114 aa overlap
(1857-1963:1-113)

      1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSSGGAGS-KLGG
      :... :..... :... :.....
gi|180 MVAMSATASLFPVSSPKPHSGAKTSDKLGG
      10      20      30

      1890      1900      1910      1920      1930      1940
frame1 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLSPPPRTFIN
      :.... :..... :..... :..... :..... :.....
gi|180 EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDDYSTSPAPRTFIN
      40      50      60      70      80

      1950      1960      1970      1980      1990
frame1 QSCRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPED
      : .. :... :..... :.
gi|180 XLPDWMXLLAAITTTIFLAAEKQWMMLEWTKRPMIADMDPFGLGRIVQDGLVFRQNFSI
      90      100      110      120      130      140

>>gi|6048397|gb|AAF02215.1|AF076535_1 palmitoyl-acyl car (413 aa)
      initn: 186 init1: 102 opt: 276 Z-score: 312.8 bits: 69.4 E(): 2.3e-08
Smith-Waterman score: 276; 50.442% identity (75.221% similar) in 113 aa overlap
(1857-1963:1-106)

      1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSSGGAGSKLGGG
      :..... :..... :..... :..... :..... :.....
gi|604 MVATAVTSAFFPVTSPDSSDSKNK----
      10      20

      1890      1900      1910      1920      1930      1940
frame1 PANLGGGLKSK-SASSGGL--KAKAQAPSKINGTTVVTSK-ESFKHDDDLSPPPRTFINQ
      :..... :..... :..... :..... :..... :.....
gi|604 --KLGSIKSKPSVSSGSLQVKANAQAPPKINGTVASTTPVEGSKNDDGASSPPRTFINQ
      30      40      50      60      70      80

      1950      1960      1970      1980      1990      2000
frame1 SCRFK--LSVFENGFMMSGKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDA
      : .. :... :..... :.

```



```
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDGGAGSKLGGG
gi|229      1890      1900      1910      1920      1930      1940
      1890      1900      1910      1920      1930      1940
frame1 -PANLG--GLKSKSASSGGLKAK--AQAPSKINGTTVVTSKESFKHDDDLSPPPRTFIN
      ...      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|229 VPSTVDARGIAKSKSASSGSLQVRTNAQAPSKVNGTKFGVV--EGLKGENDTPSHHRTIIN
      40      50      60      70      80

      1950      1960      1970      1980      1990
frame1 QSCRFK--LSVFENGFMGSKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPED
      : . . . . . : : : : : : :
gi|229 QLPDWSMLLAATTTIFLAAEKQWMMMLDWKPRPDMLTDPFGLGKIVQDGLVFSQNFIRS
      90      100      110      120      130      140

>>gi|18073133|emb|CAC80371.1| Acyl-ACP thioesterase [Hel (430 aa)
      initn: 236 initl: 106 opt: 258 Z-score: 292.0 bits: 65.6 E(): 3.3e-07
Smith-Waterman score: 258; 42.105% identity (72.807% similar) in 114 aa overlap
(1857-1963:1-113)

      1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDGGAGS-KLGG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|180      1830      1840      1850      1860      1870      1880
      1830      1840      1850      1860      1870      1880
frame1 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLSPPPRTFIN
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|180 EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH--GFKTDDYSTSPAXXTFIX
      40      50      60      70      80

      1950      1960      1970      1980      1990
frame1 QSCRFK--LSVFENGFMGSKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPED
      : . . . . . : : : : : : :
gi|180 QLPDWSMLLAATTTIFLAAEKQWMMLEWKTKRPDMIADMPFGLGRIVQDGLVFRQNFISI
      90      100      110      120      130      140

>>gi|3859830|gb|AAC72882.1| thioesterase FatB1-1 [Cuphea (414 aa)
      initn: 271 initl: 162 opt: 254 Z-score: 287.6 bits: 64.7 E(): 5.8e-07
Smith-Waterman score: 254; 42.857% identity (74.107% similar) in 112 aa overlap
(1857-1963:1-110)

      1830      1840      1850      1860      1870      1880
frame1 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDGGAGSKLGGG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|385      1830      1840      1850      1860      1870      1880
      1830      1840      1850      1860      1870      1880
frame1 PANLGGKSKSASSGGL--KAKAQAPSKINGTTVVTSKESFKHDDDLSPPPRTFINQS
      . . . . . : : : : : : : : : : : : : : : : : : : : : : :
gi|385 SSSFSPIKPKFVANGGLQVKANASAPPKINGSSVGLKCSLSLTKQEDTPSAPAPRTFINQL
      30      40      50      60      70      80

      1950      1960      1970      1980      1990      2000
frame1 CRFK--LSVFENGFMGSKSTWISKGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
      . . . . . : : : : : : :
gi|385 PDWSMLLAATTTVFLAAEKQWMMMLDWKPKRPDMLVDPFGLGSIVQHGLVFRQNFIRSIE
      90      100      110      120      130      140
```

2417 residues in 1 query sequences  
4761287459 residues in 17815538 library sequences  
Scomplib [34t26]  
start: Wed Feb 3 22:04:48 2010 done: Wed Feb 3 22:29:17 2010  
Total Scan time: 1319.250 Total Display time: 0.640  
  
Function used was FASTA [version 3.4t26 July 7, 2006]

Appendix 2. Bioinformatic analysis of polypeptide frame2

```
>frame2
SCPGNLHGSAMSMVMNEKKKE.LPIFFQFKNVDVRSVVIK.KYLILKRQITIRRIYRRKQ.TNYSNSEIFISTCLHSRPNGG
LDEKLHDCRGRVEANSQSKASTRSGYRVSKPLAKSYRRSMKNLQSK.TTVPAHASWSVSFRKRHPKKT.S.WASLKVLST
SSWLVGTRQKRNGAELLGAPTKSIFAFIAKIKQIPLVQVGNKITWKRAVLTAHSLMRMTNAVTTTKELA.AQDLAAEQIGFNQ
QGTSHITLFLKLVSPKPRRNSHPQRFRKNSISSLISEVSLLEVARGMGHIWLL.TIIVMR.F.G.YLYLFILFV.G.IND
INGS..SAFIEISNIYSREECHID.WSTINEIFGRRTCITICLNNLN.KV.LNDVCNM.Y.TFIKHNQPKRY.VFTAKQGHNG
NLKNDIILC.TLTLVSDSTL.IKPLSLLIPFIVLIDPSRYTLFCDISEVSVNVPDRSFSSADRCSEYFDDCWIRCSLLID
RSENSDLLFRFYRLLISTFHCF.TIIYS.NYFPILIDLVLRF.FVRLIVGWIIYIRVNLLICVFKMYVDLNCDWILE.HNKS
VPPFF.G.FSDCLLYIS.NCRFD.I.LA.LR..STTIFVVEIGLTPIAAFYYDCFLVKMIFIMVIYVYCFY.FNI.LFFFLLQI
C.PEIIYHGAS.QNLQWCAEPISYLQSLQSTQISLIGFSEDAAASTSLSDFVVVGIEEWDVNWNL.ASSS.GHVFCFHGVH
ASRCKQPSSNCS.VLWSFWNRPYSR.QVYLPQVLHVWRSR.R.NSYHRSFGR.RCYQHW.GYASYGCQNP.GR.YLDH.WCW.
RWTPCS.GSSRFR.RCNWLFPFDYGCWCCLRFR.HFHW.RFSA.ASNGSCVEPTSRNGCAGEV.RR.SSSSYLAWTKDSNANHL
QGTYGFRSSEVRCSACWSQHPRYHHCYRANHDS.PH.KDASRFWC.PYR.D.C.RCAYHPS.RSW.AHRSSD.CSR.SILYCF
PIGCCCLACSRFRRHHP.RFDEPNPYWSHLDSAGNGCRHRSQDPTSCWWRRRG.LACSFYFEGCYCSRRCPSFYDRRVNSRC
F.SCIR.RCYRYERFGRTPC.GKRPSFCCKRKSQAQRC.LR.R.DFSRRAWSS.R.GSR.RFWSSCRRPPSPRYELPRYGRS
R.F.KPCYC..CYDDRY.LPRVHGFDDGWSWS.DRTLRLH.GCLMSSRIRARYILRS.DPLARAFVRIIGFDNVRQVCISFIAHT
PESY.V.VLWHWENCFSTICACACNLLCFLFGFRYRTVKWKWMEKS..MIWSECFPSN.YYLFLLFVVC.I.NYKRYANILF
.VKMCQIVASNDRS.YEE.NTCSCTIMLIH.ATNIFSDLEKLQMLNTSMSSCVLDIYELFSM.FSRILVRF.SLLYNYSYTH
GFVYENIF.CIL.IAN.LTTCINRPAATRSRGPTRGPIWPGPAAAVRSTLEDPRQKHILIRI.EKMM..YIYILISILMYFL
KIYYILIN.NIFISTILHIFYQFSCVFWHI..LFFNNQSLFLHGTWCNHHMKPLHLTMWIVF.SRPPFAAY.LIW.QSVLI
SYLSFLHN.SW.SRMPQH.LVSWIIRKSQGTKEKDTQ.EYPLHSNV.VHKIQTKTQSHTVDITYPLAQDQDRRVKKKLDPKS
HAQQHVLTKVSIQPKFTNSTHHEPTHLFLTLQPTQTRILFRHLIFVYFNTRQTACHPVAKCPCMLTRPMTINICNLGPGFHH
QEPGTELEPRGKLNACISGYFMWGSYLLLRKNSRN.IMHLYILFHF.IS.RLRCRHLASSYNISTS.LY..TIGCNAAVEDD
HNTSCWLLFVLWS.FOD.TFYRRSDLRNLSLWQQLLLHHFSLLLHPRRTLVEQAANLVGLQTLDE.NPNRLLLVA.RQRRKP
LRKLMEPQLLHLKASSMMMIYLRLLPELLSTSPAGLNYQCLKMASCPGNLHGSAAKVDHHAH.SFL.M.QLWFH.FSKGLAPL
PSSHQKTQIWIWLVQLGLQAHHQVCCLLHQSPARVK.QGKMMK.QLLPP..ISKVASPVERLVKS.P.NKK.PARNVVIVLYC
GITANCLSIERSADIVATS.VTPKPSGNLEMKKDI.MHNSISGVSPQOI..TPHEISRDACTI.FTPRTRNKRVPHI.CNS
SVMFHSFAYFGLVFLVFYGLMSHMGFLY.RAYVTLPLYCIL.DIIIIFLPFEDTLFLYLKEFIVYFVWMFNGWIYSIHIYYIK
VIMLLRNVRTYKYKDRIPNFMQINLQPTQV.H.INVIG..IHVSLFNLLNLYIIRANSADIHTGGRGSHIYIAIAARQIVK
FLI.APIWT.M.TRRNKDFRIRIICLLLSPINTDRNLSFYQNVLSFYNNADIYFELKKNW.LLFLFLHIDHHTHC.SM.I
SRT.SHLQL
```

Sliding 8 amino acid window search  
Database searched = AD\_2010  
Query = frame2  
  
Start time: Wed Feb 3 22:29:19 GMT 2010 Finish time: Wed Feb 3 22:29:19 GMT 2010  
  
No 8 amino acid matches exist between frame2 and the AD\_2010 database  
  
# fasta34 frame2.pep /genedata/1/db/AD\_2010 -Q -E 1 -O frame2.pep\_ad.fasta  
FASTA searches a protein or DNA sequence data bank version 3.4t26 July 7, 2006  
Please cite:  
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

frame2, 2416 aa
vs /genedata/1/db/AD_2010 library

    opt      E()
< 20      2      0:=
22      0      0:          one = represents 3 library sequences
24      0      0:
26      0      0:
28      0      0:
30      0      2:*
32      4      8:==*
34     15     21:===== *
36     29     44:===== *
38     84     72:===== *=====
40     70    101:===== *
42     80    123:===== *
44    121    136:===== *
46    145    138:===== *=====
48    133    132:===== *=====
50    109    121:===== *
52    156    106:===== *=====
54     96     91:===== *=====
56    117     76:===== *=====
58     60     62:===== *
60     50     50:===== *
62     40     40:===== *
64     38     32:===== *=====
66     25     25:===== *
68     19     20:===== *
70     20     16:===== *=====
72      7     12:===== *
74     14     10:===== *=====
76     11      7:===== *=====
78      8      6:===== *=====
80      1      4:===== *
82      3      3:===== *
84      1      3:===== *
86      4      2:===== *=====
88      0      2:===== *
90      0      1:===== *
92      0      1:===== *
94      4      1:===== *=====
96      4      1:===== *=====
98      0      0:===== *
100     1      0:===== *=====
102     0      0:===== *
104     0      0:===== *
106     0      0:===== *
108     0      0:===== *
110     0      0:===== *
112     0      0:===== *
114     0      0:===== *
116     0      0:===== *
118     0      0:===== *
>120    0      0:===== *

331323 residues in 1471 sequences
Expectation_n fit: rho(ln(x))= 4.29340.00333; mu= 15.6806 0.173
mean_var=59.023115.580, 0's: 2 Z-trim: 2 B-trim: 0 in 0/43
Lambda= 0.166941
Kolmogorov-Smirnov statistic: 0.0733 (N=28) at 50

```

```

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 44, opt: 32, open/ext: -10/-2, width: 16
!! No sequences with E() < 1.000000

```

```

2416 residues in 1 query sequences
331323 residues in 1471 library sequences
Scomplib [34t26]
start: Wed Feb 3 22:29:18 2010 done: Wed Feb 3 22:29:18 2010
Total Scan time: 0.100 Total Display time: 0.000

```

Function used was FASTA [version 3.4t26 July 7, 2006]

```

# fasta34 frame2.pep /genedata/1/db/TOX_2010 -Q -E 1 -O frame2.pep.tx.fasta
FASTA searches a protein or DNA sequence data bank version 3.4t26 July 7, 2006
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

```

frame2, 2416 aa
vs /genedata/1/db/TOX_2010 library

```

```

    opt      E()
< 20      60      0:=====
22      0      0:          one = represents 14 library sequences
24      0      0:
26      0      0:
28      5      2:*
30     10     12:*
32     27     45:===== *
34    109    122:===== *
36    405    250:===== *=====
38    520    414:===== *=====
40    469    577:===== *
42    653    706:===== *
44    724    778:===== *
46    724    793:===== *
48    819    759:===== *=====
50    636    693:===== *
52    534    609:===== *
54    470    520:===== *
56    379    435:===== *
58    473    357:===== *=====
60    277    289:===== *
62    225    232:===== *
64    283    184:===== *=====
66    153    146:===== *
68    112    115:===== *
70    124     90:===== *=====
72     35     70:===== *
74     26     55:===== *
76     22     43:===== *
78     36     33:===== *
80     26     26:===== *
82     35     20:===== *
84     28     16:===== *
86     13     12:===== *
88     11     9:===== *
90      5      7:===== *
92      2      6:===== *
94      1      4:===== *
96      3      3:===== *

```

```

 98      2      3:*      :==*
100      1      2:*      :=*
102      0      2:*      : *
104      1      1:*      :*
106      1      1:*      :*
108      1      1:*      :*
110      0      1:*      :*
112      0      0:      *
114      0      0:      *
116      1      0:=      *==
118      0      0:      *
>120     2      0:=      *==
2069351 residues in 8448 sequences
  Expectation_n fit: rho(ln(x))= 1.38850.000647; mu= 33.0266 0.033
  mean_var=62.546714.453, 0's: 60 Z-trim: 62 B-trim: 206 in 1/61
  Lambda= 0.162171
  Kolmogorov-Smirnov statistic: 0.0276 (N=29) at 38

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
  join: 44, opt: 32, open/ext: -10/-2, width: 16
The best scores are:
gi|62693|emb|CAA31123.1| precrotoxin B (AA -16 to ( 138) 109 35.4 0.064
gi|451318|gb|AAC59674.1| prebasic subunit of Mojav ( 138) 109 35.4 0.064
gi|62697|emb|CAA34227.1| crotoxin CB2 [Crotalus du ( 138) 93 31.6 0.86

>>gi|62693|emb|CAA31123.1| precrotoxin B (AA -16 to 122) (138 aa)
  initn: 60 init1: 60 opt: 109 Z-score: 137.4 bits: 35.4 E(): 0.064
  Smith-Waterman score: 110; 28.829% identity (42.342% similar) in 111 aa overlap
  (1051-1150:33-137)

      1030      1040      1050      1060      1070
frame2 YWSHLD SAGNGCRHRS DQPTSCW WRRRG.LAC SFFYFEGCYCSR----RPC SFYDRRVSN
      : : : : : : : : : : : : : : : : : : : : : : : :
gi|626 ALWIVAVLLVGV EGHLLQFNKMIKFETRKN AIPFYAFYGCYCGWGG RPKDATDRCCFV
      10      20      30      40      50      60

      1080      1090      1100      1110      1120
frame2 SRCCSCIR.RC-----YRYE-RFGRTPC.GKRPSFCCRKR SQAQRC.LR.R.DFSRRAW
      : : : : : : : : : : : : : : : : : : : : : : : :
gi|626 HDCCYGKLA KCNTKWDIYPYSLKSGYITC-GK-GTWC---EEQICECD-RVAAECLRRSL
      70      80      90      100      110

      1130      1140      1150      1160      1170      1180
frame2 SS.R.GSR.RFWSSCRYPPRSPYRYELPRYGS RF.KPCYC..CYYDRY.LPRVHGF DGWS
      : . . : : : : : : : : : : : : : : : : : : : :
gi|626 STYKYGYMFY PDSRCRGPSETC
      120      130

>>gi|451318|gb|AAC59674.1| prebasic subunit of Mojave To (138 aa)
  initn: 60 init1: 60 opt: 109 Z-score: 137.4 bits: 35.4 E(): 0.064
  Smith-Waterman score: 110; 28.829% identity (42.342% similar) in 111 aa overlap
  (1051-1150:33-137)

      1030      1040      1050      1060      1070
frame2 YWSHLD SAGNGCRHRS DQPTSCW WRRRG.LAC SFFYFEGCYCSR----RPC SFYDRRVSN
      : : : : : : : : : : : : : : : : : : : : : : : :
gi|451 ALWIVAVLLVGV EGHLLQFNKMIKFETRKN AIPFYAFYGCYCGWGG RPKDATDRCCFV
      10      20      30      40      50      60

      1080      1090      1100      1110      1120
frame2 SRCCSCIR.RC-----YRYE-RFGRTPC.GKRPSFCCRKR SQAQRC.LR.R.DFSRRAW
      : : : : : : : : : : : : : : : : : : : : : : : :

```

```

      : : : : : : : : : : : : : : : : : : : : : : : :
gi|451 HDCCYGKLA KCNTKWDIYPYSLKSGYITC-GK-GTWC---EEQICECD-RVAAECLRRSL
      70      80      90      100      110

      1130      1140      1150      1160      1170      1180
frame2 SS.R.GSR.RFWSSCRYPPRSPYRYELPRYGS RF.KPCYC..CYYDRY.LPRVHGF DGWS
      : . . : : : : : : : : : : : : : : : : : : : :
gi|451 STYKYGYMFY PDSRCRGPSETC
      120      130

>>gi|62697|emb|CAA34227.1| crotoxin CB2 [Crotalus duriss (138 aa)
  initn: 61 init1: 61 opt: 93 Z-score: 117.2 bits: 31.6 E(): 0.86
  Smith-Waterman score: 114; 29.204% identity (45.133% similar) in 113 aa overlap
  (1051-1150:33-137)

      1030      1040      1050      1060      1070
frame2 YWSHLD SAGNGCRHRS DQPTSCW WRRRG.LAC SFFYFEGCYCSR----RPC SFYDRRVSN
      : : : : : : : : : : : : : : : : : : : : : : : :
gi|626 ALWIVAVLLVGV EGHLLQFNKMIKFETRKN AIPFYAFYGCYCGWGG QGRPKDATDRCCFV
      10      20      30      40      50      60

      1080      1090      1100      1110      1120
frame2 SRCCSCIR.RC-----YRYE-RFGRTPC.GKRPSFCCRKR SQAQRC.LR.R.DFSRRAW
      : : : : : : : : : : : : : : : : : : : : : : : :
gi|626 HDCCYGKLA KCNTKWDIYRYSLKSGYITC-GK-GTWC---KEQICECD-RVAAECLRRSL
      70      80      90      100      110

      1130      1140      1150      1160      1170      1180
frame2 SS.R.GSR.RFW--SSCRYPPRSPYRYELPRYGS RF.KPCYC..CYYDRY.LPRVHGF DG
      : . . . : : : : : : : : : : : : : : : : : : :
gi|626 STYK--NEYMFY PDSRCREPSETC
      120      130

2416 residues in 1 query sequences
2069351 residues in 8448 library sequences
  Scomplib [34t26]
  start: Wed Feb 3 22:29:19 2010 done: Wed Feb 3 22:29:20 2010
  Total Scan time: 0.550 Total Display time: 0.010

Function used was FASTA [version 3.4t26 July 7, 2006]

# fasta34 frame2.pep /genedata/1/db/PRT_2010 -Q -E 1 -O frame2.pep_prt.fasta
FASTA searches a protein or DNA sequence data bank version 3.4t26 July 7, 2006
Please cite:
  W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

frame2, 2416 aa
vs /genedata/1/db/PRT_2010 library

      opt      E()
< 20 279246 0:=====
  22 219 0:= one = represents 26125 library sequences
  24 538 17:*
  26 1014 374:*
  28 3687 4039:*
  30 19121 24535:*
  32 77200 94867:====*
  34 239312 257268:=====*

```



```

36 516072 528368:=====*
38 820022 873195:===== *
40 1132405 1218030:===== *
42 1381196 1488894:===== *
44 1504224 1642389:===== *
46 1567465 1672814:=====*
48 1529332 1601526:=====*
50 1414612 1461399:=====*
52 1259520 1284814:=====*
54 1086317 1097456:=====*
56 937210 916712:=====*
58 773231 752603:=====*=
60 635304 609653:=====*=
62 532827 488761:=====*=
64 433924 388709:=====*=
66 358597 307224:=====*=
68 284935 241656:=====*=
70 224235 189375:=====*=
72 174846 147979:=====*=
74 139694 115374:=====*=
76 101782 89798:=====*
78 82762 69799:=====*=
80 65028 54198:=====*
82 51057 41460:=====*
84 37404 32842:=====*
86 32105 25411:=====*
88 24697 19662:=====*
90 19611 15213:=====*
92 14745 11771:=====*
94 11656 9108:=====*
96 9846 7047:=====*
98 8161 5453:=====*
100 5611 4219:=====*
102 4098 3264:=====*
104 3142 2526:=====*
106 2614 1954:=====*
108 2407 1512:=====*
110 1628 1170:=====*
112 1317 905:=====*
114 985 701:=====*
116 724 542:=====*
118 536 419:=====*
>120 3004 324:=====*
4761287459 residues in 17815538 sequences
statistics sampled from 60000 to 17808441 sequences
Expectation_n fit: rho(ln(x))= 4.51700.000185; mu= 17.4384 0.010
mean_var=65.381512.987, 0's: 968 Z-trim: 973 B-trim: 0 in 0/64
Lambda= 0.158616
Kolmogorov-Smirnov statistic: 0.0398 (N=29) at 54

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 44, opt: 32, open/ext: -10/-2, width: 16
The best scores are:
gi|136192920|gb|EBM50575.1| hypothetical protein G ( 396) 284 75.5 3.2e-10
gi|143377609|gb|EDE65555.1| hypothetical protein G ( 418) 269 72.1 3.5e-09
gi|15809660|gb|AAL07277.1| ribulose-1,5-bisphospha ( 93) 197 55.0 0.00011
gi|167268|gb|AAA33035.1| ribulose-1,5-bisphosphate ( 182) 200 56.0 0.00011
gi|116109615|gb|ABJ74187.1| chloroplast ribulose-1 ( 90) 195 54.5 0.00015
gi|169156|gb|AAA33686.1| ribulose 1,5-bisphosphate ( 156) 197 55.2 0.00016
gi|149938962|gb|ABR45726.1| RBCS1 [Actinidia chine ( 176) 196 55.0 0.0002
gi|12019640|gb|AAA33685.2| ribulose 1,5 bisphospha ( 139) 192 54.0 0.00032
gi|167270|gb|AAA33036.1| ribulose 1,5-bisphosphate ( 180) 193 54.4 0.00033

```

```

gi|304442|gb|AAA03693.1| ribulose-1,5-bisphosphate ( 182) 193 54.4 0.00034
gi|167272|gb|AAA33037.1| ribulose 1,5-bisphosphate ( 183) 193 54.4 0.00034
gi|436487|gb|AAA03695.1| ribulose-1,5-bisphosphate ( 183) 193 54.4 0.00034
gi|229291525|gb|EEN62196.1| hypothetical protein B ( 560) 198 56.0 0.00034
gi|20857|emb|CAA27865.1| ribulose 1,5-bisphosphate ( 180) 192 54.1 0.00039
gi|20855|emb|CAA27864.1| ribulose bisphosphate car ( 180) 192 54.1 0.00039
gi|20859|emb|CAA25390.1| ribulose bisphosphate car ( 180) 192 54.1 0.00039
gi|436489|gb|AAA03697.1| ribulose-1,5-bisphosphate ( 182) 192 54.1 0.00039
gi|436488|gb|AAA03696.1| ribulose-1,5-bisphosphate ( 183) 192 54.1 0.00039
gi|167274|gb|AAA33038.1| ribulose 1,5-bisphosphate ( 183) 192 54.1 0.00039
gi|304447|gb|AAA03698.1| ribulose-1,5-bisphosphate ( 186) 192 54.2 0.0004
gi|7545285|gb|AAA33683.2| ribulose 1,5-bisphosphat ( 27) 182 51.1 0.00049
gi|170320|gb|AAA34111.1| ribulose-1,5-bisphosphate ( 86) 186 52.5 0.0006
gi|304444|gb|AAA03694.1| ribulose-1,5-bisphosphate ( 180) 189 53.5 0.00063
gi|170504|gb|AAA34192.1| ribulose-1,5-bisphosphate ( 180) 188 53.2 0.00074
gi|170498|gb|AAA34189.1| ribulose-1,5-bisphosphate ( 180) 188 53.2 0.00074
gi|19334|emb|CAA29402.1| ribulose 1,5-bisphosphate ( 180) 188 53.2 0.00074
gi|19336|emb|CAA29403.1| ribulose 1,5-bisphosphate ( 180) 188 53.2 0.00074
gi|218318|dbj|BAA01888.1| ribulose 1,5-bisphosphat ( 180) 188 53.2 0.00074
gi|4456641|emb|CAA29401.2| ribulose 1,5-bisphospha ( 180) 188 53.2 0.00074
gi|170500|gb|AAA34190.1| ribulose-1,5-bisphosphate ( 180) 188 53.2 0.00074
gi|19338|emb|CAA29404.1| ribulose 1,5-bisphosphate ( 180) 188 53.2 0.00074
gi|162946541|gb|ABY21255.1| ribulose-1,5-bisphosph ( 181) 188 53.2 0.00074
gi|162946537|gb|ABY21253.1| ribulose-1,5-bisphosph ( 181) 188 53.2 0.00074
gi|19706|emb|CAA31994.1| ribulose bisphosphate car ( 180) 187 53.0 0.00086
gi|170318|gb|AAA34110.1| ribulose bisphosphate car ( 180) 187 53.0 0.00086
gi|4689388|gb|AAD27881.1|AF139469_1 ribulose-1,5-b ( 181) 187 53.0 0.00087
gi|169557|gb|AAA33838.1| ribulose bisphosphate car ( 181) 187 53.0 0.00087
gi|162946539|gb|ABY21254.1| ribulose-1,5-bisphosph ( 181) 187 53.0 0.00087
gi|21563|emb|CAA49413.1| ribulose bisphosphate car ( 181) 187 53.0 0.00087
gi|21571|emb|CAA49417.1| ribulose bisphosphate car ( 181) 187 53.0 0.00087
gi|16224234|gb|AAL15646.1|AF411547_1 ribulose-1,5- ( 100) 184 52.1 0.00091
gi|21567|emb|CAA49415.1| ribulose bisphosphate car ( 180) 186 52.8 0.001
gi|20024|emb|CAA26208.1| small subunit ribulose 1, ( 180) 186 52.8 0.001
gi|21569|emb|CAA49416.1| ribulose bisphosphate car ( 180) 186 52.8 0.001
gi|21565|emb|CAA49414.1| ribulose bisphosphate car ( 180) 186 52.8 0.001
gi|19760|emb|CAA25862.1| unnamed protein product [ ( 180) 186 52.8 0.001
gi|295814|emb|CAA37516.1| NySS41 [Nicotiana sylves ( 181) 186 52.8 0.001
gi|20958|emb|CAA31774.1| ribulose bisphosphate car ( 171) 185 52.5 0.0011
gi|1079740|gb|AAA82071.1| ribulose 1,5-bisphosphat ( 178) 185 52.5 0.0012
gi|13241105|gb|AAK16229.1|AF044397_1 ribulose-1,5- ( 133) 183 52.0 0.0013
gi|18920396|gb|AAL82195.1|AF396697_1 s/s2 [Nicotia ( 113) 182 51.7 0.0014
gi|3044116|gb|AAC13293.1| ribulose-1,5-bisphosphat ( 180) 184 52.3 0.0014
gi|3928152|emb|CAA10290.1| ribulose 1,5-bisphospha ( 181) 184 52.3 0.0014
gi|170496|gb|AAA34188.1| ribulose-1,5-bisphosphate ( 181) 184 52.3 0.0014
gi|170502|gb|AAA34191.1| ribulose-1,5-bisphosphate ( 181) 184 52.3 0.0014
gi|295814|emb|CAA29400.1| ribulose 1,5-bisphosphat ( 181) 184 52.3 0.0014
gi|450505|emb|CAA38026.1| ribulose bisphosphate ca ( 182) 184 52.3 0.0014
gi|20218815|emb|CAC84492.1| putative ribulose bisp ( 148) 183 52.0 0.0014
gi|144125506|gb|EDJ04873.1| hypothetical protein G ( 626) 189 54.0 0.0015
gi|149938960|gb|ABR45725.1| RBCS1 [Actinidia delic ( 176) 183 52.1 0.0016
gi|169148|gb|AAA33684.1| ribulose-1,5-bisphosphate ( 136) 181 51.5 0.0018
gi|30013663|gb|AAP03874.1| putative ribulose bisp ( 180) 182 51.9 0.0019
gi|809069|emb|CAA42617.1| ribulose bisphosphate ca ( 135) 180 51.3 0.0021
gi|19464|emb|CAA38346.1| ribulose bisphosphate car ( 171) 181 51.6 0.0022
gi|1079736|gb|AAA82069.1| ribulose 1,5-bisphosphat ( 178) 181 51.6 0.0022
gi|20493|emb|CAA27445.1| ribulose 1,5-bisphosphate ( 180) 181 51.6 0.0022
gi|40287530|gb|AAR83879.1| Cristal-Glass1 protein ( 180) 181 51.6 0.0022
gi|217071562|gb|ACU84141.1| unknown [Medicago trun ( 180) 181 51.6 0.0022
gi|6272550|gb|AAF06100.1| ribulose 1,5-bisphosphat ( 184) 181 51.6 0.0023
gi|6272548|gb|AAF06098.1| ribulose 1,5-bisphosphat ( 184) 181 51.6 0.0023
gi|6272551|gb|AAF06101.1| ribulose 1,5-bisphosphat ( 184) 181 51.6 0.0023

```

gi 19462 emb CAA38345.1	ribulose bisphosphate car	( 80)	177	50.4	0.0024
gi 170058 gb AAA34008.1	ribulose 1,5-bisphosphate	( 178)	180	51.4	0.0026
gi 28139169 gb AAO25119.1	ribulose-1,5-bisphospha	( 179)	180	51.4	0.0026
gi 21050 emb CAA42618.1	ribulose bisphosphate car	( 180)	180	51.4	0.0026
gi 556418 gb AAA69018.1	ribulose 1,5-bisphosphate	( 180)	180	51.4	0.0026
gi 21053 emb CAA40339.1	small subunit of ribulose	( 180)	180	51.4	0.0026
gi 125380592 gb ABN41481.1	putative chloroplast r	( 93)	177	50.4	0.0026
gi 167542 gb AAA33131.1	ribulose bisphosphate car	( 189)	180	51.4	0.0027
gi 13926229 gb AAK49590.1	AF372874_1 F1019.10/F101	( 125)	178	50.8	0.0028
gi 13241101 gb AAK16227.1	AF044395_1 ribulose-1,5-	( 131)	178	50.8	0.0029
gi 224284271 gb ACN39871.1	unknown [Picea sitchen	( 136)	178	50.8	0.0029
gi 224284074 gb ACN39774.1	unknown [Picea sitchen	( 136)	178	50.8	0.0029
gi 11762156 gb AAG40356.1	AF325004_1 Atlg67090 [Ar	( 180)	179	51.2	0.0031
gi 19581 emb CAA46475.1	ribulose bisphosphate car	( 183)	179	51.2	0.0031
gi 218021 dbj BAA00450.1	RuBisCO small subunit [P	( 183)	179	51.2	0.0031
gi 116791718 gb ABK26083.1	unknown [Picea sitchen	( 159)	178	50.9	0.0033
gi 1345574 emb CAA34458.1	unnamed protein product	( 82)	175	49.9	0.0033
gi 119720808 gb ABL97974.1	ribulose bisphosphate	( 162)	178	50.9	0.0033
gi 148908844 gb ABR17527.1	unknown [Picea sitchen	( 169)	178	50.9	0.0034
gi 260024743 gb ACX04274.1	Sequence 27038 from pa	( 170)	178	50.9	0.0034
gi 260024742 gb ACX04273.1	Sequence 27037 from pa	( 176)	178	50.9	0.0035
gi 967221 gb AAB67846.1	ribulose-1,5-bisphosphate	( 178)	178	50.9	0.0036
gi 260024741 gb ACX04272.1	Sequence 27036 from pa	( 180)	178	50.9	0.0036
gi 110741475 dbj BAE98695.1	ribulose bisphosphate	( 180)	178	50.9	0.0036
gi 25083492 gb AAN72087.1	ribulose bisphosphate c	( 180)	178	50.9	0.0036
gi 110741865 dbj BAE98874.1	ribulose bisphosphate	( 180)	178	50.9	0.0036
gi 16604426 gb AAL24219.1	Atlg67090/F1019.10 [Ara	( 180)	178	50.9	0.0036
gi 20453287 gb AAM19882.1	Atlg67090/F1019.10 [Ara	( 180)	178	50.9	0.0036
gi 11762170 gb AAG40363.1	AF325011_1 000C10C11 [Ar	( 180)	178	50.9	0.0036
gi 17065270 gb AAL32789.1	ribulose bisphosphate c	( 180)	178	50.9	0.0036
gi 110741446 dbj BAE98685.1	ribulose bisphosphate	( 180)	178	50.9	0.0036
gi 110741032 dbj BAE98610.1	ribulose bisphosphate	( 180)	178	50.9	0.0036
gi 23198308 gb AAN15681.1	ribulose bisphosphate c	( 180)	178	50.9	0.0036
gi 17065072 gb AAL32690.1	ribulose bisphosphate c	( 180)	178	50.9	0.0036
gi 4204274 gb AAD10655.1	ribulose bisphosphate ca	( 180)	178	50.9	0.0036
gi 20260080 gb AAM13387.1	ribulose bisphosphate c	( 180)	178	50.9	0.0036
gi 15809842 gb AAL06849.1	Atlg67090/F1019.10 [Ara	( 180)	178	50.9	0.0036
gi 16649141 gb AAL24422.1	ribulose bisphosphate c	( 180)	178	50.9	0.0036
gi 17473628 gb AAL38277.1	ribulose bisphosphate c	( 180)	178	50.9	0.0036
gi 20260064 gb AAM13379.1	ribulose bisphosphate c	( 180)	178	50.9	0.0036
gi 15294200 gb AAK95277.1	AF410291_1 F1019.10/F101	( 180)	178	50.9	0.0036
gi 15451002 gb AAK96772.1	ribulose bisphosphate c	( 180)	178	50.9	0.0036
gi 17850 emb CAA43410.1	ribulose bisphosphate car	( 181)	178	50.9	0.0036
gi 119720784 gb ABL97962.1	ribulose-1,5-bisphosph	( 181)	178	50.9	0.0036
gi 217075090 gb ACJ85905.1	unknown [Medicago trun	( 181)	178	50.9	0.0036
gi 17855 emb CAA30290.1	rubisco ssu precursor [Br	( 181)	178	50.9	0.0036
gi 79013990 gb ABB51649.1	chloroplast ribulose-1,	( 181)	178	50.9	0.0036
gi 18413570 emb CAD21856.1	putative ribulose 1,5	( 181)	178	50.9	0.0036
gi 17852 emb CAA39402.1	ribulose bisphosphate car	( 181)	178	50.9	0.0036
gi 224284697 gb ACN40080.1	unknown [Picea sitchen	( 183)	178	50.9	0.0036
gi 295822 emb CAA34161.1	ribulose-1,5-carboxylase	( 189)	178	51.0	0.0037
gi 116782078 gb ABK22359.1	unknown [Picea sitchen	( 193)	178	51.0	0.0038
gi 116780863 gb ABK21852.1	unknown [Picea sitchen	( 193)	178	51.0	0.0038
gi 116793637 gb ABK26822.1	unknown [Picea sitchen	( 193)	178	51.0	0.0038
gi 148907673 gb ABR16965.1	unknown [Picea sitchen	( 193)	178	51.0	0.0038
gi 224285721 gb ACN40576.1	unknown [Picea sitchen	( 193)	178	51.0	0.0038
gi 116780912 gb ABK21876.1	unknown [Picea sitchen	( 193)	178	51.0	0.0038
gi 116784603 gb ABK23405.1	unknown [Picea sitchen	( 193)	178	51.0	0.0038
gi 116782420 gb ABK22499.1	unknown [Picea sitchen	( 193)	178	51.0	0.0038
gi 116794195 gb ABK27041.1	unknown [Picea sitchen	( 193)	178	51.0	0.0038
gi 116793621 gb ABK26814.1	unknown [Picea sitchen	( 193)	178	51.0	0.0038
gi 224285010 gb ACN40234.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038

gi 116779842 gb ABK21447.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 116785645 gb ABK23805.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 116782047 gb ABK22347.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 224285123 gb ACN40289.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 224285683 gb ACN40557.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 116779377 gb ABK21259.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 116790039 gb ABK25479.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 148905732 gb ABR16030.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 116791702 gb ABK26077.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 116782035 gb ABK22342.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 224284134 gb ACN39804.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 116780594 gb ABK21733.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 116779440 gb ABK21283.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 224284171 gb ACN39822.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 116784457 gb ABK23349.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 224285619 gb ACN40528.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 224285468 gb ACN40456.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 116783228 gb ABK22845.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 224284342 gb ACN39906.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 224284998 gb ACN40228.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 224284941 gb ACN40200.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 116793644 gb ABK26825.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 116782020 gb ABK22337.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 224284796 gb ABK240128.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 224285656 gb ACN40544.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 224285893 gb ACN40660.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 224284301 gb ACN39886.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 224284878 gb ACN40169.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 116786299 gb ABK24056.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 224284814 gb ACN40137.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 224284389 gb ACN39929.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 116780460 gb ABK21690.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 224285107 gb ACN40281.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 116781990 gb ABK22328.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 116779401 gb ABK21267.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 116785630 gb ABK23799.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 116786044 gb ABK23949.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 224285109 gb ACN40282.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 224284032 gb ACN39754.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 116792861 gb ABK26530.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 116785985 gb ABK22931.1	unknown [Picea sitchen	( 194)	178	51.0	0.0038
gi 224284092 gb ACN39783.1	unknown [Picea sitchen	( 195)	178	51.0	0.0038
gi 116780861 gb ABK21851.1	unknown [Picea sitchen	( 195)	178	51.0	0.0038
gi 224284864 gb ACN40162.1	unknown [Picea sitchen	( 195)	178	51.0	0.0038
gi 148905962 gb ABR16142.1	unknown [Picea sitchen	( 195)	178	51.0	0.0038
gi 116779577 gb ABK21347.1	unknown [Picea sitchen	( 195)	178	51.0	0.0038
gi 224285028 gb ACN40243.1	unknown [Picea sitchen	( 195)	178	51.0	0.0038
gi 6409426 gb AAF07985.1	AF200303_1 ribulose-1,5-b	( 168)	177	50.7	0.004
gi 143306418 gb EDE23096.1	hypothetical protein G	( 510)	182	52.3	0.004
gi 1079738 gb AAA82070.1	ribulose 1,5-bisphosphat	( 178)	177	50.7	0.0042
gi 20491 emb CAA27444.1	ribulose 1,5-bisphosphate	( 180)	177	50.7	0.0042
gi 295792 emb CAA31948.1	ribulose bisphosphate ca	( 182)	177	50.7	0.0042
gi 124269258 gb ABS72189.1	ribulose-1,5-bisphosph	( 183)	177	50.7	0.0043
gi 260037942 gb ACX07559.1	Sequence 31503 from pa	( 162)	176	50.4	0.0046
gi 260037941 gb ACX07558.1	Sequence 31502 from pa	( 177)	176	50.5	0.0049
gi 403160 gb AAA33866.1	ribulose 1,5-bisphosphate	( 179)	176	50.5	0.0049
gi 16649139 gb AAL24421.1	ribulose bisphosphate c	( 181)	176	50.5	0.005
gi 9758821 dbj BAB09355.1	ribulose bisphosphate c	( 181)	176	50.5	0.005
gi 23505787 gb AAN28753.1	At5g38430/F1019.10 [Ara	( 181)	176	50.5	0.005
gi 16194 emb CAA32701.1	ribulose bisphosphate car	( 181)	176	50.5	0.005
gi 25083689 gb AAN72105.1	ribulose bisphosphate c	( 181)	176	50.5	0.005
gi 16193 emb CAA32700.1	ribulose bisphosphate car	( 181)	176	50.5	0.005

gi 9758820 dbj BAB09354.1	ribulose bisphosphate c	( 181)	176	50.5	0.005
gi 15294246 gb AAK95300.1	AF101314_1 F1019.10/F101	( 181)	176	50.5	0.005
gi 27311897 gb AAO00914.1	ribulose bisphosphate c	( 181)	176	50.5	0.005
gi 28058980 gb AAO29974.1	ribulose bisphosphate c	( 181)	176	50.5	0.005
gi 17064934 gb AAL32621.1	ribulose bisphosphate c	( 181)	176	50.5	0.005
gi 13430424 gb AAK25834.1	AF360124_1 putative ribu	( 181)	176	50.5	0.005
gi 17064764 gb AAL32536.1	ribulose bisphosphate c	( 181)	176	50.5	0.005
gi 20466117 gb AAM19980.1	At5g38410/F1019.10 [Ara	( 181)	176	50.5	0.005
gi 17978793 gb AAL47390.1	ribulose bisphosphate c	( 181)	176	50.5	0.005
gi 18087561 gb AAL58912.1	AF462822_1 At5g38410/F10	( 181)	176	50.5	0.005
gi 23397160 gb AAN31863.1	putative ribulose bisph	( 181)	176	50.5	0.005
gi 260037940 gb ACX07557.1	Sequence 31501 from pa	( 181)	176	50.5	0.005
gi 20259880 gb AAM13287.1	ribulose bisphosphate c	( 181)	176	50.5	0.005
gi 17064722 gb AAL32515.1	ribulose bisphosphate c	( 181)	176	50.5	0.005
gi 15293183 gb AAK93702.1	putative RuBisCO small	( 181)	176	50.5	0.005
gi 16195 emb CAA32702.1	ribulose bisphosphate car	( 181)	176	50.5	0.005
gi 15294184 gb AAK95269.1	AF410283_1 F1019.10/F101	( 181)	176	50.5	0.005
gi 9758819 dbj BAB09353.1	ribulose bisphosphate c	( 181)	176	50.5	0.005
gi 15450944 gb AAK96743.1	ribulose bisphosphate c	( 181)	176	50.5	0.005
gi 13241103 gb AAK16228.1	AF044396_1 ribulose-1,5-	( 131)	174	49.9	0.0054
gi 155351442 gb ABU07047.1	Sequence 194516 from p	( 135)	174	49.9	0.0055
gi 255630492 gb ACU15604.1	unknown [Glycine max]	( 178)	175	50.2	0.0057
gi 10946377 gb AAG24883.1	AF303940_1 ribulose-1,5-	( 178)	175	50.2	0.0057
gi 10946375 gb AAG24882.1	AF303939_1 ribulose-1,5-	( 178)	175	50.2	0.0057
gi 18742 emb CAA23736.1	rubpcase [Glycine max]	( 178)	175	50.2	0.0057
gi 25562588 gb ACU13285.1	unknown [Glycine max]	( 178)	175	50.2	0.0057
gi 1055368 gb AAA81328.1	ribulose-1,5-bisphosphat	( 178)	175	50.2	0.0057
gi 2258433 gb AAB63287.1	ribulose-1,5-bisphosphat	( 180)	175	50.3	0.0058
gi 238801314 gb ACR56377.1	chloroplast ribulose-1	( 180)	175	50.3	0.0058
gi 146724107 gb ABQ42552.1	chloroplast ribulose-1	( 180)	175	50.3	0.0058
gi 76574240 gb ABA46901.1	chloroplast ribulose-1,	( 181)	175	50.3	0.0058
gi 77157637 dbj BAE64384.1	ribulose-1,5-bisphosph	( 183)	175	50.3	0.0059
gi 19466 emb CAA38347.1	ribulose bisphosphate car	( 61)	170	48.7	0.0059
gi 967223 gb AAB67847.1	ribulose-1,5-bisphosphate	( 173)	174	50.0	0.0066
gi 967227 gb AAB67849.1	ribulose-1,5-bisphosphate	( 173)	174	50.0	0.0066
gi 967231 gb AAB67851.1	ribulose-1,5-bisphosphate	( 173)	174	50.0	0.0066
gi 967225 gb AAB67848.1	ribulose-1,5-bisphosphate	( 178)	174	50.0	0.0067
gi 225905973 gb ACO35888.1	ribulose-1,5-bisphosph	( 178)	174	50.0	0.0067
gi 3152715 gb AAC17126.1	ribulose 1,5-bisphosphat	( 187)	174	50.0	0.007
gi 76574236 gb ABA46899.1	chloroplast ribulose-1,	( 180)	173	49.8	0.0079
gi 169218 gb AAA33717.1	ribulose 1,5-bisphosphate	( 68)	168	48.3	0.0087
gi 10946379 gb AAG24884.1	AF303941_1 ribulose-1,5-	( 178)	172	49.6	0.0092
gi 270383720 dbj BAI53118.1	ribulose-1,5-bisphosp	( 181)	172	49.6	0.0093
gi 388253 emb CAA37905.1	ribulose bisphosphate ca	( 75)	168	48.3	0.0094
gi 223528159 gb EEF30223.1	Ribulose bisphosphate	( 185)	172	49.6	0.0095
gi 13241109 gb AAK16231.1	AF044399_1 ribulose-1,5-	( 131)	170	49.0	0.01
gi 13241111 gb AAK16232.1	AF044400_1 ribulose-1,5-	( 131)	170	49.0	0.01
gi 13241113 gb AAK16233.1	AF044401_1 ribulose-1,5-	( 131)	170	49.0	0.01
gi 155309951 gb ABT65555.1	Sequence 153025 from p	( 135)	170	49.0	0.01
gi 134592649 gb EBC26556.1	hypothetical protein G	( 267)	173	50.0	0.011
gi 270383724 dbj BAI53120.1	ribulose-1,5-bisphosp	( 179)	171	49.3	0.011
gi 406727 emb CAA53083.1	ribulose-1,5-bisphosphat	( 181)	171	49.3	0.011
gi 136606428 gb EBP24513.1	hypothetical protein G	( 388)	174	50.3	0.012
gi 143358139 gb EDE53810.1	hypothetical protein G	( 648)	176	51.0	0.012
gi 30421141 gb AAP31053.1	ribulose-1,5-bisphospha	( 173)	170	49.1	0.012
gi 30421144 gb AAP31054.1	ribulose-1,5-bisphospha	( 173)	170	49.1	0.012
gi 18466 emb CAA28711.1	unnamed protein product [	( 173)	170	49.1	0.012
gi 967229 gb AAB67850.1	ribulose-1,5-bisphosphate	( 173)	170	49.1	0.012
gi 53748423 emb CAH59404.1	Rubisco SSU [Plantago	( 174)	170	49.1	0.012
gi 169892 gb AAB39037.1	ribulose bisphosphate car	( 177)	170	49.1	0.013
gi 223528742 gb EEF30752.1	Ribulose bisphosphate	( 180)	170	49.1	0.013
gi 155365850 gb ABU21455.1	Sequence 208924 from p	( 194)	170	49.1	0.013

gi 154000871 gb ABS57004.1	ribulose-1,5-bisphosph	( 129)	168	48.5	0.014
gi 4038695 dbj BAA35164.1	ribulose-1,5-bisphospha	( 165)	169	48.8	0.014
gi 65732606 gb AAF17591.1	AF202647_1 ribulose-1,5-b	( 168)	169	48.8	0.014
gi 6409335 gb AAF07945.1	AF192776_1 ribulose-1,5-b	( 168)	169	48.8	0.014
gi 6573202 gb AAF17589.1	AF202645_1 ribulose-1,5-b	( 168)	169	48.8	0.014
gi 6573204 gb AAF17590.1	AF202646_1 ribulose-1,5-b	( 168)	169	48.8	0.014
gi 6409337 gb AAF07946.1	AF192777_1 ribulose-1,5-b	( 168)	169	48.8	0.014
gi 6409341 gb AAF07948.1	AF192779_1 ribulose-1,5-b	( 168)	169	48.8	0.014
gi 3978515 gb AAC8374.1	Rubisco-1,5-bisphosphat	( 168)	169	48.8	0.014
gi 6573208 gb AAF17592.1	AF202648_1 ribulose-1,5-b	( 168)	169	48.8	0.014
gi 6409339 gb AAF07947.1	AF192778_1 ribulose-1,5-b	( 168)	169	48.8	0.014
gi 3790104 gb AAC67588.1	ribulose-1,5-bisphosphat	( 168)	169	48.8	0.014
gi 3901434 gb AAC78644.1	ribulose-1,5-bisphosphat	( 168)	169	48.8	0.014
gi 3901432 gb AAC78643.1	ribulose-1,5-bisphosphat	( 168)	169	48.8	0.014
gi 967219 gb AAB67845.1	ribulose-1,5-bisphosphate	( 173)	169	48.9	0.015
gi 56900734 gb AAW31667.1	ribulose-1,5-bisphospha	( 174)	169	48.9	0.015
gi 53748417 emb CAH59401.1	Rubisco SSU [Plantago	( 177)	169	48.9	0.015
gi 755803 emb CAA29801.1	carboxylase [Raphanus sa	( 181)	169	48.9	0.015
gi 242138577 gb EE24979.1	conserved hypothetical	( 439)	173	50.2	0.015
gi 155325516 gb ABT81120.1	Sequence 168590 from p	( 77)	165	47.6	0.015
gi 155318236 gb ABT73840.1	Sequence 161310 from p	( 121)	167	48.3	0.015
gi 6409343 gb AAF07949.1	AF192780_1 ribulose-1,5-b	( 168)	168	48.6	0.017
gi 755735 emb CAA28737.1	RuBisCO (SSU) [Helianthu	( 178)	168	48.6	0.017
gi 295846 emb CAA36542.1	ribulose bisphosphate ca	( 178)	168	48.6	0.017
gi 18808 emb CAA68490.1	ribulose bisphosphate car	( 178)	168	48.6	0.017
gi 19439 emb CAA24969.1	unnamed protein product [	( 120)	166	48.0	0.018
gi 6409329 gb AAF07942.1	AF192773_1 ribulose-1,5-b	( 168)	167	48.4	0.02
gi 155313763 gb ABT69367.1	Sequence 156837 from p	( 177)	167	48.4	0.02
gi 2665670 gb AAC18406.1	ribulose-1,5-bisphosphat	( 178)	167	48.4	0.02
gi 76574238 gb ABA46900.1	chloroplast ribulose-1,	( 180)	167	48.4	0.021
gi 118486788 gb ABK95229.1	unknown [Populus trich	( 182)	167	48.4	0.021
gi 118487084 gb ABK95372.1	unknown [Populus trich	( 182)	167	48.4	0.021
gi 222848126 gb EEE85673.1	predicted protein [Pop	( 182)	167	48.4	0.021
gi 118486831 gb ABK95250.1	unknown [Populus trich	( 182)	167	48.4	0.021
gi 118488041 gb ABK95841.1	unknown [Populus trich	( 182)	167	48.4	0.021
gi 126228074 gb ABN91614.1	conserved hypothetical	( 741)	173	50.4	0.022
gi 155309946 gb ABT65550.1	Sequence 153020 from p	( 81)	163	47.2	0.022
gi 13241107 gb AAK16230.1	AF044398_1 ribulose-1,5-	( 131)	165	47.8	0.022
gi 155313056 gb ABT68660.1	Sequence 156130 from p	( 133)	165	47.8	0.023
gi 19408 emb CAA35104.1	unnamed protein product [	( 173)	166	48.2	0.023
gi 295818 emb CAA35102.1	ribulose bisphosphate ca	( 177)	166	48.2	0.024
gi 295821 emb CAA35100.1	ribulose bisphosphate ca	( 177)	166	48.2	0.024
gi 295820 emb CAA35099.1	ribulose bisphosphate ca	( 177)	166	48.2	0.024
gi 295819 emb CAA35103.1	ribulose bisphosphate ca	( 177)	166	48.2	0.024
gi 295817 emb CAA35101.1	ribulose bisphosphate ca	( 177)	166	48.2	0.024
gi 118489530 gb ABK96567.1	unknown [Populus trich	( 181)	166	48.2	0.024
gi 118489251 gb ABK96431.1	unknown [Populus trich	( 182)	166	48.2	0.024
gi 20341 emb CAA30393.1	ribulose bisphosphate car	( 172)	165	47.9	0.027
gi 155346813 gb ABU02418.1	Sequence 189887 from p	( 71)	161	46.7	0.027
gi 155347481 gb ABU03086.1	Sequence 190555 from p	( 71)	161	46.7	0.027
gi 116672873 gb ABK15574.1	ribulose-1,5-bisphosph	( 176)	165	48.0	0.028
gi 155345517 gb ABU01122.1	Sequence 188591 from p	( 73)	161	46.7	0.028
gi 155353538 gb ABU09143.1	Sequence 196612 from p	( 73)	161	46.7	0.028
gi 155317126 gb ABT72730.1	Sequence 160200 from p	( 114)	163	47.3	0.028
gi 155354411 gb ABU10016.1	Sequence 197485 from p	( 74)	161	46.7	0.028
gi 155358106 gb ABU13711.1	Sequence 201180 from p	( 144)	164	47.6	0.028
gi 118488741 gb ABK96181.1	unknown [Populus trich	( 181)	165	48.0	0.028
gi 118488169 gb ABK95904.1	unknown [Populus trich	( 181)	165	48.0	0.028
gi 222870738 gb EEF07869.1	predicted protein [Pop	( 181)	165	48.0	0.028
gi 155312601 gb ABT68205.1	Sequence 155675 from p	( 77)	161	46.7	0.029
gi 155345515 gb ABU01120.1	Sequence 188589 from p	( 77)	161	46.7	0.029
gi 155349458 gb ABU05063.1	Sequence 192532 from p	( 82)	161	46.7	0.03

gi 4038725 dbj BAA35179.1	ribulose-1,5-bisphospha	( 164)	164	47.7	0.031
gi 4038717 dbj BAA35175.1	ribulose-1,5-bisphospha	( 165)	164	47.7	0.031
gi 4038667 dbj BAA35150.1	ribulose-1,5-bisphospha	( 165)	164	47.7	0.031
gi 155314472 gb ABT70076.1	Sequence 157546 from p	( 85)	161	46.7	0.031
gi 155313047 gb ABT68651.1	Sequence 156121 from p	( 86)	161	46.7	0.031
gi 155313083 gb ABT68687.1	Sequence 156157 from p	( 86)	161	46.7	0.031
gi 155330671 gb ABT86275.1	Sequence 173745 from p	( 86)	161	46.7	0.031
gi 155272683 gb ABT28287.1	Sequence 115757 from p	( 87)	161	46.7	0.032
gi 155328998 gb ABT84602.1	Sequence 172072 from p	( 87)	161	46.7	0.032
gi 155313149 gb ABT68753.1	Sequence 156223 from p	( 88)	161	46.8	0.032
gi 110224770 emb CAL07985.1	ribulose-bisphosphate	( 71)	160	46.4	0.032
gi 155354481 gb ABU10086.1	Sequence 197555 from p	( 89)	161	46.8	0.032
gi 155342898 gb ABT98502.1	Sequence 185972 from p	( 90)	161	46.8	0.032
gi 19459 emb CAA38343.1	ribulose bisphosphate car	( 30)	156	45.2	0.033
gi 155317275 gb ABT72879.1	Sequence 160349 from p	( 91)	161	46.8	0.033
gi 155313399 gb ABT69003.1	Sequence 156473 from p	( 91)	161	46.8	0.033
gi 155334135 gb ABT89739.1	Sequence 177209 from p	( 92)	161	46.8	0.033
gi 155343049 gb ABT84602.1	Sequence 186123 from p	( 92)	161	46.8	0.033
gi 155317205 gb ABT72809.1	Sequence 160279 from p	( 92)	161	46.8	0.033
gi 2641205 gb AAB86854.1	ribulose 1,5 bisphosphat	( 179)	164	47.7	0.033
gi 155339864 gb ABT95468.1	Sequence 182938 from p	( 93)	161	46.8	0.033
gi 223539597 gb EEF41184.1	Ribulose bisphosphate	( 181)	164	47.7	0.033
gi 155266920 gb ABT22524.1	Sequence 109994 from p	( 182)	164	47.7	0.033
gi 155317267 gb ABT72871.1	Sequence 160341 from p	( 95)	161	46.8	0.034
gi 3997330 gb AAC90760.1	AR015925 Sequence 12 from	( 707)	170	49.7	0.034
gi 4000314 gb AAC93746.1	AR023448 Sequence 12 from	( 707)	170	49.7	0.034
gi 155354583 gb ABU10188.1	Sequence 197657 from p	( 97)	161	46.8	0.034
gi 155313002 gb ABT68606.1	Sequence 156076 from p	( 98)	161	46.8	0.034
gi 155317207 gb ABT72811.1	Sequence 160281 from p	( 99)	161	46.8	0.035
gi 155339859 gb ABT95463.1	Sequence 182933 from p	( 99)	161	46.8	0.035
gi 155358630 gb ABU14235.1	Sequence 201704 from p	( 100)	161	46.8	0.035
gi 155314473 gb ABT70077.1	Sequence 157547 from p	( 101)	161	46.8	0.035
gi 155314360 gb ABT69964.1	Sequence 157434 from p	( 101)	161	46.8	0.035
gi 155343045 gb ABT98649.1	Sequence 186119 from p	( 101)	161	46.8	0.035
gi 155329232 gb ABT84836.1	Sequence 172306 from p	( 102)	161	46.8	0.035
gi 155339939 gb ABT95543.1	Sequence 183013 from p	( 104)	161	46.8	0.036
gi 4038665 dbj BAA35149.1	ribulose-1,5-bisphospha	( 165)	163	47.5	0.036
gi 4038697 dbj BAA35165.1	ribulose-1,5-bisphospha	( 165)	163	47.5	0.036
gi 4038659 dbj BAA35146.1	ribulose-1,5-bisphospha	( 165)	163	47.5	0.036
gi 4038701 dbj BAA35167.1	ribulose-1,5-bisphospha	( 165)	163	47.5	0.036
gi 4038657 dbj BAA35145.1	ribulose-1,5-bisphospha	( 165)	163	47.5	0.036
gi 4038683 dbj BAA35158.1	ribulose-1,5-bisphospha	( 165)	163	47.5	0.036
gi 155314381 gb ABT69985.1	Sequence 157455 from p	( 107)	161	46.8	0.037
gi 155316785 gb ABT72389.1	Sequence 159859 from p	( 107)	161	46.8	0.037
gi 155316808 gb ABT72412.1	Sequence 159882 from p	( 107)	161	46.8	0.037
gi 155345547 gb ABU01152.1	Sequence 188621 from p	( 108)	161	46.8	0.037
gi 155313746 gb ABT69350.1	Sequence 156820 from p	( 109)	161	46.8	0.037
gi 155328371 gb ABT83975.1	Sequence 171445 from p	( 110)	161	46.8	0.037
gi 155354794 gb ABU10399.1	Sequence 197868 from p	( 110)	161	46.8	0.037
gi 155314980 gb ABT70584.1	Sequence 158054 from p	( 111)	161	46.8	0.038
gi 155339625 gb ABT95229.1	Sequence 182699 from p	( 111)	161	46.8	0.038
gi 155312856 gb ABT68460.1	Sequence 155930 from p	( 112)	161	46.9	0.038
gi 155339631 gb ABT95235.1	Sequence 182705 from p	( 112)	161	46.9	0.038
gi 155290417 gb ABT46021.1	Sequence 133491 from p	( 175)	163	47.5	0.038
gi 155324230 gb ABT79834.1	Sequence 167304 from p	( 175)	163	47.5	0.038
gi 155337382 gb ABT92986.1	Sequence 180456 from p	( 175)	163	47.5	0.038
gi 155339934 gb ABT95538.1	Sequence 183008 from p	( 113)	161	46.9	0.038
gi 155339813 gb ABT95417.1	Sequence 182887 from p	( 113)	161	46.9	0.038
gi 21866 emb CAA25057.1	unnamed protein product [	( 113)	161	46.9	0.038
gi 155339827 gb ABT95431.1	Sequence 182901 from p	( 114)	161	46.9	0.038
gi 2605706 gb AAB84179.1	ribulose 1,5 bisphosphat	( 179)	163	47.5	0.039
gi 2605710 gb AAB84181.1	ribulose 1,5 bisphosphat	( 179)	163	47.5	0.039

gi 2605708 gb AAB84180.1	ribulose 1,5 bisphosphat	( 179)	163	47.5	0.039
gi 2641203 gb AAB86853.1	ribulose 1,5 bisphosphat	( 179)	163	47.5	0.039
gi 5002365 gb AAD37440.1	AF150667_1 ribulose 1,5 b	( 180)	163	47.5	0.039
gi 155354950 gb ABU10555.1	Sequence 198024 from p	( 116)	161	46.9	0.039
gi 155361098 gb ABU16703.1	Sequence 204172 from p	( 119)	161	46.9	0.04
gi 155350791 gb ABU06396.1	Sequence 193865 from p	( 120)	161	46.9	0.04
gi 155325508 gb ABT81112.1	Sequence 168582 from p	( 121)	161	46.9	0.04
gi 155339943 gb ABT95547.1	Sequence 183017 from p	( 124)	161	46.9	0.041
gi 155348334 gb ABU03939.1	Sequence 191408 from p	( 124)	161	46.9	0.041
gi 155354414 gb ABU10019.1	Sequence 197488 from p	( 126)	161	46.9	0.041
gi 228391314 gb ACQ35267.1	Sequence 27781 from pa	( 250)	164	47.9	0.042
gi 155342912 gb ABT98516.1	Sequence 185986 from p	( 129)	161	46.9	0.042
gi 155314586 gb ABT70190.1	Sequence 157660 from p	( 130)	161	46.9	0.042
gi 155328956 gb ABT84560.1	Sequence 172030 from p	( 130)	161	46.9	0.042
gi 155316739 gb ABT73243.1	Sequence 159813 from p	( 131)	161	46.9	0.042
gi 155314463 gb ABT70067.1	Sequence 157537 from p	( 131)	161	46.9	0.042
gi 155328950 gb ABT84554.1	Sequence 172024 from p	( 132)	161	46.9	0.043
gi 155328985 gb ABT84589.1	Sequence 172059 from p	( 132)	161	46.9	0.043
gi 155315886 gb ABT71490.1	Sequence 158960 from p	( 132)	161	46.9	0.043
gi 155312886 gb ABT68490.1	Sequence 155960 from p	( 132)	161	46.9	0.043
gi 4038713 dbj BAA35173.1	ribulose-1,5-bisphospha	( 165)	162	47.2	0.043
gi 4038691 dbj BAA35162.1	ribulose-1,5-bisphospha	( 165)	162	47.2	0.043
gi 155316700 gb ABT72304.1	Sequence 159774 from p	( 134)	161	46.9	0.043
gi 3978511 gb AAC83372.1	ribulose-1,5-bisphosphat	( 168)	162	47.2	0.043
gi 237506812 gb ACQ99130.1	conserved hypothetical	( 635)	168	49.2	0.043
gi 155339711 gb ABT95315.1	Sequence 182785 from p	( 135)	161	46.9	0.043
gi 155272688 gb ABT28292.1	Sequence 115762 from p	( 136)	161	46.9	0.044
gi 155339713 gb ABT95317.1	Sequence 182787 from p	( 137)	161	46.9	0.044
gi 155339928 gb ABT95532.1	Sequence 183002 from p	( 137)	161	46.9	0.044
gi 155339926 gb ABT95530.1	Sequence 183000 from p	( 138)	161	46.9	0.044
gi 155339948 gb ABT95552.1	Sequence 183022 from p	( 138)	161	46.9	0.044
gi 155314377 gb ABT69981.1	Sequence 157451 from p	( 138)	161	46.9	0.044
gi 155339942 gb ABT95546.1	Sequence 183016 from p	( 138)	161	46.9	0.044
gi 155339638 gb ABT95242.1	Sequence 182712 from p	( 138)	161	46.9	0.044
gi 155339938 gb ABT95542.1	Sequence 183012 from p	( 139)	161	46.9	0.044
gi 155339937 gb ABT95541.1	Sequence 183011 from p	( 139)	161	46.9	0.044
gi 155339946 gb ABT95550.1	Sequence 183020 from p	( 139)	161	46.9	0.044
gi 155339974 gb ABT95578.1	Sequence 183048 from p	( 139)	161	46.9	0.044
gi 1167948 gb AAA87039.1	ribulose-1,5-bisphosphat	( 174)	162	47.3	0.044
gi 155345894 gb ABU01499.1	Sequence 188968 from p	( 140)	161	46.9	0.044
gi 155317603 gb ABT73207.1	Sequence 160677 from p	( 140)	161	46.9	0.044
gi 155316495 gb ABT72099.1	Sequence 159569 from p	( 143)	161	47.0	0.045
gi 86156014 gb ABC86738.1	ribulose-1,5-bisphosphat	( 180)	162	47.3	0.045
gi 270383722 dbj BAI53119.1	ribulose-1,5-bisphosp	( 181)	162	47.3	0.046
gi 5002363 gb AAD37439.1	AF150666_1 ribulose 1,5 b	( 184)	162	47.3	0.046
gi 155339873 gb ABT95477.1	Sequence 182947 from p	( 150)	161	47.0	0.047
gi 155339634 gb ABT95238.1	Sequence 182708 from p	( 157)	161	47.0	0.048
gi 270226718 emb CBI15452.1	unnamed protein produ	( 157)	161	47.0	0.048
gi 155310684 gb ABT66288.1	Sequence 153758 from p	( 67)	157	45.7	0.049
gi 155364482 gb ABU20087.1	Sequence 207556 from p	( 164)	161	47.0	0.05
gi 4038723 dbj BAA35178.1	ribulose-1,5-bisphospha	( 165)	161	47.0	0.05
gi 4038675 dbj BAA35154.1	ribulose-1,5-bisphospha	( 165)	161	47.0	0.05
gi 4038715 dbj BAA35174.1	ribulose-1,5-bisphospha	( 165)	161	47.0	0.05
gi 4038709 dbj BAA35171.1	ribulose-1,5-bisphospha	( 165)	161	47.0	0.05
gi 4038681 dbj BAA35157.1	ribulose-1,5-bisphospha	( 165)	161	47.0	0.05
gi 4038669 dbj BAA35151.1	ribulose-1,5-bisphospha	( 165)	161	47.0	0.05
gi 4038685 dbj BAA35159.1	ribulose-1,5-bisphospha	( 165)	161	47.0	0.05
gi 4038721 dbj BAA35177.1	ribulose-1,5-bisphospha	( 165)	161	47.0	0.05
gi 4038711 dbj BAA35172.1	ribulose-1,5-bisphospha	( 165)	161	47.0	0.05
gi 4038693 dbj BAA35163.1	ribulose-1,5-bisphospha	( 165)	161	47.0	0.05
gi 4038677 dbj BAA35155.1	ribulose-1,5-bisphospha	( 165)	161	47.0	0.05
gi 4038679 dbj BAA35156.1	ribulose-1,5-bisphospha	( 165)	161	47.0	0.05

gi 4038673 dbj BAA35153.1	ribulose-1,5-bisphospha	( 165)	161	47.0	0.05
gi 4038719 dbj BAA35176.1	ribulose-1,5-bisphospha	( 165)	161	47.0	0.05
gi 4038703 dbj BAA35168.1	ribulose-1,5-bisphospha	( 165)	161	47.0	0.05
gi 4038671 dbj BAA35152.1	ribulose-1,5-bisphospha	( 165)	161	47.0	0.05
gi 155328300 gb ABT83904.1	Sequence 171374 from p	( 165)	161	47.0	0.05
gi 4038689 dbj BAA35161.1	ribulose-1,5-bisphospha	( 165)	161	47.0	0.05
gi 4038687 dbj BAA35160.1	ribulose-1,5-bisphospha	( 165)	161	47.0	0.05
gi 4038705 dbj BAA35169.1	ribulose-1,5-bisphospha	( 165)	161	47.0	0.05
gi 155350719 gb ABU06324.1	Sequence 193793 from p	( 169)	161	47.0	0.051
gi 62176930 emb CAG25595.1	putative rubisco small	( 170)	161	47.0	0.051
gi 143359527 gb EDE54618.1	hypothetical protein G	(1003)	169	49.6	0.051
gi 155317306 gb ABT72910.1	Sequence 160380 from p	( 172)	161	47.0	0.052
gi 155366238 gb ABU21843.1	Sequence 209312 from p	( 173)	161	47.0	0.052
gi 155339889 gb ABT95493.1	Sequence 182963 from p	( 173)	161	47.0	0.052
gi 155337928 gb ABT93532.1	Sequence 181002 from p	( 173)	161	47.0	0.052
gi 155339541 gb ABT95145.1	Sequence 182615 from p	( 173)	161	47.0	0.052
gi 155339820 gb ABT95424.1	Sequence 182894 from p	( 173)	161	47.0	0.052
gi 155363212 gb ABU06324.1	Sequence 206286 from p	( 174)	161	47.0	0.052
gi 155290394 gb ABT45998.1	Sequence 133468 from p	( 174)	161	47.0	0.052
gi 155339819 gb ABT95423.1	Sequence 182893 from p	( 174)	161	47.0	0.052
gi 155366221 gb ABT21826.1	Sequence 209295 from p	( 174)	161	47.0	0.052
gi 155338465 gb ABT94069.1	Sequence 181539 from p	( 174)	161	47.0	0.052
gi 155290393 gb ABT45997.1	Sequence 133467 from p	( 174)	161	47.0	0.052
gi 155290355 gb ABT45959.1	Sequence 133429 from p	( 174)	161	47.0	0.052
gi 119909033 dbj BAB19815.1	ribulose-1,5-bisphosph	( 174)	161	47.0	0.052
gi 155290439 gb ABT46043.1	Sequence 133513 from p	( 174)	161	47.0	0.052
gi 155339906 gb ABT95510.1	Sequence 182980 from p	( 174)	161	47.0	0.052
gi 155328951 gb ABT84555.1	Sequence 172025 from p	( 174)	161	47.0	0.052
gi 155367041 gb ABU22646.1	Sequence 210115 from p	( 174)	161	47.0	0.052
gi 155290395 gb ABT45999.1	Sequence 133469 from p	( 174)	161	47.0	0.052
gi 155329302 gb ABT84906.1	Sequence 172376 from p	( 174)	161	47.0	0.052
gi 155329151 gb ABT84755.1	Sequence 172225 from p	( 174)	161	47.0	0.052
gi 155317339 gb ABT72943.1	Sequence 160413 from p	( 174)	161	47.0	0.052
gi 155337948 gb ABT93552.1	Sequence 181022 from p	( 174)	161	47.0	0.052
gi 119908951 dbj BAB19811.1	ribulose-1,5-bisphosph	( 174)	161	47.0	0.052
gi 170769 gb AAA34301.1	ribulose-1,5-bisphosphate	( 174)	161	47.0	0.052
gi 155290467 gb ABT46071.1	Sequence 133541 from p	( 175)	161	47.0	0.052
gi 155364484 gb ABU20089.1	Sequence 207558 from p	( 175)	161	47.0	0.052
gi 155364432 gb ABU20037.1	Sequence 207506 from p	( 175)	161	47.0	0.052
gi 155364480 gb ABU20085.1	Sequence 207554 from p	( 175)	161	47.0	0.052
gi 155364443 gb ABU20048.1	Sequence 207517 from p	( 175)	161	47.0	0.052
gi 155365929 gb ABU21534.1	Sequence 209003 from p	( 175)	161	47.0	0.052
gi 155290398 gb ABT46002.1	Sequence 133472 from p	( 175)	161	47.0	0.052
gi 155322328 gb ABT77932.1	Sequence 165402 from p	( 175)	161	47.0	0.052
gi 155318875 gb ABT74479.1	Sequence 161949 from p	( 175)	161	47.0	0.052
gi 170771 gb AAA34302.1	ribulose-1,5-bisphosphate	( 175)	161	47.0	0.052
gi 155290411 gb ABT46015.1	Sequence 133485 from p	( 175)	161	47.0	0.052
gi 4090293 emb CAA10497.1	hypothetical protein [S	( 175)	161	47.0	0.052
gi 155317637 gb ABT73241.1	Sequence 160711 from p	( 175)	161	47.0	0.052
gi 155290438 gb ABT46042.1	Sequence 133512 from p	( 175)	161	47.0	0.052
gi 155334649 gb ABT90253.1	Sequence 177723 from p	( 175)	161	47.0	0.052
gi 11990901 dbj BAB19814.1	ribulose-1,5-bisphosph	( 175)	161	47.0	0.052
gi 155322263 gb ABT77867.1	Sequence 165337 from p	( 175)	161	47.0	0.052
gi 155290462 gb ABT46066.1	Sequence 133536 from p	( 175)	161	47.0	0.052
gi 155365920 gb ABU21525.1	Sequence 208994 from p	( 175)	161	47.0	0.052
gi 155366203 gb ABU21808.1	Sequence 209277 from p	( 175)	161	47.0	0.052
gi 155272724 gb ABT28328.1	Sequence 115798 from p	( 175)	161	47.0	0.052
gi 155317623 gb ABT73227.1	Sequence 160697 from p	( 175)	161	47.0	0.052
gi 155365813 gb ABU21418.1	Sequence 208887 from p	( 175)	161	47.0	0.052
gi 155364455 gb ABU20060.1	Sequence 207529 from p	( 175)	161	47.0	0.052
gi 155339909 gb ABT95513.1	Sequence 182983 from p	( 175)	161	47.0	0.052
gi 155364481 gb ABU20086.1	Sequence 207555 from p	( 175)	161	47.0	0.052

gi 155365815 gb ABU21420.1	Sequence 208889 from p	( 175)	161	47.0	0.052
gi 155313821 gb ABT69425.1	Sequence 156895 from p	( 175)	161	47.0	0.052
gi 155329133 gb ABT84737.1	Sequence 172207 from p	( 175)	161	47.0	0.052
gi 155364470 gb ABU20075.1	Sequence 207544 from p	( 175)	161	47.0	0.052
gi 155272723 gb ABT28327.1	Sequence 115797 from p	( 175)	161	47.0	0.052
gi 3063524 gb AAC14064.1	ribulose 1,5-bisphosphat	( 175)	161	47.0	0.052
gi 155365826 gb ABU21431.1	Sequence 208900 from p	( 175)	161	47.0	0.052
gi 155317196 gb ABT72800.1	Sequence 160270 from p	( 175)	161	47.0	0.052
gi 155272725 gb ABT28329.1	Sequence 115799 from p	( 175)	161	47.0	0.052
gi 11990893 dbj BAB19810.1	ribulose-1,5-bisphosph	( 175)	161	47.0	0.052
gi 155314391 gb ABT69995.1	Sequence 157465 from p	( 175)	161	47.0	0.052
gi 11990897 dbj BAB19812.1	ribulose-1,5-bisphosph	( 175)	161	47.0	0.052
gi 155364449 gb ABU20054.1	Sequence 207523 from p	( 175)	161	47.0	0.052
gi 155313037 gb ABT68641.1	Sequence 156111 from p	( 175)	161	47.0	0.052
gi 155364483 gb ABU20088.1	Sequence 207557 from p	( 175)	161	47.0	0.052
gi 155290274 gb ABT45878.1	Sequence 133348 from p	( 175)	161	47.0	0.052
gi 155290447 gb ABT46051.1	Sequence 133521 from p	( 175)	161	47.0	0.052
gi 155364485 gb ABU20090.1	Sequence 207559 from p	( 175)	161	47.0	0.052
gi 155364469 gb ABU20074.1	Sequence 207543 from p	( 175)	161	47.0	0.052
gi 155364436 gb ABU20041.1	Sequence 207510 from p	( 175)	161	47.0	0.052
gi 155364431 gb ABU20036.1	Sequence 207505 from p	( 175)	161	47.0	0.052
gi 4090291 emb CAA10496.1	hypothetical protein [S	( 175)	161	47.0	0.052
gi 155290427 gb ABT46031.1	Sequence 133501 from p	( 175)	161	47.0	0.052
gi 155313418 gb ABT69022.1	Sequence 156492 from p	( 175)	161	47.0	0.052
gi 155363226 gb ABU18831.1	Sequence 206300 from p	( 175)	161	47.0	0.052
gi 155290337 gb ABT45941.1	Sequence 133411 from p	( 175)	161	47.0	0.052
gi 155290423 gb ABT46027.1	Sequence 133497 from p	( 176)	161	47.0	0.052
gi 155365906 gb ABU21511.1	Sequence 208980 from p	( 176)	161	47.0	0.052
gi 155328218 gb ABT83822.1	Sequence 171292 from p	( 176)	161	47.0	0.052
gi 155290403 gb ABT46007.1	Sequence 133477 from p	( 176)	161	47.0	0.052
gi 155313035 gb ABT68639.1	Sequence 156109 from p	( 176)	161	47.0	0.052
gi 155338472 gb ABT94076.1	Sequence 181546 from p	( 176)	161	47.0	0.052
gi 155328708 gb ABT84312.1	Sequence 171782 from p	( 177)	161	47.0	0.053
gi 155313031 gb ABT68635.1	Sequence 156105 from p	( 178)	161	47.0	0.053
gi 155272721 gb ABT28325.1	Sequence 115795 from p	( 178)	161	47.0	0.053
gi 147807983 emb CAN64291.1	hypothetical protein	( 180)	161	47.0	0.053
gi 217946 dbj BAA03103.1	ribulose-1,5-bisphosphat	( 181)	161	47.0	0.053
gi 155313272 gb ABT68876.1	Sequence 156346 from p	( 99)	158	46.1	0.056
gi 149392361 gb ABR26002.1	ribulose bisphosphate	( 73)	156	45.5	0.062
gi 155337962 gb ABT93566.1	Sequence 181036 from p	( 177)	160	46.8	0.062
gi 148907958 gb ABR17099.1	unknown [Picea sitchen	( 143)	159	46.5	0.062
gi 1297052 emb CAA66201.1	ribulose-bisphosphate c	( 180)	160	46.8	0.062
gi 2529378 gb AAB81105.1	ribulose 1,5-bisphosphat	( 180)	160	46.8	0.062
gi 155313138 gb ABT68742.1	Sequence 156212 from p	( 105)	157	45.9	0.068
gi 149392158 gb ABR25937.1	ribulose bisphosphate	( 85)	156	45.6	0.069
gi 169652930 gb EDS85623.1	conserved hypothetical	( 629)	165	48.5	0.069
gi 6272547 gb AAF06097.1	ribulose 1,5-bisphosphat	( 176)	159	46.6	0.072
gi 155353027 gb ABU08632.1	Sequence 196101 from p	( 93)	156	45.6	0.073
gi 168363 gb AAA99429.1	ribulose 1,5-bisphosphate	( 182)	159	46.6	0.074
gi 6272549 gb AAF06099.1	ribulose 1,5-bisphosphat	( 182)	159	46.6	0.074
gi 809533 emb CAA60636.1	ribulose 1,5-bisphosphat	( 183)	159	46.6	0.074
gi 5002361 gb AAD37438.1	AF150665_1 ribulose 1,5 b	( 183)	159	46.6	0.074
gi 155317118 gb ABT72722.1	Sequence 160192 from p	( 97)	156	45.7	0.076
gi 149392567 gb ABR26086.1	ribulose bisphosphate	( 103)	156	45.7	0.079
gi 155343291 gb ABT98895.1	Sequence 186365 from p	( 163)	158	46.3	0.08
gi 155317286 gb ABT72890.1	Sequence 160360 from p	( 107)	156	45.7	0.081
gi 155356450 gb ABU12055.1	Sequence 199524 from p	( 87)	155	45.4	0.082
gi 671740 emb CAA59218.1	ribulose-bisphosphate ca	( 128)	156	45.8	0.092
gi 77554374 gb ABA97170.1	Ribulose bisphosphate c	( 128)	156	45.8	0.092
gi 6409333 gb AAF07944.1	AF192775_1 ribulose-1,5-b	( 168)	157	46.1	0.096
gi 3978513 gb AAC83373.1	ribulose-1,5-bisphosphat	( 168)	157	46.1	0.096
gi 6409331 gb AAF07943.1	AF192774_1 ribulose-1,5-b	( 168)	157	46.1	0.096

gi 164698711 gb ABY66908.1	chloroplast ribulose-1	( 169)	157	46.1	0.096
gi 6635351 gb AAF19793.1	AF162210_1 ribulose-1,5-b	( 181)	157	46.1	0.1
gi 147751750 gb EDK58817.1	hypothetical protein B	( 559)	162	47.7	0.1
gi 149392260 gb ABR25969.1	ribulose bisphosphate	( 157)	156	45.8	0.11
gi 155330083 gb ABT85687.1	Sequence 173157 from p	( 159)	156	45.9	0.11
gi 4038663 dbj BAA35148.1	ribulose-1,5-bisphospha	( 165)	156	45.9	0.11
gi 347451 gb AAA84592.1	ribulose 1,5-bisphosphate	( 169)	156	45.9	0.11
gi 148665944 gb EDK98360.1	mCG1038238 [Mus muscul	( 216)	157	46.2	0.11
gi 218208 dbj BAA00538.1	small subunit of ribulos	( 175)	156	45.9	0.12
gi 77554383 gb ABA97179.1	Ribulose bisphosphate c	( 175)	156	45.9	0.12
gi 215768262 dbj BAH00491.1	unnamed protein produ	( 175)	156	45.9	0.12
gi 108862509 gb ABG21975.1	Ribulose bisphosphate	( 175)	156	45.9	0.12
gi 113649111 dbj BAF29623.1	Osl2g0291100 [Oryza s	( 175)	156	45.9	0.12
gi 125536346 gb EAY82834.1	hypothetical protein O	( 175)	156	45.9	0.12
gi 222616937 gb EEE53069.1	hypothetical protein O	( 175)	156	45.9	0.12
gi 125551406 gb EAY97115.1	hypothetical protein O	( 175)	156	45.9	0.12
gi 215687290 dbj BAG91855.1	unnamed protein produ	( 175)	156	45.9	0.12
gi 113649066 dbj BAF29578.1	Osl2g0274700 [Oryza s	( 175)	156	45.9	0.12
gi 215767359 dbj BAG99587.1	unnamed protein produ	( 175)	156	45.9	0.12
gi 215692618 dbj BAG88038.1	unnamed protein produ	( 175)	156	45.9	0.12
gi 155364475 gb ABU20080.1	Sequence 207549 from p	( 175)	156	45.9	0.12
gi 125579068 gb EAA220214.1	hypothetical protein O	( 175)	156	45.9	0.12
gi 108862508 gb ABG21974.1	Ribulose bisphosphate	( 175)	156	45.9	0.12
gi 149392066 gb ABR25908.1	chloroplast ribulose b	( 175)	156	45.9	0.12
gi 77554384 gb ABA97180.1	Ribulose bisphosphate c	( 175)	156	45.9	0.12
gi 218186697 gb EEC69124.1	hypothetical protein O	( 175)	156	45.9	0.12
gi 108862507 gb ABG21973.1	Ribulose bisphosphate	( 175)	156	45.9	0.12
gi 215704569 dbj BAG94202.1	unnamed protein produ	( 175)	156	45.9	0.12
gi 77554373 gb ABA97169.1	Ribulose bisphosphate c	( 175)	156	45.9	0.12
gi 149390921 gb ABR25478.1	ribulose bisphosphate	( 175)	156	45.9	0.12
gi 113649113 dbj BAF29625.1	Osl2g0291400 [Oryza s	( 175)	156	45.9	0.12
gi 77554377 gb ABA97173.1	Ribulose bisphosphate c	( 175)	156	45.9	0.12
gi 222630752 gb EEE62884.1	hypothetical protein O	( 175)	156	45.9	0.12
gi 113649114 dbj BAF29626.1	Osl2g0292400 [Oryza s	( 175)	156	45.9	0.12
gi 155277209 gb ABT32813.1	Sequence 120283 from p	( 175)	156	45.9	0.12
gi 77554260 gb ABA97056.1	Ribulose bisphosphate c	( 175)	156	45.9	0.12
gi 149391337 gb ABR25686.1	ribulose bisphosphate	( 175)	156	45.9	0.12
gi 215687018 dbj BAG90832.1	unnamed protein produ	( 175)	156	45.9	0.12
gi 38374128 gb AAR19268.1	ribulose-1,5-bisphospha	( 175)	156	45.9	0.12
gi 125579070 gb EAA220216.1	hypothetical protein O	( 175)	156	45.9	0.12
gi 2407283 gb AAB70544.1	ribulose 1,5-bisphospha	( 175)	156	45.9	0.12
gi 155313107 gb ABT68711.1	Sequence 156181 from p	( 186)	156	45.9	0.12
gi 164698713 gb ABY66909.1	chloroplast ribulose-1	( 169)	155	45.6	0.13
gi 155323856 gb ABT79460.1	Sequence 166930 from p	( 174)	155	45.7	0.13
gi 155365891 gb ABU21496.1	Sequence 208965 from p	( 175)	155	45.7	0.14
gi 218210 dbj BAA00539.1	small subunit of ribulos	( 175)	155	45.7	0.14
gi 155360055 gb ABU15660.1	Sequence 203129 from p	( 114)	153	45.0	0.14
gi 155358557 gb ABU14162.1	Sequence 201631 from p	( 119)	153	45.0	0.14
gi 155365832 gb ABU21437.1	Sequence 208906 from p	( 192)	155	45.7	0.14
gi 155359786 gb ABU15391.1	Sequence 202860 from p	( 80)	151	44.4	0.15
gi 169220 gb AAA33718.1	ribulose 1,5-bisphosphate	( 46)	148	43.5	0.16
gi 155361939 gb ABU17544.1	Sequence 205013 from p	( 77)	150	44.2	0.17
gi 164698715 gb ABY66910.1	chloroplast ribulose-1	( 121)	152	44.8	0.17
gi 155331392 gb ABT86996.1	Sequence 174466 from p	( 162)	153	45.2	0.18
gi 4038707 dbj BAA35170.1	ribulose-1,5-bisphospha	( 164)	153	45.2	0.18
gi 4038661 dbj BAA35147.1	ribulose-1,5-bisphospha	( 165)	153	45.2	0.18
gi 155277376 gb ABT32980.1	Sequence 120450 from p	( 169)	153	45.2	0.18
gi 241936147 gb EE509292.1	hypothetical protein S	( 169)	153	45.2	0.18
gi 11990899 dbj BAB19813.1	ribulose-1,5-bisphosph	( 173)	153	45.2	0.18
gi 155364450 gb ABU20055.1	Sequence 207524 from p	( 175)	153	45.2	0.19
gi 155363127 gb ABU18732.1	Sequence 206201 from p	( 177)	153	45.2	0.19
gi 121228437 gb ABM50955.1	conserved hypothetical	( 539)	158	46.8	0.19

gi 126242379 gb ABO05472.1	conserved hypothetical	( 539)	158	46.8	0.19
gi 155312982 gb ABT68586.1	Sequence 156056 from p	( 179)	153	45.2	0.19
gi 125119921 gb ABN25683.1	Sequence 3 from patent	( 181)	153	45.2	0.19
gi 24940138 emb CAD11990.1	rubisco small subunit	( 181)	153	45.2	0.19
gi 24940140 emb CAD11991.1	rubisco small subunit	( 181)	153	45.2	0.19
gi 155350798 gb ABU06403.1	Sequence 193872 from p	( 187)	153	45.2	0.19
gi 116789830 gb ABK25403.1	unknown [Picea sitchen	( 188)	153	45.2	0.2
gi 155287895 gb ABT43499.1	Sequence 130969 from p	( 98)	150	44.3	0.2
gi 219885421 gb ACL53085.1	unknown [Zea mays]	( 153)	152	44.9	0.2
gi 221044072 dbj BAH13713.1	unnamed protein produ	( 375)	156	46.2	0.2
gi 155354220 gb ABU09825.1	Sequence 197294 from p	( 124)	151	44.6	0.2
gi 155339856 gb ABT95460.1	Sequence 182930 from p	( 104)	150	44.3	0.21
gi 140292764 gb ECK78657.1	hypothetical protein G	( 84)	149	44.0	0.21
gi 155354831 gb ABU10436.1	Sequence 197905 from p	( 68)	148	43.7	0.21
gi 194703224 gb ABF85696.1	unknown [Zea mays]	( 170)	152	45.0	0.21
gi 195636562 gb ACG37749.1	ribulose bisphosphate	( 170)	152	45.0	0.21
gi 194702782 gb ACF85475.1	unknown [Zea mays]	( 170)	152	45.0	0.21
gi 1673456 emb CAA70416.1	rubisco small subunit	( 170)	152	45.0	0.21
gi 155313389 gb ABT68993.1	Sequence 156463 from p	( 174)	152	45.0	0.22
gi 1536889 emb CAA69102.1	ribulose-bisphosphate c	( 182)	152	45.0	0.22
gi 155364463 gb ABU20068.1	Sequence 207537 from p	( 191)	152	45.0	0.23
gi 155361870 gb ABU17475.1	Sequence 204944 from p	( 79)	148	43.7	0.23
gi 155339812 gb ABT95416.1	Sequence 182886 from p	( 68)	147	43.4	0.24
gi 169216 gb AAA33716.1	ribulose 1,5-bisphosphate	( 106)	149	44.1	0.24
gi 155353314 gb ABU08919.1	Sequence 196388 from p	( 111)	149	44.1	0.25
gi 155339880 gb ABT95484.1	Sequence 182954 from p	( 139)	150	44.4	0.25
gi 155354020 gb ABU09625.1	Sequence 197094 from p	( 116)	149	44.1	0.26
gi 51949879 gb AAU14862.1	chloroplast ribulose-1,	( 181)	151	44.8	0.26
gi 2588890 dbj BAA23214.1	small subunit of ribulo	( 182)	151	44.8	0.26
gi 83375916 gb ABC17796.1	plastid ribulose-1,5-bi	( 117)	149	44.1	0.26
gi 83375908 gb ABC17792.1	plastid ribulose-1,5-bi	( 117)	149	44.1	0.26
gi 83375906 gb ABC17791.1	plastid ribulose-1,5-bi	( 117)	149	44.1	0.26
gi 83375904 gb ABC17790.1	plastid ribulose-1,5-bi	( 117)	149	44.1	0.26
gi 83375922 gb ABC17799.1	plastid ribulose-1,5-bi	( 117)	149	44.1	0.26
gi 83375918 gb ABC17797.1	plastid ribulose-1,5-bi	( 117)	149	44.1	0.26
gi 83375920 gb ABC17798.1	plastid ribulose-1,5-bi	( 117)	149	44.1	0.26
gi 83375926 gb ABC17801.1	plastid ribulose-1,5-bi	( 117)	149	44.1	0.26
gi 83375914 gb ABC17795.1	plastid ribulose-1,5-bi	( 117)	149	44.1	0.26
gi 83375912 gb ABC17794.1	plastid ribulose-1,5-bi	( 117)	149	44.1	0.26
gi 83375924 gb ABC17800.1	plastid ribulose-1,5-bi	( 117)	149	44.1	0.26
gi 83375910 gb ABC17793.1	plastid ribulose-1,5-bi	( 117)	149	44.1	0.26
gi 12240092 gb AAG49562.1	ribulose-1,5-bisphospha	( 118)	149	44.1	0.26
gi 155359816 gb ABU15421.1	Sequence 202890 from p	( 120)	149	44.1	0.27
gi 155317240 gb ABT72844.1	Sequence 160314 from p	( 122)	149	44.1	0.27
gi 155354953 gb ABU10558.1	Sequence 198027 from p	( 65)	146	43.2	0.28
gi 155348643 gb ABU04248.1	Sequence 191717 from p	( 55)	145	42.9	0.29
gi 155339866 gb ABT95470.1	Sequence 182940 from p	( 71)	146	43.2	0.3
gi 155346569 gb ABU02174.1	Sequence 189643 from p	( 57)	145	42.9	0.3
gi 155337776 gb ABT93380.1	Sequence 180850 from p	( 175)	150	44.5	0.3
gi 139746677 gb ECH18013.1	hypothetical protein G	( 273)	152	45.2	0.3
gi 155364474 gb ABU20079.1	Sequence 207548 from p	( 176)	150	44.5	0.3
gi 155356410 gb ABU12015.1	Sequence 199484 from p	( 59)	145	42.9	0.3
gi 51535980 dbj BAD38061.1	putative ribulose 1,5-	( 183)	150	44.5	0.31
gi 51535337 dbj BAD38596.1	putative ribulose 1,5-	( 183)	150	44.5	0.31
gi 113355448 dbj BAF07831.1	Os02g0152400 [Oryza s	( 183)	150	44.5	0.31
gi 218190079 gb EEC72506.1	hypothetical protein O	( 183)	150	44.5	0.31
gi 222622186 gb EEE56318.1	hypothetical protein O	( 183)	150	44.5	0.31
gi 155342787 gb ABT98391.1	Sequence 185861 from p	( 62)	145	42.9	0.31
gi 155351768 gb ABU07373.1	Sequence 194842 from p	( 62)	145	42.9	0.31
gi 155347701 gb ABU03306.1	Sequence 190775 from p	( 159)	149	44.3	0.33
gi 155290189 gb ABT45793.1	Sequence 133263 from p	( 90)	146	43.3	0.35
gi 155338033 gb ABT93637.1	Sequence 181107 from p	( 177)	149	44.3	0.35

gi|155365912|gb|ABU21517.1| Sequence 208986 from p ( 185) 149 44.3 0.36  
gi|155343299|gb|ABT98903.1| Sequence 186373 from p ( 153) 148 44.0 0.37  
gi|50978423|emb|CAH10356.1| ribulose 1,5 bisphosph ( 154) 148 44.0 0.37  
gi|2407281|gb|AAB70543.1| ribulose 1,5-bisphosphat ( 175) 148 44.1 0.41  
gi|155350738|gb|ABU06343.1| Sequence 193812 from p ( 73) 144 42.8 0.41  
gi|155336153|gb|ABT91757.1| Sequence 179227 from p ( 59) 143 42.5 0.42  
gi|168217|gb|AAA33361.1| ribulose-1,5-bisphosphate ( 183) 148 44.1 0.42  
gi|168997361|gb|ACA42439.1| ribulose-1,5-bisphosph ( 183) 148 44.1 0.42  
gi|1345545|emb|CAA37517.1| NySS42 protein [Nicotia ( 39) 141 41.8 0.42  
gi|155313776|gb|ABT69380.1| Sequence 156850 from p ( 122) 146 43.5 0.44  
gi|155315815|gb|ABT71419.1| Sequence 158889 from p ( 124) 146 43.5 0.44  
gi|184212052|gb|EDU09095.1| conserved hypothetical ( 479) 152 45.4 0.45  
gi|42736575|gb|AAS40510.1| hypothetical protein BC ( 325) 150 44.8 0.47  
gi|113649112|dbj|BAF29624.1| Osl2g0291200 [Oryza s ( 123) 145 43.2 0.51  
gi|125579067|gb|EAZ20213.1| hypothetical protein O ( 123) 145 43.2 0.51  
gi|155314254|gb|ABT69858.1| Sequence 157328 from p ( 186) 146 43.6 0.59  
gi|155353453|gb|ABU09058.1| Sequence 196527 from p ( 96) 143 42.7 0.59  
gi|50978421|emb|CAH10355.1| ribulose 1,5 bisphosph ( 153) 145 43.3 0.6  
gi|226516450|gb|AC062445.1| ribulose bisphosphate ( 191) 146 43.6 0.6  
gi|156217399|gb|EDO38316.1| predicted protein [Nem ( 154) 145 43.3 0.6  
gi|155365845|gb|ABU21450.1| Sequence 208919 from p ( 175) 145 43.4 0.66  
gi|155365922|gb|ABU21527.1| Sequence 208996 from p ( 253) 146 43.8 0.73  
gi|155312310|gb|ABT67914.1| Sequence 155384 from p ( 68) 140 41.8 0.74  
gi|155338871|gb|ABT94475.1| Sequence 181945 from p ( 35) 137 40.9 0.74  
gi|22465|emb|CAA68419.1| ribulose 1,5-bisphosphate ( 169) 144 43.1 0.76  
gi|217964|dbj|BAA00120.1| ribulose 1,5-bisphosphat ( 170) 144 43.1 0.76  
gi|194702004|gb|ACF85086.1| unknown [Zea mays] ( 170) 144 43.1 0.76  
gi|194703936|gb|ACF86052.1| unknown [Zea mays] ( 170) 144 43.1 0.76  
gi|22474|emb|CAA29784.1| ribulose-1,5-bisphosphate ( 170) 144 43.1 0.76  
gi|155290251|gb|ABT45855.1| Sequence 133325 from p ( 175) 144 43.1 0.77  
gi|155337922|gb|ABT93526.1| Sequence 180996 from p ( 176) 144 43.1 0.78  
gi|155352950|gb|ABU08555.1| Sequence 196024 from p ( 74) 140 41.9 0.79  
gi|155295149|gb|ABT50753.1| Sequence 138223 from p ( 180) 144 43.2 0.79  
gi|226454223|gb|EEH51530.1| predicted protein [Mic ( 183) 144 43.2 0.8  
gi|226454056|gb|EEH51363.1| predicted protein [Mic ( 186) 144 43.2 0.81  
gi|155339858|gb|ABT95462.1| Sequence 182932 from p ( 40) 137 40.9 0.82  
gi|142559962|gb|ECY95758.1| hypothetical protein G ( 465) 148 44.5 0.83  
gi|155363740|gb|ABU19345.1| Sequence 206814 from p ( 41) 137 41.0 0.83  
gi|155343223|gb|ABT98827.1| Sequence 186297 from p ( 158) 143 42.9 0.84  
gi|155351690|gb|ABU07295.1| Sequence 194764 from p ( 43) 137 41.0 0.86  
gi|226454222|gb|EEH51529.1| predicted protein [Mic ( 204) 144 43.2 0.86  
gi|155314429|gb|ABT70033.1| Sequence 157503 from p ( 44) 137 41.0 0.87  
gi|155365907|gb|ABU21512.1| Sequence 208981 from p ( 210) 144 43.2 0.88  
gi|155348378|gb|ABU03983.1| Sequence 191452 from p ( 36) 136 40.7 0.89  
gi|155364282|gb|ABU19887.1| Sequence 207356 from p ( 45) 137 41.0 0.89  
gi|171699531|gb|ACB52512.1| ribulose bisphosphate ( 111) 141 42.3 0.9  
gi|155317771|gb|ABT73375.1| Sequence 160845 from p ( 92) 140 42.0 0.92  
gi|155354463|gb|ABU10068.1| Sequence 197537 from p ( 49) 137 41.0 0.94  
gi|155361898|gb|ABU17503.1| Sequence 204972 from p ( 96) 140 42.0 0.95  
gi|155315222|gb|ABT70826.1| Sequence 158296 from p ( 51) 137 41.0 0.97  
gi|155317459|gb|ABT73063.1| Sequence 160533 from p ( 51) 137 41.0 0.97  
gi|155363680|gb|ABU19285.1| Sequence 206754 from p ( 41) 136 40.7 0.97  
gi|155355914|gb|ABU11519.1| Sequence 198988 from p ( 52) 137 41.0 0.99  
gi|155367257|gb|ABU22862.1| Sequence 210331 from p ( 52) 137 41.0 0.99  
gi|155290238|gb|ABT45842.1| Sequence 133312 from p ( 42) 136 40.7 0.99  
gi|155353831|gb|ABU09436.1| Sequence 196905 from p ( 82) 139 41.7 0.99  
gi|155313803|gb|ABT69407.1| Sequence 156877 from p ( 53) 137 41.1 1

>>gi|136192920|gb|EBM50575.1| hypothetical protein GOS\_8 (396 aa)  
initn: 64 init1: 64 opt: 284 Z-score: 346.2 bits: 75.5 E(): 3.2e-10

Smith-Waterman score: 331; 26.905% identity (43.095% similar) in 420 aa overlap  
(770-1149:1-394)

```

      740      750      760      770      780      790
frame2 VFCFHGVHASRCKQPSSNCS.VLWSFWNRPYSR.QVYLPQVLHV-W--RSR.R.NSYHRS
      ::: . . . . .
gi|136 YSRHSRCNRSTRHTRWSRRTRYSRHSRHSR
      10      20      30

```

```

      800      810      820      830      840      850
frame2 FGR.RCYQHW.GYASYGCQNP.GR.YLDH.WCW.RWTPCS.GSSRFR-.RCNWLPFDYGS
      .: . . . . .
gi|136 HSR-RIRSTWHTRWSR-CSRHSRYYSARSTRHTRWSRCSRCSRCSRCSRCNW-PSRHRTR
      40      50      60      70      80

```

```

      860      870      880      890
frame2 C-----WCLRFR.HFHW---.RFSH.ANSGSCVEPTSRNGCAGEV.-RR.SS
      : . . . . .
gi|136 CSRCSRCSRCSRHSRCNRSTRHTRWSRRTRYSRYSRHSRHSRHSRRIRSTWHTRWSRCSR
      90      100      110      120      130      140

```

```

      900      910      920      930      940
frame2 SSYLAWTKDSNANHLLQGTGYGFRSSEV-RCSAC-W-SQHPRYHHCYRAN----HD----S.
      : . . . . .
gi|136 HSRYYWSARST-RHTRWSRCSRCSRCSRCSRCNWPSRHRTRCSRCSRCSRHSRCNRST
      150      160      170      180      190      200

```

```

      950      960      970      980      990      1000
frame2 PH.KDASR-FW---C.PYR.D.C.RCAYHPS.RSW.AHRSSD.CSR.SILYCFPIGCCLA
      .: . . . . .
gi|136 RHTRCSRSTRSTWYTRCS--RCSRCSRCSRH-SRCNWPSRHS--RCSR---C---SRCSR
      210      220      230      240      250

```

```

      1010      1020      1030      1040      1050      1060
frame2 CSRFRRHHP.RFDEPNPYWSHLDASAGNGCRHRSDQPTSCWRRRRG.LACSFYFEGC-YC
      .: . . . . .
gi|136 CSRCSRHSRCNWSTWHTRWSRCSRCSRCSRH-----SRCNWPSRHSRCSRCSRCSRCSRC
      260      270      280      290      300

```

```

      1070      1080      1090      1100      1110      1120
frame2 SRRP-CSFYDRRVSNRCCSCIR.RCYRYERFGRTPC.GKRPSFCCRKRSQAQRC.LR.R
      .: . . . . .
gi|136 SRHSRCNWSTWHTRWSRCSRC--SRCSRHSRCNW-PSRHSRCSRCSRC-SRCSRCSRHSR
      310      320      330      340      350      360

```

```

      1130      1140      1150      1160      1170
frame2 .DFS--RRAWSS.R.GSR.RFWSGCCRYPPRSFYRYELPRYGSRF.KPCYC..CYYDRY.L
      .: . . . . .
gi|136 CNWSTWHTRWSRCSRCSRCSRHSRCNWPSRHT
      370      380      390

```

>>gi|143377609|gb|EDE65555.1| hypothetical protein GOS\_1 (418 aa)  
initn: 205 init1: 82 opt: 269 Z-score: 327.4 bits: 72.1 E(): 3.5e-09  
Smith-Waterman score: 269; 26.316% identity (45.767% similar) in 437 aa overlap  
(749-1145:12-413)

```

      720      730      740      750      760      770
frame2 VGIEEEWDDVNWL.ASSS.GHVFCFHGVHASRCKQPS-----SNCS.VLWSFWNRPYSR
      .: . . . . .
gi|143 RIPRISRNTRISRCTRTNFWIPRISRCSRCNWS-----TR
      10      20      30

```

Monsanto Company

Regulatory Product Characterization Center  
RAR-10-070





Smith-Waterman score: 202; 23.750% identity (43.750% similar) in 240 aa overlap (933-1144:315-541)

```

      910      920      930      940      950      960
frame2 AWTKDSNANHQLGTYGFRSSEVRCSACWSQHPRYHHCYRANHDS.PH.KDASRFWC.PY
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|229 VIVTGTITGGHRRHIPTVAPGATILAHGPAHTRAHHCLRNNEQRSTERYHHYPWSWCNTR
      290      300      310      320      330      340
```

```

      970      980      990      1000      1010
frame2 R.D.-C.RCAYHPS.RSW.AHRSSD.-CSR.SIL---YCFPIGCCLACSRFRRHHP.RFD
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|229 SDDHGCTR--YHHYPWSCNTRSDDHGCTRYHHYPWSCYTWSNDHGCTRYH-HYPWS--
      350      360      370      380      390
```

```

      1020      1030      1040      1050      1060
frame2 EPNPYWSHLDLSAGNGCRHRSDQP-TSCW-WRRRG.LACS-FFYFEGCYCSRRP----CSF
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|229 -----WCNTRSDDHGCARYHHYPWNRCYTWSHN--YSCNRYHYHYPWSCNTRSDHHGCTR
      400      410      420      430      440      450
```

```

      1070      1080      1090      1100      1110
frame2 YDRRVSNRCCS-CIR.RCYRYERFGRTPC.GKRPSF-C-----CRKRSQAQRC.
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|229 Y-HHYPWSCYTWSHHYGCTRYHHYPWSCNTRSHDHGCTRYHHYPWSCNTRSHDHGCT
      460      470      480      490      500      510
```

```

      1120      1130      1140      1150      1160      1170
frame2 LR.R.DFSR-RAWSS.R.GSR.R--FWSSCRYPSPRYELPRYGSRF.KPCYC..CYY
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|229 RYHNPWNRCYTWSHHYGCTRNNDYTWSWCNTWSDDHGCTSKNYVKIKI
      520      530      540      550      560
```

>>gi|20857|emb|CAA27865.1| ribulose 1.5-bisphosphate car (180 aa)  
initn: 183 init1: 183 opt: 192 Z-score: 236.9 bits: 54.1 E(): 0.00039  
Smith-Waterman score: 192; 52.308% identity (67.692% similar) in 65 aa overlap (1185-1249:119-180)

```

      1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYYDRY.LPRVHGF DGWSWS.DRTL RH.GCLMSSRIRARYRIL
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|208 YLLRKGWVPCLEFELEKGFVYREHNKSPGYDGRYWTMWK-LPMFGTTDASQVLKELDEV
      90      100      110      120      130      140
```

```

      1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSCTICACNLLCF
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|208 VAAYP--QAFVRIIGFDNVRQVCISFIAHTPESY
      150      160      170      180
```

```

      1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFFLLFVVC.I.NYKRYANILF.VKMCQ
```

>>gi|20855|emb|CAA27864.1| ribulose bisphosphate carboxy (180 aa)  
initn: 183 init1: 183 opt: 192 Z-score: 236.9 bits: 54.1 E(): 0.00039  
Smith-Waterman score: 192; 52.308% identity (67.692% similar) in 65 aa overlap (1185-1249:119-180)

```

      1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYYDRY.LPRVHGF DGWSWS.DRTL RH.GCLMSSRIRARYRIL
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|208 YLLRKGWVPCLEFELEKGFVYREHNKSPGYDGRYWTMWK-LPMFGTTDASQVLKELDEV
```

```

      90      100      110      120      130      140
      1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSCTICACNLLCF
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|208 VAAYP--QAFVRIIGFDNVRQVCISFIAHTPESY
      150      160      170      180
```

```

      1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFFLLFVVC.I.NYKRYANILF.VKMCQ
```

>>gi|20859|emb|CAA25390.1| ribulose bisphosphate carboxy (180 aa)  
initn: 183 init1: 183 opt: 192 Z-score: 236.9 bits: 54.1 E(): 0.00039  
Smith-Waterman score: 192; 52.308% identity (67.692% similar) in 65 aa overlap (1185-1249:119-180)

```

      1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYYDRY.LPRVHGF DGWSWS.DRTL RH.GCLMSSRIRARYRIL
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|208 YLLRKGWVPCLEFELEKGFVYREHNKSPGYDGRYWTMWK-LPMFGTTDASQVLKELDEV
      90      100      110      120      130      140
```

```

      1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSCTICACNLLCF
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|208 VAAYP--QAFVRIIGFDNVRQVCISFIAHTPESY
      150      160      170      180
```

```

      1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFFLLFVVC.I.NYKRYANILF.VKMCQ
```

>>gi|436489|gb|AAA03697.1| ribulose-1,5-bisphosphate car (182 aa)  
initn: 159 init1: 159 opt: 192 Z-score: 236.8 bits: 54.1 E(): 0.00039  
Smith-Waterman score: 192; 45.946% identity (63.514% similar) in 74 aa overlap (1181-1249:110-180)

```

      1160      1170      1180      1190      1200
frame2 PYRYELPRYGSRF.KPCYC..CYYDRY.LPRVHG-----FDGWSWS.DRTL RH.GCLMSS
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|436 EESLLKEVQYLLNNGWVPCLEFEPTHGFFVYREHGNTPGYDGRYWTMWK-LPMFGCTDPS
      80      90      100      110      120      130
```

```

      1210      1220      1230      1240      1250      1260
frame2 RIRARYRILRS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSCTI
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|436 QVVAELEEAKKAYP--EAFIRIIGFDNVRQVCVFIAYKPASYGA
      140      150      160      170      180
```

```

      1270      1280      1290      1300      1310      1320
frame2 CCACNLLCFLFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFFLLFVVC.I.NYKRYAN
```

>>gi|436488|gb|AAA03696.1| ribulose-1,5-bisphosphate car (183 aa)  
initn: 159 init1: 159 opt: 192 Z-score: 236.8 bits: 54.1 E(): 0.00039  
Smith-Waterman score: 192; 45.946% identity (63.514% similar) in 74 aa overlap (1181-1249:111-181)

```

      1160      1170      1180      1190      1200
frame2 PYRYELPRYGSRF.KPCYC..CYYDRY.LPRVHG-----FDGWSWS.DRTL RH.GCLMSS
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|436 EESLLKEVQYLLNNGWVPCLEFEPTHGFFVYREHGNTPGYDGRYWTMWK-LPMFGCTDPS
      90      100      110      120      130
```

```

      1210      1220      1230      1240      1250      1260
frame2 RIRARYRILRS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSCTI
.. : . . . : ..... : :
gi|436 QVVALEEEAKKAYP--EAFIRIIGFDNVRQVCVSFIAYKPASYGA
      140      150      160      170      180

      1270      1280      1290      1300      1310      1320
frame2 CCACNLLCFLFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYAN

>>gi|167274|gb|AAA33038.1| ribulose 1,5-bisphosphate car (183 aa)
  initn: 159 init1: 159 opt: 192 Z-score: 236.8 bits: 54.1 E(): 0.00039
Smith-Waterman score: 192; 45.946% identity (63.514% similar) in 74 aa overlap
(1181-1249:111-181)

      1160      1170      1180      1190      1200
frame2 PYRYELPRYGSRF.KPCYC..CYYDRY.LPRVHG----FDGWSWS.DRTL RH.GCLMSS
      : :      : :      : .      : :      : :
gi|167 EESLLKEVQYLLNNGWVPCLEFEPTHG FVYREHGNTPGYDGRYWTMWK-LPMFGCTDPS
      90      100      110      120      130

      1210      1220      1230      1240      1250      1260
frame2 RIRARYRILRS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSCTI
.. : . . . : ..... : :
gi|167 QVVALEEEAKKAYP--EAFIRIIGFDNVRQVCVSFIAYKPASYGA
      140      150      160      170      180

      1270      1280      1290      1300      1310      1320
frame2 CCACNLLCFLFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYAN

>>gi|304447|gb|AAA03698.1| ribulose-1,5-bisphosphate car (186 aa)
  initn: 159 init1: 159 opt: 192 Z-score: 236.7 bits: 54.2 E(): 0.0004
Smith-Waterman score: 192; 45.946% identity (63.514% similar) in 74 aa overlap
(1181-1249:114-184)

      1160      1170      1180      1190      1200
frame2 PYRYELPRYGSRF.KPCYC..CYYDRY.LPRVHG----FDGWSWS.DRTL RH.GCLMSS
      : :      : :      : .      : :      : :
gi|304 EESLLKEVQYLLNNGWVPCLEFEPTHG FVYREHGNTPGYDGRYWTMWK-LPMFGCTDPS
      90      100      110      120      130      140

      1210      1220      1230      1240      1250      1260
frame2 RIRARYRILRS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSCTI
.. : . . . : ..... : :
gi|304 QVVALEEEAKKAYP--EAFIRIIGFDNVRQVCVSFIAYKPASYDA
      150      160      170      180

      1270      1280      1290      1300      1310      1320
frame2 CCACNLLCFLFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYAN

>>gi|7545285|gb|AAA33683.2| ribulose 1,5-bisphosphate ca (27 aa)
  initn: 182 init1: 182 opt: 182 Z-score: 235.1 bits: 51.1 E(): 0.00049
Smith-Waterman score: 182; 100.000% identity (100.000% similar) in 27 aa overlap
(1223-1249:1-27)

      1200      1210      1220      1230      1240      1250
frame2 DRTL RH.GCLMSSRIRARYRILRS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.
      : : : : : :
gi|754 AFVRIIGFDNVRQVCISFIAHTPESY
      10      20

      1260      1270      1280      1290      1300      1310
frame2 VLWHWENCFSCTICACNLLCFLFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLF

```

```

>>gi|170320|gb|AAA34111.1| ribulose-1,5-bisphosphate car (86 aa)
  initn: 145 init1: 145 opt: 186 Z-score: 233.6 bits: 52.5 E(): 0.0006
Smith-Waterman score: 186; 46.154% identity (70.769% similar) in 65 aa overlap
(1185-1249:25-86)

```

```

      1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYYDRY.LPRVHGFDGWSWS.DRTL RH.GCLMSSRIRARYRIL
      : :      : :      : :      : :      : :
gi|170 WVPCLEFETEHG FVYRENNKSPGYDGRYWTMWK-LPMFGCTDATQVLAEEVEA
      10      20      30      40      50

```

```

      1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSCTICACNLLCF
.. : ..... : :
gi|170 KKAYP--QAWIRIIGFDNVRQVCISFIAYKPEGY
      60      70      80

```

```

      1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ

```

```

>>gi|304444|gb|AAA03694.1| ribulose-1,5-bisphosphate car (180 aa)
  initn: 156 init1: 156 opt: 189 Z-score: 233.2 bits: 53.5 E(): 0.00063
Smith-Waterman score: 189; 47.297% identity (63.514% similar) in 74 aa overlap
(1181-1249:108-178)

```

```

      1160      1170      1180      1190      1200
frame2 PYRYELPRYGSRF.KPCYC..CYYDRY.LPRVHG----FDGWSWS.DRTL RH.GCLMSS
      : :      : :      : .      : :      : :
gi|304 EESLMKEVQYLLNNGWVPCLEFEPTHG FVYREHGNTPGYDGRYWTMWK-LPMFGCTDPS
      80      90      100      110      120      130

```

```

      1210      1220      1230      1240      1250      1260
frame2 RIRARYRILRS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSCTI
.. : . . . : ..... : :
gi|304 QVVALEEEAKKAYP--EAFTRIIGFDNVRQVCISFIAYKPASYDA
      140      150      160      170      180

```

```

      1270      1280      1290      1300      1310      1320
frame2 CCACNLLCFLFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYAN

```

```

>>gi|170504|gb|AAA34192.1| ribulose-1,5-bisphosphate car (180 aa)
  initn: 150 init1: 150 opt: 188 Z-score: 231.9 bits: 53.2 E(): 0.00074
Smith-Waterman score: 188; 47.692% identity (70.769% similar) in 65 aa overlap
(1185-1249:119-180)

```

```

      1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYYDRY.LPRVHGFDGWSWS.DRTL RH.GCLMSSRIRARYRIL
      : :      : :      : :      : :      : :
gi|170 YLLKNGWVPCLEFETEHG FVYRENNKSPGYDGRYWTMWK-LPMFGCTDATQVLAEVQEA
      90      100      110      120      130      140

```

```

      1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSCTICACNLLCF
.. : ..... : :
gi|170 KKAYP--QAWVRIIGFDNVRQVCISFIAYKPEGY
      150      160      170      180

```

```

      1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ

```

```

>>gi|170498|gb|AAA34189.1| ribulose-1,5-bisphosphate carb (180 aa)

```

initn: 150 init1: 150 opt: 188 Z-score: 231.9 bits: 53.2 E(): 0.00074  
Smith-Waterman score: 188; 47.692% identity (70.769% similar) in 65 aa overlap  
(1185-1249:119-180)

```
      1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYDRIY.LPRVHGFDGWSWS.DRTLRLH.GCLMSSRIRARYRIL
      .:: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|170 YLLKNGWVPCLEFETEHGFVYRENNKSPGYDGRYWTMWK-LPMFGCTDATQVLAEVQEA
      90      100      110      120      130      140
```

```
      1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSTICACNLLCF
      .. : ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..:
gi|170 KKAYP--QAWVRIIGFDNVRQVCISFIAYKPEGY
      150      160      170      180
```

```
      1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ
```

>>gi|19334|emb|CAA29402.1| ribulose 1,5-bisphosphate car (180 aa)  
initn: 150 init1: 150 opt: 188 Z-score: 231.9 bits: 53.2 E(): 0.00074  
Smith-Waterman score: 188; 47.692% identity (70.769% similar) in 65 aa overlap  
(1185-1249:119-180)

```
      1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYDRIY.LPRVHGFDGWSWS.DRTLRLH.GCLMSSRIRARYRIL
      .:: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|193 YLLKNGWVPCLEFETEHGFVYRENNKSPGYDGRYWTMWK-LPMFGCTDATQVLAEVQEA
      90      100      110      120      130      140
```

```
      1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSTICACNLLCF
      .. : ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..:
gi|193 KKAYP--QAWVRIIGFDNVRQVCISFIAYKPEGY
      150      160      170      180
```

```
      1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ
```

>>gi|19336|emb|CAA29403.1| ribulose 1,5-bisphosphate car (180 aa)  
initn: 150 init1: 150 opt: 188 Z-score: 231.9 bits: 53.2 E(): 0.00074  
Smith-Waterman score: 188; 47.692% identity (70.769% similar) in 65 aa overlap  
(1185-1249:119-180)

```
      1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYDRIY.LPRVHGFDGWSWS.DRTLRLH.GCLMSSRIRARYRIL
      .:: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|193 YLLKNGWVPCLEFETEHGFVYRENNKSPGYDGRYWTMWK-LPMFGCTDATQVLAEVQEA
      90      100      110      120      130      140
```

```
      1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSTICACNLLCF
      .. : ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..:
gi|193 KKAYP--QAWVRIIGFDNVRQVCISFIAYKPEGY
      150      160      170      180
```

```
      1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ
```

>>gi|218318|dbj|BAA01888.1| ribulose 1,5-bisphosphate ca (180 aa)  
initn: 150 init1: 150 opt: 188 Z-score: 231.9 bits: 53.2 E(): 0.00074

Smith-Waterman score: 188; 47.692% identity (70.769% similar) in 65 aa overlap  
(1185-1249:119-180)

```
      1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYDRIY.LPRVHGFDGWSWS.DRTLRLH.GCLMSSRIRARYRIL
      .:: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|218 YLLKNGWVPCLEFETEHGFVYRENNKSPGYDGRYWTMWK-LPMFGCTDATQVLAEVQEA
      90      100      110      120      130      140
```

```
      1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSTICACNLLCF
      .. : ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..:
gi|218 KKAYP--QAWVRIIGFDNVRQVCISFIAYKPEGY
      150      160      170      180
```

```
      1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ
```

>>gi|4456641|emb|CAA29401.2| ribulose 1,5-bisphosphate c (180 aa)  
initn: 150 init1: 150 opt: 188 Z-score: 231.9 bits: 53.2 E(): 0.00074  
Smith-Waterman score: 188; 47.692% identity (70.769% similar) in 65 aa overlap  
(1185-1249:119-180)

```
      1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYDRIY.LPRVHGFDGWSWS.DRTLRLH.GCLMSSRIRARYRIL
      .:: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|445 YLLKNGWVPCLEFETEHGFVYRENNKSPGYDGRYWTMWK-LPMFGCTDATQVLAEVQEA
      90      100      110      120      130      140
```

```
      1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSTICACNLLCF
      .. : ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..:
gi|445 KKAYP--QAWVRIIGFDNVRQVCISFIAYKPEGY
      150      160      170      180
```

```
      1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ
```

>>gi|170500|gb|AAA34190.1| ribulose-1,5-bisphosphate carb (180 aa)  
initn: 150 init1: 150 opt: 188 Z-score: 231.9 bits: 53.2 E(): 0.00074  
Smith-Waterman score: 188; 47.692% identity (70.769% similar) in 65 aa overlap  
(1185-1249:119-180)

```
      1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYDRIY.LPRVHGFDGWSWS.DRTLRLH.GCLMSSRIRARYRIL
      .:: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|170 YLLKNGWVPCLEFETEHGFVYRENNKSPGYDGRYWTMWK-LPMFGCTDATQVLAEVQEA
      90      100      110      120      130      140
```

```
      1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSTICACNLLCF
      .. : ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..: ..:
gi|170 KKAYP--QAWVRIIGFDNVRQVCISFIAYKPEGY
      150      160      170      180
```

```
      1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ
```

>>gi|19338|emb|CAA29404.1| ribulose 1,5-bisphosphate car (180 aa)  
initn: 150 init1: 150 opt: 188 Z-score: 231.9 bits: 53.2 E(): 0.00074  
Smith-Waterman score: 188; 47.692% identity (70.769% similar) in 65 aa overlap  
(1185-1249:119-180)

```

1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYYDRY.LPRVHGFDGWSWS.DRTL RH.GCLMSSRIRARYRIL
      .:: .: .: .: .: .: .: .:
gi|193 YLLKNGWVPCLEFETEHGFVYRENHKSPGYDGRYWTMWK-LPMFGCTDATQVLAEVQEA
      90      100      110      120      130      140

1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVQCISFIAHTPESY.V.VLWHWENCFSCITCCACNLLCF
      .. : ..: .: .: .: .: .: .:
gi|193 KKAYP--QAWVRIIGFDNVRQVQCISFIAYKPEGY
      150      160      170      180

1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ

>>gi|162946541|gb|ABY21255.1| ribulose-1,5-bisphosphate (181 aa)
      initn: 183 initl: 145 opt: 188 Z-score: 231.9 bits: 53.2 E(): 0.00074
Smith-Waterman score: 188; 41.558% identity (66.234% similar) in 77 aa overlap
(1173-1249:110-181)

1150      1160      1170      1180      1190      1200
frame2 SCRYPPSPRYELPRYGSRF.KPCYC..CYYDRY.LPRVHGFDGWSWS.DRTL RH.GCL
      : .: .: .: .: .: .: .:
gi|162 TDEQLLKEVEYLLKNGWVPCLEFETEHGIVYREKHKSPGY--YDGRYWNMWK-LPMFGCT
      80      90      100      110      120      130

1210      1220      1230      1240      1250      1260
frame2 MSSRIRARYRILRS.DPLARAFVRIIGFDNVRQVQCISFIAHTPESY.V.VLWHWENCF
      ....: ..: .: .: .: .: .: .:
gi|162 DATQVLAEVQECKKSYP--QAWIRIIGFDNVRQVQCISFIAYKPEGY
      140      150      160      170      180

1270      1280      1290      1300      1310      1320
frame2 CTICACNLLCFLFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKR

>>gi|162946537|gb|ABY21253.1| ribulose-1,5-bisphosphate (181 aa)
      initn: 183 initl: 145 opt: 188 Z-score: 231.9 bits: 53.2 E(): 0.00074
Smith-Waterman score: 188; 46.154% identity (70.769% similar) in 65 aa overlap
(1185-1249:120-181)

1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYYDRY.LPRVHGFDGWSWS.DRTL RH.GCLMSSRIRARYRIL
      .:: .: .: .: .: .: .: .:
gi|162 YLLKNGWVPCLEFETEHGFVYRENHKSPGYDGRYWTMWK-LPMFGCTDTTQVLAEVQEC
      90      100      110      120      130      140

1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVQCISFIAHTPESY.V.VLWHWENCFSCITCCACNLLCF
      .. : ..: .: .: .: .: .: .:
gi|162 KKSYP--QAWIRIIGFDNVRQVQCISFIAYKPEGY
      150      160      170      180

1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ

>>gi|19706|emb|CAA31994.1| ribulose bisphosphate carboxy (180 aa)
      initn: 150 initl: 150 opt: 187 Z-score: 230.7 bits: 53.0 E(): 0.00086
Smith-Waterman score: 187; 47.692% identity (70.769% similar) in 65 aa overlap
(1185-1249:119-180)

1160      1170      1180      1190      1200      1210

```

```

frame2 ELPRYGSRF.KPCYC..CYYDRY.LPRVHGFDGWSWS.DRTL RH.GCLMSSRIRARYRIL
      .:: .: .: .: .: .: .: .:
gi|197 YLLKSGWVPCLEFETERGFVYREHHHSPGYDGRYWTMWK-LPMFGCTDATQVLAEEVEA
      90      100      110      120      130      140

1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVQCISFIAHTPESY.V.VLWHWENCFSCITCCACNLLCF
      .. : ..: .: .: .: .: .: .:
gi|197 KKAYP--QAWVRIIGFDNVRQVQCISFIAYKPEGY
      150      160      170      180

1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ

>>gi|170318|gb|AAA34110.1| ribulose bisphosphate carboxy (180 aa)
      initn: 150 initl: 150 opt: 187 Z-score: 230.7 bits: 53.0 E(): 0.00086
Smith-Waterman score: 187; 47.692% identity (70.769% similar) in 65 aa overlap
(1185-1249:119-180)

1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYYDRY.LPRVHGFDGWSWS.DRTL RH.GCLMSSRIRARYRIL
      .:: .: .: .: .: .: .: .:
gi|170 YLLKSGWVPCLEFETERGFVYREHHHSPGYDGRYWTMWK-LPMFGCTDATQVLAEEVEA
      90      100      110      120      130      140

1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVQCISFIAHTPESY.V.VLWHWENCFSCITCCACNLLCF
      .. : ..: .: .: .: .: .: .:
gi|170 KKAYP--QAWVRIIGFDNVRQVQCISFIAYKPEGY
      150      160      170      180

1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ

>>gi|4689388|gb|AAD27881.1|AF139469_1 ribulose-1,5-bisph (181 aa)
      initn: 144 initl: 144 opt: 187 Z-score: 230.7 bits: 53.0 E(): 0.00087
Smith-Waterman score: 187; 47.692% identity (66.154% similar) in 65 aa overlap
(1185-1249:120-181)

1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYYDRY.LPRVHGFDGWSWS.DRTL RH.GCLMSSRIRARYRIL
      .:: .: .: .: .: .: .: .:
gi|468 LRSWIPCLFELSHKDAFPHRENNRSPGYDGRYWTMWK-LPMYGCTDSSQVLKELREA
      90      100      110      120      130      140

1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVQCISFIAHTPESY.V.VLWHWENCFSCITCCACNLLCF
      .. : ..: .: .: .: .: .: .:
gi|468 RTLYP--DGFVRIVGFDVSRVQVQCISFIAYKPPGY
      150      160      170      180

1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ

>>gi|169557|gb|AAA33838.1| ribulose bisphosphate carboxy (181 aa)
      initn: 183 initl: 145 opt: 187 Z-score: 230.7 bits: 53.0 E(): 0.00087
Smith-Waterman score: 187; 46.154% identity (70.769% similar) in 65 aa overlap
(1185-1249:120-181)

1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYYDRY.LPRVHGFDGWSWS.DRTL RH.GCLMSSRIRARYRIL
      .:: .: .: .: .: .: .: .:

```

```

gi|169 YLLKNGWVPCLEFETEHGFVYREHNSSPGYYDGRYWTMWK-LPMFGCTDGTQVLAEVQEA
  90      100      110      120      130      140

      1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSCTICACNLLCF
.. : .....
gi|169 KNAYP--QAWIRIIGFDNVRQVCISFIAYKPEGY
  150      160      170      180

      1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ

>>gi|162946539|gb|ABY21254.1| ribulose-1,5-bisphosphate (181 aa)
  initn: 183 init1: 145 opt: 187 Z-score: 230.7 bits: 53.0 E(): 0.00087
Smith-Waterman score: 187; 46.154% identity (70.769% similar) in 65 aa overlap
(1185-1249:120-181)

      1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYDDRY.LPRVHGFDGWSWS.DRTL RH.GCLMSSRIRARYRIL
.. : .....
gi|162 YLLKNGWVPCLEFETEHGFVYRENHKSPGYDGRYWTMWK-LPMFGCTDATQVLAEVQEC
  90      100      110      120      130      140

      1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSCTICACNLLCF
.. : .....
gi|162 KKSYP--QAWIRIIGFDNVRQVCISFIAYKPEGY
  150      160      170      180

      1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ

>>gi|21563|emb|CAA49413.1| ribulose bisphosphate carboxy (181 aa)
  initn: 183 init1: 145 opt: 187 Z-score: 230.7 bits: 53.0 E(): 0.00087
Smith-Waterman score: 187; 46.154% identity (70.769% similar) in 65 aa overlap
(1185-1249:120-181)

      1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYDDRY.LPRVHGFDGWSWS.DRTL RH.GCLMSSRIRARYRIL
.. : .....
gi|215 YLLKNGWVPCLEFETEHGFVYRENHKSPGYDGRYWTMWK-LPMFGCTDATQVLAEVQEC
  90      100      110      120      130      140

      1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSCTICACNLLCF
.. : .....
gi|215 KKSYP--QAWIRIIGFDNVRQVCISFIAYKPEGY
  150      160      170      180

      1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ

>>gi|21571|emb|CAA49417.1| ribulose bisphosphate carboxy (181 aa)
  initn: 183 init1: 145 opt: 187 Z-score: 230.7 bits: 53.0 E(): 0.00087
Smith-Waterman score: 187; 46.154% identity (70.769% similar) in 65 aa overlap
(1185-1249:120-181)

      1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYDDRY.LPRVHGFDGWSWS.DRTL RH.GCLMSSRIRARYRIL
.. : .....
gi|215 YLLKNGWVPCLEFETEHGFVYRENHKSPGYDGRYWTMWK-LPMFGCTDATQVLAEVQEA
  90      100      110      120      130      140

```

```

      1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSCTICACNLLCF
.. : .....
gi|215 KKAYP--QAWIRIIGFDNVRQVCISFIAYKPEGY
  150      160      170      180

      1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ

>>gi|16224234|gb|AAL15646.1|AF411547_1 ribulose-1,5-bisp (100 aa)
  initn: 155 init1: 155 opt: 184 Z-score: 230.3 bits: 52.1 E(): 0.00091
Smith-Waterman score: 184; 49.231% identity (66.154% similar) in 65 aa overlap
(1185-1249:39-100)

      1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYDDRY.LPRVHGFDGWSWS.DRTL RH.GCLMSSRIRARYRIL
.. : .....
gi|162 YLIRKGWIPCLEFELEKGFVYRENNHRSPGYDGRYWTMWR-LPLFGATDSSQVLELADC
  10      20      30      40      50      60

      1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSCTICACNLLCF
.. : .....
gi|162 KAEYP--DSFIRIIGFDNVRQVCISFIAHTPKNY
  70      80      90      100

      1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ

>>gi|21567|emb|CAA49415.1| ribulose bisphosphate carboxy (180 aa)
  initn: 183 init1: 145 opt: 186 Z-score: 229.5 bits: 52.8 E(): 0.001
Smith-Waterman score: 186; 46.154% identity (70.769% similar) in 65 aa overlap
(1185-1249:119-180)

      1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYDDRY.LPRVHGFDGWSWS.DRTL RH.GCLMSSRIRARYRIL
.. : .....
gi|215 YLLKNGWVPCLEFETEHGFVYRENNKSPGYDGRYWTMWK-LPMFGCTDATQVLAEEVEEA
  90      100      110      120      130      140

      1220      1230      1240      1250      1260      1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSCTICACNLLCF
.. : .....
gi|215 KKAYP--QAWIRIIGFDNVRQVCISFIAYKPEGY
  150      160      170      180

      1280      1290      1300      1310      1320      1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ

>>gi|20024|emb|CAA26208.1| small subunit ribulose 1,5-bi (180 aa)
  initn: 145 init1: 145 opt: 186 Z-score: 229.5 bits: 52.8 E(): 0.001
Smith-Waterman score: 186; 46.154% identity (70.769% similar) in 65 aa overlap
(1185-1249:119-180)

      1160      1170      1180      1190      1200      1210
frame2 ELPRYGSRF.KPCYC..CYDDRY.LPRVHGFDGWSWS.DRTL RH.GCLMSSRIRARYRIL
.. : .....
gi|200 YLLKNGWVPCLEFETEHGFVYRENNKSPGYDGRYWTMWK-LPMFGCTDATQVLAEEVEEA
  90      100      110      120      130      140

      1220      1230      1240      1250      1260      1270

```

```

frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSTICACNLLCF
.. : .....
gi|200 KKAYP--QAWIRIIGFDNVRQVCISFIAYKPEGY
150 160 170 180

1280 1290 1300 1310 1320 1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ

>>gi|21569|emb|CAA49416.1| ribulose biphosphate carboxy (180 aa)
initn: 183 initl: 145 opt: 186 Z-score: 229.5 bits: 52.8 E(): 0.001
Smith-Waterman score: 186; 46.154% identity (70.769% similar) in 65 aa overlap
(1185-1249:119-180)

1160 1170 1180 1190 1200 1210
frame2 ELPRYGSRF.KPCYC..CYDDRY.LPRVHGFDGWSWS.DRTL RH.GCLMSSRIRARYRIL
.. : .....
gi|215 YLLKNGWVPCLEFETEHEGFVYRENNKSPGYDGRYWTMWK-LPMFGCTDATQVLAEEVEEA
90 100 110 120 130 140

1220 1230 1240 1250 1260 1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSTICACNLLCF
.. : .....
gi|215 KKAYP--QAWIRIIGFDNVRQVCISFIAYKPEGY
150 160 170 180

1280 1290 1300 1310 1320 1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ

>>gi|21565|emb|CAA49414.1| ribulose bisphosphate carboxy (180 aa)
initn: 183 initl: 145 opt: 186 Z-score: 229.5 bits: 52.8 E(): 0.001
Smith-Waterman score: 186; 46.154% identity (70.769% similar) in 65 aa overlap
(1185-1249:119-180)

1160 1170 1180 1190 1200 1210
frame2 ELPRYGSRF.KPCYC..CYDDRY.LPRVHGFDGWSWS.DRTL RH.GCLMSSRIRARYRIL
.. : .....
gi|215 YLLKNGWVPCLEFETEHEGFVYRENNKSPGYDGRYWTMWK-LPMFGCTDATQVLAEEVEEA
90 100 110 120 130 140

1220 1230 1240 1250 1260 1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSTICACNLLCF
.. : .....
gi|215 KKAYP--QAWIRIIGFDNVRQVCISFIAYKPEGY
150 160 170 180

1280 1290 1300 1310 1320 1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ

>>gi|19760|emb|CAA25862.1| unnamed protein product (Nico) (180 aa)
initn: 145 initl: 145 opt: 186 Z-score: 229.5 bits: 52.8 E(): 0.001
Smith-Waterman score: 186; 46.154% identity (70.769% similar) in 65 aa overlap
(1185-1249:119-180)

1160 1170 1180 1190 1200 1210
frame2 ELPRYGSRF.KPCYC..CYDDRY.LPRVHGFDGWSWS.DRTL RH.GCLMSSRIRARYRIL
.. : .....
gi|197 YLLKNGWVPCLEFETEHEGFVYRENNKSPGYDGRYWTMWK-LPMFGCTDATQVLAEEVEEA
90 100 110 120 130 140

1220 1230 1240 1250 1260 1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSTICACNLLCF
.. : .....

```

```

gi|197 KKAYP--QAWIRIIGFDNVRQVCISFIAYKPEGY
150 160 170 180

1280 1290 1300 1310 1320 1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ

>>gi|295823|emb|CAA37516.1| NySS41 [Nicotiana sylvestris (181 aa)
initn: 145 initl: 145 opt: 186 Z-score: 229.4 bits: 52.8 E(): 0.001
Smith-Waterman score: 186; 46.154% identity (70.769% similar) in 65 aa overlap
(1185-1249:120-181)

1160 1170 1180 1190 1200 1210
frame2 ELPRYGSRF.KPCYC..CYDDRY.LPRVHGFDGWSWS.DRTL RH.GCLMSSRIRARYRIL
.. : .....
gi|295 YLLKNGWVPCLEFETEHEGFVYRENNKSPGYDGRYWTMWK-LPMFGCTDATQVLAEEVEEA
90 100 110 120 130 140

1220 1230 1240 1250 1260 1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSTICACNLLCF
.. : .....
gi|295 KKAYP--QAWIRIIGFDNVRQVCISFIAYKPEGY
150 160 170 180

1280 1290 1300 1310 1320 1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ

>>gi|20958|emb|CAA31774.1| ribulose bisphosphate carboxy (171 aa)
initn: 198 initl: 159 opt: 185 Z-score: 228.5 bits: 52.5 E(): 0.0011
Smith-Waterman score: 185; 47.619% identity (65.079% similar) in 63 aa overlap
(1185-1247:112-171)

1160 1170 1180 1190 1200 1210
frame2 ELPRYGSRF.KPCYC..CYDDRY.LPRVHGFDGWSWS.DRTL RH.GCLMSSRIRARYRIL
.. : .....
gi|209 EYLLRNKWVPCLEFDLEGSISRKYNRSPGYDGRYWMWK-LPMFGCTEASQVINEVREC
90 100 110 120 130 140

1220 1230 1240 1250 1260 1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSTICACNLLCF
.. : .....
gi|209 AKAYP--KAFIRVIGFDNVRQVCISFIVHKPE
150 160 170

1280 1290 1300 1310 1320 1330
frame2 LFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCQ

>>gi|1079740|gb|AAA82071.1| ribulose 1,5-bisphosphate ca (178 aa)
initn: 156 initl: 156 opt: 185 Z-score: 228.3 bits: 52.5 E(): 0.0012
Smith-Waterman score: 185; 46.154% identity (66.154% similar) in 65 aa overlap
(1185-1249:117-178)

1160 1170 1180 1190 1200 1210
frame2 ELPRYGSRF.KPCYC..CYDDRY.LPRVHGFDGWSWS.DRTL RH.GCLMSSRIRARYRIL
.. : .....
gi|107 YLLRKGWIPCLEFELEHGFVYREHHRSPGYDGRYWTMWK-LPMFGCTDASQVLKELQEA
90 100 110 120 130 140

1220 1230 1240 1250 1260 1270
frame2 RS.DPLARAFVRIIGFDNVRQVCISFIAHTPESY.V.VLWHWENCFSTICACNLLCF
.. : .....
gi|107 KTAYP--NAFIRIIGFDNVRQVCISFIAYKPPSF
150 160 170

```

1280 1290 1300 1310 1320 1330  
frame2 LFGRFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKRYANILF.VKMCO

>>gi|13241105|gb|AAK16229.1|AF044397\_1 ribulose-1,5-bisphosphate (133 aa)  
initn: 134 init1: 134 opt: 183 Z-score: 227.4 bits: 52.0 E(): 0.0013  
Smith-Waterman score: 183; 42.857% identity (63.636% similar) in 77 aa overlap  
(1174-1249:62-133)

1150 1160 1170 1180 1190 1200  
frame2 CRYPPRSPYRYELPRYGSRF.KPCYC..CYYDRY.LPRVHGFDGWSWS.DRT-LRH.GCL  
:: :: :: :: :: ::  
gi|132 TQLAKEVDYLPNKNWVPCLEFELEHGFVSPYNRRS-PRY--YDGRSWTIWKLPMFMFGCT  
40 50 60 70 80

1210 1220 1230 1240 1250 1260  
frame2 MSSRIRARYRILRS.DPLARAFVRIIGFDNVRQVQCISFIAHTPESY.V.VLWHWENCFS  
::: . . . . . ::  
gi|132 DSSQVMKELGECKKEYP--QAWIRIIGFDNVRQVQCISFIASKPDGF  
90 100 110 120 130

1270 1280 1290 1300 1310 1320  
frame2 CTICCACNLLCFLFGFRYRTVKWKWMEKS..MIWSFCSFSN.YYLFLLFVVC.I.NYKR

2416 residues in 1 query sequences  
4761287459 residues in 17815538 library sequences  
Scomplib [34t26]  
start: Wed Feb 3 22:29:20 2010 done: Wed Feb 3 22:52:08 2010  
Total Scan time: 1213.360 Total Display time: 2.040

Function used was FASTA [version 3.4t26 July 7, 2006]

Appendix 3. Bioinformatic analysis of polypeptide frame3

>frame3  
HVREIYMDQQ.V.WSIWKRKRSNYQFFNSKM.MSAALL.NESTF..NDKLRSVVFIGESNKQIILIRKSLFRRVYIHWQMGA  
.MRNFTIDAAVEWKILSPKPKQQGGTESPNH.PKATGDQ.RIFNQSKLLFQHMHHGQ.VSEKDIHRRLLKVSGHL.K.SCQHR  
AAGLWGPDKKGMVQNC.AHLPKASLPLLQR.SRFL.YKMGTK.RGKELS.QPTH.CV.RTQ.RPQKN.LELRI.QHSRLGSN  
KVRAISLYSNWYRQNEGTPIILKGL.GRIRYQA.YRKFLS.GRLLVWEDTYGCYNKPFPPLS.DFEVNIYFTCSLFYLVFE.MI  
.MALDNLHSLRYQIFTLEKSVI.IDGPQSMKFLGDEHV.PFA.ITLIKRC.DMMFVTCTSKHS.NTTNPRGIEYSRLNRGIMV  
I.RMILFYVKP.HWFRIQRYK.NHSRC.FHLSFLLTLAATHFSAISLR.ALTYP.IVLFLRLLIVAHIIISMIVGFDALC.LI  
VLKILICCLDFIDC.YQRFTASKR.FIHETIPFF.SILF.DFNLF.DLLVGGSIYE.TC.FAYLRCMSI.IVIG.FWSSITNP  
VFPFSKGNRSRIVCFISLEIADLIEFSSSLSDDRAPQFLW.KSV.LR.RLFTMIVLC.R.FS.WLYMSTVFIDSIFDCSFFCFR  
VDQRSTMAQVSRICNGVQNPISLISNLSKSSQKRSPLSVSLKTQQHPRAYPISSSWGKKSGMTLIGSELRLPKVMSSVSTACM  
LHGASSRPAATARKSSGLSGTVRIPGDKSISHRSMFMGGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVGN  
GGLLAPEAPLDFGNAATGCRLLTMGLVGVYDFDSTF.IGDASLTKRPMGRVNLPLREMGVQVKSSEDGDRPLVTIRGPKTPTPTIY  
RVPMASAQKSAVLLAGLNTPGITTTVIEPIIMTRDHTKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAF  
PLVAALLVPGSDVTILNVLNMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI LAV  
AAAFAGATVMNGLEELRVKESDRLSAVANGKLKNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLV  
SENPTVDDATMIATSFPFEMDLMAGLGAKIELSDTKAA..AQEFELGTGS.DLRIL.LELFSVSSVSTTFVKFNASVSLRTH  
QNPTFEFYYGIGKTVFLVPFVVLVIYCVFYSVFAIEL.NGNGWRRVNE.YGPFVHSQINIICFFSYLLCVEFEIIRDMQTFCF  
E.KCVKSWPLMTEVNMRSKTLVVVPLCLFTRQQIYFQT.KSCKCY.IQVCPILVF.TFMNFPCLNFPESLSDSNHCFIIIVILM  
DL.LSMKIFFNAYFDLPID.QHASIDLQPLEAAQPGGLYGPVRRPRYGR.L.RIPGKNI.YVLFKKKICNNIFIF.YLFLCIF.  
KSIYI.STKIFLYLHLFCIFINFLAFFGIFNNDYSLIINHYSYMVHIVGTI.SVHCI.LCG.CFDPGLHLPLIN.FGNSPY.S  
VTPSSIIINLGSLECHNTD.SLGS.EKAKEQKTKHNESILCIAMSKFIFKQKQRNHTQWTSLIH.LIRIAASRKKNWTPKA  
MHNNTYSQRCQSSSPKHSPTQIMSPHICCF.PNLKLVFSSATSLFLISTPVKLHATPWPNVHAC.QDL.L.ISAISQVFI  
KNRVPSSSLGVN.IVPASRDISCGVHHIC.GKETPEIELCIYISFFISRFPEGLVGVT.LVATISALLSIDKQLAVMPQ.RTI

TTFRAGYFLFYGHDFKTRRSTGEATLEIHVGGNSCYFIIFPCYFTLAGLWWSRQQTWWWACKPWRTKIQICVFWWLEGKGASP  
FEN.WNHSYI.RKLQA...STFASPNFYQPVLVQ.TISV.KWLHVREIYMDQQR.IIIMLEAFAFRNCGSINFRRGLRLC  
LQATRRRRRFGF.SSKVCRPTTKFAACSTRVRRG.SNREK..SSSCCHHNEFLRSLLR.NV.S.NHDHKTKSQHEML.SSSTA  
ALQPIVYQ.REVLIL.LLARCLHLSLQEI.K.KRIYKCIQFREFLFLNKYDEPHMKYPEMQAQFNLPLGRVKECLTFDAIA  
L.CFIHLLISALFFSYSMG.CLIWDFSTREPTLLYHYIVFFETLLFFYLLRLTFECI.RNLLFILFGICLVGFIRFIYII.K  
.LCY.ET..ELTNIRIESRTSCKSIYNPHKFNKLT.LVSKFMFLCLIC.ICTL.GRILQISTITLAAAGPIYI.RSRRRAKS.S  
FSSKPPFGRECRHVEIKISELE.FVYCFRL.IRRIVICRFIKMYFHFITLRTSTFLN.KKIGNYSFFFSILTIIILIADPCRF  
PGHEAIYN

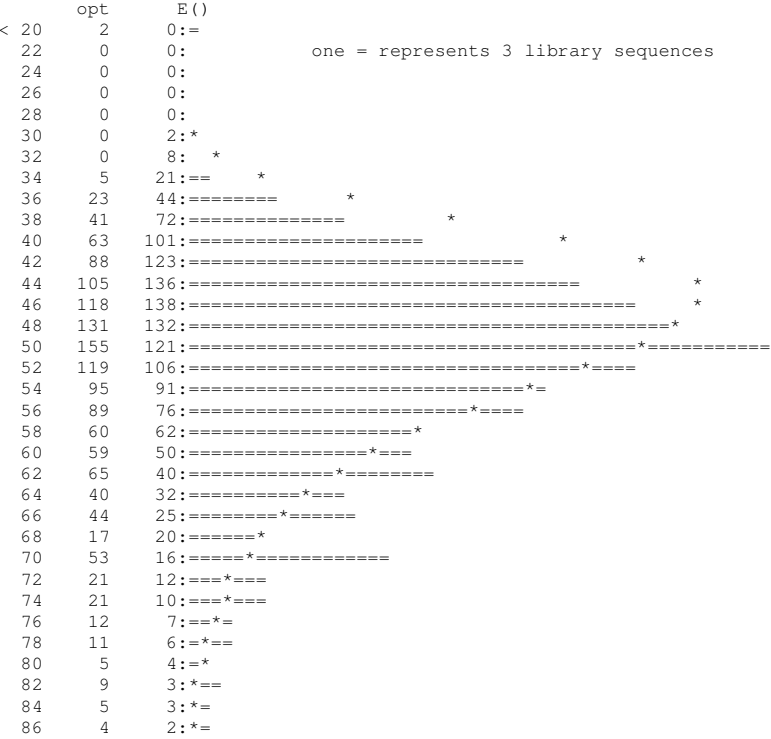
Sliding 8 amino acid window search  
Database searched = AD\_2010  
Query = frame3

Start time: Wed Feb 3 22:52:10 GMT 2010 Finish time: Wed Feb 3 22:52:10 GMT 2010

No 8 amino acid matches exist between frame3 and the AD\_2010 database

# fasta34 frame3.pep /genedata/1/db/AD\_2010 -Q -E 1 -O frame3.pep\_ad.fasta  
FASTA searches a protein or DNA sequence data bank version 3.4t26 July 7, 2006  
Please cite:  
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

frame3, 2415 aa  
vs /genedata/1/db/AD\_2010 library





```

88      0      2:*      inset = represents 1 library sequences
90      1      1:*
92      0      1:*      :*
94      2      1:*      :*
96      5      1:*      :*====
98      2      0:=      *==
100     1      0:=      *==
102     0      0:      *
104     0      0:      *
106     0      0:      *
108     0      0:      *
110     0      0:      *
112     0      0:      *
114     0      0:      *
116     0      0:      *
118     0      0:      *
>120    0      0:      *
331323 residues in 1471 sequences
Expectation_n fit: rho(ln(x))= 5.60460.00345; mu= 9.7017 0.181
mean_var=66.240918.298, 0's: 2 Z-trim: 2 B-trim: 219 in 1/42
Lambda= 0.157584
Kolmogorov-Smirnov statistic: 0.1382 (N=27) at 48

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 44, opt: 32, open/ext: -10/-2, width: 16
!! No sequences with E() < 1.000000

2415 residues in 1 query sequences
331323 residues in 1471 library sequences
Scomplib [34t26]
start: Wed Feb 3 22:52:09 2010 done: Wed Feb 3 22:52:09 2010
Total Scan time: 0.090 Total Display time: 0.000

Function used was FASTA [version 3.4t26 July 7, 2006]

# fasta34 frame3.pep /genedata/1/db/TOX_2010 -Q -E 1 -O frame3.pep_tx.fasta
FASTA searches a protein or DNA sequence data bank version 3.4t26 July 7, 2006
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

frame3, 2415 aa
vs /genedata/1/db/TOX_2010 library

< 20      opt      E()
22      0      0:      one = represents 12 library sequences
24      0      0:
26      0      0:
28      1      2:*
30      8      12:*
32      51      45:====*
34      68      122:===== *
36      340      250:=====*=====
38      515      414:=====*=====
40      487      577:===== *
42      695      706:=====*
44      632      779:===== *
46      719      793:=====*
48      691      759:===== *
50      685      693:=====*
```

```

52 528 609:===== *
54 509 520:=====*
56 328 435:===== *
58 568 357:=====*=====
60 529 289:=====*=====
62 215 232:===== *
64 170 184:=====*
66 156 146:=====*
68 76 115:===== *
70 87 90:=====*
72 48 70:===== *
74 64 55:=====*
76 46 43:=====*
78 28 33:=====*
80 15 26:=====*
82 12 20:=====*
84 10 16:=====*
86 10 12:=====*
88 5 9:*      inset = represents 1 library sequences
90 22 7:*
92 18 6:*      :=====*=====
94 16 4:*      :=====*=====
96 3 3:*      :=====*
98 2 3:*      :=====*
100 16 2:*      :=====*=====
102 2 2:*      :=====*
104 2 1:*      :=====*
106 3 1:*      :=====*
108 1 1:*      :=====*
110 0 1:*      :=====*
112 0 0:      :=====*
114 0 0:      :=====*
116 2 0:=      :=====*
118 0 0:      :=====*
>120 0 0:      :=====*
2069351 residues in 8448 sequences
Expectation_n fit: rho(ln(x))= 3.36670.000663; mu= 23.1849 0.033
mean_var=54.825911.842, 0's: 60 Z-trim: 60 B-trim: 858 in 2/60
Lambda= 0.173213
Kolmogorov-Smirnov statistic: 0.0548 (N=29) at 56

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 44, opt: 32, open/ext: -10/-2, width: 16
!! No sequences with E() < 1.000000

2415 residues in 1 query sequences
2069351 residues in 8448 library sequences
Scomplib [34t26]
start: Wed Feb 3 22:52:10 2010 done: Wed Feb 3 22:52:11 2010
Total Scan time: 0.570 Total Display time: 0.000

Function used was FASTA [version 3.4t26 July 7, 2006]

# fasta34 frame3.pep /genedata/1/db/PRT_2010 -Q -E 1 -O frame3.pep_prt.fasta
FASTA searches a protein or DNA sequence data bank version 3.4t26 July 7, 2006
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

frame3, 2415 aa
vs /genedata/1/db/PRT_2010 library
```

```

      opt      E()
< 20 279142    0:=====
22 139      0:=          one = represents 27372 library sequences
24 335      17:*
26 701      374:*
28 2734     4037:*
30 14452    24526:*
32 65009    94833:====*
34 206171   257176:===== *
36 460681   528179:===== *
38 808906   872883:===== *
40 1181531  1217595:===== *
42 1440483  1488362:===== *
44 1597749  1641802:===== *
46 1642279  1672216:===== *
48 1570088  1600954:===== *
50 1482380  1460877:===== *
52 1282871  1284354:===== *
54 1100982  1097064:===== *
56 919648   916385:===== *
58 746463   752334:===== *
60 622495   609435:===== *
62 496085   488586:===== *
64 398295   388570:===== *
66 319747   307114:===== *
68 255948   241569:===== *
70 192943   189308:===== *
72 154362   147926:===== *
74 121193   115333:===== *
76 94507    89766:===== *
78 73403    69774:===== *
80 57624    54178:===== *
82 44239    41445:===== *
84 35848    32830:===== *
86 26792    25402:===== *
88 21418    19655:===== *
90 16021    15208:===== *
92 13260    11767:===== *
94 10009    9105:===== *
96 8410     7045:===== *
98 6192     5451:===== *
100 5285     4218:===== *
102 10039    3263:===== *
104 4327     2525:===== *
106 2664     1954:===== *
108 2044     1512:===== *
110 1750     1170:===== *
112 1493     905:===== *
114 1048     700:===== *
116 841      542:===== *
118 705      419:===== *
>120 9494    324:===== *
4761287459 residues in 17815538 sequences
statistics sampled from 60000 to 17802074 sequences
Expectation_n fit: rho(ln(x))= 5.12190.000187; mu= 15.4325 0.010
mean_var=72.253214.783, 0's: 975 Z-trim: 1007 B-trim: 2760 in 2/61
Lambda= 0.150885
Kolmogorov-Smirnov statistic: 0.0235 (N=29) at 48

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 44, opt: 32, open/ext: -10/-2, width: 16

```

