

Appendix 3

Bioinformatics Analysis Of Deduced Amino Acid Sequences of Open Reading Frames Contained in the Transgene Insert of Herbicide-Tolerant Soybean BPS-CV127-9

**BIOINFORMATICS ANALYSIS OF DEDUCED AMINO ACID SEQUENCES OF
OPEN READING FRAMES CONTAINED IN THE TRANSGENE INSERT OF
HERBICIDE-TOLERANT SOYBEAN BPS-CV127-9**

AUTHOR:



STUDY COMPLETED ON: DECEMBER 17, 2008

PERFORMING LABORATORY:

**BASF PLANT SCIENCE, L. L. C.
26 DAVIS DRIVE
RESEARCH TRIANGLE PARK, NORTH CAROLINA 27709
U. S. A.**

REPORT NUMBER: BPS-015-08

OECD UNIQUE IDENTIFIER: BPS-CV127-9

SUBMITTED BY:

**BASF PLANT SCIENCE, L. L. C.
26 DAVIS DRIVE
RESEARCH TRIANGLE PARK, NORTH CAROLINA 27709
U. S. A.**



The Chemical Company
Plant Science L. L. C.

BASF Reg. Doc. No. 2008/7019371

STATEMENT OF NO DATA CONFIDENTIALITY CLAIMS

No claim of confidentiality is made for any information contained in this study on the basis of its falling within the scope of FIFRA §10(d) (1) (A), (B), or (C).

Company: BASF Plant Science, L. L. C.

Company Agent: [REDACTED] Date: 17 Dec 2008

Title: Regulatory Affairs Manager

Signature: [REDACTED]

These data are the property of BASF Plant Science, L. L. C. and, as such, are considered to be confidential for all purposes other than compliance with FIFRA §10. Submission of these data in compliance with FIFRA does not constitute a waiver of any right to confidentiality that may exist under any other statute in any other country.

STATEMENT OF COMPLIANCE

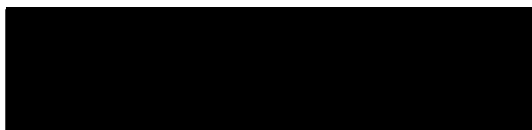
This study was not conducted in compliance with the requirements of 40 CFR Part 160.

The data generated by BASF Plant Science in support of product safety comply with generally accepted scientific procedures. Record keeping is consistent with procedures used throughout the research community. This report accurately presents the raw data developed during the study.

PRINCIPAL INVESTIGATOR:

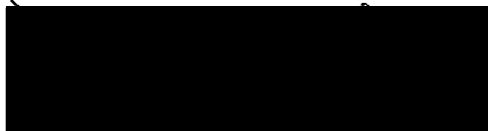
Senior Scientist, Regulatory Science
BASF Plant Science, L. L. C.

December 17, 2008
Date

SPONSOR REPRESENTATIVE:

Senior Manager, Regulatory Science
BASF Plant Science, L. L. C.

17 Dec 08
Date

SUBMITTED BY:

Regulatory Affairs Manager
BASF Plant Science, L. L. C.
26 Davis Drive
Research Triangle Park, North Carolina 27709
U. S. A.

17 Dec 2008
Date

TABLE OF CONTENTS

STATEMENT OF NO DATA CONFIDENTIALITY CLAIMS.....	2
STATEMENT OF COMPLIANCE.....	3
TABLE OF CONTENTS.....	4
LIST OF TABLES	4
LIST OF FIGURES	4
ABBREVIATIONS AND DEFINITIONS.....	5
SUMMARY	6
INTRODUCTION	6
MATERIALS AND METHODS.....	7
RESULTS AND DISCUSSION	10
CONCLUSIONS.....	10
RECORDS RETENTION.....	12
STUDY PERSONNEL	12
REFERENCES	12
APPENDIX 1. 80-AMINO ACID ALLERGEN HOMOLOGY SEARCH RESULTS..	18
APPENDIX 2. TOXIN HOMOLOGY AND GENERAL PROTEIN HOMOLOGY SEARCH RESULTS	141

LIST OF TABLES

Table 1. Summary of Open Reading Frames Present within the Transgene Insert of BPS-CV127-9.....	14
--	----

LIST OF FIGURES

Figure 1. Deduced Amino Acid Sequences of Open Reading Frames from the Transgene Insert of BPS-CV127-9.....	15
Figure 2. Graphical Representation of Open Reading Frames within the Transgene Insert of BPS-CV127-9.....	17

ABBREVIATIONS AND DEFINITIONS

AHAS	acetohydroxyacid synthase enzyme
<i>ahasl</i>	imidazolinone-tolerant alleles of <i>Arabidopsis thaliana</i> acetohydroxyacid synthase large subunit gene; includes the <i>ahasl</i> (S653N) allele found in plasmid pAC321 and the <i>ahasl</i> (R272K, S653N) allele found in BPS-CV127-9
AHASL	wild-type <i>Arabidopsis thaliana</i> acetohydroxyacid synthase large subunit gene; also referred to in the literature as <i>CSRI</i>
ahasl	imidazolinone-tolerant <i>Arabidopsis thaliana</i> acetohydroxyacid synthase large subunit protein bearing two mutations (R272K, S653N)
BLASTP	protein-protein Basic Local Alignment Search Tool
BLOSUM	<u>b</u> locks <u>s</u> ubstitution <u>m</u> atrix
FARRP	Food Allergy Research and Resource Program
E value	expect value
GI number	GenInfo identifier, or GI number, is a sequence identification number assigned by the NCBI
IgE	immunoglobulin E
NCBI	National Center for Biotechnology Information (United States of America)
ORF	open reading frame
R272K	arginine residue at position 272 of <i>Arabidopsis thaliana</i> acetohydroxyacid synthase large subunit replaced with lysine
RT-PCR	reverse-transcription polymerase chain reaction
S653N	serine residue at position 653 of <i>Arabidopsis thaliana</i> acetohydroxyacid synthase large subunit replaced with asparagine
<i>SEC61γ</i>	wild-type <i>Arabidopsis thaliana</i> <i>SEC61γ</i> -subunit gene
SEC61γ	wild-type <i>Arabidopsis thaliana</i> SEC61γ-subunit protein

BIOINFORMATICS ANALYSIS OF DEDUCED AMINO ACID SEQUENCES OF OPEN READING FRAMES CONTAINED IN THE TRANSGENE INSERT OF HERBICIDE-TOLERANT SOYBEAN BPS-CV127-9

SUMMARY

Soybean [*Glycine max* (L.) Merr.] plants have been developed that are tolerant to the imidazolinone class of agricultural herbicides. The herbicide-tolerant soybean plants, referred to as BPS-CV127-9, are derived from a single transformation event and were produced by the introduction of an imidazolinone-tolerant acetohydroxyacid synthase large subunit (*ahasl*) gene from *Arabidopsis thaliana* (L.) Heynh. into the soybean plant genome via biolistics using the PvuII fragment of transformation plasmid pAC321. The PvuII fragment of pAC321 includes what, at the time of transformation, was defined as the wild-type *Arabidopsis AHASL* promoter, the mutant *Arabidopsis hasl* (S653N) coding sequence, and the wild-type *Arabidopsis AHASL* terminator. The mutant *Arabidopsis hasl* (S653N) allele, which confers tolerance to imidazolinone herbicides, has also been referred to in the literature as *csr1-2* (Sathasivan *et al.*, 1990; Jander *et al.*, 2003). In addition to the S653N mutation, a second mutation was discovered in the *Arabidopsis hasl* coding sequence integrated in the CV127 soybean genome. This second mutation, in which arginine at position 272 of the *ahasl* protein is replaced by lysine, does not impact the enzymatic function of the acetohydroxyacid synthase (AHAS) enzyme or its herbicide-tolerance properties. Also included in the region originally annotated as the *AHASL* promoter was the majority of an *Arabidopsis* putative *SEC61 γ* -subunit gene (*Arabidopsis* Genome Initiative locus code At3g48570), including the entire *Arabidopsis SEC61 γ* coding sequence.

The transgene insert of BPS-CV127-9 was analyzed for open reading frames (ORFs) of 29 or more amino acids. In addition to the ORF corresponding to the *ahasl* coding sequence, 27 other unique ORFs were identified within the BPS-CV127-9 transgene insert sequence. In this study, the deduced amino acid sequences of these 27 ORFs were subjected to bioinformatics analyses as part of a weight-of-evidence assessment of the potential allergenicity and toxicity of any proteins potentially encoded within the BPS-CV127-9 transgene insert sequence. The bioinformatics analyses of the ORFs found within the BPS-CV127-9 transgene insert sequence do not provide any indication of a potential allergenicity or toxicity concern.

INTRODUCTION

Soybean [*Glycine max* (L.) Merr.] plants have been developed that are tolerant to the imidazolinone class of agricultural herbicides. The herbicide-tolerant soybean plants, referred to as BPS-CV127-9, are derived from a single transformation event. BPS-CV127-9 was produced by the introduction of an imidazolinone-tolerant acetohydroxyacid synthase large subunit (*ahasl*) gene from *A. thaliana* (L.) Heynh. into the soybean plant genome via biolistics using the PvuII fragment of transformation vector

pAC321. The PvuII fragment includes what, at the time of transformation, was defined as the wild-type Arabidopsis *AHASL* promoter, the herbicide-tolerant Arabidopsis *ahasl* coding sequence, and the wild-type Arabidopsis *AHASL* terminator. In addition to the S653N mutation, a second mutation was discovered in the *ahasl* coding sequence integrated in the BPS-CV127-9 soybean genome. This mutation, in which arginine at position 272 of the *ahasl* protein is replaced by lysine, does not impact the enzymatic function of the AHAS enzyme or its herbicide-tolerance properties (Stevenson Paulik, 2007). Also included in the region originally annotated as the *AHASL* promoter was the majority of the Arabidopsis putative *SEC61γ*-subunit gene (Arabidopsis Genome Initiative locus code At3g48570), including the entire Arabidopsis *SEC61γ* coding sequence.

Acetohydroxyacid synthase (AHAS) is found ubiquitously in the plant kingdom. The AHAS enzyme catalyzes the first step in the biosynthesis of branched-chain amino acids in plants. In conventional plants, inhibition of the AHAS enzyme by imidazolinone herbicides leads to a deficiency in branched-chain amino acids and other compounds derived from this pathway that are needed for plant survival. The *ahasl* gene from Arabidopsis confers tolerance to imidazolinone herbicides by encoding an AHAS catalytic subunit with altered herbicide-binding properties, while retaining its normal biosynthetic function in the soybean plant.

SEC61 is a heterotrimeric complex which forms a protein-conducting channel involved in transport across the endoplasmic reticulum membrane in eukaryotes. Two of the three subunits of SEC61, the α and γ subunits, are found in all organisms (Osborne *et al.*, 2005). Although the *SEC61γ*-subunit transcript is detectable in BPS-CV127-9 by reverse-transcription polymerase chain reaction (RT-PCR) analysis (Shen, 2007), the SEC61 γ -subunit protein product is not detectable in BPS-CV127-9 soybean leaves or grain (Jiang, 2007).

The transgene insert of BPS-CV127-9 contains the coding sequences of two proteins from *A. thaliana*: *ahasl* (R272K, S653N) and SEC61 γ . The purpose of this current study was to identify additional open reading frames (ORFs) within the BPS-CV127-9 transgene insert and conduct a bioinformatics analysis of these ORFs as part of a weight-of-evidence assessment of the potential allergenicity and toxicity of any protein potentially encoded within the BPS-CV127-9 transgene insert sequence.

MATERIALS AND METHODS

Databases. The Food Allergy Research and Resource Program (FARRP) Allergen Protein Database (version 8.00; allergenonline.com) containing 1313 entries was utilized for all bioinformatics assessments of potential allergenicity. These 1313 entries are comprised of known or putative food, respiratory, venom/salivary, or contact allergenic proteins. Glutenins and gliadins from *Triticum aestivum* L. (bread wheat) were included in the database if there was published evidence of immunoglobulin E (IgE) from sera of allergic individuals binding to the glutenin or gliadin in question. Some of these same

glutenins and gliadins are involved in Celiac disease, although that disease is a non-IgE-mediated disorder. All allergen database entries have been vetted by a panel of seven academic allergy experts based on published evidence of allergenicity.

The GenBank non-redundant peptide sequence database was utilized for all bioinformatics assessments of potential toxicity in addition to the general assessment of protein homology. The GenBank non-redundant peptide sequence data (a total of 7,155,275 sequences) was downloaded October 26, 2008 from the National Center for Biotechnology Information (NCBI) website. This database is comprised of all non-redundant GenBank coding sequence translations, protein sequences from NCBI's Reference Sequence Project, sequences derived from the Protein Data Bank, the last major release of the SWISS-PROT protein sequence database, the Protein Information Resource (PIR) Protein Sequence Database, and the Protein Research Foundation (PRF) Protein Sequence Database.

Identification of open reading frames. The entire transgene insert of BPS-CV127-9 was analyzed for open reading frames (ORFs) using Vector NTI, version 10.3.0 (Invitrogen; Carlsbad, CA). An ORF is a sequence of nucleotides which has the potential to encode a protein; whether an ORF is transcribed and ultimately translated is dependent upon having the requisite regulatory elements in the proper context. For the purposes of this study, an ORF was defined as a contiguous nucleic acid sequence which begins at a start codon (i.e. ATG) and continues until the first termination codon (i.e. TAA, TAG, or TGA) is encountered. Any ORF from any of the six potential reading frames with a deduced amino acid sequence of 29 amino acids or greater which is located either entirely or partly within the transgene insert was selected for further analysis. The minimum ORF size of 29 amino acids is appropriate because the minimum size of a polypeptide able to elicit the allergenic cascade is estimated to be 29 amino acids long (Bannon *et al.*, 2002). The deduced amino acid sequences of the 27 unique ORFs identified within the transgene insert of BPS-CV127-9 are shown in Figure 1 and the location of these ORFs within the transgene insert is represented graphically in Figure 2. The transgene insert of BPS-CV127-9 contains the coding sequences of two proteins from *A. thaliana*: ahasI (R272K, S653N) and SEC61γ; these two proteins were evaluated in BASF Plant Science report BPS-014-08 (McKean, 2008) and are not included among the 27 ORFs discussed in this current report.

Database search for eight-amino acid regions of identity to known or putative allergens. Each ORF deduced amino acid sequence was submitted to an analysis using a custom comparison (word-match) program which provides an exhaustive search of all possible eight-amino acid subsegments of the query sequence against all possible eight-amino acid segments in proteins in the FARRP Allergen Protein Database. Regions of at least eight consecutive amino acids which are identical between an ORF deduced amino acid sequence and a known allergen will be identified by this search.

This eight-amino acid search was originally suggested based on the concept that eight or more amino acids is a representative minimal size for an IgE-binding epitope (Metcalf *et al.*, 1996). Bannon and Ogawa (2006) compiled a list of characterized linear IgE-binding

epitopes from major allergens and, although one epitope from a wheat ω -5 gliadin was only four amino acids long, the majority of characterized epitopes were indeed eight amino acids or longer. However, this search does not detect conformational epitopes which are formed when non-linear amino acids are brought together by the higher-order folding of the protein. Moreover, the utility of such an eight-amino acid analysis has been questioned due to the high rate of false positives identified by this search (Silvanovich *et al.*, 2006; Hileman *et al.*, 2002).

Database search with eighty-amino acid segments of query sequence. Open reading frames with deduced amino acid sequences greater than 80 amino acids were subdivided into all possible overlapping 80-amino acid segments. Open reading frames with deduced amino acid sequences of 80 amino acids or less were analyzed as a single intact sequence. Each of these 80-amino acid segments or, in the case of ORFs with deduced amino acid sequences less than 80 amino acids, the entire protein sequence was compared *in silico* to all proteins in the FARRP Allergen Protein Database via a protein-protein FASTA (version 34.26.5; April 26, 2007) analysis. The default parameters of the FASTA program were used, including the default substitution scoring matrix of BLOSUM 50, with one exception: the threshold score for optimization was set to 20.

Since the total predicted protein sequence was analyzed incrementally in 80-amino acid segments, the query length for each of the analyses was 80 amino acids for ORFs with a deduced amino acid sequence length of greater than 80 amino acids. Thus, the percent identity for a given alignment was determined by dividing the number of identical amino acids within the alignment by 80. In the case of ORFs with a deduced amino acid sequence length of less than 80 amino acids, the number of identical amino acids was still divided by 80 to determine the percent identity over 80 amino acids. In instances where gaps were inserted into the query sequence to achieve the optimal alignment, percent identity was calculated by dividing the number of identical amino acid residues in the alignment by the alignment length of overlap if the overlap length was greater than 80. A query sequence which showed >35% identity to a known or putative allergen would be identified as potentially requiring additional studies, on a case-by-case basis, to determine the likelihood of the predicted protein being allergenic (Codex, 2003).

Database search for homology to known protein toxin sequences as well as general homology to known proteins. Each ORF deduced amino acid sequence was submitted to a protein-protein Basic Local Alignment Search Tool (BLASTP), (NCBI version 2.2.18) analysis. All parameters were left at the default setting for this BLASTP analysis with the exception of the expect (E) value cut-off which was set to one. Thus, either the top 250 unique protein alignments (the default value) or all alignments with an E value less than one, whichever is fewer, were analyzed. This BLASTP analysis will identify regions of local similarity (i.e. similarity over a portion of the protein sequence rather than the entire protein sequence) between the query sequence and proteins in the downloaded October 26, 2008 GenBank non-redundant peptide sequence database.

Descriptions of the 250 proteins from the GenBank non-redundant database which show the most significant local homology to the ORF CV127.5729_d2 deduced amino acid

sequence were manually compared to known toxins which act on humans as listed in the United States Code of Federal Regulations ([40 CFR Part 725.421](#)). Only 59 proteins in the GenBank non-redundant database had an E value less than one when the ORF CV127.1392_d3 deduced amino acid sequence was used as a query sequence. Descriptions of these 59 proteins were manually compared to known toxins which act on humans as listed in the United States Code of Federal Regulations ([40 CFR Part 725.421](#)).

RESULTS AND DISCUSSION

Database search for eight-amino acid regions of identity. The ORF deduced amino acid sequences did not share a sequence of eight or more consecutive identical amino acids with any potential allergens found in the FARRP Allergen Protein Database. Because the eight-amino acid word-match search did not identify any common identical sequences of eight or more amino acids, there are no search results to report.

Database search with eighty-amino acid segments of query sequence. The ORF deduced amino acid sequences did not show 35% or greater identity over 80 amino acids to any potential allergens found in the FARRP Allergen Protein Database. For ORFs with deduced amino acid sequences greater than 80, the allergen homology search results for the 80-amino acid segment which showed the highest homology to a known or potential allergen is shown in Appendix 1. The allergen homology search results for ORFs with deduced amino acid sequences less than 80 are also shown in Appendix 1.

Database search for homology to known protein toxin sequences as well as general homology to known proteins. The ORF deduced amino acid sequences did not show significant homology to proteins that are toxic to humans as listed in the United States Code of Federal Regulations ([40 CFR Part 725.421](#)). The toxin homology search results are shown in Appendix 2. Twenty-five ORFs did not show homology to any known protein sequence using the search parameters defined in this report. The summary of the ORF deduced amino acid sequence protein homology search is shown in Table 1.

CONCLUSIONS

The transgene insert of BPS-CV127-9 contains the coding sequences of two proteins from *A. thaliana*: ahas1 (R272K, S653N) and SEC61 γ . In this current report, the transgene insert of BPS-CV127-9 was analyzed for additional open reading frames (ORFs) of 29 or more amino acids. In this study, 27 unique ORF were identified within the BPS-CV127-9 transgene insert and the deduced amino acid sequences of these ORF were subjected to bioinformatics analyses as part of a weight-of-evidence assessment of the potential allergenicity and toxicity of any proteins potentially encoded within the BPS-CV127-9 transgene insert sequence. No ORF deduced amino acid sequence showed significant homology to a known protein toxin as defined in the United States Code of Federal Regulations ([40 CFR Part 725.421](#)). No ORF deduced amino acid sequence

showed 35% or greater identity over 80 amino acids to a potential allergen. And finally, no ORF deduced amino acid sequence shared a sequence of eight or more consecutive identical amino acids with a potential allergen. Thus, the bioinformatics analyses of the ORFs found within the BPS-CV127-9 transgene insert sequence do not provide any indication of a potential allergenicity or toxicity concern.

RECORDS RETENTION

Search records, the original copy of this report, and other relevant records are archived at BASF, 26 Davis Drive, Research Triangle Park, North Carolina 27709 U. S. A.

STUDY PERSONNEL

Bioinformatics scripts utilized for the work reported herein were written by [REDACTED] Ph.D., and [REDACTED], M.S., BASF Plant Science L. L. C., 26 Davis Drive, Research Triangle Park, North Carolina 27709.

REFERENCES

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389-3402.

Altschul, S. F., Wootton, J. C., Gertz, E. M., Agarwala, R., Morgulis, A., Schaffer, A. A., and Yu, Y.-K. (2005) Protein database searches using compositionally adjusted substitution matrices. *FEBS J.* 272:5101-5109.

Bannon, G. A., Goodman, R. E., Leach, J. N., Rice, E., Fuchs, R. L. and Astwood, J. D. (2002) Digestive stability in the context of assessing the potential allergenicity of food proteins. *Comments on Toxicology* 8:271-285.

Bannon, G. A. and Ogawa, T. (2006) Evaluation of available IgE-binding epitope data and its utility in bioinformatics. *Mol. Nutr. Food Res.* 50:638-644.

Codex Alimentarius Commission. (2003) Alinorm 03/34: Joint FAO/WHO Food Standard Programme, Codex Alimentarius Commission, Twenty-Fifth Session, Rome, Italy, 30 June-5 July, 2003. Appendix III, Guideline for the conduct of food safety assessment of foods derived from recombinant-DNA plants, and Appendix IV, Annex on the assessment of possible allergenicity, pp. 47-60.

Hileman, R. E., Silvanovich, A., Goodman, R. E., Rice, E. A., Holleschak, G., Astwood, J. D., and Hefle, S. L. (2002) Bioinformatics methods for allergenicity assessment using a comprehensive allergen database. *Int. Arch. Allergy Immunol.* 128:280-291.

Jander, G., Baerson, S. R., Hudak, J. A., Gonzalez, K. A., Gruys, K. J., and Last, R. L. (2003) Ethylmethanesulfonate saturation mutagenesis in *Arabidopsis* to determine frequency of herbicide resistance. *Plant Physiol.* 131:139-146.

Jiang, X. (2007) AtSEC61 γ subunit protein expression in Cultivance Soybean Event 127. BASF Plant Science Report Number BPS-010-07.

McKean, A. (2008) Bioinformatics analysis of deduced amino acid sequences of *Arabidopsis thaliana* ahas1 and SEC61 γ from herbicide-tolerant soybean BPS-CV127-9 for allergenicity and toxicity potential BASF Plant Science Report Number BPS-014-08.

Metcalf, D. D., Astwood, J. D., Townsend, R., Sampson, H. A., Taylor, S. L., and Fuchs, R. L. (1996) Assessment of the allergenic potential of foods derived from genetically engineered crop plants. Crit. Rev. Food Sci. Nutr. 36:S165-S186.

Osborne, A. R., Rapoport, T. A., and van den Berg, B. (2005) Protein translocation by the Sec61/SecY channel. Annu. Rev. Cell Dev. Biol. 21:529-550.

Pearson, W. R. and Lipman, D. J. (1988) Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA 85:2444-2448.

Sathasivan, K., Haughn, G. W., and Murai, N. (1990) Nucleotide sequence of a mutant acetolactate synthase gene from an imidazolinone-resistant *Arabidopsis thaliana* var. Columbia. Nucleic Acids Res. 18:2188.

Shen, Y. (2007) Molecular characterization of Cultivance Soybean Event 127. BASF Plant Science Report Number BPS-001-06.

Silvanovich, A., Nemeth, M. A., Song, P., Herman, R., Tagliani, L., and Bannon, G. A. (2006) The value of short amino acid sequence matches for prediction of protein allergenicity. Toxicol. Sci. 90:252-258.

Stevenson Paulik, J. (2007) Comparison of inhibition of AtAHAS (S653N) and AtAHAS (S653N R272K) by imidazolinones and sulfonylurea. BASF Plant Science Report Number BPS-004-06.

Table 1. Summary of Open Reading Frames Present within the Transgene Insert of BPS-CV127-9

Query Sequence	Sequence Length (aa)	Top BLAST Result (accession number)	
CV127.1392_d3	44	Protein transport protein SEC61 gamma subunit, putative.	<i>Arabidopsis thaliana</i> (NP_566909)
CV127.1365_c2	65	none	
CV127.1559_c3	41	none	
CV127.1581_c2	51	none	
CV127.1962_c2	39	none	
CV127.2041_d1	34	none	
CV127.2197_d1	45	none	
CV127.2700_d3	60	none	
CV127.2733_c2	54	none	
CV127.2704_c1	106	none	
CV127.2973_d3	36	none	
CV127.3021_c2	35	none	
CV127.3311_c3	37	none	
CV127.3783_d3	31	none	
CV127.3885_d3	47	none	
CV127.3797_c3	142	none	
CV127.3982_c1	72	none	
CV127.4164_c2	38	none	
CV127.4252_c1	72	none	
CV127.4448_c3	124	none	
CV127.4778_d2	34	none	
CV127.5086_c1	44	none	
CV127.5534_c3	51	none	
CV127.5688_d3	38	none	
CV127.5723_c3	62	none	
CV127.5811_d3	47	none	
CV127.5729_d2	166	Acetolactate synthase catalytic subunit.	<i>Arabidopsis thaliana</i> (ABJ80681)

Figure 1. Deduced Amino Acid Sequences of Open Reading Frames from the Transgene Insert of BPS-CV127-9

The deduced amino acid sequences of open reading frames that are contained either fully or partially within the transgene insert from BPS-CV127-9 are shown below.

```
>CV127.1392_d3
MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRKGNAFSQKNLISDF

>CV127.1365_c2
MISKFQKTLYRSKIGNEIFLRKGVTLAIGFVTALDETNTALSEISERIDCGINGFHDFFLDRNW

>CV127.1559_c3
MMLLIGMKTSFTKNPTNPITTNPIAVRTATFVNSDNQFNYK

>CV127.1581_c2
MRDQLPRKESTTLRRSNDVVDWDENELHEESNESHHSKSNRRTHSYFREF

>CV127.1962_c2
MTPERTNTEQTKLMGRIDIIKIQGDRSVSIAEAMETAKH

>CV127.2041_d1
MSFVCSVFVRSGVMDGVSVPFCFCNACWVVSCKKC

>CV127.2197_d1
MLRGVGGEERTKRHYQPFEEALSFVIVFVAKFSILIPFSLISCSL

>CV127.2700_d3
MFTSLFFFSVSLSLAHLSPWPWRQQQQQHLRSPSPPNHLLPPPNHHYQSPDSPSHSP

>CV127.2733_c2
MICWGLGRMGRIWRLVMVIWRRKEMVWRRRSKKMLLLLLLPPPWFRKMSKR

>CV127.2704_c1
MNVSGLVGLVGEGVVVTLVVVLSTAEMEEGELDLIPRRREEEDDLLGFRENGRENLEIGNGDLEEEGDGLV
EKEIEEDVVVVVAAAMVQEKDEQEIEKQRRRDL

>CV127.2973_d3
MSQPLPLQPNLNPKNHSSPDSLQINPAKALISSKL

>CV127.3021_c2
MHLQKRIREFLRDLVLRGYQRLCGVDLERIGR

>CV127.3311_c3
MLRNRTRYLNRSLLKRICTNHTTRDLSCDCYKRNTI

>CV127.3783_d3
MMSCRYICLECMGLCMQITLWSIVICCWRLG

>CV127.3885_d3
MIVSRVSLRLLLVGLRLFILILTRLRLGRIRLLMCLCVVMLSWLCKG
```


>CV127.3797_c3

MLPDTSTYYGFSISQLIKDLNRILWRNSFPKRLKAQRKLLFLYVQLIPPNSKIKLKLRSVLKNLIHPLQS
QLNITTHRHMRSLILPNLSRVNINMNNLSPTSLSKLTRDTIIPYPKRQQQITMLHSVICIHSPMHSKHM

>CV127.3982_c1

MAFPSVNSSRTLIAYCGGIASPNVLKLNNGNFCCTFNFLQTPKSSLSSSARFSRTLFIPTCKASLTSPHTD
T

>CV127.4164_c2

MIATALVSCNCRTPPTFDVARHQYLLWLFHQSTHQGP

>CV127.4252_c1

MPRCWLFNKSTFTGRFSTRIVASSCTFIKLPSPSISTTIASGLATDAPIAAGSPNPIAPRPPDDSHCLGF
L

>CV127.4448_c3

MHTKTIKGSNDHRFHLNVFNPAISFRYDIVESATIRDHRQHMFLVRTNHIQQVRSWCIOHCLNSFSEICFL
CHPRRWNPASCCKQQHVREYLVLLSRIPEKCVSSVSFVEPIFPLHNHAKMLVV

>CV127.4778_d2

MKPVIIRTFYGLCMHMKKLSLQFPVCFGNLSFF

>CV127.5086_c1

MWTQLILLDDIVVLNFNLETTAHMSTFQLKTNGKSIFHARFERY

>CV127.5534_c3

MLEKEMKKYKFLLFIFKFTTLILVSFLFFKSFCHLLKKKKLFEIFLLPMSD

>CV127.5688_d3

MRDLILVMMSCRYICLECMGLCMQITLWSIVICCWRL

>CV127.5723_c3

MRSILILPNLSRVNINMNNLSPTSLSKLTRDTIIPYSKRQQQITMLHSVICIHSPMHSKHM

>CV127.5811_d3

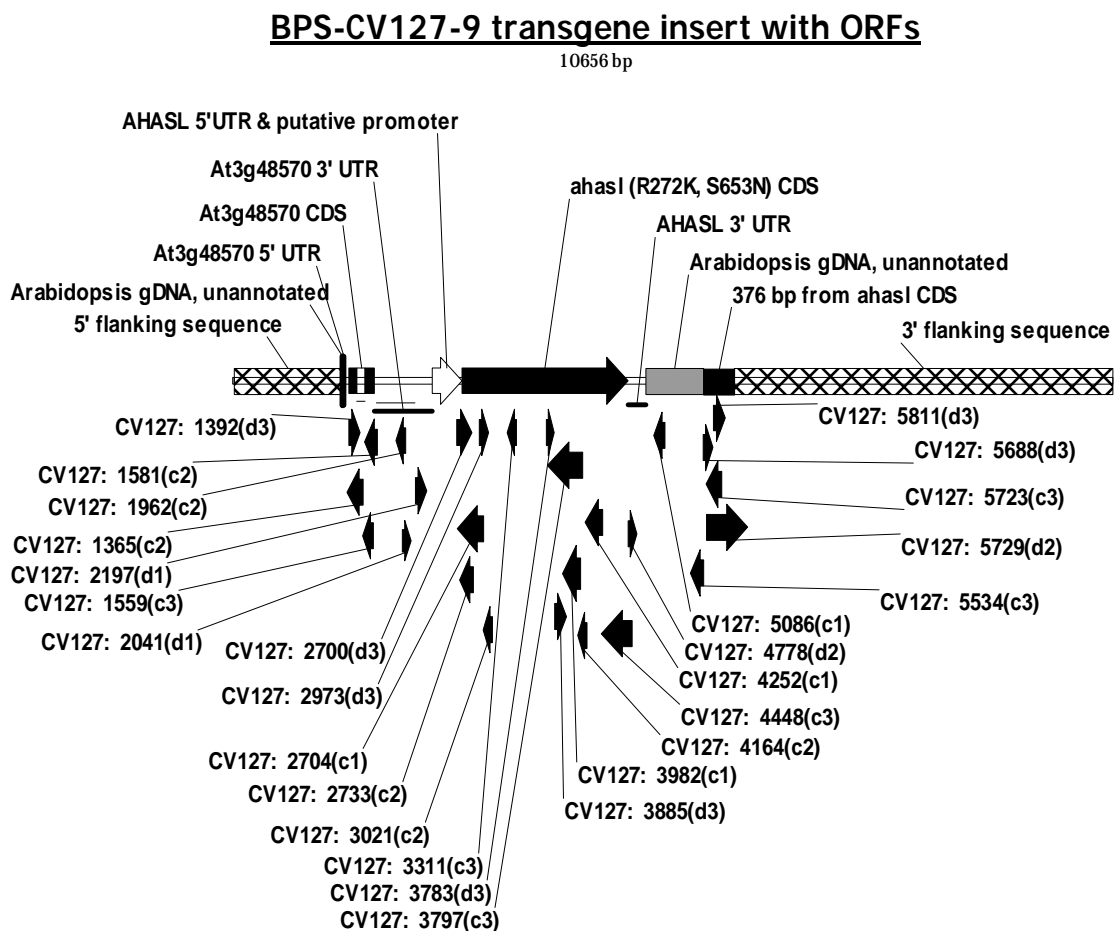
MIVSRVSLRLLLVGLRLFILILTRLRLGRIRLLMCLCVVMSWLCKG

>CV127.5729_d2

MLGMHGTVYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNKTPHVSVCQDVKLA
LQGMNKVLENRAEELKLDGFWRNELNVQKQKFLSFKTFGEAVPCPSKKTVDHPSYEWVGPVIRREGPPSP
MREDKLTQNTHLIRTYKFVFFKCI

Figure 2. Graphical Representation of Open Reading Frames within the Transgene Insert of BPS-CV127-9

The graphical representation of open reading frames of 29 or more amino acids within the transgene insert of BPS-CV127-9 is shown below.



APPENDIX 1. 80-AMINO ACID ALLERGEN HOMOLOGY SEARCH RESULTS

Typical output from the FASTA search of an allergen database using either an 80-amino acid segment from ORFs with deduced amino acid sequences greater than 80 amino acids or, if the ORF deduced amino acid sequence is less than 80, the entire ORF deduced amino acid sequence as the query sequence is shown below in Appendix 1.

CV127.5729_d2: amino acids 5-84

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.5729_d2_5-84, 80 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: rho(ln(x))= 2.5793+/-0.00366; mu= 16.8140+/- 0.191

mean_var=42.0176+/-11.266, 0's: 8 Z-trim: 8 B-trim: 101 in 2/42

Lambda= 0.197860

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2

join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:

	gi	gi	gi	gb	AA	len	opt	bits	E(1313)	%_id	%_sim	alen
gi 60280803	gi 60280803	gb AAX18294	(160)		59	22.8	2.3	0.228	0.667	57		
gi 51860756	gi 51860756	gb AAU11502	(131)		56	21.8	3.8	0.346	0.551	78		
gi 60280823	gi 60280823	gb AAX18304	(159)		56	21.9	4.2	0.228	0.632	57		
gi 60280825	gi 60280825	gb AAX18305	(159)		56	21.9	4.2	0.228	0.632	57		
gi 14285797	gi 14285797	sp O61379.1	(274)		57	22.6	4.6	0.269	0.615	52		
gi 125995167	gi 125995167	dbj BAF47	(284)		57	22.6	4.6	0.269	0.615	52		
gi 125995169	gi 125995169	dbj BAF47	(284)		57	22.6	4.6	0.269	0.615	52		
gi 2660868	gi 2660868	gb AAC48288.1	(284)		57	22.6	4.6	0.269	0.615	52		
gi 125995171	gi 125995171	dbj BAF47	(284)		57	22.6	4.6	0.269	0.615	52		
gi 14285796	gi 14285796	sp O44119.1	(284)		57	22.6	4.6	0.269	0.615	52		
gi 19338630	gi 19338630	gb AAL86739	(448)		58	23.2	4.8	0.230	0.557	61		
gi 60280827	gi 60280827	gb AAX18306	(159)		54	21.4	6.2	0.214	0.643	70		
gi 2739154	gi 2739154	gb AAC67308.1	(191)		54	21.5	6.9	0.298	0.526	57		
gi 1531589	gi 1531589	gb AAB09632.1	(631)		57	23.2	7	0.588	0.765	17		
gi 729764	gi 729764	sp P40918.1	HSP (643)		57	23.2	7	0.237	0.593	59		
gi 26190140	gi 26190140	emb CAD2055	(179)		53	21.1	8.1	0.292	0.492	65		
gi 729979	gi 729979	sp P39673.1	MAG (341)		54	21.9	9.2	0.260	0.560	50		
gi 7024506	gi 7024506	gb AAF35431.1	(264)		53	21.4	9.9	0.269	0.596	52		
gi 59800145	gi 59800145	sp P69198.1	(83)		50	19.7	9.9	0.230	0.554	74		
gi 59800146	gi 59800146	sp P69199.1	(83)		50	19.7	9.9	0.230	0.554	74		

>>>CV127.5729_d2_5-84, 80 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|60280803 gi|60280803|gb|AAX18294.1| major allergen (160 aa)

initn: 26 initl: 26 opt: 59 Z-score: 94.9 bits: 22.8 E(): 2.3

Smith-Waterman score: 59; 22.807% identity (66.667% similar) in 57 aa overlap (10-66:75-129)

gi|602 MGVTFTYETEFTSVIPAPRLFKAFILDGDNLPKIAIPQAIKSTIEIGDGGVGTIKKVTFG
10 20 30 40 50 60

```

              10      20      30      40      50
CV127.      HGTVYANYAVEHSDLLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNKT
              ..... : .. ::: .. :... :... :... :
gi|602 EGSQYGYVKQRVNGIDKDNFTYSYSMIEGDTLSDKLEKITYETKLI-ASPDGGSIIKT-T
              70      80      90      100     110

```

```

              60      70      80
CV127. PHVSVCQDVKLALQGMNKVLENRAE
              : . :...
gi|602 SHYHAKGDVEIKEEHVKAGKEKASGLFKLLEAYLVANPDAYN
              120     130     140     150     160

```

>>gi|51860756 gi|51860756|gb|AAU11502.1| fatty acid-bidi (131 aa)
 initn: 34 initl: 34 opt: 56 Z-score: 91.1 bits: 21.8 E(): 3.8
 Smith-Waterman score: 56; 34.615% identity (55.128% similar) in 78 aa overlap
 (7-80:7-76)

```

              10      20      30      40      50
CV127. HGTVYANYAVEHSDLLLA----GVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP
              .: .:..: .: : : : : : : : : : : : : : : : :
gi|518 MVQLNGSYKLEKSDNFDAFLKELGVNF---VTRNLAKSASPTVEIVDGDSTYTI---KTS
              10      20      30      40      50

```

```

              60      70      80
CV127. HVSVCQDVKLALQGMNKVLENRAE
              . . .:.. : : .. :...
gi|518 STLKNSEIKFKL-G-EEFEEDRADGKKVQTSVTKEGDNKLQVQKGDKPVTIVREFSEEG
              60      70      80      90      100     110

```

```

gi|518 LTVTATVNGVTSVRFYKRQ
              120     130

```

>>gi|60280823 gi|60280823|gb|AAX18304.1| major allergen (159 aa)
 initn: 24 initl: 24 opt: 56 Z-score: 90.3 bits: 21.9 E(): 4.2
 Smith-Waterman score: 56; 22.807% identity (63.158% similar) in 57 aa overlap
 (10-66:75-128)

```

gi|602 MGVLTYTEYASVIPPARYNALVLADNLPKIAPQAVKTVEILEGDDGGVGTIKKVSFG
              10      20      30      40      50      60

```

```

              10      20      30      40      50
CV127.      HGTVYANYAVEHSDLLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNKT
              ..... : .. ::: .. :... :... :... :
gi|602 EGSEYSYVKKHKEGIDKDNFVYSYNLIEGDAISDKIEKISYEIKLVASG--SGSIIKNIS
              70      80      90      100     110

```

```

              60      70      80
CV127. PHVSVCQDVKLALQGMNKVLENRAE
              : . :...
gi|602 -HYHTKGDVEIKEEHVKAGKERAGHLFKLIENHLVANPDAYN
              120     130     140     150

```

>>gi|60280825 gi|60280825|gb|AAX18305.1| major allergen (159 aa)
 initn: 24 initl: 24 opt: 56 Z-score: 90.3 bits: 21.9 E(): 4.2
 Smith-Waterman score: 56; 22.807% identity (63.158% similar) in 57 aa overlap
 (10-66:75-128)

```

gi|602 MGVLTYTEYASVIPPARYNALVLADNLPKIAPQAVKTVEILEGDDGGVGTIKKVSFG
              10      20      30      40      50      60

```

CV127. HGT VYANYAVEHS D L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K T
 : . . : . : . : . : . : . : .
gi|602 EGSEYSYV KHKVEGIDKNFVYSYSLIEGDAISDKIEKISYEIKLVASG--SGSI IKNIS
 70 80 90 100 110

```

          60          70          80
CV127. PHVSVCGDVKLALQGMNKVLENRAE
      :  .  : : . .
gi|602 -HYHTKGDVEIKEEHVKAGKERAGHLFKLIENHLVANPDAYN
      120      130      140      150

```

```
>>gi|14285797 gi|14285797|sp|O61379.1|TPM_PANST Tropomyo (274 aa)
  initn: 37 initl: 37 opt: 57 Z-score: 89.7 bits: 22.6 E(): 4.6
Smith-Waterman score: 57; 26.923% identity (61.538% similar) in 52 aa overlap
(29-79:73-124)
```

CV127. 10
HGTVYANYAVEHSDLL

gi | 142 MKLEKDNAMDRADTLEQQNKANIRAEKAEVEVHNLQKRMQQLENDLDQVQESLLKANTQ

10 20 30 40 50 60

CV127. LAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNK-TPHVSVCGDVKLALQGMNKKVL
gi|142 LEEKDKALSNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAADESERMRKVL

CV127. ENRAE
:::
gi | 142 ENRSLSD EERMDALENQLKEARFLAE EAD RKYDEVARKLAMVEADLERAE EERAETGESKF
130 140 150 160 170 180

gi|142 VELEEEELRVVGNNLKSLEVSEEKANQREEAYKEQIKTLTNKLKAAEARAFAERSVQKLQ
190 200 210 220 230 240

gi|142 KEVDRLDELVNEKEKYKSITDELDQTFSELSGY
250 260 270

>>gi|125995167 gi|125995167|dbj|BAF47267.1| tropomyosin (284 aa)
initn: 37 initl: 37 opt: 57 Z-score: 89.5 bits: 22.6 E(): 4.6
Smith-Waterman score: 57; 26.923% identity (61.538% similar) in 52 aa overlap
(29-79:83-134)

CV127. HGTVYA

gi | 125 MDAIKKKMQAMKLEKDNAMDKADTLEQQNKANLRAEKTEEEIRANQKKSQLVENELDHA

10 20 30 40 50 60

CV127. NYAVEHSDLLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNK-TPHVSVCADV
gi|125 QEQLSAATHKLVEKEKAFANAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAESQAA

```

              70      80
CV127. LALQGMNKVLENRAE
          . : : : : .
gi|125  DESERMKRVLENRSLSD EERMDALENQLKEARFLAE EADRKYDEVARKLAMVEADLERAE
              130      140      150      160      170      180

```

```
gi|125 ERAESGESKIVELEEEELRVVGNNLKSLEVSEEKANQREETYKEQIKTLANKLKAAAEARAE
      190          200          210          220          230          240
```

gi|125 FAERSVQKLQKEVDRLLEDELVNEKEKYKNIADEMDQAFSELSGF
250 260 270 280

```
>>gi|125995169 gi|125995169|dbj|BAF47268.1| tropomyosin (284 aa)
  initn: 37 initl: 37 opt: 57 Z-score: 89.5 bits: 22.6 E(): 4.6
Smith-Waterman score: 57; 26.923% identity (61.538% similar) in 52 aa overlap
(29-79:83-134)
```

CV127. HGTVYA

gi|125 MDAIKKKMQAMKLEKDNAMDRAHTLEQQNKEANLRAEKTEEEIRATQKKMQQIENELDHA
10 20 30 40 50 60

CV127. NYAVEHSDLLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNK-TPHVSVCGDVK
 ... : : : ... : : : : ...

gi | 125 QEQLSAANTKLDEKEKALQNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA
70 80 90 100 110 120

CV127. LALQGMNKVLENRAE
 . : : : : .

gi|125 DESERMKRVLENRSLSDEERMDALENQLKEARFLAEEDRKYDEVARKLAMVEADLERAE
130 140 150 160 170 180

```
gi|125  ERAESGESKIVELEEEELRVVGNNLKSLEVSEEKANQREETYKEQIKTLANKLKAAEARAE
      190          200          210          220          230          240
```

gi|125 FAERSVQKLQKEVDRLLEDELVNEKEKYKSITDELDQTFSELSGY
250 260 270 280

```
>>gi|2660868 gi|2660868|gb|AAC48288.1| fast tropomyosin (284 aa)
  initn: 37 initl: 37 opt: 57 Z-score: 89.5 bits: 22.6 E(): 4.6
Smith-Waterman score: 57; 26.923% identity (61.538% similar) in 52 aa overlap
(29-79:83-134)
```

CV127. HGTVYA

gi | 266 MDAIKKKMQAMKLEKDNAMDRADTLEQQNKEANIRAEKSEEEVHNLQKRMQQLENDLDQV

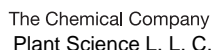
CV127. NYAVEHSDLLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNK-TPHVSVCQDVK
 :.. :. : :.. :. : :

gi|266 QESLLKANTQLEEKDKALSNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA
70 80 90 100 110 120

CV127. LALQGMNKVLENRAE
 . : : : : .

gi|266 DESERMKRVLENRSLSDEERMDALENQLKEARFLAE EADRKYDEVARKLAMVEADLERAE
130 140 150 160 170 180

```
gi|266 ERAETGESKIVELEEEELRVVGNNLKSLEVSEKANQREEAYKEQIKTLANKLKAAEARAE
      190          200          210          220          230          240
```



```
>>gi|125995171 gi|125995171|dbj|BAF47269.1| tropomyosin (284 aa)
  initn: 37 initl: 37 opt: 57 Z-score: 89.5 bits: 22.6 E(): 4.6
Smith-Waterman score: 57; 26.923% identity (61.538% similar) in 52 aa overlap
(29-79:83-134)
```

CV127. HGTVYA

gi|125 MDAIKKKMQAMKLEKDNAMDRAHTLEQQNKEANLRAEKTEEEIRATQKKMQQIENELDHA
10 20 30 40 50 60

10 20 30 40 50 60

CV127. NYAVEHSDLLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNK-TPHVSVCGDVK
: . : : . . : . : : : . .

gi | 125 QEQLSAANTKLDEKEKALQNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA
70 80 90 100 110 120

70 80

CV127. LALQGGMNKVLENRAE

```

      . : : : : : .
gi | 125 DESERMKVLNRSLSDEERMDALENQLKEARFLAEADRKYDEVARKLAMVEADLERAE

```

```
gi | 125  ERAESGESKIVELEEEELRVVGNLKSLEVSEKANQREETYKEQIKTLANKLKAEEARAE
          190          200          210          220          230          240
```

gi | 125 FAERSVQKLQKEVDRLEDELVNEKEKYKNIADEMDQAFSELSGF
250 260 270 280

```
>>gi|14285796 gi|14285796|sp|O44119.1|TPM_HOMAM Tropomyo (284 aa)
  initn: 37 initl: 37 opt: 57 Z-score: 89.5 bits: 22.6 E(): 4.6
Smith-Waterman score: 57; 26.923% identity (61.538% similar) in 52 aa overlap
(29-79:83-134)
```

CV127. HGTVYA

gi|142 MDAIKKKMQAMKLEKDNAMDRADTLEQQNKEANIRAEKTEEEIRITHKKMQQVENELDQV
10 20 30 40 50 60

10 20 30 40 50 60

CV127. NYAVEHSDLLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNK-TPHVSVCGDVK
 :.. : : :.. : : :

gi | 142 QEQLSLANTKLEEKEKALQNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA
70 80 90 100 110 120

70 80

CV127. LALQG MNKVLENRAE

gi | 142 DESERM RKVLENRSL SDEERMDALENQLKEARFLAEADRKYDEVARKLAMVEADLERAE

```
gi|142 ERAETGESKIVELEEEELRVVGNNLKSLEVSEKANQREEAYKEQIKTLANKLKAEEARAE
      190          200          210          220          230          240
```

gi | 142 FAERSVQKLQKEVDRLLEDELVNEKEKYKSITDELDQTFSELSGY
250 260 270 280

```
>>gi|19338630 gi|19338630|gb|AAL86739.1|AF441864_1 48-kD (448 aa)
  initn: 63 initl: 40 opt: 58 Z-score: 89.2 bits: 23.2 E(): 4.8
Smith-Waterman score: 58; 22.951% identity (55.738% similar) in 61 aa overlap
(2-59:277-337)
```

gi|193 MLPKEDPELKCKHKCRDERQFDEQQRRDGKQICEEKARERQQEEGNSSEESYGKEQEEN
10 20 30 40 50 60gi|193 PYVFQDEHFESRVKTEEGRVQVLENFTKRSRLLSGIENFRLAILEANPHTFISPAHFDAE
70 80 90 100 110 120gi|193 LVLVFAKGRATITMVREEKRESFNVEHGDIIRIPAGTPVYMINRDENEKLFIVKILQPVS
130 140 150 160 170 180gi|193 APGHFEAFYGAGGEDPESFYRAFSWEVLEAALKVRREQLEKVFGEQSKGSIVKASREKIR
190 200 210 220 230 240

CV127. 10 20
HGTVYANYAVEHSDLL-LAFGVRFD
: : : : : : : : : :

gi|193 ALSQHEEGPPRIWPFGGESSGPINLLHKHPSQSNQFGRLYEAHPDDHKQLQDLDMVSFA
250 260 270 280 290 300

CV127. DRVTGKLEA--FASRAKIVHIDIDSAEIGKNKTPHVSVCGDVKLALQGMNKVLENRAE
 : : : : : :
 gi|193 NITKGS MAGPYNSRATKISVVVEGEGFFEMACPHLSSSSGSYQKISARLRGVVFVAPA
 310 320 330 340 350 360

gi|193 GHPVAVIASQNNNLQVLCFEVNAHGNSRFPLAGKGNIVNEFERDAKELAFNLPSREVERI
370 380 390 400 410 420gi|193 FKNQDQAFFFPGPNKQQEEGGRGGRAFE
430 440

```
>>gi|60280827 gi|60280827|gb|AAX18306.1| major allergen (159 aa)
  initn: 24 initl: 24 opt: 54 Z-score: 87.2 bits: 21.4 E(): 6.2
Smith-Waterman score: 54; 21.429% identity (64.286% similar) in 70 aa overlap
(10-79:75-140)
```

gi|602 MGVLTYET EYASVIPP ARLYNALVLDADNLIPK IAPQAVKTVEILEG DGGVGTIKKVSFG
10 20 30 40 50 60

CV127. HGT VYANYAVEHS D L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K T
 : . : . . . : . :
gi|602 EGSEYSYVKKHKVEGIDKDNFVYSYSLIEGDAISDKIQIKISYEIKLVASG--SGSIKNIS
 70 80 90 100 110

```

          60          70          80
CV127. PHVSVC G DVK LALQGMNKVLENRAE
      :  .  : : : . . . : . : :
gi | 602 -HYHTKG DVEIKEENV-KAGKERAHGLFKLIENHLVANPDAYN
          120        130        140        150

```

```
>gi|2739154 gi|2739154|gb|AAC67308.1| 22.6 kDa tegument (191 aa)
initn: 31 initl: 31 opt: 54 Z-score: 86.5 bits: 21.5 E(): 6.9
Smith-Waterman score: 54; 29.825% identity (52.632% similar) in 57 aa overlap
(22-73:54-110)
```



```

CV127.                                10      20
HGT VYANYAVEHS D LLLA FGVR FDDRVT
                               :::
gi|273 MATTE YRLSLMEQFIRAFIEIDKDNNELIDKQELTKYCQQNQMDMKQIDPWIARFDTDKD
          10           20           30           40           50           60

          30           40           50           60           70           80
CV127. GK--LEAFASR--AKIVHIDIDS AEIGKNKTPHVS VCG-DVKLALQGMNKVLENRAE
:: :: :       . .   . .   . .   . .   . .   . .   . .   . .   . .
gi|273 GKV SLEEF CRGFGFLK VWEVRREKEELKKDK EGKVSTLPDI QIIAAATMSKAKQYNICCKF
          70           80           90           100          110          120

gi|273 KELLDKTSRTGDEVRAVANDLKAF LDSEYGRVWQVIILTGSYW MNFSHEPF LSMQFKYSN
          130          140          150          160          170          180

gi|273 YVCL LW RTPSS
          190

```

```
>>gi|1531589 gi|1531589|gb|AAB09632.1| allergen [Peripla (631 aa)
      initn: 45 initl: 45 opt: 57 Z-score: 86.3 bits: 23.2 E(): 7
Smith-Waterman score: 57; 58.824% identity (76.471% similar) in 17 aa overlap
(42-58:5-20)
```

	10	20	30	40	50	60
CV127.	HGT	VYANYAVEHSD	LLLA	FGVRFDD	RV	TGKLEAFASRAKIVHIDIDSAEIGKNKTPHVS
					:	: : . : . : . : : :
gi 153					DIGDHYDIE-ANIGHYKYPHVVK	
					10	20
	70	80				
CV127.	CGD	VKLALQGMN	KVLE	NR	AE	
gi 153	NFIS	YKKGLL	PRGEP	FSVY	EKHREQA	IKLFELFFAANDYDTFYKTACWARDRVNEG
	30	40	50	60	70	80
gi 153	MYALT	VAAFH	REDTK	DLVL	PPPYE	VNPYLFVEDDVIQQAYKYWTKESGTDKHVEHVIPVN
	90	100	110	120	130	140
gi 153	FTARS	QEDLV	AYFRED	VDLNA	FNMYF	RYIYPSWFNTTLYGKSFDRRGEQFYTYHQIYAR
	150	160	170	180	190	200
gi 153	YFLER	LSNSL	PDVKP	FQYSK	PLKTG	YNPHLRYHNGEEMP
	210	220	230	240	250	260
gi 153	YERR	VEKAID	FGYAF	DEHRT	PYSLY	HDQHGM
	270	280	290	300	310	320
gi 153	VGHV	VDPYH	KNGLA	PSALE	HPQTAL	RDPAFYQLWKRIDHIVQKYKNR
	330	340	350	360	370	380
gi 153	VKIEN	VDVGK	LYTYF	EHFEH	SLGNAM	YLGKLEDYMKASIRARHYRLNHKPF
	390	400	410	420	430	440
gi 153	KAQD	VYVRIF	LGP	KYDSL	GHECE	LDERRHYFVEMDRFVHKVEAGKT
	450	460	470	480	490	500
gi 153	SHDS	YRNL	FKKV	SDAL	QEKDQ	YYIDKSHKYCGYPENLLLPKGKKG
	510	520	530	540	550	560
gi 153	QDEH	DFEPY	HYKAF	SYCGV	HGRKYP	DDKPLGFPFDRKIH
	570	580	590	600	610	620

gi | 153 YDEVHDTVH
630

```
>>gi|729764 gi|729764|sp|P40918.1|HSP70_CLAHE Heat shock (643 aa)
  initn: 31 initl: 31 opt: 57 Z-score: 86.3 bits: 23.2 E(): 7
Smith-Waterman score: 57; 23.729% identity (59.322% similar) in 59 aa overlap
(2-60:10-65)
```

		10	20	30	40	50
CV127.		HGTVYANYAVEHSDLLLA	FGVRFDDRVTGKLEAFAS	RAKIVHIDIDS	AEIGK	
		:::: . . .:	. . .:::	. . ::::	. . .	:::
gi 729	MAPAIGIDL	GTYS	CVGIYRDDRIEII	ANDQGNRTTPSF	VAF	TDTERLIG---DS
	10	20	30	40	50	
	60	70	80			
CV127.	NKTPHVS	VC	GDV	KLALQ	GMN	KVLENRAE
	:::	:				
gi 729	AINPHNT	VFDAKRL	IGRK	FQDAE	VQADMKH	FPFKVIEKAGK
	60	70	80	90	100	110
gi 729	SSMILTK	MRETA	ESYLG	GTVNN	AVITVP	PAYFNDSQRQATK
	120	130	140	150	160	170
gi 729	AIAYGLD	KKQ	EGEKN	VLIFD	LGGGTF	DVSFLTIEEGIFEV
	180	190	200	210	220	230
gi 729	HFSNEF	KRKHKK	DLSDN	ARALRRL	RTACERAK	RTLSSSAQTSIEID
	240	250	260	270	280	290
gi 729	ARFEEV	GQDL	FRGNME	PGERTL	RDDKIDK	SSVHEIVLGGG
	300	310	320	330	340	350
gi 729	CKSINP	DEAVAY	GA	VQAAIL	SGDTS	SKSTKEILL
	360	370	380	390	400	410
gi 729	TIPTKK	SETF	STFSD	NQPGV	LIQVF	EGERARTKDIN
	420	430	440	450	460	470
gi 729	DL	DANGIM	NVSALE	KGTG	TKNKIVIT	NDKGRLSKEE
	480	490	500	510	520	530
gi 729	KNGLES	YAYS	LKNTV	SDPKV	EEKLSA	EDKETLTGAID
	540	550	560	570	580	590
gi 729	LESVAN	PVMM	KIYGA	E	GGAPG	GM
	600	610	620	630	640	

```
>>gi|26190140 gi|26190140|emb|CAD20556.1| putative inver (179 aa)
  initn: 29 initl: 29 opt: 53 Z-score: 85.2 bits: 21.1 E(): 8.1
Smith-Waterman score: 53; 29.231% identity (49.231% similar) in 65 aa overlap
(15-79:15-72)
```

```

              10          20          30          40          50          60
CV127.  HGTVYANYAVEHSDLLLAFGVRFDDRVTKGLEAFASRAKIVHIDIDSAEIGKNKTPHVSV
              ::  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|261  MKLSFSLCIIFFNLLLLLQAVISADIVQGTCKKVAQSPNVNVDYFCVKS LGADPKSHTA-
              10          20          30          40          50

              70          80
CV127.  CGDVKLALQGMNKKVLENRAE
              :  ::::  :  :  :
gi|261  --D---LQGLGVISANLAIQHGSKIQTFIGRILKSKVDPALKKYLNDCVGLYADAKSSV
              60          70          80          90          100          110

```

gi|261 QEAIADFKSKDYASANKMSAALDDSVTCEDGFKEKKGIVSPVTKENKDYVQLTAISLAI
120 130 140 150 160 170

qi | 261 TKLLGA

```
>>gi|729979 gi|729979|sp|P39673.1|MAG_DERFA Allergen Mag (341 aa)
  initn: 34 initl: 34 opt: 54 Z-score: 84.2 bits: 21.9 E(): 9.2
Smith-Waterman score: 54; 26.000% identity (56.000% similar) in 50 aa overlap
(4-53:169-214)
```

gi|729 FVMKREPLRFRDITVEGNENAYIKNGKLHLSLMDPSTLSLVTKADGKIDMTVDLISPVTK
10 20 30 40 50 60gi|729 RASLKIDSKYNYLFHEGELSASIVNPRLSWHQYTKRDSREYKSDVELSLRSSDIALKITM
70 80 90 100 110 120

CV127. 10
HGTVYANYAVEHSDL
.:.: :.

gi|729 PDYNSKIHSRQGDQINMDIDGTLIEGHAQGTIREGKIHIKGRQTDFEIESNYRYEDGKL
130 140 150 160 170 180

CV127. LLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNKTPHVSVCGDVKLALQGMNKVL

gi|729 IIE-PVKSEN--GKLEGVLSRKVPShLTLETPrVKMNMKYDRYAPVKVFkLDYDGIHFE
190 200 210 220 230

80
CV127. ENRAE

gi|729 KHTDIEYEPGVRYKIIIGNGKCLKDDGRHYSIDVQGIPRKAFNLADLMDFKLVSKPEDSN
240 250 260 270 280 290gi|729 KAQFSYTFNEYTETEEYEFDPHRAYVNWLSSIRKYIQNFIVEDN
300 310 320 330 340

```
>>gi|7024506 gi|7024506|gb|AAF35431.1| heat stable aller (264 aa)
  initn: 37 initl: 37 opt: 53 Z-score: 83.6 bits: 21.4 E(): 9.9
Smith-Waterman score: 53; 26.923% identity (59.615% similar) in 52 aa overlap
(29-79:83-134)
```

CV127. HGTVYA

gi | 702 MDAIKKKMQAMKLEKDNAMDRADTLEQQNKEANLRAEKTEEEIRATQKKMQQVENELDQA
10 20 30 40 50 60[illegible]gi|702 QEQLSAANTKLDEKEKALQNAEGEVAALNRRIQLPEEDLERSEERLNTATTKLAEASQAA
70 80 90 100 110 120

CV127. LALQGMNKVLENRAE
 . : : : : .

gi | 702 DESERMRKVLENRSLSD EERMDALENQLKEARFLAE EADRKYDEVARKLAMVEADLERAE
130 140 150 160 170 180



The Chemical Company
Plant Science L. L. C.

BASF Reg. Doc. No. 2008/7019371

gi|702 ERAESGESKIVELEEELRVVGNNLKSLEVSEEKANQREETYKEQIKTLANKLKAAEAREAE
190 200 210 220 230 240

gi|702 FAERSVQKLQKEVDRLLEDELVNEK
250 260

>>gi|59800145 gi|59800145|sp|P69198.1|POLC2_BRANA Polcal (83 aa)
initn: 31 initl: 31 opt: 50 Z-score: 83.6 bits: 19.7 E(): 9.9
Smith-Waterman score: 53; 22.973% identity (55.405% similar) in 74 aa overlap
(6-73:4-77)

10 20 30 40 50
CV127. HGT VYANYAVEHSDLLAFGVRFDDRVTGKL--EAFASRAKIVHIDIDS--AEIGKNKTP
:. .:. . . : . : :. : : : : .
gi|598 MADATEKAEHDRIFKKFDANGDGKISASELGDALKNLG SVTHDDIKRMMAEIDTDGDG
10 20 30 40 50

60 70 80
CV127. HVS V--CGDVKLALQGMNKVLENRAE
..: .: : .: :
gi|598 YISYQEFSD FASANRGLMKDVAKIF
60 70 80

>>gi|59800146 gi|59800146|sp|P69199.1|POLC2_BRARA Polcal (83 aa)
initn: 31 initl: 31 opt: 50 Z-score: 83.6 bits: 19.7 E(): 9.9
Smith-Waterman score: 53; 22.973% identity (55.405% similar) in 74 aa overlap
(6-73:4-77)

10 20 30 40 50
CV127. HGT VYANYAVEHSDLLAFGVRFDDRVTGKL--EAFASRAKIVHIDIDS--AEIGKNKTP
:. .:. . . : . : :. : : : : .
gi|598 MADATEKAEHDRIFKKFDANGDGKISASELGDALKNLG SVTHDDIKRMMAEIDTDGDG
10 20 30 40 50

60 70 80
CV127. HVS V--CGDVKLALQGMNKVLENRAE
..: .: : .: :
gi|598 YISYQEFSD FASANRGLMKDVAKIF
60 70 80

80 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:51:50 2008 done: Sun Dec 14 22:51:50 2008
Total Scan time: 0.090 Total Display time: 0.040

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.2704_c1: amino acids 14-93

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.2704_c1_14-93, 80 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 4.7385 \pm 0.00352$; $\mu = 6.1224 \pm 0.180$
mean_var=47.5167 \pm 12.275, 0's: 8 Z-trim: 10 B-trim: 5 in 1/43
Lambda= 0.186059

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:		opt	bits	E(1313)	%_id	%_sim	alen
gi 18635	gi 18635 emb CAA33215.1 g (495)	76	27.0	0.4	0.305	0.695	59
gi 18615	gi 18615 emb CAA26723.1 u (495)	76	27.0	0.4	0.305	0.695	59
gi 94468552	gi 94468552 gb ABF18125 (258)	65	23.9	1.7	0.321	0.528	53
gi 170743	gi 170743 gb AAB02788.1 (815)	70	25.4	1.9	0.309	0.647	68
gi 21743	gi 21743 emb CAA43331.1 h (830)	70	25.4	1.9	0.309	0.647	68
gi 2114497	gi 2114497 gb AAB58417.1 (253)	63	23.4	2.5	0.368	0.579	38
gi 18637	gi 18637 emb CAA33216.1 g (485)	63	23.5	4.4	0.271	0.678	59
gi 18609	gi 18609 emb CAA26575.1 u (485)	63	23.5	4.4	0.271	0.678	59
gi 83715936	gi 83715936 dbj BAE5443 (284)	60	22.6	4.8	0.358	0.585	53
gi 169927	gi 169927 gb AAA33947.1 (218)	57	21.7	6.6	0.312	0.584	77
gi 83715928	gi 83715928 dbj BAE5442 (284)	58	22.0	6.9	0.344	0.574	61
gi 2735100	gi 2735100 gb AAD13646.1 (134)	54	20.8	7.5	0.268	0.683	41
gi 18639	gi 18639 emb CAA33217.1 g (481)	60	22.7	7.6	0.274	0.645	62
gi 15419048	gi 15419048 gb AAK96889 (233)	56	21.5	8.4	0.300	0.567	60
gi 13183177	gi 13183177 gb AAK15089 (585)	60	22.7	9	0.310	0.643	42

>>>CV127.2704_c1_14-93, 80 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|18635 gi|18635|emb|CAA33215.1| glycinin subunit G1 (495 aa)
initn: 70 initl: 45 opt: 76 Z-score: 108.7 bits: 27.0 E(): 0.4
Smith-Waterman score: 76; 30.508% identity (69.492% similar) in 59 aa overlap
(3-56:160-217)

gi|186 MAKLVFSLCFLFLFSGCCFAFSSREQPQQNECQIQKLNALKPDNRIESEGGLIETWNPNNK
10 20 30 40 50 60

gi|186 PFQCAGVALSRCTLNRLNRRPSYTNQGPQEIYIQQKGIFGMIYPGCPSTFEPPQQPQQR
70 80 90 100 110 120

CV127. VVVTLVVVLSTAEMEEGELDLIP
10 20
.: : .: .: .: .: .: .:

gi|186 GQSSRPQDRHQKIYNFREGDLIAVPTGVAWWMYNNEDTPVAVSIIIDTNSLEN-QLDQMP
130 140 150 160 170

CV127. RRR----EEEDDLLGFR-ENGRENLEIGNGDLEEEGDGLVEKEIEEDVVVVVVAAMVQE
30 40 50 60 70
.: : .: .: .: .: .: .:

gi|186 RRFYLAGNQEQEFLKYQQEQGGHQSQKQKHQQEENEGGSILSGFTLEFLEHAFSVDKQI
180 190 200 210 220 230

80
CV127. KD

gi|186 AKNLQGENEGEDKGAIVTVKGGLSVIKPPTDEQQQRPQEEEEEEDEKPKQCKGKDKHCQR
240 250 260 270 280 290

gi|186 PRGSQSKSRRNGIDETICTMRLRHNIQTSSPDIYNPQAGSVTTATSLDFPALSWRLLSA
300 310 320 330 340 350

gi|186 EFGSLRKNAMEFVPHYNLNANSIIYALNGRALIQVVCNNGERVFDGELQEGRVLIIPQNFV
360 370 380 390 400 410

gi|186 VAARSQSDNFYVSFKTNDTPMIGTLAGANSLNALPEEVIQHTFNLKSQQARQIKNNNP
420 430 440 450 460 470

gi|186 FKFLVPPQESQKRAVA
480 490

>>gi|18615 gi|18615|emb|CAA26723.1| unnamed protein prod (495 aa)
initn: 70 init1: 45 opt: 76 Z-score: 108.7 bits: 27.0 E(): 0.4
Smith-Waterman score: 76; 30.508% identity (69.492% similar) in 59 aa overlap
(3-56:160-217)

gi|186 MAKLVFSLCFLLFSGCCFAFSSREQPQQNECQIQKLNALKPGNRIESEGGLIETWNPNNK
10 20 30 40 50 60

gi|186 PFQCAGVALSRCTLNARNALRRPSYTNGPQEIYIQQKGIFGMIYPGCSSTFEPPQQPQQR
70 80 90 100 110 120

CV127. VVVTLVVVLSTAEMEEGELDLIP
10 20
:. : .:. : .:. : .:

gi|186 GQSSRPQDRHQKIYNSREGDLIAVPTGVAWWMYNNEDTPVAVSIIDTNSLEN-QLDQMP
130 140 150 160 170

CV127. RRR---EEEDDLLGFR-ENGRENLEIGNGDLEEEGDGLVEKEIEEDVVVVVAAAMVQE
30 40 50 60 70
:. : .:. : .:. : .:. : .:

gi|186 RRFYLAGNQEQEFLKYQQEQGGHQSQKQKHQQEEENEGGSILSGFTLEFLEHAFSVDKQI
180 190 200 210 220 230

80
CV127. KD

gi|186 AKNLQGENEGEDKGAIVTVKGGLSVIKPPTDEQQQRPQEEEEEEDEKPKQCKGKDKHCQR
240 250 260 270 280 290

gi|186 PRGSQSKSRRNGIDETICTMRLRHNIQTSSPDIYNPQAGSVTTATSLDFPALSWRLLSA
300 310 320 330 340 350

gi|186 GFGSLRKNAMEFVPHYNLNANSIIYALNGRALIQVVCNNGERVFDGELQEGRVLIIPQNFV
360 370 380 390 400 410

gi|186 VAARSQSDNFYVSFKTNDTPMIGTLAGANSLNALPEEVIQHTFNLKSQQARQIKNNNP
420 430 440 450 460 470

gi|186 FKFLVPPQESQKRAVA
480 490

```
>>gi|94468552 gi|94468552|gb|ABF18125.1| 30 kDa salivary (258 aa)
  initn: 39 initl: 39 opt: 65 Z-score: 97.2 bits: 23.9 E(): 1.7
Smith-Waterman score: 65; 32.075% identity (52.830% similar) in 53 aa overlap
(13-65:69-119)
```

CV127. VVVT

gi | 944 MKPLVKLFLLFCLVGIVLSRPMPEDEEPPVAEGGDEETDDAGGDGGEEENEHEEHAGDED
10 20 30 40 50 60

CV127. L V V V L S T A E M E E G E L D L I P R R R E E E D D L L G F R E N G R E N L E I G N G D L E E E G D G L V E K E I E E

```
gi|944 AGGEDTGKEENTGHEDAGEEDAGEEDA--GEEDAGEEDAEKEEGEKEDAGDDAGSDDGEE
          70          80          90          100          110
```

CV127. DVVVVVVAAAMVOEKD

gi | 944 DSTGGDEGEANAEDSKGSEKNDPADTYRQVVALLDKDTKVDHIQSEYLRNALNNDLQSEV
120 130 140 150 160 170gi | 944 RVPVVEAIGRIGDYSKIQGCFKSMGKDVKKVISEEEKKFKSCMSKKKSEYQCSSEDSFAAA
180 190 200 210 220 230gi|944 KSKLSPITSKI KSCVSSKGR
240 250

```
>>gi|170743 gi|170743|gb|AAB02788.1| HMW glutenin subunit (815 aa)
initn: 35 initl: 35 opt: 70 Z-score: 96.6 bits: 25.4 E(): 1.9
Smith-Waterman score: 70; 30.882% identity (64.706% similar) in 68 aa overlap
(2-68:11-76)
```

CV127. VVVTLVVVLSTAEMEE-GELDLIPRRREEEDDLLGFRENGRENLEIGNGDL

gi|170 MTKRLVLFAAVVVALVALTAAEGEASGQLQC--ERELQEHSLKACRQVVDQQLRDVSPEC
10 20 30 40 50

CV127. EEEGDGLVEKEIEEDVVVVVVAAMVQEKD

gi|170 QPVGGGPVARQYEQQVVVPPKGGSFYPGETTPPQQLQQSILWGIPALLRRYYLSVTSPQQ

gi|170 VSYPQGASSQRPGQGQEEYLTSPQQSGQWQQPGQGQSGYYPTSPQQSGQKQPGYYPTS

gi|170 PWQPEQLQQPTQGQQRQQPGQGQQLRQGQGQSGQGQPRYYPTSSQQPGQLQQLAQGQQ

gi|170 GQQPERGQQGQSGQGQQLGQGQQGQQPGQKQSGQGQQGYYPISPQQLGQGQSGQGQL

gi|170 GGYPTSPQQSGQGSGYYPTSAQQPGQLQQSTQEQLGQEQQDQQSGQGRQGQSGQRQQ

gi|170 DQSGQGQPGQRQPGYYSTSPQQLGGQGPRIYYPTSPQQPGQEQQPRQLQQPEQGQGGQQ

gi|170 PEQGQGGQQRQGEQGQGPQGQGGQGPQGQPGYYPSTSPQSSGQGPYYPSTSPQSSGQ

```

gi|170 LQQPAQGQQPGQEQQGQQPGGQQPGGQQPGYYPTSPQQSGQEQQLEQWQQSGGQPGHY
      480      490      500      510      520      530

gi|170 PTSPLQPGGQPGYYPTSPQQIGGQQPGQLQQPTQGQQGQQPGGQQGQQPGEGQQGQQ
      540      550      560      570      580      590

gi|170 PGQGQQPGGQPGYYPTSLQQSGGQQPGQWQQPGGQQPGYYPTSSLQPEQQGQGYPTS
      600      610      620      630      640      650

gi|170 QQQPGGQPGQWQQSGGQQGYYPTSPQQSGGQQPGQWLQPGQWLQSGYYLTSPQQLG
      660      670      680      690      700      710

gi|170 QGQQPRQLQPRGQQGYYPTSPQQSGGQQLGQQGQGYPTSPQQSGGQQGYDSPYHV
      720      730      740      750      760      770

gi|170 SAEHQAASLKVAKAQQLAQLPAMCRLEGGDALLASQ
      780      790      800      810

>>gi|21743 gi|21743|emb|CAA43331.1| high molecular weigh (830 aa)
  initn: 35 initl: 35 opt: 70 Z-score: 96.5 bits: 25.4 E(): 1.9
Smith-Waterman score: 70; 30.882% identity (64.706% similar) in 68 aa overlap
(2-68:11-76)

              10              20              30              40              50
CV127.          VVVTLVVVLSTAEMEE-GELDLIPRRREEEDDLLGFRENGRENLEIGNDDL
              ::. ::.::.::.::.::.::.::.::.::.::.::.::.::.::.::.
gi|217 MTKRLVLFVAVVVALVALTAAEGEASGQLQC--ERELQEHSLKACRQVVDQQLRDVSPEC
              10              20              30              40              50

              60              70              80
CV127. EEEGDGLVEKEIEEDVVVVVVAAAMVQEKD
              . : : : . . : . : : :
gi|217 QPVGGGPVARQYEQQVVVPPKGGSFYPGETTPPQQLQQSILWGIPALLRRYYLSVTSPQQ
              60              70              80              90              100              110

gi|217 VSYPGQASSQRPGQGQQPGGQQQEYYLTSPQQSGQWQQPGGQAGYYPTSPQQSGQEQP
              120              130              140              150              160              170

gi|217 GYYPTSPWQPEQLQQPTQGQQRQQPGGQQQLRQGQQGQQSGGQPRYYPTSSQQPGQLQQ
              180              190              200              210              220              230

gi|217 LAQGQQGQQPERGQQGQQSGGQQQLGQQGQQGQQPGQKQQSGGQQGYYPISPQQLGQGQQ
              240              250              260              270              280              290

gi|217 SGQGQLGYPTSPQQSGGQSGYYPTSAQQPGQLQQSTQEQQLGQEQQDQQSGQGRQGQQ
              300              310              320              330              340              350

gi|217 SGQRQQDQQSGGQQPGQRQPGYYSTSPQQLGQGQPRYYPTSPQQPGQEQQPRQLQQPEQ
              360              370              380              390              400              410

gi|217 GQQGQQPEQGQQGQQPGQGEQGQQPGGQQGQQPGGQQPGYYPTSPQQSGGQPGYYPTS
              420              430              440              450              460              470

gi|217 PQQSGQLQQPAQGQQPGQEQQGQQPGGQQGQQPGGQQPGGQQPGYYPTSPQQSGQEQQ
              480              490              500              510              520              530

gi|217 LEQWQQSGGQPGHYPTSPQLQPGGQPGYYPTSPQQIGGQQPGQLQQPTQGQQGQQPGQ
              540              550              560              570              580              590

gi|217 GQQGQQPGGQQGQQPGGQQPGGQQPGGQPGYYPTSLQQSGGQQPGQWQQPGGGLPGYYPTS
              600              610              620              630              640              650

```


gi|217 SLQPEQGQGGYYPTSQQQPGQGPQPGWQQSGQGQGGYYPTSPQQSGQGQGPQWLQPGQ
660 670 680 690 700 710

gi|217 WLQSGYYLTSPQQLGQGQQPRQWLQPRQGQQGYYP TSPQSSGQGQQLGQGQQGYYP TSPQ

gi|217 QSGQGQGYDSPYHVS AEHQ AASLKVAKAQQLAAQLPAMCRLEGGDALLASQ
780 790 800 810 820 830

```
>>gi|2114497 gi|2114497|gb|AAB58417.1| 30 kDa salivary g (253 aa)
  initn: 39 initl: 39 opt: 63 Z-score: 94.5 bits: 23.4 E(): 2.5
Smith-Waterman score: 63; 36.842% identity (57.895% similar) in 38 aa overlap
(28-65:79-114)
```

CV127. VVVTLVVVVL

gi | 211 MKPLVKLFLFLFCLVGIVLSRPMPEDEEPPVAEGGDEETTTDDAGGDGGEENEGEHEHAGDED

10 20 30 40 50 60

CV127. STAE MEEGELD LIPRRR EEEDDL LGFREN GRENLE IGNGL EEEGDL VKEIEEDV VVV
gi|211 AGGEDT GKEENT GHEDAGEEDA--GEEDAGEEDA EKEE GEDAGDDAGSDDGEEDSTGG

CV127. VVAAAMVQEKD

gi|211 DEGEANAEDSKGSEKNDPADTYRQVVALLDKDKTKVDHIQSEYLRSA LNNDLQSEVRVPVV

gi|211 EAIGRIGDYSKIQGCFKSMGKDVKKVISEEEKFKSCMSKKKSEYQCSDFSAAAKSKLS
180 190 200 210 220 230gi|211 PITSKI KSCVSSKGR
240 250

```
>>gi|18637 gi|18637|emb|CAA33216.1| glycinin subunit G2 (485 aa)
  initn: 69 initl: 45 opt: 63 Z-score: 90.0 bits: 23.5 E(): 4.4
Smith-Waterman score: 63; 27.119% identity (67.797% similar) in 59 aa overlap
(3-56:157-214)
```

gi|186 MAKLVLSLCFLLFSGCFALREQAQQNECQIQKLNALKPDNRIESEGGFETWNPNNKPFQ
10 20 30 40 50 60gi|186 CAGVALSRCTLNRRNALRRPSYTNQPEIYIQQNGIFGMIFPGCPSTYQEPQESQQRGRS
70 80 90 100 110 120

CV127. VVVTLVVVLSTAEMEEGELDLIPRRR
 : : . . : . : . :
 gi|186 QRPQDRHQKVHRFREGDLIAVPTGVAWWMYNNEDTPVAVSIIIDTNSLEN-QLDQMPRRF
 130 140 150 160 170

```

          30          40          50          60          70          80
CV127. ----EEEDLLGFRENGRENLEIGNG-DLEEEDGLVEKEIEEDVVVVVAAAMVQEKD
          ..:::..:..:..:..:..:..:..:..:..:..:..:..:..:..:..:
gi|186  YLAGNQEQEFLKYQQQQQGGSSQKKGKQQEEENEGSNILSGFAPEFLKEAFGVNMQIVRN
          180          190          200          210          220          230

```

gi|186 LQGENEEEDSGAIVTVKGGRLVTAPAMRKPQQEEDDDDEEEQPQCVETDKGCQRQSKRSR
240 250 260 270 280 290gi|186 NGIDETICTMRLRQNIQNSSPDIYNPQAGSITTATSLDFPALWLLKLSAQYGSLRKNAM
300 310 320 330 340 350

gi|186 FVPHYTLNANSIIYALNGRALVQVNCNGERVFDGELQEGGVLIIVPQNFAVAAKSQSDNF

gi|186 EYVSFKTND RPSIGNLAGANSLLNALPEEVIQHTFNLK SQQARQVKNNNPF SFLVPPQES

gi|186_QRRAVA
480

```
>>gi|18609 gi|18609|emb|CAA26575.1| unnamed protein prod (485 aa)
  initn: 69 initl: 45 opt: 63 Z-score: 90.0 bits: 23.5 E(): 4.4
Smith-Waterman score: 63; 27.119% identity (67.797% similar) in 59 aa overlap
(3-56:157-214)
```

gi|186 MAKLVLSLCFLLFSGCFALREQAQQNECQIQKLNALKPGNRIESEGFIETWNPNNKPFQ

gi|186 CAGVALSRCTLNRNALRRPSYTNQPEIYIQQNGIGFMIFGCPSTYQEPQESQQRGRS

CV127. VVVTLVVVLSTAEMEEGELDLIPRRR

gi|186 QRPQDRHQKVHRFREGDLIAVPTGVAWWMYNNEDTPVVAVSIIDTNSLEN-QLDQMPRRF
130 140 150 160 170

CV127. ----EEEDDLLGFRENGRENLEIGNG-DLEEEGDGLVEKEIFEEDVVVVVAAAMVOEKD

gi|186 YLAGNQEQEFLKYQQQQQGGSSQSKGKQEEENEGSNILSGFAPEFLKEAFGVNMQIVRN

gi|186 LQGENEEEDSGAIVTVKGLRVTAPAMRKPQQEEDDDDEEEQPQCVETDKGCQRQSKRSR

gi | 186 NGIDETICTMRLRQNIQNSSPDIYNPQAGSITTATSLDFPALWLLKLSAQYGSLRKNAM
300 310 320 330 340 350

gi|186 FVPHYTLNANSIIYALNGRALVQVVCNGERVFDGELQEGGVLIVPQNFAVAAKSQSDNF

gi|186 EYVSFKTNDRPSIGNLAGANSLLNALPEEVIQHTFNLKSQQARQVKNNNPFSLVPPQES
420 430 440 450 460 470gi|186 QRRAVA
480

```
>>gi|83715936 gi|83715936|dbj|BAE54433.1| tropomyosin [O (284 aa)
initn: 40 initl: 40 opt: 60 Z-score: 89.3 bits: 22.6 E(): 4.8
Smith-Waterman score: 60; 35.849% identity (58.491% similar) in 53 aa overlap
(13-64:100-150)
```

gi|837 MDAIKKKMLAMKMERELATDKAEQTDQKLRTD EDNKNKLEEDLTTLQKKFSNLENDFDNA
10 20 30 40 50 60

```

CV127.          10      20      30
                VVVTLLVVVLSTAEMEEGELDLIPRRREEEDDLL
                :  :  :  :  :  :  :  :  :  :  :  :
gi|837 KEQLAEANQKLETSEKRVGECESEIAGLNRRIQLEEDLERSEERLSTAQTKLDEASKAA
                70      80      90      100      110      120