```

The best scores are:
gi|62318479|dbj|BAD94823.1| 5-enol-pyruvylshikimat ( 527) 2961 655.3 1.2e-184
gi|27549260|gb|AAO17037.1| CP4EPSPS protein [synth ( 455) 2917 645.7 8.3e-182
gi|281079467|gb|ADA36175.1| Sequence 15 from paten ( 455) 2917 645.7 8.3e-182
gi|18266432|gb|AAL67577.1|AF464188_1 CP4EPSPS [Gly ( 455) 2917 645.7 8.3e-182
gi|144974765|gb|ABP12476.1| Sequence 70 from paten ( 455) 2917 645.7 8.3e-182
gi|2484147|gb|AAB72283.1|I49176 Sequence 3 from pa ( 455) 2909 644.0 2.8e-181
gi|5957547|gb|AAE08221.1| Sequence 3 from patent U ( 455) 2909 644.0 2.8e-181
gi|217243929|gb|ACK18296.1| Sequence 1 from patent ( 455) 2909 644.0 2.8e-181
gi|197053766|gb|ACH25464.1| Sequence 16 from paten ( 455) 2909 644.0 2.8e-181
gi|144974726|gb|ABP12437.1| Sequence 3 from patent ( 455) 2909 644.0 2.8e-181
gi|2485224|gb|AAB73360.1|I44449 Sequence 3 from pa ( 455) 2909 644.0 2.8e-181
gi|15073182|emb|CAC41690.1| Putative 3-phosphoshik ( 455) 2685 595.2 1.3e-166
gi|227343087|gb|ACP27305.1| 3-phosphoshikimate 1-c ( 448) 2670 591.9 1.3e-165
gi|150030157|gb|ABR62274.1| 3-phosphoshikimate 1-c ( 456) 2660 589.7 5.8e-165
gi|115254522|emb|CAK05596.1| putative 3-phosphoshi ( 452) 2563 568.6 1.3e-158
gi|240860935|gb|ACS58602.1| 3-phosphoshikimate 1-c ( 452) 2559 567.8 2.4e-158
gi|2485226|gb|AAB73362.1|I44451 Sequence 7 from pa ( 449) 2442 542.3 1.1e-150
gi|2485225|gb|AAB73361.1|I44450 Sequence 5 from pa ( 449) 2442 542.3 1.1e-150
gi|2484148|gb|AAB72284.1|I49177 Sequence 5 from pa ( 449) 2442 542.3 1.1e-150
gi|144974728|gb|ABP12439.1| Sequence 7 from patent ( 449) 2442 542.3 1.1e-150
gi|5957549|gb|AAE08223.1| Sequence 7 from patent U ( 449) 2442 542.3 1.1e-150
gi|144974727|gb|ABP12438.1| Sequence 5 from patent ( 449) 2442 542.3 1.1e-150
gi|5957548|gb|AAE08222.1| Sequence 5 from patent U ( 449) 2442 542.3 1.1e-150
gi|2484149|gb|AAB72285.1|I49178 Sequence 7 from pa ( 449) 2442 542.3 1.1e-150
gi|151559264|gb|ABS12762.1| 3-phosphoshikimate 1-c ( 450) 2441 542.1 1.3e-150
gi|221721755|gb|ACM24911.1| 3-phosphoshikimate 1-c ( 454) 2427 539.0 1.1e-149
gi|264661798|gb|EEZ32059.1| 3-phosphoshikimate 1-c ( 450) 2424 538.4 1.7e-149
gi|23346795|gb|AAN28982.1| 3-phosphoshikimate 1-ca ( 450) 2423 538.2 1.9e-149
gi|260156845|gb|EEW91925.1| 3-phosphoshikimate 1-c ( 450) 2423 538.2 1.9e-149
gi|261745668|gb|EEX33594.1| 3-phosphoshikimate 1-c ( 450) 2423 538.2 1.9e-149
gi|161334826|gb|ABX61131.1| 3-phosphoshikimate 1-c ( 450) 2423 538.2 1.9e-149
gi|163673024|gb|ABY37135.1| 3-phosphoshikimate 1-c ( 450) 2419 537.3 3.5e-149
gi|261303782|gb|EEX07279.1| 3-phosphoshikimate 1-c ( 450) 2419 537.3 3.5e-149
gi|261293812|gb|EEX97308.1| 3-phosphoshikimate 1-c ( 450) 2419 537.3 3.5e-149
gi|260923117|gb|EEX89685.1| 3-phosphoshikimate 1-c ( 450) 2419 537.3 3.5e-149
gi|261295908|gb|EEX99404.1| 3-phosphoshikimate 1-c ( 450) 2419 537.3 3.5e-149
gi|261300126|gb|EEX03623.1| 3-phosphoshikimate 1-c ( 450) 2419 537.3 3.5e-149
gi|260919839|gb|EEX86492.1| 3-phosphoshikimate 1-c ( 450) 2419 537.3 3.5e-149
gi|262551684|gb|EEZ07674.1| 3-phosphoshikimate 1-c ( 450) 2419 537.3 3.5e-149
gi|261739124|gb|EEX27120.1| 3-phosphoshikimate 1-c ( 450) 2419 537.3 3.5e-149
gi|262764953|gb|EEZ10874.1| 3-phosphoshikimate 1-c ( 450) 2419 537.3 3.5e-149
gi|264659952|gb|EEZ30213.1| 3-phosphoshikimate 1-c ( 450) 2419 537.3 3.5e-149
gi|261741103|gb|EEX29029.1| 3-phosphoshikimate 1-c ( 450) 2419 537.3 3.5e-149
gi|225639934|gb|ACN99847.1| 3-phosphoshikimate 1-c ( 450) 2419 537.3 3.5e-149
gi|255998648|gb|ACU47047.1| 3-phosphoshikimate 1-c ( 450) 2418 537.1 4.1e-149
gi|260096414|gb|EEW80290.1| 5-enolpyruvyl shikimat ( 461) 2417 536.9 4.9e-149
gi|263003296|gb|EEZ15589.1| 3-phosphoshikimate 1-c ( 450) 2414 536.2 7.5e-149
gi|260153122|gb|EEW88214.1| 3-phosphoshikimate 1-c ( 450) 2414 536.2 7.5e-149
gi|17983963|gb|AAL53098.1| 3-phosphoshikimate 1-ca ( 480) 2414 536.2 7.9e-149
gi|260675989|gb|EEX62810.1| 5-enolpyruvylshikimate ( 450) 2413 536.0 8.7e-149
gi|82615053|emb|CAJ09979.1| ATP/GTP-binding site m ( 450) 2413 536.0 8.7e-149
gi|189018862|gb|ACD71584.1| 3-phosphoshikimate 1-c ( 450) 2413 536.0 8.7e-149
gi|260671262|gb|EEX58083.1| 5-enolpyruvylshikimate ( 450) 2413 536.0 8.7e-149
gi|260669423|gb|EEX56363.1| 5-enolpyruvylshikimate ( 450) 2413 536.0 8.7e-149
gi|260874433|gb|EEX81502.1| 3-phosphoshikimate 1-c ( 450) 2413 536.0 8.7e-149
gi|148370906|gb|ABQ60885.1| 3-phosphoshikimate 1-c ( 461) 2413 536.0 8.9e-149
gi|260916763|gb|EEX83624.1| 5-enolpyruvylshikimate ( 450) 2406 534.5 2.5e-148
gi|62195146|gb|AAX73446.1| AroA, 3-phosphoshikimat ( 450) 2404 534.0 3.4e-148
gi|13487156|gb|AAK27445.1|AF326475_1 5-enolpyruvyl ( 480) 2389 530.8 3.4e-147
gi|14025099|dbj|BAB51700.1| 3-phosphoshikimate 1-c ( 452) 2298 510.9 3e-141
gi|167285314|gb|ABZ38178.1| Sequence 12116 from pa ( 452) 2298 510.9 3e-141

```

gi|110286919|gb|ABG64978.1| 3-phosphoshikimate 1-c ( 449) 2253 501.1 2.7e-138  
gi|49239274|emb|CAF25595.1| 3-phosphoshikimate 1-c ( 442) 1949 435.0 2.2e-118  
gi|161017021|emb|CAK00579.1| 3-phosphoshikimate 1- ( 442) 1932 431.3 2.8e-117  
gi|240266876|gb|ACS50464.1| 3-phosphoshikimate 1-c ( 442) 1918 428.2 2.4e-116  
gi|49237722|emb|CAF26911.1| 3-phosphoshikimate 1-c ( 442) 1918 428.2 2.4e-116  
gi|120615007|gb|ABM45608.1| 3-phosphoshikimate 1-c ( 441) 1852 413.9 5e-112  
gi|211960244|gb|EEA95441.1| 3-phosphoshikimate 1-c ( 442) 1820 406.9 6.2e-110  
gi|154154557|gb|ABS61774.1| 3-phosphoshikimate 1-c ( 451) 1710 382.9 1e-102  
gi|179344835|gb|ACB80247.1| 3-phosphoshikimate 1-c ( 453) 1681 376.6 8.1e-101  
gi|240008687|gb|ACS39913.1| 3-enolpyruvylshikimate ( 453) 1672 374.7 3.2e-100  
gi|218522663|gb|ACK83248.1| 3-phosphoshikimate 1-c ( 453) 1666 373.4 7.8e-100  
gi|154159145|gb|ABS66361.1| 3-phosphoshikimate 1-c ( 456) 1666 373.4 7.9e-100  
gi|158332444|dbj|BAF89929.1| 3-phosphoshikimate 1- ( 471) 1664 372.9 1.1e-99  
gi|163663156|gb|ABY30523.1| 3-phosphoshikimate 1-c ( 472) 1655 371.0 4.2e-99  
gi|91681127|gb|ABE37429.1| 3-phosphoshikimate 1-ca ( 445) 1653 370.5 5.5e-99  
gi|168198858|gb|ACA20805.1| 3-phosphoshikimate 1-c ( 449) 1642 368.1 2.9e-98  
gi|209871816|gb|ACI91612.1| 3-phosphoshikimate 1-c ( 443) 1639 367.5 4.5e-98  
gi|170658506|gb|ACB27561.1| 3-phosphoshikimate 1-c ( 449) 1639 367.5 4.6e-98  
gi|254268596|emb|CAX24555.1| 3-enolpyruvylshikimat ( 436) 1638 367.3 5.2e-98  
gi|39652766|emb|CAE25505.1| 3-phosphoshikimate 1-c ( 445) 1632 366.0 1.3e-97  
gi|83574549|gb|ABC21100.1| 3-phosphoshikimate 1-ca ( 451) 1630 365.5 1.8e-97  
gi|86570797|gb|ABD05354.1| 3-phosphoshikimate 1-ca ( 445) 1629 365.3 2.1e-97  
gi|192282246|gb|ACE98626.1| 3-phosphoshikimate 1-c ( 445) 1628 365.1 2.4e-97  
gi|115516419|gb|ABJ04403.1| 3-phosphoshikimate 1-c ( 445) 1624 364.2 4.4e-97  
gi|77390011|gb|ABA81195.1| 3-phosphoshikimate 1-ca ( 445) 1620 363.4 8e-97  
gi|143260698|gb|EDD97066.1| hypothetical protein G ( 451) 1619 363.1 9.4e-97  
gi|219951457|gb|ACL61849.1| 3-phosphoshikimate 1-c ( 466) 1614 362.1 2e-96  
gi|126105843|gb|ABN78373.1| 3-phosphoshikimate 1-c ( 445) 1613 361.8 2.3e-96  
gi|119374725|gb|ABL70318.1| 3-phosphoshikimate 1-c ( 443) 1608 360.7 4.9e-96  
gi|167293964|gb|ABZ46828.1| Sequence 20766 from pa ( 430) 1606 360.3 6.4e-96  
gi|91798691|gb|ABE61066.1| 3-phosphoshikimate 1-ca ( 449) 1605 360.1 7.7e-96  
gi|98977055|gb|ABF53206.1| 3-phosphoshikimate 1-ca ( 446) 1601 359.2 1.4e-95  
gi|27348987|dbj|BAC46003.1| aroA [Bradyrhizobium j ( 469) 1599 358.8 2e-95  
gi|148500563|gb|ABQ68817.1| 3-phosphoshikimate 1-c ( 479) 1599 358.8 2e-95  
gi|221162652|gb|ACM03623.1| 3-phosphoshikimate 1-c ( 436) 1597 358.3 2.5e-95  
gi|90103938|gb|ABD85975.1| 3-phosphoshikimate 1-ca ( 467) 1595 357.9 3.6e-95  
gi|144898829|emb|CAM75693.1| 3-phosphoshikimate 1- ( 446) 1591 357.0 6.4e-95  
gi|74419278|gb|ABA03477.1| 3-phosphoshikimate 1-ca ( 449) 1591 357.0 6.4e-95  
gi|145558123|gb|ABP72735.1| 3-phosphoshikimate 1-c ( 445) 1590 356.8 7.4e-95  
gi|87135028|gb|ABD25770.1| 3-phosphoshikimate 1-ca ( 441) 1589 356.6 8.5e-95  
gi|209959239|gb|ACI99875.1| 3-phosphoshikimate 1-c ( 448) 1589 356.6 8.6e-95  
gi|258553029|gb|ACV75975.1| 3-phosphoshikimate 1-c ( 453) 1585 355.7 1.6e-94  
gi|56544266|gb|AAV90420.1| 5-enolpyruvylshikimate- ( 453) 1584 355.5 1.8e-94  
gi|88862748|gb|ABD53625.1| 3-phosphoshikimate 1-ca ( 450) 1583 355.3 2.1e-94  
gi|260421386|gb|EEX14637.1| 3-phosphoshikimate 1-c ( 442) 1582 355.1 2.5e-94  
gi|167283674|gb|ABZ36538.1| Sequence 10476 from pa ( 432) 1581 354.9 2.8e-94  
gi|109457166|gb|ABG33371.1| 3-phosphoshikimate 1-c ( 450) 1580 354.6 3.4e-94  
gi|135091817|gb|EBF43950.1| hypothetical protein G ( 451) 1575 353.6 7.2e-94  
gi|146190077|emb|CAL74069.1| 3-phosphoshikimate 1- ( 445) 1573 353.1 9.6e-94  
gi|82948203|dbj|BAE53067.1| 5-enolpyruvylshikimate ( 446) 1573 353.1 9.6e-94  
gi|214030157|gb|EEB70992.1| 3-phosphoshikimate 1-c ( 449) 1572 352.9 1.1e-93  
gi|146403899|gb|ABQ32405.1| 3-phosphoshikimate 1-c ( 445) 1569 352.3 1.8e-93  
gi|142198284|gb|ECW32705.1| hypothetical protein G ( 456) 1565 351.4 3.3e-93  
gi|255105080|gb|EET47754.1| 3-phosphoshikimate 1-c ( 496) 1565 351.4 3.5e-93  
gi|217502178|gb|ACK49587.1| 3-phosphoshikimate 1-c ( 443) 1564 351.2 3.7e-93  
gi|56680491|gb|AAV97156.1| 3-phosphoshikimate 1-ca ( 448) 1564 351.2 3.7e-93  
gi|99036320|gb|ABF62932.1| 3-phosphoshikimate 1-ca ( 450) 1562 350.7 5.1e-93  
gi|259348589|gb|EEW60351.1| 3-phosphoshikimate 1-c ( 450) 1561 350.5 5.9e-93  
gi|167281012|gb|ABZ33876.1| Sequence 7814 from pat ( 418) 1549 347.9 3.4e-92  
gi|260414998|gb|EEX08257.1| 3-phosphoshikimate 1-c ( 448) 1548 347.7 4.2e-92  
gi|256635297|dbj|BAI01266.1| 3-phosphoshikimate 1- ( 466) 1538 345.5 2e-91  
gi|256650569|dbj|BAI16503.1| 3-phosphoshikimate 1- ( 466) 1538 345.5 2e-91

gi|256653560|dbj|BAI19487.1| 3-phosphoshikimate 1- ( 466) 1538 345.5 2e-91  
gi|256641406|dbj|BAI07361.1| 3-phosphoshikimate 1- ( 466) 1538 345.5 2e-91  
gi|256632240|dbj|BAH98215.1| 3-phosphoshikimate 1- ( 466) 1538 345.5 2e-91  
gi|256644461|dbj|BAI10409.1| 3-phosphoshikimate 1- ( 466) 1538 345.5 2e-91  
gi|256638352|dbj|BAI04314.1| 3-phosphoshikimate 1- ( 466) 1538 345.5 2e-91  
gi|256647516|dbj|BAI13457.1| 3-phosphoshikimate 1- ( 466) 1538 345.5 2e-91  
gi|221536211|gb|EEE39199.1| 3-phosphoshikimate 1-c ( 448) 1537 345.3 2.2e-91  
gi|157911331|gb|ABY92764.1| 3-phosphoshikimate 1-c ( 450) 1536 345.1 2.6e-91  
gi|206686690|gb|EDZ47172.1| 3-phosphoshikimate 1-c ( 448) 1534 344.6 3.5e-91  
gi|144125505|gb|EDJ04872.1| hypothetical protein G ( 515) 1534 344.7 3.9e-91  
gi|198253480|gb|EDY77794.1| 3-phosphoshikimate 1-c ( 469) 1530 343.8 6.6e-91  
gi|206679873|gb|EDZ44360.1| 3-phosphoshikimate 1-c ( 441) 1525 342.7 1.3e-90  
gi|198269136|gb|EDY93406.1| 3-phosphoshikimate 1-c ( 450) 1521 341.8 2.5e-90  
gi|124045586|gb|EEB86224.1| 3-phosphoshikimate 1-c ( 450) 1520 341.6 2.9e-90  
gi|58001482|gb|AAW60376.1| 3-phosphoshikimate 1-ca ( 442) 1515 340.5 6e-90  
gi|146400733|gb|ABQ29260.1| 3-phosphoshikimate 1-c ( 445) 1514 340.3 7e-90  
gi|136727461|gb|EBQ00749.1| hypothetical protein G ( 459) 1498 336.8 8.1e-89  
gi|167291237|gb|ABZ44101.1| Sequence 18039 from pa ( 420) 1494 335.9 1.4e-88  
gi|254040135|gb|ACT56931.1| 3-phosphoshikimate 1-c ( 449) 1479 332.7 1.4e-87  
gi|13425333|gb|AAK25551.1| 3-phosphoshikimate 1-ca ( 443) 1477 332.2 1.9e-87  
gi|167290272|gb|ABZ43136.1| Sequence 17074 from pa ( 443) 1477 332.2 1.9e-87  
gi|220965813|gb|ACI97169.1| 3-phosphoshikimate 1-c ( 443) 1477 332.2 1.9e-87  
gi|144160862|gb|EDJ30885.1| hypothetical protein G ( 457) 1476 332.0 2.2e-87  
gi|136767257|gb|ABQ27100.1| hypothetical protein G ( 439) 1475 331.8 2.5e-87  
gi|196187174|gb|EDX82150.1| 3-phosphoshikimate 1-c ( 436) 1465 329.6 1.1e-86  
gi|144127406|gb|EDJ06257.1| hypothetical protein G ( 412) 1462 328.9 1.7e-86  
gi|134334509|gb|EBB71436.1| hypothetical protein G ( 478) 1461 328.8 2.2e-86  
gi|142444644|gb|ECY12567.1| hypothetical protein G ( 459) 1458 328.1 3.4e-86  
gi|114316863|gb|ABI62923.1| 3-phosphoshikimate 1-c ( 454) 1456 327.7 4.5e-86  
gi|167346504|gb|ABZ69239.1| 3-phosphoshikimate 1-c ( 439) 1453 327.0 6.9e-86  
gi|161787807|emb|CAP57405.1| 3-phosphoshikimate 1- ( 449) 1452 326.8 8.2e-86  
gi|209532709|gb|ACI52646.1| 3-phosphoshikimate 1-c ( 449) 1452 326.8 8.2e-86  
gi|142778383|gb|EDA51368.1| hypothetical protein G ( 450) 1449 326.1 1.3e-85  
gi|143036643|gb|EDC35022.1| hypothetical protein G ( 423) 1447 325.7 1.7e-85  
gi|142210143|gb|ECW41740.1| hypothetical protein G ( 492) 1446 325.5 2.2e-85  
gi|143821962|gb|EDG87373.1| hypothetical protein G ( 451) 1444 325.0 2.8e-85  
gi|196477074|gb|ACG76602.1| 3-phosphoshikimate 1-c ( 452) 1439 324.0 5.9e-85  
gi|142140539|gb|ECY88729.1| hypothetical protein G ( 672) 1429 321.9 3.6e-84  
gi|114339116|gb|ABI64396.1| 3-phosphoshikimate 1-c ( 447) 1421 320.0 8.8e-84  
gi|143831861|gb|EDG94516.1| hypothetical protein G ( 451) 1414 318.5 2.5e-83  
gi|136701932|gb|EBP83950.1| hypothetical protein G ( 463) 1410 317.7 4.7e-83  
gi|136594870|gb|EBP17072.1| hypothetical protein G ( 452) 1404 316.3 1.2e-82  
gi|114740446|gb|ABI78571.1| 3-phosphoshikimate 1-c ( 439) 1376 310.2 7.7e-81  
gi|142227889|gb|ECW55072.1| hypothetical protein G ( 420) 1369 308.7 2.1e-80  
gi|135799022|gb|EBJ91601.1| hypothetical protein G ( 412) 1362 307.2 6.1e-80  
gi|136155475|gb|EBM26016.1| hypothetical protein G ( 373) 1355 305.6 1.6e-79  
gi|136746874|gb|EBQ13460.1| hypothetical protein G ( 391) 1354 305.4 1.9e-79  
gi|136765878|gb|EBQ26143.1| hypothetical protein G ( 423) 1351 304.8 3.3e-79  
gi|142122134|gb|ECY75183.1| hypothetical protein G ( 406) 1339 302.2 1.9e-78  
gi|135334376|gb|EBG94098.1| hypothetical protein G ( 446) 1331 300.4 6.9e-78  
gi|143470309|gb|EDF15730.1| hypothetical protein G ( 362) 1325 299.1 1.5e-77  
gi|143106897|gb|EDC86357.1| hypothetical protein G ( 449) 1326 299.4 1.5e-77  
gi|120323802|gb|ABM18117.1| 3-phosphoshikimate 1-c ( 742) 1320 298.2 5.4e-77  
gi|117607246|gb|ABK42701.1| 3-phosphoshikimate 1-c ( 445) 1312 296.3 1.2e-76  
gi|134806831|gb|EBP56326.1| hypothetical protein G ( 435) 1311 296.1 1.4e-76  
gi|143147782|gb|EDD16387.1| hypothetical protein G ( 416) 1308 295.4 2.1e-76  
gi|134969061|gb|EBE63039.1| hypothetical protein G ( 421) 1300 293.7 7.1e-76  
gi|198263265|gb|EDY87543.1| 3-phosphoshikimate 1-c ( 442) 1297 293.0 1.2e-75  
gi|11426484|gb|ABT56283.1| 3-phosphoshikimate 1-c ( 446) 1295 292.6 1.6e-75  
gi|144019737|gb|EDI28838.1| hypothetical protein G ( 748) 1296 293.0 2e-75  
gi|144221151|gb|EDJ74707.1| hypothetical protein G ( 439) 1291 291.7 2.9e-75  
gi|77384298|gb|ABA75811.1| putative bifunctional p ( 735) 1292 292.1 3.7e-75

gi|150959179|gb|ABR81204.1| prephenate dehydrogena ( 746) 1292 292.1 3.7e-75  
gi|143774157|gb|EDG71670.1| hypothetical protein G ( 477) 1288 291.1 4.8e-75  
gi|142247445|gb|ECW69575.1| hypothetical protein G ( 400) 1283 290.0 8.9e-75  
gi|136201992|gb|EBM56643.1| hypothetical protein G ( 477) 1283 290.0 1e-74  
gi|228385087|gb|ACQ29040.1| Sequence 21554 from pa ( 782) 1283 290.2 1.5e-74  
gi|84787408|gb|ABCG63590.1| 5-enolpyruvylshikimate- ( 463) 1280 289.4 1.6e-74  
gi|218770867|emb|CAW26632.1| still frameshift 3-PH ( 746) 1282 289.9 1.7e-74  
gi|167283172|gb|ABZ36036.1| Sequence 9974 from pat ( 411) 1278 288.9 1.9e-74  
gi|237685422|gb|ACR12686.1| response regulator rec ( 742) 1281 289.7 1.9e-74  
gi|126195230|gb|EAZ59293.1| 3-phosphoshikimate 1-c ( 746) 1281 289.7 1.9e-74  
gi|145575084|gb|ABP84616.1| 3-phosphoshikimate 1-c ( 746) 1279 289.3 2.6e-74  
gi|254043237|gb|ACT60032.1| 3-phosphoshikimate 1-c ( 438) 1276 288.5 2.7e-74  
gi|262315671|gb|EEY96710.1| 3-phosphoshikimate 1-c ( 749) 1278 289.1 3.1e-74  
gi|260406041|gb|EEW95927.1| 5-enolpyruvylshikimate ( 756) 1278 289.1 3.1e-74  
gi|115586374|gb|ABJ12389.1| EPSP synthase/prephena ( 746) 1277 288.9 3.6e-74  
gi|135021201|gb|EBE98359.1| hypothetical protein G ( 431) 1273 287.8 4.3e-74  
gi|143034105|gb|EDC33172.1| hypothetical protein G ( 433) 1273 287.8 4.3e-74  
gi|148572136|gb|ABQ94195.1| 3-phosphoshikimate 1-c ( 770) 1275 288.4 4.9e-74  
gi|169148265|emb|CAM86130.1| bifunctional protein ( 748) 1274 288.2 5.6e-74  
gi|262309719|gb|EEY90849.1| 3-phosphoshikimate 1-c ( 748) 1274 288.2 5.6e-74  
gi|169151779|emb|CAP00598.1| bifunctional protein ( 748) 1274 288.2 5.6e-74  
gi|193077824|gb|ABO12699.2| hypothetical protein A ( 756) 1274 288.2 5.7e-74  
gi|213057836|gb|ACJ42738.1| hypothetical protein A ( 756) 1274 288.2 5.7e-74  
gi|262258779|gb|EEY77512.1| 3-phosphoshikimate 1-c ( 756) 1274 288.2 5.7e-74  
gi|213985849|gb|ACJ56148.1| 3-phosphoshikimate 1-c ( 756) 1274 288.2 5.7e-74  
gi|260410993|gb|EEY40290.1| 3-phosphoshikimate 1-c ( 756) 1274 288.2 5.7e-74  
gi|261836400|gb|ACX96167.1| 3-phosphoshikimate 1-c ( 456) 1271 287.4 6e-74  
gi|183210388|gb|ACC57786.1| 5-enolpyruvylshikimate ( 756) 1273 288.0 6.6e-74  
gi|143454605|gb|EDF05664.1| hypothetical protein G ( 437) 1270 287.2 6.8e-74  
gi|78363826|gb|ABBA1791.1| 3-phosphoshikimate 1-ca ( 445) 1270 287.2 6.9e-74  
gi|160878077|gb|ABX52003.1| G6 [synthetic construc ( 507) 1270 287.2 7.6e-74  
gi|53758016|gb|AAU92307.1| 3-phosphoshikimate 1-ca ( 422) 1269 286.9 7.7e-74  
gi|148513208|gb|ABQ80068.1| 3-phosphoshikimate 1-c ( 746) 1270 287.3 1e-73  
gi|219678890|gb|EED35239.1| 3-phosphoshikimate 1-c ( 433) 1267 286.5 1.1e-73  
gi|169758566|gb|ACA71882.1| 3-phosphoshikimate 1-c ( 746) 1268 286.9 1.4e-73  
gi|33703074|gb|AAQ27459.1| Sequence 4521 from pate ( 758) 1268 286.9 1.4e-73  
gi|68345959|gb|AA93565.1| prephenate dehydrogenas ( 741) 1267 286.7 1.6e-73  
gi|142060811|gb|ECY26044.1| hypothetical protein G ( 437) 1263 285.6 1.9e-73  
gi|166858861|gb|ABY97268.1| 3-phosphoshikimate 1-c ( 746) 1265 286.2 2.2e-73  
gi|95109658|emb|CAK14359.1| putative prephenate de ( 746) 1264 286.0 2.5e-73  
gi|24983266|gb|AAN67390.1| AE016365\_5 prephenate de ( 746) 1264 286.0 2.5e-73  
gi|136410014|gb|EBN98241.1| hypothetical protein G ( 445) 1261 285.2 2.7e-73  
gi|2622312455|gb|EEY93540.1| conserved hypothetical ( 749) 1259 284.9 5.4e-73  
gi|91797377|gb|ABE59516.1| 3-phosphoshikimate 1-ca ( 750) 1259 284.9 5.4e-73  
gi|219996023|gb|ACL72625.1| 3-phosphoshikimate 1-c ( 441) 1256 284.1 5.6e-73  
gi|143443468|gb|EDE98739.1| hypothetical protein G ( 422) 1255 283.9 6.3e-73  
gi|143957288|gb|EDH84919.1| hypothetical protein G ( 499) 1250 282.8 1.5e-72  
gi|121588771|gb|ABM61351.1| 3-phosphoshikimate 1-c ( 444) 1246 281.9 2.6e-72  
gi|135102048|gb|EBF50494.1| hypothetical protein G ( 383) 1244 281.5 3.1e-72  
gi|142617034|gb|ECZ35624.1| hypothetical protein G ( 402) 1243 281.3 3.7e-72  
gi|89951391|gb|ABD81406.1| prephenate dehydrogenas ( 745) 1246 282.1 3.8e-72  
gi|262299019|gb|EEY86932.1| 3-phosphoshikimate 1-c ( 748) 1246 282.1 3.8e-72  
gi|145570888|gb|ABP79994.1| 3-phosphoshikimate 1-c ( 706) 1244 281.7 5e-72  
gi|49531306|emb|CAG69018.1| bifunctional protein [ ( 748) 1244 281.7 5.2e-72  
gi|142461568|gb|ECY24891.1| hypothetical protein G ( 433) 1238 280.2 8.4e-72  
gi|229361034|emb|CAY47896.1| putative bifunctional ( 748) 1240 280.8 9.5e-72  
gi|142142726|gb|ECV90367.1| hypothetical protein G ( 524) 1238 280.3 9.7e-72  
gi|83635703|gb|ABC31670.1| 3-phosphoshikimate 1-ca ( 738) 1239 280.6 1.1e-71  
gi|219677398|gb|EED33763.1| 3-phosphoshikimate 1-c ( 436) 1236 279.8 1.1e-71  
gi|194349115|gb|ACF52238.1| 3-phosphoshikimate 1-c ( 435) 1235 279.5 1.3e-71  
gi|142215066|gb|ECW45518.1| hypothetical protein G ( 602) 1236 279.9 1.5e-71  
gi|142454573|gb|ECY19775.1| hypothetical protein G ( 736) 1237 280.1 1.5e-71

gi|190012912|emb|CAQ46544.1| putative 3-phosphoshi ( 435) 1232 278.9 2.1e-71  
gi|224463889|gb|EEF80156.1| 3-phosphoshikimate 1-c ( 439) 1230 278.5 2.8e-71  
gi|50404063|gb|AAT76791.1| 5-enolpyruvylshikimate- ( 449) 1227 277.8 4.5e-71  
gi|219718621|gb|EED37146.1| 3-phosphoshikimate 1-c ( 435) 1224 277.1 7e-71  
gi|143927391|gb|EDH63433.1| hypothetical protein G ( 433) 1219 276.1 1.5e-70  
gi|167288581|gb|ABZ41445.1| Sequence 15383 from pa ( 442) 1219 276.1 1.5e-70  
gi|167733824|emb|CAP52030.1| 3-phosphoshikimate 1- ( 438) 1218 275.8 1.7e-70  
gi|142716963|gb|EDA06672.1| hypothetical protein G ( 467) 1218 275.9 1.8e-70  
gi|136454660|gb|EBO20707.1| hypothetical protein G ( 440) 1216 275.4 2.4e-70  
gi|182632076|gb|ACB92852.1| 3-phosphoshikimate 1-c ( 442) 1216 275.4 2.4e-70  
gi|28057346|gb|AAO29203.1| 3-phosphoshikimate 1-ca ( 454) 1216 275.4 2.4e-70  
gi|110647719|emb|CAL17195.1| prephenate dehydrogen ( 760) 1216 275.6 3.6e-70  
gi|167965401|gb|ACA12411.1| 3-phosphoshikimate 1-c ( 442) 1208 273.7 7.9e-70  
gi|126168057|gb|EAE53568.1| 3-phosphoshikimate 1-c ( 747) 1210 274.3 8.8e-70  
gi|21112672|gb|AAM40886.1| 3-phosphoshikimate 1-ca ( 438) 1204 272.8 1.4e-69  
gi|66574282|gb|AA949692.1| 3-phosphoshikimate 1-ca ( 438) 1204 272.8 1.4e-69  
gi|167290928|gb|ABZ43792.1| Sequence 17730 from pa ( 454) 1198 271.5 3.6e-69  
gi|9107493|gb|AAF85123.1| AE004043\_7 3-phosphoshiki ( 454) 1198 271.5 3.6e-69  
gi|198247672|gb|ACB83265.1| 3-phosphoshikimate 1-c ( 433) 1196 271.0 4.8e-69  
gi|218517584|gb|ACK78170.1| 3-phosphoshikimate 1-c ( 433) 1196 271.0 4.8e-69  
gi|78035677|emb|CAJ23368.1| 3-phosphoshikimate 1-c ( 440) 1196 271.1 4.8e-69  
gi|21107840|gb|AAM36518.1| 3-phosphoshikimate 1-ca ( 440) 1194 270.6 6.5e-69  
gi|42681983|gb|AAS28448.1| Sequence 3590 from pate ( 773) 1196 271.2 7.4e-69  
gi|167282500|gb|AAM40886.1| Sequence 9302 from pat ( 408) 1191 269.9 1.6e-69  
gi|167282697|gb|ABZ35561.1| Sequence 9499 from pat ( 412) 1191 269.9 9.7e-69  
gi|222440562|gb|EEE47241.1| EPSP synthase (3-phosp ( 289) 1189 269.4 1e-68  
gi|167289333|gb|ABZ42197.1| Sequence 16135 from pa ( 408) 1190 269.7 1.1e-68  
gi|167288949|gb|ABZ41813.1| Sequence 15751 from pa ( 408) 1190 269.7 1.1e-68  
gi|148280723|gb|ABQ54811.1| 3-phosphoshikimate 1-c ( 433) 1189 269.5 1.4e-68  
gi|218306685|emb|CAU98758.1| unnamed protein produ ( 435) 1189 269.5 1.4e-68  
gi|52628760|gb|AAU27501.1| 3-phosphoshikimate 1-ca ( 433) 1187 269.1 1.8e-68  
gi|142559088|gb|ECY95157.1| hypothetical protein G ( 431) 1185 268.7 2.5e-68  
gi|92393712|gb|ABE74987.1| prephenate dehydrogenas ( 780) 1188 269.5 2.5e-68  
gi|143358135|gb|EDE53806.1| hypothetical protein G ( 467) 1185 268.7 2.6e-68  
gi|84367863|dbj|BAE69021.1| 3-phosphoshikimate 1-c ( 440) 1184 268.4 2.9e-68  
gi|188520832|gb|ACD58777.1| 3-phosphoshikimate 1-c ( 440) 1184 268.4 2.9e-68  
gi|58426603|gb|AAW75640.1| 3-phosphoshikimate 1-ca ( 440) 1184 268.4 2.9e-68  
gi|214037923|gb|EEB78587.1| 3-phosphoshikimate 1-c ( 434) 1182 268.0 3.9e-68  
gi|53754137|emb|CAH15610.1| 3-phosphoshikimate 1-c ( 433) 1180 267.6 5.3e-68  
gi|53751114|emb|CAH12525.1| 3-phosphoshikimate 1-c ( 433) 1176 266.7 9.7e-68  
gi|56686249|dbj|BAD79471.1| 3-phosphoshikimate 1-c ( 448) 1165 264.3 5.2e-67  
gi|81167924|gb|ABB56264.1| 3-phosphoshikimate 1-ca ( 448) 1165 264.3 5.2e-67  
gi|142228751|gb|ECW55704.1| hypothetical protein G ( 363) 1162 263.6 7e-67  
gi|71038727|gb|AAZ19035.1| 3-phosphoshikimate 1-ca ( 778) 1163 264.1 1.1e-66  
gi|147850559|emb|CAK28053.1| 3-phosphoshikimate 1- ( 437) 1160 263.2 1.1e-66  
gi|166086651|dbj|BAG01359.1| 3-phosphoshikimate 1- ( 446) 1160 263.2 1.1e-66  
gi|147847907|emb|CAK23458.1| 3-phosphoshikimate 1- ( 439) 1157 262.6 1.7e-66  
gi|169884470|gb|ACA98183.1| 3-phosphoshikimate 1-c ( 446) 1157 262.6 1.7e-66  
gi|159028618|emb|CAO90621.1| aroA [Microcystis aer ( 446) 1157 262.6 1.7e-66  
gi|144149402|gb|EDJ22378.1| hypothetical protein G ( 451) 1155 262.1 2.4e-66  
gi|78197121|gb|ABB34886.1| 3-phosphoshikimate 1-ca ( 441) 1154 261.9 2.7e-66  
gi|218173786|gb|ACK72519.1| 3-phosphoshikimate 1-c ( 448) 1154 261.9 2.8e-66  
gi|142376209|gb|ECX62405.1| hypothetical protein G ( 456) 1154 261.9 2.8e-66  
gi|143267007|gb|EDE01582.1| hypothetical protein G ( 382) 1150 261.0 4.5e-66  
gi|144974763|gb|ABP12474.1| Sequence 67 from paten ( 447) 1149 260.8 5.9e-66  
gi|515936|emb|CAA53074.1| 3-phosphoshikimate 1-car ( 447) 1149 260.8 5.9e-66  
gi|167275962|gb|ABZ28826.1| Sequence 2764 from pat ( 447) 1149 260.8 5.9e-66  
gi|5957584|gb|AAE08258.1| Sequence 67 from patent ( 447) 1149 260.8 5.9e-66  
gi|2484184|gb|AAB72320.1| I49213 Sequence 67 from p ( 447) 1149 260.8 5.9e-66  
gi|1653564|dbj|BAA18477.1| 3-phosphoshikimate 1-ca ( 447) 1149 260.8 5.9e-66  
gi|2485261|gb|AAB73397.1| I44486 Sequence 67 from p ( 447) 1149 260.8 5.9e-66  
gi|219993190|gb|ACI69793.1| 3-phosphoshikimate 1-c ( 430) 1142 259.3 1.6e-65

gi|135895733|gb|EBK54414.1| hypothetical protein G ( 435) 1142 259.3 1.6e-65  
gi|78169193|gb|ABB26290.1| 3-phosphoshikimate 1-ca ( 437) 1142 259.3 1.7e-65  
gi|167287132|gb|ABZ39996.1| Sequence 13934 from pa ( 675) 1143 259.7 2e-65  
gi|260414082|gb|EEX07378.1| 3-phosphoshikimate 1-c ( 441) 1139 258.6 2.6e-65  
gi|86557342|gb|ABDO02299.1| 3-phosphoshikimate 1-ca ( 434) 1138 258.4 3e-65  
gi|171698150|gb|ACB51131.1| 3-phosphoshikimate 1-c ( 451) 1136 258.0 4.2e-65  
gi|269468607|gb|EE280251.1| 5-enolpyruvylshikimate ( 434) 1134 257.6 5.5e-65  
gi|135056108|gb|EBF21159.1| hypothetical protein G ( 458) 1133 257.4 6.7e-65  
gi|86555472|gb|ABDO0430.1| 3-phosphoshikimate 1-ca ( 444) 1131 256.9 8.8e-65  
gi|144061985|gb|EDI59285.1| hypothetical protein G ( 451) 1131 256.9 8.9e-65  
gi|143772965|gb|EDG71093.1| hypothetical protein G ( 441) 1130 256.7 1e-64  
gi|135831629|gb|EBK12121.1| hypothetical protein G ( 339) 1128 256.2 1.1e-64  
gi|146274052|dbj|BAF59801.1| 5-enolpyruvylshikimat ( 431) 1129 256.5 1.2e-64  
gi|225645909|gb|ACO04095.1| 3-phosphoshikimate 1-c ( 432) 1127 256.0 1.6e-64  
gi|256589111|gb|ACU99997.1| 3-phosphoshikimate 1-c ( 448) 1127 256.0 1.6e-64  
gi|225643623|gb|ACN98673.1| 3-phosphoshikimate 1-c ( 431) 1125 255.6 2.1e-64  
gi|158305702|gb|ABW27319.1| 3-phosphoshikimate 1-c ( 446) 1123 255.2 3e-64  
gi|218166375|gb|ACK65112.1| 3-phosphoshikimate 1-c ( 448) 1123 255.2 3e-64  
gi|188931025|gb|ACD65655.1| 3-phosphoshikimate 1-c ( 431) 1122 254.9 3.3e-64  
gi|123963875|gb|ABM78631.1| EPSP synthase (3-phosp ( 441) 1121 254.7 4e-64  
gi|142682488|gb|ECZ81982.1| hypothetical protein G ( 441) 1121 254.7 4e-64  
gi|142815727|gb|EDA79221.1| hypothetical protein G ( 419) 1119 254.3 5.1e-64  
gi|134051714|gb|ABO49685.1| 3-phosphoshikimate 1-c ( 429) 1119 254.3 5.2e-64  
gi|167592737|gb|ABZ84485.1| 3-phosphoshikimate 1-c ( 424) 1117 253.8 7e-64  
gi|33634587|emb|CAE20573.1| EPSP synthase (3-phosp ( 441) 1116 253.6 8.4e-64  
gi|113881454|gb|ABI46412.1| 3-phosphoshikimate 1-c ( 439) 1115 253.4 9.8e-64  
gi|212013111|gb|ACJ20491.1| 3-phosphoshikimate 1-c ( 438) 1114 253.2 1.1e-63  
gi|33632708|emb|CAE07520.1| EPSP synthase (3-phosp ( 441) 1113 253.0 1.3e-63  
gi|29541129|gb|AAO90072.1| 3-phosphoshikimate 1-ca ( 438) 1111 252.5 1.8e-63  
gi|196183554|gb|EDX78537.1| 3-phosphoshikimate 1-c ( 453) 1111 252.6 1.8e-63  
gi|135817147|gb|EBK02991.1| hypothetical protein G ( 333) 1108 251.8 2.3e-63  
gi|212011386|gb|ACJ18767.1| 3-phosphoshikimate 1-c ( 438) 1109 252.1 2.4e-63  
gi|226718629|gb|ACO77800.1| 3-phosphoshikimate 1-c ( 752) 1111 252.7 2.7e-63  
gi|22294064|dbj|BAC07895.1| 3-phosphoshikimate 1-c ( 440) 1108 251.9 2.8e-63  
gi|142912654|gb|EDB47050.1| hypothetical protein G ( 437) 1106 251.5 3.8e-63  
gi|257777792|gb|ACV61686.1| 3-phosphoshikimate 1-c ( 429) 1105 251.2 4.3e-63  
gi|142097015|gb|ECV56446.1| hypothetical protein G ( 434) 1105 251.2 4.4e-63  
gi|146326549|dbj|BAF61692.1| 3-phosphoshikimate 1- ( 435) 1105 251.2 4.4e-63  
gi|167294115|gb|ABZ46979.1| Sequence 20917 from pa ( 424) 1094 248.8 2.3e-62  
gi|161762125|gb|ABX77767.1| 3-phosphoshikimate 1-c ( 438) 1093 248.6 2.7e-62  
gi|154355047|gb|ABS76509.1| 3-phosphoshikimate 1-c ( 438) 1091 248.2 3.6e-62  
gi|219863318|gb|ACL43657.1| 3-phosphoshikimate 1-c ( 451) 1090 248.0 4.3e-62  
gi|169803915|gb|ACA82533.1| 3-phosphoshikimate 1-c ( 431) 1089 247.8 4.9e-62  
gi|135045054|gb|EBF14105.1| hypothetical protein G ( 439) 1089 247.8 4.9e-62  
gi|144072878|gb|EDI67260.1| hypothetical protein G ( 416) 1085 246.9 8.7e-62  
gi|85773124|gb|ABC79961.1| 3-phosphoshikimate 1-ca ( 440) 1083 246.5 1.2e-61  
gi|135065714|gb|EBF27273.1| hypothetical protein G ( 517) 1082 246.3 1.6e-61  
gi|77996387|gb|ABB15286.1| 3-phosphoshikimate 1-ca ( 428) 1081 246.0 1.6e-61  
gi|169638186|gb|ACA59692.1| 3-phosphoshikimate 1-c ( 429) 1081 246.0 1.6e-61  
gi|143958022|gb|EDH85453.1| hypothetical protein G ( 441) 1080 245.8 1.9e-61  
gi|143018092|gb|EDC21594.1| hypothetical protein G ( 322) 1077 245.0 2.4e-61  
gi|143627605|gb|EDF93941.1| hypothetical protein G ( 436) 1078 245.4 2.6e-61  
gi|143847787|gb|EDH06009.1| hypothetical protein G ( 446) 1075 244.7 4.1e-61  
gi|166854307|gb|ABY92716.1| 3-phosphoshikimate 1-c ( 423) 1074 244.5 4.6e-61  
gi|166856233|gb|ABY94641.1| 3-phosphoshikimate 1-c ( 423) 1074 244.5 4.6e-61  
gi|141843988|gb|ECT64267.1| hypothetical protein G ( 327) 1071 243.7 5.9e-61  
gi|118567567|gb|ABL02372.1| 3-phosphoshikimate 1-c ( 435) 1071 243.8 7.4e-61  
gi|136626616|gb|EBP36954.1| hypothetical protein G ( 349) 1069 243.3 8.4e-61  
gi|22777464|dbj|BAC13736.1| 5-enolpyruvylshikimat ( 429) 1070 243.6 8.5e-61  
gi|260865161|gb|ACX52267.1| 3-phosphoshikimate 1-c ( 434) 1070 243.6 8.6e-61  
gi|219953181|gb|ACL63565.1| 3-phosphoshikimate 1-c ( 440) 1070 243.6 8.7e-61  
gi|136112196|gb|EBL99481.1| hypothetical protein G ( 353) 1067 242.9 1.2e-60

gi|20516024|gb|AAM24270.1| 5-enolpyruvylshikimate- ( 423) 1067 243.0 1.3e-60  
gi|160360305|gb|ABX31919.1| 3-phosphoshikimate 1-c ( 432) 1067 243.0 1.3e-60  
gi|142941726|gb|EDB67886.1| hypothetical protein G ( 446) 1067 243.0 1.4e-60  
gi|261374423|gb|ACX77166.1| 3-phosphoshikimate 1-c ( 427) 1066 242.7 1.6e-60  
gi|35211601|dbj|BAC88979.1| 3-phosphoshikimate 1-c ( 432) 1066 242.7 1.6e-60  
gi|223698046|gb|ACN18589.1| 3-phosphoshikimate 1-c ( 428) 1065 242.5 1.8e-60  
gi|223698186|gb|ACN18694.1| 3-phosphoshikimate 1-c ( 428) 1065 242.5 1.8e-60  
gi|223698190|gb|ACN18697.1| 3-phosphoshikimate 1-c ( 428) 1065 242.5 1.8e-60  
gi|223697797|gb|EDX88756.1| 3-phosphoshikimate 1-c ( 428) 1065 242.5 1.8e-60  
gi|223697898|gb|ACN18478.1| 3-phosphoshikimate 1-c ( 428) 1065 242.5 1.8e-60  
gi|223698218|gb|ACN18718.1| 3-phosphoshikimate 1-c ( 428) 1065 242.5 1.8e-60  
gi|223697902|gb|ACN18481.1| 3-phosphoshikimate 1-c ( 428) 1065 242.5 1.8e-60  
gi|472932|emb|CAA82544.1| enolpyruvylshikimate 3-p ( 443) 1065 242.5 1.9e-60  
gi|2485262|gb|AAB73398.1|I44487 Sequence 69 from p ( 443) 1065 242.5 1.9e-60  
gi|5957585|gb|AAE08259.1| Sequence 69 from patent ( 443) 1065 242.5 1.9e-60  
gi|144974764|gb|ABP12475.1| Sequence 69 from paten ( 443) 1065 242.5 1.9e-60  
gi|2484185|gb|AAB72321.1|I49214 Sequence 69 from p ( 443) 1065 242.5 1.9e-60  
gi|196193797|gb|EDX88756.1| 3-phosphoshikimate 1-c ( 752) 1067 243.1 2.1e-60  
gi|152026654|gb|ABS24422.1| 3-phosphoshikimate 1-c ( 430) 1064 242.3 2.1e-60  
gi|189421466|gb|ACD95864.1| 3-phosphoshikimate 1-c ( 434) 1063 242.1 2.5e-60  
gi|110169280|gb|ABG53820.1| 3-phosphoshikimate 1-c ( 462) 1062 241.9 3e-60  
gi|212560430|gb|ACU33485.1| 5-enolpyruvylshikimate ( 428) 1060 241.4 3.8e-60  
gi|143254111|gb|EDD92278.1| hypothetical protein G ( 446) 1060 241.5 4e-60  
gi|143040508|gb|ABG53820.1| hypothetical protein G ( 324) 1056 240.5 5.7e-60  
gi|197088984|gb|ACH40255.1| 3-phosphoshikimate 1-c ( 429) 1057 240.8 6.1e-60  
gi|144049341|gb|EDI49929.1| hypothetical protein G ( 434) 1057 240.8 6.1e-60  
gi|135675261|gb|EDJ14661.1| hypothetical protein G ( 312) 1055 240.2 6.4e-60  
gi|142180162|gb|ECW18900.1| hypothetical protein G ( 665) 1058 241.1 7.3e-60  
gi|144022190|gb|EDI30532.1| hypothetical protein G ( 355) 1054 240.1 8.2e-60  
gi|143998481|gb|EDJ13750.1| hypothetical protein G ( 434) 1055 240.4 8.3e-60  
gi|139614992|gb|ECG28587.1| hypothetical protein G ( 296) 1053 239.8 8.3e-60  
gi|72001739|gb|AAZ57541.1| 3-phosphoshikimate 1-ca ( 444) 1055 240.4 8.4e-60  
gi|56380573|dbj|BAD76481.1| 3-phosphoshikimate 1-c ( 427) 1054 240.1 9.5e-60  
gi|83573093|gb|ABC19645.1| 3-phosphoshikimate 1-ca ( 435) 1054 240.1 9.6e-60  
gi|196170465|gb|ACG71438.1| 3-phosphoshikimate 1-c ( 440) 1053 239.9 1.1e-59  
gi|197627301|gb|EDY39860.1| 3-phosphoshikimate 1-c ( 441) 1053 239.9 1.1e-59  
gi|136364387|gb|EBN66848.1| hypothetical protein G ( 346) 1051 239.4 1.3e-59  
gi|143140739|gb|EDD11230.1| hypothetical protein G ( 437) 1051 239.5 1.5e-59  
gi|135394319|gb|EBH34333.1| hypothetical protein G ( 446) 1051 239.5 1.5e-59  
gi|77545564|gb|ABA89126.1| 3-phosphoshikimate 1-ca ( 431) 1050 239.3 1.7e-59  
gi|146232268|gb|ABQ13246.1| 3-phosphoshikimate 1-c ( 442) 1050 239.3 1.8e-59  
gi|33238026|gb|AAQ00093.1| 5-enolpyruvylshikimate- ( 444) 1050 239.3 1.8e-59  
gi|116097111|gb|ABJ62262.1| 3-phosphoshikimate 1-c ( 434) 1048 238.8 2.4e-59  
gi|123198314|gb|ABM69955.1| EPSP synthase (3-phosp ( 436) 1048 238.8 2.4e-59  
gi|143346942|gb|EDE47625.1| hypothetical protein G ( 452) 1048 238.8 2.5e-59  
gi|557771839|dbj|BAD70280.1| 3-phosphoshikimate 1-c ( 427) 1047 238.6 2.7e-59  
gi|46196018|gb|AAS80436.1| 3-phosphoshikimate 1-ca ( 427) 1046 238.4 3.2e-59  
gi|39984592|gb|AAR35978.1| 3-phosphoshikimate 1-ca ( 429) 1046 238.4 3.2e-59  
gi|143061916|gb|EDC53506.1| hypothetical protein G ( 432) 1045 238.2 3.7e-59  
gi|159888754|gb|ABX08968.1| EPSP synthase (3-phosp ( 445) 1045 238.2 3.8e-59  
gi|226095016|dbj|BAH43458.1| 3-phosphoshikimate 1- ( 426) 1044 238.0 4.3e-59  
gi|140139249|gb|ECU80061.1| hypothetical protein G ( 293) 1042 237.4 4.3e-59  
gi|256791333|gb|ACV22003.1| 3-phosphoshikimate 1-c ( 439) 1044 238.0 4.4e-59  
gi|146397030|gb|ABQ25663.1| 3-phosphoshikimate 1-c ( 428) 1042 237.5 5.8e-59  
gi|138838604|gb|ECC09871.1| hypothetical protein G ( 340) 1038 236.6 8.9e-59  
gi|206602804|gb|EDZ39285.1| 3-phosphoshikimate 1-c ( 454) 1039 236.9 9.6e-59  
gi|135494224|gb|EBI00757.1| hypothetical protein G ( 385) 1038 236.6 9.8e-59  
gi|123960447|gb|ABM75230.1| EPSP synthase (3-phosp ( 422) 1038 236.6 1.1e-58  
gi|76882029|gb|ABA56710.1| 3-phosphoshikimate 1-ca ( 444) 1038 236.7 1.1e-58  
gi|207088576|gb|EDZ65848.1| 3-phosphoshikimate 1-c ( 444) 1038 236.7 1.1e-58  
gi|143333051|gb|EDE38631.1| hypothetical protein G ( 428) 1037 236.4 1.2e-58  
gi|134267287|gb|ABO67482.1| 3-phosphoshikimate 1-c ( 432) 1037 236.4 1.2e-58

gi|143926783|gb|EDH62981.1| hypothetical protein G ( 436) 1037 236.4 1.3e-58  
gi|251774476|gb|ACT17057.1| 3-phosphoshikimate 1-c ( 429) 1036 236.2 1.4e-58  
gi|136170720|gb|EBM35960.1| hypothetical protein G ( 435) 1036 236.2 1.5e-58  
gi|123200279|gb|ABM71887.1| EPSP synthase (3-phosp ( 438) 1036 236.2 1.5e-58  
gi|135777024|gb|EBU77664.1| hypothetical protein G ( 414) 1035 236.0 1.6e-58  
gi|141174096|gb|ECQ14403.1| hypothetical protein G ( 315) 1032 235.2 2.1e-58  
gi|167282168|gb|ABZ35032.1| Sequence 8970 from pat ( 424) 1033 235.6 2.2e-58  
gi|10174284|dbj|BAB05386.1| 5-enolpyruvoylshikimat ( 431) 1033 235.6 2.3e-58  
gi|167290440|gb|ABZ43304.1| Sequence 17242 from pa ( 431) 1033 235.6 2.3e-58  
gi|222447240|gb|ACM51506.1| 3-phosphoshikimate 1-c ( 435) 1033 235.6 2.3e-58  
gi|163666945|gb|ABY33311.1| 3-phosphoshikimate 1-c ( 435) 1033 235.6 2.3e-58  
gi|142953087|gb|EDB75980.1| hypothetical protein G ( 446) 1032 235.4 2.7e-58  
gi|143633398|gb|EDF97125.1| hypothetical protein G ( 452) 1032 235.4 2.7e-58  
gi|187426429|gb|ACD05708.1| 3-phosphoshikimate 1-c ( 435) 1031 235.1 3.1e-58  
gi|144053315|gb|ED152883.1| hypothetical protein G ( 435) 1030 234.9 3.6e-58  
gi|143212092|gb|EDD62650.1| hypothetical protein G ( 440) 1030 234.9 3.6e-58  
gi|124514310|gb|EAY55824.1| 3-phosphoshikimate 1-c ( 454) 1030 234.9 3.7e-58  
gi|142267869|gb|ECW84629.1| hypothetical protein G ( 424) 1029 234.7 4.1e-58  
gi|136251086|gb|EBM89813.1| hypothetical protein G ( 439) 1029 234.7 4.2e-58  
gi|91070245|gb|ABE11164.1| EPSP synthase [uncultur ( 436) 1028 234.5 4.9e-58  
gi|143440288|gb|EDE96880.1| hypothetical protein G ( 450) 1027 234.3 5.8e-58  
gi|257832317|gb|EEV58738.1| 3-phosphoshikimate 1-c ( 431) 1026 234.0 6.5e-58  
gi|257820270|gb|EEV47418.1| 3-phosphoshikimate 1-c ( 431) 1026 234.0 6.5e-58  
gi|257836234|gb|EEV61308.1| 3-phosphoshikimate 1-c ( 431) 1026 234.0 6.5e-58  
gi|142796331|gb|EDA64709.1| hypothetical protein G ( 436) 1026 234.0 6.6e-58  
gi|142700747|gb|ECZ95072.1| hypothetical protein G ( 438) 1026 234.0 6.6e-58  
gi|142089244|gb|ECW49966.1| hypothetical protein G ( 447) 1026 234.1 6.7e-58  
gi|134284387|gb|ABO69382.1| 5-enolpyruvylshikimate ( 429) 1025 233.8 7.6e-58  
gi|123433370|gb|ECX40386.1| hypothetical protein G ( 437) 1025 233.8 7.7e-58  
gi|91070157|gb|ABE11079.1| EPSP synthase 3-p [uncu ( 449) 1025 233.8 7.8e-58  
gi|242391389|dbj|BAH81848.1| 3-phosphoshikimate 1- ( 430) 1024 233.6 8.8e-58  
gi|257823269|gb|EEV50201.1| 3-phosphoshikimate 1-c ( 431) 1024 233.6 8.8e-58  
gi|142552593|gb|ECY90600.1| hypothetical protein G ( 432) 1024 233.6 8.9e-58  
gi|137453563|gb|EBU18911.1| hypothetical protein G ( 299) 1022 233.0 9e-58  
gi|142400398|gb|ECX79396.1| hypothetical protein G ( 449) 1024 233.6 9.1e-58  
gi|135374189|gb|EBH20771.1| hypothetical protein G ( 414) 1023 233.4 1e-57  
gi|142076285|gb|ECV39284.1| hypothetical protein G ( 430) 1023 233.4 1e-57  
gi|146426946|emb|CAM89861.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426350|emb|CAM89816.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426102|emb|CAM89537.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426094|emb|CAM89528.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426030|emb|CAM89456.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146427026|emb|CAM89951.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426038|emb|CAM89465.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146427010|emb|CAM89933.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426054|emb|CAM89483.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426358|emb|CAM89825.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146427034|emb|CAM89960.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426938|emb|CAM89852.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426978|emb|CAM89897.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426986|emb|CAM89906.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426070|emb|CAM89501.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426930|emb|CAM89843.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426922|emb|CAM89834.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426046|emb|CAM89474.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426954|emb|CAM89870.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426962|emb|CAM89879.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146427002|emb|CAM89924.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426994|emb|CAM89915.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426022|emb|CAM89447.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426078|emb|CAM89510.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426086|emb|CAM89519.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57

gi|146427018|emb|CAM89942.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|146426062|emb|CAM89492.1| 3-phosphoshikimate-1- ( 188) 1018 232.0 1.2e-57  
gi|143683282|gb|EDG25451.1| hypothetical protein G ( 290) 1020 232.6 1.2e-57  
gi|135970867|gb|EBL05592.1| hypothetical protein G ( 449) 1022 233.2 1.2e-57  
gi|83756719|gb|ABC44832.1| 3-phosphoshikimate 1-ca ( 430) 1021 233.0 1.4e-57  
gi|262335582|gb|ACV49379.1| 3-phosphoshikimate 1-c ( 434) 1021 233.0 1.4e-57  
gi|257474981|gb|ACV55301.1| 3-phosphoshikimate 1-c ( 442) 1020 232.7 1.6e-57  
gi|157387605|gb|ABV50310.1| EPSP synthase (3-phosp ( 436) 1019 232.5 1.9e-57  
gi|256583260|gb|ACU94394.1| 3-phosphoshikimate 1-c ( 443) 1019 232.5 1.9e-57  
gi|146425918|emb|CAM89330.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425950|emb|CAM89366.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427242|emb|CAM90194.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425423|emb|CAM89231.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426230|emb|CAM89681.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426142|emb|CAM89582.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425934|emb|CAM89348.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425439|emb|CAM89249.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427106|emb|CAM90041.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426310|emb|CAM89771.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426278|emb|CAM89735.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427322|emb|CAM90284.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427130|emb|CAM90068.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427210|emb|CAM90158.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427162|emb|CAM90104.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427090|emb|CAM90023.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426134|emb|CAM89573.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427946|emb|CAM90365.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427338|emb|CAM90302.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427170|emb|CAM90113.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427922|emb|CAM90338.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426254|emb|CAM89708.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427930|emb|CAM90347.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426006|emb|CAM89429.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425463|emb|CAM89276.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426246|emb|CAM89699.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427122|emb|CAM90059.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427354|emb|CAM90320.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425942|emb|CAM89357.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427290|emb|CAM90248.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425926|emb|CAM89339.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427146|emb|CAM90086.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425958|emb|CAM89375.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427066|emb|CAM89996.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427250|emb|CAM90203.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425495|emb|CAM89312.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426318|emb|CAM89780.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427282|emb|CAM90239.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426214|emb|CAM89663.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427298|emb|CAM90257.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426270|emb|CAM89726.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427258|emb|CAM90212.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427058|emb|CAM89987.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427138|emb|CAM90077.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426970|emb|CAM89888.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426150|emb|CAM89591.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425966|emb|CAM89384.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427234|emb|CAM90185.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426118|emb|CAM89555.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427330|emb|CAM90293.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426326|emb|CAM89789.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427962|emb|CAM90383.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427362|emb|CAM90329.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57

gi|146425974|emb|CAM89393.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427154|emb|CAM90095.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427074|emb|CAM90005.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425471|emb|CAM89285.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427178|emb|CAM90122.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426126|emb|CAM89564.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425487|emb|CAM89303.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427050|emb|CAM89978.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426110|emb|CAM89546.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426174|emb|CAM89618.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427202|emb|CAM90149.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425455|emb|CAM89267.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425982|emb|CAM89402.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427314|emb|CAM90275.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426262|emb|CAM89717.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426158|emb|CAM89600.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427082|emb|CAM90014.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426206|emb|CAM89546.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426182|emb|CAM89627.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427266|emb|CAM90221.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427954|emb|CAM90374.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427938|emb|CAM90356.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427186|emb|CAM90131.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427978|emb|CAM90401.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427042|emb|CAM89969.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427306|emb|CAM90266.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427274|emb|CAM90230.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427194|emb|CAM90140.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426294|emb|CAM89753.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|169402995|emb|CAO99131.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426222|emb|CAM89672.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425431|emb|CAM89240.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427114|emb|CAM90050.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426190|emb|CAM89636.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426014|emb|CAM89438.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426238|emb|CAM89690.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|169402997|emb|CAO99132.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425990|emb|CAM89411.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426302|emb|CAM89762.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425479|emb|CAM89294.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426334|emb|CAM89798.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426342|emb|CAM89807.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427218|emb|CAM90167.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426166|emb|CAM89609.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426286|emb|CAM89744.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425399|emb|CAM89204.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425407|emb|CAM89213.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425391|emb|CAM89195.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425447|emb|CAM89258.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427346|emb|CAM90311.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425415|emb|CAM89222.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427226|emb|CAM90176.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425503|emb|CAM89321.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146426198|emb|CAM89645.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427098|emb|CAM90032.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146425998|emb|CAM89420.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|146427970|emb|CAM90392.1| 3-phosphoshikimate-1- ( 188) 1014 231.2 2.1e-57  
gi|257830747|gb|EEV57354.1| 3-phosphoshikimate 1-c ( 431) 1018 232.3 2.2e-57  
gi|257825716|gb|EEV52342.1| 3-phosphoshikimate 1-c ( 431) 1018 232.3 2.2e-57  
gi|257818523|gb|EEV45851.1| 3-phosphoshikimate 1-c ( 431) 1018 232.3 2.2e-57  
gi|260073869|gb|EEW62193.1| 3-phosphoshikimate 1-c ( 431) 1018 232.3 2.2e-57  
gi|33749599|gb|AAQ46605.1| Sequence 7163 from pate ( 431) 1018 232.3 2.2e-57

gi|143005904|gb|EDC12672.1| hypothetical protein G ( 435) 1018 232.3 2.2e-57  
gi|209540793|gb|AC161369.1| 3-phosphoshikimate 1-c ( 427) 1017 232.1 2.5e-57  
gi|28810994|dbj|BAC63928.1| putative 3-phosphoshik ( 430) 1017 232.1 2.5e-57  
gi|21904763|gb|AAM79634.1| putative 3-phosphoshiki ( 430) 1017 232.1 2.5e-57  
gi|145411265|gb|ABP68269.1| 3-phosphoshikimate 1-c ( 433) 1017 232.1 2.5e-57  
gi|126543020|gb|ABO17262.1| EFSP synthase (3-phosp ( 436) 1017 232.1 2.6e-57  
gi|140350316|gb|ECL17116.1| hypothetical protein G ( 315) 1015 231.5 2.7e-57  
gi|221564743|gb|ACM20715.1| 3-phosphoshikimate 1-c ( 429) 1016 231.9 2.9e-57  
gi|134271853|gb|CAM30088.1| 3-phosphoshikimate 1- ( 430) 1016 231.9 2.9e-57  
gi|94542283|gb|ABF32332.1| 3-phosphoshikimate 1-ca ( 430) 1016 231.9 2.9e-57  
gi|94546168|gb|ABF36215.1| 3-phosphoshikimate 1-ca ( 430) 1016 231.9 2.9e-57  
gi|50903493|gb|AAT87208.1| 3-phosphoshikimate 1-ca ( 430) 1016 231.9 2.9e-57  
gi|135062405|gb|EBF25151.1| hypothetical protein G ( 284) 1013 231.1 3.4e-57  
gi|142745913|gb|EDA27571.1| hypothetical protein G ( 427) 1015 231.6 3.4e-57  
gi|167273311|gb|ABZ26175.1| Sequence 113 from pate ( 431) 1015 231.6 3.4e-57  
gi|2983892|gb|AAC07443.1| 5-enolpyruvylshikimate-3 ( 431) 1015 231.6 3.4e-57  
gi|221537862|gb|EEE40315.1| 3-phosphoshikimate 1-c ( 436) 1015 231.6 3.5e-57  
gi|160429338|gb|ABX42901.1| 3-phosphoshikimate 1-c ( 436) 1015 231.6 3.5e-57  
gi|19748517|gb|AAL97960.1| putative 3-phosphoshiki ( 427) 1014 231.4 4e-57  
gi|71802853|gb|AAX72206.1| 3-phosphoshikimate 1-ca ( 427) 1014 231.4 4e-57  
gi|94548104|gb|ABF38150.1| 3-phosphoshikimate 1-ca ( 427) 1014 231.4 4e-57  
gi|114337807|gb|ABF68655.1| 3-phosphoshikimate 1-c ( 428) 1014 231.4 4e-57  
gi|78712497|gb|ABB49674.1| 3-phosphoshikimate 1-ca ( 436) 1014 231.4 4e-57  
gi|143578905|gb|EDF73147.1| hypothetical protein G ( 438) 1014 231.4 4e-57  
gi|146425367|emb|CAM89168.1| 3-phosphoshikimate-1- ( 188) 1009 230.1 4.5e-57  
gi|146425359|emb|CAM89159.1| 3-phosphoshikimate-1- ( 188) 1009 230.1 4.5e-57  
gi|146425375|emb|CAM89177.1| 3-phosphoshikimate-1- ( 188) 1009 230.1 4.5e-57  
gi|146425383|emb|CAM89186.1| 3-phosphoshikimate-1- ( 188) 1009 230.1 4.5e-57  
gi|143387961|gb|EDE69064.1| hypothetical protein G ( 450) 1013 231.2 4.8e-57  
gi|142420835|gb|ECX94792.1| hypothetical protein G ( 408) 1012 231.0 5.2e-57  
gi|146425311|emb|CAM89105.1| 3-phosphoshikimate-1- ( 188) 1008 229.8 5.2e-57  
gi|146425207|emb|CAM88988.1| 3-phosphoshikimate-1- ( 188) 1008 229.8 5.2e-57  
gi|146425193|emb|CAM88970.1| 3-phosphoshikimate-1- ( 188) 1008 229.8 5.2e-57  
gi|146425279|emb|CAM89069.1| 3-phosphoshikimate-1- ( 188) 1008 229.8 5.2e-57  
gi|146425263|emb|CAM89051.1| 3-phosphoshikimate-1- ( 188) 1008 229.8 5.2e-57  
gi|146425303|emb|CAM89096.1| 3-phosphoshikimate-1- ( 188) 1008 229.8 5.2e-57  
gi|146425271|emb|CAM89060.1| 3-phosphoshikimate-1- ( 188) 1008 229.8 5.2e-57  
gi|146425295|emb|CAM89087.1| 3-phosphoshikimate-1- ( 188) 1008 229.8 5.2e-57  
gi|146425255|emb|CAM89042.1| 3-phosphoshikimate-1- ( 188) 1008 229.8 5.2e-57  
gi|146425335|emb|CAM89132.1| 3-phosphoshikimate-1- ( 188) 1008 229.8 5.2e-57  
gi|146425287|emb|CAM89078.1| 3-phosphoshikimate-1- ( 188) 1008 229.8 5.2e-57  
gi|146425199|emb|CAM88979.1| 3-phosphoshikimate-1- ( 188) 1008 229.8 5.2e-57  
gi|146425327|emb|CAM89123.1| 3-phosphoshikimate-1- ( 188) 1008 229.8 5.2e-57  
gi|146425215|emb|CAM88997.1| 3-phosphoshikimate-1- ( 188) 1008 229.8 5.2e-57  
gi|146425319|emb|CAM89114.1| 3-phosphoshikimate-1- ( 188) 1008 229.8 5.2e-57  
gi|146425343|emb|CAM89141.1| 3-phosphoshikimate-1- ( 188) 1008 229.8 5.2e-57  
gi|146425223|emb|CAM89006.1| 3-phosphoshikimate-1- ( 188) 1008 229.8 5.2e-57  
gi|146425231|emb|CAM89015.1| 3-phosphoshikimate-1- ( 188) 1008 229.8 5.2e-57  
gi|146425239|emb|CAM89024.1| 3-phosphoshikimate-1- ( 188) 1008 229.8 5.2e-57  
gi|146425247|emb|CAM89033.1| 3-phosphoshikimate-1- ( 188) 1008 229.8 5.2e-57  
gi|219541331|gb|ACL23069.1| 3-phosphoshikimate 1-c ( 435) 1012 231.0 5.4e-57  
gi|143006344|gb|EDC12996.1| hypothetical protein G ( 368) 1011 230.7 5.6e-57  
gi|71853696|gb|AAZ51719.1| 3-phosphoshikimate 1-ca ( 430) 1011 230.8 6.3e-57  
gi|13622461|gb|AAK34180.1| putative 3-phosphoshiki ( 430) 1011 230.8 6.3e-57  
gi|260552604|gb|EEX25604.1| 3-phosphoshikimate 1-c ( 432) 1011 230.8 6.3e-57  
gi|142511258|gb|ECV60760.1| hypothetical protein G ( 444) 1011 230.8 6.4e-57  
gi|78193339|gb|ABB31106.1| 3-phosphoshikimate 1-ca ( 429) 1010 230.6 7.3e-57  
gi|257812675|gb|EEV41433.1| 3-phosphoshikimate 1-c ( 431) 1010 230.6 7.3e-57  
gi|135521687|gb|EBI18400.1| hypothetical protein G ( 424) 1009 230.3 8.4e-57  
gi|142886839|gb|EDB28223.1| hypothetical protein G ( 437) 1009 230.3 8.6e-57  
gi|143595742|gb|EDF78097.1| hypothetical protein G ( 452) 1009 230.4 8.8e-57  
gi|16414538|emb|CAC97267.1| aroE [Listeria innocua ( 428) 1008 230.1 9.8e-57

gi|94544175|gb|ABF34223.1| 3-phosphoshikimate 1-ca ( 430) 1008 230.1 9.9e-57  
gi|170015681|emb|CAP05278.1| 3-phosphoshikimate 1- ( 188) 1003 228.8 1.1e-56  
gi|142517542|gb|ECY65289.1| hypothetical protein G ( 329) 1005 229.4 1.3e-56  
gi|142714811|gb|EDA05134.1| hypothetical protein G ( 333) 1005 229.4 1.3e-56  
gi|56909911|dbj|BAD64438.1| 3-phosphoshikimate 1-c ( 430) 1006 229.7 1.3e-56  
gi|146425351|emb|CAM89150.1| 3-phosphoshikimate 1- ( 188) 1001 228.3 1.5e-56  
gi|225702144|emb|CAW99831.1| 3-phosphoshikimate 1- ( 427) 1005 229.5 1.5e-56  
gi|143838364|gb|EDG99243.1| hypothetical protein G ( 449) 1004 229.3 1.9e-56  
gi|152023794|gb|ABS21564.1| 3-phosphoshikimate 1-c ( 424) 1003 229.0 2.1e-56  
gi|136143934|gb|EBM18523.1| hypothetical protein G ( 432) 1002 228.8 2.4e-56  
gi|159890538|gb|ABX03618.1| 3-phosphoshikimate 1-c ( 431) 1001 228.6 2.8e-56  
gi|142401851|gb|ECX80480.1| hypothetical protein G ( 435) 1001 228.6 2.9e-56  
gi|223697798|gb|ACN18403.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223698102|gb|ACN18631.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223698010|gb|ACN18562.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223698146|gb|ACN18664.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223698138|gb|ACN18658.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223697846|gb|ACN18439.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223697834|gb|ACN18430.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223698178|gb|ACN18688.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223697818|gb|ACN18418.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223697786|gb|ACN18394.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223698134|gb|ACN18655.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223698094|gb|ACN18562.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223698122|gb|ACN18646.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223697730|gb|ACN18352.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223697890|gb|ACN18472.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223697754|gb|ACN18370.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223698014|gb|ACN18565.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223697806|gb|ACN18409.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223697726|gb|ACN18349.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223698098|gb|ACN18628.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223697854|gb|ACN18445.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223697930|gb|ACN18502.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223698106|gb|ACN18634.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223697934|gb|ACN18505.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223698142|gb|ACN18661.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223697782|gb|ACN18391.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223698202|gb|ACN18706.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223697790|gb|ACN18397.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223697738|gb|ACN18358.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223698126|gb|ACN18649.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223698130|gb|ACN18652.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223697830|gb|ACN18427.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223697822|gb|ACN18421.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223697734|gb|ACN18355.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223697794|gb|ACN18400.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|223698038|gb|ACN18583.1| 3-phosphoshikimate 1-c ( 428) 1000 228.4 3.3e-56  
gi|138829093|gb|ECC05312.1| hypothetical protein G ( 299) 998 227.8 3.4e-56  
gi|138083024|gb|EBX64838.1| hypothetical protein G ( 304) 997 227.6 4e-56  
gi|183226926|dbj|BAG27442.1| 3-phosphoshikimate 1- ( 432) 998 227.9 4.5e-56  
gi|142300251|gb|ECX08327.1| hypothetical protein G ( 437) 998 227.9 4.5e-56  
gi|33639902|emb|CAE19072.1| EPSP synthase (3-phosp ( 438) 998 228.0 4.5e-56  
gi|142082676|gb|ECV44561.1| hypothetical protein G ( 438) 998 228.0 4.5e-56  
gi|143268164|gb|EDE02418.1| hypothetical protein G ( 350) 996 227.4 5.1e-56  
gi|55738600|gb|AAV62241.1| 3-phosphoshikimate 1-ca ( 427) 997 227.7 5.2e-56  
gi|223698074|gb|ACN18610.1| 3-phosphoshikimate 1-c ( 428) 997 227.7 5.2e-56  
gi|118417505|gb|ABK85924.1| 3-phosphoshikimate 1-c ( 429) 997 227.7 5.2e-56  
gi|222455696|gb|ACM59958.1| 3-phosphoshikimate 1-c ( 433) 997 227.7 5.2e-56  
gi|167282943|gb|ABZ35807.1| Sequence 9745 from pat ( 417) 996 227.5 5.9e-56  
gi|55736709|gb|AAV60351.1| 3-phosphoshikimate 1-ca ( 427) 996 227.5 6e-56  
gi|89334456|dbj|BAE84051.1| hypothetical protein [ ( 435) 996 227.5 6.1e-56

gi|219539671|gb|ACL21410.1| 3-phosphoshikimate 1-c ( 435) 996 227.5 6.1e-56  
gi|140074666|gb|ECJ36897.1| hypothetical protein G ( 299) 994 227.0 6.2e-56  
gi|156232792|gb|ABU57575.1| 3-phosphoshikimate 1-c ( 431) 995 227.3 7e-56  
gi|143019553|gb|EDC22663.1| hypothetical protein G ( 437) 995 227.3 7.1e-56  
gi|143233975|gb|EDD78447.1| hypothetical protein G ( 437) 995 227.3 7.1e-56  
gi|225699532|emb|CAW93107.1| 3-phosphoshikimate 1- ( 427) 994 227.1 8.1e-56  
gi|195974537|gb|ACG62063.1| 3-phosphoshikimate 1-c ( 427) 994 227.1 8.1e-56  
gi|223698086|gb|ACN18619.1| 3-phosphoshikimate 1-c ( 428) 994 227.1 8.1e-56  
gi|42737974|gb|AAS41905.1| 3-phosphoshikimate 1-ca ( 429) 994 227.1 8.1e-56  
gi|251849803|gb|EES77762.1| conserved hypothetical ( 424) 993 226.9 9.4e-56  
gi|222114137|emb|CAR42616.1| 3-phosphoshikimate 1- ( 427) 993 226.9 9.4e-56  
gi|223698090|gb|ACN18622.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697766|gb|ACN18379.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697918|gb|ACN18493.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698206|gb|ACN18709.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697938|gb|ACN18508.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697958|gb|ACN18523.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698150|gb|ACN18667.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698042|gb|ACN18586.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697926|gb|ACN18499.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697742|gb|ACN18361.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698162|gb|ACN18676.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698154|gb|ACN18670.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697746|gb|ACN18364.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698026|gb|ACN18574.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697878|gb|ACN18463.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697842|gb|ACN18436.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698110|gb|ACN18637.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697966|gb|ACN18529.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698210|gb|ACN18712.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697866|gb|ACN18454.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698214|gb|ACN18715.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698054|gb|ACN18595.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697950|gb|ACN18517.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|116742243|emb|CAK21367.1| aroA [Listeria welshi ( 428) 993 226.9 9.4e-56  
gi|223697882|gb|ACN18466.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697894|gb|ACN18475.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697982|gb|ACN18541.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697778|gb|ACN18388.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698222|gb|ACN18721.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698158|gb|ACN18673.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697762|gb|ACN18376.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698022|gb|ACN18571.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698174|gb|ACN18685.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697954|gb|ACN18520.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697986|gb|ACN18544.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697974|gb|ACN18535.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697994|gb|ACN18550.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698018|gb|ACN18568.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697942|gb|ACN18511.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697970|gb|ACN18532.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697862|gb|ACN18451.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697858|gb|ACN18448.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698062|gb|ACN18601.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697922|gb|ACN18496.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698002|gb|ACN18556.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698066|gb|ACN18604.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697978|gb|ACN18538.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697774|gb|ACN18385.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698006|gb|ACN18559.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698034|gb|ACN18580.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697886|gb|ACN18469.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56



gi|223697962|gb|ACN18526.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698182|gb|ACN18691.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697914|gb|ACN18490.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697850|gb|ACN18442.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698078|gb|ACN18613.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697874|gb|ACN18460.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698194|gb|ACN18700.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697770|gb|ACN18382.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697758|gb|ACN18373.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698050|gb|ACN18592.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697946|gb|ACN18514.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697910|gb|ACN18487.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697990|gb|ACN18547.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698226|gb|ACN18724.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698058|gb|ACN18598.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697998|gb|ACN18553.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697870|gb|ACN18457.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698166|gb|ACN18679.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698170|gb|ACN18682.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697838|gb|ACN18433.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698070|gb|ACN18607.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697810|gb|ACN18412.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|16411376|emb|CAD00001.1| aroE [Listeria monocyt ( 428) 993 226.9 9.4e-56  
gi|223698030|gb|ACN18571.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223698198|gb|ACN18703.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697906|gb|ACN18484.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|223697814|gb|ACN18415.1| 3-phosphoshikimate 1-c ( 428) 993 226.9 9.4e-56  
gi|135911255|gb|EBK65191.1| hypothetical protein G ( 411) 992 226.6 1.1e-55  
gi|49330610|gb|AAT61256.1| 3-phosphoshikimate 1-ca ( 429) 992 226.6 1.1e-55  
gi|136005759|gb|EBL27605.1| hypothetical protein G ( 436) 992 226.6 1.1e-55  
gi|223697826|gb|ACN18424.1| 3-phosphoshikimate 1-c ( 428) 991 226.4 1.3e-55  
gi|51976038|gb|AAU17588.1| 3-phosphoshikimate 1-ca ( 429) 991 226.4 1.3e-55  
gi|218537723|gb|ACK90121.1| 3-phosphoshikimate 1-c ( 429) 991 226.4 1.3e-55  
gi|225787055|gb|ACO27272.1| 3-phosphoshikimate 1-c ( 429) 991 226.4 1.3e-55  
gi|134964169|gb|EBE59745.1| hypothetical protein G ( 425) 990 226.2 1.5e-55  
gi|46881426|gb|AAT04722.1| 3-phosphoshikimate 1-ca ( 428) 990 226.2 1.5e-55  
gi|225876989|emb|CAS05698.1| Putative 5-enolpyruvyl ( 428) 990 226.2 1.5e-55  
gi|217333195|gb|ACK38989.1| 3-phosphoshikimate 1-c ( 428) 990 226.2 1.5e-55  
gi|139748663|gb|ECH19156.1| hypothetical protein G ( 292) 988 225.6 1.5e-55  
gi|116100807|gb|ABJ65953.1| 3-phosphoshikimate 1-c ( 427) 989 226.0 1.7e-55  
gi|29896607|gb|AAP09886.1| 3-phosphoshikimate 1-ca ( 429) 989 226.0 1.7e-55  
gi|218162110|gb|ACK62102.1| 3-phosphoshikimate 1-c ( 429) 989 226.0 1.7e-55  
gi|223698114|gb|ACN18640.1| 3-phosphoshikimate 1-c ( 428) 988 225.8 2e-55  
gi|223697802|gb|ACN18406.1| 3-phosphoshikimate 1-c ( 428) 988 225.8 2e-55  
gi|218541446|gb|ACK93840.1| 3-phosphoshikimate 1-c ( 429) 988 225.8 2e-55  
gi|136315089|gb|EBN33493.1| hypothetical protein G ( 448) 988 225.8 2.1e-55  
gi|135571448|gb|EBI50213.1| hypothetical protein G ( 418) 987 225.5 2.3e-55  
gi|118502485|gb|ABK98967.1| 3-phosphoshikimate 1-c ( 431) 987 225.6 2.3e-55  
gi|269100864|gb|AC219851.1| 3-phosphoshikimate 1-c ( 436) 987 225.6 2.4e-55  
gi|30257544|gb|AAP26774.1| 3-phosphoshikimate 1-ca ( 429) 986 225.3 2.7e-55  
gi|47551804|gb|AAT32071.2| 3-phosphoshikimate 1-ca ( 429) 986 225.3 2.7e-55  
gi|227006491|gb|ACP16234.1| 3-phosphoshikimate 1-c ( 429) 986 225.3 2.7e-55  
gi|229268332|gb|ACQ49969.1| 3-phosphoshikimate 1-c ( 429) 986 225.3 2.7e-55  
gi|49179677|gb|AAT55053.1| 3-phosphoshikimate 1-ca ( 429) 986 225.3 2.7e-55  
gi|270278481|gb|EFA24327.1| 3-phosphoshikimate 1-c ( 427) 985 225.1 3.2e-55  
gi|223698118|gb|ACN18643.1| 3-phosphoshikimate 1-c ( 428) 985 225.1 3.2e-55  
gi|217062913|gb|ACJ77163.1| 3-phosphoshikimate 1-c ( 429) 985 225.1 3.2e-55  
gi|197205480|gb|ACH47977.1| 3-phosphoshikimate 1-c ( 427) 984 224.9 3.7e-55  
gi|142478979|gb|ECY37940.1| hypothetical protein G ( 434) 984 224.9 3.7e-55  
gi|136325414|gb|EBN40361.1| hypothetical protein G ( 443) 984 224.9 3.8e-55  
gi|143475974|gb|EDF19143.1| hypothetical protein G ( 425) 983 224.7 4.2e-55  
gi|262261195|gb|EY79894.1| 3-phosphoshikimate 1-c ( 427) 983 224.7 4.3e-55

gi|139854247|gb|ECH92990.1| hypothetical protein G ( 307) 981 224.1 4.5e-55  
gi|223698082|gb|ACN18616.1| 3-phosphoshikimate 1-c ( 428) 982 224.5 5e-55  
gi|136028731|gb|EBL43002.1| hypothetical protein G ( 418) 981 224.2 5.7e-55  
gi|269788888|gb|ACZ41029.1| 3-phosphoshikimate 1-c ( 439) 981 224.3 5.9e-55  
gi|142787199|gb|EDA57892.1| hypothetical protein G ( 308) 979 223.7 6.1e-55  
gi|136256857|gb|EBM93634.1| hypothetical protein G ( 434) 980 224.0 6.8e-55  
gi|135032108|gb|EBF05757.1| hypothetical protein G ( 416) 979 223.8 7.6e-55  
gi|157075738|gb|ABV10421.1| 3-phosphoshikimate 1-c ( 427) 979 223.8 7.8e-55  
gi|257800017|gb|EEV29069.1| 3-phosphoshikimate 1-c ( 429) 979 223.8 7.8e-55  
gi|257806578|gb|EEV35400.1| 3-phosphoshikimate 1-c ( 429) 979 223.8 7.8e-55  
gi|182629665|gb|ACB90613.1| 3-phosphoshikimate 1-c ( 431) 979 223.8 7.9e-55  
gi|143656521|gb|EDG10782.1| hypothetical protein G ( 436) 979 223.8 7.9e-55  
gi|24377157|gb|AAN58503.1| AE014920\_1 5-enolpyruvyl ( 427) 978 223.6 9.1e-55  
gi|52348641|gb|AAU41275.1| AroE [Bacillus lichenif ( 428) 978 223.6 9.1e-55  
gi|145903020|gb|AAU23920.3| 5-enolpyruvylshikimat ( 428) 978 223.6 9.1e-55  
gi|257809882|gb|EEV38702.1| 3-phosphoshikimate 1-c ( 429) 978 223.6 9.1e-55  
gi|228190727|gb|ACP61997.1| Sequence 6491 from pat ( 464) 978 223.6 9.7e-55  
gi|136404114|gb|EBN94215.1| hypothetical protein G ( 389) 976 223.1 1.1e-54  
gi|163862888|gb|ABV43947.1| 3-phosphoshikimate 1-c ( 429) 976 223.2 1.2e-54  
gi|148569867|gb|ABQ92012.1| 3-phosphoshikimate 1-c ( 433) 976 223.2 1.2e-54  
gi|160221349|gb|ABX11283.1| 5-enolpyruvylshikimate ( 454) 976 223.2 1.3e-54  
gi|144215829|gb|EDU70892.1| hypothetical protein G ( 390) 975 222.9 1.3e-54  
gi|142484138|gb|ECY41805.1| hypothetical protein G ( 408) 975 222.9 1.4e-54  
gi|168995736|gb|ACA36348.1| 3-phosphoshikimate 1-c ( 427) 975 222.9 1.4e-54  
gi|134873812|gb|EBB99647.1| hypothetical protein G ( 284) 972 222.1 1.6e-54  
gi|254997648|dbj|BAH88249.1| 5-enolpyruvylshikimat ( 427) 974 222.7 1.7e-54  
gi|225722753|gb|ACO18606.1| 3-phosphoshikimate 1-c ( 427) 974 222.7 1.7e-54  
gi|225724804|gb|ACO20656.1| 3-phosphoshikimate 1-c ( 427) 974 222.7 1.7e-54  
gi|225728344|gb|ACO24195.1| 3-phosphoshikimate 1-c ( 427) 974 222.7 1.7e-54  
gi|194357091|gb|ACE55539.1| 3-phosphoshikimate 1-c ( 427) 974 222.7 1.7e-54  
gi|225720612|gb|ACO16466.1| 3-phosphoshikimate 1-c ( 427) 974 222.7 1.7e-54  
gi|220674574|emb|CAR69137.1| 3-phosphoshikimate 1- ( 427) 974 222.7 1.7e-54  
gi|56634435|gb|AAW08901.1| Sequence 4297 from pate ( 459) 974 222.7 1.8e-54  
gi|239807780|gb|ACS24845.1| 3-phosphoshikimate 1-c ( 428) 973 222.5 1.9e-54  
gi|14972856|gb|AAK75469.1| 3-phosphoshikimate 1-ca ( 427) 972 222.3 2.2e-54  
gi|135426956|gb|EBH56298.1| hypothetical protein G ( 425) 971 222.1 2.6e-54  
gi|259339911|gb|ACW47959.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|115811019|gb|ABJ32064.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|118850301|gb|ABL27658.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|125115865|gb|ABN23906.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|197042232|gb|ACH18866.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|182920657|gb|ACC11871.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|14112362|gb|AAE57779.1| Sequence 2 from patent ( 427) 971 222.1 2.6e-54  
gi|259333339|gb|ACW44685.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|47241604|gb|AAT16176.1| Sequence 3546 from pate ( 427) 971 222.1 2.6e-54  
gi|196999282|gb|ACG98297.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|189900828|gb|ACE55746.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|125085509|gb|ABN18836.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|189906152|gb|ACE58407.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|196983653|gb|ACG90840.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|118899388|gb|ABL49226.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|217237720|gb|ACK13966.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|281104388|gb|ADA50902.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|182925979|gb|ACC14532.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|189875194|gb|ACE43394.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|189895504|gb|ACE53085.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|197018713|gb|ACH07859.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|112079923|gb|ABI03507.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|197051730|gb|ACH23628.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|112096977|gb|ABI08548.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|5616525|gb|AAD45819.1| AF169483\_1 5-enolpyruvyls ( 427) 971 222.1 2.6e-54  
gi|118841667|gb|ABL24997.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54

gi|115887011|gb|ABJ49677.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|167338106|gb|ABZ66405.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|118806535|gb|ABL17420.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|182938218|gb|ACC20088.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|189885049|gb|ACE48254.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|196993017|gb|ACG95251.1| Sequence 3546 from pat ( 427) 971 222.1 2.6e-54  
gi|167291680|gb|ABZ44544.1| Sequence 18482 from pa ( 430) 971 222.1 2.6e-54  
gi|12724764|gb|AAK05842.1|AE006404\_9 3-phosphoshik ( 430) 971 222.1 2.6e-54  
gi|136611001|gb|EBP27638.1| hypothetical protein G ( 365) 970 221.8 2.7e-54  
gi|135896096|gb|EBK54670.1| hypothetical protein G ( 394) 970 221.8 2.8e-54  
gi|138507712|gb|ECA07901.1| hypothetical protein G ( 321) 968 221.3 3.3e-54  
gi|136814875|gb|EBQ58741.1| hypothetical protein G ( 434) 969 221.6 3.6e-54  
gi|140758638|gb|ECN32069.1| hypothetical protein G ( 299) 967 221.1 3.6e-54  
gi|139603541|gb|ECG20683.1| hypothetical protein G ( 307) 967 221.1 3.7e-54  
gi|116076465|gb|ABJ54185.1| 3-phosphoshikimate 1-c ( 427) 968 221.4 4.1e-54  
gi|15458865|gb|AAL00033.1| 3-Enolpyruvylshikimate- ( 431) 968 221.4 4.1e-54  
gi|14106225|gb|AAE55840.1| Sequence 2 from patent ( 427) 965 220.8 6.4e-54  
gi|134387655|gb|EBB06773.1| hypothetical protein G ( 272) 962 220.0 7.2e-54  
gi|167289692|gb|ABZ42556.1| Sequence 16494 from pa ( 417) 964 220.5 7.4e-54  
gi|149948344|gb|ABR46872.1| 3-phosphoshikimate 1-c ( 425) 963 220.3 8.7e-54  
gi|142551989|gb|ECY90178.1| hypothetical protein G ( 415) 962 220.1 9.9e-54  
gi|170287625|dbj|BAG14146.1| 3-phosphoshikimate 1- ( 430) 962 220.1 1e-53  
gi|138658485|gb|ECB10957.1| hypothetical protein G ( 299) 960 219.6 1e-53  
gi|124248218|emb|CAL26234.1| 5-enolpyruvylshikima ( 428) 961 219.9 1.2e-53  
gi|154352359|gb|ABS74438.1| AroE [Bacillus amyloli ( 428) 961 219.9 1.2e-53  
gi|251819569|emb|CAR45232.1| 3-phosphoshikimate 1- ( 426) 960 219.7 1.4e-53  
gi|251818311|emb|CBZ56122.1| 3-phosphoshikimate 1- ( 426) 960 219.7 1.4e-53  
gi|251815711|emb|CAZ51311.1| 3-phosphoshikimate 1- ( 426) 960 219.7 1.4e-53  
gi|134904510|gb|EBE20044.1| hypothetical protein G ( 325) 958 219.1 1.5e-53  
gi|167780379|gb|ABZ29867.1| 3-phosphoshikimate 1-c ( 428) 959 219.5 1.6e-53  
gi|167776701|gb|ABZ95002.1| 3-phosphoshikimate 1-c ( 428) 959 219.5 1.6e-53  
gi|206739263|gb|ACI18321.1| 3-phosphoshikimate 1-c ( 435) 959 219.5 1.6e-53  
gi|157681517|gb|ABV62661.1| 3-phosphoshikimate 1-c ( 428) 958 219.2 1.9e-53  
gi|124493516|emb|CAL98495.1| 3-phosphoshikimate 1- ( 430) 957 219.0 2.2e-53  
gi|683583|emb|CAA55180.1| 5-enolpyruvylshikimate-3 ( 430) 957 219.0 2.2e-53  
gi|139732520|gb|ECH09244.1| hypothetical protein G ( 292) 955 218.5 2.2e-53  
gi|143334178|gb|EDE39370.1| hypothetical protein G ( 484) 957 219.1 2.4e-53  
gi|195932969|gb|ACG57669.1| 3-phosphoshikimate 1-c ( 427) 955 218.6 2.9e-53  
gi|251771402|gb|EES51981.1| 3-phosphoshikimate 1-c ( 446) 955 218.6 3e-53  
gi|143130375|gb|EDD03631.1| hypothetical protein G ( 342) 953 218.1 3.3e-53  
gi|226090387|dbj|BAH38832.1| 3-phosphoshikimate 1- ( 672) 955 218.7 4.1e-53  
gi|196192722|gb|EDX87686.1| 3-phosphoshikimate 1-c ( 409) 952 217.9 4.4e-53  
gi|143517608|gb|EDF42079.1| hypothetical protein G ( 418) 952 217.9 4.5e-53  
gi|143823880|gb|EDG88784.1| hypothetical protein G ( 442) 952 217.9 4.7e-53  
gi|140283581|gb|ECK73924.1| hypothetical protein G ( 280) 949 217.1 5.2e-53  
gi|239516843|gb|EEQ56709.1| conserved hypothetical ( 425) 951 217.7 5.3e-53  
gi|24194817|gb|AAN48457.1|AE011307\_10 3-phosphoshi ( 440) 950 217.5 6.3e-53  
gi|45601534|gb|AAS71015.1| 3-phosphoshikimate 1-ca ( 440) 950 217.5 6.3e-53  
gi|189185917|gb|ACD83102.1| 5-enolpyruvylshikimate ( 695) 951 217.9 7.8e-53  
gi|125498192|gb|ABN44858.1| 3-phosphoshikimate 1-c ( 427) 948 217.1 8.4e-53  
gi|257803529|gb|EEV32351.1| 3-phosphoshikimate 1-c ( 429) 948 217.1 8.4e-53  
gi|116108251|gb|ABJ73391.1| 3-phosphoshikimate 1-c ( 430) 947 216.8 9.8e-53  
gi|140263796|gb|ECK64233.1| hypothetical protein G ( 273) 944 216.0 1.1e-52  
gi|142964063|gb|EDB83665.1| hypothetical protein G ( 387) 945 216.4 1.2e-52  
gi|141901632|gb|ECU04596.1| hypothetical protein G ( 291) 943 215.8 1.3e-52  
gi|116124307|gb|ABJ775578.1| 3-phosphoshikimate 1-c ( 440) 943 216.0 1.8e-52  
gi|116121469|gb|ABJ79512.1| 3-phosphoshikimate 1-c ( 440) 943 216.0 1.8e-52  
gi|138056886|gb|EBX50330.1| hypothetical protein G ( 317) 940 215.2 2.2e-52  
gi|143505071|gb|EDF35219.1| hypothetical protein G ( 392) 941 215.5 2.3e-52  
gi|256998702|gb|EEU85222.1| 3-phosphoshikimate 1-c ( 428) 939 215.1 3.3e-52  
gi|255968720|gb|EET99342.1| 3-phosphoshikimate 1-c ( 428) 939 215.1 3.3e-52  
gi|142439106|gb|ECY08464.1| hypothetical protein G ( 285) 936 214.3 3.7e-52

gi|40806387|gb|AAR92125.1| 5-enolpyruvylshikimate- ( 426) 938 214.9 3.8e-52  
gi|217336546|gb|ACK42339.1| 3-phosphoshikimate 1-c ( 433) 937 214.7 4.5e-52  
gi|142830212|gb|EDA90204.1| hypothetical protein G ( 434) 936 214.5 5.2e-52  
gi|135920803|gb|EBK71830.1| hypothetical protein G ( 374) 935 214.2 5.4e-52  
gi|257164438|gb|EEU94398.1| 3-phosphoshikimate 1-c ( 428) 935 214.2 6e-52  
gi|255963810|gb|EET96286.1| 3-phosphoshikimate 1-c ( 428) 935 214.2 6e-52  
gi|257157926|gb|EEU87886.1| 3-phosphoshikimate 1-c ( 428) 935 214.2 6e-52  
gi|142615026|gb|ECZ34221.1| hypothetical protein G ( 435) 935 214.2 6e-52  
gi|136339509|gb|EBN49907.1| hypothetical protein G ( 386) 934 214.0 6.4e-52  
gi|145691253|gb|ABP91758.1| 5-enolpyruvylshikimate ( 426) 934 214.0 6.9e-52  
gi|256711589|gb|EEU26627.1| 3-phosphoshikimate 1-c ( 428) 934 214.0 6.9e-52  
gi|142274829|gb|ECW89740.1| hypothetical protein G ( 354) 933 213.7 7e-52  
gi|29343591|gb|AAO81353.1| 3-phosphoshikimate 1-ca ( 428) 933 213.8 8.1e-52  
gi|256993208|gb|EEU80510.1| 3-phosphoshikimate 1-c ( 428) 933 213.8 8.1e-52  
gi|144993705|gb|ABP15206.1| Sequence 45 from paten ( 428) 933 213.8 8.1e-52  
gi|12407967|gb|AAG53678.1|AF318277\_6 3-phosphoshik ( 428) 933 213.8 8.1e-52  
gi|257161129|gb|EEU91089.1| 3-phosphoshikimate 1-c ( 428) 933 213.8 8.1e-52  
gi|256990575|gb|EBN77877.1| 3-phosphoshikimate 1-c ( 428) 933 213.8 8.1e-52  
gi|256598595|gb|EEU17771.1| 3-phosphoshikimate 1-c ( 428) 933 213.8 8.1e-52  
gi|40130701|gb|AAR58446.1| Sequence 4980 from pate ( 430) 933 213.8 8.1e-52  
gi|142661801|gb|ECZ67186.1| hypothetical protein G ( 360) 932 213.5 8.2e-52  
gi|256994769|gb|EEU82071.1| 3-phosphoshikimate 1-c ( 428) 932 213.6 9.4e-52  
gi|2484165|gb|AAB72301.1|I49194 Sequence 42 from p ( 428) 930 213.1 1.3e-51  
gi|143816|gb|AAA20869.1| AroE [Bacillus subtilis] ( 428) 930 213.1 1.3e-51  
gi|2634678|emb|CAB14176.1| 3-phosphoshikimate 1-ca ( 428) 930 213.1 1.3e-51  
gi|167296372|gb|ABZ49236.1| Sequence 23174 from pa ( 428) 930 213.1 1.3e-51  
gi|5957565|gb|AAE08239.1| Sequence 42 from patent ( 428) 930 213.1 1.3e-51  
gi|144974744|gb|ABP12455.1| Sequence 42 from paten ( 428) 930 213.1 1.3e-51  
gi|2485242|gb|AAB73378.1|I44467 Sequence 42 from p ( 428) 930 213.1 1.3e-51  
gi|197053765|gb|ACH25463.1| Sequence 15 from paten ( 428) 930 213.1 1.3e-51  
gi|135415884|gb|EBH48842.1| hypothetical protein G ( 390) 929 212.9 1.4e-51  
gi|142583289|gb|ECZ12035.1| hypothetical protein G ( 407) 929 212.9 1.4e-51  
gi|256952418|gb|EEU69050.1| 3-phosphoshikimate 1-c ( 428) 929 212.9 1.5e-51  
gi|142459624|gb|ECY23468.1| hypothetical protein G ( 287) 926 212.1 1.7e-51  
gi|256987046|gb|EEU74348.1| 3-phosphoshikimate 1-c ( 428) 928 212.7 1.7e-51  
gi|256955616|gb|EEU72248.1| 3-phosphoshikimate 1-c ( 428) 928 212.7 1.7e-51  
gi|256949481|gb|EEU66113.1| 3-phosphoshikimate 1-c ( 428) 928 212.7 1.7e-51  
gi|256683745|gb|EEU23440.1| 3-phosphoshikimate 1-c ( 428) 928 212.7 1.7e-51  
gi|142340264|gb|ECX38084.1| hypothetical protein G ( 383) 926 212.2 2.1e-51  
gi|134931800|gb|EBE38163.1| hypothetical protein G ( 352) 923 211.6 3.1e-51  
gi|142736909|gb|EBE21008.1| hypothetical protein G ( 372) 920 210.9 5.1e-51  
gi|137336724|gb|EBT52837.1| hypothetical protein G ( 305) 918 210.4 6e-51  
gi|137500819|gb|EBU43511.1| hypothetical protein G ( 307) 918 210.4 6e-51  
gi|141275643|gb|ECQ84034.1| hypothetical protein G ( 251) 915 209.7 8.1e-51  
gi|140262147|gb|ECK63065.1| hypothetical protein G ( 251) 911 208.8 1.5e-50  
gi|142241906|gb|ECW65482.1| hypothetical protein G ( 400) 913 209.4 1.6e-50  
gi|134555734|gb|EBU04319.1| hypothetical protein G ( 344) 912 209.2 1.6e-50  
gi|22533648|gb|AAM99526.1|AE014219\_18 3-phosphoshi ( 427) 912 209.2 1.9e-50  
gi|76563325|gb|ABA45909.1| 3-phosphoshikimate 1-ca ( 427) 912 209.2 1.9e-50  
gi|136313640|gb|EBN33252.1| hypothetical protein G ( 435) 912 209.2 1.9e-50  
gi|142307360|gb|ECX13668.1| hypothetical protein G ( 436) 912 209.2 1.9e-50  
gi|134892211|gb|EBE11877.1| hypothetical protein G ( 363) 911 209.0 2e-50  
gi|140523261|gb|ECM09716.1| hypothetical protein G ( 291) 909 208.4 2.2e-50  
gi|138494665|gb|EBZ99748.1| hypothetical protein G ( 269) 908 208.2 2.5e-50  
gi|135903877|gb|EBK60098.1| hypothetical protein G ( 419) 909 208.6 3e-50  
gi|23095028|emb|CAP46254.1| Unknown [Streptococcus ( 427) 907 208.1 4.1e-50  
gi|134898476|gb|EBE16051.1| hypothetical protein G ( 405) 906 207.9 4.5e-50  
gi|135056874|gb|EBF21645.1| hypothetical protein G ( 329) 903 207.2 6.1e-50  
gi|254045793|gb|ACT62586.1| 3-phosphoshikimate 1-c ( 432) 904 207.5 6.5e-50  
gi|143452946|gb|EDF04465.1| hypothetical protein G ( 281) 900 206.5 8.5e-50  
gi|28271504|emb|CAP64409.1| 3-phosphoshikimate 1-c ( 432) 901 206.8 1e-49  
gi|143396673|gb|EDE73613.1| hypothetical protein G ( 356) 899 206.3 1.2e-49

gi|138568480|gb|ECA49921.1| hypothetical protein G ( 294) 897 205.8 1.4e-49  
gi|142577643|gb|ECZ08109.1| hypothetical protein G ( 416) 898 206.2 1.6e-49  
gi|158140591|gb|ABW18903.1| 3-phosphoshikimate 1-c ( 427) 898 206.2 1.6e-49  
gi|135491965|gb|EBH99300.1| hypothetical protein G ( 289) 896 205.6 1.6e-49  
gi|142246914|gb|ECW69181.1| hypothetical protein G ( 296) 896 205.6 1.6e-49  
gi|142808318|gb|EDA73669.1| hypothetical protein G ( 313) 896 205.6 1.7e-49  
gi|134965637|gb|EBB60728.1| hypothetical protein G ( 399) 897 205.9 1.7e-49  
gi|142365697|gb|ECX55533.1| hypothetical protein G ( 395) 896 205.7 2e-49  
gi|139723826|gb|ECH03396.1| hypothetical protein G ( 275) 894 205.2 2.1e-49  
gi|143174862|gb|EDD35554.1| hypothetical protein G ( 416) 896 205.7 2.1e-49  
gi|140985559|gb|ECO85428.1| hypothetical protein G ( 269) 893 204.9 2.4e-49  
gi|143031825|gb|EDC31489.1| hypothetical protein G ( 416) 895 205.5 2.4e-49  
gi|143138582|gb|EDD09641.1| hypothetical protein G ( 313) 892 204.8 3.1e-49  
gi|136832462|gb|EBQ70477.1| hypothetical protein G ( 371) 892 204.8 3.5e-49  
gi|229263473|gb|ACAQ54506.1| 3-phosphoshikimate 1-c ( 442) 890 204.4 5.4e-49  
gi|134860307|gb|EBD90420.1| hypothetical protein G ( 369) 889 204.2 5.5e-49  
gi|221240501|gb|ACM13211.1| 3-phosphoshikimate 1-c ( 399) 889 204.2 5.8e-49  
gi|138397559|gb|EBZ03870.1| hypothetical protein G ( 283) 887 203.6 6.1e-49  
gi|152928818|gb|ABS34318.1| 3-phosphoshikimate 1-c ( 442) 887 203.8 8.6e-49  
gi|152930690|gb|ABS36189.1| 3-phosphoshikimate 1-c ( 442) 887 203.8 8.6e-49  
gi|148288929|emb|CAL83016.1| 3-phosphoshikimate 1-c ( 442) 887 203.8 8.6e-49  
gi|226841906|gb|ACO84572.1| 3-phosphoshikimate 1-c ( 442) 886 203.6 9.9e-49  
gi|142654970|gb|ECZ62388.1| hypothetical protein G ( 433) 883 202.9 1.5e-48  
gi|169120913|gb|ACA44749.1| 3-phosphoshikimate 1-c ( 442) 883 202.9 1.6e-48  
gi|143688892|gb|EDG29042.1| hypothetical protein G ( 432) 882 202.7 1.8e-48  
gi|136830361|gb|EBQ69080.1| hypothetical protein G ( 400) 880 202.2 2.3e-48  
gi|142014592|gb|ECU84573.1| hypothetical protein G ( 416) 880 202.2 2.3e-48  
gi|125714526|gb|ABN53018.1| 3-phosphoshikimate 1-c ( 423) 879 202.0 2.8e-48  
gi|116090991|gb|ABJ56145.1| 3-phosphoshikimate 1-c ( 437) 879 202.0 2.8e-48  
gi|169407871|gb|ACA56282.1| 3-phosphoshikimate 1-c ( 442) 878 201.8 3.3e-48  
gi|136052580|gb|EBL59062.1| hypothetical protein G ( 390) 877 201.6 3.5e-48  
gi|144136414|gb|EDJ12922.1| hypothetical protein G ( 485) 878 201.9 3.6e-48  
gi|140596163|gb|ECM33153.1| hypothetical protein G ( 280) 875 201.0 3.7e-48  
gi|152936571|gb|ABS42069.1| 3-phosphoshikimate 1-c ( 442) 876 201.4 4.5e-48  
gi|167041247|gb|ABZ06003.1| putative EPSP synthase ( 446) 876 201.4 4.5e-48  
gi|142703421|gb|ECZ96994.1| hypothetical protein G ( 426) 874 200.9 5.9e-48  
gi|142560595|gb|ECY96203.1| hypothetical protein G ( 308) 871 200.2 7.2e-48  
gi|134974804|gb|EBY66929.1| hypothetical protein G ( 392) 872 200.5 7.5e-48  
gi|135727138|gb|EBJ46737.1| hypothetical protein G ( 331) 871 200.2 7.6e-48  
gi|138292864|gb|EBY83547.1| hypothetical protein G ( 260) 869 199.7 8.6e-48  
gi|142707532|gb|ECZ99900.1| hypothetical protein G ( 416) 871 200.3 9.1e-48  
gi|135630574|gb|EBI87001.1| hypothetical protein G ( 283) 869 199.7 9.2e-48  
gi|143510416|gb|EDF38005.1| hypothetical protein G ( 415) 869 199.9 1.2e-47  
gi|134855951|gb|EBD87629.1| hypothetical protein G ( 254) 865 198.8 1.5e-47  
gi|143906379|gb|EDH48348.1| hypothetical protein G ( 599) 869 200.0 1.6e-47  
gi|138461639|gb|EBZ83709.1| hypothetical protein G ( 290) 865 198.9 1.7e-47  
gi|143129798|gb|EDD03205.1| hypothetical protein G ( 431) 867 199.4 1.7e-47  
gi|142417854|gb|ECX92494.1| hypothetical protein G ( 660) 868 199.8 2e-47  
gi|116229535|gb|ABJ88244.1| 3-phosphoshikimate 1-c ( 426) 865 199.0 2.3e-47  
gi|142699976|gb|ECZ94523.1| hypothetical protein G ( 326) 863 198.5 2.5e-47  
gi|142019123|gb|ECU88920.1| hypothetical protein G ( 414) 864 198.8 2.6e-47  
gi|143234451|gb|EDD78785.1| hypothetical protein G ( 366) 863 198.5 2.8e-47  
gi|137267823|gb|EBT14391.1| hypothetical protein G ( 276) 861 198.0 3e-47  
gi|206741482|gb|ACI20539.1| 3-phosphoshikimate 1-c ( 430) 863 198.6 3.1e-47  
gi|135448115|gb|EBH70436.1| hypothetical protein G ( 383) 860 197.9 4.5e-47  
gi|136710153|gb|EBP89355.1| hypothetical protein G ( 272) 856 196.9 6.3e-47  
gi|143575506|gb|EDF71556.1| hypothetical protein G ( 358) 857 197.2 6.7e-47  
gi|143445598|gb|EDE99944.1| hypothetical protein G ( 366) 856 197.0 7.9e-47  
gi|141751437|gb|ECT10385.1| hypothetical protein G ( 240) 853 196.2 9e-47  
gi|167045751|gb|ABZ10397.1| putative EPSP synthase ( 443) 856 197.0 9.2e-47  
gi|142502060|gb|ECY54068.1| hypothetical protein G ( 274) 852 196.0 1.2e-46  
gi|222120489|dbj|BAH17824.1| 3-phosphoshikimate 1-c ( 433) 854 196.6 1.2e-46

gi|135770651|gb|EBJ73691.1| hypothetical protein G ( 349) 852 196.1 1.4e-46  
gi|141716480|gb|ECZ90978.1| hypothetical protein G ( 351) 852 196.1 1.4e-46  
gi|28203719|gb|AAO36160.1| 3-phosphoshikimate 1-ca ( 439) 853 196.4 1.4e-46  
gi|161726857|emb|CAP47298.1| unnamed protein produ ( 439) 853 196.4 1.4e-46  
gi|140153247|gb|ECU87824.1| hypothetical protein G ( 307) 850 195.6 1.7e-46  
gi|163861631|gb|ABY42690.1| 3-phosphoshikimate 1-c ( 367) 850 195.7 2e-46  
gi|225793440|gb|ACO33530.1| 3-phosphoshikimate 1-c ( 433) 847 195.1 3.5e-46  
gi|138199350|gb|EBY35873.1| hypothetical protein G ( 281) 844 194.3 4e-46  
gi|142236410|gb|ECW61386.1| hypothetical protein G ( 315) 844 194.3 4.3e-46  
gi|229470142|gb|ACQ71914.1| 3-phosphoshikimate 1-c ( 416) 845 194.6 4.6e-46  
gi|142521472|gb|ECY68161.1| hypothetical protein G ( 443) 845 194.6 4.8e-46  
gi|138683081|gb|ECB28394.1| hypothetical protein G ( 246) 841 193.6 5.6e-46  
gi|135924801|gb|EBK74524.1| hypothetical protein G ( 368) 843 194.2 5.7e-46  
gi|14112363|gb|AAE57780.1| Sequence 4 from patent ( 415) 843 194.2 6.2e-46  
gi|14106226|gb|AAE55841.1| Sequence 4 from patent ( 415) 843 194.2 6.2e-46  
gi|143143221|gb|EDD13054.1| hypothetical protein G ( 343) 841 193.7 7.3e-46  
gi|171990323|gb|ACB61245.1| 3-phosphoshikimate 1-c ( 417) 842 194.0 7.3e-46  
gi|142975716|gb|ECW61669.1| hypothetical protein G ( 334) 840 193.5 8.3e-46  
gi|144047589|gb|EDT48620.1| hypothetical protein G ( 354) 840 193.5 8.7e-46  
gi|143651447|gb|EDG07727.1| hypothetical protein G ( 334) 839 193.3 9.6e-46  
gi|139093977|gb|ECB237104.1| hypothetical protein G ( 289) 838 193.0 1e-45  
gi|143260329|gb|EDD96800.1| hypothetical protein G ( 256) 837 192.7 1.1e-45  
gi|135647562|gb|EBI97552.1| hypothetical protein G ( 338) 836 192.6 1.5e-45  
gi|221571690|gb|ACM22502.1| 3-phosphoshikimate 1-c ( 421) 837 192.9 1.6e-45  
gi|147735256|gb|ABQ46596.1| 3-phosphoshikimate 1-c ( 421) 837 192.9 1.6e-45  
gi|214040543|gb|EEB81191.1| 3-phosphoshikimate 1-c ( 421) 837 192.9 1.6e-45  
gi|170175891|gb|ACB08943.1| 3-phosphoshikimate 1-c ( 421) 837 192.9 1.6e-45  
gi|138880837|gb|ECC27684.1| hypothetical protein G ( 257) 834 192.1 1.7e-45  
gi|4980846|gb|AAD35431.1|AE001715\_7 3-phosphoshiki ( 410) 835 192.4 2.1e-45  
gi|167276129|gb|ABZ28993.1| Sequence 2931 from pat ( 410) 835 192.4 2.1e-45  
gi|141303260|gb|ECQ94812.1| hypothetical protein G ( 320) 833 191.9 2.3e-45  
gi|27315616|gb|AAO04750.1|AE016747\_247 3-phosphosh ( 433) 833 192.0 2.9e-45  
gi|134562359|gb|EBC08340.1| hypothetical protein G ( 318) 831 191.5 3.1e-45  
gi|135489309|gb|EBH97580.1| hypothetical protein G ( 350) 830 191.3 3.9e-45  
gi|68447173|dbj|BAE04757.1| 3-phosphoshikimate 1-c ( 432) 830 191.4 4.6e-45  
gi|142087282|gb|ECV48343.1| hypothetical protein G ( 425) 829 191.2 5.2e-45  
gi|57637630|gb|AAW54418.1| 3-phosphoshikimate 1-ca ( 433) 829 191.2 5.3e-45  
gi|217109301|gb|ACJ93181.1| Sequence 3223 from pat ( 436) 829 191.2 5.3e-45  
gi|281043957|gb|ADA22685.1| Sequence 3223 from pat ( 436) 829 191.2 5.3e-45  
gi|259269376|gb|ACW29189.1| Sequence 3223 from pat ( 436) 829 191.2 5.3e-45  
gi|143466480|gb|EDF13416.1| hypothetical protein G ( 273) 825 190.1 6.8e-45  
gi|207084714|gb|EDZ62001.1| 3-phosphoshikimate 1-c ( 427) 827 190.7 7.1e-45  
gi|135095629|gb|EBF46390.1| hypothetical protein G ( 225) 823 189.6 7.9e-45  
gi|257271987|gb|ECV94125.1| 3-phosphoshikimate 1-c ( 432) 826 190.5 8.3e-45  
gi|257275030|gb|EEV06517.1| 3-phosphoshikimate 1-c ( 432) 826 190.5 8.3e-45  
gi|257278947|gb|EEV09566.1| 3-phosphoshikimate 1-c ( 432) 826 190.5 8.3e-45  
gi|257281670|gb|EEV11807.1| 3-phosphoshikimate 1-c ( 432) 826 190.5 8.3e-45  
gi|49241782|emb|CAG40473.1| 3-phosphoshikimate 1-c ( 432) 826 190.5 8.3e-45  
gi|257284701|gb|EEV14821.1| 3-phosphoshikimate 1-c ( 432) 826 190.5 8.3e-45  
gi|138508120|gb|ECA08043.1| hypothetical protein G ( 264) 823 189.7 9e-45  
gi|142552699|gb|ECY90674.1| hypothetical protein G ( 425) 825 190.3 9.6e-45  
gi|142973132|gb|EDB89862.1| hypothetical protein G ( 258) 822 189.5 1e-44  
gi|139178038|gb|ECV94124.1| hypothetical protein G ( 315) 822 189.5 1.2e-44  
gi|139089271|gb|ECD33698.1| hypothetical protein G ( 268) 821 189.3 1.2e-44  
gi|142941353|gb|EDB67616.1| hypothetical protein G ( 419) 823 189.8 1.3e-44  
gi|143051067|gb|EDC45621.1| hypothetical protein G ( 338) 821 189.3 1.5e-44  
gi|135388286|gb|EBH30252.1| hypothetical protein G ( 247) 819 188.8 1.6e-44  
gi|143823138|gb|EDG88250.1| hypothetical protein G ( 277) 819 188.8 1.7e-44  
gi|158604933|gb|ABW74752.1| 3-phosphoshikimate 1-c ( 428) 821 189.4 1.8e-44  
gi|143990770|gb|EDI08126.1| hypothetical protein G ( 321) 819 188.9 1.9e-44  
gi|82656594|emb|CAT81017.1| 3-phosphoshikimate 1-c ( 432) 820 189.2 2.1e-44  
gi|143505131|gb|EDF35248.1| hypothetical protein G ( 381) 819 188.9 2.2e-44

gi|135341079|gb|EBG98573.1| hypothetical protein G ( 269) 817 188.4 2.3e-44  
gi|136274212|gb|EBN05767.1| hypothetical protein G ( 343) 817 188.5 2.7e-44  
gi|136878306|gb|EBR00609.1| hypothetical protein G ( 266) 815 188.0 3e-44  
gi|139609260|gb|ECG24653.1| hypothetical protein G ( 267) 815 188.0 3e-44  
gi|142963473|gb|EDB83252.1| hypothetical protein G ( 432) 817 188.5 3.2e-44  
gi|141829663|gb|ECT53965.1| hypothetical protein G ( 306) 815 188.0 3.4e-44  
gi|140734279|gb|ECN15117.1| hypothetical protein G ( 219) 813 187.5 3.5e-44  
gi|134338102|gb|EBA73894.1| hypothetical protein G ( 291) 814 187.8 3.8e-44  
gi|139248991|gb|ECE35883.1| hypothetical protein G ( 293) 814 187.8 3.8e-44  
gi|143533533|gb|EDF49839.1| hypothetical protein G ( 445) 816 188.3 3.9e-44  
gi|137063356|gb|EBS00002.1| hypothetical protein G ( 249) 813 187.5 3.9e-44  
gi|262080109|gb|ACY16078.1| 3-phosphoshikimate 1-c ( 459) 816 188.3 4e-44  
gi|189341744|gb|ACD91147.1| 3-phosphoshikimate 1-c ( 434) 815 188.1 4.4e-44  
gi|143906650|gb|EDH48541.1| hypothetical protein G ( 662) 817 188.7 4.5e-44  
gi|138329239|gb|EBZ00337.1| hypothetical protein G ( 321) 813 187.6 4.7e-44  
gi|150374387|dbj|BAF67647.1| 3-phosphoshikimate 1- ( 432) 814 187.9 5.1e-44  
gi|269940957|emb|CBI49341.1| 3-phosphoshikimate 1- ( 432) 814 187.9 5.1e-44  
gi|87202756|gb|ABD30566.1| 3-phosphoshikimate 1-ca ( 432) 814 187.9 5.1e-44  
gi|49244747|emb|CAG43183.1| 3-phosphoshikimate 1-c ( 432) 814 187.9 5.1e-44  
gi|21204523|dbj|BAB95219.1| 3-PHOSPHOSHIKIMATE 1-C ( 432) 814 187.9 5.1e-44  
gi|87125900|gb|ABD20414.1| 3-phosphoshikimate 1-ca ( 432) 814 187.9 5.1e-44  
gi|160368441|gb|ABX29412.1| 3-phosphoshikimate 1-c ( 432) 814 187.9 5.1e-44  
gi|281032786|gb|ADA18062.1| Sequence 5610 from pat ( 435) 814 187.9 5.1e-44  
gi|144188786|gb|EDJ51597.1| hypothetical protein G ( 603) 815 188.2 5.7e-44  
gi|57284605|gb|AAW36699.1| 3-phosphoshikimate 1-ca ( 432) 813 187.7 5.9e-44  
gi|136995971|gb|EBR61979.1| hypothetical protein G ( 325) 811 187.1 6.4e-44  
gi|136354113|gb|EBN59856.1| hypothetical protein G ( 283) 810 186.9 6.7e-44  
gi|138555285|gb|ECA40675.1| hypothetical protein G ( 267) 809 186.6 7.5e-44  
gi|156721918|dbj|BAF78335.1| 3-phosphoshikimate 1- ( 432) 811 187.2 8e-44  
gi|262075379|gb|ACY11352.1| 3-phosphoshikimate 1-c ( 432) 811 187.2 8e-44  
gi|14247235|dbj|BAB57626.1| 3-phosphoshikimate 1-c ( 432) 811 187.2 8e-44  
gi|13701263|dbj|BAB42557.1| 3-phosphoshikimate 1-c ( 432) 811 187.2 8e-44  
gi|143913100|gb|EDH52995.1| hypothetical protein G ( 479) 811 187.3 8.7e-44  
gi|135176472|gb|EBF98299.1| hypothetical protein G ( 278) 808 186.4 9e-44  
gi|142344653|gb|ECX41282.1| hypothetical protein G ( 430) 809 186.8 1.1e-43  
gi|194312559|gb|ACF46954.1| 3-phosphoshikimate 1-c ( 434) 809 186.8 1.1e-43  
gi|142888789|gb|EDB29602.1| hypothetical protein G ( 430) 808 186.6 1.3e-43  
gi|189496532|gb|ACE05080.1| 3-phosphoshikimate 1-c ( 435) 808 186.6 1.3e-43  
gi|134547100|gb|EBB99141.1| hypothetical protein G ( 336) 806 186.1 1.4e-43  
gi|142309197|gb|ECX15063.1| hypothetical protein G ( 340) 806 186.1 1.4e-43  
gi|142383766|gb|ECX67542.1| hypothetical protein G ( 430) 807 186.4 1.5e-43  
gi|78497609|gb|ABB44149.1| 3-phosphoshikimate 1-ca ( 428) 806 186.1 1.7e-43  
gi|135033333|gb|EBF06615.1| hypothetical protein G ( 414) 804 185.7 2.2e-43  
gi|144199065|gb|EDJ59187.1| hypothetical protein G ( 246) 801 184.9 2.3e-43  
gi|260078368|gb|EEW66073.1| LOW QUALITY PROTEIN: 3 ( 266) 801 184.9 2.5e-43  
gi|140334519|gb|ECL05829.1| hypothetical protein G ( 270) 800 184.7 2.9e-43  
gi|138262365|gb|EBY65200.1| hypothetical protein G ( 238) 799 184.4 3.1e-43  
gi|134990285|gb|EBE77443.1| hypothetical protein G ( 300) 800 184.7 3.2e-43  
gi|137254189|gb|EBT06708.1| hypothetical protein G ( 285) 799 184.5 3.6e-43  
gi|144219781|gb|EDJ73727.1| hypothetical protein G ( 428) 800 184.8 4.2e-43  
gi|149946465|gb|ABR52401.1| 3-phosphoshikimate 1-c ( 432) 800 184.8 4.2e-43  
gi|147741019|gb|ABQ49317.1| 3-phosphoshikimate 1-c ( 432) 800 184.8 4.2e-43  
gi|151422271|dbj|BAF69775.1| 3-phosphoshikimate 1- ( 433) 800 184.8 4.2e-43  
gi|137335428|gb|EBT52086.1| hypothetical protein G ( 303) 798 184.3 4.3e-43  
gi|135781046|gb|EBJ80223.1| hypothetical protein G ( 308) 798 184.3 4.4e-43  
gi|135266853|gb|EBG52393.1| hypothetical protein G ( 424) 799 184.6 4.8e-43  
gi|139490738|gb|ECF43051.1| hypothetical protein G ( 279) 796 183.8 5.5e-43  
gi|143414697|gb|EDE83057.1| hypothetical protein G ( 429) 798 184.4 5.7e-43  
gi|136328373|gb|EBN42356.1| hypothetical protein G ( 293) 796 183.8 5.7e-43  
gi|119355398|gb|ABL66269.1| 3-phosphoshikimate 1-c ( 434) 798 184.4 5.7e-43  
gi|141355624|gb|ECR25075.1| hypothetical protein G ( 303) 795 183.6 6.8e-43  
gi|143971097|gb|EDH94528.1| hypothetical protein G ( 570) 798 184.5 7.1e-43

gi|135098087|gb|EBF47962.1| hypothetical protein G ( 222) 793 183.1 7.3e-43  
gi|137854115|gb|EBW38016.1| hypothetical protein G ( 286) 793 183.2 8.8e-43  
gi|141976830|gb|ECU56929.1| hypothetical protein G ( 330) 793 183.2 9.9e-43  
gi|142680915|gb|ECZ80867.1| hypothetical protein G ( 338) 793 183.2 1e-42  
gi|110742505|dbj|BAE99170.1| 5-enolpyruvylshikimat ( 520) 795 183.8 1e-42  
gi|2583124|gb|AAB82633.1| 5-enolpyruvylshikimate-3 ( 520) 795 183.8 1e-42  
gi|295790|emb|CAA29828.1| EPSP [Arabidopsis thalia ( 520) 795 183.8 1e-42  
gi|63003818|gb|AAZ52438.1| At2g45300 [Arabidopsis ( 520) 795 183.8 1e-42  
gi|62786929|gb|AAZ06808.1| Sequence 2 from patent ( 520) 795 183.8 1e-42  
gi|141101489|gb|ECP63549.1| hypothetical protein G ( 294) 792 183.0 1e-42  
gi|141097662|gb|ECP60818.1| hypothetical protein G ( 301) 792 183.0 1.1e-42  
gi|137395150|gb|EBT85818.1| hypothetical protein G ( 319) 792 183.0 1.1e-42  
gi|142090737|gb|ECV51207.1| hypothetical protein G ( 425) 793 183.3 1.2e-42  
gi|139921163|gb|ECI38192.1| hypothetical protein G ( 235) 789 182.3 1.4e-42  
gi|194310012|gb|ACF44712.1| 3-phosphoshikimate 1-c ( 433) 792 183.1 1.4e-42  
gi|34483387|emb|CABE10385.1| 3-PHOSPHOSHIKIMATE 1-C ( 437) 792 183.1 1.4e-42  
gi|142643653|gb|ECZ54421.1| hypothetical protein G ( 339) 790 182.6 1.6e-42  
gi|142006350|gb|ECU77442.1| hypothetical protein G ( 324) 789 182.4 1.8e-42  
gi|139420983|gb|ECP99819.1| hypothetical protein G ( 337) 788 182.2 2.1e-42  
gi|111979102|gb|ABH83339.1| Sequence 8 from patent ( 520) 790 182.7 2.2e-42  
gi|21647940|gb|AAM73137.1| 3-phosphoshikimate 1-ca ( 434) 789 182.5 2.2e-42  
gi|144105457|gb|EDT90364.1| hypothetical protein G ( 242) 786 181.6 2.2e-42  
gi|94552691|gb|ABF42615.1| 3-phosphoshikimate 1-ca ( 435) 789 182.5 2.2e-42  
gi|253510676|gb|EES89335.1| 5-enolpyruvylshikimate ( 436) 789 182.5 2.2e-42  
gi|143978542|gb|EDH99586.1| hypothetical protein G ( 260) 786 181.6 2.4e-42  
gi|140902672|gb|ECO27576.1| hypothetical protein G ( 290) 786 181.7 2.6e-42  
gi|135918291|gb|EBK70077.1| hypothetical protein G ( 307) 785 181.5 3.1e-42  
gi|142317432|gb|ECX21250.1| hypothetical protein G ( 454) 786 181.8 3.6e-42  
gi|139662994|gb|ECG60888.1| hypothetical protein G ( 279) 782 180.8 4.6e-42  
gi|135473041|gb|EBH87124.1| hypothetical protein G ( 425) 783 181.1 5.4e-42  
gi|143880166|gb|EDH29665.1| hypothetical protein G ( 333) 781 180.6 6.1e-42  
gi|142062898|gb|ECV27877.1| hypothetical protein G ( 430) 782 180.9 6.4e-42  
gi|193085440|gb|ACF10716.1| 3-phosphoshikimate 1-c ( 434) 782 180.9 6.4e-42  
gi|143263748|gb|EDD99246.1| hypothetical protein G ( 430) 781 180.7 7.4e-42  
gi|256008713|gb|ACU54280.1| 3-phosphoshikimate 1-c ( 434) 781 180.7 7.4e-42  
gi|32261686|gb|AAP76736.1| 3-phosphoshikimate 1-ca ( 438) 781 180.7 7.5e-42  
gi|136211928|gb|EBM63318.1| hypothetical protein G ( 333) 779 180.8 8.2e-42  
gi|281108575|gb|ADA52833.1| Sequence 8 from patent ( 520) 781 180.8 8.6e-42  
gi|139841346|gb|ECH83875.1| hypothetical protein G ( 258) 777 179.7 9.1e-42  
gi|139163008|gb|ECB83628.1| hypothetical protein G ( 312) 777 179.7 1.1e-41  
gi|135319394|gb|EBG83991.1| hypothetical protein G ( 333) 777 179.8 1.1e-41  
gi|138154978|gb|EBY06220.1| hypothetical protein G ( 209) 773 178.7 1.4e-41  
gi|135240518|gb|EBG36889.1| hypothetical protein G ( 311) 775 179.3 1.4e-41  
gi|138397214|gb|EBZ38470.1| hypothetical protein G ( 334) 775 179.3 1.5e-41  
gi|144025250|gb|EDI32680.1| hypothetical protein G ( 308) 774 179.1 1.6e-41  
gi|141897414|gb|ECU01571.1| hypothetical protein G ( 305) 773 178.9 1.9e-41  
gi|144075303|gb|EDT68980.1| hypothetical protein G ( 474) 775 179.4 2e-41  
gi|136622362|gb|EBP34455.1| hypothetical protein G ( 328) 772 178.7 2.3e-41  
gi|140030681|gb|ECU12144.1| hypothetical protein G ( 246) 770 178.1 2.5e-41  
gi|142893869|gb|EDB33252.1| hypothetical protein G ( 557) 774 179.3 2.6e-41  
gi|138700831|gb|ECB40892.1| hypothetical protein G ( 314) 771 178.4 2.6e-41  
gi|139949518|gb|ECI57412.1| hypothetical protein G ( 283) 770 178.2 2.8e-41  
gi|144037176|gb|EDT41070.1| hypothetical protein G ( 544) 773 179.0 3e-41  
gi|135329585|gb|EBG90868.1| hypothetical protein G ( 450) 772 178.8 3e-41  
gi|134947939|gb|EBB48938.1| hypothetical protein G ( 308) 769 178.0 3.5e-41  
gi|134644279|gb|EBC57396.1| hypothetical protein G ( 256) 768 177.7 3.5e-41  
gi|135183725|gb|EBG02913.1| hypothetical protein G ( 265) 768 177.7 3.6e-41  
gi|143093399|gb|EDC76589.1| hypothetical protein G ( 228) 767 177.3 3.7e-41  
gi|2484166|gb|AAB72302.1|I49195 Sequence 44 from p ( 430) 770 178.5 3.9e-41  
gi|144974745|gb|ABP12456.1| Sequence 44 from paten ( 430) 770 178.3 3.9e-41  
gi|5957566|gb|AAE08240.1| Sequence 44 from patent ( 430) 770 178.3 3.9e-41  
gi|2485243|gb|AAB73379.1|I44468 Sequence 44 from p ( 430) 770 178.3 3.9e-41

gi|152956|gb|AAA71897.1| 3-phosphoshikimate-1-carb ( 430) 770 178.3 3.9e-41  
gi|142011761|gb|ECU81815.1| hypothetical protein G ( 440) 770 178.3 4e-41  
gi|136955165|gb|EBR38869.1| hypothetical protein G ( 311) 767 177.6 4.8e-41  
gi|142971616|gb|EDB88788.1| hypothetical protein G ( 317) 767 177.6 4.8e-41  
gi|142331844|gb|ECX31901.1| hypothetical protein G ( 430) 768 177.9 5.3e-41  
gi|144212823|gb|EDJ68735.1| hypothetical protein G ( 445) 768 177.9 5.4e-41  
gi|143877405|gb|EDH27656.1| hypothetical protein G ( 447) 768 177.9 5.4e-41  
gi|142267558|gb|ECW84392.1| hypothetical protein G ( 313) 766 177.3 5.6e-41  
gi|142403788|gb|ECX81944.1| hypothetical protein G ( 333) 765 177.1 6.8e-41  
gi|138571481|gb|ECA52062.1| hypothetical protein G ( 245) 763 176.6 7.2e-41  
gi|135364074|gb|EBH13997.1| hypothetical protein G ( 441) 766 177.5 7.2e-41  
gi|142133901|gb|ECV83825.1| hypothetical protein G ( 430) 765 177.2 8.3e-41  
gi|145205328|gb|ABP36371.1| 3-phosphoshikimate 1-c ( 433) 765 177.2 8.3e-41  
gi|138840709|gb|ECC10949.1| hypothetical protein G ( 342) 763 176.7 9.4e-41  
gi|157700108|gb|ABV693268.1| 3-phosphoshikimate 1-c ( 425) 764 177.0 9.5e-41  
gi|142585981|gb|EC213884.1| hypothetical protein G ( 441) 764 177.0 9.8e-41  
gi|144051266|gb|EDI51354.1| hypothetical protein G ( 327) 762 176.5 1.1e-40  
gi|139284870|gb|ECE45110.1| hypothetical protein G ( 241) 760 175.9 1.1e-40  
gi|143992208|gb|EDI09178.1| hypothetical protein G ( 432) 761 176.4 1.5e-40  
gi|143696214|gb|EDG32946.1| hypothetical protein G ( 434) 761 176.4 1.5e-40  
gi|135647573|gb|EBI97559.1| hypothetical protein G ( 298) 759 175.8 1.5e-40  
gi|141935581|gb|ECU28288.1| hypothetical protein G ( 226) 757 175.3 1.7e-40  
gi|142656427|gb|ECZ63418.1| hypothetical protein G ( 445) 760 176.1 1.8e-40  
gi|143830277|gb|EBI97559.1| hypothetical protein G ( 449) 760 176.2 1.8e-40  
gi|137335317|gb|EBT52025.1| hypothetical protein G ( 306) 758 175.6 1.8e-40  
gi|141064495|gb|ECP38367.1| hypothetical protein G ( 306) 758 175.6 1.8e-40  
gi|139664047|gb|ECG61621.1| hypothetical protein G ( 258) 757 175.3 1.9e-40  
gi|134622016|gb|EBC44186.1| hypothetical protein G ( 314) 758 175.6 1.9e-40  
gi|78171788|gb|ABB28884.1| 3-phosphoshikimate 1-ca ( 435) 759 175.9 2.1e-40  
gi|134801092|gb|EBD52793.1| hypothetical protein G ( 446) 759 175.9 2.1e-40  
gi|139939562|gb|EC151111.1| hypothetical protein G ( 308) 757 175.4 2.1e-40  
gi|136821916|gb|EBQ63445.1| hypothetical protein G ( 319) 757 175.4 2.2e-40  
gi|143006221|gb|EDC12907.1| hypothetical protein G ( 448) 757 175.5 2.9e-40  
gi|134675387|gb|EBC75542.1| hypothetical protein G ( 306) 755 174.9 2.9e-40  
gi|140528515|gb|ECM11387.1| hypothetical protein G ( 318) 755 174.9 3e-40  
gi|134775377|gb|EBD34861.1| hypothetical protein G ( 406) 756 175.2 3.1e-40  
gi|136630657|gb|EBP39342.1| hypothetical protein G ( 335) 755 175.0 3.1e-40  
gi|141860322|gb|ECT75628.1| hypothetical protein G ( 283) 754 174.7 3.1e-40  
gi|136302090|gb|EBN24661.1| hypothetical protein G ( 430) 756 175.3 3.2e-40  
gi|142888313|gb|EDB29243.1| hypothetical protein G ( 321) 754 174.7 3.5e-40  
gi|141233117|gb|ECQ54347.1| hypothetical protein G ( 228) 751 174.0 4.2e-40  
gi|223588889|gb|ACM92625.1| 3-phosphoshikimate 1-c ( 431) 754 174.8 4.4e-40  
gi|135502733|gb|EBI06227.1| hypothetical protein G ( 271) 751 174.0 4.8e-40  
gi|141024369|gb|ECP11205.1| hypothetical protein G ( 274) 751 174.0 4.8e-40  
gi|138011480|gb|EBX26531.1| hypothetical protein G ( 292) 751 174.1 5.1e-40  
gi|134407854|gb|EBB17419.1| hypothetical protein G ( 293) 750 173.8 5.9e-40  
gi|138073518|gb|EBX59366.1| hypothetical protein G ( 289) 749 173.6 6.8e-40  
gi|135097328|gb|EBF47474.1| hypothetical protein G ( 314) 749 173.6 7.2e-40  
gi|141494611|gb|ECS06053.1| hypothetical protein G ( 234) 747 173.1 7.8e-40  
gi|138660172|gb|ECB12113.1| hypothetical protein G ( 276) 747 173.2 8.9e-40  
gi|141299754|gb|ECQ93545.1| hypothetical protein G ( 286) 747 173.2 9.1e-40  
gi|143240990|gb|EDD83029.1| hypothetical protein G ( 430) 749 173.7 9.2e-40  
gi|139617327|gb|ECG29769.1| hypothetical protein G ( 270) 746 172.9 1e-39  
gi|138602361|gb|ECA73437.1| hypothetical protein G ( 300) 746 173.0 1.1e-39  
gi|142294047|gb|ECX03786.1| hypothetical protein G ( 323) 746 173.0 1.2e-39  
gi|136830709|gb|EBQ69311.1| hypothetical protein G ( 327) 746 173.0 1.2e-39  
gi|135080385|gb|EBF36631.1| hypothetical protein G ( 253) 744 172.5 1.3e-39  
gi|140702862|gb|ECM93365.1| hypothetical protein G ( 286) 744 172.5 1.4e-39  
gi|138281250|gb|EBY78010.1| hypothetical protein G ( 302) 744 172.5 1.5e-39  
gi|139974108|gb|ECI74764.1| hypothetical protein G ( 304) 744 172.5 1.5e-39  
gi|143234938|gb|EDD79123.1| hypothetical protein G ( 270) 743 172.3 1.6e-39  
gi|138290616|gb|EBY82177.1| hypothetical protein G ( 295) 743 172.3 1.7e-39

gi|142666850|gb|ECZ70721.1| hypothetical protein G ( 334) 743 172.4 1.9e-39  
gi|141489620|gb|ECS04477.1| hypothetical protein G ( 301) 742 172.1 2e-39  
gi|139845582|gb|ECH86898.1| hypothetical protein G ( 305) 742 172.1 2e-39  
gi|142767462|gb|EDA43227.1| hypothetical protein G ( 445) 743 172.4 2.3e-39  
gi|138465274|gb|EB286300.1| hypothetical protein G ( 214) 739 171.3 2.4e-39  
gi|135744571|gb|EBJ57515.1| hypothetical protein G ( 301) 740 171.7 2.7e-39  
gi|140289895|gb|ECK77037.1| hypothetical protein G ( 307) 740 171.7 2.8e-39  
gi|141103301|gb|ECP64813.1| hypothetical protein G ( 309) 740 171.7 2.8e-39  
gi|136860453|gb|EBQ89188.1| hypothetical protein G ( 246) 736 170.7 4.3e-39  
gi|140955798|gb|ECO64828.1| hypothetical protein G ( 273) 736 170.8 4.6e-39  
gi|139768554|gb|ECH32813.1| hypothetical protein G ( 278) 736 170.8 4.7e-39  
gi|135103443|gb|EBF51389.1| hypothetical protein G ( 418) 738 171.3 4.8e-39  
gi|143449513|gb|EDF02359.1| hypothetical protein G ( 444) 738 171.4 5e-39  
gi|139134351|gb|ECD63724.1| hypothetical protein G ( 301) 736 170.8 5e-39  
gi|140036702|gb|ECJ16354.1| hypothetical protein G ( 301) 736 170.8 5e-39  
gi|143676349|gb|EDG21217.1| hypothetical protein G ( 446) 738 171.4 5e-39  
gi|140729862|gb|ECN11951.1| hypothetical protein G ( 304) 736 170.8 5e-39  
gi|140364603|gb|ECL26871.1| hypothetical protein G ( 219) 734 170.3 5.3e-39  
gi|139696294|gb|ECC83976.1| hypothetical protein G ( 287) 735 170.6 5.6e-39  
gi|135562484|gb|EBI44473.1| hypothetical protein G ( 293) 735 170.6 5.7e-39  
gi|142220950|gb|ECH49934.1| hypothetical protein G ( 747) 739 171.7 6.4e-39  
gi|134813827|gb|EBD60698.1| hypothetical protein G ( 286) 734 170.3 6.5e-39  
gi|222421192|emb|CAL28006.1| 3-phosphoshikimate 1- ( 432) 736 170.9 6.6e-39  
gi|142156622|gb|ECW00862.1| hypothetical protein G ( 222) 732 169.8 7.2e-39  
gi|141163007|gb|ECQ06591.1| hypothetical protein G ( 235) 732 169.8 7.5e-39  
gi|72495104|dbj|BAE18425.1| 5-enolpyruvylshikimate ( 432) 735 170.7 7.7e-39  
gi|193088545|gb|ECF13820.1| 3-phosphoshikimate 1-c ( 435) 735 170.7 7.7e-39  
gi|137000351|gb|EBR64461.1| hypothetical protein G ( 305) 733 170.1 7.9e-39  
gi|137930097|gb|EBW81399.1| hypothetical protein G ( 306) 732 169.9 9.2e-39  
gi|142971459|gb|EBD88674.1| hypothetical protein G ( 315) 732 169.9 9.4e-39  
gi|135969735|gb|EBL04825.1| hypothetical protein G ( 263) 731 169.7 9.6e-39  
gi|141382733|gb|ECR44145.1| hypothetical protein G ( 233) 730 169.4 1e-38  
gi|262223840|gb|EEY74299.1| 3-phosphoshikimate 1-c ( 665) 735 170.8 1.1e-38  
gi|141192666|gb|ECQ26680.1| hypothetical protein G ( 258) 729 169.2 1.3e-38  
gi|143031842|gb|EDC31502.1| hypothetical protein G ( 316) 730 169.5 1.3e-38  
gi|141378761|gb|ECR41291.1| hypothetical protein G ( 221) 728 169.0 1.3e-38  
gi|137889644|gb|EBW58374.1| hypothetical protein G ( 275) 729 169.2 1.3e-38  
gi|140991096|gb|ECO89414.1| hypothetical protein G ( 280) 729 169.2 1.4e-38  
gi|144082397|gb|EDI74102.1| hypothetical protein G ( 515) 732 170.1 1.4e-38  
gi|229376059|gb|EEQ026150.1| 3-phosphoshikimate 1-c ( 434) 731 169.8 1.4e-38  
gi|139558026|gb|ECF89335.1| hypothetical protein G ( 296) 729 169.3 1.4e-38  
gi|142096572|gb|ECV56061.1| hypothetical protein G ( 434) 730 169.6 1.6e-38  
gi|134767569|gb|EBD29449.1| hypothetical protein G ( 449) 730 169.6 1.7e-38  
gi|143559332|gb|EDF63130.1| hypothetical protein G ( 439) 729 169.4 1.9e-38  
gi|134325002|gb|EBA65155.1| hypothetical protein G ( 299) 727 168.8 1.9e-38  
gi|140765239|gb|ECN35318.1| hypothetical protein G ( 307) 727 168.8 2e-38  
gi|134912557|gb|EBE25355.1| hypothetical protein G ( 281) 726 168.6 2.1e-38  
gi|138267882|gb|EBY68980.1| hypothetical protein G ( 284) 726 168.6 2.2e-38  
gi|140299537|gb|ECS81908.1| hypothetical protein G ( 286) 726 168.6 2.2e-38  
gi|239523935|gb|EEQ63801.1| 3-phosphoshikimate 1-c ( 431) 728 169.2 2.2e-38  
gi|136822130|gb|EBQ63589.1| hypothetical protein G ( 444) 728 169.2 2.3e-38  
gi|112803452|gb|EAU00796.1| 3-phosphoshikimate 1-c ( 424) 727 168.9 2.5e-38  
gi|142993753|gb|EDC04217.1| hypothetical protein G ( 275) 724 168.2 2.8e-38  
gi|136028828|gb|EBL43065.1| hypothetical protein G ( 409) 726 168.7 2.9e-38  
gi|144224364|gb|EDU77028.1| hypothetical protein G ( 348) 725 168.4 2.9e-38  
gi|137891350|gb|EBW59332.1| hypothetical protein G ( 287) 723 167.9 3.4e-38  
gi|142525284|gb|ECY70920.1| hypothetical protein G ( 304) 723 168.0 3.6e-38  
gi|134510091|gb|EBB77103.1| hypothetical protein G ( 264) 722 167.7 3.7e-38  
gi|137383188|gb|EBT78926.1| hypothetical protein G ( 286) 722 167.7 4e-38  
gi|141523746|gb|ECS14701.1| hypothetical protein G ( 291) 722 167.7 4e-38  
gi|139645346|gb|ECC48218.1| hypothetical protein G ( 275) 721 167.5 4.5e-38  
gi|142113139|gb|ECV68675.1| hypothetical protein G ( 416) 723 168.1 4.6e-38

gi|135522926|gb|EBI19182.1| hypothetical protein G ( 293) 721 167.5 4.7e-38  
gi|135150889|gb|EBF81845.1| hypothetical protein G ( 293) 721 167.5 4.7e-38  
gi|140728099|gb|ECN10681.1| hypothetical protein G ( 300) 721 167.5 4.8e-38  
gi|134795852|gb|EBD49110.1| hypothetical protein G ( 412) 722 167.9 5.3e-38  
gi|109714282|emb|CAJ99290.1| 3-phosphoshikimate 1- ( 429) 722 167.9 5.4e-38  
gi|139742010|gb|ECH15842.1| hypothetical protein G ( 291) 720 167.3 5.4e-38  
gi|138415605|gb|EBZ51637.1| hypothetical protein G ( 246) 718 166.8 6.5e-38  
gi|142655804|gb|ECZ62978.1| hypothetical protein G ( 388) 720 167.4 6.8e-38  
gi|140237179|gb|ECG64758.1| hypothetical protein G ( 265) 718 166.8 6.8e-38  
gi|143683511|gb|EDG25599.1| hypothetical protein G ( 449) 720 167.4 7.6e-38  
gi|141829660|gb|ECT53963.1| hypothetical protein G ( 311) 718 166.9 7.7e-38  
gi|144036682|gb|EDI40729.1| hypothetical protein G ( 312) 718 166.9 7.8e-38  
gi|261839786|gb|ACX99551.1| 3-phosphoshikimate 1-c ( 429) 719 167.2 8.5e-38  
gi|143487900|gb|EDF25456.1| hypothetical protein G ( 443) 718 167.0 1e-37  
gi|139047127|gb|ECD04427.1| hypothetical protein G ( 259) 715 166.2 1.1e-37  
gi|142554361|gb|ECY91840.1| hypothetical protein G ( 282) 715 166.2 1.1e-37  
gi|136230222|gb|EBM75731.1| hypothetical protein G ( 279) 714 166.0 1.3e-37  
gi|142520919|gb|ECY67749.1| hypothetical protein G ( 435) 716 166.6 1.4e-37  
gi|118414572|gb|ABK82992.1| 3-phosphoshikimate 1-c ( 425) 715 166.3 1.5e-37  
gi|134831338|gb|EBD71836.1| hypothetical protein G ( 283) 712 165.6 1.8e-37  
gi|134743795|gb|EBD15255.1| hypothetical protein G ( 283) 712 165.6 1.8e-37  
gi|137700970|gb|EBV53502.1| hypothetical protein G ( 289) 712 165.6 1.8e-37  
gi|107837189|gb|ABF85058.1| 3-phosphoshikimate 1-c ( 429) 714 166.1 1.8e-37  
gi|261838386|gb|ACX98152.1| 3-phosphoshikimate 1-c ( 429) 714 166.1 1.8e-37  
gi|135540996|gb|EBI30679.1| hypothetical protein G ( 294) 712 165.6 1.8e-37  
gi|136769707|gb|EBQ28730.1| hypothetical protein G ( 259) 711 165.3 1.9e-37  
gi|141103360|gb|ECP64054.1| hypothetical protein G ( 262) 711 165.3 1.9e-37  
gi|135815796|gb|EBK02149.1| hypothetical protein G ( 389) 713 165.9 2e-37  
gi|4155563|gb|AAD06557.1| 3-PHOSPHOSHIKIMATE 1-CAR ( 429) 713 165.9 2.1e-37  
gi|140461717|gb|ECL90174.1| hypothetical protein G ( 208) 709 164.8 2.2e-37  
gi|143103963|gb|EDC84209.1| hypothetical protein G ( 262) 710 165.1 2.3e-37  
gi|139513223|gb|ECF58552.1| hypothetical protein G ( 273) 710 165.1 2.3e-37  
gi|254001519|emb|CAX29537.1| 3-phosphoshikimate 1- ( 429) 712 165.7 2.5e-37  
gi|142716191|gb|EDA06118.1| hypothetical protein G ( 299) 710 165.1 2.5e-37  
gi|208432878|gb|ACI27749.1| 3-phosphoshikimate 1-c ( 429) 711 165.5 2.9e-37  
gi|188144074|gb|ACD48491.1| 3-phosphoshikimate 1-c ( 429) 711 165.5 2.9e-37  
gi|140210178|gb|ECK27766.1| hypothetical protein G ( 216) 707 164.4 3.1e-37  
gi|141113013|gb|ECP71635.1| hypothetical protein G ( 281) 708 164.7 3.2e-37  
gi|142651615|gb|ECZ60006.1| hypothetical protein G ( 423) 710 165.2 3.3e-37  
gi|138605445|gb|ECA75646.1| hypothetical protein G ( 264) 707 164.4 3.6e-37  
gi|139129981|gb|ECD60708.1| hypothetical protein G ( 290) 707 164.5 3.9e-37  
gi|140841479|gb|ECN88442.1| hypothetical protein G ( 292) 707 164.5 3.9e-37  
gi|142239293|gb|ECW63535.1| hypothetical protein G ( 442) 709 165.0 3.9e-37  
gi|141081650|gb|ECP49597.1| hypothetical protein G ( 247) 706 164.2 4e-37  
gi|137901546|gb|EBW65266.1| hypothetical protein G ( 302) 707 164.5 4e-37  
gi|135409824|gb|EBH44760.1| hypothetical protein G ( 450) 709 165.0 4e-37  
gi|27262190|gb|AAN87376.1| 3-phosphoshikimate 1-ca ( 256) 706 164.2 4.1e-37  
gi|135830394|gb|EBK11348.1| hypothetical protein G ( 423) 708 164.8 4.4e-37  
gi|138564576|gb|ECA47108.1| hypothetical protein G ( 312) 706 164.3 4.7e-37  
gi|139434088|gb|ECF08737.1| hypothetical protein G ( 287) 705 164.0 5.2e-37  
gi|142010327|gb|ECU80421.1| hypothetical protein G ( 394) 706 164.4 5.7e-37  
gi|210133180|gb|ACJ08171.1| 3-phosphoshikimate 1-c ( 429) 706 164.4 6.1e-37  
gi|2313507|gb|AAD07470.1| 3-phosphoshikimate 1-car ( 429) 706 164.4 6.1e-37  
gi|136044945|gb|EBL53876.1| hypothetical protein G ( 439) 706 164.4 6.2e-37  
gi|135295582|gb|EBG69224.1| hypothetical protein G ( 221) 702 163.3 6.6e-37  
gi|141929650|gb|ECU24164.1| hypothetical protein G ( 278) 702 163.4 7.9e-37  
gi|142551449|gb|ECY89791.1| hypothetical protein G ( 284) 702 163.4 8.1e-37  