```

```

CV127.          40      50      60      70      80
                GFRENGRENLEI-GNGDLEEEGDGLVEKEIEEDVVVVVVAAAMVQEKD
                :  :  :  :  :  :  :  :  :  :  :  :
gi|837 DESERGRKVLNRSQGD-EERID-LLEKQLEEAKWIAEDADRFDEAARKLAITEVDLER
                130      140      150      160      170

```

```

gi|837 AEARLEAAEAKIVELEEEELKVVGNMKSLEISEQEASQREDSYEETIRDLTHRLKEAENR
                180      190      200      210      220      230

```

```

gi|837 AAEAERTVSKLQKEVDRLEDELLAEKERYKAISDELDTFAELAGY
                240      250      260      270      280

```

>>gi|169927 gi|169927|gb|AAA33947.1| beta-conglycinin-a1 (218 aa)
initn: 37 initl: 37 opt: 57 Z-score: 86.8 bits: 21.7 E(): 6.6
Smith-Waterman score: 57; 31.169% identity (58.442% similar) in 77 aa overlap
(5-80:49-117)

```

CV127.          10
                VVVTLLVVVLSTAEMEE
                :  :  :  :  :  :
gi|169 SKRAKSSSRKTISSEDKPFLNLSRDPISYKSKLKGKFFETPEKNPQLRDLDFLSIVDMNE
                10      20      30      40      50      60

```

```

CV127.          20      30      40      50      60      70
                GELDLIPRRREEEDLLGFRENGRENLE-IGNGDLEEEGDGLVEKEIEEDVVVVVVAAAM
                :  :  :  :  :  :  :  :  :  :  :  :
gi|169 GAL-LLPHFNSKAIVILVINE-GDANIELVG--LKEQQQ--EQQQEEQPLEVRKYRAE
                70      80      90      100      110

```

```

CV127.          80
                VQEKD
                :  :  :  :
gi|169 LSEQDIFVIPAGYPVVVNATSNLNFFAIGINAENNQRNFLAGSDNVISQIPSQVQELAF
                120      130      140      150      160      170

```

```

gi|169 PGSAQAVEKLLKNQRESYFVDAQPNEKEEGNKGKGPLSSILRAFY
                180      190      200      210

```

>>gi|83715928 gi|83715928|dbj|BAE54429.1| tropomyosin [S (284 aa)
initn: 40 initl: 40 opt: 58 Z-score: 86.4 bits: 22.0 E(): 6.9
Smith-Waterman score: 58; 34.426% identity (57.377% similar) in 61 aa overlap
(7-64:92-150)

```

gi|837 MDAIKKKMLAMKMEKEVATDKAEQTEQSLRDLEDAKNKTEEDLSTLQKKYSNLENDFDNA
                10      20      30      40      50      60

```

```

CV127.          10      20      30
                VVVTLLVVVLSTAEMEEGELDLIPRRREEED--DLL
                :  :  :  :  :  :  :  :  :
gi|837 NEQLTAANTNLEASEKRVACESEIQGLNRRIQLEEDLERSEERLSTAQSKLEDAASKAA
                70      80      90      100      110      120

```

```

CV127.          40      50      60      70      80
                GFRENGRENLEI-GNGDLEEEGDGLVEKEIEEDVVVVVVAAAMVQEKD
                :  :  :  :  :  :  :  :  :  :  :  :
gi|837 DESERGRKVLNRSQGD-EERID-LLEKQLEEAKWIAEDADRFDEAARKLAITEVDLER
                130      140      150      160      170

```

gi|837 AEARLEAAEAKIVELEEEELKVVGNNMKSLEISEQEASQREDSYEETIRDLTHRLKEAENR
180 190 200 210 220 230

gi|837 AAEAERTVSKLQKEVDRLEDELLAEKERYKSISDELDTFAELAGY
240 250 260 270 280

>>gi|2735100 gi|2735100|gb|AAD13646.1| ABA-1 allergen [A (134 aa)
initn: 25 initl: 25 opt: 54 Z-score: 85.8 bits: 20.8 E(): 7.5
Smith-Waterman score: 54; 26.829% identity (68.293% similar) in 41 aa overlap
(26-64:18-58)

10 20 30 40 50
CV127. VVVTLVVVLSTAEMEEGELDLIPRRREEEDDLLGFRENGRENLEIGNGDLE--EEGDGLV
.:...: : . : . : .
gi|273 HHFTLESSLDTHLKWLSQEQKDELLKMKKDGKTKKDLQAKILHYDELEGDA
10 20 30 40 50

60 70 80
CV127. EKEIEEDVVVVVAAAMVQEKD
.: :
gi|273 KKEATEHLKDGCREILKHVVGEEKEAELKKLKDSGASKEEVKAKVEEALHAVTDEEKKQY
60 70 80 90 100 110

gi|273 IADFGPACKKIFGAAHTSRRRR
120 130

>>gi|18639 gi|18639|emb|CAA33217.1| glycinin subunit G3 (481 aa)
initn: 65 initl: 41 opt: 60 Z-score: 85.7 bits: 22.7 E(): 7.6
Smith-Waterman score: 60; 27.419% identity (64.516% similar) in 62 aa overlap
(4-56:155-216)

gi|186 MAKLVLSLCFLLFSGCCFAFSFREQPQQNECQIQRLNALKPDNRIESEGGFIETWPNPNK
10 20 30 40 50 60

gi|186 PFQCAGVALSRCTLNRLNRRPSYTNAPQEIIYIQQSGIFGMIFGCPSTFEPPQKQGS
70 80 90 100 110 120

10 20
CV127. VVVTLVVVLSTAEME--EGELDLIPRRR-
: :...: . . :...: :...:
gi|186 SRPQDRHQKIYHFREGDLIAVPTGFAYWMYNNEDTPVAVSLIDTNSFQNLQDMPRRFY
130 140 150 160 170 180

30 40 50 60 70 80
CV127. ---EEEDDLLGF---RENGRENLEIGNGDLEEEGDGLVEKEIEEDVVVVVAAAMVQEKD
.....: . :...: . . :...: :...:
gi|186 LAGNQEQEFLQYQPKQKQGGTQSQKGRQEEENEGGSILSGFAPEFLEHAFVVDQRQIVR
190 200 210 220 230 240

gi|186 KLQGENEEEEKGAIVTVKGGLSVISPPTEEQQRPEEEEEKPCDEKDKHCQSQRNGIDE
250 260 270 280 290 300

gi|186 TICTMRLRHNIGQTSSPDIFNPQAGSITTATSLDFPALSWLKLSAQFGSLRKNAMEFVPHY
310 320 330 340 350 360

gi|186 NLNANSIIYALNGRALVQVNCNGERVFDGELQEGQVLIVPQNFAVAARSQSDNFEYVSF
370 380 390 400 410 420

gi|186 KTNDRPSIGNLAGANSLNALPEEVIQQTFFNLRRQQARQVKNNNPFSFLVPPKESQRRVV
430 440 450 460 470 480

gi|186 A

>>gi|15419048 gi|15419048|gb|AAK96889.1| tropomyosin [Cr (233 aa)
initn: 41 initl: 41 opt: 56 Z-score: 84.9 bits: 21.5 E(): 8.4
Smith-Waterman score: 56; 30.000% identity (56.667% similar) in 60 aa overlap
(7-64:41-99)

```

                                10          20
CV127.                          VVVTLVVVLSTAEMEERGE--LDLIPR
                                . : . : . : . : . : .
gi|154 NSARGFDTVNEKYQECQTKMEEAEKTASEAEQEIQSLNRRIQLLLEEDMERSEERLQTATE
              10          20          30          40          50          60

              30          40          50          60          70          80
CV127. RREEEDDLLGFRENGRENLEIGNGDLEEEGDGLVEKEIEEDVVVVVVAAAMVQEKD
              . : . : . : . : . : . : . : . : . : . : . : . : . : . : .
gi|154 KLEEASKAADESERNRKVLNNASEERTD-VLEKQLTEAKLIAEEADKKYDEAARKLA
              70          80          90          100          110

gi|154 ITEVDLERAEARLEAAEAKVLELEELKVVGNMKSLEISEQEASQREDSYEETIRDLTQ
              120          130          140          150          160          170

gi|154 RLKDAENRATEAERTVSKLQKEVDRLDELLAEKERYKAISDELDTFAELAGY
              180          190          200          210          220          230

```

>>gi|13183177 gi|13183177|gb|AAK15089.1|AF240006_1 7S gl (585 aa)
initn: 68 initl: 53 opt: 60 Z-score: 84.4 bits: 22.7 E(): 9
Smith-Waterman score: 61; 30.952% identity (64.286% similar) in 42 aa overlap
(36-77:270-311)

```

gi|131 MSCGGRCLCLVLFALLLASAVVASESKDPELKQCKHQCKAQQQISKEQKEACIQACKEYIR
              10          20          30          40          50          60

gi|131 QKHQGEHGRGGDILEEEVWNRKSPIERLRECSRGCEQQHGEQREECLRRCQEEFYQREKG
              70          80          90          100          110          120

gi|131 RQDDDNPTDPEKQYQQCRLQCRROGEGGFSREHCERRREEKYREQQGREGGRGEMYEGR
              130          140          150          160          170          180

CV127.                          VVVTLV

gi|131 EREEEQEEQGRGRIPYVFEDQHFITGFRTQHGRMRVLQKFTDRSELLRGIENYRVAILEA
              190          200          210          220          230          240

              10          20          30          40          50          60
CV127. VVLSTAEMEELDLIPRRREEEDDLLGFRENGRENLEIGNGDLEEEGDGLVEKEIEEDV
              . : . : . : . : . : . : . : . : . : . : . : . : . : . : .
gi|131 EPQTFIVPNHWDAESVVFVAKGRGTISLVRQDRRESLNKQGDILKINAGTTAYLINRDN
              250          260          270          280          290          300

              70          80
CV127. VVVVAAAMVQEKD
              . : . : . : . : . : . : . : . : . : . : . : . : . : . : .
gi|131 NERLVLAKLLQPVSTPGFELFFGAGGENPESFFKSFSDILEAAFNTRDRDLQRIFGQQ
              310          320          330          340          350          360

gi|131 RQGVIVKASEEQVRAMSRHEEGGIWPFGGESKGTINIYQQRPTHSNQYQQLHEVDASQYR
              370          380          390          400          410          420

gi|131 QLRDLDLTVSLANITQGAMTAPHYNSKATKIALVVDGEGYFEMACPHMSRSRGSYQGETR
              430          440          450          460          470          480

```

gi|131 GRPSYQRVASRLTRGTVVIIIPAGHPFVAVASSNQNLQVLCFEVNANNNEKFPLAGRNV
490 500 510 520 530 540

gi|131 NQLEREAKELAFGMPAREVEEVSRSQEEFFFKGPRQQQGRADA
550 560 570 580

80 residues in 1 query sequences

292011 residues in 1313 library sequences

Scomplib [34.26]

start: Sun Dec 14 22:14:24 2008 done: Sun Dec 14 22:14:24 2008

Total Scan time: 0.110 Total Display time: 0.040

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.3797_c3: amino acids 15-94

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.3797_c3_15-94, 80 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 4.8891 \pm 0.00381$; $\mu = 2.1682 \pm 0.199$

mean_var=33.1421 \pm 8.229, 0's: 12 Z-trim: 16 B-trim: 62 in 1/43

Lambda= 0.222784

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2

join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:

gi	emb	opt	bits	E(1313)	%_id	%_sim	alen
gi 1460058	gi 1460058 emb CAA46317. (211)	70	28.3	0.069	0.287	0.500	80
gi 61608445	gi 61608445 gb AA47076 (216)	70	28.3	0.07	0.287	0.500	80
gi 83754033	gi 83754033 pdb 2AS8 B (222)	70	28.3	0.072	0.287	0.500	80
gi 730036	gi 730036 sp P08176.2 PEP (320)	70	28.2	0.11	0.287	0.500	80
gi 21725572	gi 21725572 emb CAD3836 (222)	59	24.7	0.84	0.387	0.645	31
gi 21725570	gi 21725570 emb CAD3836 (222)	59	24.7	0.84	0.387	0.645	31
gi 21725566	gi 21725566 emb CAD3836 (222)	59	24.7	0.84	0.387	0.645	31
gi 21725562	gi 21725562 emb CAD3836 (222)	59	24.7	0.84	0.387	0.645	31
gi 21725564	gi 21725564 emb CAD3836 (222)	59	24.7	0.84	0.387	0.645	31
gi 21725568	gi 21725568 emb CAD3836 (222)	59	24.7	0.84	0.387	0.645	31
gi 21725560	gi 21725560 emb CAD3836 (222)	59	24.7	0.84	0.387	0.645	31
gi 62550933	gi 62550933 emb CAI7905 (326)	58	24.4	1.6	0.255	0.617	47
gi 4538529	gi 4538529 emb CAB39376. (81)	51	22.3	1.7	0.366	0.512	41
gi 32363125	gi 32363125 sp P81943_ (24)	43	19.9	2.6	0.333	0.722	18
gi 170734	gi 170734 gb AAA34287.1 (244)	53	22.8	3.5	0.355	0.677	31
gi 21773	gi 21773 emb CAA31685.1 u (307)	53	22.8	4.6	0.355	0.677	31
gi 21725580	gi 21725580 emb CAD3837 (222)	51	22.2	5	0.355	0.645	31
gi 21725576	gi 21725576 emb CAD3836 (222)	51	22.2	5	0.355	0.645	31
gi 21725574	gi 21725574 emb CAD3836 (222)	50	21.8	6.2	0.355	0.645	31
gi 21725578	gi 21725578 emb CAD3837 (222)	50	21.8	6.2	0.355	0.645	31
gi 19847822	gi 19847822 gb AAK27264 (306)	51	22.1	7.1	0.423	0.654	26
gi 1478292	gi 1478292 gb AAB35896.1 (25)	38	18.3	8.4	0.714	1.000	7

>>>CV127.3797_c3_15-94, 80 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|1460058 gi|1460058|emb|CAA46317.1| cysteine proteas (211 aa)

initn: 49 initl: 49 opt: 70 Z-score: 122.4 bits: 28.3 E(): 0.069

Smith-Waterman score: 70; 28.750% identity (50.000% similar) in 80 aa overlap (8-80:96-175)

gi|146 ETNACSINGNAPAEIDLQMRVTPTIRMQGGCGSCWAFSGVAATESAYLAYRNQSLDLAE
10 20 30 40 50 60

CV127. QLIKDLNRLWRNSFPKRLKAQRKLLFLYVQL
10 20 30
: : :

gi|146 QELVDCASQHGCHGDTIPRGIEYIQHNGVVQESYYRYVAREQSCRRPNAQRFGISNYCQI
70 80 90 100 110 120

```

40          50          60          70          80
CV127. IPPNS-KIKLKLRLSVLKNLIHPLQSQLNITTHRHMRSLIL-----PNLSRVNI
      ::. ::. : .. . .:: : :. :. :.
gi|146 YPPNANKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSN
      130      140      150      160      170      180

gi|146 AQGVVDYWIVRNSWDTNWGDNGYGYFAANIDL
      190      200      210

>>gi|61608445 gi|61608445|gb|AAx47076.1| Der p 1 allergen (216 aa)
  initn: 54 initl: 54 opt: 70 Z-score: 122.2 bits: 28.3 E(): 0.07
Smith-Waterman score: 70; 28.750% identity (50.000% similar) in 80 aa overlap
(8-80:89-168)

gi|616 NEIAXAKIDLRQMRTVTPIXMQGGCGSCWALSGVAATESAYLAYGNXSLDLAEQELVDCA
      10      20      30      40      50      60

                        10      20      30
CV127.                  QLIKDLNRLWRNSFPKRLKAQRKLLFLYVQLIPPN-SK
                        : . :.. .: :.:: . : :. :. :.
gi|616 SQHGCHGDTIPRGIEYIQHNGVVQESYYRYVAREQSCRRPNAQRFGISNYCQIYPPNVNK
      70      80      90      100      110      120

      40      50      60      70      80
CV127. IKLKLRLSVLKNLIHPLQSQLNITTHRHMRSLIL-----PNLSRVNI
      :. : .. . .:: : :. :. :.
gi|616 IREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSN AQGVVDYW
      130      140      150      160      170      180

gi|616 IVRNSWDTNWGDNGYGYFAANIDLMMIEEYPYVVIL
      190      200      210

>>gi|83754033 gi|83754033|pdb|2AS8|B Chain B, Crystal St (222 aa)
  initn: 49 initl: 49 opt: 70 Z-score: 121.9 bits: 28.3 E(): 0.072
Smith-Waterman score: 70; 28.750% identity (50.000% similar) in 80 aa overlap
(8-80:95-174)

gi|837 TNACSINGNAPAEIDLRQMRTVTPIRMQGGCGSCWAFSGVAATESAYLAYRQQSLDLAEQ
      10      20      30      40      50      60

                        10      20      30
CV127.                  QLIKDLNRLWRNSFPKRLKAQRKLLFLYVQLI
                        : . :.. .: :.:: . : :.
gi|837 ELVDCASQHGCHGDTIPRGIEYIQHNGVVQESYYRYVAREQSCRRPNAQRFGISNYCQIY
      70      80      90      100      110      120

      40      50      60      70      80
CV127. PPNS-KIKLKLRLSVLKNLIHPLQSQLNITTHRHMRSLIL-----PNLSRVNI
      ::. ::. : .. . .:: : :. :. :.
gi|837 PPNANKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSN
      130      140      150      160      170      180

gi|837 QGVVDYWIVRNSWDTNWGDNGYGYFAANIDLMMIEEYPYVVIL
      190      200      210      220

>>gi|730036 gi|730036|sp|P08176.2|PEPT1_DERPT Peptidase (320 aa)
  initn: 54 initl: 54 opt: 70 Z-score: 118.8 bits: 28.2 E(): 0.11
Smith-Waterman score: 70; 28.750% identity (50.000% similar) in 80 aa overlap
(8-80:193-272)

gi|730 MKIVLAIASLLALSAVYARPSSIKTFEYKYKAFNKSyatFEDEEAARKNFLESVKYVQSN
      10      20      30      40      50      60

```


gi|730 GGAINHLSDSLDEFKNRFLMSAEAFEHLKTQFDLNAETNACSSINGNAPAEIDLQMRVT
70 80 90 100 110 120gi|730 TPIRMQGGCGSCWAFSGVAATESAYLAYRNQSLDLAEQELVDCASQHGCHGDTIPRGIEY
130 140 150 160 170 180

CV127. QLIKDLNRILWRNSFPKRLKAQRKLLFLYVQLIPPN-SKIKLKLRSVLKNLIHP

gi|730 IQHNGVVQESYYRYVAREQSCRRPNAQRFGISNYCQIYPPNVNKKIREALAQTHTSAIAVII
190 200 210 220 230 240

CV127. LQSQLNITTHRHMRSLIL-----PNLSRVNI

gi|730 GIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNAGVDYWIVRNSWDTNWGDNGYG
250 260 270 280 290 300

```
gi|730 YFAANIDLMMIEEYPYVVIL
          310          320
```

```
>>gi|21725572 gi|21725572|emb|CAD38367.1| unnamed protei (222 aa)
  initn: 54 initl: 54 opt: 59 Z-score: 102.8 bits: 24.7 E(): 0.84
Smith-Waterman score: 59; 38.710% identity (64.516% similar) in 31 aa overlap
(17-46:104-134)
```

gi|217 TNAC SINGNAPAEIDLRQMQTVP I R M Q G G C S C W A F S G V A A T E S A Y L A V R N Q S L D L A E Q
10 20 30 40 50 60

CV127. QLIKDLNRILWRNSFPKRLKAQRKLLFLYVQLI

gi|217 ELVDCANQHGHGDTIPRGIEYIQHNGVVQESYYRYVAQEQSCRRPNAQRFGISNYCQIY
70 80 90 100 110 120

CV127. PPN-SKIKLKLLRSVLKNLIHPLQSQLNITTHRHMRSLILPNLSRVNI
::: :::: : ..

gi|217 PPNVNKIREALQTHSAIAVIIIGIKDLDAFRHYDGETIIQEDNGYQTNYHAVNIVGYSNA
130 140 150 160 170 180

gi|217 QGV¹DY²WIV³RNS⁴WDTN⁵WGD⁶NGY⁷GYFA⁸ANID⁹LMMIE¹⁰EYPY¹¹VVIL¹²

```
>>gi|21725570 gi|21725570|emb|CAD38366.1| unnamed protei (222 aa)
  initn: 54 initl: 54 opt: 59 Z-score: 102.8 bits: 24.7 E(): 0.84
Smith-Waterman score: 59; 38.710% identity (64.516% similar) in 31 aa overlap
(17-46:104-134)
```

gi|217 TNACSINGNAPAEIDLRQMETVTPIRMQGGCGSCWAFSGVAATESAYLAVRNQSLDLAEQ
10 20 30 40 50 60

CV127. QLIKDLNRILWRNSFPKRLKAQRKLLFLYVQLI
 10 20 30
 .: .: .: .: .

gi|217 ELVDCANQHGHGCHDGTIPRGIEYIQHNGVVQESYYRYVAEEQSCRRPNAQRFGISNYCQIY
70 80 90 100 110 120

```

      40      50      60      70      80
CV127. PPN-SKIKLKLRSVLKNLIHPLQSQLNITTHRHMRSLILPNLSRVNI
      ::: :::: : ..
gi|217 PPNVNKIREALAQTHSAIAVIIGIKDLDAFRHYDGQTIIQEDNGYQTNHVNIVGYNSA
      130      140      150      160      170      180

gi|217 QGVYDWIVRNSWDTNWGDNGYGYFAANIDLMMIEEYPYVVIL
      190      200      210      220

>>gi|21725566 gi|21725566|emb|CAD38364.1| unnamed protei (222 aa)
  initn: 54 initl: 54 opt: 59 Z-score: 102.8 bits: 24.7 E(): 0.84
Smith-Waterman score: 59; 38.710% identity (64.516% similar) in 31 aa overlap
(17-46:104-134)

gi|217 TNACSINGNAPAEIDLRQMOTVTPIRMQGGCGSCWAFSGVAATESAYLAVRNQSLDLAEQ
      10      20      30      40      50      60

      10      20      30
CV127. QLIKDLNRIILWRNSFPKRLKAQRKLLFLYVQLI
      .: :::: . : :.
gi|217 ELVDCASQHGCHGDTIPEGIEYIQHNGVVQESYYRYVAEEQSCRRPNAQRFGISNYCQIY
      70      80      90      100      110      120

      40      50      60      70      80
CV127. PPN-SKIKLKLRSVLKNLIHPLQSQLNITTHRHMRSLILPNLSRVNI
      ::: :::: : ..
gi|217 PPNVNKIREALAQTHSAIAVIIGIKDLDAFRHYDGQTIIQQDNGYQTNHVNIVGYNSA
      130      140      150      160      170      180

gi|217 QGVYDWIVRNSWDTNWGDNGYGYFAANIDLMMIEEYPYVVIL
      190      200      210      220

>>gi|21725562 gi|21725562|emb|CAD38362.1| unnamed protei (222 aa)
  initn: 54 initl: 54 opt: 59 Z-score: 102.8 bits: 24.7 E(): 0.84
Smith-Waterman score: 59; 38.710% identity (64.516% similar) in 31 aa overlap
(17-46:104-134)

gi|217 TNACSINGNAPAEIDLRQMRTVTTIRMQGGCGSCWAFSGVAATESAYLAVRNQSLDLAEQ
      10      20      30      40      50      60

      10      20      30
CV127. QLIKDLNRIILWRNSFPKRLKAQRKLLFLYVQLI
      .: :::: . : :.
gi|217 ELVDCASQHGCHGDTIPQGIIEYIQHNGVVQESYYRYVAEEQSCRRPNAQRFGISNYCQIY
      70      80      90      100      110      120

      40      50      60      70      80
CV127. PPN-SKIKLKLRSVLKNLIHPLQSQLNITTHRHMRSLILPNLSRVNI
      ::: :::: : ..
gi|217 PPNVNKIREALAQTHSAIAVIIGIKDLDAFRHYDGETIIQQDNGYQTNHVNIVGYNSA
      130      140      150      160      170      180

gi|217 QGVYDWIVRNSWDTNWGDNGYGYFAANIDLMMIEEYPYVVIL
      190      200      210      220

>>gi|21725564 gi|21725564|emb|CAD38363.1| unnamed protei (222 aa)
  initn: 54 initl: 54 opt: 59 Z-score: 102.8 bits: 24.7 E(): 0.84
Smith-Waterman score: 59; 38.710% identity (64.516% similar) in 31 aa overlap
(17-46:104-134)

gi|217 TNACSINGNAPAEIDLRQMETVTPIRMQGGCGSCWAFSGVAATESAYLAVRNQSLDLAEQ
      10      20      30      40      50      60

```

```

                                10      20      30
CV127.                        QLIKDLNRILWRNSFPKRLKAQRKLLFLYVQLI
                                .: .::: . : :.
gi|217 ELVDCASQHGCHGDTIPQGIIEYIQHNGVVQESYYRYVAQEQSCRRPNAQRFGISNYCQIY
                                70      80      90      100     110     120

```

```

                                40      50      60      70      80
CV127. PPN-SKIKLKLLRSVLKNLIHPLQSQLNITTHRHMRSILPNLSRVNI
                                ::: :::: : ..
gi|217 PPNVKNKIREALAQTHSAIAVIIGIKDLDAFRHYDGETIIQQDNGYQNTYHAVNIVGYSNA
                                130     140     150     160     170     180

```

```

gi|217 QGV DYWIVRNSWDTNWGDNGYGYFAANIDLMMIEEYPYVVIL
                                190     200     210     220

```

>>gi|21725568 gi|21725568|emb|CAD38365.1| unnamed protei (222 aa)
 initn: 54 initl: 54 opt: 59 Z-score: 102.8 bits: 24.7 E(): 0.84
 Smith-Waterman score: 59; 38.710% identity (64.516% similar) in 31 aa overlap
 (17-46:104-134)

```

gi|217 TNACSINGNAPAEIDLRQMRTVTTIRMQGGCGSCWAFSGVAATESAYLAVRNQSLDLAEQ
                                10      20      30      40      50      60

```

```

                                10      20      30
CV127.                        QLIKDLNRILWRNSFPKRLKAQRKLLFLYVQLI
                                .: .::: . : :.
gi|217 ELVDCANQHGCHGDTIPRGIEYIQHNGVVQESYYRYVAEEQSCRRPNAQRFGISNYCQIY
                                70      80      90      100     110     120

```

```

                                40      50      60      70      80
CV127. PPN-SKIKLKLLRSVLKNLIHPLQSQLNITTHRHMRSILPNLSRVNI
                                ::: :::: : ..
gi|217 PPNVKNKIREALAQTHSAIAVIIGIKDLDAFRHYDGQTIIQQDNGYQNTYHAVNIVGYSNA
                                130     140     150     160     170     180

```

```

gi|217 QGV DYWIVRNSWDTNWGDNGYGYFAANIDLMMIEEYPYVVIL
                                190     200     210     220

```

>>gi|21725560 gi|21725560|emb|CAD38361.1| unnamed protei (222 aa)
 initn: 54 initl: 54 opt: 59 Z-score: 102.8 bits: 24.7 E(): 0.84
 Smith-Waterman score: 59; 38.710% identity (64.516% similar) in 31 aa overlap
 (17-46:104-134)

```

gi|217 TNACSINGNAPAEIDLRQMRTVTTIRMQGGCGSCWAFSGVAATESAYLAVRNQSLDLAEQ
                                10      20      30      40      50      60

```

```

                                10      20      30
CV127.                        QLIKDLNRILWRNSFPKRLKAQRKLLFLYVQLI
                                .: .::: . : :.
gi|217 ELVDCASQHGCHGDTIPEGIEYIQHNGVVQESYYRYVAQEQSCRRPNAQRFGISNYCQIY
                                70      80      90      100     110     120

```

```

                                40      50      60      70      80
CV127. PPN-SKIKLKLLRSVLKNLIHPLQSQLNITTHRHMRSILPNLSRVNI
                                ::: :::: : ..
gi|217 PPNVKNKIREALAQTHSAIAVIIGIKDLDAFRHYDGQTIIQEDNGYQNTYHAVNIVGYSNA
                                130     140     150     160     170     180

```

```

gi|217 QGV DYWIVRNSWDTNWGDNGYGYFAANIDLMMIEEYPYVVIL
                                190     200     210     220

```

>>gi|62550933 gi|62550933|emb|CAI79052.1| putative LMW-g (326 aa)
initn: 46 initl: 46 opt: 58 Z-score: 97.8 bits: 24.4 E(): 1.6
Smith-Waterman score: 58; 25.532% identity (61.702% similar) in 47 aa overlap
(15-60:107-153)

gi|625 MSHIPGLERPSQQQPLRPQQTLSHHHHQQPIQQQPQQFFQQQPCSQQQQQPPLSQQQQPP
10 20 30 40 50 60

CV127. QLIKDLNRILWRNSFPKRLKAQRKLLFL
10 20
: : : : : :

gi|625 FSQQQQPPFSQQQQPVLPQQPSFSQQQLPFPFSQQQQPPFSQQQQPVLPQQPSFSQQQLPP
70 80 90 100 110 120

CV127. YVQLIPPNSKIKLKLRLSVLKNLIHP-LQSQLNITTHRHMRSLILPNLSRVNI
30 40 50 60 70 80
: : : : : : : : : : : :

gi|625 FSQQQLPFPFSQQQQPVLLQQQIPFVHPSILQQQLNPCKVFLQQQCSPVAMPQSLARSQMLQQ
130 140 150 160 170 180

gi|625 SSCHVMQQCCQQLPQIPQQSRYEAIRAIVYSIILQEQQQVQGSIQTQQQQPQQQLGQCVS
190 200 210 220 230 240

gi|625 QPQQQSQQQLGQCSFQQPQQQLGQQPQQQQIPQGIFLQPHQISQLEVMTSIALRTLPT
250 260 270 280 290 300

gi|625 MCGVNVPLYSSTTIMPFSIGTGVGGY
310 320

>>gi|4538529 gi|4538529|emb|CAB39376.1| Cop c1 allergen (81 aa)
initn: 45 initl: 45 opt: 51 Z-score: 97.5 bits: 22.3 E(): 1.7
Smith-Waterman score: 51; 36.585% identity (51.220% similar) in 41 aa overlap
(32-72:17-56)

CV127. QLIKDLNRILWRNSFPKRLKAQRKLLFLYVQLIPPNSKIKLKLRLSVLKNLIHPLQSQLN
10 20 30 40 50 60
: : : : : :

gi|453 RFLPSSSHLNPQHLPWLVPVAPVLLLPVLPQ-LKPVAHPLLLPL
10 20 30 40

CV127. ITTHRHMRSLILPNLSRVNI
70 80
: : : : :

gi|453 DTTTLHMPPLLLQLQLPPLLSQGNPACSPKWLQLLVP
50 60 70 80

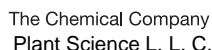
>>gi|32363125 gi|32363125|sp||P81943_3 [Segment 3 of 4] (24 aa)
initn: 31 initl: 31 opt: 43 Z-score: 93.9 bits: 19.9 E(): 2.6
Smith-Waterman score: 43; 33.333% identity (72.222% similar) in 18 aa overlap
(8-25:4-21)

CV127. QLIKDLNRILWRNSFPKRLKAQRKLLFLYVQLIPPNSKIKLKLRLSVLKNLIHPLQSQLN
10 20 30 40 50 60
: : : : : :

gi|323 IYARVLWVGNTTQKLEWIRSLHDY
10 20

CV127. ITTHRHMRSLILPNLSRVNI
70 80

>>gi|170734 gi|170734|gb|AAA34287.1| gamma gliadin B-III (244 aa)
initn: 37 initl: 37 opt: 53 Z-score: 91.6 bits: 22.8 E(): 3.5



BASF Reg. Doc. No. 2008/7019371

gi|170 PQQPFPLQPQQSFLWQSQQPFLLQPPQQPSPQPQQVVQIISPATPTTIPSAGKPTSAPFPQ

gi|170 QQQQHQLAQQQIPVVQPSILQQLNPCKVFLQQQCSPVAMPQRLARSQMLQQSSCHVMQQ
70 80 90 100 110 120

CV127. OLIKDLNRILWRNSFPKRLKAORKLLFLYVOLIPP

gi|170 QCCQQLPQIPQPSRYQAIRAIYSIILQEQQQVQGSIQSQQQQPQQLGQCVSQPQQQSQQ

CV127. 40 50 60 70 80
 NSKIKLKLRSVLKNIHPLO-SOLNITTHRHMRSLILPNLSRVNI

gi|170 QLGGQPQQQLAQGTFLQPHQIAQLEVMTSIALR--ILPTMCSVNVPLYRTTTSVPFGVG

gi|170 TGVGAY
240

```
>>gi|21773 gi|21773|emb|CAA31685.1| unnamed protein prod (307 aa)
  initn: 37 initl: 37 opt: 53 Z-score: 89.7 bits: 22.8 E(): 4.6
Smith-Waterman score: 53; 35.484% identity (67.742% similar) in 31 aa overlap
(51-80:259-287)
```

gi|217 MKTFLV FALLAVAATSAIAQMETRCIPGLERPWQQQPLPPQQTFPQQPLFSQQQQQQQLFP
10 20 30 40 50 60gi|217 QQPSFSQQQPPFWQQQPPFSQQQPILPQQPPFSQQQQLVLPQQPPFSQQQQPVLPPQQSP
70 80 90 100 110 120gi|217 FPQQQQQHQQLVQQQIPVVQPSILQQLNPKVFLQQQCSPVAMPQRLARSQMLQQSSCHV
130 140 150 160 170 180

CV127. OLIKDLNRILWRNSFPKRLKAORKLLFLYVOL

gi|217 MQQQCCQQLPQIPQQSRYEAIRAIIYSIILQEQQVVGSGIQSQQQQPQQLGQCVSQPQQQ

CV127 40 50 60 70 80
 IPPNSKIKLKLIRSVLKNLIHPLO-SQNLITTHRHMRSLILPNLSRVNT

gi|217 SQQQLGQQPQQQQLAQGTFLQPHQIAQLEVMTSIALR--ILPTMCSVNVPLYRTTTSVPF

gi | 217 GVG TGV GAY
300

```
>>gi|21725580 gi|21725580|emb|CAD38371.1| unnamed protei  (222 aa)
  initn: 42 initl: 42 opt: 51 Z-score: 88.9 bits: 22.2 E(): 5
Smith-Waterman score: 51; 35.484% identity (64.516% similar) in 31 aa overlap
(17-46:104-134)
```

gi | 217 TNACSGINGNAPASIDLRQMRTVTITIRMQGGCGSCWAFSGVAATESAYLAVRNQSLDLAEQ
10 20 30 40 50 60

```

                                10      20      30
CV127.                        QLIKDLNRILWRNSFPKRLKAQRKLLFLYVQLI
                                .: .:.: . : :.
gi|217 ELVDCASQHGCHGDTIPQGIEYIQHNGVVQESYYRYVAEEQSCRRPNADRFGISNYCQIY
                                70      80      90      100     110     120

```

```

                                40      50      60      70      80
CV127. PPN-SKIKLKLLRSVLKNLIHPLQSQLNITTHRHMRSILPNLSRVNI
                                ::: .:.: : ..
gi|217 PPNVNKIQEALAQTHSAIAVIIGIKDLDAFRHYDGETIIQQDNGYQTNHAYNIVGYNSA
                                130     140     150     160     170     180

```

```

gi|217 QGVYDWIVRNSFDTNWGDNGYGYFAANIDLMMIEEYPYVVIL
                                190     200     210     220

```

>>gi|21725576 gi|21725576|emb|CAD38369.1| unnamed protei (222 aa)
 initn: 42 initl: 42 opt: 51 Z-score: 88.9 bits: 22.2 E(): 5
 Smith-Waterman score: 51; 35.484% identity (64.516% similar) in 31 aa overlap
 (17-46:104-134)

```

gi|217 TNACSINGNAPASIDLRQMRTVTTIRMQGGCGSCWAFSGVAATESAYLAVRNQSLDLAEQ
                                10      20      30      40      50      60

```

```

                                10      20      30
CV127.                        QLIKDLNRILWRNSFPKRLKAQRKLLFLYVQLI
                                .: .:.: . : :.
gi|217 ELVDCASQHGCHGDTIPQGIEYIQHNGVVQESYYRYVAEEQSCRRPNADRFGISNYCQIY
                                70      80      90      100     110     120

```

```

                                40      50      60      70      80
CV127. PPN-SKIKLKLLRSVLKNLIHPLQSQLNITTHRHMRSILPNLSRVNI
                                ::: .:.: : ..
gi|217 PPNVNKIQEALAQTHSAIAVIIGIKDLDAFRHYDGETIIQQDNGYQTNHAYNIVGYNSA
                                130     140     150     160     170     180

```

```

gi|217 QGVYDWIVRNSWDTNWGDNGYGYFAANIDLMMIEEYPYVVIL
                                190     200     210     220

```

>>gi|21725574 gi|21725574|emb|CAD38368.1| unnamed protei (222 aa)
 initn: 42 initl: 42 opt: 50 Z-score: 87.2 bits: 21.8 E(): 6.2
 Smith-Waterman score: 50; 35.484% identity (64.516% similar) in 31 aa overlap
 (17-46:104-134)

```

gi|217 TNACSINGNAPASIDLRQMRTVTTIRMQGGCGSCWAFSGVAATESAYLAVRNQSLDLAEQ
                                10      20      30      40      50      60

```

```

                                10      20      30
CV127.                        QLIKDLNRILWRNSFPKRLKAQRKLLFLYVQLI
                                .: .:.: . : :.
gi|217 ELVDCASQHGCHGDTIPEGIEYIQHNGVVQESYYRYVAEQSCRRPNADRFGISNYCQIY
                                70      80      90      100     110     120

```

```

                                40      50      60      70      80
CV127. PPN-SKIKLKLLRSVLKNLIHPLQSQLNITTHRHMRSILPNLSRVNI
                                ::: .:.: : ..
gi|217 PPNVNKIEEALAQTHSAIAVIIGIKDLDAFRHYDGTIIQEDNGYQTNHAYNIVGYNSA
                                130     140     150     160     170     180

```

```

gi|217 QGVYDWIVRNSWDTNWGDNGYGYFAANIDLMMIEEYPYVVIL
                                190     200     210     220

```

```
>>gi|21725578 gi|21725578|emb|CAD38370.1| unnamed protei (222 aa)
  initn: 42 initl: 42 opt: 50 Z-score: 87.2 bits: 21.8 E(): 6.2
Smith-Waterman score: 50; 35.484% identity (64.516% similar) in 31 aa overlap
(17-46:104-134)
```

gi|217 TNACSINGNAPASIDLRQMRVTVTIRMQGGCGSCWAFSGVAATESAYLAVRNQSLDLAEQ

CV127. QLIKDLNRILWRNSFPKRLKAQRKLLFLYVQLI

gi|217 ELVDCASQHGHGDTIPEGIEYIQHNGVVQESYYRYVAQE_QSCRRPNADRFGISNYCQIY
70 80 90 100 110 120

CV127. PPN-SKIKLKLRLSVLKNLIHPLOSOLNITTHRHMRSLILPNLSRVNI

gi|217 PPNVNKIEEALAQTHSAIAVIIIGIKDLDAFRHYDGGTTIIQEDNGYQTNYHAVNIVGYSNA
130 140 150 160 170 180

gi|217 QGVDYWIVRNSFDTNWGDNGYGYFAANIDLMMIEEYPYVVIL
190 200 210 220

```
>>gi|19847822 gi|19847822|gb|AAK27264.1| isoflavone redu (306 aa)
  initn: 51 initl: 51 opt: 51 Z-score: 86.2 bits: 22.1 E(): 7.1
Smith-Waterman score: 51; 42.308% identity (65.385% similar) in 26 aa overlap
(21-46:25-50)
```

CV127. QLIKDLNRILWRNSFPKRLKAQRKLLFLYVQLIPPNSKIKLKLRLSVLKNLIHPLQ

gi|198 MGGSRVLIIGGTGYIGRHVTNASLAQGHPFTLLVREITPSNPEKAQLLESFTSKGATLVQ

CV127. SOLNITTHRHMRSLILPNLSRVNI

gi | 198 GSIDDHASLVAALKKVDVVISTLGAPQIADQFNLIKAIKEVGTIKRFFPSEFGNDVDKHH
70 80 90 100 110 120gi|198 AVEPMKSMFDLKI¹³⁰KLRR¹⁴⁰TIEAEGIPHTYV¹⁵⁰VP¹⁶⁰HC¹⁷⁰FAGYFLTNLAQLGLAAPP¹⁸⁰RDKIV¹⁹⁰IYGD²⁰⁰

gi|198 GTTKAVYMKEEDIGTFTIKAVDDPRTLNKTLYLKPPANTISTNDLVALWEAKIGKTLEKV

gi|198 YLSEEQVLKLLQDTPFPGTFMVSI FHTIYVKGDQTNFQIGPDGVEASALYPDVKYTTVEE
250 260 270 280 290 300

qi|198 YISAFV

```
>>gi|1478292 gi|1478292|gb|AAB35896.1| 18 kda major alle (25 aa)
  initn: 38 initl: 38 opt: 38 Z-score: 84.9 bits: 18.3 E(): 8.4
Smith-Waterman score: 38; 71.429% identity (100.000% similar) in 7 aa overlap
(29-35:9-15)
```

CV127. QLIKDLNRILWRNSFPKRLKAQRKLLFLYVQLIPPNSKIKLKLLRSVLKNLIHPLQSQNLN

gi|147 G V Y T F E N E Y T S L I P P X I L F K A F V X D
10 20



The Chemical Company
Plant Science L. L. C.

BASF Reg. Doc. No. 2008/7019371

70 80
CV127. ITTHRHMRSLILPNLSRVNI

80 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:26:31 2008 done: Sun Dec 14 22:26:31 2008
Total Scan time: 0.100 Total Display time: 0.040

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.2700_d3

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.2700_d3, 60 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 5.7665 \pm 0.00468$; $\mu = -3.6248 \pm 0.241$
mean_var=71.5377 \pm 17.471, 0's: 20 Z-trim: 26 B-trim: 69 in 1/43
Lambda= 0.151638

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2

join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:

			opt	bits	E(1313)	%_id	%_sim	alen
gi 75317968	gi 75317968 sp O22116 O	(373)	104	31.4	0.011	0.341	0.659	44
gi 21783	gi 21783 emb CAA30570.1 u	(356)	89	28.1	0.1	0.342	0.632	38
gi 62550933	gi 62550933 emb CAI7905	(326)	87	27.6	0.13	0.340	0.642	53
gi 886967	gi 886967 emb CAA59340.1	(276)	86	27.4	0.13	0.356	0.622	45
gi 21930	gi 21930 emb CAA44473.1 L	(285)	84	26.9	0.18	0.378	0.622	45
gi 76782247	gi 76782247 gb ABA54897	(134)	79	25.7	0.19	0.291	0.527	55
gi 75219081	gi 75219081 sp O22108 O	(285)	81	26.3	0.28	0.370	0.522	46
gi 21773	gi 21773 emb CAA31685.1 u	(307)	81	26.3	0.29	0.370	0.522	46
gi 886965	gi 886965 emb CAA59339.1	(261)	79	25.8	0.35	0.348	0.565	46
gi 886963	gi 886963 emb CAA59338.1	(229)	73	24.5	0.77	0.347	0.592	49
gi 170708	gi 170708 gb AAA34274.1	(291)	74	24.7	0.81	0.206	0.667	63
gi 62484809	gi 62484809 emb CAI7890	(285)	73	24.5	0.93	0.400	0.520	50
gi 1168171	gi 1168171 gb AAB35353.1	(30)	57	20.6	1.5	0.357	0.571	28
gi 21926	gi 21926 emb CAA36063.1 u	(295)	69	23.7	1.8	0.444	0.741	27
gi 170730	gi 170730 gb AAA34285.1	(304)	68	23.4	2.1	0.347	0.571	49
gi 170722	gi 170722 gb AAA34281.1	(262)	67	23.2	2.1	0.291	0.673	55
gi 27818335	gi 27818335 gb AAO24900	(132)	63	22.2	2.2	0.444	0.519	27
gi 473876	gi 473876 gb AAA17741.1	(287)	67	23.2	2.3	0.291	0.673	55
gi 170702	gi 170702 gb AAA34272.1	(302)	67	23.2	2.4	0.271	0.667	48
gi 21765	gi 21765 emb CAA26385.1 u	(313)	67	23.2	2.5	0.291	0.673	55
gi 170718	gi 170718 gb AAA34279.1	(313)	67	23.2	2.5	0.291	0.673	55
gi 162805	gi 162805 gb AAA30431.1	(224)	65	22.7	2.5	0.356	0.578	45
gi 459292	gi 459292 gb AAB29137.1	(224)	65	22.7	2.5	0.356	0.578	45
gi 162931	gi 162931 gb AAA30480.1	(224)	65	22.7	2.5	0.356	0.578	45
gi 170710	gi 170710 gb AAA34275.1	(318)	67	23.2	2.5	0.291	0.673	55
gi 170716	gi 170716 gb AAA34278.1	(319)	67	23.2	2.5	0.291	0.673	55
gi 1063270	gi 1063270 dbj BAA11251.	(279)	64	22.6	3.6	0.300	0.750	40
gi 21755	gi 21755 emb CAA25593.1 u	(286)	64	22.6	3.6	0.291	0.655	55
gi 21761	gi 21761 emb CAA26384.1 u	(286)	64	22.6	3.6	0.291	0.655	55
gi 170720	gi 170720 gb AAA34280.1	(286)	64	22.6	3.6	0.291	0.655	55
gi 162797	gi 162797 gb AAA30430.1	(224)	61	21.9	4.6	0.333	0.500	48
gi 741844	gi 741844 prf 2008179A m	(143)	58	21.1	4.9	0.323	0.613	31
gi 170732	gi 170732 gb AAA34286.1	(323)	61	21.9	6.4	0.350	0.525	40
gi 169969	gi 169969 gb AAA33964.1	(516)	61	22.0	9.6	0.333	0.741	27
gi 736002	gi 736002 emb CAA55977.1	(517)	61	22.0	9.6	0.333	0.741	27
gi 1228078	gi 1228078 emb CAA33034.	(190)	55	20.5	10	0.237	0.526	38
gi 162811	gi 162811 gb AAA30433.1	(190)	55	20.5	10	0.237	0.526	38

>>>CV127.2700_d3, 60 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

```
>>gi|75317968 gi|75317968|sp|O22116|O22116_WHEAT LMM glu (373 aa)
  initn: 225 initl: 86 opt: 104 Z-score: 136.9 bits: 31.4 E(): 0.011
Smith-Waterman score: 104; 34.091% identity (65.909% similar) in 44 aa overlap
(17-60:142-185)
```

gi|753 ASAVAQISQQQQPPPPFSQQQQPPFSQQQQPPFSQQQQSPFSQQQQPPFSQQQQPPFSQQ

gi|753 PLISQQQQLPFSQQQQPQFSQQQQPPYSQQQQPPYSQQQQPPFSQQQQPPFSQQQQPSFS
70 80 90 100 110 120

CV127. MFTSLFFFSVSLSLAHLSPPEWRRQQQQQHLLRSPSPPNHLLPPPNHHYQSPDS

gi|753 QQQQQPPFTQQQQPPFSQQSPISQQQQQQQQQQQPPFTQQQQPPFSQQPPISQQQQPPFS
130 140 150 160 170 180

CV127. PSHSP

gi|753 QQQQPPFSQQQQIPVIHPSVLQQLNPCMVFLQQQCIPVAMQRCLARSQMLQQSICHVMQR

gi|753 QCCQQLRQIPEQSRHESIRAI IYSIILQQQQQQQQQQQQQGGQSIIQYQQQPPQQLGQCV

gi|753 SQPQQQLQQQLGQQPQQQQLAHGTF LQPHKIAQLVMTSIALRTLPRMCSVNVPLYETTT
310 320 330 340 350 360gi|753 SVPLGVGIGVGVY
370

```
>>gi|21783|gi|21783|emb|CAA30570.1| unnamed protein prod (356 aa)
  initn: 286 initl: 86 opt: 89 Z-score: 119.5 bits: 28.1 E(): 0.1
Smith-Waterman score: 89; 34.211% identity (63.158% similar) in 38 aa overlap
(23-60:138-175)
```

gi|217 MKTFLVFALLALAAASAVAQISQQQQAPPFSQQQQPPFSQQQQPPFSQQQQSPFSQQQQQ

CV127. MFTSL

gi | 217 PPFAQQQQPPFSQQPPISQQQPPFSQQQQPQFSQQQQPPYSQQQQPPYSQQQQPPFSQQ

CV127. FFFSVSLSLAHLSPPEWRRQQQQQHHLLRSPSPPNHLLPPPNHHYQSPDSPSHSP

gi|217 QQPPFSQQQQQPPFTQQQQQQQQQPPFTQQQQPPFSQQPPISQQQQPPFLQQQRPPFSRQ
130 140 150 160 170 180gi|217 QQIPVIHPSVLQQLNPCKVFLQQQCIPVAMQRCLARSQMLQQSICHVMQQQCCQQLRQIP
190 200 210 220 230 240gi|217 EQSRHESIRAIISIIILQQQQQQQQQQQQQGQSIIQYQQQQPQQLGQCVSQPLQQLQQQ
250 260 270 280 290 300gi|217 LGQQPQQQLAHQIAQLVMTSIALRTLPTMCNVNVPLYETTTSVPLGVGIGVGIVY
310 320 330 340 350

>>gi|62550933 gi|62550933|emb|CAI79052.1| putative LMW-g (326 aa)
initn: 125 initl: 73 opt: 87 Z-score: 117.7 bits: 27.6 E(): 0.13
Smith-Waterman score: 87; 33.962% identity (64.151% similar) in 53 aa overlap
(14-60:53-105)

```

                                10      20
CV127.                          MFTSLFFFSVSLSLAHLSP
                                :. . . :
gi|625 MSHIPGLERPSQQQPLRPQQTLSHHHHQPIQQQPQQFFQQQPCSQQQQPPPLSQQQQPP
                                10      20      30      40      50      60

                                30      40      50      60
CV127. WRRQQQ-----QQQHLL-RSPSPPNHLLPPPNHHYQSPDSPSHSP
. .::: .::: .: .::: ..::: ...: : : : ...:
gi|625 FSQQQPPFSSQQQPVLPQQPSFSQQQLPPFSQQQPPFSSQQQPVLPQQPSFSQQQLPP
                                70      80      90      100     110     120

gi|625 FSQQQLPPFSSQQQPVLLQQQIPFVHPSILQQLNPCKVFLQQQCSPVAMPQSLARSQMLQQ
                                130     140     150     160     170     180

gi|625 SSCHVMQQCCQQLPQIPQQSRYEAIRAIVYSIILQEQQQVQGSITQQQQPQQQLGQCVS
                                190     200     210     220     230     240

gi|625 QPQQSSQQQLGQCSFQQPQQQLGQQPQQQQIPQGIFLQPHQISQLEVMTSIALRTLPT
                                250     260     270     280     290     300

gi|625 MCGVNVPLYSSTTIMPFSIGTGVGGY
                                310     320

```

>>gi|886967 gi|886967|emb|CAA59340.1| low molecular weig (276 aa)
initn: 112 initl: 69 opt: 86 Z-score: 117.6 bits: 27.4 E(): 0.13
Smith-Waterman score: 86; 35.556% identity (62.222% similar) in 45 aa overlap
(19-60:84-128)

```

gi|886 HIPSLEKPSQQQPLPLQQILWYHQQQPIQQQPQPFPPQPPCSQQQQPPPLSQQQPPFSSQ
                                10      20      30      40      50      60

                                10      20      30      40      50
CV127. MFTSLFFFSVSLSLAHLSPPEPWRQQ--QQQHLLRSPSPPNHLLPPPNHHYQS
                                :. .::: .::: .: .: .: .: .: .: .: .: .:
gi|886 QPPFSQQQLPILPQQPPFSSQQQPPFSSQQQPPFQQQQPPLLLQQPPFSSQQRPPFSSQQQQ
                                70      80      90      100     110     120

                                60
CV127. PDSPSHSP
: :. . :
gi|886 PVLPQQPPFSSQQQQPILPQQPPFSLHQQPVLPPQQIPYVQPSILQQLNPCKVFLQQQC
                                130     140     150     160     170     180

gi|886 SPVAMPQSLARSQMLWQSSCHVMQQCCQQLPRIPEQSRYSYDAIRAIYSIVLQEQQHGQG
                                190     200     210     220     230     240

gi|886 FNQPQQQPQQSVQGVSPQQQQKQLGQCSFQQPQQ
                                250     260     270

```

>>gi|21930 gi|21930|emb|CAA44473.1| LMW glutenin [Tritic (285 aa)
initn: 73 initl: 53 opt: 84 Z-score: 115.1 bits: 26.9 E(): 0.18
Smith-Waterman score: 84; 37.778% identity (62.222% similar) in 45 aa overlap
(20-60:47-91)

```

                                10      20      30
CV127.                        MFTSLFFFSVSLSLAHLSPWPWRRQQQQQHLL
                                .:  ::::: .
gi|219 MKTFLVFALLAVVATSAIAQMDTSCIPGLERPWQQQLPQQTFPQQPPFSQQQQQQPFP
                                10      20      30      40      50      60

```

```

                                40      50      60
CV127. RSPSPPNH--LLP--PPNHHYQSPDSPSHSP
                                .:: .: .: .: .: .: .: .: .:
gi|219 QQPSFSQQQPILPQGPPFPQQTQPVLPPQQSPFSQQQQQLLPPQQQQQLPQQQISIVQPSV
                                70      80      90      100     110     120

```

```

gi|219 LQQLNPKVFLQQQCSPVAIPQRLARSQMWQQSSCHVMQQQCCQQLSQIPEQSRDAIRA
                                130     140     150     160     170     180

```

```

gi|219 ITYSIILQEQQQGQSQQQQPQQSGQGVSSQQQSQQQLGQCSFQQPQQQLGQQPQQQQVQ
                                190     200     210     220     230     240

```

```

gi|219 QGTFLQPHQIAHLEVMTSIALRTLPTMCSNVNPLYSSTTSVPFGV
                                250     260     270     280

```

>>gi|76782247 gi|76782247|gb|ABA54897.1| hydrophobic see (134 aa)
 initn: 68 initl: 68 opt: 79 Z-score: 114.3 bits: 25.7 E(): 0.19
 Smith-Waterman score: 79; 29.091% identity (52.727% similar) in 55 aa overlap
 (4-57:16-63)

```

                                10      20      30      40
CV127.                        MFTSLFFFSVSLSLAHLSPWPWRRQQQQQHLLRSP-SPPNHLLPPPN
                                .:.:.: .: .: .: .: .: .: .: .:
gi|767 MGSKVVASVALLLSINILFISMVSSSSHYDPPP-----PPCYVPAPLTPPPSLSPPPS
                                10      20      30      40      50

```

```

                                50      60
CV127. HHYQSPDSPSHSP
                                .:.:.:
gi|767 LSPPPPSGPPCDLSVCLNILDGSPADDCALIADLVLDLEASVCLCIQLRVLGIVNLDLN
                                60      70      80      90      100     110

```

```

gi|767 LQLILNACGPSYPSNATCPRT
                                120     130

```

>>gi|75219081 gi|75219081|sp|O22108|O22108_WHEAT LMM glu (285 aa)
 initn: 52 initl: 52 opt: 81 Z-score: 111.5 bits: 26.3 E(): 0.28
 Smith-Waterman score: 81; 36.957% identity (52.174% similar) in 46 aa overlap
 (20-60:24-68)

```

                                10      20      30      40      50
CV127. MFTSLFFFSVSLSLAHLSPWPWRRQQQQQHLLRSPS-----PPNHLLPPPNHHYQ
                                .:  ::::: . .: .: .: .: .: .: .:
gi|752 RCIPGLERPWQQQLPQQTFPQQPLFSQQQQQQLFPQQPSFSQQQPPFWQQQPP-FSQQ
                                10      20      30      40      50

```

```

                                60
CV127. SPDSPSHSP
                                .:  .: .:
gi|752 QPILPQQPPFSQQQQLVLPQQSPFSQQQQQLLPPQQQQQLPQQQISIVQPSVLQQLNPCK
                                60      70      80      90      100     110

```

```

gi|752 VFLQQQCSPVAMPQRLARSQMWQQSSCHVMQQQCCQQLSQIPEQSRDAIRAITYSIILQ
                                120     130     140     150     160     170

```

```

gi|752 EQQQGFVQAQQQPQQSGQGVSSQQQSQQQLGQCSFQQPQQQLGQQPQQQQVQQGTFLQ
                                180     190     200     210     220     230

```

gi|752 PHQIAHLEVMTSIALRTLPMCSVNVPLYSSTTSVPFGVGTGVGAY
240 250 260 270 280

>>gi|21773 gi|21773|emb|CAA31685.1| unnamed protein prod (307 aa)
initn: 83 initl: 52 opt: 81 Z-score: 111.0 bits: 26.3 E(): 0.29
Smith-Waterman score: 81; 36.957% identity (52.174% similar) in 46 aa overlap
(20-60:47-91)

CV127. MFTSLFFFSVSLSLAHLSPWPWRQQQQQHLL
10 20 30
.: : : : : .
gi|217 MKTFLV FALLAVAATSAIAQMETRCIPGLERPWQQQLPQPQTFFPQQPLFSQQQQQLFP
10 20 30 40 50 60

CV127. RSPS----PPNHLLPPPNHHYQSPDSPSHSP
40 50 60
.: : : : : .
gi|217 QQPFSQQQPFPFQQQP-FSQQPILPQQPPFSQQQLVLPQQPPFSQQQPVLPPQQS
70 80 90 100 110

gi|217 PFPQQQQHQQQLVQQQIPVVQPSILQQLNPCKVFLQQQCSPVAMPQRLARSQMLQQSSCH
120 130 140 150 160 170

gi|217 VMQQCCQQLPQIPQQSRYEAIIRAIISIIILQEQQVQGSISQQQPQQLGQCVSQPQQ
180 190 200 210 220 230

gi|217 QSQQQLGQQPQQQLAQGTFLQPHQIAQLEVMTSIALRILPTMCSVNVPLYRTTTSVPFG
240 250 260 270 280 290

gi|217 VGTGVGAY
300

>>gi|886965 gi|886965|emb|CAA59339.1| low molecular weig (261 aa)
initn: 181 initl: 70 opt: 79 Z-score: 109.8 bits: 25.8 E(): 0.35
Smith-Waterman score: 92; 34.783% identity (56.522% similar) in 46 aa overlap
(19-60:32-77)

CV127. MFTSLFFFSVSLSLAHLSPWPWRQ---QQQQHLLRSPSPPNHLL
10 20 30 40
.: : : : : .
gi|886 HIPSLEKPLQQQLPLPQQILWYQQQPIQQQPQFPFPQQPQPCSQQQPPLSQQQPPFSQQ
10 20 30 40 50 60

CV127. PPPNHHYQSPDSPSHSP
50 60
.: : : : : .
gi|886 QPPFSQQQPILPQQPPFSQQQQQFPQQQPLLPQQPPFSQQQPFSQQQQPPFSQQQQ
70 80 90 100 110 120

gi|886 QPILLQQPPFSQHQQPVLPPQQIPSVQPSILQQLNPCKVFLQQQCSPVAMPQSLARSQML
130 140 150 160 170 180

gi|886 WQSSCHVMQQCCRQLPQIQEQRDYAIRAIISIVLQEQHGQGLNQPPQQPQQSVQG
190 200 210 220 230 240

gi|886 VSQPQQQKQLGQCSFQQPQQ
250 260

>>gi|886963 gi|886963|emb|CAA59338.1| low molecular weig (229 aa)
initn: 113 initl: 73 opt: 73 Z-score: 103.5 bits: 24.5 E(): 0.77

Smith-Waterman score: 76; 34.694% identity (59.184% similar) in 49 aa overlap
(20-60:43-91)

```

                                10      20      30
CV127.      MFTSLFFFSVSLSLAHLSPWPWRRQ---QQQHLLR
                                ... ..: .::: .
gi|886 FALIAVVATSTIAQMETSCLPGLRPWQQQPLPPQQLFPPQQPPFQQPPFSQQQPSFS
                                10      20      30      40      50      60

                                40      50      60
CV127. SPSPP---NHLLP-PPNHYYQSPDPSHSP
                                .::: .::: .: .:::
gi|886 QQQPPFSQQQPILPEPPFSLQQQPVLPQQSPFSQQQLVLPQQQQQLPQQQISIVQPSVL
                                70      80      90      100     110     120

gi|886 QQLNPCKVFLQQQCSPVAMPQRLARSQMWQSRCHVMQQCCQLSQIPEQSRDAICAI
                                130     140     150     160     170     180

gi|886 TYSIILQEQQQGFVQAQQQQPQQSGQGVQSQQQSQQQLGQCSFQQPQQ
                                190     200     210     220

```

>>gi|170708 gi|170708|gb|AAA34274.1| gamma-gliadin B pre (291 aa)
initn: 38 initl: 38 opt: 74 Z-score: 103.1 bits: 24.7 E(): 0.81
Smith-Waterman score: 78; 20.635% identity (66.667% similar) in 63 aa overlap
(1-58:5-65)

```

                                10      20      30      40      50
CV127.      MFTSLFFFSVSLSLAHLSPWPWRRQQQQHLLRSPSP---PNHLLPPPNHYYQ
                                .. ... ..: .: .::: .: .: .::: .: .:
gi|170 MKTLILITILAMAITIATANMQADPSGQVQWPQQQPFLLQPHQPFSSQQPQQIFPQPQQTF-
                                10      20      30      40      50

                                60
CV127. SPSPPSHSP
                                : .: .:
gi|170 -PHQPQQQFPQPQQPQQQLQPRQPFPPQQPYPQPQQPFPQTQQPQQPFPQSKQPQQ
                                60      70      80      90      100     110

gi|170 PFPQPQQPQQSFPQQPSLIQQSLQQQLNPCKNFLQCKPVSLVSSLWSIILPPSDCQV
                                120     130     140     150     160     170

gi|170 MRQQCCQLAQIPQQLQCAIHSVVHSIIMQQEQQEQLQGVQILVPLSQQQVQGILVQ
                                180     190     200     210     220     230

gi|170 GQGIIQPQQPAQLEVIRSLVLQTLPTMCNVYVPPYCSTIRAPFASIVASIGGQ
                                240     250     260     270     280     290

```

>>gi|62484809 gi|62484809|emb|CAI78902.1| putative gamma (285 aa)
initn: 92 initl: 43 opt: 73 Z-score: 102.1 bits: 24.5 E(): 0.93
Smith-Waterman score: 83; 40.000% identity (52.000% similar) in 50 aa overlap
(16-60:60-105)

```

                                10
CV127.      MFTSLFFFSVSLSLAH
                                :
gi|624 MQVNPSVQVQPTQQQPYPESQQPFISQSQQQFPQPQQPFPQQPFPQSQQQLQPPQH
                                10      20      30      40      50      60

                                20      30      40      50      60
CV127. LSPEPWRQQQQQHLLRSPSP---PNHLLPPPNHYYQS-PDPSHSP
                                :. :. :. :. :. :. :. :. :. :. :
gi|624 QFPQP--TQFPQRPLLPFTHPFLTFPDQLLPQPPH--QSFPQPPQSYQPPLQFPFPQP
                                70      80      90      100     110

```

gi|624 QQKYPEQPQQPFPWQQPTIQLYLQQQLNPCKEFLQQCRPVSLLSYLWSKIVQQSSCRM
120 130 140 150 160 170

gi|624 QQQCCLQLAQIPEQYKCTAIDSIVHAIFMQQGQRQGVQIVQQQPQQVGGQCVLVQGQGV
180 190 200 210 220 230

gi|624 VQPQQLAQMEAIRTLVLQSVPSMCNFNVPNCSTIKAPFVGVTGVGGQ
240 250 260 270 280

>>gi|1168171 gi|1168171|gb|AAB35353.1| allergenic peptid (30 aa)
initn: 46 initl: 46 opt: 57 Z-score: 98.5 bits: 20.6 E(): 1.5
Smith-Waterman score: 57; 35.714% identity (57.143% similar) in 28 aa overlap
(27-53:2-29)

10 20 30 40 50
CV127. MFTSLFFFSVSLSLAHLSPFPWRRQQQQHLLRSPSPF-NHLLPPPNHHYQSPDSPSHS
::: . . .: . . : . :
gi|116 SQQQPPFSSQQPPFSSQQPPFSSQQPPF
10 20 30

60
CV127. P

>>gi|21926 gi|21926|emb|CAA36063.1| unnamed protein prod (295 aa)
initn: 76 initl: 66 opt: 69 Z-score: 97.1 bits: 23.7 E(): 1.8
Smith-Waterman score: 69; 44.444% identity (74.074% similar) in 27 aa overlap
(27-51:80-106)

CV127. MFTSLFF
gi|219 MKTFLV FALLAVVATSTIAQMETSCIPGLERPWQEQLPPQHTLFPQQPPFPQQQPPFSS
10 20 30 40 50 60

10 20 30 40 50 60
CV127. FSVSLSLAHLSPFPWRRQQQQHLL--RSPSPPNHLLPPPNHHYQSPDSPSHS
::: . : .: . . : . . .:
gi|219 QQQPSFLQQPILPQLPFSQQQPVLPPQQSPFSSQQQLVLPQQYQQVLQQQIPVQPSV
70 80 90 100 110 120

gi|219 LQQNLNPKVFLQQCNPVAMPQRLARSQMLQQSSCHVMQQCCQLPQIPEQSRVDVIRA
130 140 150 160 170 180

gi|219 ITYSIILQEQQGFVQAQQQPQQLGQGVSSQQSSQQQLGQCSFQQPQQQLGQQPQQQQ
190 200 210 220 230 240

gi|219 VLQGTFLQPHQIAHLEVMTSIALRTLPTMCSVNVPLYSSTTSVPFSGVTGVGAYL
250 260 270 280 290

>>gi|170730 gi|170730|gb|AAA34285.1| gamma-gliadin B-I p (304 aa)
initn: 56 initl: 56 opt: 68 Z-score: 95.7 bits: 23.4 E(): 2.1
Smith-Waterman score: 78; 34.694% identity (57.143% similar) in 49 aa overlap
(21-60:49-97)

10 20
CV127. MFTSLFFFSVSLSLAHLSPFPWRRQQQQ--Q
: . .: . :
gi|170 MKTFLVFALIAVVATSAIAQMETSCISGLERPWQQQLPPQSFSSQQPPFSSQQQQLPQ
10 20 30 40 50 60

```

      30      40      50      60
CV127. QHLLRSPSP-PNLLP-----PPNHYYQSPDSPSHSP
      : . . .:: . . :      :: . ::: :...:
gi|170 QPSFSQQQPPFSQQQPILSQQPPFSQQQPVLPQQSPFSQQQQLVLPQQQQQQLVQQQI
      70      80      90      100      110      120

gi|170 PIVQPSVLQQLNPCKVFLQQQCSPVAMPQRLARSQMWQSSCHVMQQCCQQLQQIPEQS
      130      140      150      160      170      180

gi|170 RYEAIRAIYIISIILQEQQQGFVQPQQQPQQSGQGVSSQSSQQQLGQCSFQQPQQQLG
      190      200      210      220      230      240

gi|170 QQPQQQQQQQLVLTGTFQLQPHQIAHLEAVTSIALRTLPTMCSVNVPLYSATTSVPFVGVTG
      250      260      270      280      290      300

gi|170 VGAY

```

>>gi|170722 gi|170722|gb|AAA34281.1| pre-alpha-/beta-gli (262 aa)
 initn: 44 initl: 44 opt: 67 Z-score: 95.5 bits: 23.2 E(): 2.1
 Smith-Waterman score: 78; 29.091% identity (67.273% similar) in 55 aa overlap
 (9-60:20-74)

```

      10      20      30      40
CV127. MFTSLFFFSVSLSLAHLSP-PWRRQQQQQHLLRSPSPNHL-LP-PN
      .: . . .::: . : .: ::: :... . . . :::
gi|170 MKTFLILALLAIVATTATTAVRVVPQLQPQNPSQQQPQEQVPLVQQQFLGQQQFPFPQ
      10      20      30      40      50      60

      50      60
CV127. HHYQSPDS-PSHSP
      . : .: .:::
gi|170 QPYPQPQPFPSQQPYLQLQPFLQPQLPYSQPQPFRRPQQPYPQPQPQYSQPQQPISQQQQQ
      70      80      90      100      110      120

gi|170 QQQQQQQQQQQQQQIIQQILQQQLIPCMDVVLQQHNIVHGKSQVLQQSTYQLLQELCCQH
      130      140      150      160      170      180

gi|170 LWQIPEQSQCQAIHNVVHAILHQQQKQQQQPSSQVSFQQPLQQYPLGQGSFRPSQQNPQ
      190      200      210      220      230      240

gi|170 AQGSVQPQQLPQFEEIRNLARK
      250      260

```

>>gi|27818335 gi|27818335|gb|AA024900.1| major pollen al (132 aa)
 initn: 75 initl: 63 opt: 63 Z-score: 95.5 bits: 22.2 E(): 2.2
 Smith-Waterman score: 65; 44.444% identity (51.852% similar) in 27 aa overlap
 (35-60:100-126)

```

      10      20      30      40      50      60
gi|278 MAKCSYVFCVALLIFIVAIGEMEAAGSKLCEKTSKTYSGKCDNKKCDKKCIEWEKAQHGA

      10      20      30      40      50
CV127. MFTSLFFFSVSLSLAHLSP-PWRRQQQQQHLLRSPSP-PNLLPPPNHHYQSPD
      :::: . ::: :
gi|278 CHKREAGKESCFYFDCSKSPPGATPAPPGAAPPPAAGGSPSPADGGSPPPADGGSP
      70      80      90      100      110      120

      60
CV127. SPSHSP
      . :
gi|278 VDGGSPPPSTH
      130

```


>>gi|473876 gi|473876|gb|AAA17741.1| alpha-gliadin (287 aa)
initn: 44 initl: 44 opt: 67 Z-score: 94.9 bits: 23.2 E(): 2.3
Smith-Waterman score: 78; 29.091% identity (67.273% similar) in 55 aa overlap
(9-60:20-74)

```

              10      20      30      40
CV127.      MFTSLFFFSVSLSLAHLSP-E-PWRRQQQQQHLLRSPSPPNHLLP-PPN
              .: . . .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|473 MKTFLILALLAIVATTATTAVRVVPQLQPQNPSQQQPQEQLVPLVQQQFLGQQQPFPPQ
              10      20      30      40      50      60

              50      60
CV127. HHYQSPDS-PSHSP
              .: .: .: .: .:
gi|473 QPYQPQPFPSPQPYLQLQPFLQPQLPYSQPQPFPPQPYQPQPQYSQPQPISQQQQQ
              70      80      90      100     110     120

gi|473 QQQQQQQQQQQQQQIIQQILQQQLIPCMDVVLQQHNIVHGKSQVLQQSTYQLLQELCCQH
              130     140     150     160     170     180

gi|473 LWQIPEQSQCQAIHNVVHAILHQQQKQQQPSSQVSFQQPLQQYPLGQGSFRPSQQNPQ
              190     200     210     220     230     240

gi|473 AQGSVQPQQLPQFEEIRNLALQTLPMCNVYIPPYCTIAPFGIFGTN
              250     260     270     280

```

>>gi|170702 gi|170702|gb|AAA34272.1| gamma gliadin precu (302 aa)
initn: 38 initl: 38 opt: 67 Z-score: 94.6 bits: 23.2 E(): 2.4
Smith-Waterman score: 84; 27.083% identity (66.667% similar) in 48 aa overlap
(12-59:105-151)

```

              10      20      30      40      50      60
gi|170 MKTLLILITILAMATTIATANMQVDPSPGVQWPQQQPFPPQPPFCQQPQRTIPQPHQTFH
              10      20      30      40      50      60

              10      20
CV127.      MFTSLFFFSVSLSLAHLSP-E-PWRRQQQ
              .: . . .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|170 HQPQQTFPPQPPQTYHPQPPQFPQTQQPQQPFPQPPQTFPPQPPQLPFPQPPQPPFPQPPQ
              70      80      90      100     110     120

              30      40      50      60
CV127. QQQHLLRSPSPPNHLLPPPNHHYQSPDSPSHSP
              .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|170 PQQPFPQSQQPQQPF-PQPQQQFPQPQQPQSFPPQQQPAIQSFLQQQMNPCKNFLQQC
              130     140     150     160     170

gi|170 NHVSLVSSLVSIILPRSDCQVMQQQCCQQLAQIPQQQLQCAAIHVAHSIIMQQEQQQGVP
              180     190     200     210     220     230

gi|170 ILRPLFQLAQGLGIIQPQQPAQLEGIRSLVLKTLPTMCNVYVPPDCSTINVPYANIDAGI
              240     250     260     270     280     290

gi|170 GGQ
              300

```

>>gi|21765 gi|21765|emb|CAA26385.1| unnamed protein prod (313 aa)
initn: 217 initl: 56 opt: 67 Z-score: 94.3 bits: 23.2 E(): 2.5
Smith-Waterman score: 80; 29.091% identity (67.273% similar) in 55 aa overlap
(9-60:17-71)

```

          10          20          30          40          50
CV127.      MFTSLFFFSVSLSLAHLSP-E-PWRRQQQQQHLLRSPSPPNHLLP-PPNHYY
          .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|217 MKTFLILALVATTATTAVRVVPQLQPKNPSQQQPQEQLVPLVQQQQFPGQQQQFPPQQPY
          10          20          30          40          50          60

```

```

          60
CV127. QSPDS-PSHSP
          .: .: .: .: .:
gi|217 PQQPFPSPQQPYLQLQFPFPQPFLPQLPYQPQSFPFPQQPYPPQRPYLPQQPISQQQ
          70          80          90          100          110          120

```

```

gi|217 AQQQQQQQQQQQQQQQQQQILQQILQQQLIPCRDVLQHHNIAHASSQVLQQSTYQLLQQ
          130          140          150          160          170          180

```

```

gi|217 CCQQLLQIPEQSRCQAIHNVVHAIIMHQQEQQQQLQQQQQQQLQQQQQQQQQQQPSSQV
          190          200          210          220          230          240

```

```

gi|217 SFQQPQQQYPSSQGSFQPSQQNPQAQGSVQPQQLPQFAEIRNLALQTLPMACNVYIPPHC
          250          260          270          280          290          300

```

```

gi|217 STTIAPFGIFGTN
          310

```

>>gi|170718 gi|170718|gb|AAA34279.1| alpha/beta-gliadin (313 aa)
 initn: 217 initl: 56 opt: 67 Z-score: 94.3 bits: 23.2 E(): 2.5
 Smith-Waterman score: 80; 29.091% identity (67.273% similar) in 55 aa overlap
 (9-60:17-71)

```

          10          20          30          40          50
CV127.      MFTSLFFFSVSLSLAHLSP-E-PWRRQQQQQHLLRSPSPPNHLLP-PPNHYY
          .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|170 MKTFLILALVATTATTAVRVVPQLQPKNPSQQQPQEQLVPLVQQQQFPGQQQQFPPQQPY
          10          20          30          40          50          60

```

```

          60
CV127. QSPDS-PSHSP
          .: .: .: .: .:
gi|170 PQQPFPSPQQPYLQLQFPFPQPFLPQLPYQPQSFPFPQQPYPPQRPKYLPQQPISQQQ
          70          80          90          100          110          120

```

```

gi|170 AQQQQQQQQQQQQQQQQQQILQQILQQQLIPCRDVLQHHNIAHASSQVLQQSTYQLLQQ
          130          140          150          160          170          180

```

```

gi|170 CCQQLLQIPEQSRCQAIHNVVHAIIMHQQEQQQQLQQQQQQQLQQQQQQQQQQQPSSQV
          190          200          210          220          230          240

```

```

gi|170 SFQQPQQQYPSSQGSFQPSQQNPQAQGSVQPQQLPQFAEIRNLALQTLPMACNVYIPPHC
          250          260          270          280          290          300

```

```

gi|170 STTIAPFGIFGTN
          310

```

>>gi|162805 gi|162805|gb|AAA30431.1| beta-casein (224 aa)
 initn: 47 initl: 47 opt: 65 Z-score: 94.2 bits: 22.7 E(): 2.5
 Smith-Waterman score: 65; 35.556% identity (57.778% similar) in 45 aa overlap
 (6-45:134-174)

```

gi|162 MKVLILACLVALALARELEELNVPGEIVESLSSEESITRINKKIEKFQSEEQQQTEDEL
          10          20          30          40          50          60

```


>>gi|170710 gi|170710|gb|AAA34275.1| alpha-type gliadin (318 aa)
initn: 188 initl: 48 opt: 67 Z-score: 94.2 bits: 23.2 E(): 2.5
Smith-Waterman score: 81; 29.091% identity (67.273% similar) in 55 aa overlap
(9-60:20-74)

```

              10      20      30      40
CV127.      MFTSLFFFSVSLSLAHLSP-E-PWRRQQQQQHLLRSPSPPNHLLP-PPN
              .: . . .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|170 MKTFLILALLAIVATTATTAVRVVPQLQPQNPSQQQPQEQVPLVQQQFPGQQQFPPQ
              10      20      30      40      50      60

              50      60
CV127. HHYQSPDS-PSHSP
              .: .: .: .: .:
gi|170 QPYQPQPFPSPQPYLQLQPFPPQPFPPPLPYPPQPSFPPQPPYPPQQPQYLQPPQPIISQ
              70      80      90      100     110     120

gi|170 QQAQQQQQQQQQQQQQQQILQQILQQQLIPCRDVVLQQHNIAHASSQVLQQSTYQLLQQ
              130     140     150     160     170     180

gi|170 CCQQLLQIPEQSQCAIHNVAHAIIMHQQQQQQEQKQQLQQQQQQQQQLQQQQQQQQQQ
              190     200     210     220     230     240

gi|170 PSSQVSFQQPQQYPSSQVSFQPSQLNPQAQGSVQPQQLPQFAEIRNLALQTLPMCNVY
              250     260     270     280     290     300

gi|170 IPPHCSTTIAPFGISGTN
              310

```

>>gi|170716 gi|170716|gb|AAA34278.1| pre-alpha-/beta-gli (319 aa)
initn: 188 initl: 48 opt: 67 Z-score: 94.2 bits: 23.2 E(): 2.5
Smith-Waterman score: 81; 29.091% identity (67.273% similar) in 55 aa overlap
(9-60:20-74)

```

              10      20      30      40
CV127.      MFTSLFFFSVSLSLAHLSP-E-PWRRQQQQQHLLRSPSPPNHLLP-PPN
              .: . . .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|170 MKTFLILALLAIVATTATTAVRVVPQLQPQNPSQQQPQEQVPLVQQQFPGQQQFPPQ
              10      20      30      40      50      60

              50      60
CV127. HHYQSPDS-PSHSP
              .: .: .: .: .:
gi|170 QPYQPQPFPSPQPYLQLQPFPPQPFPPPLPYPPQPSFPPQPPYPPQQPQYLQPPQPIIS
              70      80      90      100     110     120

gi|170 QQAQQQQQQQQQQQQQQQILQQILQQQLIPCRDVVLQQHNIAHASSQVLQQSTYQLLQQ
              130     140     150     160     170     180

gi|170 LCCQQLLQIPEQSQCAIHNVAHAIIMHQQQQQQEQKQQLQQQQQQQQQLQQQQQQQQQQ
              190     200     210     220     230     240

gi|170 QPSSQVSFQQPQQYPSSQVSFQPSQLNPQAQGSVQPQQLPQFAEIRNLALQTLPMCNV
              250     260     270     280     290     300

gi|170 YIPPHCSTTIAPFGISGTN
              310

```

>>gi|1063270 gi|1063270|dbj|BAA11251.1| gamma-gliadin pr (279 aa)
initn: 43 initl: 43 opt: 64 Z-score: 91.6 bits: 22.6 E(): 3.6
Smith-Waterman score: 83; 30.000% identity (75.000% similar) in 40 aa overlap
(20-59:79-117)

CV127. M

gi|106 NIQVDPSGQVQWPQQQPFQPHQPFSSQQPQQTFPQPPQQTFFPHQPPQQFSQPPQQPQQQFIQ

CV127. FTSLFFFSVSLSLAHLSPWPWRRQQQQQHLLRSPSPPNHLLPPPNHHYQSPDPSHSP

gi|106 PQQPFPQQPQQTYPQRPQQPFPQTQQPQQPFPQS-QQPQQPFPQPPQQFPQPQQPQQSFP

gi|106 QQQPSLIQQSLQQQLNPCKNFLQCKPVSLVSSLWSMILPRSDCQVMRQQCCQQLAQIP
120 130 140 150 160 170

gi|106 QQLQCAAIHSIVHSIIMQQEQQEQRQGVQILVPLSQQQQVVGQGTLVQGQGIIPQPPAQL

gi|106 EVIRSSVLQTLATMCNVYVPPYCSTIRAPFASIVAGIGGQ
240 250 260 270

```
>>gi|21755 gi|21755|emb|CAA25593.1| unnamed protein prod (286 aa)
  initn: 44 initl: 44 opt: 64 Z-score: 91.4 bits: 22.6 E(): 3.6
Smith-Waterman score: 75; 29.091% identity (65.455% similar) in 55 aa overlap
(9-60:20-74)
```

CV127. MFTSLFFFSVSLSLAHLSP-E-PWRRQQQQQHLLRSPSPPNHLLP-P

gi|217 MKTFLILVLLAIVATTATTAVRFPVLPQLQPQNPSQQQPQEQVPLVQQQQFLGQQQPFPQQ

CV127. HHYQSPDS-PSHSP

gi|217 QPYPQPQPFPSQLPYLQLQFPQPQLPYSQPQPFRRPQQPYFPQPQPYSPQPQQPISQQQQQ

gi|217 QQQQQQQQQQQQILQQILQQQLIPCMDVVLQQHNIAHGRSQVLQQSTYQLLQELCCQHL

gi | 217 WQIPEQSQCQAILKVVAIIILHQQKQQQQPSSQVSFQQPLQQYPLGQGSFRPSQQNPQA

gi|217 QGSVQPQQLPQFEEIRNLALQTL PAMCNVYIPPYCTIAPFGIFGTN

```
>>gi|21761 gi|21761|emb|CAA26384.1| unnamed protein prod (286 aa)
initn: 44 initl: 44 opt: 64 Z-score: 91.4 bits: 22.6 E(): 3.6
Smith-Waterman score: 75; 29.091% identity (65.455% similar) in 55 aa overlap
(9-60:20-74)
```

CV127. MFTSLFFFSVLSLAHLSPE-PWRRQQQQQHLLRSPSPPNHLLP-PPN

gi|217 MKTFLILVLLAIVATTATTAVRFVPVQLQPQNPSQQQPFQEQVPLVQQQQLGQQQFPFPQ
10 20 30 40 50 60

CV127. HHYQSPDS-PSHSP

gi|217 QPYPQPQPFPSQLPYLQLQPFQQLPYSQPQQFRPQQPYFPQFPQYSQPQQPISQQQQQ

gi|217 QQQQQQQQQQQQILQQILQQQLIPCMDVVLQQHNIAGRSQVLQQSTYQLLQELCCQHL
130 140 150 160 170 180

gi|217 WQIPEQSQCQAIHNVVHAILHQQQKQQQPSSQVSFQQPLQQYPLGQGSFRPSQQNPQA
190 200 210 220 230 240

gi|217 QGSVQPQLPQFEEIRNLALQTLPAMCNVYIPPYCTIAPFGIFGTN
250 260 270 280

>>gi|170720 gi|170720|gb|AAA34280.1| alpha/beta-gliadin (286 aa)
initn: 44 initl: 44 opt: 64 Z-score: 91.4 bits: 22.6 E(): 3.6
Smith-Waterman score: 75; 29.091% identity (65.455% similar) in 55 aa overlap
(9-60:20-74)

CV127. MFTSLFFFSVSLSLAHLSP- PWRQQQQQHLRSPSPPNHLLP-PPN
10 20 30 40
gi|170 MKTFLILVLLAIVATTATTAVRFPVPLQPNPSQQQPQEQVPLVQQQFLGQQQFPFPQ
10 20 30 40 50 60

CV127. HHYQSPDS-PSHSP
50 60
gi|170 QPYQPQPFPSQLPYLQLQPFQPLPSQPFQFPQPYQPQPYQPQISQQQQQ
70 80 90 100 110 120

gi|170 QQQQQQQQQQQQILQQILQQQLIPCMDVVLQQHNIAGRSQVLQQSTYQLLQELCCQHL
130 140 150 160 170 180

gi|170 WQIPEQSQCQAIHNVVHAILHQQQKQQQPSSQVSFQQPLQQYPLGQGSFRPSQQNPQA
190 200 210 220 230 240

gi|170 QGSVQPQLPQFEEIRNLALQTLPAMCNVYIPPYCTIAPFGIFGTN
250 260 270 280

>>gi|162797 gi|162797|gb|AAA30430.1| beta-casein precurs (224 aa)
initn: 47 initl: 47 opt: 61 Z-score: 89.5 bits: 21.9 E(): 4.6
Smith-Waterman score: 61; 33.333% identity (50.000% similar) in 48 aa overlap
(6-45:134-174)

gi|162 MKVLILACLVALALARELEELNVPGEIVESLSSESITRINKKIEKFQSEEQQQTEDEL
10 20 30 40 50 60

gi|162 QDKIHPPAQTQSLVYPFGPIPNLQNIPLTQTPVVVPPFLQPEVLGVSKVKEAMAPK
70 80 90 100 110 120

CV127. MFTSLFFFSVSLSLA-----HLSP---EPWRQQQQQHLRSPSPPNHLLP
10 20 30 40
gi|162 HKEMPFPKYPVEPFTESQSLTLTDVENLHLPPLLQSWMHQPHQ-----PLPPTVMFP
130 140 150 160 170

CV127. PPNHHYQSPDSPSHSP
50 60
gi|162 PQSVLSLSQSKVLPVPQKAVPYQPRDMPQAFLLYQQPVLGPVRGPFPIIV
180 190 200 210 220

>>gi|741844 gi|741844|prf||2008179A major allergen Par j (143 aa)
initn: 64 initl: 40 opt: 58 Z-score: 89.0 bits: 21.1 E(): 4.9

Smith-Waterman score: 58; 32.258% identity (61.290% similar) in 31 aa overlap (24-53:93-123)

gi | 741 MVRALMPCLPFVQGKEKEPSKGCCSGAKRLDGETKTGPQRVHACECIQTAMKTYSDIDGK
10 20 30 40 50 60

CV127. MFTSLFFFSVSLSLAHLSPWPWRRQQQQQHLLRSP-SPPNHLLPPPNHHY

gi|741 LVSEVPKHC GIVDSKLPPI DVNMDCKTVGVVPRQPQLPVSLRHGPVTGPSRSRPPTKHGW
70 80 90 100 110 120

CV127. OSPDSPSHSP 60

gi|741 RDPRLEFRPPHRKKPNPAFSTLG
130 140

```
>>gi|170732 gi|170732|gb|AAA34286.1| gamma-gliadin (323 aa)
  initn: 114 initl: 61 opt: 61 Z-score: 87.0 bits: 21.9 E(): 6.4
Smith-Waterman score: 62; 35.000% identity (52.500% similar) in 40 aa overlap
(21-60:83-116)
```

gi|170 TITRTFPIPTISSNNHHFRSNSNHHFHNNNQFYRNNNSPGHNNPLNNNNNSPNNNSPSN
10 20 30 40 50 60

CV127. 10 20 30 40 50
MFTSLFFFSVSLSLAHLSPWPWRROOOOOOHLRSPSPNHLPPPNHHYOSPDSPSH

gi|170 HHNNSPNNNFQYHTHPSNHKNLPHTNNIQQQQ-----P-PFSQQQQPPFSQQQQPVLPPQ

CV127. SP
gi|170 SPFSQQQQQLVLPFPQQQQQLVQQQIPVQPSVLQQLNPCKVFLQQQCSPVAMPQRLARSQ

gi|170 MWQQSSCHVMQQQCCQQLQQIPEQSRYEAIRAIIYSIILQEQQQGFVQPQQQPPQQSGQG

gi|170 VSQSQQQSQQQLGQCSFQQPQQQLGQQPQQQQQQQVLQGTFLLQPHQIAHLEAVTSIALRT

gi|170 LPTMCSVNVPLYSATTSPVFGVGTGVGAY
300 310 320

```
>>gi|169969 gi|169969|gb|AAA33964.1| glycinin (516 aa)
  initn: 104 initl: 56 opt: 61 Z-score: 83.8 bits: 22.0 E(): 9.6
Smith-Waterman score: 61; 33.333% identity (74.074% similar) in 27 aa overlap
(5-31:186-210)
```

gi|169 MGKPFFTLSLSSLCLLLLSSACFAITSSKFNECQLNNLNALEPDHRVESEGGLIETWNSQ
10 20 30 40 50 60

gi|169 HPQLQCAGVTVSKRTLNRNGSHLPSYLPYPQMIIVVQKGGAIGFAFPGCPETFEKPPQQQS

gi|169 SRRGSRSQQLQDSHQKIRHFNEGDVLIPLGVPTYTYNTGDEPVVAISPLDTSNFNQQL

```

          10          20          30          40          50
CV127.  MFTSLFFFSVSLSLAHLSPWPWRQQQQQHLLRSPSPPNHLLPPPNHHYQSPDSPSHS
          . . . . . : : : . . . . .
gi|169 DQNPRVFYLAGNPDIEH--PETMQQQQQKSHGGRKQGQHRQQEEEGGSVLSGFSGHFLA
          190          200          210          220          230

```

```

          60
CV127. P

```

```

gi|169 QSFNTNEDTAEKLRSPDDERKQIVTVEGGLSVISPKWQEQEDEDDEDEEYGRTPSYPPR
          240          250          260          270          280          290

```

```

gi|169 RPSHGKHEDDEDEDEEDQPRPDHPPQRPSPRPEQQEPRGRGCQTRNGVEENICTMKLHEN
          300          310          320          330          340          350

```

```

gi|169 IARPSRADFYNNPKAGRISTLNSLTLPALRQFGLSAQYVVLRYNGIYSPDWNLNANSVTMT
          360          370          380          390          400          410

```

```

gi|169 RGKGRVVRVNCQGNVFDGELRRGQLLVVPQNPVAVAEQGGEGGLEYYVFKTHHNAVSSYI
          420          430          440          450          460          470

```

```

gi|169 KDVFVRVIPSEVLSNSYNLGQSQVRQLKYQGNSGPLVNP
          480          490          500          510

```

```

>>gi|736002 gi|736002|emb|CAA55977.1| Gy5 [Glycine soja] (517 aa)
  initn: 104 initl: 56 opt: 61 Z-score: 83.8 bits: 22.0 E(): 9.6
Smith-Waterman score: 61; 33.333% identity (74.074% similar) in 27 aa overlap
(5-31:186-210)

```

```

gi|736 MGKPFFTLSLSSCLLLSSACFAITSSKFNECQLNNLNALEPDHRVESEGGLIETWNSQ
          10          20          30          40          50          60

```

```

gi|736 HPELQCAGVTVSKRTLNRNGLHLPSYPQMIIVVQKGAIGFAFPGPCPETFEKPQQQS
          70          80          90          100          110          120

```

```

gi|736 SRRGSRSQQLQDSHQKIRHFNEGDVLVIPPVGPVWYNTGDEPVVAISLLDTSNFNQQL
          130          140          150          160          170          180

```

```

          10          20          30          40          50
CV127.  MFTSLFFFSVSLSLAHLSPWPWRQQQQQHLLRSPSPPNHLLPPPNHHYQSPDSPSHS
          . . . . . : : : . . . . .
gi|736 DQNPRVFYLAGNPDIEH--PETMQQQQQKSHGGRKQGQHQQQEEEGGSVLSGFSGHFLA
          190          200          210          220          230

```

```

          60
CV127. P

```

```

gi|736 QSFNTNEDTAEKLRSPDDERKQIVTVEGGLSVISPKWQEQEDEDDEDEEYEQTPSYPPR
          240          250          260          270          280          290

```

```

gi|736 RPSHGKHEDDEDEDEEDQPRPDHPPQRPSPRPEQQEPRGRGCQTRNGVEENICTMKLHEN
          300          310          320          330          340          350

```

```

gi|736 IARPSRADFYNNPKAGRISTLNSLTLPALRQFGLSAQYVVLRYNGIYSPHWNLNANSVIYV
          360          370          380          390          400          410

```

```

gi|736 TRGKGRVVRVNCQGNVFDGELRRGQLLVVPQNFVVAEQGGEGGLEYYVFKTHHNAVSSY
          420          430          440          450          460          470

```

```

gi|736 IKDVFRAIPSEVLSNSYNLGQSQVRQLKYQGNSGPLVNP
          480          490          500          510

```



```
>>gi|1228078 gi|1228078|emb|CAA33034.1| kappa-casein [Bo (190 aa)
  initn: 30 initl: 30 opt: 55 Z-score: 83.5 bits: 20.5 E(): 10
Smith-Waterman score: 55; 23.684% identity (52.632% similar) in 38 aa overlap
(21-57:105-142)
```

```
gi|122 MMKSFFLVVTILALTLPFLGAQEQNQEPIRCEKDERFFSDKIAKYIPIQYVLSRYPSTY
      10          20          30          40          50          60
```

```
CV127.          10          20          30
      MFTSLFFFSVSLSLAHLSPWPWRQQQQQQLLRSP
      : . : : . : :
```

```
gi|122 LNYQQKPVALLNNQFLPYYPYAKPAAVRSPAQILQWQVLSNTVPAKSCQAQPTTMRHP
      70          80          90          100          110          120
```

```
      40          50          60
CV127. SPPNHLLP-PPNHYYQSPDPSHSP
      : .. : : : . : .
```

```
gi|122 HPHLSFMAIPPKKNQDKTEIPTINTIASGEPTSTPTTEAVESTVATLEDSPEVIESPPEI
      130          140          150          160          170          180
```

```
gi|122 NTVQVTSTAV
      190
```

```
>>gi|162811 gi|162811|gb|AAA30433.1| kappa-casein precu (190 aa)
  initn: 30 initl: 30 opt: 55 Z-score: 83.5 bits: 20.5 E(): 10
Smith-Waterman score: 55; 23.684% identity (52.632% similar) in 38 aa overlap
(21-57:105-142)
```

```
gi|162 MMKSFFLVVTILALTLPFLGAQEQNQEPIRCEKDERFFSDKIAKYIPIQYVLSRYPSTY
      10          20          30          40          50          60
```

```
CV127.          10          20          30
      MFTSLFFFSVSLSLAHLSPWPWRQQQQQQLLRSP
      : . : : . : :
```

```
gi|162 LNYQQKPVALLNNQFLPYYPYAKPAAVRSPAQILQWQVLSNTVPAKSCQAQPTTMRHP
      70          80          90          100          110          120
```

```
      40          50          60
CV127. SPPNHLLP-PPNHYYQSPDPSHSP
      : .. : : : . : .
```

```
gi|162 HPHLSFMAIPPKKNQDKTEIPTINTIASGEPTSTPTTEAVESTVATLEASPEVTESPEI
      130          140          150          160          170          180
```

```
gi|162 NTVQVTSTAV
      190
```

```
60 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:10:10 2008 done: Sun Dec 14 22:10:10 2008
Total Scan time: 0.080 Total Display time: 0.040
```

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.4448_c3: amino acids 33-112

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.4448_c3_33-112, 80 aa
vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 3.2153 \pm 0.00419$; $\mu = 11.9553 \pm 0.221$
mean_var=52.2786 \pm 14.725, 0's: 14 Z-trim: 15 B-trim: 29 in 1/41
Lambda= 0.177383

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:

	opt	bits	E(1313)	%_id	%_sim	alen
gi 121308880 gi 121308880 dbj BAF43 (84)	65	23.5	0.76	0.345	0.500	58
gi 112754 gi 112754 sp P04403.2 2SS (146)	59	22.3	3	0.316	0.579	38
gi 94471624 gi 94471624 gb ABF21078 (175)	59	22.4	3.3	0.393	0.714	28
gi 75219081 gi 75219081 sp O22108 O (285)	58	22.4	5.3	0.250	0.574	68
gi 1228078 gi 1228078 emb CAA33034. (190)	56	21.7	5.9	0.294	0.588	34
gi 162811 gi 162811 gb AAA30433.1 (190)	56	21.7	5.9	0.294	0.588	34
gi 66845476 gi 66845476 gb EAL85811 (168)	55	21.3	6.6	0.293	0.610	41
gi 2769700 gi 2769700 gb AAB95638.1 (168)	55	21.3	6.6	0.293	0.610	41
gi 21930 gi 21930 emb CAA44473.1 L (285)	56	21.9	7.5	0.250	0.559	68
gi 75219561 gi 75219561 sp O49178 O (157)	54	21.0	7.6	0.545	0.818	11
gi 2827316 gi 2827316 gb AAB99797.1 (157)	54	21.0	7.6	0.545	0.818	11
gi 5326864 gi 5326864 gb AAD42074.1 (167)	54	21.1	7.9	0.268	0.634	41
gi 94471622 gi 94471622 gb ABF21077 (223)	54	21.3	9.3	0.526	0.842	19

>>>CV127.4448_c3_33-112, 80 aa vs
/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|121308880 gi|121308880|dbj|BAF43535.1| protease inh (84 aa)
initn: 38 init1: 38 opt: 65 Z-score: 103.7 bits: 23.5 E(): 0.76
Smith-Waterman score: 65; 34.483% identity (50.000% similar) in 58 aa overlap
(25-80:38-82)

```

              10      20      30      40
CV127.      ATIRDHRQHMFLVRTNHIQQVRSWCIQHCLNSFSEIC-FLCHPRRW
              : . : . : . : . : . :
gi|121 MFQSTFFLVLMVCVATARFANKDHCPPNEEYNECGNPCQEKCDNGEPVICTYQCEHR--
              10      20      30      40      50

              50      60      70      80
CV127. PASC-CKQHVREYLVLLSRIPEKCVSSVSFVEPI
              : : : : : : : : : :
gi|121 ---CFCKQGYVR-----LTEDGE-CVPE-EFCKPIHY
              60      70      80

```

>>gi|112754 gi|112754|sp|P04403.2|2SS_BEREX 2S sulfur-ri (146 aa)
initn: 42 init1: 42 opt: 59 Z-score: 92.9 bits: 22.3 E(): 3
Smith-Waterman score: 59; 31.579% identity (57.895% similar) in 38 aa overlap
(19-53:47-84)

```

              10      20      30
CV127.      ATIRDHRQHMFLVRTNHIQQVRSWCIQHCLNS
              : . : . : . : .
gi|112 MAKISVAAALLVLMALGHATAFRATVTTTVVEENQEECREQMQRQQLSHCRMVYMRQQ
              10      20      30      40      50      60

```

```

              40              50              60              70              80
CV127. FSEICFLCHPRR-WNP--ASCCKQQHVVREYLVLLSRIPEKCVSSVSFVEPI
      . : . . . . . : . . . . .
gi|112 MEESPYQTMPPRGMEPHMSECCEQLEGMDDESCRCEGLRMMMMRMQQEEMQPRGEQMRMM
              70              80              90              100             110             120

gi|112 RLAENIPSRCNLSPMRCMPMGSIAGF
              130             140

```

>>gi|94471624 gi|94471624|gb|ABF21078.1| icarapin varian (175 aa)
 initn: 42 initl: 42 opt: 59 Z-score: 92.1 bits: 22.4 E(): 3.3
 Smith-Waterman score: 59; 39.286% identity (71.429% similar) in 28 aa overlap
 (45-71:40-67)

```

              10              20              30              40              50              60
CV127. ATIRDHRQHMFVLRVTNHIQQVRSWCIQHCLNSFSEICFLCHPRRWNPPASCCKQQHVVREYL
      . . . . . : . . . . .
gi|944 ERDQMMMAATFDFPSLSFEDSDEGSNWNWNTLLRPNFLDGWYQTLQTHMKKVREQM
              10              20              30              40              50

              70              80
CV127. V-LLSRIPEKCVSSVSFVEPI
      . . . . . :
gi|944 AGILSRIPEQGVVWNWNIKEGANTTSTTKIIDGHVVTINETTYTDGSDDYSTLIRVRVID
              60              70              80              90              100             110

gi|944 VRPQNETILTTSSEADSDVTTLPTLIGKNETSTQSSRSVESVEDFDNEIPKNQGDVLTA
              120             130             140             150             160             170

```

>>gi|75219081 gi|75219081|sp|O22108|O22108_WHEAT LMM glu (285 aa)
 initn: 35 initl: 35 opt: 58 Z-score: 88.5 bits: 22.4 E(): 5.3
 Smith-Waterman score: 58; 25.000% identity (57.353% similar) in 68 aa overlap
 (7-69:101-163)

```

gi|752 RCIPGLERPWWQQPLPPQQTFFPQQPLFSQQQQQQLFPQQPSFSQQQPPFWQQQPPFSQQQ
              10              20              30              40              50              60

              10              20
CV127. ATIRDHRQHMFVLRVTNHIQQV--RS
      . . . . . : . . . . .
gi|752 PILPQQPPFSQQQQLVLPQQSPFSQQQQLILPPQQQQQLPQQQISIVQPSVLQQLNPCKV
              70              80              90              100             110             120

              30              40              50              60              70              80
CV127. WCIQHCLNSFSEICFLCHPRRWNPPASC--CKQQHVVREYLVLLSRIPEKCVSSVSFVEPI
      . : . . . : . . . . . : . . . . .
gi|752 FLQQQC-SPVAMPQRLARSQMWQQSSCHVMQQQCCQQ---LSQIPEQSRDYDAIRAITYS
              130             140             150             160             170

gi|752 IILQEQQQGFVQAQQQQPQQSGQGVSSQSSQQQLGQCSFQQPQQQLGQQPQQQQVQQG
              180             190             200             210             220             230

gi|752 TFLQPHQIAHLEVMTSIALRTLPMCSVNVPLYSSTTSVPFGVGTGVGAY
              240             250             260             270             280

```

>>gi|1228078 gi|1228078|emb|CAA33034.1| kappa-casein [Bo (190 aa)
 initn: 55 initl: 55 opt: 56 Z-score: 87.6 bits: 21.7 E(): 5.9
 Smith-Waterman score: 56; 29.412% identity (58.824% similar) in 34 aa overlap
 (38-70:76-109)

```

CV127.                                     10      20
                                         ATIRDHRQHMFLVRTNHIQQVR

gi|122 MMKSFFLVVTILALTLPFLGAQEONQEQPIRCEKDERFFSDKIAKYIPIQYVLSRYPSYG
      10      20      30      40      50      60

      30      40      50      60      70      80
CV127. SWCIQHCLNSFSEICFLCHPRRWNPASCCKQQHVREYLVLLSRIPEK-CVSSVSFVEPI
      :: :: :.:. . . . :: . : : :
gi|122 LNNYQQKPVALINNQFLPYYPYAKPAAVRSPAQILQWQVLSNTVPAKSCQAQPTTMARHP
      70      80      90      100     110     120

gi|122 HPHLSFMAIPPKNQDKTEIPTINTIASGEPTSTPTTEAVESTVATLEDSPEVIESPPEI
      130     140     150     160     170     180

gi|122 NTVQVTSTAV
      190

```

>>gi|162811 gi|162811|gb|AAA30433.1| kappa-casein precur (190 aa)
 initn: 55 initl: 55 opt: 56 Z-score: 87.6 bits: 21.7 E(): 5.9
 Smith-Waterman score: 56; 29.412% identity (58.824% similar) in 34 aa overlap
 (38-70:76-109)

```

CV127.                                     10      20
                                         ATIRDHRQHMFLVRTNHIQQVR

gi|162 MMKSFFLVVTILALTLPFLGAQEONQEQPIRCEKDERFFSDKIAKYIPIQYVLSRYPSYG
      10      20      30      40      50      60

      30      40      50      60      70      80
CV127. SWCIQHCLNSFSEICFLCHPRRWNPASCCKQQHVREYLVLLSRIPEK-CVSSVSFVEPI
      :: :: :.:. . . . :: . : : :
gi|162 LNNYQQKPVALINNQFLPYYPYAKPAAVRSPAQILQWQVLSNTVPAKSCQAQPTTMARHP
      70      80      90      100     110     120

gi|162 HPHLSFMAIPPKNQDKTEIPTINTIASGEPTSTPTIEAVESTVATLEASPEVTESPPEI
      130     140     150     160     170     180

gi|162 NTVQVTSTAV
      190

```

>>gi|66845476 gi|66845476|gb|EAL85811.1| allergen Asp F3 (168 aa)
 initn: 28 initl: 28 opt: 55 Z-score: 86.7 bits: 21.3 E(): 6.6
 Smith-Waterman score: 55; 29.268% identity (60.976% similar) in 41 aa overlap
 (35-75:47-85)

```

CV127.                                     10      20      30      40
                                         ATIRDHRQHMFLVRTNHIQQVRSWCIQHCLNSFSEICFLCHPRRWNP
                                         .. . . : .:.
gi|668 MSGLKAGDSFPSDVVFSYIPWSEDKGEITACGIPINYNASKEWADKKVILFALPGAFTPV
      10      20      30      40      50      60

      50      60      70      80
CV127. SCCKQQHVREYLVLLSRIPEKCVSSVSFVEPI
      :. :. :. :. :. :. :.
gi|668 --CSARHVPEYIEKLPEIRAKGVDDVAVLAYNDAYVMSAWGKANQVTGDDILFLSDPDAR
      70      80      90      100     110

gi|668 FSKSIGWADEEGRTKRYALVIDHGKITYAAALEPAKNHLEFSSAETVLKHL
      120     130     140     150     160

```

>>gi|2769700 gi|2769700|gb|AAB95638.1| peroxisomal-like (168 aa)
initn: 28 initl: 28 opt: 55 Z-score: 86.7 bits: 21.3 E(): 6.6
Smith-Waterman score: 55; 29.268% identity (60.976% similar) in 41 aa overlap
(35-75:47-85)

```

              10      20      30      40
CV127.          ATIRDHRQHMFLVRTNHIQQVRSWCIQHCLNSFSEICFLCHPRRWNP
              .. .. : ...
gi|276 MSGLKAGDSFPSDVVFSYIPWSEDKGEITACGIPINYNASKEWADKKVILFALPGAFTPV
              10      20      30      40      50      60

              50      60      70      80
CV127. SCCKQQHVREYLVLLSRIPEKCVSSVSFVEPI
              :. .:: :. : :. :. :.
gi|276 --CSARHVPEYIEKLPEIRAKGVDVAVLAYNDAYVMSAWGKANQVTGDDILFLSDPDAR
              70      80      90      100     110

gi|276 FSKSIGWADEEGRTKRYALVIDHGKITYAALEPAKNHLEFSSAETVLKHL
              120     130     140     150     160

```

>>gi|21930 gi|21930|emb|CAA44473.1| LMW glutenin [Tritic (285 aa)
initn: 35 initl: 35 opt: 56 Z-score: 85.8 bits: 21.9 E(): 7.5
Smith-Waterman score: 56; 25.000% identity (55.882% similar) in 68 aa overlap
(7-69:110-172)

```

              10      20      30      40      50      60
gi|219 MKTFLVFALLAVVATSAIAQMDTSCIPGLERPWQQQLPFPQQTFFPQQPPFSQQQQQQPFP
              10      20      30      40      50      60

              10
CV127.          ATIRDHRQHMFLVRTNH
              .... :. .
gi|219 QQPSFSQQQPILPQGPPFPQQTQPVLPQQSPFSQQQLILPPQQQQQLPQQQISIVQPSV
              70      80      90      100     110     120

              20      30      40      50      60      70
CV127. IQQVRSWCIQHCLNSFSEICF---LCHPRRWNPASC--CKQQHVREYLVLLSRIPEKCVS
              :. . : .. : . . : . . :. :. :. :. :. :. :. :. :. :. :. :. :.
gi|219 LQQLNPKVFLQQQCSPVAIPQRLARSQMWQQSSCHVMQQCCQQ----LSQIPEQSR
              130     140     150     160     170

              80
CV127. SVSFVEPI

gi|219 DAIRAITYSIILQEQQQQSQQQPQQSGQGVVSQSQQQSQQLGQCSFQQPQQQLGQQPQ
              180     190     200     210     220     230

gi|219 QQVQVQGTFLQPHQIAHLEVMTSIALRTLPTMCSVNVPLYSSTTSVPFGV
              240     250     260     270     280

```

>>gi|75219561 gi|75219561|sp|O49178|O49178_ORYSA Allerge (157 aa)
initn: 54 initl: 54 opt: 54 Z-score: 85.7 bits: 21.0 E(): 7.6
Smith-Waterman score: 54; 54.545% identity (81.818% similar) in 11 aa overlap
(43-53:71-81)

```

              10      20      30
CV127.          ATIRDHRQHMFLVRTNHIQQVRSWCIQHCLNS
              .. .. : ...
gi|752 MASNKVVFSVLLAVVSVLAATATMAEYHHQDQVVYTPGPLCQPGMGYPMPYPLRVAGVGE
              10      20      30      40      50      60

```

```

      40      50      60      70      80
CV127. FSEICFLCHPRRWNPASCKQQHVVREYLVLSSRIPEKCVSSVSFVEPI
      ::  :::::
gi|752 APLLGRARPRRRVPGDCCRQFPPVDYSWCRCEAISHMLGGIYRELGAPDVGHMSEVFR
      70      80      90      100     110     120

gi|752 GCRRGTWSARRRAPGVLQVDIPNGGGGVCYWLARSGY
      130     140     150

```

>>gi|2827316 gi|2827316|gb|AAB99797.1| allergenic protei (157 aa)
 initn: 54 init1: 54 opt: 54 Z-score: 85.7 bits: 21.0 E(): 7.6
 Smith-Waterman score: 54; 54.545% identity (81.818% similar) in 11 aa overlap
 (43-53:71-81)

```

                                10      20      30
CV127.                        ATIRDHRQHMFLVRTNHIQQVRSWCIQHCLNS

gi|282 MASNKVVFSVLLLAUVSVLAATATMAEYHHQDQVVYTPGPLCQPGMGYPMPYPLRVAGVGE
      10      20      30      40      50      60

      40      50      60      70      80
CV127. FSEICFLCHPRRWNPASCKQQHVVREYLVLSSRIPEKCVSSVSFVEPI
      ::  :::::
gi|282 APLLGRARPRRRVPGDCCRQFPPVDYSWCRCEAISHMLGGIYRELGAPDVGHMSEVFR
      70      80      90      100     110     120

gi|282 GCRRGTWSARRRAPGVLQVDIPNGGGGVCYWLARSGY
      130     140     150

```

>>gi|5326864 gi|5326864|gb|AAD42074.1|AF144753_1 peroxis (167 aa)
 initn: 28 init1: 28 opt: 54 Z-score: 85.4 bits: 21.1 E(): 7.9
 Smith-Waterman score: 54; 26.829% identity (63.415% similar) in 41 aa overlap
 (35-75:46-84)

```

                                10      20      30      40
CV127.                        ATIRDHRQHMFLVRTNHIQQVRSWCIQHCLNSFSEICFLCHPRRWNPAS
                                .. .. : ::.
gi|532 MSLKAGDSFPEGVTFSYIPWAEDASEITSCGIPINYNASKEFANKKVVLFPALPGAFTPV-
      10      20      30      40      50

      50      60      70      80
CV127. CCKQQHVVREYLVLSSRIPEKCVSSVSFVEPI
      :. .:: ::. : .. : :::.
gi|532 -CSANHVPEYIQKLPELRAKGVDQVAVLAYNDAYVMSAWGKANGVTGDDILFLSDPEAKF
      60      70      80      90      100     110

gi|532 SKSIGWADEEGRTYRYVLVIDNGKIIYAAKEAAKNSLELSRADHVLKQL
      120     130     140     150     160

```

>>gi|94471622 gi|94471622|gb|ABF21077.1| icarapin varian (223 aa)
 initn: 42 init1: 42 opt: 54 Z-score: 84.1 bits: 21.3 E(): 9.3
 Smith-Waterman score: 54; 52.632% identity (84.211% similar) in 19 aa overlap
 (54-71:97-115)

```

                                10
CV127.                        ATIRDHRQHMFLVRTNH

gi|944 MKTLGVLFIAAWFIACHTSFPGAHDEDSKEERKNVDTVLVLPSIERDQMMMAATFDFPSSL
      10      20      30      40      50      60

```



The Chemical Company
Plant Science L. L. C.

BASF Reg. Doc. No. 2008/7019371

```

      20      30      40      50      60      70
CV127. IQQVRSWCIQHCLNSFSEICFLCHPRRWNPASCKQQHVREYLV-LLSRIPEKCVSSVSF
      ..::: .. ::::~::~:
gi|944 FEDSDEGSNWNWNTLLRPNFLDGWYQTLQSAISAHMKKVREQMAGILSRIPEQGVVWNK
      70      80      90      100     110     120

      80
CV127. VEPI

gi|944 IPEGANTTSTTKIIDGHVVTINETTYTDGSDDYSTLIRVRVIDVRPQNETILTTSSEAD
      130     140     150     160     170     180

gi|944 SDVTTLPTLIGKNETSTQSSRSVESVEDFDNEIPKNQGDVLTA
      190     200     210     220
```

80 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:36:56 2008 done: Sun Dec 14 22:36:56 2008
Total Scan time: 0.090 Total Display time: 0.030

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.4252 c1

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.4252 c1, 72 aa

```
vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library
```

292011 residues in 1313 sequences

Expectation n fit: $\rho(\ln(x)) = 4.7932 \pm 0.00412$; $\mu = 4.1898 \pm 0.214$

```
mean_var=55.9193+/-14.820, 0's: 6 Z-trim: 7 B-trim: 93 in 1/43
```

Lambda= 0.171512

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2

```
join: 36, opt: 20, open/ext: -10/-2, width: 16
```

The best scores are:

The best scores are:					opt bits	E(1313)	%_id	%_sim	alen
gi 27818335	gi 27818335	gb AAO24900	(132)	76	26.0	0.18	0.432	0.541	37
gi 13430402	gi 13430402	gb AAK25823	(275)	65	23.5	2.2	0.533	0.733	15
gi 28630919	gi 28630919	gb AAO45607	(269)	60	22.2	5.1	0.286	0.510	49
gi 398830	gi 398830	emb CAA52753.1	(312)	60	22.3	5.8	0.308	0.590	39
gi 168314	gi 168314	gb AAA63278.1	(252)	58	21.7	6.8	0.375	0.521	48
gi 75274600	gi 75274600	sp Q9SC98	Q (263)	58	21.7	7.1	0.375	0.521	48
gi 126385	gi 126385	sp P14946.2	MPA (263)	58	21.7	7.1	0.375	0.521	48
gi 33149333	gi 33149333	gb AAP96759	(240)	57	21.5	7.8	0.375	0.521	48
gi 3309039	gi 3309039	gb AAC25994.1	(312)	58	21.8	8.1	0.371	0.629	35
gi 18093991	gi 18093991	emb CAD2040	(264)	57	21.5	8.4	0.375	0.521	48
gi 14193761	gi 14193761	gb AAK56124	(269)	57	21.5	8.5	0.281	0.484	64
gi 75139991	gi 75139991	sp Q7M1Y1	Q (35)	47	18.5	8.9	0.636	0.727	11
gi 170724	gi 170724	gb AAA34282.1	(297)	57	21.5	9.3	0.231	0.577	52
gi 162650	gi 162650	gb AA62707.1	(93)	51	19.7	10	0.615	0.692	13

>>>CV127.4252 c1, 72 aa vs

/n/na4/bioinfo/refsets/PUB AllergenOnline12Dec2007 V8.fasta library

```
>>q1|27818335 q1|27818335|qb|AA024900.1| major pollen al (132 aa)
```

```
initn: 57 initl: 57 opt: 76 Z-score: 114.7 bits: 26.0 E(): 0.18
```

Smith-Waterman score: 76; 43.243% identity (54.054% similar) in 37 aa overlap (33-65:80-115)

CV127. MPRCWLFNKSTFT 10

qi | 278 MAKCSYVFCVLLIFIVAIGEMEAAGSKLCEKTSKTYSGKCDNKKCDKKCIEWEKAOHGA

10 20 30 40 50 60

20 30 40 50 60

CV127. GRFSTRIVASSCTFIIKLPSISITTIASGLATDAPIAAGSPNPPIA----PRPPDDSHCL

: : : : : : : : : : : :

qi | 278 CHKREAGKESFCYFDCSKSPPGATPAPPG-AAPPPAAGGSPSPPADGGSPPPPADGGSP

gi|278 PVDGGSPPPPSTH

120 130


```
>>gi|13430402 gi|13430402|gb|AAK25823.1| group V allerge (275 aa)
  initn: 65 initl: 65 opt: 65 Z-score: 95.3 bits: 23.5 E(): 2.2
Smith-Waterman score: 65; 53.333% identity (73.333% similar) in 15 aa overlap
(49-63:1-15)
```

CV127. MPRCWLFNKSTFTGRFSTRIVASSCTFIKLPSPSISTTIASGLATDAPIAAGSPNPIAP
gi|134 PISVTAPPPQLP

```

              70
CV127. RPPDDSHCLGFL
      :::
gi | 134 RPPATPPPPPPQLGASPYKLGPSPKARSERPAIVPPADKYRTFVATFGAASNKAFAEGL
          20          30          40          50          60          70

gi | 134 SGEPKGGAESSKAALTSKLDAAVKLAYKTAEGATPEAKYDAYVATLSEALRIIAGTLEV
          80          90          100         110         120         130

gi | 134 HAVKPAAEVVKVIPAGELQVIEKVDAAFKVAATAANAAPANDKFTVFEEAFNNAIKAGTG
          140         150         160         170         180         190

gi | 134 GAYESYKFIPALEAAVKQAYAATVATAPEVKYTVFETALKKAITAMSEAQKAAPAAAAT
          200         210         220         230         240         250

gi | 134 ATATSAVGAPTGATTAAAGGYKV
          260         270

```

```
>>gi|28630919 gi|28630919|gb|AA045607.1| beta-expansin 9 (269 aa)
  initn: 44 initl: 44 opt: 60 Z-score: 88.8 bits: 22.2 E(): 5.1
Smith-Waterman score: 60; 28.571% identity (51.020% similar) in 49 aa overlap
(20-68:19-66)
```

```

              10              20              30              40              50              60
CV127.  MPRCWLFNKSTFTGRFSTRIVASSCTFI IKLPSPSISTT IASGLATDAP I AAGSPNPIAP
              . . . . :              : . . . . :              : . . . . :
gi | 286  MGSLANNIMVVGAVL AALVVG GSCGPPKVP PPGPNIT TNYNGKWL TARATWY GQPNG-AG
              10              20              30              40              50

              70
CV127.  RPPDDSHCLGFL
              : . . :
gi | 286  APDNGGACGIK NVNLPPYSGMTACGNVPIFKD GKGCGSCYEVRC KEKPECSGNPVTVFIT
              60              70              80              90              100             110

gi | 286  DMNYEPIAPYHFDL SGKAFGSLAKPGLNDKLRHCGIMDVEFRRVRCKYPAGQKIVFHIEK
              120             130             140             150             160             170

gi | 286  GCNPNYVAVLVK FVADDGDIVLMEIQDKLSAEWKPMKLSWGAIWRMDTAKALKGPF SIRL
              180             190             200             210             220             230

gi | 286  TSESGKKVIAKDIIPANWRPDAVYTSNVQFY
              240             250             260

```

```
>>gi|398830 gi|398830|emb|CAA52753.1| Phlp5 [Phleum prat (312 aa)
initn: 46 initl: 46 opt: 60 Z-score: 87.8 bits: 22.3 E(): 5.8
Smith-Waterman score: 60; 30.769% identity (58.974% similar) in 39 aa overlap
(28-66:16-53)
```

```

      10      20      30      40      50      60
CV127. MPRCWLFNKSTFTGRFSTRIVASSCTFIKLPSPSISTTASGLATDAPIAAGSPNPIAP
      ..  :  :  ..  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|398      MAVHQYTVLFLAVALVAGPAASYAADLGYGPATPAAPAAGY-TPATP
      10      20      30      40

```

```

      70
CV127. RPPDDSHCLGFL
      :  ..
gi|398 AAPAEAAPAGKATTEEQKLIKINAGFKAALAAAAGVQPADKYRTFVATFGAASNKAFAE
      50      60      70      80      90      100

```

```

gi|398 GLSGEPKGAAESSKAALTSKLDAAAYKLAYKTAEGATPEAKYDAYVATLSEALRIIAGTL
      110      120      130      140      150      160

```

```

gi|398 EVHAVKPAEEVKVIPAGELQVIEKVDAAFKVAATAANAAPANDKFTVFEEAFNDAIKAS
      170      180      190      200      210      220

```

```

gi|398 TGGAYESYKFIPALEAAVKQAYAATVATAPEVKYTVFETALKKAITAMSEAQKAAKPAAA
      230      240      250      260      270      280

```

```

gi|398 ATATATAAVGAATGAATAATGGYKV
      290      300      310

```

>>gi|168314 gi|168314|gb|AAA63278.1| pollen allergen (252 aa)
 initn: 71 initl: 45 opt: 58 Z-score: 86.5 bits: 21.7 E(): 6.8
 Smith-Waterman score: 58; 37.500% identity (52.083% similar) in 48 aa overlap
 (24-70:83-125)

```

CV127. M
gi|168 ALFAVFLGSAHGIKVPPGPNITAEYGDKWLDAKSTWYGKPTGAGPKDNGGACGYKDVK
      10      20      30      40      50      60

```

```

      10      20      30      40      50      60
CV127. PRCWLFNKSTFTGRFSTRIVASSCTFIKLPSPSISTTASGLATDAPIAAGSPNPIAP
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|168 APFNGMTGCGNTPIFKDGRCGSC-FEIKCTKPES----CSGEAVTVTITDDNEEPIAPY
      70      80      90      100      110

```

```

      70
CV127. PPDDS-HCLGFL
      :  :  :  :
gi|168 HFDLSGHAFGSMKKGEEQNVRSALELQFRRVKCKYPDDTKPTFHVEKGSNPNYLAIL
      120      130      140      150      160      170

```

```

gi|168 VKYVDGDDVVAVDIKEKGDKWIELKESWGA VWRIDTPDKLTGPFTVRYTTEGGTKSEV
      180      190      200      210      220      230

```

```

gi|168 EDVIPEGWKADTSYSAK
      240      250

```

>>gi|75274600 gi|75274600|sp|Q9SC98|Q9SC98_LOLPR Pollen (263 aa)
 initn: 71 initl: 45 opt: 58 Z-score: 86.2 bits: 21.7 E(): 7.1
 Smith-Waterman score: 58; 37.500% identity (52.083% similar) in 48 aa overlap
 (24-70:94-136)

```

gi|752 MASSSSVLLVVALFAVFLGSAHGIKVPPGPNITAEYGDKWLDAKSTWYGKPTGAGPKDN
      10      20      30      40      50      60

```

```

              10      20      30      40      50
CV127.      MPRCWL FNKSTFTGRFSTRIVASSCTFIIKLPSPSISTTIIASGLATDAPI
              :: : :: . : : : :
gi|752 GGACGYKDVDKAPFNGMTGCGNTPIFKDGRGCGSC-FEIKCTKPES----CSGEAVTVTI
              70      80      90      100      110

              60      70
CV127. AAGSPNPIAPRPPDDS-HCLGFL
              . . . : : : : : : :
gi|752 TDDNEEPIAPYHFDLSGHAFGSMAKKGEEQKLRSAGELELQFRRVKCKYPDGTKPTFHVE
              120      130      140      150      160      170

gi|752 KASNPNYLAILVKYVDGDGDVVAVDIKEKGDKWIELKESWGAVWRIDTPDKLTGPFTVR
              180      190      200      210      220      230

gi|752 YTTEGGTKSEVEDVIPEGWKADTSYSAK
              240      250      260

>>gi|126385 gi|126385|sp|P14946.2|MPAL1_LOLPR Pollen all (263 aa)
  initn: 71 initl: 45 opt: 58 Z-score: 86.2 bits: 21.7 E(): 7.1
Smith-Waterman score: 58; 37.500% identity (52.083% similar) in 48 aa overlap
(24-70:94-136)

gi|126 MASSSSVLLVVALFAVFLGSAHGIKVPVPGPNITAEYGDKWLDKSTWYGKPTGAGPKDN
              10      20      30      40      50      60

              10      20      30      40      50
CV127.      MPRCWL FNKSTFTGRFSTRIVASSCTFIIKLPSPSISTTIIASGLATDAPI
              :: : :: . : : : :
gi|126 GGACGYKNVDKAPFNGMTGCGNTPIFKDGRGCGSC-FEIKCTKPES----CSGEAVTVTI
              70      80      90      100      110

              60      70
CV127. AAGSPNPIAPRPPDDS-HCLGFL
              . . . : : : : : : :
gi|126 TDDNEEPIAPYHFDLSGHAFGSMAKKGEEQNVR SAGELELQFRRVKCKYPDDTKPTFHVE
              120      130      140      150      160      170

gi|126 KASNPNYLAILVKYVDGDGDVVAVDIKEKGDKWIELKESWGAVWRIDTPDKLTGPFTVR
              180      190      200      210      220      230

gi|126 YTTEGGTKSEFEDVIPEGWKADTSYSAK
              240      250      260

>>gi|33149333 gi|33149333|gb|AAP96759.1| group 1 allerge (240 aa)
  initn: 70 initl: 44 opt: 57 Z-score: 85.5 bits: 21.5 E(): 7.8
Smith-Waterman score: 57; 37.500% identity (52.083% similar) in 48 aa overlap
(24-70:71-113)

              10
CV127.      MPRCWL FNKSTFT
              10

gi|331 IPKVPPGP NITATYGDKWLDKSTWYGKPTGAGPKDNGGACGYKDVDKAPFNGMTGCGNT
              10      20      30      40      50      60

              20      30      40      50      60      70
CV127. GRFSTRIVASSCTFIIKLPSPSISTTIIASGLATDAPIAAGSPNPIAPRPPDDS-HCLGFL
              :: : :: . : : : : . : : : : : : : : : : : : : : : :
gi|331 PIFKDGRGCGSC-FEIKCTKPES----CSGEAVTVHITDDNEEPIAPYHFDLSGHAFGSM
              70      80      90      100      110

gi|331 AKKGEEQKLRSAGELELQFRRVKCKYPEGKLTTFHVEKGSNPNYLAILVKYVDGDGDVVA
              120      130      140      150      160      170

```

gi | 331 VDIKEKGKDKWIALKESWGA IWRVDT PDKLTGPFTVRYTTEGGTKSEVEDVIPEGWKADT
180 190 200 210 220 230gi | 331 SYEAK
240

```
>>gi|3309039 gi|3309039|gb|AAC25994.1| group V allergen (312 aa)
initn: 45 initl: 45 opt: 58 Z-score: 85.1 bits: 21.8 E(): 8.1
Smith-Waterman score: 58; 37.143% identity (62.857% similar) in 35 aa overlap
(28-60:16-50)
```

CV127. MPRCWLFNKSTFTGRFSTRIVASSCTFI¹⁰IKL²⁰SPSISTT³⁰IASGLATDAPIAAG-SP-NPI⁴⁰
gi|330 MAVHQYTVALFLAVALVAGPAGSYAADLGYGPATPAAPAAGYTPATPA⁵⁰

CV127. APRPPDDSHCLGFL
::
gi|330 APAGAEPAKGATTEEQKLIKINAGFKAALAAAAGVPPADKYRTFVATFGAASNKAFAEG

gi|330 LSGEPKGAAESSKAALTSKLDAAAYKLAYKTAEGATPEAKYDAYVATVSEALRI IAGTLE
110 120 130 140 150 160gi|330 VHAVKPAAEEVKVIPAGELQVIEKVDAAFKVAATAANAAPANDKFTVFEEAFNDAIKAST
170 180 190 200 210 220gi|330 GGAYESYKFIPALEAAVKQAYAATVATAPEVKYTVFETALKKAITAMSEAQKAAKPAAAA
230 240 250 260 270 280

```
gi|330 TATATAAVGAATGAATAATGGYKV
      290      300      310
```

```
>>gi|18093991 gi|18093991|emb|CAD20406.1| unnamed protei (264 aa)
initn: 70 initl: 44 opt: 57 Z-score: 84.9 bits: 21.5 E(): 8.4
Smith-Waterman score: 57; 37.500% identity (52.083% similar) in 48 aa overlap
(24-70:95-137)
```

gi|180 MASSSSSVLLVVALFAVFLGSAHGIPKVPVPGPNITATYGDKWLDKSTWYGKPTGAGPKD

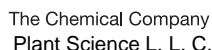
CV127. 10 20 30 40
MPRCWLFNKSTFTGRFSTRIVASSCTFI IKLPSPSISTTIIASGLATDAP
 :: : : : :
gi|180 NGGACGYKDVDKAPFNGMTGCGNTPIFKDGRGCSC- FEIKCTKPES---CSGEAVTVH
 70 80 90 100 110

```

      50          60          70
CV127. IAAGSPNPIAPRPPDDS-HCLGFL
      . . . . . : : : : .
gi|180 ITDDNEEPIAPYHFDLSGHAFGSMAKKGEEQKLRSAGELELQFRRVKCKYPEGTKVTFHV
      120          130          140          150          160          170

```

gi|180 EKGSNPNYLALLVKYVDGDGDVVAVDIKEKGKDKWIALKESWGAIWRVDTPKLTGPFVT
180 190 200 210 220 230gi|180 RYTTEGGTKSEVEDVIPEGWKADTSYEAK
240 250 260



```
>gi|14193761 gi|14193761|gb|AAK56124.1|AF332174_1 beta- (269 aa)
initn: 44 initl: 44 opt: 57 Z-score: 84.8 bits: 21.5 E(): 8.5
Smith-Waterman score: 57; 28.125% identity (48.438% similar) in 64 aa overlap
(6-68:4-66)
```

CV127. MPRCWLFNKSTFTGRFSTRIVAS-SCTFIKLPSPSISTTIIASGLATDAPIAAGSPNPPIA
gi | 141 MGSLVNNIMVVGAVLAALVAGGSCGPPKVPGPNNITNYNGKWL TARATWYGPNG-A

```

      60      70
CV127. PRPPDDSHCLGFL
      :  :  :
gi|141 GAPDNGGACGIKNVNLPPYSGMTACGNVPIFKDGKCGSCYEVRCKEKEPCSGNPFVTVYI
      60      70      80      90     100     110

```

```
gi|141 TDMNYEPIAPYHFDLSGKAFGSLAKPGLNDKIRHCGIMDVEFRRVRCKYPAGQKIVFHIE
      120          130          140          150          160          170
```

gi|141 KGCNPNYLAVLVKYVADDGDIVLMEIQDKLSAEWKPMKLSWGAIWRMDTAKALKGPFSIR
180 190 200 210 220 230

gi|141 LTSESGKKVIAKDVIPANWRPDAVYTSNVQFY
240 250 260

```
>>gi|75139991 gi|75139991|sp|Q7M1Y1|Q7M1Y1_FESAR Group I (35 aa)
  initn: 47 initl: 47 opt: 47 Z-score: 84.5 bits: 18.5 E(): 8.9
Smith-Waterman score: 47; 63.636% identity (72.727% similar) in 11 aa overlap
(5-15:18-28)
```

```

                                10      20      30      40
CV127.      MPRCWLFNKSTFTGRFSTRIVASSCTFI IKLPSPSISTTIIASGLATD
              ::  ::::  :.
gi|751  IAKVPPGPNITAEYGDKWLDAKSTFYGKPTGAGPK
              10      20      30

```

CV127. 50 60 70
APIAAGSPNPIAPRPPDDSHCLGFL

```
>>gi|170724 gi|170724|gb|AAA34282.1| pre-alpha-/beta-gli (297 aa)
  initn: 50 initl: 50 opt: 57 Z-score: 84.1 bits: 21.5 E(): 9.3
Smith-Waterman score: 57; 23.077% identity (57.692% similar) in 52 aa overlap
(20-67:12-63)
```

CV127. MPRCWLFNKSTFTGRFSTRIVASSCTFIKLPSPSISTTIIASGLATD--APIAAGS--PN
gi|170 MKTFLILALRAIVATTATIAVRVPVPQLQPQNPSQQQPQKQVPLVQQQFPFG

CV127. PIAPRPPDDSHCLGFL
: : . .
gi|170 QQQFPPQQPYQQQPFPSQQPYMQLQPFPPQLPYPPQQLPYPPQPFRRQQSYPPQPPQ
60 70 80 90 100 110

gi|170 QYSQPQQPISQQQQQQQQQQQQQQQILQQILQQQLIPCRDVVLQQHSIAHGSSQVLQOST
120 130 140 150 160 170

gi|170 YQLVQQFCCQQLWQIPEQSRCQAIHNVVHAILHQQQQQQQQQQQQQQQPLSQVCFQQSQ

gi|170 QQYPSGGGSFQPSQQNPQAQGSVQPQQLPQFEEIRNLALETLPAMCNVYIPPYCTIAPVG
240 250 260 270 280 290

gi|170 IFGTN

>>gi|162650 gi|162650|gb|AAA62707.1| alpha-s1-casein (93 aa)
initn: 51 initl: 51 opt: 51 Z-score: 83.5 bits: 19.7 E(): 10
Smith-Waterman score: 51; 61.538% identity (69.231% similar) in 13 aa overlap
(46-58:68-80)

CV127. MPRCWLFNKSTFTGRFSTRIVASSCTFIIKLPSISIT
10 20 30

gi|162 PQLEIVPNSAEERLHSMKEGIHAQQKEPMIGVNQELAYFYPELFRQFYQLDAYPSGAWYY
10 20 30 40 50 60

CV127. TIASGLATDAPIAAGSPNPIAPRPPDDSHCLGFL
40 50 60 70
:::: . ::::

gi|162 VPLGTQYTDAPSFSDIPNPIGLENSEKTTMPLW
70 80 90

72 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:34:05 2008 done: Sun Dec 14 22:34:06 2008
Total Scan time: 0.100 Total Display time: 0.030

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.1365_c2

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.1365_c2, 65 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 4.2246 \pm 0.00396$; $\mu = 7.0568 \pm 0.207$

mean_var=42.1608 \pm 11.699, 0's: 17 Z-trim: 17 B-trim: 15 in 1/43

Lambda= 0.197524

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2

join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:		opt	bits	E(1313)	%_id	%_sim	alen
gi 71057064	gi 71057064 emb CAI3879 (225)	59	23.0	2.4	0.500	0.786	14
gi 21913174	gi 21913174 gb AAM77471 (115)	56	21.9	2.4	0.364	0.576	33
gi 1842045	gi 1842045 gb AAB47552.1 (157)	56	22.0	3.2	0.364	0.576	33
gi 395407	gi 395407 emb CAA44345.1 (107)	54	21.3	3.4	0.333	0.604	48
gi 39840779	gi 39840779 emb CAD2170 (152)	52	20.9	6.8	0.314	0.657	35

>>>CV127.1365_c2, 65 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|71057064 gi|71057064|emb|CAI38795.2| thaumatin-like (225 aa)
initn: 59 initl: 59 opt: 59 Z-score: 94.8 bits: 23.0 E(): 2.4
Smith-Waterman score: 59; 50.000% identity (78.571% similar) in 14 aa overlap
(50-63:182-195)

gi|710 MSTFKSLSLFIAFLFTCARGATFNIINNCPFTVWAAAVPGGKRLDRGQNWIIINPG
10 20 30 40 50 60

gi|710 AGTKGARVWPRTGCFNFDGAGRGKCQTGDCNGLLQCQAFGQPPNTLAELYALNQFNNLDFFD
70 80 90 100 110 120

CV127. MISKFQKTLIYRSKIGNEIFLRKGVTLAIGFVTALDETNTALSEISER
10 20 30 40

gi|710 ISLVDGFNVAMEFSPTSGGCTRGIKCTADINGQCPNELRAPGGCNPCTVFKTDQYCCNS
130 140 150 160 170 180

CV127. IDCGINGFHDFFLDRNW
50 60
.:...: :: ::

gi|710 GNCGLTNFSKFFKDRCPDAYSYPKDDQTSTFTCPAGTNYKVVFPCP
190 200 210 220

>>gi|21913174 gi|21913174|gb|AAM77471.1| major allergen (115 aa)
initn: 54 initl: 54 opt: 56 Z-score: 94.5 bits: 21.9 E(): 2.4
Smith-Waterman score: 56; 36.364% identity (57.576% similar) in 33 aa overlap
(35-64:70-102)

CV127. MISKFQKTLIYRSKIGNEIFLRKGV
10 20

gi|219 MQFTTIASLFAAAGLAAAAPLESRQDNASCPVTTKGDYVWKISEFYGRKPEGTYYNLSLGF
10 20 30 40 50 60

```

          30          40          50          60
CV127. TLAIGFVTALDETNALSEISER---IDCGINGFHDFFLDRNW
          :: : .: .: : .: : .: : .: : .: : .: : .
gi|219 NIKATNGGTLDFTCSAQADKLEDHKWYSCGENSFMDFSFSDSDRSGLLLKQKVSDE
          70          80          90          100          110

>>gi|1842045 gi|1842045|gb|AAB47552.1| major allergen A1 (157 aa)
  initn: 54 initl: 54 opt: 56 Z-score: 92.5 bits: 22.0 E(): 3.2
Smith-Waterman score: 56; 36.364% identity (57.576% similar) in 33 aa overlap
(35-64:70-102)

```

```

                                10          20
CV127.                                MISKFQKTLIYRSKIGNEIFLRKGV
gi|184 MQFTTIASLFAAAGLAAAAPLESRQDTASCPVTTEGDYVWKISEFYGRKPEGTYYNLGLF
          10          20          30          40          50          60

```

```

          30          40          50          60
CV127. TLAIGFVTALDETNALSEISER---IDCGINGFHDFFLDRNW
          :: : .: .: : .: : .: : .: : .: : .: : .
gi|184 NIKATNGGTLDFTCSAQADKLEDHKWYSCGENSFMDFSFSDSDRSGLLLKQKVSDITYVA
          70          80          90          100          110          120

gi|184 TATLPNYCRAGNGPKDFVCQGVADAYITLVTLPKSS
          130          140          150

```

```

>>gi|395407 gi|395407|emb|CAA44345.1| fel d I chain 2 pr (107 aa)
  initn: 29 initl: 29 opt: 54 Z-score: 91.9 bits: 21.3 E(): 3.4
Smith-Waterman score: 54; 33.333% identity (60.417% similar) in 48 aa overlap
(16-62:37-77)

```

```

                                10          20          30
CV127.                                MISKFQKTLIYRSKIGNEIFLRKGVTLAIGFVTALDETN
                                ::::: .: .: : .: :
gi|395 MRGALLVLALLVTQALGVKMAETCPIFYDVFFAVANGNELLL----DLSLTKVNATEPER
          10          20          30          40          50

```

```

          40          50          60
CV127. TALSEISERIDCGI-NGFHDFFLDRNW
          ::::: .: .: : .: :
gi|395 TAMKKIQ---DCYVENGLISRVLDGLVMIAINEYCMGEAVQNTVEDLKLNTLGR
          60          70          80          90          100

```

```

>>gi|39840779 gi|39840779|emb|CAD21706.2| Cu /Zn super-o (152 aa)
  initn: 37 initl: 37 opt: 52 Z-score: 86.5 bits: 20.9 E(): 6.8
Smith-Waterman score: 52; 31.429% identity (65.714% similar) in 35 aa overlap
(23-57:12-45)

```

```

          10          20          30          40          50          60
CV127. MISKFQKTLIYRSKIGNEIFLRKGVTLAIGFVTALDETNALSEISERIDCGINGFHDF
          .: .: : .: : .: : .: : .: : .: : .: :
gi|398 MVKAVTVLNSSEGVTGTQEGDGPTTVTGNLSG-LKPGLHGFHVHA
          10          20          30          40

```

```

CV127. LDRNW
gi|398 LGDTTNGCMSTGPHFNPVGKEHGAPGDENRHAGDLGNITVGEDGTAAINIVDKQIPLTGP
          50          60          70          80          90          100

gi|398 HSIIGRAVVVHSDPDDLGRGGHELKSTGNAGGRVACGIIGLQG
          110          120          130          140          150

```




The Chemical Company
Plant Science L. L. C.

BASF Reg. Doc. No. 2008/7019371

65 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 21:58:07 2008 done: Sun Dec 14 21:58:08 2008
Total Scan time: 0.080 Total Display time: 0.010

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.1559_c3

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.1559_c3, 41 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 4.2049 \pm 0.00345$; $\mu = 2.3228 \pm 0.178$
mean_var=36.1942 \pm 8.920, 0's: 23 Z-trim: 23 B-trim: 45 in 1/43
Lambda= 0.213184

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:		opt	bits	E(1313)	%_id	%_sim	alen
gi 1173367	gi 1173367 sp P18632.2 S (374)	55	23.3	2	0.324	0.676	34
gi 19570315	gi 19570315 dbj BAB8628 (374)	55	23.3	2	0.324	0.676	34
gi 493634	gi 493634 dbj BAA05543.1 (374)	55	23.3	2	0.324	0.676	34
gi 51093373	gi 51093373 gb AAT95008 (346)	54	23.0	2.3	0.452	0.645	31
gi 19847822	gi 19847822 gb AAK27264 (306)	51	22.0	3.9	0.471	0.706	17
gi 22353013	gi 22353013 gb AAK97787 (195)	49	21.3	4	0.423	0.654	26
gi 14279169	gi 14279169 gb AAK58515 (460)	52	22.4	4.5	0.583	0.833	12
gi 169929	gi 169929 gb AAB01374.1 (639)	52	22.4	6.1	0.364	0.515	33
gi 2317674	gi 2317674 dbj BAA21760. (504)	51	22.1	6.1	0.448	0.586	29
gi 139002766	gi 139002766 dbj BAF51 (225)	46	20.4	8.6	0.381	0.619	21
gi 1405736	gi 1405736 emb CAA35188. (144)	44	19.8	8.8	0.323	0.581	31
gi 19009	gi 19009 emb CAA46705.1 C (148)	44	19.8	9	0.323	0.581	31

>>>CV127.1559_c3, 41 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|1173367 gi|1173367|sp|P18632.2|SBP_CRYJA Sugi basic (374 aa)

initn: 50 initl: 50 opt: 55 Z-score: 96.2 bits: 23.3 E(): 2

Smith-Waterman score: 55; 32.353% identity (67.647% similar) in 34 aa overlap (7-39:157-190)

gi|117 MDSPCLVALLVLSFVIGSCFSDNPIDSCWRGDSNWAQNRMKLADCAVGFSGSSTMGGKGGD
10 20 30 40 50 60

gi|117 LYTVTNSDDDPVNPAPGTLRYGATRDRPLWIIIFSGNMNIKLKMPMYIAGYKTFDGRGAQV
70 80 90 100 110 120

CV127. MLLIGMKTSFTKNPTNPITTNPIAVRTAT
10 20 30

gi|117 YIGNGGPCVFIKRVSNVHGLHLYGCSTSVLGNVLINESFGVEPVHPQDGDALTLRTAT
130 140 150 160 170 180

CV127. FVNSD-NQFNYK
40

gi|117 NIWIDHNSFSNSSDGLVDVTLSSSTGVTISNNLFFNHHKVMLLGHDDAYSDDKSMKVTVAF
190 200 210 220 230 240

gi|117 NQFGPNCGQRMPRARYGLVHVANNNDPWTIYAIGGSSNPITLSEGNSTAPNESYKKQV
250 260 270 280 290 300

gi|117 TIRIGCKTSSSCSNWVWQSTQDVFYNGAYFVSSGKYEGGNIYTKKEAFNVENG NATPQLT
310 320 330 340 350 360

gi|117 KNAGVLTCSLSKRC
370

>>gi|19570315 gi|19570315|dbj|BAB86286.1| Cry j 1 precu r (374 aa)
initn: 50 initl: 50 opt: 55 Z-score: 96.2 bits: 23.3 E(): 2
Smith-Waterman score: 55; 32.353% identity (67.647% similar) in 34 aa overlap
(7-39:157-190)

gi|195 MDSPCLVALLVFSFVIGSCFSDNPIDSCWRGDSNWAQNRMKLADCAVGFGSSTMGGKGGD
10 20 30 40 50 60

gi|195 LYTVTNSDDDDPVNPAPGTLRYGATRDRLWIIFSGNMNIKLKMPMYIAGYKTFDGRGAQV
70 80 90 100 110 120

CV127. MLLIGMKTSFTKNPTNPITTNPIAVRTAT
10 20 30
.. :: :...:

gi|195 YIGNGGPCVFIKRVSNVIIHGLYLYGCSTSVLGNVLINESFGVEPVHPQDGDALTLRTAT
130 140 150 160 170 180

40
CV127. FVNSD-NQFNKY
. : : . .

gi|195 NIWIDHNSFSNSSDGLVDVTLTSTGVTISNNLFFNHHKVMMLGHDDAYSDDKSMKVTVAF
190 200 210 220 230 240

gi|195 NQFGPNCGQRMPRARYGLVHVANNNDPWTIYAIGGSSNPITLSEGNSTAPNESYKKQV
250 260 270 280 290 300

gi|195 TIRIGCKTSSSCSNWVWQSTQDVFYNGAYFVSSGKYEGGNIYTKKEAFNVENG NATPQLT
310 320 330 340 350 360

gi|195 KNAGVLTCSLSKRC
370

>>gi|493634 gi|493634|dbj|BAA05543.1| Cry j IB precursor (374 aa)
initn: 50 initl: 50 opt: 55 Z-score: 96.2 bits: 23.3 E(): 2
Smith-Waterman score: 55; 32.353% identity (67.647% similar) in 34 aa overlap
(7-39:157-190)

gi|493 MDSPCLVALLVFSFVIGSCFSDNPIDSCWRGDSNWAQNRMKLADCAVGFGSSTMGGKGGD
10 20 30 40 50 60

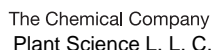
gi|493 LYTVTNSDDDDPVNPAPGTLRYGATRDRLWIIFSGNMNIKLKMPMYIAGYKTFDGRGAQV
70 80 90 100 110 120

CV127. MLLIGMKTSFTKNPTNPITTNPIAVRTAT
10 20 30
.. :: :...:

gi|493 YIGNGGPCVFIKRVSNVIIHGLYLYGCSTSVLGNVLINESFGVEPVHPQDGDALTLRTAT
130 140 150 160 170 180

40
CV127. FVNSD-NQFNKY
. : : . .

gi|493 NIWIDHNSFSNSSDGLVDVTLTSTGVTISNNLFFNHHKVMMLGHDDAYSDDKSMKVTVAF
190 200 210 220 230 240

gi|493 NQFGPNCGRMPRARYGLVHVANNYDPWTIYAIGGSSNPITLSEGSFTAPNESYKKQV
250 260 270 280 290 300gi | 493 TIRIGCKTSSSCSNVWVQSTQDVFYNGAYFVSSGKYEGGNIYTKKEAFNVENG NATPHLT
310 320 330 340 350 360gi|493 QNAGVLTCSLSKRC
370

```
>>gi|51093373 gi|51093373|gb|AAT95008.1| allergen Sol i (346 aa)
  initn: 35 initl: 35 opt: 54 Z-score: 95.0 bits: 23.0 E(): 2.3
Smith-Waterman score: 54; 45.161% identity (64.516% similar) in 31 aa overlap
(11-38:297-326)
```

gi|510 MRKFAAIFVVFVQCTHLYSLAQARAEPDPGVVEYLKQSCVYGNSSYINVYLYNSRFQGK

gi|510 NLGNQQSCQDINASLPVVFITHGFTSSAQVSTFKDLANAFVQKGHTAFIVDWSEAACTDG

gi|510 LPGVQFAEYNAAASNTYDIGQLMAKYTVDL MNKCKIPLNNIQYVGHSLGSHVCGFAAKHV
130 140 150 160 170 180gi|510 KKLINKTMPYILALDPADPSFGSNKCGERICKSDAKRIVVFKTSLGIGENIIGHLLIVF
190 200 210 220 230 240

CV127. MLLIGMKTSETKN 10

gi|510 DGGKSQLPACSWYDVPCHSHSESIYATGMVSGRCQHLAVPWTAAQQRINPIQWKFWRVFTSN
250 260 270 280 290 300

CV127. -PTNP I--TTNP IAVRTATFVNSDNQFNYK
 ::: ::: ::: ::: :::

gi|510 IPAYPTSDTTNCVVLNTNVFKN-DNTFEGEYHAFDPCARNLFKCRQQ
310 320 330 340

```
>>gi|19847822 gi|19847822|gb|AAK27264.1| isoflavone redu (306 aa)
  initn: 51 initl: 51 opt: 51 Z-score: 90.9 bits: 22.0 E(): 3.9
Smith-Waterman score: 51; 47.059% identity (70.588% similar) in 17 aa overlap
(8-24:209-225)
```

gi|198 MGGSRVLIIGGTGYIGRHVTNASLAQGHPTFLLVREITPSNPEKAQLLESFTSKGATLVQ

gi|198 GSIDDHASLVAALKKVDVVISTLGAPQIADQFNLIKAIKEVGTIKRFFPSEFGNDVDKHH
70 80 90 100 110 120gi|198 AVEPMKSMFDLKI¹³⁰KLRR¹⁴⁰TIEAEGIPHTYVVP¹⁵⁰HC¹⁶⁰FAGYFLT¹⁷⁰NLAQLGLAAPP¹⁸⁰RDKIV¹⁹⁰IYGD²⁰⁰

CV127. MLLIGMKTSETKNPTNPITTNPIAVRTATFVNSDNOFN

gi|198 GTTKAVYMKEEDIGTFTIKAVDDPRTLNKTLYLKPPANTISTNDLVALWEAKIGKTLEKV
190 200 210 220 230 240

40
CV127. YK

gi|198 YLSEEQVLKLLQDTPFPPTFMVSIFHTIYVKGDTNFQIGPDGVEASALYPDVKYTTVEE
250 260 270 280 290 300

gi|198 YISAFV

>>gi|22353013 gi|22353013|gb|AAK97787.1| allergenic prot (195 aa)
initn: 30 initl: 30 opt: 49 Z-score: 90.7 bits: 21.3 E(): 4
Smith-Waterman score: 49; 42.308% identity (65.385% similar) in 26 aa overlap
(6-31:118-142)

gi|223 GLEQAFCNLKFQNVNRPSHADVFNPGRINTVNSNNLPILFLQLSAQHVVLKNAII
10 20 30 40 50 60

CV127. MMLLIGMK
:
gi|223 GPRWNLNAHSALYVTRGEGRVQVVGDEGKSVFDDNVQRGQILVVPQGFVAVVKAGRQGLE
70 80 90 100 110 120

10 20 30 40
CV127. TSFTKNPTNPITTNPIAVRTATFVNSDNQFNKY
:
gi|223 WVELKNNDNAITS-PIAGRTSVLRAIPVEVLANSYDISTEEAYKLKNGRQEVEVFRPFQS
130 140 150 160 170

gi|223 RYEKEEEKERERFSIV
180 190

>>gi|14279169 gi|14279169|gb|AAK58515.1|AF249675_1 beta- (460 aa)
initn: 45 initl: 45 opt: 52 Z-score: 89.7 bits: 22.4 E(): 4.5
Smith-Waterman score: 52; 58.333% identity (83.333% similar) in 12 aa overlap
(12-23:353-364)

gi|142 MAANVQTSSLLFLVFLLLQNFYSANSQSFLGVNYGQLSDNLPQLATVNLKSTTIQKVR
10 20 30 40 50 60

gi|142 LFGEAPAVIKAFANTGVEIVIGFDNGDIPTLASNPNVASQFVKSNVMSFYPASNIIAIV
70 80 90 100 110 120

gi|142 GNEVLTSQDQKLISQLLPAMQNVQNALNAASLGKVKVSTVHAMAVLSQSYPPSSGVFNP
130 140 150 160 170 180

gi|142 GLGDTMKALLQFQSANDAPFMISPPYFAYKNQPTPDTLAFCLFQPNAGQVDSGNHGYT
190 200 210 220 230 240

gi|142 NMFDAQVDAVHSALNAMGFKDIEIVVAETGWPHGGDSNEVGPSLDNAKAYVGNLINHLKS
250 260 270 280 290 300

10
CV127. MMLLIGMKTSFTKNPTNP
:
gi|142 KVGTPMPGKSIDTYLFSLYDEDKKTGASSEKYFGLFKPDGSTTYDVGLLKNTQNPTTPA
310 320 330 340 350 360

20 30 40
CV127. TTNPIAVRTATFVNSDNQFNKY
:
gi|142 TPTPTPKAAGSWCVPKPGVSDDQLTGNININACGQIDCGPIQPGGACFEPNTVKAHAAYV
370 380 390 400 410 420

gi|142 MNLYYQSAGRNSWNCDFSQTATLTNTNPSYGACNFPSPGSN
430 440 450 460

>>gi|169929 gi|169929|gb|AAB01374.1| beta-conglycinin st (639 aa)
initn: 49 initl: 49 opt: 52 Z-score: 87.4 bits: 22.4 E(): 6.1
Smith-Waterman score: 52; 36.364% identity (51.515% similar) in 33 aa overlap
(9-41:308-340)

gi|169 MMRARFPLLLLGVVFLASVSVSFGIAYWEKQNP SHNKL RSCNSEKDSYRNQACHARCNL
10 20 30 40 50 60

gi|169 LKVEEEEECEEGQIPRPRPQHPERERQQHGEKEEDEGEQPRPFPPRPRPQPHQEEHEQK
70 80 90 100 110 120

gi|169 EEHEWHRKEEKHGKGSEEEQDEREHPRPHQKKEEKHEWQHKQEKHQKSEEEEEED
130 140 150 160 170 180

gi|169 QDEDEEQDKESQESEGESQREPRRHKNKNPFHFN SKRFQTLFKNQYGHVRVLQRFNKR
190 200 210 220 230 240

CV127. M

gi|169 QQLQNLRDYRILEFNSKPNTLLLP HHADADYLIVILNGTAILTLVNDDRDSYNLQSGDA
250 260 270 280 290 300

10 20 30 40
CV127. MLLIGMKTSTFTKNPTNPITNPIAVRTATFVN SDNQFNK
: . . : : . : : : : : : : .

gi|169 LRVPA GTTFYVVPNDNENLRMIAGTTFYVVPNDNENLRMITLAIPV NKPGRFESFFLS
310 320 330 340 350 360

gi|169 STQAQQSYLQGF SKNILEASYDTKFEEINKVLFGREEGQQQGEERLQESVIVEISKQIR
370 380 390 400 410 420

gi|169 ELSKHAKSSSRKTISSEDKPFLNLSRDP IYSNKLGLFEITQRNPQLRDL DVFLSVVDMN
430 440 450 460 470 480

gi|169 EGALFLPHFNSKAIVVLVINEGEANIELVG IKEQQQRQQQEEQPLEVRKYRAELSEQDIF
490 500 510 520 530 540

gi|169 VIPAGYPVMVNATSDLNFFAFGINAENNQRNFLAGSKDNVISQIP SQVQELAFPRSAKDI
550 560 570 580 590 600

gi|169 ENLIK SQSESYFVDAQPQQKEEGNKGRKGPLSSILRAFY
610 620 630

>>gi|2317674 gi|2317674|dbj|BAA21760.1| legumin-like pro (504 aa)
initn: 42 initl: 42 opt: 51 Z-score: 87.4 bits: 22.1 E(): 6.1
Smith-Waterman score: 51; 44.828% identity (58.621% similar) in 29 aa overlap
(11-36:324-351)

gi|231 MSTKLILSFS LCLMVLSCSAQLWPWQKGQGS RPHHGRQQHQFQHQC DIQRLTASEPSRRV
10 20 30 40 50 60

gi|231 RSEAGVTEIWDHDTPEFRCTGFVAVRVVIQPGGL LPSYSNAPYITFVEQGRGVQGVVIP
70 80 90 100 110 120

gi|231 GCPETFQSDSEFEYPQSQRGRHSRQSESEEESSRGDQH KIFRIREGDVIPSPAGVVQWT
130 140 150 160 170 180

gi | 231 HNDGNDLLISVTLTDANSYHKQLDENVRSFFLAGQSQRETRREGSDRQSRESDDDEALLG
190 200 210 220 230 240gi | 231 ANILSGFQDEILHELFRDVDRETISKLRGENDQRGFIVQAQDLKLRVPQDFEEYERERG
250 260 270 280 290 300

CV127. MMLLIGMKTSFTKN---PTNPITTNPIAVRTATFVNSDNQFNYK

gi|231 DRRRGQGSGRSNGVEQGFCNLKFRNFNTPNTYVFNPAGRINT-VNSNSLPILFLQ
310 320 330 340 350gi|231 LSAQHVVLYKNAIIGPRWNLNAHSALYVTRGEGRVQVVGDEGKSVFDDKVQRGQILVVPQ
360 370 380 390 400 410

gi|231 GFAVVLKAGREGLEWVELKNSGNAITSPIGGRTSVLRAIPVEVLANSYDISTKEAYKLKN
420 430 440 450 460 470

gi|231 GRQEVEVFRPFQSRDEKERERFSIV
480 490 500

```
>>gi|139002766 gi|139002766|dbj|BAF51970.1| thaumatin-li (225 aa)
  initn: 37 initl: 37 opt: 46 Z-score: 84.7 bits: 20.4 E(): 8.6
Smith-Waterman score: 46; 38.095% identity (61.905% similar) in 21 aa overlap
(3-23:122-142)
```

gi|139 MAKVSDLALLLVAGMAISLYIQETGAVKFDIKNQCGYTVWAAGLPGGGQQLTQGQTWTVN
10 20 30 40 50 60

CV127. M

gi|139 LAAGTQSARFWGRTGCSFDASGKGTCQTGDCGGQLSCTVSGAVPATLAEYTDQSDQDYDDV
70 80 90 100 110 120

CV127. MLLIGMKTSTFTKNPTNPITTNPIAVRTATFVNSDNQFNKY

gi | 139 SLVDGFNIPLSINPTNAQCTAPACKADVNAVCPAELKVDGGCKSACAAAFQTDQYCCTGT
130 140 150 160 170 180

gi|139 ANSCPATNYSMIFKNQCPQAYSYPKDDTATFACPSGTDYSIVFCP
190 200 210 220

```
>>gi|1405736 gi|1405736|emb|CAA35188.1| trypsin inhibito (144 aa)
  initn: 29 initl: 29 opt: 44 Z-score: 84.5 bits: 19.8 E(): 8.8
Smith-Waterman score: 44; 32.258% identity (58.065% similar) in 31 aa overlap
(1-28:15-45)
```

CV127. MLLIGMKTSTFTKN--PTNPITTNPI-AVRTATFVNSDNQFNYK

gi|140 MAFKYQLLLSAAVMLAILVATATSGDSCAPGDALPHNPLRACRTYVVSQICHQGPRLLT
10 20 30 40 50 60gi | 140 SDMKRRCCDELSAIPAYCRCEALRIIMQGVVTWQGAFEGAYFKDSPNCPRERQTSYAANL
70 80 90 100 110 120gi|140 VTPQECNLGTIHGSAYCPELQPAY
130 140

```
>>gi|19009 gi|19009|emb|CAA46705.1| CMe [Hordeum vulgare (148 aa)
  initn: 29 initl: 29 opt: 44 Z-score: 84.3 bits: 19.8 E(): 9
Smith-Waterman score: 44; 32.258% identity (58.065% similar) in 31 aa overlap
(1-28:15-45)
```

```

                                10      20      30      40
CV127.      MMLLIGMKTSTFTKN--PTNPITTNPI-AVRTATFVNSDNQFNYK
              . . . . . : : : . . . : : :
gi|190 MAFKYQLLLSAAVMLAILVATATSTFGDSCAPGDALPHNPLRACRTYVVSQICHQGPRLLT
              10      20      30      40      50      60

gi|190 SDMKRRCCELSAIPAYCRCEALRIIMQGVVTWQGAFAFEGAYFKDSPNCPRERQTSYAANL
              70      80      90      100     110     120

gi|190 VTPQECNLGTIHGSAYCPELQPGYGVVL
              130     140
```

```
41 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:00:18 2008 done: Sun Dec 14 22:00:18 2008
Total Scan time: 0.060 Total Display time: 0.020
```

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.1581_c2

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.1581_c2, 51 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 6.0286 \pm 0.00321$; $\mu = -6.3528 \pm 0.166$

mean_var=38.9082 \pm 8.854, 0's: 24 Z-trim: 24 B-trim: 0 in 0/44

Lambda= 0.205614

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2

join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:

	opt	bits	E(1313)	%_id	%_sim	alen
gi 169929 gi 169929 gb AAB01374.1 (639)	61	24.3	2.1	0.333	0.690	42
gi 170732 gi 170732 gb AAA34286.1 (323)	56	23.0	2.6	0.375	0.750	24
gi 5712199 gi 5712199 gb AAD47382.1 (530)	55	22.5	5.8	0.321	0.714	28
gi 1350779 gi 1350779 sp P49148.1 R (110)	45	20.1	6.5	0.333	0.600	30
gi 47117012 gi 47117012 sp Q7M4I5.1 (134)	45	20.0	8.3	0.317	0.610	41
gi 7271239 gi 7271239 emb CAA09886. (179)	46	20.2	9.7	0.476	0.667	21

>>>CV127.1581_c2, 51 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|169929 gi|169929|gb|AAB01374.1| beta-conglycinin st (639 aa)

initn: 37 init1: 37 opt: 61 Z-score: 95.5 bits: 24.3 E(): 2.1

Smith-Waterman score: 61; 33.333% identity (69.048% similar) in 42 aa overlap (7-45:165-206)

gi|169 MMRARFPLLLLVVFLASVSVSFGIAYWEKQNP SHNKL RSCNSEKDSYRNQACHARCNL
10 20 30 40 50 60

gi|169 LKVEEEEECEEGQIPRPRPQHPERERQQHGEKEEDEGEQPRPFPFPRPRPQHQEEEEHEQK
70 80 90 100 110 120

CV127. MRDQLPRKESTTLRRSNDVVD
10 20
..: ..:...

gi|169 EEHEWHRKEEKHGGKGSEEEQDEREHPRPHQPHQKEEEKHEWQHKEKHQKQKESEEEEEED
130 140 150 160 170 180

CV127. WDENELHE-ESNESH HHKSNR--RTHSYFREF
30 40 50
::: .. ::::: ::: :

gi|169 QDEDEEQDKESQSEGESQREPRRHKNKNPFHFN SKRFQTLFKNQYGHVVRVLQRFNKR
190 200 210 220 230 240

gi|169 QQLQNLRDYRILEFN SKPNTLLPHHADADYLIVILNGTAILTLVNDDRDSYNLQSGDA
250 260 270 280 290 300

gi|169 LRVPA GTTFYVVPDNDENLRMIAGTTFYVVPDNDENLRMITLAIPVNKPGRFESFFLS
310 320 330 340 350 360

gi|169 STQAQQSYLQGF SKNILEASYDTKFEEINKVLFGREEGQQQGEERLQESVIVEISKQIR
370 380 390 400 410 420

gi|169 ELSKHAKSSSRKTISSEDKPFNLGSRDPIYSNKLGLKLFETQRNPQLRDLDFVLSVVDMM
430 440 450 460 470 480

gi|169 EGALFLPHFNSKAIVVLVINEGEANIELVGIKEQQQRQQEEQPLEVRKYRAELSEQDIF
490 500 510 520 530 540

gi|169 VIPAGYPVMVNATSDLNFFAFGINAENNQRNFLAGSKDNVISQIPSQVQELAFPRSAKDI
550 560 570 580 590 600

gi|169 ENLIKSQSESYFVDAQPPQKKEEGNKGRKGPLSSILRAFY
610 620 630

>>gi|170732 gi|170732|gb|AAA34286.1| gamma-gliadin (323 aa)
initn: 54 initl: 54 opt: 56 Z-score: 94.1 bits: 23.0 E(): 2.6
Smith-Waterman score: 56; 37.500% identity (75.000% similar) in 24 aa overlap
(26-49:15-36)