gi|140715196|gb|ECN01656.1| hypothetical protein G ( 277) 701 163.1 9.2e-37  
gi|135312851|gb|EBG79579.1| hypothetical protein G ( 300) 701 163.2 9.8e-37  
gi|139189085|gb|ECE01980.1| hypothetical protein G ( 262) 700 162.9 1e-36  
gi|139889525|gb|ECT15914.1| hypothetical protein G ( 221) 699 162.6 1e-36  
gi|140284595|gb|ECK74389.1| hypothetical protein G ( 287) 700 162.9 1.1e-36

gi|142271427|gb|ECW87239.1| hypothetical protein G ( 291) 700 162.9 1.1e-36  
gi|135379194|gb|EBH24154.1| hypothetical protein G ( 443) 702 163.5 1.1e-36  
gi|136832675|gb|EBQ71286.1| hypothetical protein G ( 413) 701 163.3 1.3e-36  
gi|134894864|gb|EBE13658.1| hypothetical protein G ( 289) 699 162.7 1.3e-36  
gi|268617282|gb|ACZ11647.1| 3-phosphoshikimate 1-c ( 428) 700 163.1 1.5e-36  
gi|143151911|gb|EDD19366.1| hypothetical protein G ( 215) 696 162.0 1.6e-36  
gi|137912650|gb|EBW71582.1| hypothetical protein G ( 271) 696 162.1 1.9e-36  
gi|143273185|gb|EDE05978.1| hypothetical protein G ( 423) 698 162.6 2e-36  
gi|142397546|gb|ECX77275.1| hypothetical protein G ( 437) 698 162.6 2.1e-36  
gi|136942528|gb|EBR31747.1| hypothetical protein G ( 291) 695 161.9 2.4e-36  
gi|34558834|gb|AAQ75178.1| 3-phosphoshikimate 1-ca ( 431) 697 162.4 2.4e-36  
gi|142845608|gb|EDB01702.1| hypothetical protein G ( 443) 697 162.4 2.4e-36  
gi|134524270|gb|EBB85548.1| hypothetical protein G ( 302) 694 161.7 2.8e-36  
gi|135936082|gb|EBK82171.1| hypothetical protein G ( 254) 693 161.4 2.9e-36  
gi|142032131|gb|EDV01025.1| hypothetical protein G ( 432) 695 162.0 3.2e-36  
gi|78166074|gb|ABB23172.1| 3-phosphoshikimate 1-ca ( 433) 695 162.0 3.2e-36  
gi|137866498|gb|EBW45118.1| hypothetical protein G ( 278) 692 161.2 3.6e-36  
gi|143597077|gb|EDF78773.1| hypothetical protein G ( 439) 694 161.8 3.8e-36  
gi|140201840|gb|ECC21774.1| hypothetical protein G ( 211) 690 160.7 3.9e-36  
gi|142744348|gb|EDA26459.1| hypothetical protein G ( 383) 693 161.5 3.9e-36  
gi|140055201|gb|ECZ26753.1| hypothetical protein G ( 317) 692 161.2 4e-36  
gi|140869145|gb|ECO05746.1| hypothetical protein G ( 276) 690 160.8 4.8e-36  
gi|138176812|gb|EBY21559.1| hypothetical protein G ( 292) 690 160.8 5e-36  
gi|141310217|gb|ECQ97279.1| hypothetical protein G ( 272) 689 160.5 5.5e-36  
gi|18251236|gb|AAL65913.1|AF440389\_1 5-enolpyruvyl ( 518) 692 161.4 5.8e-36  
gi|134662035|gb|EBC67782.1| hypothetical protein G ( 298) 689 160.6 5.9e-36  
gi|134922415|gb|EBE39193.1| hypothetical protein G ( 456) 691 161.1 6.1e-36  
gi|136619536|gb|EBP32785.1| hypothetical protein G ( 324) 689 160.6 6.3e-36  
gi|136260802|gb|EBM96358.1| hypothetical protein G ( 404) 690 160.9 6.5e-36  
gi|142629081|gb|ECZ44128.1| hypothetical protein G ( 413) 690 160.9 6.6e-36  
gi|135868048|gb|EBK35027.1| hypothetical protein G ( 234) 687 160.0 6.7e-36  
gi|134888104|gb|EBE09153.1| hypothetical protein G ( 441) 690 160.9 6.9e-36  
gi|135009399|gb|EBE90409.1| hypothetical protein G ( 216) 686 159.8 7.3e-36  
gi|140327911|gb|ECL01330.1| hypothetical protein G ( 236) 686 159.8 7.8e-36  
gi|137921162|gb|EBW76361.1| hypothetical protein G ( 304) 687 160.1 8.2e-36  
gi|141089428|gb|ECP54985.1| hypothetical protein G ( 266) 686 159.9 8.6e-36  
gi|135875132|gb|EBK39974.1| hypothetical protein G ( 444) 688 160.5 9.4e-36  
gi|1347420376|gb|EBD01873.1| hypothetical protein G ( 255) 685 159.6 9.6e-36  
gi|141879530|gb|ECT88967.1| hypothetical protein G ( 278) 685 159.7 1e-35  
gi|137736100|gb|EBV72455.1| hypothetical protein G ( 294) 685 159.7 1.1e-35  
gi|137812730|gb|EBW14044.1| hypothetical protein G ( 259) 684 159.4 1.1e-35  
gi|138564577|gb|ECA47109.1| hypothetical protein G ( 321) 685 159.7 1.2e-35  
gi|136572401|gb|EBP02753.1| hypothetical protein G ( 285) 684 159.5 1.2e-35  
gi|138825029|gb|ECC03218.1| hypothetical protein G ( 287) 684 159.5 1.2e-35  
gi|135036247|gb|EBF08483.1| hypothetical protein G ( 261) 683 159.2 1.3e-35  
gi|143162244|gb|EDD26590.1| hypothetical protein G ( 394) 685 159.8 1.3e-35  
gi|134355864|gb|EBA85489.1| hypothetical protein G ( 401) 685 159.8 1.4e-35  
gi|135952956|gb|EBK93503.1| hypothetical protein G ( 444) 684 159.6 1.7e-35  
gi|135382585|gb|EBH26426.1| hypothetical protein G ( 445) 684 159.6 1.7e-35  
gi|138197142|gb|EBY34330.1| hypothetical protein G ( 194) 679 158.2 1.9e-35  
gi|136006587|gb|EBL28090.1| hypothetical protein G ( 304) 681 158.8 2e-35  
gi|140294305|gb|ECK79442.1| hypothetical protein G ( 210) 679 158.3 2.1e-35  
gi|139762490|gb|ECH28616.1| hypothetical protein G ( 258) 680 158.6 2.1e-35  
gi|136503636|gb|EBO58780.1| hypothetical protein G ( 377) 681 158.9 2.4e-35  
gi|1378777407|gb|EBW51396.1| hypothetical protein G ( 275) 679 158.4 2.5e-35  
gi|140796557|gb|ECN57125.1| hypothetical protein G ( 229) 678 158.1 2.6e-35  
gi|136080841|gb|EBL78109.1| hypothetical protein G ( 285) 679 158.4 2.6e-35  
gi|136694506|gb|EBP79131.1| hypothetical protein G ( 288) 679 158.4 2.6e-35  
gi|141425814|gb|ECC74107.1| hypothetical protein G ( 239) 678 158.1 2.6e-35  
gi|143027028|gb|EDC27946.1| hypothetical protein G ( 443) 681 158.9 2.7e-35  
gi|137295367|gb|EBT29790.1| hypothetical protein G ( 269) 678 158.1 2.9e-35  
gi|137577482|gb|EBU85349.1| hypothetical protein G ( 253) 677 157.9 3.2e-35

gi 135410539 gb EBH45247.1	hypothetical protein G ( 393)	679	158.5	3.3e-35
gi 143179993 gb EDD39182.1	hypothetical protein G ( 421)	679	158.5	3.5e-35
gi 141725526 gb ECS96576.1	hypothetical protein G ( 296)	677	157.9	3.6e-35
gi 137706785 gb EBV56572.1	hypothetical protein G ( 223)	675	157.4	3.9e-35
gi 139539653 gb ECF76318.1	hypothetical protein G ( 228)	675	157.4	4e-35
gi 134953240 gb EBE52444.1	hypothetical protein G ( 305)	676	157.7	4.3e-35
gi 136286814 gb EBN14262.1	hypothetical protein G ( 276)	675	157.5	4.6e-35
gi 136080172 gb EBL77651.1	hypothetical protein G ( 352)	676	157.8	4.8e-35
gi 136044545 gb EBL53611.1	hypothetical protein G ( 439)	677	158.1	4.9e-35
gi 142251381 gb ECW72468.1	hypothetical protein G ( 368)	676	157.8	5e-35
gi 137723714 gb EBV65818.1	hypothetical protein G ( 262)	674	157.3	5.2e-35
gi 142131720 gb ECV82212.1	hypothetical protein G ( 413)	676	157.8	5.4e-35
gi 140986458 gb ECO86073.1	hypothetical protein G ( 289)	674	157.3	5.6e-35
gi 137473914 gb EBU29669.1	hypothetical protein G ( 242)	673	157.0	5.7e-35
gi 139938553 gb ECI50387.1	hypothetical protein G ( 254)	673	157.0	5.9e-35
gi 141443180 gb ECR86430.1	hypothetical protein G ( 255)	673	157.0	5.9e-35
gi 139973384 gb ECI74252.1	hypothetical protein G ( 265)	673	157.0	6.1e-35
gi 136042971 gb EBL52555.1	hypothetical protein G ( 395)	675	157.6	6.1e-35
gi 142012462 gb ECU82503.1	hypothetical protein G ( 395)	675	157.6	6.1e-35
gi 140047126 gb ECJ22456.1	hypothetical protein G ( 272)	673	157.0	6.2e-35
gi 143907191 gb EDH48930.1	hypothetical protein G ( 418)	674	157.4	7.4e-35
gi 138520674 gb ECA16304.1	hypothetical protein G ( 265)	671	156.6	8.2e-35
gi 138932794 gb ECC47942.1	hypothetical protein G ( 272)	671	156.6	8.4e-35
gi 143099253 gb EDC80882.1	hypothetical protein G ( 274)	671	156.6	8.4e-35
gi 139543375 gb ECF78911.1	hypothetical protein G ( 236)	670	156.3	8.7e-35
gi 142970879 gb EDB88270.1	hypothetical protein G ( 299)	670	156.4	1e-34
gi 138970089 gb ECC63836.1	hypothetical protein G ( 251)	669	156.2	1.1e-34
gi 135658881 gb EBJ04539.1	hypothetical protein G ( 262)	669	156.2	1.1e-34
gi 143516129 gb EDF41123.1	hypothetical protein G ( 394)	671	156.7	1.1e-34
gi 141389048 gb ECR48542.1	hypothetical protein G ( 192)	667	155.6	1.2e-34
gi 137356768 gb EBT64149.1	hypothetical protein G ( 305)	669	156.2	1.2e-34
gi 140971289 gb ECO75475.1	hypothetical protein G ( 267)	668	156.0	1.3e-34
gi 142540256 gb ECY81791.1	hypothetical protein G ( 443)	670	156.6	1.4e-34
gi 134586542 gb EBC22942.1	hypothetical protein G ( 254)	667	155.7	1.5e-34
gi 139351418 gb ECE59394.1	hypothetical protein G ( 260)	667	155.7	1.5e-34
gi 143636057 gb EDF98691.1	hypothetical protein G ( 268)	667	155.7	1.5e-34
gi 137904507 gb EBW66910.1	hypothetical protein G ( 270)	667	155.7	1.5e-34
gi 134587137 gb EBC23302.1	hypothetical protein G ( 238)	666	155.5	1.6e-34
gi 137316119 gb EBT41367.1	hypothetical protein G ( 265)	666	155.5	1.7e-34
gi 136798838 gb EBQ48122.1	hypothetical protein G ( 227)	665	155.2	1.8e-34
gi 143869852 gb EDH22124.1	hypothetical protein G ( 441)	668	156.1	1.9e-34
gi 142945309 gb EDB70428.1	hypothetical protein G ( 330)	665	155.4	2.4e-34
gi 135007077 gb EBE88844.1	hypothetical protein G ( 407)	666	155.7	2.4e-34
gi 152939223 gb ABS43964.1	3-phosphoshikimate 1-c ( 423)	666	155.7	2.5e-34
gi 136812901 gb EBQ57436.1	hypothetical protein G ( 442)	666	155.7	2.6e-34
gi 143435399 gb EDE94044.1	hypothetical protein G ( 443)	666	155.7	2.6e-34
gi 139807489 gb ECH60285.1	hypothetical protein G ( 208)	662	154.6	2.6e-34
gi 136262713 gb EBM97691.1	hypothetical protein G ( 388)	665	155.4	2.7e-34
gi 139835846 gb ECH79917.1	hypothetical protein G ( 269)	663	154.9	2.8e-34
gi 153804313 gb ABS51320.1	3-phosphoshikimate 1-c ( 424)	665	155.5	2.9e-34
gi 57166528 gb AAW35307.1	3-phosphoshikimate 1-ca ( 428)	665	155.5	2.9e-34
gi 139775304 gb ECH37500.1	hypothetical protein G ( 292)	663	154.9	3e-34
gi 135835370 gb EBK14477.1	hypothetical protein G ( 278)	662	154.7	3.3e-34
gi 136331028 gb EBN44134.1	hypothetical protein G ( 432)	664	155.2	3.4e-34
gi 134346473 gb EBA79510.1	hypothetical protein G ( 441)	664	155.2	3.5e-34
gi 140352336 gb ECL18584.1	hypothetical protein G ( 266)	661	154.4	3.7e-34
gi 141168333 gb ECQ10384.1	hypothetical protein G ( 258)	660	154.2	4.2e-34
gi 135306712 gb EBG75775.1	hypothetical protein G ( 382)	662	154.8	4.2e-34
gi 135879561 gb EBK43089.1	hypothetical protein G ( 260)	660	154.2	4.3e-34
gi 142505282 gb ECY56360.1	hypothetical protein G ( 386)	662	154.8	4.3e-34
gi 143183876 gb EDD42018.1	hypothetical protein G ( 396)	662	154.8	4.4e-34
gi 139374135 gb ECE67784.1	hypothetical protein G ( 277)	660	154.2	4.5e-34

gi 140092149 gb ECJ47988.1	hypothetical protein G ( 279)	660	154.2	4.5e-34
gi 134797263 gb EBB50100.1	hypothetical protein G ( 398)	661	154.6	5.1e-34
gi 137852058 gb EBW36821.1	hypothetical protein G ( 276)	659	154.0	5.2e-34
gi 135390469 gb EBH31737.1	hypothetical protein G ( 441)	661	154.6	5.5e-34
gi 137457344 gb EBU20973.1	hypothetical protein G ( 239)	657	153.5	6.3e-34
gi 136053553 gb EBI59732.1	hypothetical protein G ( 362)	659	154.1	6.4e-34
gi 144040383 gb EDI43271.1	hypothetical protein G ( 442)	660	154.4	6.4e-34
gi 112360221 emb CAL35016.1	3-phosphoshikimate 1- ( 428)	659	154.1	7.3e-34
gi 136808575 gb EBQ54561.1	hypothetical protein G ( 442)	659	154.2	7.5e-34
gi 143988677 gb EDI06566.1	hypothetical protein G ( 218)	655	153.1	7.9e-34
gi 135009040 gb EBE90157.1	hypothetical protein G ( 395)	658	153.9	7.9e-34
gi 142813782 gb EDA77762.1	hypothetical protein G ( 433)	658	153.9	8.5e-34
gi 143789403 gb EDG79192.1	hypothetical protein G ( 407)	657	153.7	9.5e-34
gi 121504279 gb EAQ72379.2	3-phosphoshikimate 1-c ( 423)	657	153.7	9.7e-34
gi 143680111 gb EDG23778.1	hypothetical protein G ( 443)	657	153.7	1e-33
gi 141975165 gb ECU55743.1	hypothetical protein G ( 254)	654	152.9	1e-33
gi 135461960 gb EBH79689.1	hypothetical protein G ( 258)	654	152.9	1e-33
gi 143598405 gb EBQ79423.1	hypothetical protein G ( 394)	656	153.5	1.1e-33
gi 142675360 gb ECI276848.1	hypothetical protein G ( 395)	656	153.5	1.1e-33
gi 139088179 gb ECD32916.1	hypothetical protein G ( 239)	653	152.7	1.1e-33
gi 142684538 gb ECB283443.1	hypothetical protein G ( 293)	654	152.9	1.2e-33
gi 135845413 gb EBK20780.1	hypothetical protein G ( 213)	652	152.4	1.2e-33
gi 138398823 gb EBZ39594.1	hypothetical protein G ( 274)	653	152.7	1.3e-33
gi 144114347 gb EDG23778.1	hypothetical protein G ( 511)	656	153.6	1.3e-33
gi 141732305 gb ECT00772.1	hypothetical protein G ( 300)	653	152.7	1.4e-33
gi 140820936 gb ECN74293.1	hypothetical protein G ( 215)	651	152.2	1.4e-33
gi 141934358 gb ECU27467.1	hypothetical protein G ( 262)	652	152.5	1.4e-33
gi 139557236 gb ECF88763.1	hypothetical protein G ( 288)	652	152.5	1.5e-33
gi 143179362 gb EDD38720.1	hypothetical protein G ( 441)	654	153.1	1.6e-33
gi 139498116 gb ECF48108.1	hypothetical protein G ( 310)	652	152.5	1.6e-33
gi 140268029 gb ECK67239.1	hypothetical protein G ( 214)	650	152.0	1.7e-33
gi 139005441 gb ECC75437.1	hypothetical protein G ( 261)	651	152.2	1.7e-33
gi 139734911 gb ECH10883.1	hypothetical protein G ( 218)	650	152.0	1.7e-33
gi 143388359 gb EDE69270.1	hypothetical protein G ( 272)	651	152.3	1.7e-33
gi 138689266 gb ECB32656.1	hypothetical protein G ( 277)	651	152.3	1.7e-33
gi 140305932 gb ECK86014.1	hypothetical protein G ( 267)	650	152.0	2e-33
gi 141883047 gb ECT91467.1	hypothetical protein G ( 257)	649	151.8	2.2e-33
gi 140994978 gb ECO92196.1	hypothetical protein G ( 213)	648	151.5	2.2e-33
gi 143460504 gb EDF09805.1	hypothetical protein G ( 384)	651	152.4	2.2e-33
gi 134796922 gb EBD49853.1	hypothetical protein G ( 299)	649	151.9	2.5e-33
gi 138675605 gb ECB23156.1	hypothetical protein G ( 300)	649	151.9	2.5e-33
gi 139635959 gb ECG41765.1	hypothetical protein G ( 249)	648	151.6	2.5e-33
gi 134579288 gb EBC18564.1	hypothetical protein G ( 222)	647	151.3	2.7e-33
gi 141310782 gb ECQ97441.1	hypothetical protein G ( 276)	648	151.6	2.7e-33
gi 140306938 gb ECK86704.1	hypothetical protein G ( 289)	648	151.6	2.8e-33
gi 139645545 gb ECG48352.1	hypothetical protein G ( 259)	647	151.4	3e-33
gi 138280102 gb EBY77418.1	hypothetical protein G ( 268)	647	151.4	3.1e-33
gi 138196756 gb EBY34068.1	hypothetical protein G ( 263)	646	151.2	3.5e-33
gi 157386116 gb ABV52431.1	3-phosphoshikimate 1-c ( 423)	648	151.8	3.8e-33
gi 895778 emb CAA61554.1	5-enolpyruvylshikimate ( 428)	648	151.8	3.8e-33
gi 134704854 gb EBC93118.1	hypothetical protein G ( 242)	645	150.9	3.9e-33
gi 142537980 gb ECY80156.1	hypothetical protein G ( 442)	648	151.8	3.9e-33
gi 143980116 gb EDT00633.1	hypothetical protein G ( 379)	647	151.5	4e-33
gi 143510567 gb EDF38097.1	hypothetical protein G ( 307)	645	151.0	4.7e-33
gi 140945922 gb ECO57780.1	hypothetical protein G ( 255)	644	150.7	4.7e-33
gi 142143916 gb ECY91266.1	hypothetical protein G ( 182)	642	150.2	4.9e-33
gi 136620139 gb EBP33148.1	hypothetical protein G ( 271)	644	150.7	4.9e-33
gi 134887599 gb EBE08816.1	hypothetical protein G ( 440)	645	151.1	6.1e-33
gi 137180732 gb EBK56453.1	hypothetical protein G ( 215)	641	150.0	6.5e-33
gi 46095337 gb AAS80163.1	5-enolpyruvylshikimate- ( 514)	645	151.2	6.9e-33
gi 139689782 gb ECG79632.1	hypothetical protein G ( 240)	641	150.0	7e-33
gi 138251002 gb EBY57235.1	hypothetical protein G ( 260)	641	150.1	7.5e-33

gi|140473337|gb|ECL94305.1| hypothetical protein G ( 262) 641 150.1 7.5e-33  
gi|137295054|gb|EBT29629.1| hypothetical protein G ( 281) 641 150.1 7.9e-33  
gi|137002628|gb|EBR65772.1| hypothetical protein G ( 286) 641 150.1 8.1e-33  
gi|137369411|gb|EBT71302.1| hypothetical protein G ( 243) 640 149.8 8.3e-33  
gi|135292728|gb|EBG67583.1| hypothetical protein G ( 447) 643 150.7 8.4e-33  
gi|141387612|gb|ECR47580.1| hypothetical protein G ( 221) 639 149.6 8.9e-33  
gi|142280802|gb|ECW94104.1| hypothetical protein G ( 280) 640 149.9 9.2e-33  
gi|134781214|gb|EBD38889.1| hypothetical protein G ( 352) 641 150.2 9.5e-33  
gi|135690666|gb|EBJ24138.1| hypothetical protein G ( 283) 639 149.7 1.1e-32  
gi|138043293|gb|EBX43263.1| hypothetical protein G ( 306) 639 149.7 1.1e-32  
gi|134976758|gb|EBE68251.1| hypothetical protein G ( 386) 640 150.0 1.2e-32  
gi|134787001|gb|EBD42903.1| hypothetical protein G ( 265) 638 149.4 1.2e-32  
gi|137922726|gb|EBW77238.1| hypothetical protein G ( 281) 638 149.4 1.2e-32  
gi|135275881|gb|EBG57715.1| hypothetical protein G ( 204) 636 148.9 1.3e-32  
gi|137358507|gb|EBT65149.1| hypothetical protein G ( 271) 637 149.2 1.4e-32  
gi|137110893|gb|EBS26319.1| hypothetical protein G ( 278) 637 149.2 1.4e-32  
gi|140822463|gb|ECN75406.1| hypothetical protein G ( 229) 636 148.9 1.4e-32  
gi|207083516|gb|ED260942.1| 3-phosphoshikimate 1-c ( 439) 639 149.8 1.5e-32  
gi|136359784|gb|EBN63723.1| hypothetical protein G ( 255) 636 149.0 1.6e-32  
gi|135487726|gb|EBH96571.1| hypothetical protein G ( 381) 638 149.5 1.6e-32  
gi|139743346|gb|ECH16606.1| hypothetical protein G ( 215) 635 148.7 1.6e-32  
gi|135846862|gb|EBK21692.1| hypothetical protein G ( 387) 638 149.5 1.6e-32  
gi|141859518|gb|ECT75045.1| hypothetical protein G ( 265) 636 149.0 1.6e-32  
gi|140501840|gb|ECM03625.1| hypothetical protein G ( 284) 636 149.0 1.7e-32  
gi|135721677|gb|EBJ43342.1| hypothetical protein G ( 361) 637 149.3 1.8e-32  
gi|135666846|gb|EBJ09495.1| hypothetical protein G ( 249) 635 148.7 1.8e-32  
gi|136407897|gb|EBN96812.1| hypothetical protein G ( 213) 634 148.5 1.8e-32  
gi|143495182|gb|EDF30165.1| hypothetical protein G ( 390) 637 149.3 1.9e-32  
gi|141235790|gb|ECQ56283.1| hypothetical protein G ( 244) 634 148.5 2e-32  
gi|143253346|gb|EDD91725.1| hypothetical protein G ( 249) 634 148.5 2.1e-32  
gi|142795846|gb|EDA64359.1| hypothetical protein G ( 403) 636 149.1 2.2e-32  
gi|136098922|gb|EBL90401.1| hypothetical protein G ( 413) 636 149.1 2.3e-32  
gi|142223275|gb|ECW51657.1| hypothetical protein G ( 439) 636 149.2 2.4e-32  
gi|141580814|gb|ECS46000.1| hypothetical protein G ( 250) 633 148.3 2.4e-32  
gi|141370926|gb|ECR35647.1| hypothetical protein G ( 258) 633 148.3 2.5e-32  
gi|141226948|gb|ECQ49886.1| hypothetical protein G ( 270) 633 148.3 2.6e-32  
gi|137316621|gb|EBT41642.1| hypothetical protein G ( 296) 633 148.4 2.8e-32  
gi|137070063|gb|EBS03773.1| hypothetical protein G ( 298) 633 148.4 2.8e-32  
gi|137335226|gb|EBT51972.1| hypothetical protein G ( 276) 632 148.1 3e-32  
gi|135856354|gb|EBK27680.1| hypothetical protein G ( 408) 634 148.7 3e-32  
gi|143245592|gb|EDD86275.1| hypothetical protein G ( 435) 634 148.7 3.2e-32  
gi|141856808|gb|ECT73158.1| hypothetical protein G ( 265) 631 147.9 3.4e-32  
gi|138656624|gb|ECB09665.1| hypothetical protein G ( 305) 631 147.9 3.8e-32  
gi|141237285|gb|ECQ57366.1| hypothetical protein G ( 200) 628 147.2 4.3e-32  
gi|143886715|gb|EDH34407.1| hypothetical protein G ( 441) 632 148.3 4.4e-32  
gi|139108648|gb|ECD47280.1| hypothetical protein G ( 272) 629 147.5 4.7e-32  
gi|143115386|gb|EDC92602.1| hypothetical protein G ( 409) 631 148.0 4.8e-32  
gi|141882967|gb|ECT91409.1| hypothetical protein G ( 236) 628 147.2 4.9e-32  
gi|137078632|gb|EBS08614.1| hypothetical protein G ( 199) 627 146.9 5e-32  
gi|135216672|gb|EBG22887.1| hypothetical protein G ( 166) 626 146.7 5.1e-32  
gi|140055200|gb|ECJ26752.1| hypothetical protein G ( 180) 626 146.7 5.4e-32  
gi|137898190|gb|EBW63309.1| hypothetical protein G ( 268) 628 147.2 5.4e-32  
gi|142296833|gb|ECX05822.1| hypothetical protein G ( 403) 630 147.8 5.5e-32  
gi|141718414|gb|ECS92174.1| hypothetical protein G ( 356) 628 147.3 6.8e-32  
gi|136926330|gb|EBR22646.1| hypothetical protein G ( 328) 627 147.1 7.4e-32  
gi|143186318|gb|EDD43806.1| hypothetical protein G ( 333) 627 147.1 7.5e-32  
gi|142343455|gb|ECX40449.1| hypothetical protein G ( 351) 627 147.1 7.8e-32  
gi|139942801|gb|ECI53440.1| hypothetical protein G ( 263) 625 146.6 8.4e-32  
gi|143465264|gb|EDF12656.1| hypothetical protein G ( 377) 626 146.9 9.6e-32  
gi|141034040|gb|ECP17849.1| hypothetical protein G ( 272) 624 146.4 1e-31  
gi|134508435|gb|EBB76121.1| hypothetical protein G ( 272) 624 146.4 1e-31  
gi|139929141|gb|ECI43695.1| hypothetical protein G ( 272) 624 146.4 1e-31

gi|141181800|gb|ECQ19627.1| hypothetical protein G ( 248) 623 146.1 1.1e-31  
gi|139450633|gb|ECL18656.1| hypothetical protein G ( 261) 623 146.2 1.1e-31  
gi|142618019|gb|ECZ36310.1| hypothetical protein G ( 276) 623 146.2 1.2e-31  
gi|135008910|gb|EBE90069.1| hypothetical protein G ( 431) 625 146.7 1.2e-31  
gi|135188909|gb|EBG062440.1| hypothetical protein G ( 244) 622 145.9 1.3e-31  
gi|135301778|gb|EBG72877.1| hypothetical protein G ( 346) 623 146.2 1.4e-31  
gi|138733010|gb|ECB63289.1| hypothetical protein G ( 267) 621 145.7 1.6e-31  
gi|143710497|gb|EDG639549.1| hypothetical protein G ( 439) 623 146.3 1.7e-31  
gi|136294048|gb|EBN19181.1| hypothetical protein G ( 281) 620 145.5 1.9e-31  
gi|139100964|gb|ECD41990.1| hypothetical protein G ( 263) 619 145.3 2.1e-31  
gi|140415147|gb|ECL60719.1| hypothetical protein G ( 275) 619 145.3 2.2e-31  
gi|143916848|gb|EDH55629.1| hypothetical protein G ( 339) 620 145.6 2.2e-31  
gi|29727034|gb|AAD47362.2| AF038578\_5 cyclohexadien ( 505) 622 146.1 2.2e-31  
gi|139248676|gb|ECE35811.1| hypothetical protein G ( 200) 617 144.8 2.3e-31  
gi|17815|emb|CAA35839.1| 5-enolpyruvylshikimate-3- ( 516) 621 145.9 2.6e-31  
gi|62786940|gb|AA06819.1| Sequence 26 from patent ( 516) 621 145.9 2.6e-31  
gi|136976536|gb|EBR50991.1| hypothetical protein G ( 309) 618 145.1 2.7e-31  
gi|141216849|gb|ECQ42758.1| hypothetical protein G ( 263) 617 144.8 2.8e-31  
gi|136480496|gb|EBD43798.1| hypothetical protein G ( 404) 619 145.4 2.9e-31  
gi|136637387|gb|EBP43347.1| hypothetical protein G ( 230) 616 144.6 3e-31  
gi|142392190|gb|ECX73356.1| hypothetical protein G ( 299) 617 144.9 3.1e-31  
gi|141067369|gb|ECP39770.1| hypothetical protein G ( 206) 615 144.3 3.2e-31  
gi|143914669|gb|EDH54067.1| hypothetical protein G ( 392) 618 145.2 3.3e-31  
gi|135475610|gb|EBH88835.1| hypothetical protein G ( 279) 616 144.6 3.4e-31  
gi|134425336|gb|EBB27523.1| hypothetical protein G ( 294) 616 144.7 3.6e-31  
gi|143744226|gb|EDG56728.1| hypothetical protein G ( 398) 617 145.0 3.9e-31  
gi|134326213|gb|EBA645879.1| hypothetical protein G ( 271) 615 144.4 3.9e-31  
gi|139402158|gb|ECE86608.1| hypothetical protein G ( 239) 614 144.2 4.1e-31  
gi|135575756|gb|EBI52970.1| hypothetical protein G ( 162) 612 143.6 4.1e-31  
gi|139179376|gb|ECX730573.1| hypothetical protein G ( 247) 614 144.2 4.2e-31  
gi|143681911|gb|EDG24746.1| hypothetical protein G ( 376) 616 144.7 4.3e-31  
gi|138396080|gb|EBZ37708.1| hypothetical protein G ( 279) 614 144.2 4.6e-31  
gi|137007600|gb|EBR86805.1| hypothetical protein G ( 199) 612 143.7 4.8e-31  
gi|143823139|gb|EDG88251.1| hypothetical protein G ( 206) 612 143.7 5e-31  
gi|143592271|gb|EDF76017.1| hypothetical protein G ( 381) 615 144.5 5.1e-31  
gi|135328530|gb|EBG90150.1| hypothetical protein G ( 341) 614 144.3 5.4e-31  
gi|143227632|gb|EDD73899.1| hypothetical protein G ( 346) 614 144.3 5.5e-31  
gi|136662556|gb|EBP58336.1| hypothetical protein G ( 190) 610 143.2 6.3e-31  
gi|143764419|gb|EDG66830.1| hypothetical protein G ( 335) 612 143.8 7.2e-31  
gi|139012268|gb|ECC79849.1| hypothetical protein G ( 277) 611 143.6 7.3e-31  
gi|222539413|gb|ACM64514.1| 3-phosphoshikimate 1-c ( 428) 613 144.1 7.5e-31  
gi|143718994|gb|EDG43389.1| hypothetical protein G ( 379) 612 143.9 8e-31  
gi|140907696|gb|ECO31017.1| hypothetical protein G ( 282) 610 143.3 8.6e-31  
gi|134855060|gb|EBD87065.1| hypothetical protein G ( 394) 611 143.7 9.5e-31  
gi|137284545|gb|EBT23794.1| hypothetical protein G ( 271) 609 143.1 9.7e-31  
gi|139332513|gb|ECE53914.1| hypothetical protein G ( 273) 609 143.1 9.7e-31  
gi|142204276|gb|ECW37281.1| hypothetical protein G ( 249) 608 142.9 1.1e-30  
gi|142886453|gb|EDB28141.1| hypothetical protein G ( 379) 610 143.4 1.1e-30  
gi|139891102|gb|ECI17002.1| hypothetical protein G ( 218) 607 142.6 1.1e-30  
gi|143113848|gb|EDC91473.1| hypothetical protein G ( 334) 609 143.2 1.1e-30  
gi|135644105|gb|EBI95407.1| hypothetical protein G ( 276) 608 142.9 1.1e-30  
gi|144041930|gb|EDT44427.1| hypothetical protein G ( 441) 610 143.5 1.2e-30  
gi|140209892|gb|ECK27558.1| hypothetical protein G ( 210) 606 142.4 1.2e-30  
gi|142731161|gb|EDA16839.1| hypothetical protein G ( 379) 609 143.2 1.3e-30  
gi|138372334|gb|EBZ21784.1| hypothetical protein G ( 261) 607 142.7 1.3e-30  
gi|135303110|gb|EBG73663.1| hypothetical protein G ( 300) 607 142.7 1.4e-30  
gi|141349373|gb|ECR20756.1| hypothetical protein G ( 247) 606 142.4 1.4e-30  
gi|138288446|gb|EBY80873.1| hypothetical protein G ( 259) 606 142.4 1.5e-30  
gi|144126252|gb|EDU05420.1| hypothetical protein G ( 233) 605 142.2 1.6e-30  
gi|134722639|gb|EBD03170.1| hypothetical protein G ( 276) 605 142.3 1.8e-30  
gi|139146426|gb|ECQ72028.1| hypothetical protein G ( 284) 605 142.3 1.8e-30  
gi|143509311|gb|EDF37468.1| hypothetical protein G ( 363) 605 142.3 2.2e-30



gi|143837477|gb|EDG98597.1| hypothetical protein G ( 252) 603 141.8 2.3e-30  
gi|143644718|gb|EDG03752.1| hypothetical protein G ( 253) 603 141.8 2.3e-30  
gi|71062313|gb|AAZ21316.1| 3-phosphoshikimate 1-ca ( 395) 604 142.2 2.7e-30  
gi|143249054|gb|EDD88694.1| hypothetical protein G ( 344) 603 141.9 2.9e-30  
gi|142445939|gb|ECY13527.1| hypothetical protein G ( 199) 600 141.1 3e-30  
gi|134563621|gb|EBC09100.1| hypothetical protein G ( 348) 602 141.7 3.4e-30  
gi|136317320|gb|EBN34967.1| hypothetical protein G ( 349) 602 141.7 3.4e-30  
gi|143387950|gb|EDE69059.1| hypothetical protein G ( 364) 602 141.7 3.5e-30  
gi|142973874|gb|EDB90380.1| hypothetical protein G ( 374) 602 141.7 3.6e-30  
gi|144034658|gb|EDI39332.1| hypothetical protein G ( 222) 599 140.9 3.7e-30  
gi|134915661|gb|EBE27413.1| hypothetical protein G ( 279) 600 141.2 3.8e-30  
gi|142353791|gb|ECX47379.1| hypothetical protein G ( 253) 599 140.9 4.1e-30  
gi|134407016|gb|EBB16951.1| hypothetical protein G ( 276) 599 140.9 4.4e-30  
gi|141387610|gb|ECR47578.1| hypothetical protein G ( 286) 599 141.0 4.6e-30  
gi|141519084|gb|ECS13650.1| hypothetical protein G ( 195) 597 140.4 4.6e-30  
gi|135723033|gb|EBJ44194.1| hypothetical protein G ( 367) 600 141.3 4.7e-30  
gi|135016962|gb|EBE95506.1| hypothetical protein G ( 377) 600 141.3 4.8e-30  
gi|136764304|gb|EBQ25038.1| hypothetical protein G ( 405) 600 141.3 5.1e-30  
gi|140834228|gb|ECN83709.1| hypothetical protein G ( 234) 597 140.5 5.3e-30  
gi|135033042|gb|EBF06409.1| hypothetical protein G ( 382) 599 141.1 5.7e-30  
gi|135169736|gb|EBF93991.1| hypothetical protein G ( 337) 598 140.8 6e-30  
gi|139809562|gb|ECH61775.1| hypothetical protein G ( 236) 596 140.2 6.2e-30  
gi|138544954|gb|ECA33332.1| hypothetical protein G ( 281) 596 140.3 7.1e-30  
gi|135462160|gb|EBH79820.1| hypothetical protein G ( 241) 595 140.0 7.3e-30  
gi|140430189|gb|ECL69708.1| hypothetical protein G ( 249) 595 140.0 7.5e-30  
gi|135735463|gb|EBJ51874.1| hypothetical protein G ( 270) 595 140.1 8e-30  
gi|140007059|gb|ECI196257.1| hypothetical protein G ( 226) 593 139.6 9.4e-30  
gi|141198878|gb|ECQ30289.1| hypothetical protein G ( 227) 593 139.6 9.4e-30  
gi|140711280|gb|ECM98898.1| hypothetical protein G ( 228) 593 139.6 9.4e-30  
gi|138799428|gb|ECB91382.1| hypothetical protein G ( 302) 594 139.9 1e-29  
gi|142782139|gb|EDA54172.1| hypothetical protein G ( 227) 592 139.4 1.1e-29  
gi|138307154|gb|EBY90044.1| hypothetical protein G ( 233) 592 139.4 1.1e-29  
gi|140995634|gb|ECOC92668.1| hypothetical protein G ( 171) 590 138.8 1.2e-29  
gi|134426111|gb|EBB27979.1| hypothetical protein G ( 309) 593 139.7 1.2e-29  
gi|134961005|gb|EBE57619.1| hypothetical protein G ( 270) 592 139.4 1.3e-29  
gi|141691249|gb|ECS83238.1| hypothetical protein G ( 167) 589 138.6 1.4e-29  
gi|136211953|gb|EBM63336.1| hypothetical protein G ( 310) 592 139.5 1.4e-29  
gi|140624511|gb|ECM41483.1| hypothetical protein G ( 232) 590 138.9 1.5e-29  
gi|142202639|gb|ECW36029.1| hypothetical protein G ( 236) 590 138.9 1.5e-29  
gi|142721012|gb|EDA09566.1| hypothetical protein G ( 424) 593 139.8 1.5e-29  
gi|139811438|gb|ECH63106.1| hypothetical protein G ( 241) 590 138.9 1.6e-29  
gi|137824042|gb|EBW20669.1| hypothetical protein G ( 203) 589 138.7 1.6e-29  
gi|135649501|gb|EBI98752.1| hypothetical protein G ( 371) 592 139.5 1.6e-29  
gi|47250293|gb|AAT20236.1| Sequence 1608 from pate ( 309) 591 139.2 1.6e-29  
gi|141720170|gb|ECS93268.1| hypothetical protein G ( 346) 591 139.3 1.8e-29  
gi|135901397|gb|EBK58397.1| hypothetical protein G ( 223) 588 138.5 2e-29  
gi|137342847|gb|EBT56305.1| hypothetical protein G ( 272) 589 138.8 2e-29  
gi|140677547|gb|ECM75559.1| hypothetical protein G ( 237) 588 138.5 2.1e-29  
gi|135393586|gb|EBH33843.1| hypothetical protein G ( 379) 590 139.1 2.2e-29  
gi|143372153|gb|EDE62339.1| hypothetical protein G ( 438) 590 139.1 2.5e-29  
gi|137160598|gb|EBS54177.1| hypothetical protein G ( 225) 586 138.0 2.7e-29  
gi|135428951|gb|EBH57631.1| hypothetical protein G ( 227) 586 138.1 2.7e-29  
gi|141272822|gb|ECQ82078.1| hypothetical protein G ( 278) 587 138.3 2.7e-29  
gi|135744269|gb|EBJ57327.1| hypothetical protein G ( 285) 587 138.3 2.8e-29  
gi|139450953|gb|ECF18884.1| hypothetical protein G ( 252) 586 138.1 2.9e-29  
gi|142804270|gb|EDA70643.1| hypothetical protein G ( 256) 586 138.1 3e-29  
gi|141418805|gb|ECR69222.1| hypothetical protein G ( 257) 586 138.1 3e-29  
gi|135544424|gb|EBI32879.1| hypothetical protein G ( 396) 588 138.7 3.1e-29  
gi|137512952|gb|EBU49685.1| hypothetical protein G ( 231) 585 137.8 3.2e-29  
gi|139383908|gb|ECE74599.1| hypothetical protein G ( 201) 584 137.6 3.3e-29  
gi|139799834|gb|ECH54843.1| hypothetical protein G ( 254) 585 137.9 3.4e-29  
gi|137230117|gb|EBS93243.1| hypothetical protein G ( 236) 584 137.6 3.8e-29

gi|138264744|gb|EBY66746.1| hypothetical protein G ( 221) 583 137.4 4.2e-29  
gi|135468157|gb|EBH83851.1| hypothetical protein G ( 343) 585 138.0 4.3e-29  
gi|137606050|gb|EBV01280.1| hypothetical protein G ( 235) 583 137.4 4.4e-29  
gi|161726859|emb|CAP47299.1| unnamed protein produ ( 430) 586 138.3 4.4e-29  
gi|20905696|gb|AAM30934.1| 3-phosphoshikimate 1-ca ( 430) 586 138.3 4.4e-29  
gi|197053761|gb|ACH25459.1| Sequence 11 from paten ( 430) 586 138.3 4.4e-29  
gi|137600689|gb|EBU98255.1| hypothetical protein G ( 198) 582 137.1 4.5e-29  
gi|139197803|gb|ECE08220.1| hypothetical protein G ( 250) 583 137.4 4.6e-29  
gi|138511461|gb|ECA10116.1| hypothetical protein G ( 262) 583 137.4 4.8e-29  
gi|135500999|gb|EBI05114.1| hypothetical protein G ( 325) 584 137.7 4.8e-29  
gi|142626295|gb|ECZ42151.1| hypothetical protein G ( 270) 583 137.5 4.9e-29  
gi|135169737|gb|EBF93992.1| hypothetical protein G ( 317) 583 137.5 5.5e-29  
gi|140948059|gb|ECO59318.1| hypothetical protein G ( 222) 581 137.0 5.7e-29  
gi|137232547|gb|EBS94634.1| hypothetical protein G ( 237) 581 137.0 6e-29  
gi|141393185|gb|ECD51347.1| hypothetical protein G ( 271) 581 137.0 6.6e-29  
gi|142266870|gb|ECW83884.1| hypothetical protein G ( 185) 579 136.5 6.6e-29  
gi|137557613|gb|EBU74085.1| hypothetical protein G ( 192) 579 136.5 6.8e-29  
gi|139040947|gb|ECA00081.1| hypothetical protein G ( 204) 579 136.5 7.2e-29  
gi|136231925|gb|EBM76868.1| hypothetical protein G ( 230) 579 136.5 7.9e-29  
gi|136568435|gb|EBP00216.1| hypothetical protein G ( 379) 581 137.1 8.6e-29  
gi|136609934|gb|EBP26903.1| hypothetical protein G ( 394) 581 137.1 8.8e-29  
gi|136346125|gb|EBN54406.1| hypothetical protein G ( 228) 578 136.3 9.1e-29  
gi|137068138|gb|EBS02685.1| hypothetical protein G ( 198) 577 136.0 9.5e-29  
gi|144114617|gb|EDT96951.1| hypothetical protein G ( 245) 578 136.3 9.6e-29  
gi|136184878|gb|EBM45270.1| hypothetical protein G ( 202) 577 136.1 9.6e-29  
gi|136484986|gb|EBO46689.1| hypothetical protein G ( 259) 578 136.4 1e-28  
gi|135790786|gb|EBJ86387.1| hypothetical protein G ( 225) 577 136.1 1e-28  
gi|134492839|gb|EBB67196.1| hypothetical protein G ( 258) 577 136.1 1.2e-28  
gi|138858289|gb|ECC17909.1| hypothetical protein G ( 275) 577 136.2 1.2e-28  
gi|136253248|gb|EBM91261.1| hypothetical protein G ( 304) 577 136.2 1.3e-28  
gi|143058690|gb|EDC51173.1| hypothetical protein G ( 257) 576 135.9 1.3e-28  
gi|141132797|gb|ECP85624.1| hypothetical protein G ( 322) 577 136.2 1.4e-28  
gi|138387052|gb|EBZ31733.1| hypothetical protein G ( 294) 576 136.0 1.5e-28  
gi|134347423|gb|EBA80130.1| hypothetical protein G ( 359) 577 136.2 1.5e-28  
gi|142019072|gb|ECU88872.1| hypothetical protein G ( 248) 575 135.7 1.5e-28  
gi|14492761|gb|EBB67150.1| hypothetical protein G ( 205) 574 135.4 1.5e-28  
gi|141338252|gb|ECR14195.1| hypothetical protein G ( 265) 575 135.7 1.6e-28  
gi|136318277|gb|EBN35598.1| hypothetical protein G ( 325) 576 136.0 1.6e-28  
gi|139715639|gb|ECG97589.1| hypothetical protein G ( 197) 573 135.2 1.7e-28  
gi|141935579|gb|ECU28286.1| hypothetical protein G ( 164) 572 134.9 1.7e-28  
gi|140780154|gb|ECN45628.1| hypothetical protein G ( 245) 574 135.5 1.8e-28  
gi|143635271|gb|EDF98232.1| hypothetical protein G ( 375) 576 136.0 1.8e-28  
gi|136290187|gb|EBN16555.1| hypothetical protein G ( 258) 574 135.5 1.8e-28  
gi|141308402|gb|EBC65556.1| hypothetical protein G ( 258) 574 135.5 1.8e-28  
gi|140265299|gb|ECK65294.1| hypothetical protein G ( 178) 572 134.9 1.9e-28  
gi|136929793|gb|EBR24575.1| hypothetical protein G ( 270) 574 135.5 1.9e-28  
gi|141109595|gb|ECP69313.1| hypothetical protein G ( 187) 572 134.9 1.9e-28  
gi|135208488|gb|EBG18075.1| hypothetical protein G ( 230) 573 135.2 1.9e-28  
gi|144118385|gb|EDT99657.1| hypothetical protein G ( 344) 575 135.8 2e-28  
gi|139479977|gb|ECF36382.1| hypothetical protein G ( 233) 573 135.2 2e-28  
gi|138410663|gb|EBZ48146.1| hypothetical protein G ( 221) 572 135.0 2.2e-28  
gi|139206262|gb|ECN13862.1| hypothetical protein G ( 225) 572 135.0 2.2e-28  
gi|138473826|gb|EBZ92124.1| hypothetical protein G ( 187) 571 134.7 2.2e-28  
gi|136438974|gb|EBO16938.1| hypothetical protein G ( 254) 572 135.0 2.4e-28  
gi|140760372|gb|ECN33022.1| hypothetical protein G ( 188) 570 134.5 2.6e-28  
gi|140425672|gb|ECL67335.1| hypothetical protein G ( 252) 571 134.8 2.8e-28  
gi|139040218|gb|ECC99544.1| hypothetical protein G ( 258) 571 134.8 2.9e-28  
gi|140296649|gb|ECK80535.1| hypothetical protein G ( 267) 571 134.8 3e-28  
gi|141975411|gb|ECU55913.1| hypothetical protein G ( 246) 570 134.6 3.2e-28  
gi|136884800|gb|EBR04358.1| hypothetical protein G ( 184) 568 134.1 3.5e-28  
gi|142623014|gb|ECZ39816.1| hypothetical protein G ( 333) 571 134.9 3.5e-28  
gi|143225744|gb|EDD72527.1| hypothetical protein G ( 346) 571 134.9 3.6e-28

gi 136616380 gb EBP30906.1	hypothetical protein G ( 354)	571	134.9	3.7e-28
gi 139119268 gb ECD54583.1	hypothetical protein G ( 249)	569	134.4	3.8e-28
gi 139492061 gb ECF43960.1	hypothetical protein G ( 205)	568	134.1	3.8e-28
gi 139865703 gb ECI00815.1	hypothetical protein G ( 254)	569	134.4	3.8e-28
gi 134872290 gb EBD98647.1	hypothetical protein G ( 383)	571	135.0	3.9e-28
gi 139687054 gb ECG77816.1	hypothetical protein G ( 217)	568	134.1	4e-28
gi 136075621 gb EBL74549.1	hypothetical protein G ( 396)	571	135.0	4e-28
gi 135964875 gb EBL01570.1	hypothetical protein G ( 339)	570	134.7	4.1e-28
gi 140139580 gb ECJ80221.1	hypothetical protein G ( 184)	566	133.6	4.7e-28
gi 143392869 gb EDE71678.1	hypothetical protein G ( 331)	569	134.5	4.7e-28
gi 138159520 gb EBY09503.1	hypothetical protein G ( 227)	567	133.9	4.8e-28
gi 142653024 gb ECZ61002.1	hypothetical protein G ( 229)	567	133.9	4.8e-28
gi 135377215 gb EBH22814.1	hypothetical protein G ( 340)	569	134.5	4.8e-28
gi 142389451 gb ECX71311.1	hypothetical protein G ( 247)	567	133.9	5.1e-28
gi 137889824 gb EBW58476.1	hypothetical protein G ( 260)	567	134.0	5.3e-28
gi 143038670 gb EDC36519.1	hypothetical protein G ( 389)	569	134.5	5.3e-28
gi 140309242 gb ECK88285.1	hypothetical protein G ( 285)	567	134.0	5.7e-28
gi 135121196 gb EBF62739.1	hypothetical protein G ( 199)	565	133.4	5.8e-28
gi 134339575 gb EBA74899.1	hypothetical protein G ( 199)	565	133.4	5.8e-28
gi 135914634 gb EBK67553.1	hypothetical protein G ( 250)	566	133.7	6e-28
gi 134863258 gb EBD92485.1	hypothetical protein G ( 172)	564	133.2	6e-28
gi 138424579 gb EBZ57865.1	hypothetical protein G ( 257)	566	133.7	6.1e-28
gi 143691937 gb EDG30929.1	hypothetical protein G ( 336)	567	134.0	6.4e-28
gi 137468525 gb EBU26841.1	hypothetical protein G ( 240)	565	133.5	6.7e-28
gi 136969398 gb EBR46933.1	hypothetical protein G ( 213)	564	133.2	7.1e-28
gi 134551651 gb EBC01843.1	hypothetical protein G ( 326)	566	133.8	7.3e-28
gi 142551184 gb ECY89604.1	hypothetical protein G ( 290)	565	133.6	7.8e-28
gi 142423272 gb ECX96633.1	hypothetical protein G ( 259)	564	133.3	8.3e-28
gi 143371316 gb EDE61828.1	hypothetical protein G ( 260)	564	133.3	8.3e-28
gi 135716657 gb EBJ740232.1	hypothetical protein G ( 326)	565	133.6	8.5e-28
gi 139696462 gb ECG84095.1	hypothetical protein G ( 226)	563	133.0	8.7e-28
gi 138340797 gb EBZ07170.1	hypothetical protein G ( 188)	562	132.8	8.7e-28
gi 138908490 gb ECC37780.1	hypothetical protein G ( 241)	563	133.1	9.1e-28
gi 137948089 gb EBW91495.1	hypothetical protein G ( 201)	562	132.8	9.2e-28
gi 142922285 gb EDB53971.1	hypothetical protein G ( 362)	565	133.6	9.2e-28
gi 135941869 gb EBK80602.1	hypothetical protein G ( 402)	565	133.7	1e-27
gi 135240692 gb EBG36991.1	hypothetical protein G ( 283)	563	133.1	1e-27
gi 138795144 gb ECB889357.1	hypothetical protein G ( 240)	562	132.8	1.1e-27
gi 134868092 gb EBD95850.1	hypothetical protein G ( 246)	562	132.9	1.1e-27
gi 138512645 gb ECA10915.1	hypothetical protein G ( 247)	562	132.9	1.1e-27
gi 141158183 gb ECK03135.1	hypothetical protein G ( 233)	561	132.6	1.2e-27
gi 135590928 gb EBI62468.1	hypothetical protein G ( 298)	562	132.9	1.2e-27
gi 72395592 gb AAZ69865.1	3-phosphoshikimate 1-ca ( 443)	564	133.5	1.3e-27
gi 167284551 gb ABZ37415.1	Sequence 11353 from pa ( 424)	563	133.2	1.4e-27
gi 141894282 gb ECT99316.1	hypothetical protein G ( 237)	560	132.4	1.4e-27
gi 137705280 gb EBV55748.1	hypothetical protein G ( 244)	560	132.4	1.4e-27
gi 143559119 gb EDF62999.1	hypothetical protein G ( 319)	561	132.7	1.5e-27
gi 139477785 gb ECF35725.1	hypothetical protein G ( 179)	558	131.9	1.5e-27
gi 141656055 gb ECS67288.1	hypothetical protein G ( 242)	559	132.2	1.7e-27
gi 138863194 gb ECC19987.1	hypothetical protein G ( 245)	559	132.2	1.7e-27
gi 134961533 gb EBE57972.1	hypothetical protein G ( 364)	561	132.8	1.7e-27
gi 136107740 gb EBL96408.1	hypothetical protein G ( 251)	559	132.2	1.7e-27
gi 141118072 gb ECP75106.1	hypothetical protein G ( 266)	559	132.2	1.8e-27
gi 143233038 gb EDD77786.1	hypothetical protein G ( 304)	559	132.3	2e-27
gi 143554952 gb EDF60809.1	hypothetical protein G ( 325)	559	132.3	2.1e-27
gi 142807522 gb EDA73087.1	hypothetical protein G ( 335)	559	132.3	2.1e-27
gi 140786812 gb ECN50244.1	hypothetical protein G ( 233)	557	131.7	2.2e-27
gi 139608209 gb ECG23917.1	hypothetical protein G ( 241)	557	131.8	2.3e-27
gi 136366647 gb EBN68381.1	hypothetical protein G ( 305)	558	132.1	2.3e-27
gi 136116511 gb EBM02403.1	hypothetical protein G ( 311)	558	132.1	2.4e-27
gi 136336798 gb EBN48052.1	hypothetical protein G ( 321)	558	132.1	2.4e-27
gi 136526549 gb EB073484.1	hypothetical protein G ( 337)	558	132.1	2.5e-27

gi 139515305 gb ECF59981.1	hypothetical protein G ( 284)	557	131.8	2.6e-27
gi 141806782 gb ECT38147.1	hypothetical protein G ( 247)	556	131.5	2.7e-27
gi 140158907 gb ECJ91707.1	hypothetical protein G ( 264)	556	131.6	2.8e-27
gi 138537975 gb ECA28323.1	hypothetical protein G ( 183)	554	131.0	2.9e-27
gi 138798250 gb ECB90802.1	hypothetical protein G ( 239)	555	131.3	3e-27
gi 135373751 gb EBH20476.1	hypothetical protein G ( 359)	557	131.9	3.1e-27
gi 136321702 gb EBN37871.1	hypothetical protein G ( 322)	556	131.6	3.3e-27
gi 139196107 gb ECE07005.1	hypothetical protein G ( 218)	554	131.1	3.3e-27
gi 140882192 gb ECO14647.1	hypothetical protein G ( 183)	553	130.8	3.3e-27
gi 135576705 gb EBI53580.1	hypothetical protein G ( 346)	556	131.7	3.5e-27
gi 137023050 gb EBR77234.1	hypothetical protein G ( 197)	553	130.8	3.5e-27
gi 169815598 gb ACA90181.1	3-phosphoshikimate 1-c ( 434)	557	131.9	3.6e-27
gi 198037329 emb CAR53255.1	3-phosphoshikimate 1- ( 434)	557	131.9	3.6e-27
gi 137920165 gb EBW75787.1	hypothetical protein G ( 200)	553	130.8	3.6e-27
gi 143557400 gb EDF62059.1	hypothetical protein G ( 254)	554	131.1	3.7e-27
gi 143843065 gb EDH02621.1	hypothetical protein G ( 387)	556	131.7	3.8e-27
gi 124870769 gb EAY62485.1	5-enolpyruvylshikimate ( 479)	557	132.0	3.8e-27
gi 137968623 gb EBI03144.1	hypothetical protein G ( 233)	553	130.9	4e-27
gi 135893800 gb EBI53064.1	hypothetical protein G ( 207)	552	130.6	4.3e-27
gi 139071862 gb ECD21725.1	hypothetical protein G ( 232)	552	130.7	4.7e-27
gi 187716173 gb ACD17397.1	3-phosphoshikimate 1-c ( 434)	555	131.5	4.8e-27
gi 134551891 gb EBC01982.1	hypothetical protein G ( 329)	553	131.0	5.2e-27
gi 137566498 gb EBU79105.1	hypothetical protein G ( 225)	551	130.4	5.3e-27
gi 91688767 gb ABE31967.1	3-phosphoshikimate 1-ca ( 434)	554	131.3	5.6e-27
gi 135226702 gb EBG28755.1	hypothetical protein G ( 310)	552	130.8	5.8e-27
gi 134952781 gb EBE52135.1	hypothetical protein G ( 377)	553	131.0	5.8e-27
gi 139245948 gb ECS35232.1	hypothetical protein G ( 215)	550	130.2	5.9e-27
gi 135604263 gb EBI70715.1	hypothetical protein G ( 320)	552	130.8	6e-27
gi 135483426 gb EBH93800.1	hypothetical protein G ( 347)	552	130.8	6.4e-27
gi 105892312 gb ABF75477.1	3-phosphoshikimate 1-c ( 434)	553	131.1	6.5e-27
gi 139636870 gb ECG42399.1	hypothetical protein G ( 253)	550	130.2	6.7e-27
gi 138847143 gb ECC13990.1	hypothetical protein G ( 217)	549	130.0	6.9e-27
gi 136947697 gb EBR34653.1	hypothetical protein G ( 240)	549	130.0	7.5e-27
gi 237875728 gb ACR28061.1	Prephenate dehydrogena ( 432)	552	130.9	7.5e-27
gi 142836321 gb EDA94821.1	hypothetical protein G ( 438)	552	130.9	7.6e-27
gi 134979326 gb EBE69999.1	hypothetical protein G ( 212)	548	129.8	7.9e-27
gi 143696142 gb EDG32910.1	hypothetical protein G ( 270)	549	130.1	8.2e-27
gi 137405532 gb EBT91789.1	hypothetical protein G ( 308)	549	130.1	9.1e-27
gi 137807735 gb EBW11120.1	hypothetical protein G ( 227)	547	129.6	9.7e-27
gi 140684627 gb ECM80425.1	hypothetical protein G ( 188)	546	129.3	9.8e-27
gi 140227603 gb ECK39251.1	hypothetical protein G ( 192)	546	129.3	9.9e-27
gi 135486766 gb EBH95950.1	hypothetical protein G ( 350)	549	130.1	1e-26
gi 116647157 gb ABK07798.1	3-phosphoshikimate 1-c ( 434)	550	130.4	1e-26
gi 139664048 gb ECG61622.1	hypothetical protein G ( 174)	545	129.0	1.1e-26
gi 141904797 gb ECU06828.1	hypothetical protein G ( 225)	546	129.3	1.1e-26
gi 137657757 gb EBV28937.1	hypothetical protein G ( 277)	547	129.6	1.1e-26
gi 138799594 gb ECB91450.1	hypothetical protein G ( 188)	545	129.1	1.1e-26
gi 142502870 gb ECY54666.1	hypothetical protein G ( 244)	546	129.4	1.2e-26
gi 135619300 gb EBI80064.1	hypothetical protein G ( 216)	545	129.1	1.3e-26
gi 134905326 gb EBE20586.1	hypothetical protein G ( 322)	547	129.7	1.3e-26
gi 135336253 gb EBG95349.1	hypothetical protein G ( 222)	545	129.1	1.3e-26
gi 144092979 gb EDT81512.1	hypothetical protein G ( 224)	545	129.1	1.3e-26
gi 139718797 gb ECG99847.1	hypothetical protein G ( 282)	546	129.4	1.3e-26
gi 141939792 gb ECU31133.1	hypothetical protein G ( 241)	545	129.1	1.4e-26
gi 142106223 gb ECV63468.1	hypothetical protein G ( 260)	545	129.2	1.5e-26
gi 142174489 gb ECW14547.1	hypothetical protein G ( 215)	544	128.9	1.5e-26
gi 135961619 gb EBK99364.1	hypothetical protein G ( 334)	546	129.5	1.5e-26
gi 141370398 gb ECS35261.1	hypothetical protein G ( 229)	544	128.9	1.5e-26
gi 139693729 gb ECG82133.1	hypothetical protein G ( 231)	544	128.9	1.6e-26
gi 137265443 gb EBT13073.1	hypothetical protein G ( 285)	545	129.2	1.6e-26
gi 141829576 gb ECT53906.1	hypothetical protein G ( 246)	544	128.9	1.6e-26
gi 139685248 gb ECG76510.1	hypothetical protein G ( 210)	543	128.7	1.7e-26

gi|142659509|gb|EC265576.1| hypothetical protein G ( 383) 546 129.5 1.7e-26  
gi|136941044|gb|EBR30919.1| hypothetical protein G ( 316) 545 129.2 1.7e-26  
gi|141897485|gb|ECU01623.1| hypothetical protein G ( 270) 544 129.0 1.7e-26  
gi|141433718|gb|ECR79657.1| hypothetical protein G ( 199) 542 128.4 1.9e-26  
gi|137239164|gb|EBS98164.1| hypothetical protein G ( 242) 543 128.7 1.9e-26  
gi|141866828|gb|ECT80295.1| hypothetical protein G ( 172) 541 128.2 1.9e-26  
gi|135865254|gb|EBK33280.1| hypothetical protein G ( 313) 544 129.0 2e-26  
gi|143916119|gb|EDH55081.1| hypothetical protein G ( 216) 542 128.5 2e-26  
gi|139589323|gb|ECG18068.1| hypothetical protein G ( 225) 542 128.5 2.1e-26  
gi|141665863|gb|ECS70196.1| hypothetical protein G ( 226) 542 128.5 2.1e-26  
gi|135896558|gb|EBK54995.1| hypothetical protein G ( 345) 544 129.0 2.1e-26  
gi|138530901|gb|ECA23368.1| hypothetical protein G ( 287) 543 128.8 2.1e-26  
gi|140405936|gb|ECL55694.1| hypothetical protein G ( 232) 541 128.3 2.4e-26  
gi|134533954|gb|EBB91312.1| hypothetical protein G ( 242) 541 128.3 2.5e-26  
gi|137214519|gb|EBS84393.1| hypothetical protein G ( 295) 542 128.6 2.5e-26  
gi|134450625|gb|EBB42108.1| hypothetical protein G ( 201) 540 128.0 2.5e-26  
gi|83655416|gb|ABC39479.1| prephenate dehydrogenas ( 805) 547 130.0 2.6e-26  
gi|134849254|gb|EBD83368.1| hypothetical protein G ( 275) 541 128.3 2.8e-26  
gi|134981203|gb|EBE71267.1| hypothetical protein G ( 292) 541 128.3 2.9e-26  
gi|136206440|gb|EBM59627.1| hypothetical protein G ( 245) 540 128.1 3e-26  
gi|138641254|gb|ECA99072.1| hypothetical protein G ( 302) 541 128.3 3e-26  
gi|139967027|gb|ECL69743.1| hypothetical protein G ( 306) 541 128.4 3e-26  
gi|140654805|gb|ECM59434.1| hypothetical protein G ( 256) 540 128.1 3.1e-26  
gi|137329713|gb|EBT48905.1| hypothetical protein G ( 223) 539 127.8 3.2e-26  
gi|77966375|gb|ABBO7755.1| 3-phosphoshikimate 1-ca ( 434) 542 128.7 3.4e-26  
gi|141333425|gb|ECR11074.1| hypothetical protein G ( 296) 540 128.1 3.4e-26  
gi|135645619|gb|EBI96344.1| hypothetical protein G ( 306) 540 128.1 3.5e-26  
gi|141671849|gb|ECS73496.1| hypothetical protein G ( 314) 540 128.1 3.6e-26  
gi|140178654|gb|ECK05494.1| hypothetical protein G ( 222) 538 127.6 3.7e-26  
gi|141882968|gb|ECT91410.1| hypothetical protein G ( 185) 537 127.3 3.8e-26  
gi|141048681|gb|ECP27929.1| hypothetical protein G ( 290) 539 127.9 3.9e-26  
gi|142109868|gb|ECV66243.1| hypothetical protein G ( 524) 542 128.7 3.9e-26  
gi|135841164|gb|EBK18114.1| hypothetical protein G ( 312) 539 127.9 4.2e-26  
gi|141927970|gb|ECU22932.1| hypothetical protein G ( 327) 539 127.9 4.3e-26  
gi|142998533|gb|EDC07332.1| hypothetical protein G ( 227) 537 127.4 4.4e-26  
gi|138499014|gb|ECA02780.1| hypothetical protein G ( 237) 537 127.4 4.5e-26  
gi|189333856|dbj|BAG42926.1| 3-phosphoshikimate 1- ( 434) 540 128.2 4.6e-26  
gi|160342858|gb|ABX15944.1| 3-phosphoshikimate 1-c ( 434) 540 128.2 4.6e-26  
gi|55274195|gb|AAV48979.1| hypothetical carboxyvin ( 202) 536 127.1 4.7e-26  
gi|139629194|gb|ECG37093.1| hypothetical protein G ( 306) 538 127.7 4.8e-26  
gi|135327266|gb|EBG89302.1| hypothetical protein G ( 357) 538 127.8 5.4e-26  
gi|167281782|gb|ABZ34646.1| Sequence 8584 from pat ( 435) 539 128.0 5.4e-26  
gi|222436771|gb|EEE43450.1| hypothetical protein S ( 136) 533 126.3 5.4e-26  
gi|142105249|gb|ECG62774.1| hypothetical protein G ( 305) 537 127.5 5.5e-26  
gi|93353516|gb|ABF07605.1| 3-phosphoshikimate 1-ca ( 452) 539 128.0 5.5e-26  
gi|137591925|gb|EBU93376.1| hypothetical protein G ( 256) 536 127.2 5.6e-26  
gi|135239148|gb|EBG36086.1| hypothetical protein G ( 260) 536 127.2 5.7e-26  
gi|140346704|gb|ECL14523.1| hypothetical protein G ( 289) 536 127.2 6.2e-26  
gi|140909158|gb|ECO32037.1| hypothetical protein G ( 197) 534 126.7 6.2e-26  
gi|136287602|gb|EBN14795.1| hypothetical protein G ( 240) 535 127.0 6.2e-26  
gi|171992497|gb|ACB63416.1| 3-phosphoshikimate 1-c ( 434) 538 127.8 6.2e-26  
gi|141468959|gb|ECS00232.1| hypothetical protein G ( 280) 535 127.0 7e-26  
gi|139681527|gb|ECG73843.1| hypothetical protein G ( 233) 534 126.7 7.1e-26  
gi|222452024|gb|ACM56289.1| 3-phosphoshikimate 1-c ( 453) 537 127.6 7.5e-26  
gi|138983048|gb|ECC68255.1| hypothetical protein G ( 312) 535 127.1 7.6e-26  
gi|142861145|gb|EBD09899.1| hypothetical protein G ( 338) 535 127.1 8.1e-26  
gi|139853911|gb|ECH92771.1| hypothetical protein G ( 292) 534 126.8 8.4e-26  
gi|135288143|gb|EBG64892.1| hypothetical protein G ( 207) 532 126.3 8.7e-26  
gi|139064451|gb|ECD16608.1| hypothetical protein G ( 223) 532 126.3 9.2e-26  
gi|136661056|gb|EBP57372.1| hypothetical protein G ( 184) 531 126.0 9.2e-26  
gi|135317778|gb|EBG82900.1| hypothetical protein G ( 410) 535 127.1 9.4e-26  
gi|141709826|gb|ECS86859.1| hypothetical protein G ( 234) 532 126.3 9.6e-26

gi|139040948|gb|ECD00082.1| hypothetical protein G ( 237) 532 126.3 9.7e-26  
gi|110729299|gb|ABG88197.1| EPSP synthase [Phaseol ( 522) 536 127.4 9.7e-26  
gi|140690954|gb|ECM8480.1| hypothetical protein G ( 298) 533 126.6 9.9e-26  
gi|140343196|gb|ECL11994.1| hypothetical protein G ( 304) 533 126.6 1e-25  
gi|135393659|gb|EBH33889.1| hypothetical protein G ( 261) 532 126.3 1e-25  
gi|138276283|gb|EBY74911.1| hypothetical protein G ( 308) 532 126.4 1.2e-25  
gi|136443021|gb|EBO19562.1| hypothetical protein G ( 457) 534 127.0 1.2e-25  
gi|140028597|gb|ECU10751.1| hypothetical protein G ( 312) 532 126.4 1.2e-25  
gi|136315894|gb|EBN34015.1| hypothetical protein G ( 280) 531 126.1 1.3e-25  
gi|134967391|gb|EBE61919.1| hypothetical protein G ( 354) 532 126.4 1.3e-25  
gi|184191968|gb|ACC69933.1| 3-phosphoshikimate 1-c ( 434) 533 126.7 1.3e-25  
gi|142905356|gb|EDB41681.1| hypothetical protein G ( 199) 529 125.6 1.3e-25  
gi|143715743|gb|EDG41937.1| hypothetical protein G ( 243) 530 125.9 1.3e-25  
gi|138116427|gb|EBX83452.1| hypothetical protein G ( 297) 531 126.2 1.3e-25  
gi|141915818|gb|ECU14327.1| hypothetical protein G ( 249) 530 125.9 1.4e-25  
gi|141188377|gb|ECC24336.1| hypothetical protein G ( 306) 531 126.2 1.4e-25  
gi|135670495|gb|EBJ11737.1| hypothetical protein G ( 316) 531 126.2 1.4e-25  
gi|134138234|gb|ABC03977.1| 3-phosphoshikimate 1-c ( 434) 532 126.5 1.5e-25  
gi|137178927|gb|EBE64474.1| hypothetical protein G ( 249) 529 125.7 1.6e-25  
gi|139824132|gb|ECH71785.1| hypothetical protein G ( 238) 528 125.4 1.8e-25  
gi|115280963|gb|ABT86480.1| 3-phosphoshikimate 1-c ( 434) 531 126.3 1.8e-25  
gi|139196108|gb|ECC07006.1| hypothetical protein G ( 207) 527 125.2 1.9e-25  
gi|135564677|gb|EBI45872.1| hypothetical protein G ( 306) 529 125.7 1.9e-25  
gi|141545034|gb|ECS26839.1| hypothetical protein G ( 254) 528 125.5 1.9e-25  
gi|135871899|gb|EBK37709.1| hypothetical protein G ( 255) 528 125.5 1.9e-25  
gi|135477171|gb|EBH89817.1| hypothetical protein G ( 384) 530 126.0 1.9e-25  
gi|137171985|gb|EBE05590.1| hypothetical protein G ( 272) 528 125.5 2e-25  
gi|143039536|gb|EDC37155.1| hypothetical protein G ( 336) 529 125.8 2e-25  
gi|140788599|gb|ECN51475.1| hypothetical protein G ( 235) 527 125.2 2e-25  
gi|141712157|gb|ECS88316.1| hypothetical protein G ( 216) 526 125.0 2.2e-25  
gi|134604121|gb|EBC33250.1| hypothetical protein G ( 325) 528 125.5 2.3e-25  
gi|139732419|gb|ECH09173.1| hypothetical protein G ( 227) 526 125.0 2.3e-25  
gi|140472221|gb|ECL93917.1| hypothetical protein G ( 284) 527 125.3 2.4e-25  
gi|138558264|gb|ECA42657.1| hypothetical protein G ( 291) 527 125.3 2.4e-25  
gi|142394074|gb|ECX74728.1| hypothetical protein G ( 296) 527 125.3 2.4e-25  
gi|140118856|gb|ECU67643.1| hypothetical protein G ( 210) 525 124.7 2.5e-25  
gi|137547042|gb|EBU68034.1| hypothetical protein G ( 312) 527 125.3 2.5e-25  
gi|136980766|gb|EBR53373.1| hypothetical protein G ( 285) 526 125.1 2.8e-25  
gi|113525597|emb|CAJ91942.1| 3-Enolpyruvylshikimat ( 434) 528 125.6 2.8e-25  
gi|139494432|gb|ECF45603.1| hypothetical protein G ( 305) 526 125.1 2.9e-25  
gi|138110485|gb|EBX80097.1| hypothetical protein G ( 218) 524 124.5 3e-25  
gi|141385612|gb|ECR46209.1| hypothetical protein G ( 187) 523 124.3 3.1e-25  
gi|139454002|gb|ECF21045.1| hypothetical protein G ( 284) 525 124.8 3.2e-25  
gi|136364219|gb|EBN66733.1| hypothetical protein G ( 310) 525 124.9 3.4e-25  
gi|143591543|gb|EDF75609.1| hypothetical protein G ( 219) 523 124.3 3.5e-25  
gi|137657142|gb|EBV28587.1| hypothetical protein G ( 276) 524 124.6 3.6e-25  
gi|141543043|gb|ECS25410.1| hypothetical protein G ( 294) 524 124.6 3.8e-25  
gi|138799362|gb|ECB91357.1| hypothetical protein G ( 298) 524 124.6 3.9e-25  
gi|135489629|gb|EBH97784.1| hypothetical protein G ( 363) 525 124.9 3.9e-25  
gi|136969921|gb|EBR47225.1| hypothetical protein G ( 306) 524 124.7 3.9e-25  
gi|140748856|gb|ECN25154.1| hypothetical protein G ( 312) 524 124.7 4e-25  
gi|138613133|gb|ECA81107.1| hypothetical protein G ( 326) 524 124.7 4.1e-25  
gi|137375246|gb|EBT74514.1| hypothetical protein G ( 226) 522 124.1 4.2e-25  
gi|142865504|gb|EDB13092.1| hypothetical protein G ( 227) 522 124.1 4.2e-25  
gi|139995621|gb|ECT89691.1| hypothetical protein G ( 284) 523 124.4 4.3e-25  
gi|137691967|gb|EBV48315.1| hypothetical protein G ( 286) 523 124.4 4.3e-25  
gi|142607018|gb|ECZ28589.1| hypothetical protein G ( 202) 521 123.9 4.5e-25  
gi|135845227|gb|EBK20663.1| hypothetical protein G ( 303) 523 124.4 4.5e-25  
gi|142543985|gb|ECY84528.1| hypothetical protein G ( 268) 522 124.2 4.8e-25  
gi|135854902|gb|EBK26763.1| hypothetical protein G ( 236) 521 123.9 5.1e-25  
gi|138005019|gb|EBX23095.1| hypothetical protein G ( 288) 522 124.2 5.1e-25  
gi|138091258|gb|EBX69473.1| hypothetical protein G ( 175) 519 123.4 5.4e-25

gi 137912933 gb EBW71741.1	hypothetical protein G ( 279)	521	124.0	5.8e-25
gi 138660060 gb ECB12035.1	hypothetical protein G ( 284)	521	124.0	5.8e-25
gi 134549068 gb EBC00317.1	hypothetical protein G ( 241)	520	123.7	6e-25
gi 142411096 gb ECX87359.1	hypothetical protein G ( 436)	523	124.6	6e-25
gi 138550260 gb ECA37196.1	hypothetical protein G ( 251)	520	123.7	6.2e-25
gi 52428212 gb AAU48805.1	prephenate dehydrogenas ( 673)	525	125.1	6.2e-25
gi 138878119 gb ECC26539.1	hypothetical protein G ( 255)	520	123.7	6.3e-25
gi 137078318 gb EBS08436.1	hypothetical protein G ( 314)	521	124.0	6.3e-25
gi 76581095 gb ABA50570.1	prephenate dehydrogenas ( 714)	525	125.1	6.5e-25
gi 157805514 gb EDO82684.1	prephenate dehydrogena ( 740)	525	125.2	6.7e-25
gi 147745439 gb EDK52519.1	prephenate dehydrogena ( 749)	525	125.2	6.8e-25
gi 126243949 gb ABO07042.1	prephenate dehydrogena ( 749)	525	125.2	6.8e-25
gi 121227083 gb ABM49601.1	prephenate dehydrogena ( 749)	525	125.2	6.8e-25
gi 126219744 gb ABN83250.1	prephenate dehydrogena ( 749)	525	125.2	6.8e-25
gi 124294618 gb ABN03887.1	prephenate dehydrogena ( 749)	525	125.2	6.8e-25
gi 160697699 gb EDP87669.1	prephenate dehydrogena ( 749)	525	125.2	6.8e-25
gi 148029278 gb EDK87183.1	prephenate dehydrogena ( 749)	525	125.2	6.8e-25
gi 126226820 gb ABN90360.1	prephenate dehydrogena ( 749)	525	125.2	6.8e-25
gi 147750783 gb EDK57852.1	prephenate dehydrogena ( 749)	525	125.2	6.8e-25
gi 157987840 gb EDO95605.1	prephenate dehydrogena ( 753)	525	125.2	6.8e-25
gi 254218017 gb EET07401.1	prephenate dehydrogena ( 753)	525	125.2	6.8e-25
gi 52210541 emb CAH36524.1	putative bifunctional ( 753)	525	125.2	6.8e-25
gi 242138323 gb EES24725.1	prephenate dehydrogena ( 753)	525	125.2	6.8e-25
gi 141123386 gb ECP78845.1	hypothetical protein G ( 245)	519	123.5	7.1e-25
gi 142771693 gb EDA46374.1	hypothetical protein G ( 179)	517	123.0	7.5e-25
gi 2485229 gb AAB73365.1	I44454 Sequence 13 from p ( 105)	514	122.1	7.8e-25
gi 144974731 gb ABP12442.1	Sequence 13 from paten ( 105)	514	122.1	7.8e-25
gi 2484152 gb AAB72288.1	I49181 Sequence 13 from p ( 105)	514	122.1	7.8e-25
gi 5957552 gb AAE08226.1	Sequence 13 from patent ( 105)	514	122.1	7.8e-25
gi 184211626 gb EDU08669.1	prephenate dehydrogena ( 749)	524	124.9	7.9e-25
gi 143839650 gb EDH00163.1	hypothetical protein G ( 302)	519	123.6	8.3e-25
gi 48995007 gb AAT48249.1	chloroplast CP4-EPSPS f ( 141)	515	122.4	8.4e-25
gi 135489518 gb EBH97714.1	hypothetical protein G ( 271)	518	123.3	8.9e-25
gi 140119424 gb ECJ67138.1	hypothetical protein G ( 227)	517	123.0	9e-25
gi 136324373 gb EBN39656.1	hypothetical protein G ( 337)	519	123.6	9e-25
gi 167277472 gb ABZ20336.1	Sequence 4274 from pat ( 415)	520	123.9	9.1e-25
gi 150011975 gb ABR54427.1	3-phosphoshikimate 1-c ( 429)	520	123.9	9.4e-25
gi 139441542 gb ECF12258.1	hypothetical protein G ( 242)	517	123.1	9.4e-25
gi 138996659 gb ECC72241.1	hypothetical protein G ( 201)	516	122.8	9.5e-25
gi 135491229 gb EBH98820.1	hypothetical protein G ( 208)	516	122.8	9.8e-25
gi 140215327 gb ECK31464.1	hypothetical protein G ( 216)	516	122.8	1e-24
gi 142823924 gb EDA85437.1	hypothetical protein G ( 332)	518	123.4	1e-24
gi 142702742 gb ECZ96501.1	hypothetical protein G ( 231)	516	122.8	1.1e-24
gi 140985159 gb ECO85137.1	hypothetical protein G ( 283)	517	123.1	1.1e-24
gi 143159217 gb EDD24462.1	hypothetical protein G ( 283)	517	123.1	1.1e-24
gi 141847270 gb ECT66591.1	hypothetical protein G ( 241)	516	122.8	1.1e-24
gi 136230377 gb EBM75834.1	hypothetical protein G ( 307)	517	123.1	1.1e-24
gi 135215196 gb EBG22013.1	hypothetical protein G ( 259)	516	122.9	1.2e-24
gi 169652329 gb EDS85022.1	prephenate dehydrogena ( 749)	521	124.3	1.2e-24
gi 139643880 gb ECG47198.1	hypothetical protein G ( 301)	516	122.9	1.3e-24
gi 135686333 gb EBJ21480.1	hypothetical protein G ( 304)	516	122.9	1.3e-24
gi 237503144 gb ACQ95462.1	prephenate dehydrogena ( 749)	520	124.1	1.4e-24
gi 134442483 gb EBB37469.1	hypothetical protein G ( 285)	515	122.7	1.4e-24
gi 141110942 gb ECP70249.1	hypothetical protein G ( 196)	513	122.1	1.5e-24
gi 142656925 gb ECZ63763.1	hypothetical protein G ( 197)	513	122.1	1.5e-24
gi 138180415 gb EBY24121.1	hypothetical protein G ( 314)	515	122.7	1.6e-24
gi 139647519 gb ECG49757.1	hypothetical protein G ( 225)	513	122.2	1.6e-24
gi 137703414 gb EBV54784.1	hypothetical protein G ( 279)	514	122.4	1.7e-24
gi 139920801 gb ECI37947.1	hypothetical protein G ( 156)	511	121.6	1.7e-24
gi 139914190 gb ECI33276.1	hypothetical protein G ( 294)	514	122.5	1.7e-24
gi 143760794 gb EDG64991.1	hypothetical protein G ( 317)	514	122.5	1.8e-24
gi 138840632 gb ECC10892.1	hypothetical protein G ( 289)	513	122.2	2e-24

gi 138115352 gb EBX82843.1	hypothetical protein G ( 238)	512	122.0	2e-24
gi 135713521 gb EBJ38289.1	hypothetical protein G ( 353)	514	122.5	2e-24
gi 136509201 gb EBB062385.1	hypothetical protein G ( 297)	513	122.2	2e-24
gi 134531717 gb EBB89974.1	hypothetical protein G ( 201)	511	121.7	2e-24
gi 141139548 gb ECP90029.1	hypothetical protein G ( 249)	512	122.0	2.1e-24
gi 135795613 gb EBJ89439.1	hypothetical protein G ( 209)	511	121.7	2.1e-24
gi 136178233 gb EBM40882.1	hypothetical protein G ( 213)	511	121.7	2.1e-24
gi 135495643 gb EBI01665.1	hypothetical protein G ( 384)	514	122.5	2.1e-24
gi 138618036 gb ECA84358.1	hypothetical protein G ( 281)	512	122.0	2.3e-24
gi 141373745 gb ECR37668.1	hypothetical protein G ( 236)	511	121.7	2.3e-24
gi 139219178 gb ECE22826.1	hypothetical protein G ( 290)	512	122.0	2.3e-24
gi 140388086 gb ECL43376.1	hypothetical protein G ( 304)	512	122.0	2.4e-24
gi 143397410 gb EDE73971.1	hypothetical protein G ( 270)	511	121.8	2.5e-24
gi 136214158 gb EBM64824.1	hypothetical protein G ( 183)	509	121.2	2.5e-24
gi 134548409 gb EBB99923.1	hypothetical protein G ( 317)	511	121.8	2.9e-24
gi 142916528 gb EDB49892.1	hypothetical protein G ( 218)	509	121.3	2.9e-24
gi 179351349 gb ACB85619.1	3-phosphoshikimate 1-c ( 484)	513	122.4	3e-24
gi 135742000 gb EBU55937.1	hypothetical protein G ( 336)	511	121.9	3e-24
gi 139132650 gb ECR62535.1	hypothetical protein G ( 231)	509	121.3	3e-24
gi 142699289 gb ECZ94025.1	hypothetical protein G ( 199)	508	121.0	3.2e-24
gi 134788582 gb EBD44017.1	hypothetical protein G ( 150)	506	120.5	3.4e-24
gi 137536874 gb EBU62471.1	hypothetical protein G ( 271)	509	121.3	3.4e-24
gi 142775989 gb EDA49581.1	hypothetical protein G ( 335)	510	121.6	3.5e-24
gi 142277937 gb ECP90213.1	hypothetical protein G ( 234)	508	121.1	3.6e-24
gi 141879126 gb ECT88675.1	hypothetical protein G ( 237)	508	121.1	3.6e-24
gi 141219643 gb ECQ44682.1	hypothetical protein G ( 291)	509	121.4	3.6e-24
gi 193222745 emb CAQ268748.1	3-enolpyruvylshikimat ( 434)	511	121.9	3.7e-24
gi 134400710 gb EBB13457.1	hypothetical protein G ( 303)	509	121.4	3.8e-24
gi 137333376 gb EBT50948.1	hypothetical protein G ( 265)	508	121.1	3.9e-24
gi 136273877 gb EBN05547.1	hypothetical protein G ( 220)	507	120.8	4e-24
gi 136805390 gb EBQ52454.1	hypothetical protein G ( 225)	507	120.9	4e-24
gi 138147023 gb EBY00725.1	hypothetical protein G ( 239)	507	120.9	4.2e-24
gi 140173956 gb ECK02082.1	hypothetical protein G ( 291)	508	121.2	4.2e-24
gi 72119670 gb AAZ61933.1	3-phosphoshikimate 1-ca ( 434)	510	121.7	4.3e-24
gi 138822185 gb ECC01789.1	hypothetical protein G ( 244)	507	120.9	4.3e-24
gi 138838260 gb ECOC90710.1	hypothetical protein G ( 205)	506	120.6	4.4e-24
gi 141978626 gb ECU58234.1	hypothetical protein G ( 184)	505	120.3	4.7e-24
gi 139531150 gb ECF70548.1	hypothetical protein G ( 241)	506	120.7	5e-24
gi 124895538 gb EAY69418.1	5-enolpyruvylshikimate ( 434)	509	121.5	5e-24
gi 135888366 gb EBK49261.1	hypothetical protein G ( 208)	505	120.4	5.1e-24
gi 27311149 gb AAO00677.1	Sequence 10 from patent ( 80)	500	119.0	5.2e-24
gi 137246014 gb EBT02032.1	hypothetical protein G ( 268)	506	120.7	5.4e-24
gi 141653377 gb ECS66510.1	hypothetical protein G ( 279)	506	120.7	5.5e-24
gi 137929679 gb EBW81155.1	hypothetical protein G ( 200)	504	120.2	5.8e-24
gi 143057941 gb EDC50622.1	hypothetical protein G ( 296)	506	120.7	5.8e-24
gi 139195857 gb ECE06823.1	hypothetical protein G ( 298)	506	120.7	5.8e-24
gi 139453942 gb ECF21003.1	hypothetical protein G ( 205)	504	120.2	5.9e-24
gi 141592162 gb ECS49928.1	hypothetical protein G ( 315)	506	120.7	6.1e-24
gi 140673199 gb ECM72415.1	hypothetical protein G ( 227)	504	120.2	6.4e-24
gi 137849920 gb EBW35596.1	hypothetical protein G ( 238)	504	120.2	6.6e-24
gi 142560062 gb ECY95832.1	hypothetical protein G ( 296)	505	120.5	6.8e-24
gi 140999785 gb ECO95579.1	hypothetical protein G ( 168)	502	119.7	6.9e-24
gi 142690999 gb EBZ88084.1	hypothetical protein G ( 220)	503	120.0	7.3e-24
gi 141903494 gb ECU05918.1	hypothetical protein G ( 235)	503	120.0	7.6e-24
gi 19918689 gb AAM07883.1	3-phosphoshikimate 1-ca ( 430)	506	120.8	7.7e-24
gi 197053762 gb ACH25460.1	Sequence 12 from paten ( 430)	506	120.8	7.7e-24
gi 141723183 gb ECS95130.1	hypothetical protein G ( 197)	502	119.7	7.7e-24
gi 140248541 gb ECK53972.1	hypothetical protein G ( 295)	504	120.3	7.8e-24
gi 141081658 gb ECP49602.1	hypothetical protein G ( 200)	502	119.7	7.8e-24
gi 139559346 gb ECF90294.1	hypothetical protein G ( 223)	502	119.8	8.5e-24
gi 136628519 gb EBP38073.1	hypothetical protein G ( 229)	502	119.8	8.7e-24
gi 137863209 gb EBW43227.1	hypothetical protein G ( 283)	503	120.1	8.8e-24

gi|134737900|gb|EBD11884.1| hypothetical protein G ( 239) 502 119.8 9e-24  
gi|135427488|gb|EBH56660.1| hypothetical protein G ( 307) 503 120.1 9.4e-24  
gi|134602901|gb|EBC32497.1| hypothetical protein G ( 311) 503 120.1 9.5e-24  
gi|134518053|gb|EBB81858.1| hypothetical protein G ( 176) 500 119.2 9.6e-24  
gi|142717897|gb|EDA07347.1| hypothetical protein G ( 282) 502 119.8 1e-23  
gi|137036578|gb|EBR84858.1| hypothetical protein G ( 237) 501 119.6 1e-23  
gi|76558044|emb|CAI49630.1| 3-phosphoshikimate 1-c ( 438) 504 120.4 1.1e-23  
gi|137009311|gb|EBR69583.1| hypothetical protein G ( 300) 502 119.9 1.1e-23  
gi|144974730|gb|ABP12441.1| Sequence 11 from paten ( 77) 495 117.9 1.1e-23  
gi|2484151|gb|AAB72287.1|I49180 Sequence 11 from p ( 77) 495 117.9 1.1e-23  
gi|5957551|gb|AAE08225.1| Sequence 11 from patent ( 77) 495 117.9 1.1e-23  
gi|2485228|gb|AAB73364.1|I44453 Sequence 11 from p ( 77) 495 117.9 1.1e-23  
gi|137829749|gb|EBW24000.1| hypothetical protein G ( 258) 501 119.6 1.1e-23  
gi|141404015|gb|ECR59067.1| hypothetical protein G ( 286) 501 119.6 1.2e-23  
gi|135562509|gb|EBT44489.1| hypothetical protein G ( 348) 502 119.9 1.2e-23  
gi|187724763|gb|ACD25928.1| 3-phosphoshikimate 1-c ( 434) 503 120.2 1.2e-23  
gi|138810206|gb|ECB96158.1| hypothetical protein G ( 201) 499 119.1 1.2e-23  
gi|141673429|gb|ECS74370.1| hypothetical protein G ( 303) 501 119.6 1.3e-23  
gi|135589703|gb|EBT61715.1| hypothetical protein G ( 305) 501 119.6 1.3e-23  
gi|142592013|gb|ECZ18106.1| hypothetical protein G ( 213) 499 119.1 1.3e-23  
gi|135570872|gb|EBT49844.1| hypothetical protein G ( 275) 500 119.4 1.4e-23  
gi|140714513|gb|ECN01171.1| hypothetical protein G ( 281) 500 119.4 1.4e-23  
gi|139789418|gb|ECH47697.1| hypothetical protein G ( 204) 498 118.9 1.5e-23  
gi|136659405|gb|EBP56376.1| hypothetical protein G ( 310) 500 119.4 1.5e-23  
gi|140794555|gb|ECN55677.1| hypothetical protein G ( 215) 498 118.9 1.5e-23  
gi|135334320|gb|EBG94059.1| hypothetical protein G ( 290) 499 119.2 1.6e-23  
gi|143324668|gb|EDE33761.1| hypothetical protein G ( 198) 497 118.6 1.7e-23  
gi|138094881|gb|EBX71452.1| hypothetical protein G ( 293) 499 119.2 1.7e-23  
gi|135216633|gb|EBG22867.1| hypothetical protein G ( 147) 495 118.1 1.8e-23  
gi|137319505|gb|EBT43281.1| hypothetical protein G ( 279) 498 119.0 1.9e-23  
gi|137697594|gb|EBV51548.1| hypothetical protein G ( 239) 497 118.7 1.9e-23  
gi|136961265|gb|EBR42313.1| hypothetical protein G ( 240) 497 118.7 1.9e-23  
gi|240864488|gb|ACS62149.1| 3-phosphoshikimate 1-c ( 434) 500 119.5 1.9e-23  
gi|134404891|gb|EBB15769.1| hypothetical protein G ( 296) 498 119.0 1.9e-23  
gi|254589820|gb|ACT69182.1| putative 3-phosphoshik ( 440) 500 119.5 2e-23  
gi|138945048|gb|ECC53187.1| hypothetical protein G ( 249) 497 118.7 2e-23  
gi|135382545|gb|EBH26399.1| hypothetical protein G ( 290) 497 118.8 2.2e-23  
gi|141160798|gb|ECQ05001.1| hypothetical protein G ( 203) 495 118.2 2.3e-23  
gi|139190333|gb|ECE02883.1| hypothetical protein G ( 300) 497 118.8 2.3e-23  
gi|138445398|gb|EBZ72305.1| hypothetical protein G ( 174) 494 117.9 2.4e-23  
gi|144206042|gb|EDJ63793.1| hypothetical protein G ( 218) 495 118.2 2.4e-23  
gi|138737512|gb|ECB66482.1| hypothetical protein G ( 292) 496 118.5 2.6e-23  
gi|142325118|gb|ECX26898.1| hypothetical protein G ( 207) 494 118.0 2.7e-23  
gi|56387797|gb|AAV86384.1| 3-phosphoshikimate 1-ca ( 462) 498 119.1 2.7e-23  
gi|134674198|gb|EBC74845.1| hypothetical protein G ( 218) 494 118.0 2.8e-23  
gi|135066400|gb|EBF27712.1| hypothetical protein G ( 187) 493 117.7 2.9e-23  
gi|141235824|gb|ECQ56308.1| hypothetical protein G ( 189) 493 117.7 2.9e-23  
gi|134587656|gb|EBC23605.1| hypothetical protein G ( 190) 493 117.7 2.9e-23  
gi|139895987|gb|ECT120390.1| hypothetical protein G ( 236) 494 118.0 3e-23  
gi|141391411|gb|ECR50083.1| hypothetical protein G ( 236) 494 118.0 3e-23  
gi|140456535|gb|ECL88213.1| hypothetical protein G ( 161) 492 117.5 3e-23  
gi|281079466|gb|ADA36174.1| Sequence 11 from paten ( 76) 488 116.4 3.1e-23  
gi|138566080|gb|ECA48176.1| hypothetical protein G ( 303) 495 118.3 3.1e-23  
gi|137942810|gb|EBW88522.1| hypothetical protein G ( 173) 492 117.5 3.2e-23  
gi|137048868|gb|EBR91792.1| hypothetical protein G ( 182) 492 117.5 3.3e-23  
gi|136876301|gb|EBQ99430.1| hypothetical protein G ( 227) 493 117.8 3.4e-23  
gi|143115784|gb|EDC92899.1| hypothetical protein G ( 338) 495 118.4 3.4e-23  
gi|141104379|gb|ECP65577.1| hypothetical protein G ( 297) 494 118.1 3.6e-23  
gi|140108073|gb|ECJ59094.1| hypothetical protein G ( 122) 489 116.7 3.8e-23  
gi|141452906|gb|ECR92796.1| hypothetical protein G ( 271) 493 117.9 3.9e-23  
gi|141656057|gb|ECS67290.1| hypothetical protein G ( 224) 492 117.6 3.9e-23  
gi|137527307|gb|EBU57304.1| hypothetical protein G ( 155) 490 117.0 3.9e-23

gi|138008762|gb|EBX25078.1| hypothetical protein G ( 191) 491 117.3 4e-23  
gi|140375918|gb|ECL34759.1| hypothetical protein G ( 290) 493 117.9 4.1e-23  
gi|222419069|gb|ACM49092.1| 3-phosphoshikimate 1-c ( 462) 495 118.5 4.3e-23  
gi|137324469|gb|EBT46018.1| hypothetical protein G ( 274) 492 117.6 4.5e-23  
gi|141899038|gb|ECU02730.1| hypothetical protein G ( 285) 492 117.7 4.7e-23  
gi|143255766|gb|EDD93469.1| hypothetical protein G ( 298) 492 117.7 4.8e-23  
gi|140676988|gb|ECM75154.1| hypothetical protein G ( 137) 488 116.6 4.8e-23  
gi|139669584|gb|ECG65360.1| hypothetical protein G ( 219) 490 117.1 5.1e-23  
gi|134446918|gb|EBB39935.1| hypothetical protein G ( 234) 490 117.2 5.4e-23  
gi|137834745|gb|EBW26918.1| hypothetical protein G ( 289) 491 117.4 5.5e-23  
gi|137322978|gb|EBT45218.1| hypothetical protein G ( 293) 491 117.5 5.5e-23  
gi|140887583|gb|ECO18446.1| hypothetical protein G ( 242) 490 117.2 5.6e-23  
gi|68346293|gb|AAY93899.1| prephenate dehydrogenas ( 447) 493 118.0 5.7e-23  
gi|134804078|gb|EBD54578.1| hypothetical protein G ( 211) 489 116.9 5.8e-23  
gi|136393960|gb|EBN87219.1| hypothetical protein G ( 312) 491 117.5 5.8e-23  
gi|139998431|gb|ECT91715.1| hypothetical protein G ( 271) 490 117.2 6.1e-23  
gi|136262507|gb|EBM97543.1| hypothetical protein G ( 237) 489 116.9 6.4e-23  
gi|141955051|gb|ECU42041.1| hypothetical protein G ( 165) 487 116.4 6.5e-23  
gi|139052693|gb|ECN08205.1| hypothetical protein G ( 298) 490 117.2 6.5e-23  
gi|140838036|gb|ECN85960.1| hypothetical protein G ( 292) 489 117.0 7.5e-23  
gi|138419111|gb|EBZ53968.1| hypothetical protein G ( 247) 488 116.7 7.6e-23  
gi|141898445|gb|ECU02305.1| hypothetical protein G ( 256) 488 116.8 7.8e-23  
gi|143287006|gb|EDE12905.1| hypothetical protein G ( 145) 485 115.9 8e-23  
gi|136637899|gb|EBP43656.1| hypothetical protein G ( 273) 488 116.8 8.2e-23  
gi|134902958|gb|EBB19018.1| hypothetical protein G ( 336) 489 117.1 8.3e-23  
gi|257046992|gb|ACV36180.1| 3-phosphoshikimate 1-c ( 655) 492 117.9 8.9e-23  
gi|143069597|gb|EDC56159.1| hypothetical protein G ( 303) 488 116.8 8.9e-23  
gi|137129035|gb|EBB36550.1| hypothetical protein G ( 214) 486 116.3 9.2e-23  
gi|135579519|gb|EBI55375.1| hypothetical protein G ( 329) 488 116.8 9.5e-23  
gi|136994338|gb|EBR61028.1| hypothetical protein G ( 280) 487 116.6 9.8e-23  
gi|141927329|gb|ECU22470.1| hypothetical protein G ( 289) 487 116.6 1e-22  
gi|139717854|gb|ECG99181.1| hypothetical protein G ( 231) 485 116.1 1.1e-22  
gi|140066777|gb|ECJ32528.1| hypothetical protein G ( 282) 486 116.4 1.1e-22  
gi|139910655|gb|ECI30758.1| hypothetical protein G ( 235) 485 116.1 1.2e-22  
gi|140973901|gb|ECQ77339.1| hypothetical protein G ( 290) 486 116.4 1.2e-22  
gi|140297850|gb|ECK81100.1| hypothetical protein G ( 260) 485 116.1 1.2e-22  
gi|138602658|gb|ECA73649.1| hypothetical protein G ( 273) 485 116.1 1.3e-22  
gi|137717949|gb|EBV62551.1| hypothetical protein G ( 230) 484 115.9 1.3e-22  
gi|139026437|gb|ECC89644.1| hypothetical protein G ( 293) 485 116.1 1.4e-22  
gi|140786534|gb|ECN50058.1| hypothetical protein G ( 295) 485 116.2 1.4e-22  
gi|137195758|gb|EBS73722.1| hypothetical protein G ( 210) 483 115.6 1.4e-22  
gi|135680007|gb|EBJ17574.1| hypothetical protein G ( 311) 485 116.2 1.4e-22  
gi|161165740|emb|CAN97045.1| 3-phosphoshikimate 1- ( 461) 487 116.7 1.4e-22  
gi|141989319|gb|ECU65887.1| hypothetical protein G ( 264) 484 115.9 1.5e-22  
gi|137804261|gb|EBW09125.1| hypothetical protein G ( 273) 484 115.9 1.5e-22  
gi|136078872|gb|EBL76774.1| hypothetical protein G ( 199) 482 115.4 1.6e-22  
gi|139918049|gb|ECI36047.1| hypothetical protein G ( 260) 483 115.7 1.7e-22  
gi|142266198|gb|ECW83397.1| hypothetical protein G ( 325) 484 116.0 1.7e-22  
gi|138607890|gb|ECA77386.1| hypothetical protein G ( 181) 481 115.1 1.7e-22  
gi|140442914|gb|ECL78527.1| hypothetical protein G ( 242) 482 115.4 1.9e-22  
gi|138159519|gb|EBY09502.1| hypothetical protein G ( 205) 481 115.2 1.9e-22  
gi|136332752|gb|EBN45289.1| hypothetical protein G ( 209) 481 115.2 1.9e-22  
gi|143136067|gb|ECC28514.1| hypothetical protein G ( 313) 483 115.7 1.9e-22  
gi|139699914|gb|ECG86528.1| hypothetical protein G ( 265) 482 115.5 2e-22  
gi|139820926|gb|ECH69481.1| hypothetical protein G ( 223) 481 115.2 2e-22  
gi|137214394|gb|EBS84325.1| hypothetical protein G ( 291) 482 115.5 2.1e-22  
gi|135162903|gb|EBF89585.1| hypothetical protein G ( 163) 479 114.6 2.2e-22  
gi|134835060|gb|EBB74225.1| hypothetical protein G ( 295) 482 115.5 2.2e-22  
gi|138860211|gb|ECC18701.1| hypothetical protein G ( 259) 481 115.2 2.3e-22  
gi|142969511|gb|EDB87361.1| hypothetical protein G ( 266) 481 115.2 2.3e-22  
gi|137576321|gb|EBU84711.1| hypothetical protein G ( 292) 481 115.3 2.5e-22  
gi|141073413|gb|ECP43912.1| hypothetical protein G ( 300) 481 115.3 2.5e-22

gi 134635372 gb EBC52105.1	hypothetical protein G ( 205)	479	114.7	2.6e-22
gi 140438418 gb ECL75365.1	hypothetical protein G ( 182)	478	114.5	2.7e-22
gi 135799699 gb EBJ92019.1	hypothetical protein G ( 271)	480	115.0	2.7e-22
gi 140053257 gb ECJ25908.1	hypothetical protein G ( 225)	479	114.8	2.8e-22
gi 141713106 gb ECS88901.1	hypothetical protein G ( 346)	481	115.3	2.8e-22
gi 138586257 gb ECA62165.1	hypothetical protein G ( 288)	480	115.1	2.9e-22
gi 137030492 gb EBR81435.1	hypothetical protein G ( 171)	477	114.2	3e-22
gi 143098825 gb EDC80571.1	hypothetical protein G ( 309)	480	115.1	3e-22
gi 142218889 gb ECW48454.1	hypothetical protein G ( 223)	478	114.5	3.2e-22
gi 138251176 gb EBY57356.1	hypothetical protein G ( 233)	478	114.5	3.3e-22
gi 137014132 gb EBR72273.1	hypothetical protein G ( 244)	478	114.6	3.4e-22
gi 140523504 gb ECM09797.1	hypothetical protein G ( 245)	478	114.6	3.4e-22
gi 171193469 gb ACB44430.1	3-phosphoshikimate 1-c ( 442)	481	115.4	3.4e-22
gi 138030261 gb EBX36003.1	hypothetical protein G ( 263)	478	114.6	3.6e-22
gi 137719292 gb EBV63298.1	hypothetical protein G ( 180)	476	114.0	3.7e-22
gi 135982409 gb EBL13409.1	hypothetical protein G ( 309)	478	114.6	4.1e-22
gi 226458362 gb EEH55659.1	predicted protein [Mic ( 459)	480	115.2	4.1e-22
gi 142962907 gb EDB82861.1	hypothetical protein G ( 150)	474	113.5	4.3e-22
gi 168992012 gb ACA39552.1	3-phosphoshikimate 1-c ( 187)	475	113.8	4.4e-22
gi 139532469 gb ECF71502.1	hypothetical protein G ( 279)	477	114.4	4.4e-22
gi 141790910 gb ECT30914.1	hypothetical protein G ( 168)	474	113.6	4.7e-22
gi 137963940 gb EBX00422.1	hypothetical protein G ( 305)	477	114.4	4.7e-22
gi 141940271 gb ECU31475.1	hypothetical protein G ( 213)	475	113.9	4.8e-22
gi 139409852 gb ECF61909.1	hypothetical protein G ( 260)	476	114.1	4.9e-22
gi 142711155 gb EDA02482.1	hypothetical protein G ( 226)	475	113.9	5.1e-22
gi 138319895 gb EBY95130.1	hypothetical protein G ( 282)	476	114.2	5.2e-22
gi 134629694 gb EBC48725.1	hypothetical protein G ( 288)	476	114.2	5.3e-22
gi 17427919 emb CAD14609.1	probable 3-phosphoshik ( 436)	478	114.8	5.4e-22
gi 140656566 gb ECM60669.1	hypothetical protein G ( 250)	475	113.9	5.5e-22
gi 142100413 gb ECV59278.1	hypothetical protein G ( 308)	476	114.2	5.5e-22
gi 141687467 gb ECS81421.1	hypothetical protein G ( 262)	475	113.9	5.7e-22
gi 138135881 gb EBX94279.1	hypothetical protein G ( 268)	475	113.9	5.8e-22
gi 134908887 gb EBE22927.1	hypothetical protein G ( 308)	475	114.0	6.4e-22
gi 137259371 gb EBT09655.1	hypothetical protein G ( 221)	473	113.4	6.7e-22
gi 141106614 gb ECP67175.1	hypothetical protein G ( 270)	474	113.7	6.8e-22
gi 137243061 gb EBT00335.1	hypothetical protein G ( 245)	473	113.5	7.3e-22
gi 139122965 gb ECD57240.1	hypothetical protein G ( 306)	474	113.8	7.4e-22
gi 141407006 gb ECR61199.1	hypothetical protein G ( 262)	473	113.5	7.7e-22
gi 157321714 gb ABV40811.1	3-phosphoshikimate 1-c ( 428)	475	114.1	8.3e-22
gi 139371007 gb ECE65615.1	hypothetical protein G ( 304)	473	113.5	8.6e-22
gi 142507077 gb ECY57668.1	hypothetical protein G ( 305)	473	113.5	8.6e-22
gi 136431133 gb EB011861.1	hypothetical protein G ( 262)	472	113.3	8.9e-22
gi 137293075 gb EBT28512.1	hypothetical protein G ( 272)	472	113.3	9.2e-22
gi 139621692 gb ECG31968.1	hypothetical protein G ( 272)	472	113.3	9.2e-22
gi 134884660 gb EBE06854.1	hypothetical protein G ( 276)	472	113.3	9.3e-22
gi 137000978 gb EBR64813.1	hypothetical protein G ( 230)	471	113.0	9.4e-22
gi 140795522 gb ECN56376.1	hypothetical protein G ( 280)	472	113.3	9.4e-22
gi 141302923 gb ECQ94687.1	hypothetical protein G ( 281)	472	113.3	9.4e-22
gi 138657864 gb ECB10524.1	hypothetical protein G ( 283)	472	113.3	9.5e-22
gi 143672084 gb EDG618795.1	hypothetical protein G ( 419)	474	113.9	9.5e-22
gi 136993434 gb EBR60509.1	hypothetical protein G ( 284)	472	113.3	9.5e-22
gi 139023907 gb ECC87913.1	hypothetical protein G ( 166)	469	112.5	9.9e-22
gi 134407858 gb EBB17421.1	hypothetical protein G ( 314)	472	113.3	1e-21
gi 136268536 gb EBN01805.1	hypothetical protein G ( 315)	472	113.3	1e-21
gi 143053685 gb EDC47513.1	hypothetical protein G ( 266)	471	113.1	1.1e-21
gi 140469715 gb ECL93060.1	hypothetical protein G ( 231)	470	112.8	1.1e-21
gi 138080088 gb EBX63139.1	hypothetical protein G ( 281)	471	113.1	1.1e-21
gi 139905975 gb ECI27361.1	hypothetical protein G ( 248)	470	112.8	1.2e-21
gi 138620417 gb ECA85940.1	hypothetical protein G ( 312)	471	113.1	1.2e-21
gi 134464839 gb EBB50457.1	hypothetical protein G ( 177)	468	112.3	1.2e-21
gi 148551350 gb ABQ86478.1	EPSP synthase (3-phosp ( 438)	472	113.5	1.3e-21
gi 134728264 gb EBD06387.1	hypothetical protein G ( 253)	469	112.6	1.4e-21

gi 137182909 gb EBS66675.1	hypothetical protein G ( 260)	469	112.6	1.4e-21
gi 139654368 gb ECG54707.1	hypothetical protein G ( 147)	466	111.8	1.4e-21
gi 137424329 gb EBU02368.1	hypothetical protein G ( 278)	469	112.6	1.5e-21
gi 140787776 gb ECN50901.1	hypothetical protein G ( 287)	469	112.7	1.5e-21
gi 206590606 emb CAQ37568.1	3-phosphoshikimate 1- ( 435)	471	113.2	1.5e-21
gi 142577198 gb EC207785.1	hypothetical protein G ( 482)	471	113.3	1.7e-21
gi 136917542 gb EBR18692.1	hypothetical protein G ( 289)	468	112.4	1.8e-21
gi 134775968 gb EBB35272.1	hypothetical protein G ( 297)	468	112.5	1.8e-21
gi 140676899 gb ECM75091.1	hypothetical protein G ( 168)	465	111.6	1.8e-21
gi 139227125 gb ECE28521.1	hypothetical protein G ( 268)	467	112.2	1.9e-21
gi 138030771 gb EBX36292.1	hypothetical protein G ( 275)	467	112.2	2e-21
gi 142058446 gb ECV24053.1	hypothetical protein G ( 439)	469	112.8	2.1e-21
gi 138150939 gb EBY03425.1	hypothetical protein G ( 202)	465	111.7	2.1e-21
gi 138014097 gb EBX27907.1	hypothetical protein G ( 256)	466	112.0	2.2e-21
gi 138600955 gb ABD46423.1	putative 3-phosphoshiki ( 435)	468	112.6	2.2e-21
gi 137260499 gb EBT10295.1	hypothetical protein G ( 274)	466	112.0	2.3e-21
gi 139682895 gb ECG74820.1	hypothetical protein G ( 274)	466	112.0	2.3e-21
gi 139974702 gb ECM75189.1	hypothetical protein G ( 157)	463	111.2	2.3e-21
gi 136408371 gb EBN97136.1	hypothetical protein G ( 286)	466	112.0	2.4e-21
gi 136231923 gb EBM76866.1	hypothetical protein G ( 197)	464	111.4	2.4e-21
gi 88600955 gb ABD46423.1	putative 3-phosphoshiki ( 435)	468	112.6	2.4e-21
gi 142249738 gb ECW1259.1	hypothetical protein G ( 439)	468	112.6	2.4e-21
gi 138140452 gb EBX96828.1	hypothetical protein G ( 302)	466	112.0	2.5e-21
gi 1406214096 gb EBP98532.1	hypothetical protein G ( 256)	465	111.8	2.5e-21
gi 141056695 gb ECP33646.1	hypothetical protein G ( 179)	463	111.2	2.6e-21
gi 136146607 gb EBM20258.1	hypothetical protein G ( 190)	463	111.2	2.7e-21
gi 14822654 gb ECT49103.1	hypothetical protein G ( 293)	465	111.8	2.8e-21
gi 206595339 emb CAQ62266.1	3-phosphoshikimate 1- ( 435)	467	112.4	2.8e-21
gi 144582815 gb ABP00889.1	predicted protein [Ost ( 436)	467	112.4	2.8e-21
gi 135910595 gb EBK64732.1	hypothetical protein G ( 305)	465	111.8	2.9e-21
gi 136631260 gb EBP39693.1	hypothetical protein G ( 209)	463	111.2	2.9e-21
gi 141964658 gb ECU48472.1	hypothetical protein G ( 268)	464	111.5	3e-21
gi 139579623 gb ECG30907.1	hypothetical protein G ( 281)	464	111.6	3.2e-21
gi 141317497 gb ECR00630.1	hypothetical protein G ( 283)	464	111.6	3.2e-21
gi 135455852 gb EBH75607.1	hypothetical protein G ( 235)	463	111.3	3.2e-21
gi 1146414 gb AAA85091.1	3-phosphoshikimate 1-car ( 432)	466	112.1	3.2e-21
gi 141939310 gb ECU30795.1	hypothetical protein G ( 251)	463	111.3	3.4e-21
gi 140510912 gb ECM06926.1	hypothetical protein G ( 208)	462	111.0	3.4e-21
gi 141809387 gb ECT39920.1	hypothetical protein G ( 270)	463	111.3	3.6e-21
gi 136694505 gb EBP79130.1	hypothetical protein G ( 183)	461	110.8	3.6e-21
gi 134629207 gb EBC48440.1	hypothetical protein G ( 272)	463	111.3	3.6e-21
gi 188029313 emb CAO97190.1	3-phosphoshikimate 1- ( 428)	465	111.9	3.8e-21
gi 21555078 gb AAM63771.1	5-enolpyruvylshikimate- ( 521)	466	112.2	3.8e-21
gi 197053768 gb ACH25466.1	Sequence 18 from paten ( 521)	466	112.2	3.8e-21
gi 134464347 gb EBB50178.1	hypothetical protein G ( 197)	461	110.8	3.8e-21
gi 134557217 gb EBC05219.1	hypothetical protein G ( 296)	463	111.4	3.8e-21
gi 136996308 gb EBR62166.1	hypothetical protein G ( 181)	460	110.5	4.1e-21
gi 140328230 gb ECL01542.1	hypothetical protein G ( 182)	460	110.6	4.1e-21
gi 137176800 gb EBS63254.1	hypothetical protein G ( 271)	462	111.1	4.1e-21
gi 140032122 gb ECJ13099.1	hypothetical protein G ( 194)	460	110.6	4.3e-21
gi 139748675 gb ECH19163.1	hypothetical protein G ( 288)	462	111.1	4.3e-21
gi 135851336 gb EBK24502.1	hypothetical protein G ( 161)	459	110.0	4.4e-21
gi 14532882 gb AAK64123.1	putative 5-enolpyruvyls ( 521)	465	112.3	4.4e-21
gi 11094810 gb AAG29739.1	AC084414_7 5-enolpyruvyl ( 521)	465	112.0	4.4e-21
gi 13430624 gb AAK25934.1	AF360224_1 putative 5-en ( 521)	465	112.0	4.4e-21
gi 12321113 gb AAG50661.1	AC084242_5 5-enolpyruvyl ( 521)	465	112.0	4.4e-21
gi 135343553 gb EBH00233.1	hypothetical protein G ( 300)	462	111.1	4.5e-21
gi 141186881 gb ECQ23250.1	hypothetical protein G ( 206)	460	110.6	4.5e-21
gi 135208489 gb EBG18076.1	hypothetical protein G ( 206)	460	110.6	4.5e-21
gi 137091767 gb EBS15725.1	hypothetical protein G ( 214)	460	110.6	4.7e-21
gi 139568311 gb ECF96248.1	hypothetical protein G ( 264)	461	110.9	4.7e-21
gi 138340796 gb EBZ07169.1	hypothetical protein G ( 125)	457	109.8	4.8e-21

gi|141460123|gb|ECR97455.1| hypothetical protein G ( 278) 461 110.9 4.9e-21  
gi|135951103|gb|EBK92257.1| hypothetical protein G ( 282) 461 110.9 5e-21  
gi|140995635|gb|ECO92669.1| hypothetical protein G ( 157) 458 110.1 5e-21  
gi|140503041|gb|ECM04119.1| hypothetical protein G ( 288) 461 110.9 5.1e-21  
gi|91712081|gb|ABE52008.1| 3-phosphoshikimate 1-ca ( 427) 463 111.5 5.1e-21  
gi|134745606|gb|EBD16293.1| hypothetical protein G ( 197) 459 110.4 5.1e-21  
gi|137278631|gb|EBT20527.1| hypothetical protein G ( 136) 457 109.8 5.2e-21  
gi|141018576|gb|ECP08567.1| hypothetical protein G ( 248) 460 110.7 5.2e-21  
gi|139397227|gb|ECE83265.1| hypothetical protein G ( 305) 461 110.9 5.3e-21  
gi|142982566|gb|EDB96566.1| hypothetical protein G ( 310) 461 110.9 5.3e-21  
gi|140111307|gb|ECJ61391.1| hypothetical protein G ( 265) 460 110.7 5.5e-21  
gi|137916341|gb|EBW73633.1| hypothetical protein G ( 284) 460 110.7 5.8e-21  
gi|139806318|gb|ECH59432.1| hypothetical protein G ( 292) 460 110.7 5.9e-21  
gi|138161925|gb|EBY11146.1| hypothetical protein G ( 202) 458 110.1 6e-21  
gi|135811093|gb|EBJ99203.1| hypothetical protein G ( 247) 459 110.4 6.1e-21  
gi|143480854|gb|EDF21802.1| hypothetical protein G ( 250) 459 110.4 6.1e-21  
gi|135346262|gb|EBH02057.1| hypothetical protein G ( 304) 460 110.7 6.1e-21  
gi|134663080|gb|EBC68366.1| hypothetical protein G ( 140) 456 109.6 6.2e-21  
gi|142930364|gb|EDB59786.1| hypothetical protein G ( 262) 459 110.5 6.3e-21  
gi|139455460|gb|ECF22083.1| hypothetical protein G ( 268) 459 110.5 6.5e-21  
gi|139075902|gb|ECD24383.1| hypothetical protein G ( 281) 459 110.5 6.7e-21  
gi|141222898|gb|ECQ46942.1| hypothetical protein G ( 237) 458 110.2 6.8e-21  
gi|142575612|gb|ECZ06674.1| hypothetical protein G ( 196) 457 109.9 6.9e-21  
gi|168827279|gb|ACA32650.1| 3-phosphoshikimate 1-c ( 432) 461 111.1 6.9e-21  
gi|139188816|gb|ECE01784.1| hypothetical protein G ( 245) 458 110.2 7e-21  
gi|135804817|gb|EBJ95270.1| hypothetical protein G ( 248) 458 110.2 7.1e-21  
gi|138660385|gb|ECB12258.1| hypothetical protein G ( 221) 457 110.0 7.1e-21  
gi|137958827|gb|EBW97523.1| hypothetical protein G ( 284) 458 110.3 7.9e-21  
gi|137942191|gb|EBW88167.1| hypothetical protein G ( 200) 456 109.7 8.1e-21  
gi|141314678|gb|ECQ98633.1| hypothetical protein G ( 296) 458 110.3 8.1e-21  
gi|139068047|gb|ECD19190.1| hypothetical protein G ( 260) 457 110.0 8.5e-21  
gi|137004591|gb|EBR66887.1| hypothetical protein G ( 281) 457 110.0 9.1e-21  
gi|135477800|gb|EBH90218.1| hypothetical protein G ( 282) 457 110.0 9.1e-21  
gi|111979101|gb|ABH83338.1| Sequence 7 from patent ( 521) 460 110.9 9.3e-21  
gi|140658132|gb|ECM61746.1| hypothetical protein G ( 198) 455 109.5 9.4e-21  
gi|143251258|gb|EDD90204.1| hypothetical protein G ( 295) 457 110.1 9.4e-21  
gi|139916408|gb|ECI34864.1| hypothetical protein G ( 256) 456 109.8 9.8e-21  
gi|139076956|gb|ECD25149.1| hypothetical protein G ( 278) 456 109.8 1e-20  
gi|140772822|gb|ECN40516.1| hypothetical protein G ( 189) 454 109.3 1e-20  
gi|137755595|gb|EBV82450.1| hypothetical protein G ( 281) 456 109.8 1.1e-20  
gi|138685421|gb|ECB29910.1| hypothetical protein G ( 240) 455 109.6 1.1e-20  
gi|165876218|gb|ABY69266.1| 3-phosphoshikimate 1-c ( 432) 458 110.4 1.1e-20  
gi|189915141|gb|ACE61393.1| 3-phosphoshikimate 1-c ( 432) 458 110.4 1.1e-20  
gi|89345096|gb|ABD69299.1| 3-phosphoshikimate 1-ca ( 669) 460 111.0 1.1e-20  
gi|197698457|gb|EDY45390.1| PlmJK [Streptomyces sp ( 995) 462 111.5 1.1e-20  
gi|136650971|gb|EBP51388.1| hypothetical protein G ( 266) 455 109.6 1.2e-20  
gi|141123028|gb|ECP78588.1| hypothetical protein G ( 273) 455 109.6 1.2e-20  
gi|137030922|gb|EBR81678.1| hypothetical protein G ( 276) 455 109.6 1.2e-20  
gi|135507167|gb|EBI09057.1| hypothetical protein G ( 192) 453 109.0 1.2e-20  
gi|136830333|gb|EBQ69060.1| hypothetical protein G ( 158) 452 108.8 1.2e-20  
gi|281108574|gb|ADA52832.1| Sequence 7 from patent ( 521) 458 110.5 1.3e-20  
gi|112822801|gb|ABT24890.1| 3-phosphoshikimate 1-c ( 432) 457 110.2 1.3e-20  
gi|141749797|gb|ECT09608.1| hypothetical protein G ( 135) 451 108.5 1.3e-20  
gi|145047091|gb|ABP33718.1| 3-phosphoshikimate 1-c ( 442) 457 110.2 1.3e-20  
gi|140653578|gb|ECM58585.1| hypothetical protein G ( 171) 452 108.8 1.3e-20  
gi|139760239|gb|ECH27080.1| hypothetical protein G ( 253) 454 109.4 1.3e-20  
gi|137426729|gb|EBU03730.1| hypothetical protein G ( 261) 454 109.4 1.3e-20  
gi|138139693|gb|EBX96411.1| hypothetical protein G ( 264) 454 109.4 1.4e-20  
gi|126096971|gb|ABN73799.1| 3-phosphoshikimate 1-c ( 432) 456 110.0 1.5e-20  
gi|134886066|gb|EBE07800.1| hypothetical protein G ( 299) 454 109.4 1.5e-20  
gi|138691138|gb|ECB33980.1| hypothetical protein G ( 137) 450 108.3 1.5e-20  
gi|139051627|gb|ECD07438.1| hypothetical protein G ( 276) 453 109.2 1.6e-20

gi|144034657|gb|EDI39331.1| hypothetical protein G ( 193) 451 108.6 1.7e-20  
gi|137863910|gb|EBW43621.1| hypothetical protein G ( 253) 452 108.9 1.8e-20  
gi|142318022|gb|ECC21690.1| hypothetical protein G ( 257) 452 108.9 1.8e-20  
gi|139643546|gb|ECG46969.1| hypothetical protein G ( 125) 448 107.8 1.9e-20  
gi|137040706|gb|EBR87177.1| hypothetical protein G ( 231) 451 108.7 1.9e-20  
gi|137821808|gb|EBW19362.1| hypothetical protein G ( 191) 450 108.4 1.9e-20  
gi|136801771|gb|EBQ50060.1| hypothetical protein G ( 294) 452 109.0 2e-20  
gi|141894283|gb|ECT99317.1| hypothetical protein G ( 202) 450 108.4 2e-20  
gi|139125728|gb|ECD58839.1| hypothetical protein G ( 167) 449 108.1 2e-20  
gi|140227605|gb|ECK39253.1| hypothetical protein G ( 307) 452 109.0 2.1e-20  
gi|269848851|gb|ACZ49495.1| 3-phosphoshikimate 1-c ( 458) 454 109.5 2.1e-20  
gi|134628386|gb|EBC47961.1| hypothetical protein G ( 265) 451 108.7 2.1e-20  
gi|139408555|gb|ECE90996.1| hypothetical protein G ( 275) 451 108.7 2.2e-20  
gi|134852959|gb|EBD85720.1| hypothetical protein G ( 307) 451 108.8 2.4e-20  
gi|137060732|gb|EBR98539.1| hypothetical protein G ( 215) 449 108.2 2.5e-20  
gi|136936196|gb|EBR28143.1| hypothetical protein G ( 178) 448 107.9 2.5e-20  
gi|134872021|gb|EBD98471.1| hypothetical protein G ( 270) 450 108.5 2.5e-20  
gi|140036373|gb|ECU16119.1| hypothetical protein G ( 279) 450 108.5 2.6e-20  
gi|138956668|gb|ECC58149.1| hypothetical protein G ( 287) 450 108.5 2.6e-20  
gi|139642797|gb|ECG46468.1| hypothetical protein G ( 249) 449 108.3 2.8e-20  
gi|134872282|gb|EBN98642.1| hypothetical protein G ( 269) 449 108.3 2.9e-20  
gi|140933952|gb|ECO49463.1| hypothetical protein G ( 279) 449 108.3 3e-20  
gi|137360628|gb|EBT66371.1| hypothetical protein G ( 286) 449 108.3 3.1e-20  
gi|114307433|gb|EBR58676.1| 3-phosphoshikimate 1-c ( 433) 451 108.9 3.1e-20  
gi|139739230|gb|ECH13944.1| hypothetical protein G ( 261) 448 108.1 3.3e-20  
gi|139354446|gb|ECE60428.1| hypothetical protein G ( 325) 449 108.3 3.4e-20  
gi|136808175|gb|EBQ54293.1| hypothetical protein G ( 275) 448 108.1 3.5e-20  
gi|137557298|gb|EBU73902.1| hypothetical protein G ( 277) 448 108.1 3.5e-20  
gi|138941425|gb|ECC51619.1| hypothetical protein G ( 231) 447 107.8 3.5e-20  
gi|139152261|gb|ECB76059.1| hypothetical protein G ( 285) 448 108.1 3.6e-20  
gi|144114046|gb|EDI96547.1| hypothetical protein G ( 200) 446 107.5 3.7e-20  
gi|139991868|gb|ECI87059.1| hypothetical protein G ( 248) 447 107.8 3.7e-20  
gi|139750669|gb|ECH20532.1| hypothetical protein G ( 262) 447 107.8 3.9e-20  
gi|139188209|gb|ECE01339.1| hypothetical protein G ( 187) 445 107.3 4e-20  
gi|140114936|gb|ECJ63978.1| hypothetical protein G ( 286) 447 107.9 4.2e-20  
gi|140380447|gb|ECL38008.1| hypothetical protein G ( 292) 447 107.9 4.2e-20  
gi|137264374|gb|EBT12455.1| hypothetical protein G ( 245) 446 107.6 4.3e-20  
gi|137450070|gb|EBU16913.1| hypothetical protein G ( 251) 446 107.6 4.4e-20  
gi|138907009|gb|ECC37100.1| hypothetical protein G ( 177) 444 107.1 4.5e-20  
gi|138677357|gb|ECB24420.1| hypothetical protein G ( 279) 446 107.6 4.7e-20  
gi|1890139|gb|AAB86439.1| AroA [Mannheimia haemoly ( 432) 448 108.2 4.9e-20  
gi|135666691|gb|EBJ09399.1| hypothetical protein G ( 164) 443 106.8 4.9e-20  
gi|134673469|gb|EBC74420.1| hypothetical protein G ( 181) 443 106.8 5.3e-20  
gi|138636499|gb|ECA95841.1| hypothetical protein G ( 271) 445 107.4 5.4e-20  
gi|141113089|gb|ECP71687.1| hypothetical protein G ( 155) 442 106.6 5.5e-20  
gi|140648727|gb|ECM55237.1| hypothetical protein G ( 292) 445 107.4 5.7e-20  
gi|139498755|gb|ECF48542.1| hypothetical protein G ( 209) 443 106.9 6e-20  
gi|135640463|gb|EBI93148.1| hypothetical protein G ( 310) 445 107.5 6e-20  
gi|139454391|gb|ECF21324.1| hypothetical protein G ( 258) 444 107.2 6e-20  
gi|142005931|gb|ECU77145.1| hypothetical protein G ( 271) 444 107.2 6.3e-20  
gi|140303989|gb|ECK84652.1| hypothetical protein G ( 282) 444 107.2 6.5e-20  
gi|224964543|emb|CAX56053.1| 3-phosphoshikimate 1- ( 428) 446 107.8 6.6e-20  
gi|138335879|gb|EBZ03751.1| hypothetical protein G ( 198) 442 106.7 6.6e-20  
gi|194682094|emb|CAR41667.1| 3-phosphoshikimate 1- ( 428) 445 107.6 7.7e-20  
gi|371992501|dbj|BAC95082.1| 3-phosphoshikimate 1-c ( 428) 445 107.6 7.7e-20  
gi|153093746|gb|EDN74689.1| 3-phosphoshikimate 1-c ( 432) 445 107.6 7.7e-20  
gi|139018177|gb|ECC83955.1| hypothetical protein G ( 200) 441 106.4 7.8e-20  
gi|143035975|gb|EDC34531.1| hypothetical protein G ( 169) 440 106.2 8e-20  
gi|138430761|gb|EBZ62251.1| hypothetical protein G ( 255) 442 106.7 8.1e-20  
gi|138542053|gb|ECA31262.1| hypothetical protein G ( 262) 442 106.8 8.3e-20  
gi|141128430|gb|ECP82499.1| hypothetical protein G ( 284) 442 106.8 8.8e-20  
gi|139888767|gb|ECI15397.1| hypothetical protein G ( 161) 439 105.9 8.9e-20

gi|139450344|gb|ECF18445.1| hypothetical protein G ( 139) 438 105.7 9.2e-20  
gi|141367144|gb|ECR33007.1| hypothetical protein G ( 258) 441 106.5 9.5e-20  
gi|140940783|gb|ECO54306.1| hypothetical protein G ( 213) 440 106.2 9.5e-20  
gi|146453322|gb|EDK47578.1| pentafunctional AROM p (1551) 450 109.1 9.8e-20  
gi|138153319|gb|EBY05062.1| hypothetical protein G ( 271) 441 106.5 9.9e-20  
gi|138972884|gb|ECC65105.1| hypothetical protein G ( 184) 439 106.0 9.9e-20  
gi|138520206|gb|ECA15974.1| hypothetical protein G ( 277) 441 106.6 1e-19  
gi|140334517|gb|ECL05827.1| hypothetical protein G ( 158) 438 105.7 1e-19  
gi|28806007|dbj|BAC59283.1| 3-phosphoshikimate 1-c ( 426) 443 107.1 1e-19  
gi|56616713|gb|AAW04888.1| Sequence 2 from patent ( 434) 443 107.1 1e-19  
gi|3940967|gb|AAC81966.1|I96497 Sequence 4 from pa ( 434) 443 107.1 1e-19  
gi|3008973|gb|AAC10319.1|I72834 Sequence 4 from pa ( 434) 443 107.1 1e-19  
gi|1830395|gb|AAB44849.1| Sequence 4 from patent U ( 434) 443 107.1 1e-19  
gi|451485|gb|AAA21529.1| 5-enolpyruvylshikimate 3- ( 434) 443 107.1 1e-19  
gi|3015354|gb|AAC12234.1|I79200 Sequence 4 from pa ( 434) 443 107.1 1e-19  
gi|5972478|gb|AAE12369.1| Sequence 4 from patent U ( 434) 443 107.1 1e-19  
gi|138911546|gb|ECC39123.1| hypothetical protein G ( 256) 440 106.3 1.1e-19  
gi|135561916|gb|EBT44110.1| hypothetical protein G ( 270) 440 106.3 1.1e-19  
gi|143482336|gb|EDF22566.1| hypothetical protein G ( 275) 440 106.3 1.2e-19  
gi|137492420|gb|EBU39211.1| hypothetical protein G ( 276) 440 106.3 1.2e-19  
gi|5957579|gb|AAE08253.1| Sequence 61 from patent ( 432) 442 106.9 1.2e-19  
gi|144974758|gb|ABP12469.1| Sequence 61 from paten ( 432) 442 106.9 1.2e-19  
gi|2485256|gb|AAB73392.1|I44481 Sequence 61 from p ( 432) 442 106.9 1.2e-19  
gi|148865|gb|AAA24943.1| enolpyruvylshikimatephosp ( 432) 442 106.9 1.2e-19  
gi|2484179|gb|AAB72315.1|I49208 Sequence 61 from p ( 432) 442 106.9 1.2e-19  
gi|40083323|gb|AAR41877.1| Sequence 5147 from pate ( 434) 442 106.9 1.2e-19  
gi|137275963|gb|EBT19032.1| hypothetical protein G ( 207) 438 105.8 1.3e-19  
gi|139639483|gb|ECG44207.1| hypothetical protein G ( 255) 439 106.1 1.3e-19  
gi|134932131|gb|EBE38383.1| hypothetical protein G ( 144) 436 105.2 1.3e-19  
gi|139606942|gb|ECG23026.1| hypothetical protein G ( 269) 439 106.1 1.3e-19  
gi|138968495|gb|ECC63182.1| hypothetical protein G ( 193) 437 105.6 1.4e-19  
gi|262338316|gb|ACY52111.1| 5-Enolpyruvylshikimate ( 426) 441 106.7 1.4e-19  
gi|138376657|gb|EBZ24565.1| hypothetical protein G ( 201) 437 105.6 1.4e-19  
gi|86569313|gb|ABD13122.1| 3-phosphoshikimate 1-ca ( 446) 441 106.7 1.4e-19  
gi|140196807|gb|ECK18245.1| hypothetical protein G ( 249) 438 105.9 1.5e-19  
gi|142004810|gb|ECU76380.1| hypothetical protein G ( 306) 439 106.1 1.5e-19  
gi|146453336|gb|EDK47592.1| pentafunctional AROM p (1551) 447 108.4 1.5e-19  
gi|141291390|gb|ECQ89981.1| hypothetical protein G ( 149) 435 105.0 1.5e-19  
gi|138538470|gb|ECA28678.1| hypothetical protein G ( 278) 438 105.9 1.6e-19  
gi|136617331|gb|EBP31472.1| hypothetical protein G ( 287) 438 105.9 1.6e-19  
gi|68058025|gb|AAH88278.1| 3-phosphoshikimate 1-ca ( 432) 440 106.5 1.6e-19  
gi|148716269|gb|ABQ98479.1| 3-phosphoshikimate 1-c ( 432) 440 106.5 1.6e-19  
gi|138337633|gb|EBZ04990.1| hypothetical protein G ( 241) 437 105.6 1.6e-19  
gi|143553814|gb|EDF60517.1| hypothetical protein G ( 295) 438 105.9 1.7e-19  
gi|136829494|gb|EBQ68499.1| hypothetical protein G ( 245) 437 105.6 1.7e-19  
gi|143161428|gb|EDD26007.1| hypothetical protein G ( 246) 437 105.6 1.7e-19  
gi|137262673|gb|EBT11508.1| hypothetical protein G ( 262) 437 105.7 1.8e-19  
gi|138671510|gb|ECB20197.1| hypothetical protein G ( 268) 437 105.7 1.8e-19  
gi|136327210|gb|EBN41578.1| hypothetical protein G ( 182) 435 105.1 1.8e-19  
gi|122088811|emb|CAL11617.1| 3-phosphoshikimate 1- ( 428) 439 106.3 1.9e-19  
gi|143040406|gb|EDC37793.1| hypothetical protein G ( 431) 439 106.3 1.9e-19  
gi|135614522|gb|EBT77073.1| hypothetical protein G ( 298) 437 105.7 1.9e-19  
gi|91696897|gb|ABE43726.1| 3-phosphoshikimate 1-ca ( 668) 441 106.8 2e-19  
gi|136650301|gb|EBP50987.1| hypothetical protein G ( 254) 436 105.4 2e-19  
gi|138659889|gb|ECB11911.1| hypothetical protein G ( 255) 436 105.4 2e-19  
gi|134544087|gb|EBB97350.1| hypothetical protein G ( 226) 435 105.2 2.1e-19  
gi|138065352|gb|EBX54709.1| hypothetical protein G ( 155) 433 104.6 2.1e-19  
gi|138382720|gb|EBZ28752.1| hypothetical protein G ( 193) 434 104.9 2.2e-19  
gi|134711099|gb|EBC96621.1| hypothetical protein G ( 236) 435 105.2 2.2e-19  
gi|148719663|gb|ABR00791.1| 3-phosphoshikimate 1-c ( 432) 438 106.0 2.2e-19  
gi|144146470|gb|EDJ20240.1| hypothetical protein G ( 163) 433 104.6 2.2e-19  
gi|140083489|gb|ECJ43144.1| hypothetical protein G ( 254) 435 105.2 2.3e-19

gi|139957031|gb|ECI62788.1| hypothetical protein G ( 175) 433 104.7 2.3e-19  
gi|142459623|gb|ECY23467.1| hypothetical protein G ( 144) 432 104.4 2.4e-19  
gi|142732836|gb|EDA18046.1| hypothetical protein G ( 259) 435 105.2 2.4e-19  
gi|141987601|gb|ECU64667.1| hypothetical protein G ( 219) 434 105.0 2.4e-19  
gi|141776609|gb|ECT23341.1| hypothetical protein G ( 185) 433 104.7 2.5e-19  
gi|142596803|gb|ECZ21432.1| hypothetical protein G ( 233) 434 105.0 2.5e-19  
gi|139777013|gb|ECH38737.1| hypothetical protein G ( 202) 433 104.7 2.6e-19  
gi|140304325|gb|ECK84887.1| hypothetical protein G ( 202) 433 104.7 2.6e-19  
gi|163259926|emb|CAP42227.1| 3-phosphoshikimate 1- ( 447) 437 105.8 2.7e-19  
gi|260220855|emb|CBA28839.1| 3-phosphoshikimate 1- ( 665) 439 106.4 2.7e-19  
gi|135717215|gb|EBJ40588.1| hypothetical protein G ( 251) 434 105.0 2.7e-19  
gi|163774845|gb|EDQ88471.1| predicted protein [Mon (1520) 443 107.5 2.8e-19  
gi|135377759|gb|EBH23190.1| hypothetical protein G ( 183) 432 104.5 2.8e-19  
gi|150837793|gb|ABR71769.1| Prephenate dehydrogena ( 748) 439 106.4 2.9e-19  
gi|84372628|gb|ABC56898.1| AroA [Methanosphaera st ( 426) 436 105.6 3e-19  
gi|145317961|gb|ABP60108.1| 3-phosphoshikimate 1-c ( 427) 436 105.6 3e-19  
gi|139865705|gb|ECI00817.1| hypothetical protein G ( 165) 431 104.2 3e-19  
gi|167323336|gb|ABZ59929.1| Sequence 10728 from pa ( 443) 436 105.6 3.1e-19  
gi|141512033|gb|ECS11482.1| hypothetical protein G ( 213) 432 104.5 3.2e-19  
gi|143715077|gb|EDG41632.1| hypothetical protein G ( 266) 433 104.8 3.2e-19  
gi|139552854|gb|ECF85640.1| hypothetical protein G ( 279) 433 104.8 3.4e-19  
gi|136829266|gb|EBQ68345.1| hypothetical protein G ( 280) 433 104.8 3.4e-19  
gi|139736854|gb|ECH12243.1| hypothetical protein G ( 283) 433 104.8 3.4e-19  
gi|134548028|gb|ABY83709.1| hypothetical protein G ( 176) 430 104.0 3.7e-19  
gi|139918409|gb|ECI36300.1| hypothetical protein G ( 262) 432 104.6 3.7e-19  
gi|139807487|gb|ECH60283.1| hypothetical protein G ( 153) 429 103.7 3.9e-19  
gi|138293084|gb|ABY83709.1| hypothetical protein G ( 230) 431 104.3 3.9e-19  
gi|218322092|emb|CAV18185.1| 3-phosphoshikimate 1- ( 426) 434 105.2 4e-19  
gi|156525456|gb|ABU70542.1| hypothetical protein V ( 426) 434 105.2 4e-19  
gi|141831946|gb|ECT55589.1| hypothetical protein G ( 161) 429 103.8 4e-19  
gi|140679635|gb|ECM77044.1| hypothetical protein G ( 136) 428 103.5 4.1e-19  
gi|138544070|gb|ECA32707.1| hypothetical protein G ( 298) 432 104.6 4.1e-19  
gi|138500221|gb|ECA03654.1| hypothetical protein G ( 256) 431 104.3 4.3e-19  
gi|142699709|gb|ECZ94320.1| hypothetical protein G ( 270) 431 104.4 4.4e-19  
gi|139955470|gb|ECI61700.1| hypothetical protein G ( 160) 428 103.5 4.7e-19  
gi|139552257|gb|ECF85214.1| hypothetical protein G ( 261) 430 104.1 5e-19  
gi|141373757|gb|ECR37676.1| hypothetical protein G ( 152) 427 103.3 5.2e-19  
gi|140907504|gb|ECO30880.1| hypothetical protein G ( 287) 430 104.2 5.4e-19  
gi|139711446|gb|ECG94601.1| hypothetical protein G ( 287) 430 104.2 5.4e-19  
gi|238546281|dbj|BAH62632.1| 3-phosphoshikimate 1- ( 427) 432 104.7 5.4e-19  
gi|139974316|gb|ECI74913.1| hypothetical protein G ( 240) 429 103.9 5.5e-19  
gi|143330502|gb|EDE37121.1| hypothetical protein G ( 437) 432 104.7 5.5e-19  
gi|138558546|gb|ECA42855.1| hypothetical protein G ( 253) 429 103.9 5.7e-19  
gi|139688350|gb|ECG78743.1| hypothetical protein G ( 217) 428 103.6 5.9e-19  
gi|141561788|gb|ECS38612.1| hypothetical protein G ( 264) 429 103.9 5.9e-19  
gi|138813436|gb|ECB97627.1| hypothetical protein G ( 189) 427 103.4 6.2e-19  
gi|141100594|gb|ECP62927.1| hypothetical protein G ( 230) 428 103.7 6.2e-19  
gi|135954943|gb|EBK94850.1| hypothetical protein G ( 237) 428 103.7 6.3e-19  
gi|186697859|gb|ACC88488.1| 3-phosphoshikimate 1-c ( 428) 431 104.5 6.3e-19  
gi|152961130|gb|ABS48591.1| 3-phosphoshikimate 1-c ( 428) 431 104.5 6.3e-19  
gi|119767369|gb|ABL99939.1| 3-phosphoshikimate 1-c ( 428) 431 104.5 6.3e-19  
gi|51589037|emb|CAH20655.1| 3-phosphoshikimate 1-c ( 428) 431 104.5 6.3e-19  
gi|169751429|gb|ACA68947.1| 3-phosphoshikimate 1-c ( 428) 431 104.5 6.3e-19  
gi|119536596|gb|ABL81213.1| 3-phosphoshikimate 1-c ( 430) 431 104.5 6.4e-19  
gi|138953444|gb|ECC56703.1| hypothetical protein G ( 250) 428 103.7 6.6e-19  
gi|120594658|gb|ABM38097.1| 3-phosphoshikimate 1-c ( 667) 433 105.1 6.6e-19  
gi|160365739|gb|ABX37352.1| 3-phosphoshikimate 1-c ( 675) 433 105.1 6.7e-19  
gi|139361912|gb|ECB62605.1| hypothetical protein G ( 257) 428 103.7 6.7e-19  
gi|143973335|gb|EDH96102.1| hypothetical protein G ( 179) 426 103.1 6.9e-19  
gi|141265029|gb|ECQ76941.1| hypothetical protein G ( 220) 427 103.4 6.9e-19  
gi|1161308|gb|AAB48057.1| 5-enolpyruvylshikimate-3 ( 424) 430 104.3 7.3e-19  
gi|262365043|gb|ACV61600.1| 3-phosphoshikimate 1-c ( 428) 430 104.3 7.4e-19



gi|115347149|emb|CAL20042.1| 3-phosphoshikimate 1- ( 428) 430 104.3 7.4e-19  
gi|74056467|gb|AAZ96907.1| 3-phosphoshikimate 1-ca ( 428) 430 104.3 7.4e-19  
gi|262361421|gb|ACY58142.1| 3-phosphoshikimate 1-c ( 428) 430 104.3 7.4e-19  
gi|108778591|gb|ABG12649.1| 3-phosphoshikimate 1-c ( 428) 430 104.3 7.4e-19  
gi|162351017|gb|ABX84965.1| 3-phosphoshikimate 1-c ( 428) 430 104.3 7.4e-19  
gi|108776396|gb|ABG18915.1| 3-phosphoshikimate 1-c ( 428) 430 104.3 7.4e-19  
gi|45435889|gb|AAS61446.1| 3-phosphoshikimate 1-ca ( 428) 430 104.3 7.4e-19  
gi|145211273|gb|ABP40680.1| 3-phosphoshikimate 1-c ( 428) 430 104.3 7.4e-19  
gi|270344367|gb|ACZ77132.1| 3-phosphoshikimate 1-c ( 429) 430 104.3 7.4e-19  
gi|21959675|gb|AAM86335.1|AE013881\_5 5-enolpyruvyl ( 443) 430 104.3 7.6e-19  
gi|135460444|gb|EBH78673.1| hypothetical protein G ( 178) 425 102.9 8e-19  
gi|54208741|gb|AAV31127.1| 5-enolpyruvylshikimate- ( 393) 429 104.1 8e-19  
gi|227009768|gb|ACP05980.1| 3-phosphoshikimate 1-c ( 426) 429 104.1 8.5e-19  
gi|150420105|gb|EDN12408.1| 3-phosphoshikimate 1-c ( 426) 429 104.1 8.5e-19  
gi|254844554|gb|EET22968.1| 3-phosphoshikimate 1-c ( 426) 429 104.1 8.5e-19  
gi|229370390|gb|ACQ60813.1| 5-Enolpyruvylshikimate ( 426) 429 104.1 8.5e-19  
gi|9656253|gb|AAF94882.1| 3-phosphoshikimate 1-car ( 426) 429 104.1 8.5e-19  
gi|150954342|gb|ABR76372.1| 3-phosphoshikimate 1-c ( 427) 429 104.1 8.6e-19  
gi|137209058|gb|EBS81261.1| hypothetical protein G ( 251) 426 103.3 8.9e-19  
gi|139435660|gb|ECF09838.1| hypothetical protein G ( 252) 426 103.3 9e-19  
gi|141883060|gb|ECT91476.1| hypothetical protein G ( 277) 426 103.3 9.6e-19  
gi|61396468|gb|AAX45933.1| 5-enolpyruvylshikimate ( 413) 428 103.9 9.7e-19  
gi|61396470|gb|AAX45934.1| 5-enolpyruvylshikimate ( 413) 428 103.9 9.7e-19  
gi|140674293|gb|ECM73199.1| hypothetical protein G ( 236) 425 103.0 9.9e-19  
gi|150424595|gb|EDN16531.1| 3-phosphoshikimate 1-c ( 426) 428 103.9 9.9e-19  
gi|140734280|gb|ECN15118.1| hypothetical protein G ( 178) 423 102.5 1.1e-18  
gi|142755731|gb|EDA34694.1| hypothetical protein G ( 266) 425 103.1 1.1e-18  
gi|140315571|gb|ECK92797.1| hypothetical protein G ( 268) 425 103.1 1.1e-18  
gi|139820786|gb|ECH69383.1| hypothetical protein G ( 200) 423 102.5 1.2e-18  
gi|164415617|gb|ABY53228.1| 3-phosphoshikimate 1-c ( 437) 427 103.7 1.2e-18  
gi|140837662|gb|ECN85695.1| hypothetical protein G ( 246) 424 102.8 1.2e-18  
gi|237624944|gb|ACR01634.1| 3-phosphoshikimate 1-c ( 653) 429 104.2 1.2e-18  
gi|141162624|gb|ECQ06320.1| hypothetical protein G ( 170) 422 102.3 1.2e-18  
gi|139074825|gb|ECD23598.1| hypothetical protein G ( 255) 424 102.8 1.2e-18  
gi|139887326|gb|ECI14382.1| hypothetical protein G ( 154) 421 102.0 1.3e-18  
gi|134900083|gb|EBE17116.1| hypothetical protein G ( 278) 424 102.9 1.3e-18  
gi|142831753|gb|EDA91373.1| hypothetical protein G ( 193) 422 102.3 1.3e-18  
gi|141595259|gb|ECS50998.1| hypothetical protein G ( 198) 422 102.3 1.4e-18  
gi|143027515|gb|EDC28300.1| hypothetical protein G ( 244) 423 102.6 1.4e-18  
gi|135499961|gb|EBI04446.1| hypothetical protein G ( 168) 421 102.0 1.4e-18  
gi|138842884|gb|ECC12012.1| hypothetical protein G ( 256) 423 102.6 1.4e-18  
gi|136101807|gb|EBL92381.1| hypothetical protein G ( 216) 422 102.3 1.5e-18  
gi|136760214|gb|EBQ22187.1| hypothetical protein G ( 277) 423 102.6 1.5e-18  
gi|125622565|gb|EAA50884.1| 3-phosphoshikimate 1-c ( 426) 425 103.2 1.6e-18  
gi|155523|gb|AAA27666.1| 5-enolpyruvylshikimate 3- ( 427) 425 103.2 1.6e-18  
gi|5957578|gb|AAE08252.1| Sequence 60 from patent ( 427) 425 103.2 1.6e-18  
gi|2485255|gb|AAB73391.1|I44480 Sequence 60 from p ( 427) 425 103.2 1.6e-18  
gi|2484178|gb|AAB72314.1|I49207 Sequence 60 from p ( 427) 425 103.2 1.6e-18  
gi|144974757|gb|ABP12468.1| Sequence 60 from paten ( 427) 425 103.2 1.6e-18  
gi|1574434|gb|AAC23237.1| 3-phosphoshikimate-1-car ( 432) 425 103.2 1.6e-18  
gi|136184571|gb|EBM45065.1| hypothetical protein G ( 198) 421 102.1 1.6e-18  
gi|140331781|gb|ECL03954.1| hypothetical protein G ( 136) 419 101.5 1.6e-18  
gi|111152891|emb|CAJ64639.1| 3-phosphoshikimate 1- ( 439) 425 103.2 1.6e-18  
gi|119669775|emb|CAL93688.1| 3-phosphoshikimate 1- ( 653) 427 103.8 1.6e-18  
gi|134889936|gb|EBE10360.1| hypothetical protein G ( 137) 419 101.5 1.6e-18  
gi|137684839|gb|EBV44401.1| hypothetical protein G ( 252) 422 102.4 1.6e-18  
gi|135914933|gb|EBK67763.1| hypothetical protein G ( 261) 422 102.4 1.7e-18  
gi|138646184|gb|ECB02436.1| hypothetical protein G ( 269) 422 102.4 1.7e-18  
gi|139177896|gb|ECD94031.1| hypothetical protein G ( 269) 422 102.4 1.7e-18  
gi|140412812|gb|ECL59115.1| hypothetical protein G ( 270) 422 102.4 1.7e-18  
gi|142684103|gb|ECZ83120.1| hypothetical protein G ( 223) 421 102.1 1.7e-18  
gi|134346105|gb|EBA79260.1| hypothetical protein G ( 272) 422 102.4 1.7e-18

gi|168992013|gb|ACA39553.1| 3-phosphoshikimate 1-c ( 186) 420 101.8 1.8e-18  
gi|141109424|gb|ECP69193.1| hypothetical protein G ( 230) 421 102.1 1.8e-18  
gi|143575841|gb|EDF71719.1| hypothetical protein G ( 156) 419 101.6 1.8e-18  
gi|139452944|gb|ECF20286.1| hypothetical protein G ( 283) 422 102.4 1.8e-18  
gi|140015718|gb|ECU02171.1| hypothetical protein G ( 284) 422 102.4 1.8e-18  
gi|197317472|gb|ACH66919.1| 3-phosphoshikimate 1-c ( 426) 424 103.0 1.8e-18  
gi|109701064|gb|ABG40984.1| 3-phosphoshikimate 1-c ( 429) 424 103.0 1.8e-18  
gi|40109989|gb|AAR52369.1| Sequence 12086 from pat ( 429) 424 103.0 1.8e-18  
gi|142646574|gb|ECZ56461.1| hypothetical protein G ( 172) 419 101.6 1.9e-18  
gi|140456534|gb|ECL88212.1| hypothetical protein G ( 146) 418 101.3 2e-18  
gi|141581759|gb|ECS46281.1| hypothetical protein G ( 263) 421 102.2 2e-18  
gi|135704029|gb|EBJ32424.1| hypothetical protein G ( 267) 421 102.2 2e-18  
gi|134844834|gb|EBD80509.1| hypothetical protein G ( 270) 421 102.2 2e-18  
gi|134502363|gb|EBB72442.1| hypothetical protein G ( 273) 421 102.2 2e-18  
gi|141204390|gb|ECQ34073.1| hypothetical protein G ( 186) 419 101.6 2e-18  
gi|61396474|gb|AAX45936.1| 5-enolpyruvylshikimate ( 413) 423 102.8 2.1e-18  
gi|134676029|gb|EBC75916.1| hypothetical protein G ( 190) 419 101.6 2.1e-18  
gi|76875277|emb|ICA86498.1| 3-phosphoshikimate 1-c ( 425) 423 102.8 2.1e-18  
gi|229465653|gb|ACQ67427.1| 3-enolpyruvylshikimate ( 428) 423 102.8 2.1e-18  
gi|253781755|emb|CAQ84918.1| 3-phosphoshikimate 1- ( 431) 423 102.8 2.1e-18  
gi|138149138|gb|EBY02188.1| hypothetical protein G ( 201) 419 101.7 2.2e-18  
gi|223692065|gb|ACN15348.1| AroA [Desulfo bacterium ( 439) 423 102.8 2.2e-18  
gi|141338053|gb|ECR14097.1| hypothetical protein G ( 139) 417 101.1 2.2e-18  
gi|61396478|gb|AAX45938.1| 5-enolpyruvylshikimate ( 413) 422 102.5 2.4e-18  
gi|137874616|gb|EBW49779.1| hypothetical protein G ( 196) 418 101.4 2.5e-18  
gi|139491037|gb|ECF43255.1| hypothetical protein G ( 252) 419 101.7 2.6e-18  
gi|134542621|gb|EBB96479.1| hypothetical protein G ( 258) 419 101.7 2.6e-18  
gi|138274656|gb|EBY73771.1| hypothetical protein G ( 267) 419 101.8 2.7e-18  
gi|135809669|gb|EBJ98309.1| hypothetical protein G ( 270) 419 101.8 2.7e-18  
gi|144195827|gb|EDU56799.1| hypothetical protein G ( 405) 421 102.3 2.7e-18  
gi|139584109|gb|ECG07146.1| hypothetical protein G ( 277) 419 101.8 2.8e-18  
gi|137699902|gb|EBV52881.1| hypothetical protein G ( 188) 417 101.2 2.8e-18  
gi|143212084|gb|EDD62644.1| hypothetical protein G ( 194) 417 101.2 2.8e-18  
gi|146315739|gb|ABQ20278.1| 3-phosphoshikimate 1-c ( 426) 421 102.3 2.9e-18  
gi|2270113635|gb|ACP09845.1| 3-phosphoshikimate 1-c ( 426) 421 102.3 2.9e-18  
gi|140698605|gb|ECM90343.1| hypothetical protein G ( 240) 418 101.5 2.9e-18  
gi|71846535|gb|AAZ46031.1| cytidylate kinase / 3-p ( 643) 423 102.9 2.9e-18  
gi|139386082|gb|EEB76188.1| hypothetical protein G ( 139) 415 100.7 3e-18  
gi|36786091|emb|CAE15143.1| unnamed protein produc ( 451) 421 102.4 3e-18  
gi|134420848|gb|EBB24938.1| hypothetical protein G ( 210) 417 101.2 3e-18  
gi|140799955|gb|ECN59559.1| hypothetical protein G ( 147) 415 100.7 3.1e-18  
gi|143269600|gb|EDE03471.1| hypothetical protein G ( 283) 418 101.6 3.3e-18  
gi|122937807|gb|ABM68632.1| plastid EPS synthase ( 514) 421 102.4 3.3e-18  
gi|5957577|gb|AAE08251.1| Sequence 59 from patent ( 427) 420 102.1 3.3e-18  
gi|144974756|gb|ABP12467.1| Sequence 59 from paten ( 427) 420 102.1 3.3e-18  
gi|2485254|gb|AAB73390.1|I44479 Sequence 59 from p ( 427) 420 102.1 3.3e-18  
gi|562760|emb|CAA57812.1| 3-phosphoshikimate 1-car ( 427) 420 102.1 3.3e-18  
gi|157083610|gb|ABV13288.1| hypothetical protein C ( 427) 420 102.1 3.3e-18  
gi|2484177|gb|AAB72313.1|I49206 Sequence 59 from p ( 427) 420 102.1 3.3e-18  
gi|49612043|emb|CAG75492.1| 3-phosphoshikimate 1-c ( 429) 420 102.1 3.3e-18  
gi|137214858|gb|EBS84591.1| hypothetical protein G ( 162) 415 100.7 3.3e-18  
gi|141041963|gb|ECP23312.1| hypothetical protein G ( 163) 415 100.7 3.4e-18  
gi|135141355|gb|EBF75716.1| hypothetical protein G ( 169) 415 100.7 3.5e-18  
gi|142982479|gb|EDB96508.1| hypothetical protein G ( 215) 416 101.0 3.6e-18  
gi|135408047|gb|EBH43560.1| hypothetical protein G ( 263) 417 101.3 3.6e-18  
gi|139522890|gb|ECF65242.1| hypothetical protein G ( 218) 416 101.0 3.6e-18  
gi|140173280|gb|ECK01590.1| hypothetical protein G ( 267) 417 101.3 3.6e-18  
gi|61396476|gb|AAX45937.1| 5-enolpyruvylshikimate ( 413) 419 101.9 3.8e-18  
gi|61396458|gb|AAX45928.1| 5-enolpyruvylshikimate ( 413) 419 101.9 3.8e-18  
gi|61396466|gb|AAX45932.1| 5-enolpyruvylshikimate ( 413) 419 101.9 3.8e-18  
gi|61396472|gb|AAX45935.1| 5-enolpyruvylshikimate ( 413) 419 101.9 3.8e-18  
gi|61396460|gb|AAX45929.1| 5-enolpyruvylshikimate ( 413) 419 101.9 3.8e-18

gi|59480469|gb|AAW86256.1| 5-enolpyruvylshikimate- ( 426) 419 101.9 3.9e-18  
gi|229378861|gb|EEO28952.1| prephenate dehydrogena ( 439) 419 101.9 4e-18  
gi|35186983|gb|AAQ84158.1| PlmJK [Streptomyces sp. (1003) 423 103.1 4.1e-18  
gi|135610065|gb|EBI74288.1| hypothetical protein G ( 181) 414 100.5 4.2e-18  
gi|140301825|gb|ECK83158.1| hypothetical protein G ( 273) 416 101.1 4.3e-18  
gi|158270965|gb|ED096795.1| 5-enolpyruvylshikimate ( 512) 419 102.0 4.5e-18  
gi|18957974|gb|AAL79610.1| 3-phosphoshikimate 1-ca ( 428) 418 101.7 4.5e-18  
gi|140667355|gb|ECM68197.1| hypothetical protein G ( 175) 413 100.3 4.8e-18  
gi|135678884|gb|EBJ16882.1| hypothetical protein G ( 222) 414 100.6 5e-18  
gi|61396462|gb|AAx45930.1| 5-enolpyruvylshikimate ( 413) 417 101.5 5.1e-18  
gi|138432976|gb|EBZ63840.1| hypothetical protein G ( 232) 414 100.6 5.1e-18  
gi|247538562|gb|ACT07183.1| 3-phosphoshikimate 1-c ( 429) 417 101.5 5.3e-18  
gi|140740736|gb|ECN19673.1| hypothetical protein G ( 175) 412 100.1 5.6e-18  
gi|136039035|gb|EBL49925.1| hypothetical protein G ( 179) 412 100.1 5.7e-18  
gi|140677578|gb|ECM75581.1| hypothetical protein G ( 179) 412 100.1 5.7e-18  
gi|141587996|gb|ECS48226.1| hypothetical protein G ( 278) 414 100.7 5.9e-18  
gi|61396486|gb|AAx45942.1| 5-enolpyruvylshikimate ( 413) 416 101.2 5.9e-18  
gi|206569939|gb|ACT11715.1| 3-phosphoshikimate 1-c ( 427) 416 101.3 6.1e-18  
gi|36784993|emb|CAB13913.1| 3-phosphoshikimate 1-c ( 428) 416 101.3 6.1e-18  
gi|138872276|gb|ECC23886.1| hypothetical protein G ( 242) 413 100.4 6.2e-18  
gi|139504738|gb|ECF52676.1| hypothetical protein G ( 244) 413 100.4 6.2e-18  
gi|139938907|gb|ECI50641.1| hypothetical protein G ( 268) 413 100.4 6.7e-18  
gi|18376380|emb|CAD21207.1| probable PENTAFUNCTION (1563) 422 103.0 6.7e-18  
gi|134747467|gb|EBD17355.1| hypothetical protein G ( 183) 411 99.9 6.7e-18  
gi|140540312|gb|ECM15770.1| hypothetical protein G ( 183) 411 99.9 6.7e-18  
gi|226906411|gb|EEH92329.1| 5-enolpyruvylshikimate ( 427) 415 101.0 7.1e-18  
gi|251754698|gb|ACT12774.1| 3-phosphoshikimate 1-c ( 429) 415 101.0 7.1e-18  
gi|140385024|gb|ECL41258.1| hypothetical protein G ( 178) 410 99.7 7.7e-18  
gi|49651207|emb|CAG78144.1| YALIOF12639p [Yarrowia (1556) 421 102.8 7.8e-18  
gi|138403520|gb|EBZ43001.1| hypothetical protein G ( 183) 410 99.7 7.8e-18  
gi|61396464|gb|AAx45931.1| 5-enolpyruvylshikimate ( 413) 414 100.8 8e-18  
gi|135003079|gb|EBE86104.1| hypothetical protein G ( 194) 410 99.7 8.2e-18  
gi|111969689|gb|ABH78640.1| Sequence 8483 from pat ( 431) 414 100.8 8.3e-18  
gi|136022122|gb|EBL38615.1| hypothetical protein G ( 240) 411 100.0 8.3e-18  
gi|141161546|gb|ECQ05541.1| hypothetical protein G ( 242) 411 100.0 8.3e-18  
gi|140419894|gb|ECF63396.1| hypothetical protein G ( 243) 411 100.0 8.4e-18  
gi|142870290|gb|EDB16562.1| hypothetical protein G ( 266) 411 100.0 9e-18  
gi|140374552|gb|ECL33774.1| hypothetical protein G ( 153) 408 99.2 9.2e-18  
gi|143280338|gb|EDE09543.1| hypothetical protein G ( 155) 408 99.2 9.3e-18  
gi|167322676|gb|ABZ52969.1| Sequence 10068 from pa ( 428) 413 100.6 9.6e-18  
gi|141087172|gb|ECP53417.1| hypothetical protein G ( 162) 408 99.2 9.6e-18  
gi|139894040|gb|ECI19042.1| hypothetical protein G ( 199) 409 99.5 9.7e-18  
gi|239801199|gb|ACS18266.1| 3-phosphoshikimate 1-c ( 670) 415 101.2 1e-17  
gi|136855918|gb|EBQ86151.1| hypothetical protein G ( 254) 410 99.8 1e-17  
gi|140238212|gb|ECK46609.1| hypothetical protein G ( 211) 409 99.5 1e-17  
gi|139153582|gb|ECD76894.1| hypothetical protein G ( 219) 409 99.5 1e-17  
gi|139621088|gb|ECG31545.1| hypothetical protein G ( 182) 408 99.2 1.1e-17  
gi|136961527|gb|EBR42455.1| hypothetical protein G ( 184) 408 99.2 1.1e-17  
gi|456205|emb|CAA54317.1| 5-enolpyruvylshikimate-3 ( 409) 412 100.4 1.1e-17  
gi|135668636|gb|EBJ10592.1| hypothetical protein G ( 281) 410 99.8 1.1e-17  
gi|226714722|gb|ACO73860.1| AroA [Laribacter hongk ( 428) 412 100.4 1.1e-17  
gi|139007276|gb|ECC76408.1| hypothetical protein G ( 243) 409 99.5 1.1e-17  
gi|135856027|gb|EBK27473.1| hypothetical protein G ( 215) 408 99.3 1.2e-17  
gi|136604454|gb|EBP23203.1| hypothetical protein G ( 401) 411 100.1 1.2e-17  
gi|167294414|gb|ABZ24728.1| Sequence 21216 from pa ( 421) 411 100.2 1.3e-17  
gi|139023026|gb|ECC87309.1| hypothetical protein G ( 161) 406 98.8 1.3e-17  
gi|139604357|gb|ECG21255.1| hypothetical protein G ( 240) 408 99.3 1.3e-17  
gi|138274368|gb|EBY73567.1| hypothetical protein G ( 256) 408 99.3 1.4e-17  
gi|137282407|gb|EBT22623.1| hypothetical protein G ( 118) 404 98.2 1.4e-17  
gi|141051049|gb|ECP29554.1| hypothetical protein G ( 223) 407 99.1 1.4e-17  
gi|88187784|gb|ABD40781.1| 3-phosphoshikimate 1-ca ( 430) 410 99.9 1.5e-17  
gi|229378960|gb|EEO29051.1| prephenate dehydrogena ( 440) 410 100.0 1.5e-17