10 20 30 40 50
CV127. MRDQLPRKESTTLRRSNDDVVDWDENELHEESNESHHSNRRTHSYFREF
.: : .: : .: : .: :
gi|170 TITRTFPIPTISSNNHHFRSNSNHFFHSN--NNQFYRNNNSPGHNNPL
10 20 30 40

gi|170 NNNNSPNNNSPNHHNNSPNNNFQYHTHPSNHKNLPHTNNIQQQQPPFSQQQQPPFSQQQ
50 60 70 80 90 100

gi|170 QPVLPPQQSPFSQQQQLVLPQQQQQLVQQQIPVQPSVLQQLNPCKVFLQQQCSPVAMP
110 120 130 140 150 160

gi|170 QRLARSQMWWQSSCHVMQQQCCQQLQIQIPEQSRYEAIRAIYSIILQEQQQGFVQPQQQQ
170 180 190 200 210 220

gi|170 PQQSGQGVVSQSQQQSQQQLGQCSFQQPQQQLGQQPQQQQQQVQLQGTFLQPHQIAHLEAV
230 240 250 260 270 280

gi|170 TSIALRTLPTMCSVNVPLYSATTSVPFVGVTGVGAY
290 300 310 320

>>gi|5712199 gi|5712199|gb|AAD47382.1| glycinin [Arachis (530 aa)
initn: 74 initl: 52 opt: 55 Z-score: 87.7 bits: 22.5 E(): 5.8
Smith-Waterman score: 55; 32.143% identity (71.429% similar) in 28 aa overlap
(16-43:309-336)

gi|571 MAKLLELSFCFCFLVLGASSISFRQQPEENACQFQRLNAQRPDNRIESEGGYIETWNPNN
10 20 30 40 50 60

gi|571 QEFECAGVALSRLVLRNALRRPFFYSNAPQEIFIQQGRGYFGLIFPGCPSTYEEPAQQGR
70 80 90 100 110 120

gi|571 RYQSQRPRLQEEDQSQQQDQSHQKVHRFNEGDLIAVPTGVAFWLYNDHDTDVAVSLT
130 140 150 160 170 180

gi|571 DTNNNDNQLDQFPRRNLGNHEQEFLRYQQQSRQSRRLSLPSPYSPHSRPRREEREFR
190 200 210 220 230 240

CV127. MRDQLPR

gi|571 PRGQHSRRERAGQEEDEGGNIFSGFTPEFLEQAFQVDDRQIVQNLWGENESEEEGAIVT
250 260 270 280 290 300

```

      10      20      30      40      50
CV127. KESTTLRRSNDVVDWDENELHEESNESHHSNRRTHSYFREF
      : : . : : . : : : : : : : : : : : : : : : :
gi|571 VRGGLRILSPDGTRGADEEEYDEDQYEHQDGRGRGRSGGGNGIEETICTACVKKNI
      310      320      330      340      350      360

gi|571 GGNRSPHIYDPQRWFTQNCHDLNLLILRWLGLSAEYGNLYRNALFVPHYNTNAHSIIYAL
      370      380      390      400      410      420

gi|571 RGRAHVQVVDSSNGNRVYDEELQEGHVLVVPQNFVAVAGKSQSENFYVAFKTDSPRSIANF
      430      440      450      460      470      480

gi|571 AGENSFIDNLPEEVVANSYGLPREQARQLKNNNPFFKFFVPPFQQSPRAVA
      490      500      510      520      530

>>gi|1350779 gi|1350779|sp|P49148.1|RLA1_ALTAL 60S acidi (110 aa)
  initn: 45 init1: 45 opt: 45 Z-score: 86.9 bits: 20.1 E(): 6.5
Smith-Waterman score: 45; 33.333% identity (60.000% similar) in 30 aa overlap
(6-35:71-100)

gi|135 MSTSELATSYAALILADDGVDITADKLQSLIKAAKIEEVEPIWTTLFAKALEGKDVKDLL
      10      20      30      40      50      60

      10      20      30      40      50
CV127. MRDQLPRKESTTLRRSNDVVDWDENELHEESNESHHSNRRTHSYFREF
      : : . : : . : : : : : : : : : : : : : : : :
gi|135 LNVGSGGGAAPLPEALLRWRAADAAPAAEEKKEEKEEDEDMDGFGFLFD
      70      80      90      100      110

>>gi|47117012 gi|47117012|sp|Q7M4I5.1|PA2_APIDO Phosphol (134 aa)
  initn: 24 init1: 24 opt: 45 Z-score: 85.0 bits: 20.0 E(): 8.3
Smith-Waterman score: 45; 31.707% identity (60.976% similar) in 41 aa overlap
(3-39:19-58)

      10      20      30      40
CV127. MRDQLPR-KESTTLRRSND---DVVDWDENELHEESNESHHSNRRTHSYFREF
      : : : : : : : : : : : : : : : : : : : : : :
gi|471 IIYPGTLWCGHGNVSSSPDELGRFKHTDSCCRSHDMCPDVMSAGESK-HGLTNTASHTRL
      10      20      30      40      50

      50
CV127. NRRTHSYFREF

gi|471 SCDCDDKFYDCLKNSSDTISSYFVGEMYFNILDTKCYKLEHPVTGCGKRTEGRCLNYTVD
      60      70      80      90      100      110

gi|471 KSKPKVYQWFDLRKY
      120      130

>>gi|7271239 gi|7271239|emb|CAA09886.2| allergen [Malass (179 aa)
  initn: 31 init1: 31 opt: 46 Z-score: 83.8 bits: 20.2 E(): 9.7
Smith-Waterman score: 46; 47.619% identity (66.667% similar) in 21 aa overlap
(6-26:102-120)

gi|727 MVALKFAAVLSVVAAMAPSSMDRRASPDNQVWVTSASDYCLILPRHRESIGDSESPG
      10      20      30      40      50      60

      10      20
CV127. MRDQLPRKESTTLRRSNDVVDWD
      : : : : : : : : : : : : : : : : : : : : : :
gi|727 RMRSFCSPYPYDSSQGQINPGFWKEVHFKKTKNYVQLTGCINPRVQSTLL--SHDDGGQYD
      70      80      90      100      110

```

```
          30          40          50
CV127. ENELHEESNESHHSNRRTHSYFREF
      :
gi|727 SNGNGGVGNPEGSVCLGYSSYVELVEPSDGKACIRCCVNDKYCDVGHDEDGCEAVIPGQY
      120          130          140          150          160          170

gi|727 C
```

```
51 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:02:12 2008 done: Sun Dec 14 22:02:12 2008
Total Scan time: 0.080 Total Display time: 0.010
```

```
Function used was FASTA [version 34.26.5 April 26, 2007]
```

CV127.5534_c3

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.5534_c3, 51 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 4.7392 \pm 0.00334$; $\mu = 1.6178 \pm 0.172$

mean_var=39.4535 \pm 10.186, 0's: 29 Z-trim: 29 B-trim: 5 in 1/41

Lambda= 0.204189

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2

join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:				opt	bits	E(1313)	%_id	%_sim	alen
gi 170702	gi 170702	gb AAA34272.1	(302)	59	23.7	1.5	0.260	0.700	50
gi 295774	gi 295774	emb CAA29664.1	(142)	52	21.6	3	0.333	0.741	27
gi 162644	gi 162644	gb AAA30367.1	(142)	52	21.6	3	0.333	0.741	27
gi 27462836	gi 27462836	gb AAO15607	(219)	54	22.2	3.1	0.304	0.630	46
gi 807138	gi 807138	gb AAB32224.1	(219)	53	21.9	3.8	0.286	0.595	42
gi 62484809	gi 62484809	emb CAI7890	(285)	54	22.2	4	0.235	0.647	51
gi 320546	gi 320546	pir B45786 maj	(51)	43	18.9	7.1	0.375	0.667	24
gi 1063270	gi 1063270	dbj BAA11251.	(279)	51	21.3	7.2	0.240	0.680	50
gi 2677826	gi 2677826	gb AAB97141.1	(160)	48	20.4	7.7	0.318	0.591	22
gi 298737	gi 298737	gb AAB25851.1	(51)	42	18.6	8.7	0.458	0.708	24

>>>CV127.5534_c3, 51 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|170702 gi|170702|gb|AAA34272.1| gamma gliadin precu (302 aa)

initn: 38 initl: 38 opt: 59 Z-score: 98.3 bits: 23.7 E(): 1.5

Smith-Waterman score: 59; 26.000% identity (70.000% similar) in 50 aa overlap (2-48:164-213)

gi|170 MKTLLILITILAMATTIATANMQVDPSPGVQWPQQQPFQPPQPPFCQQPQRTIPQPHQTFH
10 20 30 40 50 60

gi|170 HQPQQTFPQPQQTYPHQPQQQFPQTQQPQQPFPQPQQTFPQQPQLPFPQQPQQPFPQPQQ
70 80 90 100 110 120

CV127. MLEKEMKKYKKFLLF-IK
: : :
gi|170 PQQPFPQSQQPQQPFPQPQQQFPQPQQPQQSFPQQQPAIQSFLQQQMNPKNFLLQQCN
130 140 150 160 170 180

CV127. FTTLLI--LVSFLLFFKSFCHLLKKKKLFEIFLLPMSD
: . . : : : :
gi|170 HVSLVSSLVSIILPRSDCQVMQQQCCQQLAQIPQQLQCAAIHSAHSSIIMQQEQQQGVPI
190 200 210 220 230 240

gi|170 LRPLFQLAQGLGIIQPQQPAQLEGIRSLVLKTLPTMCNVYVPPDCSTINVPYANIDAGIG
250 260 270 280 290 300

gi|170 GQ

>>gi|295774 gi|295774|emb|CAA29664.1| alpha-lactalbumin (142 aa)
initn: 42 init1: 42 opt: 52 Z-score: 92.8 bits: 21.6 E(): 3
Smith-Waterman score: 52; 33.333% identity (74.074% similar) in 27 aa overlap
(16-42:2-28)

```

              10      20      30      40      50
CV127. MLEKEMKKYKKFLLFIKFTTLILVSFLFFKSFCHELLKKKKLFEIFLLPMSD
              .....: . . : : .:
gi|295      MMSFVSLLLVGILFHATQAEQLTKCEVFRELKDLKGYGGVSLPEWV
              10      20      30      40

gi|295 CTAFHTSGYDTQAIVQNNDSTEYGLFQINNKIWCKDDQNPSSNICNISCDKFLDDDLTD
              50      60      70      80      90     100

gi|295 DIMCVKKILDKVGINYWLAHKALCSEKLDQWLCEKL
              110     120     130     140

```

>>gi|162644 gi|162644|gb|AAA30367.1| alpha-lactalbumin p (142 aa)
initn: 42 init1: 42 opt: 52 Z-score: 92.8 bits: 21.6 E(): 3
Smith-Waterman score: 52; 33.333% identity (74.074% similar) in 27 aa overlap
(16-42:2-28)

```

              10      20      30      40      50
CV127. MLEKEMKKYKKFLLFIKFTTLILVSFLFFKSFCHELLKKKKLFEIFLLPMSD
              .....: . . : : .:
gi|162      MMSFVSLLLVGILFHATQAEQLTKCEVFRELKDLKGYGGVSLPEWV
              10      20      30      40

gi|162 CTTFHTSGYDTQAIVQNNDSTEYGLFQINNKIWCKDDQNPSSNICNISCDKFLDDDLTD
              50      60      70      80      90     100

gi|162 DIMCVKKILDKVGINYWLAHKALCSEKLDQWLCEKL
              110     120     130     140

```

>>gi|27462836 gi|27462836|gb|AAO15607.1|AF462190_1 gluta (219 aa)
initn: 34 init1: 34 opt: 54 Z-score: 92.7 bits: 22.2 E(): 3.1
Smith-Waterman score: 54; 30.435% identity (63.043% similar) in 46 aa overlap
(2-45:133-177)

```

gi|274 MSSKPTLGWDLRGLGQSIRILLTYAGVDFVDKRYKIGSAPDFDRGEWLNDKFNGLGDFP
              10      20      30      40      50      60

gi|274 NLPYYIEGDVKTQSIAILRYLGRKHKLDGQNEQEWRRITLCEQQIMDLLMALARICYDP
              70      80      90      100     110     120

              10      20      30      40
CV127.      MLEKEMKKYKKFLLFIKFTTLILVSFLFFKSFCHELLKKKKLFEIFLL
              : ..: .....: ...: ...: : ..: ...: : : :
gi|274 NFEKLKLDLVAKLPDDLKLFSGFLGDHQFVAGTNISYIDFLVYEYLIRVK-IFAPEIFTK
              130     140     150     160     170

```

```

              50
CV127. PMSD

gi|274 FPNLNSYITRIESMPKISAYIKQQEPQLFNGPMAKWNTKY
              180     190     200     210

```

>>gi|807138 gi|807138|gb|AAB32224.1| major allergen p Dp (219 aa)
initn: 44 init1: 44 opt: 53 Z-score: 91.1 bits: 21.9 E(): 3.8
Smith-Waterman score: 53; 28.571% identity (59.524% similar) in 42 aa overlap
(4-45:137-176)

gi | 807 MSQPILGYWDIRGYAQPIRLLLTYSGVDFVDKRYQIGPAPDFDRSEWLNEKFNLGLDFPN
10 20 30 40 50 60gi|807 LPYYIDGDMKMTQTFAILRLYLGRKYKLNGSNDHEEIRISMAEQQTEDMMAAMIRVCYDAN
70 80 90 100 110 120

CV127. MLEKEMKKYKKFLLFIKFTTLILVSFLFFKSFCHLLKKKKLFEIFLL
:
gi|807 CDLKP DY LKSLPDCLKLMSKFVGEHAFIAGANISYVD FNLYEYLCHV--KVMVPEVFQG

CV127. PMSD

gi|807 FENLKRYVERMESLPRVSDYIKKQQPKTFNAPTSKWNASYA
180 190 200 210

```
>>gi|62484809 gi|62484809|emb|CAI78902.1| putative gamma (285 aa)
initn: 35 initl: 35 opt: 54 Z-score: 90.7 bits: 22.2 E(): 4
Smith-Waterman score: 54; 23.529% identity (64.706% similar) in 51 aa overlap
(2-48:139-188)
```

gi | 624 MQVNPSVQVQPTQQQPYPESQQPFISQSQQQFPQPQQPFPQQPFPQSQQQCLQQPQH
10 20 30 40 50 60gi|624 QFPQPTQQFPQRLLPFTHPFLTFFDQLLPQPPHQSFPPQSYPPQLQPPFPQPPQKY
70 80 90 100 110 120

```

                                10      20      30
CV127.      MLEKEMKKYKKFLLFIKFTTLLILVSFLFFK---SFCHLLKKK
              : : : : : : : : : : : : : : : : : : : : : :
gi|624 PEQPQQPFPWQQPTIQLYLQQQLNPCKEFL-QQCRPVSLLSYLWSKIVQQSSCRMQQQ
              130      140      150      160      170

```

CV127. KLFEIFLLPMSD

gi|624 CCLQLAQIPEQYKCTAIDSIVHAIFMQQGQRQGVQIVQQQPQPQVVGQCVLVQGGGVVQP
180 190 200 210 220 230

gi|624 QQLAQMEAIRTLVLQSVPSMCNFNVPPNCSTIKAPFVGVTGVGGO
240 250 260 270 280

```
>>gi|320546 gi|320546|pir||B45786 major pollen allergen (51 aa)
initn: 35 init1: 35 opt: 43 Z-score: 86.2 bits: 18.9 E(): 7.1
Smith-Waterman score: 43; 37.500% identity (66.667% similar) in 24 aa overlap
(19-42:9-30)
```

```

          10          20          30          40          50
CV127. MLEKEMKKYKKFLFIKFTTLILVSFLFFKSFCHELLKKKKLFEIFLLPMSD
          :: . . . . :: . :   ::
gi|320      SVFNYETETTSVIPAAMLFAF--ILDGDKLFPKVPAPQXQSIVXNXRVY
          10          20          30          40

```

```
gi | 320  XPK
      50
```

```
>>gi|1063270 gi|1063270|dbj|BAA11251.1| gamma-gliadin pr (279 aa)
  initn: 34 initl: 34 opt: 51 Z-score: 86.1 bits: 21.3 E(): 7.2
```

Smith-Waterman score: 51; 24.000% identity (68.000% similar) in 50 aa overlap (2-48:130-179)

gi | 106 NIQVDPGSGVQWPQQQFPQPHQPFSSQQPQQTFPQPQQTFPHQPQQQFSQPPQPQQQFIQ

gi|106 PQQPFPQQPQQTYPQRPQQPFPQTQQPQQPFPQSQQPQQPFPQPQQQFPQPQQPQQSFPQ
70 80 90 100 110 120

CV127. MLEKEMKKYKKFLLF-IKFTTLI--LVSFLFFKSFCHELLKKKKLFEIFLLPM
gi|106 QQPSLIQQSLQQQLNPFCKNFLQOCKPVSVLSVSSLWSMILPRSDCQVMRQCCQQLAQIPQ

10 20 30 40
130 140 150 160 170 180

CV127. SD

gi|106 QLQCAAHSIVHSIIMQQEQEQEQRGVQILVPLSQQQQVVGQGLVQGGQIIQPQQPAQLE

gi|106 VIRSSVLQTLATMCNVYVPPYCSTIRAPFASIVAGIGGQ
250 260 270

```
>>gi|2677826 gi|2677826|gb|AAB97141.1| major allergen pr (160 aa)
  initn: 34 initl: 34 opt: 48 Z-score: 85.6 bits: 20.4 E(): 7.7
Smith-Waterman score: 48; 31.818% identity (59.091% similar) in 22 aa overlap
(26-47:4-25)
```

```

          10          20          30          40          50
CV127. MLEKEMKKYKKFLLFIKFTTLILVSFLFFKSFCHLLKKKKLFEIFLLPMSD
          : . : . . . . . :
gi|267      MGVFTYETEFTSVIPPEKLFKAFILDADNLIPKVAPTA
          10          20          30

```

gi|267 VKGTEILEGDDGGVTIKKVTFGEGSQYAYVKHRVDGIDKDNLSYSYTLIEGDALSDVIEN
40 50 60 70 80 90gi|267 IAYDIKLVASPDGGSIVKTTSHYHTKGDVEIKEEQVKAGKEKAAGLFLVEAYLLANPDA
100 110 120 130 140 150gi | 267 YN
160

```
>>gi|298737 gi|298737|gb|AAB25851.1| isoallergen {N-term (51 aa)
  initn: 33 init1: 33 opt: 42 Z-score: 84.6 bits: 18.6 E(): 8.7
Smith-Waterman score: 42; 45.833% identity (70.833% similar) in 24 aa overlap
(19-42:10-30)
```

```

          10          20          30          40          50
CV127. MLEKEMKKYKKFLLFIKFTTLILVSFLFFKSFCHLLKKKKLFEIFLLPMSD
          :.: :.: :.: :.: :.:
gi|298      SVFNJETETTSVIPAAFLF-KAF--ILDGDKLFPKVPAPQXQSIVXNXRVY
          10          20          30          40

```

gi | 298 XPK
50



The Chemical Company
Plant Science L. L. C.

BASF Reg. Doc. No. 2008/7019371

51 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:43:39 2008 done: Sun Dec 14 22:43:39 2008
Total Scan time: 0.080 Total Display time: 0.010

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.5723_c3

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.5723_c3, 62 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 4.9838 \pm 0.00389$; $\mu = -0.8788 \pm 0.201$
mean_var=44.2940 \pm 11.603, 0's: 17 Z-trim: 17 B-trim: 0 in 0/44
Lambda= 0.192709

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:		opt	bits	E(1313)	%_id	%_sim	alen
gi 3097321	gi 3097321 dbj BAA25899. (379)	58	23.5	2.6	0.359	0.615	39
gi 21783	gi 21783 emb CAA30570.1 u (356)	54	22.4	5.3	0.500	0.727	22
gi 22353013	gi 22353013 gb AAK97787 (195)	51	21.5	5.3	0.345	0.655	29
gi 1199563	gi 1199563 gb AAB09252.1 (379)	54	22.4	5.6	0.364	0.606	33
gi 129353	gi 129353 sp P22895.1 P34 (379)	54	22.4	5.6	0.364	0.606	33
gi 71360928	gi 71360928 emb CAJ1970 (114)	46	20.1	8.3	0.333	0.630	27
gi 62484809	gi 62484809 emb CAI7890 (285)	50	21.3	9.3	0.250	0.625	24

>>>CV127.5723_c3, 62 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|3097321 gi|3097321|dbj|BAA25899.1| Bd 30K [Glycine (379 aa)
initn: 28 initl: 28 opt: 58 Z-score: 94.0 bits: 23.5 E(): 2.6
Smith-Waterman score: 58; 35.897% identity (61.538% similar) in 39 aa overlap
(8-43:71-109)

gi|309 MGFLVLLLFSLGLSSSSSISTHRSILDLDLTKFTTQKQVSSLFQLWKSEHGRVYHNHEE
10 20 30 40 50 60

CV127. MRSLLIPNLSRV-NINMNNLSPTSKSLKLTRDTIIPK--YSKRQQQITMLHSVICIH
10 20 30 40 50
:: . . . : : : : . . . : : : : . . . :

gi|309 EAKRLEIFKNLNYIRDNMNANRKSPhSHRLGLNKFADITPQEFSSKKYLQAPKDVSSQQIKM
70 80 90 100 110 120

60
CV127. SPMHSHKM

gi|309 ANKKMKKEQYSCDHPPASWDWRKKGVITQVKYQGGCGSGWAFSATGAIEAAHAIATGDLV
130 140 150 160 170 180

gi|309 SLSEQELVDCVEESEGCGYNGWHYQSFVWLEHGGIATDDYPYRAKEGRCKANKIQDKVT
190 200 210 220 230 240

gi|309 IDGYETLIMSDESTESETEQAFLSAILEQPIVSIDAKDFHLYTGGIYDGENCTSPYGIN
250 260 270 280 290 300

gi|309 HFVLLVGYGSADGVYWIAKNSWGEDWGEGYIWIQRNTGNLLGVCGMNYFASYPTKEES
310 320 330 340 350 360

gi|309 ETLVSARVKGHRRVDHSPL
370

>>gi|21783 gi|21783|emb|CAA30570.1| unnamed protein prod (356 aa)
initn: 72 initl: 46 opt: 54 Z-score: 88.5 bits: 22.4 E(): 5.3
Smith-Waterman score: 54; 50.000% identity (72.727% similar) in 22 aa overlap
(36-56:176-196)

gi|217 MKTFLV FALL L A A A S A V A Q I S Q Q Q Q A P P F S Q Q Q Q P P F S Q Q Q Q S P F S Q Q Q Q Q
10 20 30 40 50 60

gi|217 P P F A Q Q Q Q P P F S Q Q P P I S Q Q Q Q P P F S Q Q Q Q P F S Q Q Q Q P P Y S Q Q Q Q P P Y S Q Q Q Q P P F S Q Q
70 80 90 100 110 120

CV127. MRSLILPNLSRVNINMNNLSPTSKSLKLTRDTIIPYSKR
10 20 30 40
: : :

gi|217 Q Q P P F S Q Q Q Q P P F T Q Q Q Q Q Q Q Q P F T Q Q Q P P F S Q Q P P I S Q Q Q Q P P F L Q Q Q R P P F S - R
130 140 150 160 170

CV127. Q Q I T M L H - S V I C I H S P M H S K H M
50 60
: : : : . : . : . :

gi|217 Q Q Q I P V I H P S V L Q Q L N P C K V F L Q Q C I P V A M Q R C L A R S Q M L Q Q S I C H V M Q Q Q C C Q Q L R Q I
180 190 200 210 220 230

gi|217 P E Q S R H E S I R A I I Y S I I L Q Q Q Q Q Q Q Q Q Q Q Q Q Q S I I Q Y Q Q Q Q P Q Q L G Q C V S Q P L Q Q L Q Q
240 250 260 270 280 290

gi|217 Q L G Q Q P Q Q Q L A H Q I A Q L E V M T S I A L R T L P T M C N V N V P L Y E T T T S V P L G V G I G V G V Y
300 310 320 330 340 350

>>gi|22353013 gi|22353013|gb|AAK97787.1| allergenic prot (195 aa)
initn: 28 initl: 28 opt: 51 Z-score: 88.5 bits: 21.5 E(): 5.3
Smith-Waterman score: 51; 34.483% identity (65.517% similar) in 29 aa overlap
(7-34:26-53)

CV127. MRSLILPNLSRVN-INMNNLSPTSKSLKLTRDTIIPYSKR
10 20 30 40
: : : : . : : : . : . : . :

gi|223 G L E Q A F C N L K F R Q N V N R P S H A D V F N P R A G R I N T V N S N N L - P I L E F L Q L S A Q H V V L Y K N A I
10 20 30 40 50

CV127. Q Q I T M L H S V I C I H S P M H S K H M
50 60

gi|223 I G P R W N L N A H S A L Y V T R G E G R V Q V V G D E G K S V F D D N V Q R G Q I L V V P Q G F A V V V K A G R Q G L
60 70 80 90 100 110

gi|223 E W V E L K N N D N A I T S P I A G R T S V L R A I P V E V L A N S Y D I S T E E A Y K L K N G R Q E V E V F R P F Q S
120 130 140 150 160 170

gi|223 R Y E K E E E K E R E R F S I V
180 190

>>gi|1199563 gi|1199563|gb|AAB09252.1| 34 kDa maturing s (379 aa)
initn: 28 initl: 28 opt: 54 Z-score: 88.0 bits: 22.4 E(): 5.6
Smith-Waterman score: 54; 36.364% identity (60.606% similar) in 33 aa overlap
(13-43:77-109)

gi|119 M G F L V L L L F S L L G L S S S S I S T H R S I L D L D L T K F T T Q K Q V S S L F Q L W K S E H G R V Y H N H E E
10 20 30 40 50 60

```

      10      20      30      40      50
CV127.  MRSLILPNLSRVNINMNNLSPTSKSLKLTRDTI IKP--YSKRQQQITMLHSVICIH
      .. :  :: :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|119  EAKRLEIFKNNSNYIRDMNANRKSPHSHRLGLNKFADITPQEFSSKKYLQAPKDVSSQIKM
      70      80      90      100     110     120

```

```

      60
CV127.  SPMHSHKM

gi|119  ANKKMKKEQYSCDHPPASWDWRKKGVITQVKYQGGCGRGWAFSATGAIEAAHAIATGDLV
      130     140     150     160     170     180

```

```

gi|119  SLSEQELVDCVEESEGSYNGWQYQSFVWLEHGGIATDDDYPRAKEGRCKANKIQDKVT
      190     200     210     220     230     240

```

```

gi|119  IDGYETLIMSDESTESETEQAFLSAILEQPISVSIDAKDFHLYTGGIYDGENCTSPYGIN
      250     260     270     280     290     300

```

```

gi|119  HFVLLVGYGSADGVYWIAKNSWGEDWGEDGYIWIQRNTGNLLGVCGMNYFASYPTKEES
      310     320     330     340     350     360

```

```

gi|119  ETLVSARVKGHRRVDHSPL
      370

```

```

>>gi|129353 gi|129353|sp|P22895.1|P34_SOYBN P34 probable (379 aa)
  initn: 28 initl: 28 opt: 54 Z-score: 88.0 bits: 22.4 E(): 5.6
Smith-Waterman score: 54; 36.364% identity (60.606% similar) in 33 aa overlap
(13-43:77-109)

```

```

gi|129  MGFLVLLLFSLGLSSSSSISTHRSILDLDLTKFTTQKQVSSLFQLWKSEHGRVYHNHEE
      10      20      30      40      50      60

```

```

      10      20      30      40      50
CV127.  MRSLILPNLSRVNINMNNLSPTSKSLKLTRDTI IKP--YSKRQQQITMLHSVICIH
      .. :  :: :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|129  EAKRLEIFKNNSNYIRDMNANRKSPHSHRLGLNKFADITPQEFSSKKYLQAPKDVSSQIKM
      70      80      90      100     110     120

```

```

      60
CV127.  SPMHSHKM

gi|129  ANKKMKKEQYSCDHPPASWDWRKKGVITQVKYQGGCGRGWAFSATGAIEAAHAIATGDLV
      130     140     150     160     170     180

```

```

gi|129  SLSEQELVDCVEESEGSYNGWQYQSFVWLEHGGIATDDDYPRAKEGRCKANKIQDKVT
      190     200     210     220     230     240

```

```

gi|129  IDGYETLIMSDESTESETEQAFLSAILEQPISVSIDAKDFHLYTGGIYDGENCTSPYGIN
      250     260     270     280     290     300

```

```

gi|129  HFVLLVGYGSADGVYWIAKNSWGFWDGEDGYIWIQRNTGNLLGVCGMNYFASYPTKEES
      310     320     330     340     350     360

```

```

gi|129  ETLVSARVKGHRRVDHSPL
      370

```

```

>>gi|71360928 gi|71360928|emb|CAJ19705.1| non-specific 1 (114 aa)
  initn: 37 initl: 37 opt: 46 Z-score: 85.0 bits: 20.1 E(): 8.3
Smith-Waterman score: 46; 33.333% identity (62.963% similar) in 27 aa overlap
(2-28:88-113)

```

gi|713 MEMVNKIACFVLLCMVVVAPHAEALTCGQVTSTLAPCLPYLMNRGPLRNCCDGVKGLLGQ
10 20 30 40 50 60

CV127. MRSLLLPNLSRVNINMNNLSPTSKSLKLTRDTII
.. ::: ::: .. :::: ..
gi|713 AKTTVDRQAACTCLKSAASSFTGLNLGKAAALPNTCSVNIPYK-ISPSTDCSKVQ
70 80 90 100 110

40 50 60
CV127. KPYSKRQQQITMLHSVICIHSPMHSKM

>>gi|62484809 gi|62484809|emb|CAI78902.1| putative gamma (285 aa)
initn: 37 initl: 37 opt: 50 Z-score: 84.1 bits: 21.3 E(): 9.3
Smith-Waterman score: 50; 25.000% identity (62.500% similar) in 24 aa overlap
(35-58:37-60)

10 20 30 40 50
CV127. MRSLLLPNLSRVNINMNNLSPTSKSLKLTRDTIIKPYSKRQQQITMLHSVICIHSPMH
.:. .: .: .: .: .: .: .:
gi|624 MQVNPSVQVQPTQQQYPESQQPFISQSQQQFPQPQQPFPQQPQQPFPQSQQQCLQQPQH
10 20 30 40 50 60

60
CV127. SKHM

gi|624 QFPQPTQQFPQRPLLPFTHPFLTFPDQLLPQPPHQSFPPQSYQPPLQFPQPFPQQKY
70 80 90 100 110 120

gi|624 PEQPQQPFPWQPTIQLYLQQQLNPCKEFLQCCRVPVSLLSYLWSKIVQQSSCRMVQQQC
130 140 150 160 170 180

gi|624 CLQLAQIPEQYKCTAIDSIVHAIFMQGQRQGVQIVQQPQPQQVGQCVLVQGGQGVVQPQ
190 200 210 220 230 240

gi|624 QLAQMEAIRTLVLQSVPSMCNFNVPNCSTIKAPFVGVTGVGGQ
250 260 270 280

62 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:47:34 2008 done: Sun Dec 14 22:47:34 2008
Total Scan time: 0.080 Total Display time: 0.010

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.1392 d3

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.1392 d3, 44 aa

```
vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library
```

292011 residues in 1313 sequences

Expectation n fit: $\rho(\ln(x)) = 3.2196 \pm 0.00331$; $\mu = 8.5459 \pm 0.174$

```
mean_var=39.2699+/-11.256, 0's: 13 Z-trim: 13 B-trim: 76 in 1/43
```

Lambda= 0.204665

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2

```
join: 36, opt: 20, open/ext: -10/-2, width: 16
```

The best scores are:

The best scores are:					opt	bits	E(1313)	%_id	%_sim	alen
gi 31339067	gi 31339067	dbj BAC7715	(194)	53	22.0	2.7	0.438	0.812	16	
gi 47605452	gi 47605452	sp Q7Z1K3.1	(194)	53	22.0	2.7	0.438	0.812	16	
gi 14423846	gi 14423846	sp O82040.1	(78)	49	20.4	3.3	0.324	0.595	37	
gi 123062	gi 123062	sp P02877.2	HEV (204)	52	21.7	3.4	0.400	0.680	25	
gi 2832430	gi 2832430	emb CAA05978.	(187)	51	21.4	3.9	0.400	0.680	25	
gi 1247373	gi 1247373	emb CAA01909.	(71)	47	19.7	4.7	0.357	0.607	28	
gi 1247375	gi 1247375	emb CAA01910.	(73)	47	19.8	4.8	0.357	0.607	28	
gi 62240390	gi 62240390	gb AAX77383	(510)	53	22.5	5	0.371	0.686	35	
gi 1871507	gi 1871507	emb CAA62634.	(82)	47	19.8	5.1	0.357	0.607	28	
gi 14423643	gi 14423643	sp O24172.1	(50)	45	19.0	5.6	0.538	0.769	13	
gi 664852	gi 664852	gb AAB07620.1	(250)	49	21.0	7.1	0.394	0.636	33	
gi 217308	gi 217308	dbj BAA01241.1	(138)	47	20.1	7.2	0.233	0.605	43	
gi 546852	gi 546852	gb AAB30829.1	(142)	47	20.1	7.4	0.233	0.605	43	
gi 55859470	gi 55859470	emb CAI0585	(146)	47	20.1	7.5	0.233	0.605	43	
gi 56378069	gi 56378069	dbj BAD7406	(146)	47	20.1	7.5	0.233	0.605	43	
gi 86450747	gi 86450747	gb ABC96702	(146)	47	20.1	7.5	0.233	0.605	43	
gi 55859468	gi 55859468	emb CAI0584	(146)	47	20.1	7.5	0.233	0.605	43	
gi 217306	gi 217306	dbj BAA01240.1	(146)	47	20.1	7.5	0.233	0.605	43	
gi 66849502	gi 66849502	gb EAL89830	(304)	49	21.1	8.1	0.394	0.636	33	
gi 83300352	gi 83300352	sp P79017.2	(310)	49	21.1	8.2	0.394	0.636	33	
gi 17978844	gi 17978844	gb AAL47677	(129)	46	19.7	8.5	0.250	0.611	36	

```
>>>CV127.1392 d3, 44 aa vs
```

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

```
>>gi|31339067 gi|31339067|dbj|BAC77154.1| 21k allergen [ (194 aa)
```

```
initn: 35 init1: 35 opt: 53 Z-score: 93.9 bits: 22.0 E(): 2.7
```

Smith-Waterman score: 53; 43.750% identity (81.250% similar) in 16 aa overlap (23-38:82-97)

CV127.

M

gi|313 MASMQHFSLAALLLAASICLGADRTCEQLPLDKGTPCTQEGGVKPSVAWWHDDKSGICL
10 20 30 40 50 60

CV127. EAIDSAIDPLRDFAKSSVRLVORCHKPDRKGNAFSOKNLISDF

gi|313 SFKYTGCGGNANRFTTIKNCEQHCKMPDRGACALGKKPAEDSNGEQLVCAGMREDKCPNG
70 80 90 100 110 120

gi|313 YQCKMMAFMGLCCPTKEEELFAREYEGVCKSGKPVKMDRSGSGWMMTILGKSCDDQFCPED
130 140 150 160 170 180

gi|313 AKCERGLFANCKK
190

>>gi|47605452 gi|47605452|sp|Q7Z1K3.1|ANIS1_ANISI Major (194 aa)
initn: 35 initl: 35 opt: 53 Z-score: 93.9 bits: 22.0 E(): 2.7
Smith-Waterman score: 53; 43.750% identity (81.250% similar) in 16 aa overlap
(23-38:82-97)

CV127. M

gi|476 MASMQHFSLAALLLAASICLGADARTECQLPLDKGTPCTQEGGVKPSVAWWHDDKSGICL
10 20 30 40 50 60

10 20 30 40
CV127. EAIDSAIDPLRDFAKSSVRLVQRCHKPDRKGNFASQKNLISDF
:.

gi|476 SFKYTGCGGNANRFTTIKNCEQHCKMPDRGACALGKKPAEDSNGEQLVCAGMREDKCPNG
70 80 90 100 110 120

gi|476 YQCKMMAFMGLCCPTKEEELFAREYEGVCKSGKPVKMDRSGSGWMMTILGKSCDDQFCPED
130 140 150 160 170 180

gi|476 AKCERGLFANCKK
190

>>gi|14423846 gi|14423846|sp|O82040.1|POLC7_PHLPR Polcal (78 aa)
initn: 40 initl: 40 opt: 49 Z-score: 92.2 bits: 20.4 E(): 3.3
Smith-Waterman score: 49; 32.432% identity (59.459% similar) in 37 aa overlap
(9-44:27-63)

10 20 30 40
CV127. MEAIDSAIDPLRDFAKSSVRLVQRCHKP-DRKGNFASQKNLI
: : : : : :

gi|144 MADDMERIFKRFDTNMGDKISLSELTDALRTLSTSADEVQRMMAEIDTDGDFIDFNEF
10 20 30 40 50 60

CV127. SDF
. :

gi|144 ISFCNANPGLMKDVAKVF
70

>>gi|123062 gi|123062|sp|P02877.2|HEVE_HEVBR Pro-hevein (204 aa)
initn: 48 initl: 48 opt: 52 Z-score: 92.0 bits: 21.7 E(): 3.4
Smith-Waterman score: 52; 40.000% identity (68.000% similar) in 25 aa overlap
(15-37:141-165)

gi|123 MNIFIVVLLCLTGVAIAEQCGRQAGGKLCPPNNLCCSQWGWCGSTDEYCSPDHNCQSNCKD
10 20 30 40 50 60

gi|123 SGEGVGGGSASNVLATYHLYNSQDHGWDLNAASAYCSTWDANKPYSWRSKYGWTAFCGPV
70 80 90 100 110 120

10 20 30 40
CV127. MEAIDSAIDPLRDFAKSSVRLVQRCHKP--DRKGNFASQKNLISDF
: : : : :

gi|123 GAHGQSSCGKCLSVTNTGTGAKTTVRIVDQCSNGGLDLVDNVFRQLDTDGKGYERGHITV
130 140 150 160 170 180

gi | 123 NYQFVDCGDSFNPLFSVMKSSVIN
190 200

```
>>gi|2832430 gi|2832430|emb|CAA05978.1| prohevein [Hevea (187 aa)
  initn: 47 initl: 47 opt: 51 Z-score: 90.9 bits: 21.4 E(): 3.9
Smith-Waterman score: 51; 40.000% identity (68.000% similar) in 25 aa overlap
(15-37:124-148)
```

gi | 283 EQCGRQAGGKLCNNLCCSQWGWCGSTDEYCSPDHNCQSNCKDSGEGVGGGSASNVLATY

CV127. MEAIDSAIDPL 10

gi|283 HLYNSQDHGWDLNAASAYCSTWDANKPYSWRSKYGTAF CGPVG AHQPSCGKCLSVTNT
70 80 90 100 110 120

CV127. R D F A K S S V R L V Q R C H K P -- D R K G N A F S Q K N L I S D F

gi|283 GTGAKATVRIVDQCSNGGLDLVDNVFRQLDTDGKGYERGHLLTVNYQFVDCGDSFNPLFSV
130 140 150 160 170 180

qi | 283 MKSSVIN

```
>>gi|1247373 gi|1247373|emb|CAA01909.1| B1 protein aller (71 aa)
initn: 40 initl: 40 opt: 47 Z-score: 89.5 bits: 19.7 E(): 4.7
Smith-Waterman score: 47; 35.714% identity (60.714% similar) in 28 aa overlap
(9-35:20-47)
```

CV127. MEAIDSAIDPLRDFAKSSVRLVORCHKP-DRKGNAFSOKNLI SDF

gi|124 GTRRFDTNGDGKISLAELTDALRTLGSTADEVQRMMAEIDTDGDFIDFDEFISFCNAN
10 20 30 40 50 60gi | 124 PGLMKDVAKV F
70

```
>>gi|1247375 gi|1247375|emb|CAA01910.1| B4 protein aller (73 aa)
  initn: 40 initl: 40 opt: 47 Z-score: 89.3 bits: 19.8 E(): 4.8
Smith-Waterman score: 47; 35.714% identity (60.714% similar) in 28 aa overlap
(9-35;22-49)
```

CV127.1020 MEAIDSAIDPLRDEAKSSVRLVORCHKP-DRKGNAFSOKNLISDF

gi | 124 GTSFKRFD TNGDGKISLAELTDALRTL GSTSADEVQRMMAEIDTDGDGFIDFDEFISFCN
10 20 30 40 50 60gi | 124 ANPGLMKDVAKVF
70

```
>>gi|62240390 gi|62240390|gb|AAx77383.1| 11S globulin pr (510 aa)
  initn: 31 init1: 31 opt: 53 Z-score: 88.9 bits: 22.5 E(): 5
Smith-Waterman score: 53; 37.143% identity (68.571% similar) in 35 aa overlap
(2-33:249-283)
```

gi|622 MVKLAHLLVATVGVLVLNGCLARQSLGVPPQVKDACNLNDVLQPTTEVIKSEAGQVEY
10 20 30 40 50 60

gi|622 WDHNHPQIRCAGVSIARLVIQKGGLYLPTFFSSPFISYVVQGMGISGRVIPGCAETFMDS
70 80 90 100 110 120[illegible]gi|622 VEHVRHGDAlAMTPGSAQWIYNTGDQPLIIIVSLIDIANYQNQLDRNPRTFRLAGNNQQGS
190 200 210 220 230 240

CV127. MEAIDSAIDP--LRDFAKSSVRLVQRCR-KPDRKGNAFSQKNLISDF
gi|622 SQQQQQQQQNILSGFDPOVLAAQLKIDIVRLAQELNQDDKRGNIVRVKGPFFQVVRPPLRQ

10 20 30 40
250 260 270 280 290 300

gi|622 AYESEQWRHPRGPPQSPQDNGLEETICSMRTHENIDDPARADIYKPNLGRVTSVNSYTLF

gi|622 ILQYIRLSATRGILQGSAMVLPKYNMNAEILYCTQGQARIQVVNDNGQNVLDQQVQKGQ

gi|622 LVVIPQGFAYVVQSQNNFEWISFKTNANAMISTLAGRTSALRALPLEVITNAYQISLEEA
430 440 450 460 470 480gi|622 RKIKFNTLETTLTRARGGQPQLIEEIVEV
490 500 510

>>gi|1871507 gi|1871507|emb|CAA62634.1| calcium-binding (82 aa)
 initn: 40 initl: 40 opt: 47 Z-score: 88.7 bits: 19.8 E(): 5.1
 Smith-Waterman score: 47; 35.714% identity (60.714% similar) in 28 aa overlap
 (9-35;31-58)

CV127. MEAIDSAIDPLRDFAKSSVRLVQRCHKP-DRKGNAFSQ
:
gi|187 KTMADTGDMEHIFKRFDTNGDGKISLAELTDALRTLGSTADEVQRMMAEITDGDGDFID

40
CV127. KNLSDF

gi | 187 FDEFISFCNANPGLMKDVAKVF
70 80

```
>>gi|14423643 gi|14423643|sp|O24172.1|ALL6_OLEEU Pollen (50 aa)
  initn: 31 initl: 31 opt: 45 Z-score: 88.1 bits: 19.0 E(): 5.6
Smith-Waterman score: 45; 53.846% identity (76.923% similar) in 13 aa overlap
(25-36:12-24)
```

```

      10      20      30      40
CV127. MEAIDSAIDPLRDFAKSSVRLVQRCHKP-DRKGNAFSQKNLISDF
      :: . ::::
gi|144      DEAQFKECYDTCHEKCSDKNGGFTFCCEMKCDTDCSVKDVKEKLENYK
      10      20      30      40

```

gi | 144 PKN
50

```
>>gi|664852 gi|664852|gb|AAB07620.1| Asp FII [Aspergillu (250 aa)
initn: 37 initl: 37 opt: 49 Z-score: 86.2 bits: 21.0 E(): 7.1
```

Smith-Waterman score: 49; 39.394% identity (63.636% similar) in 33 aa overlap
(1-33:42-69)

```

                                10
CV127.                          MEAIDSAIDPLRDFAKSSV
                                .... .: : . . . :...
gi|664 SARDEAGLNEAVELARHAKAHILRWGNESEIYRKYFGNRPTMEAV-GAYDVIVNGDKANV
                                10      20      30      40      50

                                20      30      40
CV127. RLVQRCHKPDRKGNAFSQKNLISDF
                                . : : .: : : :
gi|664 LF--RCDNPD--GNCALEGWGGHWRGANATSETVICDRSYTTRRWLVSMCSQGYTVAGSE
                                60      70      80      90      100     110

gi|664 TNTFWASDLMHRLYHVPVAVGQGWVDHFDGDEVIYALAKSNGTESTHDSEAFEFYFALEAY
                                120     130     140     150     160     170

gi|664 AFDIAAPGVGCAGESHGPDPQGHDTGSASAPASTSTSSSSSGSGSGATTTPTDSPSATIDV
                                180     190     200     210     220     230

gi|664 PSNCHTHEGGQLHCT
                                240     250

```

>>gi|217308 gi|217308|dbj|BAA01241.1| mite allergen Der (138 aa)
initn: 28 initl: 28 opt: 47 Z-score: 86.0 bits: 20.1 E(): 7.2
Smith-Waterman score: 47; 23.256% identity (60.465% similar) in 43 aa overlap
(4-40:8-49)

```

                                10      20      30      40
CV127.      MEAIDSAIDPLRDFAKSSVR--LVQRCHKPD---RKGNAFSQKNLISDF
                                . . .: .: : .: . . .: : : : .: .: : .: :
gi|217 SLLVAADVADQVD-VKDCANNEIKKVMVDGCHGSDPCIHHRGKPFTEALFDANQNTKTA
                                10      20      30      40      50

gi|217 KIEIKASLDGLEIDVPGIDTNACHFMKCPLVKGGQYDAKYTNVVPKIAPKSENVVTVKL
                                60      70      80      90      100     110

gi|217 VGDNGVLACAIATHAKIRD
                                120     130

```

>>gi|546852 gi|546852|gb|AAB30829.1| Der f II [Dermatoph (142 aa)
initn: 28 initl: 28 opt: 47 Z-score: 85.9 bits: 20.1 E(): 7.4
Smith-Waterman score: 47; 23.256% identity (60.465% similar) in 43 aa overlap
(4-40:12-53)

```

                                10      20      30      40
CV127.      MEAIDSAIDPLRDFAKSSVR--LVQRCHKPD---RKGNAFSQKNLISDF
                                . . .: .: : .: . . .: : : : .: .: : .: :
gi|546 GTMVSLVAADVADQVD-VKDCANNEIKKVMVDGCHGSDPCIHHRGKPFTEALFDANQN
                                10      20      30      40      50

gi|546 TKTAKIEIKASLDGLEIDVPGIDTNACHFVKCPLVKGGQYDIKYTNVVPKIAPKSENVV
                                60      70      80      90      100     110

gi|546 TVKLIGDNGVLACAIATHGKIRD
                                120     130     140

```

>>gi|55859470 gi|55859470|emb|CAI05850.1| mite allergen (146 aa)
initn: 28 initl: 28 opt: 47 Z-score: 85.8 bits: 20.1 E(): 7.5
Smith-Waterman score: 47; 23.256% identity (60.465% similar) in 43 aa overlap
(4-40:16-57)

```

              10      20      30      40
CV127.      MEAIDSAIDPLRDFAKSSVR--LVQRCHKPD---RKGNAFSQKNLIS
              . . . : . . : . . . . : . . : . . . : . . . :
gi|558 MISKILCLSLLVAAVVADQVD-VKDCANNEIKKVMVDGCHGSDPCI IHRGKPF TLEALFD
              10      20      30      40      50

```

CV127. DF

```

gi|558 ANQNTKTAKIEIKANIDGLEVDVPGIDTNACHYIKCPLVKGGQYDAKYTNVNPKIAPKSE
      60      70      80      90     100     110

```

```

gi|558 NVVVTVKLVGDNGVLACAIATHAKIRD
      120     130     140

```

>>gi|56378069 gi|56378069|dbj|BAD74060.1| group 2 allerg (146 aa)
 initn: 28 initl: 28 opt: 47 Z-score: 85.8 bits: 20.1 E(): 7.5
 Smith-Waterman score: 47; 23.256% identity (60.465% similar) in 43 aa overlap
 (4-40:16-57)

```

              10      20      30      40
CV127.      MEAIDSAIDPLRDFAKSSVR--LVQRCHKPD---RKGNAFSQKNLIS
              . . . : . . : . . . . : . . : . . . : . . . :
gi|563 MISKILCLSLLVAAVVADQVD-VKDCANNEIKKVMVDGCHGSDPCI IHRGKPF TLEALFD
              10      20      30      40      50

```

CV127. DF

```

gi|563 ANQNTKTAKIEIKASLDGLEIDVPGIDTNACHFMKCPLVKGGQYDIKYTDVNPKIAPKSE
      60      70      80      90     100     110

```

```

gi|563 NVVVTVKLVGDNGVLACAIATHGKIRD
      120     130     140

```

>>gi|86450747 gi|86450747|gb|ABC96702.1| Der s 2 a aller (146 aa)
 initn: 28 initl: 28 opt: 47 Z-score: 85.8 bits: 20.1 E(): 7.5
 Smith-Waterman score: 47; 23.256% identity (60.465% similar) in 43 aa overlap
 (4-40:16-57)

```

              10      20      30      40
CV127.      MEAIDSAIDPLRDFAKSSVR--LVQRCHKPD---RKGNAFSQKNLIS
              . . . : . . : . . . . : . . : . . . : . . . :
gi|864 MISKILCLSLLVAAVVADQVD-VKDCANNEIKKVMVDGCHGSDPCI IHRGKPF TLEALFD
              10      20      30      40      50

```

CV127. DF

```

gi|864 ANQNTKTAKIEIKANIDGLEVDVPGIDTNACHFIKCPVLKGGQYDAKYTNVNPKIAPKSE
      60      70      80      90     100     110

```

```

gi|864 NVVVTVKLIGDNGVLACAIATHAKIRD
      120     130     140

```

>>gi|55859468 gi|55859468|emb|CAI05849.1| Der f 2 [Derma (146 aa)
 initn: 28 initl: 28 opt: 47 Z-score: 85.8 bits: 20.1 E(): 7.5
 Smith-Waterman score: 47; 23.256% identity (60.465% similar) in 43 aa overlap
 (4-40:16-57)

```

              10      20      30      40
CV127.      MEAIDSAIDPLRDFAKSSVR--LVQRCHKPD---RKGNAFSQKNLIS
              . . .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|558 MISKILCLSLVAADVADQVD-VKDCANNEIKKVMVDGCHGSDPCI IHRGKPFLEALFD
              10      20      30      40      50

```

CV127. DF

```

gi|558 ANQNTKTAKIEIKANINGLEVDVPGIDTNACHYIKCPLVKGGQYDAKYTNVVPKIAPKSE
        60      70      80      90     100     110

```

```

gi|558 NVVVTVKLIGDNGVLACAIATHAKIRD
        120     130     140

```

>>gi|217306 gi|217306|dbj|BAA01240.1| mite allergen Der (146 aa)
initn: 28 init1: 28 opt: 47 Z-score: 85.8 bits: 20.1 E(): 7.5
Smith-Waterman score: 47; 23.256% identity (60.465% similar) in 43 aa overlap
(4-40:16-57)

```

              10      20      30      40
CV127.      MEAIDSAIDPLRDFAKSSVR--LVQRCHKPD---RKGNAFSQKNLIS
              . . .: .: .: .: .: .: .: .: .: .: .: .: .: .:
gi|217 MISKILCLSLVAADVADQVD-VKDCANNEIKKVMVDGCHGSDPCI IHRGKPFLEALFD
              10      20      30      40      50

```

CV127. DF

```

gi|217 ANQNTKTAKIEIKASLDGLEIDVPGIDTNACHFMKCPLVKGGQYDIKYTNVVPKIAPKSE
        60      70      80      90     100     110

```

```

gi|217 NVVVTVKLIGDNGVLACAIATHGKIRD
        120     130     140

```

>>gi|66849502 gi|66849502|gb|EAL89830.1| major allergen (304 aa)
initn: 37 init1: 37 opt: 49 Z-score: 85.2 bits: 21.1 E(): 8.1
Smith-Waterman score: 49; 39.394% identity (63.636% similar) in 33 aa overlap
(1-33:102-129)

```

gi|668 MAALLRLAVLLPLAAPLVATLPTSPVPPIAARATPHEPVFFSWDAGAVTSFPIHSSCNATQ
              10      20      30      40      50      60

```

```

              10
CV127.      MEAIDSAIDPLRDFAKSSV
              : : : .: .: .: .: .:
gi|668 RRQIEAGLNEAVELARHAKAHILRWGNESEIYRKYFGNRPTMEAV-GAYDVIVNGDKANV
              70      80      90     100     110

```

```

              20      30      40
CV127. RLVQRCHKPDRKGNFASQKNLISDF
              . : : .: : :
gi|668 LF--RCDNPD--GNCALEGWGGHWRGANATSETVICDRSYTTRRWLVSMCSQGYTVAGSE
        120     130     140     150     160     170

```

```

gi|668 TNTFWASDLMHRLYHVPVAVGGQWVDHFDADGYDEVIALAKSNGTESTHDSEALQYFALEAY
        180     190     200     210     220     230

```

```

gi|668 AFDIAAPGVGCAGESHGPDPQGHDTGSASAPASTSTSSSSSGSGGATTTPTDSPSATIDV
        240     250     260     270     280     290

```

```

gi|668 PPVRTVRIS
        300

```

>>gi|83300352 gi|83300352|sp|P79017.2|ALL2_ASPFU Major a (310 aa)
initn: 37 initl: 37 opt: 49 Z-score: 85.1 bits: 21.1 E(): 8.2
Smith-Waterman score: 49; 39.394% identity (63.636% similar) in 33 aa overlap
(1-33:102-129)

gi|833 MAALLRLAVLLPLAAPLVATLPTSPVPPIAARATPHEPVFFSWDAGAVTSFPIHSSCNATQ
10 20 30 40 50 60

CV127. MEAIDSAIDPLRDFAKSSV
10
::: .: : . . :
gi|833 RRQIEAGLNEAVELARHAKAHILRWGNESEIYRKYFGNRPTMEAV-GAYDVIVNGDKANV
70 80 90 100 110

20 30 40
CV127. RLVQRCHKPDRKGNAFSQKNLISDF
. : : : :
gi|833 LF--RCDNPD--GNCALEGWGGHWRGANATSETVICDRSYTTRRWLVSMCSQGYTVAGSE
120 130 140 150 160 170

gi|833 TNTFWASDLMHRLYHVPVAVGGQVWDHFDGYDEVIALAKSNGTESTHDSEALQYFALEAY
180 190 200 210 220 230

gi|833 AFDIAAPGVGCAGESHGPDPQGHDTGSASAPASTSTSSSSSGSGSGATTTPTDSPSATIDV
240 250 260 270 280 290

gi|833 PSNCHTHEGGQLHCT
300 310

>>gi|17978844 gi|17978844|gb|AAL47677.1| major Der f 2 i (129 aa)
initn: 28 initl: 28 opt: 46 Z-score: 84.8 bits: 19.7 E(): 8.5
Smith-Waterman score: 46; 25.000% identity (61.111% similar) in 36 aa overlap
(11-40:5-40)

10 20 30 40
CV127. MEAIDSAIDPLRDFAKSSVR--LVQRCHKPD---RKGNAFSQKNLISDF
..: ..: ..: : : : : : : : : :
gi|179 DQVDVKDCANNEIKKVMVDGCHGSDPCIHRGKPFTEALFDANQNTKTAKIEI
10 20 30 40 50

gi|179 KASLDGLEIDVPGIDTNACHFMKCPLVKGQQYDAKYTWNVPKIAPESENVVVTVKLVGDN
60 70 80 90 100 110

gi|179 GVLACAIATHAKIRD
120

44 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 21:55:32 2008 done: Sun Dec 14 21:55:32 2008
Total Scan time: 0.080 Total Display time: 0.020

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.2973_d3

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.2973_d3, 36 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 5.2458 \pm 0.00415$; $\mu = -3.9189 \pm 0.214$

mean_var=49.2478 \pm 13.013, 0's: 20 Z-trim: 21 B-trim: 70 in 1/41

Lambda= 0.182760

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2

join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:

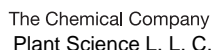
					opt	bits	E(1313)	%_id	%_sim	alen	
gi 170736	gi 170736	gb AAA34288.1	(251)		70	25.8		0.2	0.450	0.750	20
gi 170716	gi 170716	gb AAA34278.1	(319)		68	25.3		0.36	0.385	0.577	26
gi 170718	gi 170718	gb AAA34279.1	(313)		67	25.0		0.43	0.385	0.577	26
gi 21765	gi 21765	emb CAA26385.1	u (313)		67	25.0		0.43	0.385	0.577	26
gi 170738	gi 170738	gb AAA34289.1	(327)		67	25.0		0.45	0.471	0.765	17
gi 170708	gi 170708	gb AAA34274.1	(291)		65	24.5		0.57	0.471	0.706	17
gi 170724	gi 170724	gb AAA34282.1	(297)		63	24.0		0.84	0.529	0.765	17
gi 21673	gi 21673	emb CAA35238.1	u (307)		63	24.0		0.87	0.529	0.765	17
gi 1304264	gi 1304264	dbj BAA12318.	(259)		62	23.7		0.89	0.360	0.600	25
gi 170722	gi 170722	gb AAA34281.1	(262)		62	23.7		0.9	0.360	0.600	25
gi 62484809	gi 62484809	emb CAI7890	(285)		62	23.7		0.98	0.308	0.513	39
gi 170720	gi 170720	gb AAA34280.1	(286)		62	23.7		0.98	0.360	0.600	25
gi 21755	gi 21755	emb CAA25593.1	u (286)		62	23.7		0.98	0.360	0.600	25
gi 473876	gi 473876	gb AAA17741.1	(287)		62	23.7		0.98	0.360	0.600	25
gi 170726	gi 170726	gb AAA34283.1	(282)		61	23.5		1.2	0.375	0.542	24
gi 170740	gi 170740	gb AAA34290.1	(296)		61	23.5		1.2	0.412	0.765	17
gi 21757	gi 21757	emb CAA26383.1	u (296)		61	23.5		1.2	0.375	0.542	24
gi 75219081	gi 75219081	sp O22108 O	(285)		60	23.2		1.4	0.324	0.676	34
gi 21773	gi 21773	emb CAA31685.1	u (307)		60	23.2		1.5	0.324	0.676	34
gi 170712	gi 170712	gb AAA34276.1	(291)		58	22.7		2.1	0.385	0.615	26
gi 170702	gi 170702	gb AAA34272.1	(302)		58	22.7		2.1	0.438	0.750	16
gi 21930	gi 21930	emb CAA44473.1	L (285)		57	22.4		2.4	0.324	0.647	34
gi 76782247	gi 76782247	gb ABA54897	(134)		51	20.8		3.5	0.478	0.565	23
gi 21926	gi 21926	emb CAA36063.1	u (295)		55	21.9		3.6	0.367	0.600	30
gi 1063270	gi 1063270	dbj BAA11251.	(279)		53	21.3		4.9	0.370	0.630	27
gi 170734	gi 170734	gb AAA34287.1	(244)		52	21.1		5.2	0.333	0.545	33
gi 3915783	gi 3915783	sp P43217.3 N	(139)		49	20.2		5.3	0.348	0.609	23
gi 162811	gi 162811	gb AAA30433.1	(190)		50	20.5		5.9	0.333	0.667	30
gi 1228078	gi 1228078	emb CAA33034.	(190)		50	20.5		5.9	0.333	0.667	30
gi 21761	gi 21761	emb CAA26384.1	u (286)		52	21.1		6.1	0.385	0.615	26
gi 1532058	gi 1532058	emb CAA65123.	(176)		49	20.3		6.6	0.348	0.609	23
gi 886963	gi 886963	emb CAA59338.1	(229)		50	20.5		7.1	0.423	0.577	26
gi 170730	gi 170730	gb AAA34285.1	(304)		51	20.8		7.7	0.429	0.667	21
gi 886965	gi 886965	emb CAA59339.1	(261)		50	20.6		8	0.400	0.640	25
gi 62550933	gi 62550933	emb CAI7905	(326)		51	20.8		8.3	0.367	0.633	30
gi 886967	gi 886967	emb CAA59340.1	(276)		50	20.6		8.5	0.400	0.640	25

>>>CV127.2973_d3, 36 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|170736 gi|170736|gb|AAA34288.1| gamma-gliadin (251 aa)

initn: 170 init1: 70 opt: 70 Z-score: 114.0 bits: 25.8 E(): 0.2



BASF Reg. Doc. No. 2008/7019371

```
gi|170 MKTLLILTLILAMAITIGTANMQVDPSSQVQWPQQQPVPQHPQPFSSQQPQQTFPPQQTFFP
          10           20           30           40           50           60

CV127.                                     MSQPLPLQPNLPNPKHSSPDSLQINPAKALISS
          :: : :: : : . : . . . . :

gi|170 HQPQQQFPQPQQPQQQLQPQQPFPQQPQQPYQPQQPFPQTQQPQQLFPPQSQQPQQQF
          70           80           90           100          110          120

CV127. SKL

gi|170 SQPQQQFPQPQQPQQSFPQQQPPFIQPSLQQQVNPCKNFLQCKPVSVLSLWSMIWPO
          130          140          150          160          170          180

gi|170 SDCQVMRQQCCQQLAQIPQQQLQCAAIHTIIHSIIMQQEQQEQQQGMMHILLPLYQQQQVGQ
          190          200          210          220          230          240

gi|170 GTLVQGQGIIQ
          250
```

```
>>gi|170716 gi|170716|gb|AAA34278.1| pre-alpha-/beta-gli (319 aa)
initn: 95 initl: 63 opt: 68 Z-score: 109.4 bits: 25.3 E(): 0.36
Smith-Waterman score: 68; 38.462% identity (57.692% similar) in 26 aa overlap
(3-28:83-108)
```

```
gi|170 MKTFLILALLAIVATTATTAVRVVPVQLQPQNPSQQQPQEQVPLVQQQQFPGQQQQFPPQ
      10      20      30      40      50      60

CV127.      MSQPLPLQPNLPNPKHSSPDSLQINPAKALISSKL
      :: :: :::: :: :: :: ::
gi|170 QPYPQPQPFPSQQPYLQLQFPQPQPFPPQLPYPQPQSFPPQQPYPQQPQYLQPQQPIS
      70      80      90      100      110      120

gi|170 QQQAQQQQQQQQQQQQQQILQQILQQQLIPCRDVVLQQHNIHAASSQVLQQSTYQLLQQ
      130      140      150      160      170      180

gi|170 LCCQQLLQIPEQSQCQAIHNVAHAIIMHQQQQQQQEQKQQLQQQQQQQQQLQQQQQQQQQ
      190      200      210      220      230      240

gi|170 QPSSQVSFQQPQQQYPSSQVSFQPSQLNPQAQGSVQPQQLPQFAEIRNLALQTLPAMCNV
      250      260      270      280      290      300

gi|170 YIPPHCSTTIAPFGISGTN
      310
```

```
>>gi|170718 gi|170718|gb|AAA34279.1| alpha/beta-gliadin (313 aa)
  initn: 62 initl: 62 opt: 67 Z-score: 108.1 bits: 25.0 E(): 0.43
Smith-Waterman score: 67; 38.462% identity (57.692% similar) in 26 aa overlap
(3-28:80-105)
```

```
gi|170 MKTFLILALVATTATTAVRVPVQLQPKNPSQQQPQEVLVQQQQFPGQQQQFPQQPY
          10           20           30           40           50           60
CV127.      MSQPLPLQPNLPPNPKHSSPDSLQINPAKALISSSKL
              :: :. :.: : . :. : : : :
gi|170 PQPQPFPSQQPYLQLQFPFPQPFLLPQLFYPPQPSFFPQYPQQRPKYLQPQQPISQQQ
          70           80           90          100          110          120
```

gi|170 AQQQQQQQQQQQQQQQQQILQQILQQQLIPCRDVLQQHNIHASSQVLQQSTYQLLQQL
130 140 150 160 170 180

gi|170 CCQQLLQIPEQSRCQAIHNVVHAIIMHQQEQQQQLQQQQQQQLQQQQQQQQQQPSSQV
190 200 210 220 230 240

gi|170 SFQQPQQQYPSSQGSFQPSQQNPQAQGSVQPPQLPQFAEIRNLALQTLPAMCNVYIPPHC
250 260 270 280 290 300

gi|170 STTIAPFGIFGTN
310

>>gi|21765 gi|21765|emb|CAA26385.1| unnamed protein prod (313 aa)
initn: 62 initl: 62 opt: 67 Z-score: 108.1 bits: 25.0 E(): 0.43
Smith-Waterman score: 67; 38.462% identity (57.692% similar) in 26 aa overlap
(3-28:80-105)

gi|217 MKTFLILALVATTATTAVRVVPQLQPKNPSQQQPQEQVPLVQQQQFPGQQQQFPPQQPY
10 20 30 40 50 60

CV127. MSQPLPLQPNLPNPKHSSPDSLQINPAKALISSSKL
:: :: :::: :: :: : :
gi|217 PQQPFFPSQQPYLQLQFPFPQPFPLQPLYPQPSFPPQQPYQQRPMYLQPPQIPISQQQ
70 80 90 100 110 120

gi|217 AQQQQQQQQQQQQQQQQQILQQILQQQLIPCRDVLQQHNIHASSQVLQQSTYQLLQQL
130 140 150 160 170 180

gi|217 CCQQLLQIPEQSRCQAIHNVVHAIIMHQQEQQQQLQQQQQQQLQQQQQQQQQQPSSQV
190 200 210 220 230 240

gi|217 SFQQPQQQYPSSQGSFQPSQQNPQAQGSVQPPQLPQFAEIRNLALQTLPAMCNVYIPPHC
250 260 270 280 290 300

gi|217 STTIAPFGIFGTN
310

>>gi|170738 gi|170738|gb|AAA34289.1| gamma-gliadin (327 aa)
initn: 250 initl: 67 opt: 67 Z-score: 107.8 bits: 25.0 E(): 0.45
Smith-Waterman score: 67; 47.059% identity (76.471% similar) in 17 aa overlap
(3-19:82-98)

gi|170 MKTLLILTILAMAITIGTANIQVDPQSGVQWLQQQLVPLQQLPLSQPPQQTFFPQQPQTFP
10 20 30 40 50 60

CV127. MSQPLPLQPNLPNPKHSSPDSLQINPAKALISSSKL
::: ::. : : :
gi|170 HQPQQQVPQPPQPPFLQPPQPPFPQPPQPPFPQTQQPQQPFPQPPQPPFPQTQQPQQPF
70 80 90 100 110 120

gi|170 PQQPQQPFPQTQQPQQPFPQLQQPQQPFPQPPQLPQPPQPSFPQQQRPFQPSLQQQ
130 140 150 160 170 180

gi|170 LNPCKNILLQQSKPASLVSSLWSIIWPQSDCQVMRQQCCQQLAQIPQQLQCAAIHVVHS
190 200 210 220 230 240

gi|170 IIMQQQQQQQQQGIDIFLPLSQHEQVGGSLVQGGIIPQPPAQLEAIRSLVLQTLPS
250 260 270 280 290 300

gi|170 MCNVYVPPECSIMRAPFASIVAGIGGQ
310 320

```
>>gi|170708 gi|170708|gb|AAA34274.1| gamma-gliadin B pre (291 aa)
  initn: 65 initl: 65 opt: 65 Z-score: 105.8 bits: 24.5 E(): 0.57
Smith-Waterman score: 65; 47.059% identity (70.588% similar) in 17 aa overlap
(3-19:90-106)
```

gi|170 MKTLLILITILAMAITIATANMQADPSGQVQWPQQQPFLQPHQPFSSQQPQQIFPQPQQTFP

CV127.
MSQPLPLQP¹⁰NLPNPKHSSPDSLQINPAKALISS³⁰
::: :::

gi|170 HQPQQQFPQPQPQQQFLQPRQFPQPQQPYYPQQPQQPFPQTQQPQQPFPQSKQPQQPF

CV127. SKL

gi|170 PQQPQQPSFPQQPSLIQQSLQQQLNPCKNFLQQCKPVSLVSSLWSIILPPSDCQVMR
130 140 150 160 170 180

gi|170 QQCCQQLAQIPQQLQCAAIHSSVHSSIIMQQEQQEQLQGVQILVPLSQQQQVGGILVQGQ

gi|170 GIIQPQQPAQLEVIRSLVLQTLPTMCNVYVPPYCSTIRAPFASIVASIGGQ
250 260 270 280 290

```
>gi|170724 gi|170724|gb|AAA34282.1| pre-alpha-/beta-gli (297 aa)
initn: 93 initl: 42 opt: 63 Z-score: 102.8 bits: 24.0 E(): 0.84
Smith-Waterman score: 63; 52.941% identity (76.471% similar) in 17 aa overlap
(3-19:79-94)
```

gi|170 MKTFLILALRAIVATTATIAVRVPVPQLQPQNPSQQQPQKQVPLVQQQFPGQQQPFPFQ

CV127.
MSQPLPLQP¹⁰NLPNPKHSSPDSL²⁰QINPAKALISSSKL³⁰
 ::: . :::: : . :

gi|170 QPYPQQQPFPSQQPYMQLQPFQ-QPQLPYQPQLPYQPQPFQFQQSYQPQPQYSQPQQ

gi|170 PISQQQQQQQQQQQQQQQILQQILQQQLIPCRDVVLQQHSIAHGSSQVLQQSTYQLVQQF

gi|170 CCQQLWQIPEQSRCQAIHNVVHAILHQQQQQQQQQQQQQQQPLSQVCFQQSQQQYPSGQ
180 190 200 210 220 230gi|170 GSFQPSQQNPQAQGSVQPQQLPQFEEIRNLAETLPAMCNVYIPPYCTIAPVGFIGTN
240 250 260 270 280 290

```
>>gi|21673 gi|21673|emb|CAA35238.1| unnamed protein prod (307 aa)
  initn: 126 initl: 42 opt: 63 Z-score: 102.5 bits: 24.0 E(): 0.87
Smith-Waterman score: 63; 52.941% identity (76.471% similar) in 17 aa overlap
(3-19:79-94)
```

gi|216 MKTFLILALLAIVATTARIAVRVPVPLQPNPSQQQPQEQVPLVQQQFPQGGQPFPPQ
10 20 30 40 50 60

```

                                10      20      30
CV127.      MSQPLPLQPNLNPKNHSSPDSLQINPAKALISSSKL
              :::: :::: :. . :
gi|216 QPYQPQPFPSQQPYLQLQPFQ-QPQLPYQPQLPYQPQLPYQPQPFQPFPSQQPYQPSQP
              70      80      90      100      110

gi|216 QYSQPQPISQQQQQQQQQQKQQQQQQQILQQILQQQLIPCRDVVLQQHSHIAYGSSQ
              120      130      140      150      160      170

gi|216 VLQQSTYQLVQQLCCQQLWQIPEQSRCQAIHNVVHAILHQQQQQQQQQQPLSQVSFQ
              180      190      200      210      220      230

gi|216 QPQQQYPSGQGSFQPSQQNPQAQGSVQPQQLPQFEEIRNLALETLPAMCNVYIPPYCTIA
              240      250      260      270      280      290

gi|216 PVGIFGTN
              300

>>gi|1304264 gi|1304264|dbj|BAA12318.1| alpha-gliadin [T (259 aa)
  initn: 82 initl: 60 opt: 62 Z-score: 102.4 bits: 23.7 E(): 0.89
Smith-Waterman score: 62; 36.000% identity (60.000% similar) in 25 aa overlap
(2-26:69-93)

gi|130 VRVPVQLQPQNPSQQQPQEQVPLVQQQQLGQQQPFPFPQQPYQPQPFPSQQPYLQLQP
              10      20      30      40      50      60

                                10      20      30
CV127.      MSQPLPLQPNLNPKNHSSPDSLQINPAKALISSSKL
              :::: :::: :. . :
gi|130 FPQPQLPYSQPQPFQPFQPYQPQPYQPQYSQPQEPISQQQQQQQQQQQILQQILQQQLIPCM
              70      80      90      100      110      120

gi|130 DVVLQQHNIAGHRSQVLQQSTYQLLQELCCQHLWQIPEQSQCAIQNVVHAILHQQQKQ
              130      140      150      160      170      180

gi|130 QQQPSSQVSFQQPLQQYPLGQGSFRPSQQNPQDQGSVQPQQLPQFEEIRNLALQTLPAMC
              190      200      210      220      230      240

gi|130 NVYIPPYCTIAPFGIFGTN
              250

>>gi|170722 gi|170722|gb|AAA34281.1| pre-alpha-/beta-gli (262 aa)
  initn: 106 initl: 60 opt: 62 Z-score: 102.3 bits: 23.7 E(): 0.9
Smith-Waterman score: 62; 36.000% identity (60.000% similar) in 25 aa overlap
(2-26:89-113)

gi|170 MKTFLILALLAIVATTATTAVRVPVQLQPQNPSQQQPQEQVPLVQQQQLGQQQPFPFPQ
              10      20      30      40      50      60

                                10      20      30
CV127.      MSQPLPLQPNLNPKNHSSPDSLQINPAKALISS
              :::: :::: :. . :
gi|170 QPYQPQPFPSQQPYLQLQPFQQLPYSQPQPFQPFQPYQPQPYQPQYSQPQQPISQQQQQ
              70      80      90      100      110      120

CV127. SKL

gi|170 QQQQQQQQQQQQQQIIQQILQQQLIPCMDVVLQQHNIHVHGSQVLQQSTYQLLQELCCQH
              130      140      150      160      170      180

gi|170 LWQIPEQSQCAIQNVVHAILHQQQKQQQPSSQVSFQQPLQQYPLGQGSFRPSQQNPQ
              190      200      210      220      230      240

```

gi|170 AQGSVQPQQLPQFEEIRNLARK
250 260

>>gi|62484809 gi|62484809|emb|CAI78902.1| putative gamma (285 aa)
initn: 59 initl: 59 opt: 62 Z-score: 101.7 bits: 23.7 E(): 0.98
Smith-Waterman score: 66; 30.769% identity (51.282% similar) in 39 aa overlap
(3-31:37-75)

CV127. MSQPLPLQPNLP-----NPKH
::: ::: :
gi|624 MQVNPSVQVQPTQQQPYEPESQQPFISQSQQQFPQPQQPFPQPPQSPQSQCLQQPQH
10 20 30 40 50 60

CV127. SSPDSLQINPAKALISSKL
.: : : :
gi|624 QFPQPTQQFPQRPLLPFTHPFLTFPDQLLPQPPHQSFPQPPQSYQPPLQFPFPQPPQKY
70 80 90 100 110 120

gi|624 PEQPQQPFPWQQPTIQLYLQQQLNPCKEFLQCCRPVSLLSYLWSKIVQQSSCRMVQQQC
130 140 150 160 170 180

gi|624 CLQLAQIPEQYKCTAIDSIVHAIFMQGQRQGVQIVQQPQPQQVGQCVLVQGGQGVVQPQ
190 200 210 220 230 240

gi|624 QLAQMEAIRTLVLQSVPSMCNFNVPNCSTIKAPFVGVTGVGGQ
250 260 270 280

>>gi|170720 gi|170720|gb|AAA34280.1| alpha/beta-gliadin (286 aa)
initn: 108 initl: 60 opt: 62 Z-score: 101.7 bits: 23.7 E(): 0.98
Smith-Waterman score: 62; 36.000% identity (60.000% similar) in 25 aa overlap
(2-26:89-113)

gi|170 MKTFLILVLLAIVATTATTAVRFPVPQLQPQNPSQQQPQEQVPLVQQQFLGQQQFPFPQ
10 20 30 40 50 60

CV127. MSQPLPLQPNLPNPKHSSPDSLQINPAKALISS
::: :::: : : : :
gi|170 QPYPQPQPFPSQLPYLQLQFPFPQPLPYSQFPQFPQPPQYQPPQYSQPPQISQQQQQ
70 80 90 100 110 120

CV127. SKL

gi|170 QQQQQQQQQQQQILQQILQQQLIPCMDVVLQQHNIAGRSQVLQQSTYQLLQELCCQHL
130 140 150 160 170 180

gi|170 WQIPEQSQCQAIHNVVHAILHQKQQQQPSSQVSFQQPLQQYPLGQGSFRPSQQNPQA
190 200 210 220 230 240

gi|170 QGSVQPQQLPQFEEIRNLALQTLPAMCNVYIPPYCTIAPFGIFGTN
250 260 270 280

>>gi|21755 gi|21755|emb|CAA25593.1| unnamed protein prod (286 aa)
initn: 108 initl: 60 opt: 62 Z-score: 101.7 bits: 23.7 E(): 0.98
Smith-Waterman score: 62; 36.000% identity (60.000% similar) in 25 aa overlap
(2-26:89-113)

gi|217 MKTFLILVLLAIVATTATTAVRFPVPQLQPQNPSQQQPQEQVPLVQQQFLGQQQFPFPQ
10 20 30 40 50 60

```

                                10      20      30
CV127.      MSQPLPLQPNLPNPKHSSPDSLQINPAKALISS
              ::: :...: : :. . : :. :
gi|217 QPYPQPQPFPSQLPYLQLQPFQQLPYSQLPQPFQPPQYPQPPQYSQPPQPISSQQQQQ
              70      80      90      100      110      120

```

CV127. SKL

```

gi|217 QQQQQQQQQQQQQILQQILQQQLIPCMDVVLQQHNIAGHRSQVLQQSTYQLLQELCCQHL
              130      140      150      160      170      180

```

```

gi|217 WQIPEQSQCQAILKVVHAILHQQKQQQPPSSQVSFQQPLQQYPLGQGSFRPSQQNPQA
              190      200      210      220      230      240

```

```

gi|217 QGSVQPQLPQFEEIRNLALQTL PAMCNVYIPPYCTIAPFGIFGTN
              250      260      270      280

```

>>gi|473876 gi|473876|gb|AAA17741.1| alpha-gliadin (287 aa)
 initn: 106 initl: 60 opt: 62 Z-score: 101.6 bits: 23.7 E(): 0.98
 Smith-Waterman score: 62; 36.000% identity (60.000% similar) in 25 aa overlap
 (2-26:89-113)

```

gi|473 MKTFLILALLAIVATTATTAVRVVPQLQPNPSQQQPQEQVPLVQQQFLGQQQPFPPQ
              10      20      30      40      50      60

```

```

                                10      20      30
CV127.      MSQPLPLQPNLPNPKHSSPDSLQINPAKALISS
              ::: :...: : :. . : :. :
gi|473 QPYPQPQPFPSQQPYLQLQPFQQLPYSQLPQPFQPPQYPQPPQYSQPPQPISSQQQQQ
              70      80      90      100      110      120

```

CV127. SKL

```

gi|473 QQQQQQQQQQQQQIIQQILQQQLIPCMDVVLQQHNIHVHGKSQVLQQSTYQLLQELCCQH
              130      140      150      160      170      180

```

```

gi|473 LWQIPEQSQCQAIHNVVHAILHQQKQQQPPSSQVSFQQPLQQYPLGQGSFRPSQQNPQ
              190      200      210      220      230      240

```

```

gi|473 AQGSVQPQLPQFEEIRNLALQTL PAMCNVYIPPYCTIAPFGIFGTN
              250      260      270      280

```

>>gi|170726 gi|170726|gb|AAA34283.1| pre-alpha-/beta-gli (282 aa)
 initn: 109 initl: 60 opt: 61 Z-score: 100.3 bits: 23.5 E(): 1.2
 Smith-Waterman score: 61; 37.500% identity (54.167% similar) in 24 aa overlap
 (3-26:78-101)

```

gi|170 MKTFLILALLAIVATTATS AVRVVPQLQPNPSQQQPQEQVPLMQQQQFPQQEQFPF
              10      20      30      40      50      60

```

```

                                10      20      30
CV127.      MSQPLPLQPNLPNPKHSSPDSLQINPAKALISSSKL
              :: :. :...: :. . : :. :
gi|170 QQPYPHQPPFPSQQPYPPQPFPPQLPYPTQPFPPQQPYPPQPPQYPQPPQPISSQQQAQ
              70      80      90      100      110      120

```

```

gi|170 QQQQQQQQLTQQILQQQLIPCRDVVLQQHNI AHASSQVLQQSSYQQQLQQQLCCQQLFQIPEQ
              130      140      150      160      170      180

```

gi|170 SRCQAIHNVVHAILLHHHQQQQQPPSSQVSYQQPQEQYPSGQVSFQSSQQNPQAQGSVQP
190 200 210 220 230 240

gi|170 QQLPQFQEIIRNLALQTLPMCNVYIPPYCSTTIAPFGIFGTN
250 260 270 280

>>gi|170740 gi|170740|gb|AAA34290.1| gliadin (296 aa)
initn: 108 initl: 60 opt: 61 Z-score: 100.0 bits: 23.5 E(): 1.2
Smith-Waterman score: 62; 41.176% identity (76.471% similar) in 17 aa overlap
(3-19:27-43)

CV127. MSQPLPLQPNLNPKNHSSPDSLQINPAKALISSSKL
:: : .:. :.:. :
gi|170 MKTFLILALLAIVATTATTAVRVVPVQPQNPSPQPPQQRQVPLVQQQFPGQQQFPPQ
10 20 30 40 50 60

gi|170 QPYQPQPPFSSQQPYLQLQPPFPQPPFPQLPYQPPPFSPQQPYQPQPPQPPQPPIS
70 80 90 100 110 120

gi|170 QQQAQQQQQQQQQQQQQQQQILPQILQQQLIPCRDVVLQQHNIHARSQVLQQSTYQP
130 140 150 160 170 180

gi|170 LQQLCQQQLWQIPEQSRCQAIHNVVHAILHQQQQQQPPSSQVSLQQPQQQYPSGQGFFQ
190 200 210 220 230 240

gi|170 PSQQNPQAQGSVQPQQLPQFEEIRNLALQTLPRMCNVYIPPYCSTTTAPFGIFGTN
250 260 270 280 290

>>gi|21757 gi|21757|emb|CAA26383.1| unnamed protein prod (296 aa)
initn: 108 initl: 60 opt: 61 Z-score: 100.0 bits: 23.5 E(): 1.2
Smith-Waterman score: 61; 37.500% identity (54.167% similar) in 24 aa overlap
(3-26:83-106)

gi|217 MKTFLILALLAIVATTATTAVRVVPVQPQNPSPQPPQQRQVPLVQQQFPGQQQFPPQ
10 20 30 40 50 60

CV127. MSQPLPLQPNLNPKNHSSPDSLQINPAKALISSSKL
:: : .:. :.:. :
gi|217 QPYQPQPPFSSQQPYLQLQPPFPQPPFPQLPYQPPPFSPQQPYQPQPPQPPQPPIS
70 80 90 100 110 120

gi|217 QQQAQQQQQQQQQQQQQQQQILPQILQQQLIPCRDVVLQQHNIHARSQVLQQSTYQP
130 140 150 160 170 180

gi|217 LQQLCQQQLWQIPEQSRCQAIHNVVHAILHQQQQQQPPSSQVSLQQPQQQYPSGQGFFQ
190 200 210 220 230 240

gi|217 PSQQNPQAQGSVQPQQLPQFEEIRNLALQTLPRMCNVYIPPYCSTTIAPFGIFGTN
250 260 270 280 290

>>gi|75219081 gi|75219081|sp|O22108|O22108_WHEAT LMM glu (285 aa)
initn: 45 initl: 45 opt: 60 Z-score: 98.8 bits: 23.2 E(): 1.4
Smith-Waterman score: 60; 32.353% identity (67.647% similar) in 34 aa overlap
(3-35:13-46)

CV127. MSQPLPLQPNLNP-KNHSSPDSLQINPAKALISSSKL
:::: : .:. : : . . . : . : . : . :
gi|752 RCIPGLERPWQQQLPQQTFPQQPLFSQQQQQQLFPQQPSFSQQQPPFWQQQPPFSQQQ
10 20 30 40 50 60

gi|752 PILPQQPPFSQQQQLVLPQQSPFSQQQQLILPPQQQQQLPQQQISIVQPSVLQQLNPCKV
70 80 90 100 110 120

gi|752 FLQQQCSPVAMPQRLARSQMWWQSSCHVMQQQCCQQLSQIPEQSRYDAIRAITYSIILQE

gi|752 QQQGFVQAQQQQPQQSGQGVSSQSQQQSQQQLGQCSFQQPQQQLGQQPQQQQVQQGTFLQP

gi|752 HQIAHLEVMTSIALRTLPMCSVNVPLYSSTTSVPFGVGTGVGAY
250 260 270 280

```
>>gi|21773 gi|21773|emb|CAA31685.1| unnamed protein prod (307 aa)
  initn: 45 initl: 45 opt: 60 Z-score: 98.3 bits: 23.2 E(): 1.5
Smith-Waterman score: 60; 32.353% identity (67.647% similar) in 34 aa overlap
(3-35:36-69)
```

CV127. MSQPLPLQPNLPN-PKHSSPDSLQINP
 ::::: : . . . : : . . . : .

gi|217 MKTFLV FALLAVAATSAIAQMETRCIPGLERPWWQQQLPLPPQQTFFPQQPLFSQQQQQQQLFP
 10 20 30 40 50 60

CV127. AKALISSSKL
gi|217 QQPSFSQQQPPFWQQQPPFSQQQPILPQQPPFSQQQQLVLPQQPPFSQQQPVLPQQSP

gi|217 FPQQQQQHQQLVQQQIPVVQPSILQQLNPCKVFLQQQCSPVAMPQRLARSQMLQQSSCHV

gi|217 MQQQCCQQLPQIPQQSRYEAIRAIIYSIILQEQQVQVQSGSIQSQQQQPQQLGQCVSQPQQQ

gi|217 SQQQLGQQPQQQLAQGTFLQPHQIAQLVMTSIALRILPTMCSVNVPLYRTTTSVPFGV
250 260 270 280 290 300

gi | 217 GTGVGAY

```
>>gi|170712 gi|170712|gb|AAA34276.1| pre-alpha-/beta-gli (291 aa)
  initn: 65 initl: 42 opt: 58 Z-score: 95.8 bits: 22.7 E(): 2.1
Smith-Waterman score: 58; 38.462% identity (61.538% similar) in 26 aa overlap
(3-26:55-80)
```

CV127. MSQPLPLQ

gi|170 MKTFPILALLAIVATTATTAVRVPVPQLQLQNPSQQQPQEQVPLVQEQQFQGGQQPFPPQ

10 20 30 40 50 60

```

          10          20          30
CV127. PNLPNPKH--SSPDSLQINPAKALISSSKL
          :.:.: :.:.: :.:.:
gi|170  QPYPPQPFPSQQPYLQLQFPFPQPLPYPPQPFRRPQQPYPPQPPQYSQPQQPISQQQQQ
          70          80          90          100          110          120

```

gi|170 QQQQQQQQQQILQQILQQQLIPCRDVVLQQHNIAGSSQVLQESTYQLVQQQLCCQQLWQI

gi|170 PEQSRCQAIHNVVHAIILHQQHHHHQQQQQQQQQPLSQVVSFQQPQQYPSGQGFQPSQ
190 200 210 220 230 240

gi|170 QNPQAQGSFQPQLPQFEEIRNLALQTLPMCNVYIPPYCTIAPFGIFGTN
250 260 270 280 290

>>gi|170702 gi|170702|gb|AAA34272.1| gamma gliadin precu (302 aa)
initn: 74 initl: 47 opt: 58 Z-score: 95.5 bits: 22.7 E(): 2.1
Smith-Waterman score: 60; 43.750% identity (75.000% similar) in 16 aa overlap
(4-19:106-121)

gi|170 MKTLLILTILAMATTIATANMQVDPGQVQWPQQQPFQPPQPPFCQQPQRTIPQPHQTFH
10 20 30 40 50 60

CV127. MSQPLPLQPNLNPKNHSS
10
:: :: : : . .
gi|170 HQPQQTFPQPQQTYPHQPQQQFPQTQQPQQPFPQPQQTFPQQPLPFPQQPQQPFPQQ
70 80 90 100 110 120

20 30
CV127. PDSLQINPAKALISSSKL
:
gi|170 PQQPFPQSQQPQQPFPQPQQQFPQPQQPQQSFPQQQPAIQSFLQQMNPCKNFFLLQQCN
130 140 150 160 170 180

gi|170 HVSLVSSLVSIILPRSDCQVMQQQCCQQLAQIPQQLQCAAIHSAHSHIIMQQEQQQGVPI
190 200 210 220 230 240

gi|170 LRPLFQLAQGLGIIQPQQPAQLEGIRSLVLKTLPTMCNVYVPPDCSTINVPYANIDAGIG
250 260 270 280 290 300

gi|170 GQ

>>gi|21930 gi|21930|emb|CAA44473.1| LMW glutenin [Tritic (285 aa)
initn: 76 initl: 45 opt: 57 Z-score: 94.6 bits: 22.4 E(): 2.4
Smith-Waterman score: 57; 32.353% identity (64.706% similar) in 34 aa overlap
(3-35:36-69)

10 20
CV127. MSQPLPLQPNLNP-PKHSSPDSLQINP
:::: : :: : : : . . : :
gi|219 MKTFLVFALLAVVATSAIAQMDTSCIPGLERPWWQQPLPPQTFPQQPPFSQQQQQQPFP
10 20 30 40 50 60

30
CV127. AKALISSSKL
: : : :
gi|219 QQPSFSQQQPILPQGPPFPQQTQPVLPQQSPFSQQQQLILPPQQQQQLPQQQISIVQPSV
70 80 90 100 110 120

gi|219 LQQLNPKVFLQQQCSPVAIPQRLARSQMWQQSSCHVMQQCCQQLSQIPEQSRDAIRA
130 140 150 160 170 180

gi|219 ITYSIILQEQQQQSQQQQPQQSGQVSQSQQQSQQQLGQCSFQQPQQQLGQQPQQQQVQ
190 200 210 220 230 240

gi|219 QGTFLQPHQIAHLEVMTSIALRTLPTMCSNVNPLYSSTTSVPFGV
250 260 270 280

>>gi|76782247 gi|76782247|gb|ABA54897.1| hydrophobic see (134 aa)
initn: 42 initl: 42 opt: 51 Z-score: 91.6 bits: 20.8 E(): 3.5
Smith-Waterman score: 51; 47.826% identity (56.522% similar) in 23 aa overlap
(4-26:39-58)

```

                                10      20
CV127.      MSQPLPLQPNLNPKNHSSPDSLQIN
              : : : : : : : : : :
gi|767 MGSKVVASVALLLSINILFISMVSSSSHYDPPPPPCYVPAPLTP--P-PSLSPPPSLSPP
              10      20      30      40      50

```

```

          30
CV127. PAKALISSSKL
          :
gi|767 PPSGPSCPDLSVCLNILDGSPADDCCALIADLVDLEASVCLCIQLRVLGIVNLDLNLQLI
          60      70      80      90      100     110

```

```

gi|767 LNACGPSYPSNATCPRT
          120     130

```

>>gi|21926 gi|21926|emb|CAA36063.1| unnamed protein prod (295 aa)
initn: 38 initl: 38 opt: 55 Z-score: 91.4 bits: 21.9 E(): 3.6
Smith-Waterman score: 55; 36.667% identity (60.000% similar) in 30 aa overlap
(3-27:36-65)

```

                                10      20
CV127.      MSQPLPLQPNL-----PNPKHSSPDSL
              : : : : : : : : : :
gi|219 MKTFLV FALLAVVATSTIAQMETSCLIPGLERPWQEQPLPPQHTLFPQQQFPQQQPPFS
              10      20      30      40      50      60

```

```

          30
CV127. QINPAKALISSSKL
          : : :
gi|219 QQQPSFLQQQPILPQLPFSQQQPPVLPQQSPFSQQQLVLPPQQQYQQVLQQQIPIVQPSV
          70      80      90      100     110     120

```

```

gi|219 LQQLNPKCKVFLQQQCNFVAMPQRLARSQMLQQSSCHVMQQQCCQQLPQIPEQSRVDVIRA
          130     140     150     160     170     180

```

```

gi|219 ITYSIILQEQQQGFVQAQQQQPQQLGQGVQSQQQSQQQLGQCSFQQPQQQLGQQPQQQQ
          190     200     210     220     230     240

```

```

gi|219 VLQGTFLQPHQIAHLEVMTSIALRTLPTMCSNVNPLYSTTSVPFSGTGVGAYL
          250     260     270     280     290

```

>>gi|1063270 gi|1063270|dbj|BAA11251.1| gamma-gliadin pr (279 aa)
initn: 43 initl: 43 opt: 53 Z-score: 89.0 bits: 21.3 E(): 4.9
Smith-Waterman score: 57; 37.037% identity (62.963% similar) in 27 aa overlap
(3-26:16-41)

```

                                10      20      30
CV127.      MSQPLPLQPNLP---NPKHSSPDSLQINPAKALISSSKL
              : : : : : : : : : :
gi|106 NIQVDPGQVQWPQQQFPF-QPHQPFSSQQPQQTFFPQPQQTFFPHQPQQQFSQPQQPQQQFI
              10      20      30      40      50

```

```

gi|106 QPQQPFPQQPQQTYPPQRPQQPFPQTQQPQQPFPQSQQPQQPFPQPQQQFPQPQQPQQQSF
          60      70      80      90      100     110

```

```

gi|106 QQQPSLIQQSLQQQLNPKCNFLQCKPVSLVSSLWSMILPRSDCQVMRQCCQQLAQIP
          120     130     140     150     160     170

```


gi|106 QQLQCAAIHSIVHSIIMQQEQQEQRQGVQILVPLSQQQQVGGQTLVQGGIIQPQQPAQL
180 190 200 210 220 230

gi|106 EVIRSSVLQTLATMCNVYVPPYCSTIRAPFASIVAGIGGQ
240 250 260 270

>>gi|170734 gi|170734|gb|AAA34287.1| gamma gliadin B-III (244 aa)
initn: 124 initl: 52 opt: 52 Z-score: 88.6 bits: 21.1 E(): 5.2
Smith-Waterman score: 57; 33.333% identity (54.545% similar) in 33 aa overlap
(3-24:3-35)

10 20 30
CV127. MSQPLPLQPN-----LPNPKHSSPDSLQINPAKALISSSKL
:::.....: .:. :. :.
gi|170 PQQPFPLQPQQSFLWQSQQPFLQQPQQPSPQPQQVVQIISPATPTTIPSAGKPTSAPFPQ
10 20 30 40 50 60

gi|170 QQQQHQQLAQQQIPVVQPSILQQLNPKCKVFLQQQCSVPAMPQRLARSQMLQQSSCHVMQQ
70 80 90 100 110 120

gi|170 QCCQQLPQIPQQSRYQAIRAIYSIILQEQQQVQGSIQSQQQQPQQLGQCVSQPQQQSQQ
130 140 150 160 170 180

gi|170 QLGQQPQQQQLAQGTFLQPHQIAQLEVMTSIALRILPTMCSVNVPLYRTTTSVPFGVGTG
190 200 210 220 230 240

gi|170 VGAY

>>gi|3915783 gi|3915783|sp|P43217.3|NLT11_PARJU Probable (139 aa)
initn: 49 initl: 49 opt: 49 Z-score: 88.5 bits: 20.2 E(): 5.3
Smith-Waterman score: 49; 34.783% identity (60.870% similar) in 23 aa overlap
(5-27:97-119)

gi|391 QETCGTMRVRLMPCLPFVQGKEKEPSKGCCSGAKRLDGETKTGPQRVHACECIQTAMKTY
10 20 30 40 50 60

10 20
CV127. MSQPLPLQPNLPNPKHSSPDSLQINPAK
.: :. :.
gi|391 SDIDGKLVSEVPKHCGIVDSKLPPIIDVNMCKTVGVVPRQPQLPVSLRHGPPVTGSPDPAH
70 80 90 100 110 120

30
CV127. ALISSSKL

gi|391 KARLERPQIRVPPPAPEKA
130

>>gi|162811 gi|162811|gb|AAA30433.1| kappa-casein precur (190 aa)
initn: 35 initl: 35 opt: 50 Z-score: 87.6 bits: 20.5 E(): 5.9
Smith-Waterman score: 50; 33.333% identity (66.667% similar) in 30 aa overlap
(3-30:67-95)

gi|162 MMKSFFLVVTILALTLPFLGAQEQNQEPIRCEKDERFFSDKIAKYIPIQYVLSRYPSYG
10 20 30 40 50 60

10 20 30
CV127. MSQPLPLQPN--LPNPKHSSPDSLQINPAKALISSSKL
.: : : : : : : : : :
gi|162 LNNYQQKPVALINNQFLPYYPYAKPAAVR-SPAQILQWQVLSNTVPAKSCQAQPTTMARH
70 80 90 100 110

gi | 162 PHPHLSFMAIPPCKKNQDKTEIPTINTIASGEPTSTPTIEAVESTVATLEASPEVTESPP

gi|162 INTVQVTSTAV
180 190

```
>>gi|1228078 gi|1228078|emb|CAA33034.1| kappa-casein [Bo (190 aa)
  initn: 35 initl: 35 opt: 50 Z-score: 87.6 bits: 20.5 E(): 5.9
Smith-Waterman score: 50; 33.333% identity (66.667% similar) in 30 aa overlap
(3-30:67-95)
```

gi|122 MMKSFFLVVITILALTLPLGAQEQNQEQPIRCEKDERFFSDKIAKYIPIQYVLSRPSYG
10 20 30 40 50 60

CV127. MSQPLPLQPN--LPNPKHSSPDSLQINPAKALISSSKL
gi|122 LNYYQQKPVALINNQLFPYPYAKPAAVR-SPAQILQWQVLSNTVPKSCQAQPTTMARH

gi|122 PHPHLSFMAIPPKKNQDKTEIPTINTIASGEPTSTPTTEAVESTVATLEDSPEVIESPPE

gi|122 INTVQVTSTAV
180 190

```
>>gi|21761 gi|21761|emb|CAA26384.1| unnamed protein prod (286 aa)
  initn: 97 initl: 49 opt: 52 Z-score: 87.4 bits: 21.1 E(): 6.1
Smith-Waterman score: 55; 38.462% identity (61.538% similar) in 26 aa overlap
(3-26:55-80)
```

CV127. MSQPLPLQ
::: :
gi | 217 MKTFLILVLLAIVATTATTAVRFPVPQLQPQNPSQQQPQEQVPLVQQQQLGQQQFPFPQ
10 20 30 40 50 60

```

          10          20          30
CV127. PNLPNPKH--SSPDSLQINPAKALISSSKL
          :.: :.: :.: :.: :.:
gi|217 QPYPPQPFPSQLPYLQLQFFPQPQLPYSQPQQFRPQQPYFPQPQPQYSQPQQPISQQQQQ
          70          80          90          100          110          120

```

gi | 217 QQQQQQQQQQQQQILQQILQQQLIPCMDVVLQQHNIAHGRSQVLQQSTYQLLQELCCQHL
130 140 150 160 170 180gi|217 WQIPEQSQCAIHNVVHAILHQQQKQQQQPSSQVSFQQPLQQYPLGQGSFRPSQQNPQA
190 200 210 220 230 240gi|217 QGSVQPQQLPQFEEIRNLALQTLPAMCNVYIPPYCTIAPFGIFGTN
250 260 270 280

```
>>gi|1532058 gi|1532058|emb|CAA65123.1| P9 protein [Pari (176 aa)
  initn: 49 initl: 49 opt: 49 Z-score: 86.8 bits: 20.3 E(): 6.6
Smith-Waterman score: 49; 34.783% identity (60.870% similar) in 23 aa overlap
(5-27:134-156)
```

gi|153 MRTVSAPSAVALVVIVAAGLAWTSLASVAPPAPAPGSEETCGTVVRALMPCLPFVQGKEK
10 20 30 40 50 60gi|153 EPSKGCCSGAKRLDGETKTGLQRVHACECIQTAMKTYSDIDGKLVSEVPKHCGIVDSKLP
70 80 90 100 110 120

```

              10      20      30
CV127.      MSQPLPLQPNLPNPKHSSPDSLQINPAKALISSSKL
              .: :.:.: .: .: .: .:
gi|153 PIDVNMCKTLGVVPRQPQLPVSLRHGPVTGPSDBAHKARLERPQIRVPPPAPEKA
              130      140      150      160      170

>>gi|886963 gi|886963|emb|CAA59338.1| low molecular weig (229 aa)
  initn: 67 initl: 42 opt: 50 Z-score: 86.2 bits: 20.5 E(): 7.1
Smith-Waterman score: 52; 42.308% identity (57.692% similar) in 26 aa overlap
(3-23:30-55)

```

```

              10      20
CV127.      MSQPLPLQPNL-----PNPKHSSPDSLQINPAK
              .: :.:.: .: .: .: .:
gi|886 FALIAVVATSTIAQMETSCLIPGLERPWQQQLPPLPQQTLFPQQQPFPPQQPPFSQQQPSFS
              10      20      30      40      50      60

```

```

              30
CV127. ALISSSKL

gi|886 QQQPPFSQQQPILPEPPFSLQQQPVLPQQSPFSQQQLVLPPQQQQQLPQQQISIVQPSVL
              70      80      90      100      110      120

gi|886 QQLNPCKVFLQQQCSPVAMPQRLARSQMWQQSRCHVMQQCCQQLSQIPEQSRDAICAI
              130      140      150      160      170      180

gi|886 TYSIILQEQQQGFVQAQQQQPQQSGQVQSQQQSQQQLGQCSFQQPQQ
              190      200      210      220

```

```

>>gi|170730 gi|170730|gb|AAA34285.1| gamma-gliadin B-I p (304 aa)
  initn: 71 initl: 49 opt: 51 Z-score: 85.5 bits: 20.8 E(): 7.7
Smith-Waterman score: 51; 42.857% identity (66.667% similar) in 21 aa overlap
(3-23:56-74)

```

```

CV127.      MSQPLPL
              .: :.:.:
gi|170 MKTFLVFAVATSAIAQMETSCLIPGLERPWQQQLPPLPQQSFSQQPPFSQQQQQLPPLPQ
              10      20      30      40      50      60

```

```

              10      20      30
CV127. QPNLPNPKHSSPDSLQINPAKALISSSKL
              .: :.:.: .: .: .: .:
gi|170 QPSFS--QQQPPFSQQQPILSQQPPFSQQQQPVLPQQSPFSQQQLVLPPQQQQQLVQQ
              70      80      90      100      110

gi|170 QIPVQPSVLQQLNPKVFLQQQCSPVAMPQRLARSQMWQQSSCHVMQQCCQQLQIPE
              120      130      140      150      160      170

gi|170 QSRYEAIRAIYIYSIILQEQQQGFVQPQQQQPQQSGQVQSQQQSQQQLGQCSFQQPQQQ
              180      190      200      210      220      230

gi|170 LGQQPQQQQQQVQLGTFQLPHQIAHLEAVTSIALRTLPTMCSVNVPLYSATTSVPFVG
              240      250      260      270      280      290

```

```

gi|170 TGVGAY
              300

```

```

>>gi|886965 gi|886965|emb|CAA59339.1| low molecular weig (261 aa)
  initn: 73 initl: 45 opt: 50 Z-score: 85.2 bits: 20.6 E(): 8

```

Smith-Waterman score: 53; 40.000% identity (64.000% similar) in 25 aa overlap
(3-26:33-55)

```

CV127.
MSQPLPLQPNLPNPKHSSPD-SLQINPAKA
::: :: : ..... : : .
gi|886 HIPSLEKPLQQQPLPLQQILWYQQQPIQQQPQFPFPQP--PCSQQQPPPLSQQQPPFS
      10      20      30      40      50

```

```

CV127. LISSSKL
      30
gi|886 QQQPPFSQQQPIPLPQQPPFSQQQQQFPQQQPLLPQQPPFSQQQPPFSQQQQQPPFSQQ
      60      70      80      90     100     110
gi|886 QQQPILLQQPPFSQHQQPVLPQQQIPSVQPSILQQLNPCKVFLQQQCSPVAMPQSLARSQ
      120     130     140     150     160     170

```

```

gi|886 MLWQSSCHVMQQCCCRQLPQIPEQSRDYAIRAIYSIVLQEQQHGGQLNQPPQQQPQQSV
      180     190     200     210     220     230

```

```

gi|886 QGVSQPQQQKQLGQCSFQQPQQ
      240     250     260

```

>>gi|62550933 gi|62550933|emb|CAI79052.1| putative LMW-g (326 aa)
initn: 37 initl: 37 opt: 51 Z-score: 85.0 bits: 20.8 E(): 8.3
Smith-Waterman score: 51; 36.667% identity (63.333% similar) in 30 aa overlap
(3-31:133-159)

```

gi|625 MSHIPGLERPSQQQPLRPQQTLSHHHQQPIQQQPQQFPQQQPCSQQQQPPPLSQQQPP
      10      20      30      40      50      60
gi|625 FSQQQPPFSQQQPVLPQQPSFSQQQLPFPFSQQQPPFSQQQPVLPQQPSFSQQQLPP
      70      80      90     100     110     120

```

```

CV127.
MSQPLPLQPNLPNPKHSSPDSLQ-INPAKALISSSKL
::: :: : : : : : : : : : : :
gi|625 FSQQLPFPFSQQQPVLLQQQIP--FVHPSILQQLNPCKVFLQQQCSPVAMPQSLARSQM
      130     140     150     160     170

```

```

gi|625 LQSSSCHVMQQCCQQLPQIPQQSRYEAIRAIVYSIILQEQQVQGSIQTQQQPPQQLGQ
      180     190     200     210     220     230

```

```

gi|625 CVSQPQQSQQQLGQCSFQQPQQQLQQLGQQPQQQIPQGIFLQPHQISQLEVMTSIALRT
      240     250     260     270     280     290

```

```

gi|625 LPTMCGVNVPLYSSTTIMPFSGTGVGGY
      300     310     320

```

>>gi|886967 gi|886967|emb|CAA59340.1| low molecular weig (276 aa)
initn: 71 initl: 45 opt: 50 Z-score: 84.8 bits: 20.6 E(): 8.5
Smith-Waterman score: 53; 40.000% identity (64.000% similar) in 25 aa overlap
(3-26:33-55)

```

CV127.
MSQPLPLQPNLPNPKHSSPD-SLQINPAKA
::: :: : ..... : : .
gi|886 HIPSLEKPSQQQPLPLQQILWYHQQQPIQQQPQFPFPQP--PCSQQQPPPLSQQQPPFS
      10      20      30      40      50

```

30
CV127. LISSSKL

gi|886 QQQPPFSQQELPILPQQPPFSQQQQPQFSQQQQPFPQQQQPLLLQQPPFSQQRPFFSQQQ
60 70 80 90 100 110

gi|886 QQPVLPQQPPFSQQQQQPILPQQPPFSLHQQPVLPQQQIPYVQPSILQQLNPCKVFLQQ
120 130 140 150 160 170

gi|886 QCSPVAMPQSLARSQMLWQSSCHVMQQCCQQLPRIPEQSRYPDRAIRAIYSIVLQEQQHG
180 190 200 210 220 230

gi|886 QGFNQPPQQPQQSVQGVSPQQQQKQLGQCSFQQPQQ
240 250 260 270

36 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:16:33 2008 done: Sun Dec 14 22:16:33 2008
Total Scan time: 0.070 Total Display time: 0.030

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.3783_d3

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.3783_d3, 31 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 3.8021 \pm 0.00408$; $\mu = 1.1706 \pm 0.212$
mean_var=51.9844 \pm 12.802, 0's: 39 Z-trim: 40 B-trim: 50 in 1/41
Lambda= 0.177884

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:		opt	bits	E(1313)	%_id	%_sim	alen
gi 76666767	gi 76666767 emb CAA5507 (497)	70	26.3	0.24	0.387	0.613	31
gi 170708	gi 170708 gb AAA34274.1 (291)	52	21.5	4	0.261	0.652	23
gi 3913017	gi 3913017 sp P81496.1 A (27)	42	17.8	4.8	0.385	0.615	13
gi 1063270	gi 1063270 dbj BAA11251. (279)	49	20.7	6.7	0.217	0.652	23
gi 1169665	gi 1169665 sp P30438.2 F (92)	44	18.9	7.7	0.308	0.692	13
gi 256429	gi 256429 gb AAB23464.1 (216)	47	20.1	8.1	0.333	0.625	24
gi 18770	gi 18770 emb CAA45777.1 t (217)	47	20.1	8.1	0.333	0.625	24
gi 1514943	gi 1514943 dbj BAA08246. (375)	49	20.8	8.2	0.290	0.516	31
gi 170736	gi 170736 gb AAA34288.1 (251)	47	20.1	8.9	0.227	0.636	22
gi 50659891	gi 50659891 gb AAT80665 (115)	44	19.0	9	0.296	0.519	27
gi 50659889	gi 50659889 gb AAT80664 (115)	44	19.0	9	0.296	0.519	27
gi 50659885	gi 50659885 gb AAT80662 (115)	44	19.0	9	0.296	0.519	27
gi 38492338	gi 38492338 gb AAR22488 (115)	44	19.0	9	0.222	0.519	27
gi 14285595	gi 14285595 sp O04004.1 (118)	44	19.0	9.1	0.368	0.632	19

>>>CV127.3783_d3, 31 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|76666767 gi|76666767|emb|CAA55071.2| aldehyde dehyd (497 aa)
initn: 43 initl: 43 opt: 70 Z-score: 112.7 bits: 26.3 E(): 0.24
Smith-Waterman score: 70; 38.710% identity (61.290% similar) in 31 aa overlap
(3-31:146-176)

gi|766 MTSVKLSTPQTGEFEQPTGLFINNEFVKAVDGKTFDVINPSTEEVICSVQEATEKDVDIA
10 20 30 40 50 60

gi|766 VAAARKAFNGPWRKETPENRGKLLNKLADLFKKNADLIAAVEALDNGKAFSMAKNVDVPA
70 80 90 100 110 120

CV127. MMSCRYICLEMGLCMQITLWS--IVICCWRLG
: : : : : : : : : : : : : : : :
gi|766 AAGCLRYGGWADKIEGKVVDTPDSFNYIRKEPIGVCQIIPWNFPILMWSWKIGPAIA
130 140 150 160 170 180

gi|766 TGNTVVLKTAEQTPLSAYIACKLIQEAGFPFPGVINVITGFGKIAGAAMSAMHMDIDKIAFT
190 200 210 220 230 240

gi|766 GSTVVGRQIMKSAAGSNLKKVLTLELGKSPNIVFADADLDEAIHWVNFNGIYFNHGQACCA
250 260 270 280 290 300

gi|766 GSRIYVQEEIYDKFIQRFKERAQAQNAVGDPPFAADTFQGFPQVSQLQFDRIMGYIEEGKKS
310 320 330 340 350 360

gi|766 ATIETGGRKGDKGFIETIFSNVTEDMKIQEEIFGPVCTISKFKTKADVIKIGNNTT
370 380 390 400 410 420

gi|766 YGLAAAVHTSNLTTAIEVANALRAGTVWVNSYNLHWQLPFGGYKESGIGRELGEAALDN
430 440 450 460 470 480

gi|766 YIQTKTVSIRLGDVLF
490

>>gi|170708 gi|170708|gb|AAA34274.1| gamma-gliadin B pre (291 aa)
initn: 36 initl: 36 opt: 52 Z-score: 90.6 bits: 21.5 E(): 4
Smith-Waterman score: 52; 26.087% identity (65.217% similar) in 23 aa overlap
(4-25:149-171)

gi|170 MKTLLILITILAMAITIATANMQADPSGQVQWPQQPFLQPHQPFSSQPPQIFPQPQQTFP
10 20 30 40 50 60

gi|170 HQPQQQFPQPQQPQQFLQPRQPFQPPQPPQPPQPPQTQQPQQPFQSKQPQQPF
70 80 90 100 110 120

CV127. 10 20 30
MMSCR-YICLEMGLCMQITLWSIVICWRLG
:. . . : . . . :
gi|170 PQPQQPQSFQPPQPSLIQQSLQQQLNPCKNLLQCKPVSLSVSSLSIILPPSDCQVMR
130 140 150 160 170 180

gi|170 QQCCQQLAQIPQQLQCAAIHVSVHSIIMQQEQEQQLQGVQILVPLSQQQQVGQILVQGG
190 200 210 220 230 240

gi|170 GIIQPQQPAQLEVIRSLVLQTLPTMCNVVPPYPCSTIRAPFASIVASIGGQ
250 260 270 280 290

>>gi|3913017 gi|3913017|sp|P81496.1|ALCC_WHEAT Allergen (27 aa)
initn: 34 initl: 34 opt: 42 Z-score: 89.2 bits: 17.8 E(): 4.8
Smith-Waterman score: 42; 38.462% identity (61.538% similar) in 13 aa overlap
(5-17:10-22)

CV127. 10 20 30
MMSCRYICLEMGLCMQITLWSIVICWRLG
: : : . . :
gi|391 SFREQCVPGREITYECLNACAEYAVRQ
10 20

>>gi|1063270 gi|1063270|dbj|BAA11251.1| gamma-gliadin pr (279 aa)
initn: 33 initl: 33 opt: 49 Z-score: 86.6 bits: 20.7 E(): 6.7
Smith-Waterman score: 49; 21.739% identity (65.217% similar) in 23 aa overlap
(4-25:137-159)

gi|106 NIQVDPGQVQWPQQPFPQPHQPFSSQPPQTFPQPQQTFPHQPPQFSSQPPQPPQFIQ
10 20 30 40 50 60

gi|106 PQQPFPQPQQTYPQRPPQPPQTQQPQQPFPQSQQPQQPFPQPQQQFPQPQQPQSFQ
70 80 90 100 110 120

CV127. 10 20 30
MMSCR-YICLEMGLCMQITLWSIVICWRLG
:. . . : . . . :
gi|106 QQPSLIQQSLQQQLNPCKNLLQCKPVSLSVSSLSMILPRSDCQVMRQQCCQQLAQIPQ
130 140 150 160 170 180

gi|106 QLQCAAHSIVHSIIMQQEQEQRQGVQILVPLSQQQQVGQGTLVQGGI IQPQQPAQLE
190 200 210 220 230 240

gi|106 VIRSSVLQTLATMCNVYVPPYCSTIRAPFASIVAGIGGQ
250 260 270

>>gi|1169665 gi|1169665|sp|P30438.2|FEL1A_FELCA Major al (92 aa)
initn: 23 initl: 23 opt: 44 Z-score: 85.6 bits: 18.9 E(): 7.7
Smith-Waterman score: 44; 30.769% identity (69.231% similar) in 13 aa overlap
(13-25:3-15)

10 20 30
CV127. MMSCRYICLECMGLCMQITLWSIVICWRLG
: : . . : : . .
gi|116 MKGACVLVLLWAALLLISGGNCEICPAVKRDVDLFTGTPDEYVEQVAQY
10 20 30 40 50

gi|116 KALPVVLENARILKNCVDAKMTEEDKENALSVDKIYTSPLC
60 70 80 90

>>gi|256429 gi|256429|gb|AAB23464.1| Kunitz trypsin inhi (216 aa)
initn: 29 initl: 29 opt: 47 Z-score: 85.2 bits: 20.1 E(): 8.1
Smith-Waterman score: 47; 33.333% identity (62.500% similar) in 24 aa overlap
(2-24:97-120)

gi|256 MKSTIFFLFLFCAFTTSYLP SAIADFLVDNEGNPLENGGTYIILSDITAFGGIRAAPTGN
10 20 30 40 50 60

10 20
CV127. MMSCRYICLECMGLCMQI-TLWSIV
.: . . . : : : : : : : :
gi|256 ERCPLTVVQSRNELDKGIGTIISPPYRIRFIAEGHPLSLKFDSFAVIMLCVGIPTEWSV
70 80 90 100 110 120

30
CV127. ICCWRLG
gi|256 EDLPEGPAVKIGENKDAMDGWFRLEVRSDDEFNNYKLVFCPQQAEDDKCGDIGISIDHDD
130 140 150 160 170 180

gi|256 GTRRLVVSKNKPLVVQFQKLDKESLAKKNHGLSRSE
190 200 210

>>gi|18770 gi|18770|emb|CAA45777.1| trypsin inhibitor su (217 aa)
initn: 29 initl: 29 opt: 47 Z-score: 85.2 bits: 20.1 E(): 8.1
Smith-Waterman score: 47; 33.333% identity (62.500% similar) in 24 aa overlap
(2-24:98-121)

gi|187 MKSTIFFALFLFCAFTTSYLP SAIADFLVDNEGNPLENGGTYIILSDITAFGGIRAAPTG
10 20 30 40 50 60

10 20
CV127. MMSCRYICLECMGLCMQI-TLWSI
.: . . . : : : : : : : :
gi|187 NERCPLTVVQSRNELDKGIGTIISPPYRIRFIAEGHPLSLKFDSFAVIMLCVGIPTEWSV
70 80 90 100 110 120

30
CV127. VICCWRLG
:
gi|187 VEDLPEGPAVKIGENKDAMDGWFRLEVRSDDEFNNYKLVFCPQQAEDDKCGDIGISIDHD
130 140 150 160 170 180

gi|187 DGTRRLVVSKNKPLVVQFQKLDKESLAKKNHGLSRSE
190 200 210

>>gi|1514943 gi|1514943|dbj|BAA08246.1| Chao1 [Chamaecyp (375 aa)
initn: 35 initl: 35 opt: 49 Z-score: 85.1 bits: 20.8 E(): 8.2
Smith-Waterman score: 49; 29.032% identity (51.613% similar) in 31 aa overlap
(1-29:1-30)

CV127. MMSCRYICLECMGLCMQITLWS--IVICCWRLG
: : : . . . : : . : : : :
gi|151 MASCTLLAVLVF-LCAIVSCFSDNPIDSCWRGDANWDQNRMKLADCAVGFSSAMGGKGG
10 20 30 40 50

gi|151 AFYTVTSSDDDPVNPAPGTLRYGATRERSLWIIIFSKNLNIKLNMPLYIAGNKTIDGRGAE
60 70 80 90 100 110

gi|151 VHIGNGGPGCLFMRTVSHVILHGLNIHGNTSVSGNVLISEASGVVPVHAQDGDAITMRNV
120 130 140 150 160 170

gi|151 TDVWIDHNSLSDSSDGLVDVTLASTGVTISNNHFFNHHKVMLLGHSDIYDDKSMKVTV
180 190 200 210 220 230

gi|151 FNQFGPNAGQRMPPARYGLIHVANNNDPWSIYAIGGSSNPITLSEGNSTAPNDSKKE
240 250 260 270 280 290

gi|151 VTRRVGCEPSTCANVWVRSTQDSFNNGAYFVSSGKNEGNTNIYNNNEAFKVENGSAAPQL
300 310 320 330 340 350

gi|151 TKNAGVLTCLSKPCS
360 370

>>gi|170736 gi|170736|gb|AAA34288.1| gamma-gliadin (251 aa)
initn: 31 initl: 31 opt: 47 Z-score: 84.4 bits: 20.1 E(): 8.9
Smith-Waterman score: 47; 22.727% identity (63.636% similar) in 22 aa overlap
(4-24:156-177)

gi|170 MKTLLILITILAMAITIGTANMQVDPSSQVQWPQQPVPQPHQPFSSQPPQTFPQPQQTFP
10 20 30 40 50 60

gi|170 HQPQQQFPQPQQPQQQFLQPQQPFPQQPQQPYPQQPQQPFPQTQQPQQLFQPSQQPQQQF
70 80 90 100 110 120

CV127. MMSCR-YICLECMGLCMQITLWSIVICC
: . . . : . . . : : : :
gi|170 SQPQQQFPQPQQPQQSFPPQQPFFIQPSLQQQVNPCKNLLQCKPVSLSLWSMIWPQ
130 140 150 160 170 180

30
CV127. WRLG

gi|170 SDCQVMRQQCCQQLAQIPQQLQCAAIHTIIHSIIMQQEQQEQQQGMHILLPLYQQQQVQ
190 200 210 220 230 240

gi|170 GTLVQGGIIQ
250

>>gi|50659891 gi|50659891|gb|AAT80665.1| lipid transfer (115 aa)
initn: 31 initl: 31 opt: 44 Z-score: 84.4 bits: 19.0 E(): 9

Smith-Waterman score: 44; 29.630% identity (51.852% similar) in 27 aa overlap
(1-26:1-27)

```

          10          20          30
CV127. MMSCRYICLE-CMGLCMQITLWSIVICCWRLG
      : : : : : : : : : : : :
gi|506 MASSAVINLALVVALCMAVSVAHAITCGQVSSNLVPCFDYVRSGGPVPPACCNGIRTING
          10          20          30          40          50          60

gi|506 LAKTTPDRQAACNCLKSLAGSVSGVNPNGNVESESLPGKCGVNPYPYKISTSTNCATVK
          70          80          90          100          110

```

>>gi|50659889 gi|50659889|gb|AAT80664.1| lipid transfer (115 aa)
initn: 31 init1: 31 opt: 44 Z-score: 84.4 bits: 19.0 E(): 9
Smith-Waterman score: 44; 29.630% identity (51.852% similar) in 27 aa overlap
(1-26:1-27)

```

          10          20          30
CV127. MMSCRYICLE-CMGLCMQITLWSIVICCWRLG
      : : : : : : : : : : : :
gi|506 MASSAVINLALVVALCMAVSVAHAITCGQVSSNLVPCFDYVRSDGPVPPACCNGIRTING
          10          20          30          40          50          60

gi|506 LAKTTPDRQAACNCLKSLAGSVSGVNPNGNAESLPGKCGVNPYPYKISTSTNCATVK
          70          80          90          100          110

```

>>gi|50659885 gi|50659885|gb|AAT80662.1| lipid transfer (115 aa)
initn: 31 init1: 31 opt: 44 Z-score: 84.4 bits: 19.0 E(): 9
Smith-Waterman score: 44; 29.630% identity (51.852% similar) in 27 aa overlap
(1-26:1-27)

```

          10          20          30
CV127. MMSCRYICLE-CMGLCMQITLWSIVICCWRLG
      : : : : : : : : : : : :
gi|506 MASSAVINLALVVALCMAVSVAHAITCGQVSSNLVPCFDYVRSGGPVPPACCNGIRTING
          10          20          30          40          50          60

gi|506 LAKTTPDRQAACNCLKSLAGSVSGVNPNGNAESLPGKCGVNPYPYKISTSTNCATVK
          70          80          90          100          110

```

>>gi|38492338 gi|38492338|gb|AAR22488.1| allergen Mal d (115 aa)
initn: 31 init1: 31 opt: 44 Z-score: 84.4 bits: 19.0 E(): 9
Smith-Waterman score: 44; 22.222% identity (51.852% similar) in 27 aa overlap
(2-26:1-27)

```

          10          20          30
CV127. MMSCRYICLECM--GLCMQITLWSIVICCWRLG
      : : : : : : : : : : : :
gi|384 MACSAVIKLALVVALCMAVSVAHAITCGQVTSSLAPCIGYVRSGGAVPPACCNGIRTIN
          10          20          30          40          50

gi|384 GLARTTADRQTACNCLKNLAGSISGVNPNNAAGLPGKCGVNPYPYKISTSTNCATVK
          60          70          80          90          100          110

```

>>gi|14285595 gi|14285595|sp|O04004.1|NLTP6_AMBAR Non-sp (118 aa)
initn: 33 init1: 33 opt: 44 Z-score: 84.2 bits: 19.0 E(): 9.1
Smith-Waterman score: 44; 36.842% identity (63.158% similar) in 19 aa overlap
(15-29:3-20)

```

          10          20          30
CV127.  MMSCRYICLECMGLCMQITLWSIVI----CCWRLG
          :.:.: :.:.:.: :.:.
gi|142          MDCIRI-LWSVAVGLLLVSWRPTMFAASPTCDTVQNILAPCAGFLTGQ
          10          20          30          40

gi|142  EPSKACCTGVNNLNNRKTKADRVAVCNCIKELTKSIAYDPKRMPLLSTKCGVKPDFPAV
          50          60          70          80          90          100

gi|142  DKNLDCSKLPV
          110
```

```

31 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:22:21 2008 done: Sun Dec 14 22:22:21 2008
Total Scan time: 0.040 Total Display time: 0.010
```

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.1962_c2

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.1962_c2, 39 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 2.9418 \pm 0.00331$; $\mu = 10.1224 \pm 0.172$
mean_var=35.5373 \pm 9.492, 0's: 21 Z-trim: 21 B-trim: 15 in 1/43
Lambda= 0.215145

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:		opt	bits	E(1313)	%_id	%_sim	alen
gi 127533	gi 127533 sp P02761.1 MUP (181)	49	20.9	4.6	0.333	0.667	24
gi 37725377	gi 37725377 gb AA033897 (52)	45	19.0	5	0.333	0.633	30
gi 46014884	gi 46014884 pdb 1H4B A (84)	44	19.0	8.3	0.333	0.633	30
gi 14423850	gi 14423850 sp Q39419.1 (85)	44	19.0	8.4	0.333	0.633	30
gi 59800144	gi 59800144 sp P69197.1 (79)	43	18.6	9.9	0.344	0.625	32
gi 59800143	gi 59800143 sp P69196.1 (79)	43	18.6	9.9	0.344	0.625	32

>>>CV127.1962_c2, 39 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|127533 gi|127533|sp|P02761.1|MUP_RAT Major urinary (181 aa)
initn: 48 init1: 48 opt: 49 Z-score: 89.6 bits: 20.9 E(): 4.6
Smith-Waterman score: 49; 33.333% identity (66.667% similar) in 24 aa overlap
(9-31:21-44)

	10	20	30	
CV127.	MTPERTNTEQTKLMGRIDIKIQGDR-SVSIAEAMETAKH			
	: . . : . . : . . : . .			
gi 127	MKLLLLLLCLGLTLVCGHAEASSTRGNLDVAKLNGDWFSIVVASNKREKIEENGSMRVF			
	10	20	30	40 50 60
gi 127	MQHIDVLENSLGFKFRIKENGECRELYLVAYKTPEDGEYFVEYDGGNTFTILKTDYDRYV			
	70	80	90	100 110 120
gi 127	MFHLINFKNGETFQLMVLYGRKDLSSDIKEKFAKLCEAHGITRDNIIIDLTKTDRCLQAR			
	130	140	150	160 170 180

gi|127 G

>>gi|37725377 gi|37725377|gb|AA033897.1| Ole e 3 allerge (52 aa)
initn: 36 init1: 36 opt: 45 Z-score: 89.0 bits: 19.0 E(): 5
Smith-Waterman score: 45; 33.333% identity (63.333% similar) in 30 aa overlap
(9-38:1-27)

	10	20	30	
CV127.	MTPERTNTEQTKLMGRIDIKIQGDRSVSIAEAMETAKH			
	: . . . : . . : . . : . .			
gi 377	EHERIFKRFD-----KGDGKISSSELGETLKLPLGSVTLEEIQRMMAEIDTDGD			
	10	20	30	40
gi 377	GFL			
	50			

>>gi|46014884 gi|46014884|pdb|1H4B|A Chain A, Solution S (84 aa)
initn: 32 initl: 32 opt: 44 Z-score: 85.0 bits: 19.0 E(): 8.3
Smith-Waterman score: 44; 33.333% identity (63.333% similar) in 30 aa overlap
(9-38:10-36)

```

              10      20      30
CV127.  MTPERTNTEQTKLMGRIDI I KIQGDRSVSIAEAMETAKH
              . . . . . : : : : : : : : :
gi|460  XDDHPQDKAERERIFKRFDA---NGDGKISAAELGEALKTLGSITPDEVKHMMAEIDTDG
              10      20      30      40      50

gi|460  DGFISFQEFTDFGRANRGLLKDVAKIF
              60      70      80

```

>>gi|14423850 gi|14423850|sp|Q39419.1|POLC4_BETVE Polcal (85 aa)
initn: 32 initl: 32 opt: 44 Z-score: 84.9 bits: 19.0 E(): 8.4
Smith-Waterman score: 44; 33.333% identity (63.333% similar) in 30 aa overlap
(9-38:11-37)

```

              10      20      30
CV127.  MTPERTNTEQTKLMGRIDI I KIQGDRSVSIAEAMETAKH
              . . . . . : : : : : : : : :
gi|144  MADDHPQDKAERERIFKRFDA---NGDGKISAAELGEALKTLGSITPDEVKHMMAEIDTDG
              10      20      30      40      50

gi|144  GDGFISFQEFTDFGRANRGLLKDVAKIF
              60      70      80

```

>>gi|59800144 gi|59800144|sp|P69197.1|POLC1_BRARA Polcal (79 aa)
initn: 38 initl: 38 opt: 43 Z-score: 83.6 bits: 18.6 E(): 9.9
Smith-Waterman score: 51; 34.375% identity (62.500% similar) in 32 aa overlap
(6-37:37-65)

```

              10      20
CV127.  MTPERTNTEQTKLMGRIDI I KIQGDRSVS
              : . . . . . : : : :
gi|598  MADAHERIFKKFDTGDGKISAAEL E EALKKLGSVTPDDVTRMMAKID--TDGDGNIS
              10      20      30      40      50

              30
CV127.  IAEAMETAKH
              . : : :
gi|598  FQEFTEFASANPGLMKDVAKVF
              60      70

```

>>gi|59800143 gi|59800143|sp|P69196.1|POLC1_BRANA Polcal (79 aa)
initn: 38 initl: 38 opt: 43 Z-score: 83.6 bits: 18.6 E(): 9.9
Smith-Waterman score: 51; 34.375% identity (62.500% similar) in 32 aa overlap
(6-37:37-65)

```

              10      20
CV127.  MTPERTNTEQTKLMGRIDI I KIQGDRSVS
              : . . . . . : : : :
gi|598  MADAHERIFKKFDTGDGKISAAEL E EALKKLGSVTPDDVTRMMAKID--TDGDGNIS
              10      20      30      40      50

              30
CV127.  IAEAMETAKH
              . : : :
gi|598  FQEFTEFASANPGLMKDVAKVF
              60      70

```

39 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:04:09 2008 done: Sun Dec 14 22:04:09 2008
Total Scan time: 0.080 Total Display time: 0.000

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.2197_d1

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.2197_d1, 45 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 2.6069 \pm 0.00316$; $\mu = 13.7020 \pm 0.165$
mean_var=23.5410 \pm 5.414, 0's: 17 Z-trim: 17 B-trim: 9 in 1/42
Lambda= 0.264339

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2

join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:

	opt	bits	E(1313)	%_id	%_sim	alen
gi 17939981 gi 17939981 gb AAL49391 (98)	48	21.5	2	0.379	0.621	29
gi 2735118 gi 2735118 gb AAB93839.1 (134)	43	19.7	9.3	0.600	0.900	10

>>>CV127.2197_d1, 45 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|17939981 gi|17939981|gb|AAL49391.1| cystatin [Felis (98 aa)

initn: 46 initl: 46 opt: 48 Z-score: 96.1 bits: 21.5 E(): 2

Smith-Waterman score: 48; 37.931% identity (62.069% similar) in 29 aa overlap
(8-36:28-55)

	10	20	30	40
CV127.	MLRGVGGEERTKRHYQPFEEALSFVIVFVAKFSILIPFSL			
	:::.. :: :: :. .	:: :. :		
gi 179	MIPGGLSEAKPATPEIQEIANEVKPKLEEKTNETYQKFE-AIEYKTQVVAGINYYIKVQV			
	10 20 30 40 50			

CV127. ISCSL

gi 179	DDNRYIHIKVFKGLPVQDSSLTLTGYQTGKSEDELDTGF
	60 70 80 90

>>gi|2735118 gi|2735118|gb|AAB93839.1| ABA-1 allergen [A (134 aa)

initn: 32 initl: 32 opt: 43 Z-score: 84.1 bits: 19.7 E(): 9.3

Smith-Waterman score: 43; 60.000% identity (90.000% similar) in 10 aa overlap
(1-10:65-74)

gi 273	HTMEHYLKTLYLSWLTEEQKEKLKEMKEAGKTKAEIQHEVMHFYDQLHGEEKQQATEKLKV
	10 20 30 40 50 60

	10	20	30	40
CV127.	MLRGVGGEERTKRHYQPFEEALSFVIVFVAKFSILIPFSLISCSL			
	.::: :::.			
gi 273	GCKMLLKGVIKEKVVELRNMEAGADIQELQKQVEKMLSEVTDEKQKEKVHEYGPCKK			
	70 80 90 100 110 120			

gi 273	IFGATTLQHRRRR
	130



The Chemical Company
Plant Science L. L. C.

BASF Reg. Doc. No. 2008/7019371

45 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:08:06 2008 done: Sun Dec 14 22:08:06 2008
Total Scan time: 0.070 Total Display time: 0.000

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.3982_c1

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.3982_c1, 72 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 4.7725 \pm 0.00324$; $\mu = 2.7904 \pm 0.168$

mean_var=36.5555 \pm 8.725, 0's: 10 Z-trim: 10 B-trim: 0 in 0/43

Lambda= 0.212128

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2

join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:		opt	bits	E(1313)	%_id	%_sim	alen
gi 113475	gi 113475 sp P27759.1 MPA (396)	60	24.4	1.7	0.500	0.812	16
gi 62738637	gi 62738637 pdb 1YG9 A (330)	54	22.6	5	0.389	0.667	18
gi 1703445	gi 1703445 sp P54958.1 A (352)	54	22.6	5.3	0.389	0.667	18
gi 807698	gi 807698 dbj BAA06905.1 (731)	57	23.5	5.9	0.345	0.759	29
gi 1304216	gi 1304216 dbj BAA07772. (111)	48	20.8	5.9	0.306	0.556	36
gi 1359600	gi 1359600 emb CAA64868. (316)	52	22.0	7.3	0.312	0.625	32
gi 549188	gi 549188 sp P35780.1 VA5 (205)	49	21.0	8.9	0.533	0.733	15
gi 549187	gi 549187 sp P35759.1 VA5 (205)	49	21.0	8.9	0.533	0.733	15
gi 160780	gi 160780 gb AAA29793.1 (209)	49	21.0	9.1	0.533	0.733	15
gi 51093375	gi 51093375 gb AAT95009 (226)	49	21.0	9.9	0.533	0.733	15

>>>CV127.3982_c1, 72 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|113475 gi|113475|sp|P27759.1|MPA11_AMBAR Pollen all (396 aa)

initn: 60 initl: 60 opt: 60 Z-score: 97.4 bits: 24.4 E(): 1.7

Smith-Waterman score: 60; 50.000% identity (81.250% similar) in 16 aa overlap (17-32:298-313)

gi|113 MGIKHCCYILYFTLALVTLLQPVRSAEDLQEILPVNETRRLTTSGAYNIIDGCWRGKADW
10 20 30 40 50 60

gi|113 AENRKALADCAQGFQKGTGKGKGDIIYTVTSELDDVDANPKEGTLRFQAAQNRPLWIIFE
70 80 90 100 110 120

gi|113 RDMVIRLDKEMVNSDKTIDGRGAKVEIINAGFTLNGVKNVIIHNINMHDVKVNPGLIK
130 140 150 160 170 180

gi|113 SNDGPAAPRAGSDGDAISISGSSQIWDHCSLSKSDGLVDAKLGTTRLTVSNSLFTQHQ
190 200 210 220 230 240

CV127. 10
MAFPSVNSSRTLIAYCGGI
::
gi|113 FVLLFGAGDENIEDRGMLATVAFNTFTDNDQRMPCRHRGFFQVNNNYDKWGSYAIGGS
250 260 270 280 290 300

20 30 40 50 60 70
CV127. ASPNVKLKNGNFCFCTFNSFLQTPKSSLSARFSRTLFIPCKASLTSPHTDT
:::..:.. ::
gi|113 ASPTILSQGNRFCAPDERSKKNVLGRHGEAAAESMKWNWRTNKDVLENGAIFVASGVDPV
310 320 330 340 350 360

gi|113 LTPEQSAGMIPAEPEGESALSLTSSAGVLSQPGAPC
370 380 390

>>gi|62738637 gi|62738637|pdb|1YG9|A Chain A, The Struct (330 aa)
initn: 45 initl: 45 opt: 54 Z-score: 88.9 bits: 22.6 E(): 5
Smith-Waterman score: 54; 38.889% identity (66.667% similar) in 18 aa overlap
(19-36:276-293)

gi|627 GASIVPLYKLHVHVFINTQYAGITKIGNQNFLTTFDSTSCNVVVASQECVGGACVCPNLQK
10 20 30 40 50 60

gi|627 YEKLKPKYISDGNVQVKFFDTGSAVGRGIEDSLTISQLTTSQQDIVLADELSQEVCLISA
70 80 90 100 110 120

gi|627 DVVVGIAAPGCPNALKGKTVLENFVEENLIAPVFSIHARFQDGEHFGEIIFGGSDWKYV
130 140 150 160 170 180

gi|627 DGEFTYVPLVGDDSWKFRLDGVKIGDTTVAPAGTQAIIDTSKAIIVGPKAYVNPINEAIG
190 200 210 220 230 240

CV127. 10 20 30 40
MAFPSVNSSRTLIAYCGGIASPNVLKLNCFCTFNSFLQTP
:.: .. :.:.: :

gi|627 CVVEKTTTTRICKLDCSKIPSLPDVTFVINGRNFNISSQYIQQNGNLCYSGFQPGHSD
250 260 270 280 290 300

50 60 70
CV127. KSSLSSARFSRTLFIPCKASLTSPHTDT

gi|627 HFFIGDFFVDHYYSEFNWENKTMGFGRSVE
310 320 330

>>gi|1703445 gi|1703445|sp|P54958.1|ASP2_BLAG Aspartic (352 aa)
initn: 45 initl: 45 opt: 54 Z-score: 88.4 bits: 22.6 E(): 5.3
Smith-Waterman score: 54; 38.889% identity (66.667% similar) in 18 aa overlap
(19-36:296-313)

gi|170 MIGLKLVTVLFAVATITHAAELQRVPLYKLHVHVFINTQYAGITKIGNQNFLTTFDSTSCN
10 20 30 40 50 60

gi|170 VVVASQECVGGACVCPNLQKYEKLKPKYISDGNVQVKFFDTGSAVGRGIEDSLTISNLTT
70 80 90 100 110 120

gi|170 SQQDIVLADELSQEVCLISADVVGIAAPGCPNALKGKTVLENFVEENLIAPVFSIHAR
130 140 150 160 170 180

gi|170 FQDGEHFGEIIFGGSDWKYVDGEFTYVPLVGDDSWKFRLDGVKIGDTTVAPAGTQAIIDT
190 200 210 220 230 240

CV127. 10 20
MAFPSVNSSRTLIAYCGGIASPN
:.:

gi|170 SKAIIVGPKAYVNPINEAIGCVVEKTTTTRICKLDCSKIPSLPDVTFVINGRNFNISSQY
250 260 270 280 290 300

30 40 50 60 70
CV127. VLKLNCFCTFNSFLQTPKSSLSSARFSRTLFIPCKASLTSPHTDT
.. :.:.: :

gi|170 YIQQNGNLCYSGFQPCGHSDFHFFIGDFFVDHYYSEFNWENKTMGFGRSVESV
310 320 330 340 350

>>gi|807698 gi|807698|dbj|BAA06905.1| pre-pro-cucumisin (731 aa)
initn: 44 initl: 44 opt: 57 Z-score: 87.6 bits: 23.5 E(): 5.9
Smith-Waterman score: 57; 34.483% identity (75.862% similar) in 29 aa overlap
(4-29:660-688)

gi|807 MSSSLIFKLFSSLFSSNRLASRLDSDDDGKNYIVYMGRKLEDPDSAHLHHRAMLEQVV
10 20 30 40 50 60

gi|807 GSTFAPESVLHTYKRSFNGFAVKLTETEEAEKIASMEGVVSVFLNEMNELHTTRSWDFLGF
70 80 90 100 110 120

gi|807 PLTVPRRSQVESNIVVGVLDTGIWPESPSFDDEGFSPPPPKWKGTCTSNFRNCRKIIG
130 140 150 160 170 180

gi|807 ARSYHIGRPISPGDVNGPRDTNGHGTHTASTAAGGLVSQANLYGLGLGTARGGVPLARIA
190 200 210 220 230 240

gi|807 AYKVCWNDGCSDDTDLAAYDDAIADGVDIISLSVGGANPRHYFVDAIAIGSFHAVERGIL
250 260 270 280 290 300

gi|807 TSNSAGNGGPNFFTTASLSPWLLSVAASTMDRKFTVQVQIGNGQSFQGVINTFDNQYYP
310 320 330 340 350 360

gi|807 LVSGRDIPNTGFDKSTSRFCTDKSVNPNLLKGKIVVCEASFGPHEFFKSLDGAAGVLMTS
370 380 390 400 410 420

gi|807 NTRDYADSYPLPSSVLDPNDDLATLRYIYSIRSPGATIFKSTTILNASAPVVVSFSSRGP
430 440 450 460 470 480

gi|807 NRATKDVIKPDISGPGVEILAAWPSVAPVGGIRRNTLFNIISGTSMSCPHITGIATYVKT
490 500 510 520 530 540

gi|807 YNPTWSPAAIKSALMTTASPMNARFNPQAEFAYGSGHVNPLKAVRPGLVYDANESDYVKF
550 560 570 580 590 600

CV127. MAFP
:

gi|807 LCGQGYNTQAVRRITGDYSACTSGNTGRVWDLNYPSPFGLSVSPSQTFNQYFNRTLTSVAP
610 620 630 640 650 660

10 20 30 40 50 60

CV127. SVNSSRTLIAYCGGIA---SPNVKLNGNFCFCTFNSFLQTPKSSLSSSARFSRTLFIPI

.... :... :... :... :... :

gi|807 QASTYRAMISAPQGLTISVNPVLSFNGLGDRKSFTLTVRGSIKGFVVSASLVWSDGVHY
670 680 690 700 710 720

70
CV127. KASLTSPHTDT

gi|807 VRSPITITSLV
730

>>gi|1304216 gi|1304216|dbj|BAA07772.1| allergenic prote (111 aa)
initn: 37 initl: 37 opt: 48 Z-score: 87.6 bits: 20.8 E(): 5.9
Smith-Waterman score: 48; 30.556% identity (55.556% similar) in 36 aa overlap
(10-45:80-109)

gi|130 RHEVKRQCVATTHPAAPGTEQLRQDCCRQLAAVDDSWCRCALNHMVGGIYRELGATDVG
10 20 30 40 50 60

```

              10      20      30      40      50
CV127.      MAFPSVNSSRTLIAYCGGIASPNVLKLNNGFCFCTFNSFLQTPKSSLSS
              :: ::  ::  ::  : . . . : ::
gi|130 HPMAEVFPGCRGRDLERRRRSLPAFC-----NVDIPNGTGGVCYWLGYPRTPRTGH
              70      80      90      100      110

              60      70
CV127. ARFSRTLFIIPCKASLTSPHTDT

>>gi|1359600 gi|1359600|emb|CAA64868.1| chitinase Ib [Ca (316 aa)
  initn: 31 initl: 31 opt: 52 Z-score: 86.0 bits: 22.0 E(): 7.3
Smith-Waterman score: 52; 31.250% identity (62.500% similar) in 32 aa overlap
(19-50:210-241)

gi|135 MKLFSLLLFLAFLLGTSAEQCGRQAGGAACANNLCCSQFGWCGNTAEYCGAGCQSQCSSP
              10      20      30      40      50      60

gi|135 TTTTSSPTASSGGGGDVGSLISASLFDQMLKYRNDPRCKSNGFYTYNAFIAAARSFNFG
              70      80      90      100      110      120

gi|135 TTGDVTTRKRELAFLAQTSHETTGGWATAPDGPYAWGYCFVMENNKQTYCTSKSWPCVF
              130      140      150      160      170      180

              10      20      30      40
CV127.      MAFPSVNSSRTLIAYCGGIASPNVLKLNNGFCFCTFNSFLQTPKSSLSS
              : : : : : : : : : : : : : : : : : :
gi|135 GKQYYGRGPIQLTHNINYGGAGKAIGADLINNPDLVATNPTISFKTAIWFWMTQANKPS
              190      200      210      220      230      240

              50      60      70
CV127. SARFSRTLFIIPCKASLTSPHTDT
              :
gi|135 SHDVIIGNWRPSAADTSAGRVPSYGVITNIINGGLECGHGSDDRANRIGFYKRYCDTLG
              250      260      270      280      290      300

gi|135 VSYGNNLDCYNQKPFA
              310

>>gi|549188 gi|549188|sp|P35780.1|VA5_POLFU Venom allerg (205 aa)
  initn: 33 initl: 33 opt: 49 Z-score: 84.4 bits: 21.0 E(): 8.9
Smith-Waterman score: 49; 53.333% identity (73.333% similar) in 15 aa overlap
(13-27:116-130)

gi|549 VDYCKIKCSSGIHTVCQYGESTKPSKNCADKVIKSVGPTEEEKKLIVNEHNRFRQKVAQG
              10      20      30      40      50      60

              10
CV127.      MAFPSVNSSRTLIAYCG
              : : : :
gi|549 LETRGNPGPQPAASDMNNLVWDELAHIAQVWASQCQILVHDKCRNTAKYQVGQNIAYAG
              70      80      90      100      110      120

              20      30      40      50      60      70
CV127. GIASPNVLKLNNGFCFCTFNSFLQTPKSSLSSSARFSRTLFIIPCKASLTSPHTDT
              : : : : :
gi|549 GSKLPDVVSLIKLWENEVKDFNYNKGITKQNFQGVGHYTQMIWAKTKEIGCGSLKYMKN
              130      140      150      160      170      180

gi|549 MQHHYLICNYGPAGNYLGQLPYTKK
              190      200

```

>>gi|549187 gi|549187|sp|P35759.1|VA5_POLEX Venom allerg (205 aa)
initn: 33 init1: 33 opt: 49 Z-score: 84.4 bits: 21.0 E(): 8.9
Smith-Waterman score: 49; 53.333% identity (73.333% similar) in 15 aa overlap
(13-27:116-130)

gi|549 VDYCKIKCPSGIHTVCQYGESTKPSKNCAGKVIKSVGPTTEEEKKLIVSEHNRFQKVAQG
10 20 30 40 50 60

CV127. MAFPSVNSSRTLIAYCG
10
:::

gi|549 LETRGNPGPQPAASDMNDLVWNDELAHIAQVWASQCQFLVHDKCRNTAKYPVGQNIAYAG
70 80 90 100 110 120

CV127. GIASPNVLKLNNGNFCFCTFNSFLQTPKSSLSARFSRTLFIPCKASLTSPHTDT
20 30 40 50 60 70
: :.:.:

gi|549 GSKLPDVVSLIKLWENEVKDFNYNTGITKQNFQKIGHYTMVWGKTKEIGCGSLKYIENK
130 140 150 160 170 180

gi|549 MQNHYLICNYGPAGNYLGQLPYTKK
190 200

>>gi|160780 gi|160780|gb|AAA29793.1| allergen 5 (209 aa)
initn: 33 init1: 33 opt: 49 Z-score: 84.3 bits: 21.0 E(): 9.1
Smith-Waterman score: 49; 53.333% identity (73.333% similar) in 15 aa overlap
(13-27:120-134)

gi|160 SSQGVVDYCKIKCPSGIHTVCQYGESTKPSKNCAGKVIKSVGPTTEEEKKLIVSEHNRFQK
10 20 30 40 50 60

CV127. MAFPSVNSSRTLI
10
:

gi|160 VAQGLETRGNPGPQPAASDMNDLVWNDELAHIAQVWASQCQFLVHDKCRNTAKYPVGQNI
70 80 90 100 110 120

CV127. AYCGGIASPNVLKLNNGNFCFCTFNSFLQTPKSSLSARFSRTLFIPCKASLTSPHTDT
20 30 40 50 60 70
:: :.:.:

gi|160 AYAGGSNLPDVVSLIKLWENEVKDFNYNTGITKQNFQKIGHYTMVWGKTKEIGCGSLKY
130 140 150 160 170 180

gi|160 MENNMQNHYLICNYGPAGNYLGQLPYTKK
190 200

>>gi|51093375 gi|51093375|gb|AAT95009.1| allergen Pol e (226 aa)
initn: 33 init1: 33 opt: 49 Z-score: 83.6 bits: 21.0 E(): 9.9
Smith-Waterman score: 49; 53.333% identity (73.333% similar) in 15 aa overlap
(13-27:137-151)

gi|510 MEIGGLVYLVIIVVAIIHSSQGVVDYCKIRCPSGIHTVCQYGESTKPSKNCAGKVIKSVGPT
10 20 30 40 50 60

gi|510 EEEKKLIVSEHNRFQKVAQGLETRGNPGPQPAASDMNDLVWNDELAHIAQVWASQCQFL
70 80 90 100 110 120

CV127. MAFPSVNSSRTLIAYCGGIASPNVLKLNNGNFCFCTFNSFLQTPKSSLSARFSRT
10 20 30 40 50
::: :.:.:

gi|510 VHDKCRNTAKYPVGQNIAYAGGSKLPDVVSLIKLWENEVKDFNYNTGITKQNFQKIGHYT
130 140 150 160 170 180



The Chemical Company
Plant Science L. L. C.

BASF Reg. Doc. No. 2008/7019371

60 70
CV127. LFIPCKASLTSPHTDT

gi|510 QMVWGKTKEIGCGSLKYMENKMQNHLYLICNYGPAGNYLGQLPNTKK
190 200 210 220

72 residues in 1 query sequences

292011 residues in 1313 library sequences

Scomplib [34.26]

start: Sun Dec 14 22:29:51 2008 done: Sun Dec 14 22:29:51 2008

Total Scan time: 0.100 Total Display time: 0.020

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.4164_c2

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.4164_c2, 38 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 4.3779 \pm 0.00334$; $\mu = 0.9766 \pm 0.175$

mean_var=34.6864 \pm 9.332, 0's: 14 Z-trim: 14 B-trim: 104 in 2/41

Lambda= 0.217768

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2

join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:

gi	opt	bits	E(1313)	%_id	%_sim	alen
gi 999009 gi 999009 gb AAB34785.1 (117)	50	21.8	1.6	0.296	0.630	27
gi 99031759 gi 99031759 pdb 2A0A A (131)	47	20.9	3.4	0.375	0.625	24
gi 8118421 gi 8118421 gb AAF72983.1 (267)	49	21.5	4.4	0.290	0.516	31
gi 8118439 gi 8118439 gb AAF72991.1 (267)	49	21.5	4.4	0.275	0.550	40
gi 109913547 gi 109913547 sp Q40638 (267)	49	21.5	4.4	0.290	0.516	31
gi 28630919 gi 28630919 gb AAO45607 (269)	48	21.2	5.5	0.282	0.538	39
gi 28630923 gi 28630923 gb AAO45608 (269)	48	21.2	5.5	0.282	0.538	39
gi 14193761 gi 14193761 gb AAK56124 (269)	48	21.2	5.5	0.282	0.538	39
gi 111120428 gi 111120428 gb ABH063 (129)	44	19.9	6.5	0.250	0.625	32
gi 83300389 gi 83300389 sp O42799.2 (270)	47	20.9	6.8	0.333	0.708	24
gi 66849793 gi 66849793 gb EAL90121 (270)	47	20.9	6.8	0.333	0.708	24
gi 75139991 gi 75139991 sp Q7M1Y1 Q (35)	38	18.0	6.9	0.304	0.522	23
gi 111120420 gi 111120420 gb ABH063 (129)	43	19.6	8	0.250	0.594	32
gi 111120432 gi 111120432 gb ABH063 (129)	43	19.6	8	0.250	0.594	32
gi 111120424 gi 111120424 gb ABH063 (129)	43	19.6	8	0.250	0.594	32
gi 111494253 gi 111494253 gb ABH063 (129)	43	19.6	8	0.250	0.594	32
gi 2507248 gi 2507248 sp P49275.2 D (259)	46	20.6	8.1	0.412	0.706	17
gi 407609 gi 407609 gb AAB27812.1 (37)	37	17.6	9	0.310	0.517	29

>>>CV127.4164_c2, 38 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|999009 gi|999009|gb|AAB34785.1| 68 kDa allergen [Pe (117 aa)

initn: 23 initl: 23 opt: 50 Z-score: 97.8 bits: 21.8 E(): 1.6

Smith-Waterman score: 50; 29.630% identity (62.963% similar) in 27 aa overlap (1-27:90-116)

gi|999 GPYKWTQRIYDYDFLTNLTSSSEANDIIGAEAPLWSEQVDDVTVSSVFWPRAAALGELVWS
10 20 30 40 50 60

CV127. MIATALVSCNCRTPPTFDVARHQYLLWLFH
10 20 30
..... : : . : . : .
gi|999 GNRDAAGRKRRTTSFTQRIILNFREYLVANGVMAAALVPKYCLQHPHACDLYKNQTVMS
70 80 90 100 110

CV127. QSTHQGP

>>gi|99031759 gi|99031759|pdb|2A0A|A Chain A, Solution S (131 aa)

initn: 45 initl: 45 opt: 47 Z-score: 91.9 bits: 20.9 E(): 3.4

Smith-Waterman score: 47; 37.500% identity (62.500% similar) in 24 aa overlap
(6-27:28-51)

```

                                10      20      30
CV127.      MIATALVSCNCRTAPPTFDVA--RHQYLLWLFHQSTHQ
              ... ..: ::::: :...
gi|990 MASIEGKYKLEKSEKFDEFLDKLGVGFVMVKTAAKTLKPTFEVAIENDQYIFRSLSTFKNT
              10      20      30      40      50      60

```

CV127. GP

```

gi|990 EAKFKLGEEFEEDRADGKRVKTVIQEGDNKQVQTQFGDKEVKIIREFNGDEVVVTASCD
              70      80      90      100     110     120

gi|990 GVTSVRTYKRI
              130

```

>>gi|8118421 gi|8118421|gb|AAF72983.1| beta-expansin [Or (267 aa)
initn: 42 init1: 42 opt: 49 Z-score: 90.0 bits: 21.5 E(): 4.4
Smith-Waterman score: 49; 29.032% identity (51.613% similar) in 31 aa overlap
(7-37:20-50)

```

                                10      20      30
CV127.      MIATALVSCNCRTAPPTFDVARHQYLLWLFHQSTHQGP
              ::: ..: ... :: ..: :
gi|811 MASSSLLLACVVVAAMVSAVSCGPPKVPVPGPNITTSYGDKWLEAKATWYGAPKGAGPKDN
              10      20      30      40      50      60

gi|811 GGACGYKDVDKAPFLGMNSCGNDPIFKDGGKGCSCFEIKCSKPEACSDKPALIHVTDMND
              70      80      90      100     110     120

gi|811 EPIAAYHFDLSGLAFGAMAKDGKDEELRKAGIIDTQFRRVKCKYPADTKITFHIEKASNP
              130      140      150      160      170      180

gi|811 NYLALLVKYVAGDGDVVEVEIKEKGSEEWKALKESWGAIWRIDTPKPLKGPFSVRVTTEG
              190      200      210      220      230      240

gi|811 GEKIIAEDAIPDGWKADSVYKSNVQAK
              250      260

```

>>gi|8118439 gi|8118439|gb|AAF72991.1|AF261278_1 beta-ex (267 aa)
initn: 42 init1: 42 opt: 49 Z-score: 90.0 bits: 21.5 E(): 4.4
Smith-Waterman score: 49; 27.500% identity (55.000% similar) in 40 aa overlap
(1-37:11-50)

```

                                10      20      30
CV127.      MIATALVS---CNCRTAPPTFDVARHQYLLWLFHQSTHQGP
              .....: .: ... :: ..: :
gi|811 MASSCLLLACVVAAAMVSAVSCGPPKVPVPGPNITAAYGKQWLEARGTWYGKPKGAGPDDN
              10      20      30      40      50      60

gi|811 GGACGYKDIDKAPFLGMNSCGNDPIFKDGGKGCSCFEVKCSKPEACSDKPVIIHITDMNT
              70      80      90      100     110     120

gi|811 EPIAAYHFDLSGHAFGAMAKEGKDEELRKAGIIDMQFRRVRCKYPGETKVTFHVEKGSNP
              130      140      150      160      170      180

gi|811 NYFAVLVKYVGGDGDVVKVELKEKGSEEWKPLNESWGAIWRIDTPKPLKGPFSLRVTTES
              190      200      210      220      230      240

gi|811 DQQLVANDVIPDNWKANALYKSEIQVD
              250      260

```


>>gi|109913547 gi|109913547|sp|Q40638.2|EXPB1_ORYSJ Expa (267 aa)
initn: 42 initl: 42 opt: 49 Z-score: 90.0 bits: 21.5 E(): 4.4
Smith-Waterman score: 49; 29.032% identity (51.613% similar) in 31 aa overlap
(7-37:20-50)

```

              10      20      30
CV127.      MIATALVSCNCRTPPTFDVARHQYLLWLFHQSTHQGP
              ::: . .: . . . . . : : . .: :
gi|109 MASSSLLLACVVVAAMVSAVSCGPPKVPVPGPNITTSYGDKWLEAKATWYGAPKGAGPKDN
              10      20      30      40      50      60
gi|109 GGACGYKDVDKAPFLGMNSCGNDPIFKDGGKCGSCFEIKCSKPEACSDKPALIHVTDMND
              70      80      90      100     110     120
gi|109 EPIAAYHFDLSGLAFGAMAKDGKDEELRKAGIIDTQFRRVKCKYPADTKITFHIEKASNP
              130     140     150     160     170     180
gi|109 NYLALLVKYVAGDGDVVEVEIKEKGSEEWKALKESWGAIWRIDTPKPLKGPFSVRVTTEG
              190     200     210     220     230     240
gi|109 GEKIIAEDAIPDGWKADSVYKSNVQAK
              250     260

```

>>gi|28630919 gi|28630919|gb|AAO45607.1| beta-expansin 9 (269 aa)
initn: 47 initl: 47 opt: 48 Z-score: 88.3 bits: 21.2 E(): 5.5
Smith-Waterman score: 48; 28.205% identity (53.846% similar) in 39 aa overlap
(2-37:14-52)

```

              10      20      30
CV127.      MIATALV---SCNCRTPPTFDVARHQYLLWLFHQSTHQGP
              . .::: .: . .: . . . . . : : . .: :
gi|286 MGSLANNIMVVGAVLAALVVGSGCGPPKVPVPGPNITTNNGKWLRTARATWYGQPNGAGAP
              10      20      30      40      50      60
gi|286 DNGGACGKIKNVNLPYPYSGMTACGNVPIFKDGGKCGSCYEVRCCKEKEPCSGNPVTVFITDM
              70      80      90      100     110     120
gi|286 NYEPIAPYHFDLSGKAFGSLAKPGLNDKLRHCGIMDVEFRRVRCKYPAGQKIVFHIEKGC
              130     140     150     160     170     180
gi|286 NPNYVAVLVKFVADDGDIVLMEIQDKLSAEWKPMKLSWGAIWRMDTAKALKGPFSIRLTS
              190     200     210     220     230     240
gi|286 ESGKKVIAKDIIPANWRPDVYTSNVQFY
              250     260

```

>>gi|28630923 gi|28630923|gb|AAO45608.1| beta-expansin 1 (269 aa)
initn: 47 initl: 47 opt: 48 Z-score: 88.3 bits: 21.2 E(): 5.5
Smith-Waterman score: 48; 28.205% identity (53.846% similar) in 39 aa overlap
(2-37:14-52)

```

              10      20      30
CV127.      MIATALV---SCNCRTPPTFDVARHQYLLWLFHQSTHQGP
              . .::: .: . .: . . . . . : : . .: :
gi|286 MGSLANNIMVVGAVLAALVAGSGCGPPKVPVPGPNITTNNGKWLRTARATWYGQPNGAGAP
              10      20      30      40      50      60
gi|286 DNGGACGKIKNVNLPYPYSGMTACGNVPIFKDGGKCGSCYEVRCCKEKEPCSGNPVTVYITDM
              70      80      90      100     110     120
gi|286 NYEPIAPYHFDLSGKAFGSLAKPGLNDKIRHCGIMDVEFRRVRCKYPAGQKIVFHIEKGC
              130     140     150     160     170     180

```

gi|286 NPNYLAVLVKYVADDGDIVLMEIQDKLSAEWKPKLSWGAIWRMDTAKALKGPFSSIRLTS
190 200 210 220 230 240gi|286 ESGKKVIAKDVIPANWRPDVYTSNVQFY
250 260

```
>>gi|14193761 gi|14193761|gb|AAK56124.1|AF332174_1 beta- (269 aa)
  initn: 47 initl: 47 opt: 48 Z-score: 88.3 bits: 21.2 E(): 5.5
Smith-Waterman score: 48; 28.205% identity (53.846% similar) in 39 aa overlap
(2-37:14-52)
```

CV127. MIATALV---SCNCRTPAPTFDVARHQYLLWLFLHQSSTHGQP
gi|141 MGSLVNINIMVVGAVLAALVAGGSCGPVKVPNGPNIITNYNGKWLRTATWYGQPNGAGAP

gi|141 DNGGACGIKVNLPYPYSGMTACGNVPIFKDGGKCGSCYEVRCKEKEPCSGNPVTYITDM
70 80 90 100 110 120gi|141 NYEPIAPYHFDLSGKAFGSLAKPGLNDKIRHCGIMDVEFRRVRCKYPAGQKIVFHIEKGC
130 140 150 160 170 180gi|141 NPNYLAVLVKYVADDGDIVLMEIQDKLSAEWKPKLSWGAIWRMDTAKALKGPFSIRLTS
190 200 210 220 230 240gi|141 ESGKKVIAKDVIPANWRPDVYTSNVQFY
250 260

```
>>gi|111120428 gi|111120428|gb|ABH06348.1| Blo t 21 alle (129 aa)
  initn: 30 initl: 30 opt: 44 Z-score: 86.9 bits: 19.9 E(): 6.5
Smith-Waterman score: 44; 25.000% identity (62.500% similar) in 32 aa overlap
(5-36:9-39)
```

CV127. MIATALVSCNCRTAPPTFDVARHQYLLWLFLHQQSTHQGP
 ::.. : . : : ::.. .. :..
gi|111 MKFIIALAALIAVAC-ALPVSNDNFRHFEDHMIINTATQRFHEIEKFLLHITHEVDLEK
 10 20 30 40 50

gi|111 TGNKDEKARLLRELTVSEAFIEGSRGYFQRELKRTDLLLEKFNFEAALATGDL LKDLK
60 70 80 90 100 110gi|111 ALQKRVQDSE
120

```
>>gi|83300389 gi|83300389|sp|O42799.2|ALL7_ASPFU Allerge (270 aa)
  initn: 30 initl: 30 opt: 47 Z-score: 86.5 bits: 20.9 E(): 6.8
Smith-Waterman score: 47; 33.333% identity (70.833% similar) in 24 aa overlap
(3-25:230-251)
```

gi|833 MAPIFKSLALVSALFAAIISSAAPVNLDKREVDVWTTVTTVVWTTIDVTTTIYPTPQAPT
10 20 30 40 50 60gi|833 PPVVESTPTPTPSAAPEQAEPietSTQPETTKSQPTQPSVATFIPVAAAAAADSAAPIP
70 80 90 100 110 120gi|833 EEPAPQPATTAAPSTSTTTQAAPSAPPAANSGSTKAASSGYSGPCKSGSPCVGQLTTYD
130 140 150 160 170 180

```

CV127.
MIATALVSC-NCR
      10
      :... .: ...
gi|833 TATSASAPSSCGLTNDGFSENVVALPVGIMTDADCGKTVTITYNGITKTATVVDKCMGCK
      190      200      210      220      230      240
      20      30
CV127. TAPPTFDVARHQYLLWLFHQSTHQGP
      : .: .: .: .
gi|833 --PTDLASRHLFGELADFSAGRIDGMSWYFN
      250      260      270

>>gi|66849793 gi|66849793|gb|EAL90121.1| allergen Asp F7 (270 aa)
initn: 30 initl: 30 opt: 47 Z-score: 86.5 bits: 20.9 E(): 6.8
Smith-Waterman score: 47; 33.333% identity (70.833% similar) in 24 aa overlap
(3-25:230-251)

gi|668 MAPIFKSLALVSALFAAIISSAAPVNLDKREVDVVWTTVTTVVWTTIDVTTTIYPTPQAPT
      10      20      30      40      50      60
gi|668 PPVVESTPTPTPSAAPEQAEPIETSTQPETTKSQPTQPSVATFIPVAAAAAADSAAPIP
      70      80      90      100      110      120
gi|668 EEPAPQPATTAAPSTSTTTQAAPSAPPAANSGSTEKAASSGYSGPCSKGSPCVGQLTTYD
      130      140      150      160      170      180

CV127.
MIATALVSC-NCR
      10
      :... .: ...
gi|668 TATSASAPSSCGLTNDGFSENVVALPVGIMTDADCGKTVTITYNGITKTATVVDKCMGCK
      190      200      210      220      230      240
      20      30
CV127. TAPPTFDVARHQYLLWLFHQSTHQGP
      : .: .: .: .
gi|668 --PTDLASRHLFGELADFSAGRIDGMSWYFN
      250      260      270

>>gi|75139991 gi|75139991|sp|Q7M1Y1|Q7M1Y1_FESAR Group I (35 aa)
initn: 38 initl: 38 opt: 38 Z-score: 86.4 bits: 18.0 E(): 6.9
Smith-Waterman score: 38; 30.435% identity (52.174% similar) in 23 aa overlap
(15-37:5-27)

      10      20      30
CV127. MIATALVSCNCR TAPPTFDVARHQYLLWLFHQSTHQGP
      :: ... . :: .: :
gi|751 IAKVPPGPNNITA EYGDKWLD AKSTFYGKPTGAGPK
      10      20      30

>>gi|111120420 gi|111120420|gb|ABH06344.1| Blo t 21 alle (129 aa)
initn: 30 initl: 30 opt: 43 Z-score: 85.2 bits: 19.6 E(): 8
Smith-Waterman score: 43; 25.000% identity (59.375% similar) in 32 aa overlap
(5-36:9-39)

      10      20      30
CV127. MIATALVSCNCR TAPPTFDVARHQYLLWLFHQSTHQGP
      :... : . : . : :... . . :...
gi|111 MKFIIALAALIAVAC-ALPVSNDNFRHEFDHMI VNTATQRFHEIEKFL LHITHEVDDLEK
      10      20      30      40      50

gi|111 TGNKDEKARLLRELTVSEAFIEGSRGYFQRELKRTDLDLLEKFNFEAALATGDLLLLKDLK
      60      70      80      90      100      110

```

gi|111 ALQKRVQDSE
120

>>gi|111120432 gi|111120432|gb|ABH06350.1| Blo t 21 alle (129 aa)
initn: 30 initl: 30 opt: 43 Z-score: 85.2 bits: 19.6 E(): 8
Smith-Waterman score: 43; 25.000% identity (59.375% similar) in 32 aa overlap
(5-36:9-39)

	10	20	30	
CV127.	MIATALVSCNCR	TAPPTFDVARHQYLLWLFHQSTHQGP		
	::::	::::	::::	::::
gi 111	MKFIIALAALIAVAC-ALPVSNDNFRHEFDHMIVNTATQRFHEIEKFLLLHITHEVDDLEE			
	10	20	30	40 50

gi 111	TGNKDEKARLLRELTVSEAFIEGSRGYFQRELKRTDLDLLEKFNFEAALATGDLLLLKDLK
	60 70 80 90 100 110

gi|111 ALQKRVQDSE
120

>>gi|111120424 gi|111120424|gb|ABH06346.1| Blo t 21 alle (129 aa)
initn: 30 initl: 30 opt: 43 Z-score: 85.2 bits: 19.6 E(): 8
Smith-Waterman score: 43; 25.000% identity (59.375% similar) in 32 aa overlap
(5-36:9-39)

	10	20	30	
CV127.	MIATALVSCNCR	TAPPTFDVARHQYLLWLFHQSTHQGP		
	::::	::::	::::	::::
gi 111	MKFIIALAALIAVAC-ALPVSNDNFRHEFDHMIVNTATQRFHEIEKFLLLHITHEVDDLEK			
	10	20	30	40 50

gi 111	TGNKDEKARPLRELTVSEAFIEGSRGYFQRELKRTDLDLLEKFNFEAALATGDLLLLKDLK
	60 70 80 90 100 110

gi|111 ALQKRVQDSE
120

>>gi|111494253 gi|111494253|gb|ABH06347.1| Blo t 21 alle (129 aa)
initn: 30 initl: 30 opt: 43 Z-score: 85.2 bits: 19.6 E(): 8
Smith-Waterman score: 43; 25.000% identity (59.375% similar) in 32 aa overlap
(5-36:9-39)