gi|138296295|gb|EBY85353.1| hypothetical protein G ( 139) 404 98.3 1.6e-17  
gi|134914016|gb|EBE26328.1| hypothetical protein G ( 175) 405 98.6 1.6e-17  
gi|135465600|gb|EBH82121.1| hypothetical protein G ( 266) 407 99.1 1.6e-17  
gi|141069352|gb|ECP41137.1| hypothetical protein G ( 183) 405 98.6 1.7e-17  
gi|141402472|gb|ECS7949.1| hypothetical protein G ( 228) 406 98.9 1.7e-17  
gi|61396480|gb|AAx45939.1| 5-enolpyruvylshikimate ( 413) 409 99.7 1.7e-17  
gi|140169499|gb|ECJ98961.1| hypothetical protein G ( 237) 406 98.9 1.7e-17  
gi|137329660|gb|EBT48876.1| hypothetical protein G ( 161) 404 98.3 1.8e-17  
gi|140049962|gb|ECJ24138.1| hypothetical protein G ( 240) 406 98.9 1.8e-17  
gi|139132738|gb|ECD62600.1| hypothetical protein G ( 164) 404 98.3 1.8e-17  
gi|138340681|gb|EBZ07096.1| hypothetical protein G ( 136) 403 98.0 1.8e-17  
gi|135945291|gb|EBK88343.1| hypothetical protein G ( 254) 406 98.9 1.8e-17  
gi|137398175|gb|EBT87602.1| hypothetical protein G ( 257) 406 98.9 1.9e-17  
gi|143037211|gb|EDC35437.1| hypothetical protein G ( 174) 404 98.3 1.9e-17  
gi|15420539|gb|AAK97382.1| AF360730\_1 5-enolpyruvyl ( 427) 408 99.5 2e-17  
gi|88599102|gb|ABD44571.1| putative 3-phosphoshiki ( 428) 408 99.5 2e-17  
gi|56311966|emb|CAI06611.1| 3-phosphoshikimate 1-c ( 646) 410 100.1 2.1e-17  
gi|140869146|gb|ECO05747.1| hypothetical protein G ( 169) 403 98.1 2.1e-17  
gi|135370189|gb|EBH18092.1| hypothetical protein G ( 254) 405 98.7 2.1e-17  
gi|140317761|gb|ECK94388.1| hypothetical protein G ( 257) 405 98.7 2.2e-17  
gi|61396482|gb|AAx45940.1| 5-enolpyruvylshikimate ( 413) 407 99.3 2.3e-17  
gi|261605294|gb|ACT87780.1| 3-phosphoshikimate 1-c ( 429) 407 99.3 2.4e-17  
gi|140953256|gb|ECO62988.1| hypothetical protein G ( 256) 404 98.5 2.5e-17  
gi|137400223|gb|EBT88769.1| hypothetical protein G ( 262) 404 98.5 2.6e-17  
gi|61396484|gb|AAx45941.1| 5-enolpyruvylshikimate ( 413) 406 99.1 2.7e-17  
gi|61396496|gb|AAx45947.1| 5-enolpyruvylshikimate ( 413) 406 99.1 2.7e-17  
gi|134455928|gb|EBB45238.1| hypothetical protein G ( 267) 403 98.3 3e-17  
gi|142770656|gb|EDA45614.1| hypothetical protein G ( 186) 401 97.7 3.1e-17  
gi|138139300|gb|EBX96194.1| hypothetical protein G ( 275) 403 98.3 3.1e-17  
gi|139078368|gb|ECB26173.1| hypothetical protein G ( 227) 402 98.0 3.1e-17  
gi|137841823|gb|EBW30955.1| hypothetical protein G ( 230) 402 98.0 3.1e-17  
gi|139589854|gb|ECG11246.1| hypothetical protein G ( 160) 400 97.4 3.2e-17  
gi|140449002|gb|ECL82972.1| hypothetical protein G ( 288) 403 98.3 3.2e-17  
gi|157317811|gb|ABV36909.1| 3-phosphoshikimate 1-c ( 426) 405 98.9 3.2e-17  
gi|217498375|gb|ACK46568.1| 3-phosphoshikimate 1-c ( 426) 405 98.9 3.2e-17  
gi|84779489|dbj|BAE74266.1| 3-phosphoshikimate 1-c ( 435) 405 98.9 3.2e-17  
gi|138542215|gb|ECA31380.1| hypothetical protein G ( 252) 402 98.0 3.3e-17  
gi|144124603|gb|EDU04217.1| hypothetical protein G ( 180) 400 97.5 3.5e-17  
gi|140575918|gb|ECM27279.1| hypothetical protein G ( 268) 402 98.1 3.5e-17  
gi|139772528|gb|ECH35546.1| hypothetical protein G ( 269) 402 98.1 3.5e-17  
gi|61396494|gb|AAx45946.1| 5-enolpyruvylshikimate ( 413) 404 98.6 3.6e-17  
gi|46914055|emb|CAG20835.1| putative 3-phosphoshik ( 428) 404 98.6 3.7e-17  
gi|137241746|gb|EBS99587.1| hypothetical protein G ( 241) 401 97.8 3.8e-17  
gi|140283088|gb|ECK73670.1| hypothetical protein G ( 241) 401 97.8 3.8e-17  
gi|138964745|gb|ECC61414.1| hypothetical protein G ( 164) 399 97.2 3.8e-17  
gi|142199015|gb|ECW33272.1| hypothetical protein G ( 439) 404 98.6 3.8e-17  
gi|124260205|gb|ABM95199.1| 3-phosphoshikimate 1-c ( 674) 406 99.2 3.9e-17  
gi|135066978|gb|EBF28079.1| hypothetical protein G ( 119) 397 96.7 4e-17  
gi|138522325|gb|ECA17436.1| hypothetical protein G ( 215) 400 97.5 4e-17  
gi|61396488|gb|AAx45943.1| 5-enolpyruvylshikimate ( 413) 403 98.4 4.2e-17  
gi|61396490|gb|AAx45944.1| 5-enolpyruvylshikimate ( 413) 403 98.4 4.2e-17  
gi|138219010|gb|EBY49364.1| hypothetical protein G ( 193) 399 97.3 4.3e-17  
gi|71914680|gb|AAZ54582.1| 3-phosphoshikimate 1-ca ( 424) 403 98.4 4.3e-17  
gi|139917801|gb|ECI35869.1| hypothetical protein G ( 239) 400 97.6 4.3e-17  
gi|141442027|gb|ECR85618.1| hypothetical protein G ( 163) 398 97.0 4.4e-17  
gi|261372470|gb|ACX75215.1| 3-phosphoshikimate 1-c ( 442) 403 98.4 4.4e-17  
gi|137219979|gb|EBS87529.1| hypothetical protein G ( 255) 400 97.6 4.6e-17  
gi|135610492|gb|EBT74558.1| hypothetical protein G ( 211) 399 97.3 4.6e-17  
gi|136582315|gb|EBP09076.1| hypothetical protein G ( 218) 399 97.3 4.7e-17  
gi|137401562|gb|EBT89526.1| hypothetical protein G ( 229) 399 97.3 4.9e-17  
gi|167353991|gb|ABZ76604.1| 3-phosphoshikimate 1-c ( 426) 402 98.2 5e-17  
gi|57160999|emb|CAH57905.1| 3-phosphoshikimate 1-c ( 427) 402 98.2 5e-17

gi|58417478|emb|CAI26682.1| 3-phosphoshikimate 1-c ( 427) 402 98.2 5e-17  
gi|141869737|gb|ECT82397.1| hypothetical protein G ( 238) 399 97.4 5e-17  
gi|142661590|gb|EC267038.1| hypothetical protein G ( 163) 397 96.8 5.1e-17  
gi|140165244|gb|ECJ96063.1| hypothetical protein G ( 241) 399 97.4 5.1e-17  
gi|143418238|gb|EDE85044.1| hypothetical protein G ( 243) 399 97.4 5.1e-17  
gi|257477369|gb|ACV57688.1| 3-phosphoshikimate 1-c ( 450) 402 98.2 5.2e-17  
gi|88176432|gb|EAQ83900.1| hypothetical protein CH (1464) 408 99.9 5.3e-17  
gi|135563740|gb|EBI45267.1| hypothetical protein G ( 260) 399 97.4 5.4e-17  
gi|139566690|gb|ECF95076.1| hypothetical protein G ( 262) 399 97.4 5.4e-17  
gi|142002556|gb|ECU74821.1| hypothetical protein G ( 265) 399 97.4 5.5e-17  
gi|143173373|gb|EDD34522.1| hypothetical protein G ( 182) 397 96.8 5.5e-17  
gi|170945047|emb|CAP71158.1| unnamed protein produ (1567) 408 99.9 5.6e-17  
gi|134895707|gb|EBE14212.1| hypothetical protein G ( 184) 397 96.8 5.6e-17  
gi|140710317|gb|ECM98211.1| hypothetical protein G ( 227) 398 97.1 5.6e-17  
gi|135898935|gb|EBK56671.1| hypothetical protein G ( 277) 399 97.4 5.7e-17  
gi|33148677|gb|AAP96196.1| 3-phosphoshikimate 1-ca ( 435) 401 98.0 5.9e-17  
gi|141085446|gb|ECP52224.1| hypothetical protein G ( 204) 397 96.9 6e-17  
gi|141091027|gb|ECP95076.1| hypothetical protein G ( 206) 397 96.9 6.1e-17  
gi|34104358|gb|AAQ60717.1| 3-phosphoshikimate 1-ca ( 456) 401 98.0 6.2e-17  
gi|135398882|gb|EBH37393.1| hypothetical protein G ( 267) 398 97.2 6.4e-17  
gi|138749005|gb|ECB73949.1| hypothetical protein G ( 225) 397 96.9 6.5e-17  
gi|138076915|gb|EBX61317.1| hypothetical protein G ( 162) 395 96.4 6.8e-17  
gi|140570131|gb|ECM24965.1| hypothetical protein G ( 204) 396 96.7 7e-17  
gi|134767539|gb|EBD29427.1| hypothetical protein G ( 257) 397 96.9 7.2e-17  
gi|140438577|gb|ECL75479.1| hypothetical protein G ( 268) 397 97.0 7.5e-17  
gi|139560754|gb|ECF91299.1| hypothetical protein G ( 227) 396 96.7 7.6e-17  
gi|139398702|gb|ECE84262.1| hypothetical protein G ( 229) 396 96.7 7.7e-17  
gi|143544422|gb|EDF55404.1| hypothetical protein G ( 189) 395 96.4 7.7e-17  
gi|197053763|gb|ACH25461.1| Sequence 13 from paten ( 419) 399 97.5 7.8e-17  
gi|167274362|gb|ABZ27226.1| Sequence 1164 from pat ( 419) 399 97.5 7.8e-17  
gi|2621856|gb|AAB85269.1| 5-enolpyruvylshikimate 3 ( 419) 399 97.5 7.8e-17  
gi|145564830|gb|ABP75765.1| 3-phosphoshikimate 1-c ( 426) 399 97.6 7.9e-17  
gi|218355966|emb|CAQ88582.1| 5-enolpyruvylshikimat ( 434) 399 97.6 8e-17  
gi|134823747|gb|EBD66995.1| hypothetical protein G ( 255) 396 96.7 8.4e-17  
gi|120590404|gb|ABM33844.1| 3-phosphoshikimate 1-c ( 679) 401 98.1 8.4e-17  
gi|134462524|gb|EBB49131.1| hypothetical protein G ( 220) 395 96.5 8.7e-17  
gi|141988235|gb|ECU65125.1| hypothetical protein G ( 150) 393 95.9 8.7e-17  
gi|142902212|gb|EDB39353.1| hypothetical protein G ( 228) 395 96.5 8.9e-17  
gi|125997494|gb|ABN61569.1| 3-phosphoshikimate 1-c ( 426) 398 97.3 9.2e-17  
gi|33113261|gb|AAP94212.1| 3-phosphoshikimate 1-ca ( 427) 398 97.3 9.2e-17  
gi|151424006|dbj|BAF71509.1| 3-phosphoshikimate 1- ( 427) 398 97.3 9.2e-17  
gi|156532856|gb|ABU77682.1| hypothetical protein E ( 428) 398 97.3 9.2e-17  
gi|142194251|gb|ECW29626.1| hypothetical protein G ( 439) 398 97.3 9.4e-17  
gi|142196972|gb|ECW31688.1| hypothetical protein G ( 439) 398 97.3 9.4e-17  
gi|61396492|gb|AAX45945.1| 5-enolpyruvylshikimate ( 413) 397 97.1 1e-16  
gi|120558876|gb|ABM24803.1| 3-phosphoshikimate 1-c ( 426) 397 97.1 1.1e-16  
gi|157846892|gb|ABV87391.1| 3-phosphoshikimate 1-c ( 426) 397 97.1 1.1e-16  
gi|115422436|emb|CAJ48961.1| 3-phosphoshikimate 1- ( 439) 397 97.1 1.1e-16  
gi|142200197|gb|ECW34167.1| hypothetical protein G ( 439) 397 97.1 1.1e-16  
gi|142068853|gb|ECV33175.1| hypothetical protein G ( 449) 397 97.1 1.1e-16  
gi|141237520|gb|ECQ57535.1| hypothetical protein G ( 217) 393 96.0 1.2e-16  
gi|138521898|gb|ECA17149.1| hypothetical protein G ( 236) 393 96.1 1.2e-16  
gi|141399647|gb|ECR55895.1| hypothetical protein G ( 160) 391 95.5 1.2e-16  
gi|167294358|gb|ABZ47222.1| Sequence 21160 from pa ( 425) 396 96.9 1.2e-16  
gi|33504186|emb|CAD83448.1| 3-phosphoshikimate 1-c ( 433) 396 96.9 1.3e-16  
gi|142171806|gb|ECW12483.1| hypothetical protein G ( 435) 396 96.9 1.3e-16  
gi|144974761|gb|ABP12472.1| Sequence 64 from paten ( 442) 396 96.9 1.3e-16  
gi|2485259|gb|AAB73395.1| I44484 Sequence 64 from p ( 442) 396 96.9 1.3e-16  
gi|5957582|gb|AAE08256.1| Sequence 64 from patent ( 442) 396 96.9 1.3e-16  
gi|2484182|gb|AAB72318.1| I49211 Sequence 64 from p ( 442) 396 96.9 1.3e-16  
gi|144037|gb|AAA22968.1| 5-enolpyruvylshikimate-3- ( 442) 396 96.9 1.3e-16  
gi|85722750|gb|ABC77693.1| 3-phosphoshikimate-1-ca ( 446) 396 96.9 1.3e-16

gi|136122103|gb|EBM06235.1| hypothetical protein G ( 253) 393 96.1 1.3e-16  
gi|139732399|gb|ECH09160.1| hypothetical protein G ( 255) 393 96.1 1.3e-16  
gi|142099007|gb|ECV58104.1| hypothetical protein G ( 173) 391 95.5 1.3e-16  
gi|136993418|gb|EBR60500.1| hypothetical protein G ( 222) 392 95.8 1.4e-16  
gi|167281308|gb|ABZ34172.1| Sequence 8110 from pat ( 416) 395 96.7 1.4e-16  
gi|194342310|gb|EDX23276.1| 3-phosphoshikimate 1-c ( 428) 395 96.7 1.4e-16  
gi|242131724|gb|ACS86026.1| 3-phosphoshikimate 1-c ( 429) 395 96.7 1.5e-16  
gi|150840321|gb|ABR74292.1| 3-phosphoshikimate 1-c ( 433) 395 96.7 1.5e-16  
gi|141190121|gb|ECQ25416.1| hypothetical protein G ( 253) 392 95.9 1.5e-16  
gi|141918038|gb|ECU15844.1| hypothetical protein G ( 257) 392 95.9 1.5e-16  
gi|139696464|gb|ECG84097.1| hypothetical protein G ( 188) 390 95.3 1.6e-16  
gi|212556775|gb|ACJ29229.1| 3-phosphoshikimate 1-c ( 426) 394 96.5 1.7e-16  
gi|135134197|gb|EBF71145.1| hypothetical protein G ( 430) 394 96.5 1.7e-16  
gi|138904179|gb|ECC35826.1| hypothetical protein G ( 242) 391 95.6 1.7e-16  
gi|139089556|gb|ECN33905.1| hypothetical protein G ( 244) 391 95.6 1.7e-16  
gi|262091733|gb|ACY25322.1| 3-phosphoshikimate 1-c ( 370) 393 96.2 1.8e-16  
gi|126227113|gb|ABN90653.1| 3-phosphoshikimate 1-c ( 451) 394 96.5 1.8e-16  
gi|242139654|gb|ECS26056.1| 3-phosphoshikimate 1-c ( 451) 394 96.5 1.8e-16  
gi|157805653|gb|EDU82823.1| 3-phosphoshikimate 1-c ( 451) 394 96.5 1.8e-16  
gi|169653245|gb|EDS85938.1| 3-phosphoshikimate 1-c ( 451) 394 96.5 1.8e-16  
gi|76581267|gb|ABA50742.1| 3-phosphoshikimate 1-ca ( 451) 394 96.5 1.8e-16  
gi|157935342|gb|EDU91012.1| 3-phosphoshikimate 1-c ( 451) 394 96.5 1.8e-16  
gi|184212268|gb|EDU09311.1| 3-phosphoshikimate 1-c ( 451) 394 96.5 1.8e-16  
gi|254219427|gb|EET08811.1| 3-phosphoshikimate 1-c ( 451) 394 96.5 1.8e-16  
gi|126219794|gb|ABN83300.1| 3-phosphoshikimate 1-c ( 451) 394 96.5 1.8e-16  
gi|237506815|gb|ACQ99133.1| 3-phosphoshikimate 1-c ( 451) 394 96.5 1.8e-16  
gi|138376143|gb|EBZ24210.1| hypothetical protein G ( 214) 390 95.4 1.8e-16  
gi|140011396|gb|ECT99246.1| hypothetical protein G ( 226) 390 95.4 1.9e-16  
gi|81241720|gb|ABB62430.1| 5-enolpyruvylshikimate- ( 424) 393 96.2 1.9e-16  
gi|208009607|emb|CAQ79903.1| 3-phosphoshikimate 1- ( 427) 393 96.2 2e-16  
gi|138333550|gb|EBZ02342.1| hypothetical protein G ( 238) 390 95.4 2e-16  
gi|91709269|gb|ABE49197.1| 3-phosphoshikimate 1-ca ( 429) 393 96.2 2e-16  
gi|91709413|gb|ABE49341.1| 3-phosphoshikimate 1-ca ( 429) 393 96.2 2e-16  
gi|139024443|gb|ECC88277.1| hypothetical protein G ( 242) 390 95.4 2e-16  
gi|140767972|gb|ECN37198.1| hypothetical protein G ( 204) 389 95.1 2e-16  
gi|140374206|gb|ECL33521.1| hypothetical protein G ( 253) 390 95.4 2.1e-16  
gi|135683728|gb|EBU19875.1| hypothetical protein G ( 215) 389 95.1 2.1e-16  
gi|137852928|gb|EBW37332.1| hypothetical protein G ( 221) 389 95.2 2.2e-16  
gi|135593367|gb|EBI63979.1| hypothetical protein G ( 269) 390 95.4 2.2e-16  
gi|135512902|gb|EBI12761.1| hypothetical protein G ( 271) 390 95.4 2.2e-16  
gi|140889253|gb|ECN01965.1| hypothetical protein G ( 273) 390 95.4 2.2e-16  
gi|137525040|gb|EBU56039.1| hypothetical protein G ( 286) 390 95.5 2.3e-16  
gi|85816343|gb|ABC84481.1| putative 5-enolpyruvyls ( 160) 387 94.6 2.3e-16  
gi|209774898|gb|ACI85761.1| 5-enolpyruvylshikimate ( 427) 392 96.0 2.3e-16  
gi|257763436|dbj|BAI34931.1| 5-enolpyruvylshikimat ( 427) 392 96.0 2.3e-16  
gi|218431432|emb|CAR12310.1| 5-enolpyruvylshikimat ( 427) 392 96.0 2.3e-16  
gi|254591507|gb|ACF70868.1| 3-phosphoshikimate 1-c ( 427) 392 96.0 2.3e-16  
gi|209911417|dbj|BAG76491.1| 3-phosphoshikimate 1- ( 427) 392 96.0 2.3e-16  
gi|197053764|gb|ACH25462.1| Sequence 14 from paten ( 427) 392 96.0 2.3e-16  
gi|1787137|gb|AAC73994.1| 5-enolpyruvylshikimate-3 ( 427) 392 96.0 2.3e-16  
gi|209774902|gb|ACI85763.1| 5-enolpyruvylshikimate ( 427) 392 96.0 2.3e-16  
gi|209774906|gb|ACI85765.1| 5-enolpyruvylshikimate ( 427) 392 96.0 2.3e-16  
gi|15113316|gb|AAE68900.1| Sequence 7 from patent ( 427) 392 96.0 2.3e-16  
gi|1651430|dbj|BAA35643.1| 5-enolpyruvylshikimate- ( 427) 392 96.0 2.3e-16  
gi|12514074|gb|AAG55393.1| AE005280\_4 5-enolpyruvyl ( 427) 392 96.0 2.3e-16  
gi|261073316|gb|AAN79516.1| AE016758\_120 3-phosphosh ( 427) 392 96.0 2.3e-16  
gi|169888401|gb|ACB02108.1| 5-enolpyruvylshikimate ( 427) 392 96.0 2.3e-16  
gi|209161693|gb|ACI39126.1| 3-phosphoshikimate 1-c ( 427) 392 96.0 2.3e-16  
gi|238859881|gb|ACR61879.1| 5-enolpyruvylshikimate ( 427) 392 96.0 2.3e-16  
gi|81246057|gb|ABB66765.1| 5-enolpyruvylshikimate- ( 427) 392 96.0 2.3e-16  
gi|187430562|gb|ACD09836.1| 3-phosphoshikimate 1-c ( 427) 392 96.0 2.3e-16  
gi|260449946|gb|ACX40368.1| 3-phosphoshikimate 1-c ( 427) 392 96.0 2.3e-16

gi|257752848|dbj|BAI24350.1| 5-enolpyruvylshikimat ( 427) 392 96.0 2.3e-16  
gi|209774900|gb|ACI85762.1| 5-enolpyruvylshikimate ( 427) 392 96.0 2.3e-16  
gi|133604500|dbj|BAB34414.1| 5-enolpyruvylshikimate ( 427) 392 96.0 2.3e-16  
gi|218351113|emb|CAU96817.1| 5-enolpyruvylshikimat ( 427) 392 96.0 2.3e-16  
gi|157077738|gb|ABV17446.1| 3-phosphoshikimate 1-c ( 427) 392 96.0 2.3e-16  
gi|218360262|emb|CAQ97812.1| 5-enolpyruvylshikimat ( 427) 392 96.0 2.3e-16  
gi|257758305|dbj|BAI29802.1| 5-enolpyruvylshikimat ( 427) 392 96.0 2.3e-16  
gi|73854941|gb|AAZ87648.1| 5-enolpyruvylshikimate- ( 427) 392 96.0 2.3e-16  
gi|167273981|gb|ABZ26845.1| Sequence 783 from pate ( 427) 392 96.0 2.3e-16  
gi|138641134|gb|ECA98992.1| hypothetical protein G ( 197) 388 94.9 2.3e-16  
gi|82411256|gb|ABB75365.1| 3-phosphoshikimate 1-ca ( 437) 392 96.0 2.3e-16  
gi|115250880|emb|CAJ68705.1| 3-phosphoshikimate 1- ( 437) 392 96.0 2.3e-16  
gi|138831281|gb|ECC06379.1| hypothetical protein G ( 136) 386 94.3 2.3e-16  
gi|88598075|gb|ABD43545.1| putative 3-phosphoshiki ( 440) 392 96.0 2.3e-16  
gi|139365969|gb|ECE63735.1| hypothetical protein G ( 206) 388 94.9 2.4e-16  
gi|142263836|gb|ECW81652.1| hypothetical protein G ( 140) 386 94.4 2.4e-16  
gi|137098844|gb|EBS19574.1| hypothetical protein G ( 208) 388 94.9 2.4e-16  
gi|139515492|gb|ECF260106.1| hypothetical protein G ( 149) 386 94.4 2.5e-16  
gi|135036529|gb|EBF08664.1| hypothetical protein G ( 423) 391 95.8 2.6e-16  
gi|151365419|gb|ABS08419.1| 3-phosphoshikimate 1-c ( 426) 391 95.8 2.6e-16  
gi|209774904|gb|ACI85764.1| 5-enolpyruvylshikimate ( 427) 391 95.8 2.6e-16  
gi|3859585|gb|AAC72854.1| 3-enolpyruvylshikimate-5 ( 427) 391 95.8 2.6e-16  
gi|91715109|gb|ABE55035.1| 3-phosphoshikimate 1-ca ( 428) 391 95.8 2.7e-16  
gi|139449952|gb|ECF18163.1| hypothetical protein G ( 239) 388 95.0 2.7e-16  
gi|149905851|gb|ABR36684.1| 3-phosphoshikimate 1-c ( 435) 391 95.8 2.7e-16  
gi|143954156|gb|EDH82710.1| hypothetical protein G ( 136) 385 94.1 2.7e-16  
gi|140917730|gb|ECQ37955.1| hypothetical protein G ( 256) 388 95.0 2.8e-16  
gi|139203617|gb|ECE11938.1| hypothetical protein G ( 260) 388 95.0 2.8e-16  
gi|136974924|gb|EBR50068.1| hypothetical protein G ( 190) 386 94.5 3e-16  
gi|30040688|gb|AAP16419.1| 5-enolpyruvylshikimate- ( 427) 390 95.6 3.1e-16  
gi|110614463|gb|ABF03130.1| 5-enolpyruvylshikimate ( 427) 390 95.6 3.1e-16  
gi|24051176|gb|AAN42533.1| 5-enolpyruvylshikimate- ( 427) 390 95.6 3.1e-16  
gi|260216483|emb|CBA29637.1| 3-phosphoshikimate 1- ( 428) 390 95.6 3.1e-16  
gi|52307682|gb|AAU38182.1| AroA protein [Mannheimi ( 433) 390 95.6 3.1e-16  
gi|6013215|gb|AAF01290.1|AF182427\_3 5-enolpyruvyls ( 442) 390 95.6 3.2e-16  
gi|135784761|gb|EBJ82581.1| hypothetical protein G ( 250) 387 94.8 3.2e-16  
gi|142771453|gb|EDA46198.1| hypothetical protein G ( 253) 387 94.8 3.2e-16  
gi|134450405|gb|EBB41981.1| hypothetical protein G ( 174) 385 94.2 3.3e-16  
gi|141500557|gb|ECS07759.1| hypothetical protein G ( 212) 386 94.5 3.3e-16  
gi|138644794|gb|ECB01472.1| hypothetical protein G ( 262) 387 94.8 3.3e-16  
gi|139984221|gb|ECI1817.1| hypothetical protein G ( 100) 382 93.4 3.4e-16  
gi|137204744|gb|EBS78807.1| hypothetical protein G ( 155) 384 94.0 3.5e-16  
gi|138900787|gb|ECC34332.1| hypothetical protein G ( 232) 386 94.5 3.5e-16  
gi|215264105|emb|CAS08449.1| 5-enolpyruvylshikimat ( 427) 389 95.4 3.6e-16  
gi|218364582|emb|CAR02268.1| 5-enolpyruvylshikimat ( 427) 389 95.4 3.6e-16  
gi|110342697|gb|ABG68934.1| 3-phosphoshikimate 1-c ( 427) 389 95.4 3.6e-16  
gi|91071583|gb|ABE06464.1| 3-phosphoshikimate 1-ca ( 427) 389 95.4 3.6e-16  
gi|218426312|emb|CAR07137.1| 5-enolpyruvylshikimat ( 427) 389 95.4 3.6e-16  
gi|115512250|gb|ABJ700325.1| AroA [Escherichia coli ( 427) 389 95.4 3.6e-16  
gi|226898314|gb|EEH84573.1| AroA [Escherichia sp. ( 427) 389 95.4 3.6e-16  
gi|139068739|gb|ECD19686.1| hypothetical protein G ( 163) 384 94.0 3.6e-16  
gi|144114020|gb|EDI96529.1| hypothetical protein G ( 438) 389 95.4 3.6e-16  
gi|28852194|gb|AAO55268.1| prephenate dehydrogenas ( 535) 390 95.7 3.7e-16  
gi|151280371|gb|ABR88781.1| 3-phosphoshikimate 1-c ( 441) 389 95.4 3.7e-16  
gi|147750604|gb|EDK57673.1| 3-phosphoshikimate 1-c ( 451) 389 95.4 3.7e-16  
gi|121229589|gb|ABM52107.1| 3-phosphoshikimate 1-c ( 451) 389 95.4 3.7e-16  
gi|160697726|gb|EDP87696.1| 3-phosphoshikimate 1-c ( 451) 389 95.4 3.7e-16  
gi|148029460|gb|EDK87365.1| 3-phosphoshikimate 1-c ( 451) 389 95.4 3.7e-16  
gi|126241624|gb|ABO04717.1| 3-phosphoshikimate 1-c ( 451) 389 95.4 3.7e-16  
gi|124291356|gb|ABN00625.1| 3-phosphoshikimate 1-c ( 451) 389 95.4 3.7e-16  
gi|147745264|gb|EDK52344.1| 3-phosphoshikimate 1-c ( 451) 389 95.4 3.7e-16  
gi|52428112|gb|AAU48705.1| 3-phosphoshikimate-1-ca ( 451) 389 95.4 3.7e-16

gi|138258528|gb|EBY62639.1| hypothetical protein G ( 260) 386 94.6 3.8e-16  
gi|140993087|gb|ECO90844.1| hypothetical protein G ( 152) 383 93.7 4e-16  
gi|139229948|gb|ECB30111.1| hypothetical protein G ( 234) 385 94.3 4.1e-16  
gi|141003112|gb|ECO97873.1| hypothetical protein G ( 242) 385 94.3 4.2e-16  
gi|140079528|gb|ECU40341.1| hypothetical protein G ( 242) 385 94.3 4.2e-16  
gi|136855917|gb|EBQ86150.1| hypothetical protein G ( 243) 385 94.3 4.2e-16  
gi|137366429|gb|EBT69661.1| hypothetical protein G ( 245) 385 94.3 4.3e-16  
gi|133576881|emb|CAE33961.1| 3-phosphoshikimate 1-c ( 442) 388 95.2 4.3e-16  
gi|13571749|emb|CAE41250.1| 3-phosphoshikimate 1-c ( 442) 388 95.2 4.3e-16  
gi|52208740|emb|CAH34676.1| putative 3-phosphoshik ( 451) 388 95.2 4.3e-16  
gi|142161969|gb|ECW04955.1| hypothetical protein G ( 372) 387 94.9 4.3e-16  
gi|135772951|gb|EBJ75107.1| hypothetical protein G ( 255) 385 94.3 4.4e-16  
gi|138122745|gb|EBX87000.1| hypothetical protein G ( 262) 385 94.3 4.5e-16  
gi|140676897|gb|ECM75089.1| hypothetical protein G ( 127) 381 93.2 4.7e-16  
gi|218370560|emb|CBR18367.1| 5-enolpyruvylshikimat ( 427) 387 94.9 4.8e-16  
gi|253972923|gb|ACT38594.1| 3-phosphoshikimate 1-c ( 427) 387 94.9 4.8e-16  
gi|142851614|gb|ABO89935.1| 3-phosphoshikimate 1-c ( 427) 387 94.9 4.8e-16  
gi|169755618|gb|ACA78317.1| 3-phosphoshikimate 1-c ( 427) 387 94.9 4.8e-16  
gi|170521635|gb|ACB19813.1| 3-phosphoshikimate 1-c ( 427) 387 94.9 4.8e-16  
gi|253325107|gb|ACT29709.1| 3-phosphoshikimate 1-c ( 427) 387 94.9 4.8e-16  
gi|242376723|emb|CAQ31436.1| aroA [Escherichia col ( 427) 387 94.9 4.8e-16  
gi|253977137|gb|ACT42807.1| 3-phosphoshikimate 1-c ( 427) 387 94.9 4.8e-16  
gi|157066110|gb|ABV05365.1| 3-phosphoshikimate 1-c ( 427) 387 94.9 4.8e-16  
gi|281178039|dbj|BAI54369.1| 3-phosphoshikimate 1- ( 427) 387 94.9 4.8e-16  
gi|140345209|gb|ECI13444.1| hypothetical protein G ( 196) 383 93.8 4.8e-16  
gi|21623205|gb|AAM67855.1| 3-phosphoshikimate 1-ca ( 428) 387 94.9 4.8e-16  
gi|29470333|gb|AAC05428.1| 3-enolpyruvylshikimate-5 ( 428) 387 94.9 4.8e-16  
gi|142367329|gb|ECX56585.1| hypothetical protein G ( 438) 387 94.9 4.9e-16  
gi|133574090|emb|CAE38415.1| 3-phosphoshikimate 1-c ( 442) 387 95.0 5e-16  
gi|143576125|gb|EDF71819.1| hypothetical protein G ( 172) 382 93.6 5.1e-16  
gi|134610603|gb|EBC37222.1| hypothetical protein G ( 258) 384 94.1 5.2e-16  
gi|137006884|gb|EBR68201.1| hypothetical protein G ( 215) 383 93.8 5.2e-16  
gi|134963519|gb|EBE59309.1| hypothetical protein G ( 263) 384 94.1 5.2e-16  
gi|135563570|gb|EBI45165.1| hypothetical protein G ( 189) 382 93.6 5.5e-16  
gi|5957550|gb|AAE08224.1| Sequence 8 from patent U ( 423) 386 94.7 5.6e-16  
gi|2485227|gb|AAB73363.1|I44452 Sequence 8 from pa ( 423) 386 94.7 5.6e-16  
gi|144974729|gb|ABP12440.1| Sequence 8 from patent ( 423) 386 94.7 5.6e-16  
gi|2484150|gb|AAB72286.1|I49179 Sequence 8 from pa ( 423) 386 94.7 5.6e-16  
gi|409661|emb|CAA25223.1| unnamed protein product [ ( 427) 386 94.7 5.6e-16  
gi|4731170|gb|AAD28375.1|AF110153\_2 5-enolpyruvoyl ( 428) 386 94.7 5.6e-16  
gi|238869916|gb|ACR69627.1| 3-phosphoshikimate 1-c ( 428) 386 94.7 5.6e-16  
gi|137822467|gb|EBW19744.1| hypothetical protein G ( 110) 379 92.8 5.7e-16  
gi|140310203|gb|ECK88940.1| hypothetical protein G ( 137) 380 93.0 5.8e-16  
gi|71554514|gb|AAZ33725.1| prephenate dehydrogenas ( 366) 385 94.5 5.8e-16  
gi|139774292|gb|ECH36762.1| hypothetical protein G ( 177) 381 93.3 6.1e-16  
gi|141329021|gb|ECR07965.1| hypothetical protein G ( 264) 383 93.9 6.1e-16  
gi|141395406|gb|ECR52919.1| hypothetical protein G ( 125) 379 92.8 6.3e-16  
gi|125595950|gb|EAZ35730.1| hypothetical protein O ( 506) 386 94.8 6.4e-16  
gi|169212099|gb|ACA49855.1| enoylpyrovyl-shikimate ( 427) 385 94.5 6.5e-16  
gi|143270564|gb|EDE04150.1| hypothetical protein G ( 428) 385 94.5 6.6e-16  
gi|226935015|gb|ACO92617.1| AroA [Edwardsiella tar ( 428) 385 94.5 6.6e-16  
gi|140659527|gb|ECM62702.1| hypothetical protein G ( 253) 382 93.7 6.9e-16  
gi|221730036|gb|ACM32856.1| 3-phosphoshikimate 1-c ( 673) 387 95.1 6.9e-16  
gi|141168332|gb|ECQ10383.1| hypothetical protein G ( 147) 379 92.8 7.1e-16  
gi|134685548|gb|EBC81580.1| hypothetical protein G ( 151) 379 92.9 7.2e-16  
gi|125715790|gb|ABN54282.1| 3-phosphoshikimate 1-c ( 423) 384 94.3 7.6e-16  
gi|24348389|gb|AAN55438.1|AE015681\_2 3-phosphoshik ( 426) 384 94.3 7.6e-16  
gi|139503952|gb|ECF52132.1| hypothetical protein G ( 240) 381 93.4 7.7e-16  
gi|142089203|gb|ECV49931.1| hypothetical protein G ( 198) 380 93.2 7.7e-16  
gi|135563091|gb|EBI44861.1| hypothetical protein G ( 243) 381 93.4 7.7e-16  
gi|63257580|gb|AAY38676.1| Prephenate dehydrogenas ( 534) 385 94.6 7.8e-16  
gi|137396460|gb|EBT86593.1| hypothetical protein G ( 247) 381 93.5 7.8e-16

gi|137245288|gb|EBT01614.1| hypothetical protein G ( 215) 380 93.2 8.2e-16  
gi|138880582|gb|ECC27596.1| hypothetical protein G ( 270) 381 93.5 8.4e-16  
gi|138615794|gb|ECA82879.1| hypothetical protein G ( 225) 380 93.2 8.5e-16  
gi|167273444|gb|ABZ26308.1| Sequence 246 from pate ( 341) 382 93.8 8.6e-16  
gi|138512522|gb|ECA10830.1| hypothetical protein G ( 235) 380 93.2 8.8e-16  
gi|17978913|gb|AAL47682.1| 5-enolpyruvylshikimate ( 426) 383 94.1 8.8e-16  
gi|160865242|gb|ABX21865.1| hypothetical protein S ( 427) 383 94.1 8.8e-16  
gi|253985497|gb|ACT50354.1| 3-phosphoshikimate 1-c ( 432) 383 94.1 8.9e-16  
gi|260209657|emb|CBA63357.1| 3-phosphoshikimate 1- ( 437) 383 94.1 9e-16  
gi|260213106|emb|CBE04506.1| 3-phosphoshikimate 1- ( 437) 383 94.1 9e-16  
gi|247534368|gb|ACS97614.1| 3-phosphoshikimate 1-c ( 441) 383 94.1 9.1e-16  
gi|134893001|gb|EBE12405.1| hypothetical protein G ( 137) 377 92.4 9.1e-16  
gi|137545103|gb|EBU66926.1| hypothetical protein G ( 116) 376 92.1 9.3e-16  
gi|262109647|gb|EEY67699.1| pentafunctional AROM p (1548) 389 95.8 9.7e-16  
gi|138795121|gb|ECB89349.1| hypothetical protein G ( 235) 379 93.0 1e-15  
gi|136238252|gb|EBM81150.1| hypothetical protein G ( 195) 378 92.7 1e-15  
gi|111913995|gb|ABH66278.1| Sequence 10 from paten ( 426) 382 93.8 1e-15  
gi|155109881|gb|ABT06556.1| Sequence 10 from paten ( 426) 382 93.8 1e-15  
gi|3452007|gb|AAC32745.1| EPSF synthase AroA [Shig ( 427) 382 93.9 1e-15  
gi|190687453|gb|ACE85131.1| 3-phosphoshikimate 1-c ( 428) 382 93.9 1e-15  
gi|140301223|gb|ECK82759.1| hypothetical protein G ( 210) 378 92.7 1.1e-15  
gi|140343211|gb|ECL12005.1| hypothetical protein G ( 262) 379 93.0 1.1e-15  
gi|137447190|gb|EBU15268.1| hypothetical protein G ( 268) 379 93.0 1.1e-15  
gi|139532179|gb|ECF71291.1| hypothetical protein G ( 222) 378 92.8 1.1e-15  
gi|261354892|gb|EEY17320.1| pentafunctional AROM p (1574) 388 95.6 1.1e-15  
gi|160861032|gb|ABX49566.1| 3-phosphoshikimate 1-c ( 426) 381 93.6 1.2e-15  
gi|267985183|gb|ACY85012.1| 3-phosphoshikimate 1-c ( 428) 381 93.6 1.2e-15  
gi|229374383|gb|EE024774.1| 3-phosphoshikimate 1-c ( 443) 381 93.6 1.2e-15  
gi|140209686|gb|ECK27413.1| hypothetical protein G ( 176) 376 92.3 1.3e-15  
gi|134570021|gb|EBC13001.1| hypothetical protein G ( 262) 378 92.8 1.3e-15  
gi|593745|gb|AAA55352.1| Sequence 4 from Patent EP ( 427) 380 93.4 1.4e-15  
gi|139589855|gb|ECG11247.1| hypothetical protein G ( 162) 375 92.0 1.4e-15  
gi|141407657|gb|ECR61667.1| hypothetical protein G ( 257) 377 92.6 1.5e-15  
gi|135392272|gb|EBH32957.1| hypothetical protein G ( 215) 376 92.3 1.5e-15  
gi|134567711|gb|EBC11607.1| hypothetical protein G ( 226) 376 92.3 1.6e-15  
gi|139590611|gb|ECG11788.1| hypothetical protein G ( 126) 373 91.5 1.6e-15  
gi|138839333|gb|ECC10257.1| hypothetical protein G ( 156) 374 91.8 1.6e-15  
gi|140376762|gb|ECL35372.1| hypothetical protein G ( 191) 375 92.1 1.6e-15  
gi|194712899|gb|ACF92120.1| 3-phosphoshikimate 1-c ( 427) 379 93.2 1.6e-15  
gi|194407735|gb|ACF67954.1| 3-phosphoshikimate 1-c ( 427) 379 93.2 1.6e-15  
gi|141828747|gb|ECT53330.1| hypothetical protein G ( 161) 374 91.8 1.6e-15  
gi|140259294|gb|ECK61125.1| hypothetical protein G ( 244) 376 92.4 1.7e-15  
gi|136202243|gb|EBM56814.1| hypothetical protein G ( 203) 375 92.1 1.7e-15  
gi|1356831953|gb|EBJ18772.1| hypothetical protein G ( 170) 374 91.8 1.7e-15  
gi|136386119|gb|EBN81787.1| hypothetical protein G ( 171) 374 91.8 1.7e-15  
gi|142607846|gb|ECZ29157.1| hypothetical protein G ( 192) 374 91.8 1.9e-15  
gi|164194901|gb|AAL19912.1| 3-enolpyruvylshikimate- ( 427) 378 93.0 1.9e-15  
gi|267992715|gb|ACY87600.1| 3-phosphoshikimate 1-c ( 427) 378 93.0 1.9e-15  
gi|261246194|emb|CBG23998.1| 3-phosphoshikimate 1- ( 427) 378 93.0 1.9e-15  
gi|137244587|gb|EBT01212.1| hypothetical protein G ( 238) 375 92.1 1.9e-15  
gi|135597997|gb|EBI66840.1| hypothetical protein G ( 244) 375 92.1 1.9e-15  
gi|140146263|gb|ECJ83031.1| hypothetical protein G ( 167) 373 91.6 1.9e-15  
gi|256363216|gb|ACU76713.1| 3-phosphoshikimate 1-c ( 454) 378 93.0 2e-15  
gi|142746219|gb|EDA27786.1| hypothetical protein G ( 211) 374 91.9 2e-15  
gi|141443182|gb|ECR86432.1| hypothetical protein G ( 146) 372 91.3 2e-15  
gi|141055602|gb|ECP32851.1| hypothetical protein G ( 179) 373 91.6 2e-15  
gi|139521800|gb|ECF64621.1| hypothetical protein G ( 271) 375 92.2 2.1e-15  
gi|137798212|gb|EBW05651.1| hypothetical protein G ( 232) 374 91.9 2.1e-15  
gi|140433318|gb|ECL71874.1| hypothetical protein G ( 233) 374 91.9 2.2e-15  
gi|141889995|gb|ECT96217.1| hypothetical protein G ( 237) 374 91.9 2.2e-15  
gi|161364150|gb|ABX67918.1| hypothetical protein S ( 427) 377 92.8 2.2e-15  
gi|224467314|gb|ACN45144.1| 3-phosphoshikimate 1-c ( 427) 377 92.8 2.2e-15

gi|194404814|gb|ACF65036.1| 3-phosphoshikimate 1-c ( 427) 377 92.8 2.2e-15  
gi|62127135|gb|AAK64838.1| 3-enolpyruvylshikimate- ( 427) 377 92.8 2.2e-15  
gi|138264743|gb|EBY66745.1| hypothetical protein G ( 161) 372 91.4 2.2e-15  
gi|137234142|gb|EBS95536.1| hypothetical protein G ( 242) 374 91.9 2.2e-15  
gi|140182536|gb|ECK08290.1| hypothetical protein G ( 206) 373 91.7 2.3e-15  
gi|136315848|gb|EBN33986.1| hypothetical protein G ( 172) 372 91.4 2.3e-15  
gi|140890066|gb|ECO20239.1| hypothetical protein G ( 143) 371 91.1 2.3e-15  
gi|142342306|gb|ECX39599.1| hypothetical protein G ( 214) 373 91.7 2.3e-15  
gi|142942923|gb|EDB68713.1| hypothetical protein G ( 264) 374 92.0 2.4e-15  
gi|220975650|gb|EED93978.1| predicted protein [Tha ( 486) 377 92.8 2.4e-15  
gi|143676232|gb|EDG21149.1| hypothetical protein G ( 152) 371 91.1 2.4e-15  
gi|137786059|gb|EBV98729.1| hypothetical protein G ( 156) 371 91.1 2.5e-15  
gi|126638175|gb|ABO23818.1| 3-phosphoshikimate 1-c ( 426) 376 92.5 2.5e-15  
gi|261413785|gb|ACX83156.1| 3-phosphoshikimate 1-c ( 443) 376 92.6 2.6e-15  
gi|140013850|gb|ECJ00937.1| hypothetical protein G ( 203) 372 91.4 2.6e-15  
gi|183580816|dbj|BAG29287.1| 3-phosphoshikimate 1- ( 445) 376 92.6 2.6e-15  
gi|194343512|gb|EDX24478.1| 3-phosphoshikimate 1-c ( 450) 376 92.6 2.6e-15  
gi|142317060|gb|ECR20980.1| hypothetical protein G ( 119) 369 90.6 2.7e-15  
gi|139996187|gb|ECT90104.1| hypothetical protein G ( 158) 370 90.9 2.9e-15  
gi|197213620|gb|ACH51017.1| 3-phosphoshikimate 1-c ( 427) 375 92.3 3e-15  
gi|143912469|gb|EDH52565.1| hypothetical protein G ( 434) 375 92.3 3e-15  
gi|134745039|gb|EBD15968.1| hypothetical protein G ( 248) 372 91.5 3.1e-15  
gi|137056263|gb|EBR96016.1| hypothetical protein G ( 254) 372 91.5 3.1e-15  
gi|135581832|gb|EBI56856.1| hypothetical protein G ( 211) 371 91.2 3.1e-15  
gi|137553390|gb|EBU71649.1| hypothetical protein G ( 224) 371 91.2 3.3e-15  
gi|142778570|gb|EDA51511.1| hypothetical protein G ( 154) 369 90.7 3.3e-15  
gi|143182082|gb|EDD40705.1| hypothetical protein G ( 188) 370 91.0 3.3e-15  
gi|150837332|gb|ABR71308.1| 3-phosphoshikimate 1-c ( 425) 374 92.1 3.4e-15  
gi|269413|gb|AAA01407.1| Sequence 3 from Patent US ( 427) 374 92.1 3.4e-15  
gi|139503982|gb|ECF52152.1| hypothetical protein G ( 240) 371 91.3 3.5e-15  
gi|140215326|gb|ECK31463.1| hypothetical protein G ( 168) 369 90.7 3.6e-15  
gi|141237519|gb|ECQ57534.1| hypothetical protein G ( 211) 370 91.0 3.6e-15  
gi|136199323|gb|EBM54861.1| hypothetical protein G ( 175) 369 90.7 3.7e-15  
gi|138754029|gb|ECB75799.1| hypothetical protein G ( 263) 371 91.3 3.7e-15  
gi|140386793|gb|ECL42479.1| hypothetical protein G ( 147) 368 90.5 3.7e-15  
gi|142827467|gb|EDL88112.1| hypothetical protein G ( 265) 371 91.3 3.7e-15  
gi|140428372|gb|ECL68838.1| hypothetical protein G ( 236) 370 91.0 4e-15  
gi|141350069|gb|ECR21238.1| hypothetical protein G ( 160) 368 90.5 4e-15  
gi|113884944|gb|ABI38996.1| 3-phosphoshikimate 1-c ( 426) 373 91.9 4e-15  
gi|117612752|gb|ABK48206.1| 3-phosphoshikimate 1-c ( 426) 373 91.9 4e-15  
gi|142038696|gb|ECV07294.1| hypothetical protein G ( 426) 373 91.9 4e-15  
gi|113888993|gb|ABI43044.1| 3-phosphoshikimate 1-c ( 426) 373 91.9 4e-15  
gi|120606889|gb|ABM42629.1| 3-phosphoshikimate 1-c ( 673) 375 92.5 4.2e-15  
gi|141550593|gb|ECS30851.1| hypothetical protein G ( 209) 369 90.8 4.2e-15  
gi|137239733|gb|EBS98468.1| hypothetical protein G ( 174) 368 90.5 4.3e-15  
gi|135964545|gb|EBL01341.1| hypothetical protein G ( 178) 368 90.5 4.3e-15  
gi|137987089|gb|EBX13684.1| hypothetical protein G ( 221) 369 90.8 4.4e-15  
gi|143454987|gb|EDF05938.1| hypothetical protein G ( 186) 368 90.5 4.5e-15  
gi|137016135|gb|EBR73396.1| hypothetical protein G ( 230) 369 90.8 4.5e-15  
gi|135052511|gb|EBF18868.1| hypothetical protein G ( 422) 372 91.7 4.6e-15  
gi|71147195|gb|AAZ27668.1| 3-phosphoshikimate 1-ca ( 426) 372 91.7 4.6e-15  
gi|1781351|emb|CAA71382.1| aroA [Salmonella typhim ( 427) 372 91.7 4.7e-15  
gi|153876|gb|AAA27028.1| 5-enolpyruvylshikimate-3- ( 427) 372 91.7 4.7e-15  
gi|140303484|gb|ECK84310.1| hypothetical protein G ( 256) 369 90.9 4.9e-15  
gi|138907006|gb|ECC37097.1| hypothetical protein G ( 149) 366 90.0 5.1e-15  
gi|140064591|gb|ECJ31472.1| hypothetical protein G ( 222) 368 90.6 5.1e-15  
gi|593743|gb|AAA55350.1| Sequence 2 from Patent EP ( 427) 371 91.5 5.4e-15  
gi|139934982|gb|EC147822.1| hypothetical protein G ( 239) 368 90.6 5.4e-15  
gi|135273602|gb|EBG56384.1| hypothetical protein G ( 263) 368 90.6 5.8e-15  
gi|167277430|gb|ABZ30294.1| Sequence 4232 from pat (1542) 377 93.2 5.9e-15  
gi|143181940|gb|EDD40611.1| hypothetical protein G ( 223) 367 90.4 6e-15  
gi|137056967|gb|EBR96418.1| hypothetical protein G ( 226) 367 90.4 6e-15

gi|138215696|gb|EBY47061.1| hypothetical protein G ( 128) 364 89.5 6.1e-15  
gi|219621962|gb|ACL30118.1| 3-phosphoshikimate 1-c ( 427) 370 91.2 6.3e-15  
gi|205271966|emb|CAR36810.1| 3-phosphoshikimate 1- ( 427) 370 91.2 6.3e-15  
gi|15113317|gb|AAE68901.1| Sequence 8 from patent ( 427) 370 91.2 6.3e-15  
gi|197940547|gb|ACH77880.1| 3-phosphoshikimate 1-c ( 427) 370 91.2 6.3e-15  
gi|154366|gb|AAA27223.1| 5-enolpyruvylshikimate 3- ( 427) 370 91.2 6.3e-15  
gi|144974754|gb|ABP12465.1| Sequence 57 from paten ( 427) 370 91.2 6.3e-15  
gi|2485252|gb|AAB73388.1|I44477 Sequence 57 from p ( 427) 370 91.2 6.3e-15  
gi|206708172|emb|CAR32465.1| 3-phosphoshikimate 1- ( 427) 370 91.2 6.3e-15  
gi|2484175|gb|AAB72311.1|I49204 Sequence 57 from p ( 427) 370 91.2 6.3e-15  
gi|5957575|gb|AAE08249.1| Sequence 57 from patent ( 427) 370 91.2 6.3e-15  
gi|141381885|gb|ECR43532.1| hypothetical protein G ( 198) 366 90.1 6.4e-15  
gi|141159028|gb|ECQ03748.1| hypothetical protein G ( 245) 367 90.4 6.4e-15  
gi|139161830|gb|ECD82757.1| hypothetical protein G ( 174) 365 89.9 6.7e-15  
gi|136877422|gb|EBR00100.1| hypothetical protein G ( 262) 367 90.4 6.8e-15  
gi|140604008|gb|ECM35576.1| hypothetical protein G ( 183) 365 89.9 6.9e-15  
gi|136301860|gb|EBN24503.1| hypothetical protein G ( 236) 366 90.2 7.3e-15  
gi|137656038|gb|EBV27947.1| hypothetical protein G ( 160) 364 89.6 7.3e-15  
gi|143926832|gb|EDH63014.1| hypothetical protein G ( 426) 369 91.0 7.3e-15  
gi|219624517|gb|ACL30672.1| 3-phosphoshikimate 1-c ( 427) 369 91.0 7.3e-15  
gi|10038984|dbj|BAB13019.1| 3-phosphoshikimate 1-c ( 427) 369 91.0 7.3e-15  
gi|134379630|gb|EBB01558.1| hypothetical protein G ( 434) 369 91.0 7.4e-15  
gi|135336773|gb|EBG95699.1| hypothetical protein G ( 253) 366 90.2 7.7e-15  
gi|137340742|gb|EBT55118.1| hypothetical protein G ( 210) 365 89.9 7.7e-15  
gi|139222176|gb|ECE24995.1| hypothetical protein G ( 256) 366 90.2 7.7e-15  
gi|138153842|gb|EBY05412.1| hypothetical protein G ( 179) 364 89.6 7.9e-15  
gi|142950535|gb|EDB74156.1| hypothetical protein G ( 159) 363 89.4 8.4e-15  
gi|142105911|gb|ECV63243.1| hypothetical protein G ( 131) 362 89.1 8.4e-15  
gi|140955222|gb|ECO64409.1| hypothetical protein G ( 240) 365 90.0 8.6e-15  
gi|145689057|gb|ABP89563.1| 5-enolpyruvylshikimate ( 164) 363 89.4 8.6e-15  
gi|197702692|gb|EDY48504.1| 3-phosphoshikimate 1-c ( 440) 368 90.8 8.7e-15  
gi|193222394|emb|CAL62698.2| 3-phosphoshikimate 1- ( 442) 368 90.8 8.7e-15  
gi|142660287|gb|ECZ66125.1| hypothetical protein G ( 171) 363 89.4 8.9e-15  
gi|135256647|gb|EBG46395.1| hypothetical protein G ( 263) 365 90.0 9.2e-15  
gi|136044801|gb|EBL53783.1| hypothetical protein G ( 269) 365 90.0 9.4e-15  
gi|140310865|gb|ECK89393.1| hypothetical protein G ( 226) 364 89.7 9.5e-15  
gi|138037253|gb|EBX39901.1| hypothetical protein G ( 232) 364 89.7 9.7e-15  
gi|253983562|gb|ACT48420.1| 3-phosphoshikimate 1-c ( 426) 367 90.6 9.9e-15  
gi|137883640|gb|EBW54936.1| hypothetical protein G ( 243) 364 89.7 1e-14  
gi|256686562|gb|ACV09455.1| 3-phosphoshikimate 1-c ( 446) 367 90.6 1e-14  
gi|139462417|gb|ECF27014.1| hypothetical protein G ( 229) 363 89.5 1.1e-14  
gi|142757194|gb|EDA35755.1| hypothetical protein G ( 229) 363 89.5 1.1e-14  
gi|240857213|gb|ACS54880.1| 3-phosphoshikimate 1-c ( 420) 366 90.4 1.1e-14  
gi|124364003|gb|ABN07811.1| 3-phosphoshikimate 1-c ( 422) 366 90.4 1.1e-14  
gi|139532178|gb|ECF71290.1| hypothetical protein G ( 235) 363 89.5 1.1e-14  
gi|140315884|gb|ECK93023.1| hypothetical protein G ( 242) 363 89.5 1.2e-14  
gi|137524207|gb|EBU55586.1| hypothetical protein G ( 164) 361 89.0 1.2e-14  
gi|142544985|gb|ECY85240.1| hypothetical protein G ( 137) 360 88.7 1.2e-14  
gi|136460127|gb|EB030630.1| hypothetical protein G ( 383) 365 90.1 1.2e-14  
gi|256729198|gb|EEU42552.1| predicted protein [Nec (1569) 372 92.1 1.3e-14  
gi|141682065|gb|ECS79110.1| hypothetical protein G ( 223) 362 89.3 1.3e-14  
gi|138393430|gb|EBZ36056.1| hypothetical protein G ( 152) 360 88.7 1.3e-14  
gi|136145599|gb|EBM19604.1| hypothetical protein G ( 409) 365 90.1 1.3e-14  
gi|135427661|gb|EBH56773.1| hypothetical protein G ( 127) 359 88.4 1.3e-14  
gi|169812004|gb|ACA85688.1| 3-phosphoshikimate 1-c ( 426) 365 90.1 1.3e-14  
gi|16502141|emb|CAD05378.1| 3-phosphoshikimate 1-c ( 427) 365 90.1 1.3e-14  
gi|29138009|gb|AAO69570.1| 3-phosphoshikimate 1-ca ( 427) 365 90.1 1.3e-14  
gi|56128229|gb|AAV77735.1| 3-phosphoshikimate 1-ca ( 427) 365 90.1 1.3e-14  
gi|197094372|emb|CAR59885.1| 3-phosphoshikimate 1- ( 427) 365 90.1 1.3e-14  
gi|193084121|gb|ACF09788.1| 3-phosphoshikimate 1-c ( 429) 365 90.2 1.3e-14  
gi|135253730|gb|EBG44676.1| hypothetical protein G ( 442) 365 90.2 1.4e-14  
gi|108765521|gb|ABG04403.1| 3-phosphoshikimate 1-c ( 447) 365 90.2 1.4e-14

gi|197720164|gb|EDY64072.1| 3-phosphoshikimate 1-c ( 453) 365 90.2 1.4e-14  
gi|117648862|gb|ABK52964.1| 3-phosphoshikimate 1-c (1004) 369 91.3 1.4e-14  
gi|140421747|gb|ECL64640.1| hypothetical protein G ( 175) 360 88.8 1.4e-14  
gi|140302847|gb|ECK83869.1| hypothetical protein G ( 227) 361 89.1 1.5e-14  
gi|141365559|gb|ECR31937.1| hypothetical protein G ( 235) 361 89.1 1.5e-14  
gi|138406922|gb|EBZ45453.1| hypothetical protein G ( 237) 361 89.1 1.6e-14  
gi|136106430|gb|EBL95515.1| hypothetical protein G ( 110) 357 88.0 1.6e-14  
gi|143923175|gb|EDH60320.1| hypothetical protein G ( 244) 361 89.1 1.6e-14  
gi|139154621|gb|ECD77605.1| hypothetical protein G ( 301) 362 89.4 1.6e-14  
gi|134799090|gb|EBD51377.1| hypothetical protein G ( 141) 358 88.3 1.6e-14  
gi|137794484|gb|EBW03536.1| hypothetical protein G ( 220) 360 88.8 1.7e-14  
gi|141205783|gb|ECQ35040.1| hypothetical protein G ( 221) 360 88.8 1.7e-14  
gi|137796550|gb|EBW04720.1| hypothetical protein G ( 222) 360 88.8 1.7e-14  
gi|144012226|gb|EDI23677.1| hypothetical protein G ( 272) 361 89.1 1.7e-14  
gi|138629106|gb|ECS52762.1| hypothetical protein G ( 224) 360 88.9 1.7e-14  
gi|142554251|gb|ECY91763.1| hypothetical protein G ( 188) 359 88.6 1.8e-14  
gi|219547426|gb|ACL17876.1| 3-phosphoshikimate 1-c ( 423) 363 89.7 1.8e-14  
gi|142080005|gb|ECD742334.1| hypothetical protein G ( 352) 362 89.4 1.8e-14  
gi|136866671|gb|EBD93382.1| hypothetical protein G ( 246) 360 88.9 1.9e-14  
gi|138109040|gb|EBX79294.1| hypothetical protein G ( 246) 360 88.9 1.9e-14  
gi|135744502|gb|EBU57473.1| hypothetical protein G ( 211) 360 88.8 1.9e-14  
gi|141398247|gb|ECR54893.1| hypothetical protein G ( 119) 356 87.8 1.9e-14  
gi|142010326|gb|ECU80420.1| hypothetical protein G ( 180) 358 88.3 2e-14  
gi|137165195|gb|EBS56765.1| hypothetical protein G ( 182) 358 88.3 2e-14  
gi|135179383|gb|EBG00170.1| hypothetical protein G ( 155) 357 88.1 2e-14  
gi|227453530|gb|ACP32283.1| 3-phosphoshikimate 1-c ( 434) 362 89.5 2.1e-14  
gi|135951288|gb|EBR92380.1| hypothetical protein G ( 205) 358 88.4 2.2e-14  
gi|137326809|gb|EBT47329.1| hypothetical protein G ( 250) 359 88.7 2.2e-14  
gi|136647641|gb|EBP49413.1| hypothetical protein G ( 257) 359 88.7 2.2e-14  
gi|140376968|gb|ECL35520.1| hypothetical protein G ( 219) 358 88.4 2.3e-14  
gi|135600560|gb|EBT68434.1| hypothetical protein G ( 235) 358 88.4 2.4e-14  
gi|135035087|gb|EBF07763.1| hypothetical protein G ( 426) 361 89.3 2.4e-14  
gi|2484176|gb|AAB72312.1|I49205 Sequence 58 from p ( 427) 361 89.3 2.4e-14  
gi|2485253|gb|AAB73389.1|I44478 Sequence 58 from p ( 427) 361 89.3 2.4e-14  
gi|144974762|gb|ABP12473.1| Sequence 65 from paten ( 427) 361 89.3 2.4e-14  
gi|2484183|gb|AAB72319.1|I49212 Sequence 65 from p ( 427) 361 89.3 2.4e-14  
gi|5957583|gb|AAE08257.1| Sequence 65 from patent ( 427) 361 89.3 2.4e-14  
gi|5957576|gb|AAE08250.1| Sequence 58 from patent ( 427) 361 89.3 2.4e-14  
gi|47637|emb|CAA38417.1| unnamed protein product [ ( 427) 361 89.3 2.4e-14  
gi|144974755|gb|ABP12466.1| Sequence 58 from paten ( 427) 361 89.3 2.4e-14  
gi|237500387|gb|ACQ92980.1| 3-phosphoshikimate 1-c ( 427) 361 89.3 2.4e-14  
gi|2485260|gb|AAB73396.1|I44485 Sequence 65 from p ( 427) 361 89.3 2.4e-14  
gi|138112315|gb|EBX81121.1| hypothetical protein G ( 133) 355 87.6 2.5e-14  
gi|142221372|gb|ECW50540.1| hypothetical protein G ( 434) 361 89.3 2.5e-14  
gi|137840358|gb|EBW30125.1| hypothetical protein G ( 142) 355 87.6 2.6e-14  
gi|136940074|gb|EBR30361.1| hypothetical protein G ( 217) 357 88.2 2.7e-14  
gi|138737749|gb|ECB66651.1| hypothetical protein G ( 182) 356 87.9 2.7e-14  
gi|135145339|gb|EBF78282.1| hypothetical protein G ( 410) 360 89.0 2.8e-14  
gi|32397219|emb|CAD75256.1| 3-phosphoshikimate 1-c ( 501) 361 89.3 2.8e-14  
gi|255924589|gb|ACU40100.1| 3-phosphoshikimate 1-c ( 420) 360 89.1 2.8e-14  
gi|140996322|gb|ECO93156.1| hypothetical protein G ( 194) 356 87.9 2.8e-14  
gi|72393962|gb|AAZ68239.1| 3-phosphoshikimate 1-ca ( 428) 360 89.1 2.8e-14  
gi|137413372|gb|EBT96140.1| hypothetical protein G ( 205) 356 88.0 2.9e-14  
gi|142660118|gb|ECZ66009.1| hypothetical protein G ( 143) 354 87.4 3e-14  
gi|141051576|gb|ECP29931.1| hypothetical protein G ( 263) 357 88.2 3.1e-14  
gi|138368034|gb|EBT19603.1| hypothetical protein G ( 266) 357 88.3 3.1e-14  
gi|140238213|gb|ECK46610.1| hypothetical protein G ( 219) 356 88.0 3.1e-14  
gi|2171410917|gb|EEC50846.1| 3-phosphoshikimate 1-c ( 483) 360 89.1 3.1e-14  
gi|139077328|gb|ECM25425.1| hypothetical protein G ( 226) 356 88.0 3.2e-14  
gi|143362591|gb|EDE56421.1| hypothetical protein G ( 410) 359 88.8 3.2e-14  
gi|140972874|gb|ECO76607.1| hypothetical protein G ( 236) 356 88.0 3.3e-14  
gi|140263797|gb|ECK64234.1| hypothetical protein G ( 134) 353 87.2 3.3e-14

gi|145305567|gb|ABP56149.1| 3-phosphoshikimate 1-c ( 433) 359 88.8 3.3e-14  
gi|12721147|gb|AAK02923.1| AroA [Pasteurella multo ( 440) 359 88.9 3.4e-14  
gi|135240693|gb|EBG36992.1| hypothetical protein G ( 166) 354 87.4 3.4e-14  
gi|197710217|gb|EDY54251.1| 3-phosphoshikimate 1-c ( 441) 359 88.9 3.4e-14  
gi|142532130|gb|ECY75910.1| hypothetical protein G ( 137) 353 87.2 3.4e-14  
gi|135111798|gb|EBF56743.1| hypothetical protein G ( 138) 353 87.2 3.4e-14  
gi|142175064|gb|ECW14980.1| hypothetical protein G ( 138) 353 87.2 3.4e-14  
gi|139355526|gb|ECE60735.1| hypothetical protein G ( 205) 355 87.7 3.4e-14  
gi|143855585|gb|EDH11623.1| hypothetical protein G ( 118) 352 86.9 3.5e-14  
gi|140733658|gb|ECN14670.1| hypothetical protein G ( 260) 356 88.0 3.5e-14  
gi|138490939|gb|EBZ98202.1| hypothetical protein G ( 220) 355 87.8 3.6e-14  
gi|134912975|gb|EBE25634.1| hypothetical protein G ( 223) 355 87.8 3.7e-14  
gi|140482318|gb|ECL96816.1| hypothetical protein G ( 226) 355 87.8 3.7e-14  
gi|139459145|gb|ECF24711.1| hypothetical protein G ( 228) 355 87.8 3.7e-14  
gi|137634728|gb|EBV17335.1| hypothetical protein G ( 232) 355 87.8 3.8e-14  
gi|138811671|gb|ECB96825.1| hypothetical protein G ( 232) 355 87.8 3.8e-14  
gi|115255367|emb|CAK06442.1| putative 3-phosphoshi ( 420) 358 88.6 3.8e-14  
gi|137115528|gb|EBS28958.1| hypothetical protein G ( 160) 353 87.2 3.8e-14  
gi|138740065|gb|ECB68288.1| hypothetical protein G ( 160) 353 87.2 3.8e-14  
gi|8546878|emb|CAB94597.1| 3-phosphoshikimate 1-ca ( 438) 358 88.6 3.9e-14  
gi|209533895|gb|ACI53830.1| 3-phosphoshikimate 1-c ( 420) 357 88.4 4.4e-14  
gi|135777069|gb|EBJ77694.1| hypothetical protein G ( 237) 354 87.6 4.5e-14  
gi|137500712|gb|EBU43450.1| hypothetical protein G ( 138) 351 86.7 4.6e-14  
gi|140820545|gb|ECN74010.1| hypothetical protein G ( 206) 353 87.3 4.7e-14  
gi|141429392|gb|ECR76550.1| hypothetical protein G ( 251) 354 87.6 4.7e-14  
gi|141573502|gb|ECS42616.1| hypothetical protein G ( 141) 351 86.7 4.7e-14  
gi|135965679|gb|EBL02113.1| hypothetical protein G ( 174) 352 87.0 4.8e-14  
gi|144102999|gb|EDI88610.1| hypothetical protein G ( 118) 350 86.5 4.8e-14  
gi|137072304|gb|EBS05031.1| hypothetical protein G ( 224) 353 87.3 5e-14  
gi|141156551|gb|ECQ01946.1| hypothetical protein G ( 227) 353 87.3 5e-14  
gi|167292847|gb|ABZ45711.1| Sequence 19649 from pa ( 426) 356 88.2 5.2e-14  
gi|30180948|emb|CAD85875.1| EPSP synthase (3-phosp ( 431) 356 88.2 5.2e-14  
gi|260647047|emb|CBG70146.1| 3-phosphoshikimate 1- ( 443) 356 88.2 5.4e-14  
gi|141794130|gb|ECT32615.1| hypothetical protein G ( 145) 350 86.5 5.6e-14  
gi|140395644|gb|ECL48451.1| hypothetical protein G ( 235) 352 87.1 6e-14  
gi|114334589|gb|ABT17197.1| 3-phosphoshikimate 1-c ( 426) 355 88.0 6e-14  
gi|119863478|gb|ABM02955.1| 3-phosphoshikimate 1-c ( 428) 355 88.0 6.1e-14  
gi|238871683|gb|ACR71393.1| 3-phosphoshikimate 1-c ( 431) 355 88.0 6.1e-14  
gi|143892353|gb|EDH38461.1| hypothetical protein G ( 433) 355 88.0 6.1e-14  
gi|256796396|gb|ACV27052.1| 3-phosphoshikimate 1-c ( 442) 355 88.0 6.2e-14  
gi|143570931|gb|EDF69343.1| hypothetical protein G ( 138) 349 86.3 6.2e-14  
gi|140960254|gb|ECO68025.1| hypothetical protein G ( 254) 352 87.1 6.4e-14  
gi|135347203|gb|EBH02679.1| hypothetical protein G ( 222) 351 86.9 6.7e-14  
gi|135950277|gb|EBK91698.1| hypothetical protein G ( 230) 351 86.9 6.9e-14  
gi|137594212|gb|EBU94648.1| hypothetical protein G ( 190) 350 86.6 6.9e-14  
gi|140932293|gb|ECO48252.1| hypothetical protein G ( 195) 350 86.6 7e-14  
gi|137655461|gb|EBV27620.1| hypothetical protein G ( 201) 350 86.6 7.2e-14  
gi|140442678|gb|ECL78355.1| hypothetical protein G ( 201) 350 86.6 7.2e-14  
gi|140915160|gb|ECO36180.1| hypothetical protein G ( 142) 348 86.1 7.4e-14  
gi|139613416|gb|ECG27637.1| hypothetical protein G ( 211) 350 86.7 7.5e-14  
gi|138983487|gb|ECC68508.1| hypothetical protein G ( 179) 349 86.4 7.6e-14  
gi|135466250|gb|EBH82554.1| hypothetical protein G ( 477) 354 87.8 7.7e-14  
gi|142329282|gb|ECX29996.1| hypothetical protein G ( 185) 349 86.4 7.8e-14  
gi|137854155|gb|EBW38038.1| hypothetical protein G ( 227) 350 86.7 7.9e-14  
gi|140188819|gb|ECK12796.1| hypothetical protein G ( 191) 349 86.4 8e-14  
gi|137525237|gb|EBU56150.1| hypothetical protein G ( 234) 350 86.7 8.1e-14  
gi|141822648|gb|ECT49099.1| hypothetical protein G ( 194) 349 86.4 8.1e-14  
gi|76782198|gb|ABA54869.1| putative 5-enolpyruvyls ( 520) 354 87.8 8.2e-14  
gi|46400157|emb|CAF23606.1| putative 3-phosphoshik ( 939) 357 88.7 8.2e-14  
gi|38199572|emb|CAE49223.1| 3-phosphoshikimate 1-c ( 431) 353 87.5 8.2e-14  
gi|137080837|gb|EBS09819.1| hypothetical protein G ( 205) 349 86.4 8.5e-14  
gi|141447794|gb|ECR89501.1| hypothetical protein G ( 250) 350 86.7 8.5e-14

gi|141931434|gb|ECU25456.1| hypothetical protein G ( 251) 350 86.7 8.5e-14  
gi|143142368|gb|EDD12423.1| hypothetical protein G ( 383) 352 87.3 8.7e-14  
gi|143973504|gb|EDH96220.1| hypothetical protein G ( 125) 346 85.6 9.1e-14  
gi|135841015|gb|EBK18020.1| hypothetical protein G ( 196) 348 86.2 9.5e-14  
gi|141547228|gb|ECS28434.1| hypothetical protein G ( 199) 348 86.2 9.6e-14  
gi|139065978|gb|ECD17692.1| hypothetical protein G ( 247) 349 86.5 9.8e-14  
gi|136452097|gb|EBO25416.1| hypothetical protein G ( 204) 348 86.2 9.8e-14  
gi|141823036|gb|ECT49373.1| hypothetical protein G ( 211) 348 86.2 1e-13  
gi|141358226|gb|ECR26863.1| hypothetical protein G ( 218) 348 86.2 1e-13  
gi|86280632|gb|ABC89695.1| 3-phosphoshikimate 1-ca ( 420) 351 87.1 1.1e-13  
gi|118135516|gb|ABK62560.1| 3-phosphoshikimate 1-c ( 435) 351 87.1 1.1e-13  
gi|138464932|gb|EBZ86060.1| hypothetical protein G ( 165) 346 85.7 1.1e-13  
gi|218301274|emb|CAU98622.1| unnamed protein produ ( 450) 351 87.1 1.2e-13  
gi|135660988|gb|EBJ05847.1| hypothetical protein G ( 216) 347 86.0 1.2e-13  
gi|135957341|gb|EBK96478.1| hypothetical protein G ( 180) 346 85.7 1.2e-13  
gi|190695875|gb|ACE89960.1| 3-phosphoshikimate 1-c ( 420) 350 86.9 1.3e-13  
gi|135311115|gb|EBG78396.1| hypothetical protein G ( 162) 345 85.5 1.3e-13  
gi|135475974|gb|EBH89065.1| hypothetical protein G ( 242) 347 86.0 1.3e-13  
gi|140555709|gb|ECM20146.1| hypothetical protein G ( 207) 346 85.8 1.3e-13  
gi|139395594|gb|ECE82263.1| hypothetical protein G ( 198) 345 85.5 1.5e-13  
gi|136291219|gb|EBN17253.1| hypothetical protein G ( 252) 346 85.8 1.6e-13  
gi|138997581|gb|ECC72530.1| hypothetical protein G ( 230) 345 85.6 1.7e-13  
gi|136446532|gb|EBO21828.1| hypothetical protein G ( 191) 344 85.3 1.7e-13  
gi|135667198|gb|EBJ09710.1| hypothetical protein G ( 238) 345 85.6 1.7e-13  
gi|143377822|gb|EDE65681.1| hypothetical protein G ( 428) 348 86.5 1.7e-13  
gi|136805392|gb|EBQ52456.1| hypothetical protein G ( 197) 344 85.3 1.7e-13  
gi|135324770|gb|EBG87629.1| hypothetical protein G ( 135) 342 84.8 1.8e-13  
gi|189428021|gb|ACD98169.1| 5-enolpyruvylshikimate ( 445) 348 86.5 1.8e-13  
gi|137513759|gb|EBU50081.1| hypothetical protein G ( 179) 343 85.1 1.9e-13  
gi|137196438|gb|EBK96478.1| hypothetical protein G ( 149) 342 84.8 1.9e-13  
gi|136113324|gb|EBM00254.1| hypothetical protein G ( 182) 343 85.1 1.9e-13  
gi|137214083|gb|EBS84150.1| hypothetical protein G ( 222) 344 85.4 1.9e-13  
gi|161985195|gb|ABX80844.1| 3-phosphoshikimate 1-c ( 418) 347 86.2 2e-13  
gi|139407942|gb|ECE90556.1| hypothetical protein G ( 132) 341 84.5 2e-13  
gi|138935690|gb|ECC49156.1| hypothetical protein G ( 162) 342 84.8 2e-13  
gi|221155485|gb|ACM04612.1| 3-phosphoshikimate 1-c ( 434) 347 86.2 2e-13  
gi|138501049|gb|ECA04254.1| hypothetical protein G ( 242) 344 85.4 2.1e-13  
gi|41817616|gb|AAS12201.1| 3-phosphoshikimate 1-ca ( 449) 347 86.2 2.1e-13  
gi|138224534|gb|EBY53274.1| hypothetical protein G ( 212) 343 85.1 2.2e-13  
gi|135569263|gb|EBT48826.1| hypothetical protein G ( 149) 341 84.6 2.2e-13  
gi|141045191|gb|ECP25528.1| hypothetical protein G ( 149) 341 84.6 2.2e-13  
gi|138146322|gb|EBY00264.1| hypothetical protein G ( 150) 341 84.6 2.2e-13  
gi|143324225|gb|EDE33520.1| hypothetical protein G ( 402) 346 86.0 2.2e-13  
gi|142404151|gb|ECX82212.1| hypothetical protein G ( 228) 343 85.2 2.3e-13  
gi|142160570|gb|ECW03887.1| hypothetical protein G ( 512) 347 86.3 2.3e-13  
gi|197053760|gb|ACH25458.1| Sequence 10 from paten ( 428) 346 86.0 2.4e-13  
gi|19887052|gb|AAM01843.1| 5-enolpyruvylshikimate- ( 428) 346 86.0 2.4e-13  
gi|161726853|emb|CAP47297.1| unnamed protein produ ( 428) 346 86.0 2.4e-13  
gi|134378823|gb|EBB01013.1| hypothetical protein G ( 359) 345 85.7 2.4e-13  
gi|63334427|gb|AAY40476.1| 5-enol-pyruvylshikimate ( 444) 346 86.0 2.4e-13  
gi|141709852|gb|ECS86876.1| hypothetical protein G ( 143) 340 84.3 2.5e-13  
gi|141398250|gb|ECR54896.1| hypothetical protein G ( 122) 339 84.1 2.6e-13  
gi|136365694|gb|EBN67728.1| hypothetical protein G ( 231) 342 84.9 2.7e-13  
gi|143017653|gb|EDC21271.1| hypothetical protein G ( 195) 341 84.7 2.7e-13  
gi|142080142|gb|ECV42444.1| hypothetical protein G ( 428) 345 85.8 2.7e-13  
gi|140120371|gb|ECJ67812.1| hypothetical protein G ( 242) 342 85.0 2.8e-13  
gi|139972695|gb|ECI73769.1| hypothetical protein G ( 243) 342 85.0 2.8e-13  
gi|139798315|gb|ECH53759.1| hypothetical protein G ( 310) 343 85.3 2.9e-13  
gi|137897622|gb|EBW62970.1| hypothetical protein G ( 175) 340 84.4 2.9e-13  
gi|134491078|gb|EBB66101.1| hypothetical protein G ( 153) 339 84.2 3.1e-13  
gi|138213505|gb|EBY45565.1| hypothetical protein G ( 190) 340 84.4 3.1e-13  
gi|270512698|gb|ACZ90976.1| 3-phosphoshikimate 1-c ( 425) 344 85.6 3.2e-13

gi|207102480|emb|CAR82020.1| unnamed protein produ ( 425) 344 85.6 3.2e-13  
gi|75702201|gb|ABA21877.1| 3-phosphoshikimate 1-ca ( 426) 344 85.6 3.2e-13  
gi|138071564|gb|EBX58228.1| hypothetical protein G ( 239) 341 84.7 3.2e-13  
gi|71796266|gb|AAZ41017.1| 3-phosphoshikimate 1-ca ( 432) 344 85.6 3.2e-13  
gi|140119422|gb|ECJ67136.1| hypothetical protein G ( 165) 339 84.2 3.2e-13  
gi|2485257|gb|AAB73393.1| I44482 Sequence 62 from p ( 441) 344 85.6 3.3e-13  
gi|144974759|gb|ABP12470.1| Sequence 62 from paten ( 441) 344 85.6 3.3e-13  
gi|5957580|gb|AAE08254.1| Sequence 62 from patent ( 441) 344 85.6 3.3e-13  
gi|2484180|gb|AAB72316.1| I49209 Sequence 62 from p ( 441) 344 85.6 3.3e-13  
gi|581456|emb|CAA78480.1| aroA [Pasteurella multoc ( 441) 344 85.6 3.3e-13  
gi|135434940|gb|EBH61643.1| hypothetical protein G ( 209) 340 84.5 3.3e-13  
gi|136401218|gb|EBN92222.1| hypothetical protein G ( 98) 336 83.4 3.4e-13  
gi|136293903|gb|EBN19081.1| hypothetical protein G ( 193) 339 84.2 3.7e-13  
gi|136662565|gb|EBP58343.1| hypothetical protein G ( 159) 338 83.9 3.7e-13  
gi|139775941|gb|ECH37969.1| hypothetical protein G ( 159) 338 83.9 3.7e-13  
gi|138088359|gb|EBX67891.1| hypothetical protein G ( 241) 340 84.5 3.7e-13  
gi|137102829|gb|EBS21784.1| hypothetical protein G ( 112) 336 83.4 3.8e-13  
gi|198260919|gb|EDY85227.1| 3-phosphoshikimate 1-c ( 446) 343 85.4 3.8e-13  
gi|23326204|gb|AAN24780.1| 3-phosphoshikimate 1-ca ( 455) 343 85.4 3.9e-13  
gi|135035458|gb|EBF07996.1| hypothetical protein G ( 405) 342 85.1 4.1e-13  
gi|136951000|gb|EBR36509.1| hypothetical protein G ( 276) 340 84.6 4.2e-13  
gi|139395595|gb|ECE82264.1| hypothetical protein G ( 230) 339 84.3 4.2e-13  
gi|135635389|gb|EBI89991.1| hypothetical protein G ( 129) 336 83.4 4.2e-13  
gi|91202313|emb|CAJ75373.1| strongly similar to 3- ( 424) 342 85.1 4.3e-13  
gi|139117329|gb|ECD53219.1| hypothetical protein G ( 196) 338 84.0 4.3e-13  
gi|237882706|gb|EEP71534.1| 3-phosphoshikimate 1-c ( 433) 342 85.1 4.3e-13  
gi|135109078|gb|EBF55010.1| hypothetical protein G ( 452) 342 85.2 4.5e-13  
gi|137893657|gb|EBW60674.1| hypothetical protein G ( 95) 334 82.9 4.5e-13  
gi|134841298|gb|EBD78239.1| hypothetical protein G ( 225) 338 84.1 4.8e-13  
gi|138800588|gb|ECB91870.1| hypothetical protein G ( 156) 336 83.5 4.9e-13  
gi|135522927|gb|EBI19183.1| hypothetical protein G ( 129) 335 83.2 4.9e-13  
gi|164564335|gb|ABY61050.1| 5-enolpyruvylshikimate ( 521) 342 85.2 5e-13  
gi|157142998|gb|ABV24481.1| 5-enolpyruvylshikimate ( 521) 342 85.2 5e-13  
gi|193230774|gb|ACF16410.1| 5-enolpyruvylshikimate ( 521) 342 85.2 5e-13  
gi|213504188|emb|CAS92805.1| unnamed protein produ ( 431) 341 84.9 5e-13  
gi|137271986|gb|EBT16720.1| hypothetical protein G ( 202) 337 83.8 5.1e-13  
gi|269305959|gb|ACZ31509.1| 3-phosphoshikimate 1-c ( 444) 341 84.9 5.2e-13  
gi|239515101|gb|EEQ54968.1| 3-phosphoshikimate 1-c ( 445) 341 84.9 5.2e-13  
gi|141879835|gb|ECT89188.1| hypothetical protein G ( 207) 337 83.8 5.2e-13  
gi|137527344|gb|EBU57324.1| hypothetical protein G ( 217) 337 83.8 5.4e-13  
gi|137475361|gb|EBU30443.1| hypothetical protein G ( 121) 334 83.0 5.4e-13  
gi|118193930|gb|ABK76848.1| 5-enolpyruvylshikimate ( 414) 340 84.7 5.7e-13  
gi|141808595|gb|ECT39392.1| hypothetical protein G ( 134) 334 83.0 5.9e-13  
gi|134362434|gb|EBA89872.1| hypothetical protein G ( 163) 335 83.3 5.9e-13  
gi|136288662|gb|EBN15525.1| hypothetical protein G ( 202) 336 83.6 6e-13  
gi|142504418|gb|ECY55742.1| hypothetical protein G ( 249) 337 83.9 6e-13  
gi|139540931|gb|ECF77151.1| hypothetical protein G ( 177) 335 83.3 6.3e-13  
gi|137873963|gb|EBW49399.1| hypothetical protein G ( 262) 337 83.9 6.3e-13  
gi|136799237|gb|EBQ48383.1| hypothetical protein G ( 263) 337 83.9 6.3e-13  
gi|137927027|gb|EBW79646.1| hypothetical protein G ( 150) 334 83.1 6.4e-13  
gi|135280343|gb|EBG60330.1| hypothetical protein G ( 188) 335 83.4 6.6e-13  
gi|137961021|gb|EBW98740.1| hypothetical protein G ( 230) 336 83.6 6.6e-13  
gi|142028027|gb|ECU97300.1| hypothetical protein G ( 420) 339 84.5 6.7e-13  
gi|143735046|gb|EDG51254.1| hypothetical protein G ( 237) 336 83.6 6.7e-13  
gi|140469806|gb|ECL93098.1| hypothetical protein G ( 240) 336 83.6 6.8e-13  
gi|50951940|gb|AAT89641.1| 3-phosphoshikimate 1-ca ( 436) 339 84.5 6.9e-13  
gi|142007150|gb|ECU78011.1| hypothetical protein G ( 248) 336 83.7 7e-13  
gi|144197942|gb|EDJ58366.1| hypothetical protein G ( 482) 339 84.5 7.4e-13  
gi|137268762|gb|EBT14925.1| hypothetical protein G ( 187) 334 83.1 7.6e-13  
gi|68264311|emb|CAI37799.1| 3-phosphoshikimate 1-c ( 429) 338 84.3 7.9e-13  
gi|213524060|gb|ACJ52807.1| 3-phosphoshikimate 1-c ( 461) 338 84.3 8.3e-13  
gi|140758228|gb|ECN31786.1| hypothetical protein G ( 220) 334 83.2 8.6e-13

gi|135852837|gb|EBK25457.1| hypothetical protein G ( 226) 334 83.2 8.8e-13  
gi|136645114|gb|EBP47908.1| hypothetical protein G ( 132) 331 82.4 9.1e-13  
gi|6855375|emb|CAB71266.1| 3-phosphoshikimate 1-ca ( 440) 337 84.1 9.4e-13  
gi|197697271|gb|EDY44204.1| 3-phosphoshikimate 1-c ( 443) 337 84.1 9.4e-13  
gi|137113514|gb|EBS27820.1| hypothetical protein G ( 301) 335 83.5 9.4e-13  
gi|137606007|gb|EBV01255.1| hypothetical protein G ( 208) 333 82.9 9.6e-13  
gi|137020450|gb|EBR75785.1| hypothetical protein G ( 120) 330 82.1 9.9e-13  
gi|140674336|gb|ECM73229.1| hypothetical protein G ( 220) 333 83.0 1e-12  
gi|139752541|gb|ECH2814.1| hypothetical protein G ( 182) 332 82.7 1e-12  
gi|134412896|gb|EBB20265.1| hypothetical protein G ( 152) 331 82.4 1e-12  
gi|139695641|gb|ECG83504.1| hypothetical protein G ( 235) 333 83.0 1.1e-12  
gi|137065670|gb|ECC58151.1| hypothetical protein G ( 134) 330 82.2 1.1e-12  
gi|142221795|gb|ECW50550.1| hypothetical protein G ( 356) 335 83.6 1.1e-12  
gi|142176534|gb|ECW16103.1| hypothetical protein G ( 441) 336 83.8 1.1e-12  
gi|134510774|gb|EBB77501.1| hypothetical protein G ( 120) 329 81.9 1.1e-12  
gi|136108377|gb|EBL96850.1| hypothetical protein G ( 216) 332 82.7 1.1e-12  
gi|138908492|gb|ECC37782.1| hypothetical protein G ( 123) 329 81.9 1.2e-12  
gi|143051684|gb|EDC46069.1| hypothetical protein G ( 151) 330 82.2 1.2e-12  
gi|136377067|gb|EBN75527.1| hypothetical protein G ( 152) 330 82.2 1.2e-12  
gi|140538314|gb|ECM15202.1| hypothetical protein G ( 153) 330 82.2 1.2e-12  
gi|171341571|dbj|BAB76718.1| 3-phosphoshikimate 1-c ( 425) 333 83.6 1.2e-12  
gi|138685422|gb|ECW29911.1| hypothetical protein G ( 288) 333 83.1 1.2e-12  
gi|137249529|gb|EBT04003.1| hypothetical protein G ( 132) 329 81.9 1.2e-12  
gi|141561787|gb|EBB73611.1| hypothetical protein G ( 196) 331 82.5 1.2e-12  
gi|213504190|emb|CAS92806.1| unnamed protein produ ( 431) 335 83.6 1.2e-12  
gi|218301288|emb|CAU98636.1| unnamed protein produ ( 431) 335 83.6 1.2e-12  
gi|141357056|gb|ECR26055.1| hypothetical protein G ( 198) 331 82.5 1.2e-12  
gi|138599473|gb|ECA71366.1| hypothetical protein G ( 200) 331 82.5 1.3e-12  
gi|139957029|gb|ECI62786.1| hypothetical protein G ( 165) 330 82.2 1.3e-12  
gi|143120591|gb|ECB96442.1| hypothetical protein G ( 210) 331 82.5 1.3e-12  
gi|135966631|gb|EBL02759.1| hypothetical protein G ( 214) 331 82.5 1.3e-12  
gi|139776743|gb|ECH38540.1| hypothetical protein G ( 236) 331 82.6 1.4e-12  
gi|135013362|gb|EBE93072.1| hypothetical protein G ( 165) 329 82.0 1.5e-12  
gi|139543335|gb|ECF78883.1| hypothetical protein G ( 141) 328 81.7 1.5e-12  
gi|140194355|gb|ECU16491.1| hypothetical protein G ( 142) 328 81.7 1.5e-12  
gi|141975410|gb|ECU55912.1| hypothetical protein G ( 107) 326 81.2 1.6e-12  
gi|140674338|gb|ECM73231.1| hypothetical protein G ( 236) 330 82.3 1.7e-12  
gi|137625100|gb|EBV12215.1| hypothetical protein G ( 161) 328 81.8 1.7e-12  
gi|142806346|gb|EDA72212.1| hypothetical protein G ( 240) 330 82.3 1.7e-12  
gi|269786272|gb|ACB38415.1| 3-phosphoshikimate 1-c ( 435) 333 83.2 1.7e-12  
gi|51856246|dbj|BAD40404.1| 3-phosphoshikimate-1-c ( 449) 333 83.2 1.7e-12  
gi|136051934|gb|EBL58617.1| hypothetical protein G ( 207) 329 82.1 1.7e-12  
gi|140878389|gb|ECO12061.1| hypothetical protein G ( 211) 329 82.1 1.8e-12  
gi|137601619|gb|EBU98761.1| hypothetical protein G ( 146) 327 81.5 1.8e-12  
gi|142981403|gb|EDB95732.1| hypothetical protein G ( 122) 326 81.2 1.8e-12  
gi|142914613|gb|EDB48485.1| hypothetical protein G ( 106) 325 81.0 1.9e-12  
gi|134820000|gb|EBD64619.1| hypothetical protein G ( 130) 326 81.3 1.9e-12  
gi|139913576|gb|ECI32833.1| hypothetical protein G ( 214) 328 81.9 2.1e-12  
gi|137698489|gb|EBV52072.1| hypothetical protein G ( 151) 326 81.3 2.2e-12  
gi|143904377|gb|EDH46982.1| hypothetical protein G ( 343) 330 82.5 2.2e-12  
gi|140309628|gb|ECK88544.1| hypothetical protein G ( 98) 323 80.5 2.4e-12  
gi|140182762|gb|ECK08451.1| hypothetical protein G ( 184) 326 81.4 2.5e-12  
gi|141954027|gb|ECU41298.1| hypothetical protein G ( 225) 327 81.7 2.5e-12  
gi|53854589|gb|AAU95677.1| arom [Sclerotinia scler (1590) 337 84.5 2.5e-12  
gi|135943313|gb|EBK87033.1| hypothetical protein G ( 190) 326 81.4 2.6e-12  
gi|134380743|gb|EBB02262.1| hypothetical protein G ( 200) 326 81.4 2.7e-12  
gi|136357207|gb|EBN61951.1| hypothetical protein G ( 204) 326 81.4 2.7e-12  
gi|137887915|gb|EBW57387.1| hypothetical protein G ( 168) 325 81.1 2.7e-12  
gi|134702029|gb|EBC90411.1| hypothetical protein G ( 169) 325 81.1 2.7e-12  
gi|167288349|gb|ABZ41213.1| Sequence 15151 from pa ( 400) 329 82.3 2.9e-12  
gi|167284960|gb|ABZ37824.1| Sequence 11762 from pa ( 410) 329 82.3 3e-12  
gi|167287866|gb|ABZ40730.1| Sequence 14668 from pa ( 410) 329 82.3 3e-12



gi|167287577|gb|ABZ40441.1| Sequence 14379 from pa ( 410) 329 82.3 3e-12  
gi|138839332|gb|ECC10256.1| hypothetical protein G ( 155) 324 80.9 3e-12  
gi|138510310|gb|EACA09339.1| hypothetical protein G ( 160) 324 80.9 3e-12  
gi|15155591|gb|AAK86449.1| 3-phosphoshikimate 1-ca ( 425) 329 82.3 3e-12  
gi|209409579|emb|CAR82016.1| unnamed protein produ ( 425) 329 82.3 3e-12  
gi|141666967|gb|ECS70781.1| hypothetical protein G ( 238) 326 81.5 3.1e-12  
gi|138006824|gb|EBX24048.1| hypothetical protein G ( 164) 324 80.9 3.1e-12  
gi|237878972|gb|ACR31304.1| 5-enolpyruvylshikimate ( 451) 329 82.3 3.2e-12  
gi|142858079|gb|EDB07673.1| hypothetical protein G ( 141) 323 80.6 3.2e-12  
gi|136532719|gb|EB077403.1| hypothetical protein G ( 375) 328 82.1 3.2e-12  
gi|138288445|gb|EBY80872.1| hypothetical protein G ( 143) 323 80.6 3.2e-12  
gi|141201113|gb|ECQ31813.1| hypothetical protein G ( 184) 324 80.9 3.4e-12  
gi|119693519|gb|ABL90592.1| 3-phosphoshikimate 1-c ( 438) 328 82.1 3.6e-12  
gi|108768752|gb|ABG07474.1| 3-phosphoshikimate 1-c ( 438) 328 82.1 3.6e-12  
gi|139751825|gb|ECD61537.1| hypothetical protein G ( 201) 324 81.0 3.6e-12  
gi|142196767|gb|ECW31538.1| hypothetical protein G ( 325) 326 81.6 3.9e-12  
gi|135506999|gb|EBI08954.1| hypothetical protein G ( 220) 324 81.0 3.9e-12  
gi|117576131|emb|CAL68900.1| unnamed protein produ ( 419) 327 81.9 4.1e-12  
gi|239679344|gb|ACS07594.1| Sequence 9 from patent ( 419) 327 81.9 4.1e-12  
gi|135220220|gb|EBG24982.1| hypothetical protein G ( 234) 324 81.0 4.1e-12  
gi|94220102|gb|ABF14261.1| 3-phosphoshikimate 1-ca ( 428) 327 81.9 4.1e-12  
gi|134549606|gb|EBC00642.1| hypothetical protein G ( 111) 320 79.9 4.2e-12  
gi|144123245|gb|EDJ03229.1| hypothetical protein G ( 438) 327 81.9 4.2e-12  
gi|139137061|gb|ECP65571.1| hypothetical protein G ( 114) 320 79.9 4.3e-12  
gi|137815339|gb|EBW15583.1| hypothetical protein G ( 172) 322 80.5 4.4e-12  
gi|137007196|gb|EBR68377.1| hypothetical protein G ( 98) 319 79.7 4.4e-12  
gi|154698953|gb|EDN98691.1| hypothetical protein S (1576) 333 83.6 4.6e-12  
gi|126233798|gb|ABN97198.1| 3-phosphoshikimate 1-c ( 438) 326 81.7 4.9e-12  
gi|137439907|gb|EBU11139.1| hypothetical protein G ( 139) 320 80.0 5e-12  
gi|135301369|gb|EBG72638.1| hypothetical protein G ( 142) 320 80.0 5.1e-12  
gi|138189835|gb|EBY30590.1| hypothetical protein G ( 213) 322 80.6 5.1e-12  
gi|134789146|gb|EBD44416.1| hypothetical protein G ( 147) 320 80.0 5.2e-12  
gi|141436523|gb|ECR81656.1| hypothetical protein G ( 179) 321 80.3 5.2e-12  
gi|139445388|gb|ECP14877.1| hypothetical protein G ( 219) 322 80.6 5.2e-12  
gi|140647368|gb|ECM54294.1| hypothetical protein G ( 191) 321 80.3 5.5e-12  
gi|218086784|emb|CAT03395.1| unnamed protein produ ( 419) 325 81.4 5.5e-12  
gi|141076390|gb|ECP45929.1| hypothetical protein G ( 195) 321 80.3 5.6e-12  
gi|135798487|gb|EBJ91258.1| hypothetical protein G ( 132) 319 79.8 5.6e-12  
gi|117561405|gb|ABK38353.1| 3-phosphoshikimate 1-c ( 455) 325 81.5 5.9e-12  
gi|141784517|gb|ECT27521.1| hypothetical protein G ( 218) 321 80.4 6.1e-12  
gi|137434272|gb|EBU08031.1| hypothetical protein G ( 231) 321 80.4 6.4e-12  
gi|138267880|gb|EBY68978.1| hypothetical protein G ( 130) 318 79.5 6.4e-12  
gi|137514998|gb|EBU50711.1| hypothetical protein G ( 194) 320 80.1 6.5e-12  
gi|135767474|gb|EBJ71737.1| hypothetical protein G ( 237) 321 80.4 6.5e-12  
gi|137551557|gb|EBU70622.1| hypothetical protein G ( 135) 318 79.5 6.6e-12  
gi|134678975|gb|EBC77637.1| hypothetical protein G ( 200) 320 80.1 6.6e-12  
gi|136051935|gb|EBL58618.1| hypothetical protein G ( 206) 320 80.1 6.8e-12  
gi|213504180|emb|CAS92801.1| unnamed protein produ ( 467) 324 81.3 7e-12  
gi|140278343|gb|ECK71234.1| hypothetical protein G ( 145) 318 79.6 7e-12  
gi|135848766|gb|EBK22892.1| hypothetical protein G ( 148) 318 79.6 7.1e-12  
gi|76782132|gb|ABA54838.1| 3-phosphoshikimate 1-ca ( 403) 323 81.0 7.2e-12  
gi|143489039|gb|EDF26103.1| hypothetical protein G ( 153) 318 79.6 7.3e-12  
gi|239679345|gb|ACS07595.1| Sequence 11 from paten ( 419) 323 81.0 7.5e-12  
gi|117576125|emb|CAL68897.1| unnamed protein produ ( 419) 323 81.0 7.5e-12  
gi|239679343|gb|ACS07593.1| Sequence 7 from patent ( 419) 323 81.0 7.5e-12  
gi|209571207|emb|CAR82018.1| unnamed protein produ ( 419) 323 81.0 7.5e-12  
gi|117576123|emb|CAL68666.1| unnamed protein produ ( 419) 323 81.0 7.5e-12  
gi|239679341|gb|ACS07591.1| Sequence 3 from patent ( 419) 323 81.0 7.5e-12  
gi|117576129|emb|CAL68899.1| unnamed protein produ ( 419) 323 81.0 7.5e-12  
gi|223548872|gb|EEF50361.1| 3-phosphoshikimate 1-c ( 518) 324 81.3 7.5e-12  
gi|140945347|gb|ECO57367.1| hypothetical protein G ( 198) 319 79.9 7.6e-12  
gi|135582507|gb|EBI57266.1| hypothetical protein G ( 202) 319 79.9 7.8e-12

gi|137183149|gb|EBS66811.1| hypothetical protein G ( 140) 317 79.3 7.9e-12  
gi|117576145|emb|CAL68907.1| unnamed protein produ ( 419) 322 80.8 8.7e-12  
gi|239679351|gb|ACS07601.1| Sequence 23 from paten ( 419) 322 80.8 8.7e-12  
gi|226911720|gb|EEH96921.1| 3-phosphoshikimate 1-c ( 431) 322 80.8 8.9e-12  
gi|134740074|gb|EBD13133.1| hypothetical protein G ( 240) 319 79.9 8.9e-12  
gi|218301271|emb|CAU98620.1| unnamed protein produ ( 442) 322 80.8 9e-12  
gi|126219902|gb|ABN83408.1| 5-enolpyruvylshikimate ( 448) 322 80.8 9.1e-12  
gi|27361607|gb|AAO10514.1|AE016804\_24 3-phosphoshi ( 376) 321 80.5 9.3e-12  
gi|139300510|gb|ECE48466.1| hypothetical protein G ( 180) 317 79.4 9.6e-12  
gi|140922142|gb|ECO40954.1| hypothetical protein G ( 127) 315 78.9 9.9e-12  
gi|239679347|gb|ACS07597.1| Sequence 15 from paten ( 419) 321 80.6 1e-11  
gi|117576143|emb|CAL68906.1| unnamed protein produ ( 419) 321 80.6 1e-11  
gi|117576137|emb|CAL68903.1| unnamed protein produ ( 419) 321 80.6 1e-11  
gi|239679350|gb|ACS07600.1| Sequence 21 from paten ( 419) 321 80.6 1e-11  
gi|140050235|gb|ECG24530.1| hypothetical protein G ( 107) 314 78.6 1e-11  
gi|141531454|gb|ECS17387.1| hypothetical protein G ( 162) 316 79.2 1e-11  
gi|139125726|gb|ECD58837.1| hypothetical protein G ( 114) 314 78.6 1.1e-11  
gi|141358068|gb|ECE26755.1| hypothetical protein G ( 172) 316 79.2 1.1e-11  
gi|138538608|gb|ECA28773.1| hypothetical protein G ( 174) 316 79.2 1.1e-11  
gi|139459143|gb|ECF24709.1| hypothetical protein G ( 179) 316 79.2 1.1e-11  
gi|139512332|gb|ECF57947.1| hypothetical protein G ( 122) 314 78.6 1.1e-11  
gi|141571591|gb|ECS42100.1| hypothetical protein G ( 124) 314 78.6 1.1e-11  
gi|143612207|gb|EDF86144.1| hypothetical protein G ( 334) 319 80.1 1.1e-11  
gi|143683392|gb|EDG25520.1| hypothetical protein G ( 155) 315 78.9 1.2e-11  
gi|182634487|gb|ACB95261.1| 3-phosphoshikimate 1-c ( 424) 320 80.4 1.2e-11  
gi|281079464|gb|ADA36172.1| Sequence 5 from patent ( 525) 321 80.6 1.2e-11  
gi|134974650|gb|EBE66821.1| hypothetical protein G ( 198) 316 79.2 1.2e-11  
gi|137082794|gb|EBE10909.1| hypothetical protein G ( 242) 317 79.5 1.2e-11  
gi|140242114|gb|ECK49397.1| hypothetical protein G ( 201) 316 79.2 1.2e-11  
gi|136310120|gb|EBN30164.1| hypothetical protein G ( 390) 319 80.1 1.3e-11  
gi|135819476|gb|EBK04465.1| hypothetical protein G ( 182) 315 79.0 1.3e-11  
gi|139753549|gb|ECH22504.1| hypothetical protein G ( 227) 316 79.3 1.3e-11  
gi|138575184|gb|ECA54711.1| hypothetical protein G ( 193) 315 79.0 1.4e-11  
gi|229383027|gb|EEO33118.1| 3-phosphoshikimate 1-c ( 426) 319 80.1 1.4e-11  
gi|141796891|gb|ECT33585.1| hypothetical protein G ( 161) 314 78.7 1.4e-11  
gi|157918462|gb|ABV99889.1| 3-phosphoshikimate 1-c ( 433) 319 80.1 1.4e-11  
gi|136047742|gb|EBL55766.1| hypothetical protein G ( 199) 315 79.0 1.4e-11  
gi|143086208|gb|EDC71322.1| hypothetical protein G ( 200) 315 79.0 1.4e-11  
gi|138033807|gb|EBX38001.1| hypothetical protein G ( 97) 311 77.9 1.5e-11  
gi|136446530|gb|EBO021826.1| hypothetical protein G ( 187) 314 78.8 1.6e-11  
gi|268312873|gb|ACV99247.1| 3-phosphoshikimate 1-c ( 425) 318 79.9 1.6e-11  
gi|270259770|emb|CBI38903.1| unnamed protein produ ( 521) 319 80.2 1.6e-11  
gi|137251448|gb|EBT05121.1| hypothetical protein G ( 198) 314 78.8 1.6e-11  
gi|140795564|gb|ECN50407.1| hypothetical protein G ( 216) 314 78.8 1.7e-11  
gi|135696052|gb|EBJ27471.1| hypothetical protein G ( 190) 313 78.6 1.8e-11  
gi|141070657|gb|ECP42029.1| hypothetical protein G ( 157) 312 78.3 1.8e-11  
gi|117576135|emb|CAL68902.1| unnamed protein produ ( 419) 317 79.7 1.8e-11  
gi|239679346|gb|ACS07596.1| Sequence 13 from paten ( 419) 317 79.7 1.8e-11  
gi|158510786|gb|ABW67753.1| 3-phosphoshikimate 1-c ( 423) 317 79.7 1.9e-11  
gi|257050668|gb|ACV39852.1| 3-phosphoshikimate 1-c ( 431) 317 79.7 1.9e-11  
gi|139909706|gb|ECI30075.1| hypothetical protein G ( 205) 313 78.6 1.9e-11  
gi|143157320|gb|EDD23165.1| hypothetical protein G ( 205) 313 78.6 1.9e-11  
gi|134843697|gb|EBN79778.1| hypothetical protein G ( 122) 310 77.8 2e-11  
gi|143344597|gb|EDE46223.1| hypothetical protein G ( 328) 315 79.2 2.1e-11  
gi|137183088|gb|EBS66775.1| hypothetical protein G ( 187) 312 78.3 2.1e-11  
gi|160339357|gb|ABX12444.1| 3-phosphoshikimate 1-c ( 422) 316 79.9 2.2e-11  
gi|135111849|gb|EBF56773.1| hypothetical protein G ( 294) 314 78.9 2.2e-11  
gi|135933536|gb|EBR80451.1| hypothetical protein G ( 168) 311 78.1 2.2e-11  
gi|139308929|gb|ECE49564.1| hypothetical protein G ( 120) 309 77.5 2.3e-11  
gi|138324259|gb|EBY97937.1| hypothetical protein G ( 218) 312 78.4 2.4e-11  
gi|135162902|gb|EBF89584.1| hypothetical protein G ( 124) 309 77.6 2.4e-11  
gi|142836951|gb|EDA95298.1| hypothetical protein G ( 184) 311 78.1 2.4e-11

gi|117576133|emb|CAL68901.1| unnamed protein produ ( 419) 315 79.3 2.5e-11  
gi|138052107|gb|EBX47852.1| hypothetical protein G ( 130) 309 77.6 2.5e-11  
gi|259027640|gb|ACV91084.1| 5-enolpyruvylshikimate ( 427) 315 79.3 2.5e-11  
gi|143437003|gb|EDE94972.1| hypothetical protein G ( 635) 317 79.8 2.5e-11  
gi|137533283|gb|EBU60530.1| hypothetical protein G ( 162) 310 77.9 2.5e-11  
gi|121052443|emb|CAM08779.1| 5-enolpyruvylshikimate ( 433) 315 79.3 2.6e-11  
gi|184212201|gb|EDU09244.1| putative 3-phosphoshik ( 448) 315 79.3 2.6e-11  
gi|143355421|gb|EDE52319.1| hypothetical protein G ( 448) 315 79.3 2.6e-11  
gi|137996781|gb|EBX18771.1| hypothetical protein G ( 210) 311 78.2 2.7e-11  
gi|141116592|gb|ECP74054.1| hypothetical protein G ( 147) 309 77.6 2.7e-11  
gi|134917544|gb|EBE28661.1| hypothetical protein G ( 180) 310 77.9 2.8e-11  
gi|145689056|gb|ABP89562.1| 5-enolpyruvylshikimate ( 101) 307 77.1 2.8e-11  
gi|140659905|gb|ECM62959.1| hypothetical protein G ( 124) 308 77.3 2.8e-11  
gi|140082999|gb|ECJ42830.1| hypothetical protein G ( 191) 310 77.9 2.9e-11  
gi|239679349|gb|ACS07599.1| Sequence 19 from paten ( 419) 314 79.0 2.9e-11  
gi|117576127|emb|CAL68898.1| unnamed protein produ ( 419) 314 79.0 2.9e-11  
gi|239679342|gb|ACS07592.1| Sequence 5 from patent ( 419) 314 79.0 2.9e-11  
gi|117576141|emb|CAL68905.1| unnamed protein produ ( 419) 314 79.0 2.9e-11  
gi|222843005|gb|EEB80552.1| 3-phosphoshikimate 1-c ( 518) 315 79.3 2.9e-11  
gi|261392307|emb|CAX49833.1| 3-phosphoshikimate 1- ( 433) 314 79.1 3e-11  
gi|135491966|gb|EBH99301.1| hypothetical protein G ( 112) 307 77.1 3e-11  
gi|141725760|gb|ECS96726.1| hypothetical protein G ( 144) 308 77.4 3.1e-11  
gi|169653664|gb|EDS86357.1| putative 3-phosphoshik ( 469) 314 79.1 3.2e-11  
gi|135594207|gb|EBI64500.1| hypothetical protein G ( 99) 306 76.8 3.2e-11  
gi|134753000|gb|EBD20520.1| hypothetical protein G ( 180) 309 77.7 3.2e-11  
gi|137658590|gb|EBV29424.1| hypothetical protein G ( 101) 306 76.8 3.2e-11  
gi|141920457|gb|ECU17551.1| hypothetical protein G ( 221) 310 78.0 3.2e-11  
gi|141515914|gb|ECS12618.1| hypothetical protein G ( 194) 309 77.7 3.4e-11  
gi|257228991|gb|ACV53022.1| 5-enolpyruvylshikimate ( 518) 314 79.1 3.4e-11  
gi|109626980|emb|CAJ53455.1| 3-phosphoshikimate 1- ( 432) 313 78.8 3.4e-11  
gi|138733009|gb|ECB63288.1| hypothetical protein G ( 135) 307 77.1 3.5e-11  
gi|135882330|gb|EBK45024.1| hypothetical protein G ( 201) 309 77.7 3.5e-11  
gi|138717870|gb|ECB52618.1| hypothetical protein G ( 311) 311 78.3 3.6e-11  
gi|143243535|gb|EDD84840.1| hypothetical protein G ( 177) 308 77.5 3.7e-11  
gi|137453946|gb|EBU19127.1| hypothetical protein G ( 182) 308 77.5 3.8e-11  
gi|161595853|gb|ABX713513.1| 5-enolpyruvylshikimate ( 399) 312 78.6 3.8e-11  
gi|207102484|emb|CAR82022.1| unnamed protein produ ( 418) 312 78.6 3.9e-11  
gi|71556208|gb|AAZ35419.1| 3-phosphoshikimate 1-ca ( 418) 312 78.6 3.9e-11  
gi|229362102|emb|CAY49004.1| putative 3-phosphoshik ( 425) 312 78.6 4e-11  
gi|136508745|gb|EBO62095.1| hypothetical protein G ( 195) 308 77.5 4e-11  
gi|141182843|gb|ECQ20365.1| hypothetical protein G ( 134) 306 76.9 4e-11  
gi|136050609|gb|EBL57707.1| hypothetical protein G ( 389) 311 78.4 4.3e-11  
gi|140269326|gb|ECK68180.1| hypothetical protein G ( 123) 305 76.7 4.4e-11  
gi|137751638|gb|EBV80445.1| hypothetical protein G ( 131) 305 76.7 4.6e-11  
gi|140226600|gb|ECK38548.1| hypothetical protein G ( 236) 308 77.5 4.6e-11  
gi|257792900|gb|ACV67278.1| 5-enolpyruvylshikimate ( 518) 312 78.7 4.6e-11  
gi|508779001|emb|CAG37740.1| related to 3-phosphoshik ( 451) 311 78.4 4.8e-11  
gi|270259768|emb|CBI38901.1| unnamed protein produ ( 463) 311 78.4 4.9e-11  
gi|139731566|gb|ECH08565.1| hypothetical protein G ( 180) 306 77.0 5e-11  
gi|137485375|gb|EBU35622.1| hypothetical protein G ( 102) 303 76.2 5.1e-11  
gi|140064590|gb|ECJ31471.1| hypothetical protein G ( 184) 306 77.0 5.1e-11  
gi|134392972|gb|EBB09817.1| hypothetical protein G ( 185) 306 77.0 5.1e-11  
gi|207102476|emb|CAR82015.1| unnamed protein produ ( 418) 310 78.2 5.3e-11  
gi|28851495|gb|AAO54572.1| 3-phosphoshikimate 1-ca ( 418) 310 78.2 5.3e-11  
gi|134385480|gb|EBB05348.1| hypothetical protein G ( 107) 303 76.2 5.3e-11  
gi|138081070|gb|EBX63704.1| hypothetical protein G ( 235) 307 77.3 5.3e-11  
gi|142545689|gb|ECY85744.1| hypothetical protein G ( 188) 305 76.8 6.1e-11  
gi|218301285|emb|CAU98633.1| unnamed protein produ ( 412) 309 77.9 6.1e-11  
gi|137813443|gb|EBW14469.1| hypothetical protein G ( 191) 305 76.8 6.1e-11  
gi|207102488|emb|CAR82024.1| unnamed protein produ ( 418) 309 78.0 6.1e-11  
gi|221722512|gb|ACM25668.1| 3-phosphoshikimate 1-c ( 418) 309 78.0 6.1e-11  
gi|63254850|gb|AAY35946.1| 3-phosphoshikimate 1-ca ( 418) 309 78.0 6.1e-11

gi|207102482|emb|CAR82021.1| unnamed protein produ ( 418) 309 78.0 6.1e-11  
gi|170229|gb|AAA34071.1| 5-enolpyruvylshikimate-3- ( 518) 310 78.2 6.2e-11  
gi|137272029|gb|EBT16745.1| hypothetical protein G ( 162) 304 76.6 6.3e-11  
gi|226184786|dbj|BAH32890.1| 3-phosphoshikimate 1- ( 437) 309 78.0 6.4e-11  
gi|134457901|gb|EBB46412.1| hypothetical protein G ( 201) 305 76.8 6.4e-11  
gi|142078992|gb|ECV41483.1| hypothetical protein G ( 439) 309 78.0 6.4e-11  
gi|117576139|emb|CAL68904.1| unnamed protein produ ( 419) 308 77.7 7.2e-11  
gi|239679348|gb|ACS07598.1| Sequence 17 from paten ( 419) 308 77.7 7.2e-11  
gi|186464230|gb|ACC80031.1| 3-phosphoshikimate 1-c ( 426) 308 77.7 7.3e-11  
gi|257792898|gb|ACV67277.1| 5-enolpyruvylshikimate ( 518) 309 78.0 7.3e-11  
gi|143081824|gb|EDC68127.1| hypothetical protein G ( 196) 304 76.6 7.3e-11  
gi|139814833|gb|ECH65564.1| hypothetical protein G ( 196) 304 76.6 7.3e-11  
gi|138014336|gb|EBX28033.1| hypothetical protein G ( 294) 306 77.2 7.4e-11  
gi|118765287|dbj|BAF39466.1| 3-phosphoshikimate 1- ( 450) 308 77.8 7.6e-11  
gi|137819489|gb|EBW18017.1| hypothetical protein G ( 170) 303 76.4 7.6e-11  
gi|138151161|gb|EBY03576.1| hypothetical protein G ( 118) 301 75.8 7.7e-11  
gi|140182535|gb|ECK08289.1| hypothetical protein G ( 148) 302 76.1 7.9e-11  
gi|169157002|emb|CAQ02172.1| 3-phosphoshikimate 1- ( 487) 308 77.8 8e-11  
gi|209571209|emb|CAR82019.1| unnamed protein produ ( 412) 307 77.5 8.2e-11  
gi|209571213|emb|CAR82026.1| unnamed protein produ ( 412) 307 77.5 8.2e-11  
gi|136360811|gb|EBN64427.1| hypothetical protein G ( 131) 301 75.8 8.4e-11  
gi|189170087|gb|ACD80082.1| 5-enolpyruvylshikimate ( 520) 308 77.8 8.5e-11  
gi|209571211|emb|CAR82025.1| unnamed protein produ ( 444) 307 77.5 8.7e-11  
gi|138697010|gb|ECB38164.1| hypothetical protein G ( 208) 303 76.4 8.9e-11  
gi|134676963|gb|EBC76457.1| hypothetical protein G ( 179) 302 76.2 9.2e-11  
gi|170374|gb|AAA34136.1| 5-enolpyruvylshikimate-3- ( 520) 307 77.6 9.8e-11  
gi|135603657|gb|EBT170337.1| hypothetical protein G ( 197) 302 76.2 9.9e-11  
gi|226512689|gb|EEH62034.1| 3-phosphoshikimate 1-c ( 433) 306 77.3 9.9e-11  
gi|268624509|gb|EEZ56909.1| 3-phosphoshikimate 1-c ( 433) 306 77.3 9.9e-11  
gi|256559915|gb|ACU85762.1| 3-phosphoshikimate 1-c ( 441) 306 77.3 1e-10  
gi|136066746|gb|EBL68557.1| hypothetical protein G ( 177) 301 75.9 1.1e-10  
gi|141993398|gb|ECU68756.1| hypothetical protein G ( 178) 301 75.9 1.1e-10  
gi|136291221|gb|EBN17255.1| hypothetical protein G ( 181) 301 75.9 1.1e-10  
gi|143277548|gb|EDE08112.1| hypothetical protein G ( 149) 300 75.7 1.1e-10  
gi|135811094|gb|EBJ99204.1| hypothetical protein G ( 153) 300 75.7 1.1e-10  
gi|167293329|gb|ABZ46193.1| Sequence 20131 from pa ( 411) 305 77.1 1.1e-10  
gi|144119343|gb|EDU00333.1| hypothetical protein G ( 133) 299 75.4 1.1e-10  
gi|268622298|gb|EEZ54698.1| 3-phosphoshikimate 1-c ( 433) 305 77.1 1.2e-10  
gi|7226672|gb|AAF41793.1| 3-phosphoshikimate 1-car ( 433) 305 77.1 1.2e-10  
gi|268550754|gb|EEZ54773.1| 3-phosphoshikimate 1-c ( 433) 305 77.1 1.2e-10  
gi|268587665|gb|EEZ52341.1| 3-phosphoshikimate 1-c ( 433) 305 77.1 1.2e-10  
gi|268585344|gb|EEZ50020.1| 3-phosphoshikimate 1-c ( 433) 305 77.1 1.2e-10  
gi|268548093|gb|EEZ43511.1| 3-phosphoshikimate 1-c ( 433) 305 77.1 1.2e-10  
gi|138737167|gb|ECB6241.1| hypothetical protein G ( 206) 301 76.0 1.2e-10  
gi|134471594|gb|EBB54403.1| hypothetical protein G ( 209) 301 76.0 1.2e-10  
gi|139521188|gb|ECF64171.1| hypothetical protein G ( 155) 299 75.5 1.3e-10  
gi|141594961|gb|ECS50891.1| hypothetical protein G ( 194) 300 75.7 1.3e-10  
gi|257228989|gb|ACV53021.1| 5-enolpyruvylshikimate ( 518) 305 77.2 1.3e-10  
gi|557407691|gb|AAV64030.1| 5-enolpyruvylshikimate ( 519) 305 77.2 1.3e-10  
gi|138262361|gb|EBY65197.1| hypothetical protein G ( 166) 299 75.5 1.4e-10  
gi|151359694|gb|ABS02697.1| 3-phosphoshikimate 1-c ( 449) 304 76.9 1.4e-10  
gi|136405337|gb|EBN95052.1| hypothetical protein G ( 97) 296 74.6 1.4e-10  
gi|137396650|gb|EBT86710.1| hypothetical protein G ( 129) 297 75.0 1.5e-10  
gi|247545418|gb|ACT02437.1| 3-phosphoshikimate 1-c ( 430) 303 76.7 1.6e-10  
gi|254668754|emb|CBA06624.1| 3-phosphoshikimate 1- ( 433) 303 76.7 1.6e-10  
gi|254672959|emb|CBA07392.1| 3-phosphoshikimate 1- ( 433) 303 76.7 1.6e-10  
gi|139642795|gb|ECG46466.1| hypothetical protein G ( 199) 299 75.5 1.6e-10  
gi|207102486|emb|CAR82023.1| unnamed protein produ ( 444) 303 76.7 1.6e-10  
gi|120400573|gb|ABM21481.1| 5'-enolpyruvylshikimate ( 444) 303 76.7 1.6e-10  
gi|157272154|gb|ABV26712.1| 5-enolpyruvylshikimate ( 444) 303 76.7 1.6e-10  
gi|135849858|gb|EBK23583.1| hypothetical protein G ( 95) 295 74.4 1.6e-10  
gi|136430441|gb|EBB11415.1| hypothetical protein G ( 256) 300 75.8 1.6e-10

gi|138011315|gb|EBX26442.1| hypothetical protein G ( 174) 298 75.3 1.6e-10  
gi|135898933|gb|EBK56669.1| hypothetical protein G ( 181) 298 75.3 1.7e-10  
gi|135680289|gb|EBJ17747.1| hypothetical protein G ( 126) 296 74.7 1.7e-10  
gi|268582986|gb|EEZ47662.1| 3-phosphoshikimate 1-c ( 433) 302 76.4 1.8e-10  
gi|268626765|gb|EEZ59165.1| 3-phosphoshikimate 1-c ( 433) 302 76.4 1.8e-10  
gi|137446725|gb|EBU15001.1| hypothetical protein G ( 115) 295 74.5 1.9e-10  
gi|141787258|gb|ECT28985.1| hypothetical protein G ( 140) 296 74.8 1.9e-10  
gi|134136777|gb|ABO57891.1| 3-phosphoshikimate 1-c ( 460) 302 76.5 1.9e-10  
gi|197628303|gb|EDY40847.1| 3-phosphoshikimate 1-c ( 403) 301 76.2 2e-10  
gi|138502658|gb|ECA05423.1| hypothetical protein G ( 152) 296 74.8 2e-10  
gi|197710309|gb|EDY54343.1| 3-phosphoshikimate 1-c ( 415) 301 76.2 2e-10  
gi|254670251|emb|CBA05487.1| 3-phosphoshikimate 1- ( 433) 301 76.2 2.1e-10  
gi|134977338|gb|EBE68641.1| hypothetical protein G ( 114) 294 74.3 2.2e-10  
gi|134357023|gb|EBA86222.1| hypothetical protein G ( 126) 294 74.3 2.3e-10  
gi|14302313|gb|EDC25299.1| hypothetical protein G ( 131) 294 74.3 2.4e-10  
gi|139859200|gb|ECH96267.1| hypothetical protein G ( 195) 296 74.9 2.4e-10  
gi|142130241|gb|ECV81111.1| hypothetical protein G ( 133) 294 74.3 2.4e-10  
gi|116256981|gb|ABY90663.1| 3-phosphoshikimate 1-c ( 435) 300 76.0 2.5e-10  
gi|139968357|gb|ECI70702.1| hypothetical protein G ( 201) 296 74.9 2.5e-10  
gi|142419593|gb|ECX93837.1| hypothetical protein G ( 420) 299 75.8 2.8e-10  
gi|139934981|gb|ECI47821.1| hypothetical protein G ( 184) 294 74.4 3.1e-10  
gi|141235823|gb|ECQ56307.1| hypothetical protein G ( 131) 292 73.9 3.3e-10  
gi|136414763|gb|EBO01321.1| hypothetical protein G ( 423) 298 75.6 3.3e-10  
gi|139449963|gb|ECF18170.1| hypothetical protein G ( 194) 294 74.4 3.3e-10  
gi|257170988|gb|ACV48747.1| 3-phosphoshikimate 1-c ( 430) 298 75.6 3.3e-10  
gi|218301277|emb|CAU98625.1| unnamed protein produ ( 447) 298 75.6 3.4e-10  
gi|259420526|emb|CBF57577.1| unnamed protein produ ( 447) 298 75.6 3.4e-10  
gi|141805848|gb|ECT37496.1| hypothetical protein G ( 174) 293 74.2 3.5e-10  
gi|138461383|gb|EBZ83526.1| hypothetical protein G ( 178) 293 74.2 3.6e-10  
gi|139740652|gb|ECH14977.1| hypothetical protein G ( 183) 293 74.2 3.6e-10  
gi|147830428|emb|CAN01363.1| 3-phosphoshikimate 1- ( 487) 298 75.6 3.6e-10  
gi|218675784|gb|ACL00597.1| 5-enolpyruvylshikimate ( 410) 297 75.3 3.7e-10  
gi|141778677|gb|ECT24444.1| hypothetical protein G ( 147) 291 73.7 4.1e-10  
gi|141399981|gb|ECR56140.1| hypothetical protein G ( 182) 292 74.0 4.2e-10  
gi|116699210|gb|ABK18398.1| 3-phosphoshikimate 1-c ( 423) 296 75.1 4.4e-10  
gi|140604009|gb|ECM35577.1| hypothetical protein G ( 194) 292 74.0 4.4e-10  
gi|56179475|gb|AAV82197.1| 5-enolpyruvylshikimate- ( 429) 296 75.1 4.5e-10  
gi|193933758|gb|ACF29582.1| 3-phosphoshikimate 1-c ( 433) 296 75.1 4.5e-10  
gi|59718193|gb|AAW89598.1| putative 5-enolpyruvoyl ( 433) 296 75.1 4.5e-10  
gi|120866826|emb|CAM10584.1| 5-enolpyruvoylshikima ( 433) 296 75.1 4.5e-10  
gi|262181543|gb|ACY29662.1| 5-enolpyruvylshikimate ( 248) 293 74.3 4.6e-10  
gi|141952739|gb|ECU40371.1| hypothetical protein G ( 171) 291 73.7 4.7e-10  
gi|137016683|gb|EBR73704.1| hypothetical protein G ( 99) 288 72.9 4.8e-10  
gi|138587156|gb|ECA62809.1| hypothetical protein G ( 181) 291 73.8 4.9e-10  
gi|139832401|gb|ECH77500.1| hypothetical protein G ( 289) 293 74.3 5.2e-10  
gi|136634439|gb|EBP41578.1| hypothetical protein G ( 136) 289 73.2 5.3e-10  
gi|137394998|gb|EBT85728.1| hypothetical protein G ( 159) 289 73.3 6e-10  
gi|139968913|gb|ECI71100.1| hypothetical protein G ( 132) 288 73.0 6e-10  
gi|170783792|gb|ACB37380.1| 5-enolpyruvylshikimate ( 520) 295 75.0 6e-10  
gi|161161889|emb|CAN93194.1| 3-phosphoshikimate 1- ( 428) 294 74.7 6e-10  
gi|142712954|gb|EDA03778.1| hypothetical protein G ( 170) 289 73.3 6.3e-10  
gi|136829493|gb|EBQ68498.1| hypothetical protein G ( 307) 292 74.2 6.3e-10  
gi|135833976|gb|EBK13589.1| hypothetical protein G ( 143) 288 73.0 6.4e-10  
gi|141325309|gb|ECR05885.1| hypothetical protein G ( 97) 286 72.5 6.4e-10  
gi|139552759|gb|ECF85571.1| hypothetical protein G ( 178) 289 73.3 6.5e-10  
gi|136580598|gb|EBP07978.1| hypothetical protein G ( 329) 292 74.2 6.6e-10  
gi|141711640|gb|ECS87995.1| hypothetical protein G ( 184) 289 73.3 6.7e-10  
gi|141393186|gb|ECR51348.1| hypothetical protein G ( 103) 286 72.5 6.7e-10  
gi|134397597|gb|EBB11730.1| hypothetical protein G ( 186) 289 73.3 6.7e-10  
gi|139936889|gb|ECI49196.1| hypothetical protein G ( 192) 289 73.3 6.9e-10  
gi|136634441|gb|EBP41580.1| hypothetical protein G ( 132) 287 72.8 7e-10  
gi|142319596|gb|ECX22834.1| hypothetical protein G ( 133) 287 72.8 7e-10

gi|142053474|gb|ECV20051.1| hypothetical protein G ( 362) 292 74.2 7.2e-10  
gi|136931619|gb|EBR25604.1| hypothetical protein G ( 203) 289 73.4 7.2e-10  
gi|135714631|gb|EBU38978.1| hypothetical protein G ( 115) 286 72.5 7.3e-10  
gi|140795452|gb|ECN56325.1| hypothetical protein G ( 172) 288 73.1 7.4e-10  
gi|138381148|gb|EBZ27658.1| hypothetical protein G ( 103) 285 72.3 7.8e-10  
gi|136715732|gb|EBP93040.1| hypothetical protein G ( 417) 292 74.3 8e-10  
gi|137380099|gb|EBT77156.1| hypothetical protein G ( 246) 289 73.4 8.3e-10  
gi|141038491|gb|ECP20912.1| hypothetical protein G ( 180) 287 72.9 8.9e-10  
gi|3893083|emb|CAA10164.1| 5-enolpyruvylshikimate ( 332) 290 73.7 9e-10  
gi|141831272|gb|ECT55097.1| hypothetical protein G ( 190) 287 72.9 9.2e-10  
gi|138274369|gb|EBY73568.1| hypothetical protein G ( 192) 287 72.9 9.3e-10  
gi|138621767|gb|ECA86898.1| hypothetical protein G ( 193) 287 72.9 9.4e-10  
gi|135327870|gb|EBG89709.1| hypothetical protein G ( 196) 287 72.9 9.5e-10  
gi|139417854|gb|ECE97583.1| hypothetical protein G ( 137) 285 72.4 9.7e-10  
gi|170231|gb|AAA34072.1| 5-enolpyruvylshikimate-3- ( 338) 289 73.5 1.1e-09  
gi|119955588|gb|ABM12593.1| 3-phosphoshikimate 1-c ( 446) 290 73.8 1.1e-09  
gi|139214786|gb|ECE19726.1| hypothetical protein G ( 310) 288 73.3 1.2e-09  
gi|141668834|gb|ECS71793.1| hypothetical protein G ( 150) 284 72.2 1.2e-09  
gi|141342655|gb|ECR16143.1| hypothetical protein G ( 187) 285 72.5 1.2e-09  
gi|142629045|gb|ECZ44103.1| hypothetical protein G ( 188) 285 72.5 1.2e-09  
gi|141309322|gb|ECQ7031.1| hypothetical protein G ( 190) 285 72.5 1.2e-09  
gi|218760993|gb|ACU03459.1| 3-phosphoshikimate 1-c ( 419) 289 73.6 1.3e-09  
gi|221734706|gb|ACM35669.1| 3-phosphoshikimate 1-c ( 423) 289 73.6 1.3e-09  
gi|137283778|gb|EBT25443.1| hypothetical protein G ( 89) 281 71.4 1.3e-09  
gi|140448493|gb|ECJ82608.1| hypothetical protein G ( 109) 282 71.6 1.3e-09  
gi|261282220|gb|ACX64191.1| 3-phosphoshikimate 1-c ( 430) 289 73.6 1.3e-09  
gi|142537102|gb|EBD56578.1| hypothetical protein G ( 165) 284 72.2 1.3e-09  
gi|140904031|gb|ECO28501.1| hypothetical protein G ( 174) 284 72.2 1.4e-09  
gi|141771908|gb|ECT20846.1| hypothetical protein G ( 118) 282 71.7 1.4e-09  
gi|137287625|gb|EBT25443.1| hypothetical protein G ( 179) 284 72.2 1.4e-09  
gi|3410961|dbj|BAA32276.1| 3-phosphoshikimate 1-ca ( 391) 288 73.4 1.4e-09  
gi|138333549|gb|EBZ02341.1| hypothetical protein G ( 184) 284 72.2 1.4e-09  
gi|136109563|gb|EBL97675.1| hypothetical protein G ( 158) 283 72.0 1.5e-09  
gi|144179424|gb|EDJ44599.1| hypothetical protein G ( 236) 285 72.5 1.5e-09  
gi|251844766|gb|EES72779.1| 3-phosphoshikimate 1-c ( 430) 288 73.4 1.5e-09  
gi|143911526|gb|EDH51939.1| hypothetical protein G ( 366) 287 73.1 1.5e-09  
gi|138211244|gb|EBY44003.1| hypothetical protein G ( 126) 281 71.5 1.7e-09  
gi|138316610|gb|EBY93784.1| hypothetical protein G ( 154) 282 71.7 1.7e-09  
gi|134641468|gb|EBC55730.1| hypothetical protein G ( 128) 281 71.5 1.7e-09  
gi|135193457|gb|EBG09189.1| hypothetical protein G ( 195) 283 72.0 1.7e-09  
gi|262193245|gb|ACY30625.1| EPSP synthase [Ginkgo ( 353) 286 72.9 1.7e-09  
gi|134857212|gb|EBD88436.1| hypothetical protein G ( 166) 282 71.8 1.8e-09  
gi|141711572|gb|ECS87954.1| hypothetical protein G ( 181) 282 71.8 1.9e-09  
gi|136863779|gb|EBQ91435.1| hypothetical protein G ( 124) 280 71.2 1.9e-09  
gi|262086810|gb|ACY22778.1| 3-phosphoshikimate 1-c ( 421) 286 72.9 2e-09  
gi|139026901|gb|ECC89959.1| hypothetical protein G ( 131) 280 71.3 2e-09  
gi|145217759|gb|ABP47163.1| 3-phosphoshikimate 1-c ( 446) 286 73.0 2.1e-09  
gi|138859647|gb|ECC18477.1| hypothetical protein G ( 188) 281 71.6 2.3e-09  
gi|218602333|emb|CAV32035.1| unnamed protein produ ( 413) 285 72.7 2.3e-09  
gi|158064957|emb|CAP11440.1| unnamed protein produ ( 413) 285 72.7 2.3e-09  
gi|254257746|emb|CAZ91482.1| unnamed protein produ ( 413) 285 72.7 2.3e-09  
gi|137439612|gb|EBU10974.1| hypothetical protein G ( 194) 281 71.6 2.3e-09  
gi|141103300|gb|ECP64812.1| hypothetical protein G ( 132) 279 71.0 2.3e-09  
gi|136894582|gb|EBR09947.1| hypothetical protein G ( 161) 280 71.3 2.3e-09  
gi|135246787|gb|EBG40574.1| hypothetical protein G ( 396) 284 72.5 2.6e-09  
gi|137613393|gb|EBV05490.1| hypothetical protein G ( 110) 277 70.5 2.7e-09  
gi|120563746|gb|ABM29490.1| 3-phosphoshikimate 1-c ( 439) 284 72.5 2.8e-09  
gi|256586404|gb|ACU97537.1| 3-phosphoshikimate 1-c ( 444) 284 72.5 2.8e-09  
gi|138115862|gb|EBX83138.1| hypothetical protein G ( 205) 280 71.4 2.8e-09  
gi|141718967|gb|ECS92519.1| hypothetical protein G ( 177) 279 71.1 2.9e-09  
gi|138705585|gb|ECB44304.1| hypothetical protein G ( 183) 279 71.2 3e-09  
gi|139308621|gb|ECE49455.1| hypothetical protein G ( 185) 279 71.2 3e-09

gi|73660021|emb|CAI82628.1| 3-phosphoshikimate 1-c ( 420) 283 72.3 3.1e-09  
gi|146270033|gb|ABQ17025.1| 3-phosphoshikimate 1-c ( 420) 283 72.3 3.1e-09  
gi|183174668|gb|ACC39778.1| 3-phosphoshikimate 1-c ( 431) 283 72.3 3.2e-09  
gi|158064938|emb|CAP11431.1| unnamed protein produ ( 436) 283 72.3 3.2e-09  
gi|254257728|emb|CAZ91473.1| unnamed protein produ ( 436) 283 72.3 3.2e-09  
gi|256356896|gb|ACU70393.1| UDP-N-acetylglucosamin ( 450) 283 72.3 3.3e-09  
gi|139600607|gb|ECG18735.1| hypothetical protein G ( 115) 276 70.3 3.3e-09  
gi|136149188|gb|EBM21932.1| hypothetical protein G ( 313) 281 71.8 3.4e-09  
gi|135076564|gb|EBF34204.1| hypothetical protein G ( 215) 279 71.2 3.4e-09  
gi|135862908|gb|EBK31807.1| hypothetical protein G ( 181) 278 70.9 3.5e-09  
gi|135775967|gb|EBJ77004.1| hypothetical protein G ( 181) 278 70.9 3.5e-09  
gi|140340030|gb|ECL09713.1| hypothetical protein G ( 151) 277 70.7 3.5e-09  
gi|146407374|gb|ABQ35880.1| 3-phosphoshikimate 1-c ( 418) 282 72.1 3.6e-09  
gi|138212148|gb|EBY44623.1| hypothetical protein G ( 131) 276 70.4 3.6e-09  
gi|261370679|gb|ACX73428.1| 3-phosphoshikimate 1-c ( 426) 282 72.1 3.7e-09  
gi|167727088|emb|CAP13874.1| 3-phosphoshikimate 1- ( 430) 282 72.1 3.7e-09  
gi|54018070|dbj|BAD59440.1| putative 5-enolpyruvyl ( 437) 282 72.1 3.7e-09  
gi|197053758|gb|ACH25456.1| Sequence 8 from patent ( 439) 282 72.1 3.8e-09  
gi|167291814|gb|ABZ44678.1| Sequence 18616 from pa ( 439) 282 72.1 3.8e-09  
gi|46448291|gb|AAS94946.1| 3-phosphoshikimate 1-ca ( 439) 282 72.1 3.8e-09  
gi|105807601|gb|AAG19594.1| 3-phosphoshikimate 1-ca ( 439) 282 72.1 3.8e-09  
gi|137322029|gb|EBT44686.1| hypothetical protein G ( 178) 277 70.7 4e-09  
gi|137129293|gb|EBS36687.1| hypothetical protein G ( 225) 278 71.0 4.1e-09  
gi|141714720|gb|ECS89901.1| hypothetical protein G ( 186) 277 70.7 4.1e-09  
gi|137504641|gb|EBU45451.1| hypothetical protein G ( 188) 277 70.7 4.1e-09  
gi|143304386|gb|EDE21860.1| hypothetical protein G ( 416) 281 71.9 4.2e-09  
gi|167043753|gb|ABZ08445.1| putative EPSP synthase ( 422) 281 71.9 4.2e-09  
gi|140867415|gb|EC004516.1| hypothetical protein G ( 165) 276 70.5 4.4e-09  
gi|226244298|dbj|BAH54646.1| 3-phosphoshikimate 1- ( 438) 281 71.9 4.4e-09  
gi|29606692|dbj|BAC70753.1| putative 3-phosphoshik ( 446) 281 71.9 4.4e-09  
gi|269096461|gb|ACZ20897.1| 3-phosphoshikimate 1-c ( 448) 281 71.9 4.4e-09  
gi|137062259|gb|EBR99387.1| hypothetical protein G ( 190) 276 70.5 4.9e-09  
gi|135601197|gb|EBI68833.1| hypothetical protein G ( 195) 276 70.5 5e-09  
gi|137781568|gb|EBV96306.1| hypothetical protein G ( 289) 278 71.1 5e-09  
gi|167294666|gb|ABZ47530.1| Sequence 21468 from pa ( 429) 280 71.6 5e-09  
gi|197053759|gb|ACH25457.1| Sequence 9 from patent ( 429) 280 71.6 5e-09  
gi|2826287|gb|AAB98493.1| 3-phosphoshikimate-1-car ( 429) 280 71.6 5e-09  
gi|142307382|gb|ECX13686.1| hypothetical protein G ( 167) 275 70.3 5.1e-09  
gi|140677546|gb|ECM75558.1| hypothetical protein G ( 117) 273 69.7 5.3e-09  
gi|141055410|gb|ECP32712.1| hypothetical protein G ( 177) 275 70.3 5.3e-09  
gi|143120588|gb|EDC96439.1| hypothetical protein G ( 178) 275 70.3 5.4e-09  
gi|141347116|gb|ECR19224.1| hypothetical protein G ( 100) 272 69.4 5.4e-09  
gi|23492616|dbj|BAC17589.1| 5-enolpyruvylshikimate ( 408) 279 71.4 5.6e-09  
gi|136997668|gb|EBR62935.1| hypothetical protein G ( 189) 275 70.3 5.6e-09  
gi|136122790|gb|EBM06705.1| hypothetical protein G ( 163) 274 70.0 5.8e-09  
gi|8163275|gb|AAF73581.1| 3-phosphoshikimate 1-car ( 441) 279 71.4 5.9e-09  
gi|137991869|gb|EBX16169.1| hypothetical protein G ( 167) 274 70.0 5.9e-09  
gi|143421907|gb|EDE86867.1| hypothetical protein G ( 252) 276 70.6 6e-09  
gi|140878388|gb|ECO12060.1| hypothetical protein G ( 171) 274 70.0 6.1e-09  
gi|134993483|gb|EBE79617.1| hypothetical protein G ( 174) 274 70.0 6.1e-09  
gi|139557235|gb|ECF88762.1| hypothetical protein G ( 120) 272 69.5 6.2e-09  
gi|140266570|gb|ECK66195.1| hypothetical protein G ( 223) 275 70.3 6.4e-09  
gi|143286973|gb|EDE12890.1| hypothetical protein G ( 330) 277 70.9 6.4e-09  
gi|218602347|emb|CAV32047.1| unnamed protein produ ( 413) 278 71.2 6.5e-09  
gi|218602343|emb|CAV32045.1| unnamed protein produ ( 413) 278 71.2 6.5e-09  
gi|139742666|gb|ECH16215.1| hypothetical protein G ( 130) 272 69.5 6.6e-09  
gi|141281733|gb|ECQ86790.1| hypothetical protein G ( 239) 275 70.4 6.7e-09  
gi|110822831|gb|ABG98115.1| 3-phosphoshikimate 1-c ( 438) 278 71.2 6.9e-09  
gi|269096620|gb|ACZ21056.1| UDP-N-acetylglucosamin ( 438) 278 71.2 6.9e-09  
gi|134403526|gb|EBB15016.1| hypothetical protein G ( 171) 273 69.8 7e-09  
gi|158064955|emb|CAP11439.1| unnamed protein produ ( 413) 277 71.0 7.6e-09  
gi|254257744|emb|CAZ91481.1| unnamed protein produ ( 413) 277 71.0 7.6e-09

gi|218602331|emb|CAV32033.1| unnamed protein produ ( 413) 277 71.0 7.6e-09  
gi|144221934|gb|EDU75270.1| hypothetical protein G ( 415) 277 71.0 7.6e-09  
gi|139983349|gb|ECI81241.1| hypothetical protein G ( 284) 275 70.4 7.7e-09  
gi|55229780|gb|AAV45199.1| 3-phosphoshikimate 1-ca ( 429) 277 71.0 7.8e-09  
gi|140795343|gb|ECN56244.1| hypothetical protein G ( 162) 272 69.6 7.9e-09  
gi|142965654|gb|EDB84756.1| hypothetical protein G ( 134) 271 69.3 7.9e-09  
gi|135937496|gb|EBK83135.1| hypothetical protein G ( 169) 272 69.6 8.1e-09  
gi|88783188|gb|EAR14361.1| 3-phosphoshikimate 1-ca ( 409) 276 70.8 8.8e-09  
gi|140962513|gb|ECO69651.1| hypothetical protein G ( 127) 270 69.1 8.8e-09  
gi|158064953|emb|CAP11438.1| unnamed protein produ ( 413) 276 70.8 8.9e-09  
gi|254257732|emb|CAZ91475.1| unnamed protein produ ( 413) 276 70.8 8.9e-09  
gi|218602319|emb|CAV32025.1| unnamed protein produ ( 413) 276 70.8 8.9e-09  
gi|218602349|emb|CAV32048.1| unnamed protein produ ( 413) 276 70.8 8.9e-09  
gi|254257740|emb|CAZ91479.1| unnamed protein produ ( 413) 276 70.8 8.9e-09  
gi|218602327|emb|CAV32029.1| unnamed protein produ ( 413) 276 70.8 8.9e-09  
gi|158064951|emb|CAP11437.1| unnamed protein produ ( 413) 276 70.8 8.9e-09  
gi|158064943|emb|CAP11433.1| unnamed protein produ ( 413) 276 70.8 8.9e-09  
gi|218602329|emb|CAV32031.1| unnamed protein produ ( 413) 276 70.8 8.9e-09  
gi|254257742|emb|CAZ91480.1| unnamed protein produ ( 413) 276 70.8 8.9e-09  
gi|140607098|gb|ECM36324.1| hypothetical protein G ( 232) 273 69.9 8.9e-09  
gi|137598092|gb|EBU96799.1| hypothetical protein G ( 88) 268 68.5 9e-09  
gi|193083953|gb|ACF09630.1| 3-phosphoshikimate 1-c ( 426) 276 70.8 9.1e-09  
gi|139916023|gb|ECI34592.1| hypothetical protein G ( 133) 270 69.1 9.1e-09  
gi|260078369|gb|EEW66204.1| 3-phosphoshikimate 1-c ( 135) 270 69.1 9.2e-09  
gi|134563127|gb|EBC08803.1| hypothetical protein G ( 99) 268 68.6 9.8e-09  
gi|134557848|gb|EBC05598.1| hypothetical protein G ( 407) 275 70.5 1e-08  
gi|110620656|emb|CAJ35934.1| putative 3-phosphoshi ( 422) 275 70.6 1e-08  
gi|256690345|gb|ACV10682.1| 3-phosphoshikimate 1-c ( 428) 275 70.6 1.1e-08  
gi|118171644|gb|ABK72540.1| 3-phosphoshikimate 1-c ( 443) 275 70.6 1.1e-08  
gi|261373400|gb|ACX76145.1| EPSP synthase (3-phosp ( 455) 275 70.6 1.1e-08  
gi|219860009|gb|ACL40351.1| 3-phosphoshikimate 1-c ( 475) 275 70.6 1.1e-08  
gi|139189020|gb|ECE01933.1| hypothetical protein G ( 182) 270 69.2 1.2e-08  
gi|134423121|gb|EBB26248.1| hypothetical protein G ( 182) 270 69.2 1.2e-08  
gi|143138581|gb|EDD09640.1| hypothetical protein G ( 84) 266 68.1 1.2e-08  
gi|254257748|emb|CAZ91483.1| unnamed protein produ ( 413) 274 70.3 1.2e-08  
gi|218602335|emb|CAV32038.1| unnamed protein produ ( 413) 274 70.3 1.2e-08  
gi|158064959|emb|CAP11441.1| unnamed protein produ ( 413) 274 70.3 1.2e-08  
gi|114337300|gb|ABT68148.1| UDP-N-acetylglucosamin ( 416) 274 70.3 1.2e-08  
gi|141573503|gb|ECS42617.1| hypothetical protein G ( 114) 267 68.4 1.3e-08  
gi|178464619|dbj|BAG19139.1| putative 5-enolpyruvyl ( 450) 274 70.4 1.3e-08  
gi|164454854|dbj|BAF96972.1| putative 5-enolpyruvyl ( 122) 267 68.4 1.3e-08  
gi|218602321|emb|CAV32026.1| unnamed protein produ ( 413) 273 70.1 1.4e-08  
gi|218602325|emb|CAV32028.1| unnamed protein produ ( 413) 273 70.1 1.4e-08  
gi|218602341|emb|CAV32044.1| unnamed protein produ ( 413) 273 70.1 1.4e-08  
gi|254257730|emb|CAZ91474.1| unnamed protein produ ( 413) 273 70.1 1.4e-08  
gi|254257734|emb|CAZ91476.1| unnamed protein produ ( 413) 273 70.1 1.4e-08  
gi|158064940|emb|CAP11432.1| unnamed protein produ ( 413) 273 70.1 1.4e-08  
gi|254257752|emb|CAZ91485.1| unnamed protein produ ( 413) 273 70.1 1.4e-08  
gi|218602339|emb|CAV32043.1| unnamed protein produ ( 413) 273 70.1 1.4e-08  
gi|254257754|emb|CAZ91486.1| unnamed protein produ ( 413) 273 70.1 1.4e-08  
gi|158064945|emb|CAP11434.1| unnamed protein produ ( 413) 273 70.1 1.4e-08  
gi|158064965|emb|CAP11444.1| unnamed protein produ ( 413) 273 70.1 1.4e-08  
gi|158064949|emb|CAP11436.1| unnamed protein produ ( 413) 273 70.1 1.4e-08  
gi|158064963|emb|CAP11443.1| unnamed protein produ ( 413) 273 70.1 1.4e-08  
gi|254257738|emb|CAZ91478.1| unnamed protein produ ( 413) 273 70.1 1.4e-08  
gi|134322213|gb|EBA63460.1| hypothetical protein G ( 283) 271 69.6 1.4e-08  
gi|144012227|gb|EDI23678.1| hypothetical protein G ( 173) 268 68.7 1.5e-08  
gi|222136865|gb|ACM45082.1| 5-enolpyruvylshikimate ( 317) 271 69.6 1.5e-08  
gi|145751285|gb|ABP96766.1| putative 5-enolpyruvyl ( 122) 266 68.2 1.6e-08  
gi|27651526|emb|CAD36875.1| putative 5-enolpyruvyl ( 122) 266 68.2 1.6e-08  
gi|138393082|gb|EBZ35835.1| hypothetical protein G ( 183) 268 68.8 1.6e-08  
gi|41398263|gb|AAS05884.1| AroA [Mycobacterium avi ( 435) 272 69.9 1.7e-08

gi|141443969|gb|ECR86984.1| hypothetical protein G ( 137) 266 68.2 1.7e-08  
gi|139715640|gb|ECG97590.1| hypothetical protein G ( 94) 264 67.7 1.7e-08  
gi|136937780|gb|EBR29049.1| hypothetical protein G ( 169) 267 68.5 1.7e-08  
gi|139801494|gb|ECH56019.1| hypothetical protein G ( 172) 267 68.5 1.7e-08  
gi|143754925|gb|EDG61826.1| hypothetical protein G ( 177) 267 68.5 1.8e-08  
gi|138181735|gb|EBY25008.1| hypothetical protein G ( 180) 267 68.5 1.8e-08  
gi|143715076|gb|EDG41631.1| hypothetical protein G ( 150) 266 68.3 1.8e-08  
gi|139765372|gb|ECH30618.1| hypothetical protein G ( 126) 265 68.0 1.9e-08  
gi|135654046|gb|EBJ01552.1| hypothetical protein G ( 187) 267 68.5 1.9e-08  
gi|57225151|gb|AAW40208.1| 3-phosphoshikimate 1-ca ( 420) 271 69.7 1.9e-08  
gi|143840587|gb|EDH00849.1| hypothetical protein G ( 90) 263 67.4 1.9e-08  
gi|135805402|gb|EBJ95639.1| hypothetical protein G ( 168) 266 68.3 2e-08  
gi|137599818|gb|EBU97767.1| hypothetical protein G ( 78) 262 67.2 2e-08  
gi|139602942|gb|ECG20278.1| hypothetical protein G ( 266) 268 68.9 2.1e-08  
gi|164454856|dbj|BAF96973.1| putative 5-enolpyruvyl ( 122) 264 67.8 2.1e-08  
gi|167044926|gb|ABZ09592.1| putative EPSP synthase ( 394) 270 69.4 2.1e-08  
gi|138589959|gb|ECA64753.1| hypothetical protein G ( 273) 268 68.9 2.2e-08  
gi|135117332|gb|EBF60283.1| hypothetical protein G ( 86) 262 67.2 2.2e-08  
gi|256793860|gb|ACV24529.1| 3-phosphoshikimate 1-c ( 427) 270 69.5 2.2e-08  
gi|139675536|gb|ECG69534.1| hypothetical protein G ( 92) 262 67.2 2.3e-08  
gi|139349149|gb|ECE58652.1| hypothetical protein G ( 136) 264 67.8 2.3e-08  
gi|134554900|gb|EBC03811.1| hypothetical protein G ( 173) 265 68.1 2.4e-08  
gi|262208075|gb|ACY32173.1| 3-phosphoshikimate 1-c ( 690) 272 70.1 2.4e-08  
gi|27651528|emb|CAD36876.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|164454852|dbj|BAF96971.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|27651568|emb|CAD36896.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|27651538|emb|CAD36881.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|145750944|gb|ABP96750.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|27651564|emb|CAD36894.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|27651546|emb|CAD36885.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|145751141|gb|ABP96759.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|27651530|emb|CAD36877.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|27651508|emb|CAD36866.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|27651570|emb|CAD36897.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|27651522|emb|CAD36873.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|27651524|emb|CAD36874.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|27651534|emb|CAD36879.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|27651536|emb|CAD36880.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|27651558|emb|CAD36891.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|27651506|emb|CAD36865.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|27651544|emb|CAD36884.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|27651532|emb|CAD36878.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|27651514|emb|CAD36869.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|27651542|emb|CAD36883.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|27651548|emb|CAD36886.1| putative 5-enolpyruvyl ( 122) 263 67.5 2.5e-08  
gi|136187309|gb|EBM46881.1| hypothetical protein G ( 279) 267 68.7 2.5e-08  
gi|158064961|emb|CAP11442.1| unnamed protein produ ( 413) 269 69.2 2.5e-08  
gi|254257750|emb|CAZ91484.1| unnamed protein produ ( 413) 269 69.2 2.5e-08  
gi|218602337|emb|CAV32040.1| unnamed protein produ ( 413) 269 69.2 2.5e-08  
gi|270153723|gb|ACZ61561.1| 3-phosphoshikimate 1-c ( 420) 269 69.2 2.6e-08  
gi|238876015|gb|ACR75722.1| 3-phosphoshikimate 1-c ( 427) 269 69.3 2.6e-08  
gi|27651494|emb|CAD36948.1| putative 5-enolpyruvyl ( 122) 262 67.3 2.9e-08  
gi|27651554|emb|CAD36889.1| putative 5-enolpyruvyl ( 122) 262 67.3 2.9e-08  
gi|27651510|emb|CAD36867.1| putative 5-enolpyruvyl ( 122) 262 67.3 2.9e-08  
gi|145750758|gb|ABP96741.1| putative 5-enolpyruvyl ( 122) 262 67.3 2.9e-08  
gi|62948560|gb|AAZ22964.1| putative 5-enolpyruvyls ( 122) 262 67.3 2.9e-08  
gi|27651498|emb|CAD36861.1| putative 5-enolpyruvyl ( 122) 262 67.3 2.9e-08  
gi|155369179|dbj|BAF75638.1| putative 5-enolpyruvyl ( 122) 262 67.3 2.9e-08  
gi|27651552|emb|CAD36888.1| putative 5-enolpyruvyl ( 122) 262 67.3 2.9e-08  
gi|27651562|emb|CAD36893.1| putative 5-enolpyruvyl ( 122) 262 67.3 2.9e-08  
gi|76577727|gb|ABA54134.1| putative 5-enolpyruvyls ( 122) 262 67.3 2.9e-08  
gi|145751418|gb|ABP96773.1| putative 5-enolpyruvyl ( 122) 262 67.3 2.9e-08

gi|137383382|gb|EBT79041.1| hypothetical protein G ( 151) 263 67.6 2.9e-08  
gi|254257736|emb|CAZ91477.1| unnamed protein produ ( 413) 268 69.0 3e-08  
gi|158064947|emb|CAP11435.1| unnamed protein produ ( 413) 268 69.0 3e-08  
gi|218602323|emb|CAV32027.1| unnamed protein produ ( 413) 268 69.0 3e-08  
gi|143832006|gb|EDG94622.1| hypothetical protein G ( 162) 263 67.6 3.1e-08  
gi|135286490|gb|EBG63923.1| hypothetical protein G ( 164) 263 67.6 3.1e-08  
gi|135097052|gb|EBF47290.1| hypothetical protein G ( 367) 267 68.8 3.1e-08  
gi|140482317|gb|ECL96815.1| hypothetical protein G ( 171) 263 67.6 3.2e-08  
gi|139901711|gb|ECI24311.1| hypothetical protein G ( 118) 261 67.1 3.2e-08  
gi|27651520|emb|CAD36872.1| putative 5-enolpyruvyl ( 122) 261 67.1 3.3e-08  
gi|134801291|gb|EBD52889.1| hypothetical protein G ( 398) 267 68.8 3.3e-08  
gi|218602345|emb|CAV32046.1| unnamed protein produ ( 413) 267 68.8 3.4e-08  
gi|110682538|gb|ABG85908.1| 3-phosphoshikimate 1-c ( 424) 267 68.8 3.5e-08  
gi|165970311|gb|ABY76172.1| chloroplast CTP/EPSPS ( 41) 255 65.4 3.5e-08  
gi|118570151|gb|ABA04902.1| 3-phosphoshikimate 1-c ( 431) 267 68.8 3.6e-08  
gi|137240909|gb|EB599122.1| hypothetical protein G ( 111) 260 66.9 3.6e-08  
gi|137289947|gb|EBT26741.1| hypothetical protein G ( 121) 260 66.9 3.8e-08  
gi|27651566|emb|CAD36895.1| putative 5-enolpyruvyl ( 122) 260 66.9 3.9e-08  
gi|27651550|emb|CAD36887.1| putative 5-enolpyruvyl ( 122) 260 66.9 3.9e-08  
gi|27651496|emb|CAD36949.1| putative 5-enolpyruvyl ( 122) 260 66.9 3.9e-08  
gi|208968957|dbj|BAG74299.1| putative 5-enolpyruvyl ( 122) 260 66.9 3.9e-08  
gi|27651500|emb|CAD36862.1| putative 5-enolpyruvyl ( 122) 260 66.9 3.9e-08  
gi|140356997|gb|ECL21447.1| hypothetical protein G ( 126) 260 66.9 4e-08  
gi|138607893|dbj|BAG74389.1| hypothetical protein G ( 127) 260 66.9 4e-08  
gi|110673799|gb|ABG82786.1| 3-phosphoshikimate 1-c ( 424) 266 68.6 4.1e-08  
gi|136501410|gb|EBO57343.1| hypothetical protein G ( 287) 264 68.0 4.1e-08  
gi|142748549|gb|EDA29468.1| hypothetical protein G ( 238) 263 67.8 4.1e-08  
gi|197053755|gb|ACH25453.1| Sequence 5 from patent ( 428) 266 68.6 4.1e-08  
gi|15023793|gb|AAK78871.1|AE007605\_4 5-enolpyruvyl ( 428) 266 68.6 4.1e-08  
gi|188499339|gb|ACD52475.1| 3-phosphoshikimate 1-c ( 433) 266 68.6 4.2e-08  
gi|137572843|gb|EBU82720.1| hypothetical protein G ( 91) 258 66.4 4.2e-08  
gi|139117331|gb|ECD53221.1| hypothetical protein G ( 201) 262 67.5 4.2e-08  
gi|140222712|gb|ECK35845.1| hypothetical protein G ( 169) 261 67.2 4.3e-08  
gi|137440713|gb|EBU11589.1| hypothetical protein G ( 140) 260 66.9 4.3e-08  
gi|143380471|gb|EDE67259.1| hypothetical protein G ( 212) 262 67.5 4.4e-08  
gi|27651502|emb|CAD36863.1| putative 5-enolpyruvyl ( 122) 259 66.7 4.5e-08  
gi|27651540|emb|CAD36882.1| putative 5-enolpyruvyl ( 122) 259 66.7 4.5e-08  
gi|27651512|emb|CAD36868.1| putative 5-enolpyruvyl ( 122) 259 66.7 4.5e-08  
gi|136076477|gb|EBL75135.1| hypothetical protein G ( 149) 260 66.9 4.5e-08  
gi|142613449|gb|ECZ33107.1| hypothetical protein G ( 101) 258 66.4 4.5e-08  
gi|142081421|gb|ECV43511.1| hypothetical protein G ( 106) 258 66.4 4.7e-08  
gi|153999656|gb|ABS56079.1| 3-phosphoshikimate 1-c ( 421) 265 68.4 4.7e-08  
gi|197053756|gb|ACH25454.1| Sequence 6 from patent ( 424) 265 68.4 4.8e-08  
gi|18144355|dbj|BAB80402.1| 3-phosphoshikimate 1-c ( 424) 265 68.4 4.8e-08  
gi|16751567|gb|AAL27697.1| 5-enolpyruvylshikimate- ( 516) 266 68.7 4.8e-08  
gi|136649531|gb|EBP50530.1| hypothetical protein G ( 110) 258 66.4 4.8e-08  
gi|140161022|gb|ECJ93154.1| hypothetical protein G ( 171) 260 67.0 5e-08  
gi|136661055|gb|EBP57371.1| hypothetical protein G ( 102) 257 66.2 5.3e-08  
gi|137626582|gb|EBV13018.1| hypothetical protein G ( 188) 260 67.0 5.4e-08  
gi|139710378|gb|ECG93826.1| hypothetical protein G ( 131) 258 66.5 5.5e-08  
gi|197053753|gb|ACH25451.1| Sequence 3 from patent ( 427) 264 68.2 5.6e-08  
gi|167296048|gb|ABZ48912.1| Sequence 22850 from pa ( 427) 264 68.2 5.6e-08  
gi|5104222|dbj|BAA79537.1| 3-phosphoshikimate 1-ca ( 427) 264 68.2 5.6e-08  
gi|135307869|gb|EBG76462.1| hypothetical protein G ( 167) 259 66.8 5.7e-08  
gi|143352053|gb|EDE50399.1| hypothetical protein G ( 300) 262 67.6 5.7e-08  
gi|140653357|gb|ECM58430.1| hypothetical protein G ( 304) 262 67.6 5.8e-08  
gi|253321772|gb|ACT26375.1| 3-phosphoshikimate 1-c ( 450) 264 68.2 5.8e-08  
gi|148722956|gb|ABR07581.1| 3-phosphoshikimate 1-c ( 450) 264 68.2 5.8e-08  
gi|149928|gb|AAA25356.1| key enzyme from the shiki ( 450) 264 68.2 5.8e-08  
gi|224774724|dbj|BAH27530.1| 3-phosphoshikimate 1- ( 450) 264 68.2 5.8e-08  
gi|13883144|gb|AAK47667.1| 3-phosphoshikimate 1-ca ( 450) 264 68.2 5.8e-08  
gi|1061198|emb|CAA36510.1| unnamed protein product ( 450) 264 68.2 5.8e-08

gi|121494857|emb|CAL73339.1| 3-phosphoshikimate 1- ( 450) 264 68.2 5.8e-08  
gi|148507242|gb|ABQ75051.1| 3-phosphoshikimate 1-c ( 450) 264 68.2 5.8e-08  
gi|2072694|emb|CAB08328.1| 3-PHOSPHOSHIKIMATE 1-CA ( 450) 264 68.2 5.8e-08  
gi|121494765|emb|CAL73246.1| 3-phosphoshikimate 1- ( 450) 264 68.2 5.8e-08  
gi|124599370|gb|EAY58474.1| 3-phosphoshikimate 1-c ( 450) 264 68.2 5.8e-08  
gi|27651560|emb|CAD36892.1| putative 5-enolpyruvyl ( 122) 257 66.2 6.1e-08  
gi|27651518|emb|CAD36871.1| putative 5-enolpyruvyl ( 122) 257 66.2 6.1e-08  
gi|136265415|gb|EBM99606.1| hypothetical protein G ( 129) 257 66.2 6.3e-08  
gi|229419569|gb|EEO34616.1| 3-phosphoshikimate 1-c ( 422) 263 67.9 6.4e-08  
gi|137897728|gb|EBW63033.1| hypothetical protein G ( 112) 256 66.0 6.6e-08  
gi|141690760|gb|ECS82975.1| hypothetical protein G ( 166) 258 66.5 6.6e-08  
gi|135563245|gb|EBI44958.1| hypothetical protein G ( 321) 261 67.4 7e-08  
gi|27651556|emb|CAD36890.1| putative 5-enolpyruvyl ( 122) 256 66.0 7.1e-08  
gi|37359246|gb|AAN77867.1| 5-enolpyruvylshikimate- ( 330) 261 67.4 7.2e-08  
gi|49645101|emb|CAG98673.1| KLLA0F19712p [Kluyvero (1578) 269 69.7 7.2e-08  
gi|143123363|gb|EDC98470.1| hypothetical protein G ( 407) 262 67.7 7.2e-08  
gi|135856457|gb|EBK27747.1| hypothetical protein G ( 157) 257 66.3 7.4e-08  
gi|137691378|gb|EBV47987.1| hypothetical protein G ( 135) 256 66.0 7.6e-08  
gi|137063681|gb|EBS00190.1| hypothetical protein G ( 112) 255 65.8 7.7e-08  
gi|137064853|gb|EBS00850.1| hypothetical protein G ( 172) 257 66.3 7.9e-08  
gi|138955423|gb|ECC57612.1| hypothetical protein G ( 312) 260 67.2 8e-08  
gi|135501503|gb|EBI05440.1| hypothetical protein G ( 144) 256 66.1 8e-08  
gi|136331388|gb|EBN44372.1| hypothetical protein G ( 264) 259 66.9 8.1e-08  
gi|159887343|gb|ABX02280.1| 3-phosphoshikimate 1-c ( 429) 261 67.5 8.8e-08  
gi|138588318|gb|ECA63615.1| hypothetical protein G ( 170) 256 66.1 9.1e-08  
gi|219690818|gb|ACL32041.1| 3-phosphoshikimate 1-c ( 385) 260 67.3 9.4e-08  
gi|140006705|gb|ECI96016.1| hypothetical protein G ( 180) 256 66.1 9.5e-08  
gi|27651504|emb|CAD36864.1| putative 5-enolpyruvyl ( 122) 254 65.6 9.5e-08  
gi|135293978|gb|EBG68304.1| hypothetical protein G ( 397) 260 67.3 9.6e-08  
gi|150013290|gb|ABR55741.1| 3-phosphoshikimate 1-c ( 433) 260 67.3 1e-07  
gi|143336668|gb|EBE41053.1| hypothetical protein G ( 252) 257 66.5 1.1e-07  
gi|141584608|gb|ECS47173.1| hypothetical protein G ( 171) 255 65.9 1.1e-07  
gi|167273761|gb|ABZ26625.1| Sequence 563 from pate ( 462) 260 67.3 1.1e-07  
gi|6458826|gb|AAF10666.1|AE001959\_6 3-phosphoshiki ( 462) 260 67.3 1.1e-07  
gi|136534380|gb|EBO78463.1| hypothetical protein G ( 175) 255 65.9 1.1e-07  
gi|140047125|gb|ECJ22455.1| hypothetical protein G ( 119) 253 65.4 1.1e-07  
gi|134806237|gb|EBD55969.1| hypothetical protein G ( 176) 255 65.9 1.1e-07  
gi|134811217|gb|EBD59044.1| hypothetical protein G ( 470) 260 67.3 1.1e-07  
gi|141988236|gb|ECU65126.1| hypothetical protein G ( 102) 252 65.1 1.1e-07  
gi|136179936|gb|EBM42208.1| hypothetical protein G ( 177) 257 66.5 1.2e-07  
gi|125862039|gb|ABN52003.1| 3-phosphoshikimate 1-c ( 422) 259 67.1 1.2e-07  
gi|139775605|gb|ECH37721.1| hypothetical protein G ( 160) 254 65.7 1.2e-07  
gi|137831221|gb|EBW24864.1| hypothetical protein G ( 164) 254 65.7 1.2e-07  
gi|136423872|gb|EBR007160.1| hypothetical protein G ( 378) 258 66.8 1.3e-07  
gi|138145249|gb|EBX99526.1| hypothetical protein G ( 144) 253 65.4 1.3e-07  
gi|138516940|gb|ECA13844.1| hypothetical protein G ( 177) 254 65.7 1.3e-07  
gi|138175700|gb|EBY20773.1| hypothetical protein G ( 216) 255 66.0 1.3e-07  
gi|135401199|gb|EBH38962.1| hypothetical protein G ( 391) 258 66.8 1.3e-07  
gi|139864844|gb|ECI00209.1| hypothetical protein G ( 265) 256 66.3 1.3e-07  
gi|140133038|gb|ECJ76373.1| hypothetical protein G ( 100) 251 64.9 1.3e-07  
gi|135740077|gb|EBJ54747.1| hypothetical protein G ( 182) 254 65.7 1.3e-07  
gi|139813738|gb|ECH64771.1| hypothetical protein G ( 156) 253 65.4 1.3e-07  
gi|134563693|gb|EBC09142.1| hypothetical protein G ( 284) 256 66.3 1.4e-07  
gi|138206680|gb|EBY40866.1| hypothetical protein G ( 293) 256 66.3 1.4e-07  
gi|136514553|gb|EBO65808.1| hypothetical protein G ( 111) 251 64.9 1.4e-07  
gi|94555387|gb|ABF45301.1| 3-phosphoshikimate 1-ca ( 440) 258 66.9 1.4e-07  
gi|134606507|gb|EBC34706.1| hypothetical protein G ( 118) 251 64.9 1.5e-07  
gi|141081778|gb|ECP49687.1| hypothetical protein G ( 102) 250 64.6 1.5e-07  
gi|136650521|gb|EBP51121.1| hypothetical protein G ( 152) 252 65.2 1.5e-07  
gi|118164840|gb|ABK65737.1| 3-phosphoshikimate 1-c ( 419) 257 66.6 1.6e-07  
gi|150033408|gb|ABR65521.1| 3-phosphoshikimate 1-c ( 429) 257 66.6 1.6e-07  
gi|140638840|gb|ECM48477.1| hypothetical protein G ( 134) 251 65.0 1.6e-07

gi|137155531|gb|EBS51327.1| hypothetical protein G ( 166) 252 65.2 1.6e-07  
gi|138088436|gb|EBX67931.1| hypothetical protein G ( 93) 249 64.4 1.6e-07  
gi|31620004|emb|CAD95348.1| 3-PHOSPHOSHIKIMATE 1-C ( 450) 257 66.7 1.7e-07  
gi|140136930|gb|ECJ78639.1| hypothetical protein G ( 82) 248 64.1 1.7e-07  
gi|27651516|emb|CAD36870.1| putative 5-enolpyruvyl ( 122) 250 64.7 1.7e-07  
gi|141842424|gb|ECI763134.1| hypothetical protein G ( 165) 251 65.0 1.9e-07  