	10	20	30	
CV127.	MIATALVSCNCR	TAPPTFDVARHQYLLWLFHQSTHQGP		
	::::	::::	::::	::::
gi 111	MKFIIALAALIAVAC-ALPVSNDNFRHEFDHMIVNAATQRFHEIEKFLLLHITHEVDDLEK			
	10	20	30	40 50

gi 111	TGNKDEKARLLRELTVSEAFIEGSRGYFQRELKRTDLDLLEKFNFEAALATGDLLLLKDLK
	60 70 80 90 100 110

gi|111 ALQKRVQDSE
120

>>gi|2507248 gi|2507248|sp|P49275.2|DERF3_DERFA Mite all (259 aa)
initn: 46 initl: 46 opt: 46 Z-score: 85.1 bits: 20.6 E(): 8.1
Smith-Waterman score: 46; 41.176% identity (70.588% similar) in 17 aa overlap
(8-24:207-223)

gi 250	MMILTIVVLLAANILATPILPSSPNATIVGGVKAQAGDCPYQISLQSSSHFCGGSILDEY
	10 20 30 40 50 60

gi|250 WILTAACHVNGQSAKKLSIRYNTLKHASGGEKIQVAEIQHENYDSMTIDNDVALIKLKT
70 80 90 100 110 120

gi|250 PMTLDQTNAPVPLPAQGSVDKVGDKIRVSGWGYLQEGSYSLPSELQRVDIDVVSREQCD
130 140 150 160 170 180

CV127. MIATALVSCNCRTPPTFDVARHQYLLWLFHQSTHQGP
10 20 30
:: .. :: :: ::

gi|250 QLYSKAGADVSENMICGGDVANGGVDSQGDSSGGPVVDVATKQIVGIVSWGYGCAKRGYP
190 200 210 220 230 240

gi|250 GYTRVGNFVDWIESKRSQ
250

>>gi|407609 gi|407609|gb|AAB27812.1| Bra j IE small chai (37 aa)
initn: 35 initl: 35 opt: 37 Z-score: 84.3 bits: 17.6 E(): 9
Smith-Waterman score: 37; 31.034% identity (51.724% similar) in 29 aa overlap
(11-37:9-34)

CV127. MIATALVSCNCRTPPTFDVARHQYLL--WLFHQSTHQGP
10 20 30
:: .. :: :: ::
gi|407 AGPFRFPRCRK--EFQQAQHLRACQQLHKQAMQSGSGP
10 20 30

38 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:32:00 2008 done: Sun Dec 14 22:32:00 2008
Total Scan time: 0.070 Total Display time: 0.020

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.5086_c1

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.5086_c1, 44 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 3.1453 \pm 0.00329$; $\mu = 7.5936 \pm 0.171$

mean_var=32.0458 \pm 9.962, 0's: 25 Z-trim: 25 B-trim: 52 in 1/42

Lambda= 0.226563

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2

join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:

	opt	bits	E(1313)	%_id	%_sim	alen
gi 34495278 gi 34495278 gb AAQ73486 (141)	50	22.2	1.7	0.364	0.727	22
gi 34495286 gi 34495286 gb AAQ73490 (141)	48	21.6	2.7	0.364	0.727	22
gi 4572592 gi 4572592 gb AAD13530.2 (412)	50	22.6	3.6	0.611	0.778	18
gi 4138173 gi 4138173 emb CAA09884. (162)	47	21.3	3.7	0.636	0.909	11
gi 34495282 gi 34495282 gb AAQ73488 (141)	46	20.9	4.2	0.364	0.727	22
gi 34495280 gi 34495280 gb AAQ73487 (140)	45	20.6	5.2	0.364	0.682	22
gi 34495290 gi 34495290 gb AAQ73492 (141)	45	20.6	5.2	0.364	0.682	22
gi 34495288 gi 34495288 gb AAQ73491 (141)	45	20.6	5.2	0.364	0.682	22
gi 33772588 gi 33772588 gb AAQ54603 (141)	45	20.6	5.2	0.364	0.682	22
gi 34495284 gi 34495284 gb AAQ73489 (141)	45	20.6	5.2	0.364	0.682	22
gi 34495274 gi 34495274 gb AAQ73484 (141)	44	20.2	6.6	0.364	0.682	22
gi 129641 gi 129641 sp P15476 PAT1_ (377)	47	21.6	6.7	0.444	0.722	18
gi 21510 gi 21510 emb CAA31575.1 u (386)	47	21.6	6.8	0.444	0.722	18
gi 21514 gi 21514 emb CAA27588.1 p (386)	47	21.6	6.8	0.444	0.722	18
gi 21512 gi 21512 emb CAA27571.1 p (386)	47	21.6	6.8	0.444	0.722	18
gi 169500 gi 169500 gb AAA33819.1 (386)	47	21.6	6.8	0.444	0.722	18
gi 2253610 gi 2253610 gb AAC34737.1 (274)	45	20.8	8.4	0.417	0.750	24
gi 2580504 gi 2580504 gb AAB82404.1 (395)	46	21.3	8.7	0.257	0.629	35
gi 729764 gi 729764 sp P40918.1 HSP (643)	47	21.9	9.8	0.367	0.667	30

>>>CV127.5086_c1, 44 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>gi|34495278 gi|34495278|gb|AAQ73486.1| type 2 allergen (141 aa)
initn: 50 initl: 50 opt: 50 Z-score: 97.4 bits: 22.2 E(): 1.7
Smith-Waterman score: 50; 36.364% identity (72.727% similar) in 22 aa overlap
(13-34:5-26)

```

              10      20      30      40
CV127. MWTQLILLDDIVVLFNLETTAHMSTFQLKTNGKSIFHARFERY
              . . . . .
gi|344      MMKFIVLFALDAVASAGKMKFKDCGKGEVTELDITDCSGDFCVLHRGKSVTL
              10      20      30      40      50

gi|344 DAKFVANQDSAKATIKVLAKVAGTQIQVPGFDTDGCKIIKCPICKGDPIDFKYSGTIPAI
              60      70      80      90      100      110

gi|344 TPKIKAEVTAELIGDHGILACGTVNGQVE
              120      130      140

```

>>gi|34495286 gi|34495286|gb|AAQ73490.1| type 2 allergen (141 aa)
initn: 48 initl: 48 opt: 48 Z-score: 93.9 bits: 21.6 E(): 2.7

Smith-Waterman score: 48; 36.364% identity (72.727% similar) in 22 aa overlap (13-34:5-26)

```

      10      20      30      40
CV127. MWTQLILLLLDDIVVLFNLETTAHMSTFQLKTNGKSIFHARFERY
      .:.: .:.: .:.: .:.:
gi|344      MMKFIVLFALNAVASAGKMKFKDCGKGEVTELDITDCSGDFCVLHRGKSVTL
      10      20      30      40      50

gi|344 DAKFVANQDSAKATIKVLAKVAGTQIQVPGFDTDGCKIIKCPKKGDPIDFKYSGTIPAI
      60      70      80      90      100      110

gi|344 TPKIKA EVTAE LIGDHGILACGTVNGQVE
      120      130      140

```

>>gi|4572592 gi|4572592|gb|AAD13530.2| major allergen B1 (412 aa)
initn: 46 initl: 46 opt: 50 Z-score: 91.5 bits: 22.6 E(): 3.6
Smith-Waterman score: 50; 61.111% identity (77.778% similar) in 18 aa overlap (6-21:338-355)

```

gi|457 NLEKLREKGV DVDKIIELIRALFGLTLNAKASRN LQDDLQDFLALIPVDQIIAIATDYL
      10      20      30      40      50      60

gi|457 ANDAEVQA AAVAYLQSDEFETIVVALDALPELQNFLNFLEANGLNAIDFLNGIHDL LGIPH
      70      80      90      100      110      120

gi|457 IPVSGRKYHIRRGV GITGLIDDVLAILPIEDLKALFNEKLETSPDFLALYNAIRSP EFQS
      130      140      150      160      170      180

gi|457 IVQTLNAMPEYQNLLQKLREKGV DVDKIIELIRALFGLTLNGKASRN LQDDLQDFLALIP
      190      200      210      220      230      240

gi|457 VDQIIAIATDYLANDAEVQA AAVAYLQSDEFETIVVTLDALPELQNFLNFLEANGLNAIDF
      250      260      270      280      290      300

      10      20
CV127.      MWTQLILLLLDDIVVLFN--LETTAHMST
      .:.: .:.: .:.: .:.:
gi|457 LNGIHDL LGIPHIPVSGRKYHIRRGV GITGLIDDVLAILPLDDLKALFNEKLETSPDFLA
      310      320      330      340      350      360

      30      40
CV127. FQLKTNGKSIFHARFERY

gi|457 LYNAIKSP EFQSIVQTLNAMPEYQNLEKLREKGV DVDKIIELIRALFGLTH
      370      380      390      400      410

```

>>gi|4138173 gi|4138173|emb|CAA09884.1| allergen [Malass (162 aa)
initn: 42 initl: 42 opt: 47 Z-score: 91.3 bits: 21.3 E(): 3.7
Smith-Waterman score: 47; 63.636% identity (90.909% similar) in 11 aa overlap (31-41:71-81)

```

      10      20
CV127.      MWTQLILLLLDDIVVLFNLET

gi|413 MSNVFFDITKNGAPLGTIKFKLFDDVVPKTAANFRALCTGEKGFYAGSHFHRVIPDFML
      10      20      30      40      50      60

      30      40
CV127. TAHMSTFQLKTNGKSIFHARFERY
      .:.: .:.: .:.:
gi|413 QGGDFTAGNGTGKSIYGAKFADENFQLKHNPGLLSMANAGPNTNGSQFFITTVVTSWL
      70      80      90      100      110      120

```

gi|413 DGKHVVVFGEVIDGMNVVKAIEAEGSGSGKPRSRIEIAKCGVC
130 140 150 160

>>gi|34495282 gi|34495282|gb|AAQ73488.1| type 2 allergen (141 aa)
initn: 46 initl: 46 opt: 46 Z-score: 90.3 bits: 20.9 E(): 4.2
Smith-Waterman score: 46; 36.364% identity (72.727% similar) in 22 aa overlap
(13-34:5-26)

CV127. MWTQLILLDDIVVLFNLETTAHMSTFQLKTNGKSIFHARFERY
. : : : : : : : : : : : :
gi|344 MMKFIVLFALVAVASAGNMKFKDCGKGVEITDITDCSGDFCVLHRGKSVTL
10 20 30 40 50

gi|344 DAKFVANQDSAKATIKVLAKVAGTQIQVPGFDTGCKIIKCPPIKKGDPIDFKYSGTIPAI
60 70 80 90 100 110

gi|344 TPKIKAEVTAELIGDHGILACGTVNGQVE
120 130 140

>>gi|34495280 gi|34495280|gb|AAQ73487.1| type 2 allergen (140 aa)
initn: 45 initl: 45 opt: 45 Z-score: 88.6 bits: 20.6 E(): 5.2
Smith-Waterman score: 45; 36.364% identity (68.182% similar) in 22 aa overlap
(13-34:4-25)

CV127. MWTQLILLDDIVVLFNLETTAHMSTFQLKTNGKSIFHARFERY
. : : : : : : : : : : : :
gi|344 MKFIVLFALVAVASAGKMKFKDCGKGVEITDITDCSGDFCVLHRGQSVTL
10 20 30 40 50

gi|344 DAKFVANQDSAKATIKVLAKVAGTQIQVPGFDTGCKIIKCPPIKKGDPIDFKYSGTIPAI
60 70 80 90 100 110

gi|344 TPKIKAEVTAELIGDHGILACGTVNGQVE
120 130 140

>>gi|34495290 gi|34495290|gb|AAQ73492.1| type 2 allergen (141 aa)
initn: 45 initl: 45 opt: 45 Z-score: 88.6 bits: 20.6 E(): 5.2
Smith-Waterman score: 45; 36.364% identity (68.182% similar) in 22 aa overlap
(13-34:5-26)

CV127. MWTQLILLDDIVVLFNLETTAHMSTFQLKTNGKSIFHARFERY
. : : : : : : : : : : : :
gi|344 MMKFIVLFALVAVASAGKMKFKDCGKGVEITDITDCSGDFCVLHRGKSVTL
10 20 30 40 50

gi|344 DAKFVANQDSAKATIKVLAKVVGTTQIQVPGFDTGCKIIKCPPIKKGDPIDFKYSGTIPAI
60 70 80 90 100 110

gi|344 TPKIKAEVTAELIGDHGILACGTVNGTVE
120 130 140

>>gi|34495288 gi|34495288|gb|AAQ73491.1| type 2 allergen (141 aa)
initn: 45 initl: 45 opt: 45 Z-score: 88.6 bits: 20.6 E(): 5.2
Smith-Waterman score: 45; 36.364% identity (68.182% similar) in 22 aa overlap
(13-34:5-26)


```

      10      20      30      40
CV127. MWTQLILLLLDDIVVLFNLETTAHMSTFQLKTNGKSIFHARFERY
      . . . . : . . . : . . . . : . .
gi|344      MMKFIVLFALVAVASAGKMKFKDCGKGVEITELDITDCSGDFCVLHRGKSVTL
              10      20      30      40      50

gi|344 DAKFVANQDSAKATIKVLAKVAGTQIQVPGFDTDGCKIIKCPPIKKGDPIDFKYSGTIPAI
      60      70      80      90      100      110

gi|344 TPKIKAEVTAELIGDHGILACGTVNGQVE
      120      130      140

>>gi|33772588 gi|33772588|gb|AAQ54603.1| Gly d 2.03 [Gly (141 aa)
  initn: 45 initl: 45 opt: 45 Z-score: 88.6 bits: 20.6 E(): 5.2
Smith-Waterman score: 45; 36.364% identity (68.182% similar) in 22 aa overlap
(13-34:5-26)

```

```

      10      20      30      40
CV127. MWTQLILLLLDDIVVLFNLETTAHMSTFQLKTNGKSIFHARFERY
      . . . . : . . . : . . . . : . .
gi|337      MMKFIVLFALVAVASAGKMKFKDCGKGVEITELDITDCSGDFCVLHRGKSVTL
              10      20      30      40      50

gi|337 DAKFVANQDSAKATIKVLAKVAGTQIQVPGFDTDGCKIIKCPPIKKGDPIDFKYSGTIPAI
      60      70      80      90      100      110

gi|337 TPKIKAEVTAELIGDHGILACGTVNGQVE
      120      130      140

>>gi|34495284 gi|34495284|gb|AAQ73489.1| type 2 allergen (141 aa)
  initn: 45 initl: 45 opt: 45 Z-score: 88.6 bits: 20.6 E(): 5.2
Smith-Waterman score: 45; 36.364% identity (68.182% similar) in 22 aa overlap
(13-34:5-26)

```

```

      10      20      30      40
CV127. MWTQLILLLLDDIVVLFNLETTAHMSTFQLKTNGKSIFHARFERY
      . . . . : . . . : . . . . : . .
gi|344      MMKFIVLFALVAVASAGKMKFKDCGKGVEITELDITDCSGGFCVLHRGKSVTL
              10      20      30      40      50

gi|344 DAKFVANQDSAKATIKVLAKVAGTQIQVPGFDTDGCKIIKCPPIKKGDPIDFKYSGTIPAI
      60      70      80      90      100      110

gi|344 TPKIKAEVTAELIGDHGILACGTVNGQVE
      120      130      140

>>gi|34495274 gi|34495274|gb|AAQ73484.1| type 2 allergen (141 aa)
  initn: 44 initl: 44 opt: 44 Z-score: 86.8 bits: 20.2 E(): 6.6
Smith-Waterman score: 44; 36.364% identity (68.182% similar) in 22 aa overlap
(13-34:5-26)

```

```

      10      20      30      40
CV127. MWTQLILLLLDDIVVLFNLETTAHMSTFQLKTNGKSIFHARFERY
      . . . . : . . . : . . . . : . .
gi|344      MMKFIVLFALIAVASAGKMKFKDCGKGVEITELDITDCSGDFCVLHRGKSVTL
              10      20      30      40      50

gi|344 DAKFVANQDSAKATIKVLAKVAGTQIQVPGFDTDGCKIIKCPPIKKGDPIDFKYSGTIPAI
      60      70      80      90      100      110

gi|344 TPKIKAEVTAELIGDHGILACGTVNGQVE
      120      130      140

```

>>gi|129641 gi|129641|sp|P15476|PAT1_SOLTU Patatin B1 pr (377 aa)
initn: 47 initl: 47 opt: 47 Z-score: 86.7 bits: 21.6 E(): 6.7
Smith-Waterman score: 47; 44.444% identity (72.222% similar) in 18 aa overlap
(20-37:139-156)

gi|129 LFFMILATTSSTCAKLEEMVTVLSIDGGGIKGIIPAIILEFLEGQLQEVDNNKDARLADY
10 20 30 40 50 60

CV127. M

gi|129 FDVIGGTSTGGLLTAMITTPNENNRPFAAAKDIVPFYFEHGPHIFNYSGSIFGPRYDGKY
70 80 90 100 110 120

10 20 30 40
CV127. WTQLILLDDIVVLFNLETTAHMSTFQLKTNGKSIFHARFERY
: : :

gi|129 LLQVLQEKLGGETRVHQALTEVAISSFDIKTNKPVIFTKSNLAKSPELDAKMYDICYIAA
130 140 150 160 170 180

gi|129 APIYFPPHHFVTHTSNGATYEFNLVDGGVATVGDPALLSLSVATRLAQEDPAFSSIKSLD
190 200 210 220 230 240

gi|129 YKQMLLLSLGTGTNSEFDKTYTAEAAKWGPLRWMLAIQQLTNAASSYMTDYYISTVFQA
250 260 270 280 290 300

gi|129 RHSQNNYLRVQENALTGTTTEMDDASEANMELLVQVGETLLKKPVSKDSPETYEELKRF
310 320 330 340 350 360

gi|129 AKLLSNRKKLRANKASY
370

>>gi|21510 gi|21510|emb|CAA31575.1| unnamed protein prod (386 aa)
initn: 47 initl: 47 opt: 47 Z-score: 86.5 bits: 21.6 E(): 6.8
Smith-Waterman score: 47; 44.444% identity (72.222% similar) in 18 aa overlap
(20-37:148-165)

gi|215 MATTKSFLILFFMILATTSSTCAKLEEMVTVLSIDGGGIKGIIPAIILEFLEGQLQEVDN
10 20 30 40 50 60

gi|215 NKDARLADYFDVIGGTSTGGLLTAMITTPNENNRPFAAAKDIVPFYFEHGPHIFNYSGSI
70 80 90 100 110 120

10 20 30 40
CV127. MWTQLILLDDIVVLFNLETTAHMSTFQLKTNGKSIFHARFERY
: : :

gi|215 LGPMYDGKYLQVLQEKLGGETRVHQALTEVAISSFDIKTNKPVIFTKSNLAKSPELDAKM
130 140 150 160 170 180

gi|215 YDICYSTAAAPIYFPPHHFVTHTSNGARYEFNLVDGAVATVGDPALLSLSVATRLAQEDP
190 200 210 220 230 240

gi|215 AFSSIKSLDYKQMLLLSLGTGTNSEFDKTYTAEAAKWGPLRWMLAIQQMTNAASSYMTD
250 260 270 280 290 300

gi|215 YYISTVFAQRHQSQNNYLRVQENALNGTTTEMDDASEANMELLVQVGETLLKKPVSKDSPE
310 320 330 340 350 360

gi|215 TYEELKRFKLLSDRKKLRANKASH
370 380

>>gi|21514 gi|21514|emb|CAA27588.1| patatin [Solanum tub (386 aa)
initn: 47 initl: 47 opt: 47 Z-score: 86.5 bits: 21.6 E(): 6.8
Smith-Waterman score: 47; 44.444% identity (72.222% similar) in 18 aa overlap
(20-37:148-165)

gi|215 MATTKSFLILFFMILATTSSTCATLGEMVTVLSIDGGGIKGIIPAILLEFLEGQLQEVDN
10 20 30 40 50 60

gi|215 NKDARLADYFDVIGGTSTGGLLTAMITTPNENNRPFAAAKDIVPFYFEHGPHIFNYSGSI
70 80 90 100 110 120

10 20 30 40
CV127. MWTQLILLDDIVVLFNLETTAHMSTFQLKTNGKSIFHARFERY
: : :

gi|215 FGPRYDGKYLQVLQEKLGETRVHQALTEVAISSFDIKTNKPVIFTKSNLAESPQLDAKM
130 140 150 160 170 180

gi|215 YDICYSTAAAPIYFPPHHFVTHTSNGATYEFNLVDGAVATVGDPALLSLSVATRLAQDDP
190 200 210 220 230 240

gi|215 AFSSIKSLDYKQMLLLSLGTGTNSEFDKTYTAEAAKWGPLRWMLAIQQMTNAASSYMTD
250 260 270 280 290 300

gi|215 YYISTVFQARHSQNNYLRVQENALTGTTTEMDDASEANMELLVQVGETLLKKPVSKDSPE
310 320 330 340 350 360

gi|215 TYEEALKRFAKLLSDRKKLRANKASH
370 380

>>gi|21512 gi|21512|emb|CAA27571.1| patatin [Solanum tub (386 aa)
initn: 47 initl: 47 opt: 47 Z-score: 86.5 bits: 21.6 E(): 6.8
Smith-Waterman score: 47; 44.444% identity (72.222% similar) in 18 aa overlap
(20-37:148-165)

gi|215 MATTNSFTILIFMILATTSSTFATLGEMVTVLSIDGGGIKGIIPATILEFLEGQLQEVDN
10 20 30 40 50 60

gi|215 NTDARLADYFDVIGGTSTGGLLTAMITTPNETNRPFAAAKDIVPFYFEHGPKIFQSSGSI
70 80 90 100 110 120

10 20 30 40
CV127. MWTQLILLDDIVVLFNLETTAHMSTFQLKTNGKSIFHARFERY
: : :

gi|215 FGPKYDGKYLQVLQEKLGETRVHQALTEVAISSFDIKTNKPVIFTKSNLAKSPELDAKM
130 140 150 160 170 180

gi|215 YDICYSTAAAPTFFPPHYFATNTSNGDKYEFNLVDGAVATVDDPALLSISVATKLAQVDP
190 200 210 220 230 240

gi|215 KFASIKSLNYKQMLLLSLGTGTSEFDKTYTAEETAKWGRTARWMLVIQKMTSAASSYMTD
250 260 270 280 290 300

gi|215 YYLSTAFQALDSQNNYLRVQENALTGTTTELDDASEANMQLLVQVGEDLLKKSVDKNPE
310 320 330 340 350 360

gi|215 TYEEALKRFAKLLSDRKKLRANKASY
370 380

>>gi|169500 gi|169500|gb|AAA33819.1| patatin (386 aa)
initn: 47 initl: 47 opt: 47 Z-score: 86.5 bits: 21.6 E(): 6.8
Smith-Waterman score: 47; 44.444% identity (72.222% similar) in 18 aa overlap
(20-37:148-165)

gi|169 MATTKSFLILFFMILATTSSTCAKLEEMVTVLSIDGGGIKGIIPAILLEFLEGQLQEVDN
10 20 30 40 50 60

gi|169 NKDARLADYFDVIGGTSTGGLLTAMITTPNENNRPFAAAKDIVPFYFEHGPHIFNYSGSI
70 80 90 100 110 120

CV127. MWTQLILLDDIVVLFNLETTAHMSTFQLKTNGKSIFHARFERY
10 20 30 40
: : :

gi|169 IGPMYDGKYLQLVLEKLGETRVRHQALETEVAISSFDIKTNKPVIFTKSNLAKSPELDAKM
130 140 150 160 170 180

gi|169 YDICYSTAAAPIYFPPHYFITHTSNGDIYEFNLVDGGVATVGDPAALLSLSVATRLAQEDP
190 200 210 220 230 240

gi|169 AFSSIKSLDYKQMLLSLGTGTNSEFDKTYTAQEAAKWGPLRWMLAIQQMTNAASSYMTD
250 260 270 280 290 300

gi|169 YYISTVFQARHSQNNYLRVQENALTGTTTEMDDASEANMELLVQVGETLLKKPVSKDSPE
310 320 330 340 350 360

gi|169 TYEEALKRFAKLLSDRKKLRANKASY
370 380

>>gi|2253610 gi|2253610|gb|AAC34737.1| Cr-PII allergen [(274 aa)
initn: 36 initl: 36 opt: 45 Z-score: 84.9 bits: 20.8 E(): 8.4
Smith-Waterman score: 45; 41.667% identity (75.000% similar) in 24 aa overlap
(6-27:13-36)

CV127. MWTQLILLDDIVVLFN--LETTAHMSTFQLKTNGKSIFHARFERY
10 20 30 40
: : :

gi|225 VGVDGLIDDIAILPIDDLKALFQEKLETSLDFKAFYDAVRSPEFQSIVQTLNAMPEYQD
10 20 30 40 50 60

gi|225 LLQKLRLDKGVDVDHYIELIRALFGLTREARNLQDDLNDFLALIPTDQILAIAMDYLANDA
70 80 90 100 110 120

gi|225 EVQELVAYLQSDDFHKIINTIEALPEFANFYNFLKGHGLDVANYINEIHSIIGLPPFVPP
130 140 150 160 170 180

gi|225 SRRHARRGVINGLIDDVIAILPVDELKTLFQEKLETSPDFKALYDAIRSPEFQSIISTL
190 200 210 220 230 240

gi|225 NAMPEYQELLQNLRDKGVDVDHFIELIRSWFGLP
250 260 270

>>gi|2580504 gi|2580504|gb|AAB82404.1| Cr-PII [Periplane (395 aa)
initn: 36 initl: 36 opt: 46 Z-score: 84.6 bits: 21.3 E(): 8.7
Smith-Waterman score: 47; 25.714% identity (62.857% similar) in 35 aa overlap
(8-42:127-159)

gi|258 MKLPIMILAVLGVAFGKSLPNRNLQDDLNDFLALLPVDEITAIVMDYLANDAEVQEAVAY
10 20 30 40 50 60

CV127. M

gi|258 LQGEFHKIVFTVEGLQEFGNFVQFLEDHGLDAVGYNRLHVSFVGWDPYVPSSKRKHTRR
70 80 90 100 110 120

```

      10          20          30          40
CV127. WTQLILLLLDDIVVLFNLETTAHMSTFQLKTNKGSIFHARFERY
      .:.::..... ..   . :: : . . . :.: .
gi|258 GVGVDGLIDDIIAILPIDDLKAL--FQEKLSTSPDFKAFYDAVRSPFEQSIVQTLNAMPE
      130          140          150          160          170

gi|258 YQDLLQKLRLDKGVDVDHYIELIRALFGLTRAARNLQDDLNDFLALIPTDQILAIAMDYLA
      180          190          200          210          220          230

gi|258 NDAEVQELVAYLQSDDFHKIINTIEALPEFANFYNFLKGHGLDVADYINEIHSHIIIGLPFF
      240          250          260          270          280          290

gi|258 VPPSRRHARRGVGINGLIDDVIAILPVDELKALFQEKLSTSPDFKALYDAIRSPEFQSII
      300          310          320          330          340          350

gi|258 STLNAMPEYQELLQNLRDKAVDVDHFIELIRSLFGLP
      360          370          380          390


>gi|729764 gi|729764|sp|P40918.1|HSP70_CLAHE Heat shock (643 aa)
initn: 34 initl: 34 opt: 47 Z-score: 83.7 bits: 21.9 E(): 9.8
Smith-Waterman score: 47; 36.667% identity (66.667% similar) in 30 aa overlap
(3-32:387-416)

gi|729 MAPAIGIDLGTTYSCVGIYRDDRIEIIANDQGNRTTPSFVAFTDTERLIGDSAKNQVAIN
      10          20          30          40          50          60

gi|729 PHNTVFDAKRLIGRKFQDAEVQADMKHFPFKVIEKAGKPVTQVEFKGETKDFTPEEISSM
      70          80          90          100         110         120

gi|729 ILTKMRETAESYLGGTVNNNAVITVPAYFNDSQRQATKDAGLIAGLNLVLRINEPTAAAIA
      130          140          150          160          170          180

gi|729 YGLDKKQEGEKNVLIFDLGGGTFDVSFLTIEEGIFEVKSTAGDTHLGGEDFDNRLNVNHFS
      190          200          210          220          230          240

gi|729 NEFKRKHHKKDLSDNARALRRLRTACERAKRTLSSSAQTSIEIDSLEFEGIDFFTSTNRARF
      250          260          270          280          290          300

gi|729 EEVGQDLFRGNMEPGERTLRDDKIDKSSVHEIVLGGGSTRIPKVQKLVSDFFNGKEPCKS
      310          320          330          340          350          360



      10          20          30
CV127. MWTQLILLLLDDIVVLFNLETTAHMSTFQLKTNKSI
      .: .::::: . ....:.. . : .:
gi|729 INPDEAVAYGA AVQAAILSGDTSSKSTKEIILLLDVAPLSLGIETAGGVMTALIKRNTTIP
      370          380          390          400          410          420




      40
CV127. FHARFERY

gi|729 TKKSETFSTFSDNPQGVLIQVFEGERARTKDINLMGFELSGIRPAPRGVPQIEVTFDLD
      430          440          450          460          470          480

gi|729 ANGIMNVSALEKGTGKTNKIVITNDKGRLSKEEIERMLADAKEYKEEDEAEAGRIQAKNG
      490          500          510          520          530          540

gi|729 LESYAYSLKNTVSDPKVEEKLSAEDKETLTGAIDKTVAWIDENQTATKEEYEAEQKQLES
      550          560          570          580          590          600

gi|729 VANPVMMKIYGAEQGAPGGMPGQGAGAPPPGAGDDGPTEVEVD
      610          620          630          640

```



The Chemical Company
Plant Science L. L. C.

BASF Reg. Doc. No. 2008/7019371

44 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:41:41 2008 done: Sun Dec 14 22:41:41 2008
Total Scan time: 0.070 Total Display time: 0.030

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.5688 d3

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.5688 d3, 38 aa

```
vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library
```

292011 residues in 1313 sequences

Expectation n fit: $\rho(\ln(x)) = 4.8710 \pm 0.00414$; $\mu = -4.2852 \pm 0.215$

```
mean_var=53.4008+/-13.453, 0's: 39 Z-trim: 40 B-trim: 50 in 1/41
```

Lambda= 0.175509

FASTA (3.5 Sept. 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2

```
join: 36, opt: 20, open/ext: -10/-2, width: 16
```

The best scores are:

The best scores are:					opt	bits	E(1313)	%_id	%_sim	alen
gi 76666767	gi 76666767	emb CAA5507	(497)	67	25.4	0.55	0.423	0.577	26	
gi 3913017	gi 3913017	sp P81496.1	A (27)	42	18.5	3.7	0.385	0.615	13	
gi 170708	gi 170708	gb AAA34274.1	(291)	52	21.5	4.8	0.261	0.652	23	
gi 1169665	gi 1169665	sp P30438.2	F (92)	44	19.3	7.3	0.308	0.692	13	
gi 50659889	gi 50659889	gb AAT80664	(115)	44	19.3	8.9	0.296	0.519	27	
gi 50659885	gi 50659885	gb AAT80662	(115)	44	19.3	8.9	0.296	0.519	27	
gi 38492338	gi 38492338	gb AAR22488	(115)	44	19.3	8.9	0.222	0.519	27	
gi 50659891	gi 50659891	gb AAT80665	(115)	44	19.3	8.9	0.296	0.519	27	
gi 256429	gi 256429	gb AAB23464.1	(216)	47	20.2	9	0.333	0.625	24	
gi 18770	gi 18770	emb CAA45777.1	t (217)	47	20.2	9	0.333	0.625	24	
gi 14285595	gi 14285595	sp O04004.1	(118)	44	19.3	9.1	0.368	0.632	19	

```
>>>CV127.5688_d3, 38 aa vs
```

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

```
>>gi|76666767|gi|76666767|emb|CAA55071.2| aldehyde dehyd (497 aa)
```

```
initn: 43 init1: 43 opt: 67 Z-score: 106.2 bits: 25.4 E(): 0.55
```

Smith-Waterman score: 67; 42.308% identity (57.692% similar) in 26 aa overlap (10-35:146-171)

gi|766 MTSVKLSTPQTGEFEQPTGLFINNEFVKAVDGTKTFDVINPSTEEVICSVQEATEKDVDIA
10 20 30 40 50 60gi|766 VAAARKAFNGPWRKETPENRGKLLNKLADLFENADLIAAVEALDNGKAFSMAKNVDVPA
70 80 90 100 110 120

CV127. MRDLILVMMSCRYICLECMGLCMQITLWSIVICCWRLE
gi|766 AAGCLRYGGWADKIEGKVVDTPDSFNIRKEPIGVCGQIIPWNFPILMWSWKIGPAIA

gi|766 TGNTVVVLKTAEQTPLSAYIACKLIQEAGFP PGVIN VITGFGKIAGAAMSAHMDIDKIAFT

gi|766 GSTVVGRQIMKSAAGSNLKKVTLELGGKSPNIVFADADLDEAIHWVNFGLIYFNHGQACCA
250 260 270 280 290 300

gi|766 GSRIYVQEEIYDKFIQRFKERAQAQNAVGDPAADTFQGPQVSQLQFDRIMGYIEEGKKS

gi|766 ATIETGGRNRKDGKGYFIEPTIFSNTVEDMKIQEEIFGPVCTISKFKTKADVIGNNNTT
370 380 390 400 410 420

gi|766 YGLAAAVHTSNLTTAIEVANALRAGTVWVNSYNTLHWQLPFGGYKESGIGRELGEAALDN
430 440 450 460 470 480

gi|766 YIQTKTVSIRLGDVLFQ
490

>>gi|3913017 gi|3913017|sp|P81496.1|ALCC_WHEAT Allergen (27 aa)
initn: 34 initl: 34 opt: 42 Z-score: 91.4 bits: 18.5 E(): 3.7
Smith-Waterman score: 42; 38.462% identity (61.538% similar) in 13 aa overlap
(12-24:10-22)

10 20 30
CV127. MRDLILVMMSCRYICLECMGLCMQITLWSIVICWRL
: : : : :
gi|391 SFREQCVPGREITYECLNACAAYAVRQ
10 20

>>gi|170708 gi|170708|gb|AAA34274.1| gamma-gliadin B pre (291 aa)
initn: 36 initl: 36 opt: 52 Z-score: 89.2 bits: 21.5 E(): 4.8
Smith-Waterman score: 52; 26.087% identity (65.217% similar) in 23 aa overlap
(11-32:149-171)

gi|170 MKTLLILTILAMAITIATANMQADPSGQVQWPQQQPFLLQPHQPFSSQQPQQIFPQPQQTFP
10 20 30 40 50 60

gi|170 HQPQQQFPQPQQPQQQFLQPRQPFPQPQQPYPQQPQQPFPQTQQPQQPFPQSKQPQQPF
70 80 90 100 110 120

10 20 30
CV127. MRDLILVMMSCR-YICLECMGLCMQITLWSIVICWRL
: . . . : . . . : : : :
gi|170 PQQPQQPQQSFPQQPQSLIQQSLQQQLNPKNFFLLQQCKPVSLVSSLWSIILPPSDCQVMR
130 140 150 160 170 180

gi|170 QQQCQQLAQIPQQQLQCAAIHSSVHSHIIMQQEQEQQLQGVQILVPLSQQQQVGQILVQGG
190 200 210 220 230 240

gi|170 GIIQPQQPAQLEVIRSLVLQTLPTMCNVYVPPYCSITRAPFASIVASIGGQ
250 260 270 280 290

>>gi|1169665 gi|1169665|sp|P30438.2|FEL1A_FELCA Major al (92 aa)
initn: 23 initl: 23 opt: 44 Z-score: 85.9 bits: 19.3 E(): 7.3
Smith-Waterman score: 44; 30.769% identity (69.231% similar) in 13 aa overlap
(20-32:3-15)

10 20 30
CV127. MRDLILVMMSCRYICLECMGLCMQITLWSIVICWRL
: : . . : : . .
gi|116 MKGACVLVLLWAALLLISGGNCEICPAVKRDVDLFLTGTPEY
10 20 30 40

gi|116 VEQVAQYKALPVVLENARILKNCVDKMTTEEDKENALSVDKIYTSPLC
50 60 70 80 90

>>gi|50659889 gi|50659889|gb|AAT80664.1| lipid transfer (115 aa)
initn: 31 initl: 31 opt: 44 Z-score: 84.4 bits: 19.3 E(): 8.9
Smith-Waterman score: 44; 29.630% identity (51.852% similar) in 27 aa overlap
(8-33:1-27)


```
          10          20          30
CV127. MRDLILVMMSCRYICLE-CMGLCMQITLWSIVICWRLE
      : : : : : : : : : : : :
gi|506      MASSAVINLALVVALCMAVSVVAHAITCGQVSSNLVPCFDYVRSDGVPVPACCN
          10          20          30          40          50

gi|506 GIRTINGLAKTTPDRQAACNCLKSLAGSVSGVNPNGNAESLPGKCGVNPYPYKISTSTNCAT
      60          70          80          90         100         110

gi|506 VK
```

>>gi|50659885 gi|50659885|gb|AAT80662.1| lipid transfer (115 aa)
initn: 31 init1: 31 opt: 44 Z-score: 84.4 bits: 19.3 E(): 8.9
Smith-Waterman score: 44; 29.630% identity (51.852% similar) in 27 aa overlap
(8-33:1-27)

```
          10          20          30
CV127. MRDLILVMMSCRYICLE-CMGLCMQITLWSIVICWRLE
      : : : : : : : : : : : :
gi|506      MASSAVINLALVVALCMAVSVVAHAITCGQVSSNLVPCFDYVRSGGPVPACCN
          10          20          30          40          50

gi|506 GIRTINGLAKTTPDRQAACNCLKSLAGSVSGVNPNGNAESLPGKCGVNPYPYKISTSTNCAT
      60          70          80          90         100         110

gi|506 VK
```

>>gi|38492338 gi|38492338|gb|AAR22488.1| allergen Mal d (115 aa)
initn: 31 init1: 31 opt: 44 Z-score: 84.4 bits: 19.3 E(): 8.9
Smith-Waterman score: 44; 22.222% identity (51.852% similar) in 27 aa overlap
(9-33:1-27)

```
          10          20          30
CV127. MRDLILVMMSCRYICLECM--GLCMQITLWSIVICWRLE
      : : . . . : : : : : : : : :
gi|384      MACSAVIKLALVVALCMAVSVVAHAITCGQVTSSLAPCIGYVRSGGAVPPACC
          10          20          30          40          50

gi|384 NGIRTINGLARTTADRQTACNCLKNLAGSISGVNPNNAAGLPGKCGVNPYPYKISTSTNCA
      60          70          80          90         100         110

gi|384 TVK
```

>>gi|50659891 gi|50659891|gb|AAT80665.1| lipid transfer (115 aa)
initn: 31 init1: 31 opt: 44 Z-score: 84.4 bits: 19.3 E(): 8.9
Smith-Waterman score: 44; 29.630% identity (51.852% similar) in 27 aa overlap
(8-33:1-27)

```
          10          20          30
CV127. MRDLILVMMSCRYICLE-CMGLCMQITLWSIVICWRLE
      : : : : : : : : : : : :
gi|506      MASSAVINLALVVALCMAVSVVAHAITCGQVSSNLVPCFDYVRSGGPVPACCN
          10          20          30          40          50

gi|506 GIRTINGLAKTTPDRQAACNCLKSLAGSVSGVNPNGNVESLPGKCGVNPYPYKISTSTNCAT
      60          70          80          90         100         110

gi|506 VK
```

>>gi|256429 gi|256429|gb|AAB23464.1| Kunitz trypsin inhi (216 aa)
initn: 29 initl: 29 opt: 47 Z-score: 84.4 bits: 20.2 E(): 9
Smith-Waterman score: 47; 33.333% identity (62.500% similar) in 24 aa overlap
(9-31:97-120)

gi|256 MKSTIFFLFLFCFTTSYLPSAIADFVLDNEGNPLENGGTYIILSDITAFGGIRAAPTGN
10 20 30 40 50 60

CV127. MRDLILVMMSCRYICLECMGLCMQI-TLWSIV
10 20 30
.: : : : : : :

gi|256 ERCPLTVVQSRNELDKGIGTIISSPYRIRFIAEGHPLSLKFDSFAVIMLCVGIPTEWSV
70 80 90 100 110 120

CV127. ICCWRLE

gi|256 EDLPEGPAVKIGENKDAMDGWFRLEVRSDDEFNNYKLVFCPQQAEDDKCGDIGISIDHDD
130 140 150 160 170 180

gi|256 GTRRLVVSKNKPLVVQFQKLDKESLAKKNHGLSRSE
190 200 210

>>gi|187770 gi|187770|emb|CAA45777.1| trypsin inhibitor su (217 aa)
initn: 29 initl: 29 opt: 47 Z-score: 84.3 bits: 20.2 E(): 9
Smith-Waterman score: 47; 33.333% identity (62.500% similar) in 24 aa overlap
(9-31:98-121)

gi|187 MKSTIFFALFLFCFTTSYLPSAIADFVLDNEGNPLENGGTYIILSDITAFGGIRAAPTG
10 20 30 40 50 60

CV127. MRDLILVMMSCRYICLECMGLCMQI-TLWSI
10 20 30
.: : : : : : :

gi|187 NERCPLTVVQSRNELDKGIGTIISSPYRIRFIAEGHPLSLKFDSFAVIMLCVGIPTEWSV
70 80 90 100 110 120

CV127. VICWRLE

gi|187 VEDLPEGPAVKIGENKDAMDGWFRLEVRSDDEFNNYKLVFCPQQAEDDKCGDIGISIDHD
130 140 150 160 170 180

gi|187 DGTRRLVVSKNKPLVVQFQKLDKESLAKKNHGLSRSE
190 200 210

>>gi|14285595 gi|14285595|sp|O04004.1|NLTP6_AMBAR Non-sp (118 aa)
initn: 33 initl: 33 opt: 44 Z-score: 84.3 bits: 19.3 E(): 9.1
Smith-Waterman score: 44; 36.842% identity (63.158% similar) in 19 aa overlap
(22-36:3-20)

CV127. MRDLILVMMSCRYICLECMGLCMQITLWSIVI----CCWRLE
10 20 30
: : : : : : : : : : :

gi|142 MDCIRI-LWSVAVGLLLVSWRPTMFAASPTCDTVQNILAPC
10 20 30 40

gi|142 AGFLTGTQEPSKACCTGVNNLNNSRKTADRVAVCNCIKELTKSIAYDPKRMPLSTKCGV
50 60 70 80 90 100

gi|142 KPDPFAVDKNLDCSKLPV
110



The Chemical Company
Plant Science L. L. C.

BASF Reg. Doc. No. 2008/7019371

38 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:45:32 2008 done: Sun Dec 14 22:45:32 2008
Total Scan time: 0.070 Total Display time: 0.000

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.2041_d1

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.2041_d1, 34 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: rho(ln(x))= 3.5071+/-0.0039; mu= 7.4621+/- 0.204

mean_var=63.4947+/-18.147, 0's: 43 Z-trim: 43 B-trim: 29 in 1/41

Lambda= 0.160955

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2

join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:		opt	bits	E(1313)	%_id	%_sim	alen
gi 50199132	gi 50199132 emb CAH0379 (91)	62	21.4	1.4	0.500	0.650	20
gi 114794319	gi 114794319 pdb 2HCZ (245)	56	20.6	6.7	0.357	0.571	28
gi 28630919	gi 28630919 gb AAO45607 (269)	56	20.7	7	0.357	0.571	28
gi 28630923	gi 28630923 gb AAO45608 (269)	56	20.7	7	0.357	0.571	28
gi 14193761	gi 14193761 gb AAK56124 (269)	56	20.7	7	0.357	0.571	28

>>>CV127.2041_d1, 34 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|50199132 gi|50199132|emb|CAH03799.1| lipid transfer (91 aa)

initn: 40 initl: 40 opt: 62 Z-score: 98.6 bits: 21.4 E(): 1.4

Smith-Waterman score: 62; 50.000% identity (65.000% similar) in 20 aa overlap (5-24:13-29)

	10	20	30	
CV127.	MSFVCSVFVRSGVMDGVS	VPCFCNACWVVSCKKC		
	:	..:::	.	:::
gi 501	ITCGQVTGSLAPCIVYLRSG--GPIPVP	C-CNGVRS	LNAAARTTPDRQTACNCLKQAAGS	
	10	20	30	40
gi 501	IPNLPNNNAVGLPRACGV	SIPYKISISTDCSKVR		
	60	70	80	90

>>gi|114794319 gi|114794319|pdb|2HCZ|X Chain X, Crystal (245 aa)

initn: 32 initl: 32 opt: 56 Z-score: 86.7 bits: 20.6 E(): 6.7

Smith-Waterman score: 56; 35.714% identity (57.143% similar) in 28 aa overlap (11-33:53-80)

	10	20	30	40	50	60
CV127.	MSFVCSVFVRSGVMDGVS					
	:	..	.			
gi 114	GPPKVPPGXNITTN	YNGKWL	TARATWYGQ	PNGAGAPD	NGGACGIKNV	LPYPYSGMTACGN
	10	20	30	40	50	60
	20	30				
CV127.	VPCF----	CNACWVVSCKKC				
	::	:	..:::	.	:::	
gi 114	VPIFKD	GKGC	SGCYEVRCKE	KPECSGNP	VTYITDMN	YEPIAPYHFDLSGKAFGLAKPG
	70	80	90	100	110	120
gi 114	LNDKIRHCGIMD	VEFRRVRCKYPAGQ	KIVFHIEKGCNPNYLAVLVKYVADDG	DIIVLMEIQ		
	130	140	150	160	170	180

gi | 114 DKL^{SA}E^WK^{PM}KL^{SW}GA^IW^{RM}DTAK^{AL}K^{GP}FS^{IR}LT^{SE}SG^{KK}VI^{AK}D^{VI}PAN^{WR}PD^{AV}Y^{TS}
190 200 210 220 230 240

qi | 114 NVOFY

```
>>gi|28630919 gi|28630919|gb|AA045607.1| beta-expansin 9 (269 aa)
  initn: 32 initl: 32 opt: 56 Z-score: 86.3 bits: 20.7 E(): 7
Smith-Waterman score: 56; 35.714% identity (57.143% similar) in 28 aa overlap
(11-33:77-104)
```

gi|286 MGSLANNIMVVGAVLAALVVGSCGPPKVPVPGPNITTNYGKWL TARATWYGQPNGAGAP
10 20 30 40 50 60

```

              10      20      30
CV127.      MSFVCSVFVRSGVMDGVSVPCF-----CNACWVVVSKKC
              :::  :::  :::  :::  :::
gi|286 DNGGACGIKNVNLPPYSGMTACGNVPIFKDGKGCSCGYEVRCKEKEPECSGNPVTVFITDM
              70      80      90      100     110     120

```

gi|286 NYEPIAPYHFDLSGKAFGSLAKPGLNDKLRHCGIMDVEFRVRCKYPAGQKIVFHIEKGC
130 140 150 160 170 180gi|286 NPNYVAVLVKFVADDGDIVLMEIQDKLSAEWKPKMLSWGAIWRMDTAKALKGPFSSIRLTS
190 200 210 220 230 240gi|286 ESGKKVIAKDIIIPANWRPDVYTSNVQFY
250 260

```
>>gi|28630923 gi|28630923|gb|AA045608.1| beta-expansin 1 (269 aa)
  initn: 32 initl: 32 opt: 56 Z-score: 86.3 bits: 20.7 E(): 7
Smith-Waterman score: 56; 35.714% identity (57.143% similar) in 28 aa overlap
(11-33:77-104)
```

gi|286 MGSLANNIMVVGAVLAALVAGGSCGPPKVPVPGPNITTNYGKWL TARATWYGQPNGAGAP
10 20 30 40 50 60

```

              10      20      30
CV127.      MSFVCSVFVRSGVMDGVSVPCF-----CNACWVVVSKKC
              :::  :::  :::  :::  :::
gi|286 DNGGACGIKNVNLPPYSGMTACGNVPIFKDGKGCSCGYEVRCKEKEPECSGNPVTYITDM
              70      80      90      100      110      120

```

gi | 286 NYEPIAPYHFDLSGKAFGSLAKPGLNDKIRHCGIMDVEFRRVRCKYPAGQKIVFHIEKGC

gi|286 NPNYLAVLVKYVADDGDIVLMEIQDKLSAEWKPMKLSWGAIWRMDTAKALKGPFSSIRLTS
190 200 210 220 230 240gi|286 ESGKKVIAKDVIPANWRPDVYTSNVQFY
250 260

```
>>gi|14193761 gi|14193761|gb|AAK56124.1|AF332174_1 beta- (269 aa)
initn: 32 initl: 32 opt: 56 Z-score: 86.3 bits: 20.7 E(): 7
Smith-Waterman score: 56; 35.714% identity (57.143% similar) in 28 aa overlap
(11-33:77-104)
```

gi|141 MGSLVNNIMVVGAVLAALVAGGSCGPPKVPVPGPNITTNYGKWL TARATWYGQPNGAGAP
10 20 30 40 50 60

```

      10      20      30
CV127.  MSFVCSVFVRSGVMDGVSVPFC-----CNACWVVSCKKC
      ::.  .:: :  :... : ::.
gi|141 DNGGACGIKNVNLPPYSGMTACGNVPIFKDGKGCSCYEVRCKEKPECSGNPVTVYITDM
      70      80      90      100      110      120

gi|141 NYEPIAPYHFDLSGKAFGSLAKPGLNDKIRHCGIMDVEFRRVRCKYPAGQKIVFHIEKGC
      130      140      150      160      170      180

gi|141 NPNYLAVLVKYVADDGDIVLMEIQDKLSAEWKPMKLSWGAIWRMDTAKALKGPFSIRLTS
      190      200      210      220      230      240

gi|141 ESGKKVIAKDVIPANWRPDVYTSNVQFY
      250      260
```

34 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:06:06 2008 done: Sun Dec 14 22:06:06 2008
Total Scan time: 0.060 Total Display time: 0.000

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.4778_d2

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.4778_d2, 34 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 4.1657 \pm 0.00359$; $\mu = 1.2040 \pm 0.186$

mean_var=30.2246 \pm 7.518, 0's: 25 Z-trim: 25 B-trim: 5 in 1/42

Lambda= 0.233289

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2

join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:			opt	bits	E(1313)	%_id	%_sim	alen
gi 20135538	gi 20135538	gb AAL25839 (116)	45	20.9	2.7	0.259	0.556	27
gi 13507262	gi 13507262	gb AAK28533 (115)	44	20.5	3.4	0.417	0.542	24
gi 21725604	gi 21725604	emb CAD3838 (129)	42	19.8	6.1	0.438	0.625	16
gi 21725592	gi 21725592	emb CAD3837 (129)	42	19.8	6.1	0.438	0.625	16
gi 45108967	gi 45108967	emb CAF3256 (500)	46	21.2	9	0.294	0.647	17
gi 45108973	gi 45108973	emb CAF3256 (500)	46	21.2	9	0.294	0.647	17
gi 54144332	gi 54144332	emb CAD5467 (508)	46	21.3	9.1	0.294	0.647	17
gi 10189816	gi 10189816	emb CAC0923 (213)	42	19.9	9.9	0.370	0.593	27
gi 2498299	gi 2498299	sp Q26456.1 A (213)	42	19.9	9.9	0.370	0.593	27

>>>CV127.4778_d2, 34 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|20135538 gi|20135538|gb|AAL25839.1| lipid transfer (116 aa)

initn: 31 init1: 31 opt: 45 Z-score: 93.6 bits: 20.9 E(): 2.7

Smith-Waterman score: 45; 25.926% identity (55.556% similar) in 27 aa overlap (1-27:1-27)

	10	20	30	
CV127.	MKPVI	IRTFYGLCMH	MVKKLSLQFPV	CFGNLSFF
	:	:	:	:
gi 201	MAALKMVS	FLVLCMLVA	APMTAQAITCG	QVQSALVPCLSYLKTTGPTPPATCCNGVRTIN
	10	20	30	40 50 60
gi 201	NAAKTTAD	RRRTACQCL	KSAAAGSVKGLN	PPTVAGLPKCGVNIPYKISLSTNCATVK
	70	80	90	100 110

>>gi|13507262 gi|13507262|gb|AAK28533.1|AF329829_1 lipid (115 aa)

initn: 30 init1: 30 opt: 44 Z-score: 91.9 bits: 20.5 E(): 3.4

Smith-Waterman score: 44; 41.667% identity (54.167% similar) in 24 aa overlap (12-32:12-35)

	10	20	30	
CV127.	MKPVI	IRTFYGLCMH	M---VKKLSLQFPV	CFGNLSFF
	:	:	:	:
gi 135	MGSLKL	VCVLLCMMVA	APVARASLTCPQ	IKGNLTPCVLYLKNGGVLPSCCKGVRAVND
	10	20	30	40 50 60
gi 135	ASRTTSD	RQSACNCLK	DTAKGIAGLN	PNLAAGLPKCGVNIPYKISPSTNCNNVK
	70	80	90	100 110

>>gi|21725604 gi|21725604|emb|CAD38383.1| unnamed protei (129 aa)
initn: 33 initl: 33 opt: 42 Z-score: 87.4 bits: 19.8 E(): 6.1
Smith-Waterman score: 42; 43.750% identity (62.500% similar) in 16 aa overlap
(15-30:11-24)

```

          10          20          30
CV127. MKPVIIRTFYGLCMHMKKLSLQFPVCFGNLSFF
          : .:. . . : : :
gi|217   DQVDVKDCANHEIKKVLV--PGCHGNEPCIHSGKPFQLEALFEANQNSATAKIEI
          10          20          30          40          50

gi|217 KASIDGLSVDVPGIDPNACNYMKCPLVNGQQYDIKYTWNVPKIAPNSENVVVTVKVLGDN
          60          70          80          90         100         110

gi|217 GVLACAIATHAKIQD
          120

```

>>gi|21725592 gi|21725592|emb|CAD38377.1| unnamed protei (129 aa)
initn: 33 initl: 33 opt: 42 Z-score: 87.4 bits: 19.8 E(): 6.1
Smith-Waterman score: 42; 43.750% identity (62.500% similar) in 16 aa overlap
(15-30:11-24)

```

          10          20          30
CV127. MKPVIIRTFYGLCMHMKKLSLQFPVCFGNLSFF
          : .:. . . : : :
gi|217   DQVDVKDCANHEIKKVLV--PGCHGNEPCIIGRGKPFQLEALFEANQNSATAKIEI
          10          20          30          40          50

gi|217 KASIDGLSVDVPGIDPNACHYMNCPVNGQQYDIKYTWNVPKIAPNSENVVVTVKVLGDN
          60          70          80          90         100         110

gi|217 GVLACAIATHAKIQD
          120

```

>>gi|45108967 gi|45108967|emb|CAF32566.2| unnamed protei (500 aa)
initn: 46 initl: 46 opt: 46 Z-score: 84.4 bits: 21.2 E(): 9
Smith-Waterman score: 46; 29.412% identity (64.706% similar) in 17 aa overlap
(9-25:277-293)

```

gi|451 YFPPPAAKEDFLGCLVKEIPRLLYAKSSPAYPSVLGQTIRNSRWSSPDNVKPIYIVTPT
          10          20          30          40          50          60

gi|451 NASHIQSAVVCGRRHGVRIRVRSGGHDYEGLSYRSLQPEEFAVVDLSKMRVWVDGKART
          70          80          90         100         110         120

gi|451 AWVDSGAQLGELYAIIHKASTVLAFPAGVCPTIGVGGNFAGGGFGMLLRKYGIAAENVID
          130         140         150         160         170         180

gi|451 VKLVDANGTLHDKKSMGDDHFWAVRGGGESFSGIVVAWKVRLLPVPPTVTTFKIPKKASE
          190         200         210         220         230         240

```

```

          10          20          30
CV127. MKPVIIRTFYGLCMHMKKLSLQFPVCFGNLS
          . : . . . . : : :
gi|451 GAVDIINRWQVQVAPQLPDDLIRVIAQGPATFEAMYLGTCTLTTPMSSKFPELGMNAS
          250         260         270         280         290         300

```

CV127. FF

```

gi|451 HCNEMSWIQSIPFVHLGHRDNIEDDLLNRNNTFKPFAEYKSDYVYEPFPRVWEQIFSTW
          310         320         330         340         350         360

```


gi|451 LLKPGAGIMIFDPYGATISATPEWATPFPHRKGVLFNIIQYVNYWFAPGAGAAPLSWSKEI
370 380 390 400 410 420gi|451 YNYMEPYVSKNPRQAYANYRDIDLGRNEVVNDVSTFSSGLVWGQKYFKGNFQRLAITKKGK
430 440 450 460 470 480gi|451 VDPTDYFRNEQSIPPLIKKY
490 500

```
>>gi|45108973 gi|45108973|emb|CAF32567.2| unnamed protei  (500 aa)
  initn: 46 initl: 46 opt: 46 Z-score: 84.4 bits: 21.2 E(): 9
Smith-Waterman score: 46; 29.412% identity (64.706% similar) in 17 aa overlap
(9-25:277-293)
```

gi|451 YFPPPAAKEDFLGCLVKEIPPRLLYAKSSPAYPSVLGQTIRNSRWSSPDNVKPIYIVTPT
10 20 30 40 50 60gi | 451 NASHIQSAVVCGRRHGVRIRVRSGGHDYEGLSYRSLQPEEFAVVDLSKMRVWVDGKART
70 80 90 100 110 120gi|451 AWVDSGAQLGELYAIIHKASPVLAFAPAGVCPTIGVGGNFAGGGFGMLLRKYGIAAENVID
130 140 150 160 170 180gi|451 VKLVDANGTLHDKKSMGDDHFWAVRGGGGESFGIVVAVKVRLLPVPPTVTVFVKIPKKASE
190 200 210 220 230 240

CV127. MKPVIIRTFYGLCMHMKKLSLOFPVCFGNLS

gi|451 GAVDIINRWQVAPQLPDDLIRVIAQGPTATFEAMYLGTCQTLTPMSSSKFPGLGMNAS
250 260 270 280 290 300

CV127. FF

gi | 451 HCNEMSWIQSIPFVHLGHRDNIEDDLLNRNNTFKPFAEYKSDYVYEPFPKKEVWEQIFSTW
310 320 330 340 350 360

gi|451 LLKPGAGIMIFDPYGATISATPEWATPFPHRKGVLFNIIQYVNYWFAPGAGAAPLSWSKEI
370 380 390 400 410 420gi | 451 YNYMEPYVSKNPRQAYANYRDLGRNEVVNDVSTFSSGLVWGQKYFKGNFQRLAITKKGK
430 440 450 460 470 480gi|451 VDPTDYFRNEQSIPPLIKKY
490 500

```
>>gi|54144332 gi|54144332|emb|CAD54670.2| pollen allerge (508 aa)
  initn: 46 initl: 46 opt: 46 Z-score: 84.3 bits: 21.3 E(): 9.1
Smith-Waterman score: 46; 29.412% identity (64.706% similar) in 17 aa overlap
(9-25;285-301)
```

gi|541 SSCEVALSYIPTPLAKEDFLRCLVKEIPRLLLYAKSSPAYPSVLGQTIRNSRWSSPDNVK
10 20 30 40 50 60gi|541 PIYIVTPTNASHIQSAVVCGRRHGVRIRVRSGGHDYEGLSYRSLQPEEFVAVDL SKMRAV
70 80 90 100 110 120gi | 541 WVDGKARTAWVDSGAQLGELYAIHKASPVLAFFAGVCPTIGVGGNFAGGGFGMLLRKYG
130 140 150 160 170 180

gi|541 IAAENVIDVKLV DANGTLHDKKSMGDDHFWAVRGGGGESFGIVVAVKVRLLPVPPTVTVF
190 200 210 220 230 240

CV127. MKPVIIRTFYGLCMHVMKKLSLQF
10 20
.: :. :. :.
gi|541 KIPKKASEGAVDIINRWQV VAPQLPDDLIRVIAQGP TATFEAMYLGTCTLTPMMSSKF
250 260 270 280 290 300

CV127. PVCFGNLSFF
30
:
gi|541 PELGMNASHCNEMSWIQSIPFVHLGHRDNIEDDLLNRNNTFKPFAEYKSDYVYEPFPKEV
310 320 330 340 350 360

gi|541 WEQIFSTWLLKPGAGIMIFDPYGATISATPEWATPFPHRKGVLFNIQYVNYWFAPGAGAA
370 380 390 400 410 420

gi|541 PLSWSKEIYNYMEPYVSKNPRQAYANYRIDLGRNEVVNDVSTFSSGLVWGQKYFKGNFQ
430 440 450 460 470 480

gi|541 RLAITKGKVDPTDYFRNEQSIPPLIQKY
490 500

>>gi|10189816 gi|10189816|emb|CAC09235.1| unnamed protei (213 aa)
initn: 23 init1: 23 opt: 42 Z-score: 83.6 bits: 19.9 E(): 9.9
Smith-Waterman score: 42; 37.037% identity (59.259% similar) in 27 aa overlap
(6-32:159-183)

gi|101 MMKFLLLIAAVAFVAVSADPIHYDKITEEINKAIDDAIAAIEQSETIDPMKVPDHDADKFER
10 20 30 40 50 60

gi|101 HVGIVDFKGE LAMRNIEARGLKQMKRQGDANVKGE EGIVKAHLLIGVHDDIVSMEYDLAY
70 80 90 100 110 120

CV127. MKPVIIRTFYGLCMHVMKKLSLQFPVC
10 20
.: :. :. :.
gi|101 KLGLDHPTTHVISDIQDFVVALSLEISDEGNITMTSFEVRQFANVVNH-IGGLSILDPI-
130 140 150 160 170

CV127. FGNLSFF
30
.: :.
gi|101 FGVLSVDVLT AIFQD TVRKEMTKVLAPAFKRELEKN
180 190 200 210

>>gi|2498299 gi|2498299|sp|Q26456.1|ALL7_DERFA Mite alle (213 aa)
initn: 23 init1: 23 opt: 42 Z-score: 83.6 bits: 19.9 E(): 9.9
Smith-Waterman score: 42; 37.037% identity (59.259% similar) in 27 aa overlap
(6-32:159-183)

gi|249 MMKFLLLIAAVAFVAVSADPIHYDKITEEINKAIDDAIAAIEQSETIDPMKVPDHDADKFER
10 20 30 40 50 60

gi|249 HVGIVDFKGE LAMRNIEARGLKQMKRQGDANVKGE EGIVKAHLLIGVHDDIVSMEYDLAY
70 80 90 100 110 120

CV127. MKPVIIRTFYGLCMHVMKKLSLQFPVC
10 20
.: :. :. :.
gi|249 KLGLDHPTTHVISDIQDFVVALSLEISDEGNITMTSFEVRQFANVVNH-IGGLSILDPI-
130 140 150 160 170



The Chemical Company
Plant Science L. L. C.

BASF Reg. Doc. No. 2008/7019371

```

      30
CV127. FGNLSFF
      :: ::
gi|249 FGVLSVDLTAIFQDTRKEMTKVLAPAFKRELEKN
      180          190          200          210
```

```

34 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:39:44 2008 done: Sun Dec 14 22:39:44 2008
Total Scan time: 0.060 Total Display time: 0.000

Function used was FASTA [version 34.26.5 April 26, 2007]
```

CV127.3311_c3

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.3311_c3, 37 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 2.6434 \pm 0.00377$; $\mu = 10.0790 \pm 0.198$
mean_var=38.2772 \pm 10.748, 0's: 40 Z-trim: 41 B-trim: 18 in 1/41
Lambda= 0.207302

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2

join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:		opt	bits	E(1313)	%_id	%_sim	alen
gi 60280839	gi 60280839 gb AAX18312 (159)	61	24.6	0.31	0.500	0.667	18
gi 999009	gi 999009 gb AAB34785.1 (117)	53	22.0	1.4	0.320	0.600	25
gi 60280837	gi 60280837 gb AAX18311 (159)	50	21.3	3	0.467	0.667	15
gi 7435005	gi 7435005 pir A59055 p (134)	45	19.7	7.7	0.636	0.727	11
gi 47117012	gi 47117012 sp Q7M4I5.1 (134)	45	19.7	7.7	0.636	0.727	11
gi 24638082	gi 24638082 sp Q9BMK4.1 (134)	45	19.7	7.7	0.636	0.727	11
gi 24418862	gi 24418862 sp P00630.3 (167)	45	19.8	8.7	0.636	0.727	11

>>>CV127.3311_c3, 37 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|60280839 gi|60280839|gb|AAX18312.1| major allergen (159 aa)
initn: 61 initl: 61 opt: 61 Z-score: 110.6 bits: 24.6 E(): 0.31
Smith-Waterman score: 61; 50.000% identity (66.667% similar) in 18 aa overlap
(10-27:109-126)

gi|602 MGVTYSESESTSVIPPARLFNATALDGDGLIAKLAPQAVKSVEILEGGAGTIMKISFG
10 20 30 40 50 60

CV127. MLNRNTRYLNRRSLKRICTN
10 20
:: :::::

gi|602 ESSTYGYVKKRIDAIDKENFVYKYSMIEGDAISETIEKISYETMLVASNNGSIIKRTCHY
70 80 90 100 110 120

30
CV127. HTTRDLSCDCYKRNTI
:: ::

gi|602 HTKGDVEIKEEHLKAGKEKASQLLKLVENYLLEHQDAYN
130 140 150

>>gi|999009 gi|999009|gb|AAB34785.1| 68 kDa allergen [Pe (117 aa)
initn: 41 initl: 41 opt: 53 Z-score: 99.0 bits: 22.0 E(): 1.4
Smith-Waterman score: 53; 32.000% identity (60.000% similar) in 25 aa overlap
(13-37:94-115)

gi|999 GPYKTWQRIYDYDFLTNLTSSSEANDIIGAEAPLWSEQVDDVTVSSVFWPRAAALGELVWS
10 20 30 40 50 60

```

                                10      20      30
CV127.      MLRNRTRYLNNRSLKRICTNHTTRDLSCDCYKRNTI
                                ... : : : ... : : :
gi|999 GNRDAAGRKRRTTSFTQRIILNFREYLVANGVMAAALVPKYCLQHPH---ACDLYKNQTVMS
                                70      80      90      100      110

>>gi|60280837 gi|60280837|gb|AA18311.1| major allergen (159 aa)
  initn: 41 initl: 41 opt: 50 Z-score: 92.9 bits: 21.3 E(): 3
Smith-Waterman score: 50; 46.667% identity (66.667% similar) in 15 aa overlap
(13-27:112-126)

gi|602 MGVFTYESETTSVIPPARLFNATALDGDDELIAPQAVKSIEILEGGVGVTVQKIIFG
                                10      20      30      40      50      60

                                10      20
CV127.      MLRNRTRYLNNRSLKRICTN
                                ... : : :
gi|602 EGSTNGYVKKRIDVIDKDNFVYKYSMIEGDAISETIEKISYETTLVASGSGSIKRTCHY
                                70      80      90      100      110      120

                                30
CV127. HTTRDLSCDCYKRNTI
                                : : :
gi|602 HTKGDVEINEEHLKASKEKSSHLLKLVENYLLQDAYN
                                130      140      150

>>gi|7435005 gi|7435005|pir||A59055 phospholipase A2 (EC (134 aa)
  initn: 44 initl: 44 opt: 45 Z-score: 85.5 bits: 19.7 E(): 7.7
Smith-Waterman score: 45; 63.636% identity (72.727% similar) in 11 aa overlap
(21-31:55-63)

                                10      20
CV127.      MLRNRTRYLNNRSLKRICTNHTTRD
                                ... : : :
gi|743 IIYPGTLWCGHGNVSSGPNELGRFKHTDACCRTDHMCPCDVMSAGESKHGLTNTASHTR--
                                10      20      30      40      50

                                30
CV127. LSCDCYKRNTI
                                : : : :
gi|743 LSCDCDDTFYDCLKNSGEKISSYFVGKMYFNILDTKCYKLEHPVTGCGERTEGRCLRYTV
                                60      70      80      90      100      110

gi|743 DKSKPKVYQWFDLRKY
                                120      130

>>gi|47117012 gi|47117012|sp|Q7M4I5.1|PA2_APIDO Phosphol (134 aa)
  initn: 44 initl: 44 opt: 45 Z-score: 85.5 bits: 19.7 E(): 7.7
Smith-Waterman score: 45; 63.636% identity (72.727% similar) in 11 aa overlap
(21-31:55-63)

                                10      20
CV127.      MLRNRTRYLNNRSLKRICTNHTTRD
                                ... : : :
gi|471 IIYPGTLWCGHGNVSSSPDELGRFKHTDSCCRSHDMCPDVMSAGESKHGLTNTASHTR--
                                10      20      30      40      50

                                30
CV127. LSCDCYKRNTI
                                : : : :
gi|471 LSCDCDDKFDYDCLKNSSDTISSYFVGEMYFNILDTKCYKLEHPVTGCGKRTEGRCLNYTV
                                60      70      80      90      100      110

```

gi|471 DSKPKVYQWFDLRKY
120 130

>>gi|24638082 gi|24638082|sp|Q9BMK4.1|PA2_APICC Phosphol (134 aa)
initn: 44 initl: 44 opt: 45 Z-score: 85.5 bits: 19.7 E(): 7.7
Smith-Waterman score: 45; 63.636% identity (72.727% similar) in 11 aa overlap
(21-31:55-63)

CV127. MLNRNTRYLNNRSLKRICTNHTTRD
.::
gi|246 IIYPGTLWCGHGNVSSGPNELGRFKHTDACCRTHDMCPDVMSAGESKHGLTNTASHTR--
10 20 30 40 50

30
CV127. LSCDCYKRNTI
:::::
gi|246 LSCDCDDTFYDCLKNSGDKISSYFVGKMYFNLIDTKCYKLEHPVTGCGERTEGRCLRYTV
60 70 80 90 100 110

gi|246 DSKPKAYQWFDLRKY
120 130

>>gi|24418862 gi|24418862|sp|P00630.3|PA2_APIME Phosphol (167 aa)
initn: 44 initl: 44 opt: 45 Z-score: 84.6 bits: 19.8 E(): 8.7
Smith-Waterman score: 45; 63.636% identity (72.727% similar) in 11 aa overlap
(21-31:88-96)

gi|244 MQVVLGSLFLLLLSTSHGWQIRDRIGDNELEERIIYPGTLWCGHGNKSSGPNELGRFKHT
10 20 30 40 50 60
CV127. MLNRNTRYLNNRSLKRICTNHTTRDLSCDCYKRNTI
.:: :::::
gi|244 DACCRTHDMCPDVMSAGESKHGLTNTASHTR--LSCDCDDKFYDCLKNSADTISSYFVGK
70 80 90 100 110

gi|244 MYFNLIDTKCYKLEHPVTGCGERTEGRCLHYTVDKSKPKVYQWFDLRKY
120 130 140 150 160

37 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:20:24 2008 done: Sun Dec 14 22:20:24 2008
Total Scan time: 0.070 Total Display time: 0.010

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.2733_c2

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.2733_c2, 54 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 6.0761 \pm 0.00354$; $\mu = -7.8405 \pm 0.184$
mean_var=34.7071 \pm 7.506, 0's: 31 Z-trim: 31 B-trim: 5 in 1/43
Lambda= 0.217703

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:

	opt	bits	E(1313)	%_id	%_sim	alen
gi 56550550 gi 56550550 dbj BAD7793 (281)	52	22.8	2.8	0.308	0.615	26
gi 1587177 gi 1587177 prf 2206305A (75)	43	20.5	3.5	0.333	0.619	21
gi 2498604 gi 2498604 sp Q26464.1 M (75)	43	20.5	3.5	0.333	0.619	21
gi 1311510 gi 1311510 gb AAB36009.1 (15)	33	18.2	3.6	0.750	1.000	4
gi 2506771 gi 2506771 sp P16968.2 I (146)	45	20.9	5.4	0.429	0.714	14
gi 3608493 gi 3608493 gb AAC35986.1 (218)	46	21.0	7.4	0.250	0.607	28
gi 1311509 gi 1311509 gb AAB36008.1 (17)	30	17.2	8.3	0.375	0.750	8
gi 94400907 gi 94400907 ref NP_0010 (92)	40	19.5	8.8	0.625	0.875	8
gi 163825 gi 163825 gb AAC37318.1 (92)	40	19.5	8.8	0.214	0.714	14
gi 23894240 gi 23894240 emb CAD2361 (292)	47	21.2	8.8	0.556	1.000	9
gi 94400875 gi 94400875 ref NP_0010 (94)	40	19.5	9	0.625	0.875	8
gi 1911819 gi 1911819 gb AAB50883.1 (112)	41	19.7	9.1	0.462	0.769	13
gi 730091 gi 730091 sp Q07932.1 MYR (112)	41	19.7	9.1	0.462	0.769	13
gi 4538529 gi 4538529 emb CAB39376. (81)	39	19.3	9.2	0.667	0.778	9

>>>CV127.2733_c2, 54 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|56550550 gi|56550550|dbj|BAD77932.1| class IV chiti (281 aa)
initn: 27 initl: 27 opt: 52 Z-score: 93.4 bits: 22.8 E(): 2.8
Smith-Waterman score: 52; 30.769% identity (61.538% similar) in 26 aa overlap
(17-42:3-28)

	10	20	30	40	50
CV127.	MICWGLGRMGGR	IWRLVMVIWRRK	EMVWVRRRSK	MLLLLLPPPW	FRRKM
		..	:	..:	..:..:::
gi 565		MQIMATQNSKSN	IFWSSSASVVL	LLLLVDVGVC	QNCGCNGLCCSQ
		10	20	30	40
gi 565	YGYCGSGEAYC	GAGCKEGPCSS	SSPSTGTGVGS	IVSSDVFN	SIVGGAASGCAGNGFYTY
	50	60	70	80	90
gi 565	DSFISAANAFNG	FGTSGSSDVNK	REIAAFAFANA	AHETGGFCYIEE	QNPTSIYCDASNTQY
	110	120	130	140	150
gi 565	PCASGKTYHGR	GPLQLSWNYNY	GAGSYIQFDGL	NNPEIVGTD	STISFKTAVWFWMVNSN
	170	180	190	200	210
gi 565	CHTAITSGQG	FGATIRAINSM	ECDDGNAATV	ASRVNYYQKFC	QQLNVDTG
	230	240	250	260	270
					280

>>gi|1587177 gi|1587177|prf||2206305A major allergen Myr (75 aa)
initn: 35 initl: 35 opt: 43 Z-score: 91.8 bits: 20.5 E(): 3.5
Smith-Waterman score: 43; 33.333% identity (61.905% similar) in 21 aa overlap
(21-41:51-68)

```

                                10      20      30
CV127.                        MICWGLGRMGRIWRLVMVIWRRKEMVWWR
                                ...      : ..
gi|158 MKLSCLLLTLAIFVLTIVHAPNVEAKALADPESDAVGFDAGVGADPIDWKK---VDWK
                                10      20      30      40      50

                                40      50
CV127. RRSKMLLLLLLLLPPPWFRRKMSKR
. ::: ..:
gi|158 KVSCKTKVMLKACKFLG
                                60      70

```

>>gi|2498604 gi|2498604|sp|Q26464.1|MYR2_MYRPI Pilosulin (75 aa)
initn: 35 initl: 35 opt: 43 Z-score: 91.8 bits: 20.5 E(): 3.5
Smith-Waterman score: 43; 33.333% identity (61.905% similar) in 21 aa overlap
(21-41:51-68)

```

                                10      20      30
CV127.                        MICWGLGRMGRIWRLVMVIWRRKEMVWWR
                                ...      : ..
gi|249 MKLSCLLLTLAIFVLTIVHAPNVEAKALADPESDAVGFDAGVGADPIDWKK---VDWK
                                10      20      30      40      50

                                40      50
CV127. RRSKMLLLLLLLLPPPWFRRKMSKR
. ::: ..:
gi|249 KVSCKTKVMLKACKFLG
                                60      70

```

>>gi|1311510 gi|1311510|gb|AAB36009.1| mAb 8C7-reactive (15 aa)
initn: 33 initl: 33 opt: 33 Z-score: 91.4 bits: 18.2 E(): 3.6
Smith-Waterman score: 33; 75.000% identity (100.000% similar) in 4 aa overlap
(42-45:11-14)

```

                                10      20      30      40      50
CV127. MICWGLGRMGRIWRLVMVIWRRKEMVWWRRRSKMLLLLLLLLPPPWFRRKMSKR
                                .:::
gi|131                        APAGGVVVAAMPPPL
                                10

```

>>gi|2506771 gi|2506771|sp|P16968.2|IAA1_HORVU Alpha-amy (146 aa)
initn: 45 initl: 45 opt: 45 Z-score: 88.3 bits: 20.9 E(): 5.4
Smith-Waterman score: 45; 42.857% identity (71.429% similar) in 14 aa overlap
(2-15:68-81)

```

                                10      20      30      40      50      60
gi|250 PTSVAVDQGSMSVNSPGEWCWPGMGYPVYFPRCRALVKSQCAGGQVVESIQKDCCRQIA
                                10      20      30      40      50      60

                                10      20      30      40      50
CV127.                        MICWGLGRMGRIWRLVMVIWRRKEMVWWRRRSKMLLLLLLLLPPPWFRRKMSKR
                                :: :: : : ...
gi|250 AIGDEWCICGALGSMRGSMYKELGVALADDKATVAEVFPGCRTEVMDRAVASLPAVCNQY
                                70      80      90      100     110     120

gi|250 IPNTNGTDGVCYWLSYYQPPRQMSSR
                                130     140

```



```
>>gi|3608493 gi|3608493|gb|AAC35986.1| secreted protein (218 aa)
  initn: 29 initl: 29 opt: 46 Z-score: 85.9 bits: 21.0 E(): 7.4
Smith-Waterman score: 46; 25.000% identity (60.714% similar) in 28 aa overlap
(4-28:127-154)
```

gi|360 MLVLVPLLALLAVSVHGNSMRCGNNGMTDEARQKFLDVHNSYRSMVAKGQAKDAISGNAP
10 20 30 40 50 60gi|360 KAAKMKKMIYDCNVESTAMQNAKKCVFAHSHRKGVGGENIWMSTARQMDKAQAAQQASDGW
70 80 90 100 110 120

CV127. MICWGLG---RMGGRIWRLVMVIWRRKEMVWVRRRSKMLLLLLLPPPWFRRKMSKR

gi|360 FSELAKYGVGQENKLTQLWNRGVMIGHYTMVWQESYKLGCIYVEWCSSMTYGVCGYSPQ
130 140 150 160 170 180gi|360 GNMMNSLIYEKGNPCTKSDCGSNASC SAGEALCVVRG
190 200 210

```
>>gi|1311509 gi|1311509|gb|AAB36008.1| mAb 2F9-reactive (17 aa)
  initn: 27 initl: 27 opt: 30 Z-score: 85.0 bits: 17.2 E(): 8.3
Smith-Waterman score: 30; 37.500% identity (75.000% similar) in 8 aa overlap
(40-47:9-16)
```

CV127. MICWGLGRMGGRIVRLVMVIWRRKEMVWRRRSKMLLLLLPPPPWFRKMSKR
gi|131 APAGGVVPIIMPPXXFF

```
>>gi|94400907 gi|94400907|ref|NP_001035360.1| allergen A (92 aa)
  initn: 40 initl: 40 opt: 40 Z-score: 84.6 bits: 19.5 E(): 8.8
Smith-Waterman score: 40; 62.500% identity (87.500% similar) in 8 aa overlap
(5-12;27-34)
```

CV127.