gi|139635409|gb|ECG41393.1| hypothetical protein G ( 171) 251 65.0 1.9e-07  
gi|141008389|gb|ECP01391.1| hypothetical protein G ( 255) 253 65.6 2e-07  
gi|137153403|gb|EBS50130.1| hypothetical protein G ( 97) 248 64.2 2e-07  
gi|143030018|gb|EDC30145.1| hypothetical protein G ( 269) 253 65.6 2e-07  
gi|140108183|gb|ECJ59174.1| hypothetical protein G ( 124) 249 64.5 2.1e-07  
gi|144018831|gb|EDI28211.1| hypothetical protein G ( 104) 248 64.2 2.1e-07  
gi|140449001|gb|ECL82971.1| hypothetical protein G ( 154) 250 64.8 2.1e-07  
gi|256581891|gb|ACU93026.1| 3-phosphoshikimate 1-c ( 409) 255 66.2 2.1e-07  
gi|110279818|gb|ABG58004.1| 3-phosphoshikimate 1-c ( 415) 255 66.2 2.1e-07  
gi|16751569|gb|AAL27698.1|AF371966\_1 5-enolpyruvyl ( 516) 256 66.5 2.2e-07  
gi|89144043|emb|CAJ79291.1| 3-phosphoshikimate 1-c ( 425) 255 66.2 2.2e-07  
gi|156252874|gb|ABU61380.1| 3-phosphoshikimate 1-c ( 425) 255 66.2 2.2e-07  
gi|151570774|gb|EDN36428.1| 3-phosphoshikimate 1-c ( 425) 255 66.2 2.2e-07  
gi|115129578|gb|ABI82765.1| 3-phosphoshikimate 1-c ( 425) 255 66.2 2.2e-07  
gi|134253354|gb|EBA52448.1| 3-phosphoshikimate 1-c ( 425) 255 66.2 2.2e-07  
gi|142143769|gb|ECV91152.1| hypothetical protein G ( 132) 249 64.5 2.2e-07  
gi|141952437|gb|ECU40153.1| hypothetical protein G ( 166) 250 64.8 2.2e-07  
gi|137279818|gb|EBS97769.1| hypothetical protein G ( 172) 250 64.8 2.3e-07  
gi|135861318|gb|EBN30804.1| hypothetical protein G ( 101) 247 64.0 2.4e-07  
gi|139853511|gb|ECH92492.1| hypothetical protein G ( 184) 250 64.8 2.4e-07  
gi|143596894|gb|EDF78687.1| hypothetical protein G ( 331) 253 65.7 2.4e-07  
gi|139510525|gb|ECF56688.1| hypothetical protein G ( 155) 249 64.6 2.4e-07  
gi|167294559|gb|ABZ47423.1| Sequence 21361 from pa ( 416) 254 66.0 2.5e-07  
gi|197053754|gb|ACH25452.1| Sequence 4 from patent ( 416) 254 66.0 2.5e-07  
gi|2649064|gb|AAB89746.1| 5-enolpyruvylshikimate 3 ( 416) 254 66.0 2.5e-07  
gi|137391662|gb|EBT83813.1| hypothetical protein G ( 108) 247 64.0 2.5e-07  
gi|118423587|gb|ABK89977.1| 3-phosphoshikimate 1-c ( 425) 254 66.0 2.5e-07  
gi|151572218|gb|EDN37872.1| 3-phosphoshikimate 1-c ( 425) 254 66.0 2.5e-07  
gi|134049878|gb|ABO46949.1| 3-phosphoshikimate 1-c ( 425) 254 66.0 2.5e-07  
gi|135475905|gb|EBH89025.1| hypothetical protein G ( 298) 252 65.4 2.6e-07  
gi|141430667|gb|ECR77456.1| hypothetical protein G ( 64) 244 63.2 2.6e-07  
gi|142239667|gb|ECW63816.1| hypothetical protein G ( 172) 249 64.6 2.6e-07  
gi|142647091|gb|ECZ56827.1| hypothetical protein G ( 117) 247 64.0 2.7e-07  
gi|137267092|gb|EBT13999.1| hypothetical protein G ( 117) 247 64.0 2.7e-07  
gi|141074770|gb|ECP44839.1| hypothetical protein G ( 150) 248 64.3 2.8e-07  
gi|141358911|gb|ECR27338.1| hypothetical protein G ( 231) 250 64.9 2.9e-07  
gi|220000157|gb|ACL76758.1| 3-phosphoshikimate 1-c ( 422) 253 65.8 2.9e-07  
gi|157122156|gb|EDO66296.1| 3-phosphoshikimate 1-c ( 425) 253 65.8 2.9e-07  
gi|254842865|gb|EET21279.1| 3-phosphoshikimate 1-c ( 426) 253 65.8 2.9e-07  
gi|167597744|gb|ABZ87742.1| 3-phosphoshikimate 1-c ( 426) 253 65.8 2.9e-07  
gi|136895041|gb|EBR10177.1| hypothetical protein G ( 75) 244 63.2 3e-07  
gi|143429820|gb|EDE90925.1| hypothetical protein G ( 165) 248 64.4 3e-07  
gi|138433557|gb|EBZ64253.1| hypothetical protein G ( 169) 248 64.4 3e-07  
gi|140986024|gb|ECO85762.1| hypothetical protein G ( 141) 247 64.1 3.1e-07  
gi|239838503|gb|ACS30300.1| 3-phosphoshikimate 1-c ( 476) 253 65.8 3.2e-07  
gi|138143615|gb|EBX98782.1| hypothetical protein G ( 272) 250 65.0 3.2e-07  
gi|142029705|gb|ECU98902.1| hypothetical protein G ( 107) 245 63.6 3.4e-07  
gi|116665787|gb|ABK14814.1| 3-phosphoshikimate 1-c ( 421) 252 65.5 3.4e-07  
gi|142847470|gb|EDB03074.1| hypothetical protein G ( 89) 244 63.3 3.4e-07  
gi|62786941|gb|AAY06820.1| Sequence 27 from patent ( 516) 253 65.8 3.4e-07  
gi|169191|gb|AAA33699.1| 5-enolpyruvylshikimate-3- ( 516) 253 65.8 3.4e-07  
gi|6732247|emb|CAB69241.1| unnamed protein product ( 516) 253 65.8 3.4e-07  
gi|56604207|emb|CAG45221.1| 3-phosphoshikimate 1-c ( 425) 252 65.6 3.4e-07  
gi|254840528|gb|EET18964.1| 3-phosphoshikimate 1-c ( 425) 252 65.6 3.4e-07  
gi|110320519|emb|CAL08604.1| 3-phosphoshikimate 1- ( 425) 252 65.6 3.4e-07  
gi|151568465|gb|EDN34119.1| hypothetical protein F ( 425) 252 65.6 3.4e-07

gi|58416522|emb|CAI27635.1| 3-phosphoshikimate 1-c ( 427) 252 65.6 3.4e-07  
gi|132663055|gb|ABO34701.1| 3-phosphoshikimate 1-c ( 429) 252 65.6 3.4e-07  
gi|1226317774|gb|ACO45770.1| putative 3-phosphoshik ( 446) 252 65.6 3.5e-07  
gi|15113313|gb|AAE68897.1| Sequence 4 from patent ( 446) 252 65.6 3.5e-07  
gi|139910522|gb|ECI30664.1| hypothetical protein G ( 171) 247 64.2 3.6e-07  
gi|135519158|gb|EBI16781.1| hypothetical protein G ( 142) 246 63.9 3.6e-07  
gi|142052471|gb|ECV19283.1| hypothetical protein G ( 117) 245 63.6 3.6e-07  
gi|218301283|emb|CAU98630.1| unnamed protein produ ( 463) 252 65.6 3.6e-07  
gi|218301280|emb|CAU98627.1| unnamed protein produ ( 468) 252 65.6 3.6e-07  
gi|138638232|gb|ECA97023.1| hypothetical protein G ( 150) 246 63.9 3.7e-07  
gi|237907951|gb|EEP82352.1| pentafunctional AROM p (1580) 258 67.3 3.8e-07  
gi|136951562|gb|EBR36833.1| hypothetical protein G ( 85) 243 63.1 3.8e-07  
gi|139588193|gb|ECG10061.1| hypothetical protein G ( 155) 246 63.9 3.8e-07  
gi|141936386|gb|ECU28820.1| hypothetical protein G ( 105) 244 63.4 3.8e-07  
gi|135097282|gb|EBF47442.1| hypothetical protein G ( 106) 244 63.4 3.9e-07  
gi|141196850|gb|ECQ28892.1| hypothetical protein G ( 129) 245 63.6 3.9e-07  
gi|138209136|gb|EBY42538.1| hypothetical protein G ( 160) 246 63.9 3.9e-07  
gi|48526086|gb|AAT45244.1| 5-enol-pyruvylshikimate ( 523) 252 65.6 4e-07  
gi|139526428|gb|ECF67652.1| hypothetical protein G ( 165) 246 63.9 4e-07  
gi|63334331|gb|AAY40472.1| 5-enol-pyruvylshikimate ( 447) 251 65.3 4.1e-07  
gi|139975788|gb|ECI75973.1| hypothetical protein G ( 210) 247 64.2 4.2e-07  
gi|134453802|gb|EBB43977.1| hypothetical protein G ( 119) 244 63.4 4.2e-07  
gi|139774290|gb|ECH36760.1| hypothetical protein G ( 120) 244 63.4 4.3e-07  
gi|141063525|gb|ECP37936.1| hypothetical protein G ( 319) 249 64.8 4.3e-07  
gi|167283499|gb|ABZ36363.1| Sequence 10301 from pa ( 392) 250 65.1 4.3e-07  
gi|138636289|gb|ECA95687.1| hypothetical protein G ( 150) 245 63.7 4.4e-07  
gi|219867914|gb|ACL48249.1| 3-phosphoshikimate 1-c ( 484) 251 65.4 4.4e-07  
gi|137707930|gb|EBV57186.1| hypothetical protein G ( 72) 241 62.6 4.5e-07  
gi|140359360|gb|ECL2139.1| hypothetical protein G ( 111) 243 63.2 4.7e-07  
gi|144974750|gb|ABP12461.1| Sequence 53 from paten ( 444) 250 65.1 4.7e-07  
gi|2485248|gb|AAB73384.1|I44473 Sequence 53 from p ( 444) 250 65.1 4.7e-07  
gi|2484171|gb|AAB72307.1|I49200 Sequence 53 from p ( 444) 250 65.1 4.7e-07  
gi|5957571|gb|AAE08245.1| Sequence 53 from patent ( 444) 250 65.1 4.7e-07  
gi|63334354|gb|AAY40473.1| 5-enol-pyruvylshikimate ( 447) 250 65.1 4.8e-07  
gi|142291970|gb|ECX02262.1| hypothetical protein G ( 116) 243 63.2 4.8e-07  
gi|138705584|gb|ECB44303.1| hypothetical protein G ( 172) 245 63.7 4.8e-07  
gi|136354624|gb|EBN60202.1| hypothetical protein G ( 99) 242 62.9 5e-07  
gi|137776675|gb|EBV93594.1| hypothetical protein G ( 264) 247 64.3 5e-07  
gi|139761003|gb|ECH27597.1| hypothetical protein G ( 154) 244 63.5 5.2e-07  
gi|135666238|gb|EBJ09113.1| hypothetical protein G ( 157) 244 63.5 5.2e-07  
gi|136491500|gb|EBO50896.1| hypothetical protein G ( 299) 247 64.3 5.5e-07  
gi|256579750|gb|ACU90886.1| 3-phosphoshikimate 1-c ( 448) 249 64.9 5.5e-07  
gi|143922947|gb|EDH60152.1| hypothetical protein G ( 96) 241 62.7 5.6e-07  
gi|116054912|emb|CAL56989.1| 5-enolpyruvylshikimat ( 316) 247 64.4 5.7e-07  
gi|145689058|gb|ABP89564.1| 5-enolpyruvylshikimate ( 151) 243 63.3 5.9e-07  
gi|475972|gb|AAA17839.1| pentafunctional enzyme [P (1581) 255 66.6 5.9e-07  
gi|141006545|gb|ECP00155.1| hypothetical protein G ( 280) 246 64.1 6.1e-07  
gi|197068384|gb|ACH29559.1| Sequence 178 from pate ( 129) 242 63.0 6.1e-07  
gi|140055797|gb|ECJ27025.1| hypothetical protein G ( 90) 240 62.4 6.2e-07  
gi|187723900|gb|ACD25121.1| 3-phosphoshikimate 1-c ( 433) 248 64.7 6.3e-07  
gi|143676556|gb|EDG21358.1| hypothetical protein G ( 443) 248 64.7 6.4e-07  
gi|218757794|gb|ACL08693.1| 3-phosphoshikimate 1-c ( 450) 248 64.7 6.5e-07  
gi|213504186|emb|CAS92804.1| unnamed protein produ ( 461) 248 64.7 6.6e-07  
gi|138379540|gb|EBZ26536.1| hypothetical protein G ( 118) 241 62.7 6.6e-07  
gi|134557324|gb|EBC05284.1| hypothetical protein G ( 212) 244 63.6 6.6e-07  
gi|137262334|gb|EBT11318.1| hypothetical protein G ( 99) 240 62.5 6.7e-07  
gi|170775096|gb|ACB33235.1| 3-phosphoshikimate 1-c ( 699) 250 65.3 6.7e-07  
gi|143445595|gb|EDE99941.1| hypothetical protein G ( 84) 239 62.2 6.9e-07  
gi|142394073|gb|ECX74727.1| hypothetical protein G ( 153) 242 63.0 6.9e-07  
gi|141867175|gb|ECT80548.1| hypothetical protein G ( 107) 240 62.5 7.1e-07  
gi|140342741|gb|ECL11664.1| hypothetical protein G ( 193) 243 63.3 7.1e-07  
gi|219567771|dbj|BAH05755.1| hypothetical protein ( 426) 247 64.5 7.2e-07

gi|146346304|gb|EDK32840.1| AroA [Clostridium kluy ( 426) 247 64.5 7.2e-07  
gi|45047634|emb|CAP30761.1| 3-phosphoshikimate-1-c ( 429) 247 64.5 7.2e-07  
gi|139588191|gb|ECC10059.1| hypothetical protein G ( 169) 242 63.1 7.5e-07  
gi|140213869|gb|ECK30436.1| hypothetical protein G ( 100) 239 62.2 7.9e-07  
gi|137313236|gb|EBT39796.1| hypothetical protein G ( 148) 241 62.8 7.9e-07  
gi|142092265|gb|ECV52480.1| hypothetical protein G ( 149) 241 62.8 7.9e-07  
gi|137181885|gb|EBS66113.1| hypothetical protein G ( 151) 241 62.8 8e-07  
gi|145017901|gb|EDK02180.1| hypothetical protein M (1575) 253 66.2 8e-07  
gi|241914647|gb|EER87791.1| hypothetical protein S ( 506) 247 64.5 8.2e-07  
gi|187712349|gb|ACD30646.1| 3-phosphoshikimate 1-c ( 425) 246 64.2 8.4e-07  
gi|139022251|gb|ECC86765.1| hypothetical protein G ( 161) 241 62.8 8.4e-07  
gi|394748|emb|CAA51291.1| 3-phosphoshikimate 1-car ( 92) 238 62.0 8.6e-07  
gi|136246949|gb|EBM87035.1| hypothetical protein G ( 115) 239 62.3 8.8e-07  
gi|144157096|gb|EDJ28064.1| hypothetical protein G ( 460) 246 64.3 8.9e-07  
gi|167294770|gb|ABZ247634.1| Sequence 21572 from pa ( 410) 245 64.0 9.5e-07  
gi|5457888|emb|CAB49378.1| aroA 3-phosphoshikimate ( 410) 245 64.0 9.5e-07  
gi|140786127|gb|ECN49774.1| hypothetical protein G ( 166) 240 62.6 1e-06  
gi|134899027|gb|EBE16415.1| hypothetical protein G ( 170) 240 62.6 1e-06  
gi|141358868|gb|ECD27311.1| hypothetical protein G ( 146) 239 62.4 1.1e-06  
gi|142798898|gb|EDA66645.1| hypothetical protein G ( 179) 240 62.7 1.1e-06  
gi|138386372|gb|EBZ31255.1| hypothetical protein G ( 222) 241 62.9 1.1e-06  
gi|137034169|gb|EBM83523.1| hypothetical protein G ( 127) 238 62.1 1.1e-06  
gi|139597835|gb|ECG16899.1| hypothetical protein G ( 157) 239 62.4 1.1e-06  
gi|140365320|gb|ECL27383.1| hypothetical protein G ( 107) 237 61.8 1.1e-06  
gi|50840349|gb|AAT83016.1| 3-phosphoshikimate 1-ca ( 429) 244 63.8 1.1e-06  
gi|139225737|gb|ECE27527.1| hypothetical protein G ( 299) 242 63.3 1.2e-06  
gi|63334403|gb|AAY40475.1| 5-enol-pyruvylshikimate ( 446) 244 63.8 1.2e-06  
gi|137956336|gb|EBW96154.1| hypothetical protein G ( 142) 238 62.1 1.2e-06  
gi|139920800|gb|ECI37946.1| hypothetical protein G ( 174) 239 62.4 1.2e-06  
gi|142194442|gb|ECW29770.1| hypothetical protein G ( 339) 242 63.3 1.3e-06  
gi|140436263|gb|ECL73891.1| hypothetical protein G ( 287) 241 63.0 1.3e-06  
gi|126226989|gb|ABN90529.1| 3-phosphoshikimate 1-c ( 435) 243 63.6 1.3e-06  
gi|157935341|gb|EDO91011.1| 3-phosphoshikimate 1-c ( 435) 243 63.6 1.3e-06  
gi|242139779|gb|EES26181.1| 3-phosphoshikimate 1-c ( 435) 243 63.6 1.3e-06  
gi|126220435|gb|ABN83941.1| 3-phosphoshikimate 1-c ( 435) 243 63.6 1.3e-06  
gi|157805890|gb|EDC83060.1| 3-phosphoshikimate 1-c ( 435) 243 63.6 1.3e-06  
gi|169653402|gb|EDS86095.1| 3-phosphoshikimate 1-c ( 435) 243 63.6 1.3e-06  
gi|237505533|gb|ACQ97851.1| 3-phosphoshikimate 1-c ( 435) 243 63.6 1.3e-06  
gi|136844649|gb|EBQ78580.1| hypothetical protein G ( 165) 238 62.2 1.3e-06  
gi|15113311|gb|AAE68895.1| Sequence 2 from patent ( 444) 243 63.6 1.4e-06  
gi|141003111|gb|ECN97872.1| hypothetical protein G ( 138) 237 61.9 1.4e-06  
gi|63334366|gb|AAY40474.1| 5-enol-pyruvylshikimate ( 446) 243 63.6 1.4e-06  
gi|137645515|gb|EBV22535.1| hypothetical protein G ( 99) 235 61.4 1.4e-06  
gi|178464133|dbj|BAG18653.1| putative 5-enolpyruvy ( 415) 242 63.4 1.5e-06  
gi|77862401|gb|ABB04469.1| 5-enolpyruvylshikimate ( 235) 239 62.5 1.5e-06  
gi|140728122|gb|ECN10699.1| hypothetical protein G ( 90) 234 61.1 1.5e-06  
gi|141920103|gb|ECU17301.1| hypothetical protein G ( 162) 237 62.0 1.5e-06  
gi|143084014|gb|EDC69721.1| hypothetical protein G ( 165) 237 62.0 1.6e-06  
gi|137626704|gb|EBV13083.1| hypothetical protein G ( 166) 237 62.0 1.6e-06  
gi|139813737|gb|ECH64770.1| hypothetical protein G ( 175) 237 62.0 1.6e-06  
gi|135591171|gb|EBI62617.1| hypothetical protein G ( 176) 237 62.0 1.6e-06  
gi|141562291|gb|ECS38975.1| hypothetical protein G ( 122) 235 61.4 1.7e-06  
gi|57158522|dbj|BAB84452.1| 5-enolpyruvylshikimate ( 399) 241 63.1 1.7e-06  
gi|141309321|gb|ECQ97030.1| hypothetical protein G ( 153) 236 61.7 1.7e-06  
gi|139756385|gb|ECH24426.1| hypothetical protein G ( 187) 237 62.0 1.7e-06  
gi|85821331|gb|EAQ42478.1| 3-phosphoshikimate 1-ca ( 409) 241 63.1 1.7e-06  
gi|218093860|emb|CAT71432.1| unnamed protein produ ( 415) 241 63.1 1.7e-06  
gi|197053757|gb|ACH25455.1| Sequence 7 from patent ( 424) 241 63.2 1.8e-06  
gi|19714500|gb|AAL95129.1| 3-phosphoshikimate 1-ca ( 424) 241 63.2 1.8e-06  
gi|161726851|emb|CAP47296.1| unnamed protein produ ( 424) 241 63.2 1.8e-06  
gi|135648868|gb|EBT98362.1| hypothetical protein G ( 133) 235 61.5 1.8e-06  
gi|229468460|gb|ACQ70232.1| UDP-N-acetylglucosamin ( 433) 241 63.2 1.8e-06

gi|2485249|gb|AAB73385.1|I44474 Sequence 54 from p ( 444) 241 63.2 1.8e-06  
gi|144974751|gb|ABP12462.1| Sequence 54 from paten ( 444) 241 63.2 1.8e-06  
gi|5957572|gb|AAE08246.1| Sequence 54 from patent ( 444) 241 63.2 1.8e-06  
gi|2484172|gb|AAB72308.1|I49201 Sequence 54 from p ( 444) 241 63.2 1.8e-06  
gi|48526088|gb|AAT45245.1| 5-enol-pyruvylshikimate ( 454) 241 63.2 1.9e-06  
gi|143836029|gb|EDG97541.1| hypothetical protein G ( 173) 236 61.8 1.9e-06  
gi|137935241|gb|EBW84278.1| hypothetical protein G ( 260) 238 62.3 1.9e-06  
gi|56569933|gb|AAV99867.1| Sequence 5 from patent ( 570) 242 63.5 1.9e-06  
gi|167249310|gb|ABZ25917.1| Sequence 5 from patent ( 570) 242 63.5 1.9e-06  
gi|134840320|gb|EBD77599.1| hypothetical protein G ( 147) 235 61.5 1.9e-06  
gi|134846736|gb|EBD81731.1| hypothetical protein G ( 275) 238 62.4 2e-06  
gi|142178523|gb|ECW17637.1| hypothetical protein G ( 87) 232 60.7 2e-06  
gi|135039257|gb|EBF10408.1| hypothetical protein G ( 342) 239 62.6 2e-06  
gi|140431938|gb|ECL70931.1| hypothetical protein G ( 295) 238 62.4 2.1e-06  
gi|139575354|gb|ECG01163.1| hypothetical protein G ( 245) 237 62.1 2.1e-06  
gi|141894557|gb|ECT99515.1| hypothetical protein G ( 140) 234 61.3 2.2e-06  
gi|139453944|gb|ECF21005.1| hypothetical protein G ( 97) 232 60.7 2.2e-06  
gi|141040121|gb|ECP22036.1| hypothetical protein G ( 213) 236 61.8 2.2e-06  
gi|143575554|gb|EDF71579.1| hypothetical protein G ( 145) 234 61.3 2.2e-06  
gi|141074771|gb|ECP44840.1| hypothetical protein G ( 147) 234 61.3 2.3e-06  
gi|57117457|gb|AAW33954.1| AROM pentafunctional en (1539) 246 64.7 2.3e-06  
gi|238032531|emb|CAY70554.1| Pentafunctional arom (1545) 246 64.7 2.3e-06  
gi|223643106|emb|CAX41980.1| pentafunctional AROM (1550) 246 64.7 2.3e-06  
gi|136243804|gb|EBM84904.1| hypothetical protein G ( 85) 231 60.5 2.3e-06  
gi|138377687|gb|EBZ25278.1| hypothetical protein G ( 154) 234 61.3 2.3e-06  
gi|143289772|gb|EDE14254.1| hypothetical protein G ( 415) 239 62.7 2.4e-06  
gi|124417311|emb|CAK63956.1| unnamed protein produ (1118) 244 64.1 2.4e-06  
gi|13375567|gb|AAK20397.1|AF349754\_1 5-enolpyruvyl ( 347) 238 62.4 2.4e-06  
gi|135418683|gb|EBH50720.1| hypothetical protein G ( 425) 239 62.7 2.4e-06  
gi|143619097|gb|EDF89452.1| hypothetical protein G ( 350) 238 62.4 2.4e-06  
gi|137839327|gb|EBW29556.1| hypothetical protein G ( 110) 232 60.8 2.4e-06  
gi|137771645|gb|EBV90851.1| hypothetical protein G ( 246) 236 61.9 2.5e-06  
gi|136810426|gb|EBQ55796.1| hypothetical protein G ( 142) 233 61.1 2.6e-06  
gi|140029257|gb|ECJ11191.1| hypothetical protein G ( 98) 231 60.5 2.6e-06  
gi|167324427|gb|ABZ61020.1| Sequence 11819 from pa ( 392) 238 62.5 2.6e-06  
gi|135696355|gb|EBJ27660.1| hypothetical protein G ( 151) 233 61.1 2.7e-06  
gi|140710291|gb|ECM98191.1| hypothetical protein G ( 152) 233 61.1 2.7e-06  
gi|134923875|gb|EBB32883.1| hypothetical protein G ( 103) 231 60.5 2.7e-06  
gi|139178037|gb|ECD94123.1| hypothetical protein G ( 274) 236 61.9 2.7e-06  
gi|139483444|gb|ECF39330.1| hypothetical protein G ( 104) 231 60.5 2.7e-06  
gi|143131822|gb|EDD04677.1| hypothetical protein G ( 408) 238 62.5 2.7e-06  
gi|138345236|gb|EBZ10124.1| hypothetical protein G ( 155) 233 61.1 2.7e-06  
gi|184212267|gb|EDU09310.1| 3-phosphoshikimate 1-c ( 435) 238 62.5 2.8e-06  
gi|135670534|gb|EBJ11760.1| hypothetical protein G ( 111) 231 60.5 2.9e-06  
gi|5957573|gb|AAE08247.1| Sequence 55 from patent ( 444) 238 62.5 2.9e-06  
gi|2485250|gb|AAB73386.1|I44475 Sequence 55 from p ( 444) 238 62.5 2.9e-06  
gi|15113310|gb|AAE68894.1| Sequence 1 from patent ( 444) 238 62.5 2.9e-06  
gi|144974752|gb|ABP12463.1| Sequence 55 from paten ( 444) 238 62.5 2.9e-06  
gi|2484173|gb|AAB72309.1|I49202 Sequence 55 from p ( 444) 238 62.5 2.9e-06  
gi|138846965|gb|ECC13935.1| hypothetical protein G ( 205) 234 61.4 2.9e-06  
gi|134661181|gb|EBC67295.1| hypothetical protein G ( 210) 234 61.4 3e-06  
gi|140455535|gb|ECL87542.1| hypothetical protein G ( 259) 235 61.7 3e-06  
gi|140103800|gb|ECJ56094.1| hypothetical protein G ( 147) 232 60.8 3e-06  
gi|144061091|gb|EDI58614.1| hypothetical protein G ( 72) 228 59.7 3.2e-06  
gi|2485258|gb|AAB73394.1|I44483 Sequence 63 from p ( 426) 237 62.3 3.3e-06  
gi|144974760|gb|ABP12471.1| Sequence 63 from paten ( 426) 237 62.3 3.3e-06  
gi|141872|gb|AAA21937.1| 3-phosphoshikimate-1-carb ( 426) 237 62.3 3.3e-06  
gi|2484181|gb|AAB72317.1|I49210 Sequence 63 from p ( 426) 237 62.3 3.3e-06  
gi|5957581|gb|AAE08255.1| Sequence 63 from patent ( 426) 237 62.3 3.3e-06  
gi|48526084|gb|AAT45243.1| 5-enol-pyruvylshikimate ( 519) 238 62.6 3.3e-06  
gi|217229978|gb|ACK11087.1| Sequence 412 from pate ( 430) 237 62.3 3.3e-06  
gi|140844847|dbj|BAF53846.1| hypothetical protein ( 430) 237 62.3 3.3e-06

gi|12541842|emb|CAC25305.1| unnamed protein produc ( 430) 237 62.3 3.3e-06  
gi|21323530|dbj|BAB98157.1| 5-enolpyruvylshikimate ( 430) 237 62.3 3.3e-06  
gi|41324987|emb|CAF19468.1| 3-PHOSPHOSHIKIMATE 1-c ( 430) 237 62.3 3.3e-06  
gi|137677308|gb|EBV40060.1| hypothetical protein G ( 112) 230 60.3 3.3e-06  
gi|135107136|gb|EBF53752.1| hypothetical protein G ( 441) 237 62.3 3.3e-06  
gi|140465850|gb|ECL91459.1| hypothetical protein G ( 302) 235 61.7 3.4e-06  
gi|116611308|gb|ABK04032.1| 3-phosphoshikimate 1-c ( 460) 237 62.3 3.5e-06  
gi|138456986|gb|EBZ80350.1| hypothetical protein G ( 174) 232 60.9 3.5e-06  
gi|141550051|gb|ECC30458.1| hypothetical protein G ( 144) 231 60.6 3.5e-06  
gi|182908332|gb|ACC04216.1| Sequence 4351 from pat ( 471) 237 62.3 3.5e-06  
gi|134675150|gb|EBC75402.1| hypothetical protein G ( 83) 228 59.8 3.6e-06  
gi|140849292|gb|ECN91757.1| hypothetical protein G ( 268) 234 61.5 3.6e-06  
gi|135805587|gb|EBJ95757.1| hypothetical protein G ( 102) 229 60.1 3.6e-06  
gi|135939139|gb|EBK84238.1| hypothetical protein G ( 332) 235 61.8 3.6e-06  
gi|135040418|gb|EBF11141.1| hypothetical protein G ( 335) 235 61.8 3.7e-06  
gi|42235097|gb|AAD13108.1| 5-enolpyruvylshikimate 3 ( 498) 237 62.3 3.7e-06  
gi|139559348|gb|ECF90296.1| hypothetical protein G ( 105) 229 60.1 3.7e-06  
gi|229431757|gb|EEC041969.1| 3-phosphoshikimate 1-c ( 424) 236 62.1 3.8e-06  
gi|256035894|gb|ACU59438.1| 3-phosphoshikimate 1-c ( 426) 236 62.1 3.8e-06  
gi|142748550|gb|EDA29469.1| hypothetical protein G ( 198) 232 60.9 3.8e-06  
gi|139849794|gb|ECH89899.1| hypothetical protein G ( 163) 231 60.7 3.8e-06  
gi|254219555|gb|EET08939.1| 3-phosphoshikimate 1-c ( 435) 236 62.1 3.9e-06  
gi|52208741|emb|CAH34677.1| putative 3-phosphoskim ( 435) 236 62.1 3.9e-06  
gi|76579593|gb|ABA49068.1| putative 3-phosphoskimi ( 435) 236 62.1 3.9e-06  
gi|138831381|gb|ECC06430.1| hypothetical protein G ( 139) 230 60.4 3.9e-06  
gi|124402111|emb|CAK67586.1| unnamed protein produ (1468) 242 63.8 4e-06  
gi|135710441|gb|EBJ36390.1| hypothetical protein G ( 116) 229 60.1 4e-06  
gi|139820928|gb|ECH69483.1| hypothetical protein G ( 116) 229 60.1 4e-06  
gi|138910924|gb|ECC38856.1| hypothetical protein G ( 103) 228 59.9 4.2e-06  
gi|137781153|gb|EBV96078.1| hypothetical protein G ( 154) 230 60.4 4.3e-06  
gi|155262583|gb|ABT18187.1| Sequence 105657 from p ( 411) 235 61.8 4.3e-06  
gi|197709166|gb|ACH72672.1| 5-enolpyruvylshikimate ( 510) 236 62.1 4.4e-06  
gi|260198061|gb|EEW95577.1| 3-phosphoshikimate 1-c ( 421) 235 61.8 4.4e-06  
gi|268614908|gb|ACZ09276.1| 3-phosphoshikimate 1-c ( 424) 235 61.8 4.4e-06  
gi|207087579|gb|EDZ64862.1| 3-phosphoshikimate 1-c ( 432) 235 61.9 4.5e-06  
gi|137256208|gb|EBT07853.1| hypothetical protein G ( 166) 230 60.5 4.5e-06  
gi|141225757|gb|ECQ49018.1| hypothetical protein G ( 113) 228 59.9 4.5e-06  
gi|257796482|gb|ACV67419.1| 3-phosphoshikimate 1-c ( 444) 235 61.9 4.6e-06  
gi|136153838|gb|EBM24965.1| hypothetical protein G ( 173) 230 60.5 4.7e-06  
gi|137080022|gb|EBS09377.1| hypothetical protein G ( 121) 228 59.9 4.8e-06  
gi|139813261|gb|ECH64420.1| hypothetical protein G ( 104) 227 59.6 5e-06  
gi|143421721|gb|EDE86769.1| hypothetical protein G ( 352) 233 61.4 5.1e-06  
gi|261377205|gb|ACX79948.1| UDP-N-acetylglucosamin ( 434) 234 61.6 5.2e-06  
gi|147750605|gb|EDK57674.1| 3-phosphoshikimate 1-c ( 435) 234 61.6 5.2e-06  
gi|121228841|gb|ABM51359.1| putative 3-phosphoshik ( 435) 234 61.6 5.2e-06  
gi|147745265|gb|EDK52345.1| 3-phosphoshikimate 1-c ( 435) 234 61.6 5.2e-06  
gi|148029459|gb|EDK87364.1| 3-phosphoshikimate 1-c ( 435) 234 61.6 5.2e-06  
gi|126242614|gb|ABO05707.1| putative 3-phosphoshik ( 435) 234 61.6 5.2e-06  
gi|124291405|gb|ABN00674.1| putative 3-phosphoshik ( 435) 234 61.6 5.2e-06  
gi|160697166|gb|EDP87886.1| putative 3-phosphoshik ( 435) 234 61.6 5.2e-06  
gi|141801247|gb|ECT35781.1| hypothetical protein G ( 165) 229 60.2 5.2e-06  
gi|138609960|gb|ECA78837.1| hypothetical protein G ( 299) 232 61.1 5.3e-06  
gi|135886415|gb|EBK47897.1| hypothetical protein G ( 168) 229 60.2 5.3e-06  
gi|142813513|gb|EDA77564.1| hypothetical protein G ( 303) 232 61.1 5.3e-06  
gi|136500904|gb|EB057015.1| hypothetical protein G ( 370) 233 61.4 5.3e-06  
gi|50951139|gb|AAT88840.1| UDP-N-acetylglucosamine ( 456) 234 61.7 5.4e-06  
gi|139807552|gb|ECH60329.1| hypothetical protein G ( 173) 229 60.2 5.4e-06  
gi|142393793|gb|ECX74526.1| hypothetical protein G ( 466) 234 61.7 5.5e-06  
gi|142028895|gb|ECU98128.1| hypothetical protein G ( 392) 233 61.4 5.6e-06  
gi|143729176|gb|EDG48487.1| hypothetical protein G ( 338) 232 61.1 5.8e-06  
gi|135290345|gb|EBG66194.1| hypothetical protein G ( 132) 227 59.7 6e-06  
gi|48526066|gb|AAT45234.1| 5-enol-pyruvylshikimate ( 357) 232 61.1 6e-06



gi|229565672|gb|ACQ79523.1| 3-phosphoshikimate 1-c ( 434) 233 61.4 6e-06  
gi|48526070|gb|AAT45236.1| 5-enol-pyruvylshikimate ( 358) 232 61.1 6.1e-06  
gi|48526072|gb|AAT45237.1| 5-enol-pyruvylshikimate ( 360) 232 61.1 6.1e-06  
gi|136968507|gb|EBR46423.1| hypothetical protein G ( 113) 226 59.5 6.1e-06  
gi|136974431|gb|EBR49789.1| hypothetical protein G ( 113) 226 59.5 6.1e-06  
gi|143196767|gb|EDD51467.1| hypothetical protein G ( 64) 223 58.6 6.2e-06  
gi|137498472|gb|EBU42283.1| hypothetical protein G ( 64) 223 58.6 6.2e-06  
gi|142229932|gb|ECW56578.1| hypothetical protein G ( 144) 227 59.8 6.4e-06  
gi|141533718|gb|ECS18963.1| hypothetical protein G ( 148) 227 59.8 6.5e-06  
gi|135733489|gb|EBJ50660.1| hypothetical protein G ( 123) 226 59.5 6.6e-06  
gi|30268138|emb|CAC82655.1| 5-enolpyruvylshikimate ( 409) 232 61.2 6.7e-06  
gi|149770891|emb|CAL42356.1| 3-phosphoshikimate 1- ( 409) 232 61.2 6.7e-06  
gi|143159253|gb|EDD24485.1| hypothetical protein G ( 279) 230 60.6 6.8e-06  
gi|135588646|gb|EBI61063.1| hypothetical protein G ( 91) 224 59.0 7e-06  
gi|143663997|gb|EDG14715.1| hypothetical protein G ( 112) 225 59.2 7.1e-06  
gi|228266275|gb|ACP96948.1| Sequence 29166 from pa ( 535) 233 61.5 7.1e-06  
gi|2484174|gb|AAB72310.1|I49203 Sequence 56 from p ( 444) 232 61.2 7.2e-06  
gi|144974753|gb|ABP12464.1| Sequence 56 from paten ( 444) 232 61.2 7.2e-06  
gi|2485251|gb|AAB73387.1|I44476 Sequence 56 from p ( 444) 232 61.2 7.2e-06  
gi|5957574|gb|AAE08248.1| Sequence 56 from patent ( 444) 232 61.2 7.2e-06  
gi|139938554|gb|EC150388.1| hypothetical protein G ( 82) 223 58.7 7.5e-06  
gi|136514494|gb|EBD65771.1| hypothetical protein G ( 270) 229 60.4 7.7e-06  
gi|167042884|gb|ABZ07600.1| putative EPSP synthase ( 183) 227 59.8 7.7e-06  
gi|138146448|gb|EBY00346.1| hypothetical protein G ( 153) 226 59.6 7.8e-06  
gi|142514368|gb|ECY62998.1| hypothetical protein G ( 105) 224 59.0 7.9e-06  
gi|136927795|gb|EBR23477.1| hypothetical protein G ( 157) 226 59.6 7.9e-06  
gi|135571691|gb|EBI50372.1| hypothetical protein G ( 232) 228 60.1 7.9e-06  
gi|138740714|gb|ECB68755.1| hypothetical protein G ( 90) 223 58.7 8.1e-06  
gi|134894270|gb|EBE13265.1| hypothetical protein G ( 135) 225 59.3 8.2e-06  
gi|213504182|emb|CAS92802.1| unnamed protein produ ( 464) 231 61.0 8.6e-06  
gi|138887693|gb|ECC30776.1| hypothetical protein G ( 146) 225 59.3 8.7e-06  
gi|136717696|gb|EBP94335.1| hypothetical protein G ( 225) 227 59.9 9e-06  
gi|123965218|gb|ABE77393.4| EPSP synthase [Allium ( 522) 231 61.0 9.4e-06  
gi|21886542|emb|CAD42174.1| unnamed protein produc ( 444) 230 60.8 9.7e-06  
gi|1524383|emb|CAA44974.1| 3-phosphoshikimate 1-ca ( 444) 230 60.8 9.7e-06  
gi|15113314|gb|AAE68898.1| Sequence 5 from patent ( 444) 230 60.8 9.7e-06  
gi|3714738|emb|CAA03538.1| unnamed protein product ( 444) 230 60.8 9.7e-06  
gi|20226493|gb|AAE88870.1| Sequence 3 from patent ( 444) 230 60.8 9.7e-06  
gi|62786942|gb|AAY06821.1| Sequence 28 from patent ( 444) 230 60.8 9.7e-06  
gi|197053767|gb|ACH25465.1| Sequence 17 from paten ( 444) 230 60.8 9.7e-06  
gi|3714672|gb|CAA03525.1| unnamed protein product ( 444) 230 60.8 9.7e-06  
gi|33724442|gb|AAQ32651.1| Sequence 3 from patent ( 444) 230 60.8 9.7e-06  
gi|4774185|emb|CAB42493.1| unnamed protein product ( 445) 230 60.8 9.7e-06  
gi|136124879|gb|EBM08135.1| hypothetical protein G ( 80) 221 58.3 1e-05  
gi|187973283|gb|EDU40782.1| pentafunctional AROM p ( 864) 233 61.6 1e-05  
gi|143177389|gb|EDD37310.1| hypothetical protein G ( 399) 229 60.5 1e-05  
gi|238936984|emb|CAR25163.1| KLTH0G13090p [Lachanc (1579) 236 62.5 1e-05  
gi|137729195|gb|EBV68871.1| hypothetical protein G ( 226) 226 59.7 1.1e-05  
gi|142533801|gb|ECY77123.1| hypothetical protein G ( 106) 222 58.6 1.1e-05  
gi|251836138|gb|EES64675.1| 3-phosphoshikimate 1-c ( 423) 229 60.5 1.1e-05  
gi|141298247|gb|ECQ92973.1| hypothetical protein G ( 139) 223 58.9 1.1e-05  
gi|143940857|gb|EDH73332.1| hypothetical protein G ( 371) 228 60.3 1.1e-05  
gi|137643716|gb|EBV21662.1| hypothetical protein G ( 95) 221 58.3 1.1e-05  
gi|136595080|gb|EBP17208.1| hypothetical protein G ( 472) 229 60.6 1.2e-05  
gi|135654047|gb|EBJ01553.1| hypothetical protein G ( 122) 222 58.6 1.2e-05  
gi|226520702|gb|ACO66690.1| predicted protein [Mic ( 229) 225 59.5 1.2e-05  
gi|144033885|gb|EDI38787.1| hypothetical protein G ( 155) 223 58.9 1.2e-05  
gi|142606038|gb|EC227895.1| hypothetical protein G ( 413) 228 60.3 1.2e-05  
gi|14549196|dbj|BAB61062.1| 3-phosphoshikimate 1-c ( 511) 229 60.6 1.3e-05  
gi|15778436|gb|AAL07437.1|AF413082\_1 EPSP synthase ( 511) 229 60.6 1.3e-05  
gi|15724392|gb|AAL06593.1| 5-enolpyruvylshikimate ( 511) 229 60.6 1.3e-05  
gi|254256567|emb|CAZ91430.1| unnamed protein produ ( 511) 229 60.6 1.3e-05

gi|125553934|gb|EAY99539.1| hypothetical protein O ( 515) 229 60.6 1.3e-05  
gi|113594752|dbj|BAF18626.1| Os06g0133900 [Oryza s ( 515) 229 60.6 1.3e-05  
gi|55297191|dbj|BAD68865.1| putative 5-enolpyruvyl ( 515) 229 60.6 1.3e-05  
gi|55296169|dbj|BAD67887.1| putative 5-enolpyruvyl ( 515) 229 60.6 1.3e-05  
gi|143209192|gb|EDD60527.1| hypothetical protein G ( 357) 227 60.1 1.3e-05  
gi|239808863|gb|ACS25928.1| UDP-N-acetylglucosamin ( 435) 228 60.3 1.3e-05  
gi|138035180|gb|EBX38754.1| hypothetical protein G ( 136) 222 58.6 1.3e-05  
gi|140645644|gb|ECM53119.1| hypothetical protein G ( 297) 226 59.8 1.3e-05  
gi|5957569|gb|AAE08243.1| Sequence 51 from patent ( 444) 228 60.3 1.3e-05  
gi|144974748|gb|ABP12459.1| Sequence 51 from paten ( 444) 228 60.3 1.3e-05  
gi|2484169|gb|AAB72305.1|I49198 Sequence 51 from p ( 444) 228 60.3 1.3e-05  
gi|2485246|gb|AAB73382.1|I44471 Sequence 51 from p ( 444) 228 60.3 1.3e-05  
gi|138952139|gb|ECC56131.1| hypothetical protein G ( 141) 222 58.7 1.3e-05  
gi|140094177|gb|ECJ49446.1| hypothetical protein G ( 142) 222 58.7 1.3e-05  
gi|143692754|gb|EDG31345.1| hypothetical protein G ( 260) 225 59.5 1.4e-05  
gi|143204891|gb|EDD57396.1| hypothetical protein G ( 389) 227 60.1 1.4e-05  
gi|140662460|gb|ECM64704.1| hypothetical protein G ( 148) 222 58.7 1.4e-05  
gi|210071512|gb|EEA25601.1| 3-dehydroquinate synth (1571) 234 62.1 1.4e-05  
gi|143435517|gb|EDP94112.1| hypothetical protein G ( 332) 226 59.8 1.4e-05  
gi|139701374|gb|ECG87541.1| hypothetical protein G ( 154) 222 58.7 1.4e-05  
gi|256718776|gb|EEU32331.1| 3-phosphoshikimate 1-c ( 421) 227 60.1 1.5e-05  
gi|229429522|gb|EEO39734.1| 3-phosphoshikimate 1-c ( 424) 227 60.1 1.5e-05  
gi|512545|emb|CAA01426.1| aroA [Aeromonas salmonic ( 427) 227 60.1 1.5e-05  
gi|141110776|gb|ECH2701345.1| hypothetical protein G ( 163) 222 58.7 1.5e-05  
gi|15113312|gb|AAE68896.1| Sequence 3 from patent ( 444) 227 60.1 1.5e-05  
gi|144974749|gb|ABP12460.1| Sequence 52 from paten ( 444) 227 60.1 1.5e-05  
gi|5957570|gb|AAE08244.1| Sequence 52 from patent ( 444) 227 60.1 1.5e-05  
gi|2484170|gb|AAB72306.1|I49199 Sequence 52 from p ( 444) 227 60.1 1.5e-05  
gi|2485247|gb|AAB73383.1|I44472 Sequence 52 from p ( 444) 227 60.1 1.5e-05  
gi|134867908|gb|EBD95726.1| hypothetical protein G ( 369) 226 59.8 1.5e-05  
gi|136143958|gb|EBM18538.1| hypothetical protein G ( 370) 226 59.8 1.5e-05  
gi|134406690|gb|EBB16764.1| hypothetical protein G ( 95) 219 57.9 1.5e-05  
gi|139764010|gb|ECH29671.1| hypothetical protein G ( 143) 221 58.4 1.6e-05  
gi|134880266|gb|EBE03915.1| hypothetical protein G ( 399) 226 59.9 1.6e-05  
gi|4488967|emb|CAD29607.2| pentafunctional arom p (1582) 233 61.8 1.6e-05  
gi|134600018|gb|EBC30786.1| hypothetical protein G ( 275) 224 59.3 1.7e-05  
gi|159131499|gb|EDP56612.1| pentafunctional polype (1605) 233 61.9 1.7e-05  
gi|56909232|dbj|BAD63759.1| 3-phosphoshikimate 1-c ( 446) 226 59.9 1.8e-05  
gi|138065413|gb|EBX54745.1| hypothetical protein G ( 64) 216 57.1 1.8e-05  
gi|141325310|gb|ECR05886.1| hypothetical protein G ( 95) 218 57.7 1.8e-05  
gi|140614574|gb|ECM38495.1| hypothetical protein G ( 314) 224 59.4 1.8e-05  
gi|135651047|gb|EBI99696.1| hypothetical protein G ( 100) 218 57.7 1.9e-05  
gi|121554705|gb|ABM58854.1| 3-phosphoshikimate 1-c ( 705) 228 60.5 1.9e-05  
gi|170934465|gb|ACB39726.1| 3-phosphoshikimate 1-c ( 399) 225 59.7 1.9e-05  
gi|140240697|gb|ECK48383.1| hypothetical protein G ( 273) 223 59.1 1.9e-05  
gi|143010366|gb|EDC15947.1| hypothetical protein G ( 406) 225 59.7 1.9e-05  
gi|56381718|dbj|BAD77626.1| UDP-N-acetylglucosamin ( 434) 225 59.7 2e-05  
gi|143337878|gb|EDE41942.1| hypothetical protein G ( 165) 220 58.3 2e-05  
gi|139041852|gb|ECD00736.1| hypothetical protein G ( 165) 220 58.3 2e-05  
gi|3714740|emb|CAA03539.1| unnamed protein product ( 444) 225 59.7 2.1e-05  
gi|33724443|gb|AAQ32652.1| Sequence 5 from patent ( 444) 225 59.7 2.1e-05  
gi|21886544|emb|CAD42175.1| unnamed protein produc ( 444) 225 59.7 2.1e-05  
gi|20226494|gb|AAE88871.1| Sequence 5 from patent ( 444) 225 59.7 2.1e-05  
gi|3714674|emb|CAA03526.1| unnamed protein product ( 444) 225 59.7 2.1e-05  
gi|4774187|emb|CAB42494.1| unnamed protein product ( 445) 225 59.7 2.1e-05  
gi|141888249|gb|ECT95030.1| hypothetical protein G ( 115) 218 57.7 2.1e-05  
gi|262272507|gb|ACY40415.1| 3-phosphoshikimate 1-c ( 385) 224 59.4 2.1e-05  
gi|117576868|emb|CAL65337.1| 3-phosphoshikimate 1- ( 405) 224 59.4 2.2e-05  
gi|137372418|gb|EBT72999.1| hypothetical protein G ( 128) 218 57.8 2.3e-05  
gi|111979103|gb|ABH83340.1| Sequence 9 from patent ( 506) 225 59.7 2.3e-05  
gi|281079465|gb|ADA36173.1| Sequence 8 from patent ( 506) 225 59.7 2.3e-05  
gi|138445397|gb|EBZ72304.1| hypothetical protein G ( 76) 215 56.9 2.4e-05

gi|141196760|gb|ECQ28828.1| hypothetical protein G ( 137) 218 57.8 2.4e-05  
gi|142492368|gb|ECY47510.1| hypothetical protein G ( 147) 218 57.8 2.5e-05  
gi|119412546|gb|EAW22487.1| pentafunctional polypeptide (1578) 230 61.2 2.6e-05  
gi|49529154|emb|CAG62820.1| unnamed protein product (1579) 230 61.2 2.6e-05  
gi|140435310|gb|ECL73227.1| hypothetical protein G ( 125) 217 57.5 2.6e-05  
gi|281108576|gb|ADA52834.1| Sequence 9 from patent ( 506) 224 59.5 2.6e-05  
gi|139194383|gb|ECE05764.1| hypothetical protein G ( 131) 217 57.5 2.7e-05  
gi|281079463|gb|ADA36171.1| Sequence 1 from patent ( 515) 224 59.5 2.7e-05  
gi|138383307|gb|EBZ29157.1| hypothetical protein G ( 133) 217 57.6 2.7e-05  
gi|134887116|gb|EBE08499.1| hypothetical protein G ( 136) 217 57.6 2.8e-05  
gi|139524990|gb|ECF66675.1| hypothetical protein G ( 94) 215 57.0 2.8e-05  
gi|119950797|gb|ABM09708.1| putative UDP-N-acetylglucosaminase ( 467) 223 59.3 2.9e-05  
gi|140099644|gb|ECJ53238.1| hypothetical protein G ( 269) 220 58.4 3e-05  
gi|135556587|gb|EBI40658.1| hypothetical protein G ( 151) 217 57.6 3e-05  
gi|210064464|gb|EEA18560.1| pentafunctional polypeptide (1573) 229 61.0 3e-05  
gi|44985944|gb|AAS54555.1| AGR066Wp [Ashbya gossypii] (1577) 229 61.0 3e-05  
gi|136069640|gb|EBL70519.1| hypothetical protein G ( 154) 217 57.6 3e-05  
gi|137127816|gb|EBS35887.1| hypothetical protein G ( 133) 216 57.3 3.2e-05  
gi|139064301|gb|ECD16501.1| hypothetical protein G ( 254) 219 58.2 3.3e-05  
gi|144065936|gb|EDI62214.1| hypothetical protein G ( 121) 215 57.1 3.4e-05  
gi|144189617|gb|EDJ52210.1| hypothetical protein G ( 180) 217 57.6 3.4e-05  
gi|135507166|gb|EBI09056.1| hypothetical protein G ( 101) 214 56.8 3.4e-05  
gi|137552102|gb|EBU70915.1| hypothetical protein G ( 124) 215 57.1 3.5e-05  
gi|141443548|gb|ECR86688.1| hypothetical protein G ( 153) 216 57.4 3.5e-05  
gi|225214557|gb|ACN83291.1| 3-phosphoshikimate 1-c ( 420) 221 58.8 3.6e-05  
gi|143026908|gb|EDC27859.1| hypothetical protein G ( 159) 216 57.4 3.6e-05  
gi|139610443|gb|ECG25482.1| hypothetical protein G ( 286) 219 58.2 3.6e-05  
gi|136654155|gb|EBP53271.1| hypothetical protein G ( 348) 220 58.5 3.6e-05  
gi|136459333|gb|EBO30109.1| hypothetical protein G ( 244) 218 58.0 3.7e-05  
gi|136049990|gb|EBL57228.1| hypothetical protein G ( 202) 217 57.7 3.7e-05  
gi|137722143|gb|EBV64909.1| hypothetical protein G ( 115) 214 56.9 3.8e-05  
gi|141676308|gb|ECS75854.1| hypothetical protein G ( 151) 215 57.2 4e-05  
gi|137478171|gb|EBU31918.1| hypothetical protein G ( 154) 215 57.2 4.1e-05  
gi|134912558|gb|EBE25356.1| hypothetical protein G ( 128) 214 56.9 4.1e-05  
gi|138640927|gb|ECA98853.1| hypothetical protein G ( 171) 215 57.2 4.5e-05  
gi|143907468|gb|EDH49131.1| hypothetical protein G ( 209) 216 57.5 4.5e-05  
gi|143364680|gb|EDE57647.1| hypothetical protein G ( 123) 213 56.7 4.7e-05  
gi|136692287|gb|EBP77687.1| hypothetical protein G ( 273) 217 57.8 4.7e-05  
gi|138388666|gb|EBZ32841.1| hypothetical protein G ( 127) 213 56.7 4.8e-05  
gi|255342376|gb|ACU08489.1| 5-Enolpyruvylshikimate ( 410) 219 58.4 4.8e-05  
gi|269095297|gb|ACC25288.1| 3-phosphoshikimate 1-c ( 422) 219 58.4 4.9e-05  
gi|51858127|dbj|BAD42285.1| UDP-N-acetylglucosaminase ( 424) 219 58.4 4.9e-05  
gi|141883580|gb|ECT91843.1| hypothetical protein G ( 241) 216 57.5 5e-05  
gi|113376729|gb|AAZ79230.2| plastid 5-enolpyruvylshikimate 3-phosphoshikimate 1-c ( 437) 219 58.4 5e-05  
gi|239795036|dbj|BAH74025.1| 3-phosphoshikimate 1-c ( 445) 219 58.4 5.1e-05  
gi|40557115|gb|AAR87845.1| 5-enolpyruvylshikimate-3-phosphoshikimate 1-c ( 445) 219 58.4 5.1e-05  
gi|143294031|gb|EDE16453.1| hypothetical protein G ( 211) 215 57.3 5.2e-05  
gi|144092537|gb|EDI81200.1| hypothetical protein G ( 81) 210 55.9 5.3e-05  
gi|134077312|emb|CAK45651.1| unnamed protein product (1031) 223 59.5 5.3e-05  
gi|218722059|gb|EED21477.1| pentafunctional polypeptide (1577) 225 60.1 5.5e-05  
gi|135513466|gb|EBI13124.1| hypothetical protein G ( 152) 213 56.7 5.5e-05  
gi|134985130|gb|EBE73942.1| hypothetical protein G ( 152) 213 56.7 5.5e-05  
gi|143054617|gb|EDC48197.1| hypothetical protein G ( 335) 217 57.9 5.5e-05  
gi|256578504|gb|ACU89640.1| UDP-N-acetylglucosaminase ( 416) 218 58.1 5.6e-05  
gi|143881431|gb|EDH30581.1| hypothetical protein G ( 107) 211 56.2 5.7e-05  
gi|256559742|gb|ACU85589.1| UDP-N-acetylglucosaminase ( 517) 219 58.4 5.7e-05  
gi|141329022|gb|ECR07966.1| hypothetical protein G ( 161) 213 56.7 5.7e-05  
gi|143352907|gb|EDE50882.1| hypothetical protein G ( 429) 218 58.2 5.8e-05  
gi|24850309|gb|AAN63155.1| 5-enolpyruvylshikimate-3-phosphoshikimate 1-c ( 445) 218 58.2 5.9e-05  
gi|16415786|emb|CAD01096.1| 5-enolpyruvylshikimate-3-phosphoshikimate 1-c ( 445) 218 58.2 5.9e-05  
gi|136028772|gb|EBL43028.1| hypothetical protein G ( 143) 212 56.5 6.1e-05  
gi|141447793|gb|ECR89500.1| hypothetical protein G ( 147) 212 56.5 6.2e-05

gi|142887871|gb|EDB28923.1| hypothetical protein G ( 394) 217 57.9 6.3e-05  
gi|135290648|gb|EBG66366.1| hypothetical protein G ( 220) 214 57.1 6.3e-05  
gi|135802331|gb|EBU93691.1| hypothetical protein G ( 75) 208 55.4 6.8e-05  
gi|134840321|gb|EBD77600.1| hypothetical protein G ( 135) 211 56.2 6.8e-05  
gi|144098861|gb|EDI85676.1| hypothetical protein G ( 363) 216 57.7 6.8e-05  
gi|40557113|gb|AAR87844.1| 5-enolpyruvylshikimate-3-phosphoshikimate 1-c ( 445) 217 57.9 6.9e-05  
gi|139497159|gb|ECF47452.1| hypothetical protein G ( 94) 209 55.7 6.9e-05  
gi|141795168|gb|ECT33018.1| hypothetical protein G ( 139) 211 56.3 6.9e-05  
gi|140152196|gb|ECJ87114.1| hypothetical protein G ( 140) 211 56.3 7e-05  
gi|139335289|gb|ECE54589.1| hypothetical protein G ( 252) 214 57.1 7e-05  
gi|137648475|gb|EBV23980.1| hypothetical protein G ( 173) 212 56.5 7.1e-05  
gi|139755777|gb|ECH24015.1| hypothetical protein G ( 118) 210 56.0 7.1e-05  
gi|141081748|gb|ECP49664.1| hypothetical protein G ( 152) 211 56.3 7.4e-05  
gi|238939610|emb|CAR27785.1| ZYR00D06578p [Zygosaccharomyces] (1589) 223 59.7 7.5e-05  
gi|135766604|gb|EBU71201.1| hypothetical protein G ( 155) 211 56.3 7.5e-05  
gi|136447507|gb|EBO22455.1| hypothetical protein G ( 232) 213 56.9 7.6e-05  
gi|148323442|gb|EDK88692.1| 3-phosphoshikimate 1-c ( 420) 216 57.7 7.7e-05  
gi|27904180|gb|AAO27013.1| 3-phosphoshikimate 1-ca ( 427) 216 57.7 7.8e-05  
gi|142003656|gb|ECU75568.1| hypothetical protein G ( 238) 213 56.9 7.8e-05  
gi|136465730|gb|EBO34255.1| hypothetical protein G ( 136) 210 56.0 7.9e-05  
gi|138618826|gb|ECA84865.1| hypothetical protein G ( 76) 207 55.2 8e-05  
gi|141209606|gb|ECQ37721.1| hypothetical protein G ( 137) 210 56.0 8e-05  
gi|135261549|gb|EBG49250.1| hypothetical protein G ( 248) 213 56.9 8e-05  
gi|140096140|gb|ECJ50823.1| hypothetical protein G ( 115) 209 55.8 8.1e-05  
gi|134364163|gb|EBA91033.1| hypothetical protein G ( 102) 208 55.5 8.6e-05  
gi|137761629|gb|EBV85571.1| hypothetical protein G ( 278) 213 56.9 8.8e-05  
gi|138825030|gb|ECG30219.1| hypothetical protein G ( 105) 208 55.5 8.8e-05  
gi|261497413|gb|ACX83863.1| 3-phosphoshikimate 1-c ( 416) 215 57.5 8.8e-05  
gi|140486597|gb|ECL97657.1| hypothetical protein G ( 74) 206 55.0 9.1e-05  
gi|29605907|dbj|BAC69971.1| putative UDP-N-acetylglucosaminase ( 437) 215 57.5 9.2e-05  
gi|135319190|gb|EBG83853.1| hypothetical protein G ( 77) 206 55.0 9.3e-05  
gi|142582794|gb|ECZ11702.1| hypothetical protein G ( 114) 208 55.5 9.4e-05  
gi|213504184|emb|CAS92803.1| unnamed protein product ( 455) 215 57.5 9.5e-05  
gi|143365722|gb|EDE58318.1| hypothetical protein G ( 375) 214 57.2 9.5e-05  
gi|138629390|gb|ECA92270.1| hypothetical protein G ( 99) 207 55.3 9.8e-05  
gi|138258529|gb|EBV62640.1| hypothetical protein G ( 147) 209 55.8 9.8e-05  
gi|116792410|gb|ABK26355.1| unknown [Picea sitchensis] ( 148) 209 55.8 9.8e-05  
gi|114197527|gb|EAU39227.1| hypothetical protein A (1581) 221 59.2 0.0001  
gi|135733488|gb|EBJ50659.1| hypothetical protein G ( 160) 209 55.9 0.0001  
gi|143335933|gb|EDE40536.1| hypothetical protein G ( 430) 214 57.3 0.00011  
gi|141847271|gb|ECT66592.1| hypothetical protein G ( 112) 207 55.3 0.00011  
gi|10175334|dbj|BAB06432.1| 3-phosphoshikimate 1-c ( 447) 214 57.3 0.00011  
gi|167290545|gb|ABZ43409.1| Sequence 17347 from pABZ43409 ( 447) 214 57.3 0.00011  
gi|138609959|dbj|BAC78836.1| hypothetical protein G ( 115) 207 55.3 0.00011  
gi|141892373|gb|ECT97923.1| hypothetical protein G ( 308) 212 56.7 0.00011  
gi|136116509|gb|EBM02401.1| hypothetical protein G ( 118) 207 55.3 0.00011  
gi|137414679|gb|EBT96858.1| hypothetical protein G ( 154) 208 55.6 0.00012  
gi|135659439|gb|EBJ04885.1| hypothetical protein G ( 127) 207 55.4 0.00012  
gi|140810164|gb|ECN66826.1| hypothetical protein G ( 278) 211 56.5 0.00012  
gi|140771259|gb|ECN39459.1| hypothetical protein G ( 130) 207 55.4 0.00012  
gi|138611211|gb|ECA79728.1| hypothetical protein G ( 286) 211 56.5 0.00012  
gi|170174741|gb|ACB07794.1| 3-phosphoshikimate 1-c ( 423) 213 57.1 0.00012  
gi|137589221|gb|EBU91870.1| hypothetical protein G ( 134) 207 55.4 0.00012  
gi|138629391|gb|ECA92271.1| hypothetical protein G ( 91) 205 54.8 0.00012  
gi|137848316|gb|EBW34665.1| hypothetical protein G ( 135) 207 55.4 0.00012  
gi|140413732|gb|ECL59758.1| hypothetical protein G ( 137) 207 55.4 0.00013  
gi|140455536|gb|ECL87543.1| hypothetical protein G ( 247) 210 56.2 0.00013  
gi|143268167|gb|EDE02421.1| hypothetical protein G ( 77) 204 54.5 0.00013  
gi|140672153|gb|ECM71666.1| hypothetical protein G ( 141) 207 55.4 0.00013  
gi|141798088|gb|ECT34142.1| hypothetical protein G ( 144) 207 55.4 0.00013  
gi|141992988|gb|ECU68466.1| hypothetical protein G ( 149) 207 55.4 0.00013  
gi|141546046|gb|ECS27575.1| hypothetical protein G ( 124) 206 55.1 0.00014

gi|143270309|gb|EDE03964.1| hypothetical protein G ( 408) 212 56.8 0.00014  
gi|140236302|gb|ECK45224.1| hypothetical protein G ( 154) 207 55.4 0.00014  
gi|140413370|gb|ECL59501.1| hypothetical protein G ( 231) 209 56.0 0.00014  
gi|142893931|gb|EDB33295.1| hypothetical protein G ( 427) 212 56.8 0.00014  
gi|140654303|gb|ECM59089.1| hypothetical protein G ( 112) 205 54.9 0.00015  
gi|142142776|gb|ECV90405.1| hypothetical protein G ( 138) 206 55.2 0.00015  
gi|139498756|gb|ECF48543.1| hypothetical protein G ( 64) 202 54.0 0.00015  
gi|143205281|gb|EDD57686.1| hypothetical protein G ( 140) 206 55.2 0.00015  
gi|119948400|gb|ABM07311.1| 3-Phosphoshikimate-1-c ( 456) 212 56.9 0.00015  
gi|136292074|gb|EBN17837.1| hypothetical protein G ( 259) 209 56.0 0.00015  
gi|593744|gb|AAA55351.1| Sequence 3 from Patent EP ( 469) 212 56.9 0.00015  
gi|138878118|gb|ECC26538.1| hypothetical protein G ( 120) 205 54.9 0.00015  
gi|2484167|gb|AAB72303.1| I49196 Sequence 49 from p ( 480) 212 56.9 0.00016  
gi|144974746|gb|ABP12457.1| Sequence 49 from paten ( 480) 212 56.9 0.00016  
gi|5957567|gb|AAE08241.1| Sequence 49 from patent ( 480) 212 56.9 0.00016  
gi|2485244|gb|AAB73380.1| I44469 Sequence 49 from p ( 480) 212 56.9 0.00016  
gi|136531748|gb|EBO76792.1| hypothetical protein G ( 185) 207 55.5 0.00016  
gi|197053769|gb|ACH25467.1| Sequence 19 from paten ( 506) 212 56.9 0.00016  
gi|229422320|gb|EE037367.1| 3-phosphoshikimate 1-c ( 418) 211 56.6 0.00016  
gi|142608246|gb|ECZ29452.1| hypothetical protein G ( 419) 211 56.6 0.00016  
gi|136314634|gb|EBN33187.1| hypothetical protein G ( 514) 212 56.9 0.00016  
gi|143970801|gb|EDH94315.1| hypothetical protein G ( 110) 204 54.7 0.00017  
gi|143855993|gb|EDH11923.1| hypothetical protein G ( 648) 213 57.2 0.00017  
gi|137781507|gb|EBV96275.1| hypothetical protein G ( 206) 207 55.5 0.00017  
gi|119674666|gb|ABL88922.1| 3-phosphoshikimate 1-c ( 399) 210 56.4 0.00018  
gi|137013318|gb|EBR71823.1| hypothetical protein G ( 126) 204 54.7 0.00018  
gi|142582261|gb|EC211328.1| hypothetical protein G ( 104) 203 54.4 0.00019  
gi|144194317|gb|EDJ55692.1| hypothetical protein G ( 232) 207 55.6 0.00019  
gi|116062764|dbj|BAA80401.2| 3-phosphoshikimate 1- ( 419) 210 56.4 0.00019  
gi|160707495|gb|EAT91093.2| hypothetical protein S (1661) 217 58.4 0.00019  
gi|167296094|gb|ABZ48958.1| Sequence 22896 from pa ( 427) 210 56.4 0.00019  
gi|139191265|gb|ECE03543.1| hypothetical protein G ( 166) 205 55.0 0.0002  
gi|142478397|gb|ECY37505.1| hypothetical protein G ( 63) 200 53.6 0.0002  
gi|56638357|gb|AAW10629.1| Sequence 7 from patent ( 445) 210 56.4 0.0002  
gi|16415784|emb|CAD01095.1| 5-enolpyruvylshikimate ( 445) 210 56.4 0.0002  
gi|24850311|gb|AAN63156.1| 5-enolpyruvylshikimate- ( 445) 210 56.4 0.0002  
gi|135662774|gb|EBJ06959.1| hypothetical protein G ( 303) 208 55.9 0.0002  
gi|135651046|gb|EBI99695.1| hypothetical protein G ( 117) 203 54.5 0.0002  
gi|144974747|gb|ABP12458.1| Sequence 50 from paten ( 460) 210 56.4 0.0002  
gi|2484168|gb|AAB72304.1| I49197 Sequence 50 from p ( 460) 210 56.4 0.0002  
gi|2485245|gb|AAB73381.1| I44470 Sequence 50 from p ( 460) 210 56.4 0.0002  
gi|5957568|gb|AAE08242.1| Sequence 50 from patent ( 460) 210 56.4 0.0002  
gi|136950192|gb|EBR36051.1| hypothetical protein G ( 143) 204 54.7 0.0002  
gi|142857706|gb|EDB07392.1| hypothetical protein G ( 486) 210 56.5 0.00021  
gi|7019762|emb|CAB75770.1| pentafunctional aromati (1573) 216 58.1 0.00021  
gi|137646588|gb|EBV23065.1| hypothetical protein G ( 224) 206 55.3 0.00021  
gi|141954719|gb|ECU41802.1| hypothetical protein G ( 349) 208 55.9 0.00022  
gi|139177681|gb|ECD93905.1| hypothetical protein G ( 288) 207 55.6 0.00022  
gi|143866695|gb|EDH19810.1| hypothetical protein G ( 352) 208 55.9 0.00022  
gi|219932929|emb|CAR70886.1| putative 3-phosphoshi ( 430) 209 56.2 0.00022  
gi|13092893|emb|CAC30301.1| putative 3-phosphoshik ( 430) 209 56.2 0.00022  
gi|135856458|gb|EBK27748.1| hypothetical protein G ( 110) 202 54.2 0.00023  
gi|137313646|gb|EBT40020.1| hypothetical protein G ( 140) 203 54.5 0.00023  
gi|137327459|gb|EBT47689.1| hypothetical protein G ( 143) 203 54.5 0.00024  
gi|138437383|gb|EBZ67000.1| hypothetical protein G ( 221) 205 55.1 0.00025  
gi|238851221|gb|EEQ40685.1| hypothetical protein C (1568) 215 57.9 0.00025  
gi|3834343|emb|CAA28836.1| arom polypeptide [Emeri (1586) 215 57.9 0.00025  
gi|119397025|gb|EAW07456.1| pentafunctional polype (1587) 215 57.9 0.00025  
gi|29339497|gb|AAO77293.1| 3-phosphoshikimate 1-ca ( 410) 208 56.0 0.00025  
gi|135343554|gb|EBH00234.1| hypothetical protein G ( 155) 203 54.6 0.00025  
gi|76876364|emb|CAI87586.1| UDP-N-acetylglucosamin ( 419) 208 56.0 0.00026  
gi|136570079|gb|EBP01271.1| hypothetical protein G ( 195) 204 54.8 0.00026

gi|257472973|gb|ACV51092.1| 3-phosphoshikimate 1-c ( 426) 208 56.0 0.00026  
gi|135693637|gb|EBU25980.1| hypothetical protein G ( 200) 204 54.9 0.00026  
gi|135913705|gb|EBK66907.1| hypothetical protein G ( 136) 202 54.3 0.00027  
gi|143659900|gb|EDG12514.1| hypothetical protein G ( 64) 198 53.2 0.00027  
gi|139968256|gb|ECI70629.1| hypothetical protein G ( 140) 202 54.3 0.00027  
gi|141580813|gb|ECS45999.1| hypothetical protein G ( 78) 199 53.5 0.00027  
gi|139125658|gb|ECD58797.1| hypothetical protein G ( 181) 203 54.6 0.00028  
gi|143231890|gb|EDD76949.1| hypothetical protein G ( 398) 207 55.7 0.00029  
gi|144034478|gb|EDI39203.1| hypothetical protein G ( 69) 198 53.2 0.00029  
gi|139101794|gb|ECD42558.1| hypothetical protein G ( 132) 201 54.1 0.0003  
gi|140473336|gb|ECL94304.1| hypothetical protein G ( 75) 198 53.2 0.00031  
gi|256686438|gb|ACV09331.1| UDP-N-acetylglucosamin ( 439) 207 55.8 0.00031  
gi|141179292|gb|ECQ17891.1| hypothetical protein G ( 112) 200 53.8 0.00031  
gi|135589340|gb|EBI61491.1| hypothetical protein G ( 68) 197 53.0 0.00033  
gi|240135769|gb|EER35322.1| pentafunctional polype (1551) 213 57.5 0.00033  
gi|142456675|gb|ECY21312.1| hypothetical protein G ( 405) 206 55.5 0.00034  
gi|141670245|gb|ECS72593.1| hypothetical protein G ( 275) 204 55.0 0.00034  
gi|143081548|gb|EDC67931.1| hypothetical protein G ( 411) 206 55.5 0.00034  
gi|15113315|gb|AAE68899.1| Sequence 6 from patent ( 423) 206 55.5 0.00035  
gi|136481378|gb|EBO44369.1| hypothetical protein G ( 108) 199 53.6 0.00035  
gi|258554844|gb|ACV77786.1| 3-phosphoshikimate 1-c ( 428) 206 55.5 0.00035  
gi|134719113|gb|EBD01143.1| hypothetical protein G ( 136) 200 53.9 0.00036  
gi|143094736|gb|EDC77565.1| hypothetical protein G ( 306) 204 55.0 0.00037  
gi|194037323|gb|EDR33952.2| hypothetical protein P (1210) 211 57.0 0.00037  
gi|139902023|gb|ECT24529.1| hypothetical protein G ( 211) 202 54.4 0.00037  
gi|137252226|gb|EBT05562.1| hypothetical protein G ( 123) 199 53.6 0.00039  
gi|238842364|gb|EDC32026.1| pentafunctional AROM p (1571) 212 57.3 0.00039  
gi|665659|emb|CAA88208.1| Aro1p [Saccharomyces cer (1588) 212 57.3 0.00039  
gi|167274833|gb|ABZ27697.1| Sequence 1635 from pat (1588) 212 57.3 0.00039  
gi|13381|emb|CAA29458.1| unnamed protein product [S (1588) 212 57.3 0.00039  
gi|141537327|gb|ECS21320.1| hypothetical protein G ( 70) 196 52.8 0.00039  
gi|134759239|gb|EBD23750.1| hypothetical protein G ( 416) 205 55.3 0.0004  
gi|144203151|gb|EDU361880.1| hypothetical protein G ( 108) 198 53.3 0.00041  
gi|140981503|gb|EC082557.1| hypothetical protein G ( 237) 202 54.5 0.00041  
gi|78220930|gb|ABB40279.1| 3-phosphoshikimate 1-ca ( 442) 205 55.3 0.00042  
gi|136132486|gb|EBM13199.1| hypothetical protein G ( 256) 202 54.5 0.00043  
gi|142257629|gb|ECW77085.1| hypothetical protein G ( 410) 204 55.1 0.00046  
gi|157921240|gb|ABW02667.1| 3-phosphoshikimate 1-c ( 414) 204 55.1 0.00046  
gi|171991199|gb|ACB62121.1| UDP-N-acetylglucosamin ( 434) 204 55.1 0.00048  
gi|136333268|gb|EBN45633.1| hypothetical protein G ( 136) 198 53.4 0.00048  
gi|140268788|gb|ECK67789.1| hypothetical protein G ( 137) 198 53.4 0.00049  
gi|136136795|gb|EBM15292.1| hypothetical protein G ( 255) 201 54.3 0.0005  
gi|136253249|gb|EBM91262.1| hypothetical protein G ( 143) 198 53.4 0.0005  
gi|138953443|gb|ECC56702.1| hypothetical protein G ( 120) 197 53.2 0.00051  
gi|136574586|gb|EBP04142.1| hypothetical protein G ( 67) 194 52.3 0.00051  
gi|142876278|gb|EDB20907.1| hypothetical protein G ( 151) 198 53.5 0.00053  
gi|220698176|gb|EED54516.1| pentafunctional polype (1578) 210 56.8 0.00053  
gi|211586138|emb|CAP93886.1| Pcl6g12160 [Penicilli (1586) 210 56.8 0.00053  
gi|83770536|dbj|BAE60669.1| unnamed protein produc (1595) 210 56.8 0.00053  
gi|239593894|gb|EEQ76475.1| 3-dehydroquinone synth (1597) 210 56.8 0.00053  
gi|239607236|gb|EEQ84223.1| 3-dehydroquinone synth (1597) 210 56.8 0.00053  
gi|229452487|gb|EEO58278.1| 3-phosphoshikimate 1-c ( 409) 203 54.9 0.00053  
gi|143444815|gb|EDN99481.1| hypothetical protein G ( 419) 203 54.9 0.00054  
gi|134562746|gb|EBC08575.1| hypothetical protein G ( 164) 198 53.5 0.00056  
gi|135669319|gb|EBU11015.1| hypothetical protein G ( 136) 197 53.2 0.00056  
gi|139197802|gb|ECN08219.1| hypothetical protein G ( 80) 194 52.4 0.00059  
gi|137954036|gb|EBW94853.1| hypothetical protein G ( 178) 198 53.5 0.0006  
gi|226289074|gb|EEH44586.1| pentafunctional AROM p (1538) 209 56.6 0.0006  
gi|150414167|gb|EDN09532.1| 3-dehydroquinone synth (1538) 209 56.6 0.0006  
gi|143141346|gb|EDD11677.1| hypothetical protein G ( 100) 195 52.7 0.0006  
gi|140376969|gb|ECL35521.1| hypothetical protein G ( 149) 197 53.2 0.0006  
gi|136760213|gb|EBQ22186.1| hypothetical protein G ( 150) 197 53.2 0.00061

gi 140235542 gb ECK44673.1	hypothetical protein G ( 273)	200	54.1	0.00061
gi 138404715 gb EBZ43876.1	hypothetical protein G ( 229)	199	53.8	0.00062
gi 142704948 gb ECZ98083.1	hypothetical protein G ( 107)	195	52.7	0.00063
gi 141970069 gb ECU52289.1	hypothetical protein G ( 286)	200	54.1	0.00064
gi 137468274 gb EBU26710.1	hypothetical protein G ( 131)	196	53.0	0.00064
gi 116096414 gb ABJ61565.1	UDP-N-acetylglucosamin ( 428)	202	54.7	0.00064
gi 138719578 gb ECB53769.1	hypothetical protein G ( 137)	196	53.0	0.00066
gi 135010860 gb EBE91396.1	hypothetical protein G ( 93)	194	52.4	0.00066
gi 136789549 gb EBQ41934.1	hypothetical protein G ( 63)	192	51.9	0.00066
gi 139957028 gb ECI62785.1	hypothetical protein G ( 95)	194	52.4	0.00067
gi 135347204 gb EBH02680.1	hypothetical protein G ( 95)	194	52.4	0.00067
gi 139593930 gb ECG14149.1	hypothetical protein G ( 141)	196	53.0	0.00067
gi 226281118 gb EEH36684.1	pentafunctional AROM p (1523)	208	56.4	0.00069
gi 138838259 gb ECC09709.1	hypothetical protein G ( 99)	194	52.4	0.00069
gi 135099842 gb EBF49079.1	hypothetical protein G ( 322)	200	54.1	0.0007
gi 181604701 gb AAL63821.1	3-phosphoshikimate 1-ca ( 400)	201	54.4	0.00071
gi 137153595 gb EBS50236.1	hypothetical protein G ( 103)	194	52.5	0.00072
gi 225681904 gb EEH20184.1	pentafunctional AROM p (1603)	208	56.4	0.00072
gi 145304842 gb ABP55424.1	3-phosphoshikimate 1-c ( 414)	201	54.4	0.00073
gi 138452474 gb EBZ77315.1	hypothetical protein G ( 157)	196	53.0	0.00073
gi 140722456 gb ECN06632.1	hypothetical protein G ( 157)	196	53.0	0.00073
gi 141051574 gb ECP29929.1	hypothetical protein G ( 158)	196	53.0	0.00074
gi 138254743 gb EBY59922.1	hypothetical protein G ( 90)	193	52.2	0.00075
gi 257801515 gb EEV30445.1	UDP-N-acetylglucosamin ( 430)	201	54.5	0.00075
gi 257807952 gb EEV36774.1	UDP-N-acetylglucosamin ( 430)	201	54.5	0.00075
gi 169242594 emb CAM63622.1	3-phosphoshikimate 1- ( 431)	201	54.5	0.00075
gi 137958190 gb EBW97170.1	hypothetical protein G ( 111)	194	52.5	0.00076
gi 135163783 gb EBF90152.1	hypothetical protein G ( 165)	196	53.1	0.00076
gi 142110315 gb ECV66589.1	hypothetical protein G ( 366)	200	54.2	0.00077
gi 135657735 gb EBJ03832.1	hypothetical protein G ( 94)	193	52.2	0.00077
gi 142201089 gb ECW34849.1	hypothetical protein G ( 453)	201	54.5	0.00078
gi 136947037 gb EBR34275.1	hypothetical protein G ( 143)	195	52.8	0.00079
gi 141916054 gb ECU14493.1	hypothetical protein G ( 82)	192	52.0	0.00081
gi 263254749 gb EEZ26183.1	3-phosphoshikimate 1-c ( 410)	200	54.2	0.00084
gi 142559146 gb ECY95199.1	hypothetical protein G ( 146)	194	52.6	0.00094
gi 139664636 gb ECG62011.1	hypothetical protein G ( 263)	197	53.4	0.00094
gi 149386230 gb ABN65979.2	predicted protein [Pic (1571)	206	56.0	0.00096
gi 259145368 emb CAY78632.1	Arolp [Saccharomyces (1588)	206	56.0	0.00097
gi 190404911 gb EDV08178.1	3-dehydroquinate dehyd (1588)	206	56.0	0.00097
gi 142072287 gb ECV36024.1	hypothetical protein G ( 419)	199	54.0	0.00099
gi 136640296 gb EBP45062.1	hypothetical protein G ( 132)	193	52.3	0.001
gi 143868004 gb EDH20779.1	hypothetical protein G ( 167)	194	52.6	0.001
gi 242123418 gb ACS81114.1	3-phosphoshikimate 1-c ( 444)	199	54.0	0.001
gi 134553253 gb EBC02793.1	hypothetical protein G ( 114)	192	52.1	0.001
gi 140706516 gb ECM95928.1	hypothetical protein G ( 65)	189	51.2	0.0011
gi 138687742 gb ECB31567.1	hypothetical protein G ( 215)	195	52.9	0.0011
gi 139379779 gb ECE71675.1	hypothetical protein G ( 120)	192	52.1	0.0011
gi 141113012 gb ECP71634.1	hypothetical protein G ( 83)	190	51.5	0.0011
gi 137964772 gb EBX00908.1	hypothetical protein G ( 151)	193	52.4	0.0011
gi 138751797 gb ECB75063.1	hypothetical protein G ( 273)	196	53.2	0.0011
gi 240278751 gb EER42257.1	pentafunctional AROM p (1595)	205	55.8	0.0011
gi 225560322 gb EEH08604.1	pentafunctional AROM p (1595)	205	55.8	0.0011
gi 229445205 gb EBC050996.1	3-phosphoshikimate 1-c ( 409)	198	53.8	0.0011
gi 262355555 gb EE204646.1	3-phosphoshikimate 1-c ( 410)	198	53.8	0.0011
gi 251947714 gb EES87996.1	3-phosphoshikimate 1-c ( 412)	198	53.8	0.0011
gi 143307135 gb EDE23454.1	hypothetical protein G ( 419)	198	53.8	0.0012
gi 135739268 gb EBJ54246.1	hypothetical protein G ( 131)	192	52.1	0.0012
gi 140653063 gb ECM58227.1	hypothetical protein G ( 132)	192	52.1	0.0012
gi 148571397 gb ABQ93456.1	UDP-N-acetylglucosamin ( 432)	198	53.8	0.0012
gi 139727823 gb ECH06066.1	hypothetical protein G ( 163)	193	52.4	0.0012
gi 143690268 gb EDG29963.1	hypothetical protein G ( 256)	195	53.0	0.0012
gi 138114114 gb EBX82134.1	hypothetical protein G ( 144)	192	52.1	0.0013

gi 140033355 gb ECJ13975.1	hypothetical protein G ( 98)	190	51.6	0.0013
gi 141588775 gb ECS48489.1	hypothetical protein G ( 262)	195	53.0	0.0013
gi 143462007 gb EDF10728.1	hypothetical protein G ( 178)	193	52.4	0.0013
gi 141303147 gb ECQ94769.1	hypothetical protein G ( 264)	195	53.0	0.0013
gi 139363992 gb ECE63192.1	hypothetical protein G ( 82)	189	51.3	0.0013
gi 142562795 gb ECY97747.1	hypothetical protein G ( 151)	192	52.2	0.0013
gi 126249231 gb ABO08322.1	3-phosphoshikimate 1-c ( 403)	197	53.6	0.0013
gi 140222713 gb ECK35846.1	hypothetical protein G ( 160)	192	52.2	0.0014
gi 134469062 gb EBB52900.1	hypothetical protein G ( 74)	188	51.0	0.0014
gi 134454271 gb EBB44255.1	hypothetical protein G ( 133)	191	51.9	0.0014
gi 134945150 gb EBE47064.1	hypothetical protein G ( 133)	191	51.9	0.0014
gi 135692503 gb EBJ25287.1	hypothetical protein G ( 133)	191	51.9	0.0014
gi 139589325 gb ECG10870.1	hypothetical protein G ( 113)	190	51.6	0.0014
gi 138198542 gb EBY35312.1	hypothetical protein G ( 63)	187	50.8	0.0014
gi 240249395 gb ACS46335.1	3-phosphoshikimate 1-c ( 448)	197	53.6	0.0014
gi 240250963 gb ACS47902.1	3-phosphoshikimate 1-c ( 448)	197	53.6	0.0014
gi 136051312 gb EBL58189.1	hypothetical protein G ( 370)	196	53.3	0.0014
gi 143709062 gb EDG38906.1	hypothetical protein G ( 376)	196	53.3	0.0014
gi 219621611 gb ACL29768.1	3-phosphoshikimate 1-c ( 462)	197	53.6	0.0015
gi 139199645 gb ECE09358.1	hypothetical protein G ( 322)	195	53.1	0.0015
gi 141118077 gb ECF75109.1	hypothetical protein G ( 69)	187	50.8	0.0015
gi 134859754 gb EBB90063.1	hypothetical protein G ( 62)	186	50.6	0.0016
gi 140905310 gb ECO29373.1	hypothetical protein G ( 298)	194	52.8	0.0016
gi 138684318 gb ECB29207.1	hypothetical protein G ( 137)	190	51.7	0.0016
gi 139957153 gb ECI62875.1	hypothetical protein G ( 203)	192	52.2	0.0016
gi 135540756 gb ECI30520.1	hypothetical protein G ( 371)	195	53.1	0.0017
gi 141117651 gb EBJ74806.1	hypothetical protein G ( 267)	193	52.6	0.0017
gi 138463778 gb EBZ85248.1	hypothetical protein G ( 327)	194	52.8	0.0017
gi 141272693 gb ECQ81987.1	hypothetical protein G ( 152)	190	51.7	0.0018
gi 241995660 gb ACS75029.1	AroA [Methylophilus me ( 126)	189	51.4	0.0018
gi 60491659 emb CAH06411.1	putative 3-phosphoshik ( 410)	195	53.1	0.0018
gi 116612514 gb ABK05238.1	UDP-N-acetylglucosamin ( 507)	196	53.4	0.0018
gi 195933610 gb ACG58310.1	UDP-N-acetylglucosamin ( 421)	195	53.1	0.0018
gi 53956243 gb AAV06240.1	Sequence 18197 from pat ( 421)	195	53.1	0.0018
gi 196170661 gb ACG71634.1	UDP-N-acetylglucosamin ( 422)	195	53.1	0.0018
gi 117648206 gb ABK52308.1	3-phosphoshikimate 1-c ( 423)	195	53.1	0.0018
gi 193089610 gb ACF14885.1	UDP-N-acetylglucosamin ( 429)	195	53.1	0.0019
gi 115789082 gb ABJ22152.1	Sequence 7120 from pat ( 430)	195	53.1	0.0019
gi 135364141 gb EBH14043.1	hypothetical protein G ( 91)	187	50.9	0.0019
gi 139013401 gb ECC80641.1	hypothetical protein G ( 248)	192	52.3	0.0019
gi 143911586 gb EDH51981.1	hypothetical protein G ( 261)	192	52.3	0.002
gi 142827402 gb EDA88063.1	hypothetical protein G ( 124)	188	51.2	0.002
gi 143610523 gb EDF85316.1	hypothetical protein G ( 419)	194	52.9	0.0021
gi 212561985 gb ACJ30543.1	UDP-N-acetylglucosamin ( 439)	194	52.9	0.0022
gi 239838602 gb ACS30399.1	UDP-N-acetylglucosamin ( 442)	194	52.9	0.0022
gi 138307153 gb EBY90043.1	hypothetical protein G ( 96)	186	50.7	0.0023
gi 135681712 gb EBJ18625.1	hypothetical protein G ( 124)	187	51.0	0.0024
gi 143224629 gb EDD71741.1	hypothetical protein G ( 413)	193	52.7	0.0024
gi 219953367 gb ACL63751.1	UDP-N-acetylglucosamin ( 422)	193	52.7	0.0025
gi 139689781 gb ECG79631.1	hypothetical protein G ( 94)	185	50.5	0.0026
gi 137532653 gb EBU60183.1	hypothetical protein G ( 140)	187	51.0	0.0026
gi 143644117 gb EDG03371.1	hypothetical protein G ( 96)	185	50.5	0.0026
gi 137415456 gb EBT97288.1	hypothetical protein G ( 255)	190	51.9	0.0026
gi 142882449 gb EDB25460.1	hypothetical protein G ( 80)	184	50.2	0.0027
gi 134382056 gb EBB03106.1	hypothetical protein G ( 67)	183	49.9	0.0027
gi 137268607 gb EBT14833.1	hypothetical protein G ( 269)	190	51.9	0.0027
gi 149390731 gb ABR25383.1	5-enolpyruvylshikimate ( 273)	190	51.9	0.0028
gi 142495350 gb ECY49520.1	hypothetical protein G ( 276)	190	51.9	0.0028
gi 146152689 gb ABQ03543.1	3-phosphoshikimate 1-c ( 409)	192	52.5	0.0028
gi 254835623 gb EET15932.1	3-phosphoshikimate 1-c ( 409)	192	52.5	0.0028
gi 167291322 gb ABZ44186.1	Sequence 18124 from pa ( 410)	192	52.5	0.0028
gi 10639425 emb CAC11427.1	3-phosphoshikimate 1-c ( 410)	192	52.5	0.0028

gi|223512251|gb|EEF23969.1| 3-phosphoshikimate 1-c ( 193) 188 51.4 0.0029  
gi|162953230|gb|ABY22745.1| UDP-N-acetylglucosamin ( 515) 193 52.8 0.0029  
gi|136087910|gb|EBL82913.1| hypothetical protein G ( 134) 186 50.8 0.0029  
gi|140607099|gb|ECM36325.1| hypothetical protein G ( 136) 186 50.8 0.003  
gi|134455209|gb|EBB44828.1| hypothetical protein G ( 137) 186 50.8 0.003  
gi|136139207|gb|EBM16318.1| hypothetical protein G ( 169) 187 51.1 0.003  
gi|162953604|gb|ABY23119.1| 3-phosphoshikimate 1-c ( 461) 192 52.5 0.0031  
gi|137462153|gb|EBU23533.1| hypothetical protein G ( 120) 185 50.6 0.0031  
gi|139030135|gb|ECC92270.1| hypothetical protein G ( 100) 184 50.3 0.0032  
gi|15623398|dbj|BAB67386.1| 408aa long hypothetica ( 408) 191 52.3 0.0033  
gi|155262582|gb|ABT18186.1| Sequence 105656 from p ( 155) 186 50.9 0.0033  
gi|141745506|gb|ECT07595.1| hypothetical protein G ( 128) 185 50.6 0.0033  
gi|157917592|gb|ABV99019.1| 3-phosphoshikimate 1-c ( 414) 191 52.3 0.0033  
gi|143916751|gb|EDH55551.1| hypothetical protein G ( 418) 191 52.3 0.0033  
gi|114338910|gb|ABH56758.1| UDP-N-acetylglucosamin ( 419) 191 52.3 0.0033  
gi|135390495|gb|EBH31755.1| hypothetical protein G ( 133) 185 50.6 0.0034  
gi|134677879|gb|EBC77000.1| hypothetical protein G ( 243) 188 51.4 0.0034  
gi|143133776|gb|ECC06103.1| hypothetical protein G ( 296) 189 51.7 0.0034  
gi|134764306|gb|EBD27211.1| hypothetical protein G ( 246) 188 51.4 0.0035  
gi|136357609|gb|EBN62230.1| hypothetical protein G ( 303) 189 51.7 0.0035  
gi|135692660|gb|EBJ25383.1| hypothetical protein G ( 140) 185 50.6 0.0035  
gi|145284076|gb|ABP51658.1| 3-phosphoshikimate 1-c ( 398) 190 52.0 0.0037  
gi|139513765|gb|ECF58913.1| hypothetical protein G ( 124) 184 50.3 0.0037  
gi|137393308|gb|EBT84756.1| hypothetical protein G ( 335) 189 51.8 0.0038  
gi|142452999|gb|ECY18640.1| hypothetical protein G ( 417) 190 52.0 0.0039  
gi|89335099|dbj|BAE84694.1| hypothetical protein [ ( 419) 190 52.0 0.0039  
gi|142253022|gb|ECW73688.1| hypothetical protein G ( 284) 188 51.5 0.0039  
gi|218086749|emb|CAT03393.1| unnamed protein produ ( 431) 190 52.1 0.004  
gi|51587625|emb|CAH19218.1| 3-phosphoshikimate 1-c ( 431) 190 52.1 0.004  
gi|197053752|gb|ACH25450.1| Sequence 2 from patent ( 431) 190 52.1 0.004  
gi|218301240|emb|CAU98560.1| unnamed protein produ ( 431) 190 52.1 0.004  
gi|207102478|emb|CAR82017.1| unnamed protein produ ( 431) 190 52.1 0.004  
gi|218301238|emb|CAU98558.1| unnamed protein produ ( 431) 190 52.1 0.004  
gi|213501389|emb|CAS92773.1| unnamed protein produ ( 431) 190 52.1 0.004  
gi|213501391|emb|CAS92774.1| unnamed protein produ ( 431) 190 52.1 0.004  
gi|161726847|emb|CAP47295.1| unnamed protein produ ( 431) 190 52.1 0.004  
gi|197030981|gb|ACH14982.1| Sequence 2 from patent ( 431) 190 52.1 0.004  
gi|213501387|emb|CAS92772.1| unnamed protein produ ( 431) 190 52.1 0.004  
gi|140904032|gb|ECO28502.1| hypothetical protein G ( 135) 184 50.4 0.004  
gi|270507581|gb|AC285859.1| UDP-N-acetylglucosamin ( 438) 190 52.1 0.004  
gi|135009318|gb|EBE90352.1| hypothetical protein G ( 362) 189 51.8 0.004  
gi|136629195|gb|EBP38477.1| hypothetical protein G ( 76) 181 49.5 0.004  
gi|138371238|gb|EBZ21040.1| hypothetical protein G ( 76) 181 49.5 0.004  
gi|138176813|gb|EBY21560.1| hypothetical protein G ( 79) 181 49.5 0.0041  
gi|142900876|gb|EDB38366.1| hypothetical protein G ( 317) 188 51.5 0.0042  
gi|142331408|gb|ECX31573.1| hypothetical protein G ( 324) 188 51.5 0.0043  
gi|149932870|gb|ABR39568.1| 3-phosphoshikimate 1-c ( 409) 189 51.8 0.0044  
gi|138544955|gb|ECA33333.1| hypothetical protein G ( 86) 181 49.6 0.0044  
gi|138700454|gb|ECB40624.1| hypothetical protein G ( 105) 182 49.9 0.0044  
gi|46914771|emb|CAG21548.1| putative UDP-N-acetylgl ( 420) 189 51.8 0.0045  
gi|108461643|gb|ABF86828.1| putative 3-phosphoshik ( 430) 189 51.8 0.0046  
gi|135215041|gb|EBG21926.1| hypothetical protein G ( 76) 180 49.3 0.0047  
gi|119947437|gb|ABM06348.1| UDP-N-acetylglucosamin ( 441) 189 51.8 0.0047  
gi|138138563|gb|EBX95779.1| hypothetical protein G ( 140) 183 50.2 0.0048  
gi|141808592|gb|ECT39389.1| hypothetical protein G ( 117) 182 49.9 0.0048  
gi|141433721|gb|ECR79660.1| hypothetical protein G ( 80) 180 49.3 0.0049  
gi|256689097|gb|ACV06899.1| 3-phosphoshikimate 1-c ( 476) 189 51.9 0.005  
gi|20239787|gb|AAE91809.1| Sequence 2 from patent (1551) 195 53.6 0.005  
gi|15103454|gb|AAE65493.1| Sequence 2 from patent (1551) 195 53.6 0.005  
gi|142476987|gb|ECY36457.1| hypothetical protein G ( 149) 183 50.2 0.005  
gi|212009244|gb|ACJ16626.1| 5-enolpyruvylshikimate ( 402) 188 51.6 0.0051  
gi|143620960|gb|EDF90404.1| hypothetical protein G ( 331) 187 51.3 0.0051

gi|134468282|gb|EBB52463.1| hypothetical protein G ( 70) 179 49.1 0.0051  
gi|166857315|gb|ABY95723.1| UDP-N-acetylglucosamin ( 416) 188 51.6 0.0052  
gi|166853016|gb|ABY91425.1| UDP-N-acetylglucosamin ( 416) 188 51.6 0.0052  
gi|269096834|gb|ACZ21270.1| UDP-N-acetylglucosamin ( 509) 189 51.9 0.0052  
gi|219540331|gb|ACL22070.1| UDP-N-acetylglucosamin ( 419) 188 51.6 0.0052  
gi|167283017|gb|ABZ35881.1| Sequence 9819 from pat ( 420) 188 51.6 0.0052  
gi|133915492|emb|CAM05605.1| 3-phosphoshikimate 1- ( 422) 188 51.6 0.0053  
gi|134598959|gb|EBC30184.1| hypothetical protein G ( 160) 183 50.2 0.0053  
gi|13740525|gb|EBV74756.1| hypothetical protein G ( 133) 182 49.9 0.0053  
gi|144091022|gb|EDI80155.1| hypothetical protein G ( 365) 187 51.4 0.0055  
gi|29834829|gb|AAP05464.1| 3-phosphoshikimate 1-ca ( 446) 188 51.6 0.0055  
gi|141983563|gb|ECU61776.1| hypothetical protein G ( 64) 178 48.8 0.0055  
gi|140208878|gb|ECK26833.1| hypothetical protein G ( 79) 179 49.1 0.0056  
gi|143317823|gb|EDE29836.1| hypothetical protein G ( 176) 183 50.2 0.0057  
gi|139513224|gb|ECF58553.1| hypothetical protein G ( 69) 178 48.8 0.0059  
gi|141885008|gb|ECT92840.1| hypothetical protein G ( 192) 183 50.3 0.0061  
gi|92394696|gb|ABE75971.1| UDP-N-acetylglucosamine ( 422) 187 51.4 0.0061  
gi|140284195|gb|EBK74219.1| hypothetical protein G ( 110) 180 49.4 0.0062  
gi|135121308|gb|EBF62811.1| hypothetical protein G ( 296) 185 50.8 0.0063  
gi|140624510|gb|ECM41482.1| hypothetical protein G ( 76) 178 48.9 0.0063  
gi|142953368|gb|EDB77607.1| hypothetical protein G ( 95) 179 49.2 0.0065  
gi|138764060|gb|ECK78675.1| hypothetical protein G ( 117) 180 49.5 0.0065  
gi|141648244|gb|ECS65129.1| hypothetical protein G ( 97) 179 49.2 0.0066  
gi|143042478|gb|EDC39310.1| hypothetical protein G ( 212) 183 50.3 0.0066  
gi|139487309|gb|ECF41573.1| hypothetical protein G ( 98) 179 49.2 0.0066  
gi|139739266|gb|ECH13970.1| hypothetical protein G ( 317) 185 50.9 0.0066  
gi|141432679|gb|ECW78915.1| hypothetical protein G ( 67) 177 48.6 0.0067  
gi|143745983|gb|EDG57607.1| hypothetical protein G ( 150) 181 49.8 0.0068  
gi|142166301|gb|ECW08274.1| hypothetical protein G ( 400) 186 51.2 0.0068  
gi|256689084|gb|EDB76886.1| 5-enolpyruvylshikimate ( 348) 185 50.9 0.0071  
gi|137218459|gb|EBS86662.1| hypothetical protein G ( 109) 179 49.2 0.0072  
gi|116742768|emb|CAK21892.1| murA-1 [Listeria wels ( 430) 186 51.2 0.0072  
gi|164152006|emb|CAC97896.1| UDP-N-acetylglucosamin ( 430) 186 51.2 0.0072  
gi|217332641|gb|ACK38435.1| UDP-N-acetylglucosamin ( 430) 186 51.2 0.0072  
gi|16412014|emb|CAD00604.1| UDP-N-acetylglucosamin ( 430) 186 51.2 0.0072  
gi|21702575|gb|AAM75972.1|AF481102\_8 5-enolpyruvyl ( 432) 186 51.2 0.0072  
gi|254046068|gb|ACT62861.1| UDP-N-acetylglucosamin ( 438) 186 51.2 0.0073  
gi|28271773|emb|CAD64679.1| UDP-N-acetylglucosamin ( 438) 186 51.2 0.0073  
gi|138140870|gb|EBX97063.1| hypothetical protein G ( 136) 180 49.5 0.0073  
gi|139188211|gb|ECE01341.1| hypothetical protein G ( 63) 176 48.4 0.0074  
gi|62148370|emb|CAH64137.1| 3-phosphoshikimate 1-c ( 445) 186 51.2 0.0074  
gi|260649506|emb|CBG72621.1| UDP-N-acetylglucosami ( 448) 186 51.2 0.0074  
gi|142637013|gb|ECZ49717.1| hypothetical protein G ( 369) 185 50.9 0.0075  
gi|140383974|gb|ECZ140532.1| hypothetical protein G ( 251) 183 50.4 0.0075  
gi|140045594|gb|ECJ21510.1| hypothetical protein G ( 78) 177 48.7 0.0075  
gi|139770468|gb|ECH34129.1| hypothetical protein G ( 121) 179 49.2 0.0078  
gi|142928836|gb|EDB58683.1| hypothetical protein G ( 322) 184 50.7 0.0078  
gi|136211954|gb|EBM63337.1| hypothetical protein G ( 122) 179 49.3 0.0078  
gi|52214895|dbj|BAD47488.1| 3-phosphoshikimate 1-c ( 410) 185 51.0 0.0081  
gi|142018338|gb|ECU88169.1| hypothetical protein G ( 158) 180 49.6 0.0082  
gi|190688437|gb|ACE86115.1| UDP-N-acetylglucosamin ( 420) 185 51.0 0.0082  
gi|135320669|gb|EBG84854.1| hypothetical protein G ( 132) 179 49.3 0.0083  
gi|139391913|gb|ECR79640.1| hypothetical protein G ( 133) 179 49.3 0.0084  
gi|142351900|gb|ECX46155.1| hypothetical protein G ( 91) 177 48.7 0.0084  
gi|137419753|gb|EBT99762.1| hypothetical protein G ( 111) 178 49.0 0.0085  
gi|140808226|gb|ECN65512.1| hypothetical protein G ( 136) 179 49.3 0.0085  
gi|140185524|gb|ECK10449.1| hypothetical protein G ( 137) 179 49.3 0.0086  
gi|162954008|gb|ABY23523.1| UDP-N-acetylglucosamin ( 444) 185 51.0 0.0086  
gi|178466894|dbj|BAG21414.1| putative UDP-N-acetyl ( 446) 185 51.0 0.0086  
gi|135556520|gb|EBI40613.1| hypothetical protein G ( 94) 177 48.7 0.0087  
gi|138212658|gb|EBY44975.1| hypothetical protein G ( 169) 180 49.6 0.0087  
gi|136285546|gb|EBN13410.1| hypothetical protein G ( 153) 179 49.3 0.0093

gi 139501414 gb ECF50392.1	hypothetical protein G ( 285)	182	50.2	0.0096
gi 136171835 gb EBM36678.1	hypothetical protein G ( 422)	184	50.7	0.0096
gi 169803576 gb ACA82194.1	UDP-N-acetylglucosamin ( 428)	184	50.8	0.0097
gi 139023559 gb ECC87682.1	hypothetical protein G ( 133)	178	49.1	0.0097
gi 140352867 gb ECL18954.1	hypothetical protein G ( 133)	178	49.1	0.0097
gi 139883257 gb ECT11584.1	hypothetical protein G ( 137)	178	49.1	0.01
gi 142291045 gb ECX01581.1	hypothetical protein G ( 146)	178	49.1	0.01
gi 144100806 gb EDI87031.1	hypothetical protein G ( 180)	179	49.4	0.011
gi 134374036 gb EBA97727.1	hypothetical protein G ( 122)	177	48.8	0.011
gi 167831743 dbj BAG08659.1	3-phosphoshikimate 1- ( 401)	183	50.5	0.011
gi 45775290 gb AAS77246.1	putative 3-phosphoshiki ( 272)	181	49.9	0.011
gi 136287502 gb EBN14730.1	hypothetical protein G ( 411)	183	50.5	0.011
gi 219861261 gb ACL41603.1	transcriptional regula ( 507)	184	50.8	0.011
gi 212555054 gb ACJ27508.1	UDP-N-acetylglucosamin ( 419)	183	50.5	0.011
gi 135323535 gb EBG86796.1	hypothetical protein G ( 357)	182	50.3	0.011
gi 116099232 gb ABJ64381.1	UDP-N-acetylglucosamin ( 435)	183	50.5	0.011
gi 140243474 gb ECK50349.1	hypothetical protein G ( 166)	178	49.1	0.012
gi 134400389 gb EBB13277.1	hypothetical protein G ( 77)	174	48.0	0.012
gi 141252115 gb ECQ67624.1	hypothetical protein G ( 117)	176	48.6	0.012
gi 144219177 gb EDJ73294.1	hypothetical protein G ( 182)	178	49.2	0.012
gi 263233991 gb EEZ19592.1	3-phosphoshikimate 1-c ( 409)	182	50.3	0.013
gi 229437449 gb EEE047526.1	3-phosphoshikimate 1-c ( 409)	182	50.3	0.013
gi 229457254 gb EEE062975.1	3-phosphoshikimate 1-c ( 409)	182	50.3	0.013
gi 136904146 gb EBR14194.1	hypothetical protein G ( 86)	174	48.0	0.013
gi 143907330 gb EDH49031.1	hypothetical protein G ( 338)	181	50.0	0.013
gi 13813452 gb AAK04646.1	3-phosphoshikimate 1-ca ( 414)	182	50.3	0.013
gi 261601917 gb ACX91520.1	3-phosphoshikimate 1-c ( 414)	182	50.3	0.013
gi 119768707 gb ABM01278.1	UDP-N-acetylglucosamin ( 418)	182	50.3	0.013
gi 34397864 gb AAQ66925.1	3-phosphoshikimate 1-ca ( 419)	182	50.3	0.013
gi 134720968 gb EBD02202.1	hypothetical protein G ( 192)	178	49.2	0.013
gi 226098022 dbj BAH46464.1	UDP-N-acetylglucosami ( 427)	182	50.3	0.013
gi 139978480 gb ECI77873.1	hypothetical protein G ( 90)	174	48.1	0.013
gi 140931170 gb ECO47450.1	hypothetical protein G ( 134)	176	48.6	0.013
gi 135441282 gb EBH65877.1	hypothetical protein G ( 136)	176	48.6	0.013
gi 28410485 emb CAD66871.1	3-phosphoshikimate 1-c ( 443)	182	50.3	0.014
gi 89331472 dbj BAE81065.1	3-phosphoshikimate 1-c ( 445)	182	50.3	0.014
gi 140343210 gb ECL12004.1	hypothetical protein G ( 141)	176	48.6	0.014
gi 137616702 gb EBV07380.1	hypothetical protein G ( 147)	176	48.7	0.014
gi 141831273 gb ECT55098.1	hypothetical protein G ( 149)	176	48.7	0.014
gi 28476576 gb AAO44664.1	3-phosphoshikimate 1-ca ( 486)	182	50.4	0.015
gi 138558547 gb ECA42856.1	hypothetical protein G ( 124)	175	48.4	0.015
gi 136808146 gb EBQ54274.1	hypothetical protein G ( 406)	181	50.1	0.015
gi 142230728 gb ECW57166.1	hypothetical protein G ( 409)	181	50.1	0.015
gi 142857632 gb EDB07334.1	hypothetical protein G ( 128)	175	48.4	0.015
gi 136031855 gb EBL45103.1	hypothetical protein G ( 414)	181	50.1	0.015
gi 134051246 gb ABO49217.1	UDP-N-acetylglucosamin ( 417)	181	50.1	0.015
gi 247541801 gb ACS98819.1	UDP-N-acetylglucosamin ( 417)	181	50.1	0.015
gi 160859372 gb ABX47906.1	UDP-N-acetylglucosamin ( 419)	181	50.1	0.015
gi 151363855 gb ABS06855.1	UDP-N-acetylglucosamin ( 419)	181	50.1	0.015
gi 217497045 gb ACK45238.1	UDP-N-acetylglucosamin ( 419)	181	50.1	0.015
gi 125999056 gb ABN63131.1	UDP-N-acetylglucosamin ( 419)	181	50.1	0.015
gi 71039444 gb AAZ19752.1	UDP-N-acetylglucosamine ( 422)	181	50.1	0.015
gi 139190332 gb ECE02882.1	hypothetical protein G ( 108)	174	48.1	0.015
gi 141795829 gb ECT33174.1	hypothetical protein G ( 288)	179	49.5	0.015
gi 156224224 gb EDO45052.1	predicted protein [Nem ( 428)	181	50.1	0.015
gi 22294501 dbj BAC08331.1	UDP-N-acetylglucosamin ( 439)	181	50.1	0.016
gi 143566017 gb EDF66785.1	hypothetical protein G ( 76)	172	47.6	0.016
gi 136707484 gb EBP87611.1	hypothetical protein G ( 204)	177	49.0	0.016
gi 141223561 gb ECQ47426.1	hypothetical protein G ( 95)	173	47.9	0.016
gi 143314137 gb EDE27504.1	hypothetical protein G ( 210)	177	49.0	0.016
gi 136096902 gb EBL89017.1	hypothetical protein G ( 102)	173	47.9	0.017
gi 141686809 gb ECS81156.1	hypothetical protein G ( 124)	174	48.2	0.017

gi 139395280 gb ECE82035.1	hypothetical protein G ( 271)	178	49.3	0.017
gi 211999671 gb EEB05331.1	pentafunctional AROM p (1584)	187	51.8	0.017
gi 137037500 gb EBB85370.1	hypothetical protein G ( 126)	174	48.2	0.017
gi 136179199 gb EBM41522.1	hypothetical protein G ( 408)	180	49.9	0.017
gi 2274557241 gb ACP35928.1	3-phosphoshikimate 1-c ( 414)	180	49.9	0.017
gi 228020375 gb ACP55782.1	3-phosphoshikimate 1-c ( 414)	180	49.9	0.017
gi 228012361 gb ACP48122.1	3-phosphoshikimate 1-c ( 414)	180	49.9	0.017
gi 228010404 gb ACP46166.1	3-phosphoshikimate 1-c ( 414)	180	49.9	0.017
gi 143724011 gb EDG45785.1	hypothetical protein G ( 414)	180	49.9	0.017
gi 139360258 gb ECE62143.1	hypothetical protein G ( 235)	177	49.0	0.018
gi 134478916 gb EBB58719.1	hypothetical protein G ( 90)	172	47.6	0.018
gi 46881970 gb AAT05264.1	UDP-N-acetylglucosamine ( 430)	180	49.9	0.018
gi 225877536 emb CAS06250.1	Putative UDP-N-acetyl ( 430)	180	49.9	0.018
gi 177841144 gb ACB75396.1	3-phosphoshikimate 1-c ( 431)	180	49.9	0.018
gi 4377372 gb AAD19176.1	Phosphoshikimate Vinyltr ( 445)	180	49.9	0.018
gi 8979412 dbj BAA99246.1	phosphoshikimate vinyltr ( 445)	180	49.9	0.018
gi 33236921 gb AAP99008.1	3-phosphoshikimate 1-ca ( 445)	180	49.9	0.018
gi 8163504 gb AAF73706.1	3-phosphoshikimate 1-car ( 445)	180	49.9	0.018
gi 269302841 gb ACE32941.1	3-phosphoshikimate 1-c ( 445)	180	49.9	0.018
gi 143904828 gb EDH47285.1	hypothetical protein G ( 302)	178	49.3	0.018
gi 56660946 gb AAW16305.1	Sequence 1040 from pate ( 449)	180	49.9	0.018
gi 137543749 gb EBU66153.1	hypothetical protein G ( 116)	173	47.9	0.019
gi 139601152 gb ECG19101.1	hypothetical protein G ( 308)	178	49.3	0.019
gi 136799044 gb EBQ48260.1	hypothetical protein G ( 173)	175	48.5	0.019
gi 142380263 gb ECX65159.1	hypothetical protein G ( 395)	179	49.6	0.019
gi 140119616 gb ECJ67272.1	hypothetical protein G ( 222)	176	48.8	0.02
gi 144212288 gb EDJ68348.1	hypothetical protein G ( 331)	178	49.4	0.02
gi 135420409 gb EBH51870.1	hypothetical protein G ( 127)	173	48.0	0.02
gi 213504192 emb CAS92807.1	unnamed protein produ ( 416)	179	49.7	0.02
gi 259420524 emb CBF57576.1	unnamed protein produ ( 416)	179	49.7	0.02
gi 160429271 gb ABX42834.1	UDP-N-acetylglucosamin ( 418)	179	49.7	0.02
gi 120560331 gb ABM26258.1	UDP-N-acetylglucosamin ( 419)	179	49.7	0.02
gi 145563522 gb ABP74457.1	UDP-N-acetylglucosamin ( 419)	179	49.7	0.02
gi 139143879 gb ECD70262.1	hypothetical protein G ( 192)	175	48.5	0.02
gi 138081133 gb EBX63740.1	hypothetical protein G ( 130)	173	48.0	0.02
gi 145690079 gb ABP90585.1	UDP-N-acetylglucosamin ( 423)	179	49.7	0.02
gi 145692283 gb ABP92788.1	UDP-N-acetylglucosamin ( 423)	179	49.7	0.02
gi 55739074 gb AAV62715.1	UDP-N-acetylglucosamine ( 428)	179	49.7	0.021
gi 281012168 gb ADA07818.1	Sequence 1114 from pat ( 449)	179	49.7	0.021
gi 115825398 gb ABU37297.1	Sequence 1114 from pat ( 449)	179	49.7	0.021
gi 31710845 gb AAP67810.1	Sequence 1114 from pate ( 449)	179	49.7	0.021
gi 167292185 gb ABZ45049.1	Sequence 18987 from pa ( 373)	178	49.4	0.022
gi 139819050 gb ECH68273.1	hypothetical protein G ( 96)	171	47.4	0.022
gi 140328259 gb ECL01562.1	hypothetical protein G ( 98)	171	47.4	0.022
gi 140340029 gb ECL09712.1	hypothetical protein G ( 120)	172	47.7	0.022
gi 138198137 gb EBY35030.1	hypothetical protein G ( 180)	174	48.3	0.023
gi 137616572 gb EBV07310.1	hypothetical protein G ( 122)	172	47.7	0.023
gi 135507526 gb EBI09289.1	hypothetical protein G ( 272)	176	48.9	0.023
gi 152206293 gb ABS30603.1	3-phosphoshikimate 1-c ( 413)	178	49.4	0.023
gi 138829636 gb ECC05571.1	hypothetical protein G ( 128)	172	47.7	0.023
gi 137192869 gb EBS72151.1	hypothetical protein G ( 128)	172	47.7	0.023
gi 227459852 gb ACP38538.1	3-phosphoshikimate 1-c ( 414)	178	49.4	0.023
gi 238381356 gb ACR42444.1	3-phosphoshikimate 1-c ( 414)	178	49.4	0.023
gi 136384109 gb EBN80400.1	hypothetical protein G ( 415)	178	49.4	0.023
gi 225702160 emb CAW99858.1	UDP-N-acetylglucosamin ( 419)	178	49.4	0.024
gi 557371651 gb AAV60803.1	UDP-N-acetylglucosamine ( 428)	178	49.4	0.024
gi 156567059 gb ABU82464.1	3-phosphoshikimate 1-c ( 443)	178	49.5	0.025
gi 94544194 gb ABF34242.1	UDP-N-acetylglucosamine ( 419)	177	49.2	0.027
gi 13622467 gb AAK34186.1	putative UDP-N-acetylgl ( 419)	177	49.2	0.027
gi 251820383 emb CAR46971.1	UDP-N-acetylglucosami ( 419)	177	49.2	0.027
gi 134271847 emb CAM30082.1	UDP-N-acetylglucosami ( 419)	177	49.2	0.027
gi 242391406 dbj BAH81865.1	UDP-N-acetylglucosami ( 419)	177	49.2	0.027

gi|94546174|gb|ABF36221.1| UDP-N-acetylglucosamine ( 419) 177 49.2 0.027  
gi|71853702|gb|AAZ51725.1| UDP-N-acetylglucosamine ( 419) 177 49.2 0.027  
gi|28810988|dbj|BAC63922.1| putative UDP-N-acetylgl ( 419) 177 49.2 0.027  
gi|94542289|gb|ABF32338.1| UDP-N-acetylglucosamine ( 419) 177 49.2 0.027  
gi|251816600|emb|CAZ52238.1| UDP-N-acetylglucosami ( 419) 177 49.2 0.027  
gi|251818535|emb|CAZ56366.1| UDP-N-acetylglucosami ( 419) 177 49.2 0.027  
gi|209540799|gb|AC161375.1| UDP-N-acetylglucosamin ( 419) 177 49.2 0.027  
gi|71802859|gb|AA72212.1| UDP-N-acetylglucosamine ( 419) 177 49.2 0.027  
gi|143565502|gb|EDF66515.1| hypothetical protein G ( 419) 177 49.2 0.027  
gi|138671508|gb|ECB20195.1| hypothetical protein G ( 130) 171 47.5 0.028  
gi|14517938|gb|AAK64441.1|AF377339\_2 UDP-GlcNAc 1- ( 420) 177 49.2 0.028  
gi|108465795|gb|ABF90980.1| UDP-N-acetylglucosamin ( 420) 177 49.2 0.028  
gi|116101170|gb|ABJ66316.1| UDP-N-acetylglucosamin ( 423) 177 49.2 0.028  
gi|142787608|gb|EDA58191.1| hypothetical protein G ( 355) 176 48.9 0.028  
gi|136476208|gb|EBO41023.1| hypothetical protein G ( 243) 174 48.4 0.028  
gi|139146809|gb|ECD72295.1| hypothetical protein G ( 76) 168 46.7 0.029  
gi|136121233|gb|EBM05633.1| hypothetical protein G ( 371) 176 49.0 0.029  
gi|138699474|gb|ECB39929.1| hypothetical protein G ( 98) 169 47.0 0.03  
gi|140180092|gb|ECK06544.1| hypothetical protein G ( 69) 167 46.5 0.031  
gi|136738289|gb|EBQ07820.1| hypothetical protein G ( 124) 170 47.3 0.031  
gi|136634440|gb|EBF41579.1| hypothetical protein G ( 86) 168 46.7 0.031  
gi|137127351|gb|EBJ35626.1| hypothetical protein G ( 106) 169 47.0 0.032  
gi|166856063|gb|ABY94471.1| UDP-N-acetylglucosamin ( 417) 176 49.0 0.032  
gi|19748524|gb|AAL97966.1| putative UDP-N-acetylgl ( 419) 176 49.0 0.032  
gi|22533863|gb|AAM99730.1|AE014230\_10 UDP-N-acetyl ( 419) 176 49.0 0.032  
gi|195974521|gb|ACG62047.1| UDP-N-acetylglucosamin ( 419) 176 49.0 0.032  
gi|50903501|gb|AAT87216.1| UDP-N-acetylglucosamine ( 419) 176 49.0 0.032  
gi|76562914|gb|ABA45498.1| UDP-N-acetylglucosamine ( 419) 176 49.0 0.032  
gi|158141998|gb|ABW20310.1| UDP-N-acetylglucosamin ( 420) 176 49.0 0.032  
gi|229430270|gb|EEO40482.1| UDP-N-acetylglucosamin ( 423) 176 49.0 0.032  
gi|136363919|gb|EBN66534.1| hypothetical protein G ( 73) 167 46.5 0.032  
gi|140723067|gb|ECN07050.1| hypothetical protein G ( 89) 168 46.8 0.032  
gi|136683126|gb|EBP71689.1| hypothetical protein G ( 351) 175 48.7 0.032  
gi|257804381|gb|EEV33203.1| UDP-N-acetylglucosamin ( 429) 176 49.0 0.033  
gi|2484155|gb|AAB72291.1|I49184 Sequence 18 from p ( 28) 162 45.1 0.033  
gi|2485232|gb|AAB73368.1|I44457 Sequence 18 from p ( 28) 162 45.1 0.033  
gi|5957555|gb|AAE08229.1| Sequence 18 from patent ( 28) 162 45.1 0.033  
gi|144974734|gb|ABP12445.1| Sequence 18 from paten ( 28) 162 45.1 0.033  
gi|140152195|gb|ECJ87113.1| hypothetical protein G ( 134) 170 47.3 0.033  
gi|143151259|gb|EDD18893.1| hypothetical protein G ( 92) 168 46.8 0.033  
gi|140874660|gb|ECO09587.1| hypothetical protein G ( 115) 169 47.1 0.034  
gi|136049666|gb|EBL57068.1| hypothetical protein G ( 314) 174 48.5 0.035  
gi|270230871|emb|CBI20198.1| unnamed protein produ ( 475) 176 49.0 0.035  
gi|140724657|gb|ECN08179.1| hypothetical protein G ( 102) 168 46.8 0.036  
gi|136459335|gb|EBO30111.1| hypothetical protein G ( 228) 172 47.9 0.037  
gi|137081208|gb|EBS10031.1| hypothetical protein G ( 71) 166 46.2 0.037  
gi|142027336|gb|ECU96678.1| hypothetical protein G ( 87) 167 46.5 0.037  
gi|144079806|gb|EDI72256.1| hypothetical protein G ( 87) 167 46.5 0.037  
gi|143126581|gb|EDD00824.1| hypothetical protein G ( 191) 171 47.7 0.037  
gi|282702301|emb|CAD63131.1| UDP-N-acetylglucosamin ( 426) 175 48.8 0.038  
gi|254044508|gb|ACT61301.1| UDP-N-acetylglucosamin ( 426) 175 48.8 0.038  
gi|143182820|gb|EDD41234.1| hypothetical protein G ( 133) 169 47.1 0.038  
gi|257164739|gb|EEU94699.1| UDP-N-acetylglucosamin ( 430) 175 48.8 0.038  
gi|255969024|gb|EET99646.1| UDP-N-acetylglucosamin ( 430) 175 48.8 0.038  
gi|256952732|gb|EEU69364.1| UDP-N-acetylglucosamin ( 430) 175 48.8 0.038  
gi|256598280|gb|EEU17456.1| UDP-N-acetylglucosamin ( 430) 175 48.8 0.038  
gi|256994553|gb|EEU81855.1| UDP-N-acetylglucosamin ( 430) 175 48.8 0.038  
gi|256990899|gb|EEU78201.1| UDP-N-acetylglucosamin ( 430) 175 48.8 0.038  
gi|256998392|gb|EEU84912.1| UDP-N-acetylglucosamin ( 430) 175 48.8 0.038  
gi|255964127|gb|EET96603.1| UDP-N-acetylglucosamin ( 430) 175 48.8 0.038  
gi|256993601|gb|EEU80903.1| UDP-N-acetylglucosamin ( 430) 175 48.8 0.038  
gi|29343206|gb|AAO80969.1| UDP-N-acetylglucosamine ( 430) 175 48.8 0.038

gi|257157615|gb|EEU87575.1| UDP-N-acetylglucosamin ( 430) 175 48.8 0.038  
gi|256683428|gb|EEU23123.1| UDP-N-acetylglucosamin ( 430) 175 48.8 0.038  
gi|257160824|gb|EEU90784.1| UDP-N-acetylglucosamin ( 430) 175 48.8 0.038  
gi|256955926|gb|EEU72558.1| UDP-N-acetylglucosamin ( 430) 175 48.8 0.038  
gi|256986010|gb|EEU73312.1| UDP-N-acetylglucosamin ( 430) 175 48.8 0.038  
gi|256949162|gb|EEU65794.1| UDP-N-acetylglucosamin ( 430) 175 48.8 0.038  
gi|137151382|gb|EBS49008.1| hypothetical protein G ( 164) 170 47.4 0.038  
gi|138582552|gb|ECA59501.1| hypothetical protein G ( 248) 172 48.0 0.039  
gi|197710518|gb|EDY54552.1| UDP-N-acetylglucosamin ( 448) 175 48.8 0.039  
gi|143906361|gb|EDH48335.1| hypothetical protein G ( 117) 168 46.8 0.04  
gi|141310784|gb|ECQ97443.1| hypothetical protein G ( 66) 165 46.0 0.04  
gi|196191655|gb|EDX86619.1| UDP-N-acetylglucosamin ( 477) 175 48.8 0.041  
gi|138345862|gb|EBZ10558.1| hypothetical protein G ( 83) 166 46.3 0.041  
gi|142838038|gb|ECN85962.1| hypothetical protein G ( 101) 167 46.6 0.041  
gi|140265664|gb|ECW83000.1| hypothetical protein G ( 189) 170 47.4 0.043  
gi|143127849|gb|EDD01758.1| hypothetical protein G ( 282) 172 48.0 0.043  
gi|188595418|dbj|BAG34393.1| putative 3-phosphoshi ( 419) 174 48.6 0.043  
gi|23095295|emb|CAD46505.1| Unknown [Streptococcus ( 419) 174 48.6 0.043  
gi|140891397|gb|ECO21011.1| hypothetical protein G ( 289) 172 48.0 0.044  
gi|6911979|emb|CAB72195.1| UDP-N-acetylglucosamine ( 448) 174 48.6 0.046  
gi|138633487|gb|ECX94615.1| hypothetical protein G ( 119) 167 46.6 0.047  
gi|136511372|gb|EBO63787.1| hypothetical protein G ( 214) 170 47.5 0.047  
gi|142134144|gb|ECV84003.1| hypothetical protein G ( 215) 170 47.5 0.047  
gi|59746623|gb|AAW97135.1| Sequence 14698 from pat ( 263) 171 47.8 0.047  
gi|143298282|gb|EDB18600.1| hypothetical protein G ( 268) 171 47.8 0.048  
gi|141085359|gb|ECP52162.1| hypothetical protein G ( 69) 164 45.8 0.049  
gi|139799833|gb|ECH54842.1| hypothetical protein G ( 84) 165 46.1 0.049  
gi|136971600|gb|EBR48170.1| hypothetical protein G ( 128) 167 46.7 0.05  
gi|94548110|gb|ABF38156.1| UDP-N-acetylglucosamine ( 419) 173 48.3 0.05  
gi|140437930|gb|ECL75035.1| hypothetical protein G ( 194) 169 47.2 0.051  
gi|138970500|gb|ECC64000.1| hypothetical protein G ( 240) 170 47.5 0.051  
gi|143636058|gb|EDF98692.1| hypothetical protein G ( 75) 164 45.8 0.052  
gi|143542777|gb|EDF54584.1| hypothetical protein G ( 441) 173 48.4 0.052  
gi|141883049|gb|ECT91469.1| hypothetical protein G ( 77) 164 45.8 0.053  
gi|144009555|gb|EDT21805.1| hypothetical protein G ( 316) 171 47.8 0.055  
gi|137829606|gb|EBW23918.1| hypothetical protein G ( 99) 165 46.1 0.055  
gi|140438989|gb|ECL75771.1| hypothetical protein G ( 121) 166 46.4 0.055  
gi|135591094|gb|EBT62571.1| hypothetical protein G ( 69) 163 45.6 0.056  
gi|55420321|gb|AAV52046.1| AroA [Haemophilus paras ( 329) 171 47.8 0.056  
gi|137693394|gb|EBV49130.1| hypothetical protein G ( 125) 166 46.4 0.057  
gi|135188539|gb|EBG06001.1| hypothetical protein G ( 342) 171 47.8 0.058  
gi|139370234|gb|ECE65074.1| hypothetical protein G ( 106) 165 46.2 0.058  
gi|142718885|gb|EDA08059.1| hypothetical protein G ( 417) 172 48.1 0.058  
gi|138487694|gb|ECC30777.1| hypothetical protein G ( 130) 166 46.4 0.059  
gi|135669959|gb|EBJ11406.1| hypothetical protein G ( 159) 167 46.7 0.059  
gi|138926726|gb|ECC45302.1| hypothetical protein G ( 133) 166 46.4 0.06  
gi|256714009|gb|EEU28997.1| UDP-N-acetylglucosamin ( 432) 172 48.1 0.06  
gi|256613445|gb|EEU18648.1| UDP-N-acetylglucosamin ( 432) 172 48.1 0.06  
gi|135734076|gb|EBU51018.1| hypothetical protein G ( 134) 166 46.5 0.06  
gi|135059368|gb|EBF23231.1| hypothetical protein G ( 364) 171 47.9 0.061  
gi|139375434|gb|ECE68685.1| hypothetical protein G ( 67) 162 45.4 0.064  
gi|139132736|gb|ECB62598.1| hypothetical protein G ( 82) 163 45.6 0.064  
gi|134338116|gb|EBA73904.1| hypothetical protein G ( 413) 171 47.9 0.067  
gi|135327958|gb|EBG89770.1| hypothetical protein G ( 418) 171 47.9 0.068  
gi|262261230|gb|EEV79929.1| UDP-N-acetylglucosamin ( 420) 171 47.9 0.068  
gi|116094484|gb|ABJ59636.1| UDP-N-acetylglucosamin ( 421) 171 47.9 0.068

gi 135254630 gb EBG45208.1	hypothetical protein G ( 422)	171	47.9	0.068
gi 134737951 gb EBD11912.1	hypothetical protein G ( 235)	168	47.1	0.068
gi 257473601 gb ACV51720.1	UDP-N-acetylglucosamin ( 427)	171	47.9	0.069
gi 135167020 gb EBF92253.1	hypothetical protein G ( 352)	170	47.6	0.069
gi 142831308 gb EDA91032.1	hypothetical protein G ( 291)	169	47.4	0.069
gi 134600428 gb EBC31021.1	hypothetical protein G ( 242)	168	47.1	0.07
gi 137755492 gb EBV82397.1	hypothetical protein G ( 298)	169	47.4	0.071
gi 142168248 gb ECW09775.1	hypothetical protein G ( 449)	171	47.9	0.072
gi 135096128 gb EBF46697.1	hypothetical protein G ( 454)	171	47.9	0.072
gi 135113525 gb EBF57833.1	hypothetical protein G ( 454)	171	47.9	0.072
gi 135111290 gb EBF56415.1	hypothetical protein G ( 454)	171	47.9	0.072
gi 139434090 gb ECF08739.1	hypothetical protein G ( 96)	163	45.7	0.073
gi 138120911 gb EBX85976.1	hypothetical protein G ( 260)	168	47.1	0.074
gi 137174025 gb EBS61743.1	hypothetical protein G ( 103)	163	45.7	0.077
gi 141041491 gb ECP22985.1	hypothetical protein G ( 104)	163	45.7	0.077
gi 166854871 gb ABY93280.1	UDP-N-acetylglucosamin ( 417)	170	47.7	0.079
gi 142898535 gb EDB36622.1	hypothetical protein G ( 287)	168	47.1	0.08
gi 212549346 dbj BAG84014.1	3-phosphoshikimate 1- ( 426)	170	47.7	0.08
gi 256711277 gb EEU26315.1	UDP-N-acetylglucosamin ( 430)	170	47.7	0.081
gi 142861353 gb EDB10051.1	hypothetical protein G ( 432)	170	47.7	0.081
gi 29608787 dbj BAC72840.1	putative UDP-N-acetylgl ( 448)	170	47.7	0.083
gi 137104776 gb EBS22894.1	hypothetical protein G ( 250)	167	46.9	0.083
gi 142158366 gb ECW02190.1	hypothetical protein G ( 454)	170	47.7	0.084
gi 135046225 gb EBF14853.1	hypothetical protein G ( 257)	167	46.9	0.085
gi 135597220 gb EBT66356.1	hypothetical protein G ( 119)	163	45.8	0.086
gi 137253093 gb EBT06079.1	hypothetical protein G ( 119)	163	45.8	0.086
gi 136573786 gb EBP03634.1	hypothetical protein G ( 222)	166	46.6	0.089
gi 219047701 gb AAM79640.1	putative UDP-N-acetylgl ( 404)	169	47.5	0.089
gi 140673198 gb ECM72414.1	hypothetical protein G ( 70)	160	44.9	0.09
gi 55420325 gb AAV52048.1	AroA [Actinobacillus ur ( 338)	168	47.2	0.091
gi 228228929 gb ACP86594.1	Sequence 8450 from pat ( 417)	169	47.5	0.092
gi 268621293 gb EEZ53693.1	UDP-N-acetylglucosamin ( 417)	169	47.5	0.092
gi 268588887 gb EEZ53563.1	UDP-N-acetylglucosamin ( 417)	169	47.5	0.092
gi 59719126 gb AAW90531.1	putative UDP-N-acetylgl ( 417)	169	47.5	0.092
gi 193935115 gb ACF30939.1	UDP-N-acetylglucosamin ( 417)	169	47.5	0.092
gi 226512474 gb EEH61819.1	UDP-N-acetylglucosamin ( 417)	169	47.5	0.092
gi 268551781 gb EEZ46800.1	UDP-N-acetylglucosamin ( 417)	169	47.5	0.092
gi 268584379 gb EEZ49055.1	UDP-N-acetylglucosamin ( 417)	169	47.5	0.092
gi 268586619 gb EEZ51295.1	UDP-N-acetylglucosamin ( 417)	169	47.5	0.092
gi 268549193 gb EEZ44611.1	UDP-N-acetylglucosamin ( 417)	169	47.5	0.092
gi 268625593 gb EEZ57993.1	UDP-N-acetylglucosamin ( 417)	169	47.5	0.092
gi 157075899 gb ABV10582.1	UDP-N-acetylglucosamin ( 420)	169	47.5	0.092
gi 218519485 gb ACK80071.1	UDP-N-acetylglucosamin ( 423)	169	47.5	0.093
gi 198249249 gb ACH84842.1	UDP-N-acetylglucosamin ( 423)	169	47.5	0.093
gi 141827987 gb ECT52798.1	hypothetical protein G ( 135)	163	45.8	0.095
gi 139811437 gb ECH63105.1	hypothetical protein G ( 62)	159	44.7	0.095
gi 6063409 dbj BAA85335.1	UDP-N-acetylglucosamine ( 446)	169	47.5	0.096
gi 142973737 gb EDB90282.1	hypothetical protein G ( 96)	161	45.3	0.098
gi 143528632 gb EDF47496.1	hypothetical protein G ( 379)	168	47.2	0.099
gi 134464232 gb EBB50113.1	hypothetical protein G ( 118)	162	45.5	0.099
gi 137574225 gb EBU83505.1	hypothetical protein G ( 119)	162	45.5	0.1
gi 142725044 gb EDA12427.1	hypothetical protein G ( 83)	160	45.0	0.1
gi 139819051 gb ECH68274.1	hypothetical protein G ( 70)	159	44.7	0.1
gi 141468301 gb ECR99942.1	hypothetical protein G ( 275)	166	46.7	0.1
gi 139116258 gb ECD52455.1	hypothetical protein G ( 127)	162	45.6	0.11
gi 73912408 dbj BAE20403.1	5-enolpyruvylshikimate ( 231)	165	46.4	0.11
gi 140788486 gb ECN51392.1	hypothetical protein G ( 106)	161	45.3	0.11
gi 25166285 dbj BAC24475.1	murA [Wigglesworthia g ( 418)	168	47.3	0.11
gi 125498222 gb ABN44888.1	Conserved uncharacteri ( 419)	168	47.3	0.11
gi 16412040 emb CAD00630.1	murZ [Listeria monocyt ( 423)	168	47.3	0.11
gi 160347816 gb ABX26490.1	UDP-N-acetylglucosamin ( 431)	168	47.3	0.11
gi 141775381 gb ECT22687.1	hypothetical protein G ( 135)	162	45.6	0.11

gi 143630447 gb EDF95521.1	hypothetical protein G ( 135)	162	45.6	0.11
gi 197720926 gb EDY64834.1	UDP-N-acetylglucosamin ( 446)	168	47.3	0.11
gi 141040532 gb ECF22321.1	hypothetical protein G ( 96)	160	45.0	0.11
gi 140395645 gb ECL48452.1	hypothetical protein G ( 79)	159	44.8	0.11
gi 136612405 gb EBP28570.1	hypothetical protein G ( 174)	163	45.9	0.12
gi 171851474 emb CAQ04450.1	3-phosphoshikimate 1- ( 467)	168	47.3	0.12
gi 137848322 gb EBW34669.1	hypothetical protein G ( 83)	159	44.8	0.12
gi 143736357 gb EDG51926.1	hypothetical protein G ( 155)	162	45.6	0.12
gi 228227429 gb ACX85094.1	Sequence 5450 from pat ( 282)	165	46.5	0.12
gi 225699514 emb CAW93074.1	UDP-N-acetylglucosami ( 419)	167	47.0	0.12
gi 137764462 gb EBV87009.1	hypothetical protein G ( 107)	160	45.1	0.12
gi 134410709 gb EBB19029.1	hypothetical protein G ( 285)	165	46.5	0.12
gi 143091256 gb EDC75017.1	hypothetical protein G ( 73)	158	44.5	0.13
gi 135591110 gb EBI62581.1	hypothetical protein G ( 132)	161	45.4	0.13
gi 142900568 gb EDB38135.1	hypothetical protein G ( 352)	166	46.8	0.13
gi 136440761 gb EB018088.1	hypothetical protein G ( 111)	160	45.1	0.13
gi 137935986 gb EBW84686.1	hypothetical protein G ( 201)	163	45.9	0.13
gi 76167624 gb AAXP50632.1	3-phosphoshikimate 1-ca ( 440)	167	47.1	0.13
gi 231274025 emb CAX10818.1	3-phosphoshikimate 1- ( 440)	167	47.1	0.13
gi 231273013 emb CAX09925.1	3-phosphoshikimate 1- ( 440)	167	47.1	0.13
gi 137233043 gb EBS94913.1	hypothetical protein G ( 249)	164	46.2	0.13
gi 137259628 gb EBT09802.1	hypothetical protein G ( 261)	164	46.2	0.14
gi 137867122 gb EBW45478.1	hypothetical protein G ( 262)	164	46.2	0.14
gi 137233843 gb EDB35366.1	hypothetical protein G ( 102)	159	44.8	0.14
gi 137036035 gb EBR84572.1	hypothetical protein G ( 70)	157	44.3	0.14
gi 142888082 gb EDB29077.1	hypothetical protein G ( 341)	165	46.5	0.14
gi 20515136 gb AAM23459.1	UDP-N-acetylglucosamine ( 415)	166	46.8	0.14
gi 254671262 emb CBA08549.1	UDP-N-acetylglucosami ( 417)	166	46.8	0.14
gi 161594583 gb ABX72243.1	UDP-N-acetylglucosamin ( 417)	166	46.8	0.14
gi 139812606 gb EBH63945.1	hypothetical protein G ( 130)	160	45.1	0.14
gi 219679913 gb EED36262.1	UDP-N-acetylglucosamin ( 420)	166	46.8	0.14
gi 68446636 dbj BAE04220.1	UDP-N-acetylglucosamin ( 423)	166	46.8	0.15
gi 135629141 gb EBT86117.1	hypothetical protein G ( 89)	158	44.6	0.15
gi 141200123 gb ECQ31132.1	hypothetical protein G ( 237)	163	46.0	0.15
gi 137560711 gb EBU75799.1	hypothetical protein G ( 109)	159	44.9	0.15
gi 260572493 gb EEEX29055.1	UDP-N-acetylglucosamin ( 431)	166	46.8	0.15
gi 165930565 emb CAP04061.1	3-phosphoshikimate 1- ( 440)	166	46.8	0.15
gi 111957970 gb ABH73822.1	Sequence 59 from paten ( 440)	166	46.8	0.15
gi 165931440 emb CAP07015.1	3-phosphoshikimate 1- ( 440)	166	46.8	0.15
gi 259306889 gb ACW37700.1	Sequence 59 from paten ( 440)	166	46.8	0.15
gi 136430523 gb EB011467.1	hypothetical protein G ( 249)	163	46.0	0.15
gi 141756443 gb ECT12702.1	hypothetical protein G ( 208)	162	45.7	0.15
gi 136637968 gb EBP43697.1	hypothetical protein G ( 326)	164	46.3	0.16
gi 135135012 gb EBF71681.1	hypothetical protein G ( 402)	165	46.6	0.16
gi 143875110 gb EDH25993.1	hypothetical protein G ( 152)	160	45.2	0.16
gi 214034673 gb EEB75413.1	UDP-N-acetylglucosamin ( 415)	165	46.6	0.17
gi 151280568 gb ABR88978.1	UDP-N-acetylglucosamin ( 416)	165	46.6	0.17
gi 135561634 gb EBI43929.1	hypothetical protein G ( 418)	165	46.6	0.17
gi 142316922 gb ECK20878.1	hypothetical protein G ( 418)	165	46.6	0.17
gi 140339709 gb ECL09481.1	hypothetical protein G ( 88)	157	44.4	0.17
gi 167591717 gb ABZ83465.1	udp-n-acetylglucosamin ( 420)	165	46.6	0.17
gi 256583776 gb ACU94910.1	UDP-N-acetylglucosamin ( 425)	165	46.6	0.17
gi 140369628 gb ECL30304.1	hypothetical protein G ( 294)	163	46.1	0.17
gi 140421466 gb ECL64457.1	hypothetical protein G ( 111)	158	44.6	0.17
gi 137052829 gb EBR94060.1	hypothetical protein G ( 112)	158	44.7	0.17
gi 138650547 gb ECB05452.1	hypothetical protein G ( 168)	160	45.2	0.18
gi 144188789 gb EDJ51600.1	hypothetical protein G ( 94)	157	44.4	0.18
gi 136491502 gb EB050898.1	hypothetical protein G ( 169)	160	45.2	0.18
gi 135044669 gb EBF13863.1	hypothetical protein G ( 170)	160	45.2	0.18
gi 140152257 gb ECJ87157.1	hypothetical protein G ( 97)	157	44.4	0.18
gi 239911204 gb ACS34095.1	3-phosphoshikimate 1-c ( 398)	164	46.4	0.19
gi 142900460 gb EDB38056.1	hypothetical protein G ( 332)	163	46.1	0.19



gi 143591526 gb EDF75600.1	hypothetical protein G ( 226)	161	45.5	0.19
gi 134521141 gb EBB83682.1	hypothetical protein G ( 104)	157	44.4	0.19
gi 134893000 gb EBE12404.1	hypothetical protein G ( 86)	156	44.1	0.19
gi 229380775 gb EE030866.1	UDP-N-acetylglucosamin ( 417)	164	46.4	0.19
gi 138619891 gb ECA85573.1	hypothetical protein G ( 158)	159	45.0	0.2
gi 137172829 gb EBS61062.1	hypothetical protein G ( 89)	156	44.1	0.2
gi 3328790 gb AAC67962.1	Phosphoshikimate 1-carbo ( 440)	164	46.4	0.2
gi 137945478 gb EBW90050.1	hypothetical protein G ( 300)	162	45.8	0.2
gi 134891434 gb EBE11362.1	hypothetical protein G ( 138)	158	44.7	0.2
gi 33748212 gb AAQ45218.1	Sequence 5776 from pate ( 446)	164	46.4	0.21
gi 135513173 gb EBI12931.1	hypothetical protein G ( 94)	156	44.2	0.21
gi 136086971 gb EBL82277.1	hypothetical protein G ( 145)	158	44.7	0.21
gi 134767538 gb EBD29426.1	hypothetical protein G ( 120)	157	44.5	0.21
gi 137520362 gb EBU53503.1	hypothetical protein G ( 69)	154	43.6	0.22
gi 141041104 gb ECP22722.1	hypothetical protein G ( 102)	156	44.2	0.22
gi 143014918 gb EDC19274.1	hypothetical protein G ( 154)	158	44.8	0.22
gi 138345237 gb EBZ10125.1	hypothetical protein G ( 106)	156	44.2	0.23
gi 133740129 emb CAL61380.1	UDP-N-acetylglucosami ( 416)	163	46.2	0.23
gi 167283169 gb ABZ36033.1	Sequence 9971 from pat ( 417)	163	46.2	0.23
gi 117608277 gb ABK43732.1	UDP-N-acetylglucosamin ( 419)	163	46.2	0.23
gi 57638265 gb AAW55053.1	UDP-N-acetylglucosamine ( 419)	163	46.2	0.23
gi 262298484 gb EEY86397.1	UDP-N-acetylglucosamin ( 419)	163	46.2	0.23
gi 41582594 gb AAS08206.1	UDP-N-acetylglucosamine ( 421)	163	46.2	0.23
gi 116103448 gb ABJ68591.1	UDP-N-acetylglucosamin ( 423)	163	46.2	0.23
gi 136905237 gb EBR14604.1	hypothetical protein G ( 287)	161	45.6	0.23
gi 217110295 gb ACJ94175.1	Sequence 4217 from pat ( 431)	163	46.2	0.23
gi 259270370 gb ACW30183.1	Sequence 4217 from pat ( 431)	163	46.2	0.23
gi 144970267 gb ABP08557.1	Sequence 3841 from pat ( 431)	163	46.2	0.23
gi 112037166 gb ABH88468.1	Sequence 3841 from pat ( 431)	163	46.2	0.23
gi 259365177 gb ACW59123.1	Sequence 3841 from pat ( 431)	163	46.2	0.23
gi 281044951 gb ADA23679.1	Sequence 4217 from pat ( 431)	163	46.2	0.23
gi 142433170 gb ECY04151.1	hypothetical protein G ( 91)	155	43.9	0.23
gi 143435502 gb EDE94103.1	hypothetical protein G ( 295)	161	45.6	0.23
gi 136764602 gb EBQ25258.1	hypothetical protein G ( 437)	163	46.2	0.23
gi 139795572 gb ECH51845.1	hypothetical protein G ( 136)	157	44.5	0.24
gi 142812888 gb EDA77094.1	hypothetical protein G ( 444)	163	46.2	0.24
gi 136512696 gb EB064623.1	hypothetical protein G ( 369)	162	45.9	0.24
gi 138967682 gb ECC62769.1	hypothetical protein G ( 66)	153	43.4	0.25
gi 140811098 gb ECN67441.1	hypothetical protein G ( 262)	160	45.4	0.25
gi 137859016 gb EBW40859.1	hypothetical protein G ( 69)	153	43.4	0.26
gi 141741379 gb ECT05533.1	hypothetical protein G ( 151)	157	44.5	0.26
gi 140887232 gb ECO18192.1	hypothetical protein G ( 273)	160	45.4	0.26
gi 140777409 gb ECN43706.1	hypothetical protein G ( 71)	153	43.4	0.26
gi 136556862 gb EB092852.1	hypothetical protein G ( 339)	161	45.7	0.26
gi 138016246 gb EBX28950.1	hypothetical protein G ( 156)	157	44.5	0.26
gi 268625775 gb EEZ58175.1	UDP-N-acetylglucosamin ( 417)	162	46.0	0.26
gi 143569979 gb EDF68808.1	hypothetical protein G ( 417)	162	46.0	0.26
gi 139046784 gb ECD04192.1	hypothetical protein G ( 191)	158	44.8	0.26
gi 27316187 gb AAO05321.1	AE016749_267 UDP-N-acety ( 419)	162	46.0	0.26
gi 134708780 gb EBC95330.1	hypothetical protein G ( 158)	157	44.5	0.26
gi 135174336 gb EBF96923.1	hypothetical protein G ( 284)	160	45.4	0.26
gi 142003655 gb ECU75567.1	hypothetical protein G ( 235)	159	45.1	0.27
gi 141800948 gb ECT35640.1	hypothetical protein G ( 93)	154	43.7	0.28
gi 135671323 gb EBJ12248.1	hypothetical protein G ( 301)	160	45.4	0.28
gi 194345478 gb EDX26444.1	UDP-N-acetylglucosamin ( 448)	162	46.0	0.28
gi 136628582 gb EBP38113.1	hypothetical protein G ( 115)	155	44.0	0.28
gi 140462588 gb ECL90454.1	hypothetical protein G ( 170)	157	44.6	0.28
gi 138672413 gb ECB20843.1	hypothetical protein G ( 309)	160	45.4	0.28
gi 139391915 gb ECE79642.1	hypothetical protein G ( 117)	155	44.0	0.28
gi 140266572 gb ECK66197.1	hypothetical protein G ( 179)	157	44.6	0.29
gi 135393658 gb EBH33888.1	hypothetical protein G ( 82)	153	43.5	0.29
gi 141399980 gb ECR56139.1	hypothetical protein G ( 122)	155	44.0	0.29

gi 142382350 gb ECX66588.1	hypothetical protein G ( 334)	160	45.4	0.3
gi 142664031 gb ECZ68739.1	hypothetical protein G ( 72)	152	43.2	0.31
gi 262397170 emb CAX66184.1	UDP-N-acetylglucosami ( 421)	161	45.7	0.31
gi 218323683 emb CAV20010.1	UDP-N-acetylglucosami ( 422)	161	45.7	0.31
gi 140434033 gb ECI72356.1	hypothetical protein G ( 60)	151	42.9	0.31
gi 138945420 gb ECC53340.1	hypothetical protein G ( 73)	152	43.2	0.31
gi 1049100 gb AAA97400.1	encodes EPSP synthase do ( 108)	154	43.8	0.31
gi 138201102 gb EBY37054.1	hypothetical protein G ( 108)	154	43.8	0.31
gi 9136181 gb AAB32920.1	enolpyruvylshikimate-3-ph ( 108)	154	43.8	0.31
gi 40134109 gb AAR60243.1	Sequence 6777 from pate ( 237)	158	44.9	0.31
gi 138381113 gb EBZ27633.1	hypothetical protein G ( 110)	154	43.8	0.31
gi 148530739 gb ABQ82738.1	UDP-N-acetylglucosamin ( 438)	161	45.8	0.32
gi 183224458 dbj BAG24975.1	UDP-N-acetylglucosami ( 438)	161	45.8	0.32
gi 138764059 gb ECB78674.1	hypothetical protein G ( 201)	157	44.6	0.32
gi 140728797 gb ECN11886.1	hypothetical protein G ( 300)	159	45.2	0.32
gi 139167173 gb ECB86616.1	hypothetical protein G ( 302)	159	45.2	0.32
gi 75701382 gb ABA21058.1	UDP-N-acetylglucosamine ( 447)	161	45.8	0.32
gi 142118410 gb ECV72506.1	hypothetical protein G ( 369)	160	45.5	0.32
gi 143558889 gb EDF62870.1	hypothetical protein G ( 78)	152	43.2	0.33
gi 134904511 gb EBE20045.1	hypothetical protein G ( 79)	152	43.2	0.33
gi 140241779 gb ECK49161.1	hypothetical protein G ( 142)	155	44.1	0.33
gi 138737511 gb ECB66481.1	hypothetical protein G ( 118)	154	43.8	0.33
gi 142767344 gb EDA43135.1	hypothetical protein G ( 119)	154	43.8	0.33
gi 135078388 gb EBF35366.1	hypothetical protein G ( 388)	160	45.5	0.34
gi 134333526 gb EBA70769.1	hypothetical protein G ( 219)	157	44.7	0.34
gi 142204617 gb ECW37535.1	hypothetical protein G ( 396)	160	45.5	0.34
gi 135058427 gb ECV72506.1	hypothetical protein G ( 396)	160	45.5	0.34
gi 135139598 gb EBF74601.1	hypothetical protein G ( 396)	160	45.5	0.34
gi 143080046 gb EDC66841.1	hypothetical protein G ( 69)	151	43.0	0.34
gi 142693403 gb ECZ89811.1	hypothetical protein G ( 102)	153	43.5	0.34
gi 136442255 gb EB019063.1	hypothetical protein G ( 272)	158	44.9	0.35
gi 140977470 gb ECO79842.1	hypothetical protein G ( 276)	158	44.9	0.35
gi 136868207 gb EBQ94410.1	hypothetical protein G ( 338)	159	45.2	0.35
gi 262295835 gb EEY83766.1	3-phosphoshikimate 1-c ( 411)	160	45.5	0.35
gi 254951509 gb ACT96209.1	3-phosphoshikimate 1-c ( 411)	160	45.5	0.35
gi 136884865 gb EBR04394.1	hypothetical protein G ( 87)	152	43.3	0.35
gi 267983528 gb ACY83357.1	UDP-N-acetylglucosamin ( 418)	160	45.5	0.36
gi 262316857 gb EEY97895.1	UDP-N-acetylglucosamin ( 419)	160	45.5	0.36
gi 55737521 gb AAV61163.1	UDP-N-acetylglucosamine ( 423)	160	45.5	0.36
gi 55739449 gb AAV63090.1	UDP-N-acetylglucosamine ( 423)	160	45.5	0.36
gi 140010510 gb ECT98634.1	hypothetical protein G ( 132)	154	43.8	0.36
gi 142341681 gb ECX39143.1	hypothetical protein G ( 351)	159	45.2	0.36
gi 142539260 gb ECY81081.1	hypothetical protein G ( 352)	159	45.2	0.36
gi 58253892 gb AAV42129.1	udp-n-acetylglucosamine ( 431)	160	45.5	0.37
gi 239684576 gb ACS09904.1	Sequence 120 from pate ( 441)	160	45.5	0.37
gi 239678519 gb ACS07433.1	Sequence 120 from pate ( 441)	160	45.5	0.37
gi 182996761 gb ACC31203.1	Sequence 120 from pate ( 441)	160	45.5	0.37
gi 143330349 gb EDE37054.1	hypothetical protein G ( 376)	159	45.3	0.38
gi 142007843 gb ECU78491.1	hypothetical protein G ( 79)	151	43.0	0.38
gi 135188364 gb EBG05889.1	hypothetical protein G ( 177)	155	44.1	0.39
gi 134973646 gb EBE66138.1	hypothetical protein G ( 321)	158	45.0	0.39
gi 136360997 gb EBN64555.1	hypothetical protein G ( 394)	159	45.3	0.4
gi 141974299 gb ECU55148.1	hypothetical protein G ( 83)	151	43.0	0.4
gi 135138671 gb EBF74016.1	hypothetical protein G ( 396)	159	45.3	0.4
gi 135136907 gb EBF72904.1	hypothetical protein G ( 396)	159	45.3	0.4
gi 141226028 gb ECC05067.1	hypothetical protein G ( 85)	151	43.0	0.41
gi 135127298 gb EBF66688.1	hypothetical protein G ( 276)	157	44.7	0.41
gi 135459070 gb EBH77757.1	hypothetical protein G ( 86)	151	43.0	0.41
gi 142185192 gb ECW22719.1	hypothetical protein G ( 278)	157	44.7	0.41
gi 142121563 gb ECV74769.1	hypothetical protein G ( 411)	159	45.3	0.41
gi 138911545 gb ECC39122.1	hypothetical protein G ( 105)	152	43.3	0.41
gi 134698699 gb EBC89437.1	hypothetical protein G ( 282)	157	44.7	0.41

gi 135359014 gb EBH10607.1	hypothetical protein G ( 419)	159	45.3	0.42
gi 142169325 gb ECW10606.1	hypothetical protein G ( 419)	159	45.3	0.42
gi 15638978 gb JAA86297.2	UDP-N-acetylglucosamine ( 423)	159	45.3	0.42
gi 46881995 gb AAT05289.1	UDP-N-acetylglucosamine ( 423)	159	45.3	0.42
gi 225877561 emb CAS06275.1	Putative UDP-N-acetyl ( 423)	159	45.3	0.42
gi 16415233 emb CAC97923.1	murZ [Listeria innocua ( 423)	159	45.3	0.42
gi 134366442 gb EBA92583.1	hypothetical protein G ( 194)	155	44.2	0.42
gi 139444518 gb ECF14274.1	hypothetical protein G ( 160)	154	43.9	0.42
gi 135104860 gb EBF52299.1	hypothetical protein G ( 288)	157	44.7	0.42
gi 143380233 gb EDB67130.1	hypothetical protein G ( 430)	159	45.3	0.42
gi 143230676 gb EDD76083.1	hypothetical protein G ( 110)	152	43.3	0.43
gi 140674676 gb ECM73477.1	hypothetical protein G ( 110)	152	43.3	0.43
gi 117648040 gb ABK52142.1	UDP-N-acetylglucosamin ( 432)	159	45.3	0.43
gi 137615022 gb EBV06424.1	hypothetical protein G ( 293)	157	44.7	0.43
gi 137662455 gb EBV31632.1	hypothetical protein G ( 112)	152	43.3	0.43
gi 197703480 gb EDY49292.1	UDP-N-acetylglucosamin ( 445)	159	45.3	0.44
gi 139341795 gb ECE56379.1	hypothetical protein G ( 114)	152	43.4	0.44
gi 137667452 gb EBV34442.1	hypothetical protein G ( 169)	154	43.9	0.44
gi 134532404 gb EBB90389.1	hypothetical protein G ( 64)	149	42.5	0.44
gi 138296296 gb EBY85354.1	hypothetical protein G ( 78)	150	42.8	0.44
gi 162688145 gb EDQ74524.1	predicted protein [Phy ( 454)	159	45.3	0.44
gi 135489195 gb EBH97504.1	hypothetical protein G ( 119)	152	43.4	0.45
gi 139579622 gb ECG03906.1	hypothetical protein G ( 120)	152	43.4	0.45
gi 135176976 gb EBF98626.1	hypothetical protein G ( 324)	157	44.8	0.46
gi 141068175 gb ECP40329.1	hypothetical protein G ( 101)	151	43.1	0.46
gi 140383936 gb ECL40504.1	hypothetical protein G ( 123)	152	43.4	0.46
gi 142094653 gb ECV54445.1	hypothetical protein G ( 282)	156	44.5	0.48
gi 142278045 gb ECW92094.1	hypothetical protein G ( 418)	158	45.1	0.48
gi 189420145 gb ACD94543.1	UDP-N-acetylglucosamin ( 419)	158	45.1	0.48
gi 115424369 emb CAJ50922.1	UDP-N-acetylglucosami ( 421)	158	45.1	0.49
gi 217332615 gb ACK38409.1	UDP-N-acetylglucosamin ( 439)	158	45.1	0.5
gi 142610502 gb ECZ31048.1	hypothetical protein G ( 440)	158	45.1	0.5
gi 124515207 gb EAY56718.1	UDP-N-acetylglucosamin ( 442)	158	45.1	0.5
gi 140159165 gb ECJ91878.1	hypothetical protein G ( 300)	156	44.5	0.51
gi 136457490 gb EBO28915.1	hypothetical protein G ( 248)	155	44.3	0.51
gi 17135152 dbj BAB77698.1	UDP-N-acetylglucosamin ( 447)	158	45.1	0.51
gi 139586669 gb ECG08967.1	hypothetical protein G ( 117)	151	43.1	0.52
gi 141825909 gb ECT51361.1	hypothetical protein G ( 122)	151	43.2	0.54
gi 135611902 gb EBI75428.1	hypothetical protein G ( 125)	151	43.2	0.55
gi 141048527 gb ECP27825.1	hypothetical protein G ( 125)	151	43.2	0.55
gi 136659917 gb EBP56687.1	hypothetical protein G ( 126)	151	43.2	0.55
gi 256739508 gb EEU52832.1	3-phosphoshikimate 1-c ( 411)	157	44.9	0.55
gi 28190677 gb AAO33155.1	putative 3-phosphoshiki ( 190)	153	43.7	0.56
gi 142185596 gb ECW23027.1	hypothetical protein G ( 419)	157	44.9	0.56
gi 150956962 gb ABR78992.1	UDP-N-acetylglucosamin ( 419)	157	44.9	0.56
gi 238549023 dbj BAH65374.1	UDP-N-acetylglucosami ( 419)	157	44.9	0.56
gi 77995958 gb ABB14857.1	UDP-N-acetylglucosamine ( 420)	157	44.9	0.56
gi 270281718 gb EFA27550.1	UDP-N-acetylglucosamin ( 423)	157	44.9	0.57
gi 40111601 gb AAR53881.1	Sequence 13598 from pat ( 423)	157	44.9	0.57
gi 138264746 gb EBY66748.1	hypothetical protein G ( 61)	147	42.1	0.57
gi 138263743 gb EBY66096.1	hypothetical protein G ( 162)	152	43.5	0.57
gi 142642921 gb ECZ53921.1	hypothetical protein G ( 91)	149	42.6	0.58
gi 142115059 gb ECV70074.1	hypothetical protein G ( 437)	157	44.9	0.58
gi 137997686 gb EBX19252.1	hypothetical protein G ( 94)	149	42.6	0.59
gi 141350070 gb ECR21239.1	hypothetical protein G ( 65)	147	42.1	0.6
gi 141429393 gb ECR76551.1	hypothetical protein G ( 121)	150	42.9	0.62
gi 140885888 gb ECO17231.1	hypothetical protein G ( 68)	147	42.1	0.62
gi 135625189 gb EBI83682.1	hypothetical protein G ( 101)	149	42.7	0.63
gi 254672604 emb CBA06320.1	UDP-N-acetylglucosami ( 398)	156	44.6	0.63
gi 142197920 gb ECW32422.1	hypothetical protein G ( 400)	156	44.6	0.63
gi 140854433 gb ECN95237.1	hypothetical protein G ( 276)	154	44.1	0.64
gi 137217195 gb EBS85944.1	hypothetical protein G ( 104)	149	42.7	0.64

gi 143351054 gb EDE49890.1	hypothetical protein G ( 228)	153	43.8	0.64
gi 149938155 gb ABR44852.1	3-phosphoshikimate 1-c ( 411)	156	44.6	0.64
gi 139540894 gb ECF77125.1	hypothetical protein G ( 156)	151	43.2	0.65
gi 73912410 dbj BAE20404.1	5-enolpyruvylshikimate ( 231)	153	43.8	0.65
gi 135159134 gb EBF87155.1	hypothetical protein G ( 342)	155	44.4	0.65
gi 261391570 emb CAX49003.1	UDP-N-acetylglucosami ( 417)	156	44.6	0.65
gi 149950500 gb ABR49028.1	UDP-N-acetylglucosamin ( 417)	156	44.6	0.65
gi 7225236 gb AAF40490.1	UDP-N-acetylglucosamine ( 417)	156	44.6	0.65
gi 260404638 gb EEW98154.1	UDP-N-acetylglucosamin ( 418)	156	44.6	0.65
gi 143112917 gb EDC90785.1	hypothetical protein G ( 418)	156	44.6	0.65
gi 262313451 gb EEY94536.1	UDP-N-acetylglucosamin ( 419)	156	44.6	0.65
gi 142171308 gb ECW12107.1	hypothetical protein G ( 419)	156	44.6	0.65
gi 158140612 gb ABW18924.1	UDP-N-acetylglucosamin ( 421)	156	44.6	0.66
gi 135223693 gb EBG26997.1	hypothetical protein G ( 421)	156	44.6	0.66
gi 206739060 gb ACI18138.1	UDP-N-acetylglucosamin ( 425)	156	44.7	0.66
gi 167293014 gb ABZ45878.1	Sequence 19816 from pa ( 425)	156	44.7	0.66
gi 217502985 gb ACK50394.1	UDP-N-acetylglucosamin ( 429)	156	44.7	0.67
gi 137370618 gb EBT71981.1	hypothetical protein G ( 292)	154	44.1	0.67
gi 140528342 gb ECM11325.1	hypothetical protein G ( 75)	147	42.1	0.67
gi 137182770 gb EBS66594.1	hypothetical protein G ( 75)	147	42.1	0.67
gi 186468218 gb ACC84019.1	UDP-N-acetylglucosamin ( 435)	156	44.7	0.67
gi 139577870 gb ECG02772.1	hypothetical protein G ( 299)	154	44.1	0.68
gi 135372780 gb EBH19824.1	hypothetical protein G ( 93)	148	42.4	0.68
gi 135835202 gb ACI18138.1	hypothetical protein G ( 214)	152	43.6	0.71
gi 136266331 gb EBN00256.1	hypothetical protein G ( 263)	153	43.8	0.72
gi 228229599 gb ACP87264.1	Sequence 9206 from pat ( 397)	155	44.4	0.73
gi 138012622 gb EBX21133.1	hypothetical protein G ( 188)	151	43.3	0.75
gi 121051288 emb CAM07564.1	UDP-N-acetylglucosami ( 417)	155	44.4	0.76
gi 228228558 gb ACP86223.1	Sequence 7708 from pat ( 417)	155	44.4	0.76
gi 120867520 emb CAM11297.1	UDP-N-acetylglucosami ( 417)	155	44.4	0.76
gi 254667575 emb CBA03311.1	UDP-N-acetylglucosami ( 417)	155	44.4	0.76
gi 135022987 gb EBE99542.1	hypothetical protein G ( 419)	155	44.4	0.76
gi 143574769 gb EDF71190.1	hypothetical protein G ( 421)	155	44.4	0.76
gi 143651789 gb EDG07916.1	hypothetical protein G ( 421)	155	44.4	0.76
gi 225699712 emb CAW93447.1	UDP-N-acetylglucosami ( 423)	155	44.4	0.77
gi 242390773 dbj BAH81232.1	UDP-N-acetylglucosami ( 423)	155	44.4	0.77
gi 142908315 gb EDB43847.1	hypothetical protein G ( 90)	147	42.2	0.77
gi 257814094 gb EEV42852.1	UDP-N-acetylglucosamin ( 433)	155	44.4	0.78
gi 260074292 gb EEW62614.1	UDP-N-acetylglucosamin ( 433)	155	44.4	0.78
gi 260077191 gb EEW64911.1	UDP-N-acetylglucosamin ( 433)	155	44.4	0.78
gi 257819862 gb EEV47010.1	UDP-N-acetylglucosamin ( 433)	155	44.4	0.78
gi 134656392 gb EBC64520.1	hypothetical protein G ( 91)	147	42.2	0.78
gi 140108185 gb ECJ59176.1	hypothetical protein G ( 63)	145	41.6	0.79
gi 139398701 gb ECE84261.1	hypothetical protein G ( 94)	147	42.2	0.8
gi 143758365 gb EDG63465.1	hypothetical protein G ( 306)	153	43.9	0.81
gi 139102418 gb ECD42987.1	hypothetical protein G ( 212)	151	43.3	0.82
gi 138714387 gb EBC50183.1	hypothetical protein G ( 260)	152	43.6	0.83
gi 141298287 gb ECQ92990.1	hypothetical protein G ( 100)	147	42.2	0.84
gi 138900786 gb ECC34331.1	hypothetical protein G ( 123)	148	42.5	0.85
gi 141941218 gb ECU32154.1	hypothetical protein G ( 102)	147	42.2	0.85
gi 141594479 gb ECS50739.1	hypothetical protein G ( 85)	146	41.9	0.86
gi 48430322 gb AAT43187.1	3-phosphoshikimate 1-ca ( 411)	154	44.2	0.87
gi 49529866 emb CAG67578.1	UDP-N-acetylglucosamin ( 418)	154	44.2	0.88
gi 213055248 gb ACJ40150.1	UDP-N-acetylglucosamin ( 418)	154	44.2	0.88
gi 193076481 gb ABO11129.2	UDP-N-acetylglucosamin ( 418)	154	44.2	0.88
gi 183208544 gb ACC55942.1	UDP-N-acetylglucosamin ( 418)	154	44.2	0.88
gi 169150049 emb CAM87943.1	UDP-N-acetylglucosami ( 418)	154	44.2	0.88
gi 213986105 gb ACJ56404.1	UDP-N-acetylglucosamin ( 418)	154	44.2	0.88
gi 143251316 gb EDD90247.1	hypothetical protein G ( 418)	154	44.2	0.88
gi 262081987 gb ACY17956.1	UDP-N-acetylglucosamin ( 419)	154	44.2	0.88
gi 139514508 gb ECF59424.1	hypothetical protein G ( 107)	147	42.2	0.89
gi 33704550 gb AAQ29025.1	Sequence 5997 from pate ( 420)	154	44.2	0.89