MICWGLGRMGRIWRLVMVIWRRKEMVWWRSSKKMLL

10 20 30

⋮ ⋮ ⋮

gi | 944 MSRLVLASFLLLAIFSMVLGGFGGFGGGLGGRGKCPSEIFSRCDGRCQRFPCNVVVK
10 20 30 40 50 60

CV127. LLLLPPWFRRKMSKR

gi | 944 PLCIKICAPGCVCRLGYLRNKKKVCVPRSKCG
70 80 90

```
>>gi|163825 gi|163825|gb|AAC37318.1| major allergen I      (92 aa)
  initn: 38 initl: 38 opt: 40 Z-score: 84.6 bits: 19.5 E(): 8.8
Smith-Waterman score: 40; 21.429% identity (71.429% similar) in 14 aa overlap
(15-28:5-18)
```

CV127. MICWGLGRMGGRIVRLVMVIRRRKEMVWVRRRSKMLLLLLPPPWFRRKMSKR

gi | 163 MKGARVLVLLWAALLLIWGGNCEICPAVKRDVDLFLTGTPDEYVEQVAQY
10 20 30 40 50gi|163 KALPVVLENARILKNCVDAKMTEEDKENALSLLDKIYTSPLC
60 70 80 90

>>gi|23894240 gi|23894240|emb|CAD23613.1| trim 2 allerg (292 aa)
initn: 47 initl: 47 opt: 47 Z-score: 84.5 bits: 21.2 E(): 8.8
Smith-Waterman score: 47; 55.556% identity (100.000% similar) in 9 aa overlap
(4-12:32-40)

CV127. MICWGLGRMGRIWRLVMVIWRRKEMVWRRR
:::..:::
gi|238 NHDDVDYIEPDFVVRTSTNGTNLTRQENVPSWGLARVGSKKAGGTTYDSSAGKGV
10 20 30 40 50 60

CV127. SKKMLLLLLLLLPPWFRRKMSKR
gi|238 IIDTGIDINHEDFGGRKWKGNFVDKMDKNGHGHVAGTVGGTKYGLAKGVTLVAVKV
70 80 90 100 110 120
gi|238 LDCDGSNSGVIEGMEWAMREASGGNGTAKAAGKAVMNMSLGGPRSQASNDAAKAISD
130 140 150 160 170 180
gi|238 AGIFMAVAAGNENMDAQHSSPASEPSVCTVAASTEDDGKAEFSNYGAVVDVYAPGKDITS
190 200 210 220 230 240

gi|238 LKPGGSTDTLSGTSMASPHVCGLGAYLIGLGKQGGPGLCDTIKEMANEAIQR
250 260 270 280 290

>>gi|94400875 gi|94400875|ref|NP_001035358.1| allergen A (94 aa)
initn: 40 initl: 40 opt: 40 Z-score: 84.3 bits: 19.5 E(): 9
Smith-Waterman score: 40; 62.500% identity (87.500% similar) in 8 aa overlap
(5-12:27-34)

CV127. MICWGLGRMGRIWRLVMVIWRRKEMVWRRRSKKMLL
:::..:::
gi|944 MSRLVLASFLLAVFSLVGGFGGFGGLGGRGKCPSEIFSRCDGRCQRFPCPNVVPK
10 20 30 40 50 60

CV127. LLLLPPPWFRKMSKR
gi|944 PLCIKICAPGCVCRLLGYLRNKKKVCVPRSKCLPG
70 80 90

>>gi|1911819 gi|1911819|gb|AAB50883.1| Myr p I=allergeni (112 aa)
initn: 41 initl: 41 opt: 41 Z-score: 84.2 bits: 19.7 E(): 9.1
Smith-Waterman score: 41; 46.154% identity (76.923% similar) in 13 aa overlap
(5-17:57-69)

CV127. MICWGLGR
:::
gi|191 MKLSCLLLTLAIFVLTIVHAPNVEAKDLADPESEAVGFADAFGEADAVGEADPNAGLGS
10 20 30 40 50 60

CV127. MGRIWRLVMVIWRRKEMVWRRRSKKMLLLLLLLLPPPWFRKMSKR
..:::..
gi|191 VFGRLARILGRVIPKVAKKLGPKVAKVLPKVMKEAIPMAVEMAKSQEEQQPQ
70 80 90 100 110

>>gi|730091 gi|730091|sp|Q07932.1|MYR1_MYRPI Pilosulin-1 (112 aa)
initn: 41 initl: 41 opt: 41 Z-score: 84.2 bits: 19.7 E(): 9.1

Smith-Waterman score: 41; 46.154% identity (76.923% similar) in 13 aa overlap
(5-17:57-69)

```
CV127.                                     MICWGLGR
                                     :::
gi|730 MKLSCLLLTLTIIFVLTIVHAPNVEAKDLADPESEAVGFADAFGEADAVGEADPNAGLGS
          10          20          30          40          50          60
```

```
          10          20          30          40          50
CV127. MGGRIWRLVMVIWRRKEMVWRRRSKKMLLLLLPPPWFRKMSKR
      . ::. :..
gi|730 VFGRILARILGRVIPKVAKKLGPKVAKVLPKVMKEAIPMAVEMAKSQEEQQPQ
          70          80          90         100         110
```

>>gi|4538529 gi|4538529|emb|CAB39376.1| Cop c1 allergen (81 aa)
initn: 39 initl: 39 opt: 39 Z-score: 84.2 bits: 19.3 E(): 9.2
Smith-Waterman score: 39; 66.667% identity (77.778% similar) in 9 aa overlap
(36-44:54-62)

```
          10          20          30          40
CV127.             MICWGLGRMGRIWRLVMVIWRRKEMVWRRRSKKMLLLLL
                                     .:: : :
gi|453 RFLPSSSHLNPQHLPWLVHPAPVLLLPVLPQLKPVAPHLLLLPLDTTTLHMPPLLLQLQL
          10          20          30          40          50          60

          50
CV127. PPPWFRKMSKR
      ::
gi|453 PPLLSQGNPACSPKWLQLLVP
          70          80
```

54 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:12:05 2008 done: Sun Dec 14 22:12:05 2008
Total Scan time: 0.080 Total Display time: 0.010

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.3021_c2

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.3021_c2, 35 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 3.8720 \pm 0.00285$; $\mu = 3.1501 \pm 0.151$
mean_var=31.6686 \pm 8.405, 0's: 26 Z-trim: 26 B-trim: 191 in 1/42
Lambda= 0.227908

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:

gi	1304216	gi	1304216	dbj	BAA07772.1	(111)	opt	bits	E(1313)	%_id	%_sim	alen
gi	1304216	gi	1304216	dbj	BAA07772.1	(111)	48	21.5	1.7	0.368	0.684	19
gi	3914426	gi	3914426	sp	O24169.1	P	(134)	46	20.9	3.2	0.200	0.571	35
gi	3914428	gi	3914428	sp	O24171.1	P	(134)	46	20.9	3.2	0.200	0.571	35
gi	3914427	gi	3914427	sp	O24170.1	P	(134)	46	20.9	3.2	0.200	0.571	35
gi	1942360	gi	1942360	pdb	1CQA	Bi	(133)	44	20.2	5	0.200	0.571	35
gi	130975	gi	130975	sp	P25816.1	PRO	(133)	44	20.2	5	0.200	0.571	35
gi	14422363	gi	14422363	emb	CAC4163	(131)	43	19.9	6.2	0.400	0.667	15
gi	14347293	gi	14347293	emb	CAC4120	(207)	44	20.3	7.4	0.500	0.688	16
gi	60418858	gi	60418858	gb	AAX19856	(131)	42	19.6	7.8	0.229	0.543	35
gi	60418854	gi	60418854	gb	AAX19854	(131)	42	19.6	7.8	0.229	0.543	35
gi	14423860	gi	14423860	sp	Q9M7N0.1	(131)	42	19.6	7.8	0.200	0.543	35
gi	14423856	gi	14423856	sp	Q9LEI8.1	(131)	42	19.6	7.8	0.200	0.543	35
gi	14423858	gi	14423858	sp	Q9M7M8.1	(131)	42	19.6	7.8	0.200	0.543	35
gi	14423873	gi	14423873	sp	Q9XF40.1	(131)	42	19.6	7.8	0.229	0.543	35
gi	28881453	gi	28881453	emb	CAD4655	(131)	42	19.6	7.8	0.229	0.543	35
gi	18652049	gi	18652049	gb	AAL76933	(134)	42	19.6	7.9	0.200	0.571	35
gi	1168696	gi	1168696	sp	P43187.1	A	(205)	43	20.0	9.2	0.500	0.786	14

>>>CV127.3021_c2, 35 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|1304216 gi|1304216|dbj|BAA07772.1| allergenic prote (111 aa)

initn: 39 initl: 39 opt: 48 Z-score: 97.3 bits: 21.5 E(): 1.7

Smith-Waterman score: 48; 36.842% identity (68.421% similar) in 19 aa overlap (12-30:72-90)

gi|130 RHEVKRQCVATTHPAAPGTEQLRQDCCRQLAAVDDSWCRCSALNHMVGGIYRELGATDVG

10 20 30 40 50 60

10 20 30

CV127. MHLQGKRIRFLRLDVLKLRGYPQLCGVDLERIGR

: : . . : : . . : : . .

gi|130 HPMAEVFPGCRGRGDLERRRRSLPAFCNVDPNGTGGVCYWLGYPRTPRTGH

70 80 90 100 110

>>gi|3914426 gi|3914426|sp|O24169.1|PROF1_OLEEU Profilin (134 aa)

initn: 30 initl: 30 opt: 46 Z-score: 92.4 bits: 20.9 E(): 3.2

Smith-Waterman score: 46; 20.000% identity (57.143% similar) in 35 aa overlap (1-35:68-102)

gi|391 MSWQAYVDDHLMCDIEGHEDHRLTAAAIIVGHDGSVWAQSATFPQFKPEEMNGIMTDFNEP

10 20 30 40 50 60

```

              10      20      30
CV127.      MHLQGKRIRFLRLDVLKLRRGYQRLCGVDLERIGR
              .:: :. . . . . :. . :. . . . :.
gi|391 GH LAPTGLHLGGTKYMQEAGAVIRGKKGSGGITIKKTGQALVFGIYEETPTPGQCNM
              70      80      90      100      110      120

gi|391 VVERLGDYLVLEQGM
              130

```

>>gi|3914428 gi|3914428|sp|O24171.1|PROF3_OLEEU Profilin (134 aa)
 initn: 30 initl: 30 opt: 46 Z-score: 92.4 bits: 20.9 E(): 3.2
 Smith-Waterman score: 46; 20.000% identity (57.143% similar) in 35 aa overlap
 (1-35:68-102)

```

gi|391 MSWQAYVDDHLMCDIEGHEGHRLTAAAI VGH DGSVWAQSATFPQFKPEEMNGIMTDFNEP
              10      20      30      40      50      60

              10      20      30
CV127.      MHLQGKRIRFLRLDVLKLRRGYQRLCGVDLERIGR
              .:: :. . . . . :. . :. . . . :.
gi|391 GH LAPTGLHLGGTKYMQEAGAVIRGKKGSGGITIKKTGQALVFGIYEETPTPGQCNM
              70      80      90      100      110      120

gi|391 VAERLGDYLLLEQGL
              130

```

>>gi|3914427 gi|3914427|sp|O24170.1|PROF2_OLEEU Profilin (134 aa)
 initn: 30 initl: 30 opt: 46 Z-score: 92.4 bits: 20.9 E(): 3.2
 Smith-Waterman score: 46; 20.000% identity (57.143% similar) in 35 aa overlap
 (1-35:68-102)

```

gi|391 MSWQAYVDDHLMCDIEGHEGHRLTAAAI VGH DGSVWAQSATFPQFKPEEMNGIMTDFNEP
              10      20      30      40      50      60

              10      20      30
CV127.      MHLQGKRIRFLRLDVLKLRRGYQRLCGVDLERIGR
              .:: :. . . . . :. . :. . . . :.
gi|391 GH LAPTGLHLGGTKYMQEAGAVIRGKKGSGGITIKKTGQALVFGIYEETPTPGQCNM
              70      80      90      100      110      120

gi|391 VVERLGDYLLLEQGL
              130

```

>>gi|1942360 gi|1942360|pdb|1CQA| Birch Pollen Profilin (133 aa)
 initn: 28 initl: 28 opt: 44 Z-score: 88.9 bits: 20.2 E(): 5
 Smith-Waterman score: 44; 20.000% identity (57.143% similar) in 35 aa overlap
 (1-35:67-101)

```

gi|194 MSWQTYVDEHLMCDIDGQGEELAASAI VGH DGSVWAQSSSFPQFKPQEITGIMKDFEETPG
              10      20      30      40      50      60

              10      20      30
CV127.      MHLQGKRIRFLRLDVLKLRRGYQRLCGVDLERIGR
              .:: :. . . . . :. . :. . . . :.
gi|194 HLAPTGLHLGGIKYMQEAGAVIRGKKGSGGITIKKTGQALVFGIYEETPTPGQCNMV
              70      80      90      100      110      120

gi|194 VERLGDYLLIDQGL
              130

```

>>gi|130975 gi|130975|sp|P25816.1|PROF_BETVE Profilin (P (133 aa)
initn: 28 initl: 28 opt: 44 Z-score: 88.9 bits: 20.2 E(): 5
Smith-Waterman score: 44; 20.000% identity (57.143% similar) in 35 aa overlap
(1-35:67-101)

gi|130 MSWQTYVDEHLMCDIDGQASNSLASAIVGHGDSVWAQSSSFPQFKPQEITGIMKDFEEPG
10 20 30 40 50 60

CV127. MHLQGKRIRFLRLDVLKLRGGYQRLCGVDLERIGR
::: : :: . :.

gi|130 HLAPTGLHLGGIKYMWIQGEAGAVIRGKKGSGGITIKKTGQALVFGIYEPPVTPGQCNMV
70 80 90 100 110 120

gi|130 VERLGDYLIDQGL
130

>>gi|14422363 gi|14422363|emb|CAC41635.1| plantain polle (131 aa)
initn: 43 initl: 43 opt: 43 Z-score: 87.3 bits: 19.9 E(): 6.2
Smith-Waterman score: 43; 40.000% identity (66.667% similar) in 15 aa overlap
(21-35:69-83)

CV127. MHLQGKRIRFLR
10

gi|144 TQTSHPAKFHVEGEVYCNVCHSRNLINELSERMAGAQQVQLDCKDDSKKVIYSIGGETGQD
10 20 30 40 50 60

CV127. LDVLKLRGGYQRLCGVDLERIGR
: . . . : . : . : .

gi|144 GVYRLPVVGYHEDCEIKLVKSGRPDCSEIPKLAKGTIQTSKVDLSKNTTITEKTRHVKPL
70 80 90 100 110 120

gi|144 SFRAKTDAPGC
130

>>gi|14347293 gi|14347293|emb|CAC41202.1| unnamed protei (207 aa)
initn: 38 initl: 38 opt: 44 Z-score: 85.9 bits: 20.3 E(): 7.4
Smith-Waterman score: 44; 50.000% identity (68.750% similar) in 16 aa overlap
(20-35:143-157)

gi|143 MAKLTILVALALFLLAAHASARQQWELQGDRCQSQLERANLRPCEQHLMQKIQRDEDSY
10 20 30 40 50 60

gi|143 ERDPYSPSQDPYSPSPYDRRGAGSSQHQRCCNELNEFENNQRCMCEALQQIMENQSDRL
70 80 90 100 110 120

CV127. MHLQGKRIRFLRLDVLKLRGGYQRLCGVDLERIGR
: . :: : . :: : . ::

gi|143 QGRQQEQQFKRELRLNLPQQCGLRAPQR-CDLDVESGGRRPRIPPIITGSRSRRHQSPYGN
130 140 150 160 170

gi|143 RRYSAMCLLPRAADGDGWFPSVAVDCSG
180 190 200

>>gi|60418858 gi|60418858|gb|AA19856.1| profilin 1 [Mal (131 aa)
initn: 39 initl: 39 opt: 42 Z-score: 85.5 bits: 19.6 E(): 7.8
Smith-Waterman score: 42; 22.857% identity (54.286% similar) in 35 aa overlap
(1-35:65-99)

gi|604 MSWQAYVDDHLMCDIDGHHLTAAAILGHGDSVWAQSSSTFPKFKPEEITAIMKDFDEPGSL
10 20 30 40 50 60

CV127. MHLQGKRIRFLRLDVLKLRGYQRLCGVDLERIGR
.: : : : : :

gi|604 APTGLHLGGTKYMVIQEGGAVIRGKKGSGGVTVKKTGQDLVFGIYEEPLTPGQCNMIVE
70 80 90 100 110 120

gi|604 RLGDYLIDQGL
130

>>gi|60418854 gi|60418854|gb|AA19854.1| profilin 1 [Mal (131 aa)
initn: 39 initl: 39 opt: 42 Z-score: 85.5 bits: 19.6 E(): 7.8
Smith-Waterman score: 42; 22.857% identity (54.286% similar) in 35 aa overlap
(1-35:65-99)

gi|604 MSWQAYVDDHLMCDIDGHHLTAAAILGHGDSVWAHSSTFPKFKPEEITAIMKDFDEPGSL
10 20 30 40 50 60

CV127. MHLQGKRIRFLRLDVLKLRGYQRLCGVDLERIGR
.: : : : : :

gi|604 APTGLHLGGTKYMVIQEGGAVIRGKKGSGGVTVKKTGQALVFGIYEETLTPGQCNMIVE
70 80 90 100 110 120

gi|604 RLGDYLIDQGL
130

>>gi|14423860 gi|14423860|sp|Q9M7N0.1|PROF3_HEVBR Profil (131 aa)
initn: 27 initl: 27 opt: 42 Z-score: 85.5 bits: 19.6 E(): 7.8
Smith-Waterman score: 42; 20.000% identity (54.286% similar) in 35 aa overlap
(1-35:65-99)

gi|144 MSWQTYVDEHLMCDIDGHHLTAAAIIGHGDSVWAQSSSFPQFKPEEVAAIMKDFDEPGSL
10 20 30 40 50 60

CV127. MHLQGKRIRFLRLDVLKLRGYQRLCGVDLERIGR
.: : : : : :

gi|144 APTGLHLGGTKYMVIQGEPGAVIRGKKGSGGITVKKTGQALIIGIYDEPLTPGQCNMIVE
70 80 90 100 110 120

gi|144 RLGDYLLDQGM
130

>>gi|14423856 gi|14423856|sp|Q9LEI8.1|PROF6_HEVBR Profil (131 aa)
initn: 27 initl: 27 opt: 42 Z-score: 85.5 bits: 19.6 E(): 7.8
Smith-Waterman score: 42; 20.000% identity (54.286% similar) in 35 aa overlap
(1-35:65-99)

gi|144 MSWQTYVDDHLMCDIDGHHLTAAAIIGHGDSVWAQSSSFPQFKSDEVAAVMKDFDEPGSL
10 20 30 40 50 60

CV127. MHLQGKRIRFLRLDVLKLRGYQRLCGVDLERIGR
.: : : : : :

gi|144 APTGLHLGGTKYMVIQGEPGAVIRGKKGSGGITVKKTGQALIIGIYDEPLTPGQCNMIVE
70 80 90 100 110 120

gi|144 RLGDYLLDQGL
130

>>gi|14423858 gi|14423858|sp|Q9M7M8.1|PROF5_HEVBR Profil (131 aa)
initn: 27 initl: 27 opt: 42 Z-score: 85.5 bits: 19.6 E(): 7.8
Smith-Waterman score: 42; 20.000% identity (54.286% similar) in 35 aa overlap
(1-35:65-99)

gi|144 MSWQTYVDDHLMCDIDGHRLTAAAIIGHDGSVWAQSSGFPQFKSDEVAAVMKDFDEPGSL
10 20 30 40 50 60

CV127. MHLQGKRIRFLRLDVLKLRGGYQRLCGVDLERIGR
10 20 30
.: : : : . :

gi|144 APTGLHLGGTKYQVQGEPAVIRGKKGSGGITVKKTGQALIIIGIYDEPLTPGQCNMIVE
70 80 90 100 110 120

gi|144 RLGDYLLLEQGM
130

>>gi|14423873 gi|14423873|sp|Q9XF40.1|PROF1_MALDO Profil (131 aa)
initn: 39 initl: 39 opt: 42 Z-score: 85.5 bits: 19.6 E(): 7.8
Smith-Waterman score: 42; 22.857% identity (54.286% similar) in 35 aa overlap
(1-35:65-99)

gi|144 MSWQAYVDDRLMCDIDGHHLTAAAILGHGDSVWAHSSTFPKFKPEEITAIMKDFDEPGSL
10 20 30 40 50 60

CV127. MHLQGKRIRFLRLDVLKLRGGYQRLCGVDLERIGR
10 20 30
.: : : : . :

gi|144 APTGLHLGGTKYQVQGEPAVIRGKKGSGGVTVKKTGQALVFGIYEEPLTPGQCNMIVE
70 80 90 100 110 120

gi|144 RLGDYLLIDQGL
130

>>gi|28881453 gi|28881453|emb|CAD46559.1|profilin [Malu (131 aa)
initn: 39 initl: 39 opt: 42 Z-score: 85.5 bits: 19.6 E(): 7.8
Smith-Waterman score: 42; 22.857% identity (54.286% similar) in 35 aa overlap
(1-35:65-99)

gi|288 MSWQAYVDDHLMCDIDGHHLTAAAILGHGDSVWAHSSTFPKFKPEEITAIMKDFDEPGSL
10 20 30 40 50 60

CV127. MHLQGKRIRFLRLDVLKLRGGYQRLCGVDLERIGR
10 20 30
.: : : : . :

gi|288 APTGLHLGGTKYQVQGEPAVIRGKKGSGGVTVKKTGQALVFGIYEEPLTPGQCNMIVE
70 80 90 100 110 120

gi|288 RLGDYLLIDQGL
130

>>gi|18652049 gi|18652049|gb|AAL76933.1|AF456482_1 minor (134 aa)
initn: 32 initl: 32 opt: 42 Z-score: 85.3 bits: 19.6 E(): 7.9
Smith-Waterman score: 42; 20.000% identity (57.143% similar) in 35 aa overlap
(1-35:68-102)

gi|186 MSWQTYVDDHLMCEVDGNPGQQLSAAAIIGHDGSVWAQSSSTFPKFKPEEITGIMKNFDEP
10 20 30 40 50 60


```

              10      20      30
CV127.      MHLQGKRIRFLRLDVLKLRGYQRLCGVDLERIGR
              ..: : . . . . . : : . : : . . . : .
gi|186 GH LAPTGLYLGGTKYMWIQGEPIAVIRGKKGSGGVTIKKTGQALVFGVYDEPVTGQC�NL
              70      80      90      100      110      120

gi|186 IVERLGDYDLIEQGL
              130

>>gi|1168696 gi|1168696|sp|P43187.1|ALLB3_BETVE Calcium- (205 aa)
  initn: 34 initl: 34 opt: 43 Z-score: 84.2 bits: 20.0 E(): 9.2
Smith-Waterman score: 43; 50.000% identity (78.571% similar) in 14 aa overlap
(7-20:30-43)

```

```

              10      20      30
CV127.      MHLQGKRIRFLRLDVLKLRGYQRLCGVDLERIGR
              :. : . . . . . : : : : :
gi|116 MPCSTEAMEKAGHGHASTPRKRSLSNSSFRLRSESLNTRLRLRRIFDLFDKNSDGIITVDE
              10      20      30      40      50      60

gi|116 LSRALNLLGLETDLSELESTVKSFTREGNIGLQFEDFISLHQSLNDSYFAYGGEDEDDNE
              70      80      90      100      110      120

gi|116 EDMRKSILSQEEADSFGGFKVFDEDDGDGYISARELQMVLGKLGFSSEGSEIDRVEKMIVSV
              130      140      150      160      170      180

gi|116 DSNRDGRVDFFEFKDMMRSVLVRSS
              190      200

```

```

35 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:18:26 2008 done: Sun Dec 14 22:18:26 2008
Total Scan time: 0.080 Total Display time: 0.020

```

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.3885_d3

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.3885_d3, 47 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 3.6113 \pm 0.00303$; $\mu = 4.9944 \pm 0.157$

mean_var=30.9678 \pm 7.690, 0's: 27 Z-trim: 28 B-trim: 0 in 0/43

Lambda= 0.230473

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2

join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:		opt	bits	E(1313)	%_id	%_sim	alen
gi 76782249	gi 76782249 gb ABA54898 (119)	52	23.2	0.78	0.500	0.833	12
gi 1588669	gi 1588669 prf 2209273A (170)	50	22.6	1.7	0.556	1.000	9
gi 21413	gi 21413 emb CAA45723.1 a (217)	46	21.3	5.1	0.625	1.000	8
gi 20141714	gi 20141714 sp P30941.2 (221)	46	21.3	5.2	0.625	1.000	8
gi 169971	gi 169971 gb AAA33965.1 (240)	44	20.7	8.8	0.353	0.765	17
gi 25991543	gi 25991543 gb AAN76862 (457)	46	21.5	9.5	0.389	0.667	18

>>>CV127.3885_d3, 47 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|76782249 gi|76782249|gb|ABA54898.1| hydrophobic see (119 aa)

initn: 52 initl: 52 opt: 52 Z-score: 103.5 bits: 23.2 E(): 0.78

Smith-Waterman score: 52; 50.000% identity (83.333% similar) in 12 aa overlap (27-38:74-85)

					10
CV127.					MIVSRVSLRLLLV

gi 767	MGSKVVASVALLLSINILFISMVSSSSHYDPQPQPSHVTALITRPSCPDLSICLNILGGS
	10 20 30 40 50 60

	20	30	40
CV127.	GLRLFILILTRLRLGRIRLLMCLCVVMSWLCKG		

gi 767	LGTVDCCALIGGLGDIEAIVCLCIQLRALGILNLNRNLQLILNSCGRSYPSNATCPRT				
	70 80 90 100 110				

>>gi|1588669 gi|1588669|prf||2209273A Zml3 (170 aa)

initn: 44 initl: 44 opt: 50 Z-score: 97.5 bits: 22.6 E(): 1.7

Smith-Waterman score: 50; 55.556% identity (100.000% similar) in 9 aa overlap (32-40:14-22)

	10	20	30	40
CV127.	MIVSRVSLRLLLVGLRLFILILTRLRLGRIRLLMCLCVVMSWLCKG			

gi 158	MASVPAPATTTAAVILCLCVVLSCAAADDPNLPDYVIQGRVY				
	10 20 30 40				

gi 158	CDTCRAGFVTNVTEYIAGAKVRLECKHFGTGKLERAI DGVT DATGT YTI ELKDSHEEDIC
	50 60 70 80 90 100

gi 158	QVVLVASPRKDCDEVQALRDRAGVLLTRNVGISDSL RPANPLGYFKDVPLPVCAALLKQL
	110 120 130 140 150 160

gi|158 DSDDDDQ
170

>>gi|21413 gi|21413|emb|CAA45723.1| aspartic proteinase (217 aa)
initn: 37 initl: 37 opt: 46 Z-score: 88.8 bits: 21.3 E(): 5.1
Smith-Waterman score: 46; 62.500% identity (100.000% similar) in 8 aa overlap
(32-39:5-12)

```

      10      20      30      40
CV127. MIVSRVSLRLLLVGLRFLFILILTRLRLGRIRLLMCLCVVMSWLCKG
      . . . . .
gi|214      MKCLFLLCLCLVPIVVFSSFTTSKNPINLPSDA
      10      20      30

gi|214 TPVLDVAGKELDSRLSYRIISTFWGALGGDVYLGKSPNSDAPCANGIFRYNSDVGPSGTP
      40      50      60      70      80      90

gi|214 VRFSHFGQGIFENELLNIQFAISTSKLCVSYTIWKVGDYDASLGTMLLETGGTIGQADSS
      100      110      120      130      140      150

gi|214 WFKIVKSSQFGYNLLYCPVTSTMSCPFSDDQFCLKVGVVHQNGKRRLALVKDNPLDVSF
      160      170      180      190      200      210

gi|214 KQVQ

```

>>gi|20141714 gi|20141714|sp|P30941.2|SPI7_SOLTU Serine (221 aa)
initn: 37 initl: 37 opt: 46 Z-score: 88.7 bits: 21.3 E(): 5.2
Smith-Waterman score: 46; 62.500% identity (100.000% similar) in 8 aa overlap
(32-39:5-12)

```

      10      20      30      40
CV127. MIVSRVSLRLLLVGLRFLFILILTRLRLGRIRLLMCLCVVMSWLCKG
      . . . . .
gi|201      MKCLFLLCLCLVPIVVFSSFTTSKNPINLPSDA
      10      20      30

gi|201 TPVLDVAGKELDSRLSYRIISTFWGALGGDVYLGKSPNSDAPCANGIFRYNSDVGPSGTP
      40      50      60      70      80      90

gi|201 VRFIGSSSHFGQGIFENELLNIQFAISTSKLCVSYTIWKVGDYDASLGTMLLETGGTIGQ
      100      110      120      130      140      150

gi|201 ADSSWFKIVKSSQFGYNLLYCPVTSTMSCPFSDDQFCLKVGVVHQNGKRRLALVKDNPL
      160      170      180      190      200      210

gi|201 DVSFKQVQ
      220

```

>>gi|169971 gi|169971|gb|AAA33965.1| glycinin precursor (240 aa)
initn: 39 initl: 39 opt: 44 Z-score: 84.5 bits: 20.7 E(): 8.8
Smith-Waterman score: 44; 35.294% identity (76.471% similar) in 17 aa overlap
(19-35:138-153)

```

gi|169 QEDEDDEDEEYEQTPSYPPRRPSHGKHEDDEDEDEEHQPRPDHPPQRPSRPEQQEPRG
      10      20      30      40      50      60

CV127.
      M

gi|169 RGCQTRNGVEENICTMKLHENIARPSRADFYNPAGRISTLNSLTLPALRQFGLSAQYLV
      70      80      90      100      110      120

```

```

              10          20          30          40
CV127.  IVSRVSLRLLLVLRLFILILTRLRLGRIRLLMCLCVVMLSWLCKG
              .. .:: . ::... :
gi|169  LYRNGIYSPHWNLNANSVIYVTRGK-GRVRVVNCQGNPVFDGDLTRGQLLLVPQNFVVAD
              130          140          150          160          170

gi|169  QGGKQGLEIYVFKTQHNAVSSYIKDLFRAIPSEVLSNSYNLGQSQVRQLKYQGNSGPLLN
              180          190          200          210          220          230

gi|169  P
              240

>>gi|25991543 gi|25991543|gb|AAN76862.1|AF453947_1 aller (457 aa)
  initn: 46 initl: 46 opt: 46 Z-score: 83.9 bits: 21.5 E(): 9.5
Smith-Waterman score: 46; 38.889% identity (66.667% similar) in 18 aa overlap
(27-44:299-316)

gi|259  LSVCFILILFHGCLASRQEWQQQDECQIDRLDALEPDNRVEYEAGTVEAWDPNHEQFRCAG
              10          20          30          40          50          60

gi|259  VALVRHTIQPNGLLLPQYSNAPQLIYVVGEGMTGISYPGCPETYQAPQQGRQQQSGGRF
              70          80          90          100          110          120

gi|259  QDRHQKIRRFRRGDIIAIPAGVAHWCYNEGNSPVVTVTLDDVSNSQNQLDRTPRKFHLAG
              130          140          150          160          170          180

gi|259  NPKDVFQQQQQHQSRRNLFSGFDTELLAEAFQVDERLIKQLKSEDNRGGIVKVKDDELRL
              190          200          210          220          230          240

              10          20
CV127.  MIVSRVSLRLLLVLRLFILILTRLRLG
              .:
gi|259  VIRPSRSQSERGSESEEESEDEKRRWGQRDNGIEETICTMRLKENINDPARADIYTPEVG
              250          260          270          280          290          300

              30          40
CV127.  RIRLLMCLCVVMLSWLCKG
              :. : : . :.:
gi|259  RLTTLSLNLPLKWLQLSVEKGVLYKNALVLPHWNLNSHSIIYGCKGKGQVQVVDNFGN
              310          320          330          340          350          360

gi|259  RVFDGEVREGQMLVVPQNFAVVKRAREERFEWISFKTNDRAMTSPLAGRTSVLGGMPEEV
              370          380          390          400          410          420

gi|259  LANAFQISREDARKIKFNNQQTTLTSGESSHMRDDA
              430          440          450
```

47 residues in 1 query sequences

292011 residues in 1313 library sequences

Scomplib [34.26]

start: Sun Dec 14 22:24:19 2008 done: Sun Dec 14 22:24:20 2008

Total Scan time: 0.070 Total Display time: 0.000

Function used was FASTA [version 34.26.5 April 26, 2007]

CV127.5811_d3

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007

Please cite:

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

CV127.5811_d3, 47 aa

vs /n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

292011 residues in 1313 sequences

Expectation_n fit: $\rho(\ln(x)) = 3.6113 \pm 0.00303$; $\mu = 4.9944 \pm 0.157$
mean_var=30.9678 \pm 7.690, 0's: 27 Z-trim: 28 B-trim: 0 in 0/43
Lambda= 0.230473

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2

join: 36, opt: 20, open/ext: -10/-2, width: 16

The best scores are:		opt	bits	E(1313)	%_id	%_sim	alen
gi 76782249	gi 76782249 gb ABA54898 (119)	52	23.2	0.78	0.500	0.833	12
gi 1588669	gi 1588669 prf 2209273A (170)	50	22.6	1.7	0.556	1.000	9
gi 21413	gi 21413 emb CAA45723.1 a (217)	46	21.3	5.1	0.625	1.000	8
gi 20141714	gi 20141714 sp P30941.2 (221)	46	21.3	5.2	0.625	1.000	8
gi 169971	gi 169971 gb AAA33965.1 (240)	44	20.7	8.8	0.353	0.765	17
gi 25991543	gi 25991543 gb AAN76862 (457)	46	21.5	9.5	0.389	0.667	18

>>>CV127.5811_d3, 47 aa vs

/n/na4/bioinfo/refsets/PUB_AllergenOnline12Dec2007_V8.fasta library

>>gi|76782249 gi|76782249|gb|ABA54898.1| hydrophobic see (119 aa)
initn: 52 initl: 52 opt: 52 Z-score: 103.5 bits: 23.2 E(): 0.78
Smith-Waterman score: 52; 50.000% identity (83.333% similar) in 12 aa overlap
(27-38:74-85)

					10
CV127.					MIVSRVSLRLLLV

gi 767	MGSKVVASVALLLSINILFISMVSSSSHYDPQPQPSHVTALITRPSCPDLSICLNILGGS
	10 20 30 40 50 60

	20	30	40
CV127.	GLRLFILILTRLRLGRIRLLMCLCVVMSWLCKG		
	:: :: ..:::		

gi 767	LGTVDCCALIGGLGDIEAIVCLCIQLRALGILNLNRNLQLILNSCGRSYPSNATCPRT
	70 80 90 100 110

>>gi|1588669 gi|1588669|prf||2209273A Zml3 (170 aa)
initn: 44 initl: 44 opt: 50 Z-score: 97.5 bits: 22.6 E(): 1.7
Smith-Waterman score: 50; 55.556% identity (100.000% similar) in 9 aa overlap
(32-40:14-22)

	10	20	30	40
CV127.	MIVSRVSLRLLLVGLRLFILILTRLRLGRIRLLMCLCVVMSWLCKG			
			
gi 158	MASVPAPATTTAAVILCLCVVLSCAAADDPNLPDYVIQGRVY			
	10 20 30 40			

gi 158	CDTCRAGFVTNVTEYIAGAKVRLECKHFGTGKLERAI DGVT DATGT YTI ELKDSHEEDIC
	50 60 70 80 90 100

gi 158	QVVLVASPRKDCDEVQALRDRAGVLLTRNVGISDSL RPANPLGYFKDVPLPVCAALLKQL
	110 120 130 140 150 160

gi|158 DSDDDDQ
170

>>gi|21413 gi|21413|emb|CAA45723.1| aspartic proteinase (217 aa)
initn: 37 initl: 37 opt: 46 Z-score: 88.8 bits: 21.3 E(): 5.1
Smith-Waterman score: 46; 62.500% identity (100.000% similar) in 8 aa overlap
(32-39:5-12)

```

      10      20      30      40
CV127. MIVSRVSLRLLLVGLRFLFILILTRLRLGRIRLLMCLCVVMSWLCKG
      . . . . .
gi|214      MKCLFLLCLCLVPIVVFSSFTTSKNPINLPSDA
      10      20      30

gi|214 TPVLDVAGKELDSRLSYRIISTFWGALGGDVYLGKSPNSDAPCANGIFRYNSDVGPSGTP
      40      50      60      70      80      90

gi|214 VRFSHFQGQIFENELLNIQFAISTSKLCVSYTIWKVGDYDASLGTMLLETGGTIGQADSS
      100      110      120      130      140      150

gi|214 WFKIVKSSQFGYNLLYCPVTSTMSCPFSSDDQFCLKVGVVHQNGKRRLALVKDNPLDVSF
      160      170      180      190      200      210

gi|214 KQVQ

```

>>gi|20141714 gi|20141714|sp|P30941.2|SPI7_SOLTU Serine (221 aa)
initn: 37 initl: 37 opt: 46 Z-score: 88.7 bits: 21.3 E(): 5.2
Smith-Waterman score: 46; 62.500% identity (100.000% similar) in 8 aa overlap
(32-39:5-12)

```

      10      20      30      40
CV127. MIVSRVSLRLLLVGLRFLFILILTRLRLGRIRLLMCLCVVMSWLCKG
      . . . . .
gi|201      MKCLFLLCLCLVPIVVFSSFTTSKNPINLPSDA
      10      20      30

gi|201 TPVLDVAGKELDSRLSYRIISTFWGALGGDVYLGKSPNSDAPCANGIFRYNSDVGPSGTP
      40      50      60      70      80      90

gi|201 VRFIGSSSHFGQIFENELLNIQFAISTSKLCVSYTIWKVGDYDASLGTMLLETGGTIGQ
      100      110      120      130      140      150

gi|201 ADSSWFKIVKSSQFGYNLLYCPVTSTMSCPFSSDDQFCLKVGVVHQNGKRRLALVKDNPL
      160      170      180      190      200      210

gi|201 DVSFKQVQ
      220

```

>>gi|169971 gi|169971|gb|AAA33965.1| glycinin precursor (240 aa)
initn: 39 initl: 39 opt: 44 Z-score: 84.5 bits: 20.7 E(): 8.8
Smith-Waterman score: 44; 35.294% identity (76.471% similar) in 17 aa overlap
(19-35:138-153)

```

gi|169 QEDEDEDEDEEYEQTPSYPPRRPSHGKHEDDEDEDEEHQPRPDHPPQRPSRPEQQEPRG
      10      20      30      40      50      60

CV127.
      M

gi|169 RGCQTRNGVEENICTMKLHENIARPSRADFYNPKAGRISTLNSLTLPALRQFGLSAQYLV
      70      80      90      100      110      120

```

```

                10         20         30         40
CV127.  IVSRVSLRLLLVLRLFILILTRLRLGRIRLLMCLCVVMSWLCKG
                .. .:: . ::... :
gi|169  LYRNGIYSPHWNLNANSVIYVTRGK-GRVRVVNCQGNPVFDGDLTRGQLLLVPQNFVVAD
                130         140         150         160         170

gi|169  QGGKQGLEIYVFKTQHNAVSSYIKDLFRAIPSEVLSNSYNLGQSQVRQLKYQGNSGPLLN
                180         190         200         210         220         230

gi|169  P
                240

>>gi|25991543 gi|25991543|gb|AAN76862.1|AF453947_1 aller (457 aa)
  initn: 46 initl: 46 opt: 46 Z-score: 83.9 bits: 21.5 E(): 9.5
Smith-Waterman score: 46; 38.889% identity (66.667% similar) in 18 aa overlap
(27-44:299-316)

gi|259  LSVCFILILFHGCLASRQEWQQQDECQIDRLDALEPDNRVEYEAGTVEAWDPNHEQFRCAG
                10         20         30         40         50         60

gi|259  VALVRHTIQPNGLLLPQYSNAPQLIYVVGEGMTGISYPGCPETYQAPQQGRQQQSGGRF
                70         80         90         100        110        120

gi|259  QDRHQKIRRFRRGDIIAIPAGVAHWCYNEGNSPVTVTLLDVSNSQNQLDRTPRKFHLAG
                130        140        150        160        170        180

gi|259  NPKDVFQQQQQHQSRRNLFSGFDTELLAEAFQVDERLIKQLKSEDNRGGIVKVKDDEL
                190        200        210        220        230        240

                10         20
CV127.  MIVSRVSLRLLLVLRLFILILTRLRLG
                ..
gi|259  VIRPSRSQSERGSESEEESEDEKRRWGQRDNGIEETICTMRLKENINDPARADIYTPEVG
                250        260        270        280        290        300

                30         40
CV127.  RIRLLMCLCVVMSWLCKG
                .. : : . .:::
gi|259  RLTTLSLNLPLKWLQLSVEKGVLYKNALVLPHWNLNSHSIIYGCKGKGQVQVVDNFGN
                310        320        330        340        350        360

gi|259  RVFDGEVREGQMLVVPQNFAVVKRAREERFEWISFKTNDRAMTSPLAGRTSVLGGMPEEV
                370        380        390        400        410        420

gi|259  LANAFQISREDARKIKFNNQQTTLTSGESSHMRDDA
                430        440        450

```

47 residues in 1 query sequences
292011 residues in 1313 library sequences
Scomplib [34.26]
start: Sun Dec 14 22:49:43 2008 done: Sun Dec 14 22:49:44 2008
Total Scan time: 0.070 Total Display time: 0.010

Function used was FASTA [version 34.26.5 April 26, 2007]

APPENDIX 2. TOXIN HOMOLOGY AND GENERAL PROTEIN HOMOLOGY SEARCH RESULTS

The BLASTP results for the search of the GenBank non-redundant database with the deduced amino acid sequences of ORFs CV127.1392_d3 and CV127.5729_d2 are shown below. All descriptions of proteins from the GenBank non-redundant database which show the most significant local homology to these ORFs were manually compared to known toxins which act on humans as listed in the United States Code of Federal Regulations ([40 CFR Part 725.421](#)) to ensure that neither shared significant homology to a known toxin. Twenty-five of the 27 ORFs present within the BPS-CV127-9 transgene insert did not show homology to any known protein sequence using the search parameters defined in this report; thus no results are reported for those 25 ORFs in this appendix. The transgene insert of BPS-CV127-9 contains the coding sequences of two proteins from *A. thaliana*: ahasl (R272K, S653N) and SEC61γ; these two proteins were evaluated in BASF Plant Science report BPS-014-08 (McKean, 2008) and are not included among the 27 ORFs discussed in this current report.

BLASTP 2.2.18 [Mar-02-2008]

Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.* 25:3389-3402.

Reference for compositional score matrix adjustment: Altschul, Stephen F., John C. Wootton, E. Michael Gertz, Richa Agarwala, Aleksandr Morgulis, Alejandro A. Schaffer, and Yi-Kuo Yu (2005) "Protein database searches using compositionally adjusted substitution matrices", *FEBS J.* 272:5101-5109.

Query= CV127.1392_d3
(44 letters)

Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF excluding environmental samples
from WGS projects
7,155,275 sequences; 2,468,333,968 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
ref NP_566909.1 protein transport protein SEC61 gamma subunit, ...	67	3e-10
gb EAA37855.1 hypothetical protein OsJ_021338 [Oryza sativa (ja...	62	2e-08
ref NP_194222.1 protein transport protein SEC61 gamma subunit, ...	60	7e-08
emb CAN62208.1 hypothetical protein [Vitis vinifera]	59	1e-07
ref NP_001058242.1 Os06g0653900 [Oryza sativa (japonica cultiva...	59	2e-07
ref NP_001046076.1 Os02g0178400 [Oryza sativa (japonica cultiva...	58	2e-07
gb ABK93099.1 unknown [Populus trichocarpa]	57	5e-07
gb ABK93166.1 unknown [Populus trichocarpa]	57	5e-07
emb CAN67771.1 hypothetical protein [Vitis vinifera] >gi 157341...	55	1e-06
ref XP_001783762.1 predicted protein [Physcomitrella patens sub...	55	2e-06
ref XP_001771807.1 predicted protein [Physcomitrella patens sub...	55	2e-06
gb ABK21891.1 unknown [Picea sitchensis]	54	3e-06
gb ABK23495.1 unknown [Picea sitchensis]	54	3e-06

ref XP_001752211.1	predicted protein [Physcomitrella patens sub...	53	5e-06
gb AAR01761.1	putative transport protein [Oryza sativa (japonic...	49	1e-04
gb EAY90536.1	hypothetical protein OsI_011769 [Oryza sativa (in...	47	4e-04
ref XP_002116543.1	conserved hypothetical protein [Trichoplax a...	45	0.002
ref XP_797211.1	PREDICTED: hypothetical protein [Strongylocentr...	43	0.006
ref XP_001698939.1	SEC61-gamma subunit or ER translocon [Chlamy...	43	0.007
ref XP_001203964.1	PREDICTED: similar to Sec61 gamma subunit, p...	43	0.008
ref XP_001892530.1	Protein transport protein SEC61 gamma subuni...	43	0.009
gb ABX75489.1	protein transport protein SEC61 subunit gamma [Ly...	42	0.011
sp Q7Z1B8.1	S61G1_GRYOR Protein transport protein Sec61 subunit ...	42	0.012
ref XP_001421202.1	predicted protein [Ostreococcus lucimarinus ...	42	0.018
ref XP_001949988.1	PREDICTED: similar to Sec61p gamma subunit [...	41	0.028
ref XP_001606506.1	PREDICTED: similar to Sec61p gamma subunit [...	41	0.029
ref XP_001634305.1	predicted protein [Nematostella vectensis] >...	41	0.031
ref NP_001027676.1	sec61 protein [Ciona intestinalis] >gi 19843...	41	0.035
ref XP_975621.1	PREDICTED: similar to Sec61p gamma subunit [Tri...	40	0.046
sp Q7SZU9.1	SC61G_GADMO Protein transport protein Sec61 subunit ...	40	0.047
ref XP_001476947.1	PREDICTED: similar to Sec61-complex gamma-su...	40	0.049
ref NP_505778.1	EndoMitotic Oocytes family member (emo-1) [Caen...	40	0.052
ref NP_001037674.1	transport protein Sec61 gamma subunit [Bomby...	40	0.054
gb ACT64116.1	predicted protein [Thalassiosira pseudonana CCMP1...	40	0.065
ref XP_001120892.1	PREDICTED: similar to CG14214-PA [Apis melli...	40	0.065
emb CAF91108.1	unnamed protein product [Tetraodon nigroviridis]...	40	0.068
sp Q962X7.1	SC61G_BRABE Protein transport protein Sec61 subunit ...	40	0.069
sp Q7T207.1	SC61G_HARAN Protein transport protein Sec61 subunit ...	40	0.071
gb ACG76276.1	mitochondrial import inner membrane translocase s...	40	0.072
ref NP_055117.1	Sec61 gamma subunit [Homo sapiens] >gi 50979172...	40	0.072
ref NP_035473.2	SEC61, gamma subunit isoform 1 [Mus musculus] >...	40	0.075
gb ACT68444.1	transport protein Sec61 subunit gamma [Salmo salar]	40	0.076
sp Q66KU2.1	SC61G_XENLA Protein transport protein Sec61 subunit ...	40	0.076
ref XP_001367112.1	PREDICTED: similar to Sec61-complex gamma-su...	40	0.079
ref XP_001650928.1	Sec61 protein complex gamma subunit, putativ...	39	0.12
gb ABR23409.1	mitochondrial import inner membrane translocase s...	39	0.15
gb EDL06017.1	mCG140712 [Mus musculus]	39	0.16
gb AAI53755.1	LOC100127634 protein [Xenopus tropicalis]	39	0.16
ref XP_001163709.1	PREDICTED: similar to SEC61 gamma, partial [...	39	0.16
ref NP_610738.1	CG8860 CG8860-PA [Drosophila melanogaster] >gi ...	39	0.16
gb EDL03260.1	mCG1026143 [Mus musculus]	38	0.18
ref XP_001355499.1	GA12829 [Drosophila pseudoobscura pseudoobsc...	38	0.18
ref NP_608337.1	Sec61gamma CG14214-PA [Drosophila melanogaster]...	38	0.19
ref XP_001977803.1	GG19242 [Drosophila erecta] >gi 190649452 gb...	38	0.22
ref XP_346041.3	PREDICTED: similar to Protein transport protein...	38	0.22
ref XP_637046.1	protein transport protein SEC61 gamma subunit [...	38	0.23
ref XP_001480688.1	PREDICTED: similar to Sec61-complex gamma-su...	38	0.28
ref XP_001473751.1	PREDICTED: similar to Sec61-complex gamma-su...	37	0.46
gb AAQ96232.1	LRRGT00019 [Rattus norvegicus]	36	0.75

>ref|NP_566909.1| protein transport protein SEC61 gamma subunit, putative
[Arabidopsis thaliana]
sp|Q9SMP2.1|S61G3_ARATH Protein transport protein Sec61 subunit gamma-3
emb|CAB62346.1| protein translocation complex sec61 gamma chain-like protein
[Arabidopsis thaliana]
gb|AAM62573.1| protein translocation complex Sec61 gamma chain (pir T05513)
[Arabidopsis thaliana]
dbj|BAC42969.1| putative protein translocation complex sec61 gamma chain
[Arabidopsis thaliana]
gb|ABD38856.1| At3g48570 [Arabidopsis thaliana]
Length = 69

Score = 67.4 bits (163), Expect = 3e-10, Method: Compositional matrix adjust.
Identities = 31/31 (100%), Positives = 31/31 (100%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK
Sbjct: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31

>gb|EAZ37855.1| hypothetical protein OsJ_021338 [Oryza sativa (japonica
cultivar-group)]
Length = 415

Score = 61.6 bits (148), Expect = 2e-08, Method: Composition-based stats.
Identities = 26/32 (81%), Positives = 30/32 (93%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRKG 32
M+AIDS +DPLR+FAK SVRLV+RCHKPDRKG
Sbjct: 1 MDAIDSVVDPLREFAKDSVRLVKRCHKPDRKG 32

>ref|NP_194222.1| protein transport protein SEC61 gamma subunit, putative
[Arabidopsis thaliana]
ref|NP_568728.1| protein transport protein SEC61 gamma subunit, putative
[Arabidopsis thaliana]
sp|Q9SW34.1|S61G1_ARATH Protein transport protein Sec61 subunit gamma-1
gb|AAK43984.1|AF370169_1 putative protein translocation complex Sec61 gamma chain
[Arabidopsis thaliana]
emb|CAB36734.1| PROTEIN TRANSPORT PROTEIN SEC61 GAMMA SUBUNIT-like [Arabidopsis
thaliana]
emb|CAB79401.1| PROTEIN TRANSPORT PROTEIN SEC61 GAMMA SUBUNIT-like [Arabidopsis
thaliana]
dbj|BAB09131.1| protein translocation complex Sec61 gamma chain [Arabidopsis
thaliana]
gb|AAL15237.1| putative protein translocation complex Sec61 gamma chain
[Arabidopsis thaliana]
gb|AAM20513.1| protein transport protein SEC61 gamma subunit-like [Arabidopsis
thaliana]
gb|AAM91250.1| protein transport protein SEC61 gamma subunit-like [Arabidopsis
thaliana]
Length = 69

Score = 59.7 bits (143), Expect = 7e-08, Method: Compositional matrix adjust.
Identities = 25/31 (80%), Positives = 29/31 (93%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+AIDS +DPLRDFAK S+RLV+RCHKPDRK
Sbjct: 1 MDAIDSVVDPLRDFAKDSIRLVKRCHKPDRK 31

>emb|CAN62208.1| hypothetical protein [Vitis vinifera]
Length = 69

Score = 58.9 bits (141), Expect = 1e-07, Method: Compositional matrix adjust.
Identities = 25/31 (80%), Positives = 30/31 (96%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+AIDSA+DPLR+F+K SVRLV+RCHKPDRK
Sbjct: 1 MDAIDSAVDPLREFSKDSVRLVKRCHKPDRK 31

>ref|NP_001058242.1| Os06g0653900 [Oryza sativa (japonica cultivar-group)]
dbj|BAD37419.1| putative transport protein SEC61 [Oryza sativa Japonica Group]
dbj|BAD37541.1| putative transport protein SEC61 [Oryza sativa Japonica Group]
dbj|BAF20156.1| Os06g0653900 [Oryza sativa (japonica cultivar-group)]
gb|EAZ01925.1| hypothetical protein OsI_023157 [Oryza sativa (indica
cultivar-group)]
Length = 69

Score = 58.5 bits (140), Expect = 2e-07, Method: Compositional matrix adjust.
Identities = 25/31 (80%), Positives = 29/31 (93%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+AIDS +DPLR+FAK SVRLV+RCHKPDRK
Sbjct: 1 MDAIDSVVDPLREFAKDSVRLVKRCHKPDRK 31

>ref|NP_001046076.1| Os02g0178400 [Oryza sativa (japonica cultivar-group)]
sp|P38385.1|SC61G_ORYSJ Protein transport protein Sec61 subunit gamma
dbj|BAD27997.1| transport protein SEC61 [Oryza sativa Japonica Group]
dbj|BAF07990.1| Os02g0178400 [Oryza sativa (japonica cultivar-group)]
gb|EAY84716.1| hypothetical protein OsI_005949 [Oryza sativa (indica

```
cultivar-group)]
gb|EAZ21956.1| hypothetical protein OsJ_005439 [Oryza sativa (japonica
cultivar-group)]
gb|ABR26051.1| protein transport protein sec61 gamma subunit [Oryza sativa
(indica cultivar-group)]
Length = 69

Score = 57.8 bits (138), Expect = 2e-07, Method: Compositional matrix adjust.
Identities = 24/31 (77%), Positives = 29/31 (93%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+A+DS +DPLR+FAK SVRLV+RCHKPDRK
Sbjct: 1 MDAVDSVVDPLREFAKDSVRLVKRCHKPDRK 31

>gb|ABK93099.1| unknown [Populus trichocarpa]
Length = 69

Score = 57.0 bits (136), Expect = 5e-07, Method: Compositional matrix adjust.
Identities = 24/31 (77%), Positives = 29/31 (93%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+AIDS +DPLR+FAK SVRLV+RCHKPD+K
Sbjct: 1 MDAIDSVVDPLREFAKDSVRLVKRCHKPDQK 31

>gb|ABK93166.1| unknown [Populus trichocarpa]
Length = 69

Score = 56.6 bits (135), Expect = 5e-07, Method: Compositional matrix adjust.
Identities = 24/31 (77%), Positives = 29/31 (93%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+AIDS +DPLR+FAK SVRLV+RCHKPD+K
Sbjct: 1 MDAIDSVVDPLREFAKDSVRLVKRCHKPDQK 31

>emb|CAN67771.1| hypothetical protein [Vitis vinifera]
emb|CAO49023.1| unnamed protein product [Vitis vinifera]
Length = 69

Score = 55.5 bits (132), Expect = 1e-06, Method: Compositional matrix adjust.
Identities = 23/31 (74%), Positives = 28/31 (90%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+AID+ DP+R+FAK SVRLV+RCHKPDRK
Sbjct: 1 MDAIDNVFDPMREFAKDSVRLVKRCHKPDRK 31

>ref|XP_001783762.1| predicted protein [Physcomitrella patens subsp. patens]
gb|EDQ51414.1| predicted protein [Physcomitrella patens subsp. patens]
Length = 69

Score = 55.1 bits (131), Expect = 2e-06, Method: Compositional matrix adjust.
Identities = 22/31 (70%), Positives = 28/31 (90%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
MEA+D+ + P++DFAK SVRLV+RCHKPDRK
Sbjct: 1 MEAVDTVVRPVKDFAKDSVRLVKRCHKPDRK 31

>ref|XP_001771807.1| predicted protein [Physcomitrella patens subsp. patens]
gb|EDQ63415.1| predicted protein [Physcomitrella patens subsp. patens]
Length = 69

Score = 54.7 bits (130), Expect = 2e-06, Method: Compositional matrix adjust.
Identities = 21/31 (67%), Positives = 28/31 (90%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
MEA+D+ + P++DFAK S+RLV+RCHKPDRK
```

Sbjct: 1 MEAVDTVVRPVKDFAKDSIRLVKRCHKPDRK 31

>gb|ABK21891.1| unknown [Picea sitchensis]
Length = 69

Score = 54.3 bits (129), Expect = 3e-06, Method: Compositional matrix adjust.
Identities = 22/31 (70%), Positives = 28/31 (90%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+A+D+ + PL+DFAK SVRLV+RCHKPDRK
Sbjct: 1 MDAVDTVVRPLKDFAKDSVRLVVRCHKPDRK 31

>gb|ABK23495.1| unknown [Picea sitchensis]
Length = 69

Score = 54.3 bits (129), Expect = 3e-06, Method: Compositional matrix adjust.
Identities = 22/31 (70%), Positives = 28/31 (90%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
MEA+D+ + PL++FAK SVRLV+RCHKPDRK
Sbjct: 1 MEAVDAVVRPLQEFKDSVRLVVRCHKPDRK 31

>ref|XP_001752211.1| predicted protein [Physcomitrella patens subsp. patens]
ref|XP_001769079.1| predicted protein [Physcomitrella patens subsp. patens]
gb|EDQ66157.1| predicted protein [Physcomitrella patens subsp. patens]
gb|EDQ82944.1| predicted protein [Physcomitrella patens subsp. patens]
Length = 69

Score = 53.1 bits (126), Expect = 5e-06, Method: Compositional matrix adjust.
Identities = 20/31 (64%), Positives = 28/31 (90%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
MEA+D+ + P++DFAK S+RLV+RCHKPDR+
Sbjct: 1 MEAVDTVVRPVKDFAKDSIRLVKRCHKPDRR 31

>gb|AAR01761.1| putative transport protein [Oryza sativa (japonica
cultivar-group)]
gb|ABF96729.1| Protein transport protein SEC61 gamma subunit, putative [Oryza
sativa (japonica cultivar-group)]
gb|EAAZ27400.1| hypothetical protein OsJ_010883 [Oryza sativa (japonica
cultivar-group)]
Length = 71

Score = 48.5 bits (114), Expect = 1e-04, Method: Compositional matrix adjust.
Identities = 20/31 (64%), Positives = 26/31 (83%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+A+DS +D LR+FAK S+ LV+ CHKPDRK
Sbjct: 1 MDAVDSVVDLREFAKDSICLVKHCHKPDRK 31

>gb|EAY90536.1| hypothetical protein OsI_011769 [Oryza sativa (indica
cultivar-group)]
Length = 118

Score = 47.4 bits (111), Expect = 4e-04, Method: Compositional matrix adjust.
Identities = 20/31 (64%), Positives = 26/31 (83%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+A+DS +D LR+FAK S+ LV+ CHKPDRK
Sbjct: 72 MDAVDSVVDLREFAKDSICLVKHCHKPDRK 102

>ref|XP_002116543.1| conserved hypothetical protein [Trichoplax adhaerens]
gb|EDV20899.1| conserved hypothetical protein [Trichoplax adhaerens]
Length = 68

Score = 44.7 bits (104), Expect = 0.002, Method: Compositional matrix adjust.
Identities = 18/31 (58%), Positives = 24/31 (77%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + + +DPL+ F K SVRLV+RC KPDRK
Sbjct: 1 MDQVTTLVDPLKVFVKDSVRLVKRCTKPDRK 31

>ref|XP_797211.1| PREDICTED: hypothetical protein [Strongylocentrotus purpuratus]
ref|XP_001176408.1| PREDICTED: hypothetical protein [Strongylocentrotus purpuratus]
Length = 68

Score = 43.1 bits (100), Expect = 0.006, Method: Compositional matrix adjust.
Identities = 18/31 (58%), Positives = 23/31 (74%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ I I+P + FAK S+RLV+RC KPDRK
Sbjct: 1 MDQIQQTIEPAKQFAKDSIRLVKRCTKPDRK 31

>ref|XP_001698939.1| SEC61-gamma subunit or ER translocon [Chlamydomonas reinhardtii]
dbj|BAF46295.1| putative protein transport protein SEC61 [Chlamydomonas
reinhardtii]
gb|ED099221.1| SEC61-gamma subunit or ER translocon [Chlamydomonas reinhardtii]
Length = 68

Score = 43.1 bits (100), Expect = 0.007, Method: Compositional matrix adjust.
Identities = 18/31 (58%), Positives = 24/31 (77%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ D + P++DFAK+S RLV+RC KPDRK
Sbjct: 1 MDLGDVVVKPVKDFAKNSARLVKRCTKPDRK 31

>ref|XP_001203964.1| PREDICTED: similar to Sec61 gamma subunit, partial
[Strongylocentrotus purpuratus]
ref|XP_001199784.1| PREDICTED: similar to Sec61 gamma subunit, partial
[Strongylocentrotus purpuratus]
Length = 31

Score = 42.7 bits (99), Expect = 0.008, Method: Compositional matrix adjust.
Identities = 18/31 (58%), Positives = 23/31 (74%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ I I+P + FAK S+RLV+RC KPDRK
Sbjct: 1 MDQIQQTIEPAKQFAKDSIRLVKRCTKPDRK 31

>ref|XP_001892530.1| Protein transport protein SEC61 gamma subunit [Brugia malayi]
gb|EDP38639.1| Protein transport protein SEC61 gamma subunit, putative [Brugia
malayi]
Length = 71

Score = 42.7 bits (99), Expect = 0.009, Method: Compositional matrix adjust.
Identities = 18/31 (58%), Positives = 24/31 (77%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + + IDP + FAK S+RLV+RC KPDRK
Sbjct: 4 MDHLQALIDPSKQFAKDSIRLVKRCTKPDRK 34

>gb|ABX75489.1| protein transport protein SEC61 subunit gamma [Lycosa
singoriensis]
Length = 68

Score = 42.4 bits (98), Expect = 0.011, Method: Compositional matrix adjust.
Identities = 19/31 (61%), Positives = 24/31 (77%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31

M+ + S I+P R FAK S+RLV+RC KPDRK
Sbjct: 1 MDQVMSFIEPDRQFAKDSIRLVKRCTKPDRK 31

>sp|Q7Z1B8.1|S61G1_GRYOR Protein transport protein Sec61 subunit gamma
gb|AAP47228.1| Sec61p gamma subunit [Gryllotalpa orientalis]
Length = 68

Score = 42.4 bits (98), Expect = 0.012, Method: Compositional matrix adjust.
Identities = 18/31 (58%), Positives = 23/31 (74%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + I+P R FAK S+RLV+RC KPDRK
Sbjct: 1 MDQVTKFIEPGRQFAKDSIRLVKRCTKPDRK 31

>ref|XP_001421202.1| predicted protein [Ostreococcus lucimarinus CCE9901]
gb|ABO99495.1| predicted protein [Ostreococcus lucimarinus CCE9901]
Length = 68

Score = 41.6 bits (96), Expect = 0.018, Method: Compositional matrix adjust.
Identities = 17/31 (54%), Positives = 23/31 (74%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
ME D A+ P+RDFA+ S RLV++C KPD +
Sbjct: 1 MEVADIAVKPVRDFARDSARLVKKCAKPDAR 31

>ref|XP_001949988.1| PREDICTED: similar to Sec61p gamma subunit [Acyrtosiphon pisum]
Length = 66

Score = 40.8 bits (94), Expect = 0.028, Method: Compositional matrix adjust.
Identities = 17/31 (54%), Positives = 23/31 (74%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + ++P R F+K SVRLV+RC KPDRK
Sbjct: 1 MDQVTKVLEPGRQFSKDSVRLVVRCTKPDRK 31

>ref|XP_001606506.1| PREDICTED: similar to Sec61p gamma subunit [Nasonia vitripennis]
Length = 68

Score = 40.8 bits (94), Expect = 0.029, Method: Compositional matrix adjust.
Identities = 18/31 (58%), Positives = 23/31 (74%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ I ++P R FAK S+RLV+RC KPDRK
Sbjct: 1 MDQIKMFVEPSRQFAKDSMRLVVRCTKPDRK 31

>ref|XP_001634305.1| predicted protein [Nematostella vectensis]
gb|ED042242.1| predicted protein [Nematostella vectensis]
Length = 67

Score = 40.8 bits (94), Expect = 0.031, Method: Compositional matrix adjust.
Identities = 17/31 (54%), Positives = 24/31 (77%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ I + ++P + FAK S+RLV+RC KPDRK
Sbjct: 1 MDQIVNTLEPCKQFAKDSMRLVVRCTKPDRK 31

>ref|NP_001027676.1| sec61 protein [Ciona intestinalis]
ref|XP_002125574.1| PREDICTED: similar to putative transport protein sec61 gamma
subunit [Ciona intestinalis]
ref|XP_002119112.1| PREDICTED: hypothetical protein [Ciona intestinalis]
sp|Q8I7D9.1|SC61G_CIOIN Protein transport protein Sec61 subunit gamma
emb|CAC82549.1| putative transport protein sec61 gamma subunit [Ciona
intestinalis]
Length = 68

Score = 40.8 bits (94), Expect = 0.035, Method: Compositional matrix adjust.
Identities = 16/31 (51%), Positives = 24/31 (77%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + + ++P + FAK S+RLV+RC KPDRK
Sbjct: 1 MDQVMAWVEPGKQFAKDSIRLVKRCTKPDRK 31

>ref|XP_975621.1| PREDICTED: similar to Sec61p gamma subunit [Tribolium castaneum]
Length = 68

Score = 40.4 bits (93), Expect = 0.046, Method: Compositional matrix adjust.
Identities = 17/31 (54%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ I I+P R F K S+RLV+RC KPDR+
Sbjct: 1 MDQITKFIEPGRQFTKDSIRLVKRCTKPDRR 31

>sp|Q7SZU9.1|SC61G_GADMO Protein transport protein Sec61 subunit gamma
gb|AAQ18693.1| Sec61 gamma subunit [Gadus morhua]
Length = 68

Score = 40.4 bits (93), Expect = 0.047, Method: Compositional matrix adjust.
Identities = 17/31 (54%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ I ++P R F K S+RLV+RC KPDRK
Sbjct: 1 MDQIMQFVEPSRQFVKDSIRLVKRCTKPDRK 31

>ref|XP_001476947.1| PREDICTED: similar to Sec61-complex gamma-subunit [Mus musculus]
ref|XP_001473383.1| PREDICTED: similar to Sec61-complex gamma-subunit [Mus musculus]
gb|EDL20549.1| MCG8406 [Mus musculus]
Length = 68

Score = 40.0 bits (92), Expect = 0.049, Method: Compositional matrix adjust.
Identities = 16/31 (51%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + ++P R F K S+RLV+RC KPDRK
Sbjct: 1 MDQVMQFVEPSRQFVKESIRLVKRCTKPDRK 31

>ref|NP_505778.1| EndoMitotic Oocytes family member (emo-1) [Caenorhabditis elegans]
ref|XP_001673886.1| Hypothetical protein CBG09741 [Caenorhabditis briggsae AF16]
ref|XP_001673897.1| Hypothetical protein CBG09730 [Caenorhabditis briggsae AF16]
sp|Q19967.1|SC61G_CAEEL Protein transport protein Sec61 subunit gamma
gb|AAC47274.1| Sec61p gamma homolog
emb|CAA98458.1| C. elegans protein F32D8.6, confirmed by transcript evidence
[Caenorhabditis elegans]
emb|CAP28905.1| C. briggsae CBR-EMO-1 protein [Caenorhabditis briggsae]
Length = 68

Score = 40.0 bits (92), Expect = 0.052, Method: Compositional matrix adjust.
Identities = 17/31 (54%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + I+P R F+K S RLV+RC KPDRK
Sbjct: 1 MDQFQALIEPARQFSKDSYRLVKRCTKPDRK 31

>ref|NP_001037674.1| transport protein Sec61 gamma subunit [Bombyx mori]
gb|ABF51395.1| transport protein SEC61 [Bombyx mori]
gb|ABF85697.1| transport protein Sec61 gamma subunit [Bombyx mori]
Length = 68

Score = 40.0 bits (92), Expect = 0.054, Method: Compositional matrix adjust.

Identities = 17/31 (54%), Positives = 23/31 (74%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ I ++P + FAK S+RLV+RC KPDRK
Sbjct: 1 MDQIAKFVEPGKQFAKDSIRLVRRCTKPDRK 31

>gb|ACI64116.1| predicted protein [Thalassiosira pseudonana CCMP1335]
Length = 72

Score = 39.7 bits (91), Expect = 0.065, Method: Compositional matrix adjust.
Identities = 16/27 (59%), Positives = 20/27 (74%)

Query: 5 DSAIDPLRDFAKSSVRLVQRCHKPDRK 31
+ I PLR FAK S+ LV++C KPDRK
Sbjct: 9 EVLIQPLRQFAKDSIHLVKKCTKPDRK 35

>ref|XP_001120892.1| PREDICTED: similar to CG14214-PA [Apis mellifera]
Length = 68

Score = 39.7 bits (91), Expect = 0.065, Method: Compositional matrix adjust.
Identities = 16/31 (51%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + ++P R FAK S+RL++RC KPDRK
Sbjct: 1 MDQVKKLTEPGRQFAKDSIRLIKRCTKPDRK 31

>emb|CAF91108.1| unnamed protein product [Tetraodon nigroviridis]
emb|CAF97954.1| unnamed protein product [Tetraodon nigroviridis]
Length = 68

Score = 39.7 bits (91), Expect = 0.068, Method: Compositional matrix adjust.
Identities = 17/31 (54%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + ++P R F K SVRLV+RC KPDRK
Sbjct: 1 MDQVMQFVEPGRQFIKDSVRLVKRCTKPDRK 31

>sp|Q962X7.1|SC61G_BRABE Protein transport protein Sec61 subunit gamma
gb|AAK72260.1|AF395915_1 SEC61 [Branchiostoma belcheri]
Length = 68

Score = 39.7 bits (91), Expect = 0.069, Method: Compositional matrix adjust.
Identities = 17/31 (54%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + ++P R FAK S RLV+RC KPDRK
Sbjct: 1 MDQLVQFVEPARQFAKDSYRLVKRCTKPDRK 31

>sp|Q7T207.1|SC61G_HARAN Protein transport protein Sec61 subunit gamma
gb|AAP74549.1| SEC61 gamma [Harpagifer antarcticus]
Length = 68

Score = 39.7 bits (91), Expect = 0.071, Method: Compositional matrix adjust.
Identities = 16/31 (51%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + ++P R F K S+RLV+RC KPDRK
Sbjct: 1 MDQVMQFVEPSRQFVKDSIRLVKRCTKPDRK 31

>gb|ACG76276.1| mitochondrial import inner membrane translocase subunit TIM8
[Amblyomma americanum]
Length = 68

Score = 39.7 bits (91), Expect = 0.072, Method: Compositional matrix adjust.

Identities = 15/31 (48%), Positives = 23/31 (74%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + + IDP + +K S+RL++RC KPDRK
Sbjct: 1 MDQVMAFIDPFKQLSKDSIRLIKRCTKPDRK 31

```
>ref|NP_055117.1| Sec61 gamma subunit [Homo sapiens]
ref|NP_001003325.1| Sec61 gamma subunit [Canis lupus familiaris]
ref|NP_001012474.1| Sec61 gamma subunit [Homo sapiens]
ref|NP_001035676.1| Sec61 gamma subunit [Bos taurus]
ref|NP_001103441.1| SEC61, gamma subunit isoform 2 [Mus musculus]
ref|NP_001103442.1| SEC61, gamma subunit isoform 2 [Mus musculus]
ref|NP_001128492.1| SEC61, gamma subunit [Rattus norvegicus]
ref|XP_001107920.1| PREDICTED: similar to Protein transport protein SEC61 gamma
subunit [Macaca mulatta]
ref|XP_001152525.1| PREDICTED: similar to SEC61 gamma [Pan troglodytes]
ref|XP_001236152.1| PREDICTED: similar to SEC61 gamma [Gallus gallus]
ref|XP_001479532.1| PREDICTED: similar to SEC61 gamma [Mus musculus]
ref|XP_001475666.1| PREDICTED: similar to SEC61 gamma [Mus musculus]
ref|XP_001478183.1| PREDICTED: similar to SEC61 gamma [Mus musculus]
ref|XP_001478570.1| PREDICTED: similar to SEC61 gamma [Mus musculus]
ref|XP_001480211.1| PREDICTED: similar to SEC61 gamma [Mus musculus]
sp|P60058.1|SC61G_CANFA Protein transport protein Sec61 subunit gamma
sp|P60059.1|SC61G_HUMAN Protein transport protein Sec61 subunit gamma
sp|P60060.1|SC61G_MOUSE RecName: Full=Protein transport protein Sec61 subunit gamma
sp|Q3T104.1|SC61G_BOVIN Protein transport protein Sec61 subunit gamma
gb|AAA19705.1| Sec61-complex gamma-subunit
gb|AAA19431.1| Sec61 protein complex gamma subunit
gb|AAC99401.1| Sec61 gamma [Homo sapiens]
dbj|BAB28210.1| unnamed protein product [Mus musculus]
dbj|BAB28376.1| unnamed protein product [Mus musculus]
gb|AAH09480.1| Sec61 gamma subunit [Homo sapiens]
gb|AAH19158.1| SEC61, gamma subunit [Mus musculus]
emb|CAG33260.1| SEC61G [Homo sapiens]
gb|AAH81456.1| SEC61, gamma subunit [Mus musculus]
gb|AAH51840.1| Sec61 gamma subunit [Homo sapiens]
emb|CAI24345.1| SEC61, gamma subunit (S. cerevisiae) [Mus musculus]
dbj|BAE35762.1| unnamed protein product [Mus musculus]
gb|AAI02187.1| Sec61 gamma subunit [Bos taurus]
emb|CAL38284.1| hypothetical protein [synthetic construct]
gb|EAW50960.1| Sec61 gamma subunit, isoform CRA_a [Homo sapiens]
gb|EAW50961.1| Sec61 gamma subunit, isoform CRA_a [Homo sapiens]
gb|EDL00154.1| mCG116534 [Mus musculus]
gb|EDL25421.1| mCG4124 [Mus musculus]
gb|EDL40660.1| mCG7641, isoform CRA_a [Mus musculus]
gb|EDL40661.1| mCG7641, isoform CRA_a [Mus musculus]
gb|EDL97891.1| rCG23282, isoform CRA_a [Rattus norvegicus]
gb|EDL97892.1| rCG23282, isoform CRA_a [Rattus norvegicus]
gb|EDL97893.1| rCG23282, isoform CRA_a [Rattus norvegicus]
dbj|BAG34787.1| unnamed protein product [Homo sapiens]
gb|AAI68720.1| Unknown (protein for MGC:188791) [Rattus norvegicus]
gb|AAI68779.1| Unknown (protein for MGC:188933) [Rattus norvegicus]
prf||2005371B Sec61 protein:SUBUNIT=gamma
Length = 68
```

Score = 39.7 bits (91), Expect = 0.072, Method: Compositional matrix adjust.
Identities = 16/31 (51%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + ++P R F K S+RLV+RC KPDRK
Sbjct: 1 MDQVMQFVEPSRQFVKDSIRLVKRCTKPDRK 31

```
>ref|NP_035473.2| SEC61, gamma subunit isoform 1 [Mus musculus]
gb|EDL40662.1| mCG7641, isoform CRA_b [Mus musculus]
Length = 100
```

Score = 39.7 bits (91), Expect = 0.075, Method: Compositional matrix adjust.
Identities = 16/31 (51%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + ++P R F K S+RLV+RC KPDRK
Sbjct: 33 MDQVMQFVEPSRQFVKDSIRLVKRCTKPDRK 63

>gb|ACI68444.1| transport protein Sec61 subunit gamma [Salmo salar]
Length = 68

Score = 39.7 bits (91), Expect = 0.076, Method: Compositional matrix adjust.
Identities = 16/31 (51%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + ++P R F K S+RLV+RC KPDRK
Sbjct: 1 MDQVMQFVEPSRQFVKDSIRLVKRCTKPDRK 31

>sp|Q66KU2.1|SC61G_XENLA Protein transport protein Sec61 subunit gamma
gb|AAH78558.1| Unknown (protein for MGC:85435) [Xenopus laevis]
Length = 68

Score = 39.7 bits (91), Expect = 0.076, Method: Compositional matrix adjust.
Identities = 16/31 (51%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + ++P R F K S+RLV+RC KPDRK
Sbjct: 1 MDQVMQFVEPSRQFVKDSIRLVKRCTKPDRK 31

>ref|XP_001367112.1| PREDICTED: similar to Sec61-complex gamma-subunit [Monodelphis domestica]
ref|XP_001505526.1| PREDICTED: similar to Sec61-complex gamma-subunit [Ornithorhynchus anatinus]
Length = 66

Score = 39.7 bits (91), Expect = 0.079, Method: Compositional matrix adjust.
Identities = 16/31 (51%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + ++P R F K S+RLV+RC KPDRK
Sbjct: 1 MDQVMQFVEPSRQFVKDSIRLVKRCTKPDRK 31

>ref|XP_001650928.1| Sec61 protein complex gamma subunit, putative [Aedes aegypti]
ref|XP_001846718.1| Sec61 protein complex gamma subunit [Culex quinquefasciatus]
gb|ABF18076.1| preprotein translocase gamma subunit [Aedes aegypti]
gb|EAT43047.1| Sec61 protein complex gamma subunit, putative [Aedes aegypti]
gb|EDS44447.1| Sec61 protein complex gamma subunit [Culex quinquefasciatus]
Length = 68

Score = 38.9 bits (89), Expect = 0.12, Method: Compositional matrix adjust.
Identities = 16/31 (51%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ I +P R FAK S+RL++RC KPDR+
Sbjct: 1 MDQIAKIYEPGRSFAKDSIRLIKRCTKPDRR 31

>gb|ABR23409.1| mitochondrial import inner membrane translocase subunit TIM8 [Ornithodoros parkeri]
Length = 68

Score = 38.5 bits (88), Expect = 0.15, Method: Compositional matrix adjust.
Identities = 15/31 (48%), Positives = 23/31 (74%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + + +PL+ +K S+RLV+RC KPDRK
Sbjct: 1 MDQVMAFFELKQLSKDSIRLVKRCTKPDRK 31

>gb|EDL06017.1| mCG140712 [Mus musculus]
Length = 84

Score = 38.5 bits (88), Expect = 0.16, Method: Compositional matrix adjust.
Identities = 15/31 (48%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + ++P R F K S+R+V+RC KPDRK
Sbjct: 17 MDQVMQFVEPSRQFVKDSIRLVKRCCTKPDRK 47

>gb|AAI53755.1| LOC100127634 protein [Xenopus tropicalis]
Length = 69

Score = 38.5 bits (88), Expect = 0.16, Method: Compositional matrix adjust.
Identities = 15/31 (48%), Positives = 23/31 (74%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ ++ PL+DFAK+SVRL ++C KPD +
Sbjct: 1 MDILEETAAPLKDFAKNSVRLFKKCTKPDAQ 31

>ref|XP_001163709.1| PREDICTED: similar to SEC61 gamma, partial [Pan troglodytes]
Length = 31

Score = 38.5 bits (88), Expect = 0.16, Method: Compositional matrix adjust.
Identities = 16/31 (51%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + ++P R F K S+RLV+RC KPDRK
Sbjct: 1 MDQVMQFVEPSRQFVKDSIRLVKRCCTKPDRK 31

>ref|NP_610738.1| CG8860 CG8860-PA [Drosophila melanogaster]
ref|XP_001975973.1| GG22600 [Drosophila erecta]
ref|XP_002033494.1| GM20379 [Drosophila sechellia]
ref|XP_002091116.1| GE13468 [Drosophila yakuba]
ref|XP_002076308.1| GD15244 [Drosophila simulans]
sp|Q9V668.1|S61G1_DROME Protein transport protein Sec61 gamma-1 subunit
gb|AAF58563.1| CG8860-PA [Drosophila melanogaster]
gb|AAM50706.1| GM14157p [Drosophila melanogaster]
gb|EDV56373.1| GG22600 [Drosophila erecta]
gb|EDW47507.1| GM20379 [Drosophila sechellia]
gb|EDW90828.1| GE13468 [Drosophila yakuba]
gb|EDX15533.1| GD15244 [Drosophila simulans]
Length = 68

Score = 38.5 bits (88), Expect = 0.16, Method: Compositional matrix adjust.
Identities = 17/31 (54%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + +P R FAK S+RLV+RC KPDRK
Sbjct: 1 MDKVVKFAEPGRAFAKDSIRLVKRCCTKPDRK 31

>gb|EDL03260.1| mCG1026143 [Mus musculus]
Length = 68

Score = 38.1 bits (87), Expect = 0.18, Method: Compositional matrix adjust.
Identities = 16/31 (51%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + ++P R F K SV+LV+RC KPDRK
Sbjct: 1 MDQVMQFVEPSRQFVKDSVQLVKRCCTKPDRK 31

>ref|XP_001355499.1| GA12829 [Drosophila pseudoobscura pseudoobscura]
ref|XP_001963467.1| GF20416 [Drosophila ananassae]
ref|XP_002010749.1| GI21533 [Drosophila mojavensis]
ref|XP_002027916.1| GL27062 [Drosophila persimilis]

```

ref|XP_002058289.1| GJ16006 [Drosophila virilis]
ref|XP_002067352.1| GK16225 [Drosophila willistoni]
gb|EAL32558.1| GA12829 [Drosophila pseudoobscura pseudoobscura]
gb|EDV44543.1| GF20416 [Drosophila ananassae]
gb|EDW06404.1| GI21533 [Drosophila mojavensis]
gb|EDW37648.1| GL27062 [Drosophila persimilis]
gb|EDW66397.1| GJ16006 [Drosophila virilis]
gb|EDW78338.1| GK16225 [Drosophila willistoni]
      Length = 68

Score = 38.1 bits (87), Expect = 0.18, Method: Compositional matrix adjust.
Identities = 17/31 (54%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
      M+ + +P R FAK S+RLV+RC KPDRK
Sbjct: 1 MDKVVKFAEPGRAFAKDSIRLVKRCTKPDRK 31

>ref|NP_608337.1| Sec61gamma CG14214-PA [Drosophila melanogaster]
ref|XP_002039446.1| GM22977 [Drosophila sechellia]
ref|XP_002100996.1| GE15861 [Drosophila yakuba]
sp|Q9VWE9.1|S61G2_DROME Protein transport protein Sec61 gamma-2 subunit
gb|AAF48993.1| CG14214-PA [Drosophila melanogaster]
gb|AAL48065.1| RE69515p [Drosophila melanogaster]
gb|EDW56188.1| GM22977 [Drosophila sechellia]
gb|EDX02104.1| GE15861 [Drosophila yakuba]
      Length = 68

Score = 38.1 bits (87), Expect = 0.19, Method: Compositional matrix adjust.
Identities = 17/31 (54%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
      M+ + +P R FAK S+RLV+RC KPDRK
Sbjct: 1 MDKVVKFAEPGRAFAKDSIRLVKRCTKPDRK 31

>ref|XP_001977803.1| GG19242 [Drosophila erecta]
gb|EDV46730.1| GG19242 [Drosophila erecta]
      Length = 68

Score = 38.1 bits (87), Expect = 0.22, Method: Compositional matrix adjust.
Identities = 17/31 (54%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
      M+ + +P R FAK S+RLV+RC KPDRK
Sbjct: 1 MDKVVKFAEPGRVFAKDSIRLVKRCTKPDRK 31

>ref|XP_346041.3| PREDICTED: similar to Protein transport protein SEC61 gamma
      subunit [Rattus norvegicus]
      Length = 68

Score = 38.1 bits (87), Expect = 0.22, Method: Compositional matrix adjust.
Identities = 15/31 (48%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
      M+ + ++P + F K S+RLV+RC KPDRK
Sbjct: 1 MDQVMQFVEPSQQFVKDSIRLVKRCTKPDRK 31

>ref|XP_637046.1| protein transport protein SEC61 gamma subunit [Dictyostelium
      discoideum AX4]
sp|Q54JV6.1|SC61G_DICDI RecName: Full=Protein transport protein Sec61 subunit gamma
gb|EAL63535.1| protein transport protein SEC61 gamma subunit [Dictyostelium
      discoideum AX4]
      Length = 69

Score = 38.1 bits (87), Expect = 0.23, Method: Compositional matrix adjust.
Identities = 14/31 (45%), Positives = 23/31 (74%)

```



The Chemical Company
Plant Science L. L. C.

BASF Reg. Doc. No. 2008/7019371

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ ++ PL+DFAK+S+RL ++C KPD +
Sbjct: 1 MDILEETAAPLKDFAKNSIRLFKKCTKPDAQ 31

>ref|XP_001480688.1| PREDICTED: similar to Sec61-complex gamma-subunit [Mus musculus]
ref|XP_001476807.1| PREDICTED: similar to Sec61-complex gamma-subunit [Mus musculus]
Length = 68

Score = 37.7 bits (86), Expect = 0.28, Method: Compositional matrix adjust.
Identities = 16/31 (51%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
ME + + ++P R F K SVRLV+R KPDR+
Sbjct: 1 MEQVMNFVEPSRQFVKDSVRLVKRSTKPDRR 31

>ref|XP_001473751.1| PREDICTED: similar to Sec61-complex gamma-subunit [Mus musculus]
Length = 267

Score = 37.0 bits (84), Expect = 0.46, Method: Compositional matrix adjust.
Identities = 16/31 (51%), Positives = 22/31 (70%)

Query: 1 MEAIDSAIDPLRDFAKSSVRLVQRCHKPDRK 31
M+ + ++P R F K SV+LV+RC KPDRK
Sbjct: 1 MDQVMQFVEPSRQFVKDSVQLVKRCTKPDRK 31

>gb|AAQ96232.1| LRRGT00019 [Rattus norvegicus]
Length = 353

Score = 36.2 bits (82), Expect = 0.75, Method: Composition-based stats.
Identities = 14/24 (58%), Positives = 19/24 (79%)

Query: 8 IDPLRDFAKSSVRLVQRCHKPDRK 31
++P + F K S+RLV+RC KPDRK
Sbjct: 138 VEPSQQFVKDSIRLVKRCTKPDRK 161

Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF excluding environmental samples
from WGS projects
Posted date: Oct 26, 2008 5:54 PM
Number of letters in database: 2,468,333,968
Number of sequences in database: 7,155,275

Lambda	K	H
0.321	0.134	0.384

Gapped Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 7155275
Number of Hits to DB: 171,863,387
Number of extensions: 3369367
Number of successful extensions: 6783
Number of sequences better than 1.0: 59
Number of HSP's gapped: 6783
Number of HSP's successfully gapped: 59
Length of query: 44
Length of database: 2,468,333,968
Length adjustment: 18
Effective length of query: 26
Effective length of database: 2,339,539,018
Effective search space: 60828014468
Effective search space used: 60828014468

Neighboring words threshold: 11
Window for multiple hits: 40
X1: 16 (7.4 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (21.9 bits)
S2: 82 (36.2 bits)

BLASTP 2.2.18 [Mar-02-2008]

Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

Reference for compositional score matrix adjustment: Altschul, Stephen F., John C. Wootton, E. Michael Gertz, Richa Agarwala, Aleksandr Morgulis, Alejandro A. Schaffer, and Yi-Kuo Yu (2005) "Protein database searches using compositionally adjusted substitution matrices", FEBS J. 272:5101-5109.