```
>>gi|62318479|dbj|BAD94823.1| 5-enol-pyruvylshikimate-3- (527 aa)
  initn: 2968 initl: 2927 opt: 2961 Z-score: 3477.5 bits: 655.3 E(): 1.2e-184
Smith-Waterman score: 2961; 88.701% identity (93.974% similar) in 531 aa overlap
(671-1201:1-527)
```

```

          1010      1020      1030      1040      1050      1060
frame3 ALLVPGSDVTILNVLNMPTRTGLILTLQEMGADIEVINRLAGGEDVADLRVRSSTLKG
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
gi|623  ALLVPGSDVTILNVLNMPTRTGLILTLQEMGADIEVINRLAGGEDVADLRVRSSTLKG
          330      340      350      360      370      380

```

```
>>gi|27549260|gb|AAO17037.1| CP4EPSPS protein [synthetic (455 aa)
  initn: 2917 init1: 2917 opt: 2917 Z-score: 3426.7 bits: 645.7 E(): 8.3e-182
Smith-Waterman score: 2917; 100.000 identity (100.000 similar) in 455 aa overlap
(747-1201:1-455)
```

```

      1080      1090      1100      1110      1120      1130
frame3 LAVAAAF AEGATV MNGLEEL RVKESDR LSAVANGL KLN GVDCDEGETSL VVRGRPDGKGL
      :

```

```

1080      1090      1100      1110      1120      1130
frame3  LAVAAAFAGGATVMNGLEELRVKESDRLSAVANGKLKNGVDCDEGETSLVVRGRPDGKGL
       .....
gi|281  LAVAAAFAGGATVMNGLEELRVKESDRLSAVANGKLKNGVDCDEGETSLVVRGRPDGKGL
       340      350      360      370      380      390

```

```

      1140      1150      1160      1170      1180      1190
frame3 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDATMIATSFPFMDLMAGLGAKIELS
       ::::::::::::::::::::::::::::::::::::::::::::::::::::
```

```

gi|182 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
      400      410      420      430      440      450

      1200      1210      1220      1230      1240      1250
frame3 DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEFYYGI
      :::::
gi|182 DTKAA

>>gi|144974765|gb|ABP12476.1| Sequence 70 from patent US (455 aa)
      initn: 2917 initl: 2917 opt: 2917 Z-score: 3426.7 bits: 645.7 E(): 8.3e-182
Smith-Waterman score: 2917; 100.000% identity (100.000% similar) in 455 aa overlap
(747-1201:1-455)

      720      730      740      750      760      770
frame3 SSWGLKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
      ::::::::::::::::::::::::::::::::::::::
gi|144      MLHGASSRPATARKSSGLSGTVRIPGDKSI
      10      20      30

      780      790      800      810      820      830
frame3 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVGNGGLLAP
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
gi|144 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVGNGGLLAP
      40      50      60      70      80      90

      840      850      860      870      880      890
frame3 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEDGD
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
gi|144 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEDGD
      100      110      120      130      140      150

      900      910      920      930      940      950
frame3 RLPVTLRGPKTPTPTITYRVPMSAQVKSAVLLAGLNTPGITTVIEPIMTRDHTKMLQGF
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
gi|144 RLPVTLRGPKTPTPTITYRVPMSAQVKSAVLLAGLNTPGITTVIEPIMTRDHTKMLQGF
      160      170      180      190      200      210

      960      970      980      990      1000      1010
frame3 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFFPLVAALLVPGSDVTILNVLM
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
gi|144 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFFPLVAALLVPGSDVTILNVLM
      220      230      240      250      260      270

      1020      1030      1040      1050      1060      1070
frame3 NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
gi|144 NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
      280      290      300      310      320      330

      1080      1090      1100      1110      1120      1130
frame3 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGLKLNVDCEGETSLVVRGRPDGKGL
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
gi|144 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGLKLNVDCEGETSLVVRGRPDGKGL
      340      350      360      370      380      390

      1140      1150      1160      1170      1180      1190
frame3 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
gi|144 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
      400      410      420      430      440      450

```

```

      1200      1210      1220      1230      1240      1250
frame3 DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEFYYGI
      :::::
gi|144 DTKAA

>>gi|2484147|gb|AAB72283.1|I49176 Sequence 3 from patent (455 aa)
      initn: 2909 initl: 2909 opt: 2909 Z-score: 3417.2 bits: 644.0 E(): 2.8e-181
Smith-Waterman score: 2909; 99.780% identity (99.780% similar) in 455 aa overlap
(747-1201:1-455)

      720      730      740      750      760      770
frame3 SSWGLKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
      ::::::::::::::::::::::::::::::::::::::
gi|248      MSHGASSRPATARKSSGLSGTVRIPGDKSI
      10      20      30

      780      790      800      810      820      830
frame3 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVGNGGLLAP
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
gi|248 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVGNGGLLAP
      40      50      60      70      80      90

      840      850      860      870      880      890
frame3 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEDGD
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
gi|248 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEDGD
      100      110      120      130      140      150

      900      910      920      930      940      950
frame3 RLPVTLRGPKTPTPTITYRVPMSAQVKSAVLLAGLNTPGITTVIEPIMTRDHTKMLQGF
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
gi|248 RLPVTLRGPKTPTPTITYRVPMSAQVKSAVLLAGLNTPGITTVIEPIMTRDHTKMLQGF
      160      170      180      190      200      210

      960      970      980      990      1000      1010
frame3 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFFPLVAALLVPGSDVTILNVLM
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
gi|248 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFFPLVAALLVPGSDVTILNVLM
      220      230      240      250      260      270

      1020      1030      1040      1050      1060      1070
frame3 NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
gi|248 NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
      280      290      300      310      320      330

      1080      1090      1100      1110      1120      1130
frame3 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGLKLNVDCEGETSLVVRGRPDGKGL
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
gi|248 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGLKLNVDCEGETSLVVRGRPDGKGL
      340      350      360      370      380      390

      1140      1150      1160      1170      1180      1190
frame3 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
gi|248 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
      400      410      420      430      440      450

      1200      1210      1220      1230      1240      1250
frame3 DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEFYYGI
      :::::

```

gi|248 DTKAA

>>gi|5957547|gb|AAE08221.1| Sequence 3 from patent US 58 (455 aa)  
initn: 2909 initl: 2909 opt: 2909 Z-score: 3417.2 bits: 644.0 E(): 2.8e-181  
Smith-Waterman score: 2909; 99.780% identity (99.780% similar) in 455 aa overlap  
(747-1201:1-455)

	720	730	740	750	760	770
frame3	SSWGLKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI					
gi 595						

	780	790	800	810	820	830
frame3	SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVGNGGLLAP					
gi 595						

	840	850	860	870	880	890
frame3	EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEDGD					
gi 595						

	900	910	920	930	940	950
frame3	RLPVTLRGPKTPTPTITYRVPMSAQVKSALLAGLNTPGITTVIEPIIMTRDHTEKMLQGF					
gi 595						

	960	970	980	990	1000	1010
frame3	GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM					
gi 595						

	1020	1030	1040	1050	1060	1070
frame3	NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI					
gi 595						

	1080	1090	1100	1110	1120	1130
frame3	LAVAAAFAGATVMNGLEELRVKESDRLSAVANGLKLNVDCEGETSLVVRGRPDGKGL					
gi 595						

	1140	1150	1160	1170	1180	1190
frame3	GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS					
gi 595						

	1200	1210	1220	1230	1240	1250
frame3	DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEYYGI					
gi 595						

>>gi|217243929|gb|ACK18296.1| Sequence 1 from patent US (455 aa)  
initn: 2909 initl: 2909 opt: 2909 Z-score: 3417.2 bits: 644.0 E(): 2.8e-181  
Smith-Waterman score: 2909; 99.780% identity (99.780% similar) in 455 aa overlap  
(747-1201:1-455)

	720	730	740	750	760	770
frame3	SSWGLKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI					
gi 217						

	780	790	800	810	820	830
frame3	SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVGNGGLLAP					
gi 217						

	840	850	860	870	880	890
frame3	EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEDGD					
gi 217						

	900	910	920	930	940	950
frame3	RLPVTLRGPKTPTPTITYRVPMSAQVKSALLAGLNTPGITTVIEPIIMTRDHTEKMLQGF					
gi 217						

	960	970	980	990	1000	1010
frame3	GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM					
gi 217						

	1020	1030	1040	1050	1060	1070
frame3	NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI					
gi 217						

	1080	1090	1100	1110	1120	1130
frame3	LAVAAAFAGATVMNGLEELRVKESDRLSAVANGLKLNVDCEGETSLVVRGRPDGKGL					
gi 217						

	1140	1150	1160	1170	1180	1190
frame3	GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS					
gi 217						

	1200	1210	1220	1230	1240	1250
frame3	DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEYYGI					
gi 217						

>>gi|197053766|gb|ACH25464.1| Sequence 16 from patent US (455 aa)  
initn: 2909 initl: 2909 opt: 2909 Z-score: 3417.2 bits: 644.0 E(): 2.8e-181

Smith-Waterman score: 2909; 99.780% identity (99.780% similar) in 455 aa overlap (747-1201:1-455)

```

      720      730      740      750      760      770
frame3 SSWGLKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
      : .....
gi|197      MSHGASSRPATARKSSGLSGTVRIPGDKSI
              10      20      30

      780      790      800      810      820      830
frame3 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVGNGGLLAP
      : .....
gi|197 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVGNGGLLAP
      40      50      60      70      80      90

      840      850      860      870      880      890
frame3 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEGD
      : .....
gi|197 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEGD
      100     110     120     130     140     150

      900      910      920      930      940      950
frame3 RLPVTLRGPKTPTPITYRVPMASAVLLAGLNTPGITTVIEPIMTRDHTKMLQGF
      : .....
gi|197 RLPVTLRGPKTPTPITYRVPMASAVLLAGLNTPGITTVIEPIMTRDHTKMLQGF
      160     170     180     190     200     210

      960      970      980      990     1000     1010
frame3 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM
      : .....
gi|197 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM
      220     230     240     250     260     270

      1020     1030     1040     1050     1060     1070
frame3 NPTRTGLLILTQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
      : .....
gi|197 NPTRTGLLILTQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
      280     290     300     310     320     330

      1080     1090     1100     1110     1120     1130
frame3 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGKLNGVDCDEGETSLVVGRPDGKGL
      : .....
gi|197 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGKLNGVDCDEGETSLVVGRPDGKGL
      340     350     360     370     380     390

      1140     1150     1160     1170     1180     1190
frame3 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
      : .....
gi|197 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
      400     410     420     430     440     450

      1200     1210     1220     1230     1240     1250
frame3 DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEFYGI
      : : : : :
gi|197 DTKAA
```

>>gi|144974726|gb|ABP12437.1| Sequence 3 from patent US (455 aa)  
initn: 2909 initl: 2909 opt: 2909 Z-score: 3417.2 bits: 644.0 E(): 2.8e-181  
Smith-Waterman score: 2909; 99.780% identity (99.780% similar) in 455 aa overlap (747-1201:1-455)

```

      720      730      740      750      760      770
frame3 SSWGLKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
      : .....
gi|144      MSHGASSRPATARKSSGLSGTVRIPGDKSI
              10      20      30

      780      790      800      810      820      830
frame3 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVGNGGLLAP
      : .....
gi|144 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVGNGGLLAP
      40      50      60      70      80      90

      840      850      860      870      880      890
frame3 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEGD
      : .....
gi|144 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEGD
      100     110     120     130     140     150

      900      910      920      930      940      950
frame3 RLPVTLRGPKTPTPITYRVPMASAVLLAGLNTPGITTVIEPIMTRDHTKMLQGF
      : .....
gi|144 RLPVTLRGPKTPTPITYRVPMASAVLLAGLNTPGITTVIEPIMTRDHTKMLQGF
      160     170     180     190     200     210

      960      970      980      990     1000     1010
frame3 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM
      : .....
gi|144 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM
      220     230     240     250     260     270

      1020     1030     1040     1050     1060     1070
frame3 NPTRTGLLILTQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
      : .....
gi|144 NPTRTGLLILTQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
      280     290     300     310     320     330

      1080     1090     1100     1110     1120     1130
frame3 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGKLNGVDCDEGETSLVVGRPDGKGL
      : .....
gi|144 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGKLNGVDCDEGETSLVVGRPDGKGL
      340     350     360     370     380     390

      1140     1150     1160     1170     1180     1190
frame3 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
      : .....
gi|144 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
      400     410     420     430     440     450

      1200     1210     1220     1230     1240     1250
frame3 DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEFYGI
      : : : : :
gi|144 DTKAA
```

>>gi|2485224|gb|AAB73360.1|I44449 Sequence 3 from patent (455 aa)  
initn: 2909 initl: 2909 opt: 2909 Z-score: 3417.2 bits: 644.0 E(): 2.8e-181  
Smith-Waterman score: 2909; 99.780% identity (99.780% similar) in 455 aa overlap (747-1201:1-455)

```

      720      730      740      750      760      770
frame3 SSWGLKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
      : .....

```

```

gi|248                               MSHGASSRPATARKSSGLSGTVRIPGDKSI
                               10         20         30

      780      790      800      810      820      830
frame3 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGGGLLAP
      .....
gi|248 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGGGLLAP
      40         50         60         70         80         90

      840      850      860      870      880      890
frame3 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEDGD
      .....
gi|248 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEDGD
      100        110        120        130        140        150

      900      910      920      930      940      950
frame3 RLPVTLRGPKTPTPITYRVPMSAQVKSAVLLAGLNTPGITTVIEPIMTRDHTKMLQGF
      .....
gi|248 RLPVTLRGPKTPTPITYRVPMSAQVKSAVLLAGLNTPGITTVIEPIMTRDHTKMLQGF
      160        170        180        190        200        210

      960      970      980      990      1000     1010
frame3 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM
      .....
gi|248 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM
      220        230        240        250        260        270

      1020     1030     1040     1050     1060     1070
frame3 NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
      .....
gi|248 NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
      280        290        300        310        320        330

      1080     1090     1100     1110     1120     1130
frame3 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGKLNGVDCDEGETSLVVRGRPDGKGL
      .....
gi|248 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGKLNGVDCDEGETSLVVRGRPDGKGL
      340        350        360        370        380        390

      1140     1150     1160     1170     1180     1190
frame3 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
      .....
gi|248 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
      400        410        420        430        440        450

      1200     1210     1220     1230     1240     1250
frame3 DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEFYGYI
      .....
gi|248 DTKAA

>>gi|15073182|emb|CAC41690.1| Putative 3-phosphoshikimat (455 aa)
  initn: 2685 initl: 2685 opt: 2685 Z-score: 3153.7 bits: 595.2 E(): 1.3e-166
Smith-Waterman score: 2685; 90.549% identity (97.143% similar) in 455 aa overlap
(747-1201:1-455)

      720      730      740      750      760      770
frame3 SSWGLKKSGMTLIGSELRLPLKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|150 MSHGNSNRPATARKSSDLKGTLRIPGDKSI
      10         20         30

```

```

      780      790      800      810      820      830
frame3 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGGGLLAP
      .....
gi|150 SHRSFMFGGLAAGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGGALLAP
      40         50         60         70         80         90

      840      850      860      870      880      890
frame3 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEDGD
      .....
gi|150 EAPLDFGNAAGTGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLDPLREMGVQVKSAGD
      100        110        120        130        140        150

      900      910      920      930      940      950
frame3 RLPVTLRGPKTPTPITYRVPMSAQVKSAVLLAGLNTPGITTVIEPIMTRDHTKMLQGF
      .....
gi|150 RLPVTLRGPKTPNPITYRVPMSAQVKSAVLLAGLNTPGITTVIEPIMTRDHTKMLQGF
      160        170        180        190        200        210

      960      970      980      990      1000     1010
frame3 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM
      .....
gi|150 GANLTVETDAEGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAGLIVPGSDITILNVLM
      220        230        240        250        260        270

      1020     1030     1040     1050     1060     1070
frame3 NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
      .....
gi|150 NPTRTGLILTLQEMGANIEVMNKRLAGGEDVADLRVRHSELKGVTVPEDRAPSMIDEYPV
      280        290        300        310        320        330

      1080     1090     1100     1110     1120     1130
frame3 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGKLNGVDCDEGETSLVVRGRPDGKGL
      .....
gi|150 LAVAAAFAGETTVMNGLEELRVKESDRLSAVADGLKNGVDCDEGEASLVVRGRPGGKGL
      340        350        360        370        380        390

      1140     1150     1160     1170     1180     1190
frame3 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
      : . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|150 GKISGGQVKTLDHRIAMSFLVMGLASEHPVTVDDATMIATSFPEFMGLMTGLGAKIEEA
      400        410        420        430        440        450

      1200     1210     1220     1230     1240     1250
frame3 DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEFYGYI
      : : : : :
gi|150 ENKAA

>>gi|227343087|gb|ACP27305.1| 3-phosphoshikimate 1-carbo (448 aa)
  initn: 2670 initl: 2670 opt: 2670 Z-score: 3136.2 bits: 591.9 E(): 1.3e-165
Smith-Waterman score: 2670; 91.518% identity (97.321% similar) in 448 aa overlap
(747-1194:1-448)

      720      730      740      750      760      770
frame3 SSWGLKKSGMTLIGSELRLPLKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|227 MSHGLSPRPATAKKSADLKGTVRIPGDKSI
      10         20         30

      780      790      800      810      820      830
frame3 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGGGLLAP
      .....

```



```

gi|227 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGAKIRKEGDTWIINGVNGALLAP
      40      50      60      70      80      90

      840      850      860      870      880      890
frame3 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEGD
      .....
gi|227 EAPLDFGNAGTGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLDPLREMGVQVKSAGD
      100      110      120      130      140      150

      900      910      920      930      940      950
frame3 RLPVTLRGPKTPITITYRVPMSAQVKSALLAGLNTPGITTVIEPIMTRDHTKMLQGF
      .....
gi|227 RLPVTLRGPKTPNPITYRVPMSAQVKSALLAGLNTPGITTVIEPVMTRDHTKMLQGF
      160      170      180      190      200      210

      960      970      980      990      1000      1010
frame3 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM
      .....
gi|227 GANLSVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM
      220      230      240      250      260      270

      1020      1030      1040      1050      1060      1070
frame3 NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
      .....
gi|227 NPTRTGLILTLQEMGANIEVMNPRLAGGEDVADLRVRYSELKGVTVPEERAPSMIDEYPV
      280      290      300      310      320      330

      1080      1090      1100      1110      1120      1130
frame3 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGKLNGVDCDEGETSLVVRGRPDGKGL
      .....
gi|227 LAVAAAFAGATVMNGLEELRVKESDRLSAVADGKLNGVDCDEGEASLVVRGRPGGKGL
      340      350      360      370      380      390

      1140      1150      1160      1170      1180      1190
frame3 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
      .....
gi|227 GNAAGGQVKTHLDHRIAMSFLVLGLASEHPVTVDATMIATSFPEFMDLMTGLGATIE
      400      410      420      430      440

      1200      1210      1220      1230      1240      1250
frame3 DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEFYGI

>>gi|150030157|gb|ABR62274.1| 3-phosphoshikimate 1-carbo (456 aa)
      initn: 2660 initl: 2660 opt: 2660 Z-score: 3124.3 bits: 589.7 E(): 5.8e-165
      Smith-Waterman score: 2660; 89.890% identity (96.264% similar) in 455 aa overlap
      (747-1201:1-455)

      720      730      740      750      760      770
frame3 SSWGLKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
      .....
gi|150 SSWGLKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
      .....
      MSHGNSNRPATARKSSDLKGTVRIPGDKSI
      10      20      30

      780      790      800      810      820      830
frame3 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGGGLLAP
      .....
gi|150 SHRSFMFGGLASGETRITGLLEGEDVINTGRAMQAMGAKIRKEGDTWIIDGVNGGALLAP
      40      50      60      70      80      90

      840      850      860      870      880      890
frame3 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEGD
      .....

```

```

gi|150 EAPLDFGNAGTGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLDPLREMGVQVKSAGD
      100      110      120      130      140      150

      900      910      920      930      940      950
frame3 RLPVTLRGPKTPITITYRVPMSAQVKSALLAGLNTPGITTVIEPIMTRDHTKMLQGF
      .....
gi|150 RLPVTLRGPKTPNPITYRVPMSAQVKSALLAGLNTPGITTVIEPVMTRDHTKMLQGF
      160      170      180      190      200      210

      960      970      980      990      1000      1010
frame3 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM
      .....
gi|150 GADLSVETDRDGVRTIRLEGRGKLRGQVIDVPGDPSSTAFPLVAALLVPGSDLSIFNVLM
      220      230      240      250      260      270

      1020      1030      1040      1050      1060      1070
frame3 NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
      .....
gi|150 NPTRTGLILTLQEMGARIEVLSSRLAGGEDVADLRVRYSELKGVTVPEERAPSMIDEYPV
      280      290      300      310      320      330

      1080      1090      1100      1110      1120      1130
frame3 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGKLNGVDCDEGETSLVVRGRPDGKGL
      .....
gi|150 LAVAAAFAGATVMNGLEELRVKESDRLSAVAEGLKNGVDCDEGEASLVVRGRPGGKGL
      340      350      360      370      380      390

      1140      1150      1160      1170      1180      1190
frame3 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
      .....
gi|150 GNDSGGQVKTHLDHRIAMSFLVMGLASERPVTVDATMIATSFPEFMGLMTGLGAKIET
      400      410      420      430      440      450

      1200      1210      1220      1230      1240      1250
frame3 DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEFYGI
      .....
gi|150 ENKAAL

>>gi|115254522|emb|CAK05596.1| putative 3-phosphoshikima (452 aa)
      initn: 2563 initl: 2563 opt: 2563 Z-score: 3010.2 bits: 568.6 E(): 1.3e-158
      Smith-Waterman score: 2563; 86.637% identity (96.214% similar) in 449 aa overlap
      (747-1195:1-449)

      720      730      740      750      760      770
frame3 SSWGLKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
      .....
gi|115 SSWGLKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
      .....
      MLNGSASKPATARKSAGLTGSRIPGDKSI
      10      20      30

      780      790      800      810      820      830
frame3 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGGGLLAP
      .....
gi|115 SHRSFMIGGLASGETRITGLLEGEDVINTGRAMQAMGARIRKEGAQWVIEGTGNGALLAP
      40      50      60      70      80      90

      840      850      860      870      880      890
frame3 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEGD
      .....
gi|115 DAPLDFGNAGTGVRLTMGLVGYDFHSTFIGDASLSKRPMPGRVLNPLREMGVQVSASEGD
      100      110      120      130      140      150

```

```
>>gi|240860935|gb|ACS58602.1| 3-phosphoshikimate 1-carbo (452 aa)
  initn: 2559 initl: 2559 opt: 2559 Z-score: 3005.5 bits: 567.8 E(): 2.4e-158
Smith-Waterman score: 2559; 86.801% identity (95.973% similar) in 447 aa overlap
(747-1193:1-447)
```

```
>>gi|2485226|gb|AAB73362.1|I44451 Sequence 7 from patent (449 aa)
  initn: 2170 initl: 2139 opt: 2442 Z-score: 2867.9 bits: 542.3 E(): 1.1e-150
Smith-Waterman score: 2442; 82.889% identity (93.556% similar) in 450 aa overlap
(747-1196:1-447)
```



```
>>gi|144974728|gb|ABP12439.1| Sequence 7 from patent US (449 aa)
  initn: 2170 initl: 2139 opt: 2442 Z-score: 2867.9 bits: 542.3 E(): 1.1e-150
Smith-Waterman score: 2442; 82.889% identity (93.556% similar) in 450 aa overlap
(747-1196:1-447)
```

```
>>gi|5957549|gb|AAE08223.1| Sequence 7 from patent US 58 (449 aa)
  initn: 2170 initl: 2139 opt: 2442 Z-score: 2867.9 bits: 542.3 E(): 1.1e-150
  Smith-Waterman score: 2442; 82.889% identity (93.556% similar) in 450 aa overlap
  (747-1196:1-447)
```

Monsanto Company

Regulatory Product Characterization Center

Page 109 of 190

Page 109 of 190

```

      1140      1150      1160      1170      1180      1190
frame3 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
:   .....
gi|595 G---GGTVATHLDHRIAMSFLVMGLAAEKPVTVDDSNMIATSFPEFMDMMPGLGAKIELS
      400      410      420      430      440

      1200      1210      1220      1230      1240      1250
frame3 DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEFYGYI

gi|595 IL

>>gi|2484149|gb|AAB72285.1|I49178 Sequence 7 from patent (449 aa)
  initn: 2170 initl: 2139 opt: 2442 Z-score: 2867.9 bits: 542.3 E(): 1.1e-150
Smith-Waterman score: 2442; 82.889% identity (93.556% similar) in 450 aa overlap
(747-1196:1-447)

      720      730      740      750      760      770
frame3 SSWGLKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
:   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|248      MSHSASPKPATARRSEALTGEIRIPGDKSI
      10      20      30

      780      790      800      810      820      830
frame3 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGGGLLAP
:   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|248 SHRSFMFGGLASGETRITGLLEGEDVINTGRAMQAMGAKIRKEGDVWIINGVNGCCLLQP
      40      50      60      70      80      90

      840      850      860      870      880      890
frame3 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEGD
:   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|248 EAALDFGNAGTGARLTMGLVGTYDMKTSFIGDASLSKRPMGRVLNPLREMGVQVEAADGD
      100      110      120      130      140      150

      900      910      920      930      940      950
frame3 RLPVTLRGPKTPTPITYRVPMSAQVKSAVLLAGLNTPGITTIVIEPIMTRDHTEKMLQGF
:   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|248 RMPLTLIGPKTANPITYRVPMSAQVKSAVLLAGLNTPGVTTIVIEPVMTRDHTEKMLQGF
      160      170      180      190      200      210

      960      970      980      990      1000      1010
frame3 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM
:   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|248 GADLTVETDKDGVRHIRITGQGKLVGQTIDVPGDPSSTAFPLVAALLVEGSDVTIRNVLM
      220      230      240      250      260      270

      1020      1030      1040      1050      1060      1070
frame3 NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
:   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|248 NPTRTGLILTLQEMGADIEVLNARLAGGEDVADLRVRASKLKGVVVPPERAPSMIDEYPV
      280      290      300      310      320      330

      1080      1090      1100      1110      1120      1130
frame3 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGLKLNVDCEGETSLVVRGRPDGKGL
:   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|248 LAIAAFAEGETVMDGLDELRVKESDRLAAVARGLEANGVDCTEGEMSLTVRGRPDGKGL
      340      350      360      370      380      390

      1140      1150      1160      1170      1180      1190
frame3 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
:   .....

```

```

gi|248 G---GGTVATHLDHRIAMSFLVMGLAAEKPVTVDDSNMIATSFPEFMDMMPGLGAKIELS
      400      410      420      430      440

      1200      1210      1220      1230      1240      1250
frame3 DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEFYGYI

gi|248 IL

>>gi|151559264|gb|ABS12762.1| 3-phosphoshikimate 1-carbo (450 aa)
  initn: 2418 initl: 2112 opt: 2441 Z-score: 2866.7 bits: 542.1 E(): 1.3e-150
Smith-Waterman score: 2441; 82.119% identity (93.157% similar) in 453 aa overlap
(747-1199:1-450)

      720      730      740      750      760      770
frame3 SSWGLKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
:   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|151      MSHSAPPKPATARRSEALTGEIRIPGDKSI
      10      20      30

      780      790      800      810      820      830
frame3 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGGGLLAP
:   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|151 SHRSFMFGGLASGETRITGLLEGEDVINTGRAMQAMGAKIRKDGDAWIINGVNGCCLLQP
      40      50      60      70      80      90

      840      850      860      870      880      890
frame3 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEGD
:   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|151 EAALDFGNAGTGARLTMGLVGTYDMRTSFIGDASLSKRPMGRVLNPLREMGVQVEAAEGD
      100      110      120      130      140      150

      900      910      920      930      940      950
frame3 RLPVTLRGPKTPTPITYRVPMSAQVKSAVLLAGLNTPGITTIVIEPIMTRDHTEKMLQGF
:   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|151 RMPLTLIGPKTANPITYRVPMSAQVKSAVLLAGLNTPGVTTIVIEPVMTRDHTEKMLQGF
      160      170      180      190      200      210

      960      970      980      990      1000      1010
frame3 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM
:   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|151 GADLSVETDKDGVRHIRITGQGKLVGQTIDVPGDPSSTAFPLVAALLVEGSDVTIRNVLM
      220      230      240      250      260      270

      1020      1030      1040      1050      1060      1070
frame3 NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
:   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|151 NPTRTGLILTLQEMGADIEVLNARLAGGEDVADLRVKASKLKGVVVPPERAPSMIDEYPV
      280      290      300      310      320      330

      1080      1090      1100      1110      1120      1130
frame3 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGLKLNVDCEGETSLVVRGRPDGKGL
:   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|151 LAIAAFAEGETVMDGLDELRVKESDRLAAVARGLEANGVDCTEGEMSLTVRGRPDGKGL
      340      350      360      370      380      390

      1140      1150      1160      1170      1180      1190
frame3 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
:   .....
gi|151 G---GGTVATHLDHRIAMSFLVMGLASEKPVTVDDSTMIATSFPEFMDMMPGLGAKIELS
      400      410      420      430      440

```

```

1200      1210      1220      1230      1240      1250
frame3 DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEFYGYI
...
gi|151 DAR
450

>>gi|221721755|gb|ACM24911.1| 3-phosphoshikimate 1-carbo (454 aa)
  initn: 2411 initl: 2169 opt: 2427 Z-score: 2850.2 bits: 539.0 E(): 1.1e-149
Smith-Waterman score: 2427; 82.628% identity (93.987% similar) in 449 aa overlap
(747-1195:1-448)

      720      730      740      750      760      770
frame3 SSWGLKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
      : : . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|221      MPHDTHSRPATARRSTNLSGTIIRVPGDKSI
      10      20      30

      780      790      800      810      820      830
frame3 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGGGLLAP
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|221 SHRAFMFGLASGETRITGLLEGEDVLTGKAMRAMGAKIEKAGDEWIIGVNGGALLAP
      40      50      60      70      80      90

      840      850      860      870      880      890
frame3 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEGD
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|221 EAPLDFGNAGTGSRLTMGLVGVIHFESTFIGDASLSKRPMGRILDPLRQMGVQVKSSEGD
      100     110     120     130     140     150

      900      910      920      930      940      950
frame3 RLPVTLRGPKTPTPTITYRVPMSAQVKSAVLLAGLNTPGITTVIEPIMTRDHTKMLQGF
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|221 RLPVTLHGPEANPTITYRVPMSAQVKSAVLLAGLNAPGITSVIEPVMTRDHTKMLAGF
      160     170     180     190     200     210

      960      970      980      990      1000     1010
frame3 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|221 GAELSVETDRDGVRIIRLQGQKLFQGTIDVPGDPSSTAFPLVAALLVAGSDITIRNVLM
      220     230     240     250     260     270

      1020     1030     1040     1050     1060     1070
frame3 NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSSTLKGVTVPEDRAPSMIDEYPI
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|221 NPTRTGLILTLQEMGADIEILDKRLAGGEDVADLRVRSSALKGVTVPAERAPSMIDEYPV
      280     290     300     310     320     330

      1080     1090     1100     1110     1120     1130
frame3 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGLKLNVDCEGETSLVVRGRPDGKGL
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|221 LAVAAAFAGETTVMLGLEELRVKESDRLSAVAEGTLNVDCTEGEASLTVRGRPDGKGL
      340     350     360     370     380     390

      1140     1150     1160     1170     1180     1190
frame3 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|221 GNAGGE-VKTHLDHRIAMSFLVMGLASEHPVKVDDSAMIAATSFPEFFDLMAELGADMVDE
      400     410     420     430     440

      1200     1210     1220     1230     1240     1250
frame3 DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEFYGYI

```

```

gi|221 GSISP
450

>>gi|264661798|gb|EEZ32059.1| 3-phosphoshikimate 1-carbo (450 aa)
  initn: 2401 initl: 2124 opt: 2424 Z-score: 2846.7 bits: 538.4 E(): 1.7e-149
Smith-Waterman score: 2424; 82.444% identity (93.111% similar) in 450 aa overlap
(747-1196:1-447)

      720      730      740      750      760      770
frame3 SSWGLKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|264      MSHSACPKPATARHSQALTGEIRIPGDKSI
      10      20      30

      780      790      800      810      820      830
frame3 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGGGLLAP
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|264 SHRSFMFGGLASGETRITGLLEGEDVINTGRAMQAMGARIRKEGDVWIINGVNGGCLLP
      40      50      60      70      80      90

      840      850      860      870      880      890
frame3 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEGD
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|264 EAPLDFGNAGTGARLTMGLVGTYDMKTSFIGDASLSKRPMGRVLNPLREMGVQVEAAEGD
      100     110     120     130     140     150

      900      910      920      930      940      950
frame3 RLPVTLRGPKTPTPTITYRVPMSAQVKSAVLLAGLNTPGITTVIEPIMTRDHTKMLQGF
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|264 RMPLTLIGPRTANPTITYRVPMSAQVKSAVLLAGLNTPGVTTVIEPVMTRDHTKMLQGF
      160     170     180     190     200     210

      960      970      980      990      1000     1010
frame3 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|264 GADLTVETDKDGVRIHIVGQKLTGQITDVPGDPSSTAFPLVAALLVEGSDVTIRNVLM
      220     230     240     250     260     270

      1020     1030     1040     1050     1060     1070
frame3 NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSSTLKGVTVPEDRAPSMIDEYPI
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|264 NPTRTGLILTLQEMGADIEIIDPRLAGGEDVADLRVRASKLKGVVVPPERAPSMIDEYPV
      280     290     300     310     320     330

      1080     1090     1100     1110     1120     1130
frame3 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGLKLNVDCEGETSLVVRGRPDGKGL
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|264 LAIAAPFAEGETVMDGLDELRVKESDRLLAAVARGLEANGVDCTEGEMSLTVRGRPGGKGL
      340     350     360     370     380     390

      1140     1150     1160     1170     1180     1190
frame3 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|264 G---GGTVATHLDHRIAMSFLVMGLASEKPVTVDDSTMIATSFPEFMGMAGLGAKIAES
      400     410     420     430     440

      1200     1210     1220     1230     1240     1250
frame3 DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEFYGYI

gi|264 GAE
450

```







```

      780      790      800      810      820      830
frame3 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGGLLAP
      .....
gi|261 SHRSFMFGGLASGKTRITGLLEGEDVINTGRAMQAMGARIRKEGDVWIINGVNGCLLP
      40      50      60      70      80      90

```

Regulatory Product Characterization Center  
RAR-10-070

Page 115 of 190

```
>>gi|261300126|gb|EEY03623.1| 3-phosphoshikimate 1-carbo (450 aa)
  initn: 2396 initl: 2119 opt: 2419 Z-score: 2840.9 bits: 537.3 E(): 3.5e-149
Smith-Waterman score: 2419; 82.000% identity (93.333% similar) in 450 aa overlap
(747-1196:1-447)
```

```
>>gi|260919839|gb|EEH86492.1| 3-phosphoshikimate 1-carbo (450 aa)
  initn: 2396 initl: 2119 opt: 2419 Z-score: 2840.9 bits: 537.3 E(): 3.5e-149
Smith-Waterman score: 2419; 82.000% identity (93.333% similar) in 450 aa overlap
(747-1196:1-447)
```



```

      1020      1030      1040      1050      1060      1070
frame3 NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
      .....
gi|261 NPTRTGLILTLQEMGADIEIIDPRLAGGEDVADLRVKASKLKGVVVPPERAPSMIDEYPV
      280      290      300      310      320      330

      1080      1090      1100      1110      1120      1130
frame3 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGLKLNQVDCDEGETSLVVRGRPDGKGL
      .....
gi|261 LAIAASFAGETVMDGLDELRVKESDRLAAVARGLEANGVDCTEGEMSLTVRGRPGGKGL
      340      350      360      370      380      390

      1140      1150      1160      1170      1180      1190
frame3 GNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDATMIATSFPEFMDLMAGLGAKIELS
      : .....
gi|261 G---GGTVATHLDHRIAMSFLVMGLASEKPVTVDDSTMIATSFPEFMGMAGLGAKIAES
      400      410      420      430      440

      1200      1210      1220      1230      1240      1250
frame3 DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEFYGI

gi|261 GAE
      450

>>gi|262764953|gb|EEZ10874.1| 3-phosphoshikimate 1-carbo (450 aa)
      initn: 2396 initl: 2119 opt: 2419 Z-score: 2840.9 bits: 537.3 E(): 3.5e-149
Smith-Waterman score: 2419; 82.000% identity (93.333% similar) in 450 aa overlap
(747-1196:1-447)

      720      730      740      750      760      770
frame3 SSWGLKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
      : .. .....
gi|262 MSHSACPKPATARHSQALTGEIRIPGDKSI
      10      20      30

      780      790      800      810      820      830
frame3 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGGGLLAP
      .....
gi|262 SHRSFMFGGLASGKTRITGLLEGEDVINTGRAMQAMGARIRKEGDVWIINGVNGCQLLP
      40      50      60      70      80      90

      840      850      860      870      880      890
frame3 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEGD
      .....
gi|262 EAPLDFGNAGTGARLTMGLVGTYDMKTSFIGDASLSKRPMGRVLNPLREMGVQVEAAEGD
      100      110      120      130      140      150

      900      910      920      930      940      950
frame3 RLPVTLRGPKTPTPITYRVPMSAQVKSAYLLAGLNTPGITTVIEPIMTRDHTKMLQGF
      .....
gi|262 RMPLTLIGPRTANPIAYRVPMSAQVKSAYLLAGLNTPGVTTVIEPVMTRDHTKMLQGF
      160      170      180      190      200      210

      960      970      980      990      1000      1010
frame3 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM
      .....
gi|262 GADLTVETDKDGVRHIRIVGQGKLTGQTIDVPGDPSSTAFPLVAALLVEGSDVTIRNVLM
      220      230      240      250      260      270

      1020      1030      1040      1050      1060      1070
frame3 NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
      .....

```

```

gi|262 NPTRTGLILTLQEMGADIEIIDPRLAGGEDVADLRVKASKLKGVVVPPERAPSMIDEYPV
      280      290      300      310      320      330

      1080      1090      1100      1110      1120      1130
frame3 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGLKLNQVDCDEGETSLVVRGRPDGKGL
      .....
gi|262 LAIAASFAGETVMDGLDELRVKESDRLAAVARGLEANGVDCTEGEMSLTVRGRPGGKGL
      340      350      360      370      380      390

      1140      1150      1160      1170      1180      1190
frame3 GNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDATMIATSFPEFMDLMAGLGAKIELS
      : .....
gi|262 G---GGTVATHLDHRIAMSFLVMGLASEKPVTVDDSTMIATSFPEFMGMAGLGAKIAES
      400      410      420      430      440

      1200      1210      1220      1230      1240      1250
frame3 DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEFYGI

gi|262 GAE
      450

>>gi|264659952|gb|EEZ30213.1| 3-phosphoshikimate 1-carbo (450 aa)
      initn: 2396 initl: 2119 opt: 2419 Z-score: 2840.9 bits: 537.3 E(): 3.5e-149
Smith-Waterman score: 2419; 82.000% identity (93.333% similar) in 450 aa overlap
(747-1196:1-447)

      720      730      740      750      760      770
frame3 SSWGLKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
      : .. .....
gi|264 MSHSACPKPATARHSQALTGEIRIPGDKSI
      10      20      30

      780      790      800      810      820      830
frame3 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGGGLLAP
      .....
gi|264 SHRSFMFGGLASGKTRITGLLEGEDVINTGRAMQAMGARIRKEGDVWIINGVNGCQLLP
      40      50      60      70      80      90

      840      850      860      870      880      890
frame3 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSEGD
      .....
gi|264 EAPLDFGNAGTGARLTMGLVGTYDMKTSFIGDASLSKRPMGRVLNPLREMGVQVEAAEGD
      100      110      120      130      140      150

      900      910      920      930      940      950
frame3 RLPVTLRGPKTPTPITYRVPMSAQVKSAYLLAGLNTPGITTVIEPIMTRDHTKMLQGF
      .....
gi|264 RMPLTLIGPRTANPIAYRVPMSAQVKSAYLLAGLNTPGVTTVIEPVMTRDHTKMLQGF
      160      170      180      190      200      210

      960      970      980      990      1000      1010
frame3 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM
      .....
gi|264 GADLTVETDKDGVRHIRIVGQGKLTGQTIDVPGDPSSTAFPLVAALLVEGSDVTIRNVLM
      220      230      240      250      260      270

      1020      1030      1040      1050      1060      1070
frame3 NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
      .....
gi|264 NPTRTGLILTLQEMGADIEIIDPRLAGGEDVADLRVKASKLKGVVVPPERAPSMIDEYPV
      280      290      300      310      320      330

```

Monsanto Company

Regulatory Product Characterization Center

Page 119 of 190

```

      1140      1150      1160      1170      1180      1190
frame3 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
:   .....
gi|225 G---GGTVATHLDHRIAMSFLVMGLASEKPVTVDDSTMIATSFPEFMGMAGLGAKIAES
      400      410      420      430      440

      1200      1210      1220      1230      1240      1250
frame3 DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEFYGYI

gi|225 GAE
      450

>>gi|255998648|gb|ACU47047.1| 3-phosphoshikimate 1-carbo (450 aa)
      initn: 2395 initl: 2118 opt: 2418 Z-score: 2839.7 bits: 537.1 E(): 4.1e-149
Smith-Waterman score: 2418; 81.778% identity (93.333% similar) in 450 aa overlap
(747-1196:1-447)

      720      730      740      750      760      770
frame3 SSWGLKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
:   .. :..... :.....
gi|255 MSHSACPKPATARHSQALTGEIRIPGDKSI
      10      20      30

      780      790      800      810      820      830
frame3 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGGGLLAP
:   .....
gi|255 SHRSFMFGGLASGKTRITGLLEGEDVINTGRAMQAMGARIRKEGDVWIINGVNGCLLQP
      40      50      60      70      80      90

      840      850      860      870      880      890
frame3 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSSEDGD
:   .....
gi|255 EAPLDFGNAGTGARLTMGLVGTYDMKTSFIGDASLSKRPMGRVLNPLREMGVQVEAAEGD
      100      110      120      130      140      150

      900      910      920      930      940      950
frame3 RLPVTLRGPKTPTPITYRVPMASAVKSAVLLAGLNTPGIITVIEPIMTRDHTEKMLQGF
:   .....
gi|255 RMPLTLIGPRTANPIAYRVPMASAVKSAVLLAGLNTPGVTTVIEPVMTRDHTEKMLQGF
      160      170      180      190      200      210

      960      970      980      990      1000      1010
frame3 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM
:   .....
gi|255 GADLTVETDKDGVRRHIVGQKLTGQTIDVPGDPSSTAFPLVAALLVEGSDVTIRNVLM
      220      230      240      250      260      270

      1020      1030      1040      1050      1060      1070
frame3 NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
:   .....
gi|255 NPTRTGLILTLQEMGADIEIIDPRLAGGEDVADLRVKASKLKGVVPPERAPSMIDEYPV
      280      290      300      310      320      330

      1080      1090      1100      1110      1120      1130
frame3 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGLKLNVDCEGETSLVVRGRPDGKGL
:   .....
gi|255 LAIAASFAEGETVMDGLDELRVKESDRLAAVARGLEANGVDCTEGEMSLTVRGRPGGKGL
      340      350      360      370      380      390

      1140      1150      1160      1170      1180      1190
frame3 GNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELS
:   .....

```

```

gi|255 G---GGTVATHLDHRIAMSFLVMGLASEKPVTVDDSTMIATSFPEFMGMAGLGAKIAES
      400      410      420      430      440

      1200      1210      1220      1230      1240      1250
frame3 DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEFYGYI

gi|255 GAE
      450

>>gi|260096414|gb|EEW80290.1| 5-enolpyruvyl shikimate 3- (461 aa)
      initn: 2390 initl: 2113 opt: 2417 Z-score: 2838.4 bits: 536.9 E(): 4.9e-149
Smith-Waterman score: 2417; 80.477% identity (92.191% similar) in 461 aa overlap
(739-1196:1-458)

      710      720      730      740      750      760
frame3 HPRAYPISSSWGLKKSGMTLIGSELRLPKVMSSVSTAC--MLHGASSRPATARKSSGLS
:   . . : : : : :
gi|260 MSAPKCDCEKSMSSHACPKPATARHSQALT
      10      20      30

      770      780      790      800      810      820
frame3 GTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWII
:   .....
gi|260 GEIRIPGDKSISHRSFMFGGLASGKTRITGLLEGEDVINTGRAMQAMGARIRKEGDVWII
      40      50      60      70      80      90

      830      840      850      860      870      880
frame3 DGVNGGGLLAPEAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLRE
:   .....
gi|260 NGVNGCLLQPEAPLDFGNAGTGARLTMGLVGTYDMKTSFIGDASLSKRPMGRVLNPLRE
      100      110      120      130      140      150

      890      900      910      920      930      940
frame3 MGTVQKSEGDGRLPVTLRGPKTPTPITYRVPMASAVKSAVLLAGLNTPGIITVIEPIMT
:   .....
gi|260 MGTVQVEAAEGDRMPLTLIGPRTANPIAYRVPMASAVKSAVLLAGLNTPGVTTVIEPVM
      160      170      180      190      200      210

      950      960      970      980      990      1000
frame3 RDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVP
:   .....
gi|260 RDHTEKMLQGFGADLTVETDKDGVRRHIVGQKLTGQTIDVPGDPSSTAFPLVAALLVE
      220      230      240      250      260      270

      1010      1020      1030      1040      1050      1060
frame3 GSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPED
:   .....
gi|260 GSEVTIRNVLMNPTRTGLILTLQEMGADIEIIDPRLAGGEDVADLRVKASKLKGVVVPPE
      280      290      300      310      320      330

      1070      1080      1090      1100      1110      1120
frame3 RAPSMIDEYPI LAVAAAFAGATVMNGLEELRVKESDRLSAVANGLKLNVDCEGETSL
:   .....
gi|260 RAPSMIDEYPI LAIAASFAEGETVMDGLDELRVKESDRLAAVARGLEANGVDCTEGEMSL
      340      350      360      370      380      390

      1130      1140      1150      1160      1170      1180
frame3 VVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPTVDDATMIATSFPEFMDL
:   .....
gi|260 TVRGRPGGKGLG---GGTVATHLDHRIAMSFLVMGLASEKPVTVDDSTMIATSFPEFMGM
      400      410      420      430      440

```



```

1190      1200      1210      1220      1230      1240
frame3  MAGLGAKIELSDTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTH
:         :         :         :         :         :
gi|260  MAGLGAKIAESGAE
      450      460

>>gi|263003296|gb|EEZ15589.1| 3-phosphoshikimate 1-carbo (450 aa)
initn: 2385 initl: 2118 opt: 2414 Z-score: 2835.0 bits: 536.2 E(): 7.5e-149
Smith-Waterman score: 2414; 81.778% identity (93.333% similar) in 450 aa overlap
(747-1196:1-447)

      720      730      740      750      760      770
frame3  SSWG LKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
:         :         :         :         :         :
gi|263  MSHSACPKPATARHSQALTGEIRIPGDKSI
      10      20      30

      780      790      800      810      820      830
frame3  SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGGGLLAP
:         :         :         :         :         :
gi|263  SHRSFMFGGLASGKTRITGLLEGEDVINTGRAMQAMGARIRKEGDVWIINGVNGGLLIQP
      40      50      60      70      80      90

      840      850      860      870      880      890
frame3  EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSSEDGD
:         :         :         :         :         :
gi|263  EAPLDFGNAGTGARLTMGLVGTYDMKTSFIGDASLSKRPMGRVLNPLREMGVQVEAAEGD
      100     110     120     130     140     150

      900      910      920      930      940      950
frame3  RLPVTLRGPKTPTPIYRVPMASAVKSAVLLAGLNTPGITTVIEPIMTRDHTEKMLQGF
:         :         :         :         :         :
gi|263  RMPLTLIGPRTANPIAYRVPMASAVKSAVLLAGLNTPGVTTVIEPVMTRDHTEKMLQGF
      160     170     180     190     200     210

      960      970      980      990      1000     1010
frame3  GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM
:         :         :         :         :         :
gi|263  GADLTVETDKDQVGRHIRIVGQGKLTGQTIDVPGDPSSTAFPLVAALLVEGSDVTIRNVLM
      220     230     240     250     260     270

      1020     1030     1040     1050     1060     1070
frame3  NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLVRVRSSTLKGTVVPEDRAPSMIDEYPI
:         :         :         :         :         :
gi|263  NPTRTGLILTLQEMGADIEIIDPRLAGGEDVADLRVKASKLKGVVPPERAPSMIDEYYPV
      280     290     300     310     320     330

      1080     1090     1100     1110     1120     1130
frame3  LAVAAAFAGGATVMNGLEELRVKESDRLSAVANGLKLVGDCDEGETSLVVRGRPDGKGL
:         :         :         :         :         :
gi|263  LAIAASFAGETVMDGLDELRVKESDRLAAVARGLEANGVDCTEGEMSLTVRGRPGGKGL
      340     350     360     370     380     390

      1140     1150     1160     1170     1180     1190
frame3  GNASGAAVATHLDHRIAMSFLVMGLVSENPVVDATMIATSFPEFMDLMAGLGAKIELS
:         :         :         :         :         :
gi|263  G---GGTVGTHLDHRIAMSFLVMGLASEKPVTVDDSTMIATSFPEFGMMAGLGAKIAES
      400     410     420     430     440

      1200     1210     1220     1230     1240     1250
frame3  DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHONPTEFEYYGI

```

```
>>gi|260153122|gb|EEW88214.1| 3-phosphoshikimate 1-carbo (450 aa)
initn: 2385 initl: 2118 opt: 2414 Z-score: 2835.0 bits: 536.2 E(): 7.5e-149
Smith-Waterman score: 2414; 81.778% identity (93.333% similar) in 450 aa overlap
(747-1196:1-447)
```

	720	730	740	750	760	770
frame3	SSWGLKKSGMTLIGSELRLPKVMSSSVSTACMLHGASSRPARATARKSSGSLSGTVRI	PDKSI				
gi 260				:	:	:
				MHSACPKPATARHQSALTGEIRIPGD	KSI	
				10	20	30
	780	790	800	810	820	830
frame3	SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVNGNGLLP					
gi 260	SHRSFMFGGLASGKTRITGLLEGEDVINTGRAMQAMGARIRKEGDTVIIINGVNGCLLP					
	40	50	60	70	80	90
	840	850	860	870	880	890
frame3	EAPLDFGNAATGCRLTMGLVGVDYFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSE	DG				
gi 260	EAPLDFGNAGTGARLTMGLVGTYDMKTSFIGDASLSKRPMGRVLNPLREMGVQVEA	AEGD				
	100	110	120	130	140	150
	900	910	920	930	940	950
frame3	RLPVTLRGPKTPTPIITYRVPMASAVQKSALLAGLNTPGITTVIEPIMTRDHTEKMLQGF					
gi 260	RMPLTLIGPRNTANPIAYRVPMASAVQKSALLAGLNTPGVTTVIEPVMTRDHTEKMLQGF					
	160	170	180	190	200	210
	960	970	980	990	1000	1010
frame3	GANTLVETDADGVRTIRLEGRGKLTKGVIDVPDGDPSSAFPLVAALLVPGSDVTILNVLM					
gi 260	GADTLVETDDKGVRHIRIVGGKLTGQTIDVPDGDPSSAFPLVAALLVEGSDVTIRNVLM					
	220	230	240	250	260	270
	1020	1030	1040	1050	1060	1070
frame3	NPTRTGLIILTQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTPEDRAPSMIDEYP	I				
gi 260	NPTRTGLIILTQEMGADIEIIDPRLAGGEDVADLRVKASKLGKVVPPERAPSMIDEY	PV				
	280	290	300	310	320	330
	1080	1090	1100	1110	1120	1130
frame3	LAVAAFAEAGATVMNGLEELRVKESDRLSAVANGKLNGVDCDEGETSLVVRRPDGKGL					
gi 260	LAIAASFAGETVMDDLDELRVKESDRLLAAVARGLEANGVDCDEGMTSLTVRRPDGKGL					
	340	350	360	370	380	390
	1140	1150	1160	1170	1180	1190
frame3	GNASGAATHLDHRIAMSFLVMGLVSENPNVTVDATMIATSFPEFMDLMAGLGAKIELS					
gi 260	G---GGTVGTHLDHRIAMSFLVMGLASEKPNVTVDSTMIATSFPEFMGMAGLGAKIAES					
	400	410	420	430	440	
	1200	1210	1220	1230	1240	1250
frame3	DIKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPT	EFYYGI				
gi 260	GAE					
	450					

>>gi|17983963|gb|AAL53098.1| 3-phosphoshikimate 1-carbox (480 aa)  
initn: 2385 initl: 2118 opt: 2414 Z-score: 2834.6 bits: 536.2 E(): 7.9e-149  
Smith-Waterman score: 2414; 81.778% identity (93.333% similar) in 450 aa overlap  
(747-1196:31-477)

```

      720      730      740      750      760      770
frame3 SSWGLKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
gi|179 MTTQYYYARKTALISQSRVVSAPKCDCEKSMHSACPKPATARHSQALTGEIRIPGDKSI
      10      20      30      40      50      60

      780      790      800      810      820      830
frame3 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVGNGGLLAP
gi|179 SHRSFMFGGLASGKTRITGLLEGEDVINTGRAMQAMGARIRKEGDVWIINGVNGCCLLP
      70      80      90      100     110     120

      840      850      860      870      880      890
frame3 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSSEDGD
gi|179 EAPLDFGNAGTGARLTMGLVGTIDMKTSTFIGDASLSKRPGRVLNPLREMGVQVEAAEGD
      130     140     150     160     170     180

      900      910      920      930      940      950
frame3 RLPVTLRGPKTPTPITYRVPMSAQVKSALLAGLNTPGITTVIEPIMTRDHTKMLQGF
gi|179 RMPPLTLIGPRTANPIAYRVPMSAQVKSALLAGLNTPGVTTVIEPVMTRDHTKMLQGF
      190     200     210     220     230     240

      960      970      980      990     1000     1010
frame3 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM
gi|179 GADLTVETDKDGVRRHIVGQKLTGQVIDVPGDPSSTAFPLVAALLVEGSDVTIRNVLM
      250     260     270     280     290     300

     1020     1030     1040     1050     1060     1070
frame3 NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
gi|179 NPTRTGLILTLQEMGADIEIDPRLAGGEDVADLRVKASKLKGVVPPPERAPSMIDEYPV
      310     320     330     340     350     360

     1080     1090     1100     1110     1120     1130
frame3 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGLKLNVDCEGETSLVVRGRPDGKGL
gi|179 LAIAASFAEGETVMDGLDELRVKESDRLAAVARGLEANGVDCTEGEMSLTVRGRPGGKGL
      370     380     390     400     410     420

     1140     1150     1160     1170     1180     1190
frame3 GNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDATMIATSFPEFMDLMAGLGAKIELS
gi|179 G---GGTVGTHLDHRIAMSFLVMGLASEKPVTVDDSTMIATSFPEFMGMAGLGAKIAES
      430     440     450     460     470

     1200     1210     1220     1230     1240     1250
frame3 DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEFYYGI
gi|179 GAE
      480
```

>>gi|260675989|gb|EEX62810.1| 5-enolpyruvylshikimate-3-p (450 aa)  
initn: 2390 initl: 2113 opt: 2413 Z-score: 2833.8 bits: 536.0 E(): 8.7e-149

Smith-Waterman score: 2413; 81.778% identity (93.333% similar) in 450 aa overlap  
(747-1196:1-447)

```

      720      730      740      750      760      770
frame3 SSWGLKKSGMTLIGSELRLPKVMSSVSTACMLHGASSRPATARKSSGLSGTVRIPGDKSI
gi|260 MSHSACPKPATARHSQALTGEIRIPGDKSI
      10      20      30

      780      790      800      810      820      830
frame3 SHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVGNGGLLAP
gi|260 SHRSFMFGGLASGKTRITGLLEGEDVINTGRAMQAMGARIRKEGDVWIINGVNGCCLLP
      40      50      60      70      80      90

      840      850      860      870      880      890
frame3 EAPLDFGNAATGCRLTMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPLREMGVQVKSSEDGD
gi|260 EAPLDFGNAGTGARLTMGLVGTIDMKTSTFIGDASLSKRPGRVLNPLREMGVQVEAAEGD
      100     110     120     130     140     150

      900      910      920      930      940      950
frame3 RLPVTLRGPKTPTPITYRVPMSAQVKSALLAGLNTPGITTVIEPIMTRDHTKMLQGF
gi|260 RMPPLTLIGPRTANPIAYRVPMSAQVKSALLAGLNTPGVTTVIEPVMTRDHTKMLQGF
      160     170     180     190     200     210

      960      970      980      990     1000     1010
frame3 GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDVTILNVLM
gi|260 GADLTVETDKDGVRRHIVGQKLTGQVIDVPGDPSSTAFPLVAALLVEGSEVTIRNVLM
      220     230     240     250     260     270

     1020     1030     1040     1050     1060     1070
frame3 NPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPI
gi|260 NPTRTGLILTLQEMGADIEIDPRLAGGEDVADLRVKASKLKGVVPPPERAPSMIDEYPV
      280     290     300     310     320     330

     1080     1090     1100     1110     1120     1130
frame3 LAVAAAFAGATVMNGLEELRVKESDRLSAVANGLKLNVDCEGETSLVVRGRPDGKGL
gi|260 LAIAASFAEGETVMDGLDELRVKESDRLAAVARGLEANGVDCTEGEMSLTVRGRPGGKGL
      340     350     360     370     380     390

     1140     1150     1160     1170     1180     1190
frame3 GNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDATMIATSFPEFMDLMAGLGAKIELS
gi|260 G---GGTVATHLDHRIAMSFLVMGLASEKPVTVDDSTMIATSFPEFMGMAGLGAKIAES
      400     410     420     430     440

     1200     1210     1220     1230     1240     1250
frame3 DTKAA..AQEFELGTGS.DLRIL.LELSFVSSVSTTFVKFNASVSLRTHQNPTFEFYYGI
gi|260 GAE
      450
```

2415 residues in 1 query sequences  
4761287459 residues in 17815538 library sequences

Scomplib [34t26]  
start: Wed Feb 3 22:52:11 2010 done: Wed Feb 3 23:17:08 2010  
Total Scan time: 1283.080 Total Display time: 43.800

Function used was FASTA [version 3.4t26 July 7, 2006]

## Appendix 4. Bioinformatic analysis of polypeptide frame4

```
>frame4
FNCKWLHVREIYMDQQ.V.WSIWRKRKSNYQFFFN SKM.MSAALL.NESTF..NDKLRSVVFIGESNKQIILIRKSLFRRVYI
HVQMGMA.MRNFITWRAAIAIYIWDPRPPV.WISAEFALIMYKFKNKLNRET.IY.PITLI.C.TCVGCKLICMKFGIRSLYL.V
LTTFLNNIITFI.YI.IE.IQPNIFQTK.TINSFKYKKRVSSKGGKIIIMSQRIOYNGKVT.AL..KSPI.DISP.NTRKTRP
K.ANE.NITELLHQ.M.GTLLLRVLGVN.IVPASRDISCGVHHIC.GKETPEIELCIYISFFISRFPEGLGVGT.LVATISALL
SIDKQLAVMPQ.RTITTFRAGYFLFYGHDFKTRRSTGEATLEIHYGGNSCYFIIFPCYFTLAGLWWSRQQTWWWACKPWRTKI
QICVFWWLEGKGASPFEN.WNHSCYI.RKLQA...STFADPCRFPGHEAIFKH..FKPAGLVDKSSGGRRR.IIIMLEAFFRC
NNCGSINFRRGLRLCLQATRRRRRFGF.SSKVCRPTTKFAACSTRVRRG.SNREK...SSSCCHHNEFLRSLLR.NV.S.NHDHK
TKSNQHEML.SSSTAALQPIVYQ.REVLIL.LLARCLHLSLQEI.K.KRIYKCIQFREFLFLNKYDEPHMKYPEMQAQFNLP
LGSSSVPGS...KPGPRLQIFIVIGLVNMHGHLATGWHAV.RVLK.TKMRWRKRIRV.GWVRNNKCVGS.WVELVNVLGCSID
TEVSTCCCAWLLGSSFFFLTRRS.SASG.VMSTVCDVFEV.IL.T.TLCKGYSHCVLSSFPWFLMIQETSQCCGIRDYQD
.L.WRKDK.LISTDCYQIN..AANGGLDQNTIHIVKCNHFIWFQYVPCNNND.LLKN SHY.ICQKTQEN..CKKISVDIKI
F.LINI..IFKKYIRIDIKI.IYYYIFFS.IIRIKCFCRGSSSRVDRATAAGPGHIGPRVGPLRVAAGRLMHVNVNQLASHKMH.
KIFSYSTTNP.V.L.L.SND.NLTRILENYIKESS.MSKTQEDILVFSNICSFSSSENIFVA..ISIMVQLQVYSSY.LRSL
EATI.HIFTQNMKFAYLL.FQIQHTTNKRKNK.Y.FENEQKDHIIH.LFSIHFHFTVR.RKPNKKHSLQAQQMVEKQFSQC
HNTQTQ.DSGVCAMKLMH.T.RTLSPKPIRTKALAGS.DLRIRYRARILELIKQP.CRRVRS.LQDQPSNP.TLGS..RS..
HHQQ.QGFQKRD.P.RGSS.RYGDRGG.RQLLQKRYRDPYRQDDHARREKSHLHRNQHR.A.DRLRQKQDGRFP.HGVLPNRS.
R.HLQRMQLQQRELDTRRS.KEHGLLEQ.HPSK.KNEHASQPRLLHQQDVG.SLRCRHFFPAESR.DQYGLGSSKR.GW.RRN
LEQARQQPMGKQ.RMDHLEHQSLDR.AYHDLQDGWYAHRHQHSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.DQQAQRT
SLERKP.VPCR.LALESVHAR.LEDDHRLQTSAPHPFREVGSTHDPDLA..EKRHQ.KCYRNRKHQQDP.SNGNQLQRYRNR
EEFPQEQGVHRYQHQQ.SKYHLPYGFWHF.LA.PYQC..HLHLPKDR.YEFHR.RDLQT.RTCGR.TCHLEYGRFQKDQRTYEQ
LLDGCLHREACTPWKQKT.P.EDEAQSQLTSSHSSSIPITTKSDKLVDAASSEKPIREICVDWISRDRW.EMGSAHHCRCF.
LAPW.ISGQIQCKKNNQILNQ.KQ.TYITIMKIIFNTKQS..KAAIGVKPISTTKIVVLYHLS.AS.IQSNRQFQEI.SKQS
ENYP.KKGTLDDLCCYSRITQSQFKSTYILNTQINKFTRI.IHQPTINRTN.NLKTRSI RMGK.FHE.IIV.KQ.NVDINN.R.N
LNNRSEFSEERSINKEHRIQQSSK.YEQRSADKEKERSKGTLLTSEISQKSV.RLGSIRTINGISNESGFIYSVESETNVRV
.HKIISFFKLPLCPLAVNTQYLLGWLCFMNV.YYMLQTSFNHTF.LRLFKQMVIHVRLPKISLIVDHQSI.HSSLE.IFDIS
MNADYQEPFISFIQTPNKIMNK.SIY.PQNLMTEMVYNNHMCPIPRATSLKRETSIDIKLDIEFFLTNL.GWEFLLGFGDTN
LNKVIWLVPC.LNPIWNAKS.AQANSFVVVTA FVIRISEWAVRTALFHVILFPTCTRGICFIFAIAKAKMLLVGAPNNSAPFL
FCLVPTSQLLDVDKITFKDAH.L.VFGGCLFLKLTLDHDACAGTVVYFD.RFFIDL.LLANGLETLYPDLVEALD.ELASTRR
PHRS.SFSSKPPFGRECRHVEIKISELE.FVYCFRL.IRRIVICRFIKMYFHFIIITLRTSTFLN.KKIGNYSEFFSILTIILI
ADPCRFPGHE
```

Sliding 8 amino acid window search  
Database searched = AD\_2010  
Query = frame4

Start time: Wed Feb 3 23:17:15 GMT 2010 Finish time: Wed Feb 3 23:17:15 GMT 2010

No 8 amino acid matches exist between frame4 and the AD\_2010 database

# fasta34 frame4.pep /genedata/1/db/AD\_2010 -Q -E 1 -O frame4.pep\_ad.fasta  
FASTA searches a protein or DNA sequence data bank version 3.4t26 July 7, 2006  
Please cite:

W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

frame4, 2417 aa  
vs /genedata/1/db/AD\_2010 library

	opt	E()
< 20	2	0: =

```
22 0 0: one = represents 3 library sequences
24 0 0:
26 0 0:
28 0 0:
30 0 2:*
32 3 8:= *
34 10 21:===== *
36 13 44:===== *
38 59 72:===== *
40 77 100:===== *
42 110 123:===== *
44 125 135:===== *
46 121 138:===== *
48 161 132:===== *
50 128 120:===== *
52 91 106:===== *
54 101 90:===== *
56 103 76:===== *
58 87 62:===== *
60 60 50:===== *
62 43 40:===== *
64 31 32:===== *
66 23 25:===== *
68 17 20:===== *
70 18 16:===== *
72 13 12:===== *
74 6 10:== *
76 4 7:==*
78 5 6:==*
80 4 4:==*
82 6 3:==*
84 4 3:==*
86 4 2:==*
88 5 2:==* inset = represents 1 library sequences
90 1 1:*
92 8 1:*== :*=====
94 5 1:*== :*=====
96 2 1:* :*=
98 2 0:= **==
100 2 0:= **==
102 1 0:= **==
104 6 0:= **=====
106 5 0:= **=====
108 1 0:= **==
110 0 0: *
112 1 0:= **==
114 0 0: *
116 0 0: *
118 0 0: *
>120 3 0:= **==
331323 residues in 1471 sequences
Expectation_n fit: rho(ln(x))= 5.96820.00446; mu= 7.6799 0.228
mean_var=64.733316.400, 0's: 2 Z-trim: 5 B-trim: 124 in 1/42
Lambda= 0.159408
Kolmogorov-Smirnov statistic: 0.0853 (N=29) at 46

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 44, opt: 32, open/ext: -10/-2, width: 16
The best scores are:
gi|73912496|dbj|BAE20328.1| omega-5 gliadin [Triti ( 439) 132 41.1 0.00067
gi|208605346|emb|CAR82266.1| D-type LMW glutenin s ( 272) 101 33.9 0.059
gi|170743|gb|AAB02788.1| HMW glutenin subunit Ax2* ( 815) 106 35.2 0.076
```

```

gi|13183177|gb|AAK15089.1|AF240006_1 7S globulin [ ( 585) 96 32.8 0.27
gi|19338630|gb|AAL86739.1|AF441864_1 48-kDa glycop ( 448) 92 31.9 0.4
gi|14285797|sp|O61379.1|TPM_PANST RecName: Full=Tr ( 274) 86 30.5 0.65
gi|6094504|sp|Q25456.1|TPM_METEN RecName: Full=Tr ( 274) 86 30.5 0.65
gi|125995159|dbj|BAF47263.1| tropomyosin fast isof ( 284) 86 30.5 0.67
gi|125995157|dbj|BAF47262.1| tropomyosin fast isof ( 284) 86 30.5 0.67
gi|125995163|dbj|BAF47265.1| tropomyosin fast isof ( 284) 86 30.5 0.67
gi|148615631|gb|ABQ96644.1| tropomyosin [Tyrophagu ( 284) 86 30.5 0.67
gi|2660868|gb|AAC48288.1| fast tropomyosin isoform ( 284) 86 30.5 0.67
gi|125995169|dbj|BAF47268.1| tropomyosin slow-twit ( 284) 86 30.5 0.67
gi|125995161|dbj|BAF47264.1| tropomyosin fast isof ( 284) 86 30.5 0.67
gi|14285796|sp|O44119.1|TPM_HOMAM RecName: Full=Tr ( 284) 86 30.5 0.67
gi|73532979|gb|AAZ76743.1| Pen a 1 allergen [Farfa ( 284) 86 30.5 0.67
gi|170710|gb|AAA34275.1| alpha-type gliadin precur ( 318) 85 30.3 0.88

```

```

>>gi|73912496|dbj|BAE20328.1| omega-5 gliadin [Triticum (439 aa)
initn: 44 init1: 44 opt: 132 Z-score: 159.4 bits: 41.1 E(): 0.00067
Smith-Waterman score: 134; 17.370% identity (47.395% similar) in 403 aa overlap
(1228-1612:40-421)

```

```

1200 1210 1220 1230 1240 1250
frame4 RGS.DLRIRYRARILELIKQP.CRRVRS.LQDQPSNP.TLGS..RS..HH--QQ.QGF-Q
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|739 LAMAMNIASASRLLSPRGKELHTPQEQFPQQQFPQPQQFPQQQIPQHQIPQQPQQFPQ
      10 20 30 40 50 60

```

```

1260 1270 1280 1290 1300 1310
frame4 KRDP.RGSS.RYGDRGG.RQLLQKRYRDPYRQDDHARREKSHLHRNQHR.A.DRLRQQKD
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|739 QQQFLQQQQIPQQQIPQHQIPQQPQQFPQQQFPQQQHQSPQQQFPQQQFPQQKLPQQE-
      70 80 90 100 110 120

```

```

1320 1330 1340 1350 1360
frame4 GRFP.HGV-----LPNRS.R.HLQRMQLQQRELDTRRS.KEHGLLEQ.HPSK.KNEHAS
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|739 --FPQQISQQFPQQLPQQQQIPQQPQQFLQQQQFPQQQFPQQHQFPQQQLP---QQQQIP
      130 140 150 160 170 180

```

```

1370 1380 1390 1400 1410 1420
frame4 QPRLHQQDVG.SLRCRHPF-PAESR.DQYGLGSSKR.GW.RRNLEQARQQPMGKQ.RMD
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|739 QQQQIPQQFPQQIPQQQQIPQQPQQFPQQQFPQQQFPQQQFPQQQFPQQQFPQQQIAR--
      190 200 210 220 230 240

```

```

1430 1440 1450 1460 1470 1480
frame4 HLEHQSLDR.AYHDLQDGWYAHQRHQSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.D
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|739 --QPQLPQQQQIPQQPQQFPQQQFPQQQSPQQQFPQQQFPQQQQLPQKQFP----QP
      250 260 270 280 290

```

```

1490 1500 1510 1520 1530 1540
frame4 QQAEQRTSLERKP.VPCR.LALESLVHAR.LEDDHRLQTSPAHPFREVGSHTDPLDA..E
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|739 QQIPQQQIPQQP---QQFPQQQFPQQQFPQQ--QEFPQQQFPQQQFPQQQLPQQQF
      300 310 320 330 340

```

```

1550 1560 1570 1580 1590
frame4 KRHQ.KCYRNRKHQQDP.SNG-NQLQRYRNREEPQEGVHRYQH HQ.-----SKYHLP
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|739 PQQQFPQQQFPQQQFPQQQQLTQQQFPRPQQSPQQQFPQQQFPQQPQQFPQQQFPPI
      350 360 370 380 390 400

```

```

1600 1610 1620 1630 1640 1650
frame4 YGFWHP.LA.PYQC..HLHLPKDR.YEFHR.RDLQT.RTCGR.TCHLEYGRFQKDQRTYE
      : . :.:
gi|739 YPPQQSEEPSYQQYPPQQQPSGSDVISISGL
      410 420 430

```

```

>>gi|208605346|emb|CAR82266.1| D-type LMW glutenin subun (272 aa)
initn: 75 init1: 75 opt: 101 Z-score: 124.4 bits: 33.9 E(): 0.059
Smith-Waterman score: 101; 17.692% identity (51.154% similar) in 260 aa overlap
(1361-1612:8-254)

```

```

1340 1350 1360 1370 1380 1390
frame4 HLQRMQLQQRELDTRRS.KEHGLLEQ.HPSK.KNEHASQPRLLHQQDVG.SLRCRHPFPA
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|208 MGRLLSPRGKELHTPQEQFPQQQFPQPQ---QQFPQ
      10 20 30

```

```

1400 1410 1420 1430 1440 1450
frame4 ESR.DQYGLGSSKR.GW.RRNLEQARQQPMGKQ.RMDHLEHQSLDR.AYHDLQDGWYAH
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|208 QQILQQHQIPQQPQQFPQQQFLQQQQIPQ-QQIPQQHQIPQQPQQFPQQQFPQQQFPQ
      40 50 60 70 80 90

```

```

1460 1470 1480 1490 1500 1510
frame4 QHQSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.DQQAQRTSLERKP.VPCR.LALE
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|208 QHQSPQQQFPQQQFPQQQLPQQE-FSQQQ-----ISQQPQQLPQQQQIPQQPQQFLQQQ
      100 110 120 130 140

```

```

1520 1530 1540 1550 1560
frame4 SLVHAR.LEDDHRL--QTSPA---PFRE-VGSTHDPLDA..EKRHQ.KCYRNRK--HQQ
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|208 QFPQQQP-PQQHFPQQQLPQQQQIPQQQQIPQQPQQIPQQQIPQQPQQFPQQQFPQQQ
      150 160 170 180 190 200

```

```

1570 1580 1590 1600 1610 1620
frame4 DP.SNGNQLQRYRNREEPQEGVHRYQH HQ.SKYHLPYGFWHP.LA.PYQC..HLHLPKD
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|208 FPQQQFPQQEFPQQQFPQQQIAQQPQLPQQQFPPIPYPPQSQSEEPSYQQYPPQQQPSGS
      210 220 230 240 250 260

```

```

1630 1640 1650 1660 1670 1680
frame4 R.YEFHR.RDLQT.RTCGR.TCHLEYGRFQKDQRTYEQLLDGCLHREACTPWKQKT.P.E
gi|208 DVISICGL
      270

```

```

>>gi|170743|gb|AAB02788.1| HMW glutenin subunit Ax2* [Tr (815 aa)
initn: 98 init1: 63 opt: 106 Z-score: 122.5 bits: 35.2 E(): 0.076
Smith-Waterman score: 123; 20.984% identity (49.223% similar) in 386 aa overlap
(1352-1719:124-491)

```

```

1330 1340 1350 1360 1370
frame4 VLPNRS.R.HLQRMQLQQRELDTRRS.KEHGLLEQ.HPSK.KNEHA-SQPRLLHQ-QDVG
      : .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|170 LQQSILWGIPALLRRYYLSVTSPPQVSYPGQASSQRPGGQQEYYLTSPQQSGWQQPG
      100 110 120 130 140 150

```

```

1380 1390 1400 1410 1420 1430
frame4 .SLRCRHPF-PAESR.DQYGLGSSKR.GW.RRNLEQA-----RQQP-MGKQ.RMDHLEHQ
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:

```

```

gi|170 QGQSGYYPTSPQQSGQKQPGYYPTS--PWQPEQLQOPTQGQQRQPGQGQQLRQGQQGQQ
      160      170      180      190      200      210

      1440      1450      1460      1470      1480      1490
frame4 S-LDR.AYHDLQDGWYAHRRQHSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.DQQA-
      :      :      :      :      :      :
gi|170 SGQGQPRYYPTSSQPGQLQQLAQ--GQGGQQPERGQGGQGS--GQGQQLGQGQGGQPG
      220      230      240      250      260

      1500      1510      1520      1530      1540
frame4 -EQRSLERKP.VPCR.LALESLVHAR.LEDDHRLQTSPAHPFREVQ---STHDPLDA..
      :      :      :      :      :
gi|170 QKQSGQGQGGYYYP---ISPQQLGQGQ--SGQGQLGYPTSP-QQSGQGSGYYPTSAQQ
      270      280      290      300      310      320

      1550      1560      1570      1580      1590      1600
frame4 EKRHQ.KCYRNR--KHQDP.SN-GNQLQRYRNREEPQEGVHRYQHHQ.SKYHLPYGFW
      :      :      :      :      :
gi|170 PGQLQSSQTEQQLGQEQDQSGQGRQGGQSGQRQDDQSG---QGQPGQRQPGYYST
      330      340      350      360      370

      1610      1620      1630      1640      1650      1660
frame4 HP.LA.PYQC..HLHLPKDR.YEFHR.RDLQT.RTCGR.TCHLEYGRFQKDQRTYEQLLD
      :      :      :      :      :
gi|170 SPQQLGQGPYYPTSPQPGQE--QPRQLQPEQ--GQGGQPEQGQGGQQRQGEQGGQ
      380      390      400      410      420      430

      1670      1680      1690      1700      1710      1720
frame4 GCLHREACTPWKQKT.P.EDEAQSQLTSSSSSIPTTTKSKLDVAAASSEKPIREICVD
      :      :      :      :      :
gi|170 PGQGQGGQPGQGQPGYYPTSPQ--QSGQGPGYYPTSPQSGQLQPPAQGGQPGQEQQGQ
      440      450      460      470      480      490

      1730      1740      1750      1760      1770      1780
frame4 WISRDRW.EMGSAHHCRCF.LAPW.ISGQQICKKKNNQILNQ.KQ.TYITIMKIIFNTKQ

gi|170 QPGQGQPGQGQPGYYPTSPQSGQEQLQWQQSGQGQPGHYPTSPLPGQGQPGYYPT
      500      510      520      530      540      550

>>gi|13183177|gb|AAK15089.1|AF240006_1 7S globulin [Sesa (585 aa)
      initn: 87 initl: 51 opt: 96 Z-score: 112.5 bits: 32.8 E(): 0.27
Smith-Waterman score: 121; 20.419% identity (51.309% similar) in 191 aa overlap
(1139-1310:26-215)

      1110      1120      1130      1140      1150      1160
frame4 KNK.Y.FENEQKDHIIH.LFSIHFHTVR.RKPNKKHSLQ--AQQMVEKQFSQCHNT-
      :      :      :      :      :
gi|131 MSCGGRCLVLV FALLASAVVASESKDPELKQCKHQCKAQQQISKEQKEACIQAC
      :      :      :      :      :
      10      20      30      40      50

      1170      1180      1190      1200      1210
frame4 -----QTQ.DSGVCAMLMH.T.RTLSPMIRTKALARGS.DLRIRYRARILELIKQP.
      :      :      :      :      :
gi|131 KEYIRQKHQGEHGRGGDILEEEVWNRKSPIERLRECSRGCQQHGEQREELRRCQEEY
      60      70      80      90      100      110

      1220      1230      1240      1250      1260      1270
frame4 CRRVRS.LQDQPSNP.TLGS..RS..HHQ-Q.QGFQKRD.P.RGSS.RY---GDRGG.RQ
      :      :      :      :      :
gi|131 QREKGRQDDNPTDPEKQYQQCRLQCRRGEGGGFSREHCERRREEKYREQQGREGGRGE
      120      130      140      150      160      170

```

```

      1280      1290      1300      1310      1320
frame4 LLQKRYRDPYRQDDHARREKSHLRNQH-----R.A.DRLRQQKDRFP.HGVLPNRS.R
      :      :      :      :      :
gi|131 MYEGREEREE-EQEEQGRGRIPYVFEDQHFITGFRTOHGMRVLQKFTDRSELLRGIENYR
      180      190      200      210      220      230

      1330      1340      1350      1360      1370      1380
frame4 .HLQRMQLQRELDTRRS.KEHGLLEQ.HPSK.KNEHASQPRLHHQQDVG.SLRCRHPFP

gi|131 VAILEAEPQTFIVPNHWDAESVVFVAKGRGTISLVRQDRRESLNKQGDILKINAGTTAY
      240      250      260      270      280      290

>>gi|19338630|gb|AAL86739.1|AF441864_1 48-kDa glycoprote (448 aa)
      initn: 84 initl: 84 opt: 92 Z-score: 109.5 bits: 31.9 E(): 0.4
Smith-Waterman score: 92; 24.194% identity (54.839% similar) in 62 aa overlap
(1541-1602:3-64)

      1520      1530      1540      1550      1560      1570
frame4 SLVHAR.LEDDHRLQTSPAHPFREVGSTHDPLDA..EKRHQ.KCYRNRKHQQDP.SNGNQ
      :      :      :      :      :
gi|193 MLPKEDPELKKCKHKCRDERQFDEQQRDQKQ
      10      20      30

      1580      1590      1600      1610      1620      1630
frame4 LQRYRNREEPQEGVHRYQHHQ.SKYHLPYGFWHP.LA.PYQC..HLHLPKDR.YEFHR.
      :      :      :      :      :
gi|193 ICEEKARERQEEGNSSEESYGKEQEEENPYVFQDEHFESRVKTEEGRVQVLENFTKRSRL
      40      50      60      70      80      90

      1640      1650      1660      1670      1680      1690
frame4 RDLQT.RTCGR.TCHLEYGRFQKDQRTYEQLLDGCLHREACTPWKQKT.P.EDEAQSQLT

gi|193 LSGIENFRLAILEANPHTFISPAHFDAELVLFVAKGRATITMVREEKRESFNVEHGDIIIR
      100      110      120      130      140      150

>>gi|14285797|sp|O61379.1|TPM_PANST RecName: Full=Tropom (274 aa)
      initn: 61 initl: 61 opt: 86 Z-score: 105.7 bits: 30.5 E(): 0.65
Smith-Waterman score: 86; 28.049% identity (57.317% similar) in 82 aa overlap
(1878-1959:193-266)

      1850      1860      1870      1880      1890      1900
frame4 QFKSTYILNTQINKFTRI.IHQPTINRTN.NLKTRSIRMGK.FHE.IIV.KQ.NVDINNR
      :      :      :      :      :
gi|142 EADLERAEERAETGESKFVELEEBELRVVGNLKSLEVSEEKANQREEAYKEQIKTLTNKL
      170      180      190      200      210      220

      1910      1920      1930      1940      1950      1960
frame4 .NLNNRSEFSERSINKEHRIQQSSK.YEQRSADKEKERSKGTLTLTSEISQKSV.RLGS
      :      :      :      :      :
gi|142 KAAEARAFAERSVQK---LQKEVDRLEDELVNEKEYKS-----ITDELQDTFSELSGY
      230      240      250      260      270

      1970      1980      1990      2000      2010      2020
frame4 IRTINGISNESGFIYSVESETNVRV.HKIIISFFKLPLCPCLAVNTQYLLGWLCFMNV.YY

>>gi|6094504|sp|Q25456.1|TPM_METEN RecName: Full=Tropomy (274 aa)
      initn: 61 initl: 61 opt: 86 Z-score: 105.7 bits: 30.5 E(): 0.65
Smith-Waterman score: 86; 28.049% identity (57.317% similar) in 82 aa overlap
(1878-1959:193-266)

      1850      1860      1870      1880      1890      1900
frame4 QFKSTYILNTQINKFTRI.IHQPTINRTN.NLKTRSIRMGK.FHE.IIV.KQ.NVDINNR

```

```

      : : : : . : . : . : . :
gi|609 EADLERAEERAETGESKIVELEEELRVVGNLNKSLSEVSEEEKANQREEAYKEQIKTLTNKL
      170      180      190      200      210      220

      1910      1920      1930      1940      1950      1960
frame4 .NLNNRSEFSERSINKEHRIQQSSK.YEQRSADKEKERSKGTTLTSEISQKSV.RLGS
      . : : : : : : : . : . : . : : : . : : : :
gi|609 KAAEARAEFAERSVQK---LQKEVDRLEDELVNEKEYKS-----ITDELDTFSELSGY
      230      240      250      260      270

      1970      1980      1990      2000      2010      2020
frame4 IRTINGISNESGFIYSVESETNVRV.HKIIISFFKLPLCPCLAVNTQYLLGWLCFMNV.YY

>>gi|125995159|dbj|BAF47263.1| tropomyosin fast isoform (284 aa)
      initn: 61 init1: 61 opt: 86 Z-score: 105.4 bits: 30.5 E(): 0.67
Smith-Waterman score: 86; 28.049% identity (57.317% similar) in 82 aa overlap
(1878-1959:203-276)

      1850      1860      1870      1880      1890      1900
frame4 QFKSTYILNTQINKFTRI.IHQPTINRTN.NLKTRSIRMGK.FHE.IIV.KQ.NVDINNR
      : : : : . : : : . : : : . : : : . : : :
gi|125 EADLERAEERAETGESKIVELEEELRVVGNLNKSLSEVSEEEKANQREEAYKEQIKTLTNKL
      180      190      200      210      220      230

      1910      1920      1930      1940      1950      1960
frame4 .NLNNRSEFSERSINKEHRIQQSSK.YEQRSADKEKERSKGTTLTSEISQKSV.RLGS
      . : : : : : : : . : . : . : : : . : : : :
gi|125 KAAEARAEFAERSVQK---LQKEVDRLEDELVNEKEYKS-----ITDELDTFSELSGY
      240      250      260      270      280

      1970      1980      1990      2000      2010      2020
frame4 IRTINGISNESGFIYSVESETNVRV.HKIIISFFKLPLCPCLAVNTQYLLGWLCFMNV.YY

>>gi|125995157|dbj|BAF47262.1| tropomyosin fast isoform (284 aa)
      initn: 61 init1: 61 opt: 86 Z-score: 105.4 bits: 30.5 E(): 0.67
Smith-Waterman score: 86; 28.049% identity (57.317% similar) in 82 aa overlap
(1878-1959:203-276)

      1850      1860      1870      1880      1890      1900
frame4 QFKSTYILNTQINKFTRI.IHQPTINRTN.NLKTRSIRMGK.FHE.IIV.KQ.NVDINNR
      : : : : . : : : . : : : . : : : . : : :
gi|125 EADLERAEERAETGESKIVELEEELRVVGNLNKSLSEVSEEEKANQREEAYKEQIKTLTNKL
      180      190      200      210      220      230

      1910      1920      1930      1940      1950      1960
frame4 .NLNNRSEFSERSINKEHRIQQSSK.YEQRSADKEKERSKGTTLTSEISQKSV.RLGS
      . : : : : : : : . : . : . : : : . : : : :
gi|125 KAAEARAEFAERSVQK---LQKEVDRLEDELVNEKEYKS-----ITDELDTFSELSGY
      240      250      260      270      280

      1970      1980      1990      2000      2010      2020
frame4 IRTINGISNESGFIYSVESETNVRV.HKIIISFFKLPLCPCLAVNTQYLLGWLCFMNV.YY

>>gi|125995163|dbj|BAF47265.1| tropomyosin fast isoform (284 aa)
      initn: 61 init1: 61 opt: 86 Z-score: 105.4 bits: 30.5 E(): 0.67
Smith-Waterman score: 86; 28.049% identity (57.317% similar) in 82 aa overlap
(1878-1959:203-276)

      1850      1860      1870      1880      1890      1900
frame4 QFKSTYILNTQINKFTRI.IHQPTINRTN.NLKTRSIRMGK.FHE.IIV.KQ.NVDINNR
      : : : : . : : : . : : : . : : : . : : :
gi|125 EADLERAEERAETGESKIVELEEELRVVGNLNKSLSEVSEEEKANQREEAYKEQIKTLANKL
      180      190      200      210      220      230
```

```

      180      190      200      210      220      230

      1910      1920      1930      1940      1950      1960
frame4 .NLNNRSEFSERSINKEHRIQQSSK.YEQRSADKEKERSKGTTLTSEISQKSV.RLGS
      . : : : : : : : . : . : . : : : . : : : :
gi|125 KAAEARAEFAERSVQK---LQKEVDRLEDELVNEKEYKS-----ITDELDTFSELSGY
      240      250      260      270      280

      1970      1980      1990      2000      2010      2020
frame4 IRTINGISNESGFIYSVESETNVRV.HKIIISFFKLPLCPCLAVNTQYLLGWLCFMNV.YY

>>gi|148615631|gb|ABQ96644.1| tropomyosin [Tyrophagus pu (284 aa)
      initn: 61 init1: 61 opt: 86 Z-score: 105.4 bits: 30.5 E(): 0.67
Smith-Waterman score: 86; 28.049% identity (57.317% similar) in 82 aa overlap
(1878-1959:203-276)

      1850      1860      1870      1880      1890      1900
frame4 QFKSTYILNTQINKFTRI.IHQPTINRTN.NLKTRSIRMGK.FHE.IIV.KQ.NVDINNR
      : : : : . : : : . : : : . : : : . : : :
gi|148 EADLERAEERAETGESKIVELEEELRVVGNLNKSLSEVSEEEKANQREEAYKEQIKTLTNKL
      180      190      200      210      220      230

      1910      1920      1930      1940      1950      1960
frame4 .NLNNRSEFSERSINKEHRIQQSSK.YEQRSADKEKERSKGTTLTSEISQKSV.RLGS
      . : : : : : : : . : . : . : : : . : : : :
gi|148 KAAEARAEFAERSVQK---LQKEVDRLEDELVNEKEYKS-----ITDELDTFSELTGY
      240      250      260      270      280

      1970      1980      1990      2000      2010      2020
frame4 IRTINGISNESGFIYSVESETNVRV.HKIIISFFKLPLCPCLAVNTQYLLGWLCFMNV.YY

>>gi|2660868|gb|AAC48288.1| fast tropomyosin isoform [Ho (284 aa)
      initn: 61 init1: 61 opt: 86 Z-score: 105.4 bits: 30.5 E(): 0.67
Smith-Waterman score: 86; 28.049% identity (57.317% similar) in 82 aa overlap
(1878-1959:203-276)

      1850      1860      1870      1880      1890      1900
frame4 QFKSTYILNTQINKFTRI.IHQPTINRTN.NLKTRSIRMGK.FHE.IIV.KQ.NVDINNR
      : : : : . : : : . : : : . : : : . : : :
gi|266 EADLERAEERAETGESKIVELEEELRVVGNLNKSLSEVSEEEKANQREEAYKEQIKTLANKL
      180      190      200      210      220      230

      1910      1920      1930      1940      1950      1960
frame4 .NLNNRSEFSERSINKEHRIQQSSK.YEQRSADKEKERSKGTTLTSEISQKSV.RLGS
      . : : : : : : : . : . : . : : : . : : : :
gi|266 KAAEARAEFAERSVQK---LQKEVDRLEDELVNEKEYKS-----ITDELDTFSELSGY
      240      250      260      270      280

      1970      1980      1990      2000      2010      2020
frame4 IRTINGISNESGFIYSVESETNVRV.HKIIISFFKLPLCPCLAVNTQYLLGWLCFMNV.YY

>>gi|125995169|dbj|BAF47268.1| tropomyosin slow-twitch i (284 aa)
      initn: 61 init1: 61 opt: 86 Z-score: 105.4 bits: 30.5 E(): 0.67
Smith-Waterman score: 86; 28.049% identity (57.317% similar) in 82 aa overlap
(1878-1959:203-276)

      1850      1860      1870      1880      1890      1900
frame4 QFKSTYILNTQINKFTRI.IHQPTINRTN.NLKTRSIRMGK.FHE.IIV.KQ.NVDINNR
      : : : : . : : : . : : : . : : : . : : :
gi|125 EADLERAEERAESGESKIVELEEELRVVGNLNKSLSEVSEEEKANQREETYKEQIKTLANKL
      180      190      200      210      220      230
```



```

30 56 12:***=
32 238 45:***=*****
34 227 122:*****=*****
36 302 250:*****=*****
38 393 414:*****=*****
40 401 577:*****=*****
42 624 706:*****=*****
44 632 779:*****=*****
46 701 793:*****=*****
48 821 759:*****=*****
50 651 693:*****=*****
52 543 609:*****=*****
54 402 520:*****=*****
56 392 435:*****=*****
58 311 357:*****=*****
60 294 289:*****=*****
62 289 232:*****=*****
64 267 184:*****=*****
66 161 146:*****=*****
68 347 115:*****=*****
70 62 90:*****=*****
72 65 70:*****=*****
74 55 55:*****=*****
76 35 43:*****=*****
78 33 33:*****=*****
80 18 26:*****=*****
82 16 20:*****=*****
84 15 16:*****=*****
86 7 12:*****=*****
88 9 9:*****=*****
90 1 7:*****=*****
92 2 6:*****=*****
94 2 4:*****=*****
96 2 3:*****=*****
98 2 3:*****=*****
100 0 2:*****=*****
102 0 2:*****=*****
104 0 1:*****=*****
106 0 1:*****=*****
108 0 1:*****=*****
110 0 1:*****=*****
112 0 0:*****=*****
114 0 0:*****=*****
116 1 0:*****=*****
118 0 0:*****=*****
>120 0 0:*****=*****

2069351 residues in 8448 sequences
Expectation_n fit: rho(ln(x))= 3.16360.000675; mu= 23.4716 0.034
mean_var=49.510010.159, 0's: 60 Z-trim: 60 B-trim: 1355 in 2/60
Lambda= 0.182275
Kolmogorov-Smirnov statistic: 0.0475 (N=29) at 36

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 44, opt: 32, open/ext: -10/-2, width: 16
The best scores are:
gi|28194389|gb|AAO27454.1| myotoxic phospholipase (138) 86 31.5 0.91

>>gi|28194389|gb|AAO27454.1| myotoxic phospholipase A2 [ (138 aa)
initn: 40 init1: 40 opt: 86 Z-score: 116.7 bits: 31.5 E(): 0.91
Smith-Waterman score: 86; 24.793% identity (52.066% similar) in 121 aa overlap
(423-535:26-137)

```

```

400 410 420 430 440 450
frame4 LAGLWWSRQQTWWWACKPWRTKIQICVFWWLEGKGASPFEN.WNHSCYI.RKLQA...ST
gi|281 MRTLWIMAVLLVGVEGDLWQFGQMILKETGKLPFPYYTYGCGYCGWGGQGPCKDA
10 20 30 40 50

460 470 480 490 500
frame4 FADPCRFPGHEAIF-KH..FKPAGLVDKSSGGRRR.IIIMLEA-----FFRCNNCGSIN
gi|281 -TDRCCFV-HDCCYGKLTNCKPK--TDRYSYSRENGVIICGEGTPCEKQICECDKAAAVC
60 70 80 90 100 110

510 520 530 540 550 560
frame4 FRRGLRLCLQATRRRRFGEF.SSKVCR-PTTKFAACSTRVRRG.SNREK..SSSCCHNEF
gi|281 FRENLR----TYKKRYMAYPDVLCCKPAEKC
120 130

570 580 590 600 610 620
frame4 LRSLLR.NV.S.NHDHKTKSNQHEML.SSSTAALQPIVYQ.REVLIL.LLARCLHLSLQE

```

2417 residues in 1 query sequences  
2069351 residues in 8448 library sequences  
Scomplib [34t26]  
start: Wed Feb 3 23:17:15 2010 done: Wed Feb 3 23:17:16 2010  
Total Scan time: 0.540 Total Display time: 0.000

Function used was FASTA [version 3.4t26 July 7, 2006]

# fasta34 frame4.pep /genedata/1/db/PRT\_2010 -Q -E 1 -O frame4.pep\_prt.fasta  
FASTA searches a protein or DNA sequence data bank version 3.4t26 July 7, 2006  
Please cite:  
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

frame4, 2417 aa  
vs /genedata/1/db/PRT\_2010 library

```

< 20 opt E()
20 279245 0:*****=
22 206 0:*****=
24 370 17:*****=
26 905 374:*****=
28 4028 4039:*****=
30 20695 24533:*****=
32 80296 94863:*****=
34 238430 257256:*****=
36 513797 528343:*****=
38 833737 873155:*****=
40 1149426 1217974:*****=
42 1390433 1488825:*****=
44 1512715 1642313:*****=
46 1542050 1672736:*****=
48 1530092 1601452:*****=
50 1391367 1461331:*****=
52 1268871 1284754:*****=
54 1106914 1097405:*****=
56 947403 916670:*****=

```



```

58 792722 752569:=====*=
60 662195 609625:=====*=
62 522585 488738:=====*=
64 415440 388691:=====*=
66 336949 307209:=====*=
68 268597 241644:=====*=
70 215720 189367:=====*=
72 169915 147972:=====*=
74 132999 115369:=====*=
76 110307 89794:=====*=
78 83169 69796:=====*=
80 65186 54195:=====*=
82 49148 41458:*=
84 37629 32840:*=
86 29641 25410:*=
88 22603 19661:*          inset = represents 286 library sequences
90 17946 15213:*          :=====*=
92 14268 11771:*          :=====*=
94 10987 9108:*          :=====*=
96 8371 7047:*          :=====*=
98 6962 5453:*          :=====*=
100 5323 4219:*          :=====*=
102 3909 3264:*          :=====*=
104 3121 2526:*          :=====*=
106 2601 1954:*          :=====*=
108 2053 1512:*          :=====*=
110 1804 1170:*          :=====*=
112 1405 905:*          :=====*=
114 1116 700:*          :=====*=
116 943 542:*          :=====*=
118 717 419:*          :=====*=
>120 3914 324:*          :=====*=
4761287459 residues in 17815538 sequences
statistics sampled from 60000 to 17807614 sequences
Expectation_n fit: rho(ln(x))= 4.69230.000184; mu= 16.0857 0.010
mean_var=59.838912.041, 0's: 927 Z-trim: 933 B-trim: 2504 in 1/64
Lambda= 0.165799
Kolmogorov-Smirnov statistic: 0.0385 (N=29) at 52

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 44, opt: 32, open/ext: -10/-2, width: 16
The best scores are:
opt bits E(17815538)
gi|193900736|gb|EDV99602.1| GH12332 [Drosophila gr (1317) 219 62.9 6.5e-06
gi|193907442|gb|EDW06309.1| GI21589 [Drosophila mo ( 790) 202 58.7 7.4e-05
gi|194158266|gb|EDW73167.1| GK17402 [Drosophila wi (2328) 196 57.6 0.00046
gi|28569857|dbj|BAC57901.1| gag-like protein [Anop ( 724) 189 55.5 0.00059
gi|113530477|emb|CAJ96824.1| Hypothetical protein ( 941) 188 55.4 0.00086
gi|5737842|gb|AAD50121.1|AF153362_1 adenylyl cycla (2123) 190 56.1 0.0012
gi|190654481|gb|EDV51724.1| GG13740 [Drosophila er (1835) 186 55.1 0.002
gi|145020705|gb|EDK04834.1| hypothetical protein M ( 704) 179 53.1 0.0031
gi|193893136|gb|EDV92002.1| GH24674 [Drosophila gr ( 883) 179 53.2 0.0036
gi|50417567|gb|AAH75588.1| K14 protein [Xenopus la (1320) 178 53.1 0.0059
gi|198145733|gb|EDY72268.1| GA22701 [Drosophila ps ( 520) 169 50.7 0.013
gi|198150617|gb|EAL29745.2| GA14630 [Drosophila ps ( 917) 169 50.8 0.02
gi|193907175|gb|EDW06042.1| GI16140 [Drosophila mo ( 900) 167 50.3 0.027
gi|12383113|gb|AAG24941.2| unknown [Frankia sp. Ar ( 498) 163 49.2 0.033
gi|135224193|gb|EBG27286.1| hypothetical protein G ( 423) 162 48.9 0.034
gi|144191545|gb|EDJ53627.1| hypothetical protein G ( 293) 160 48.3 0.036
gi|116697810|gb|ABK16998.1| putative chromosome se ( 848) 164 49.6 0.042
gi|137509825|gb|EBU48087.1| hypothetical protein G ( 215) 157 47.5 0.047
gi|170945019|emb|CAP71130.1| unnamed protein produ (1329) 165 50.0 0.051
gi|221488672|gb|EEE26886.1| hypothetical protein T (1948) 165 50.1 0.069

```

```

gi|176558|gb|AAA35375.1| involucrin (large allele) ( 544) 159 48.3 0.069
gi|148498882|gb|ABQ67136.1| short-chain dehydrogen ( 553) 158 48.0 0.082
gi|158601726|gb|EDP38557.1| Low-density lipoprotei ( 923) 160 48.7 0.088
gi|124415991|emb|CAK81050.1| unnamed protein produ ( 504) 157 47.8 0.09
gi|193919917|gb|EDW18784.1| GI11885 [Drosophila mo ( 967) 160 48.7 0.091
gi|193907429|gb|EDW06296.1| GI21598 [Drosophila mo ( 532) 157 47.8 0.094
gi|218675676|gb|AAI69292.2| repetin [synthetic con ( 383) 155 47.2 0.1
gi|50057436|emb|CAH03420.1| hypothetical protein f ( 2301) 163 49.7 0.11
gi|124390320|emb|CAK55862.1| unnamed protein produ (2301) 163 49.7 0.11
gi|76580731|gb|ABA50206.1| adenosine deaminase [Bu ( 726) 157 47.9 0.12
gi|190623993|gb|EDV39517.1| GF24442 [Drosophila an ( 734) 157 47.9 0.12
gi|158271402|gb|EDO97222.1| RWP-RK transcription f (2146) 162 49.4 0.12
gi|241944366|gb|EES17511.1| hypothetical protein S ( 958) 158 48.2 0.13
gi|89303110|gb|EAS01098.1| hypothetical protein TT (3459) 164 50.1 0.13
gi|2623367|gb|AAC53441.1| sex determining protein ( 420) 154 47.0 0.13
gi|9916|emb|CAA39663.1| liver stage antigen [Plasm (1909) 161 49.2 0.13
gi|142534199|gb|ECY77410.1| hypothetical protein G ( 445) 154 47.0 0.13
gi|124403363|emb|CAK68827.1| unnamed protein produ ( 517) 154 47.1 0.15
gi|193897155|gb|EDV96021.1| GH15437 [Drosophila gr ( 490) 153 46.8 0.17
gi|143313824|gb|EDE27300.1| hypothetical protein G ( 286) 150 45.9 0.19
gi|134083327|emb|CAK42894.1| unnamed protein produ ( 409) 151 46.3 0.21
gi|219884915|gb|ACL52832.1| unknown [Zea mays] ( 375) 150 46.0 0.23
gi|148668340|gb|EDL00666.1| repetin [Mus musculus] (1118) 155 47.5 0.23
gi|228383859|gb|ACQ27812.1| Sequence 20326 from pa ( 405) 150 46.0 0.24
gi|239594646|gb|EEQ77227.1| conserved hypothetical ( 811) 153 47.0 0.25
gi|146143822|gb|EAS04465.2| hypothetical protein T ( 723) 152 46.7 0.27
gi|1806132|emb|CAA67624.1| repetin [Mus musculus] (1130) 154 47.3 0.28
gi|194190700|gb|EDX04276.1| GD23558 [Drosophila si (1025) 153 47.0 0.3
gi|124871096|gb|EAY62812.1| Outer membrane autotra (1356) 154 47.4 0.32
gi|56313409|emb|CAI08054.1| conserved hypothetical ( 728) 151 46.5 0.33
gi|194107557|gb|EDW29600.1| GL22673 [Drosophila pe ( 884) 151 46.5 0.38
gi|44985804|gb|AAS54428.1| AGL062Cp [Ashbya gossyp (1110) 152 46.8 0.38
gi|164646317|gb|EDR10563.1| predicted protein [Lac (1141) 152 46.8 0.39
gi|140190328|gb|ECK13821.1| hypothetical protein G ( 260) 145 44.7 0.39
gi|17862888|gb|AAL39921.1| SD01663p [Drosophila me (1026) 151 46.6 0.42
gi|220947526|gb|ACL86306.1| Glt-PA [synthetic cons (1026) 151 46.6 0.42
gi|7297439|gb|AAF52697.1| glutactin, isoform C [Dr (1026) 151 46.6 0.42
gi|22945980|gb|AAN10678.1| glutactin, isoform A [D (1026) 151 46.6 0.42
gi|22945981|gb|AAN10679.1| glutactin, isoform B [D (1026) 151 46.6 0.42
gi|193898586|gb|EDV97452.1| GH16879 [Drosophila gr ( 747) 149 46.0 0.46
gi|156216089|gb|EDO37033.1| predicted protein [Nem ( 263) 144 44.5 0.47
gi|135168194|gb|EBF93005.1| hypothetical protein G ( 327) 145 44.8 0.47
gi|136919510|gb|EBR19291.1| hypothetical protein G ( 221) 143 44.2 0.48
gi|297085|emb|CAA37380.1| glutactin [Drosophila me (1023) 150 46.3 0.5
gi|156226676|gb|EDO47484.1| predicted protein [Nem ( 387) 145 44.8 0.54
gi|163773562|gb|EDQ87200.1| predicted protein [Mon ( 953) 149 46.1 0.56
gi|162677944|gb|EDQ64408.1| predicted protein [Phy ( 407) 145 44.8 0.56
gi|140908983|gb|ECO31917.1| hypothetical protein G ( 268) 143 44.2 0.56
gi|193907914|gb|EDW06781.1| GI15215 [Drosophila mo (1037) 149 46.1 0.6
gi|44984214|gb|AAS53180.1| AFL194Wp [Ashbya gossyp (1168) 149 46.1 0.65
gi|210065384|gb|EEA19478.1| conserved hypothetical ( 403) 144 44.6 0.66
gi|124423289|emb|CAK88084.1| unnamed protein produ (1135) 148 45.9 0.76
gi|156213492|gb|EDO34509.1| predicted protein [Nem ( 601) 145 45.0 0.76
gi|223539286|gb|EEF40879.1| U1 small nuclear ribon ( 506) 144 44.7 0.78
gi|228386947|gb|ACQ30900.1| Sequence 23414 from pa ( 409) 143 44.4 0.78
gi|142624704|gb|ECC241030.1| hypothetical protein G ( 333) 142 44.1 0.79
gi|135074547|gb|EBF32912.1| hypothetical protein G ( 472) 143 44.4 0.87
gi|193908644|gb|EDW07511.1| GI14833 [Drosophila mo ( 751) 145 45.0 0.9
gi|190619641|gb|EDY35165.1| GF22341 [Drosophila an ( 933) 146 45.3 0.9
gi|228272477|gb|ACQ00931.1| Sequence 33148 from pa ( 756) 145 45.0 0.91
gi|167881043|gb|EDS44426.1| conserved hypothetical ( 331) 141 43.8 0.92
gi|156215499|gb|EDO36457.1| predicted protein [Nem ( 416) 142 44.1 0.94

```

gi|228388133|gb|ACQ32086.1| Sequence 24600 from pa ( 552) 143 44.5 0.99

>>gi|193900736|gb|EDV99602.1| GH12332 [Drosophila grimsh (1317 aa)  
initn: 87 initl: 87 opt: 219 Z-score: 268.7 bits: 62.9 E(): 6.5e-06  
Smith-Waterman score: 271; 16.446% identity (52.552% similar) in 529 aa overlap  
(1218-1742:217-725)

```
1190      1200      1210      1220      1230      1240
frame4 KPMIRTKALARGS.DLRIRYRARIILELIKQP.CRRVRS.LQDQPSNP.TLGS..RS..HH
gi|193 VPGPVVVLVPNDYWSDMERQGAAPQPQLVLPAREYWQDVVQQRPHQQLPLQQQEQLQQQQ
190      200      210      220      230      240
```

```
1250      1260      1270      1280      1290      1300
frame4 QQ.QGFQKRDP.RGSS.RYGDRGG.RQLLQKRYRDPYRQDDHARREKSHLHRNQH-R.A.
gi|193 QQQQQQQQQHQQQQEQQQQQQLQQQEQQQEQHQQHQQQQQQQQQQQLQQQLQQQLQQQE
250      260      270      280      290      300
```

```
1310      1320      1330      1340      1350      1360
frame4 DRLRQQKDGFRP.HGVLPNRS.R.HLQRMQLQORELDTRRS.KEHGLEQ.HPSK.KNEH
gi|193 QQQHQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQHQQYQHQQQQHQQQQQLQQQLQQQLQQQH
310      320      330      340      350      360
```

```
1370      1380      1390      1400      1410      1420
frame4 ASQPRLLHQQDVG.SLRCRHFPFAESR.DQYGLGSSKR.GW.RRNLEQARQQPMGKQ.RM
gi|193 QQQQQQQQQQLQQQE-----QQEQQQEQQQEQQQQQQQQQQQQQQQQQMQQMQQQQQL
370      380      390      400      410      420
```

```
1430      1440      1450      1460      1470      1480
frame4 DHLEHQSLDR.AYHDLQDGYAHRQHQSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.
gi|193 HHQQQLQQDHPRQQEQQLQQQEQQRQHQQHQLHQ-----QQLHQQLLQREQQQLLQLQ
430      440      450      460      470
```

```
1490      1500      1510      1520      1530      1540
frame4 DQQAQRTSLERKP.VPCR.LALESLVHAR.LEDDHRLQTSPAHPFREVSTHDPDA..
gi|193 QQHQCQEQQLRQQQQIQ---LHQQQQQLHQQLR--HQ-QLQLHQ-EQMHHQQNEQQQLQQ
480      490      500      510      520
```

```
1550      1560      1570      1580      1590      1600
frame4 EKRHQ.KCYRNR--KHQQDP.SNGNQLQRYRNREEPQEGVHRYQHQQ.SKYHLPYGFWH
gi|193 QRLHQQRQLQQEHPRHQEHPRHQEQLTQQEQEQQQHQQQQQQQQSHQ-QQQQLQQQL-H
530      540      550      560      570      580
```

```
1610      1620      1630      1640      1650      1660
frame4 P.LA.PYQC..H-LHLPKDR.YEFHR.RDLQT.RTCGR.TCHLEYGRFQKDQRTYEQLLD
gi|193 QQLQLQLQLKMHYHQQQLQQQLQQQLQQQLQQQLHQQLLYQQQLHQHQLHQHQLH
590      600      610      620      630      640
```

```
1670      1680      1690      1700      1710      1720
frame4 GCLHREACTPWKQKT.P.EDEAQSQLTSSSSSIPTTTKSKDLVDAAASSEKPIREICVD
gi|193 QYHHQQ-QKEQQKEQQKQQEQEQEQQQQQQALRQSVLLDLNLEQQDAVVHPQKHTEIE
650      660      670      680      690      700
```

```
1730      1740      1750      1760      1770      1780
frame4 WISRDWR.EMGSAHHCRCF.LAPW.ISGQQICKKKNNQILNQ.KQ.TYITIMKIIFNNTKQ
gi|193 LQDIGFEAQMESTTQEDWCEREQPTDAITRQPDTTQPTLSPNDKDEVLAASSFAETRKQKQ
710      720      730      740      750      760
```

>>gi|193907442|gb|EDW06309.1| GI21589 [Drosophila mojave (790 aa)  
initn: 146 initl: 69 opt: 202 Z-score: 249.9 bits: 58.7 E(): 7.4e-05  
Smith-Waterman score: 222; 20.259% identity (52.371% similar) in 464 aa overlap  
(1265-1718:12-445)

```
1240      1250      1260      1270      1280      1290
frame4 .TLGS..RS..HHQQ.QGFQKRDP.RGSS.RYGDRGG.RQLLQKRYRDPYRQ-DDHARRE
gi|193 MAKCVDYKWRKRCMNKWQQEQQLQQQQEQLEQQQLQQQLRRE
10      20      30      40
```

```
1300      1310      1320      1330      1340
frame4 KSHLHRNQH-R.A.DRLRQQKDGFRP.HGVLPNRS.R.HLQRMQ---LQORELDTRRS.K
gi|193 QEQLQQQLHQLWEQERQREQ-----QQILQQQLQQQLQWEQEQQLLREQEQLLQREQE
50      60      70      80      90
```

```
1350      1360      1370      1380      1390      1400
frame4 EHGLEQ.HPSK.KNEHASQPRLLHQQDVG.SLRCRH-PFPAESR.DQYGL--GSSKR.G
gi|193 QQQLLQQQLLQQLR-EHEQQ--LLREQEHQQLQQQLQWEQEQQRQEQQLREHEQQQLQ
100      110      120      130      140      150
```

```
1410      1420      1430      1440      1450      1460
frame4 W.RRNLEQARQQPMGKQ.RMDHLEHQSLDR.AYHDLQDGYAHRQHQSQR.G.H-QNLEA
gi|193 WEQDQQRQKQLWEHEQLLLREQEQQLKREQEQQLLQQ-QQLHQQQLQQQLLQQLLREQEQ
160      170      180      190      200      210
```

```
1470      1480      1490      1500      1510      1520
frame4 SFQCGHES.LAR.QW.YLGC.DQQAQRTSLERKP.VPCR.LALESLVHAR.LEDDHRLQ
gi|193 QLLREQEQQLLREQEQQL--LREQEQQQQLLREQEQQL-QQLQQQLLREQEQEQEQQRQRQ
220      230      240      250      260
```

```
1530      1540      1550      1560      1570      1580
frame4 TSPAHPFREVSTHDPDA..EKRHQ.KCYRNRKHKQDP.SNGNQLQRYRNREEPQEQGV
gi|193 KQQKQQLRE-----QKRQKQLQRQQQKQQRQQK-QQRQQKQQRQQKQQL
270      280      290      300      310
```

```
1590      1600      1610      1620      1630      1640
frame4 HRYQHQQ.SKYHLPYGFWHP.LA.PYQC..HLHLPKDR.YEFHR.RDLQT.RTCGR.TCH
gi|193 RDQQEQQLRLQIQKQLQEQEQQLQLQRQLKLOREQEQQLKLOREQEQQLKLOREQDQ
320      330      340      350      360      370
```

```
1650      1660      1670      1680      1690      1700
frame4 LEYGRFQ-KDQRTYEQLLDGCLHREACTPWKQKT.P.EDEAQSQLTSSSSSIPTTTKSD
gi|193 QLQLQQLKLOREQEQQLK--LREQE---QQLKLOREQEQQLKLOREQEQQLKLOREQE
380      390      400      410      420      430
```

```
1710      1720      1730      1740      1750      1760
frame4 KLVDAASSEKPIREICVDWISRDWR.EMGSAHHCRCF.LAPW.ISGQQICKKKNNQILN
..
```

```

gi|193 QQLKLQREQEQQKLQREQEQQQLQLQIQKQLQEEQEQQEQQLQLQQQQLQLQHQQQKQLQQ
      440      450      460      470      480      490

>>gi|194158266|gb|EDW73167.1| GK17402 [Drosophila willis (2328 aa)
  initn: 137 initl: 66 opt: 196 Z-score: 235.6 bits: 57.6 E(): 0.00046
Smith-Waterman score: 242; 17.784% identity (48.251% similar) in 686 aa overlap
(1139-1765:220-897)

      1110      1120      1130      1140      1150      1160
frame4 KNK.Y.FENEQKDHIH.LFSIHFHFTVR.RKPNKKHKSLLQAQMVQEKQFSQCHNTQTQ
      .. .. . : : : : : : : : : : : : : : : : : : : : : : : : :
gi|194 PTEKIRSQKHHQQQQQQQQQQQQQQQQHQQQLQQQQQQQQQQQQQQQQQQQQQLQVQVQQQ
      190      200      210      220      230      240

      1170      1180      1190      1200      1210      1220
frame4 .DSGVCAMKLMH.T.RTSLKPMIRTKALARGS.DLRIRYRARIKLIKQP.CRR--VRS.
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|194 QQSQPTAATSTLSEILSFGEDSAAAKSIANIKSQLQAS--AEQMQQQMPATQQQVVFVR
      250      260      270      280      290      300

      1230      1240      1250      1260      1270
frame4 LQDQPSNP.TLGS..RS..HHQQ.QGF--QKRD.PRGSS.RYGDRGG.RQ-----LLQK
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|194 PQAQPQQQVQVQVQVATSHPPQQQIIIIQQKILPGNLQQQQQQQPNQQQQIKNILQQQ
      310      320      330      340      350      360

      1280      1290      1300      1310      1320      1330
frame4 RYRDPYRQDDHARREKSHLHRNQHR.A.DRLRQQKDGRFP.HGVLPNRS.R.HLQRMQLQ
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|194 RQMKSSQQQQQQQQHQQHLHQQQQQQQQQQQQQQQQQQQQIQHQIQVLLQQQQQ
      370      380      390      400      410      420

      1340      1350      1360      1370      1380      1390
frame4 QRELDTRRS.KEHGLLEQ.HPSK.KNEHASQPR--LLHQQDVG.SLRCRHPFPAESR.DQ
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|194 QHPEPQQQQPQQQSPRLLATPTGRLTPTPTPTTGTGVTTPPQQQQQVQVQVQVQVQVQ
      430      440      450      460      470      480

      1400      1410      1420      1430      1440
frame4 YGLGSSKR.GW.RRNLEQARQQPMGKQ.RMDHL--EHQSLDR.--AYHDL-QDGWY---A
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|194 QQVVQQQQVQQQQVQQQQQQQSSISPQ--MVHLIRQQQQQLQQQHMQYQQLIQSGQTLSPQ
      490      500      510      520      530      540

      1450      1460      1470      1480      1490      1500
frame4 HRQHQSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.DQQAQ--QRTSLERKP.VPCR
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|194 QLQHQRIAQQQQQLQMQLVQQQQQLQMQLVQQQQSQGLQQQQSPRMVQNHSPGMPSTPSP
      550      560      570      580      590      600

      1510      1520      1530      1540      1550
frame4 .LALESVHA-----R.LEDDHRLQTSAPHPFREVGSTHDLPLDA..EKRHQ.KC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|194 SLQQQQQHASSNMLPPPPQSPRQLQSPAPLMTPPPPPSQSAQHRLRQQQQQQQQQQQQQL
      610      620      630      640      650      660

      1560      1570      1580      1590
frame4 YRNRKH-----QQDP.SNGNQLQRYR-----NREEPQEQGVHRYQ-----HHQ.S
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|194 QQNRQHMIGSPQQPPMMSFQQQQQIQPLQQQQQMPFFQQPVHQQRMQLQQQMAQQQQQS
      670      680      690      700      710      720

```

```

      1600      1610      1620      1630      1640
frame4 KYHLPYGFWHP.LA-.PYQC..HLHLPKDR.YEFHR.RD---LQT.RTCGR.TCHLEYG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|194 PQHIS-APQSPQISQTPPMQSKHQHTVTGQQTLQKPIDPTDFVQVAQVLSRSTLSSNQD
      730      740      750      760      770      780

      1650      1660      1670      1680      1690      1700
frame4 RFQKDQRTYEQLLDGCLHREACTPWKQKT.P.EDEAQSQLTSSSSSIPTTTKSKDLVDA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|194 SLIMRQQQLKQQQQQQLQQSQIAPQQQPQQPQTPTLQQQQQQQQQPQQ-PQQQPPQTQPQT
      790      800      810      820      830      840

      1710      1720      1730      1740      1750      1760
frame4 AASSEKPIREICVDWISRDWR.EMGSAHHCRCF.LAPW.ISGQQICKKKNNQILNQ.KQ.
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|194 TTLQQQPQPQLQQQQHPQQQQQTLAPQQQQQ--ALAPQQQQQQQLATGQQQQVMPQRHVI
      850      860      870      880      890      900

      1770      1780      1790      1800      1810      1820
frame4 TYITIMKIIFNTKQS..KAAIGVKPISTTKIVVLYHLS.AS.IQSNRQFQEI.SKQSENY
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|194 NTSTAQQGQIIQSHMMNLQQKQQQQQQVHLHQQQPQQQIPQQQQQQPQLQQQQQPQL
      910      920      930      940      950      960

>>gi|28569857|dbj|BAC57901.1| gag-like protein [Anophele (724 aa)
  initn: 120 initl: 75 opt: 189 Z-score: 233.6 bits: 55.5 E(): 0.00059
Smith-Waterman score: 227; 19.697% identity (53.636% similar) in 330 aa overlap
(1269-1592:165-477)

      1240      1250      1260      1270      1280      1290
frame4 S..RS..HHQQ.QGFQKRD.PRGSS.RYGDRGG.RQLLKRYRDPYRQDDHARREKSHLH
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|285 RRENELLTGTRSVLELQTAANATLQQSSGGGNRETARKRQQLRRRE----RERQQQQ
      140      150      160      170      180      190

      1300      1310      1320      1330      1340      1350
frame4 RNQHR.A.DRLRQQKDGRFP.HGVLPNRS.R.HLQRMQLQQRELDTR---RS.KEHGLL
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|285 QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQLQWTTTVVRGPRSPRHRQP
      200      210      220      230      240      250

      1360      1370      1380      1390      1400      1410
frame4 EQ.HPSK.KNEHASQPRLLHQQDVG.SLRCRHPFPAESR.DQYGLGSSKR.GW.RRNLEQ
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|285 QQQQQQQQQQQGGERYVPPQLRQQ-----RQQQQRPQQQQQQQQQQQQGGERYVPPQLRQ
      260      270      280      290      300

      1420      1430      1440      1450      1460      1470
frame4 ARQQPMGKQ.RMDHLEH-QSLDR.AYHDLQDGWYAHRRQHSQR.G.HQNLEASFQCGHES
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|285 QRQQQQHQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQHQ
      310      320      330      340      350      360

      1480      1490      1500      1510      1520      1530
frame4 .LAR.QW.YLGC.DQQAQRTSLERKP.VPCR.LALESVHAR.LEDDHRLQTSAPHPFR
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|285 --QQQW-----QQQQQQQQPQSLPHRKQTQLQLSPRLQQQQQQQQSQQQQQQPQQ
      370      380      390      400      410

      1540      1550      1560      1570      1580      1590
frame4 EVGSTH-DPLDA..EKRHQ.KCYRNRKHQQDP.SNGNQLQRYRNREEPQEQGVHRYQH
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

gi|285 LLWTTVVRSCPSQRQRQLQQQQQQQQQQGGERYVPPQLRQQRQQQQPQQQQQRPQQQR
      420      430      440      450      460      470

      1600      1610      1620      1630      1640      1650
frame4 .SKYHLPYGFWHP.LA.PYQC..HLHLPKDR.YEFHR.RDLQT.RTCGR.TCHLEYGRFQ

gi|285 PQQQRPPQQRSQQRKPAKPELIEVSPNEGQDWESLLLVQTAVTRDERYKPLKDHVVLRG
      480      490      500      510      520      530

>>gi|113530477|emb|CAJ96824.1| Hypothetical protein [Ral (941 aa)
      initn: 128 initl: 88 opt: 188 Z-score: 230.7 bits: 55.4 E(): 0.00086
Smith-Waterman score: 235; 22.105% identity (52.632% similar) in 380 aa overlap
(1230-1595:574-940)

      1200      1210      1220      1230      1240      1250
frame4 S.DLRIRYRARILELIKQP.CRRVRS.LQDQPSNP.TLGS..RS..HHQQ.QGFQKRD.P.
      540      550      560      570      580      590      600
gi|113 SNAPGEPWRGNGRGSSRPDEARGFVGQPGGADPQRRAGMQREQQ-QVLRQQQME
      550      560      570      580      590      600

      1260      1270      1280      1290      1300      1310
frame4 RGSS.RYGDRGG.RQL--LQKRYRDPYRQDDHARREKSHLHRNQHR.A.DRLRQKQ---D
      610      620      630      640      650      660
gi|113 QQREQQRALREQRQAPDLQROQQEA-RQRDAQQRQWQDQARDQQRQDQRMQQRQWAD
      610      620      630      640      650      660

      1320      1330      1340      1350      1360      1370
frame4 GR-FP.HGVLPNRS.R.HLQRMQLQQRELDLT--RRS.KEHGLLEQ.HPSK.KNEHASQP
      670      680      690      700      710
gi|113 GQDRPDGP---RQSQEMHRQRFE-QQRQMQDAQQQQMEAQQQQESRQRQAQDAQR
      670      680      690      700      710

      1380      1390      1400      1410      1420
frame4 RLLHQDDVG.SLRCRHPFPAESR.DQYGLGSSKR.GW.RRN-LEQARQQPMGKQ.RMDHL
      720      730      740      750      760      770
gi|113 QQMEQQRQSEAAQQRQM--QESRQRQAQDIQRQQMEQQRQSEAAQQRQMQEQAQDI
      720      730      740      750      760      770

      1430      1440      1450      1460      1470      1480
frame4 EHQSILDR.AYHDLQDGWYAHQRHQHSQR.G.HQNLEASFQCGHES.LAR-.QW.YLGC.DQ
      780      790      800      810      820      830
gi|113 QRQQMEQQRQSEAAQQRQAADQQRQ-QAERQRQAQDIQRQQMEQQRQSEAAQQRQMQEAAQ
      780      790      800      810      820      830

      1490      1500      1510      1520      1530      1540
frame4 -QAE--QRTSLERKP.VPCR.LALESLVHAR.LEDDHRLQTSPAHPFREVGSSTHPLDA.
      840      850      860      870      880      890
gi|113 RQAEAAQRQMAQQRQMQEQQRQMQE--QQRQVQEQQRQAMQRQQ-AEQQRHMQEQQR
      840      850      860      870      880      890

      1550      1560      1570      1580      1590      1600
frame4 .EKRRHQ.KCYRNRKHQQDP.SNGNQLQRYRNREEPQEGVHRYQH HQ.SKYHLPYGFWHP
      900      910      920      930      940
gi|113 AAQEQQRAMQEQQRHQQDQQRQMQERQMQEQQRQQQQQQQQQQQRAQN
      900      910      920      930      940

      1610      1620      1630      1640      1650      1660
frame4 .LA.PYQC..HLHLPKDR.YEFHR.RDLQT.RTCGR.TCHLEYGRFQKDQRTYEQLLDGC

>>gi|5737842|gb|AAD50121.1|AF153362_1 adenylyl cyclase [ (2123 aa)
      initn: 190 initl: 130 opt: 190 Z-score: 228.4 bits: 56.1 E(): 0.0012

```

```

Smith-Waterman score: 190; 16.236% identity (57.934% similar) in 271 aa overlap
(1210-1479:1824-2088)

      1180      1190      1200      1210      1220      1230
frame4 H.T.RTLSKPMIRTKALARGS.DLRIRYRARILELIKQP.CRRVRS.LQDQPSNP.TLGS
      1800      1810      1820      1830      1840      1850
gi|573 EALSYFQLCTEIMPNDKPTLIYIQRCIQNIKTLELQQIQLQLQRLQQLQQQQQLLLQQ
      1800      1810      1820      1830      1840      1850

      1240      1250      1260      1270      1280      1290
frame4 ..RS..HHQQ.QGFQKRD.P.RGSS.RYGDRGG.RQLLQKRYRDPYRQDDHARREKSHLHR
      1860      1870      1880      1890      1900      1910
gi|573 QLQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQSQNIQPPQS--QQ
      1860      1870      1880      1890      1900      1910

      1300      1310      1320      1330      1340      1350
frame4 NQH-R.A.DRLRQKQDGRFP.HGVLPNRS.R.HLQRMQLQQRELDTRRS.KEHGLLEQ.H
      1920      1930      1940      1950      1960      1970
gi|573 SQYVQQPQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQPQQQQQLQQQQQHQQQKQ
      1920      1930      1940      1950      1960      1970

      1360      1370      1380      1390      1400      1410
frame4 PSK.KNEHASQPRLLHQQDVG.SLRCRHPFPAESR.DQYGLGSSKR.GW.RRNLEQARQQ
      2000      2010      2020
gi|573 PSPQQQQQPQQPQQQQQQQIQNQYQHQLQYQRQQQQQQ---QQQQQQQQQQQQQQQQQ
      1980      1990      2000      2010      2020

      1420      1430      1440      1450      1460      1470
frame4 PMGKQ.RMDHLEHQSLDR.AYHDLQDGWYAHQRHQHSQR.G.HQNLEASFQCGHES.LAR.
      2030      2040      2050      2060      2070      2080
gi|573 QQQQQQQQQQQQQQQQQQQHHHHHHQQQQFQQQSQSQSQSQSQQQQQQQQQSQSQSQSQ
      2030      2040      2050      2060      2070      2080

      1480      1490      1500      1510      1520      1530
frame4 QW.YLGC.DQQAERTSLERKP.VPCR.LALESLVHAR.LEDDHRLQTSPAHPFREVGST
      2090      2100      2110      2120
gi|573 QIQKKSQHPSQQIQQSQRHQSQPQNVDTNVKTKPQQ
      2090      2100      2110      2120

>>gi|190654481|gb|EDV51724.1| GG13740 [Drosophila erecta (1835 aa)
      initn: 116 initl: 49 opt: 186 Z-score: 224.1 bits: 55.1 E(): 0.002
Smith-Waterman score: 186; 18.143% identity (51.188% similar) in 463 aa overlap
(1139-1582:230-681)

      1110      1120      1130      1140      1150      1160
frame4 KNK.Y.FENEQKDHIH.LFSIHFFHTVR.RKPNKKHSLKLQAQQMVQEKFSSQCHNTQTQ
      200      210      220      230      240      250
gi|190 KQATQQQQIQQQQLQHQQQQMQQQQQQQQQQQQQQQQQQQQQQQHQHQQIQQQATATS
      200      210      220      230      240      250

      1170      1180      1190      1200      1210      1220
frame4 .DSGVCAMKLMH.T.RTLS--KPMIRTKALARGS.DLRIRYRARILELIKQP.CRRV--R
      260      270      280      290      300      310
gi|190 ALSDILGFGEDSVAAKSIASIKSQLQASAEQIQQQQLPVPAAQQQVVFVRPQPQPQPAQQ
      260      270      280      290      300      310

      1230      1240      1250      1260      1270      1280
frame4 S.LQDQPSNP.TLGS..RS..HHQQ.QGFQKRD.P.RGS-S.RYGDRGG.RQLLQKRYRDP
      320      330      340      350      360      370
gi|190 SPHQLQGGGQPA--PTPATLPQHQQQLIIQQQKIISANLQQQTQNNQQQIKNILLQQRQMK
      320      330      340      350      360      370

```

```
>>gi|145020705|gb|EDK04834.1| hypothetical protein MGG_1 (704 aa)
  initn: 84 initl: 84 opt: 179 Z-score: 220.8 bits: 53.1 E(): 0.0031
Smith-Waterman score: 213; 20.822% identity (49.863% similar) in 365 aa overlap
(1275-1629:36-366)
```

```
>>gi|193893136|gb|EDV92002.1| GH24674 [Drosophila grimsh (883 aa)
  initn: 147 initl: 73 opt: 179 Z-score: 219.5 bits: 53.2 E(): 0.0036
Smith-Waterman score: 210; 16.438% identity (51.370% similar) in 584 aa overlap
(1061-1619:169-733)
```

```
>>gi|50417567|gb|AAH77588.1| K14 protein [Xenopus laevis (1320 aa)
  initn: 67 initl: 67 opt: 178 Z-score: 215.7 bits: 53.1 E(): 0.0059
Smith-Waterman score: 178; 19.165% identity (53.071% similar) in 407 aa overlap
  (1274-1660:446-832)
```

```

      1370      1380      1390      1400      1410      1420
frame4 KNEHASQPRLLHQQDVG.SLRCRHFFPAESR.DQYGLGSSKR.GW.RRNLEQARQQPMGK
      .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:.
gi|198 QQENQKEDQKEHQEQEQEQEHQEQ--KEHQEQEHQEQEHQEQEHQEQEQEQEHQEQDQLQ
      220      230      240      250      260

```

```

      1430      1440      1450      1460      1470
frame4 Q.RMDHLEHQS LDR.AYHDLQDGWY AHRQHQS---QR.G.HQNLEASFQCGHES.LAR.Q
      . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|198 DQLQDHLQEQ-LQDNQQENQDNQKEHQKQKEEHQQEYQQENQKEHQKEQPDQLQEQQ
      270      280      290      300      310      320

      1480      1490      1500      1510      1520      1530
frame4 W.YLGC.DQQAERTSLERKP.VPCR.LALES LV-HAR.LEDDHRLQTSPAHPFREVGST
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|198 QEYQQ-ENQKEHQQE HQKQKEEHQQEYQKENQKEHQKEHQKEHQ-QE QPDQLQEQQEQE
      330      340      350      360      370      380

      1540      1550      1560      1570      1580      1590
frame4 HDPLDA..EKR HQ.KCYRNRK--HQQDP.SNGNQ---LQRYRNREEPQEQGVHRYQH HQ
      . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|198 YQEQENQENQKEHQQE HQKQKEEQEQE-YQQGNQKEHQQEYQQENQKEHQQE HQKEHQ
      390      400      410      420      430      440

      1600      1610      1620      1630      1640      1650
frame4 .SKYHLPYGFWHP.LA.PYQC..HLHLPKDR.YEFHR.RDLQT.RTCGR.TCHLEYGRFQ

gi|198 EQEQEATILEEEDDMELVEVKLESQVEKKR FCLHTLCGLYDSDRRLSHFPWGIDVQRV
      450      460      470      480      490      500

>>gi|198150617|gb|EAL29745.2| GA14630 [Drosophila pseudo (917 aa)
      initn: 220 initl: 83 opt: 169 Z-score: 206.3 bits: 50.8 E(): 0.02
Smith-Waterman score: 177; 19.802% identity (50.495% similar) in 303 aa overlap
(1308-1600:487-769)

      1280      1290      1300      1310      1320      1330
frame4 KRYRDPYRQDDHARREKSHLHRNQHR.A.DRLRQOKDGRF---P.HGVLPNRS.R.HLQR
      . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|198 SAQH HQVLCCQNILESIGLEKLKVMHQNLKQEEKQARFFHTALTITLNERADDEHMFQ
      460      470      480      490      500      510

      1340      1350      1360      1370      1380
frame4 MQLQORE----LDTRRS.KEHG LLE--Q.HPSK.KNEHASQPRLLHQQDVG.SLRCRHP
      . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|198 YKMKQSKKAEDKLNADKEKQQRKLKADLQYLDVIRNRQRILSWRQHQQ---QLLLQNP
      520      530      540      550      560      570

      1390      1400      1410      1420      1430      1440
frame4 FPAESR.DQYGLGSSKR.GW.RRNLEQARQQPMGKQ.RMDHLEHQS LDR.AYHDLQDGWY
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|198 FHQQQPLQQQH HQPEQHQPQQHQPKHHQHQP EHHQHQP EHHQHQP--PEQH HQHQPEQH HQ
      580      590      600      610      620      630

      1450      1460      1470      1480      1490      1500
frame4 AHRQHQSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.DQQAERTSLERKP.VPCR.L
      . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|198 QPQQHQPPQ---HQPPQHQPQ-QHQPEQHQP EQ-HLPHQHQTEQQORDILYHYPQRPV--V
      640      650      660      670      680

      1510      1520      1530      1540      1550      1560
frame4 ALES LVHAR.LEDDHRLQTSPAHPFREVGSTHDPLDA..EKR HQ.KCYRNRKHQQDP.SN
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|198 VEES--HKQYHQKQQQQQLQHQK KQ---HQQEQLQQQKQQQQQQQKQQQKQKEH
      690      700      710      720      730

      1570      1580      1590      1600      1610      1620
frame4 GNQLQRYRNREEPQEQGVHRYQH HQ.SKYHLPYGFWHP.LA.PYQC..HLHLPKDR.YEF
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

gi|198 QQQQQKQQQHQQQHQQRVHK-QHHTALFEPPPIYPSSTS YRLPLNVATTAASGSEFQS
      740      750      760      770      780      790

      1630      1640      1650      1660      1670      1680
frame4 HR.RDLQT.RTCGR.TCHLEYGRFQKDQRTYEQLLDGCLHREACTPWKQKT.P.EDEAQS

gi|198 KLMLPLPLLRKPPFSSCGNAKKKTKTQDVEPREINERRGVAYPTPSVEAHYPPAAAATV
      800      810      820      830      840      850

>>gi|193907175|gb|EDW06042.1| GI16140 [Drosophila mojave (900 aa)
      initn: 112 initl: 77 opt: 167 Z-score: 203.8 bits: 50.3 E(): 0.027
Smith-Waterman score: 193; 19.037% identity (51.641% similar) in 457 aa overlap
(1066-1498:322-749)

      1040      1050      1060      1070      1080      1090
frame4 TQEDILVFSNICSF SRSENIFVA..ISIMVQLQVFYSSY.LRSLEATI.HIFTQNKMFAY
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|193 QEQQQQQLKEPQLQQQQLQQQPLQQQQH LQQQQQLQEQQQLQQQQH LQQQQQQYQY
      300      310      320      330      340      350

      1100      1110      1120      1130      1140      1150
frame4 LL.FQI QHTNKRKNK.Y.FENEQKDH I.H.LFSIH FHTVR.RKPNKKH SKLQAQMVQ
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|193 --QHQQHHQQQLQQQQQLQE EPQQQLQQQQQLQEQQQL-----QEQQQH DQLQVQQQLQ
      360      370      380      390      400

      1160      1170      1180      1190      1200      1210
frame4 EKQFSQCHNTQTQ.DSGVCAMKLMH.T.RTLSKPMIRT-KALARGS.DLRIRYARILEL
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|193 EQQQHHQLQLQQQQQE HQLAHQHYQQQQQLQDQKQLHQGESLSQSQPQPQKQLTAPTVMW
      410      420      430      440      450      460

      1220      1230      1240      1250      1260      1270
frame4 IK---QP.CRRVRS.LQDQPSNP.TLGS..RS..HHQQ.QGFQKRDP.RGSS.RYGDRG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|193 IPIAPVAPTSTQPYVLVRPIQYPKQL-QQSASTLPPTQSKVQLKRD-----QRTDHP
      470      480      490      500      510

      1280      1290      1300      1310      1320
frame4 G.RQLLQKRYRDPYRQDDHARREKSH--LHRNQHR.A.DRLRQOK--DGRFP.HGV--LP
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|193 NYQQ--QQQQQLPNKQPLHQQQQE QYQQQLQQQQQPQELQQLQQQQPLQQQLQQLQ
      520      530      540      550      560      570

      1330      1340      1350      1360      1370      1380
frame4 NRS.R.HLQRMQLQQRELDTRRS.KE-HGLLEQ.HPSK.KNEHASQPRLLHQQDVG.SLR
      . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|193 QPQPLQQLQQLQ-QQQQLQQQQPLQQLQQLQQQPHEQQQKNAYPQQP---HRFQLSRAFQ
      580      590      600      610      620

      1390      1400      1410      1420      1430
frame4 CR-HPF-----PAESR.DQYGLGSSKR.GW.RRNLEQARQQPMGKQ.RMD--HLEH
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|193 SRTHEFSHYQPDMPGYLGQNL YNEQAGGDIGQPSKSLAYKPKQENHDMMQPPPPHMQQ
      630      640      650      660      670      680

      1440      1450      1460      1470      1480      1490
frame4 QSLDR.AYHDLQDGWY AHRQHQSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.DQQA E
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|193 QQLQQ--HFQQQQHYQQQQQQQQQHYQQQQQLHQHFQQQQQQHYQQQQQ---HCQQQQHC
      690      700      710      720      730      740

```

```

      1500      1510      1520      1530      1540      1550
frame4 QRTSLERKP.VPCR.LALESLVHAR.LEDDHRLQTSAPHPFREVGSTHDPLDA..EKRHQ
: . . . .
gi|193 QQQQLQQYYQQQLQQHFQQQQQQPQQYEPDQARNHGLNSLAYKAKRNPVNMDTIYYTN
      750      760      770      780      790      800

>>gi|12383113|gb|AAG24941.2| unknown [Frankia sp. Ar15] (498 aa)
  initn: 108 init1: 108 opt: 163 Z-score: 202.2 bits: 49.2 E(): 0.033
Smith-Waterman score: 163; 26.697% identity (47.964% similar) in 221 aa overlap
(1252-1454:220-432)

      1230      1240      1250      1260      1270      1280
frame4 RVRS.LQDQPSNP.TLGS..RS..HHQQ.QGFQK-RDP.RGSS.RYGDRGG.RQLLQKRY
: . . . . :
gi|123 ARQKRRLEIAVGPLDDTLIFRVPWRRKPDLGQRQRSEPRRRRRHPAGPPDRFPPIPHQRL
      190      200      210      220      230      240

      1290      1300      1310      1320      1330
frame4 RDP---YRQDDHARREKSHLHRNQHR.A.DRLRQQKDGFRP.HGVLPNRS.R.H-LQRMQ
: : . . . . : . . . . : : : : : : : :
gi|123 RDPDPADQQPHARENVTLPRGNHRRGQEPGERQRHHQHRKHPFLP-RTHRNHRLREPQ
      250      260      270      280      290      300

      1340      1350      1360      1370      1380
frame4 LQQRELDTRRS.KEHGLEQ.HPSK.KNEHASQPRLLHQQ-----DVG.SLRCRHFPF--
. . . . : . . . . : . . . . : . . . . : . . . . :
gi|123 ITMRELP--RLI--HHPVHRVDPDILRPDH-RQSLQRRQIRPADPLRDHRRRHPRLR
      310      320      330      340      350      360

      1390      1400      1410      1420      1430      1440
frame4 ---AESR.DQYGLGSSKR.GW.RRNLEQARQ---PMGKQ.RMDHLEHQSLDR.AYHDLQ
: . . . . : . . . . : : : : : : : :
gi|123 QQRPDHRLERIDHRPARRPFLVLRPVRQRPHRVPRHTQPSSDLLDRNALSTQPADLR
      370      380      390      400      410      420

      1450      1460      1470      1480      1490      1500
frame4 DGWYAHRHQSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.DQQAQRTSLERKP.VP
: : . . . .
gi|123 P--VLHLQHPSYTSRGSVFAPWGVSFQEAATVVVDVDPDGVVDYQDVRGGVHAGGGVVAV
      430      440      450      460      470      480

>>gi|135224193|gb|EBG27286.1| hypothetical protein GOS_9 (423 aa)
  initn: 98 init1: 84 opt: 162 Z-score: 201.9 bits: 48.9 E(): 0.034
Smith-Waterman score: 179; 20.323% identity (47.344% similar) in 433 aa overlap
(1242-1664:11-413)

      1220      1230      1240      1250      1260      1270
frame4 LELIKQP.CRRVRS.LQDQPSNP.TLGS..RS..HHQQ.QGFQKRD.PRGSS.RY-GDRG
: . . . . : . . . . : . . . . :
gi|135 AARQTDRTQTDRTETSDSQRTYDRQTDRTQTDRTQTDRTQ
      10      20      30      40

      1280      1290      1300      1310      1320      1330
frame4 G.RQLLQKRYRDPYRQDDHARREKSHLHRNQHR.A.DRLRQQKDGFRP.HGVLPNRS.R.
: : . . . . : : . . . . : : . . . . :
gi|135 TDRQTDRTQTDRTQTD--RQTGRRADRTQGRQADRTQTERQDRQ-----TDRQTD
      50      60      70      80      90

      1340      1350      1360      1370      1380
frame4 HLQRMQLQQRELDTRRS.KEHGLEQ.HPSK.KNEH-ASQPRLLHQQDVG.SLRCRHFPF
. . . . : : . . . . : . . . . : . . . . :
gi|135 QTDRTQTDRTQRETD-RETDRQADRTQDRHTDTQTHRHTASQPG--RQTDITQDRHTDRTQ

```

```

      100      110      120      130      140
frame4 AESR.DQYGLGSSKR.GW.RRNLEQARQ---QPMGKQ.RMDHLEHQSLDR.AYHDLQDGWY
: . . . . : . . . . : . . . . : . . . . :
gi|135 TDRQTDRTQTDRTQTDRTQTDRTQTDRTQTDRTQTDRTQTDRTQTDRTQTDRTQTDRTQ
      150      160      170      180      190      200

      1450      1460      1470      1480      1490      1500
frame4 AHRQHQSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.DQQAQRTSLERKP.VPC
: : : : : : : : : : : : : : : :
gi|135 RQRQSQSQRQQRQR--QRQQRQQRQAGRQPDRTQTDRTQTDRTQTDRTQTDRTQTDRTQTDRTQ
      210      220      230      240      250      260

      1510      1520      1530      1540      1550      1560
frame4 R.LALESLVHAR.LEDDHRLQTSAPHPFREVGSTHDPLDA..EKRHQ.KCYRNRKHQQDP
: . . . . : . . . . : . . . . : . . . . :
gi|135 RQTDRTQTDRTQTDRTQTDRTQTDRTQTDRTQTDRTQTDRTQTDRTQTDRTQTDRTQTDRTQ
      270      280      290      300      310      320

      1570      1580      1590      1600      1610      1620
frame4 .SNGNQLQRYRNREEPQ---EQGVHRYQH HQ.SKYHLPYGFWHP.LA.PYQC..HLHLPKD
: : : : : : : : : : : : : : : :
gi|135 RQAGRQAGRQADRTQTHRRTDTQTHRRTDSQPASQPDRTQTDRTQTDRTQTDRTQTDRTQTDRTQ
      330      340      350      360      370

      1630      1640      1650      1660      1670      1680
frame4 R.YEFHR.RDLQT.RTCGR.TCHLEYGRFQKDQRT-YEQLLDGCLHREACTPWKQKT.P.
: . . . : : : : : : : : : : : :
gi|135 R--QTYRQTDRTQTDRTQTDRTQTDRTQTDRTQTDRTQTDRTQTDRTQTDRTQTDRTQTDRTQ
      380      390      400      410      420

      1690      1700      1710      1720      1730      1740
frame4 EDEAQSQLTSSSSSIPTTKSDKLVDAASSEKPIREICVDWISRDW.EMGSAHHCRF

>>gi|144191545|gb|EDJ53627.1| hypothetical protein GOS_1 (293 aa)
  initn: 41 init1: 41 opt: 160 Z-score: 201.6 bits: 48.3 E(): 0.036
Smith-Waterman score: 170; 24.609% identity (43.750% similar) in 256 aa overlap
(1221-1458:16-267)

      1200      1210      1220      1230      1240      1250
frame4 IRTKALARGS.DLRIYRARILELIKQP.CRRVRS.LQDQPSNP.TLGS..RS..HHQQ.
: : . . . . : : . . . . :
gi|144 VTQLRRQGRHRRRRRPRPLSCPGPRSPRREGRGQRPRFHGRH
      10      20      30      40

      1260      1270      1280      1290      1300
frame4 QGFQKRD.PRGSS.RYGDRGG.RQLLQKRYRDPYRQDDHARREK--SHLHRNQHR.A.DR
: . . . : : : : : : : . . . . : . . . . :
gi|144 RRLLRSGAGRGRDRGGGRHGHRRRQRLRREGREVGHRGHRERLGP RRHPDQRRHPAR
      50      60      70      80      90      100

      1310      1320      1330      1340      1350      1360
frame4 --LRQQKDGFRP.-HGVLPNRS.R.HLQRMQLQQR-ELDTRRS.K-EHGLEQ.HPSK.K
: : : : : : : : : : : : : : : :
gi|144 QVLRQGDPLPHGRGRAPHGLRHGEGSVADHARAELRPHRHDDLALGFLRLADQLR
      110      120      130      140      150      160

      1370      1380      1390      1400      1410
frame4 NEHASQPRLLHQ--QDVG.SLRCRH-----PFPAESR.DQYGLGSSKR.GW.RRNLEQ
: : : : : : : : : : : : : : : :
gi|144 RRQA-RPRGLHEHAQDRGREQHPHERHRPGRPHADREPHRGSGEPRP GAHHAGRDL

```



```

170      180      190      200      210      220
frame4  1420      1430      1440      1450      1460      1470
frame4  ARQQPMGKQ.RMDHLEHQSLDR.AYHD---LQDGWYAHRHQSQR.G.HQNLEASFQCGHE
gi|144   PVQRGRAERCHRAGRRRQLLRRRDRREPGRPGW---RGHGRRRRGSLVEDRRSRGRQAA
230      240      250      260      270      280

1480      1490      1500      1510      1520      1530
frame4  S.LAR.QW.YLGC.DQQAEQRTSLERKP.VPCR.LALESLVHAR.LEDDHRLQTSAPHPF
gi|144   QHAPARRLIACP
290

>>gi|116697810|gb|ABK16998.1| putative chromosome segreg (848 aa)
initn: 160 initl: 74 opt: 164 Z-score: 200.3 bits: 49.6 E(): 0.042
Smith-Waterman score: 219; 20.388% identity (50.728% similar) in 412 aa overlap
(1184-1587:477-847)

1160      1170      1180      1190      1200      1210
frame4  VQEKQFSQCHNTQTQ.DSGVCAMKLMH.T.RTLSPMIRTKALARGS.DLRIRY-RARIL
gi|116   ANNKQKYNFNKALVHNKPHNDIVRRINNNRQIAKQSQRNLNARTLENRTNRIRTAQAGKA
450      460      470      480      490      500

1220      1230      1240      1250      1260      1270
frame4  ELIKQP.CRRVRS.LQDQPSNP.TLGS..RS..HHQQ.QGFQKRDP.RGSS.RYGDRGG.
gi|116   GRIQQPKATN-RLVSANERNKPRTGTPGQQRNKGNVKSPPRQISMPGRQPTDRTGTREGR
510      520      530      540      550      560

1280      1290      1300      1310      1320
frame4  RQLLQKRYRD-----PYRQDDHARREKSHLHRNQHR.A.DRLRQKDGGRFP.HGVLPNRS
gi|116   QGIRSPDRDGRIGPATPQQRGDRQQQLQQQRQER---QQQMQQQRQ-----RG
570      580      590      600      610

1330      1340      1350      1360      1370      1380
frame4  .R.HLQRMQLQQRELDTRRS.KEHGLLEQ.HPSK.KNEHASQPRLLHQDDVG.SLRCRHP
gi|116   EQQQQRLQQQQRQQQRQ-----QQMEQQKORREQQQQRLQQQQRQQQQMQQQRQ
620      630      640      650      660

1390      1400      1410      1420      1430      1440
frame4  FPAESR.DQYGLGSSKR.GW.RRNLEQARQQPMGKQ.RMDHLEHQSLDR.AYHDLQDGWY
gi|116   -QQQQRQQQLQQQKQQQQQRQQQKMQQQRQ---QQQRQQQLQQQR--QQQQQRQQQM
670      680      690      700      710

1450      1460      1470      1480      1490      1500
frame4  AHRQHQSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.DQ--QAEQRTSLERKP.VPCR
gi|116   QQRQQQQQR---QQQLQQQRQQQRQQQQRQQQLQQQKQQQRQQQMRQQQQRQQQQR
720      730      740      750      760      770

1510      1520      1530      1540      1550      1560
frame4  .LAESLVHAR.LEDDHRLQTSAPHPFREVGSTHDPDA..EKRHQ.KCYRNRKHKQDP.
gi|116   QQQMQQQRQRQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQRAQQQQRQQQ---
780      790      800      810      820      830

1570      1580      1590      1600      1610      1620

```

```

frame4  SNGNQLQRYRNREEPQEQGVHRYQHHQ.SKYHLPYGFWHP.LA.PYQC..HLHLPKDR.Y
gi|116   ----LQQRRQQQQQQQQ--HRR
840

>>gi|137509825|gb|EBU48087.1| hypothetical protein GOS_7 (215 aa)
initn: 70 initl: 70 opt: 157 Z-score: 199.6 bits: 47.5 E(): 0.047
Smith-Waterman score: 157; 27.189% identity (48.387% similar) in 217 aa overlap
(1184-1393:10-209)

1160      1170      1180      1190      1200      1210
frame4  VQEKQFSQCHNTQTQ.DSGVCAMKLMH.T.RTLSPMIRTKALARGS.DLRIRYRARILE
gi|137   ERAAHGRRRLTLGQPAPRRSA--RGARPRAARRRRREGD
10      20      30

1220      1230      1240      1250      1260      1270
frame4  LIKQP.CRRVRS-.LQDQPSNP.TLGS..RS..HHQQ.QGFQKRDP.RGSS.RYGDRG-G
gi|137   RPTHGTCARRRGPARLHRPHR-----WQQRHRRPRLRALRARDRPRHARPHLPAGRG
40      50      60      70      80      90

1280      1290      1300      1310      1320
frame4  .RQLLQK-----RYRDPYRQDDHARREKSHLHRNQHR.A.DRLRQKDGGRFP.HGVLPNR
gi|137   ARHLLRRHPGGPRRRPRLRGARHARRRHQHLHEGRRAPARDRGDAGRRGAHA-HRAGPGH
100      110      120      130      140

1330      1340      1350      1360      1370      1380
frame4  S.R.HLQRMQLQQRELDTRRS.KEHGLLEQ.HPSK.KNEHASQPRLLHQDDVG.SLRCRH
gi|137   RARCGDRGRPPRRGL--RHAHGSHG--GGRAPRAHGGRAAGPGLLRQ---GHGRTPRH
150      160      170      180      190      200

1390      1400      1410      1420      1430      1440
frame4  PFPAESR.DQYGLGSSKR.GW.RRNLEQARQQPMGKQ.RMDHLEHQSLDR.AYHDLQDGW
gi|137   DPPRLRRGRRRG
210

>>gi|170945019|emb|CAP71130.1| unnamed protein product [ (1329 aa)
initn: 261 initl: 65 opt: 165 Z-score: 198.9 bits: 50.0 E(): 0.051
Smith-Waterman score: 199; 20.388% identity (50.485% similar) in 309 aa overlap
(1217-1507:963-1252)

1190      1200      1210      1220      1230      1240
frame4  SKPMIRTKALARGS.DLRIRYRARILELIKQP.CRRVRS.LQDQPSNP.TLGS..RS..H
gi|170   VKSEPQEHGHLPKVTSFSPSIAPAPAPSTSMQPPSKPQYHHLQPQPSSPHHQAQQHQAHP
940      950      960      970      980      990

1250      1260      1270      1280      1290
frame4  HQ--Q.QGFQKRDP.RGSS.RYGDRGG.-----RQLLQKRYRDPYRQD--DHARRE-KS
gi|170   HQPQQNQAAQQQQQQNRQSPQQNQGGQHAQQHQAQQHVQQHVQQPQQQQQQGRRQTQS
1000      1010      1020      1030      1040      1050

1300      1310      1320      1330      1340
frame4  HLHRNQHR.A.DRLR-----QKDGGRFP.HGVLPNRS.R.HLQRMQLQQRELDTRRS.K
gi|170   QLHQAQQQLQAQQYQAPQHQAQQQPAPQRQQTMPHQPTAQHQHAPQAQQAQQTQQAQQA
1060      1070      1080      1090      1100      1110

```

```

      1350      1360      1370      1380      1390      1400
frame4 EHGLLEQ.HPSK.KNEHASQPRLLHQQDVG.SLRCRHPFPAESR.DQYGLGSSKR.GW.R
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|170 QHQQAQQQLQ--QVPQHHQGGQQVQHQH-----VQHSRPAQQQQQQ-----AQQLQQVQ
      1120      1130      1140      1150

      1410      1420      1430      1440      1450      1460
frame4 RNLEQARQQPMGKQ.RMDHLEHQSLDR.AYHDLQDGWYAHRRHQHSQR.G.HQNLEASFQC
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|170 QNQQAQRQHSQQHVQQA--PQQLQQHQQQQHQQQQHQQQQHQQQQHQQQQH-----HQQQQQQQQH
      1160      1170      1180      1190      1200      1210

      1470      1480      1490      1500      1510      1520
frame4 GHES.LAR.QW.YLGC.DQQAQRTSLERKP.V-PCR.LALESLVHAR.LEDDHRLQTSP
      .. .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|170 QAQAPQPPQPPQQQTHHQHRAQQQQVQQQAQPQAQPHRHPSPTTSQPSAPPRPAQPAQT
      1220      1230      1240      1250      1260      1270

      1530      1540      1550      1560      1570      1580
frame4 AHPFREVGSTHDLDA..EKRHQ.KCYRNRKHQQDP.SNGNLQRYRNRREEPQEQGVHRY
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:

gi|170 AYQPQPQASTAQPSAPPAVPLPIPPNHHSTPDHQDQELLESLOAALAVQPN
      1280      1290      1300      1310      1320

>>gi|221488672|gb|EEE26886.1| hypothetical protein TGGT1 (1948 aa)
      initn: 148 init1: 62 opt: 165 Z-score: 196.6 bits: 50.1 E(): 0.069
Smith-Waterman score: 167; 23.077% identity (48.558% similar) in 416 aa overlap
(1214-1599:704-1093)

      1190      1200      1210      1220      1230      1240
frame4 RTLSKPMIRTKALARGS.DLRIRYRARILELIKQP.CRRVRS.LQDQPSNP.TLGS..RS
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|221 PQGSLVHQPLQQQPQGTVVHQLQLQPPQGSVVHQLQQQTQGSVLVHQLPQMPPQGNVVHQ
      680      690      700      710      720      730

      1250      1260      1270      1280      1290      1300
frame4 ..HHQQ.QGFQKRD.PRGSS.RYGDRGG.RQLLQKRYRDPYRQDDHARREKSHLHRNQHR
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|221 PLQQQQQGSVLH--HRPQHQPQGNV--VHQLQQQ-----QQQGSVLVHHRPQHQPQGN
      740      750      760      770      780

      1310      1320      1330      1340      1350      1360
frame4 .A.DRLRQQKDGRFP.HGVL---PNRS.R.HLQRMQLQRELDTRRS.KEHGLLEQ.HPS
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|221 VVHQLPQQQPQGSV-VHQLPQQPPQGSVLVHHRPQHQPQGSVVHQLPQQQPQGSVLVHHRPQ
      790      800      810      820      830      840

      1370      1380      1390      1400      1410
frame4 -K.KNEHASQPRLLHQQDVG.SLRCRHPFPAESR.DQYGLGSSKR.GW.RRNLEQ-ARQ-
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|221 HQPQGNVVHQP--LQQQPQGSVLVHHRPQHQPQGNVVHQLPQQQQQGSVLVHHRPQHQRQG
      850      860      870      880      890      900

      1420      1430      1440      1450      1460
frame4 ---QPMGKQ.RMD--H--LEHQSLDR.AYHDLQ--DGWYAH--QHQSQR.G.HQNLE
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|221 NVVYQPLQQQPQGNVVHQLPQQQPQGSVVYQPLQHQPPQGAIVHHRPQHQPQGSVVHQLPQ
      910      920      930      940      950      960

      1470      1480      1490      1500      1510
frame4 ASFCQG--HES.LAR.QW.YLGC.DQQAQRTSLERKP.VPCR.LALE----SLVHAR.L

```

```

      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|221 QQPQGNFVHQPLQQQQQ---GSLVHHLHHRPQHQPQGNVVVQPLQQQQPQGAIVHHR--
      970      980      990      1000      1010

      1520      1530      1540      1550      1560      1570
frame4 EDDHRLQTSPA--PFREVGSTHDLDA..EKRHQ.KCYRNRKHQ--QDP.SNGNLQRYR
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|221 -PQHQPQGNVVHQLPQ-----QPQGSVLVHHRPQHQPQGNVVHQLPQQQQQGS-LVHHR
      1020      1030      1040      1050      1060

      1580      1590      1600      1610      1620      1630
frame4 NREEPQEQGVHR--YQHHQ.SKYHLPYGFWHP.LA.PYQC..HLHLPKDR.YEFHR.RDL
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|221 PQHQPQGNVVHQLPQQQPQGSVVHQLPQQQPQGSFVHQLPQQQPQGSVLVHQLPQHQQAQGS
      1070      1080      1090      1100      1110      1120

>>gi|176558|gb|AAA35375.1| involucrin (large allele) [Ao (544 aa)
      initn: 78 init1: 48 opt: 159 Z-score: 196.5 bits: 48.3 E(): 0.069
Smith-Waterman score: 162; 18.689% identity (51.214% similar) in 412 aa overlap
(1213-1599:76-474)

      1190      1200      1210      1220      1230
frame4 .RTLSKPMIRTKALARGS.DLRIRYRARILELIKQP.CRRVRS.LQDQPS-----NP.T
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|176 EVPVELPVEGSPKHEEKHMTIVKGAPEQECEQQQPQEQKLQQQHWQDEEHQKAENPEQ
      50      60      70      80      90      100

      1240      1250      1260      1270      1280      1290
frame4 LGS..RS..HHQQ.QG-FQKRD.PRGSS.RYGDRGG.RQLLQKRYRD---PYRQDDHAR-
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|176 QLKQEKAKREKQQLQGQLEEEKLLDQPDHELAKSDEQLGTTKKEQLLEFPQEQEGQLKC
      110      120      130      140      150      160

      1300      1310      1320      1330      1340
frame4 --REKSHLHRNQHR.A.DRLRQQKDGRFP.HGVLPNRS.R.--HLQRMQLQRELDTRRS
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|176 LEQQEGHLELPEQQEGQLKCLEQQEG---HQELPEQQEGQLKHLEQQEGQLKHLEQQEG
      170      180      190      200      210      220

      1350      1360      1370      1380      1390      1400
frame4 .KEHGLLEQ.HPSK.KNEHASQPRLLHQQDVG.SLRCRHPFPAESR.DQYGLGSSKR.G
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|176 QVKH-LEQQEKQSELPEQQRGQPKYLEQQEGQLKHLEEQKGQLKHLEHQEGQLLEPQVG
      230      240      250      260      270      280

      1410      1420      1430      1440      1450      1460
frame4 W.RRNLEQARQQ---PMGKQ.RMDHLEHQSLDR.AYHDLQDGWYAHRRHQHSQR.G.HQNL
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|176 QPK-HLEQLEKLEHPEQQEGQLKQLEEQE-GQVKHLEQQEEQLKHLEQQEGQPKHPEQL
      290      300      310      320      330

      1470      1480      1490      1500      1510
frame4 EASF-----QCGHES.LAR.QW.YLGC.DQQAQRTSLERKP.VPCR.LALESLVHAR.L
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|176 EKQLEHPEQQEGQLKQLEE-QEGQVKHLEQQEEQLKHLEQQEGQPKHLEQLE-----KQL
      340      350      360      370      380      390

      1520      1530      1540      1550      1560      1570
frame4 EDDHRLQTSPAHPFREVGSTHDLDA..EKRHQ.KCYRNRKHQQDP.SNGNLQRYRNR
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|176 EHLEQQEGQLKHLEQREEQLELPEQQVQSKHLEQEEQLEHPEQQEGQLKHLEQEAQL
      400      410      420      430      440      450

```

```

1580      1590      1600      1610      1620      1630
frame4 E-PQEQGVHRYQHHQ.SKYHLPYGFWHP.LA.PYQC..HLHLPKDR.YEFHR.RDLQT.R
      : . . : . . : . . : . . : . . : . . : . . : . . : . . : . . :
gi|176 ELPEQVGQPKHLEQQEKQLEHPEQQEQGQLKPQEQEQGQLKGLEQQERQLEQPVFAPAPGQ
      460      470      480      490      500      510

>>gi|148498882|gb|ABQ67136.1| short-chain dehydrogenase/ (553 aa)
initn: 129 initl: 84 opt: 158 Z-score: 195.1 bits: 48.0 E(): 0.082
Smith-Waterman score: 158; 30.070% identity (48.951% similar) in 143 aa overlap
(1231-1370:21-158)

      1210      1220      1230      1240      1250
frame4 .DLRIRYRARILELIKQP.CRRVRS.LQDQPSNP.-TLGS..RS..HHQQ.QGFQKRPD.
      : . . : . . : . . : . . : . . : . . : . . : . . : . . : . . :
gi|148 MTGNTYARRLRRPEWFPLFAPPRPLPSLGPDRSGRKRDDGAPEGQDRDRH
      10      20      30      40      50

      1260      1270      1280      1290      1300      1310
frame4 RGSS.RYGDREGG.RQLLQKRYRDPYRQDDHARREKSHLHRNQHR.A.D--RLRQQKDGFRF
      : . . : . . : . . : . . : . . : . . : . . : . . : . . : . . :
gi|148 R---RRVGHRPGDGADLRPRGR-PRRRRRH-RRGRRRAHRRPDRRAGDGDRLRRRRRSRL
      60      70      80      90      100

      1320      1330      1340      1350      1360      1370
frame4 P.HGVLPNRS.R.HLQRMQLQQRELDTRRS.KEHGLLEQ.HPSK.KNEHASQPRLLHQQD
      : . . : . . : . . : . . : . . : . . : . . : . . : . . : . . :
gi|148 GRRRAAHGRSLRPARHPRQYRRDRLSGHDPRPDHGPVEQDDRGQPDRRHARLPGGDPRDH
      110      120      130      140      150      160

      1380      1390      1400      1410      1420      1430
frame4 VG.SLRCRHPPAESR.DQYGLGSSKR.GW.RRNLEQARQQPMGKQ.RMDHLEHQSLDR.
      : . . : . . : . . : . . : . . : . . : . . : . . : . . : . . :
gi|148 RQRQRDRDRQLLARRAGRHLLRRRLRLRVEGRDDAQQVGRAVLRRARPADPLRLGPSHL
      170      180      190      200      210      220

>>gi|158601726|gb|EDP38557.1| Low-density lipoprotein re (923 aa)
initn: 186 initl: 76 opt: 160 Z-score: 194.6 bits: 48.7 E(): 0.088
Smith-Waterman score: 194; 20.465% identity (49.535% similar) in 430 aa overlap
(1242-1660:416-799)

      1220      1230      1240      1250      1260      1270
frame4 LELIKQP.CRRVRS.LQDQPSNP.TLGS..RS..HHQQ.QGFQKRPD.RGSS.RYGDREGG
      : . . : . . : . . : . . : . . : . . : . . : . . : . . : . . :
gi|158 ESTRKEEHERQHEEEERLRQEQRQEHERRREEHHQQ---EERERLRQKQEQQEEQ--
      390      400      410      420      430

      1280      1290      1300      1310      1320      1330
frame4 .RQLLQKRYRDPYRQDDHARREKSHLHRNQHR.A.DRLRQQKDGFRF.HGVLPNRS.R.H
      : . . : . . : . . : . . : . . : . . : . . : . . : . . : . . :
gi|158 ----HERLRQEQRQEHERRREEHHQQEER---ERLRQKQEQQEQRHERLRQEQRQE
      440      450      460      470      480      490

      1340      1350      1360      1370      1380
frame4 LQRMQLQQRELDTRRS.KEHGLLEQ.HPSK.KNEH-----ASQPRLLHQQDVG.S
      : . . : . . : . . : . . : . . : . . : . . : . . : . . : . . :
gi|158 HERRREEERE--RLKQKEHQEILRQEQRERERQEHERRREEERERAEQERI--RQEYYES
      500      510      520      530      540

      1390      1400      1410      1420      1430      1440
frame4 LRCRHPPAESR.DQYGLGSSKR.GW.RRN-LEQARQQPMGKQ.RMDHLEHQSLDR.AYH
      : . . : . . : . . : . . : . . : . . : . . : . . : . . : . . :

```

```

gi|158      ERQKQEIQQRRKGEEYEEGQEGRQGDERSRQEQERIRQSEIIRREQKE--KIEEDEGH
          550          560          570          580          590          600

          1450      1460      1470      1480      1490      1500
frame4      DLQDGWYAHRRHQHSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.DQQAQEQRTSLERKP
          .. : . . . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|158      EIFD----KKQESYEEEEQHQR----QPEEEEELERKRKE-----EEEKEDRRRLLEDR
          610          620          630          640          650

          1510      1520      1530      1540      1550      1560
frame4      .VPCR.LALESLVHAR.LEDDHRLQTSPAHPFREVGGSTHDPLDA..EKRRHQ.KCYRNRKH
          : : . : : : : : : : : : : : : : : : : : : : : : : : : :
gi|158      ---RRLEEDR---RRLEEDHRRLEELSRTDQEMKPEKERGKDEYEK-HQLT-IDGVRI
          660          670          680          690          700

          1570      1580      1590      1600      1610      1620
frame4      QQDP.SNGNQLQRYRNRREEPQEQGVHRYQHHQ.SKYHLPGYFWHP.LA.PYQC..HLHLF
          .. : . . . . . : : : : : : : : : : : : : : : : : : : : : :
gi|158      KEDEQRRRLELEHYGLDKNNQEEDVEKHERWESTRKE-EHERQHEEEERLQQEQERQEHE
          710          720          730          740          750          760

          1630      1640      1650      1660      1670      1680
frame4      KDR.YEFHR.RDLQT.RTCGR.TCHLEYGRFQKDQRTYEQLLDGLCHREACTPWQKQT.P
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|158      RRREEHHQQEERERLRQ-QKEQEEEEHERLRQEQERQEHERRREERERLQKEQHEQEI
          770          780          790          800          810

          1690      1700      1710      1720      1730      1740
frame4      .EDEAQSQLTSSHSSSIPTTTSKDKLVDAASSEKPIREICVDWISRDRW.EMGSAHCR
          820          830          840          850          860          870

gi|158      LRQEQERERQEHERRREEEEQRQREEXRERAEQERIKQEYYERERQKQEIQQRRKGEEYE
          820          830          840          850          860          870

>>gi|124415991|emb|CAK81050.1| unnamed protein product [ (504 aa)
      initn: 161 init1: 74 opt: 157 Z-score: 194.4 bits: 47.8 E(): 0.0
      Smith-Waterman score: 158; 17.647% identity (51.662% similar) in 35
      (1246-1619:106-477)

          1220      1230      1240      1250      1260
frame4      KQP.CRRVRS.LQDQPSNP.TLGS..RS..HHQQ.QGFQKRDP.RGS--S.RYG---DR
          .. : . . . . . : : : : : : : : : : : : : : : : : : : : : :
gi|124      DGEWITNENEKSQSKSNKYRKNSNKDQDKNNRKQSDQYQKQQEQRNTKNNERYGYNQKF
          80          90          100          110          120          130

          1270      1280      1290      1300      1310      1320
frame4      GG.RQLLQKRYRDPYR-QDDHARREKSHLH----RNQHR.A.DRLRQQKDGRFP.HGVLP
          .. : . . . . . : : : : : : : : : : : : : : : : : : : : : :
gi|124      NSSQNNYNNRNYRDDARNQDDDDYDQDQDNQNYRTDRGQQRYQKNTYNEGWNQRRNTNQYQY
          140          150          160          170          180          190

          1330      1340      1350      1360      1370      1380
frame4      NRS.R.HLQRMQLQORELDTRRS.KEHGLLEQ.HPSK.KNEHASQPRLLHQQDVG.SLRC
          : : . . . : : : : : : : : : : : : : : : : : : : : : : :
gi|124      NNNNNKYNQRRQQRNRNDDVFQ--QPQYVLKDQQAASE---QQQSQDQAQKQ--VSDSVN-
          200          210          220          230          240

          1390      1400      1410      1420      1430      1440
frame4      RHPFFAESR.DQYGLGSSKR.GW.RRNLEQARQCFPMGKQ.RMDHLEHQSLDR.AYHDLQD
          .. : . . . . . : : : : : : : : : : : : : : : : : : : : : :
gi|124      QQVQDSQKRKQEQDQKSSNQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQ
          250          260          270          280          290          300

```

```
>>gi|193919917|gb|EDW18784.1| GI11885 [Drosophila mojavae (967 aa)
  initn: 109 initl: 76 opt: 160 Z-score: 194.3 bits: 48.7 E(): 0.091
Smith-Waterman score: 160; 20.924% identity (48.370% similar) in 368 aa overlap
(1231-1584:24-376)
```

```
>>gi|193907429|gb|EDW06296.1| GI21598 [Drosophila mojavae (532 aa)
  initn: 198 initl: 84 opt: 157 Z-score: 194.1 bits: 47.8 E(): 0.094
Smith-Waterman score: 206; 18.465% identity (53.47% similar) in 417 aa overlap
(1189-1595:102-500)
```

```

1570      1580      1590      1600      1610      1620
frame4 NQLQRYRNREEPQEGVHRYQHHQ.SKYHLPGFWHP.LA.PYQC..HLHLPKDR.YEFH
.: : .....: : : : :
gi|193 QQQQLMHSQQQPHQQQLQQQQQKQPQQQQCLPIAQQVLSRFLAELFDELLLCAE
480      490      500      510      520      530

>>gi|218675676|gb|AAI69292.2| repetin [synthetic constru (383 aa)
initn: 103 init1: 65 opt: 155 Z-score: 193.5 bits: 47.2 E(): 0.1
Smith-Waterman score: 175; 19.693% identity (47.570% similar) in 391 aa overlap
(1228-1597:10-360)

1200      1210      1220      1230      1240      1250
frame4 RGS.DLRIRYRARILELIKQP.CRRVRS.LQDQPSNP.TLGS..RS..HHQQ.QGFQK--
.: : : : : : : : : :
gi|218 SPHRGQKGRQDQSPHGQKGRQDQSS-HQGQREGQDQNS
10      20      30

1260      1270      1280      1290      1300
frame4 ---RDP.RGSS.RYGDRGG.RQLLKRYRDPYRQDDHARREKSH---LHRNQ-----
: : : : : : : : : :
gi|218 HWHRTDRQGQSFHYGQTGGQGLSSHQGGTDSQGQNSQWHRDTSQGQSFHFDQAGREGQSS
40      50      60      70      80      90

1310      1320      1330      1340      1350
frame4 HR.A.DRLRQKDGRRFP.HGVLPNRS.R.H---LQRMQLQQRELDTRRS.K--EHGLLEQ
.: : : : : : : : : :
gi|218 HHGQTDTRQSQSSHCQSEIGKTENQGQNRHSLGTDTRRDSYVEQSGRSVKLSQQNSREE
100      110      120      130      140      150

1360      1370      1380      1390      1400      1410
frame4 .HPSK.KNEHASQPRLHQQDVG.SLRCRHPFPAESR.DQYGLGSSKR.GW.RRNLEQAR
.: : : : : : : : : :
gi|218 VRQTQSQRSHDRREQQIQQTWKPKEDNQHKLLAQVQQEPEY---SYEEYDWQSQSSEQ--
160      170      180      190      200      210

1420      1430      1440      1450      1460      1470
frame4 QQPMGKQ.RMDHLEHQSLDR.AYHDLQDGWYAHRRHQHSQR.G.HQNLEASFQCGHES.LA
.: : : : : : : : : :
gi|218 -DHCGEERYQDWDRHSVEDQENLYEMQN-WQTHEEEQSHQTSDRQ-----
220      230      240      250

1480      1490      1500      1510      1520      1530
frame4 R.QW.YLGC.DQAEQRTSLERKP.VPCR.LALESLVHAR.LEDDHRLQTSPAHPFREVG
.: : : : : : : : : :
gi|218 ----THVDEQNQRQHRQTHEEN-----HDHQHGRHHEDEHNHRRQDHHQQRER-
260      270      280      290      300

1540      1550      1560      1570      1580      1590
frame4 STHDPLDA..EKRRHQ.KCYRNRKHQQDP.SNGNQLQRYRNREEPQEGVHRYQHHQ.SKY
.: : : : : : : : : :
gi|218 QTHEEKEKYGGQDQSRSPFNREKSH--MSEDDQCEGPQGRFHPHTGGGKSQRREKSGN
310      320      330      340      350

1600      1610      1620      1630      1640      1650
frame4 HLPYGFWHP.LA.PYQC..HLHLPKDR.YEFHR.RDLQT.RTCGR.TCHLEYGRFQKDQR
:
gi|218 HPTKPANYSSPLYDYVQEQAAYQY
360      370      380

>>gi|50057436|emb|CAH03420.1| hypothetical protein [Para (2301 aa)
initn: 199 init1: 103 opt: 163 Z-score: 193.0 bits: 49.7 E(): 0.11

```

Smith-Waterman score: 172; 18.912% identity (48.964% similar) in 386 aa overlap (1228-1596:1448-1812)

```

1200      1210      1220      1230      1240      1250
frame4 RGS.DLRIRYRARILELIKQP.CRRVRS.LQDQPSNP.TLGS..RS..HHQQ.QGFQKRD
.: : : : : : : : : :
gi|500 IDVQDQNHKEQVQQQDQSHQEQQDQYQDHEEQDEDHQEIQIEDHQNQDEVHQEQIEDHQEQQD
1420      1430      1440      1450      1460      1470

1260      1270      1280      1290      1300      1310
frame4 P.RGSS.RYGDRGG.RQLLKRYRDPYRQD-DHARREKSHLHRNQHR.A.DRLRQKDGRR
.: : : : : : : : : :
gi|500 ---DNHQKQDDDH-----QKQDEDDQQQDLDHQKQDQDHQKPNEDQKQGENHQKQDGD
1480      1490      1500      1510      1520

1320      1330      1340      1350      1360      1370
frame4 FP.HGVLPNRS.R.HLQRMQLQQRELDTRRS.KEHGLLEQ.HPS--K.KNEHASQPRLH
.: : : : : : : : : :
gi|500 HQEQNEEYQEQQEDHKEQIENQQ-EQDGNHQQQEDGNHQQQEESLQEQQDGDHQEQQEQIHQ
1530      1540      1550      1560      1570      1580

1380      1390      1400      1410      1420      1430
frame4 QQDVG.SLRCR-HPFPAESR.DQYGLGSSKR.GW.RRNLEQARQQPMGKQ.RMDHLEHQS
.: : : : : : : : : :
gi|500 EQDRDHQEQQEDHQEQQDETPQQQDNEHQEQVETPLQQDNDHQEQQDETPQQDNDHQEQQE-
1590      1600      1610      1620      1630      1640

1440      1450      1460      1470      1480      1490
frame4 LDR.AYHDLQDGWYAHRRHQHSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.DQAEQR
.: : : : : : : : : :
gi|500 -ENQEQEENQDKNHQEQQHQ-EYDLDHQENDQIHQEQQEEDNLKKD-----LDHQAQDK
1650      1660      1670      1680      1690

1500      1510      1520      1530      1540
frame4 TSLERKP.VPCR.LALESLVHAR.LEDD-HRLQTSPAHPFREVGSTH--DPLDA.----
.: : : : : : : : : :
gi|500 SLQDQDQVDHQEQQDLQDQVENYQREDENHQEQQDQVQLEVQVQVQQQDEDHQDEKEVIQL
1700      1710      1720      1730      1740      1750

1550      1560      1570      1580      1590      1600
frame4 -----EKRRHQ.KCYRNRKHQQDP.SNGNQLQRYRNREEPQEGVHRYQHHQ.SKYHLPY
.: : : : : : : : : :
gi|500 KEEDNFEKKEQFQDY-DEINYQDSKCNNNDENQQHDDENLNDQDKNEFKDDDSQKHIHQE
1760      1770      1780      1790      1800      1810

1610      1620      1630      1640      1650      1660
frame4 GFWHP.LA.PYQC..HLHLPKDR.YEFHR.RDLQT.RTCGR.TCHLEYGRFQKDQRTYEQ

```

```

>>gi|124390320|emb|CAK55862.1| unnamed protein product [ (2301 aa)
initn: 199 init1: 103 opt: 163 Z-score: 193.0 bits: 49.7 E(): 0.11
Smith-Waterman score: 172; 18.912% identity (48.964% similar) in 386 aa overlap
(1228-1596:1448-1812)

```

```

1200      1210      1220      1230      1240      1250
frame4 RGS.DLRIRYRARILELIKQP.CRRVRS.LQDQPSNP.TLGS..RS..HHQQ.QGFQKRD
.: : : : : : : : : :
gi|124 IDVQDQNHKEQVQQQDQSHQEQQDQYQDHEEQDEDHQEIQIEDHQNQDEVHQEQIEDHQEQQD
1420      1430      1440      1450      1460      1470

```

```

1260      1270      1280      1290      1300      1310
frame4 P.RGSS.RYGD RGG.RQLLQKRYRDPYRQD-DHARREKSHLHRNQHR.A.DRLRQQKDGR
      . . . . .
gi|124 ---DNHQKQDDDH-----QKQDEDDQQQDLHDHQQDQDQDHQKPNEDQKQQGENHQKQDGD
      1480      1490      1500      1510      1520

1320      1330      1340      1350      1360      1370
frame4 FP.HGVLPNRS.R.HLQRMQLQQRELDTRRS.KEHGLLEQ.HPS--K.KNEHASQPRLLH
      . . . . .
gi|124 HQEQNEEYQEQQEDDHKEQIENQQ-EQDGNHQQQEDGNHQQQEESLQEQDGDHQEQEQIHQ
      1530      1540      1550      1560      1570      1580

1380      1390      1400      1410      1420      1430
frame4 QQDVG.SLRCR-HPFPAESR.DQYGLGSSKR.GW.RRNLEQARQQPMGKQ.RMDHLEHQS
      . . . . .
gi|124 EQDRDHQEQQEDDHQEQQDETTPQQDNEHQEQVETPLQQDNDHQEQDETPLQQDNDHQEQE-
      1590      1600      1610      1620      1630      1640

1440      1450      1460      1470      1480      1490
frame4 LDR.AYHDLQDGWYAHQRHQSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.DQQAQQR
      . . . . .
gi|124 -ENQEQQEENQDKNHQEQQGHQ-EYDLDHQENDQIHQEQQEEDNLKKD-----LDHQAQDK
      1650      1660      1670      1680      1690

1500      1510      1520      1530      1540
frame4 TSLERKP.VPCR.LALESVHAR.LEDD-HRLQTSPAHPFREVVGSTH---DPLDA..----
      . . . . .
gi|124 SLQDQDQVDHQEQLDLYQDQVENYQREDENHQEQDQVQLEVEVQEVQQQEQEDHQDEKEVIQL
      1700      1710      1720      1730      1740      1750

1550      1560      1570      1580      1590      1600
frame4 -----EKRHQ.KCYRNRKHQQDP.SNGNQLQRYRNREEPQEQGVHRYQHHQ.SKYHLPY
      . . . . .
gi|124 KEEDNFEKKEQFDY-DEINYSKCNNDENQOHDDENLNDQDKNEFKDDDSQKHIHQE
      1760      1770      1780      1790      1800      1810

1610      1620      1630      1640      1650      1660
frame4 GFWHP.LA.PYQC..HLHLPKDR.YEFHR.RDLQT.RTCGR.TCHLEYGRFQKDQRTYEQ

gi|124 ENQIKDTSQKHSVKSQKSQKSIPKNTQENVASHQNESEVDFNCNDIPSRYTQELEKEI
      1820      1830      1840      1850      1860      1870

>>gi|76580731|gb|ABA50206.1| adenosine deaminase [Burkho (726 aa)
      initn: 59 initl: 59 opt: 157 Z-score: 192.2 bits: 47.9 E(): 0.12
Smith-Waterman score: 157; 22.986% identity (44.550% similar) in 422 aa overlap
(1218-1619:3-406)

1190      1200      1210      1220      1230      1240
frame4 KPMIRTKALARGS.DLRIRYRARILELIKQP.CRRVRS.LQDQPSNP.T---LGS..RS
      . . . . .
gi|765 MPPPPRRARAARQPPARPSPRSDRSGETPDA
      10      20      30

1250      1260      1270      1280      1290      1300
frame4 ..HHQQ.QGFQKRDP.RGSS.RYGD RGG.RQLLQKRYRDPYRQDDHARREKSHLHRNQHR
      . . . . .
gi|765 NLAHRPAASARPR--RRGRS-RHGR--ARRRLRAACR--HEDARHARRRAPHDRRRPSR
      40      50      60      70      80

1310      1320      1330      1340      1350      1360
frame4 .A.DRLRQQ-KDGRFP.HGVLPNRS.R.HLQRMQLQQRELDTRRS.KEHGLLEQ.HPSK.
      . . . . .

```

```

gi|765 MEGDRHRTAIAEGRRAYAARAPARAARARPEPRPVLRRRGRARVRAARHRRARLGHIA-
      90      100      110      120      130      140

1370      1380      1390      1400      1410
frame4 KNEHASQPRLLHQQDVG.SLRCRHPFPAESR.DQYG----LGSSKR.GW.RRNLEQARQ
      . . . . .
gi|765 --EARGRGRADRAQRV-VRPCRHRGDAERARAGRDAFRLPAVGHGRRAPADRDD-RAARV
      150      160      170      180      190      200

1420      1430      1440      1450      1460      1470
frame4 QPMGKQ.RMDHLEHQS LDR.AYHDLQDGWYAHQRHQSQR--G.HQNLEASFQCGHES.L
      . . . . .
gi|765 PGRAVRRRRARGRGARARARHAALPRALGGRARRAVSAARGARGHPQSVARRERRARRSHR
      210      220      230      240      250      260

1480      1490      1500      1510      1520      1530
frame4 AR.QW.YLGC.DQQAQEQRTSLERKP.VPCR.LALESVHAR.LEDDHRLQTSPAHPFREV
      . . . . .
gi|765 RSRPRCVLRRDDAQSRPR-SRARAPDSRARRLCV--LRHDLRAHEAHAVR-APARRARHR
      270      280      290      300      310

1540      1550      1560      1570      1580
frame4 GSTHDP L D A . . EKR--HQ.KCYR---NRKHQQDP.SNGNQL-QRYRNREEPQEQGVHRYQ
      . . . . .
gi|765 -SAADRADAMPDRRARHRRQGARGDRDRRRRANPADRRRAARERASGRVARRERV LNRH
      320      330      340      350      360      370

1590      1600      1610      1620      1630      1640
frame4 H---HQ.SKYHLPYGFWHP.LA.PYQC..HLHLPKDR.YEFHR.RDLQT.RTCGR.TCHLE
      . . . . .
gi|765 PNPBSTQKTTMTPTF-KDKIARAPKAELHIHIEGSLEPELIFELAQRNGVKLAYESIDAL
      380      390      400      410      420      430

1650      1660      1670      1680      1690      1700
frame4 YGRFQKDQRTYEQLLDGCLHREACTPWQKT.P.EDEAQSQLTSSSSSIPTTTKSDKLV

gi|765 RAAYFTDLQSF LDIYYAGASVLLTEQDFYDMTAA YVERALADHVVAEIFFDPQTHTER
      440      450      460      470      480      490

```

```

>>gi|190623993|gb|EDV39517.1| GF24442 [Drosophila ananas (734 aa)
      initn: 77 initl: 46 opt: 157 Z-score: 192.1 bits: 47.9 E(): 0.12
Smith-Waterman score: 176; 24.194% identity (52.016% similar) in 248 aa overlap
(1218-1458:263-489)

```

```

1190      1200      1210      1220      1230      1240
frame4 KPMIRTKALARGS.DLRIRYRARILELIKQP.CRRVRS.LQDQPSNP.TLGS..RS..HH
      . . . . .
gi|190 PPSFRRSKMKTQRSGNDSGLGSPSMGHAYSDEPELRTMGRGYDDREFGHRSSGGNRNRQ-QH
      240      250      260      270      280      290

1250      1260      1270      1280      1290      1300
frame4 QQ.QGFQKRDP.RGSS.RYGD RGG.RQLLQKRYRDPYRQDDHARREKSHLHRNQHR.A.D
      . . . . .
gi|190 QQNQHHNPNYHDHRQRHTIYGDEY-ETQTLNRSHRYSSRQ-----RQRDRM-RDRHTVERE
      300      310      320      330      340

1310      1320      1330      1340      1350      1360
frame4 RL-RQQKD-GRFP.HGVLPNRS.R.HLQRMQLQQRELDTRRS.KEHGLLEQ.HPSK.KNE
      . . . . .
gi|190 RYPRSHLDAGSIEDFSDVPPPTKRAHSVRSVRSPhNQDSSRASKE---LQRLHSAFGRSR
      350      360      370      380      390      400

```

```

      1370      1380      1390      1400      1410      1420
frame4 HAS-QPRLL--HQQDVG.SLRCRHPF-PAESR.DQYGLGSSKR.GW.RRNLEQARQQPMG
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 AKSVDPRHVSSHIEDVDEVDVRKTP.IIPNRYAMDDFG-GNGHR----RQSAPRSQSM--
      410      420      430      440      450

      1430      1440      1450      1460      1470      1480
frame4 KQ.RMDHLEHQSLDR.AYHDLQDGWY-AHRQHQSQR.G.HQNLEASFQCGHES.LAR.QW
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 ---RSAHQRRHQDRERERSPPPPHNSHRQNNGRAGSLGRQTSRYGNHLELPDYDAPP
      460      470      480      490      500      510

      1490      1500      1510      1520      1530      1540
frame4 .YLGC.DQQAQRTSLERKP.VPCR.LALESLVHAR.LEDDHRLQTSAPHPFREVGSTHD

gi|190 RGRDRRRDHSGHSQGRYDDYVEESFDEGSLYGEHYEDYPPQRSRSRSREPKRERRRRRS
      520      530      540      550      560      570

>>gi|158271402|gb|EDO97222.1| RWP-RK transcription facto (2146 aa)
  initn: 122 initl: 55 opt: 162 Z-score: 192.1 bits: 49.4 E(): 0.12
Smith-Waterman score: 168; 21.171% identity (48.198% similar) in 444 aa overlap
(1186-1601:576-997)

      1160      1170      1180      1190      1200      1210
frame4 EKQFSQCHNTQTQ.DSGVCAMKLMH.T.RTLSKPMIRTKALARGS.DLRIRYRARILELI
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|158 PPAEPHPFFSPAAAAAAGPGIESSLGLGSSVPRSGALEHGVEDWEWRW----EEG
      550      560      570      580      590      600

      1220      1230      1240      1250      1260      1270
frame4 KQP.CRRVRS.LQDQPSNP.TLGS..RS..HHQQ.QGFQK---RDP.RGSS.RYGDRGG
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|158 QQ---RHQRALLQQQQQQEYMPAPHHMPHEQQQRHQQLLQSQSQSQSQSQSQSQSQ
      610      620      630      640      650

      1280      1290      1300      1310      1320
frame4 .RQLLQKRYRDPYRQDDHARREKSHLHRNQHR.A.DRLR---QQKDGRF--P.HGVLPN
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|158 QRHDLHREQLPLPPQPEALSRQQQEQRQQEHFELQQLQLLQQQQQQSQQLQQPQHLMPS
      660      670      680      690      700      710

      1330      1340      1350      1360      1370
frame4 RS.--R.HLQRMQLQORELDT-RRS.KEHGLEQ.HPSK.KNEHASQPRLL---HQQDV
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|158 SSSFQRDWIPPPQPDQPPLPSWRQSEPLHA----EPQPRASRSQQQPQQLQSPGHVQQ--
      720      730      740      750      760      770

      1380      1390      1400      1410      1420      1430
frame4 G.SLRCRHPFPAESR.DQYGLGSSKR.GW.--RRNLEQARQQPMGKQ.RMDHLEHQSLDR
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|158 GQPLQHGGASGAPTYPYRGGGFGESELDGGWHWQQQQQQQQQQQLLQQQQAAASLSHQLQER
      780      790      800      810      820      830

      1440      1450      1460      1470      1480      1490
frame4 .AYHDLQDGWYAHRHQ---QSQR.G.--HQNLEASFQCGHES-.LAR.QW.YLGC.DQQA
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|158 EQ-QELQRQQAARHHMQLQQQQGAPPSSSWHGQYPPLHQRPDQAHQYPHL--QQPQE
      840      850      860      870      880

      1500      1510      1520      1530      1540      1550
frame4 EQRTSLERKP.VPCR.LALESLVHAR.LEDDHRLQTSAPHPFREVGSTHDPDLD..EKRH
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

gi|158 HQHPQQERYP--PLHQQTQQQL---QQQQQQPQLQQQQQAQWPGPGSAPPSTSAGHPTLQ
      890      900      910      920      930      940

      1560      1570      1580      1590      1600
frame4 Q.---KCYRNRKHQQDP.SNGNLQRYRNRREEPQEQGVHRYQHHQ.SKYHLPYGFWHP.L
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|158 QQLYPFMYQQQTQQQPQPQQQNAQQPAYHHHPLLSP-HSPQYQRMQASAPVSGGGAHR
      950      960      970      980      990      1000

      1610      1620      1630      1640      1650      1660
frame4 A.PYQC..HLHLPKDR.YEFHR.RDLQT.RTCGR.TCHLEYGRFQKDQRTYEQLLDGCLH

gi|158 QSQSIIPGPPSALLSLSPASNEPSPRDFAYPSPLPPHQPHPHHSLSGSGFDSE
      1010      1020      1030      1040      1050      1060

>>gi|241944366|gb|EES17511.1| hypothetical protein SORBI (958 aa)
  initn: 151 initl: 82 opt: 158 Z-score: 191.8 bits: 48.2 E(): 0.13
Smith-Waterman score: 159; 19.783% identity (51.762% similar) in 369 aa overlap
(1139-1500:160-498)

      1110      1120      1130      1140      1150      1160
frame4 KNK.Y.FENEQKDHIH.LFSIHFHFTVR.RKPNKKHSLQAQMQVQEQFSQCHNTQTQ
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|241 KDGRRLDGAPESSEGGILSSDAENKKGHEQKQRQSRKSPNPRREKERRHTDEHHSSSS
      130      140      150      160      170      180

      1170      1180      1190      1200      1210      1220
frame4 .DSGVCAMKLMH.T.RTLSKPMIRTKALARGS.DLRIRYRARILELIKQP.CRRVRS.LQ
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|241 KDH-----YSKNHSRTSPYSRHQSEAH-SRDQHLRSRER-----GDDTNGSRASLR
      190      200      210      220      230

      1230      1240      1250      1260      1270      1280
frame4 DQPSNP.TLGS..RS..HHQQ.QGFQK-RDP.RGSS.RYGDRGG.RQLLQKRYRDPYRQD
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|241 DG-SNRESHDRNGKSGRHTIRTQESERERSSRGIHDRHGDRHCDRRGQERYRD-----
      240      250      260      270      280

      1290      1300      1310      1320      1330      1340
frame4 DHARREK-SHLHRNQHR.A.DRLRQQKDGRFP.HGVLPNRS.R.HLQRMQLQORELDTTR
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|241 DRIDRKIDSLAETRRHRSRSYSRSDLRESTR--LCDQSRERERQSGSLRHRDHDSKS
      290      300      310      320      330      340

      1350      1360      1370      1380      1390      1400
frame4 S.-----KEHGLEQ.HPSK.KNEHASQPRLLHQQDVG.SLRCRHPFPAESR.DQYGLGS
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|241 DTSKDRHRESGRVSSAHE-RERGRDA-RDREWHRVKGSETHRAKEGRDKVSDGDRHRDST
      350      360      370      380      390      400

      1410      1420      1430      1440      1450      1460
frame4 SKR.GW.RRNLEQARQQPMGKQ.RMDHLEHQSLDR.AYHDLQDGWYAHRHQHSQR.G.HQ
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|241 RSKYSVSDGYKERTRSGEKGRD--VDH-KNRKFEETKENSLSKEDEEEYQEKIEQLAMQ
      410      420      430      440      450      460

      1470      1480      1490      1500      1510      1520
frame4 NLEASFQCGHES.LAR.QW.YLGC.DQQAQRTSLERKP.VPCR.LALESLVHAR.LEDD
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|241 EEDDPKIKIEEAR--RRKEAIMAKYRQQSQKQDMESKPSSNDEEVVRAMDGNETIYQKDD
      470      480      490      500      510

```

```

      1530      1540      1550      1560      1570      1580
frame4 HRLQTS PAHPFREV GSTH DPLDA..EK RHQ.KCYRNRKH QQDP.SNGNLQRYRNRREEPQ

gi|241 ND SGSTGNDEAENKHSSEVFDDKTDFTVGKSPA HNDASTSAGAF TDERTIGV SGLGEGS
      520      530      540      550      560      570

>>gi|89303110|gb|EAS01098.1| hypothetical protein THERM (3459 aa)
      initn: 53 initl: 53 opt: 164 Z-score: 191.8 bits: 50.1 E(): 0.13
Smith-Waterman score: 212; 17.595% identity (50.111% similar) in 898 aa overlap
(1088-1937:1750-2587)

      1060      1070      1080      1090      1100      1110
frame4 A..ISIMVQLQVFYSSY.LRSLEATI.HIFTQNKMFAYLL.FQIQHTTNKRKNK.Y.FEN
      .....
gi|893 KQSTQVQKAPSSRDIVEGKNEQAIIKQNSSQNQNQSQKTQAQTPTSNKPQQKQEESEY
      1720      1730      1740      1750      1760      1770

      1120      1130      1140      1150      1160      1170
frame4 EQKDHIH.LFSIH FHTVR.RKPNKKH SKLQAQMVQEKQFSQCHNTQTQ.D---SGVC
      .....
gi|893 EEESEYETE-----QKQQQPSQAQKSQIKQDQQAQSQNSKNQAQVQSSQVQ
      1780      1790      1800      1810      1820

      1180      1190      1200      1210      1220      1230
frame4 AMKLMH.T.RTL SKPMIRTKALARG.S.DLRIRYRARILELIKQP.CRRVRS.LQDQPSNP
      .....
gi|893 QPTIKQQKQEQFEYQVQEQSEKTNQKQPNQKSQEKQEFNQKYDQDV-NKQNGKQKNQ
      1830      1840      1850      1860      1870      1880

      1240      1250      1260      1270      1280
frame4 .TLGS..RS..HHQQ.QGFQKRDP.RGSS.RYGDRGG.RQLLQKRYRD-----PYR
      .....
gi|893 VTSTQAQKPANHQKQEQEASEYEESEYETEQAQKAPQQQQAQKTKQDQAQKNVVPK
      1890      1900      1910      1920      1930      1940

      1290      1300      1310      1320      1330      1340
frame4 QDDHARREKSHLRNQHR.A.DRLRQQKDGRFP.HGVLPNRS.R.-HLQRMQLQOREL--
      .....
gi|893 QINHSE-NKSQTQNNQEQQPLVK-SQQKDSQTEEYKQKVSSSQSSQVQQQPKQEEIKQ
      1950      1960      1970      1980      1990      2000

      1350      1360      1370      1380      1390      1400
frame4 -DTRRS.KEHGLLEQ.HPSK.KNEHASQPRLLHQQDVG.SLRCRHPPFAESR.DQYGLGS
      .....
gi|893 NDLQSQEQKQNN--QKNQSQQTQKAQTQQPVVKTQQKQNESEYEESEYETEQPQNQ-----
      2010      2020      2030      2040      2050

      1410      1420      1430      1440      1450      1460
frame4 SKR.GW.RRNLEQARQQPMGKQ.RMDHLEHQSLDR.AYHDLQDGGYAHRRHQHSQR.G.HQ
      .....
gi|893 -----KQTNAQQQALKQKQ-----QQSNNQDASKNKQT---SQNTQVQKTQAQQ
      2060      2070      2080      2090

      1470      1480      1490      1500      1510      1520
frame4 NLEASFQCGHES.LAR.-QW.YLGC.DQQAQRTSLERKP.VPCR.LALESLVHAR.LED
      .....
gi|893 PTIKSQQNQEESYEESQETEQNKIKQQPQSQNSIV-KQELKKQSSQNGQVQNEQKQN
      2100      2110      2120      2130      2140      2150

      1530      1540      1550      1560
frame4 DHRLQTS----PA-----HPF----REV GSTH DPLDA..EK---RHQ.KCYRNRKHQ
      ... :::

```

```

gi|893 EKQPQTSQFQQPASKPQQTQHEIVKNQQQVKQENSNIKNSQEKNNNQNQTQVQQSSQSQKK
      2160      2170      2180      2190      2200      2210

      1570      1580      1590      1600      1610      1620
frame4 DP.SNGNLQRYRNRREEPQEQGVHRYQH HQ.SKYHLPYGFWHP.LA.PYQC..HLHLPKD
      :: .....
gi|893 DPQVQQNK--QESEYEESEEEPIKNVQQKQIEQSKVKQNSTQQKLEPVIQQKDQLKSQQD
      2220      2230      2240      2250      2260      2270

      1630      1640      1650      1660      1670      1680
frame4 R.YEFHR.RDLQT.RTCGR.TCHLEYGRFQKDQRTYEQLLDGCLHREACTPWKQKT.P.E
      .....
gi|893 QVAQSQKV-DQKQPNFQNQNTQNLNQSQGQKSQPDQTAVQQQQQQQQQKSQEQPKQGLE
      2280      2290      2300      2310      2320      2330

      1690      1700      1710      1720      1730      1740
frame4 DEAQSQLTSSSSSIPTTKSDKLVDAAASSEKPIREICVDWISR DWR.EMGSAHHCRFC
      .....
gi|893 KDVQQQKKNASNEA-----QDQNAKKSTQQVNPQQNQATIQNDQVQHQIQSQKALEKA
      2340      2350      2360      2370      2380      2390

      1750      1760      1770      1780      1790      1800
frame4 .LAPW.ISGQICKKKNNQILNQ.KQ.TYITIMKIIFNTKQS..KAAIGVKPISTT-KIV
      .....
gi|893 TNQQNQVQNSQI---KQGVLEKQKSVADQQSQPAIEQQSKSETKQEKKLQQDNQSHKEQ
      2400      2410      2420      2430      2440

      1810      1820      1830      1840
frame4 VLYHLS.AS.IQSNRQ-----FQEI.SKQSENYP.KKGTLDLLCY--SRITQSOF
      .....
gi|893 VIQKNVNQTANQQNQKQPQONQEESYDESESESQQKITFQKNQVDQQKHLDKTVAAQEN
      2450      2460      2470      2480      2490      2500

      1850      1860      1870      1880      1890      1900
frame4 KSTYIINTQINKFTRI.IHQPTINRTN.NLKTRSIRMGK.FHE.IIV.KQ.NVDINN R.N
      .....
gi|893 KTNSSQNQEIQKNSAV---KSLKQNNSTLENQSKQSEVKLQTEELASQKQSP EIKN---
      2510      2520      2530      2540      2550      2560

      1910      1920      1930      1940      1950      1960
frame4 LNNRSEFSERSINKEHRIQQSSK.YEQRSAD EKEKERSKGTLTLTSEISQKSV.RLGSIR
      .....
gi|893 -NNEQNKLEQSKVNQHS-QNPSKPEQQQILPQNNQGLSSSQKIQQQPKQNI PNKHQQNE
      2570      2580      2590      2600      2610

>>gi|2623367|gb|AAC53441.1| sex determining protein [Mus (420 aa)
      initn: 187 initl: 81 opt: 154 Z-score: 191.6 bits: 47.0 E(): 0.13
Smith-Waterman score: 194; 19.231% identity (55.983% similar) in 234 aa overlap
(1246-1468:160-388)

      1220      1230      1240      1250      1260      1270
frame4 KQP.CRRVRS.LQDQPSNP.TLGS..RS..HHQQ.QGFQKRDP.RGSS.RYGDRGG.RQL
      .....
gi|262 SFYWQVDIPTGHLQQQQQQQQQQQQFHNHHQQQQQFYDHHHQQQQQQQQQQQQQQQ
      130      140      150      160      170      180

      1280      1290      1300      1310      1320      1330
frame4 LQKRYRDPYRQDDHARREKSHLRNQHR.A.DRLRQQKDGRFP.HGVLPNRS.R.HLQRM
      .....
gi|262 QQQQF-----HDHHQQKQFHDHHQQQQQFHDHHHHHQQQQFHDHHQQQQQFHDHQQQQQ
      190      200      210      220      230      240

```



```

      1340      1350      1360      1370      1380
frame4 QLQQRELDTRRS.KEHGLEEQ.HPSK.K-NEHASQPRLLH----QQDVG.SLRCRHPPF
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|262 QQQQQQFHDHQQKQKFHDHQQQQQQFHDHQQQQQQFHDHQQQQQQFHDHQQQQHQFH
      250      260      270      280      290      300

      1390      1400      1410      1420      1430      1440
frame4 AESR.DQ--YGLGSSKR.GW.RRNLEQARQQ--PMGKQ.RMDHLEHQSLDR-.AYHDLQD
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|262 DHPQQKQKFHDHPQQQQQFHDHQQQQQQKQKFHDHQQKQKFHDHQQKQKFHDHQQQQQ
      310      320      330      340      350      360

      1450      1460      1470      1480      1490      1500
frame4 GWYHRQHQSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.DQQAQRTSLERKP.VPC
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|262 QFHDHQQQQQQQQQQQQQQQQFHDQQLTYLLTADITGEHTPYQEHLSTALWLAVS
      370      380      390      400      410      420

>>gi|9916|emb|CAA39663.1| liver stage antigen [Plasmodiu (1909 aa)
  initn: 41 initl: 41 opt: 161 Z-score: 191.5 bits: 49.2 E(): 0.13
Smith-Waterman score: 184; 20.726% identity (51.282% similar) in 468 aa overlap
(1147-1592:1109-1556)

      1120      1130      1140      1150      1160      1170
frame4 NEQKDHIH.LFSIHFTVR.RKPNKKHSLQAQQ--MVQEKQFSQ-CHNTQTQ.DSGV
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|991 QQSDLEQERLAKEKLQGGQSDLEQERLAKEKLQGGQSDLEQERLAKEKLQGGQSDLEQER
      1080      1090      1100      1110      1120      1130

      1180      1190      1200      1210      1220      1230
frame4 CAMKLMH.T.RTLSKPMIRTKALARGS.DLRIRYRARILELIKQP.CRRVRS---LQDQ
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|991 LAKEKLQEQSDLEQERLAKEKLQEQSDLEQERRAKEKLQEQSDLEKASKETLQEQ
      1140      1150      1160      1170      1180      1190

      1240      1250      1260      1270      1280      1290
frame4 PSNP.TLGS..RS..HHQQ.QGFQKRD.PRGSS.RYGDRGG.RQLLQKRYRDPYRQDDHA
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|991 QSD-LEQERLAKEKLQEQSDLEQERRAKEKLQEQSDLEQER-LAKEKLQEQSDLEQE
      1200      1210      1220      1230      1240      1250

      1300      1310      1320      1330      1340
frame4 RREKSHLHRNQ-----HR.A.DRLRQQKDGRFP.HGVLPNRS.R.HLQRMQLQQRELDTR
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|991 RRAKEKLQEQSDLEQERRAKEKLQEQQSDL-----EQERRAKEKLQEQSDLEQE
      1260      1270      1280      1290      1300

      1350      1360      1370      1380      1390      1400
frame4 RS.KEHGLEEQ.HPSK.KNEHASQPRLLHQQ-DVG.SLRCRHPPFAE-SR.DQYGLGSSK
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|991 RLAKEK-LQEQQ--QSDLEQERLAKEKLQEQSDLEQERRAKEKLQEQSDLEQERLAKEK
      1310      1320      1330      1340      1350      1360

      1410      1420      1430      1440      1450
frame4 R.GW.RRNLEQARQQPMGKQ.RMDHLEHQSLDR.AYH----DLQDGWYHRQHQSQR.G.
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|991 -LQEQQSDLEQERRAKEKLQEQSDLEQDRLAKEKLQEQQRDLEQERRAKEKLQEQQSDL
      1370      1380      1390      1400      1410      1420

      1460      1470      1480      1490      1500      1510
frame4 HQNLEASFQCG-HES.LAR.QW.YLGC.DQQA--EQRTSLERKP.VPCR.LALESVHAR
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

gi|991 EQERRAKEKLQEQQSDLEQERRAKEKLQEQQSDLEQERRAKEKLQEQQSDLEQERLAKEK
      1430      1440      1450      1460      1470      1480

      1520      1530      1540      1550      1560      1570
frame4 .LEDDHRLQTS--PAHPFREVGSTHDPDA..EKRHQ.KCYRNRKHQQDP.SNGNLQRY
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|991 LQEQQRDLEQERRAKEKLQEQQS----DLEQERRAKEKLQEQQSDLEQERLANEKLQEQ
      1490      1500      1510      1520      1530

      1580      1590      1600      1610      1620      1630
frame4 RNREEPQEQGVHRYQHHQ.SKYHLPYGFWHP.LA.PYQC..HLHLPKDR.YEFHR.RDLQ
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|991 QRDLEQERRAKEKLQEQQSDLEQERRAKEKLQEQQSDLEQERRAKEKLQEQQSDLEQERL
      1540      1550      1560      1570      1580      1590

>>gi|142534199|gb|ECY77410.1| hypothetical protein GOS_2 (445 aa)
  initn: 54 initl: 54 opt: 154 Z-score: 191.3 bits: 47.0 E(): 0.13
Smith-Waterman score: 154; 22.121% identity (53.030% similar) in 330 aa overlap
(1287-1592:6-315)

      1260      1270      1280      1290      1300      1310
frame4 DP.RGSS.RYGDRGG.RQLLQKRYRDPYRQDDHARREKSHLH--RNQHR.A.D---RLR
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|142 MRQHWDHDAQQHEDMISSMKDEHQGVIDGHASALA
      10      20      30

      1320      1330      1340      1350      1360
frame4 Q---QKDGRFP.HGVLPNRS.R.HLQRMQLQQR-ELDT-RRS.KEH-GLLEQ.HPSK.KN
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|142 QADANHDSTKAQHSADVDR-LTEELDQMRTRYDEMQLMKQSHANHVDDIEKQHQLKALE
      40      50      60      70      80      90

      1370      1380      1390      1400      1410      1420
frame4 -EHASQPRLLHQQDVG.SLRCRHPPFAESR.DQYGLGSSKR.GW.RRNLEQARQQPMGKQ
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|142 LQHAITQG---HESDIA-SMKQQHTLALESKTQQHAE-SMQRLEDEAGNMKQSHRDELARM
      100      110      120      130      140

      1430      1440      1450      1460      1470      1480
frame4 .RMDHLEHQSL-DR.AYHD--LQDGWYHRQHQSQR.G.HQNLEASFQCGHES.LAR.QW
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|142 -KQDH--HDSMSDRLEDHQSAMQRLKDDHMETLSQQAREHEDALSTMAQGHEMDHAR---
      150      160      170      180      190      200

      1490      1500      1510      1520      1530
frame4 .YLGC.DQQAQ-RTSLERKP.VPCR.LALESLV--HAR.LEDDHRLQTSAPHPFREVGS
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|142 ---ALEGHGAELERLRREHGERVDSLKIKHETAISEHQATLEEHVQRHEDAVSSLK--GE
      210      220      230      240      250

      1540      1550      1560      1570      1580      1590
frame4 THDPLDA..EKRHQ.KCYRNRKHQQDP.SNGNLQRYRN----REEPQEQGVHRYQHHQ
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|142 HQAALDDHVQRQEVMLSLNNEEQS---ALDDQVQRHQDMVSSLKDEHQAALDDHAQQHE
      260      270      280      290      300      310

      1600      1610      1620      1630      1640      1650
frame4 .SKYHLPYGFWHP.LA.PYQC..HLHLPKDR.YEFHR.RDLQT.RTCGR.TCHLEYGRFQ
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|142 DMISSMKDEHQGVIDGHASALAQADANHDSTKAQHSADVDRLTEELDQMRTHYTDMDMLM
      320      330      340      350      360      370

```

>>gi|124403363|emb|CAK68827.1| unnamed protein product [ (517 aa)  
initn: 49 init1: 49 opt: 154 Z-score: 190.4 bits: 47.1 E(): 0.15  
Smith-Waterman score: 172; 18.338% identity (52.722% similar) in 349 aa overlap  
(1140-1468:31-370)

```
1110      1120      1130      1140      1150      1160
frame4 NK.Y.FENEQKDHIIH.LFSIHFHFTVR.RKPNKKHSLQAQMQMVEKQFSQCHNTQTQ.
      . : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|124 MDSNDQNDLKKRLELKRKILQEQRKRAKEEQKKKKKEDEEQSQMQMKQEQQSSSENVKKQS
      10      20      30      40      50      60
```

```
1170      1180      1190      1200      1210      1220
frame4 DSGVCA---MKLMH.T.RTLSKPMIRTKALARGS.DLRIRYRARILELIKQP.CRRVRS.
      . : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|124 DNQSIIDNKLKQQLNEQQKQKEQQIKSEALRKQS-----EQSNQMSLQKKENSQSTQNN
      70      80      90      100      110
```

```
1230      1240      1250      1260      1270      1280
frame4 LQDQPSNP.TLGS..RS..HHQQ.QGFQKRDP.RGSS.RYGDRGG.RQLLKRYRDPYRQ
      . : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|124 QKPQKQDDQKVIQITNKQINNQQ-KTFASDQKGKKNKFIEDNNRN-QKEELQKKNDPE--K
      120      130      140      150      160      170
```

```
1290      1300      1310      1320      1330      1340
frame4 DDHARREKSHLHRNQHR.A.DRLRQQKDG---RFP.HGVLPNRS.R.HLQRMQLQQREL
      . : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|124 PDQIRQDKEQQVQEENKKQLLLINKKTDNEKQIRDQKHIKEQDEKDKQNKQRLESEKIEK
      180      190      200      210      220      230
```

```
1350      1360      1370      1380      1390
frame4 DT-RRS.KEHGLLEQ.HPSK.KNEHASQPRLLH--QQDVG.SLRCRHPFPAESR.DQYGL
      . : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|124 DRIDREKKEKERLEKEKQERLLRERQKEKEKQLERDKQEKDRQLREKQKEKDKQEKMEKLR
      240      250      260      270      280      290
```

```
1400      1410      1420      1430      1440
frame4 GSSKR.GW.RRNLEQ-----ARQQPMGKQ.RMDHLEHQS LDR.AYHDLQ--DGWYAH
      . : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|124 EREKQEKLQRQEKEKQDKLEREKREKQERLDRQERQERLERERQEKYDERKEKERQE
      300      310      320      330      340      350
```

```
1450      1460      1470      1480      1490      1500
frame4 RQHQSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.DQQAQRTSLERKP.VPCR.LAL
      . : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|124 RERQEQRKEKQLLNSKRQLERTQVSQKSDSIQTKKTQERIEQEFDNSQQQANCNIYE
      360      370      380      390      400      410
```

>>gi|193897155|gb|EDV96021.1| GH15437 [Drosophila grimsh (490 aa)  
initn: 131 init1: 76 opt: 153 Z-score: 189.4 bits: 46.8 E(): 0.17  
Smith-Waterman score: 153; 19.615% identity (49.231% similar) in 260 aa overlap  
(1355-1600:128-373)

```
1330      1340      1350      1360      1370      1380
frame4 NRS.R.HLQRMQLQQRELDTRRS.KEHGLLEQ.HPSK.KN---EHASQPRLLHQDVG.S
      . : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|193 HQQQPTQNVYGQETTNYATNGYAPATVAPTQNTYPSNNYNYPQQQQQQPQIQTQAQFNR
      100      110      120      130      140      150
```

```
1390      1400      1410      1420      1430      1440
frame4 LRCRHPFPAESR.DQYGLGSSKR.GW.RRNLEQARQQPMGKQ.RMDHLEHQS LDR.AYHD
      . : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|193 QPSREPAPATATAD-----NQQYAHYPQQQQQQQQQOPHQQQYQPSYQQQQHQPTPHPHN
```

```
160      170      180      190      200      210
      1450      1460      1470      1480      1490
frame4 LQDGWY----AHRQHQSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.DQQAQRTSL
      . : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|193 VGGGWKHIGAPAPKAHNDFTAGGVAPATNAYPNYQQQQQPQQQYPQ---QQQHQHQHQY
      220      230      240      250      260
```

```
1500      1510      1520      1530      1540      1550
frame4 ERKP.VPCR.LALESVLVHAR.LEDDHRLQTSAPHPFREVSTHDP LDA..EKRHQ.KCYR
      . : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|193 QQQPQQQ-QYRAPSNDSHQQQ-QQQWQQQQQQQQPQQQL---PPQMQQYQAPYQTSYQ
      270      280      290      300      310      320
```

```
1560      1570      1580      1590      1600      1610
frame4 NRKHQQDP.---SNG--NQLQ-RYRNREEPQEQGVHRYQH HQ.SKYHLPGFWHP.LA.P
      . : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|193 PQQQQQQPSYPQQGGGANYAQPQYNSYSQPQQPQQQQQLHQQQQLPYSQDQTDQQQG
      330      340      350      360      370      380
```

```
1620      1630      1640      1650      1660      1670
frame4 YQC..HLHLPKDR.YEFHR.RDLQT.RTCGR.TCHLEYGRFQKDQRTYEQLLDGCLHREA
      . : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|193 YRGASPGIITLRKEAPVSTPQPVYTSQPAAVSYQGGSKLRGDLKWPPEYKEAAVKENE
      390      400      410      420      430      440
```

>>gi|143313824|gb|EDE27300.1| hypothetical protein GOS\_1 (286 aa)  
initn: 161 init1: 79 opt: 150 Z-score: 188.8 bits: 45.9 E(): 0.19  
Smith-Waterman score: 155; 30.952% identity (48.571% similar) in 210 aa overlap  
(1261-1455:66-256)

```
1240      1250      1260      1270      1280
frame4 PSNP.TLGS..RS..HHQQ.QGFQKRDP.RGSS.RYG-DRGG.RQLLKRYRDPYRQDDH
      . : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|143 QTHPLHDFGWTGDQAHHPAFLRAPGALFPAGSRARQGRPRPGG---LRGR-RGPRLHAER
      40      50      60      70      80      90
```

```
1290      1300      1310      1320      1330      1340
frame4 ARREKSHLHRNQHR.A.DRLRQQKDGRFP.HGVLPNR---S.R.HLQRMQLQQRELDTRR
      . : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|143 AGRHDAHPEEPAHR-----LRPQPDHRLPHGGVLP RRVLPGRGDHGQRRRRDQR-LEGRQ
      100      110      120      130      140
```

```
1350      1360      1370      1380      1390
frame4 S.KE---HGLLEQ.HPSK.KNEHASQPRLLHQDVG.SLRCRHPF---PAESR.DQYGL
      . : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|143 HREVLSSGRGRARQHLRQ-RTRLVDQPAA-QPPDGRQRLRAHQDQMRGGPA-TRPAQ--L
      150      160      170      180      190      200
```

```
1400      1410      1420      1430      1440      1450
frame4 GSSKR.GW.RRNLEQARQQPMGKQ.R---MDHLEHQS LDR.AYHDLQDGWYAHRQHQSQ
      . : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|143 PDRWLRIIRRDVLHHLRS---GRHLRGRGAAEQSVHQRPERQACHDLHAGRIQHR-HPRQ
      210      220      230      240      250
```

```
1460      1470      1480      1490      1500      1510
frame4 R.G.HQNLEASFQCGHES.LAR.QW.YLGC.DQQAQRTSLERKP.VPCR.LALESLVHA
      . : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|143 PLRRPRPGHPSPGRLLRRRHLRQPRADPLT
      260      270      280
```

>>gi|134083327|emb|CAK42894.1| unnamed protein product [ (409 aa)

initn: 81 init1: 50 opt: 151 Z-score: 187.9 bits: 46.3 E(): 0.21  
Smith-Waterman score: 151; 20.165% identity (52.675% similar) in 243 aa overlap  
(1221-1453:152-388)

```

      1200      1210      1220      1230      1240      1250
frame4 IRTKALARGS.DLRIRYRARILELIKQP.CRRVRS.LQDQPSNP.TLGS..RS..HHQQ.
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|134 ATLGGGGAVAQANPRDTMAKVLADPLLEIRKREQAALGEMVREEVRKSSSSKGEKRDRDR
      130      140      150      160      170      180

      1260      1270      1280      1290      1300      1310
frame4 QGFQKRDP.RGSS.RYGDRGG.RQLLQKRYRDPYRQDDHARREKSHLHRNQHR.A.DRLR
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|134 DGHRR--RRERERRHRSRSPERRRDTLDRDDRRSERDHRREDS-DRDYHRSSSRHRDH
      190      200      210      220      230
```

```

      1320      1330      1340      1350      1360
frame4 QQKDGFRP.HGVLPNRS--.R.HLQRMQLQ-QRELDTRRS.KEHGLLEQ.HPSK.KNEHA
      .... .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|134 RDREDRYDRRESRQDRSPADRHHSDDRRADDRRESDGHRDSRDRDRERDRDRDRSSRYH
      240      250      260      270      280      290
```

```

      1370      1380      1390      1400      1410      1420
frame4 SQPRLLHQDDVG.SLRCR-HPFPAESR.DQYGLGSSKR.GW.RRN---LEQARQQPMGKQ
      .. .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|134 QDKRDYRPRDRDITYTRDKPSPEEAEQKRKQEEERQKLAEMQANASEMEDARRQRIAEV
      300      310      320      330      340      350
```

```

      1430      1440      1450      1460      1470      1480
frame4 .RMDHLEHQSLDR.AYHDLQDGWYA---HRQHQSQR.G.HQNLEASFQCGHES.LAR.QW
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|134 TAMEEKQHEDEK---HRSEKGRFVSGLHKQLQEDSLDERIKRSRGGLARMDDED
      360      370      380      390      400
```

```

      1490      1500      1510      1520      1530      1540
frame4 .YLGC.DQQAQRTSLERKP.VPCR.LALESLVHAR.LEDDHRLQTSPAHPFREVGGSTHD
```

>>gi|219884915|gb|ACL52832.1| unknown [Zea mays] (375 aa)  
initn: 74 init1: 74 opt: 150 Z-score: 187.2 bits: 46.0 E(): 0.23  
Smith-Waterman score: 150; 26.481% identity (47.038% similar) in 287 aa overlap  
(1189-1461:73-340)

```

      1160      1170      1180      1190      1200      1210
frame4 FSQCHNTQTQ.DSGVCAMKLMH.T.RTLSPKPMIRTKALARGS.DLRIRYRARILELIKQP
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|219 LPPLHRPLLRPPRRRLPSASVRPAAAAAVPAVPRQGPAAGR-QLRRRRRRHRHGSAL-LP
      50      60      70      80      90      100
```

```

      1220      1230      1240      1250      1260      1270
frame4 .CRRVRS.L-QDQPSNP.TLGS..HH--QQ.QGFQKRDP.RGSS.RYGDRGG.RQL
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|219 GDRRVQALDQVPVQPARLVRAAQAVALQLTERVQGVLPVLPGRDR-GERLQLR-L
      110      120      130      140      150
```

```

      1280      1290      1300      1310      1320
frame4 LQKRYRDPYRQDD---HARREKSHLHRNQHR.A.DRLRQKQDGR-FP.HGVL---PNRS.
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|219 LQGQ--DPGRRQDLRPHGRRRG---HRC-HREADQGRRHAPGGAGEPAHGLLVGVPDPAP
      160      170      180      190      200      210
```

```

      1330      1340      1350      1360      1370      1380
frame4 R.HLQRMQLQQRELDTRRS.KEHGLLEQ.HPSK.KNEHASQPRLLHQQDVG.SLRCRHPF
```

```

      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|219 GPQRQLRRRRRLPQDVQRLRAAP---QRCAPAETAGAPGEV--AGEDHVRRLLRRRHVL
      220      230      240      250      260
```

```

      1390      1400      1410      1420      1430      1440
frame4 PAESR.DQYGLGSSKR.GW.RRNLEQARQQPMGKQ.R--MDHLEHQSLDR.AYHDLQDGW
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|219 RQEPQAVRVHAGAAADVLRRRRAVQL---QPEGELRRAGVQVRGVPVGRVQLGRRAPDGG
      270      280      290      300      310      320
```

```

      1450      1460      1470      1480      1490      1500
frame4 YA--HRQHQSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.DQQAQRTSLERKP.VPC
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|219 RLPRHRRQHQPQR-ALHQPPAPLIRTPVVEARRPSIGSRSSSYNCSGANNVAGL
      330      340      350      360      370
```

>>gi|148668340|gb|EDL00666.1| repetin [Mus musculus] (1118 aa)  
initn: 65 init1: 65 opt: 155 Z-score: 187.0 bits: 47.5 E(): 0.23  
Smith-Waterman score: 175; 19.693% identity (47.570% similar) in 391 aa overlap  
(1228-1597:745-1095)

```

      1200      1210      1220      1230      1240      1250
frame4 RGS.DLRIRYRARILELIKQP.CRRVRS.LQDQPSNP.TLGS..RS..HHQQ.QGFQK--
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|148 QQKGKGRDQSPHLGQKGRHDQSPHRGQKGRDQSPHQGQKGRDQDQSS-HQQGREGQDQNS
      720      730      740      750      760      770
```

```

      1260      1270      1280      1290      1300
frame4 ---RDP.RGSS.RYGDRGG.RQLLQKRYRDPYRQDDHARREKSH---LHRNQ-----
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|148 HWHRTDRQGQSFHYGQTGGQGLSSHQGTDSQGQNSQWHRTDSQGQSFHFDQAGREGQSS
      780      790      800      810      820      830
```

```

      1310      1320      1330      1340      1350
frame4 HR.A.DRLRQKQDGRFP.HGVLPNRS.R.H---LQRMQLQQRELDTRRS.K---EHGLLEQ
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|148 HHGQTDQRQSSSHCGQSEIGKTENQGQNRHSLGTDRTTRDSYVEQSGRSVKLSQQNSREE
      840      850      860      870      880      890
```

```

      1360      1370      1380      1390      1400      1410
frame4 .HPSK.KNEHASQPRLLHQDDVG.SLRCRHPPFAESR.DQYGLGSSKR.GW.RRNLEQAR
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|148 VRQTQSQRSHDRREQQIQQTWPKPEDNQHKLLAQVQQEPY---SYEEYDWQSQSSEQ--
      900      910      920      930      940
```

```

      1420      1430      1440      1450      1460      1470
frame4 QQPMGKQ.RMDHLEHQSLDR.AYHDLQDGWYAHRHQSQR.G.HQNLEASFQCGHES.LA
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|148 -DHCGEEEYQDWDHRSVEDQENLYEMQN-WQTHEEQSHQTSDRQ-----
      950      960      970      980      990
```

```

      1480      1490      1500      1510      1520      1530
frame4 R.QW.YLGC.DQQAQRTSLERKP.VPCR.LALESLVHAR.LEDDHRLQTSPAHPFREV
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|148 ----THVDEQNQQRRQHRQTHEEN-----HDHQHGRHHEDEHNHRRQDHHQQRER-
      1000      1010      1020      1030
```

```

      1540      1550      1560      1570      1580      1590
frame4 STHDPLDA..EKRHG.KCYNRKHKQQDP.SNGNQLQRYRNREEPEQGVHRYQHQQ.SKY
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|148 QTHEEKEKYQGQDQSRSPNREKSH--MSEDDQCEGPGRRFHPHGGGKSQRREKSGN
      1040      1050      1060      1070      1080      1090
```

```

1600      1610      1620      1630      1640      1650
frame4 HLPYGFWHP.LA.PYQC..HLHLPKDR.YEFHR.RDLQT.RTCGR.TCHLEYGRFQKQDR
:
gi|148 HPTKPANYSSPLYDYVQEQAAQYQ
1100      1110

>>gi|228383859|gb|ACQ27812.1| Sequence 20326 from patent (405 aa)
  initn: 66 initl: 66 opt: 150 Z-score: 186.7 bits: 46.0 E(): 0.24
Smith-Waterman score: 150; 28.649% identity (45.405% similar) in 185 aa overlap
(1194-1370:197-374)

1170      1180      1190      1200      1210      1220
frame4 NTQTQ.DSGVCAMKLMH.T.RTLSKPMIRTKALARGS.DLRIRYRARILELIKQP.CRRV
.: : : : . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|228 GVHQPAQQRARRGWPERYPVALGRGHGPGPRAGPRRSRQRPVRLRP---ALLRRRTEHRV
170      180      190      200      210      220

1230      1240      1250      1260      1270      1280
frame4 RS.LQDQPSNP.TLGS..RS..HHQQ.QGFQKRD.P---GSS.RYGDRGG.RQLLQKRY
: : : : . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|228 RRLLELHPQPPADLQRPGRLGQHGGDQ--RPRHPRRDQPGTGRPGRGGGHRRLQLQLLH
230      240      250      260      270      280

1290      1300      1310      1320      1330
frame4 RDPYRQDDHARREKSHLHRNQHR.A.-DRLRQQKDGFRP.HGVLPNRS.R.HLQRMQ--L
: : : : . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|228 RVPGRYP-HVRPELLHHVAQRHVRERRSQLPAQRAGAVEHHR-HPLRRPRPGEHRRQPPL
290      300      310      320      330

1340      1350      1360      1370      1380      1390
frame4 QQRELDTR--RS.KEHGLLEQ.HPSK.KNEHASQPRLHLHQDVG.SLRCRHFFPAESR.D
.: : : : . : : : . : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|228 RQRLGAERPGRAGPARLSAQGDHPTADHPDALLRPGDGRRAPDPGGRDRLPDRRRPGKHL
340      350      360      370      380      390

1400      1410      1420      1430      1440      1450
frame4 QYGLGSSKR.GW.RRNLEQARQQPMGKQ.RMDHLEHQSLDR.AYHDLQDGWYAHRRQHQSQ

gi|228 QGTLRA
400

>>gi|239594646|gb|EEQ77227.1| conserved hypothetical pro (811 aa)
  initn: 54 initl: 54 opt: 153 Z-score: 186.4 bits: 47.0 E(): 0.25
Smith-Waterman score: 163; 19.322% identity (51.864% similar) in 295 aa overlap
(1247-1534:54-327)

1220      1230      1240      1250      1260      1270
frame4 QP.CRRVRS.LQDQPSNP.TLGS..RS..HHQQ.QGFQKRD.PRG---SS.RYGDRGG.R
.: : : : . : : : . : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|239 SIYQTVNKLTKPPSSKELSALRMQYEQQLQQQQQGRQPSEPSPAFPPSNYSPLSYDGGQ
30      40      50      60      70      80

1280      1290      1300      1310      1320      1330
frame4 QLLQKRYRDPYRQDDHARREKSHLHRNQHR.A.DRLRQQKDGFRP.HGVLPNRS.R.HLQ
: : : . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|239 QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ-----QQQQQQQQ
90      100      110      120      130

1340      1350      1360      1370      1380      1390
frame4 RMQLQQRELDTRRS.KEHGLLEQ.HPSK.KNEHASQPRLHLHQDVG.SLRCRHFFPAES-
.: : : : . : : : . : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

gi|239 QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQHHSQSSQLPTPQLSRISYQQQPVKKEKE
140      150      160      170      180      190

1400      1410      1420      1430      1440      1450
frame4 R.DQYGLGSSKR.GW.RRNLEQARQQPMGKQ.RMDHLEHQSLDR.AYHDLQDGWYAHRRQH
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|239 RFFQWG----RRERKRREEEQQRQKQEQEEQPHQRLKQEQKDRERK---QREKLEREQQ
200      210      220      230      240

1460      1470      1480      1490      1500      1510
frame4 QSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.DQQAQRTSLERKP.VPCR.LALES
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|239 QRKQEQMQQKEEHFQRQNRQKQREEHKKRKEEPDQYQRRRKQEPKQEQEEQQLQ--
250      260      270      280      290      300

1520      1530      1540      1550      1560
frame4 VHAR.LEDDHRLQTSAPAH---PFREVGSTHDPDLDA..EKRHQ.KCYRNRKHQQDP.SNGN
: : : : . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|239 ---RQLKQQRQEEELQSQHYQFPFHEKRTQRQKPSIDWFDGPAHFSDDDNEGGGYDDNAE
310      320      330      340      350      360

>>gi|146143822|gb|EAS04465.2| hypothetical protein THER (723 aa)
  initn: 73 initl: 73 opt: 152 Z-score: 185.8 bits: 46.7 E(): 0.27
Smith-Waterman score: 159; 25.446% identity (56.250% similar) in 224 aa overlap
(1756-1969:449-655)

1730      1740      1750      1760      1770      1780
frame4 SRDWR.EMGSAHHCRFC.LAPW.ISGQQICKKKNNQILNQ.KQ.TYITIMKIIIFNTKQS.
.: : : : . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|146 KNQNILSDMEELPKHCFKIHFHNDKLKEVKKLYDQIVKQDENE-----VKYIEYTDADI
420      430      440      450      460      470

1790      1800      1810      1820      1830
frame4 .KAAIGVKPIS-----TTKIVVLYHLS.AS.IQSNRQFQEI.SKQSENYP.-KKGTLD
... : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|146 KQITVEVLKPLSEHFEHNSTTEFEILFFLIEDVKQLNDKFFQSSYSKQENINQFKKSYED
480      490      500      510      520      530

1840      1850      1860      1870      1880      1890
frame4 L--LCYSRITQSQFKSTYILNTQINKFTRI.IHQPTINRTN.NLKTRSIRMKG.FHE.II
. : : : . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|146 IRNFYYQRL-QKILELKYQNNVTQKLMRIVI--PTNSKVNNAILTDGLESIQYVDKQL-
540      550      560      570      580

1900      1910      1920      1930      1940      1950
frame4 V.KQ.NVDINNRLNLNNRSEFSERSINKEHRIQQSSK.YEQRSADKEKERSKGLTLT
: : : : . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|146 --DQVKVDLNQ--PMSEGSYFQIKLYEEHKAHQIIEFENEIASLKQVFQTVTQI----
590      600      610      620      630      640

1960      1970      1980      1990      2000      2010
frame4 EISQKSV.RLGSIRTINGISNESGFIYSVESETNVRV.HKIISFFKLPLCPCLAVNTQYL
.. : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|146 DMMEKTIERLLKREEFESDEIEEEDKQIAEQYKQLITIVEKYTEDMVALIEISPVFRS
650      660      670      680      690      700

>>gi|1806132|emb|CAA67624.1| repetin [Mus musculus] (1130 aa)
  initn: 65 initl: 65 opt: 154 Z-score: 185.6 bits: 47.3 E(): 0.28
Smith-Waterman score: 179; 19.841% identity (49.471% similar) in 378 aa overlap
(1228-1597:757-1107)

1200      1210      1220      1230      1240      1250

```

```
frame4 RGS.DLRIRYRARILELIKQP.CRRVRS.LQDQPSNP.TLGS..RS..HHQQ.QGFQK--
      1260      1270      1280      1290      1300      1310
gi|180 QGQKGRQDQSPHLGQKGRHDQSPHRGQKGRQDQSPHQGQKGRQDQSS--HQGQREGQEQNS
      730      740      750      760      770      780

      1260      1270      1280      1290      1300      1310
frame4 ---RDP.RGSS.RYGDRGG.RQLLQKRYRDPYRQDDHARREKSHLHRNQHR.A.DRLRQQ
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
gi|180 HWHRTDRQGQSFHYGQTGG---QGLSSHQ---GQTD-SQGQNSQWHRTDSQGQSFHFDQA
      790      800      810      820      830

      1320      1330      1340      1350      1360      1370
frame4 -KDGRFP.HGVLPNRS.R.HLQRMQLQQRELDTRRS.KEHGLLEQ.HPSK.KNEHASQPR
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
gi|180 GREGQSSHGQTDRQSQSSHCGQSEIGKTE---NQGQNRHSLGTDRTRRDSYVEHSGRSG
      840      850      860      870      880      890

      1380      1390      1400      1410      1420
frame4 LLHQQDVG.SLRCRHPFPAESR.DQYGLGSS--KR.GW.RRNLEQARQQPMGKQ.RMDHL
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
gi|180 KLSQNSREEVVRQTQSRSHDRREQQIQQTWKPEDNQHKLLAQVQQEPYSYEEYDWQS
      900      910      920      930      940      950

      1430      1440      1450      1460      1470      1480
frame4 EHQS LDR.AYHDLQDGWYAHRHQSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.DQQ
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
gi|180 QSSSEQDHCGEEYQD-WDRHSVEDQENLYEMQNWQTHEE--EQSHQTSDRQTHVDEQNQQ
      960      970      980      990      1000      1010

      1490      1500      1510      1520      1530      1540
frame4 AEQRTSLERKP.VPCR.LALESVHAR.LEDDHRLQTSPAHPFREVGS THDPLDA..EK
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
gi|180 RQHRQTHEEN-----HDHQHGRHHEDEHNHRRQDHHQQQERER-QTHEEKEKYQGQ
      1020      1030      1040      1050      1060

      1550      1560      1570      1580      1590      1600
frame4 HQ.KCYRNRKHKQDP.SNGNQLQRYRNREEPQEQGVHRYQHHQ.SKYHLPYGFWHP.LA.
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
gi|180 DQSRSFPPNREKSH--MSEDDQCEGPQGRRFHPTHGGGKSQREKSGNHTPKPANYSSPLY
      1070      1080      1090      1100      1110

      1610      1620      1630      1640      1650      1660
frame4 PYQC..HLHLPKDR.YEFHR.RDLQT.RTCGR.TCHLEYGRFQKDQRTYEQLLDGCLHRE

gi|180 DYVQEQAAYQY
      1120      1130

>>gi|194190700|gb|EDX04276.1| GD23558 [Drosophila simula (1025 aa)
  initn: 142 init1: 74 opt: 153 Z-score: 184.9 bits: 47.0 E(): 0.3
Smith-Waterman score: 172; 21.842% identity (48.947% similar) in 380 aa overlap
(1248-1598:630-984)

      1220      1230      1240      1250      1260      1270
frame4 P.CRRVRS.LQDQPSNP.TLGS..RS..HHQQ.QGFQKRD.PRGSS.RYGDRGG.RQLLQ
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
gi|194 PITTTTTTTTTTTTSPYAYNPYANWQNRPTQQHPNWHPADPEYVRAQEARQQEFIREREQ
      600      610      620      630      640      650

      1280      1290      1300      1310      1320      1330
frame4 KRYRDPYR-QDDHAR---REKSHLHRNQHR.A.DRLRQQKDGRFP.HGVLPNRS.R.HLQ
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
gi|194 RRREQQLRQDQRAHQEQEPREQQDERIRQOREQEERLRQOREQEERL-----RQOREQEE
```

```
      660      670      680      690      700      710

      1340      1350      1360      1370      1380
frame4 RMQLQQRELDTR-RS.KEHGLLE---Q.HPSK.KNEHASQPRLLHQDQDVG.SLRC-----
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
gi|194 RL-LEQRELEERIRQQQEREQYEREQQEREQEREREQEREREQEQPEQQPEYNPEPA
      720      730      740      750      760      770

      1390      1400      1410      1420
frame4 -----RHPFPAESR.D-----QYGLGSSK---R.GW.RRNLEQARQQPMGKQ.RMDHL
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
gi|194 NPWGYPEREPQPDENPEDVPSYEQYGPEGNEILPESDANRNYSEEDREQQREQLRREQQ
      780      790      800      810      820      830

      1430      1440      1450      1460      1470      1480
frame4 EHQS LDR.AYHDLQDGWYAHRHQSQR.G.HQNLEASFQCGHES.LAR.QW.YLGC.DQQ
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
gi|194 EKQERE---YQLQMEREQQEREQQER-GQQEPGPEEYPSYEE--YSRALQEKYAERDRI
      840      850      860      870      880

      1490      1500      1510      1520      1530      1540
frame4 AEQRTSLERKP.VPCR.LALESVHAR.LEDDHRL-QTSPAHPFREVGS THDPLDA..EK
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
gi|194 YAEQERERKQQ-----ETLLQENHQYPEQSLPEEQPTHPNYEAADRSYAEQER
      890      900      910      920      930

      1550      1560      1570      1580      1590      1600
frame4 RHQ.KCYRNRKHKQDP.SNGNQLQRYRNREEPQEQGVHRYQHHQ.SKYHLPYGFWHP.LA
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
gi|194 EQQRRDQVEQEREQQP---GEEYERSPDEEEAAEQDVLKVEDFATYEAYLEAATKLREEQ
      940      950      960      970      980      990

      1610      1620      1630      1640      1650      1660
frame4 .PYQC..HLHLPKDR.YEFHR.RDLQT.RTCGR.TCHLEYGRFQKDQRTYEQLLDGCLHR

gi|194 EEQEKL E EIRYRAQQEEEDRIQAERERNSHN
      1000      1010      1020

>>gi|124871096|gb|EAY62812.1| Outer membrane autotranspo (1356 aa)
  initn: 110 init1: 64 opt: 154 Z-score: 184.5 bits: 47.4 E(): 0.32
Smith-Waterman score: 165; 22.777% identity (43.601% similar) in 461 aa overlap
(1202-1625:841-1272)

      1180      1190      1200      1210      1220      1230
frame4 GVCAMKLMH.T.RTLSKPMIRTKALARGS.DLRIR-YRARILELIKQP.CRRVRS.LQDQ
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
gi|124 RRQRERRRRDRGRWRERHDPGHDDQPERDRDRGHRRRPGARPRRRGCRTRSRDDADQ
      820      830      840      850      860      870

      1240      1250      1260      1270      1280
frame4 PSNP.TLGS..RS..HHQQ.QGFQKRD.PRGSS.RY---GDRGG.RQLLQKRYRDPYRQD
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
gi|124 -SRRRHVDRRRRLHRAEPR--HAREPQHGVHARRRLDGRRGHARHGQQRVDDP-RLR
      880      890      900      910      920

      1290      1300      1310      1320      1330      1340
frame4 DHARREKSHLHRNQHR.A.DRLRQQKDGRFP.HG---VLPNRS.R.HLQRMQLQQRELDT-
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
gi|124 RH-RRRAGCVGHARQYGCDSRRRRDRGRAP-HGRRICGRAVGHGHRRRERQRRRRADRL
      930      940      950      960      970      980

      1350      1360      1370      1380      1390      1400
```

```

          1370      1380      1390      1400      1410      1420
frame4  NEHASQPRLLHQQDVG.SLRCRHFFPAESR.DQYGLGSSKR.GW.RRNLEQARQQPMGKQ
      ..  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi 1563  ORDLEQDRIQQQQLGOORRQAMERQEFRLRQ----SQERERQLRH---QSSQQQ-----

```

```
Sliding 8 amino acid window search
Database searched = AD_2010
Query = frame5
```



```

24 1 0:=
26 0 0:
28 0 2:*
30 39 12:*==
32 74 45:====*==
34 140 122:====*==
36 216 250:====*
38 469 414:====*====
40 476 577:====*
42 612 706:====*
44 761 779:====*
46 786 793:====*
48 610 759:====*
50 550 693:====*
52 428 609:====*
54 455 520:====*
56 390 435:====*
58 498 357:====*
60 587 289:====*
62 213 232:====*
64 165 184:====*
66 170 146:====*==
68 100 115:====*
70 250 90:====*====
72 79 70:====*
74 69 55:====*
76 83 43:====*==
78 36 33:==*
80 20 26:=*
82 27 20:=*
84 22 16:=*
86 6 12:*
88 11 9:*      inset = represents 1 library sequences
90 11 7:*
92 8 6:*      :====*==
94 9 4:*      :==*====
96 6 3:*      :==*====
98 0 3:*      : *
100 0 2:*      : *
102 2 2:*      :=*
104 0 1:*      :*
106 1 1:*      :*
108 1 1:*      :*
110 0 1:*      :*
112 1 0:=      *=
114 0 0:      *
116 0 0:      *
118 0 0:      *
>120 1 0:=      *=
2069351 residues in 8448 sequences
Expectation_n fit: rho(ln(x))= 4.19250.000782; mu= 20.9354 0.038
mean_var=65.320913.950, 0's: 60 Z-trim: 60 B-trim: 970 in 2/60
Lambda= 0.158689
Kolmogorov-Smirnov statistic: 0.0846 (N=29) at 56

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 44, opt: 32, open/ext: -10/-2, width: 16
The best scores are:
gi|6224908|gb|AAF06006.1| enterotoxin [Bacillus ce ( 431) 103 33.8 0.59

```

>>gi|6224908|gb|AAF06006.1| enterotoxin [Bacillus cereus (431 aa)

initn: 126 initl: 64 opt: 103 Z-score: 120.1 bits: 33.8 E(): 0.59  
Smith-Waterman score: 103; 26.708% identity (55.280% similar) in 161 aa overlap  
(1355-1507:37-191)

```

1330 1340 1350 1360 1370 1380
frame5 TVHNGSTFSECSCNSENWILVDHRRSTVFWNSNTLQSRRTNQVSHVFSTSKTWVDHFDV
      :...: . :...: .... : . . :
gi|622 GLAAASVAGVAVPGMDSAHAQVSNEALKEINGQTQTQTQTQTETKTETSELKYTV
      10      20      30      40      50      60

1390 1400 1410 1420 1430 1440
frame5 GTHFLQSQDETSTGW-VHQNVKDGVDGTWKNQGSNQWESSRGWITWNINHLTGELTTTFK
      . :. . . : : :...: : : :...: : : :...: :
gi|622 TADVLNVRSAGTGHNVISKVSGQV---LQVS--WDKKNGWFKVNVNGQTGYVSGDFV
      70      80      90      100      110      120

1450 1460 1470 1480 1490
frame5 TDGTHTVSISLNGK---VSTKTLKHLFSVVTSHDWLD--NSGDTWGVETSKQN-SGLHL
      : : . : : : :...: : : : : : : : : : :
gi|622 TTGGKTGTTVQQTGTYYTVNVSSLNVRTGPSTSHTVLGSVNGKKTQVVGEVQDWFKINF
      130      140      150      160      170      180

1500 1510 1520 1530 1540 1550
frame5 SGSHRYPVGDWRWSLWSTQGNWKTITVFRLLHLTHFAKWVQHTTHWTLSESVTNEAIE
      . : : : :
gi|622 NGGTGYVSKDFVTKGGSAVSNETQQPTTNNNTTVQTGGSYVNTGALKVRTGPATYNAV
      190      200      210      220      230      240

```

2416 residues in 1 query sequences  
2069351 residues in 8448 library sequences  
Scomplib [34t26]  
start: Wed Feb 3 23:39:16 2010 done: Wed Feb 3 23:39:16 2010  
Total Scan time: 0.580 Total Display time: 0.010

Function used was FASTA [version 3.4t26 July 7, 2006]

# fasta34 frame5.pep /genedata/1/db/PRT\_2010 -Q -E 1 -O frame5.pep\_prt.fasta  
FASTA searches a protein or DNA sequence data bank version 3.4t26 July 7, 2006  
Please cite:  
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

frame5, 2416 aa  
vs /genedata/1/db/PRT\_2010 library

```

      opt      E()
< 20 279140 0:=====
22 98 0:=      one = represents 26541 library sequences
24 198 17:*
26 638 374:*
28 2211 4039:*
30 13781 24535:*
32 63681 94869:====*
34 210831 257272:===== *
36 463480 528377:===== *
38 771123 873210:===== *
40 1103898 1218051:===== *
42 1367522 1488920:===== *
44 1544487 1642417:===== *

```



```

46 1592457 1672843:=====
48 1551973 1601554:=====
50 1424476 1461424:===== *
52 1253455 1284836:=====
54 1087495 1097475:=====
56 932186 916728:=====
58 777110 752616:=====
60 647455 609663:=====
62 528075 488769:=====
64 433943 388715:=====
66 355448 307229:=====
68 290487 241660:=====
70 230839 189379:=====
72 183375 147982:=====
74 147308 115376:=====
76 116077 89800:=====
78 90206 69800:=====
80 70566 54199:=====
82 58056 41461:=====
84 46127 32842:=====
86 37182 25412:=====
88 33331 19662:===== inset = represents 367 library sequences
90 22045 15213:=====
92 18328 11771:=====
94 14217 9108:=====
96 10364 7047:=====
98 8055 5453:=====
100 6427 4219:=====
102 4569 3265:=====
104 3809 2526:=====
106 3619 1954:=====
108 2413 1512:=====
110 1876 1170:=====
112 1535 905:=====
114 1054 701:=====
116 834 542:=====
118 604 419:=====
>120 2761 325:=====
4761287459 residues in 17815538 sequences
statistics sampled from 60000 to 17808747 sequences
Expectation_n fit: rho(ln(x))= 5.22320.000188; mu= 15.2050 0.010
mean_var=75.603515.055, 0's: 930 Z-trim: 932 B-trim: 0 in 0/64
Lambda= 0.147504
Kolmogorov-Smirnov statistic: 0.0455 (N=29) at 54

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 44, opt: 32, open/ext: -10/-2, width: 16
The best scores are:
opt bits E(17815538)
gi|90192131|gb|ABD91726.1| acyl-ACP thioesterase [ ( 416) 486 114.3 7.2e-22
gi|229410505|gb|ACQ57188.1| acyl acyl-carrier-prot ( 436) 310 76.8 1.4e-10
gi|133754638|gb|ABO38558.1| acyl ACP-thioesterase ( 413) 308 76.4 1.8e-10
gi|133754636|gb|ABO38557.1| acyl ACP-thioesterase ( 413) 308 76.4 1.8e-10
gi|133754630|gb|ABO38554.1| acyl ACP-thioesterase ( 413) 308 76.4 1.8e-10
gi|133754634|gb|ABO38556.1| acyl ACP-thioesterase ( 413) 308 76.4 1.8e-10
gi|133754632|gb|ABO38555.1| acyl ACP-thioesterase ( 413) 308 76.4 1.8e-10
gi|229410507|gb|ACQ57190.1| acyl acyl-carrier-prot ( 434) 301 74.9 5.3e-10
gi|229358082|gb|ACQ57189.1| acyl acyl-carrier-prot ( 420) 295 73.6 1.3e-09
gi|229410506|gb|ACQ63293.1| acyl acyl-carrier-prot ( 434) 295 73.6 1.3e-09
gi|217072522|gb|ACJ84621.1| unknown [Medicago trun ( 181) 273 68.7 1.7e-08
gi|248570279|gb|ACT09366.1| stearoyl acyl-ACP-thio ( 418) 269 68.1 5.8e-08
gi|156900676|gb|ABU96744.1| chloroplast acyl-ACP t ( 418) 269 68.1 5.8e-08
gi|83778888|gb|ABC47311.1| FATB [Populus tomentosa ( 421) 260 66.2 2.2e-07

```

```

gi|222866396|gb|EEF03527.1| predicted protein [Pop ( 409) 250 64.0 9.4e-07
gi|3410682|gb|AAC31104.1|I90742 Sequence 23 from p ( 417) 249 63.8 1.1e-06
gi|228385038|gb|ACQ28991.1| Sequence 21505 from pa ( 554) 241 62.2 4.5e-06
gi|147858246|emb|CAN81819.1| hypothetical protein ( 400) 232 60.2 1.3e-05
gi|4529965|emb|CAA02764.1| unnamed protein product ( 177) 227 58.9 1.5e-05
gi|222855220|gb|EEE92767.1| predicted protein [Pop ( 418) 231 60.0 1.6e-05
gi|118489518|gb|ABK96561.1| unknown [Populus trich ( 418) 231 60.0 1.6e-05
gi|81362039|gb|ABB71581.1| acyl-acyl carrier prote ( 414) 226 58.9 3.3e-05
gi|18073131|emb|CAC80370.1| Acyl-ACP thioesterase ( 430) 225 58.7 3.9e-05
gi|33325193|gb|AAQ08202.1| acyl-ACP thioesterase F ( 430) 224 58.5 4.5e-05
gi|60392558|gb|AAX19385.1| acyl-ACP thioesterase [ ( 430) 224 58.5 4.5e-05
gi|60392562|gb|AAX19387.1| acyl-ACP thioesterase [ ( 430) 224 58.5 4.5e-05
gi|60392550|gb|AAX19381.1| acyl-ACP thioesterase [ ( 430) 224 58.5 4.5e-05
gi|60392546|gb|AAX19379.1| acyl-ACP thioesterase [ ( 430) 224 58.5 4.5e-05
gi|60392560|gb|AAX19386.1| acyl-ACP thioesterase [ ( 430) 224 58.5 4.5e-05
gi|60392544|gb|AAX19378.1| acyl-ACP thioesterase [ ( 430) 224 58.5 4.5e-05
gi|60392548|gb|AAX19380.1| acyl-ACP thioesterase [ ( 430) 224 58.5 4.5e-05
gi|2687425|gb|AAB88824.1| FatB thioesterase [Helia ( 430) 224 58.5 4.5e-05
gi|60392542|gb|AAX19377.1| acyl-ACP thioesterase [ ( 430) 224 58.5 4.5e-05
gi|60392552|gb|AAX19382.1| acyl-ACP thioesterase [ ( 430) 224 58.5 4.5e-05
gi|60392554|gb|AAX19383.1| acyl-ACP thioesterase [ ( 430) 224 58.5 4.5e-05
gi|60392556|gb|AAX19384.1| acyl-ACP thioesterase [ ( 430) 224 58.5 4.5e-05
gi|3410692|gb|AAC31114.1|I90752 Sequence 33 from p ( 416) 223 58.3 5.1e-05
gi|18073133|emb|CAC80371.1| Acyl-ACP thioesterase ( 430) 222 58.1 6.1e-05
gi|229358077|gb|ACQ57187.1| acyl acyl-carrier-prot ( 423) 220 57.7 8e-05
gi|223545508|gb|EEF47013.1| palmitoyl-acyl carrier ( 419) 217 57.0 0.00012
gi|157417724|gb|ABV54795.1| acyl-ACP thioesterase ( 419) 217 57.0 0.00012
gi|495227|emb|CAA54060.1| acyl-ACP thioesterase [C ( 414) 216 56.8 0.00014
gi|12057063|emb|CAC19933.1| thioesterase type B [C ( 414) 216 56.8 0.00014
gi|2298907|emb|CAA02760.1| unnamed protein product ( 414) 216 56.8 0.00014
gi|2298917|emb|CAA02765.1| unnamed protein product ( 414) 216 56.8 0.00014
gi|222855219|gb|EEE92766.1| predicted protein [Pop ( 407) 214 56.4 0.00019
gi|1930081|gb|AAB51525.1| acyl-ACP thioesterase [G ( 406) 209 55.3 0.00039
gi|3410687|gb|AAC31109.1|I90747 Sequence 28 from p ( 415) 207 54.9 0.00054
gi|758701|gb|AAC48990.1| 16:0-ACP thioesterase pre ( 415) 207 54.9 0.00054
gi|270240896|emb|CBI28125.1| unnamed protein produ ( 408) 201 53.6 0.0013
gi|2459529|gb|AAB71729.1| acyl-ACP thioesterase [M ( 429) 201 53.6 0.0013
gi|6048397|gb|AAF02215.1|AF076535_1 palmitoyl-acyl ( 413) 198 53.0 0.002
gi|3859830|gb|AAC72882.1| thioesterase FatB1-1 [Cu ( 414) 195 52.3 0.0032
gi|223271795|emb|CAX36780.1| unnamed protein produ ( 166) 188 50.5 0.0044
gi|804948|emb|CAA85388.1| acyl-(acyl carrier prote ( 412) 189 51.0 0.0076
gi|16648915|gb|AAL24309.1| putative oleoyl-[acyl-c ( 412) 188 50.8 0.0088
gi|17978739|gb|AAL47363.1| putative oleoyl-[acyl-c ( 412) 188 50.8 0.0088
gi|15451198|gb|AAK96870.1| putative oleoyl-[acyl-c ( 412) 188 50.8 0.0088
gi|30725684|gb|AAP37864.1| Atlg08510 [Arabidopsis ( 412) 188 50.8 0.0088
gi|634003|emb|CAA85387.1| acyl-(acyl carrier prote ( 412) 188 50.8 0.0088
gi|6664317|gb|AAF22899.1|AC006932_16 T2G7.19 [Ara ( 412) 188 50.8 0.0088
gi|223271145|emb|CAX36777.1| unnamed protein produ ( 169) 178 48.4 0.019
gi|223271121|emb|CAX36759.1| unnamed protein produ ( 169) 178 48.4 0.019
gi|223271109|emb|CAX36750.1| unnamed protein produ ( 169) 178 48.4 0.019
gi|4104242|gb|AAD01982.1| palmitoyl-acyl carrier p ( 410) 181 49.3 0.025
gi|223271133|emb|CAX36768.1| unnamed protein produ ( 169) 172 47.1 0.047
gi|193290692|gb|ACF17654.1| putative acyl-ACP thio ( 421) 175 48.1 0.061
gi|223271129|emb|CAX36765.1| unnamed protein produ ( 168) 168 46.3 0.084
gi|223271105|emb|CAX36747.1| unnamed protein produ ( 168) 168 46.3 0.084
gi|223271101|emb|CAX36744.1| unnamed protein produ ( 168) 160 44.6 0.27
gi|223271117|emb|CAX36756.1| unnamed protein produ ( 168) 159 44.4 0.32
gi|223271141|emb|CAX36774.1| unnamed protein produ ( 168) 159 44.4 0.32
gi|135944700|gb|EBK87954.1| hypothetical protein G ( 525) 164 45.8 0.37
gi|90018255|gb|ABD83939.1| palmitoyl-ACP thioester ( 417) 157 44.2 0.86
gi|223271125|emb|CAX36762.1| unnamed protein produ ( 167) 152 42.9 0.89
gi|111610624|gb|ABH11710.1| palmitoyl-ACP thioeste ( 414) 156 44.0 0.99

```

>>gi|90192131|gb|ABD91726.1| acyl-ACP thioesterase [Glyc (416 aa)  
initn: 486 initl: 486 opt: 486 Z-score: 555.2 bits: 114.3 E(): 7.2e-22  
Smith-Waterman score: 490; 82.000% identity (84.000% similar) in 100 aa overlap  
(377-476:1-90)

```

      350      360      370      380      390      400
frame5 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
gi|901      MVATAATSSFFPVTSPSPDSGGAGSKLGGG
              10      20      30
```

```

      410      420      430      440      450      460
frame5 PANLGLKSKSASSGGLKAKAQAPSKINGTTVVTSKESFKHDDDLPLLIHVDFPDMKPF
gi|901 PANLGLKSKSASSGGLKAKAQAPSKINGTTVVTSKESFKHDDDLPS-----PPPRTFI
              40      50      60      70      80
```

```

      470      480      490      500      510      520
frame5 NTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADLDSP
gi|901 N---QLPDWSMLLAAITIFLAAEKQWMMLDWKPRRPDMLIDPFGIGKIVQDGLVFR
              90      100      110      120      130      140
```

>>gi|229410505|gb|ACQ57188.1| acyl acyl-carrier-protein (436 aa)  
initn: 329 initl: 153 opt: 310 Z-score: 352.5 bits: 76.8 E(): 1.4e-10  
Smith-Waterman score: 310; 60.000% identity (78.947% similar) in 95 aa overlap  
(375-464:13-106)

```

      350      360      370      380      390      400
frame5 GRSQHFLVLTFCFVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGG--AGSKL
gi|229      VDFCSELEGDSFFIMVATAATSAFFPVASPPPDGSAKTSKSL
              10      20      30      40
```

```

      410      420      430      440      450
frame5 GGGPANLG--GLKSKSASSGGL--KAKAQAPSKINGTTVVTSKESFKHDDDLPLLIHVDF
gi|229 GGPPANVDARGIKSKSAATGGLQVKANAQAPSKVNGTKVGVM--EGLKSEDDSPSLHQRTF
              50      60      70      80      90      100
```

```

      460      470      480      490      500      510
frame5 PDMKPFSENTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPED
gi|229 INQLPDWSMLLAAVITIFLAAEKQWMMLDWKPRRPDMLVDPPFGFGKIVEDGFVFRQNF
              110      120      130      140      150      160
```

>>gi|133754638|gb|ABO38558.1| acyl ACP-thioesterase [Ara (413 aa)  
initn: 270 initl: 216 opt: 308 Z-score: 350.6 bits: 76.4 E(): 1.8e-10  
Smith-Waterman score: 312; 58.416% identity (71.287% similar) in 101 aa overlap  
(378-476:1-87)

```

      350      360      370      380      390      400
frame5 QHFVLVLTFCFVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGGP
gi|133      MATAATASIFPVSPSPDAGADGNKLVGGS
              10      20      30
```

```

      410      420      430      440      450      460
frame5 ANLGLKSKSASSGGL--KAKAQAPSKINGTTVVTSKESFKHDDDLPLLIHVDFPDMKPF
              .. : : : : : : : : : : : : : : : : : : : : : : : : : : : :
              .. : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```

gi|133 VKLQGLSKSHASSGGLQVKAHAQAPPKINGSTV----ESLKHDDDLPS-----PPPRTF
              40      50      60      70      80
```

```

      470      480      490      500      510      520
frame5 SNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADLD
gi|133 IN---QLPDWSMLLAAITIFLAAEKQWMMLDWKPRRSDMLIDPFGIGRIVQDGLVFRQ
              90      100      110      120      130
```

>>gi|133754636|gb|ABO38557.1| acyl ACP-thioesterase [Ara (413 aa)  
initn: 270 initl: 216 opt: 308 Z-score: 350.6 bits: 76.4 E(): 1.8e-10  
Smith-Waterman score: 312; 58.416% identity (71.287% similar) in 101 aa overlap  
(378-476:1-87)

```

      350      360      370      380      390      400
frame5 QHFVLVLTFCFVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGGP
gi|133      MATAATASIFPVSPSPDAGADGNKLVGGS
              10      20      30
```

```

      410      420      430      440      450      460
frame5 ANLGLKSKSASSGGL--KAKAQAPSKINGTTVVTSKESFKHDDDLPLLIHVDFPDMKPF
gi|133 VKLQGLSKSHASSGGLQVKAHAQAPPKINGSTV----ESLKHDDDLPS-----PPPRTF
              40      50      60      70      80
```

```

      470      480      490      500      510      520
frame5 SNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADLD
gi|133 IN---QLPDWSMLLAAITIFLAAEKQWMMLDWKPRRSDMLIDPFGIGRIVQDGLVFRQ
              90      100      110      120      130
```

>>gi|133754630|gb|ABO38554.1| acyl ACP-thioesterase [Ara (413 aa)  
initn: 270 initl: 216 opt: 308 Z-score: 350.6 bits: 76.4 E(): 1.8e-10  
Smith-Waterman score: 312; 58.416% identity (71.287% similar) in 101 aa overlap  
(378-476:1-87)

```

      350      360      370      380      390      400
frame5 QHFVLVLTFCFVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGGP
gi|133      MATAATASIFPVSPSPDAGADGNKLVGGS
              10      20      30
```

```

      410      420      430      440      450      460
frame5 ANLGLKSKSASSGGL--KAKAQAPSKINGTTVVTSKESFKHDDDLPLLIHVDFPDMKPF
gi|133 VKLQGLSKSHASSGGLQVKAHAQAPPKINGSTV----ESLKHDDDLPS-----PPPRTF
              40      50      60      70      80
```

```

      470      480      490      500      510      520
frame5 SNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADLD
gi|133 IN---QLPDWSMLLAAITIFLAAEKQWMMLDWKPRRSDMLIDPFGIGRIVQDGLVFRQ
              90      100      110      120      130
```

>>gi|133754634|gb|ABO38556.1| acyl ACP-thioesterase [Ara (413 aa)  
initn: 270 initl: 216 opt: 308 Z-score: 350.6 bits: 76.4 E(): 1.8e-10  
Smith-Waterman score: 312; 58.416% identity (71.287% similar) in 101 aa overlap  
(378-476:1-87)

```

      350      360      370      380      390      400
frame5 QHFVLVLTFCFVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGGP
```



```

      470      480      490      500      510      520
frame5  SNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADLDFS

gi|217  DWSMLLAAITTIIFLAAEKQWMLDVKPRRSDMLIDPFGIGKIVQDGLVFSNFSSIRSYEI
      90      100      110      120      130      140

>>gi|248570279|gb|ACT09366.1| stearoyl acyl-ACP-thioeste (418 aa)
  initn: 196 init1: 155 opt: 269  Z-score: 305.6  bits: 68.1 E(): 5.8e-08
Smith-Waterman score: 269; 52.885% identity (67.308% similar) in 104 aa overlap
(377-476:1-92)

      350      360      370      380      390      400
frame5  SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|248  MVATAATSSFFPVTSSADS--KSTKIGSG
      10      20

      410      420      430      440      450      460
frame5  PANLGGGLKSKSASSGGL--KAKAQAPSKINGTTV--VTSKESFKHDDDLPLLIHVDFPDM
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|248  SASLGGIKSKPASSGGLQVKANAQAPPKINGSTVGYTTPVDSVKNEGDTPS-----PPP
      30      40      50      60      70      80

      470      480      490      500      510      520
frame5  KPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADL
      . : : . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|248  RTFIN----QLPDWSMLLAAITTIIFLAAEKQWMLDVKPRRSDMLIDPFLGLRIVQDGLV
      90      100      110      120      130

>>gi|156900676|gb|ABU96744.1| chloroplast acyl-ACP thioe (418 aa)
  initn: 196 init1: 155 opt: 269  Z-score: 305.6  bits: 68.1 E(): 5.8e-08
Smith-Waterman score: 269; 52.885% identity (67.308% similar) in 104 aa overlap
(377-476:1-92)

      350      360      370      380      390      400
frame5  SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|156  MVATAATSSFFPVTSSADS--KSTKIGSG
      10      20

      410      420      430      440      450      460
frame5  PANLGGGLKSKSASSGGL--KAKAQAPSKINGTTV--VTSKESFKHDDDLPLLIHVDFPDM
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|156  SASLGGIKSKPASSGGLQVKANAQAPPKINGSTVGYTTPVDSVKNEGDTPS-----PPP
      30      40      50      60      70      80

      470      480      490      500      510      520
frame5  KPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADL
      . : : . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|156  RTFIN----QLPDWSMLLAAITTIIFLAAEKQWMLDVKPRRSDMLIDPFLGLRIVQDGLV
      90      100      110      120      130

>>gi|83778888|gb|ABC47311.1| FATB [Populus tomentosa] (421 aa)
  initn: 293 init1: 161 opt: 260  Z-score: 295.2  bits: 66.2 E(): 2.2e-07
Smith-Waterman score: 260; 52.885% identity (70.192% similar) in 104 aa overlap
(377-476:1-92)

      350      360      370      380      390      400
frame5  SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|837  MVATAATSSFFPV--PSPPGDAKSSKVGSG

```

```

      10      20

      410      420      430      440      450      460
frame5  PANLGGGLKSKSASSGGL--KAKAQAPSKINGTTV--VTSKESFKHDDDLPLLIHVDFPDM
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|837  SASLGGIKSKSASSGALQVKANAQAPPKINGSPVGLTASVETAKKEDV-----VSSPAP
      30      40      50      60      70      80

      470      480      490      500      510      520
frame5  KPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADL
      . : : . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|837  RTFIN----QLPDWSMLLAAITTMFLAAEKQWMLDVKPKRADMLIDPFGIGRIVQDGLV
      90      100      110      120      130

>>gi|222866396|gb|EEF03527.1| predicted protein [Populus (409 aa)
  initn: 289 init1: 157 opt: 250  Z-score: 283.9  bits: 64.0 E(): 9.4e-07
Smith-Waterman score: 250; 60.000% identity (76.250% similar) in 80 aa overlap
(377-452:1-78)

      350      360      370      380      390      400
frame5  SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|222  MVATAATSSFFPV--PSPPGDAKSSKVGSG
      10      20

      410      420      430      440      450      460
frame5  PANLGGGLKSKSASSGGL--KAKAQAPSKINGTT--VVTSKESFKHDDDLPLLIHVDFPDM
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|222  SASLGGIKSKSASSGALQVKANAQAPPKINGSPEDVVSSPAPRTFINQLPDWSMLLAAIT
      30      40      50      60      70      80

      470      480      490      500      510      520
frame5  KPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADL

gi|222  TMFLAAEKQWMLDVKPKRADMLIDPFGIGRIVQDGLVFSQNFSIRSSEIGADRTASIEI
      90      100      110      120      130      140

>>gi|3410682|gb|AAC31104.1|I90742 Sequence 23 from paten (417 aa)
  initn: 264 init1: 96 opt: 249  Z-score: 282.6  bits: 63.8 E(): 1.1e-06
Smith-Waterman score: 249; 57.500% identity (82.500% similar) in 80 aa overlap
(377-452:1-77)

      350      360      370      380      390      400
frame5  SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|341  MASTAATACFFPVSSPSSSDSVAKTKNIGS-
      10      20

      410      420      430      440      450      460
frame5  PANLGGGLKSKSASSGGL--KAKAQAPSKINGTTVVTSK--ESFKHDDDLPLLIHVDFPDM
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|341  -ASLGGMKAKS--SSGGLQVKASAQAPSKINGTSVGLTKPSESLKNEDEMPSSHPRTFINQ
      30      40      50      60      70      80

      470      480      490      500      510      520
frame5  KPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADL

gi|341  LPDWSMLLAAITTIIFLAAEKQWMLDVKPRRSDMLIDPFGIGRIVQDGLIFRQNFIRSISY
      90      100      110      120      130      140

>>gi|228385038|gb|ACQ28991.1| Sequence 21505 from patent (554 aa)
  initn: 222 init1: 183 opt: 241  Z-score: 271.7  bits: 62.2 E(): 4.5e-06

```

Smith-Waterman score: 241; 27.746% identity (65.318% similar) in 173 aa overlap (1480-1651:1-173)

```

      1450      1460      1470      1480      1490      1500
frame5 VSISLNGKVSTKTLKHLFSVVTSHDWLNSGDTWGVETSKQNSGLHLSGSHRYPVGDWRW
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|228                               GLTRGVPEGEQQAGLHLGAGHRHVVMHTGE
                               10      20      30

      1510      1520      1530      1540      1550      1560
frame5 SLWSTQGNWKTITVFRHLHLHTFAKWVQHTTHWTLSESVTNESAIEIVNTNKTHTSQATAT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|228 ALAAANGHRRRTAFRPRLDYRAHLAQRLGYAIHRPLGKGSVAGQGAVERLGRQQTAEQTHG
      40      50      60      70      80      90

      1570      1580      1590      1600      1610      1620
frame5 SCSVTEIERSLSRKESTVTNTIN-DPSITFLTDSGTHSLHSLTSDNIFTFKTGDTSTFT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|228 STGVTQVDRPRRLEAVQANAMDGDAAVLRPLDHHSHAERLDGCQCVALAEEALHFGGA
      100     110     120     130     140     150

      1630      1640      1650      1660      1670      1680
frame5 ASETSKHEGPVGDRVLVTWNTDGSRKTRGLTSSCWTAACTVKHARRGNRRHLKRTKLRAN
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|228 FGQGAEHDRAMGNRLVARYPDGSGQAAARLGQINQVVVHVGHIGPTQGDFAEMLTGYPG
      160     170     180     190     200     210
```

>>gi|147858246|emb|CAN81819.1| hypothetical protein [Vit (400 aa)  
initn: 186 init1: 90 opt: 232 Z-score: 263.3 bits: 60.2 E(): 1.3e-05  
Smith-Waterman score: 232; 49.524% identity (64.762% similar) in 105 aa overlap (377-476:1-92)

```

      350      360      370      380      390      400
frame5 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|147                               MVATAATSAFFAVASPSDDPAKPKPGX
                               10      20      30

      410      420      430      440      450      460
frame5 GPANLGGGLKSKSASSGGL--KAKAQAPSKINGTTV--VTSKESFKHDDDLPLLIHVDFPD
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|147 GSAILRGIKSRNAPSGSLQVKANAQAPPKINGTTVGYTSSAEGVKIEDDM-----SSPP
      40      50      60      70      80

      470      480      490      500      510      520
frame5 MKPFNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|147 PRTFIN---QLPDWSMLLAAITTFILAAEKQWMMLDWKPRRSDMLIDPFLGKIVQDGL
      90      100     110     120     130     140
```

>>gi|4529965|emb|CAA02764.1| unnamed protein product [Cu (177 aa)  
initn: 240 init1: 137 opt: 227 Z-score: 262.5 bits: 58.9 E(): 1.5e-05  
Smith-Waterman score: 233; 45.098% identity (67.647% similar) in 102 aa overlap (377-476:1-91)

```

      350      360      370      380      390      400
frame5 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|452                               MVATAASSAFFPV--PSADTSSRPGLGNG
                               10      20

      410      420      430      440      450      460
```

```

frame5 PANLGGGLKSKSASSGGL--KAKAQAPSKINGTTVVTSKESFKHDDDLPLLIHVDFPDMKP
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|452 PSSFSPLPKSIPNGGLQVKASASAPPKINGSSVGLKSGGLKTHDDAP-----SAPPPRT
      30      40      50      60      70      80

      470      480      490      500      510      520
frame5 FSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADLDF
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|452 FIN---QLPDWSMLLAAITTAFLAAEKQWMMLDWKPKRDLMDLEDPFGLGRIVQDGLVFR
      90      100     110     120     130
```

>>gi|222855220|gb|EEE92767.1| predicted protein [Populus (418 aa)  
initn: 277 init1: 132 opt: 231 Z-score: 261.9 bits: 60.0 E(): 1.6e-05  
Smith-Waterman score: 231; 48.077% identity (69.231% similar) in 104 aa overlap (377-476:1-92)

```

      350      360      370      380      390      400
frame5 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|222                               MVAASASSFFPVPSPSGDA--KASKFGSV
                               10      20

      410      420      430      440      450      460
frame5 PANLGGGLKSKSASSGGLKAK--AQAPSKINGTTV--VTSKESFKHDDDLPLLIHVDFPDM
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|222 SASLGGIKTKSASSGALQVNTNAQAPPKINGPPVGLTASVETLKNEVDV-----VSSPAP
      30      40      50      60      70      80

      470      480      490      500      510      520
frame5 KPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADL
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|222 RTFIN---QLPDWSMLLAAITTMFLAAEKQWMMLDWKPKRPDMLIDPFGIGRIVQDGLV
      90      100     110     120     130
```

>>gi|118489518|gb|ABK96561.1| unknown [Populus trichocar (418 aa)  
initn: 272 init1: 132 opt: 231 Z-score: 261.9 bits: 60.0 E(): 1.6e-05  
Smith-Waterman score: 231; 48.077% identity (69.231% similar) in 104 aa overlap (377-476:1-92)

```

      350      360      370      380      390      400
frame5 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|118                               MVATAAASSFFPVPSPSGDA--KASKFGSV
                               10      20

      410      420      430      440      450      460
frame5 PANLGGGLKSKSASSGGLKAKA--QAPSKINGTTV--VTSKESFKHDDDLPLLIHVDFPDM
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|118 SASLGGIKTKSASSGALQVNTNGQAPPKINGPPVGLAASVETLKNEVDV-----VSSPAP
      30      40      50      60      70      80

      470      480      490      500      510      520
frame5 KPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADL
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|118 RTFIN---QLPDWSMLLAAITTMFLAAEKQWMMLDWKPKRPDMLIDPFGIGRIVQDGLV
      90      100     110     120     130
```

>>gi|81362039|gb|ABB71581.1| acyl-acyl carrier protein t (414 aa)  
initn: 263 init1: 126 opt: 226 Z-score: 256.2 bits: 58.9 E(): 3.3e-05  
Smith-Waterman score: 232; 45.098% identity (68.627% similar) in 102 aa overlap (377-476:1-91)

```

      350      360      370      380      390      400
frame5 SQHFVLVTF CFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
      :::::::::::::: ::::: :::::
gi|181      MVATAASSAFFPV--PSPDTSSRPGKLGNG
              10              20

      410      420      430      440      450      460
frame5 PANLGLKSKSASSGGL--KAKAQAPSKINGTTVVTSKESFKHDDDLPLLIHVDFPDMKP
      :.:.: :.:.: :.:.: :.:.: :.:.: :.:.:
gi|181 SSSLSP LKPKLMANGGLQVKANASAPPKINGSSVGLKSGSLKTQEDTP-----SAPPPRT
      30      40      50      60      70      80

      470      480      490      500      510      520
frame5 FSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADLDF
      :.:.: :.:.:
gi|181 FIN---QLPDWSMLLAIAITTVFLAAEKQWMMMLDWKPKRPDMLVDPFGLGRIVQDGLVFR
      90      100      110      120      130

>>gi|18073131|emb|CAC80370.1| Acyl-ACP thioesterase [Hel (430 aa)
      initn: 196 init1: 106 opt: 225 Z-score: 254.9 bits: 58.7 E(): 3.9e-05
Smith-Waterman score: 225; 42.056% identity (67.290% similar) in 107 aa overlap
(377-478:1-96)

      350      360      370      380      390      400
frame5 SQHFVLVTF CFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
      :: :::::::::::::: ::. : ::::
gi|180      MVAMSATASLFPVSSPKPHSGAKTSDKLGG
              10              20              30

      410      420      430      440      450      460
frame5 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLPLLIHVDFPD
      :.:.: :.:.: :.:.: :.:.: :.:.: :.:.:
gi|180 EPGSVAVRGIKTKSVNSGGMKV KANAQAPTEVNGSRSRITH-GFKTDD-----YSTSPA
      40      50      60      70      80

      470      480      490      500      510      520
frame5 MKPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
      :.:.: :.:.:
gi|180 PRTFINX---LPDWXMLLAIAITTVFLAAEKQWMMLEWTKRPDMIADMDPFGLGRIVQD
      90      100      110      120      130

>>gi|33325193|gb|AAQ08202.1| acyl-ACP thioesterase FATB (430 aa)
      initn: 196 init1: 106 opt: 224 Z-score: 253.7 bits: 58.5 E(): 4.5e-05
Smith-Waterman score: 224; 42.857% identity (67.619% similar) in 105 aa overlap
(377-476:1-94)

      350      360      370      380      390      400
frame5 SQHFVLVTF CFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
      :: :::::::::::::: ::. : ::::
gi|333      MVAMSATASLFPVSSPKPHSGAKTSDKLGG
              10              20              30

      410      420      430      440      450      460
frame5 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLPLLIHVDFPD
      :.:.: :.:.: :.:.: :.:.: :.:.: :.:.:
gi|333 EPGSVAVRGIKTKSVNSGGMKV KANAQAPTEVNGSRSRITH-GFKTDD-----YSTSPA
      40      50      60      70      80

      470      480      490      500      510      520
frame5 MKPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
      :.:.: :.:.:
gi|333 PRTFIN---QLPDWSMLLAIAITTVFLAAEKQWMMLEWTKRPDMIADMDPFGLGRIVQD
      90      100      110      120      130

```

```

      90      100      110      120      130

>>gi|60392558|gb|AAX19385.1| acyl-ACP thioesterase [Heli (430 aa)
      initn: 196 init1: 106 opt: 224 Z-score: 253.7 bits: 58.5 E(): 4.5e-05
Smith-Waterman score: 224; 42.857% identity (67.619% similar) in 105 aa overlap
(377-476:1-94)

      350      360      370      380      390      400
frame5 SQHFVLVTF CFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
      :: :::::::::::::: ::. : ::::
gi|603      MVAMSATASLFPVSSPKPHSGAKTSDKLGG
              10              20              30

      410      420      430      440      450      460
frame5 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLPLLIHVDFPD
      :.:.: :.:.: :.:.: :.:.: :.:.: :.:.:
gi|603 EPGSVAVRGIKTKSVNSGGMKV KANAQAPTEVNGSRSRITH-GFKTDD-----YSTSPA
      40      50      60      70      80

      470      480      490      500      510      520
frame5 MKPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
      :.:.: :.:.:
gi|603 PRTFIN---QLPDWSMLLAIAITTVFLAAEKQWMMLEWTKRPDMIADMDPFGLGRIVQD
      90      100      110      120      130

>>gi|60392562|gb|AAX19387.1| acyl-ACP thioesterase [Heli (430 aa)
      initn: 196 init1: 106 opt: 224 Z-score: 253.7 bits: 58.5 E(): 4.5e-05
Smith-Waterman score: 224; 42.857% identity (67.619% similar) in 105 aa overlap
(377-476:1-94)

      350      360      370      380      390      400
frame5 SQHFVLVTF CFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
      :: :::::::::::::: ::. : ::::
gi|603      MVAMSATASLFPVSSPKPHSGAKTSDKLGG
              10              20              30

      410      420      430      440      450      460
frame5 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLPLLIHVDFPD
      :.:.: :.:.: :.:.: :.:.: :.:.: :.:.:
gi|603 EPGSVAVRGIKTKSVNSGGMKV KANAQAPTEVNGSRSRITH-GFKTDD-----YSTSPA
      40      50      60      70      80

      470      480      490      500      510      520
frame5 MKPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
      :.:.: :.:.:
gi|603 PRTFIN---QLPDWSMLLAIAITTVFLAAEKQWMMLEWTKRPDMIADMDPFGLGRIVQD
      90      100      110      120      130

>>gi|60392550|gb|AAX19381.1| acyl-ACP thioesterase [Heli (430 aa)
      initn: 196 init1: 106 opt: 224 Z-score: 253.7 bits: 58.5 E(): 4.5e-05
Smith-Waterman score: 224; 42.857% identity (67.619% similar) in 105 aa overlap
(377-476:1-94)

      350      360      370      380      390      400
frame5 SQHFVLVTF CFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
      :: :::::::::::::: ::. : ::::
gi|603      MVAMSATASLFPVSSPKPHSGAKTSDKLGG
              10              20              30

      410      420      430      440      450      460
frame5 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLPLLIHVDFPD
      :.:.: :.:.: :.:.: :.:.: :.:.: :.:.:

```

```

gi|603 EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDD-----YSTSPA
          40      50      60      70      80

frame5 MKPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
      . : :      . : :
gi|603 PRTFIN---QLPDWSMLLAAITTIFLAAEKQWMMLEWTKRPDMIADMDPFGGLGRIVQD
          90      100     110     120     130

>>gi|60392546|gb|AAx19379.1| acyl-ACP thioesterase [Heli (430 aa)
  initn: 196 initl: 106 opt: 224 Z-score: 253.7 bits: 58.5 E(): 4.5e-05
Smith-Waterman score: 224; 42.857% identity (67.619% similar) in 105 aa overlap
(377-476:1-94)

          350      360      370      380      390      400
frame5 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
      . : :      . : :      . : :      . : :      . : :
gi|603 MVAMSATASLFPVSSPKPHSGAKTSDKLGG
          10      20      30

          410      420      430      440      450      460
frame5 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDLPLLIHVDFPD
      . : :      . : :      . : :      . : :      . : :
gi|603 EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDD-----YSTSPA
          40      50      60      70      80

frame5 MKPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
      . : :      . : :
gi|603 PRTFIN---QLPDWSMLLAAITTIFLAAEKQWMMLEWTKRPDMIADMDPFGGLGRIVQD
          90      100     110     120     130

>>gi|60392560|gb|AAx19386.1| acyl-ACP thioesterase [Heli (430 aa)
  initn: 196 initl: 106 opt: 224 Z-score: 253.7 bits: 58.5 E(): 4.5e-05
Smith-Waterman score: 224; 42.857% identity (67.619% similar) in 105 aa overlap
(377-476:1-94)

          350      360      370      380      390      400
frame5 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
      . : :      . : :      . : :      . : :      . : :
gi|603 MVAMSATASLFPVSSPKPHSGAKTSDKLGG
          10      20      30

          410      420      430      440      450      460
frame5 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDLPLLIHVDFPD
      . : :      . : :      . : :      . : :      . : :
gi|603 EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDD-----YSTSPA
          40      50      60      70      80

frame5 MKPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
      . : :      . : :
gi|603 PRTFIN---QLPDWSMLLAAITTIFLAAEKQWMMLEWTKRPDMIADMDPFGGLGRIVQD
          90      100     110     120     130

>>gi|60392544|gb|AAx19378.1| acyl-ACP thioesterase [Heli (430 aa)
  initn: 196 initl: 106 opt: 224 Z-score: 253.7 bits: 58.5 E(): 4.5e-05
Smith-Waterman score: 224; 42.857% identity (67.619% similar) in 105 aa overlap
(377-476:1-94)

          350      360      370      380      390      400
frame5 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG

```

```

          410      420      430      440      450      460
frame5 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDLPLLIHVDFPD
      . : :      . : :      . : :      . : :      . : :
gi|603 EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDD-----YSTSPA
          40      50      60      70      80

frame5 MKPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
      . : :      . : :
gi|603 PRTFIN---QLPDWSMLLAAITTIFLAAEKQWMMLEWTKRPDMIADMDPFGGLGRIVQD
          90      100     110     120     130

>>gi|60392548|gb|AAx19380.1| acyl-ACP thioesterase [Heli (430 aa)
  initn: 196 initl: 106 opt: 224 Z-score: 253.7 bits: 58.5 E(): 4.5e-05
Smith-Waterman score: 224; 42.857% identity (67.619% similar) in 105 aa overlap
(377-476:1-94)

          350      360      370      380      390      400
frame5 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
      . : :      . : :      . : :      . : :      . : :
gi|603 MVAMSATASLFPVSSPKPHSGAKTSDKLGG
          10      20      30

          410      420      430      440      450      460
frame5 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDLPLLIHVDFPD
      . : :      . : :      . : :      . : :      . : :
gi|603 EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDD-----YSTSPA
          40      50      60      70      80

frame5 MKPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
      . : :      . : :
gi|603 PRTFIN---QLPDWSMLLAAITTIFLAAEKQWMMLEWTKRPDMIADMDPFGGLGRIVQD
          90      100     110     120     130

>>gi|2687425|gb|AAB88824.1| FatB thioesterase [Helianthu (430 aa)
  initn: 196 initl: 106 opt: 224 Z-score: 253.7 bits: 58.5 E(): 4.5e-05
Smith-Waterman score: 224; 42.857% identity (67.619% similar) in 105 aa overlap
(377-476:1-94)

          350      360      370      380      390      400
frame5 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
      . : :      . : :      . : :      . : :      . : :
gi|268 MVAMSATASLFPVSSPKPHSGAKTSDKLGG
          10      20      30

          410      420      430      440      450      460
frame5 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDLPLLIHVDFPD
      . : :      . : :      . : :      . : :      . : :
gi|268 EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGSRSRITH-GFKTDD-----YSTSPA
          40      50      60      70      80

frame5 MKPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
      . : :      . : :
gi|268 PRTFIN---QLPDWSMLLAAITTIFLAAEKQWMMLEWTKRPDMIADMDPFGGLGRIVQD
          90      100     110     120     130

```

>>gi|60392542|gb|AA19377.1| acyl-ACP thioesterase [Heli (430 aa)  
initn: 196 init1: 106 opt: 224 Z-score: 253.7 bits: 58.5 E(): 4.5e-05  
Smith-Waterman score: 224; 42.857% identity (67.619% similar) in 105 aa overlap  
(377-476:1-94)

```
350      360      370      380      390      400
frame5 SQHFVLVTFCFMMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
gi|603      MVAMSATASLFPVSSPKPHSGAKTSDKLGG
10      20      30
```

```
410      420      430      440      450      460
frame5 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLPLLIHVDFPD
gi|603      EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGRSRITH-GFKTDD-----YSTSPA
40      50      60      70      80
```

```
470      480      490      500      510      520
frame5 MKPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
gi|603      PRTFIN---QLPDWSMLLAIAITTFILAAEKQWMMLEWTKRPMIADMDPFGLGRIVQD
90      100      110      120      130
```

>>gi|60392552|gb|AA19382.1| acyl-ACP thioesterase [Heli (430 aa)  
initn: 196 init1: 106 opt: 224 Z-score: 253.7 bits: 58.5 E(): 4.5e-05  
Smith-Waterman score: 224; 42.857% identity (67.619% similar) in 105 aa overlap  
(377-476:1-94)

```
350      360      370      380      390      400
frame5 SQHFVLVTFCFMMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
gi|603      MVAMSATASLFPVSSPKPHSGAKTSDKLGG
10      20      30
```

```
410      420      430      440      450      460
frame5 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLPLLIHVDFPD
gi|603      EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGRSRITH-GFKTDD-----YSTSPA
40      50      60      70      80
```

```
470      480      490      500      510      520
frame5 MKPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
gi|603      PRTFIN---QLPDWSMLLAIAITTFILAAEKQWMMLEWTKRPMIADMDPFGLGRIVQD
90      100      110      120      130
```

>>gi|60392554|gb|AA19383.1| acyl-ACP thioesterase [Heli (430 aa)  
initn: 196 init1: 106 opt: 224 Z-score: 253.7 bits: 58.5 E(): 4.5e-05  
Smith-Waterman score: 224; 42.857% identity (67.619% similar) in 105 aa overlap  
(377-476:1-94)

```
350      360      370      380      390      400
frame5 SQHFVLVTFCFMMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
gi|603      MVAMSATASLFPVSSPKPHSGAKTSDKLGG
10      20      30
```

```
410      420      430      440      450      460
frame5 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLPLLIHVDFPD
gi|603      EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGRSRITH-GFKTDD-----YSTSPA
40      50      60      70      80
```

```
470      480      490      500      510      520
frame5 MKPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
gi|603      PRTFIN---QLPDWSMLLAIAITTFILAAEKQWMMLEWTKRPMIADMDPFGLGRIVQD
90      100      110      120      130
```

>>gi|60392556|gb|AA19384.1| acyl-ACP thioesterase [Heli (430 aa)  
initn: 196 init1: 106 opt: 224 Z-score: 253.7 bits: 58.5 E(): 4.5e-05  
Smith-Waterman score: 224; 42.857% identity (67.619% similar) in 105 aa overlap  
(377-476:1-94)

```
350      360      370      380      390      400
frame5 SQHFVLVTFCFMMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
gi|603      MVAMSATASLFPVSSPKPHSGAKTSDKLGG
10      20      30
```

```
410      420      430      440      450      460
frame5 GPANLG--GLKSKSASSGGLKAKA--QAPSKINGTTVVTSKESFKHDDDLPLLIHVDFPD
gi|603      EPGSVAVRGIKTKSVNSGGMKVKANAQAPTEVNGRSRITH-GFKTDD-----YSTSPA
40      50      60      70      80
```

```
470      480      490      500      510      520
frame5 MKPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
gi|603      PRTFIN---QLPDWSMLLAIAITTFILAAEKQWMMLEWTKRPMIADMDPFGLGRIVQD
90      100      110      120      130
```

>>gi|3410692|gb|AAC31114.1|I90752 Sequence 33 from paten (416 aa)  
initn: 191 init1: 85 opt: 223 Z-score: 252.8 bits: 58.3 E(): 5.1e-05  
Smith-Waterman score: 227; 50.000% identity (69.231% similar) in 104 aa overlap  
(377-476:1-90)

```
350      360      370      380      390      400
frame5 SQHFVLVTFCFMMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
gi|341      MASTAATAAFFPVSS-STDSVAKPKNIGS-
10      20
```

```
410      420      430      440      450      460
frame5 PANLGLKSKSASSGGL--KAKAQAPSKINGTTVVTSK--ESFKHDDDLPLLIHVDFPDM
gi|341      -AGLGLKSKS--SSGRLQVKATAQAPSKINGTSVGLTKPVEGLKNEDDMPS-----PPP
30      40      50      60      70      80
```

```
470      480      490      500      510      520
frame5 KPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADL
gi|341      RTFIN---QLPDWSMLLAIAITTFILAAEKQWMLLDWKPRSDMLDIPFGIGRIVQDGLI
90      100      110      120      130
```

>>gi|18073133|emb|CAC80371.1| Acyl-ACP thioesterase [Hel (430 aa)  
initn: 196 init1: 106 opt: 222 Z-score: 251.4 bits: 58.1 E(): 6.1e-05  
Smith-Waterman score: 222; 50.000% identity (79.487% similar) in 78 aa overlap  
(377-449:1-77)

```
350      360      370      380      390      400
frame5 SQHFVLVTFCFMMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
gi|180      MVAMSATASLFPVSSPKPHSGAKTSDKLGG
```



```

                10      20      30
frame5   410      420      430      440      450      460
        :....  :.....:.....:.....:  ..  :..  :
gi|180    EPGSVAVRGIKTSVNSGGMKVKANQAPTENVNGRSRITH-GFKTDDYSTSPAXXTFIX
                40      50      60      70      80

frame5   470      480      490      500      510      520
        MKPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
gi|180    QLPDWSMLLAAITTIFLAAEKQWMMLEWTKRPMIADMDPFGLRIVQDGLVFRQNFSI
        90      100      110      120      130      140

>>gi|229358077|gb|ACQ57187.1| acyl acyl-carrier-protein (423 aa)
  initn: 208 initl: 110 opt: 220 Z-score: 249.2 bits: 57.7 E(): 8e-05
Smith-Waterman score: 220; 53.086% identity (76.543% similar) in 81 aa overlap
(377-452:1-80)

        350      360      370      380      390      400
frame5    SQHFVLVTF CFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDGGAGSKLGGG
        :.....:.....:.....:.....:  :  :..  :
gi|229      MVATAATSAFFAVASLQPESGAKTSGKGGG
                10      20      30

        410      420      430      440      450      460
frame5    -PANLG--GLKSKSASSGGLKAK--AQAPSKINGTTVVTSKESFKHDDDLPLLIHVDFPD
        :...  :.....:.....:.....:  .  :..  :..  :
gi|229    VPSTVDARGIKSKSASSGSLQVRTNAQAPSKVNGTKFGVV-EGLKGENDTPSHHERTIIN
                40      50      60      70      80

frame5   470      480      490      500      510      520
        MKPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
gi|229    QLPDWSMLLAAITTIFLAAEKQWMLLDWKPRRPDMLTDPFGLGKIVQDGLVFSQNFIRS
        90      100      110      120      130      140

>>gi|223545508|gb|EEF47013.1| palmitoyl-acyl carrier pro (419 aa)
  initn: 152 initl: 114 opt: 217 Z-score: 245.8 bits: 57.0 E(): 0.00012
Smith-Waterman score: 222; 47.170% identity (65.094% similar) in 106 aa overlap
(376-476:2-93)

        350      360      370      380      390      400
frame5    RSQHFVLVTF CFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDGGAGSKLGG
        .  :.....:.....:.....:  :  :  :  :
gi|223      MVATAAAATSSFFPVPSQSAD-----ANFDK
                10      20

        410      420      430      440      450      460
frame5    GPANLGGLKSKSAS-SGGL--KAKAQAPSKINGTTV--VTSKESFKHDDDLPLLIHVDFDP
        :.....:.....:.....:.....:  :  :  :  :  :  :
gi|223    APASLGGIKLKSTSCSRGLQVKANAQAPPKINGSSVGFTTSVETVKNDGDMPLP-----P
        30      40      50      60      70      80

frame5   470      480      490      500      510      520
        DMKPFSTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDA
        .  :  :  :  :  :
gi|223    PPRTFIN----QLPDWSMLLAAITTIFLAAEKQWMLLDWKPRRPDMLIDPFGIGRIVQDG
        90      100      110      120      130

>>gi|157417724|gb|ABV54795.1| acyl-ACP thioesterase [Ric (419 aa)
  initn: 152 initl: 114 opt: 217 Z-score: 245.8 bits: 57.0 E(): 0.00012

```

Smith-Waterman score: 222; 47.170% identity (65.094% similar) in 106 aa overlap (376-476:2-93)

```

        350      360      370      380      390      400
frame5    RSQHFLVLTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDGGAGSKLGG
        .  :.....:.....:.....:  :  :  :  :
gi|157      MVATAAAATSSFFPVPSQSAD-----ANFDK
                10      20

        410      420      430      440      450      460
frame5    GPANLGGLKSKSAS-SGGL--KAKAQAPSKINGTTV--VTSKESFKHDDDLPLLIHVDFP
        :.....:.....:.....:.....:  :  :  :  :  :  :
gi|157    APASLGGIKLKSTSCSRGLQVKANAQAPPKINGSSVGFTTSVETVKNDGDMPLP-----P
        30      40      50      60      70      80

frame5   470      480      490      500      510      520
        DMKPFSTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDA
        .  :  :  :  :  :
gi|157    PPRTFIN----QLPDWSMLLAAITTIFLAAEKQWMLLDWKPRRPDMLIDPFGIGRIVQDG
        90      100      110      120      130

```

>>gi|495227|emb|CAA54060.1| acyl-ACP thioesterase [Cuphe (414 aa)  
 initn: 220 initl: 122 opt: 216 Z-score: 244.7 bits: 56.8 E(): 0.00014  
 Smith-Waterman score: 224; 43.137% identity (69.608% similar) in 102 aa overlap  
 (377-476:1-91)

```

        350      360      370      380      390      400
frame5    SQHFVLVLTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDGGAGSKLGGG
        :.....:.....:.....:.....:  :  :..  :
gi|495      MVATAASSAFFPL--PSPDTSSRPGLKNG
                10      20

        410      420      430      440      450      460
frame5    PANLGLKSKSASSGGLKAKAQA--PSKINGTTVVTSKESFKHDDDLPLLIHVDFPDMKP
        ....  :  :  :.....:.....:.....:  :  :  :  :  :  :
gi|495    SSSLSPKPKFVANAGLKVKASASAPPKINGSSVGLKSGSLKTQEDTP-----SVPPPRT
        30      40      50      60      70      80

frame5   470      480      490      500      510      520
        FSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADLDF
        :  :  :  :  :
gi|495    FIN----QLPDWSMLLAAITTVFLAAEKQWMLLDWKPRPDMLVDPFGLGSIVQGLVFR
        90      100      110      120      130

```

>>gi|12057063|emb|CAC19933.1| thioesterase type B [Cuphe (414 aa)  
 initn: 220 initl: 122 opt: 216 Z-score: 244.7 bits: 56.8 E(): 0.00014  
 Smith-Waterman score: 224; 43.137% identity (69.608% similar) in 102 aa overlap  
 (377-476:1-91)

```

        350      360      370      380      390      400
frame5    SQHFVLVLTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDGGAGSKLGGG
        :.....:.....:.....:.....:  :  :..  :
gi|120      MVATAASSAFFPL--PSPDTSSRPGLKNG
                10      20

        410      420      430      440      450      460
frame5    PANLGLKSKSASSGGLKAKAQA--PSKINGTTVVTSKESFKHDDDLPLLIHVDFPDMKP
        ....  :  :  :.....:.....:.....:  :  :  :  :  :  :
gi|120    SSSLSPKPKFVANAGLKVKASASAPPKINGSSVGLKSGSLKTQEDTP-----SVPPPRT
        30      40      50      60      70      80

        470      480      490      500      510      520

```

```
frame5 FSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADLDF
: : .: :
gi|120 FIN---QLPDWSMLLAAITTVFLAAEKQWMMMLDWKPKRPDMLVDPFGLGSIVQGGLVFR
          90          100          110          120          130
```

>>gi|2298907|emb|CAA02760.1| unnamed protein product [Cu (414 aa)  
initn: 220 init1: 122 opt: 216 Z-score: 244.7 bits: 56.8 E(): 0.00014  
Smith-Waterman score: 224; 43.137% identity (69.608% similar) in 102 aa overlap  
(377-476:1-91)

```
          350          360          370          380          390          400
frame5 SQHFVLVTF CFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
          : : : : : : : : : : : : : : : : : : : : : :
gi|229          MVATAASSAFFPL--PSPDTSSRPGLKNG
          10          20
```

```
          410          420          430          440          450          460
frame5 PANLGGGLKSKSASSGGLKAKAQA--PSKINGTTVTSKESFKHDDDLPLLIHVDFPDMKP
          . . . : : : : : : : : : : : : : : : : : : :
gi|229 SSSLSPKPKPFVANAGLKVKASASAPPKINGSSVGLKSGSLKTQEDTP-----SVPPRPRT
          30          40          50          60          70          80
```

```
          470          480          490          500          510          520
frame5 FSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADLDF
: : .: :
gi|229 FIN---QLPDWSMLLAAITTVFLAAEKQWMMMLDWKPKRPDMLVDPFGLGSIVQGGLVFR
          90          100          110          120          130
```

>>gi|2298917|emb|CAA02765.1| unnamed protein product [Cu (414 aa)  
initn: 220 init1: 122 opt: 216 Z-score: 244.7 bits: 56.8 E(): 0.00014  
Smith-Waterman score: 224; 43.137% identity (69.608% similar) in 102 aa overlap  
(377-476:1-91)

```
          350          360          370          380          390          400
frame5 SQHFVLVTF CFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
          : : : : : : : : : : : : : : : : : : : : : :
gi|229          MVATAASSAFFPL--PSPDTSSRPGLKNG
          10          20
```

```
          410          420          430          440          450          460
frame5 PANLGGGLKSKSASSGGLKAKAQA--PSKINGTTVTSKESFKHDDDLPLLIHVDFPDMKP
          . . . : : : : : : : : : : : : : : : : : : .
gi|229 SSSLSPKPKPFVANAGLKVKASASAPPKINGSSVGLKSGSLKTQEDTP-----SVPPRPRT
          30          40          50          60          70          80
```

```
          470          480          490          500          510          520
frame5 FSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADLDF
: : .: :
gi|229 FIN---QLPDWSMLLAAITTVFLAAEKQWMMMLDWKPKRPDMLVDPFGLGSIVQGGLVFR
          90          100          110          120          130
```

>>gi|222855219|gb|EEE92766.1| predicted protein [Populus (407 aa)  
initn: 129 init1: 129 opt: 214 Z-score: 242.5 bits: 56.4 E(): 0.00019  
Smith-Waterman score: 214; 53.086% identity (72.840% similar) in 81 aa overlap  
(377-452:1-79)

```
          350          360          370          380          390          400
frame5 SQHFVLVTF CFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
          : : : : : : : : : : : : : : : : : : : : : :
gi|222          MVAAAAASSFFVPSPSGDA--KASKFGSV
          10          20
```

```
          410          420          430          440          450          460
frame5 PANLGGGLKSKSASSGGLKAK--AQAPSKINGTT---VVTSKESFKHDDDLPLLIHVDFPD
          : : : : : : : : : : : : : : : : : : : : : :
gi|222 SASLGGGIKTKSASSGALQVNTNAQAPPKINGPPNEDVVSSPAPRTFINQLPDWSMLLAAI
          30          40          50          60          70          80
```

```
          470          480          490          500          510          520
frame5 MKPFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDAD
          : : : : : : : : : : : : : : : : : : : : : :
gi|222 TTMFLAAEKQWMMMLDWKPKRPDMLIDPFGIGRIVQDGLVFRQNFIRSSEIGADRTASIE
          90          100          110          120          130          140
```

>>gi|1930081|gb|AAB51525.1| acyl-ACP thioesterase [Garci (406 aa)  
initn: 207 init1: 106 opt: 209 Z-score: 236.8 bits: 55.3 E(): 0.00039  
Smith-Waterman score: 209; 64.062% identity (82.812% similar) in 64 aa overlap  
(377-437:1-62)

```
          350          360          370          380          390          400
frame5 SQHFVLVTF CFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
          : : : : : : : : : : : : : : : : : : : : : :
gi|193          MVATAATSSFFPLTSPSGDAK-SGNP-GKG
          10          20
```

```
          410          420          430          440          450          460
frame5 PANLGGGLKSKSA--SSGGL--KAKAQAPSKINGTTVTSKESFKHDDDLPLLIHVDFPDMK
          . . . : : : : : : : : : : : : : : : : : :
gi|193 SVSFGSMKSKSAASSRGLQVKANAQAPT KINGSTDDAQLPAPRTFINQLPDWSMLLAAIT
          30          40          50          60          70          80
```

```
          470          480          490          500          510          520
frame5 PFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADLD
          : : : : : : : : : : : : : : : : : : : : : :
gi|193 TVFLAAEKQWMMMLDWKPRRPDMLIDTFGLGRIVQDGLVFRQNFIRSSEIGADRTASIE
          90          100          110          120          130          140
```

>>gi|3410687|gb|AAC31109.1|I90747 Sequence 28 from paten (415 aa)  
initn: 201 init1: 112 opt: 207 Z-score: 234.4 bits: 54.9 E(): 0.00054  
Smith-Waterman score: 207; 43.137% identity (67.647% similar) in 102 aa overlap  
(377-476:1-91)

```
          350          360          370          380          390          400
frame5 SQHFVLVTF CFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
          : : : : : : : : : : : : : : : : : : : : : :
gi|341          MVATAASSAFFPL--PSADTSSRPGLGNK
          10          20
```

```
          410          420          430          440          450          460
frame5 PANLGGGLKSKSASSGGL--KAKAQAPSKINGTTVTSKESFKHDDDLPLLIHVDFPDMKP
          . . . : : : : : : : : : : : : : : : : : : .
gi|341 PSSLSPLKPKSTPNGGLQVKANASAPPKINGSPVGLKSGGLKTQED----AH-SAPPPRT
          30          40          50          60          70          80
```

```
          470          480          490          500          510          520
frame5 FSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADLDF
: : .: :
gi|341 FIN---QLPDWSMLLAAITTVFLAAEKQWMMMLDWKPKRPDMLVDPFGLGSIVQDGLVFR
          90          100          110          120          130
```

>>gi|758701|gb|AAC48990.1| 16:0-ACP thioesterase preprot (415 aa)  
initn: 201 init1: 112 opt: 207 Z-score: 234.4 bits: 54.9 E(): 0.00054  
Smith-Waterman score: 207; 43.137% identity (67.647% similar) in 102 aa overlap  
(377-476:1-91)

```

      350      360      370      380      390      400
frame5 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGSKLGGG
gi|758                MVATAAASSAFFPL--PSADTSSRPKGKLGNK
                        10      20

      410      420      430      440      450      460
frame5 PANLGGKSKSASSGGL--KAKAQAPSKINGTTVVTSKESFKHDDDLPLLIHVDFPDMKP
gi|758 PSSLSPLPKPSTPNGGLQVKANASAPPKINGSPVGLKSGGLKTQED---AH-SAPPPT
      30      40      50      60      70      80

      470      480      490      500      510      520
frame5 FSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADLDF
gi|758 FIN---QLPDWSMLLAAITTVFLAAEKQWMMLDWKPKRPDMLVDPFGLGSIVQDGLVFR
      90      100      110      120      130

>>gi|270240896|emb|CBI28125.1| unnamed protein product [ (408 aa)
  initn: 181 init1: 90 opt: 201 Z-score: 227.6 bits: 53.6 E(): 0.0013
Smith-Waterman score: 201; 62.500% identity (75.000% similar) in 64 aa overlap
(377-437:1-64)

      350      360      370      380      390      400
frame5 SQHFVLVTFCFMVMISRLDVLPEKRP.KFIMVATAATSSFFPVTSPSPDSGGAGS-KLGG
gi|270                MVATAATAFFAVASPSDDPAKPKSTKPGV
                        10      20      30

      410      420      430      440      450      460
frame5 GPANLGGKSKSASSGGL--KAKAQAPSKINGTTVVTSKESFKHDDDLPLLIHVDFPDMK
gi|270 GSAILRGIKSRNAPSGSLQVKANAQAPPKINGTTDDMSPPPTFINQLPDWSMLLAAIT
      40      50      60      70      80      90

      470      480      490      500      510      520
frame5 PFSNTDSLNLQDWLIKVLGGGEGRSSSCLKLSLDVTTVVPLIFEGACAFKPPEDADLD
gi|270 TIFLAAEKQWMMLDWKPKRRSMDLIDPFGLGKIVQDGLVFRQNFSIRSYPEIGADRTASLET
      100      110      120      130      140      150
```

2416 residues in 1 query sequences  
4761287459 residues in 17815538 library sequences  
Scomplib [34t26]  
start: Wed Feb 3 23:39:17 2010 done: Thu Feb 4 00:02:45 2010  
Total Scan time: 1292.010 Total Display time: 0.390

Function used was FASTA [version 3.4t26 July 7, 2006]

Appendix 6. Bioinformatic analysis of polypeptide frame6

```
>frame6
QL.MASCPGNLHGSAMSMVMNMEKKKE.LPIFFQFKNVDVRSVVIK.KYILIKRQITIRRIYRRKQ.TNYSNSEIFISTCLHS
RPNGLDEKLHDLARRDRYIYMGPAAASVMDICRIPYNVQIQQIKQRNMMLLTNHNVLMLNLCLGL.IDLHEVRDSILIFVSS
YYVS..HNYFYIIMNRINPTKHIPNKINNKFLOIQKKSVLKR.KNNNNVSKNTI.W.SNVSGLVEKSHMRHQPIEYEKNKAE
ISK.MKHYRAIASNVHRHSFVTRPRGKLNACISGYFMWGSSYLLRKRNSRN.IMHLYILFHF.IS.RLCRHLASSYNISTSL
```

```
Y..TIGCNAAVEDDHNISCWLLFVLWS.FQD.TFYRRSDLRNSLWWQQLLLHHFSLLLHPRTLVEQAANLVVGLQTLED.NP
NLRLLLVA.RQRRKPLRKLMEPQLLHLKKASSMMMIYLC.SM.ISRT.SHFQTLIV.TCRTG..KFWGEAKVDHHA.SFL.M.
QLWFW.FSKGLAPLPSSHOKTOIWLVLVQGLQAHHQVCCLLHQSPARVK.QGKMMK.QLLPP..ISKVASPVERLVLS.P.N
KK.PARNVVIVLYCGITANCLSIERSADIVATS.VPTPKPSGNLEMKKDI.MHNSISGVSPFQOI..TPHEISRDAAGTI.FTP
RLELGRTRFLMMKTIWAEIADIYSHRSC.HAWTFGHGVACSLTGVEINKNEVAENTSLRLG.QQKMCGLMMG.VGECFGLLD.H
LCEYVLLCMAGFVGQFFFLDAAILIS.WISDVHCV.LRFCLNFMNLDIAMQRILSLCFVFFCSLAFSYDPRD.SVLWHSRLPRL
IMMEEG.VTD.YGLLPN.LISGKWRPGSKHYPHYSQMOWTLHMVPTICTM.E..L.IIKE.SLLNMPKNARKLIKMQNKCRIKNI
LVDQYIIDE.KIHKNRY.NINILLHIFFLNNTY.MFLPGIL.SRPYRGRRTPGYRPPGWAASSGCRSIDACCSIGKS.NALK
NIFILNYKMSMTIITIIKQ.LESDKDSGKLHGKFINV.NTRGHTCIQ.HLQLF.V.KYICCLVNKHNGTTTSVLLLIILTSVIR
GHDLTHFYSKQNVCSILIIISNSTHNK.EKKQIILI.E.TKGPYHSLTLHPFFHSSIAKTE.KTQ.ITSTTNGTRKTVFPM
.YSNSVGFWCVRNETDALNLTNVVETDDTNESSS.RILRS.DPVPSNS.AHQAALVSESSILAPRAIKSMNSGKLVAIIVA
SSTVTGFSRTRPITRKLIAIR.SRWVATAAPEALPRPLPSGRPRTTREVSPSSQSTPLSLRPFATAERRSLSLTRSSSKPFI
VAPSANAAATARIGYSSIIEGARSSGTVPFPKVEERTRSATSSPPARRGLITSMASIPISCRVKMRPVRVGFIKTLRMVTSE
GTSKAATNGKAVEDGSGPGTSIT.PVSLPRPSRRMVRTPSASVSTVRLAPKP.SIFSVWSRVMIGSITVIVPGVLRPASRTADF
T.AEAITGL.VIGVGFGPRKVTGRRSPSSDFTCTPISRSGFNTRPIGRLVREASPMKVLKS.TPTRPIVKRQPVAAALPKSR
GASGARSPPLPTPSMQVSPSLRILAPIACIALPVLITSSPSKRPVIRVSPLARPPNMKDLWEIDLSPGIRTVPERPEDLRAV
AGRLLAP.SMHAVETEDMTLRGSSEPINVIPLFFNPHDDEIG.ARGCCCVFRETDKGDLR.LDFERLEIRDGFCTPLQILLT
CAMVDLWSTNLQKKEGNSIEIKTVDIYNHYENHL.HKTIIIVKSRYRSQTDIFYHKNCGLSSELSELNSIKSAISRDIKQTI
ELPLEKGTGTFVMLLQNYPIITIQIDIHLYANQQVHSYIDPPTNNQSNKLKSNKIDQNGKIVS.INYRLEAVKR.Y.QSIKS
KQIRIFRTINQQRASNPITIEII.ATISRRKRRTI.GYVNAYLRDIAEKCAARVNKNKDNQWQQREWFYL.R.IRNQC.GL
T.NNII.LTITIMPLFSREYSIPLGLVVFYECLVLHVNTII.SHLLIKVIQANGYTCCSPKNFIDCGPSIYMTLFSRVNI.YLN
ECRLSRAIYIYSNTK.NNEQVKYILTSHDNGNGLL.QPYVSHSTSNLPQERNFRYQA.YRILPYKPLRMGVPVSWFWRYQF
E.SDMARTLLIEPNLECC.ILSSS.FFCGRHCVRHTH..VGCQDSSFPYFVPHLY.RNLLYLCNKGKDAFGRCA.QFCTIPF
LSGPHKPAARC.QDYFQRCPLTSLRWMSFSETY.P.CMCWNSLL.LKILH.SPVAFG.WFGDSVP.PC.GFGLRISFHSTA
ASIVKFLI.APIWT.M.TRRNKDFRIRIICLLSPINTTDRNLSFYQNVLSFYNNAAADIYIFELKKNW.LLFLFLHIDHHTHC
.SM.ISRT
```

Sliding 8 amino acid window search  
Database searched = AD\_2010  
Query = frame6

Start time: Thu Feb 4 00:02:47 GMT 2010 Finish time: Thu Feb 4 00:02:47 GMT 2010

No 8 amino acid matches exist between frame6 and the AD\_2010 database

# fasta34 frame6.pep /genedata/1/db/AD\_2010 -Q -E 1 -O frame6.pep\_ad.fasta  
FASTA searches a protein or DNA sequence data bank version 3.4t26 July 7, 2006  
Please cite:  
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

frame6, 2415 aa  
vs /genedata/1/db/AD\_2010 library

<	20	opt	E()
22	0	0:	one = represents 4 library sequences
24	0	0:	
26	0	0:	
28	0	0:	
30	0	2:*	
32	5	8:*	
34	9	21:=== *	
36	17	44:===== *	
38	54	72:===== *	
40	141	101:=====*	
42	97	123:===== *	
44	96	136:===== *	
46	124	138:===== *	
48	148	132:=====*	
50	183	121:=====*	

```

52 103 106:=====*
54 100 91:=====*=
56 81 76:=====*=
58 49 62:===== *
60 40 50:===== *
62 37 40:=====*
64 26 32:=====*
66 30 25:=====*=
68 30 20:=====*=
70 25 16:=====*=
72 19 12:=====*=
74 12 10:=====*=
76 13 7:=====*=
78 7 6:=====*=
80 8 4:=====*=
82 1 3:=====*=
84 3 3:=====*=
86 1 2:=====*=
88 2 2:=====*=
90 1 1:=====*=
92 1 1:=====*=
94 1 1:=====*=
96 1 1:=====*=
98 0 0:=====*=
100 1 0:=====*=
102 0 0:=====*=
104 1 0:=====*=
106 0 0:=====*=
108 2 0:=====*=
110 0 0:=====*=
112 0 0:=====*=
114 0 0:=====*=
116 0 0:=====*=
118 0 0:=====*=
>120 0 0:=====*=

331323 residues in 1471 sequences
Expectation_n fit: rho(ln(x))= 5.57300.00413; mu= 11.0574 0.215
mean_var=62.482318.269, 0's: 2 Z-trim: 4 B-trim: 11 in 1/42
Lambda= 0.162254
Kolmogorov-Smirnov statistic: 0.0692 (N=29) at 46

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 44, opt: 32, open/ext: -10/-2, width: 16
The best scores are:
gi|66849793|gb|EAL90121.1| allergen Asp F7 [Asperg ( 270) 89 31.1 0.42
gi|83300389|sp|O42799.2|ALL7_ASPFU RecName: Full=A ( 270) 89 31.1 0.42
gi|3915783|sp|P43217.3|NLT11_PARJU RecName: Full=P ( 139) 81 29.1 0.84

>>gi|66849793|gb|EAL90121.1| allergen Asp F7 [Aspergillu (270 aa)
initn: 44 initl: 44 opt: 89 Z-score: 109.1 bits: 31.1 E(): 0.42
Smith-Waterman score: 104; 21.101% identity (49.541% similar) in 218 aa overlap
(1229-1438:3-209)

1200 1210 1220 1230 1240 1250
frame6 ILRS.DPVPSSNS.AHQAALVSESSILAPRPAIKSMN-SGKLVAIIVASSTVGFSETRP
: . . . . .
gi|668 MAPIFKSLALVSALFAAIISSAAPVN--LDKRE
10 20 30

1260 1270 1280 1290 1300 1310
frame6 ITRKLIAR.SRWVATAAPEAL-PRPLPSGRPRTTREVSPSSQSTPLSLRPFATAERRSL
. . . . .

```

```

. . . . .
gi|668 VDVVWTVTVTVVWTTIDVTTTIYPTPQAPTTPPVVESTPTPTPSAAPEQAEPJETSTQPET
40 50 60 70 80 90

1320 1330 1340 1350 1360 1370
frame6 SLTRSSSKPFITVAPSANAAATARIGYSSIEGARSSGVTPFKVEERTKRSATSSPPA-
. . . . .
gi|668 TKSQPTQPSVATFIPVAAAAAAD-SAAPIPEEPAPQAPATTAAPSTSTTTQAAPSAPPAA
100 110 120 130 140

1380 1390 1400 1410 1420 1430
frame6 ----RRGLITMSAPISCRVKMRPVVRVGFIKTLRMVTSEPGTSKAATNGKAVEDGSPGT
. . . . .
gi|668 NSGSTKAASSGYSGPCS---KGSPC-VGQLTYDYDTATS---ASAPSSCGLTNDGFSEN
150 160 170 180 190 200

1440 1450 1460 1470 1480 1490
frame6 SIT.PVSLPRPSRRMVRTPSASVSTVRLAPKP.SIFSVWSRVMIGSITVVIPGVLRPASR
. . . . .
gi|668 VVALPVGIMTDADCGKTVTITYNGITKTATVVDKCMGCKPTDLDASRHLFGELADFSAGR
210 220 230 240 250 260

>>gi|83300389|sp|O42799.2|ALL7_ASPFU RecName: Full=Aller (270 aa)
initn: 44 initl: 44 opt: 89 Z-score: 109.1 bits: 31.1 E(): 0.42
Smith-Waterman score: 104; 21.101% identity (49.541% similar) in 218 aa overlap
(1229-1438:3-209)

1200 1210 1220 1230 1240 1250
frame6 ILRS.DPVPSSNS.AHQAALVSESSILAPRPAIKSMN-SGKLVAIIVASSTVGFSETRP
: . . . . .
gi|833 MAPIFKSLALVSALFAAIISSAAPVN--LDKRE
10 20 30

1260 1270 1280 1290 1300 1310
frame6 ITRKLIAR.SRWVATAAPEAL-PRPLPSGRPRTTREVSPSSQSTPLSLRPFATAERRSL
. . . . .
gi|833 VDVVWTVTVTVVWTTIDVTTTIYPTPQAPTTPPVVESTPTPTPSAAPEQAEPJETSTQPET
40 50 60 70 80 90

1320 1330 1340 1350 1360 1370
frame6 SLTRSSSKPFITVAPSANAAATARIGYSSIEGARSSGVTPFKVEERTKRSATSSPPA-
. . . . .
gi|833 TKSQPTQPSVATFIPVAAAAAAD-SAAPIPEEPAPQAPATTAAPSTSTTTQAAPSAPPAA
100 110 120 130 140

1380 1390 1400 1410 1420 1430
frame6 ----RRGLITMSAPISCRVKMRPVVRVGFIKTLRMVTSEPGTSKAATNGKAVEDGSPGT
. . . . .
gi|833 NSGSTKAASSGYSGPCS---KGSPC-VGQLTYDYDTATS---ASAPSSCGLTNDGFSEN
150 160 170 180 190 200

1440 1450 1460 1470 1480 1490
frame6 SIT.PVSLPRPSRRMVRTPSASVSTVRLAPKP.SIFSVWSRVMIGSITVVIPGVLRPASR
. . . . .
gi|833 VVALPVGIMTDADCGKTVTITYNGITKTATVVDKCMGCKPTDLDASRHLFGELADFSAGR
210 220 230 240 250 260

>>gi|3915783|sp|P43217.3|NLT11_PARJU RecName: Full=Proba (139 aa)
initn: 41 initl: 41 opt: 81 Z-score: 103.7 bits: 29.1 E(): 0.84
Smith-Waterman score: 81; 23.577% identity (47.154% similar) in 123 aa overlap
(1516-1630:18-135)

```

```

      1490      1500      1510      1520      1530      1540
frame6 RPASRTADFT.AEAIGTL.VIGVGVFGRKVTGRRSPSSDFTCTPISRSGFNTRPIGRLV
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|391      QETCGTMVRALMPCLPFVQGEKEPSKGCSCGAKRLDGETKTGPQRV
      10      20      30      40

      1550      1560      1570      1580      1590      1600
frame6 REASPMKVLVSKS.TPTRPIVKRQPVAALPKSRGASGARSPPLPTP---SMIQVSPSLRIL
      . . . . . : : : : : : : : : : : : : : : : : : : : : :
gi|391 HACECIQTAMKIYSD---IDGKLVSEVPKHCGIVDSKLPPIIDVNMCKTVGVVPRQPQL
      50      60      70      80      90      100

      1610      1620      1630      1640      1650
frame6 APIACIALPVLITSSPS-----KRPVIRVSPLARPPNMKDLWEIDLSPGIRTVPERPEDL
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|391 -PVSLRHGFVTGPSDEPAHKARLERPQIRVPPPAPEKA
      110      120      130

      1660      1670      1680      1690      1700      1710
frame6 RAVAGRLLAP.SMHAVETEDMTLGRSSEPINVIPLFFNPHDDEIG.ARGCCCVFRETDK

2415 residues in 1 query sequences
331323 residues in 1471 library sequences
Scomplib [34t26]
start: Thu Feb 4 00:02:46 2010 done: Thu Feb 4 00:02:46 2010
Total Scan time: 0.140 Total Display time: 0.010

Function used was FASTA [version 3.4t26 July 7, 2006]

# fasta34 frame6.pep /genedata/1/db/TOX_2010 -Q -E 1 -O frame6.pep_tx.fasta
FASTA searches a protein or DNA sequence data bank version 3.4t26 July 7, 2006
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

frame6, 2415 aa
vs /genedata/1/db/TOX_2010 library

      opt      E()
< 20      61      0:====
22      0      0:
24      1      0:=
26      0      0:
28      1      2:*
30      19     12:*
32      20     45:*
34      80     122:===== *
36      384    250:===== *=====
38      250    414:===== *
40      568    577:===== *
42      585    706:===== *
44      614    779:===== *
46      763    793:===== *
48      1079   759:===== *=====
50      820    693:===== *=====
52      637    609:===== *
54      447    520:===== *
56      364    435:===== *
58      310    357:===== *
```

```

60      240      289:===== *
62      229      232:===== *
64      195      184:===== *
66      180      146:===== *
68      101      115:===== *
70      95      90:===== *
72      66      70:===== *
74      104      55:===== *
76      77      43:===== *
78      33      33:===== *
80      23      26:===== *
82      14      20:===== *
84      20      16:===== *
86      15      12:===== *
88      5      9:===== *
90      9      7:===== *
92      6      6:===== *
94      15      4:===== *
96      2      3:===== *
98      3      3:===== *
100     2      2:===== *
102     0      2:===== *
104     1      1:===== *
106     0      1:===== *
108     1      1:===== *
110     1      1:===== *
112     2      0:===== *
114     0      0:===== *
116     0      0:===== *
118     0      0:===== *
>120     1      0:===== *

2069351 residues in 8448 sequences
Expectation_n fit: rho(ln(x))= 4.64900.000656; mu= 16.6197 0.033
mean_var=56.571212.164, 0's: 60 Z-trim: 62 B-trim: 581 in 1/61
Lambda= 0.170520
Kolmogorov-Smirnov statistic: 0.0495 (N=29) at 46

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 44, opt: 32, open/ext: -10/-2, width: 16
The best scores are:
gi|197625399|gb|EDY37958.1| structural toxin prote (2003) 116 38.8 0.086

>>gi|197625399|gb|EDY37958.1| structural toxin protein R (2003 aa)
initn: 71 initl: 71 opt: 116 Z-score: 135.1 bits: 38.8 E(): 0.086
Smith-Waterman score: 120; 25.989% identity (52.542% similar) in 177 aa overlap
(1482-1646:54-217)

      1460      1470      1480      1490      1500      1510
frame6 SVSTVRLAPKP.SIFSVWSRVMIGSITVPIGVLRPASRTADFT.AEAIGTL.VIGVGVF
      : : : : : : : : : : : : : : : : : : : : : :
gi|197 LGASSEREKDATGNISDAALHDPGVRDNPGGHTGFGDAVDHTPVETYSPhdVgALGVS
      30      40      50      60      70      80

      1520      1530      1540      1550      1560      1570
frame6 GPRKVTGRRSPSSDFTCTPISRSGFNTRPIGRLVREASPMKVLVSKS.TPTRPIVKRQPVA
      . : . : : : : : : : : : : : : : : : : : : :
gi|197 SP--MPQPPGPSVDEAGSGGGPGIN--P-----APLQPLSSGDVPTAPLAAAALIA
      90      100      110      120      130

      1580      1590      1600      1610      1620
frame6 ALPKSRGASGAR-SPPLPT--PSMIQVSPSLRILAP--IACIALPV-LITS-SPSKRPVI
```



gi|77967877|gb|ABB09257.1| DNA translocase FtsK [B (1673) 187 50.7 0.04  
gi|111307761|gb|AAI21193.1| Mki67 protein [Xenopus (2080) 188 51.0 0.041  
gi|134065504|emb|CAM43271.1| proteophosphoglycan p (5384) 193 52.3 0.041  
gi|70905641|gb|AAZ14280.1| proteophosphoglycan ppg (7194) 194 52.6 0.045  
gi|190346874|gb|EDK39060.2| hypothetical protein P ( 961) 183 49.7 0.046  
gi|38260630|gb|AAR15447.1| pollen coat oleosin-gly (1356) 183 49.8 0.061  
gi|270236229|emb|CBI25498.1| unnamed protein produ ( 545) 178 48.4 0.062  
gi|190624521|gb|EDV40045.1| GF24155 [Drosophila an (1337) 182 49.6 0.07  
gi|190617179|gb|EDV32703.1| GF22010 [Drosophila an ( 847) 179 48.8 0.075  
gi|32452989|gb|AAP82647.1| Hypothetical protein K0 ( 825) 177 48.4 0.099  
gi|212001255|gb|EEB06915.1| predicted protein [Sch (1678) 180 49.2 0.11  
gi|123962409|gb|ABM77165.1| Translation initiation (1124) 177 48.5 0.13  
gi|211583314|emb|CAP91322.1| Pcl3g02530 [Penicilli ( 643) 174 47.6 0.13  
gi|135120742|gb|EBF62441.1| hypothetical protein G ( 307) 170 46.6 0.13  
gi|186909484|gb|ACC94277.1| mucin [Bos taurus] ( 549) 173 47.4 0.13  
gi|145017502|gb|EDK01865.1| predicted protein [Mag ( 577) 173 47.4 0.13  
gi|198150947|gb|EDY73984.1| GA28499 [Drosophila ps (1221) 177 48.5 0.13  
gi|49651333|emb|CAG78272.1| YAL10F15653p [Yarrowia (1051) 176 48.2 0.14  
gi|156232020|gb|ABU56803.1| conserved hypothetical ( 625) 173 47.4 0.14  
gi|34328647|gb|AAO83650.1| putative protein Roco5 (2800) 181 49.6 0.14  
gi|228851254|gb|EEQ40718.1| hypothetical protein C (1728) 178 48.8 0.15  
gi|221482859|gb|EEE21190.1| conserved hypothetical ( 849) 174 47.7 0.16  
gi|3834293|gb|AAC70889.1| Hypothetical protein K06 (1032) 175 48.0 0.16  
gi|228266868|gb|ACP97541.1| Sequence 29759 from pa ( 482) 169 46.5 0.21  
gi|226444262|gb|ACO57709.1| mucin 1 [Bos grunniens ( 502) 169 46.5 0.21  
gi|143178533|gb|EDD38130.1| hypothetical protein G ( 381) 167 46.0 0.23  
gi|54015468|dbj|BAD56838.1| hypothetical protein [ ( 404) 167 46.0 0.24  
gi|218196980|gb|EEG79407.1| hypothetical protein O ( 590) 169 46.6 0.24  
gi|46576017|gb|AAT01378.1| unknown protein [Oryza ( 590) 169 46.6 0.24  
gi|113579407|dbj|BAF17770.1| Os05g0480600 [Oryza s ( 590) 169 46.6 0.24  
gi|57863811|gb|AAW56864.1| unknown protein [Oryza ( 590) 169 46.6 0.24  
gi|215734844|dbj|BAG95566.1| unnamed protein produ ( 590) 169 46.6 0.24  
gi|222631981|gb|EEE64113.1| hypothetical protein O ( 590) 169 46.6 0.24  
gi|143272064|gb|EDE05223.1| hypothetical protein G ( 500) 168 46.3 0.25  
gi|20380014|gb|AAH27781.1| Srrm2 protein [Mus musc ( 506) 168 46.3 0.25  
gi|33635378|emb|CAE21703.1| Translation initiation (1125) 172 47.4 0.26  
gi|124393847|emb|CAK59372.1| unnamed protein produ ( 542) 168 46.3 0.26  
gi|140401443|gb|ECL52473.1| hypothetical protein G ( 277) 163 45.1 0.32  
gi|223533796|gb|EEF35528.1| ATP binding protein, p ( 568) 166 45.9 0.37  
gi|114188606|gb|EAU30306.1| conserved hypothetical (1205) 170 47.0 0.37  
gi|190654733|gb|EDV51976.1| GG15816 [Drosophila er ( 601) 166 45.9 0.38  
gi|194189999|gb|EDX03575.1| GD22776 [Drosophila si ( 351) 163 45.1 0.39  
gi|211966694|gb|EEB01890.1| hypothetical protein T ( 434) 164 45.4 0.4  
gi|40456135|gb|AAR86113.1| UL49 tegument protein [ ( 267) 161 44.6 0.42  
gi|143344296|gb|EDE46033.1| hypothetical protein G ( 574) 165 45.7 0.43  
gi|144106949|gb|EDI91441.1| hypothetical protein G ( 197) 159 44.1 0.44  
gi|134084593|emb|CAK97469.1| unnamed protein produ ( 616) 165 45.7 0.45  
gi|136457362|gb|EBO28834.1| hypothetical protein G ( 518) 164 45.5 0.46  
gi|52354619|gb|AAH82855.1| LOC494754 protein [Xeno ( 796) 166 46.0 0.48  
gi|66910752|gb|AAH97656.1| LOC494754 protein [Xeno ( 798) 166 46.0 0.48  
gi|86572881|gb|ABD07438.1| Peptidase M23B [Rhodops ( 474) 163 45.2 0.49  
gi|238851258|gb|EEQ40722.1| hypothetical protein C (1008) 166 46.1 0.57  
gi|22945364|gb|AAF51145.2| toucan, isoform A [Dros (2162) 170 47.2 0.58  
gi|2760522|emb|CAA74574.1| Toucan protein [Drosoph (2176) 170 47.2 0.58  
gi|12311816|emb|CAC22634.1| hypothetical protein L (1892) 169 46.9 0.6  
gi|221507814|gb|EEE33401.1| conserved hypothetical ( 434) 161 44.8 0.61  
gi|114196989|gb|EAU38689.1| conserved hypothetical (1121) 166 46.1 0.62  
gi|115281178|gb|ABI86695.1| conserved hypothetical ( 392) 160 44.5 0.66  
gi|135183063|gb|EBG02505.1| hypothetical protein G ( 186) 156 43.5 0.66  
gi|7290181|gb|AAF45644.1| mucin related 2B [Drosop (1795) 168 46.7 0.67  
gi|4691241|emb|CAA19845.2| EG:56G7.1 [Drosophila m (1795) 168 46.7 0.67  
gi|14018379|emb|CAC38347.1| cell surface glycoprot (3971) 172 47.8 0.69

gi|136493029|gb|EBO51894.1| hypothetical protein G ( 423) 160 44.6 0.7  
gi|186909486|gb|ACC94278.1| mucin [Bos indicus] ( 429) 160 44.6 0.7  
gi|211584486|emb|CAP92529.1| Pcl3g14600 [Penicilli ( 793) 163 45.4 0.73  
gi|119397654|gb|EAW08085.1| conserved hypothetical ( 577) 161 44.9 0.77  
gi|161167388|emb|CAN98693.1| protein kinase [Soran (1219) 165 46.0 0.77  
gi|256358685|gb|ACU72182.1| serine/threonine prote ( 855) 163 45.4 0.78  
gi|162955311|gb|ABY24826.1| hypothetical protein R ( 428) 159 44.4 0.81  
gi|183985550|gb|AAI66052.1| LOC494754 protein [Xen ( 774) 162 45.2 0.83  
gi|19571560|emb|CAD27470.1| sequence orphan [Schiz ( 800) 162 45.2 0.85  
gi|194116647|gb|EDW38690.1| GL16241 [Drosophila pe (1709) 166 46.3 0.86  
gi|215704255|dbj|BAG93095.1| unnamed protein produ ( 392) 158 44.1 0.88  
gi|111308074|gb|AAI21353.1| LOC779458 protein [Xen (1006) 163 45.5 0.88  
gi|166796751|gb|AAI59126.1| LOC779458 protein [Xen (1008) 163 45.5 0.88  
gi|154757609|gb|AAI51763.1| LOC100137827 protein [ ( 695) 161 44.9 0.88  
gi|137470381|gb|EBU27837.1| hypothetical protein G ( 194) 154 43.1 0.91  
gi|122890548|emb|CAM14118.1| novel protein contain (1044) 163 45.5 0.91  
gi|86568965|gb|ABD12774.1| hypothetical protein Fr ( 613) 160 44.7 0.93  
gi|223532500|gb|EEF34290.1| conserved hypothetical ( 513) 159 44.4 0.93  
gi|220976661|gb|EED94988.1| predicted protein [Tha (1964) 166 46.3 0.96

>>gi|228385049|gb|ACQ29002.1| Sequence 21516 from patent (331 aa)  
initn: 413 initl: 366 opt: 424 Z-score: 479.2 bits: 99.9 E(): 1.2e-17  
Smith-Waterman score: 428; 40.909% identity (58.678% similar) in 242 aa overlap  
(1211-1452:99-329)

	1190	1200	1210	1220	1230	1240
frame6	NLTNVVETDDTNESSS.RILRS.DPVPSSNS.AHQAAVLSESSILAPRAIKSMNSGKLV					
	:	:	:	:	:	:
gi 228	QQVPAQAPGEQPGDRALAGATGAIDGHEHRSVHQ--LFSTATRI--PVWAHRARKLGKEV					
	70	80	90	100	110	120

	1250	1260	1270	1280	1290	1300
frame6	AIIVASSTVTGFSETRPITRKLIAR.SRWVATAAPEALPRPLPSGRPRTTREVSPSSQS					
	:	:	:	:	:	:
gi 228	ATLAQSWMRIGPLARSEATLKDMAIR---W---SPCAQTSPPPNAPPSMTMPSPGVGSFAF					
	130	140	150	160	170	

	1310	1320	1330	1340	1350	1360
frame6	TPLSLRPFATAERRSLSLTRSSSKPFITVAPSANAAATARIGYSSIIIEGARSSGVTVPFK					
	:	:	:	:	:	:
gi 228	TPRAFRPSAMTWIRSDSLTRSSAPRRVTRPSAQAAATKSTGNSSMASGTRSSGMSMPLS					
	180	190	200	210	220	230

	1370	1380	1390	1400	1410	1420
frame6	VEERTRKSATSSPPARRGLITMSAPISCRVKMRPVRVGFIKTLRMVTSEPGTSKAATNG					
	:	:	:	:	:	:
gi 228	RAERTMSATSGPPTSRWFSRLRSPPIRRRISMTPTRVGLMPTCCRTSSEPSAILAATRK					
	240	250	260	270	280	290

	1430	1440	1450	1460	1470	1480
frame6	KAVEDGSPGTSIT.PVSLPRPSRRMVRTPSASVSTVRLAPKP.SIFSVWSRVMIGSITVV					
	:	:	:	:	:	:
gi 228	KAAEDMSAGTSIWVALNLC-PDSTLAVLPSTTTG					
	300	310	320	330		

	1490	1500	1510	1520	1530	1540
frame6	IPGVLRPASRTADFT.AEAIGTL.VIGVGVGPRKVTGRRSPSSDFTCTPIRSRSGFNTRP					

>>gi|269098125|gb|ACZ22561.1| hypothetical protein Sked\_ (1104 aa)  
initn: 153 initl: 81 opt: 246 Z-score: 270.2 bits: 62.9 E(): 5.5e-06

Smith-Waterman score: 297; 27.097% identity (51.828% similar) in 465 aa overlap (1257-1693:520-924)

```

      1230      1240      1250      1260      1270      1280
frame6 PRPAIKSMNSGKLVAIIVASSTVTGFSETRPIITRKLIAIR.SRWVATAAPEALPRPLPSG
      :: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
gi|269 SHDRSPSPAPASRQEAAEREPSRRRALPVVPLRRTAARRTGE---VRPAAQPRPAPAL
      490      500      510      520      530      540

      1290      1300      1310      1320      1330
frame6 RPRTTREVS---PSSQSTPLSLRPFATAERRSLSLTRSS-SKPFIT---TVAPSANAAAT
      :: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
gi|269 RPALTRRTALVPRPTPVARPVAV-PKATAGRPSTALRPTSPTRPASPQARTVVPARTAAV-
      550      560      570      580      590      600

      1340      1350      1360      1370      1380      1390
frame6 ARIGYSSIIIEGARSSGTVTPFKVEERTKRSATISSPPARRG---LITSMSAPISCRVKMRP
      :: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
gi|269 --VGLS--VPASRRRGVV-PLVAPARTLRLA---PPAVVGAARVLPVAAALLSTRT-LLP
      610      620      630      640      650

      1400      1410      1420      1430      1440      1450
frame6 VRVGFIKTLRMVTISEPGTSKAATNGKAVEDGSPGTSIT.PVSLPRPSRRMVRTPSASVST
      :: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
gi|269 A-VGRLEPAR--TARPLARAVP-GRSTLV-SPAR-VVPPRTVPVAVSRVVPARSTPLPA
      660      670      680      690      700

      1460      1470      1480      1490      1500      1510
frame6 VRLAPKP.SIFSVWSRMIGSITVVIPGVLRPASRTADFT.AEAIGTL.VIGGVFGPRK
      :: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
gi|269 ARVVP-----ARSAPLAAARVVP-----ARTALVSAPRAVPTR---GTPVSAARA
      710      720      730      740      750

      1520      1530      1540      1550      1560      1570
frame6 VTGRRSPSSDFTCTPIISRSGFNTRPIGRLVREASPMKVLKSKS.TPTRP--IVKQPVAAAL
      :: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
gi|269 VPARR-----VLAPVA---LPTRPAPR-VHASPPALTARRTAVPSAPRAVPSRRLAPV
      760      770      780      790      800

      1580      1590      1600      1610      1620
frame6 P-KSRGASGARS--PPLPT----PSMIQVSPSLRILAPIACIALPVLITSSPSKRVPVIR
      :: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
gi|269 PLPTSPAPRVQPAVPPFPARRAAAPGAPRSVPTRRVLAPVALPARPAPRVHPAAPAPAAAR
      810      820      830      840      850      860

      1630      1640      1650      1660      1670
frame6 --VSPLARP---PNMKDLWEIDL--PGIRTVPERPEDLRVAGRLLAP.SMHAVETEDM
      :: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
gi|269 RTVVVPA-PRLVPGRRALVPVPLPLRPAPRALPTSP-----GRAVPAARAH---TPVA
      870      880      890      900

      1680      1690      1700      1710      1720      1730
frame6 TLRGRSSEPINVIPLEFFNPHDDEIG.ARGCCCVFRETDKGDLR.LDFERLEIRDGFCTPL
      ::::: :::::
gi|269 PLRGRASLAPQAAPVAPPARRHTAIPRIITAPAATTPPVVVVPAARRPASLAPAGPALV
      910      920      930      940      950      960
```

>>gi|222447388|gb|ACM51654.1| autotransporter-associated (1320 aa)  
initn: 157 initl: 78 opt: 242 Z-score: 264.5 bits: 62.1 E(): 1.1e-05  
Smith-Waterman score: 259; 24.340% identity (48.491% similar) in 530 aa overlap (1118-1629:637-1129)

```

      1090      1100      1110      1120      1130      1140
frame6 SKQNVCSILIIISNSTHNK.EKKQIILI.E.TKGPYHSLTLLHPPFFHSSIAKTE.KTQ.I
      :: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
gi|222 EPTASVTPEPTASATPEPTASVTPEPTASATPEPTASVT---PEPTAST-TPSPSATASV
      610      620      630      640      650      660

      1150      1160      1170      1180      1190      1200
frame6 TSTTNGTRKTVPFMP.YSNSVGFWCVRNETDALNLTNNVETDDTNESSS.RILRS.DPVP
      :: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
gi|222 TPSPSAT-ASVTTPSPSATAS-----ATPEPTA---STTPSPSATASTTPEPTASTTPSP
      670      680      690      700      710

      1210      1220      1230      1240      1250      1260
frame6 SSNS.AHQAAALVSESSILAPRPAIKSMNSGKLVAIIVASSTVTGFSETRPIITRKLIAIR.
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
gi|222 SATASTTPSPSATASA--TPEPTASTTPSPSATASVTTPSPSATASTTPSPSATASVTTPSP
      720      730      740      750      760      770

      1270      1280      1290      1300      1310      1320
frame6 SRWVATAAPEALPRPLPSGRPRTTREVSPS-SQSTPLSLRPFATAERRSLSLTRSSSKPF
      :: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
gi|222 S---ATASMTTPSPSATASATPEPTASTTPSPSATASVTTPSPSATASTTPSPSATASTTP-
      780      790      800      810      820

      1330      1340      1350      1360      1370      1380
frame6 ITVAPSANAAATARIGYS-SIIEGARSSGTVTPFKVEERTRK---SATSSPPARRGLITS
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
gi|222 ---SPSATASATTPSPSATASVTPEPTASTTPSPSATASATPEPTASVTTPSPSATASVTTPS
      830      840      850      860      870      880

      1390      1400      1410      1420      1430
frame6 MSAPISCRVKMRPVVRVGFIKTLRMVTISEPGTSKAATNG-----KAVEDGSPGTSIT.PVS
      :: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
gi|222 PSATAS--VTPSPSATASTTPSPSVTASTTPSPSATASTTPSPSATASTTPSPSATASTT
      890      900      910      920      930      940

      1440      1450      1460      1470      1480      1490
frame6 LPRPSRRMVRTPSASVSTVRLAPKP.SIFSVWSRMIGSITVVIPGVLRPASRTADFT.A
      :: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
gi|222 -PSPSATASTTPSPS-ATASTTPSPSATASA-TPEPTASVT---PSPSATASTTPSPSAT
      950      960      970      980      990

      1500      1510      1520      1530      1540      1550
frame6 EAIGTL.VIGGVFGPRKVTGRRSPSSDFTCTPIISRSGFNTRPIGRLVREASPMKVLKSKS
      :: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
gi|222 ASVTTPSPSATASVTTPSPSATASTTPEPTASTTPSPSATASTTPSPSATASTTPSPSATAS
      1000      1010      1020      1030      1040      1050

      1560      1570      1580      1590      1600      1610
frame6 .TPTRPIVKRQP---VAALPKSRGASGARSPLPTPSMIQVSPSLRILA---PIACIAL
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
gi|222 TTP-EPTASVTTPSPSATASVTTPSPSAT-ASTTPEPTASTTP-SPSATASATPEPTASTS-
      1060      1070      1080      1090      1100      1110

      1620      1630      1640      1650      1660
frame6 PV-LITSSPSKRPIRVSPPLARPPNMKDLWEIDLSPGIRTVPERPEDLRVAGRLLAP.S
      :: ::::: :::::
gi|222 PVSTVTITSPV--PTVTITSPVPTVTITSPVPTVTITSPVPTVTITSPVPTVTITSPV
      1120      1130      1140      1150      1160
```

>>gi|163667088|gb|ABY33454.1| autotransporter-associated (1320 aa)  
initn: 157 initl: 78 opt: 242 Z-score: 264.5 bits: 62.1 E(): 1.1e-05



Smith-Waterman score: 259; 24.340% identity (48.491% similar) in 530 aa overlap (1118-1629:637-1129)

```

      1090      1100      1110      1120      1130      1140
frame6 SKQNVCIISNSTHNK.EKKQIILI.E.TKGPYHSLTLLHPPFFHSSIAKTE.KTQ.I
      : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|163 EPTASVTPEPTASATPEPTASVTPEPTASATPEPTASVT---PEPTAST-TPSPSATASV
      610      620      630      640      650      660

      1150      1160      1170      1180      1190      1200
frame6 TSTTNGTRKTVFPM.P.YNSVGFWCVRNETDALNLTNVVETDDTNESSS.RILRS.DPVP
      : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|163 TPSPSAT-ASVTTPSPSATAS-----ATPEPTA----STTPSPSATASTTPEPTASTTSPS
      670      680      690      700      710

      1210      1220      1230      1240      1250      1260
frame6 SSNS.AHQAAVLSESSILAPRPAIKSMNSGKLVAIIVASSTVIGFSETRPITRKLIAIR.
      : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|163 SATASTTPSPSATASA--TPEPTASTTPSPSATASVTTPSPSATASTTPSPSATASVTTPSP
      720      730      740      750      760      770

      1270      1280      1290      1300      1310      1320
frame6 SRWVATAAPEALRPLPSGRPRRTTREVSPS-SQSTPLSLRPFATAERRSLSLTRSSSKPF
      : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|163 S---ATASMTTPSPSATASATPEPTASTTPSPSATASVTTPSPSATASTTPSPSATASTTP-
      780      790      800      810      820

      1330      1340      1350      1360      1370      1380
frame6 ITVAPSANAAATARIGYS-SIIEGARSSGTVIPFKVEERTRK---SATSSPFARRGLITS
      : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|163 ---SPSATASATPSPSATASVTPEPTASTTPSPSATASATPEPTASVTTPSPSATASVTTPS
      830      840      850      860      870      880

      1390      1400      1410      1420      1430
frame6 MSAPISCRVKMRPVVRVGFIKTLRMVTSEPGTSKAATNG-----KAVEDGSPGTSIT.PVS
      : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|163 PSATAS--VTPSPSATASTTPSPSVTASTTPSPSATASTTPSPSATASTTPSPSATASTT
      890      900      910      920      930      940

      1440      1450      1460      1470      1480      1490
frame6 LPRPSRRMVRTPSASVSTVRLAPKP.SIFSVWSRVMIGSITVVIPGVLRPASRTADFT.A
      : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|163 -PSPSATASTTPSPS-ATASTTPSPSATASA-TPEPTASVT---PSPSATASTTPSPSAT
      950      960      970      980      990

      1500      1510      1520      1530      1540      1550
frame6 EAIGTL.VIGVGVFGRKVTGRRSPSSDFTCTPISRSGFNTRPIGRLVREASPMKVLKS
      : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|163 ASVTTPSPSATASVTTPSPSATASTTPEPTASTTPSPSATASTTPSPSATASTTPSPSATAS
      1000      1010      1020      1030      1040      1050

      1560      1570      1580      1590      1600      1610
frame6 .TPTRPIVKRQP---VAALPKSRGASGARSPLPTPSMIQVSPSLRILA---PIACIAL
      : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|163 TTP-EPTASVTTPSPSATASVTTPSPSAT-ASTTPEPTASTTP-SPSATASATPEPTASTS-
      1060      1070      1080      1090      1100      1110

      1620      1630      1640      1650      1660
frame6 PV-LITSSPSKRPIVRVSPPLARPPNMKDLWEIDLSPGIRTVPERPEDLRAGRLAP.S
      : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|163 PVSTVTTSPV--PTVTTSPVPTVTTSPVPTVTTSPVPTVTTSPVPTVTTSPVPTVTTSPV
      1120      1130      1140      1150      1160
```

>>gi|259147278|emb|CAY80531.1| Muc1p [Saccharomyces cere (1576 aa)  
initn: 60 initl: 60 opt: 240 Z-score: 261.2 bits: 61.8 E(): 1.7e-05  
Smith-Waterman score: 259; 25.052% identity (49.687% similar) in 479 aa overlap (1132-1591:257-699)

```

      1110      1120      1130      1140      1150      1160
frame6 THNK.EKKQIILI.E.TKGPYHSLTLLHPPFFHSSIAKTE.KTQ.ITSTTNGTRKTVFPM
      : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|259 TTTSESSTSSSTTAPATPTTTSCSKEKPTPPHHGTTCTCKKKT---TTSKTCCTKTTTPV
      230      240      250      260      270      280

      1170      1180      1190      1200      1210
frame6 P.YNSVGFWCVRNETDALNLTNVVETDDTNESSS.RILRS.DVPVSSNS.AHQAA----
      : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|259 PTPSSSTTPVPTPSSSTTESSTTPV-TSSTTESSSAPV-----PTPSSSTTESSSAPAPT
      290      300      310      320      330

      1220      1230      1240      1250      1260      1270
frame6 ---LVSESSILAPRPAIKSMNSGKLVAIIVASSTVIGFSETRPITRKLIAIR.SRWVAT
      : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|259 PSSSTTESSSAPAPTSSSTTESSSAPAPTSSSTTESSSAPVPTPSS-----STTESS
      340      350      360      370      380      390

      1280      1290      1300      1310      1320      1330
frame6 AAPEALRPLPSGRPRRTTREV-SPSSQSTPLSLRPFATAERRSLSLTRSSSKPFIIVAPS
      : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|259 SAPA--PTPSSSTTESSSAPVPTPSSCTTESSSAPVPTP---SSSTTESSSAPAPTSSSS
      400      410      420      430      440

      1340      1350      1360      1370      1380
frame6 ANAAATARIGYSSIIEGARSSGTV-TPFKVEERTRKSATSSPFARRGLITSMASAPI----
      : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|259 TTESSSAPVPTPSSCTTESSSAPVPTP-----SSSTTESSSPTVTSSTTESSSAPVPTPS
      450      460      470      480      490      500

      1390      1400      1410      1420      1430      1440
frame6 SCRVKMRPVVRVGFIKTLRMVTSEPGTSKAATNGKAVEDGSPGTSIT.PVSLRPSRRMVR
      : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|259 SSTTESSSPTVTSSTTESSSAPVPTPSSSTTESSSAPAPTSSSTTESSSAPVPTPSSCT
      510      520      530      540      550      560

      1450      1460      1470      1480      1490      1500
frame6 TPSASVSTVRLAPKP.SIFSVWSRVMIGSITVVIPG--VLRPASRTADFT.AEAI--GTL
      : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|259 TESSAPV---PTPSSSTTESSSPTVTSSTTESSSAPVPTPSSSTTESSSAPAPTSSSS
      570      580      590      600      610

      1510      1520      1530      1540      1550      1560
frame6 .VIGVGVFGRKVTGRRSPSSDFTCTPISRSGFNTR-PIGRLVREASPMKVLKS.TPTR
      : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|259 TTESSSAPAPTSSSTTESSSAPVPTPSSSTTESSSPTVTSSTTESSSVPVPTPSSSTTE
      620      630      640      650      660      670

      1570      1580      1590      1600      1610      1620
frame6 PIVKRQPVAAALPKSRGASGARSPLPTPSMIQVSPSLRILAPIACIALPVLITSSPSKRP
      : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|259 S-----SSAPVSSSTTESSSAPVPTPSSSTTESSSPTVTSSTTESSSVPVPTPSSSTT
      680      690      700      710      720      730

>>gi|154816266|gb|ABS87372.1| flocculin [Saccharomyces c (1360 aa)  
initn: 122 initl: 52 opt: 226 Z-score: 246.2 bits: 58.8 E(): 0.00012
```

Smith-Waterman score: 245; 24.112% identity (49.533% similar) in 535 aa overlap (1130-1645:441-927)

```

      1100      1110      1120      1130      1140      1150
frame6 NSTHNK.EKKQIILI.E.TKGPYHSLTLLHPFPFHSSIAKTE.KTQ.ITSTNGTRKTVF
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|154 SSAPVTSSTTESSAPVPTPCSSSTTESSAPAPTSS--STTESSAPATSST--TESSA
      420      430      440      450      460

      1160      1170      1180      1190      1200      1210
frame6 PMP.YSNSVGFWCVRNETDALNLTNNVETDDTNESSS.RILRS.DPVPSNS.AHQAAVLV
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|154 PVPTPSSST-----TESSAPAPTSSSTTESSAPVTSST--TESSAPAPTSSS
      470      480      490      500      510

      1220      1230      1240      1250      1260      1270
frame6 SESSILAPRPAIKSMNSGKLVAIIVASSTVTGFSRTPITRKLIAIR.SRWVATAAPEAL
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|154 TTESSAPAPAPSSSTTESSAPVTSSTTE---SSAPVTSSTTE---SSAPVTSSTTESS
      520      530      540      550      560      570

      1280      1290      1300      1310      1320
frame6 PRPLPSGRPRTTREV-----PSSQSTPLSLRPFATAERRSL-----LTRSSSKPFIT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|154 SAPAPTSSSTTESSAPAPTSSSTTESSAPVTSSTTESSAPVTSSTTESSAP-VT
      580      590      600      610      620      630

      1330      1340      1350      1360      1370      1380
frame6 VAPSANAAATARIGYSSIEGARSSGTVTPFKVEERTRKSATSSPPARRGLITSMASAPI-
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|154 SSTTESSAPVPTPCSSSTTESS-SAPAPT-----SSSTTESSAPATSSTTESSAPVP
      640      650      660      670      680

      1390      1400      1410      1420      1430      1440
frame6 -----SCRVKMRPVRVGFIKTLRMVTSEPGTSKAATNGKAVEDGSPGTSIT.PVSLPRPS
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|154 TPSSSTTESSAPAPTSSSTTES-SSAPVTS-STTESSAPVPTPSSSTTESSAPVP-
      690      700      710      720      730      740

      1450      1460      1470      1480      1490      1500
frame6 RRMVTPSASVSTVRLAPKP.SIFSVWSRVMIGSITVVIPGVLRPASRTADFT.AEAIGT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|154 -----TPCSSSTTESSAPVPT-----SSSTTESSAPVP---TPSSSTTESSAPVTS
      750      760      770      780      790

      1510      1520      1530      1540      1550      1560
frame6 L.VIGVGFGPRKVTGRRSPSSDFTCTPISRSGFNTR-PIGRLVREASPMKVLKS.TPT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|154 TTESSVAPV-PTPSSSSNITSSAPSSSTPFSSSTTESSVPVPT---PSSSTTESSAPAPT
      800      810      820      830      840

      1570      1580      1590      1600      1610      1620
frame6 RPIVKRQVVAALPKSRGASGASPPPLPTPSMIQVSPSLRILAPIAC-IALPVLITSSPSK
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|154 -PSSSTTESSAPVSSSTTESSVAPVPTPSS---SSNITSSAPSSSTPFSSSTTESSVPVP
      850      860      870      880      890      900

      1630      1640      1650      1660      1670      1680
frame6 RPVIRVSPPLARPPNMKDLWEIDLSPGIRTVPERPEDLRVAGRLLAP.SMHAVETEDMTL
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|154 TPSSSTTESSAPVSSSTTESSVAPVPTPSSSSNITSSAPSSSKYPGSQTETSVSSTTET
      910      920      930      940      950      960
```

>>gi|119589378|gb|EAW68972.1| hCG22561, isoform CRA\_b [H (897 aa)  
initn: 118 initl: 72 opt: 211 Z-score: 231.7 bits: 55.5 E(): 0.00075  
Smith-Waterman score: 221; 24.123% identity (47.807% similar) in 456 aa overlap (1257-1670:169-594)

```

      1230      1240      1250      1260      1270      1280
frame6 PRPAIKSMNSGKLVAIIVASSTVTGFSRTPITRKLIAIR.SRWVATAAPEALPRPLPSG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|119 EETVARGKATEAPKRSALSAGARRDTSRTPGTPSPAMARRSR--AAGTEVLPRPAPSA
      140      150      160      170      180      190

      1290      1300      1310      1320      1330      1340
frame6 RPRTTRE---VSPSSQSTPLSLRPFATAERRSL---LTRSSSKPFITVAPSANAAATAR
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|119 RRPPTTEGRKSVSSASEHSTTEPSAARRRPSAGGGLQRPASRLSSSATPLSSP--AR
      200      210      220      230      240      250

      1350      1360      1370      1380      1390      1400
frame6 IGYSSIIEGARSSGTVTPFKVEERTRKSATSSPPARRGLITSMASAPISCRVKMRPVRVGF
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|119 SGPS-----ARG-----TP-----RAPAHPSQPKPKGL-QALRPP-----QVTTPRKDA
      260      270      280      290

      1410      1420      1430      1440      1450
frame6 IKTLRMVTSEPGTSKAATNGKAVEDGSPGTSIT-.PVSLPRPSRRMVRTPSASVS-----
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|119 APALGPLSSSPLATPSPSGTKARPVPPPDNAATPLPATLP-PSPPVTPPPPALQSQAPP
      300      310      320      330      340      350

      1460      1470      1480      1490      1500
frame6 TVRLAPKP.SIF-SVWSRVMIGSITVVIPGVLR----PASRTADFT.AEAIGTL.VIGV
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|119 TLPATPHSSSLTCQLATPLPLAPPSAPSLSQTLSPSPATPPSQVPPTQLIMSFPEAGV
      360      370      380      390      400      410

      1510      1520      1530      1540      1550      1560
frame6 VGFGRKVTGRRSPS--SDFTCTPISRSGFNTRPIGRLVREAS--PMKVLKS.TPTRPI
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|119 SSLATAAFVASVSPSVSSPLQSMPTQANPALPSLPTLLSPLATPLSAMSPLQGFVSPA
      420      430      440      450      460      470

      1570      1580      1590      1600      1610
frame6 VKRQ---PVAALPKSRGASGARSPP----LTPSMIQVSPSLRILAPIA----CIALP
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|119 TSLGNSAFPLAALPQP-GLSALTTPPPQASPSPSLQATPHTLATLPLQDSPLLATLP
      480      490      500      510      520      530

      1620      1630      1640      1650      1660
frame6 VLITSSP-----SKRPVIRVSPPLARPPNMKDLWEIDLSPGIRTV-ERPEDLRVAGRL
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|119 LQASPSPLTTVSLQDPLVSPSLLASPP-LQAPHPQAPPSMTTPPMQAPPSLQTIPT--I
      540      550      560      570      580

      1670      1680      1690      1700      1710      1720
frame6 LAP.SMHAVETEDMTLRGRSSEFINVIPLFFNPHDDEIG.ARGCCCVFRETDKGLR.LD
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|119 QVPHSLTSPSLQAPPSPALSSSLQATTSLGSPTLQATHSFLTMSPRQTQASLTSPSRPAS
      590      600      610      620      630      640

>>gi|190406204|gb|EDV09471.1| conserved hypothetical pro (751 aa)  
initn: 171 initl: 60 opt: 207 Z-score: 228.3 bits: 54.6 E(): 0.0012
```

Smith-Waterman score: 238; 24.805% identity (50.000% similar) in 512 aa overlap (1132-1623:267-716)

```

      1110      1120      1130      1140      1150      1160
frame6 THNK.EKKQIILI.E.TKGPYHSLTLLHPPFPFHSSIAKTE.KTQ.ITSTTNGTRKTVFPM
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 TTTSESSTSSSTTAPATPTTTTSCSKEKPTPPHHDTPCTPKKKT---TTSKTCTKKTTPV
      240      250      260      270      280      290

      1170      1180      1190      1200      1210      1220
frame6 P.YSNSVGFWCVRNETDALNLTNVVETDDTNESSS.RILRS.DVPVSSNS.AHQAALVSE
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 PTPSSST-----TESSTPVT-----STTESSA-----PVPTPSSSTTES-----
      300      310      320      330

      1230      1240      1250      1260      1270      1280
frame6 SSILAPRPAIKSMNSGKLVAIIVASSTVTFSETRPITRKLIAR.SRWVATAAPEALPR
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 SSAPAPTSSSTTES---STPVTSTTESSSAPVPTPSS-----STTESSSAPA
      340      350      360      370

      1290      1300      1310      1320      1330      1340
frame6 PLPSGRPRTTREVSPSSQSTPLSLRPFATAERRSLSLTRSSSKPFITVAPSANAAATARI
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 PTPSSSTTESSTPVTSTTESSSAPVPTP---SSSTTESSSAPAPTSSSTTESSSSTPV
      380      390      400      410      420      430

      1350      1360      1370      1380      1390
frame6 GYSSIIEGARSSGTVTPFKVEERTRKSATSSPPARRGLITMSAPI-----SCRVKMRP
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 TSST---TESSSAPVPTPSSSTTESSSAPAPTSSSTTESSSAPAPTSSSTTESSSAP
      440      450      460      470      480      490

      1400      1410      1420      1430      1440      1450
frame6 VRVGFIKTLRMVITSEPGTSKAATNGKAVEDGSPGTSIT.PVSLPRPSRRMVRTPSASV--
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 VPTPSSSTTES--SSTPVTSTTESSSAPVPTPSSSTTESSSAPAP-----TPSSSTTE
      500      510      520      530      540

      1460      1470      1480      1490      1500      1510
frame6 STVRLAPKP.SIFSVWSRVMIGSITVVIPG--VLRPASRTADFT.AEAIGTL.VIGVGVF
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 SSSAPAPTSSSTTESSTPVTSTTESSSAPVPTPSSSTTE--SSSTPVTSTTESSSAP
      550      560      570      580      590      600

      1520      1530      1540      1550      1560
frame6 GPRKVTGRRSPSSDFTCTPISRSGFNTR-PIGRLVREAS--PMKVLSKS.T-----PT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 VPTPSSSTTESSSAPAPTSSSTTESSTPVTSTTESSSAPVPTPSSSTTESSSAPAPT
      610      620      630      640      650      660

      1570      1580      1590      1600      1610      1620
frame6 RPIVQRQPVAALPKSRGASGARSPPLPTPSMIQVSPSLRILAPIACIALPVLITSSPSKR
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 -PSSSTTESSTPVTSTTESSSAPVPTPSS--STTESSYAPAPT-----PSSSTTESST
      670      680      690      700      710

      1630      1640      1650      1660      1670      1680
frame6 PVIRVSPLARPPNMKDLWEIDLSPGIRTVPERPEDLRVAGRLLAP.SMHAVETEDMTLR
      : :
gi|190 PVTSTTESSFCSSNSTNSIKLYRKLLCTSSNSIQLYY
      720      730      740      750
```

>>gi|190620456|gb|EDV35980.1| GF12744 [Drosophila ananas (994 aa)  
initn: 74 initl: 44 opt: 208 Z-score: 227.7 bits: 54.9 E(): 0.0013  
Smith-Waterman score: 218; 22.638% identity (49.016% similar) in 508 aa overlap (1135-1627:213-692)

```

      1110      1120      1130      1140      1150      1160
frame6 K.EKKQIILI.E.TKGPYHSLTLLHPPFPFHSSIAKTE.KTQ.ITSTTNGTRKTVFP--MP
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 NPTESTSTQSSICVLFPNLPLCQTSTTSSSTTAAPTSTTAIPTSTTTTQATTSTAIPT
      190      200      210      220      230      240

      1170      1180      1190      1200      1210      1220
frame6 .YSNSVGFWCVRNETDALNLTNVVETDDTNESSS.RILRS.DVPVSSNS.AHQAALVSE-
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 TSTTEAPTSTTAIPTSSSSTTDSATTSTAIPSTTTEAPTSTAIPTTTTTQAPTSTTE
      250      260      270      280      290      300

      1230      1240      1250      1260      1270
frame6 ---SSILAPRP--AIKSMNSGKLVAIIVASSTVTFSETRPITRKLIAR.SRWVATAAP
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 IPTSSTEAPTSTTAIPTPSSSTTDSPTTSTAIPSTTEA-PTTS--TAIPTSTTTTQAP
      310      320      330      340      350

      1280      1290      1300      1310      1320      1330
frame6 E-ALPRPLPSGRPRTTREVSPSSQSTPLSLRPFATAERRSLSL--TRSSSKPFITVAPSA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 TTSTAIPSTSTEAPTSTTAIPTSSSSTTDSPTTSTAIPSTTTEAPTSTAIPTTTTTTQ
      360      370      380      390      400      410

      1340      1350      1360      1370      1380      1390
frame6 NAAATARIGYSSIIEGARSSGTVTPFKVEERTRKSATSSPPAR-RGLITMSAPISCRVK
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 APTTSTEIPTSTEAPTSTTAIPTPSSSSTTDSPTTSTAIPSTTTEAPTSTTAIPTSTTTT
      420      430      440      450      460      470

      1400      1410      1420      1430      1440      1450
frame6 MR-PVRVGFIKTLRMVITSEPGTSKAATNGKAVEDGSPGTSIT.PVSLPRPSRRMVRTPSA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 TQAPTSTTAIPT--SSTEAPTSTTAIPTSSSSTTDSPTTSTAIPSTTTEAPTSTTAIPTST
      480      490      500      510      520      530

      1460      1470      1480      1490      1500
frame6 SVSTVRLAPKP.SIFSVWSRVMIGSITVVIPGVLRPASRTADFT.AEAI--GTL.VIGVG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 TTTTQ-APTTSTAIPST-SSTEAPTSTTAIP--TSSSSTTASATTSTAIPSTTTEAPTST
      540      550      560      570      580      590

      1510      1520      1530      1540      1550      1560
frame6 VFGPRKVTGRRSPSSDFTCTPISRSGFNTRPIGRLVREASPMKVLSKS.TPTRPIVKRQP
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 TAIPSTSTTTQATTSTTAIPTSTTDAPT-----TSTAIPSTSSSSTTDSPTTS-----
      600      610      620      630      640

      1570      1580      1590      1600      1610      1620
frame6 VAALPKSRGASGARSPPLPTPSMIQVSPSLRILAPIACIALPVLITSSPSKRPVIRVSPL
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 -TAIPSTTTEAPTSTTAIPT-----STTTTQAPTSTTAIPSTTTEAPTSTTAIPTSSS
      650      660      670      680      690

      1630      1640      1650      1660      1670      1680
frame6 ARPPNMKDLWEIDLSPGIRTVPERPEDLRVAGRLLAP.SMHAVETEDMTLRGRSSEPIN
```

Monsanto Company

Regulatory Product Characterization Center

Page 172 of 190

RAR-10-070

Monsanto Company

Regulatory Product Characterization Center  
RAR-10-070

```

          1290      1300      1310      1320      1330      1340
frame6  SGRPRTTREVSPSSQ---STPLSLRPFATAERRSLSTRSSSKPFTITVAPSANAAATA-R
      :  : . . . . : . : . : . : . : . : . : . : . : . : . : . :
gi|134  SITPQQQQQRPQRNFVAVSV-DHPPWATATVNVFVSL---PSKTF-DVKPFKNAAWLAEQ
          330      340      350      360      370

```

```

          1540      1550      1560      1570      1580
frame6  -----SGFNTRPIGRLVREASPMKVLKS.TPTRPIVKRQPVAALPKSRGASGARSPPLP
          :...  .::  :  .  ::  .::  :...  .  .  .::  :  ::

```

```

gi|143 SVSIVSALSSGKASRLSLTPSSSESASKGSVPA-----AVSAVSSSRVSPSSSVSALLP
320      330      340      350      360      370

1590      1600      1610      1620      1630      1640
frame6 TPMSIQVSPSLRILAPIACIALPVLITSSPSKRPVIRVSPLARPPNMKDLWEIDLSPGIR
.:.:.
gi|143 VPSSLVSIVSALSSGKASRLSLTPSSSESASKGSVPAAVSAVSSMVSPSSSVSALLPVPS
380      390      400      410      420      430

>>gi|49648025|emb|CAG82478.1| YALI0C22924p [Yarrowia lip (780 aa)
initn: 96 initl: 66 opt: 193 Z-score: 212.2 bits: 51.7 E(): 0.0092
Smith-Waterman score: 223; 23.441% identity (53.367% similar) in 401 aa overlap
(1209-1586:11-397)

1180      1190      1200      1210      1220      1230
frame6 ALNLTNVVETDDTNESSS.RILRS.DPVFSSNS.AHQAAALVSESSILAPRPAIKSMNSGK
.:.:. .:.:. .:.:. .:.:. .:.:. .:.:.
gi|496 MRFHIAVALLATSVVAQGPFFANSTSAVVPQSS-QPANSSS
10      20      30

1240      1250      1260      1270      1280      1290
frame6 LVAIIVASSTVTGFSETRPITRKLIAR.SRWVATAAPEALPRPLPSGRPRITREVPSS
.:.:. .:.:. .:.:. .:.:. .:.:. .:.:.
gi|496 SAPQSSAQSTTPLPVSSAPVSSSAVPSSSAVPSSSAAPVS-SVPSSSAAPVSS---VPSS
40      50      60      70      80      90

1300      1310      1320      1330      1340      1350
frame6 QSTPLSLRPFATAERRSLSLTRSSSKPFITVAPSANAAATARIGYSSIIEGARSSGTVP
.:.:. .:.:. .:.:. .:.:. .:.:. .:.:.
gi|496 SAAPVSSVPSSSAAPVS-SVPSSSAAP-VSSAPSSSAAPVSSVPSSSAAG---SASEAP
100      110      120      130      140      150

1360      1370      1380      1390      1400      1410
frame6 FKVEERTRKSTSSPPARRGLITSMSAPISCRVKMRPVRVGFIKTLRMVTSEPGTS----
.:.:. .:.:. .:.:. .:.:. .:.:. .:.:.
gi|496 --VAANSTSPVASSAPVSSSTTPVSSSTTPVSSVAPSSAPVSSSTSPASSAPVSSSTTPV
160      170      180      190      200

1420      1430      1440      1450      1460
frame6 -KAATNGKA---VEDGSPGTSIT.PVS---LPRPSRRMVRTPASVSTVRLAPKP.SIF
.:.:. .:.:. .:.:. .:.:. .:.:. .:.:.
gi|496 SSAAPSSSEAPVAANSTTPAASSAAPVSSQATTPLPSSAPANSTAPASSSAAAAAGTDAPV
210      220      230      240      250      260

1470      1480      1490      1500      1510
frame6 SVWSRVMIGSI-----TVVIPGVLRP-ASRTADFT.AEAIGTL.VIGGVGFGRPKV
.:.:. .:.:. .:.:. .:.:. .:.:. .:.:.
gi|496 AANSTTPVASSAAPTTPLSSTAAAAGTDAPVAANSTTPVASSAAPTTPLSAATASAVEP
270      280      290      300      310      320

1520      1530      1540      1550      1560      1570
frame6 TGRSPSSDFTCTPISRSGFNTRPIGRVREASPMKVLKSKS.TPTRPI-VKRQPPAALPK
.:.:. .:.:. .:.:. .:.:. .:.:. .:.:.
gi|496 ASSAAPVANST-TPVVPVGSSTTPVGSSVDAASSAAPVANSTTPVVPVGSSTTPVGSSVD
330      340      350      360      370      380

1580      1590      1600      1610      1620      1630
frame6 SRGASGARSPPLTPSMIQVSPSLRILAPIACIALPVLITSSPSKRPVIRVSPLARPPNM
.:.:. .:.:. .:.:. .:.:. .:.:. .:.:.
gi|496 S-ASSGASAPASSTNGTAPVPVTSABDSSAASSATAASSAAAAASKDGAASGTAAPVVT
390      400      410      420      430      440

```

```

>>gi|134078510|emb|CAK40432.1| unnamed protein product [ (1257 aa)
initn: 72 initl: 72 opt: 195 Z-score: 211.6 bits: 52.3 E(): 0.01
Smith-Waterman score: 236; 25.367% identity (49.476% similar) in 477 aa overlap
(1195-1624:349-820)

```

```

1170      1180      1190      1200      1210      1220
frame6 SNSVGFWCVRNETDALNLTNVVETDDTNESSS.RILRS.DPVFSSNS.AHQAAALVSES-S
.:.:. .:.:. .:.:. .:.:. .:.:. .:.:.
gi|134 DENTFSGTLFADVIKIDILVAVDPSPVPSPSSSVIASSTPVASSTPVASSAPASSTPIS
320      330      340      350      360      370

1230      1240      1250      1260      1270
frame6 ILAPRPAIKSMNSGKLVAIIVASST--VTGF--SETRPITRKLIAR.SRWVATAAPEA
.:.:. .:.:. .:.:. .:.:. .:.:. .:.:.
gi|134 SGSPVPSSSAVSSPAVSSSTTESSTQVVSGSVSASSSPITSSPVA--SSTPVASSAPSA
380      390      400      410      420      430

1280      1290      1300      1310      1320      1330
frame6 LPRPLPSGRP-RITREVPSSQSTPLSLRPFATAERRSLSLTRSSSKPFIT-VAPSANAA
.:.:. .:.:. .:.:. .:.:. .:.:. .:.:.
gi|134 TSSAVASSPIAPSSPVASSSAIASSSAIASSSAIASSSAIASSSAIASSSAIASSSAI
440      450      460      470      480      490

1340      1350      1360      1370      1380      1390
frame6 ATARIGYSSIIIEGAR---SSGVTVPFKVEERTRKSTSSPPARRGLITSMSAPISCR--V
.:.:. .:.:. .:.:. .:.:. .:.:. .:.:.
gi|134 SSSAIASSSAIASSSAIASSSAIASSSAIASSSAIASSSPVPVAPSSSPVASSSAIV
500      510      520      530      540      550

1400      1410      1420      1430      1440
frame6 KMRPVRVGFIKTLRMVTSEPG-----TSKAATNGKAVEDGSPGTSIT.PVSLPRPS
.:.:. .:.:. .:.:. .:.:. .:.:. .:.:.
gi|134 SSTPAVSTPVASSIPVISSPAIAGSIAIASSSHVASSSTPAASSSPAVS-SSPVASSSPA
560      570      580      590      600      610

1450      1460      1470      1480      1490
frame6 RRMVRTPASVSTVRLAPKP.SIFSVWSRVMIGSITVVIPGVL---RPASRTADFT.AE
.:.:. .:.:. .:.:. .:.:. .:.:. .:.:.
gi|134 --LSSSPSASASTPIIPSSTASSAVVSSSPTPSSSVVRSSSLSSSSPALSSSTRTPSNP
620      630      640      650      660      670

1500      1510      1520      1530      1540
frame6 AIGTL.VIGGVGF-PRKVT-----GRRS-----PSSDFTCTPISRSGF---NTRPIG
.:.:. .:.:. .:.:. .:.:. .:.:. .:.:.
gi|134 VIPSSSAISITPSSTPVRSTSSVAPGKSSAPVIPKPSSTVIATFTSSSGSLPSSAPAG
680      690      700      710      720      730

1550      1560      1570      1580      1590
frame6 RLVREASPM-KVLKSKS.TPTRPIVKRQPPVA---ALPKSRGASGARSPPLTPSMIQVSPS
.:.:. .:.:. .:.:. .:.:. .:.:. .:.:.
gi|134 SGVPSSSTLPHPSSTSLSSSPVSSAEPVSSSSAVGTSVGSSSNVVTVGVSTRSSSVVPS
740      750      760      770      780      790

1600      1610      1620      1630      1640      1650
frame6 LRILAPIACIALP-VLITSSPSKRPVIRVSPLARPPNMKDLWEIDLSPGIRTVPERPEDL
.:.:. .:.:. .:.:. .:.:. .:.:. .:.:.
gi|134 GTPIPPVSGTATESVTSSSSGSGSPTVPSSINTSSTDASSSSSASSVEPTSSSGSVITSVT
800      810      820      830      840      850

```

```

>>gi|170936811|emb|CAP61469.1| unnamed protein product [ (543 aa)
initn: 57 initl: 57 opt: 190 Z-score: 211.0 bits: 50.9 E(): 0.011

```

Smith-Waterman score: 190; 23.153% identity (47.044% similar) in 406 aa overlap (1276-1661:44-425)

```

      1250      1260      1270      1280      1290      1300
frame6 SSTVTGFSETRPITRKLIAIR.SRWVATAAPEALPRPLPSGRPRITREVSPS--SQSTPL
      . . . . .
gi|170 VAEPTTMLMIEPAPVFTVTPSQQLVLFESMPAAQPAFTPAAGPSKMDVLCSSLGALAVQC
      20      30      40      50      60      70

      1310      1320      1330      1340      1350      1360
frame6 SLRPFATAERRSLSLTRSSSKPFITVAPSANAAATARIGYSSIIIEGARSSGTVTPFKV-E
      . . . . .
gi|170 AHTPMASIRQESWAVQRFV-RPAFDLQPAALMPAPRLPGLSSLVQPTVVASVVTSSSEVME
      80      90      100      110      120      130

      1370      1380      1390      1400      1410      1420
frame6 ERTRKSATSSPPARRG-LITSMSAPISCRVKMRPVRVGFIKTLRMVTSEPGTSKAATNGK
      . . . . .
gi|170 TEIVRPAVQLPPKPAQQLPLPLAPIVPVPAVPTP---LAATDNVAVAPPAPEMEVRPQQ
      140      150      160      170      180

      1430      1440      1450      1460      1470
frame6 AVEDGSPGTSIT.PVSLPRPSRRMV-RTPSASVSTVRLAPKP.SIFSVWSRVVMIG
      . . . . .
gi|170 AAVQPLKRNRRVAVPNLVIQQVPQPELVVPPRRPAAPVNLVRVPPAP-----VQTS
      190      200      210      220      230      240

      1480      1490      1500      1510      1520      1530
frame6 SITTVIPGVLRPASRTADFT.AEAIGTL.VIGGVGFGRKVTGRRSPSSDFTCTPISRSG
      . . . . .
gi|170 PQQPLPPVQSPPPARVPRNTPAGPLN-LVIAPPTLASSSRPQQAQAPKQOTPPAPASRAP
      250      260      270      280      290

      1540      1550      1560      1570      1580
frame6 ---FNTRPIGRLVREASPMKVLKS.TPTR--PIV-KRQP-VAALPKSRGASGARSPPPL
      . . . . .
gi|170 SPPPIPSRPHWKLVKKLSPSEAQAKLTLPALPSVPKVTPEVAPLPQP--APVVPVPA
      300      310      320      330      340      350

      1590      1600      1610      1620      1630      1640
frame6 PTPSMIQVSPSLRILAPIACIALPVLITS--SPSKRPVIRVSPLARPPNMKDLWEIDLSP
      . . . . .
gi|170 PAPVAVPLPPA---AAPAISDAVPEQQAIPPPGSRP---VAPMRRRRAFGNT--VPLLP
      360      370      380      390      400
```

```

      1650      1660      1670      1680      1690      1700
frame6 GIRTVPERPEDLRAVAGRLAP.SMHAVETEDMTLGRSSEPINVIPLFFNPHDDEIG.A
      . . . . .
gi|170 DASVLEQKPLPATVAAAAESSGSMKAPAPTPLPPARTAHWAGLELRTLADRLIKECVD
      410      420      430      440      450      460
```

>>gi|219885071|gb|ACL52910.1| unknown [Zea mays] (578 aa)  
initn: 112 initl: 77 opt: 190 Z-score: 210.6 bits: 51.0 E(): 0.011  
Smith-Waterman score: 203; 22.670% identity (50.126% similar) in 397 aa overlap (1284-1667:20-373)

```

      1260      1270      1280      1290      1300
frame6 ETRPITRKLIAIR.SRWVATAAPEALPRPLPSGRPRITR---EVSPSSQSITPLSLRPFA
      . . . . .
gi|219 MEDLLGSEIGKNDYDWLLTPPGTPRVPALEGAETPSSNIPKRTTTRS
      10      20      30      40
```

```

      1310      1320      1330      1340      1350      1360
frame6 TAERRSLSLTRSSSKPFITVAPSANAAATARIGYSSIIIEGARSSGTVTPFKVEERTRKSA
      . . . . .
gi|219 SSTTRASRLSISQTENGNSTIPT-RPARSNSVSRSSI-QSALMNGNNRT-SVLNTSISSV
      50      60      70      80      90      100
```

```

      1370      1380      1390      1400      1410      1420
frame6 TSSP--PARRGLITSM---SAPISCRVKMR---PVRVGFIKTLRMVTSEPGTSKAATNGK
      . . . . .
gi|219 TSRPTTPSRRSSTVGTPKHSAPAARPVPARSSTPVNTRPSTPAKTRPSTPTRTRPTAPNS
      110      120      130      140      150      160
```

```

      1430      1440      1450      1460      1470      1480
frame6 AVEDGSPGTSIT.PVSLPRPSRRMVRTPSASVSTVRLAPKP.SIFSVWSRVVMIGSITVVI
      . . . . .
gi|219 STDSAAAKTFTTHNARPSTPNRSRIVPNSSSSATHVISRPSASLSATSRP--GSSSGNV
      170      180      190      200      210      220
```

```

      1490      1500      1510      1520      1530      1540
frame6 PGVLRPASRTADFT.AEAIGTL.VIGGVGFGRKVTGRRSPSSDFTCTPISRSGFNTRPI
      . . . . .
gi|219 PGISR-----ATSLSSSVSPSMSRSSRSSTPTRQPAIRSSAPAIGRSP----SV
      230      240      250      260      270
```

```

      1550      1560      1570      1580      1590      1600
frame6 GRLVREASPMKVLKS.TPTRPIVKRQPVAAALPKSRGASGARSPPPLTPSMIQVSPSLRI
      . . . . .
gi|219 GR---TSSINSFTTSS-----RSAASTGRNSAPSSAPSSRPSSPNPRP
      280      290      300      310
```

```

      1610      1620      1630      1640      1650      1660
frame6 LAPIACIALPVLITSSPKRPVIRVSPLARPPNMKDLWEIDLSPGIRTVPE-RPEDLRAP
      . . . . .
gi|219 RAPVRPLDIPDPFNETP---PNLRTKLPERPPSAGRS-RPGMALGIRSTPNTPEPSAASAP
      320      330      340      350      360
```

```

      1670      1680      1690      1700      1710      1720
frame6 AGRLLAP.SMHAVETEDMTLGRSSEPINVIPLFFNPHDDEIG.ARGCCCVFRETDKGDL
      . . . . .
gi|219 VKKISVPAVSRSKFSDASSKTSTLTNGHQNRQADRSNIDSQANRPSRPSTGTITDNGFGRS
      370      380      390      400      410      420
```

>>gi|195615020|gb|ACG29340.1| transposon protein CACTA, (578 aa)  
initn: 112 initl: 77 opt: 190 Z-score: 210.6 bits: 51.0 E(): 0.011  
Smith-Waterman score: 203; 29.000% identity (51.000% similar) in 300 aa overlap (1183-1449:38-320)

```

      1160      1170      1180      1190      1200      1210
frame6 GTRKTVFPM.P.YSNVGVFWCVRNETDALNLTNVVETDDTNESSS.-RILR-S.DVPVSSN
      . . . . .
gi|195 EIGKNDYDWLLTPPGTPRVPALEGAETPPTNIPKRTTTRSSSTTRASRLSISQTENG
      10      20      30      40      50      60
```

```

      1220      1230      1240      1250      1260
frame6 S.AHQALVSESSILAPRAIKS--MNSGKLVAIIVAS-STVTGFSETRPITRKLIAIR.
      . . . . .
gi|195 STIPTRPARSNS---VSRSSIQSALMNGNNRTSVLNTSISSVT---SRPTTPS---RR
      70      80      90      100      110
```

```

      1270      1280      1290      1300
frame6 SRWVAT---AAPEALPRPLPSGRPRITREVSPSSQ--STPLSLRP-----
      . . . . .
```



```
>>gi|260057409|gb|ACX27026.1| Sequence 60583 from patent (578 aa)
  initn: 112 initl: 77 opt: 190 Z-score: 210.6 bits: 51.0 E(): 0.011
Smith-Waterman score: 203; 29.000% identity (51.000% similar) in 300 aa overlap
(1183-1449:38-320)
```

```

1100      1110      1120      1130      1140      1150
frame6  NVCISLIISNSTHNK.EKKQIILI.E.TKGPYHSLLTLHPFFPHSSIAKTE.KTQ.ITST
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|990  VPTPSSSTTESSTPVTSSSTTESSSAPAPTSSSSTTESSSAPVTSST--TESSSAPVSS
      590      600      610      620      630      640

1160      1170      1180      1190      1200
frame6  TNGTRKTVPFMP.YSNSVGFVCWRNETDALNLTNVETDDTNESSS.RILRS.D-----
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|990  T--TESSVAPVPTPSSST-----TESSSAPVPTPSSSSTTESSTPVTSSSTTESSSAP
      650      660      670      680      690

1210      1220      1230      1240      1250      1260
frame6  -PVPSSNS.AHQAAVLSES---SILAPRPAIKSMNSGKLVAIIVASSTVTGFSETRPITR
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|990  APTPSSSTTESSSAPVTSSTTESSSAPVPTPSS-SITESSTPVTSSSTTESSSAPAPTPS
      700      710      720      730      740      750

1270      1280      1290      1300      1310
frame6  KLIAIR.SRWVATAAPEALPRPLPSGPRRTREVS-PSSQSTPLSLRPFATAERRSLSLT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|990  SSTTESSSAPVTSSTTESSSAPVPTPSSSSTTESSTPVTSSSTTESSS--APAPTSSSTT
      760      770      780      790      800      810

1320      1330      1340      1350      1360      1370
frame6  RSSSKPFITVAPSANAAATARIGYSSIIIEGARSSGTVTPFKVEERTKRSATSPPARGL
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|990  ESSSAP-VTSSTTESSSAPAPTSSSTTES--SSAPVTS---STTESSSAPVPTPSSSTT
      820      830      840      850      860

1380      1390      1400      1410      1420      1430
frame6  ITSMSAPISCRVKMRPVRVGFIKTLRMVTSSEPGTSKAATNGKAVEDGS--PGTSIT.PVS
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|990  TESSSAPVSSSTTESSV--APVPTPSSSSTTESSSAPVPTPSSSSTTESSTPVTSSSTESS
      870      880      890      900      910      920

1440      1450      1460      1470      1480
frame6  ---LPRPSRRMVRTPSASV-----STVRLAPKP.SIFSVWSRVMIGSITV---VIPGLV
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|990  SAPAPTSSSSTTESSSAPVTSSTTESSSAQVPTPSSSSTTESSSAPVSSSTTESSVAP-VP
      930      940      950      960      970      980

1490      1500      1510      1520      1530      1540
frame6  RPASRTADFT.AEAIGTL.VIGVGVGPRKVTGRRSPSSDFTCTPISRSGFNTRPIGRLV
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|990  TPSSSSTTESSSAPVPTPSSSSTTESSTPVTSSSTTESSSAPAPTSSSSTTESSSAPVSSST
      990      1000      1010      1020      1030      1040

1550      1560      1570      1580      1590
frame6  REAS--PMKVLSKS.T-----PTRPIVKRQPVAAALPKSRGASGARSPPLTPISM----
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|990  TESSVAPVPTPSSSSTTESSSAPVPT-PSSSTTESSTPVTSSSTTESSVAPVPTPSSSTTE
      1050      1060      1070      1080      1090      1100

1600      1610      1620      1630      1640
frame6  -----IQVSPSLRILAPIACIALPVLITSS-PSKRPVIRVSPLARPPNMKDLWEIDLSPG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

gi|990 SSSAPVSSSTTESSVAPVPTPSSSSNITSSAPSSSTPFSSSTESSVAPVPTPSSSTTESSS
      1110      1120      1130      1140      1150      1160

      1650      1660      1670      1680      1690      1700
frame6 IRTVPERPEDLRAVAGRLAP.SMHAVETEDMTLRGRSSEPINVIPLEFFNPHDDEIG.AR

gi|990 APVSSSTTESSVAPVPTPSSSSNITSSAPSSSTPFSSSTTESFSTGTVTVPSSSKYPGSQTE
      1170      1180      1190      1200      1210      1220

>>gi|53964938|gb|AAV14935.1| Sequence 26892 from patent (557 aa)
      initn: 56 initl: 56 opt: 188 Z-score: 208.6 bits: 50.5 E(): 0.015
Smith-Waterman score: 198; 21.593% identity (47.799% similar) in 477 aa overlap
(1177-1627:93-532)

      1150      1160      1170      1180      1190      1200
frame6 ITSTNGTRKTVFPMF.YSNSVGFWCVRNETDALNLTNVVETDDTNESSS.RILRS.DPV
      ... ..
gi|539 TESTPATESTPATESTPATESTPATESTPCTTSTETTPATESTASTETASSTPV
      70      80      90      100      110      120

      1210      1220      1230      1240      1250      1260
frame6 PSSNS.AHQAAALVSESSILAPRPAIKSMNSGKLVAIIVASSTVTGFSETRPITRKLIAR
      ..
gi|539 EST-VIVPSTTVITVSSCYEDKCSVSSVTG----VVTISSEETIYTTYCPITSS-ITIP
      130      140      150      160      170

      1270      1280      1290      1300      1310      1320
frame6 .SRWVATAAPEALPRPLPSGRPRITREVPSSQSTPLSLRPFATAERRSLSLTRSSSKPF
      . : : :
gi|539 VPNTSTPAAP-----GTPVESQPVIPGTETTPAA--PGTPVE-----SQPV
      180      190      200      210

      1330      1340      1350      1360      1370      1380
frame6 ITVAPSANAAATARIGYSSIEGARSSGTVTPFKVEERTRKSATSSPPARRGLITMSAP
      . : : :
gi|539 IPGTETTPAAPGTPVESQPVIPGTETTPAAPGTPVE----SQPATTPVAPGTETTPAAP
      220      230      240      250      260      270

      1390      1400      1410      1420      1430      1440
frame6 ISCRVKMRPVRVGFIKTLRMVTSEPGT---SKAATNGKAVEDGSPGTSI-T.PVSLP-RP
      . : : :
gi|539 -GTPVESQPATTPVAPGTETTPAAPGTPVESQPVIPGTETTPAAPGTPVESQPATTPVAP
      280      290      300      310      320

      1450      1460      1470      1480      1490
frame6 SRRMVRTPSASVSTVRLAPK--P.SIFSVWSRVMIGSITVVIPGVLRPASRTADFT.AEA
      . : : :
gi|539 GTE--TTPAAPGTPVESQPATTPVAPGTETTPAAPGTPVESQPATTPVAPGTETTPAAPG
      330      340      350      360      370      380

      1500      1510      1520      1530      1540      1550
frame6 IGTL.VIGVGVFGRKVTGRRSPSSDFTCTPISRSGFNTRPI--GRLVREASPMKVLSSK
      . : : :
gi|539 TPVESQPATTPVTPGTETTPAAPGTPVDSQPVT-PGTETTPAAPGTPGTEATPVTTQPV
      390      400      410      420      430      440

      1560      1570      1580      1590      1600
frame6 .TPTRPIV-----KRQPVAAALPKS---RGASGARSPLPTPSMIQVSPSLRI
      : :
gi|539 VLSTLQVVTASGEFSTVTAHSTSIIVASCPEGGCVPEGQQTETSPSVPTNGP-EVEASSV
      450      460      470      480      490      500

```

```

      1610      1620      1630      1640      1650      1660
frame6 LA-PIACIALPVLITSSPSKRPVIRVSPILARPNNMKDLWEIDLSPGIRTVPERPEDLRAV
      : : : :
gi|539 LSIPVSSVTSTTIASSSETSVPPAQVSTFEGSGSALKPKPYGLAVAALVYFM
      510      520      530      540      550

>>gi|19526645|gb|AAL89737.1|AF430017_1 intestinal membra (1029 aa)
      initn: 142 initl: 65 opt: 190 Z-score: 207.1 bits: 51.1 E(): 0.018
Smith-Waterman score: 217; 23.146% identity (51.011% similar) in 445 aa overlap
(1205-1624:236-665)

      1180      1190      1200      1210      1220      1230
frame6 NETDALNLTNVVETDDTNESSS.RILRS.DVPSSNS.AHQAAALVSESSILAPRPAIKSM
      : : : :
gi|195 TQVSSSPVTPEGTTMPWIWTPSEGSTPLTTPVSTTRVTSSEGSTLSTPSVVTSTPTVTTST
      210      220      230      240      250      260

      1240      1250      1260      1270      1280      1290
frame6 NSGKLVAIIVASSTVTGFSETRPITRKLIAR.SRWVATAAPEALPRPLPSGRPRITREV
      . : : :
gi|195 EAISSSATL--DSTTMSVSMPEISTLGTTLVSTTPVTRFPESSTPSIPSVYTSMSMTT
      270      280      290      300      310      320

      1300      1310      1320      1330      1340
frame6 SPSSQSTPLSLR----PFATAERRSLSLTRSSSKPFITVAPSANAAATARIGYSS--II
      . : : :
gi|195 ASEGSSSPITLEGTTMPMSTTSERSTLLTTLVISPISVMSPEASTLSTPPGDTSTPLL
      330      340      350      360      370      380

      1350      1360      1370      1380      1390      1400
frame6 EGARSSGTVTPFKVEERTKRSATS--SPARRGLITSM---SAPISCRVKMRPVRVGFVK
      . : : :
gi|195 TSTKAGSFIPAEVTT-IRISITSERSTPLTLLVSTTLTSPFGASIASTPPLDTSTTF
      390      400      410      420      430      440

      1410      1420      1430      1440      1450
frame6 TLRMVT-SEPGTSKAATNGKAV--EDGSPGTSIT.PVSLPRPSRRMVRTPSAS--VSTVR
      : : : :
gi|195 TPSTDIASPTIPVATTISVSVITEGSTPGTTIFIP-STPVTSSADVFPAATGAVSTPV
      450      460      470      480      490      500

      1460      1470      1480      1490      1500      1510
frame6 LAPKP.SIFSVWSRVMIGSITVVIPGVLRPASRTADFT.AEAIGTL.VIGVGVFGRKVT
      . : : :
gi|195 ITSTELNTPSTSSSSTTTSFSTTKE-FTTPAMTTA-----APLTYVTMSTAPSTPRITS
      510      520      530      540      550

      1520      1530      1540      1550      1560      1570
frame6 -GRRSPSSDF--TCTPISRSGFNTRPIGRLVREASPMKVLSSK.TPTRPIVKRQPVAAALP
      . : : :
gi|195 RGCTTSASTLSATSPHTSTSVTRPVTPSESSSRPSTITSHTIPPTFPPAH----SSTP
      560      570      580      590      600      610

      1580      1590      1600      1610      1620
frame6 KSRGASGARSPLPTPSMI-QVPSRLILA-P-IACIALPV--LITSSPSKRPVIRVSP
      . : : :
gi|195 PTTSASSTVNPEAVTMTTRTKPSTRITTSFPPTVTTTAVPTNTTIKSNPTSTPTVPRIT
      620      630      640      650      660      670

      1630      1640      1650      1660      1670      1680
frame6 ARPPNNMKDLWEIDLSPGIRTVPERPEDLRAVAGRLAP.SMHAVETEDMTLRGRSSEPIN

```

```

gi|195 CFGDGCQNTASRCKNGGTWDLKLCQCPNLYYGELCEEVSSIDIGPPETISAQMELTVTV
      680      690      700      710      720      730

>>gi|118835615|gb|AAI26316.1| MUC17 protein [Homo sapien (1224 aa)
  initn: 142 initl: 65 opt: 190 Z-score: 206.1 bits: 51.2 E(): 0.02
Smith-Waterman score: 231; 22.629% identity (50.250% similar) in 601 aa overlap
(1118-1691:4-558)

      1090      1100      1110      1120      1130      1140
frame6 SKQNVCIISLIISNSTHNK.EKKQIILI.E.TKGPHYSLTLLHPFFHSS-IAKTE.KTQ.
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|118      SMPSTSTPSEGSTPLTSMVPVSTTTVASSETSTLS
      10      20      30

      1150      1160      1170      1180      1190
frame6 ITSTNGTRKTVF-----P-----MP.YSNSVGFWCVRNETDALNLTINVETDDTNE
      : . . . : . . : : : : : : : : : : : : : : : : : : : :
gi|118 TTPADTSTPVTIYSQASSSPPIADGTSMPSTSTYSEGSTPLTNMS--FSTTPVVSSEASTL
      40      50      60      70      80      90

      1200      1210      1220      1230      1240      1250
frame6 SSS.RILRS.DPVPSNS.AHQALVSESSI--LAPRPAIKSMNSGKLVAIIVASSTVTG
      : . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|118 STTP--VDTSTPVTSTEASLSPTTAEGTSIPTSSPSEGTTPLASMPVSTTPVVSSEVNT
      100      110      120      130      140

      1260      1270      1280      1290      1300      1310
frame6 FSETRPITRKLIAR.SRWVATAAPE-ALPRPLPSGRPRRTTREVSPSSQSTPLSLRPFAT
      : : : . . . . . : : : : : : : : : : : : : : : : : : :
gi|118 LSTT-PVDSNTLVTSTE--ASSSPTIAEGTSLPTS---TTSEGSTPLSIMPLSTTPVAS
      150      160      170      180      190      200

      1320      1330      1340      1350      1360
frame6 AERRSLSLTR-SSSKPFITVAPSANAAATARIGY---SSIIEGAR--SSGVTVPFKVEER
      : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|118 SEASTLSTTPVDSTPVTSTPTNGSPTTAEVTSMPSTSTAGEGSTPLTNMPVSTTPVASS
      210      220      230      240      250      260

      1370      1380      1390      1400      1410      1420
frame6 TRKSATSSPPARRGLITMSAPISCRVKMRPVRVGFIKTLRMVTSEPGTSKAATN--GKA
      : . . . : : : : : : : : : : : : : : : : : : : : : :
gi|118 EASTLSTTPVDSNTFVTSSSQASSSPATLQ-----VTTMRMSTPSEGSSSLTLMLSST
      270      280      290      300      310

      1430      1440      1450      1460      1470      1480
frame6 VEDGSPGTSIT.PVSLPRPSRRMVRTPSASVSTVRLAPKP.SIFSVWSRVMIGSITVVIP
      : . . . : : : : : : : : : : : : : : : : : : : : : :
gi|118 YVTSSEASTPSTP-SVDRSTP--VTTSTQSNST---PTPPEVITL---PMSTPSEVSTP
      320      330      340      350      360

      1490      1500      1510      1520      1530      1540
frame6 GVLRPASRTADFT.AEAIGTL.VIGGVFGPRKVTGRSPSSDFCTCTPISRSGFNTRPIG
      : . . . : : : : : : : : : : : : : : : : : : : : : :
gi|118 LTIMPVS-TTSVTISEA-GTASTLPVDTSTPV-IT----STQVSSSPVTPEG-TTMPIW
      370      380      390      400      410

      1550      1560      1570      1580      1590      1600
frame6 RLVREASPMKVLSKS.TPTRPIVKRQPAALPKSRGASGARSPPLPTPSMIQVSPSLRIL
      : . . . : : : : : : : : : : : : : : : : : : : : : :
gi|118 TPSEGSTPLTTPVST--TRVTSSEGSTLSTP-----SVVSTSTPVT--STEAISSSATLD
      420      430      440      450      460      470

```

```

      1610      1620      1630      1640      1650      1660
frame6 APIACIALPVLITSSPSKRPVIRVSPPLARPP--NMKDLWEIDLSPGIRIVPERPEDLRAV
      : . . . : : : : : : : : : : : : : : : : : : : : : :
gi|118 STTMSVSMPEI-STLGTTLVSTTPVTRFPESSTPSIPSVYTSMSMTTASEGSSSPTTL
      480      490      500      510      520

      1670      1680      1690      1700      1710      1720
frame6 AGRLLAP.SMHAVETEDMTLRGRSSEPINVIPLFFNPHDEIG.ARGCCCVFRETDKGDGL
      : : : . . . : : : : : : : : : : : : : : : : : : :
gi|118 EGTITMPMTTTERSTLLTTLVLI--PISVMSPEASTLSTPPGDTSTPLLTSTKAGSFS
      530      540      550      560      570      580

      1730      1740      1750      1760      1770      1780
frame6 R.LDFERLEIRDGFCTPLQILLTCAMVDLWSTNLQKKEQSNIESIKTVDIYNHYENHL.H

gi|118 IPA EVTIRISITSERSTPLTTLVSTTLPTSPFGASIASTPPLDTSTTFTPTSTDTASTP
      590      600      610      620      630      640

>>gi|187981249|gb|EDU47875.1| conserved hypothetical pro (812 aa)
  initn: 95 initl: 64 opt: 187 Z-score: 205.2 bits: 50.4 E(): 0.023
Smith-Waterman score: 187; 23.936% identity (50.266% similar) in 376 aa overlap
(1276-1635:444-796)

      1250      1260      1270      1280      1290      1300
frame6 SSVTVGFSETRPITRKLIAR.SRWVATAAPEALPRPLPSGRPRRTTREVSPSSQSTPLSL
      : : : : : : : : : : : : : : : : : : : : : : : : :
gi|187 SLRSFHMNPANTMSSRRRSMQPLPITDENFPQAIKTP-PNSKRHSL--ASPTSATTPELG
      420      430      440      450      460      470

      1310      1320      1330      1340      1350      1360
frame6 RPFATAERRSLSLTRSSSKPFIT-VAPSANAAATARIGYSSI-IE-GARSSGTVTPFKVE
      : : : : : : : : : : : : : : : : : : : : : : : : :
gi|187 APFVVRP---PMPKVSTPPAIEEVAPVSEPITAARASPPQVKIDVPEQPSDENVPTRED
      480      490      500      510      520

      1370      1380      1390      1400      1410
frame6 ETRTKSA-TSSPPARRGLITMSAPISCRVKMRP--VRVGFIKTLRMVTSEPGTSKAATN
      : : : : : : : : : : : : : : : : : : : : : : : : :
gi|187 SPVRFNAPTYTAPPRDVI---SPPPKDRAPVPPQARQPLAGSSPTFVAGPAPLYASTN
      530      540      550      560      570      580

      1420      1430      1440      1450      1460      1470
frame6 GKAVEDGSPGTSIT-.PVSLPRPSRRMVRTPSASVSTVRLAPKP.SIFSVWSRVMIGSIT
      : . . : : : : : : : : : : : : : : : : : : : : : : :
gi|187 ---VDRSARRTSASPKPFLRPFPVRPQAQHGDSITTIVERRQSSLTAPFTLSLPKVN-R
      590      600      610      620      630

      1480      1490      1500      1510      1520      1530
frame6 VVIPGVLRPASRTADFT.AEAIGTL.VIGGVFGPRKVTGRSPSSDFCTCT-PI-SRSGF
      : : : : : : : : : : : : : : : : : : : : : : : : :
gi|187 VTAP--IRPPSAASNMTPISKPHSLQVGPTVLRPPASVQIRPANSAAFTSSRPVLAISST
      640      650      660      670      680      690

      1540      1550      1560      1570      1580
frame6 NTRPIGRLVREASPMKVLSKS.TPTRPIVKR-----QPVAALPKSRGASGARSPPLPT
      : : : . . . : : : : : : : : : : : : : : : : : : :
gi|187 PSFVPGRRMSNVNPPATTAKPTPSVEMLRQKAFQEQQLQORTITPKRSFASMGPPAPP
      700      710      720      730      740      750

      1590      1600      1610      1620      1630      1640
frame6 PSMIQVSPSLRILAPIACIALPVLITSSPSKRPVIRVSPPLARPPNMKDLWEIDLSPGIRT
      : . : : : : : : : : : : : : : : : : : : : : : : :

```

```

gi|187 PTVPLPSP-RLA-----LTPPTMPLPPPPQSMGLPAPPPNMPLPPTPPEVALRGT
      760      770      780      790      800      810

      1650      1660      1670      1680      1690      1700
frame6 VPERPEDLRVAGRLLAP.SMHAVETEDMTLRGRSSEPINVIPLEFFNPHDDEIG.ARGCC

gi|187 AI

>>gi|33310026|gb|AAQ03243.1|AF414112_1 putative cell wal (1121 aa)
  initn: 56 initl: 56 opt: 188 Z-score: 204.3 bits: 50.8 E(): 0.025
Smith-Waterman score: 214; 22.407% identity (51.867% similar) in 482 aa overlap
(1188-1654:405-858)

      1160      1170      1180      1190      1200      1210
frame6 VFPMP.YSNSVGFWCVRNETDALNLTNVVETDDTNESSS.RILRS.DVPSSNS.-AHQA
      380      390      400      410      420      430

      1220      1230      1240      1250      1260      1270
frame6 ALVSESS-ILAPRPAIKSMNSGKLVAIIVASSTVTGFSSTRPITRKLIAIR.SRWVATAA
      440      450      460      470      480      490
gi|333 TPATESTPATESTPATESTPATESTPCTTSTESTPATESTPATE-STPATEST
      440      450      460      470      480      490

      1280      1290      1300      1310      1320      1330
frame6 PEALPRPLPSGRPRTR-EVSPSSQSTPLSLR-PFATAERRSLSLTRSSSKPFITVAPSA
      500      510      520      530      540      550
gi|333 PATESTPCTTSTESTPATESTPATESTPATESTPATESTPATESTPAT
      500      510      520      530      540      550

      1340      1350      1360      1370      1380      1390
frame6 NAA-ATARIGYSSIEGARSSGTVTPFKVEERTRKSATSSPPARRGLITMSAPISCRVK
      560      570      580      590      600      610
gi|333 ESTPATESTPATESTPATESTPATEST--PATESTPATESTPCTTSTET
      560      570      580      590      600      610

      1400      1410      1420      1430      1440
frame6 MRPVRVGFIKTLRMVISE--PGT-SKAATNGKAVEDGSPGTSIT.PVSLPRPSRRMVTRP
      620      630      640      650      660
gi|333 TTPATESTPATESTPATESTPATESTPATESTPATESTPATE-STPATESTPCTTSTETT
      620      630      640      650      660

      1450      1460      1470      1480      1490      1500
frame6 SASVSTVRLAPKP.SIFSVWSRVMIGSITVVIPGVLRLPASRTADFT.AEAIGTL.VIGVG
      670      680      690      700      710
gi|333 PATESTA--STETASSTPVESTVIVPSTTVI-----TVSSCYEDKCSVSSVTTGTVT
      670      680      690      700      710

      1510      1520      1530      1540      1550      1560
frame6 VFGPRKVTGRRSPSSDFTCTPISRSGFNTRPIGRVLRASPMKVLKSKS.TPTRP--IVKR
      720      730      740      750      760      770
gi|333 ISSEETIYTYCPIITSSITIPVNTSTPAAP-GTPV-ESQPV-IPGTETTPAAGTPVES
      720      730      740      750      760      770

      1570      1580      1590      1600      1610      1620
frame6 QPVAALPKSRGASGARSPLPTPSMIQVSPSLRILAPIACIALPVLITSSPSKRPVI---
      780      790      800      810      820
gi|333 QPV--IPGTETTPAA--PGTPVESQ-PVIPGTET-TP-AAGTPV--ESQPATTPVAPGT
      780      790      800      810      820

```

```

      1630      1640      1650      1660      1670      1680
frame6 RVSPLARP--PNMKDLWEIDLSPGIRTVPERPEDLRVAGRLLAP.SMHAVETEDMTLRG
      830      840      850      860      870      880

      1690      1700      1710      1720      1730      1740
frame6 RSSEPINVIPLEFFNPHDDEIG.ARGCCCVFRETDKGDRLR.LDFERLEIRDGFCTPLQILL

gi|333 ATTPVAPGTETTPAAGTPVESQPATTPVAPGTETTPAAGTPVESQPATTPVAPGTETT
      890      900      910      920      930      940

>>gi|155709149|gb|ABU33860.1| Sequence 299 from patent U (1121 aa)
  initn: 56 initl: 56 opt: 188 Z-score: 204.3 bits: 50.8 E(): 0.025
Smith-Waterman score: 214; 22.407% identity (51.867% similar) in 482 aa overlap
(1188-1654:405-858)

      1160      1170      1180      1190      1200      1210
frame6 VFPMP.YSNSVGFWCVRNETDALNLTNVVETDDTNESSS.RILRS.DVPSSNS.-AHQA
      380      390      400      410      420      430

gi|155 TESTPATESTPATESTPATESTPCTTSTESTPATESTPATE-STPATEST
      380      390      400      410      420      430

      1220      1230      1240      1250      1260      1270
frame6 ALVSESS-ILAPRPAIKSMNSGKLVAIIVASSTVTGFSSTRPITRKLIAIR.SRWVATAA
      440      450      460      470      480      490
gi|155 TPATESTPATESTPATESTPC-TTSTESTPATESTPATE-STPATEST
      440      450      460      470      480      490

      1280      1290      1300      1310      1320      1330
frame6 PEALPRPLPSGRPRTR-EVSPSSQSTPLSLR-PFATAERRSLSLTRSSSKPFITVAPSA
      500      510      520      530      540      550
gi|155 PATESTPCTTSTESTPATESTPATESTPATESTPATESTPATESTPAT
      500      510      520      530      540      550

      1340      1350      1360      1370      1380      1390
frame6 NAA-ATARIGYSSIEGARSSGTVTPFKVEERTRKSATSSPPARRGLITMSAPISCRVK
      560      570      580      590      600      610
gi|155 ESTPATESTPATESTPATESTPATEST--PATESTPATESTPCTTSTET
      560      570      580      590      600      610

      1400      1410      1420      1430      1440
frame6 MRPVRVGFIKTLRMVISE--PGT-SKAATNGKAVEDGSPGTSIT.PVSLPRPSRRMVTRP
      620      630      640      650      660
gi|155 TTPATESTPATESTPATESTPATESTPATESTPATESTPATE-STPATESTPCTTSTETT
      620      630      640      650      660

      1450      1460      1470      1480      1490      1500
frame6 SASVSTVRLAPKP.SIFSVWSRVMIGSITVVIPGVLRLPASRTADFT.AEAIGTL.VIGVG
      670      680      690      700      710
gi|155 PATESTA--STETASSTPVESTVIVPSTTVI-----TVSSCYEDKCSVSSVTTGTVT
      670      680      690      700      710

      1510      1520      1530      1540      1550      1560
frame6 VFGPRKVTGRRSPSSDFTCTPISRSGFNTRPIGRVLRASPMKVLKSKS.TPTRP--IVKR
      720      730      740      750      760      770
gi|155 ISSEETIYTYCPIITSSITIPVNTSTPAAP-GTPV-ESQPV-IPGTETTPAAGTPVES
      720      730      740      750      760      770

      1570      1580      1590      1600      1610      1620
frame6 QPVAALPKSRGASGARSPLPTPSMIQVSPSLRILAPIACIALPVLITSSPSKRPVI---
      780      790      800      810      820

```



Smith-Waterman score: 183; 25.503% identity (47.987% similar) in 298 aa overlap (1276-1565:7-246)

```

      1250      1260      1270      1280      1290      1300
frame6 SSTVTGFSETRPITRKLIAIR.SRWVATAAPEALPRPLPSGRPRITREVSPSSQSTPLSL
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|136      MRGRIVPLALSTPLPTASPPSPAAPAAQAQPPSSSL
      10      20      30

      1310      1320      1330      1340      1350      1360
frame6 RPFATAERRSLSLTRSSSKPFITVAPSANAAATARIGYS--SIIEGARSSSGTVTPFKVEE
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|136 PPLAAIPSRASGTTTRSTT---TFSPASRSPTPSRASVTLRSPLTASSASRSRTPSRASV
      40      50      60      70      80      90

      1370      1380      1390      1400      1410      1420
frame6 RTRKSATSSPPARRGLITMSAPISCRVKMRPVRVGFIKTLRMVTSEPGTСКАATNGKAV
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|136 PTRSTPTASPASRSRTPSQASVPPRSKPTASPA-----SRSPTPSRASVTLRSP
      100      110      120      130      140

      1430      1440      1450      1460      1470      1480
frame6 EDGSPGTSIT.PVSLPRPSRRMV--RTPSASVSTVRLAPKP.SIFSVSRVMIGSITVVI
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|136 LTASPASRS-----PTPSRASVTLRSPPTASSASRSAT-P-----SRASVSTRSEIT
      150      160      170      180

      1490      1500      1510      1520      1530
frame6 PGVLRPASRTADFT.AEAIGTL.VIGVGVGPRKVTGR--RSPSSDFTCTPISRSGFNTR
      . : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|136 AS---PASRSRTPSRA-----SVTGRSQRTPS-----PASRSPTPSR
      190      200      210      220

      1540      1550      1560      1570      1580      1590
frame6 PIGRLVREASPMKV--LSKS.TPTRPIVKRQPVAAALPKSRGASGARSPPLTPSMIQVSP
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|136 AS---VTLRSPLTASSASRSATPSRALVNTRSTP
      230      240      250
```

>>gi|77967877|gb|ABB09257.1| DNA translocase FtsK [Burkh (1673 aa)  
initn: 51 initl: 51 opt: 187 Z-score: 200.8 bits: 50.7 E(): 0.04  
Smith-Waterman score: 210; 25.249% identity (50.497% similar) in 503 aa overlap (1187-1672:564-993)

```

      1160      1170      1180      1190      1200      1210
frame6 TVFPMP.YSNVGVFWCVRNETDALNLTNVVETDD--TNESSS.RILRS.DP-VPSSNS.A
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|779 TGSEPAHAFAQPNTAAVHDIAAGNPPASFDTDDQSTAAQPAADRTASARVPQGIADA
      540      550      560      570      580      590

      1220      1230      1240      1250      1260      1270
frame6 HQAALVSESSILAPRPAIKSMNSGKLVAIIVASSTVTGSETRPITRKLIAIR.SRWVAT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|779 HADA--KPAGNIAPFAA-----LPASPIADTTSRTSAETRPF----AAASAPAAAN
      600      610      620      630

      1280      1290      1300      1310      1320      1330
frame6 AAPEALPRP-LPS-GRPRTTRE-VSPSSQSTPLSLRPFATAERRSLSLTRSSSKPFITVA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|779 AAPAAAAWTWPSTAQPAATTNPPVTAELPKSPVADQPFAAAQATASAS---ATSPSVGTV
      640      650      660      670      680      690
```

```

      1340      1350      1360      1370      1380
frame6 PSANAAATARI-GYSSIIIEGARSSSGTVTPFKVEERTRKSATSSPPARRGLITMSAPISC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|779 HGSTAQAPASVSGIAPVTVPASSDGVIG-----TASSSLAQPAASISAVASV-APAS-
      700      710      720      730      740
```

```

      1390      1400      1410      1420      1430      1440
frame6 RVKMRPVRVGFIKTLRMVTSEPGTСКАATNGKAVEDGSPGTSIT.PVSLPRPSRRMVRT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|779 -----PI--GVTGTASAPLSQPAAS--ASNAASVAPASP-IGVTGTASAP-----LSQP
      750      760      770      780
```

```

      1450      1460      1470      1480      1490      1500
frame6 SASVSTVRLAPKP.SIFSVSRVMIGSITVVIPGVLRPASRTADFT.AEAIGTL.VIGVG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|779 AASASTVA-AVVPASPIGV----TGTASA--P-LLQPATSASN---AASVAPASPIGV-
      790      800      810      820      830
```

```

      1510      1520      1530      1540      1550      1560
frame6 VFGPRKVTGRRSPSSDFTCTPISRSGFNTRPIGRLVREASPMKVLKS.TP-TRPIVKRQ
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|779 -----TGTAS-----APLSQPAASASTVAAVVS-ASPIGVGTGTASAPLSQPAASAS
      840      850      860      870      880
```

```

      1570      1580      1590      1600      1610      1620
frame6 PVAAL-PKSR-GASGARSPPLTPSM-IQVSPSLRILAPIACIALPVLITSSPSKRPVIR
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|779 NAASVIPASPIGVGTGTASTSLSQPATASASTVASVASAAPVG-VAGTIYGSFPRASVPASS
      890      900      910      920      930      940
```

```

      1630      1640      1650      1660      1670
frame6 VSLARPPNMKDLWEIDLSPGIRTVPERPEDLRVA-----GRLLAP.SMHAVETEDMT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|779 IAPAVIPSTVGTASSTSGSTAPAPAPTATPAIAAVAPANTVSGSITAPSSAPAAATLQANP
      950      960      970      980      990      1000
```

```

      1680      1690      1700      1710      1720      1730
frame6 LGRSSEPINVIPLFFNPHDDEIG.ARGCCCVFRETDKGDLR.LDFERLEIRDGFCTPLQ
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|779 ATSSGTTFGTGTSSAVAPQPAASLPVPPAASDTATTPIATAAATPAQPIANSVPTPTG
      1010      1020      1030      1040      1050      1060
```

>>gi|111307761|gb|AAI21193.1| Mki67 protein [Xenopus lae (2080 aa)  
initn: 308 initl: 73 opt: 188 Z-score: 200.6 bits: 51.0 E(): 0.041  
Smith-Waterman score: 243; 22.847% identity (49.912% similar) in 569 aa overlap (1079-1633:162-691)

```

      1050      1060      1070      1080      1090      1100
frame6 F.V.KYICLVNKHNGTTTSLVLLILTSVIRGH-DLTHFYSKQNVCSISLIISNSTHNK.E
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|111 SKLTRKSEGNVSRTHTRLRLVSPNRRSSRGKKDLSPFGELYEMLKSKVDLKQEHAKTP
      140      150      160      170      180      190
```

```

      1110      1120      1130      1140      1150      1160
frame6 KKQIILI.E.TKGP--YHSLTLLHPFFHSSIAKTE.KTQ.ITSTNGTRKTVFPMP.YS
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|111 AKQNMENGDNVTPTTTTRRSLGTVKATP-ELSIGEFESQELHVESATL-SEKTASALK--S
      200      210      220      230      240
```

```

      1170      1180      1190      1200      1210      1220
frame6 NSVGFWCVRNETDALNLTNVVETDDTNESSS.RILRS.DPVPSSNS.AHQAAALVSESSIL
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```

gi|111 PSRGRQSPKQEQRLSIKRDSPAK-SNDKLTPRRSTSEKPAQERTPAKSGGSVIPRSRNL
250      260      270      280      290      300

1230      1240      1250      1260      1270      1280
frame6 APRPAI-KSMNSGKLVAIIVASSTVTGFSETRPITRKLIAIR.SRWVATAAPEALPRPLP
.: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|111 SPRSASP KDLHKNVSKVPQQKRRSSELELPEPTPKR----KRVSGFGHLSPELPDKRLP
310      320      330      340      350      360

1290      1300      1310      1320      1330      1340
frame6 SGRPRTTREVSPSSQSTPLSLRPFATAERRSLSLTRSSSKPFI TVAPSANAAATARIGYS
.: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|111 PNSP-LKRGAAPARRSLSLAT-PLAVIRKSFGGFKQSAIKE--AFEPGTDLALFKR-SPA
370      380      390      400      410

1350      1360      1370      1380      1390      1400
frame6 SIIEGARSSGTVTPFKVEERTRKSATSSP---PARRGLITSMSAPISCRVKMRPVRVGF
.: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|111 KATPAKRSPAKGSPAKATPAKRSPAKGSPAKLTPAKRSPAKASPAKRSP-PAKASPAKRSP
420      430      440      450      460      470

1410      1420      1430      1440      1450
frame6 IKTLRMVITSEPGTSKAATNGKAVEDGSPGTSIT.PVSLPRPSRRM-VRTPSASVSTVRLA
.: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|111 AKA---SPAKRSPAKASPAKRSPAKVSPAKRSPAKVS---PAKRSPAKVSPAKRSPAKVS
480      490      500      510      520      530

1460      1470      1480      1490      1500      1510
frame6 PKP.SIFSVWSRVMIGSITVVIPGVLRPASRTADFT.AEAIGTL.VIGVGVFGRKVT-G
.: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|111 PAKRSPAKV-----SPAKRSPAKVSPAKRSPA-KVSPAKRSPAKVSPAKRSPAKVSPA
540      550      560      570      580

1520      1530      1540      1550      1560      1570
frame6 RRSPPSDFCTCTPISRSGFNTRPIGRLVREASPMKVLKSKS.TPTRPIVKRQPVAAALPKSRG
.: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|111 KRSPAK---ASPAKRSPAKASPAKRSPAKASPAK---RSPAKASP-AKRSPAKASPAKRSP
590      600      610      620      630

1580      1590      1600      1610      1620      1630
frame6 ASGA----RSPPLPTPSMIQVSPSLRILAPIACIALPVLITSSPSKRPVIRVSPLARPPN
.: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|111 PAKATPAKSPAKGSPA--KVTPSKRSPAK-ASPSKRSPAKASPSKRSPAKASPAKRSPA
640      650      660      670      680      690

1640      1650      1660      1670      1680      1690
frame6 MKDLWEIDLSPGIRTVPERPEDLRAVAGRLLAP.SMHAVETEDMTLGRSSEPINVIPLF

gi|111 KGSPAKVTPAKRSPAKGFSAKVTPAKRSPAKASPAKRSPAKVTPAKRSPAKGSPAKVTPA
700      710      720      730      740      750

>>gi|134065504|emb|CAM43271.1| proteophosphoglycan ppg4 (5384 aa)
  initn: 299 init1: 93 opt: 193 Z-score: 200.5 bits: 52.3 E(): 0.041
Smith-Waterman score: 206; 22.000% identity (48.800% similar) in 500 aa overlap
(1132-1622:653-1129)

1110      1120      1130      1140      1150      1160
frame6 THNK.EKKQIILI.E.TKGPYHSLTLLHPFPFHSSIAKTE.KTQ.ITSTTNGTRKTVPFM
.: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|134 PSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSS
630      640      650      660      670      680

```

```

1170      1180      1190      1200      1210
frame6 P.YSNSVGFWCVRNETDALNLNTNVETDDTNE--SSS.RILRS.DPVPSSNS.AHQAALV
.: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|134 PSSSSSAP---SSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSS
690      700      710      720      730

1220      1230      1240      1250      1260      1270
frame6 SE--SSILAPRPAIKSMNSGKLVAIIVASSTVTGFSETRPITRKLIAIR.SRWVATAAPE
.: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|134 SAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSS
740      750      760      770      780      790

1280      1290      1300      1310      1320      1330
frame6 ALPRPLPSGRPRTTREVSPSSQST-PLSLRPFATAERRSLSLTRSSSKPFI TV-APSAN-
.: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|134 SAPSS-SSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSS
800      810      820      830      840      850

1340      1350      1360      1370      1380      1390
frame6 --AAATARIGYSSIIEGARSSGTVTPFKVEERTRKSATSSPPARRGLITSMSAPISCRVK
.: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|134 SSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSS
860      870      880      890      900      910

1400      1410      1420      1430      1440      1450
frame6 MRPVVRVGFIKTLRMVITSEPGTSKAATNGKAVEDGSPGTSIT.PVSLPRPSRRMVRTPSAS
.: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|134 SAPSSSSSAPS-SSSSSAPSSSSSAPSSSS--SSAPSSSSSAPSSSSSAPSSSSSAPSSSS
920      930      940      950      960      970

1460      1470      1480      1490      1500      1510
frame6 VSTVRLAPKP.SIFSVWSRVMIGSITVVIPGVLRPASRTADFT.AEAIGTL.VIGVGVFGR
.: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|134 SSS---APSSSSSAPSSSSSAPSSSSSAPS--SSSSSAPSSSSSAPSS--SSSSA
980      990      1000      1010      1020

1520      1530      1540      1550      1560      1570
frame6 PRKVTGRRSPSSDFCTCTPISRSGFNTRPIGRLVREASPMKVLKSKS.TPTRPIVKRQPVAA
.: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|134 PS--SSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSS
1030      1040      1050      1060      1070      1080

1580      1590      1600      1610      1620      1630
frame6 LPKSRGASGARSPPLPTPSMIQVSPSLRILAPIACIALPVLITSSPSKRPVIRVSPLARP
.: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|134 SAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSS
1090      1100      1110      1120      1130

1640      1650      1660      1670      1680      1690
frame6 PNMKDLWEIDLSPGIRTVPERPEDLRAVAGRLLAP.SMHAVETEDMTLGRSSEPINVIP

gi|134 SSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSSSAPSSSS
1140      1150      1160      1170      1180      1190

>>gi|70905641|gb|AAZ14280.1| proteophosphoglycan ppg4 [L (7194 aa)
  initn: 438 init1: 117 opt: 194 Z-score: 199.9 bits: 52.6 E(): 0.045
Smith-Waterman score: 223; 22.490% identity (48.996% similar) in 498 aa overlap
(1132-1619:608-1092)

1110      1120      1130      1140      1150      1160
frame6 THNK.EKKQIILI.E.TKGPYHSLTLLHPFPFHSSIAKTE.KTQ.ITSTTNGTRKTVPFM
.: : : : . . . . .

```

```

gi|709 PSSSSSAPSASSSSAPSSSSSAPSASSSSAPSSSSSAPSASSSSAPSSSSSAPSASSSSA
580      590      600      610      620      630

      1170      1180      1190      1200      1210      1220
frame6 P.YNSVGFWCVRNETDALNLTNNVETDDTNESSS.RILRS.DVPSSNS.AHQAAALVSE
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|709 PSSSSSAPSASSSSAPSSSSSAPSASSSSAPSSSSSAPSASSSSAPSSSSSAPSASSSSA
640      650      660      670      680      690

      1230      1240      1250      1260      1270
frame6 --SSILAPRPAIKSMNSGKLVAIIVASSTVTGFSETRPIITRKLIAIR.SRWVATAAPEAL
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|709 PSSSSSAPSASSSSAPSSSSSAPSASSSSAPSSSSSAPSASSSSAPSSSSSAPSASSSSA
700      710      720      730      740      750

      1280      1290      1300      1310      1320      1330
frame6 PRPLPSGRPRTTREVSPSSQSTPLSLRPFATAERRSLSLTRSSSKPFITVAPSANAAATA
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|709 PSSSSSAPSASSSSAPSSSS--SSAPSASSSSAPSSSSSAPSASSSSAPSSSSSAP
760      770      780      790      800      810

      1340      1350      1360      1370      1380      1390
frame6 RIGYSSIIEGARSSGTVTPFKVEERTRKSATSSPPARRGL---ITSMSAPISCRVKMRPV
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|709 SASSSS---APSSSSSAPSSSSSSPSSSSSAPSASSSSPSTSSSSAP--SASSSSAPS
820      830      840      850      860

      1400      1410      1420      1430      1440      1450
frame6 RVGFIKTLRMVISEPGTSKAATNGKAVEDGSPGTSIT.PVSLPRPSRRMVRTPSASVSTV
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|709 SSSSSAPSASSSSAPSSSSSAPSASSSSAPSSSSSAPSASSSSAPSSSSSAPSASSSS-
870      880      890      900      910      920

      1460      1470      1480      1490      1500      1510
frame6 RLAPKP.S-IFSVWSRVMIGSITVVIPGVLRLPASRTADFT.AEAIGTL.VIGVGVFGPRK
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|709 --APSSSSSAPSASSSSAPSSSSSAPSA--SSSSAPSSSSSAPSASSSSAPSSSSSAP-S
930      940      950      960      970      980

      1520      1530      1540      1550      1560      1570
frame6 VTGRRSPSSDFTCTPISRSGFNTRPIGRLVREASPMKVLKS.TPTRPIVKRQPVAA--L
... : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|709 ASSSSAPSSSSSAPSASSSSAPSSSSSAPSASSSSAPSSSSSAPSASSSSAPSSSSA
990      1000      1010      1020      1030      1040

      1580      1590      1600      1610      1620      1630
frame6 PKSRGASGARSPPLPTPSMIQVS-PSLRILAPIAC-IALPVLITSSPSKRPVIRVSPLAR
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|709 PSASSSSAPSSSSSAPSASSSSAPSSSSSAPSASSSSAPSSSSSAPSASSSSAPSSSS
1050      1060      1070      1080      1090      1100

      1640      1650      1660      1670      1680      1690
frame6 PPNMKDLWEIDLSPGIRTVPERPEDLRAVAGRL LAP.SMHAVETEDMTLGRSSEPINVI

gi|709 SALSASSSSAPSSSSSAPSASSSSAPSSSSSALSASSSSAPSSSSSAPSASSSSAPSSS
1110      1120      1130      1140      1150      1160

>>gi|190346874|gb|EDK39060.2| hypothetical protein PGUG_ (961 aa)
  initn: 78 init1: 48 opt: 183 Z-score: 199.6 bits: 49.7 E(): 0.046
Smith-Waterman score: 199; 21.382% identity (48.520% similar) in 608 aa overlap
(1066-1627:8-589)

```

```

      1040      1050      1060      1070      1080      1090
frame6 TRGHTCIQ.HLQLF.V.KYICCLVKNHNGTTTSVLLLILTSVIRGHDLTHFYSKQNVCS
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190      MLPKLLWSSSALMLIATVLA AEHDKSN-NGKAKVESS
      10      20      30

      1100      1110      1120      1130      1140      1150
frame6 LIISNSTHNK.EKKQIILI.E.-TKGPYHSLTLLHFPFPHSSIAKTE.---KTQ.ITSTT
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 ---TSSDNNPLDFLSRLLATNGFPASAPTSTIDENTGF-FNNPFAQAQTAPSSSQPTTVA
40      50      60      70      80      90

      1160      1170      1180      1190      1200
frame6 NGRKTRTVFMP.P.YNSVGFWCVRNETDALNL--TNVETDDTNESSS.RILRS.D---PV
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 ETKTPSTPTTSTAENSPTSDPTTEPTTELTPAPTATTSTDAEKS VANTPSTEDITSPT
100      110      120      130      140      150

      1210      1220      1230      1240      1250      1260
frame6 PSSNS.AHQAAALVSESSILAPRPAIKSMNSGKLVAIIVASSTVTGFSETRPIITRKLIA--
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 KTADTQSPITTEPTTAATSSAPAAAEAKSTEKASTAMLADGPPTTSATTPTTSPSTSSGT
160      170      180      190      200      210

      1270      1280      1290      1300      1310      1320
frame6 ---IR.SRWVATAAPEALPRPLPSGRPRTTREVSPSSQSTPLSLRPFATAERRSLSLTRS
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 TTPTTSSSTASETTPTTTTSPSTTSTSPPTTS-TNPTSAAVDVIVSGSSSVVAPSTATAED
220      230      240      250      260      270

      1330      1340      1350      1360      1370
frame6 SSKPFITVAPSANAAATARIGYSSIIEGARS-----SGTVTPFKVEERTRKSATSSPP
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 STSTPTTTSPSKSATPTVEPSSSQVVGSEPTTQPKVESSTVAPSDSVIATPTDSTTSP
280      290      300      310      320      330

      1380      1390      1400      1410      1420      1430
frame6 ARRGLITSM SAPISCRVKMRPV RVGFIKTLRMVITSEPGT--SKAATNGKAVED--GSPGT
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 SGTESTTSIQQP-----STAELATSAPSSVASETASASTAPADFIATSGA
340      350      360      370

      1440      1450      1460      1470      1480
frame6 SIT.PVS---LPRPSRRMVRTPSASVSTVRLAPKP.SIFSVWSRVMIGSITVVIPGVLRLP
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 SSTLPQSGSTTPVPKATSSPTTVSEST SALAPVPTS-----QQSSVTSAAQVIPASSVP
380      390      400      410      420      430

      1490      1500      1510      1520      1530
frame6 ASRTADFT.AEAIGTL.VIGVGV----FGPRKVTGRRSPSSDFTCTPISRSGFNTR----
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 QTSSTLPSTSLVSPTESTLPVSSSTPPLPPTAISSSVSQSSVPPTIPVSPSTVATTSSTS
440      450      460      470      480      490

      1540      1550      1560      1570      1580      1590
frame6 --PIGRLVREASPMKVLKS.-TPTRPIVKRQPVAA LP-KSRGASGARSPPLPTPSMIQV
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|190 SVPIVSTPVTSTTPVVSPTVESSTPPTTVSQAVVSTPVQSESSSTPVVPTVTSESAVAP
500      510      520      530      540      550

      1600      1610      1620      1630      1640
frame6 SPSLRIL-----APIACIALPVLITS-SPSKRPVIRVSPLARPPNMKDLWEIDLSPGIRT

```



```

      : . . : . . : . : . :
gi|190 SAETSVIPVSSAETSVIPVPSAVTSETPEETTTSAVSAAVTVSPETSASPTTTAIQPSV
      560      570      580      590      600      610

      1650      1660      1670      1680      1690      1700
frame6 VPERPEDLRAVAGRLAP.SMHAVETEDMTLRGRSSEPINVIPLFFNPHDDEIG.ARGCC

gi|190 VQESSAPATSNVVPSTQTTPTVTSTIPVASNVEQTTSFEIQTPSTTEVPLQETPSSHQRPA
      620      630      640      650      660      670

>>gi|38260630|gb|AAR15447.1| pollen coat oleosin-glycine (1356 aa)
      initn: 40 init1: 40 opt: 183 Z-score: 197.5 bits: 49.8 E(): 0.061
      Smith-Waterman score: 231; 23.326% identity (50.808% similar) in 433 aa overlap
      (1272-1687:950-1340)

      1250      1260      1270      1280      1290      1300
frame6 IIVASSTVTGFSETRPITRKLIAR.SRWVATAAPEALPRPLPSGRPRTTREVSPSSQST
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|382 FPGGGASPDKLVPTKTSNKPPTTKPGDKSAKSSSKPAAPKSTKPTSKPSTKSASKPSPKPT
      920      930      940      950      960      970

      1310      1320      1330      1340      1350
frame6 PL-SLRPF--ATAERRSLSLTRSSSKPFITVAPSANAAATARIGYSSIIIE-GARSSGTVT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|382 SKPSTKPTFKPTAKPTKSTAKPSTKPTTKPITKPTITKPTSKPTTKTVAKPSSKEAGKPT
      980      990      1000      1010      1020      1030

      1360      1370      1380      1390      1400      1410
frame6 PFKVEERTKSA---TSSPPAR---RGLITMSAPISCRVKMRPVRVGFIKTLRMVTSEP
      . . . . . : : : : : . . . . . : : : : : . . . . . : : : : :
gi|382 TKPTTKSTSKPSKPTTKPPTKPTKPIAKPTSKPIA-KPSTKPIKSKPETKPIKSTSKP
      1040      1050      1060      1070      1080      1090

      1420      1430      1440      1450      1460      1470
frame6 GTSKAATNGKAVEDGSPGTSIT.PVSLPRPSRRMVRTPSASVSTVRLAPKP.SIFSVWSR
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|382 KI-KTVAKPSSKEEGKPTTKPTTKSS-SKPTAKPVSKPAAK-ST----PKPTS-----K
      1100      1110      1120      1130      1140

      1480      1490      1500      1510      1520      1530
frame6 VMIGSITVVPVGLRPASRTADFT.AEAIGTL.VIGVGVFGRKVTGRRSPSSDFTCTPI
      . . . . . : : : : : . . . . . : : : : : . . . . . : : : : :
gi|382 PIAKPVTK--P-TAKPTSNPTKKPAAKPPSKLTV-----KPVT---KPTEKPTSKPI
      1150      1160      1170      1180      1190

      1540      1550      1560      1570      1580
frame6 SRSGFN--TRPIGRLVREASPM-KVLSKS.TP--TRPIVKRQPVAAALPKSRGASGARSPP
      . . . . . : : : : : . . . . . : : : : : . . . . . : : : : :
gi|382 AKPAVKPTSKPIAKPV--AKPVSKPIAKSTVKQTSKPVGK--PVTK-PTAKPAGKLASKP
      1200      1210      1220      1230      1240

      1590      1600      1610      1620      1630      1640
frame6 LPTPSMIQVSPSLRILAPIACIALPVLITSSPSKRPVIRV--SPLARPPNMXDLWEIDL
      : . . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|382 TAKPV---AKPTAKPVAKPA--AKPV---AKPAAKPTSKLISKPVAKPAS-KPASKPTTK
      1250      1260      1270      1280      1290

      1650      1660      1670      1680      1690      1700
frame6 PGIRTVPERPEDLRAVAGRLAP.SMHAVETEDMTLRGRSSEPINVIPLFFNPHDDEIG.
      : . : . . : : : : : : : : : : : : : : : : : : : : : :
gi|382 PTSKPKPAAKSTSKPIAKPAVKPASKPAAKPTSKPITKPTSKPKPATKSAAPKPAVKPKK
      1300      1310      1320      1330      1340      1350

```

```

      1710      1720      1730      1740      1750      1760
frame6 ARGCCCVFRETDKGLR.LDFERLEIRDGFCTPLQILLTCAMVDLWSTNLQKEQSNIES

>>gi|270236229|emb|CBI25498.1| unnamed protein product [ (545 aa)
      initn: 69 init1: 69 opt: 178 Z-score: 197.4 bits: 48.4 E(): 0.062
      Smith-Waterman score: 192; 25.180% identity (48.201% similar) in 417 aa overlap
      (1205-1591:113-505)

      1180      1190      1200      1210      1220      1230
frame6 NETDALNLTNVETDDTNESSS.RILRS.DVPVSSNS.AHQAALVSESSILAPRP-AIKS
      . . : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|270 PPRKTVADDFLNSDNKNDYDWLLTPPGTPLFPSLEMSQKTVMSRIGTPKARPTALKS
      90      100      110      120      130      140

      1240      1250      1260      1270      1280
frame6 MNSGKLVAIIVASSTVTGFSETRP-ITRKLIAR.SRWVATAAPEALPRPLPSGRP----
      . : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|270 RLANPQPEPAARSSLVSKHPASSPGLNSSSAGIR--RPSSSGPGSGRP-ATPTGRPTLTA
      150      160      170      180      190

      1290      1300      1310      1320      1330      1340
frame6 --RTTREVSPSSQSTPLSLRPFATAERRSLSLTRSSSKPFITVAPSANAAATARIGYSSI
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|270 TTRPSRPATPTSRATLTSTKPTVPA-RSSTPVTRSTAR---SSTPTSRPSIPASKEVSR
      200      210      220      230      240      250

      1350      1360      1370      1380      1390      1400
frame6 IEGARSSGTVTPFKVEERTKRSATSSPPARRGLITMSAPISCRVKMRPVRVGFIKTLRM
      . : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|270 ATPTRPST--PSSV-----SNLSAPPIKSTSSVTKSGPATSRNPV-PSR-GSSPT---
      260      270      280      290      300

      1410      1420      1430      1440      1450      1460
frame6 VTSEPGTСКАATNGKAVEDGSPGTSIT.PVSLPRPSRRMVRTPSASVSTVRLAP--KP.-
      : : : : : : : : : : . . : : : : : : : : : : : : : : : :
gi|270 VKSRPW-KPLEMPGYSL-DAPPNLRTSAPERPVASRGRPSAPIARSSSVDAAPNGRPRR
      310      320      330      340      350      360

      1470      1480      1490      1500      1510
frame6 -SIFSVWSRVMIGSITVV---IPGVLRPASRTADFT.AEAIGTL.V---IGVGVFGRK
      . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|270 QSCSPARGRPNGIIHTSGSSVPAVSRGLSKVSDNVSPVLIGTKMVERVINMRKLVPPKQ
      370      380      390      400      410      420

      1520      1530      1540      1550      1560      1570
frame6 TGRRSF---SSDFTCTPISRSGFNTRPIGRLVREASPMKVLSKS.TP--TRPIVKRQPV
      . : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|270 EDKHSRGNLSSGLASSPDS-SGFG-RTLSKKSOLDMAIRHMDIRRSIPGNLRPLMTNIPA
      430      440      450      460      470

      1580      1590      1600      1610      1620
frame6 AALPKSR-GASGARS-----PPLPTPSMIQVSPSLRILAPIACIALPVLITSSPSKRPVI
      . . . : : : : : : : : : : : : : : : : : : : : : : :
gi|270 SSMYSVRSGPTRSRTVSVDSPLATSSNASSESVVNNNALCVDGGEIEDDIGSERGGRPS
      480      490      500      510      520      530

>>gi|190624521|gb|EDV40045.1| GF24155 [Drosophila ananas (1337 aa)
      initn: 117 init1: 52 opt: 182 Z-score: 196.5 bits: 49.6 E(): 0.07
      Smith-Waterman score: 184; 23.958% identity (48.698% similar) in 384 aa overlap
      (1271-1622:889-1254)

```

```
>>gi|190617179|gb|EDV32703.1| GF22010 [Drosophila ananas (847 aa)
  initn: 40 initl: 40 opt: 179 Z-score: 195.8 bits: 48.8 E(): 0.075
Smith-Waterman score: 179; 19.672% identity (52.459% similar) in 366 aa overlap
(1112-1464:443-796)
```

```
>>gi|32452989|gb|AAP82647.1| Hypothetical protein K06A9. (825 aa)
  initn: 88 initl: 57 opt: 177 Z-score: 193.7 bits: 48.4 E(): 0.099
Smith-Waterman score: 242; 23.909% identity (50.285% similar) in 527 aa overlap
(1135-1628:12-515)
```

```

      1270      1280      1290      1300      1310      1320
frame6 R.SRWVATAAPEALPRPLPSGRPRTTREVSPSSQSTPLSLRPFATAERR-SLSLTRSSSK
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|324 SSSTF-ASSTPIA-----SSSSPGSTVTIVAPGSSSTYGSSTPSASSSSSGTMTNSGSGT
      160      170      180      190      200      210

      1330      1340      1350      1360      1370      1380
frame6 PFITVAPSANAAATARIGYSSIIIEGARSSG-TVTPFKVEERTRKSAT--SSPPARRGLITS
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|324 STVTIVAP----VSSSTFGSSTPIASSSSSGSTVTVVSGSSSTYGSSTPSASSSSAGTAST
      220      230      240      250      260

      1390      1400      1410      1420      1430
frame6 MSAPISCRVKMRPVRVGFII--KTLRMVTSEPGTСКАATN--GKAVEDGSPGTSIT-.PVS
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|324 ISGSTGSTATIVPGSSSSVGSSTQASPSSTPGTMTSTVSGTGTSTVTV--VPGSSTSPAPSS
      270      280      290      300      310      320

      1440      1450      1460      1470      1480      1490
frame6 LPRPSRRMVRTPSASVSTVRLAPKP.SIFSVWSRVMIGSITVVIIPGVLRPASRTADFT.A
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|324 SPNPSS----SPASTGSTITISGSSSIIVSTVSGSTVSGTGTSTL--ASSTATPGSS
      330      340      350      360      370      380

      1500      1510      1520      1530      1540      1550
frame6 EAIGIL.VIGVGVFGRKVTGRRSPSSDFTCTPI--SRSGFNTRPIGRLVREASPMKVLK
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|324 STVPSSSSPQPSSQSPAPNTGSTITPSQTSSQSPSPMNPSSSTPTGSSQSTITPEG--ST
      390      400      410      420      430

      1560      1570      1580      1590      1600      1610
frame6 S.TPTRPIVKRQPVAAALPKSR-----GASGARSPLPTPSMIQVSPSLRILAPIACIALP
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|324 ASSPTGSTGSTFSVATEVTSQSTVPSGSSSLGTQSTNSSPSPSSSLSPSTSGMSTLTSESP
      440      450      460      470      480      490

      1620      1630      1640      1650      1660      1670
frame6 VLITSSPSKRPVIRVSPPLARPPNMKDLWEIDLSPGIRTVPERPEDLRVAGRLLAP.SMH
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|324 SSTQSSGAQSTLTTTPSPNPSSQSTSSLESSTSGATTSSGSAGTTMTSPSQSSSVGSSQGST
      500      510      520      530      540      550

>>gi|212001255|gb|EEB06915.1| predicted protein [Schizos (1678 aa)
  initn: 103 initl: 69 opt: 180 Z-score: 192.8 bits: 49.2 E(): 0.11
Smith-Waterman score: 246; 21.456% identity (48.851% similar) in 522 aa overlap
(1207-1697:817-1309)

      1180      1190      1200      1210      1220      1230
frame6 TDALNLTNVVETDDTNESSS.RILRS.DPVPSSNS.AHQAAALVSESSILAPRPAIKSMNS
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|212 VTSIPVSSAPVSSVYANGSVPATSVPVTSAPASSVYANSSVPVSSVPVSSVPATSVPV
      790      800      810      820      830      840

      1240      1250      1260      1270      1280      1290
frame6 GKLVAIIVASSVTGFSRTPITRKLIA--IR.SRWVATAAPEALPRPLPS----GRPRT
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|212 SAPVSSAYANSSV-----PVSSVLVSSVPATSVPVSSAPVNSSIPVSSVPATSVPAT
      850      860      870      880      890

      1300      1310      1320      1330      1340
frame6 TREVS--PSSQSTPLSLRPFATAERRSLSLTRS---SSKPFITVAPSANAAATARIGYSS
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : :

```

```

gi|212 SAPASSAPVNSSVPATSVPVTSAPASSAPVNSSIPVSSVP-VTSIPVSSAPASSVYANSS
      900      910      920      930      940      950

      1350      1360      1370      1380      1390
frame6 IIEGARSSGTVTVPFKVEERTRKSATSSPPARRGLITS----MSAPI-----SCRVKMR
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|212 VPVSSVSVSSVPFTSAPASSAPVNSSVPVSSVPVSSSIPATSPATSIYVNSSVPVSSA
      960      970      980      990      1000      1010

      1400      1410      1420      1430      1440      1450
frame6 PVRVGFIKTLRMVTSEPGTСКАATNGKAVEDGSPGTSIT.PV-SLPRPSRRMVRT-PSAS
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|212 PVSSVPATSVPVSSAPVSSAYANSSVPV--SSVLVSSSIPVTSAPASSAPMNSSIPVSS
      1020      1030      1040      1050      1060      1070

      1460      1470      1480      1490      1500
frame6 V--STVRLAPKP.SIFSVWSRVMIGSITVV--IPGVLRPASRTADFT.AEAIGTL.VIGV
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|212 VPATSVPVSSAPASSVYANSSVPVSSVPVSSVPATSVPVSS--SAPVNSSVPVSSVPVTSI
      1080      1090      1100      1110      1120      1130

      1510      1520      1530      1540      1550      1560
frame6 VGFGRKVTGRRSPSSDFTCTPISRSGFNTRPIGRL-VREASPMKVLKSKS.TPTRPIVKR
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|212 PISSAPASSVYANSSVPVSSVPATSVPVSSAPVSSIPVSSSAPVSSAPVSSAPSNSV--
      1140      1150      1160      1170      1180      1190

      1570      1580      1590      1600      1610      1620
frame6 QPVAAALPKSRGASGARSPLPTPSMIQVSPSLRILAPIACIAL--PVLITSSPSKRPVIR
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|212 -PVSSVPAT--SAPVSSAPVSSSVIVSSAPATS--APVSSVYANSSVLVSSAPATSVPV
      1200      1210      1220      1230      1240

      1630      1640      1650      1660      1670      1680
frame6 VSLARPPNMKDLWEIDLSPGIRTVPERPEDLRVAGRLLAP.SMHAVETEDMTLRGRSS
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|212 SAPVSSIP-----VSSAPVSSAPVSSAPSNSVPVSSVPATSVPVSSAPVSS--SS
      1250      1260      1270      1280      1290

      1690      1700      1710      1720      1730      1740
frame6 EPINVIPLFFNPHDDEIG.ARGCCCVFRETDKGDLR.LDFERLEIRDGFCTPLQILLTCA
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|212 VPVSSVPATSVPVSSSVIVSSAPATSVPVSSVYANSSVPVSSVPVSSASSSSILVPDMPG
      1300      1310      1320      1330      1340      1350

>>gi|123962409|gb|ABM77165.1| Translation initiation fac (1124 aa)
  initn: 55 initl: 55 opt: 177 Z-score: 191.8 bits: 48.5 E(): 0.13
Smith-Waterman score: 190; 24.505% identity (46.535% similar) in 404 aa overlap
(1218-1591:52-400)

      1190      1200      1210      1220      1230      1240
frame6 TDDTNESSS.RILRS.DPVPSSNS.AHQAAALVSESSILAPRPAIKSMNSGKLVAIIVASS
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|123 DVLHAAEKLISIAKSHSSSISDDEAKRIRGLLRQGSAAANSAPPSKS-EPGKTI-LSVKKA
      30      40      50      60      70

      1250      1260      1270      1280      1290
frame6 TVTGFSETRPIITRKLIAIR.SRWVATAAP-----EALPRPLPSGRPR--TTR
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : :
gi|123 APTAIKDVAAPMRKATSSSEISQVKPSAPANPTPTSPERLSRESVAHPAPPTRPVNPTTT
      80      90      100      110      120      130

```

```

      1300      1310      1320      1330      1340
frame6 EVS--PSSQSTPLSL---RPFATAERRSLSLTRSSSKPFITVAPSANAAATARIGYSSII
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|123 PTSSPPKTAARPVNAPISRP-ATPSRPSAPTTPRSANKPSSPVPPSTGSK-DPRAGQTSTS
      140      150      160      170      180      190

      1350      1360      1370      1380      1390      1400
frame6 EGAR---SSGTVTPFKVEERTKRSAT---SSPPARGLITSMASAPISCRVKMRPVVGVFIK
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|123 SKATTVSGGGRPP-KIISRPQSPAAPGRSAPPAKPSIPSDRKAP-----KPELVGRPK
      200      210      220      230      240

      1410      1420      1430      1440      1450      1460
frame6 TLRMTVSEPGTSKAATNGKAVEDGSPGTSIT.PVSLPRPSRRMVRTPSASVSTVRLAPKP
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|123 PKRPVVAPP---SRPEPEGQRDPDKKRPGIS-----PRP---IGGPNQARANTPQRPGAP
      250      260      270      280      290

      1470      1480      1490      1500      1510      1520
frame6 .SIFSVWSRVMIGSITVVIPGVLRLPAS-RTADFT.AEAIGTL.VIGGVGFGRKVTGRRS
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|123 -----IRQGKTRPGQPRASG-----NTLELVGKPIRRDRSDAGSAG
      300      310      320      330

      1530      1540      1550      1560      1570
frame6 PSSDFCTPISRSGFENTRP--IGRLVREASPMKVLKSKS.TPTRPIVK--RQPVAALPKSR
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|123 RDSN-----NRPGAPTRPGMPAGMRKPVAPGELMQLQKPTGRPGTPPPRRPDGTSVGTR
      340      350      360      370      380

      1580      1590      1600      1610      1620      1630
frame6 GASGARSPPLTPSMIQVSPSLRILAPIACIALPVLITSSPSKRPVIRVSPLARPPNMKD
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|123 GGSEGATPPVERPASPTAPKRPGHRPAQAPAAGAPRRRGRPDWDSDAKLEALRNKSPQKQ
      390      400      410      420      430      440

>>gi|211583314|emb|CAP91322.1| Pc13g02530 [Penicillium c (643 aa)
  initn: 47 initl: 47 opt: 174 Z-score: 191.8 bits: 47.6 E(): 0.13
Smith-Waterman score: 196; 23.716% identity (46.944% similar) in 409 aa overlap
(1221-1622:33-392)

      1200      1210      1220      1230      1240
frame6 TNESSS.RILRS.DVPVSSNS.AHQAAALVSESSILAPRPAIKSMNSGKLVAI--VASSTV
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|211 RFQFFLFLLLPWWASAGPIFRWLSPSLESPEPTSLPLLPDIERPEGGITAITPVAAPTQ
      10      20      30      40      50      60

      1250      1260      1270      1280      1290      1300
frame6 TGFSET-RPITRKLIAR.SRWVATAAPEALPRPLPSGRPRTTREVSPSSQSTPLSLRPF
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|211 LPFTPTIKPTKLTIFPKPTTTPEPTATPEPTTTTPEPTATPEPT--TTPEPTTTP--EPT
      70      80      90      100      110

      1310      1320      1330      1340      1350      1360
frame6 ATAERRSLSLTRSSSKPFITVAPSANAAATARIGYSSIIEGARSSGTVTPFKVEERTKRS
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|211 ATPEPTATPEPTTTTPEPTTTTPEPTTTTPEPTAT-----PEPTTTP--EPTTTPE
      120      130      140      150      160

      1370      1380      1390      1400      1410      1420
frame6 ATSSP-PARRGLITSMASAPISCRVKMRPVVGVFIKTLRMVTSEPGTSKAATNGKAVEDGS
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:

```

```

gi|211 PTATPEPTTTPEPTATPEPTTTTQEQATTKKSS---TTISHASTTTRLSISSTTPGVLHSS
      170      180      190      200      210      220

      1430      1440      1450      1460      1470      1480
frame6 PGTSIT.PVSLPRPSRRM--VRTPSASVSTVRLAPKP.SIFSVWSRVMIGSITVVIPGVL
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|211 PITS-SLPVHVPGSTSSLSGLVIPSSAQSPSDSLPGGAT--SI-----KATTATPSVS
      230      240      250      260      270

      1490      1500      1510      1520      1530      1540
frame6 RPASRTADFT.AEAIGTL.VIGGVGFGRKVTGRRSPSSDFTCTPISRSGFNTRPIGRLV
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|211 TSTSTQPTTTPVPTVPTSSSIAVGAIAP--LGGASSSSSSTSTPVIPTSNPIQPAKATQ
      280      290      300      310      320

      1550      1560      1570      1580      1590      1600
frame6 REASPMKVLKSKS.TPTRPIVK-RQPVAALPKSRGASGARSPPLP-TPSMIQVSPSLRILA
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|211 AQSTSLQQISTSSS--KPVVTVTEPVIVYV-TKGA-----PKPFSTSVIHHTPSVNN-S
      330      340      350      360      370

      1610      1620      1630      1640      1650      1660
frame6 PIACIALPVLITSSPSKRPVIRVSPLARPPNMKDLWEIDLSPGIRTVPERPEDLRAVAGR
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|211 PVSS-AKP---TVKPSVEPTTTSLMEVTIHSSTATSVVRVTVPSSSSSSSSSSSSSSSSSS
      380      390      400      410      420      430

>>gi|135120742|gb|EBF62441.1| hypothetical protein GOS_9 (307 aa)
  initn: 108 initl: 71 opt: 170 Z-score: 191.8 bits: 46.6 E(): 0.13
Smith-Waterman score: 199; 25.275% identity (50.183% similar) in 273 aa overlap
(1200-1460:1-264)

      1170      1180      1190      1200      1210      1220
frame6 FWCVRNETDALNLTNVVETDDTNESSS.RILRS.DVPVSSNS.AHQAAALVSESSILAPR-
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|135                                     MSSATTCTSRKPLGRNSRRSKADVSSRT
      10      20      30

      1230      1240      1250      1260      1270      1280
frame6 PAIKSMNSGKLVAIIVASSTVTG--FSETRPITRKLIAR.SRWVATAAPEALPRPLPSG
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|135 PPAKSSAVARPRRSCTSPSTTPGSAVSASKPKS----SARSPAPSCSASPSTPRSTPSS
      40      50      60      70      80

      1290      1300      1310      1320      1330
frame6 RPRTTREVSP--SSQSTPLSLRPFAT----AERRSLSLTRSSSKPFITVAPSANAAATA
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|135 RSATSTCTSSATSSPSTTPRRPPSPSPATQAHRTSRWTTTSSSTPLPCVRPTSSAEVTW
      90      100      110      120      130      140

      1340      1350      1360      1370      1380      1390
frame6 RIGYSSIIIEGARSSGTVTPFKVEERTKRSATSSP-PARRGLITSMAS-PISCRVKMRPVR
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|135 PWPFTDRLASSTSIPTPSNTHVTQTSGRSVTSLPYPPRPVPPSPASKRPFSSSTISSPGAR
      150      160      170      180      190      200

      1400      1410      1420      1430      1440      1450
frame6 VGFIKTLRMVTSEPGTSKAATNGKAVEDGSPGTSIT.PVSLPRPSRRMVRTPSASVSTVR
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|135 IRVPVR-RIMATPPARSSPDTAASSPN-SPTATRTSP---PSPSTRRRNAGTCTCSSAT
      210      220      230      240      250      260

```

```

1460      1470      1480      1490      1500      1510
frame6 LAPKP.SIFSVMVSRMIGSITVVIPIGVLRPASRTADFT.AEAIGTL.VIGGVGFPRKVT
:
gi|135 CFPPTGRACSAASPEPVRRPARRPHRPRATSPPLRDVASRGTVAP
270      280      290      300

>>gi|186909484|gb|ACC94277.1| mucin [Bos taurus] (549 aa)
initn: 68 initl: 68 opt: 173 Z-score: 191.7 bits: 47.4 E(): 0.13
Smith-Waterman score: 205; 25.094% identity (44.944% similar) in 534 aa overlap
(1188-1685:21-509)

1160      1170      1180      1190      1200      1210
frame6 VFPMP.YSNSVGFWCVRNETDALNLTNVVETDDTNESSS.RILRS.DPVPS-SNS.AHQA
:
gi|186 INPRRTTPVSTTQSSPTSSPTKETSWSTTTLLTASSPAPSPAASPGHDG
10      20      30      40      50

1220      1230      1240      1250      1260      1270
frame6 ALVSESSILAPRPAIKSMNSGKLVAIIVASSVTVGFSERTRPIRKLIAIR.SRWVATAAP
:
gi|186 ASTPTSSP-APSPAASPGHNG-----TSSPTGSPAPSP-----AASPGHDGAST-P
60      70      80      90

1280      1290      1300      1310      1320      1330
frame6 EALPRPLPSGRPRTTREVSPSSQSTPLSLRPFATAERRSLSLTRSSSKPFIIVAPSANAA
:
gi|186 TSSPAPSPAASPGHDGASTPTSSPAP---SPAASPGHDGASTPTSSPAPSPAASPGHN--
100     110     120     130     140

1340      1350      1360      1370      1380      1390
frame6 ATARIGYSSIIEGARSSSGTVTFPKVEERTRKSATSSPPARRGLITSMASIPSCRVKMRPV
:
gi|186 -----GTSSPTGSPAPSPAASPGHDGASTPTSSPAPSPAASPGHNGTSSPTGSPAPSPA
150     160     170     180     190     200

1400      1410      1420      1430      1440
frame6 RVGFIKTLRMVTSEPGTСКААТ---NGKAVEDGSPGTS-----IT.PVSLPRPSRR
:
gi|186 SPGH-DGASTPTSSPAPSPAASPGHNGTSSPTGSPAPSPAASPGHDGASTPTSSPAPSP-
210     220     230     240     250     260

1450      1460      1470      1480      1490
frame6 MVRTPS---ASVSTVRLAPKP.-SIFSVMVSRMIGSIT---VVIPIG---VLRPASRTADF
:
gi|186 -AASPGHDGASTPTSSPAPSPAASPGHNGTSSLTGSPAPSPAASPGHDGASTPTSSPAPS
270     280     290     300     310     320

1500      1510      1520      1530      1540
frame6 T.AEAIGTL.VIGGVGFPRKVT-----GRRSPSSDFTCTPISRSGFN--TRPIGRLVRE
:
gi|186 PAASPGHDGASTPTSSPAPSPAASPGHDGASTPTSSPAPSPAASPGHNGTSSPTGSPAPS
330     340     350     360     370     380

1550      1560      1570      1580      1590      1600
frame6 --ASPMKVLSKS.TPTRPIVKRQPVAAALPKSRGASGARSPPLTPSMIQ---VSPSLRI
:
gi|186 PAASPGH--DGASTPTSSPAP-SP-AASPGHNGTSSPTGSPAPSPAASPGHDGASTPTSS
390     400     410     420     430

1610      1620      1630      1640      1650      1660
frame6 LAPAIACIALPVLITSSPSKRPVIRVSPPLARPPNMKDLWEIDLSPGIRTVPERPEDLRAVA
:

```

```

gi|186 PAPSPAASPGHNGTSSPTGSPA--PSPTASPGH-----DSAPSLTSSPA-PSPT-ASP
440      450      460      470      480

1670      1680      1690      1700      1710      1720
frame6 GRLLAP.SMHAVETEDMTLGRSSEPINVIPLFFNPHDDEIG.ARGCCCVFRETDKGLR
:
gi|186 GQHGAS-SPTGSDTSSMTTRSMSSSMVTSAHKGTSSRATMTVPVSKGTPSSVPSSETAPTA
490      500      510      520      530      540

>>gi|145017502|gb|EDK01865.1| predicted protein [Magnapo (577 aa)
initn: 99 initl: 60 opt: 173 Z-score: 191.4 bits: 47.4 E(): 0.13
Smith-Waterman score: 173; 19.467% identity (50.400% similar) in 375 aa overlap
(1268-1632:219-573)

1240      1250      1260      1270      1280      1290
frame6 KLVAIIVASSTVTGFSERTRPIRKLIAIR.SRWVATAAPEALPRPLPSGRPRTTREVSPS
:
gi|145 NNPA AAAE GPGAFIMVKRDHSRPAASSTCTSTLATVVKPTSTMAPLAS----SSTCTSTT
190      200      210      220      230      240

1300      1310      1320      1330      1340      1350
frame6 SQSTPLSLRPFATAERRSLSLTRSSSKPFIIVAPSANAAATARIGYSSIIEGARSSGTVT
:
gi|145 TARPTTSLRPVSSSTTCTTTLATSKVPETPKAPSSSVPTTTKASVTPVA-SSTSTCTTT
250      260      270      280      290      300

1360      1370      1380      1390      1400      1410
frame6 PFKVEERTKRSATSSPPARRGLITSMASIPSCRVKMRPVRVGFIKTLRMVTSEPGTСКАА
:
gi|145 ATPVAETSK-AVPTTTARTTRYVTLN-PKTSTTPAAPSSSSSVTPVAESSTAPASTTTR
310      320      330      340      350      360

1420      1430      1440      1450      1460      1470
frame6 TNGKAVEDGSPGTSIT.PVSLPRPSRRMVRTPSASVSTVRLAPKP.SIFSVMVSRMIGSI
:
gi|145 ATRYVTITGKSSSTA AVPTTSSVEPAYETSTPTQPQTТА-APQSSSVSTIYATQKV-TV
370      380      390      400      410

1480      1490      1500      1510      1520      1530
frame6 TVVIPGVLRPASRTADFT.AEAIGTL.VI-GGVGFPRKVTGRRS-----PSSDFTCTPI
:
gi|145 TSCAGGQKQCTSQAQTFVITHIVPIYTTVCPTSALHPVKPTHQAAYPSVKPSSAASVKPV
420      430      440      450      460      470

1540      1550      1560      1570      1580
frame6 SRSGFNTRPIGRLVREASPMKVLSKS.TPTRPIVKR--QPV-AALPKSRGASGARSPPLP
:
gi|145 PTASYNQTSAAPTA-ASSSYKVVPKLNLTLTSSAKKVHAPTYAPTQKVGNGTTSVAVPTGP
480      490      500      510      520      530

1590      1600      1610      1620      1630      1640
frame6 TPSMIQ-VSPSLRILAPIACIALPVLITSSPSKRPVIRVSPPLARPPNMKDLWEIDLSPGI
:
gi|145 TAGATTGVAPSK---SP-----MPTM--GSGADRTVVGLAALAVPLLLFL
540      550      560      570

1650      1660      1670      1680      1690      1700
frame6 RTVPERPEDLRAVAGRLLAP.SMHAVETEDMTLGRSSEPINVIPLFFNPHDDEIG.ARG

>>gi|198150947|gb|EDY73984.1| GA28499 [Drosophila pseudo (1221 aa)
initn: 112 initl: 65 opt: 177 Z-score: 191.3 bits: 48.5 E(): 0.13

```

Smith-Waterman score: 195; 22.067% identity (53.352% similar) in 358 aa overlap (1118-1451:830-1163)

```

      1090      1100      1110      1120      1130      1140
frame6 SKQNVCI SLIISNSTHNK.EKKQIILI.E.TKGPYHSLTLLHPFFHSSIAKTE.KTQ.I
      . . . . .
gi|198 NNL SYAQNPFLLSLARPIV PARAPVKPSVATKSPQH---VTRPYVSSDSFRNEVIK---Y
      800      810      820      830      840      850

      1150      1160      1170      1180      1190      1200
frame6 TSTTNGTRKTVFPMP.YNSVGFWCVRNETDALNLTNVVETD-----DTNESSS.RILR
      . . . . .
gi|198 TATSSGTTTTTTKHPIITSTTT-----KTTTKSTAKATVRPTSFPPVWVQTSAKTTTTRTKP
      860      870      880      890      900

      1210      1220      1230      1240      1250      1260
frame6 S.DPVPSNS.AHQALVSESSILAPRPAIKSMNSGKLVAIIVASSTVTGFSERPIPTRK
      . . . . .
gi|198 TTRPPSTTKATHYTTVTKPTS---RPTLTITQRS---AATTRPTTIS---TTRPITRY
      910      920      930      940      950

      1270      1280      1290      1300      1310
frame6 LIAIR-.SRWVATAAPEALPRPLPSGRPRTT-REVS-----PSSQSTPLSLRPFATAE
      . . . . .
gi|198 IAVTKPTTRRTSTTRPATLRLTVPTTRPSTTARSVTHYTTATKPLPRATPTT-RPTTTAR
      960      970      980      990      1000      1010

      1320      1330      1340      1350      1360      1370
frame6 RRSLSLTRSSSKPFTIVAPSANAAATARIGYSSIEGARSSSGTVTPFKVEERTRK SATSS
      . . . . .
gi|198 PATTTRTVSPSTRPTASTKTTNHPAPTTRPTITT--RRTLTRSTTRYTATARPTQSTTR
      1020      1030      1040      1050      1060      1070

      1380      1390      1400      1410      1420
frame6 PPARRGLITMSA-PISCRV--KMRPVVRVGFIKTLR---MVTSEPGTСКАATNGKAVED
      . . . . .
gi|198 QTTRYSAITTRATRPSTSKTLPSTTPRPLYFTSTTTTPQKSSTTADPFSSHPPFFQKFAKFE
      1080      1090      1100      1110      1120      1130

      1430      1440      1450      1460      1470      1480
frame6 GSPGT--SIT.PVSLPRPSRRMVRTPSASVSTVRLAPKP.SIFSVWSRVMIGSITVVIPG
      . . . . .
gi|198 KTPSPVLQIKTPTTTAAPS RQNSTSSPYAKYQENKVKVAAKT VFSYDFQGNRTTTRK
      1140      1150      1160      1170      1180      1190
```

>>gi|49651333|emb|CAG78272.1| YALIOF15653p [Yarrowia lip (1051 aa)  
initn: 127 initl: 64 opt: 176 Z-score: 191.1 bits: 48.2 E(): 0.14  
Smith-Waterman score: 221; 22.750% identity (51.250% similar) in 400 aa overlap (1243-1622:351-732)

```

      1220      1230      1240      1250      1260      1270
frame6 AHQALVSESSILAPRPAIKSMNSGKLVAIIVASSTVTGFSERPITRK LIAIR.SRWVA
      . . . . .
gi|496 DACPPEGGSASVTEIPSTSTEPITTPEEQITPSSTEPSTTEDPTTTPESTIPSTSEEP
      330      340      350      360      370      380

      1280      1290      1300      1310      1320      1330
frame6 TAAPEALPRPLPSGRPRTTREVSPSSQSTPLSLRPFATAERRSLSLTRSSSKPFTIVAPS
      . . . . .
gi|496 TSTPESTIE--PSSTPETTEPSSTESSSTPVITEPSSTEP----SSTEPSTPESTTAPS
      390      400      410      420      430
```

```

      1340      1350      1360      1370      1380
frame6 ANAAATARIGYSSIEGARSSSGTV---TPFKVEERTRKA---TSSPPARRGLITSM-
      . . . . .
gi|496 TTDDVTTTDDVTTKPESTTEKPTTDLSSSPVTTEPSSTESSPIPETTEPSTLTLDVTTTNE
      440      450      460      470      480      490

      1390      1400      1410      1420      1430
frame6 -SAPISCRVKMRPVVRVGFIKTLRMVITSEPGTСКАATNGKAVEDGSPGT---SIT.PVSLP
      . . . . .
gi|496 STTPKSTTEKPTTDPSSSPETTEPSSTEPSSSTPETTEPSSTEPSSPETTEPSSTEPSSSTP
      500      510      520      530      540      550

      1440      1450      1460      1470      1480      1490
frame6 RPSRRMVRT-PSASVSTV---RLAPKP.SIFSVWSRVMIGSITVVIPGVLPRPASRTADFT
      . . . . .
gi|496 EPTTEQLTTEPSTTETTTTPEKQTTTPSSTPQTDDPCATTSVITSTPPVGPPIITTIV--T
      560      570      580      590      600      610

      1500      1510      1520      1530      1540      1550
frame6 .AEAIGTL.VIGVGVFGPRKVTGRRSPSSDFTCTPISRSGFNTRPIGRLVREASPMKVL
      . . . . .
gi|496 TNCSSSVPEPSTIPITSETFTTETPSTSEPIITDVTVSTD-PCESVTTIV--STPPSGEP
      620      630      640      650      660

      1560      1570      1580      1590      1600      1610
frame6 KS.TPTRPIVKRQPVAAAL---PKSRGASGARSPPLPTPSMIQVSPSLRILAPIACIALPV
      . . . . .
gi|496 STYTIITDICTSKPVTIPMTSVPTQLPQ--TSDPISEPSTLSTTPG----PSSTVCVIT
      670      680      690      700      710      720

      1620      1630      1640      1650      1660      1670
frame6 LITSSPSKRPVIRVSPRLRPPNMKDLWEIDLSPGIRTVPERPEDLRAVAGRLLAP.SMHA
      . . . . .
gi|496 LTITDTPDGNPSTTTTVDYCPHGPTTYTTTIVTCEEECTIPITLTHVPPNPSTTVIT
      730      740      750      760      770      780
```

2415 residues in 1 query sequences  
4761287459 residues in 17815538 library sequences  
Scomplib [34t26]  
start: Thu Feb 4 00:02:48 2010 done: Thu Feb 4 00:27:03 2010  
Total Scan time: 1286.070 Total Display time: 1.930

Function used was FASTA [version 3.4t26 July 7, 2006]

Database checksum values:	
Thu Feb 4 00:27:04 GMT 2010 /genedata/1/db/AD_2010	a184245745a6ed8c6ecde45b26637bba
Thu Feb 4 00:27:04 GMT 2010 /genedata/1/db/TOX_2010	17c3a19148dfb0163e270cf41e2aa437
Thu Feb 4 00:27:52 GMT 2010 /genedata/1/db/PRT_2010	e657d3127c1aad11f9f7df8dcc5e448c