Query= CV127.5729_d2
(166 letters)

Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF excluding environmental samples
from WGS projects
7,155,275 sequences; 2,468,333,968 total letters

Searching.....done

Sequences producing significant alignments:			Score (bits)	E Value
gb ABJ80681.1	acetolactate synthase catalytic subunit [Arabidop...	242	6e-63	
gb AAK68759.1	acetolactate synthase [Arabidopsis thaliana]	242	6e-63	
ref NP_190425.1	CSR1 (CHLORSULFURON/IMIDAZOLINONE RESISTANT 1) ...	242	6e-63	
gb AAM92569.1	acetolactate synthase [Arabidopsis thaliana]	242	6e-63	
emb CAA35887.1	unnamed protein product [Arabidopsis thaliana]	242	6e-63	
gb AAT72502.1	AT3G48560 [Arabidopsis lyrata subsp. lyrata]	240	2e-62	
pdb 1YBH A Chain A, Crystal Structure Of Arabidopsis Thaliana Ac...		239	4e-62	
gb AAR07632.1	acetolactate synthase 1 [Camelina microcarpa]	236	6e-61	
sp P27818 ILV1_BRANA Acetolactate synthase 1, chloroplast precur...		235	7e-61	
sp P27819 ILV3_BRANA Acetolactate synthase 3, chloroplast precur...		235	7e-61	
gb AAR07633.1	acetolactate synthase 1 [Camelina microcarpa]	235	7e-61	
gb AAA62705.1	acetolactate synthase	234	1e-60	
gb AAY46005.1	putative acetolactate synthase [Sinapis arvensis]	234	2e-60	
gb AAY46012.1	putative acetolactate synthase [Sinapis arvensis]	234	2e-60	
gb AAY46013.1	putative acetolactate synthase [Sinapis arvensis]	234	2e-60	
gb AAY45995.1	putative acetolactate synthase [Sinapis arvensis]...	234	2e-60	
gb AAY45998.1	putative acetolactate synthase [Sinapis arvensis]...	234	2e-60	
gb AAY46007.1	putative acetolactate synthase [Sinapis arvensis]	234	2e-60	
gb AAY46008.1	putative acetolactate synthase [Sinapis arvensis]	234	2e-60	
gb AAY46006.1	putative acetolactate synthase [Sinapis arvensis]...	234	2e-60	
gb AAY46004.1	putative acetolactate synthase [Sinapis arvensis]	234	2e-60	
gb AAY46010.1	putative acetolactate synthase [Sinapis arvensis]	234	2e-60	
gb AAY46011.1	putative acetolactate synthase [Sinapis arvensis]	234	2e-60	
emb CAC86692.1	putative acetolactate synthase [Raphanus raphani...	234	2e-60	
gb ACB12189.1	acetolactate synthase [Descurainia sophia]	234	2e-60	
gb ACB12190.1	acetolactate synthase [Descurainia sophia]	234	2e-60	
gb ACB12188.1	acetolactate synthase [Descurainia sophia]	234	2e-60	
gb AAR06607.1	acetolactate synthase 2 [Camelina microcarpa]	233	2e-60	
emb CAC86696.1	putative acetolactate synthase [Raphanus raphani...	233	3e-60	
emb CAC86703.1	putative acetolactate synthase [Raphanus raphani...	233	4e-60	
emb CAC86697.1	putative acetolactate synthase [Raphanus raphani...	232	6e-60	
emb CAC86701.1	putative acetolactate synthase [Raphanus raphani...	231	1e-59	
emb CAC86698.1	putative acetolactate synthase [Raphanus raphani...	231	1e-59	

emb CAC86695.1	putative acetolactate synthase [Raphanus raphani...	231	1e-59
emb CAC86700.1	putative acetolactate synthase [Raphanus raphani...	231	1e-59
emb CAC86694.1	putative acetolactate synthase [Raphanus raphani...	231	1e-59
emb CAC86702.1	putative acetolactate synthase [Raphanus raphani...	230	3e-59
emb CAC86699.1	putative acetolactate synthase [Raphanus raphani...	229	4e-59
emb CAC86693.1	putative acetolactate synthase [Raphanus raphani...	229	5e-59
gb ABM53021.1	acetolactate synthase [Amaranthus tuberculatus]	207	3e-52
gb AAK50820.1	AF363369_1 acetolactate synthase [Amaranthus retro...	206	3e-52
gb AAK50821.1	AF363370_1 acetolactate synthase [Amaranthus powel...	206	4e-52
gb ABS72164.1	acetolactate synthase [Amaranthus hypochondriacus]	206	4e-52
gb AAB67839.1	acetolactate synthase precursor [Amaranthus sp.]	204	2e-51
emb CAA87084.1	acetohydroxyacid synthase [Gossypium hirsutum]	203	3e-51
gb ABM53018.1	acetolactate synthase [Amaranthus tuberculatus]	202	5e-51
gb ABM53019.1	acetolactate synthase [Amaranthus tuberculatus]	202	5e-51
prf 1407140B	acetolactate synthase SuRB	202	7e-51
sp P09114 ILV2	TOBAC Acetolactate synthase 2, chloroplast precu...	202	7e-51
emb CAA87083.1	acetohydroxyacid synthase [Gossypium hirsutum]	202	1e-50
gb ABM53020.1	acetolactate synthase [Amaranthus tuberculatus]	201	2e-50
gb ABR68867.1	acetohydroxyacid synthase [Solanum ptychanthum]	201	2e-50
gb AAG40279.1	AF308648_1 acetolactate synthase [Solanum ptychant...	200	3e-50
gb AAG40280.1	AF308649_1 acetolactate synthase [Solanum ptychant...	200	3e-50
gb AAC69629.1	herbicide resistant acetolactate synthase precurs...	200	3e-50
gb ABR68866.1	acetohydroxyacid synthase [Solanum ptychanthum]	200	3e-50
gb ABR68865.1	acetohydroxyacid synthase [Solanum ptychanthum]	200	3e-50
gb AAT07322.1	acetohydroxyacid synthase 1 [Helianthus annuus]	200	3e-50
gb ABN08612.1	Thiamine pyrophosphate enzyme, central region [Me...	198	9e-50
sp P14874 ILV2	BRANA Acetolactate synthase 2, chloroplast precu...	198	1e-49
gb AAB60297.1	acetolactate synthase precursor	198	1e-49
gb AAT07324.1	acetohydroxyacid synthase 1 [Helianthus annuus]	198	1e-49
gb AAA74913.1	acetolactate synthase precursor	198	1e-49
gb AAT07325.1	acetohydroxyacid synthase 1 [Helianthus annuus]	198	1e-49
gb AAT07323.1	acetohydroxyacid synthase 1 [Helianthus annuus]	198	1e-49
gb AAT07326.1	acetohydroxyacid synthase 1 [Helianthus annuus]	197	2e-49
sp P09342 ILV1	TOBAC Acetolactate synthase 1, chloroplast precu...	197	2e-49
prf 1407140A	acetolactate synthase SuRA	197	2e-49
gb ACF47582.1	acetolactate synthase 1 [Sonchus asper]	196	7e-49
gb ACF47583.1	acetolactate synthase 1 [Sonchus asper]	195	8e-49
gb ACF17639.1	putative acetolactate synthase [Capsicum annum]	195	9e-49
dbj BAF57909.1	acetolactate synthase [Sagittaria trifolia]	195	1e-48
gb AAT07327.1	acetohydroxyacid synthase 2 [Helianthus annuus]	194	1e-48
gb AAT07328.1	acetohydroxyacid synthase 2 [Helianthus annuus]	194	1e-48
gb ABY57317.1	acetolactate synthase [Medicago truncatula]	193	5e-48
gb ABY57316.1	acetolactate synthase [Medicago littoralis]	193	5e-48
gb ABY57315.1	acetolactate synthase [Medicago littoralis]	192	5e-48
gb ABY57318.1	acetolactate synthase [Medicago truncatula]	192	6e-48
dbj BAE97677.1	acetolactate synthase [Schoenoplectus juncooides]...	189	7e-47
gb AAO53549.1	acetohydroxyacid synthase [Triticum aestivum]	188	1e-46
gb AAO53551.1	acetohydroxyacid synthase [Triticum aestivum]	188	1e-46
emb CAE18088.1	acetolactate synthase [Papaver rhoeas]	188	1e-46
dbj BAE97675.1	acetolactate synthase [Schoenoplectus juncooides]...	188	1e-46
gb AAO53548.1	acetohydroxyacid synthase [Triticum aestivum]	187	2e-46
gb AAO53550.1	acetohydroxyacid synthase [Triticum aestivum]	187	2e-46
gb ABM92357.2	acetolactate synthase [Cyperus difformis]	187	3e-46
gb AAC14572.1	acetohydroxyacid synthase [Hordeum vulgare]	187	3e-46
dbj BAE53610.1	acetolactate synthase [Monochoria vaginalis]	187	3e-46
dbj BAE53611.1	acetolactate synthase [Monochoria vaginalis]	187	3e-46
dbj BAE53612.1	acetolactate synthase [Monochoria vaginalis]	187	3e-46
dbj BAE53605.1	acetolactate synthase [Monochoria vaginalis] >gi...	187	3e-46
dbj BAF37288.1	acetolactate synthase [Monochoria vaginalis]	187	3e-46
dbj BAE53608.1	acetolactate synthase [Monochoria vaginalis]	187	3e-46
gb AAM03119.1	AF488771_1 acetolactate synthase [Bromus tectorum]	186	4e-46
gb AAL93207.1	AF487459_1 acetolactate synthase [Bromus tectorum]	186	4e-46
dbj BAE53591.1	acetolactate synthase [Monochoria vaginalis] >gi...	186	4e-46
dbj BAE53593.1	acetolactate synthase [Monochoria vaginalis]	186	4e-46
dbj BAE53587.1	acetolactate synthase [Monochoria vaginalis] >gi...	186	4e-46
dbj BAE53588.1	acetolactate synthase [Monochoria vaginalis]	186	5e-46
gb EAZ23185.1	hypothetical protein OsJ_006668 [Oryza sativa (ja...	186	6e-46
gb EAY86003.1	hypothetical protein OsI_007236 [Oryza sativa (in...	186	7e-46
dbj BAB20813.1	acetolactate synthase [Oryza sativa (japonica cu...	186	7e-46
gb AAX14282.1	acetolactate synthase [Oryza sativa (japonica cul...	186	7e-46

ref NP_001046931.1	Os02g0510200 [Oryza sativa (japonica cultivar...]	185	8e-46
gb ABF66052.1	acetolactate synthase [Oryza sativa (indica cultivar...]	185	8e-46
gb AA14281.1	acetolactate synthase [Oryza sativa] >gi 99079641...	185	8e-46
gb ACD74789.1	acetolactate synthase [Oryza sativa Indica Group]...	185	8e-46
gb ABF66048.1	acetolactate synthase [Oryza sativa (indica cultivar...]	185	8e-46
gb ABF66051.1	acetolactate synthase [Oryza sativa (indica cultivar...]	185	8e-46
dbj BAF37287.1	acetolactate synthase [Monochoria vaginalis]	184	2e-45
gb AA14283.1	acetolactate synthase [Oryza sativa]	184	2e-45
gb AAT07329.1	acetohydroxyacid synthase 3 [Helianthus annuus]	184	2e-45
gb ABQ85869.1	acetolactate synthase [Lolium rigidum]	184	2e-45
gb AAG30931.1	AF310684_1 acetolactate synthase precursor [Lolium...]	183	3e-45
gb ABQ85870.1	acetolactate synthase [Lolium rigidum]	183	3e-45
dbj BAE53604.1	acetolactate synthase [Monochoria vaginalis]	182	1e-44
emb CAD24801.2	acetolactate synthase [Alopecurus myosuroides]	181	1e-44
dbj BAE53597.1	acetolactate synthase [Monochoria vaginalis] >gi...	181	2e-44
dbj BAE53596.1	acetolactate synthase [Monochoria vaginalis]	181	2e-44
dbj BAE53598.1	acetolactate synthase [Monochoria vaginalis] >gi...	181	2e-44
dbj BAE53602.1	acetolactate synthase [Monochoria vaginalis]	181	2e-44
sp Q41769 ILV2_MAIZE	Acetolactate synthase 2, chloroplast precu...	179	5e-44
gb ACF87819.1	unknown [Zea mays]	179	5e-44
gb EAY93888.1	hypothetical protein OsI_015121 [Oryza sativa (in...]	178	1e-43
sp Q41768 ILV1_MAIZE	Acetolactate synthase 1, chloroplast precu...	178	1e-43
gb EAZ30535.1	hypothetical protein OsJ_014018 [Oryza sativa (ja...]	178	1e-43
gb EAZ30530.1	hypothetical protein OsJ_014013 [Oryza sativa (ja...]	178	1e-43
sp Q7XKQ8 ILV2_ORYSJ	Probable acetolactate synthase 2, chloropla...	178	2e-43
emb CAH66432.1	OSIGBa0096P03.6 [Oryza sativa (indica cultivar-g...]	177	2e-43
gb ABF66049.1	acetolactate synthase [Oryza sativa (indica cultivar...]	176	7e-43
emb CAO49020.1	unnamed protein product [Vitis vinifera]	175	8e-43
gb ABR26179.1	acetolactate synthase III [Oryza sativa (indica c...]	173	4e-42
emb CAP09635.1	acetolactate-synthase-N-DnaE intein-N fusion pro...	171	1e-41
ref XP_001759950.1	predicted protein [Physcomitrella patens sub...]	162	7e-39
emb CAN73760.1	hypothetical protein [Vitis vinifera]	160	3e-38
ref XP_001758473.1	predicted protein [Physcomitrella patens sub...]	159	9e-38
dbj BAE53595.1	acetolactate synthase [Monochoria vaginalis]	150	4e-35
gb AAC04854.1	acetolactate synthase [Volvox carterii]	149	1e-34
gb AAB88296.1	acetolactate synthase [Volvox carterii]	148	1e-34
ref XP_001695168.1	acetolactate synthase, large subunit [Chlamy...	145	9e-34
gb AAC03784.1	acetolactate synthase [Chlamydomonas reinhardtii]	145	9e-34
gb AAB88292.1	acetolactate synthase [Chlamydomonas reinhardtii]	145	1e-33
emb CAO45071.1	unnamed protein product [Vitis vinifera]	144	2e-33
emb CAO17032.1	unnamed protein product [Vitis vinifera]	144	2e-33
emb CAL58226.1	acetolactate synthase 1 (ISS) [Ostreococcus tauri]	134	3e-30
ref XP_001421626.1	predicted protein [Ostreococcus lucimarinus ...]	132	1e-29
ref YP_001941018.1	Acetolactate synthase large subunit or other...	125	1e-27
ref NP_682086.1	acetolactate synthase 3 catalytic subunit [Ther...	125	2e-27
gb EDX84416.1	acetolactate synthase, large subunit, biosynthe...	122	1e-26
ref YP_502706.1	acetolactate synthase, large subunit, biosynthe...	119	8e-26
ref ZP_02963213.1	acetolactate synthase 1 catalytic subunit [Bi...	119	1e-25
ref YP_001293618.1	acetohydroxyacid synthetase large subunit [R...	118	1e-25
ref ZP_01731805.1	acetolactate synthase III large subunit [Cyan...	118	2e-25
ref YP_001936332.1	acetolactate synthase large subunit [Heteros...	118	2e-25
ref ZP_01857674.1	acetolactate synthase III (Precursor) [Planct...	117	2e-25
ref NP_849033.1	acetohydroxyacid synthase large subunit [Cyanid...	117	3e-25
ref YP_399158.1	acetolactate synthase 3 catalytic subunit [Syne...	117	4e-25
ref YP_172076.1	acetolactate synthase 3 catalytic subunit [Syne...	117	4e-25
ref NP_926225.1	acetolactate synthase 3 catalytic subunit [Gloe...	117	4e-25
ref ZP_01621589.1	acetolactate synthase III large subunit [Lyng...	116	5e-25
ref YP_001805892.1	acetolactate synthase [Cyanotheca sp. ATCC 5...	116	5e-25
ref ZP_01079763.1	acetolactate synthase [Synechococcus sp. RS99...	116	5e-25
emb CAO42659.1	unnamed protein product [Vitis vinifera]	116	6e-25
ref NP_897837.1	acetolactate synthase 3 catalytic subunit [Syne...	116	6e-25
ref ZP_02971818.1	acetolactate synthase, large subunit, biosynt...	116	6e-25
ref ZP_01085428.1	acetolactate synthase III large subunit [Syne...	115	1e-24
ref ZP_01469460.1	acetolactate synthase III large subunit [Syne...	115	1e-24
ref NP_488653.1	acetolactate synthase 3 catalytic subunit [Nost...	115	1e-24
ref YP_323265.1	acetolactate synthase 3 catalytic subunit [Anab...	115	1e-24
ref ZP_03272318.1	acetolactate synthase, large subunit, biosynt...	115	1e-24
ref YP_731202.1	acetolactate synthase 3 catalytic subunit [Syne...	115	1e-24
ref YP_001111655.1	acetolactate synthase, large subunit, biosyn...	115	1e-24
ref ZP_00514506.1	Acetolactate synthase, large subunit, biosynt...	114	2e-24

ref ZP_02737021.1	acetolactate synthase III [Gemmata obscuriglo...	114	2e-24
ref YP_001046985.1	acetolactate synthase, large subunit, biosyn...	114	2e-24
ref ZP_03142166.1	acetolactate synthase, large subunit, biosynt...	114	2e-24
ref ZP_02942745.1	acetolactate synthase, large subunit, biosynt...	114	3e-24
ref NP_441297.1	acetolactate synthase 3 catalytic subunit [Syne...	114	3e-24
ref YP_001517397.1	acetolactate synthase 3 catalytic subunit [A...	114	3e-24
ref NP_053940.1	acetohydroxyacid synthase large subunit [Porphy...	114	3e-24
ref NP_050806.1	acetohydroxyacid synthase large subunit [Guilla...	114	3e-24
ref NP_870771.1	acetolactate synthase III [precursor] [Rhodopir...	114	3e-24
ref YP_063568.1	acetohydroxyacid synthase large subunit [Gracil...	113	4e-24
ref ZP_03137434.1	acetolactate synthase, large subunit, biosynt...	113	5e-24
ref YP_001736021.1	acetolactate synthase, large subunit, biosyn...	113	5e-24
ref NP_618663.1	acetolactate synthase 3 catalytic subunit [Meth...	113	5e-24
ref YP_377644.1	acetolactate synthase 3 catalytic subunit [Syne...	113	5e-24
ref YP_722490.1	acetolactate synthase 3 catalytic subunit [Tric...	112	7e-24
ref YP_537012.1	acetohydroxyacid synthase large subunit [Porphy...	112	8e-24
ref YP_001930430.1	acetolactate synthase, large subunit, biosyn...	112	8e-24
ref YP_359378.1	acetolactate synthase, large subunit, biosynthe...	112	9e-24
ref YP_001227109.1	acetolactate synthase 3 catalytic subunit [S...	112	1e-23
ref YP_002249089.1	acetolactate synthase, large subunit, biosyn...	112	1e-23
ref YP_002250909.1	acetolactate synthase, large subunit, biosyn...	111	1e-23
ref YP_001211077.1	thiamine pyrophosphate-requiring enzymes [Pe...	111	2e-23
ref YP_303782.1	acetolactate synthase 3 catalytic subunit [Meth...	111	2e-23
gb EDX75556.1	acetolactate synthase, large subunit, biosyntheti...	111	2e-23
ref YP_754808.1	acetolactate synthase large subunit [Syntrophom...	111	2e-23
ref ZP_01472584.1	acetolactate synthase III large subunit [Syne...	110	3e-23
ref NP_045101.1	acetohydroxyacid synthase large subunit [Cyanid...	110	3e-23
ref ZP_01089063.1	acetolactate synthase III (Precursor) [Blasto...	110	3e-23
ref NP_632694.1	acetolactate synthase 3 catalytic subunit [Meth...	110	4e-23
ref YP_001660084.1	acetolactate synthase 3 catalytic subunit [M...	110	4e-23
gb AAA26594.1	acetohydroxy acid synthase (AHAS)	110	4e-23
ref ZP_01123205.1	acetolactate synthase III large subunit [Syne...	110	4e-23
emb CAA12081.1	acetohydroxy acid synthase [Porphyridium sp.]	110	4e-23
ref ZP_00206563.1	COG0028: Thiamine pyrophosphate-requiring enz...	110	5e-23
ref ZP_02028034.1	hypothetical protein BIFADO_00444 [Bifidobact...	110	5e-23
ref ZP_03155476.1	acetolactate synthase, large subunit, biosynt...	110	5e-23
ref ZP_03286153.1	acetolactate synthase, large subunit, biosynt...	110	5e-23
ref NP_695500.1	acetolactate synthase 1 catalytic subunit [Bifi...	109	5e-23
ref YP_909084.1	acetolactate synthase 1 catalytic subunit [Bifi...	109	5e-23
ref ZP_03129981.1	acetolactate synthase, large subunit, biosynt...	109	5e-23
ref YP_001865343.1	acetolactate synthase, large subunit, biosyn...	109	5e-23
ref YP_001379064.1	acetolactate synthase, large subunit, biosyn...	109	6e-23
ref ZP_02850743.1	acetolactate synthase, large subunit, biosynt...	109	7e-23
ref NP_789495.1	acetolactate synthase 1 catalytic subunit [Trop...	109	7e-23
emb CAO90567.1	ilvG [Microcystis aeruginosa PCC 7806]	109	9e-23
ref YP_002049511.1	acetolactate synthase III large subunit [Pau...	108	1e-22
ref ZP_01628200.1	acetolactate synthase III large subunit [Nodu...	108	1e-22
ref YP_475476.1	acetolactate synthase 3 catalytic subunit [Syne...	108	1e-22
ref YP_476898.1	acetolactate synthase 3 catalytic subunit [Syne...	108	1e-22
ref YP_001243976.1	acetolactate synthase, large subunit, biosyn...	108	1e-22
ref NP_895067.1	acetolactate synthase 3 catalytic subunit [Proc...	108	2e-22
ref YP_001224368.1	acetolactate synthase 3 catalytic subunit [S...	108	2e-22
ref YP_001030046.1	acetolactate synthase, large subunit [Methan...	108	2e-22
gb ABD96281.1	acetolactate synthase [uncultured marine type-A S...	107	2e-22
ref NP_787333.1	acetolactate synthase 1 catalytic subunit [Trop...	107	2e-22
ref ZP_01188450.1	Acetolactate synthase, large subunit, biosynt...	107	2e-22
ref ZP_02917198.1	hypothetical protein BIFDEN_00474 [Bifidobact...	107	2e-22
ref ZP_00828201.1	COG0028: Thiamine pyrophosphate-requiring enz...	107	2e-22
ref YP_001016784.1	acetolactate synthase 3 catalytic subunit [P...	107	3e-22
ref YP_431088.1	acetolactate synthase, large subunit [Moorella ...	107	3e-22
ref YP_001716542.1	acetolactate synthase, large subunit, biosyn...	107	4e-22
ref ZP_01389342.1	acetolactate synthase, large subunit, biosynt...	107	4e-22
sp Q59498 ILVB_MYCAV	Acetolactate synthase (Acetohydroxy-acid sy...	107	4e-22
ref YP_001004540.1	acetolactate synthase 2 catalytic subunit [Y...	106	5e-22
ref YP_842552.1	acetolactate synthase 3 catalytic subunit [Meth...	106	5e-22
ref YP_381046.1	acetolactate synthase 3 catalytic subunit [Syne...	106	5e-22
ref ZP_03185165.1	acetolactate synthase 1 catalytic subunit [St...	106	5e-22
ref YP_001232446.1	acetolactate synthase, large subunit, biosyn...	106	6e-22
ref ZP_03275274.1	acetolactate synthase, large subunit, biosynt...	106	6e-22
ref YP_002121866.1	acetolactate synthase, large subunit, biosyn...	106	6e-22

```
ref|NP_246567.1| acetolactate synthase 2 catalytic subunit [Past... 106 7e-22
ref|YP_001738428.1| acetolactate synthase, large subunit, biosyn... 106 7e-22
ref|NP_228358.1| acetolactate synthase, large subunit [Thermotog... 106 7e-22
ref|YP_001014410.1| acetolactate synthase 3 catalytic subunit [P... 106 7e-22
ref|YP_684797.1| acetolactate synthase, large subunit [unculture... 105 8e-22
ref|YP_293049.1| acetolactate synthase 3 catalytic subunit [Proc... 105 8e-22
ref|NP_961972.1| acetolactate synthase 1 catalytic subunit [Myco... 105 8e-22
```

```
>gb|ABJ80681.1| acetolactate synthase catalytic subunit [Arabidopsis thaliana]
Length = 670
```

Score = 242 bits (618), Expect = 6e-63, Method: Compositional matrix adjust.
Identities = 115/116 (99%), Positives = 116/116 (100%)

```
Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
          MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 348 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 407
```

```
Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
          HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEA+P
Sbjct: 408 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAIP 463
```

```
>gb|AAK68759.1| acetolactate synthase [Arabidopsis thaliana]
Length = 670
```

Score = 242 bits (618), Expect = 6e-63, Method: Compositional matrix adjust.
Identities = 115/116 (99%), Positives = 116/116 (100%)

```
Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
          MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 348 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 407
```

```
Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
          HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEA+P
Sbjct: 408 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAIP 463
```

```
>ref|NP_190425.1| CSR1 (CHLORSULFURON/IMIDAZOLINONE RESISTANT 1) [Arabidopsis
thaliana]
sp|P17597|ILVB_ARATH Acetolactate synthase, chloroplast precursor (Acetohydroxy-acid
synthase) (ALS)
emb|CAB62345.1| acetolactate synthase [Arabidopsis thaliana]
gb|AAW70386.1| At3g48560 [Arabidopsis thaliana]
prf||1501386B acetolactate synthase
Length = 670
```

Score = 242 bits (618), Expect = 6e-63, Method: Compositional matrix adjust.
Identities = 115/116 (99%), Positives = 116/116 (100%)

```
Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
          MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 348 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 407
```

```
Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
          HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEA+P
Sbjct: 408 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAIP 463
```

```
>gb|AAM92569.1| acetolactate synthase [Arabidopsis thaliana]
Length = 670
```

Score = 242 bits (618), Expect = 6e-63, Method: Compositional matrix adjust.
Identities = 115/116 (99%), Positives = 116/116 (100%)

```
Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
          MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 348 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 407
```

```
Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
```

HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEA+P
Sbjct: 408 HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAIP 463

>emb|CAA35887.1| unnamed protein product [Arabidopsis thaliana]
Length = 670

Score = 242 bits (618), Expect = 6e-63, Method: Compositional matrix adjust.
Identities = 115/116 (99%), Positives = 116/116 (100%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAVEHSDLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 348 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 407

Query: 61 HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEA+P
Sbjct: 408 HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAIP 463

>gb|AAT72502.1| AT3G48560 [Arabidopsis lyrata subsp. lyrata]
Length = 344

Score = 240 bits (613), Expect = 2e-62, Method: Compositional matrix adjust.
Identities = 115/116 (99%), Positives = 116/116 (100%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAVEHSDLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 197 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 256

Query: 61 HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEA+P
Sbjct: 257 HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAIP 312

>pdb|1YBH|A Chain A, Crystal Structure Of Arabidopsis Thaliana Acetohydroxyacid
Synthase In Complex With A Sulfonyleurea Herbicide
Chlorimuron Ethyl
pdb|1YHY|A Chain A, Crystal Structure Of Arabidopsis Thaliana Acetohydroxyacid
Synthase In Complex With A Sulfonyleurea Herbicide,
Metsulfuron Methyl
pdb|1YHZ|A Chain A, Crystal Structure Of Arabidopsis Thaliana Acetohydroxyacid
Synthase In Complex With A Sulfonyleurea Herbicide,
Chlorsulfuron
pdb|1YI0|A Chain A, Crystal Structure Of Arabidopsis Thaliana Acetohydroxyacid
Synthase In Complex With A Sulfonyleurea Herbicide,
Sulfometuron Methyl
pdb|1YI1|A Chain A, Crystal Structure Of Arabidopsis Thaliana Acetohydroxyacid
Synthase In Complex With A Sulfonyleurea Herbicide,
Tribenuron Methyl
pdb|1Z8N|A Chain A, Crystal Structure Of Arabidopsis Thaliana Acetohydroxyacid
Synthase In Complex With An Imidazolinone Herbicide,
Imazaquin
Length = 590

Score = 239 bits (611), Expect = 4e-62, Method: Compositional matrix adjust.
Identities = 115/116 (99%), Positives = 116/116 (100%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAVEHSDLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 263 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 322

Query: 61 HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEA+P
Sbjct: 323 HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAIP 378

>gb|AAR07632.1| acetolactate synthase 1 [Camelina microcarpa]
Length = 668

Score = 236 bits (601), Expect = 6e-61, Method: Compositional matrix adjust.

Identities = 112/116 (96%), Positives = 115/116 (99%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANY+VEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 346 MLGMHGT VYANYSV EHSDDL LA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 405

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWR+ELN QKQKFPLSFKTFGEA+P
Sbjct: 406 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRSELNEQKQKFPLSFKTFGEAIP 461

>sp|P27818|ILV1_BRANA Acetolactate synthase 1, chloroplast precursor (Acetolactate synthase I) (Acetohydroxy-acid synthase I) (ALS I)
emb|CAA77613.1| acetohydroxyacid synthase I [Brassica napus]
Length = 655

Score = 235 bits (600), Expect = 7e-61, Method: Compositional matrix adjust.
Identities = 112/116 (96%), Positives = 115/116 (99%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 333 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 392

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWR+EL+ QKQKFPLSFKTFGEA+P
Sbjct: 393 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRSELSEQKQKFPLSFKTFGEAIP 448

>sp|P27819|ILV3_BRANA Acetolactate synthase 3, chloroplast precursor (Acetolactate synthase III) (Acetohydroxy-acid synthase III) (ALS III)
emb|CAA77615.1| acetohydroxyacid synthase III [Brassica napus]
Length = 652

Score = 235 bits (600), Expect = 7e-61, Method: Compositional matrix adjust.
Identities = 112/116 (96%), Positives = 115/116 (99%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 330 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 389

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWR+EL+ QKQKFPLSFKTFGEA+P
Sbjct: 390 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRSELSEQKQKFPLSFKTFGEAIP 445

>gb|AAR07633.1| acetolactate synthase 1 [Camelina microcarpa]
Length = 668

Score = 235 bits (600), Expect = 7e-61, Method: Compositional matrix adjust.
Identities = 112/116 (96%), Positives = 115/116 (99%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANY+VEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 346 MLGMHGT VYANYSV EHSDDL LA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 405

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWR+ELN QKQKFPLSFKTFGEA+P
Sbjct: 406 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRSELNEQKQKFPLSFKTFGEAIP 461

>gb|AAA62705.1| acetolactate synthase
Length = 599

Score = 234 bits (598), Expect = 1e-60, Method: Compositional matrix adjust.
Identities = 112/116 (96%), Positives = 115/116 (99%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 277 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 336

Query: 61 HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWR+EL+ QKQKFPLSFKTFGEA+P
Sbjct: 337 HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRSELSEQKQKFPLSFKTFGEAIP 392

>gb|AA46005.1| putative acetolactate synthase [Sinapis arvensis]
Length = 573

Score = 234 bits (597), Expect = 2e-60, Method: Compositional matrix adjust.
Identities = 112/116 (96%), Positives = 115/116 (99%)

Query: 1 MLGMHGTIVYANYAVEHSDLLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAVEHSDLLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 254 MLGMHGTIVYANYAVEHSDLLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 313

Query: 61 HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWR+EL+ QKQKFPLSFKTFGEA+P
Sbjct: 314 HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRSELSEQKQKFPLSFKTFGEAIP 369

>gb|AA46012.1| putative acetolactate synthase [Sinapis arvensis]
Length = 577

Score = 234 bits (597), Expect = 2e-60, Method: Compositional matrix adjust.
Identities = 112/116 (96%), Positives = 115/116 (99%)

Query: 1 MLGMHGTIVYANYAVEHSDLLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAVEHSDLLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 258 MLGMHGTIVYANYAVEHSDLLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 317

Query: 61 HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWR+EL+ QKQKFPLSFKTFGEA+P
Sbjct: 318 HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRSELSEQKQKFPLSFKTFGEAIP 373

>gb|AA46013.1| putative acetolactate synthase [Sinapis arvensis]
Length = 575

Score = 234 bits (597), Expect = 2e-60, Method: Compositional matrix adjust.
Identities = 112/116 (96%), Positives = 115/116 (99%)

Query: 1 MLGMHGTIVYANYAVEHSDLLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAVEHSDLLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 258 MLGMHGTIVYANYAVEHSDLLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 317

Query: 61 HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWR+EL+ QKQKFPLSFKTFGEA+P
Sbjct: 318 HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRSELSEQKQKFPLSFKTFGEAIP 373

>gb|AA45995.1| putative acetolactate synthase [Sinapis arvensis]
gb|AA45996.1| putative acetolactate synthase [Sinapis arvensis]
gb|AA45997.1| putative acetolactate synthase [Sinapis arvensis]
gb|AA46000.1| putative acetolactate synthase [Sinapis arvensis]
gb|AA46001.1| putative acetolactate synthase [Sinapis arvensis]
gb|AA46002.1| putative acetolactate synthase [Sinapis arvensis]
gb|AA46003.1| putative acetolactate synthase [Sinapis arvensis]
Length = 574

Score = 234 bits (597), Expect = 2e-60, Method: Compositional matrix adjust.
Identities = 112/116 (96%), Positives = 115/116 (99%)

Query: 1 MLGMHGTIVYANYAVEHSDLLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAVEHSDLLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 258 MLGMHGTIVYANYAVEHSDLLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 317

Query: 61 HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWR+EL+ QKQKFPLSFKTFGEA+P
Sbjct: 318 HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRSELSEQKQKFPLSFKTFGEAIP 373

>gb|AA45998.1| putative acetolactate synthase [Sinapis arvensis]
gb|AA45999.1| putative acetolactate synthase [Sinapis arvensis]
Length = 574

Score = 234 bits (597), Expect = 2e-60, Method: Compositional matrix adjust.
Identities = 112/116 (96%), Positives = 115/116 (99%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 258 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 317

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRN ELNVQKQKFPLSF KTFGEAVP 116
HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWR+EL+ QKQKFPLSF KTFGEA+P
Sbjct: 318 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRSELSEQKQKFPLSF KTFGEAIP 373

>gb|AA46007.1| putative acetolactate synthase [Sinapis arvensis]
Length = 574

Score = 234 bits (597), Expect = 2e-60, Method: Compositional matrix adjust.
Identities = 112/116 (96%), Positives = 115/116 (99%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 256 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 315

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRN ELNVQKQKFPLSF KTFGEAVP 116
HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWR+EL+ QKQKFPLSF KTFGEA+P
Sbjct: 316 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRSELSEQKQKFPLSF KTFGEAIP 371

>gb|AA46008.1| putative acetolactate synthase [Sinapis arvensis]
Length = 575

Score = 234 bits (597), Expect = 2e-60, Method: Compositional matrix adjust.
Identities = 112/116 (96%), Positives = 115/116 (99%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 256 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 315

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRN ELNVQKQKFPLSF KTFGEAVP 116
HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWR+EL+ QKQKFPLSF KTFGEA+P
Sbjct: 316 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRSELSEQKQKFPLSF KTFGEAIP 371

>gb|AA46006.1| putative acetolactate synthase [Sinapis arvensis]
gb|AA46009.1| putative acetolactate synthase [Sinapis arvensis]
Length = 567

Score = 234 bits (597), Expect = 2e-60, Method: Compositional matrix adjust.
Identities = 112/116 (96%), Positives = 115/116 (99%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 248 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 307

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRN ELNVQKQKFPLSF KTFGEAVP 116
HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWR+EL+ QKQKFPLSF KTFGEA+P
Sbjct: 308 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRSELSEQKQKFPLSF KTFGEAIP 363

>gb|AA46004.1| putative acetolactate synthase [Sinapis arvensis]
Length = 576

Score = 234 bits (597), Expect = 2e-60, Method: Compositional matrix adjust.
Identities = 112/116 (96%), Positives = 115/116 (99%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60

MLGMHGTGYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 258 MLGMHGTGYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 317

Query: 61 HVSVCADVKLALQGMNVLENRAEELKLDGVRNENLVQKQKPLSFKTFGEAVP 116
HVSVCADVKLALQGMNVLENRAEELKLDGVRN+EL+ QKQKPLSFKTFGEA+P
Sbjct: 318 HVSVCADVKLALQGMNVLENRAEELKLDGVRNSELSEQKQKPLSFKTFGEAIP 373

>gb|AA46010.1| putative acetolactate synthase [Sinapis arvensis]
Length = 576

Score = 234 bits (596), Expect = 2e-60, Method: Compositional matrix adjust.
Identities = 112/116 (96%), Positives = 115/116 (99%)

Query: 1 MLGMHGTGYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTGYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 257 MLGMHGTGYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 316

Query: 61 HVSVCADVKLALQGMNVLENRAEELKLDGVRNENLVQKQKPLSFKTFGEAVP 116
HVSVCADVKLALQGMNVLENRAEELKLDGVRN+EL+ QKQKPLSFKTFGEA+P
Sbjct: 317 HVSVCADVKLALQGMNVLENRAEELKLDGVRNSELSEQKQKPLSFKTFGEAIP 372

>gb|AA46011.1| putative acetolactate synthase [Sinapis arvensis]
Length = 577

Score = 234 bits (596), Expect = 2e-60, Method: Compositional matrix adjust.
Identities = 112/116 (96%), Positives = 115/116 (99%)

Query: 1 MLGMHGTGYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTGYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 258 MLGMHGTGYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 317

Query: 61 HVSVCADVKLALQGMNVLENRAEELKLDGVRNENLVQKQKPLSFKTFGEAVP 116
HVSVCADVKLALQGMNVLENRAEELKLDGVRN+EL+ QKQKPLSFKTFGEA+P
Sbjct: 318 HVSVCADVKLALQGMNVLENRAEELKLDGVRNSELSEQKQKPLSFKTFGEAIP 373

>emb|CAC86692.1| putative acetolactate synthase [Raphanus raphanistrum]
Length = 585

Score = 234 bits (596), Expect = 2e-60, Method: Compositional matrix adjust.
Identities = 111/116 (95%), Positives = 115/116 (99%)

Query: 1 MLGMHGTGYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTGYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 263 MLGMHGTGYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 322

Query: 61 HVSVCADVKLALQGMNVLENRAEELKLDGVRNENLVQKQKPLSFKTFGEAVP 116
HVSVCADVKLALQGMNVLENRAEELKLDGVRN+EL+ QKQKPLSFKTFGEA+P
Sbjct: 323 HVSVCADVKLALQGMNVLENRAEELKLDGVRNSELSEQKQKPLSFKTFGEAIP 378

>gb|ACB12189.1| acetolactate synthase [Descurainia sophia]
Length = 566

Score = 234 bits (596), Expect = 2e-60, Method: Compositional matrix adjust.
Identities = 112/116 (96%), Positives = 114/116 (98%)

Query: 1 MLGMHGTGYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTGYANY+VEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 255 MLGMHGTGYANYVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 314

Query: 61 HVSVCADVKLALQGMNVLENRAEELKLDGVRNENLVQKQKPLSFKTFGEAVP 116
HVSVCADVKLALQGMNVLE RAEELKLDGVRNENLV QKQKPLSFKTFGEA+P
Sbjct: 315 HVSVCADVKLALQGMNVLEKRAEELKLDGVRNENLVQKQKPLSFKTFGEAIP 370

>gb|ACB12190.1| acetolactate synthase [Descurainia sophia]
Length = 566

Score = 234 bits (596), Expect = 2e-60, Method: Compositional matrix adjust.
Identities = 112/116 (96%), Positives = 114/116 (98%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANY+VEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 255 MLGMHGT VYANYSVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 314

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FG VWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVC GDVKLALQGMNKVLE RAEELKLD FG VWRNELN QKQKFPLSFKTFGEA+P
Sbjct: 315 HVSVC GDVKLALQGMNKVLEKRAEELKLD FG VWRNELNEQKQKFPLSFKTFGEAIP 370

>gb|ACB12188.1| acetolactate synthase [Descurainia sophia]
Length = 566

Score = 234 bits (596), Expect = 2e-60, Method: Compositional matrix adjust.
Identities = 112/116 (96%), Positives = 114/116 (98%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANY+VEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 255 MLGMHGT VYANYSVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 314

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FG VWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVC GDVKLALQGMNKVLE RAEELKLD FG VWRNELN QKQKFPLSFKTFGEA+P
Sbjct: 315 HVSVC GDVKLALQGMNKVLEKRAEELKLD FG VWRNELNEQKQKFPLSFKTFGEAIP 370

>gb|AAR06607.1| acetolactate synthase 2 [Camelina microcarpa]
Length = 665

Score = 233 bits (595), Expect = 2e-60, Method: Compositional matrix adjust.
Identities = 111/116 (95%), Positives = 114/116 (98%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANY+VEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 343 MLGMHGT VYANYSVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 402

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FG VWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVC GDVKLALQGMNKVLENR EELKLD FG VWR+ELN QKQKFPLSFKTFGEA+P
Sbjct: 403 HVSVC GDVKLALQGMNKVLENRGEELKLD FG VWRSELNEQKQKFPLSFKTFGEAIP 458

>emb|CAC86696.1| putative acetolactate synthase [Raphanus raphanistrum]
Length = 585

Score = 233 bits (595), Expect = 3e-60, Method: Compositional matrix adjust.
Identities = 111/116 (95%), Positives = 114/116 (98%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 263 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 322

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FG VWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVC GDVKLALQGMNK+LENRAEELKLD FG VWR+EL QKQKFPLSFKTFGEA+P
Sbjct: 323 HVSVC GDVKLALQGMNKILENRAEELKLD FG VWRSELKEQKQKFPLSFKTFGEAIP 378

>emb|CAC86703.1| putative acetolactate synthase [Raphanus raphanistrum]
Length = 585

Score = 233 bits (594), Expect = 4e-60, Method: Compositional matrix adjust.
Identities = 110/116 (94%), Positives = 114/116 (98%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANY+VEHSD LLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 263 MLGMHGT VYANYSVEHSDXLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 322

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FG VWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVC GDVKLALQGMNK+LENRAEELKLD FG VWR+ELN QKQKFPLSFKTFGEA+P

Sbjct: 323 HVSVCADVKLALQGMNKILENRAEELKLDGFGVWRSELNEQKQKFPLSFKTFGEAIP 378

>emb|CAC86697.1| putative acetolactate synthase [Raphanus raphanistrum]
Length = 409

Score = 232 bits (592), Expect = 6e-60, Method: Compositional matrix adjust.
Identities = 111/116 (95%), Positives = 114/116 (98%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 87 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 146

Query: 61 HVSVCADVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVCADVKLALQGMNK+LENRAEELKLDGFGVWR+EL QKQKFPLSFKTFGEA+P
Sbjct: 147 HVSVCADVKLALQGMNKILENRAEELKLDGFGVWRSELKEQKQKFPLSFKTFGEAIP 202

>emb|CAC86701.1| putative acetolactate synthase [Raphanus raphanistrum]
Length = 585

Score = 231 bits (590), Expect = 1e-59, Method: Compositional matrix adjust.
Identities = 109/116 (93%), Positives = 115/116 (99%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANY+VEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 263 MLGMHGT VYANY+VEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 322

Query: 61 HVSVCADVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVCADVKLALQGMN++LENRAEELKLDGFGVWR+EL+ QKQKFPLSFKTFGEA+P
Sbjct: 323 HVSVCADVKLALQGMNEILENRAEELKLDGFGVWRSELSEQKQKFPLSFKTFGEAIP 378

>emb|CAC86698.1| putative acetolactate synthase [Raphanus raphanistrum]
Length = 585

Score = 231 bits (589), Expect = 1e-59, Method: Compositional matrix adjust.
Identities = 109/116 (93%), Positives = 115/116 (99%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANY+VEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 263 MLGMHGT VYANY+VEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 322

Query: 61 HVSVCADVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVCADVKLALQGMN++LENRAEELKLDGFGVWR+EL+ QKQKFPLSFKTFGEA+P
Sbjct: 323 HVSVCADVKLALQGMNEILENRAEELKLDGFGVWRSELSEQKQKFPLSFKTFGEAIP 378

>emb|CAC86695.1| putative acetolactate synthase [Raphanus raphanistrum]
Length = 585

Score = 231 bits (589), Expect = 1e-59, Method: Compositional matrix adjust.
Identities = 109/116 (93%), Positives = 114/116 (98%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANY+VEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 263 MLGMHGT VYANY+VEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 322

Query: 61 HVSVCADVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVCADVKLALQGMNK+LENR EELKLDGFGVWR+EL+ QKQKFPLSFKTFGEA+P
Sbjct: 323 HVSVCADVKLALQGMNKILENRGEELKLDGFGVWRSELSEQKQKFPLSFKTFGEAIP 378

>emb|CAC86700.1| putative acetolactate synthase [Raphanus raphanistrum]
Length = 585

Score = 231 bits (589), Expect = 1e-59, Method: Compositional matrix adjust.
Identities = 109/116 (93%), Positives = 115/116 (99%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60

MLGMHGTIVYANY+VEHSDLLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 263 MLGMHGTIVYANYVEHSDLLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 322

Query: 61 HVSVCADVKLALQGMNKVLENRAEELKLDGFWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVCADVKLALQGMN++LENRAEELKLDGFWR+EL+ QKQKFPLSFKTFGEA+P
Sbjct: 323 HVSVCADVKLALQGMNILENRAEELKLDGFWRSELSEQKQKFPLSFKTFGEAIP 378

>emb|CAC86694.1| putative acetolactate synthase [Raphanus raphanistrum]
Length = 585

Score = 231 bits (589), Expect = 1e-59, Method: Compositional matrix adjust.
Identities = 109/116 (93%), Positives = 114/116 (98%)

Query: 1 MLGMHGTIVYANYAVEHSDLLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANY+VEHSDLLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 263 MLGMHGTIVYANYVEHSDLLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 322

Query: 61 HVSVCADVKLALQGMNKVLENRAEELKLDGFWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVCADVKLALQGMNK+LENR EELKLDGFWR+EL+ QKQKFPLSFKTFGEA+P
Sbjct: 323 HVSVCADVKLALQGMNILENRAEELKLDGFWRSELSEQKQKFPLSFKTFGEAIP 378

>emb|CAC86702.1| putative acetolactate synthase [Raphanus raphanistrum]
Length = 585

Score = 230 bits (586), Expect = 3e-59, Method: Compositional matrix adjust.
Identities = 109/116 (93%), Positives = 114/116 (98%)

Query: 1 MLGMHGTIVYANYAVEHSDLLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANY+VEHSDLLLAFGVRFDDRTGKLEA ASRAKIVHIDIDSAEIGKNKTP
Sbjct: 263 MLGMHGTIVYANYVEHSDLLLAFGVRFDDRTGKLEAXASRAKIVHIDIDSAEIGKNKTP 322

Query: 61 HVSVCADVKLALQGMNKVLENRAEELKLDGFWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVCADVKLALQGMNK+LENRAEELKLDGFWR+EL+ QKQKFPLSFKTFGEA+P
Sbjct: 323 HVSVCADVKLALQGMNILENRAEELKLDGFWRSELSEQKQKFPLSFKTFGEAIP 378

>emb|CAC86699.1| putative acetolactate synthase [Raphanus raphanistrum]
Length = 585

Score = 229 bits (585), Expect = 4e-59, Method: Compositional matrix adjust.
Identities = 111/125 (88%), Positives = 116/125 (92%)

Query: 1 MLGMHGTIVYANYAVEHSDLLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANY+VEHSDLLLAFGVRFDDRTGKLEA ASRAKIVHIDIDSAEIGKNKTP
Sbjct: 263 MLGMHGTIVYANYVEHSDLLLAFGVRFDDRTGKLEAXASRAKIVHIDIDSAEIGKNKTP 322

Query: 61 HVSVCADVKLALQGMNKVLENRAEELKLDGFWRNELNVQKQKFPLSFKTFGEAVPCPSK 120
HVSVCADVKLALQGMNK+LENRAEELKLDGFWR+EL+ QKQKFPLSFKTFGEA+P
Sbjct: 323 HVSVCADVKLALQGMNILENRAEELKLDGFWRSELSEQKQKFPLSFKTFGEAIPQYA 382

Query: 121 KTVHD 125
V D
Sbjct: 383 XQVLD 387

>emb|CAC86693.1| putative acetolactate synthase [Raphanus raphanistrum]
Length = 585

Score = 229 bits (584), Expect = 5e-59, Method: Compositional matrix adjust.
Identities = 108/116 (93%), Positives = 114/116 (98%)

Query: 1 MLGMHGTIVYANYAVEHSDLLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANY+VEHSDLLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 263 MLGMHGTIVYANYVEHSDLLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 322

Query: 61 HVSVCADVKLALQGMNKVLENRAEELKLDGFWRNELNVQKQKFPLSFKTFGEAVP 116
HVSVCADVKLALQGMN++LENRAEELKLDGFWR+EL+ QKQKFPLSFKTFGEA+P
Sbjct: 323 HVSVCADVKLALQGMNILENRAEELKLDGFWRSELSEQKQKFPLSFKTFGEAIP 378

>gb|ABM53021.1| acetolactate synthase [Amaranthus tuberculatus]
Length = 671

Score = 207 bits (526), Expect = 3e-52, Method: Compositional matrix adjust.
Identities = 96/116 (82%), Positives = 108/116 (93%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNK P
Sbjct: 349 MLGMHGT VYANYAVDKADLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKQP 408

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+CGD+K+ALQG+NK+LE+R +KLDF WR ELN QK+KFPLSFKTFG+A+P
Sbjct: 409 HVSICGDIKVALQGLNKILES SRKGKLD FSNWREELNEQKKKPLSFKTFGDAIP 464

>gb|AAK50820.1|AF363369_1 acetolactate synthase [Amaranthus retroflexus]
Length = 669

Score = 206 bits (525), Expect = 3e-52, Method: Compositional matrix adjust.
Identities = 96/116 (82%), Positives = 108/116 (93%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNK P
Sbjct: 347 MLGMHGT VYANYAVDKADLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKQP 406

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+CGDVK+ALQG+NK+LE+R ++KLDF WR ELN QK+KFPLSFKTFG+A+P
Sbjct: 407 HVSICGDVKVALQGLNKILES SRKGKLD FSNWREELNEQKKKPLSFKTFGDAIP 462

>gb|AAK50821.1|AF363370_1 acetolactate synthase [Amaranthus powellii]
Length = 669

Score = 206 bits (525), Expect = 4e-52, Method: Compositional matrix adjust.
Identities = 96/116 (82%), Positives = 108/116 (93%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNK P
Sbjct: 347 MLGMHGT VYANYAVDKADLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKQP 406

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+CGDVK+ALQG+NK+LE+R ++KLDF WR ELN QK+KFPLSFKTFG+A+P
Sbjct: 407 HVSICGDVKVALQGLNKILES SRKGKLD FSNWREELNEQKKKPLSFKTFGDAIP 462

>gb|ABS72164.1| acetolactate synthase [Amaranthus hypochondriacus]
Length = 669

Score = 206 bits (524), Expect = 4e-52, Method: Compositional matrix adjust.
Identities = 96/116 (82%), Positives = 108/116 (93%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNK P
Sbjct: 347 MLGMHGT VYANYAVDKADLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKQP 406

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+CGDVK+ALQG+NK+LE+R ++KLDF WR ELN QK+KFPLSFKTFG+A+P
Sbjct: 407 HVSICGDVKVALQGLNKILES SRKGKLD FSNWREELNEQKKKPLSFKTFGDAIP 462

>gb|AAB67839.1| acetolactate synthase precursor [Amaranthus sp.]
Length = 665

Score = 204 bits (519), Expect = 2e-51, Method: Compositional matrix adjust.
Identities = 94/116 (81%), Positives = 108/116 (93%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNK P

Sbjct: 343 MLGMHGT VYANYAVDKADLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKQP 402

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF KTFGEAVP 116
HVS+CG++K+ALQG+NK+LE+R +LKLD F WR ELN QK+KFPLSFK+FG+A+P

Sbjct: 403 HVSICGEIKVALQGLNKILES RKGK LKLD FSNWREELNEQKKKFPLSFKSFGDAIP 458

>emb|CAA87084.1| acetohydroxyacid synthase [Gossypium hirsutum]
Length = 659

Score = 203 bits (517), Expect = 3e-51, Method: Compositional matrix adjust.
Identities = 96/116 (82%), Positives = 104/116 (89%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ SDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNK P

Sbjct: 337 MLGMHGT VYANYAVDKADLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKQP 396

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF KTFGEAVP 116
HVSVC DVKLALQG+NK+LE + +L LD+ WR ELN QK KFPLS+KTFGEA+P

Sbjct: 397 HVSVCSDVKLALQGIN KILETKVAKLNL DYSEWRQELNEQKLKFPLSYKTFGEAIP 452

>gb|ABM53018.1| acetolactate synthase [Amaranthus tuberculatus]
Length = 668

Score = 202 bits (515), Expect = 5e-51, Method: Compositional matrix adjust.
Identities = 94/116 (81%), Positives = 107/116 (92%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNK P

Sbjct: 346 MLGMHGT VYANYAVDKADLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKQP 405

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF KTFGEAVP 116
HVS+CGDVK+AL+G+N +LE+R ++KLDF WR ELN QK+KFPLSFKTFG+A+P

Sbjct: 406 HVSICGDVKVALRGLNNILES RKGKVKLD FSNWREELNEQKKKFPLSFKTFGDAIP 461

>gb|ABM53019.1| acetolactate synthase [Amaranthus tuberculatus]
Length = 669

Score = 202 bits (515), Expect = 5e-51, Method: Compositional matrix adjust.
Identities = 94/116 (81%), Positives = 107/116 (92%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNK P

Sbjct: 347 MLGMHGT VYANYAVDKADLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKQP 406

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF KTFGEAVP 116
HVS+CGDVK+AL+G+N +LE+R ++KLDF WR ELN QK+KFPLSFKTFG+A+P

Sbjct: 407 HVSICGDVKVALRGLNNILES RKGKVKLD FSNWREELNEQKKKFPLSFKTFGDAIP 462

>prf||1407140B acetolactate synthase SuRB
Length = 664

Score = 202 bits (514), Expect = 7e-51, Method: Compositional matrix adjust.
Identities = 93/116 (80%), Positives = 105/116 (90%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ SDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNK P

Sbjct: 342 MLGMHGT VYANYAVDSSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKQP 401

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF KTFGEAVP 116
HVS+C D+KLALQG+N +LE++ +LKLD F WR EL VQK K+PL+FKTFG+A+P

Sbjct: 402 HVSICADIKLALQGLN SILESKEGK LKLD FSAWRQELTVQKV KYP LNF KTFGDAIP 457

>sp|P09114| ILV2_TOBAC Acetolactate synthase 2, chloroplast precursor (Acetolactate
synthase II) (Acetohydroxy-acid synthase II) (ALS II)
emb|CAA30485.1| unnamed protein product [Nicotiana tabacum]

Length = 664

Score = 202 bits (514), Expect = 7e-51, Method: Compositional matrix adjust.
Identities = 93/116 (80%), Positives = 105/116 (90%)Query: 1 MLGMHGT VYANYAVEHS D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K T P 60
MLGMHGT VYANYAV+ SD L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K P
Sbjct: 342 MLGMHGT VYANYAV D S S D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K Q P 401Query: 61 HVS V C G D V K L A L Q G M N K V L E N R A E E L K L D F G V W R N E L N V Q K Q K F P L S F K T F G E A V P 116
HVS+C D+K L A L Q G+N +L E++ +L K L D F W R E L V Q K K+P L+F K T F G+A+P
Sbjct: 402 HVS I C A D I K L A L Q G L N S I L E S K E G K L K L D F S A W R Q E L T V Q K V K Y P L N F K T F G D A I P 457>emb|CAA87083.1| acetohydroxyacid synthase [Gossypium hirsutum]
Length = 659Score = 202 bits (513), Expect = 1e-50, Method: Compositional matrix adjust.
Identities = 95/116 (81%), Positives = 103/116 (88%)Query: 1 MLGMHGT VYANYAVEHS D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K T P 60
MLGMHGT VYANYAV+ SD L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K P
Sbjct: 337 MLGMHGT VYANYAV D K S D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K Q P 396Query: 61 HVS V C G D V K L A L Q G M N K V L E N R A E E L K L D F G V W R N E L N V Q K Q K F P L S F K T F G E A V P 116
H+SVC D V K L A L Q G+N K+L E +L L D+ W R E L N Q K K F P L S+K T F G E A+P
Sbjct: 397 H M S V C S D V K L A L Q G I N K I L E T T G A K L N L D Y S E W R Q E L N E Q K L K F P L S Y K T F G E A I P 452>gb|ABM53020.1| acetolactate synthase [Amaranthus tuberculatus]
Length = 669Score = 201 bits (511), Expect = 2e-50, Method: Compositional matrix adjust.
Identities = 93/116 (80%), Positives = 106/116 (91%)Query: 1 MLGMHGT VYANYAVEHS D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K T P 60
MLGMHGT VYANYAV+ +D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K P
Sbjct: 347 MLGMHGT VYANYAV D K A D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K Q P 406Query: 61 HVS V C G D V K L A L Q G M N K V L E N R A E E L K L D F G V W R N E L N V Q K Q K F P L S F K T F G E A V P 116
HVS+CGD K+AL+G+N +L E+R ++K L D F W R E L N Q K+K F P L S F K T F G+A+P
Sbjct: 407 HVS I C G D X K V A L R G L N N I L E S R K G K V K L D F S N W R E E L N E Q K K K F P L S F K T F G D A I P 462>gb|ABR68867.1| acetohydroxyacid synthase [Solanum ptychanthum]
Length = 554Score = 201 bits (510), Expect = 2e-50, Method: Compositional matrix adjust.
Identities = 93/116 (80%), Positives = 103/116 (88%)Query: 1 MLGMHGT VYANYAVEHS D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K T P 60
MLGMHGT VYANYAV+ SD L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K P
Sbjct: 242 MLGMHGT VYANYAV D S S D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K L P 301Query: 61 HVS V C G D V K L A L Q G M N K V L E N R A E E L K L D F G V W R N E L N V Q K Q K F P L S F K T F G E A V P 116
H S+C D+K L A L Q G+N +L E + +L K L D F W R E L N Q K K+P L+F K T F G E A+P
Sbjct: 302 H A S I C A D I K L A L Q G L N S I L E G K E G K L K L D F S A W R Q E L N E Q K V K Y P L N F K T F G E A I P 357>gb|AAG40279.1|AF308648_1 acetolactate synthase [Solanum ptychanthum]
Length = 567Score = 200 bits (509), Expect = 3e-50, Method: Compositional matrix adjust.
Identities = 93/116 (80%), Positives = 103/116 (88%)Query: 1 MLGMHGT VYANYAVEHS D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K T P 60
MLGMHGT VYANYAV+ SD L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K P
Sbjct: 253 MLGMHGT VYANYAV D S S D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K L P 312

Query: 61 HVS V C G D V K L A L Q G M N K V L E N R A E E L K L D F G V W R N E L N V Q K Q K F P L S F K T F G E A V P 116

H S+C D+KLALQG+N +LE + +LKLDF WR ELN QK K+PL+FKTFGEA+P
Sbjct: 313 HASICADIKLALQGLNSILEGKEGKCLKLDFSAWRQELNEQVKYPLNFKTFGEAIP 368

>gb|AAG40280.1|AF308649_1 acetolactate synthase [Solanum ptychanthum]
gb|AAG40281.1|AF308650_1 acetolactate synthase [Solanum ptychanthum]
Length = 567

Score = 200 bits (509), Expect = 3e-50, Method: Compositional matrix adjust.
Identities = 93/116 (80%), Positives = 103/116 (88%)

Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV+ SDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNK P
Sbjct: 253 MLGMHGTIVYANYAVDSSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKLP 312

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGFWRNELNVQKQKFPPLSFKTFGEAVP 116
H S+C D+KLALQG+N +LE + +LKLDF WR ELN QK K+PL+FKTFGEA+P
Sbjct: 313 HASICADIKLALQGLNSILEGKEGKCLKLDFSAWRQELNEQVKYPLNFKTFGEAIP 368

>gb|AAC69629.1| herbicide resistant acetolactate synthase precursor [Bassia
scoparia]
Length = 666

Score = 200 bits (509), Expect = 3e-50, Method: Compositional matrix adjust.
Identities = 94/116 (81%), Positives = 105/116 (90%)

Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV+ +DLLLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNK P
Sbjct: 344 MLGMHGTIVYANYAVDKADLLLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKQP 403

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGFWRNELNVQKQKFPPLSFKTFGEAVP 116
HVS+C DVK AL+GMNK+LE+R +LKL++ WR EL QK+KFPLSFKTFGEA+P
Sbjct: 404 HVSICADVYALKGMNKILESRKGLKLNYSWREELGEQKKKFPPLSFKTFGEAIP 459

>gb|ABR68866.1| acetohydroxyacid synthase [Solanum ptychanthum]
Length = 565

Score = 200 bits (509), Expect = 3e-50, Method: Compositional matrix adjust.
Identities = 93/116 (80%), Positives = 103/116 (88%)

Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV+ SDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNK P
Sbjct: 253 MLGMHGTIVYANYAVDSSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKLP 312

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGFWRNELNVQKQKFPPLSFKTFGEAVP 116
H S+C D+KLALQG+N +LE + +LKLDF WR ELN QK K+PL+FKTFGEA+P
Sbjct: 313 HASICADIKLALQGLNSILEGKEGKCLKLDFSAWRQELNEQVKYPLNFKTFGEAIP 368

>gb|ABR68865.1| acetohydroxyacid synthase [Solanum ptychanthum]
Length = 565

Score = 200 bits (509), Expect = 3e-50, Method: Compositional matrix adjust.
Identities = 93/116 (80%), Positives = 103/116 (88%)

Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV+ SDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNK P
Sbjct: 253 MLGMHGTIVYANYAVDSSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKLP 312

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGFWRNELNVQKQKFPPLSFKTFGEAVP 116
H S+C D+KLALQG+N +LE + +LKLDF WR ELN QK K+PL+FKTFGEA+P
Sbjct: 313 HASICADIKLALQGLNSILEGKEGKCLKLDFSAWRQELNEQVKYPLNFKTFGEAIP 368

>gb|AAT07322.1| acetohydroxyacid synthase 1 [Helianthus annuus]
Length = 655

Score = 200 bits (508), Expect = 3e-50, Method: Compositional matrix adjust.

Identities = 93/116 (80%), Positives = 102/116 (87%)

Query: 1 MLGMHGT VYANYAVEHS D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K T P 60
MLGMHGT VYANYAV+ S D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D A E I G K N K P
Sbjct: 333 MLGMHGT VYANYAV D K S D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D P A E I G K N K Q P 392

Query: 61 H V S V C G D V K L A L Q G M N K V L E N R A E E L K L D F G V W R N E L N V Q K Q K F P L S F K T F G E A V P 116
H V S + C G D + K + A L Q G + N K + L E + L D F W R E L + Q K K F P L S F K T F G E A + P
Sbjct: 393 H V S I C G D I K V A L Q G L N K I L E E K N S V T N L D F S T W R K E L D E Q K M K F P L S F K T F G E A I P 448

>gb|ABN08612.1| Thiamine pyrophosphate enzyme, central region [Medicago truncatula]
Length = 655

Score = 198 bits (504), Expect = 9e-50, Method: Compositional matrix adjust.
Identities = 93/116 (80%), Positives = 105/116 (90%)

Query: 1 MLGMHGT VYANYAVEHS D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K T P 60
MLGMHGT VYANYAV+ S D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D A E I G K N K P
Sbjct: 333 MLGMHGT VYANYAV D K S D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D A A E I G K N K Q P 392

Query: 61 H V S V C G D V K L A L Q G M N K V L E N R A E E L K L D F G V W R N E L N V Q K Q K F P L S F K T F G E A V P 116
H V S V C G D + K L A L + G + N + + L E + E K L D F G W R E L N Q K + F P + S + K T F E A + P
Sbjct: 393 H V S V C G D L K L A L K G I N Q I L E S N G I E R K L D F G G W R E E L N D Q K V R F P M S Y K T F D E A I P 448

>sp|P14874| ILV2_BRANA Acetolactate synthase 2, chloroplast precursor (Acetolactate synthase II) (Acetohydroxy-acid synthase II) (ALS II)
emb|CAA77614.1| acetohydroxyacid synthase II [Brassica napus]
emb|CAA34680.1| unnamed protein product [Brassica napus]
Length = 637

Score = 198 bits (504), Expect = 1e-49, Method: Compositional matrix adjust.
Identities = 97/116 (83%), Positives = 104/116 (89%), Gaps = 2/116 (1%)

Query: 1 MLGMHGT VYANYAVEHS D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K T P 60
MLGMHGT VYANYAVE+ S D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S E I G K N K T P
Sbjct: 325 MLGMHGT VYANYAVE Y S D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S T E I G K N K T P 384

Query: 61 H V S V C G D V K L A L Q G M N K V L E N R A E E L K L D F G V W R N E L N V Q K Q K F P L S F K T F G E A V P 116
H V S V C D V + L A L Q G M N + V L E N R + L D F G W R E L N Q + K F P L + K T F G E + P
Sbjct: 385 H V S V C C D V Q L A L Q G M N E V L E N R R D - - V L D F G E W R C E L N E Q R L K F P L R Y K T F G E E I P 438

>gb|AAB60297.1| acetolactate synthase precursor
Length = 648

Score = 198 bits (503), Expect = 1e-49, Method: Compositional matrix adjust.
Identities = 93/116 (80%), Positives = 103/116 (88%)

Query: 1 MLGMHGT VYANYAVEHS D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K T P 60
MLGMHGT VYANYAV+ S D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K P
Sbjct: 326 MLGMHGT VYANYAV D K S D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K Q P 385

Query: 61 H V S V C G D V K L A L Q G M N K V L E N R A E E L K L D F G V W R N E L N V Q K Q K F P L S F K T F G E A V P 116
H V S + C G D + K + A L Q G + N K + L E + L D F W R E L + Q K K + P L S F K T F G E A + P
Sbjct: 386 H V S I C G D I K V A L Q G L N K I L E V K N S V T N L D F S N W R K E L D E Q K V K Y P L S F K T F G E A I P 441

>gb|AAT07324.1| acetohydroxyacid synthase 1 [Helianthus annuus]
Length = 658

Score = 198 bits (503), Expect = 1e-49, Method: Compositional matrix adjust.
Identities = 93/116 (80%), Positives = 102/116 (87%)

Query: 1 MLGMHGT VYANYAVEHS D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K T P 60
MLGMHGT VYANYAV+ S D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D A E I G K N K P
Sbjct: 336 MLGMHGT VYANYAV D K S D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D P A E I G K N K Q P 395

Query: 61 H V S V C G D V K L A L Q G M N K V L E N R A E E L K L D F G V W R N E L N V Q K Q K F P L S F K T F G E A V P 116

HVS+CGD+K+ALQG+NK+LE + LDF WR EL+ QK KFPLSFKTFGEA+P
Sbjct: 396 HVSICGDIKVALQGLNKILEEKNSVTNLDNFSNWRKELDEQVKVPLSFKTFGEAIP 451

>gb|AAA74913.1| acetolactate synthase precursor
Length = 648

Score = 198 bits (503), Expect = 1e-49, Method: Compositional matrix adjust.
Identities = 93/116 (80%), Positives = 103/116 (88%)

Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV+ SDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNK P
Sbjct: 326 MLGMHGTIVYANYAVDKSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKQP 385

Query: 61 HVSVCADVKLALQGMNVLENRAEELKLDGFWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+CGD+K+ALQG+NK+LE + LDF WR EL+ QK K+PLSFKTFGEA+P
Sbjct: 386 HVSICGDIKVALQGLNKILEVKNSTNLDNFSNWRKELDEQVKVPLSFKTFGEAIP 441

>gb|AAT07325.1| acetohydroxyacid synthase 1 [Helianthus annuus]
Length = 652

Score = 198 bits (503), Expect = 1e-49, Method: Compositional matrix adjust.
Identities = 93/116 (80%), Positives = 102/116 (87%)

Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV+ SDLLAFGVRFDDRTVGKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 330 MLGMHGTIVYANYAVDKSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDPAEIGKNKQP 389

Query: 61 HVSVCADVKLALQGMNVLENRAEELKLDGFWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+CGD+K+ALQG+NK+LE + LDF WR EL+ QK KFPLSFKTFGEA+P
Sbjct: 390 HVSICGDIKVALQGLNKILEEKNSVTNLDNFSNWRKELDEQVKVPLSFKTFGEAIP 445

>gb|AAT07323.1| acetohydroxyacid synthase 1 [Helianthus annuus]
Length = 654

Score = 198 bits (503), Expect = 1e-49, Method: Compositional matrix adjust.
Identities = 93/116 (80%), Positives = 102/116 (87%)

Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV+ SDLLAFGVRFDDRTVGKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 332 MLGMHGTIVYANYAVDKSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDPAEIGKNKQP 391

Query: 61 HVSVCADVKLALQGMNVLENRAEELKLDGFWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+CGD+K+ALQG+NK+LE + LDF WR EL+ QK KFPLSFKTFGEA+P
Sbjct: 392 HVSICGDIKVALQGLNKILEEKNSVTNLDNFSNWRKELDEQVKVPLSFKTFGEAIP 447

>gb|AAT07326.1| acetohydroxyacid synthase 1 [Helianthus annuus]
Length = 652

Score = 197 bits (502), Expect = 2e-49, Method: Compositional matrix adjust.
Identities = 93/116 (80%), Positives = 102/116 (87%)

Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV+ SDLLAFGVRFDDRTVGKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 330 MLGMHGTIVYANYAVDKSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDPAEIGKNKQP 389

Query: 61 HVSVCADVKLALQGMNVLENRAEELKLDGFWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+CGD+K+ALQG+NK+LE + LDF WR EL+ QK KFPLSFKTFGEA+P
Sbjct: 390 HVSICGDIKVALQGLNKILEEKNSVTNLDNFSNWRKELDEQVKVPLSFKTFGEAIP 445

>sp|P09342|ILV1_TOBAC Acetolactate synthase 1, chloroplast precursor (Acetolactate
synthase I) (Acetohydroxy-acid synthase I) (ALS I)
emb|CAA30484.1| unnamed protein product [Nicotiana tabacum]
prf|1501386A acetolactate synthase
Length = 667

Score = 197 bits (501), Expect = 2e-49, Method: Compositional matrix adjust.
Identities = 92/116 (79%), Positives = 103/116 (88%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ SDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNK P
Sbjct: 345 MLGMHGT VYANYAVDSSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKQP 404

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+C D+KLALQG+N +LE++ +LKLD F WR EL QK K PL+FKTFG+A+P
Sbjct: 405 HVSICADIKLALQGLNSILESKEGKLKLD FSAWRQELTEQKVKHPLNFKTFGDAIP 460

>prf||1407140A acetolactate synthase SuRA
Length = 667

Score = 197 bits (501), Expect = 2e-49, Method: Compositional matrix adjust.
Identities = 92/116 (79%), Positives = 103/116 (88%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ SDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNK P
Sbjct: 345 MLGMHGT VYANYAVDSSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKQP 404

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+C D+KLALQG+N +LE++ +LKLD F WR EL QK K PL+FKTFG+A+P
Sbjct: 405 HVSICADIKLALQGLNSILESKEGKLKLD FSAWRQELTEQKVKHPLNFKTFGDAIP 460

>gb|ACF47582.1| acetolactate synthase 1 [Sonchus asper]
Length = 658

Score = 196 bits (497), Expect = 7e-49, Method: Compositional matrix adjust.
Identities = 91/116 (78%), Positives = 103/116 (88%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ SDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNK P
Sbjct: 336 MLGMHGT VYANYAVDKSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKQP 395

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+CGD+K+AL+G+N++LE R+E LDF WR EL+ QK PLSFKTFG+A+P
Sbjct: 396 HVSICGDIKIALKGLNEILEKRSEMRNLDFSSWRQELDEQKLTHPLSFKTFGDAIP 451

>gb|ACF47583.1| acetolactate synthase 1 [Sonchus asper]
Length = 658

Score = 195 bits (496), Expect = 8e-49, Method: Compositional matrix adjust.
Identities = 91/116 (78%), Positives = 103/116 (88%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ SDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNK P
Sbjct: 336 MLGMHGT VYANYAVDKSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKQP 395

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+CGD+K+AL+G+N++LE R+E LDF WR EL+ QK PLSFKTFG+A+P
Sbjct: 396 HVSICGDIKIALKGLNEILEKRSEMRNLDFSSWRQELDEQKLTHPLSFKTFGDAIP 451

>gb|ACF17639.1| putative acetolactate synthase [Capsicum annuum]
Length = 653

Score = 195 bits (496), Expect = 9e-49, Method: Compositional matrix adjust.
Identities = 92/117 (78%), Positives = 104/117 (88%), Gaps = 1/117 (0%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ SDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNK P
Sbjct: 330 MLGMHGT VYANYAVDSSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKQP 389

Query: 61 HVSVC G DVKLALQGMNKVLENRAEEL-KLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+C D+KLALQG+N +LE + +L KLD F WR ELN QK K+PL++KTFG+A+P
Sbjct: 390 HVSICADIKLALQGLNSILEGKEAKLKLDFSAWRQELNEQVKYPLNYKTFGDAIP 446

>dbj|BAF57909.1| acetolactate synthase [Sagittaria trifolia]
Length = 684

Score = 195 bits (495), Expect = 1e-48, Method: Compositional matrix adjust.
Identities = 91/116 (78%), Positives = 103/116 (88%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 362 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDPAEIGKNKQP 421

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVRN E LNVQKQKFPLSFKTFGEAVP 116
HVS+CGD+KLAL+G+N++LE +LDF WR EL+ QK+KFPLS+K FG+A+P
Sbjct: 422 HVSICGDLKLAL EGINELLEETKIHEQLDFSSWRGELDEQKRKFPLSYKKFGDAIP 477

>gb|AAT07327.1| acetohydroxyacid synthase 2 [Helianthus annuus]
Length = 646

Score = 194 bits (494), Expect = 1e-48, Method: Compositional matrix adjust.
Identities = 94/116 (81%), Positives = 103/116 (88%), Gaps = 3/116 (2%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ SDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 327 MLGMHGT VYANYAVDKSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDPAEIGKNKQP 386

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVRN E LNVQKQKFPLSFKTFGEAVP 116
HVS+CGD+K+ALQG+NK+LE E+ LDF WR EL+ QK KFPLSFKTFGEA+P
Sbjct: 387 HVSICGDIKVALQGLNKILE---EKNSLDFS NWRKELDEQKV KFP L SFKTFGEAIP 439

>gb|AAT07328.1| acetohydroxyacid synthase 2 [Helianthus annuus]
Length = 648

Score = 194 bits (494), Expect = 1e-48, Method: Compositional matrix adjust.
Identities = 94/116 (81%), Positives = 103/116 (88%), Gaps = 3/116 (2%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ SDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 329 MLGMHGT VYANYAVDKSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDPAEIGKNKQP 388

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVRN E LNVQKQKFPLSFKTFGEAVP 116
HVS+CGD+K+ALQG+NK+LE E+ LDF WR EL+ QK KFPLSFKTFGEA+P
Sbjct: 389 HVSICGDIKVALQGLNKILE---EKNSLDFS NWRKELDEQKV KFP L SFKTFGEAIP 441

>gb|ABY57317.1| acetolactate synthase [Medicago truncatula]
Length = 651

Score = 193 bits (490), Expect = 5e-48, Method: Compositional matrix adjust.
Identities = 90/115 (78%), Positives = 104/115 (90%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV++SDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNK P
Sbjct: 329 MLGMHGT VYANYAVDNSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKIP 388

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVRN E LNVQKQKFPLSFKTFGEAV 115
H+S+C D+K+AL+G+N+VLE++ +KLDF WR ELNVQK KFPL FKTF +A+
Sbjct: 389 HLSICADMKVALEGLNRVLESKGIKGLDFEAWRQELNVQK LKFP LGFKTFEDAI 443

>gb|ABY57316.1| acetolactate synthase [Medicago littoralis]
Length = 651

Score = 193 bits (490), Expect = 5e-48, Method: Compositional matrix adjust.
Identities = 90/115 (78%), Positives = 104/115 (90%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV++SDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNK P

Sbjct: 329 MLGMHGTIVYANYAVDNSDLLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKIP 388

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGVRNENLVQKQKFPPLSFKTFGEAV 115
H+S+C D+K+AL+G+N+VLE++ + KLDF WR ELNVQK KFPL FKTF +A+

Sbjct: 389 HLSICADMKVALEGLNRVLESKGIKGLDFEAWRQELNVQKLFPLGFKTFEDAI 443

>gb|ABY57315.1| acetolactate synthase [Medicago littoralis]
Length = 651

Score = 192 bits (489), Expect = 5e-48, Method: Compositional matrix adjust.
Identities = 90/115 (78%), Positives = 104/115 (90%)

Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV++SDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNK P

Sbjct: 329 MLGMHGTIVYANYAVDNSDLLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKIP 388

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGVRNENLVQKQKFPPLSFKTFGEAV 115
H+S+C D+K+AL+G+N+VLE++ + KLDF WR ELNVQK KFPL FKTF +A+

Sbjct: 389 HLSICADMKVALEGLNRVLESKGIKGLDFEAWRQELNVQKLFPLGFKTFEDAI 443

>gb|ABY57318.1| acetolactate synthase [Medicago truncatula]
Length = 651

Score = 192 bits (489), Expect = 6e-48, Method: Compositional matrix adjust.
Identities = 90/115 (78%), Positives = 104/115 (90%)

Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV++SDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNK P

Sbjct: 329 MLGMHGTIVYANYAVDNSDLLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKIP 388

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGVRNENLVQKQKFPPLSFKTFGEAV 115
H+S+C D+K+AL+G+N+VLE++ + KLDF WR ELNVQK KFPL FKTF +A+

Sbjct: 389 HLSICADMKVALEGLNRVLESKGIKGLDFEAWRQELNVQKLFPLGFKTFEDAI 443

>dbj|BAE97677.1| acetolactate synthase [Schoenoplectus juncoideus]
dbj|BAE97678.1| acetolactate synthase [Schoenoplectus juncoideus]
Length = 645

Score = 189 bits (479), Expect = 7e-47, Method: Compositional matrix adjust.
Identities = 87/116 (75%), Positives = 101/116 (87%)

Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV+ +DLLLLAFGVRFDDRTVGKLEAFASR+KIVHIDIDSAEIGKNK P

Sbjct: 320 MLGMHGTIVYANYAVDKADLLLLAFGVRFDDRTVGKLEAFASRSKIVHIDIDSAEIGKNKQP 379

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGVRNENLVQKQKFPPLSFKTFGEAVP 116
H+S+C DVK +LQGMNK+LE+ KLDF WR EL+ QK+ +PL++KTFGE +P

Sbjct: 380 HLSICADVPSLQGMNKILESTGLHRKLDFFSWRAELDEQKAYPLNYKTFGEAIP 435

>gb|AA053549.1| acetohydroxyacid synthase [Triticum aestivum]
Length = 598

Score = 188 bits (478), Expect = 1e-46, Method: Compositional matrix adjust.
Identities = 87/116 (75%), Positives = 99/116 (85%)

Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV+ +DLLLLAFGVRFDDRTVGK+EAFASR+KIVHIDID AEIGKNK P

Sbjct: 276 MLGMHGTIVYANYAVDKADLLLLAFGVRFDDRTVGKIEAFASRSKIVHIDIDPAEIGKNKQP 335

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGVRNENLVQKQKFPPLSFKTFGEAVP 116
HVS+C DVKLALQG+N +L + LDFG W EL+ QK++FPL FKTFGEA+P

Sbjct: 336 HVSICADVKLALQGLNALLNGSKAQQLDFGPWHKELDQKREFPLGFKTFGEAIP 391

>gb|AA053551.1| acetohydroxyacid synthase [Triticum aestivum]
Length = 598

Score = 188 bits (478), Expect = 1e-46, Method: Compositional matrix adjust.
Identities = 87/116 (75%), Positives = 99/116 (85%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLAFGVRFDDRVTK+EAFASR+KIVHIDID AEIGKNK P
Sbjct: 276 MLGMHGT VYANYAVDKADLLAFGVRFDDRVTKIEAFASRSKIVHIDIDPAEIGKNKQP 335

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLDGFWVRNENLVQKQKFPLSFKTFGEAVP 116
HVS+C DVKLALQG+N +L + LDFG W EL+ QK++FPL FKTFGEA+P
Sbjct: 336 HVSICADVKLALQGLNALLNGSKAQQLDFGPWHKELDQKREFPLGFKTFGEAIP 391

>emb|CAE18088.1| acetolactate synthase [Papaver rhoeas]
Length = 662

Score = 188 bits (478), Expect = 1e-46, Method: Compositional matrix adjust.
Identities = 86/116 (74%), Positives = 102/116 (87%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLAFGVRFDDRVTK+EAFASR+KIVHIDID AEIGKNK P
Sbjct: 340 MLGMHGT VYANYAVDKADLLAFGVRFDDRVTKIEAFASRSKIVHIDIDPAEIGKNKQP 399

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLDGFWVRNENLVQKQKFPLSFKTFGEAVP 116
HVS+C D+ +AL+G+NK+LE + +LKLD F WR EL QK+ +PLS+KTFGEA+P
Sbjct: 400 HVSICADMGVALKGLNKLLEEKMSKLDLDFVSWREELMEQKKNYPLSYKTFGEAIP 455

>dbj|BAE97675.1| acetolactate synthase [Schoenoplectus juncoideus]
dbj|BAE97676.1| acetolactate synthase [Schoenoplectus juncoideus]
Length = 645

Score = 188 bits (477), Expect = 1e-46, Method: Compositional matrix adjust.
Identities = 86/116 (74%), Positives = 101/116 (87%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLAFGVRFDDRVTKLEAFASR+KIVHIDIDSAEIGKNK P
Sbjct: 323 MLGMHGT VYANYAVDKADLLAFGVRFDDRVTKLEAFASRSKIVHIDIDSAEIGKNKQP 382

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLDGFWVRNENLVQKQKFPLSFKTFGEAVP 116
H+S+C DVK +LQGMN++LE+ KLDF WR EL+ Q++ +PLS+KTFGE +P
Sbjct: 383 HLSICADVKPSLQGMNEILESTGLHRKLD FSSWRAELDEQRKAYPLSYKTFGEEIP 438

>gb|AA053548.1| acetohydroxyacid synthase [Triticum aestivum]
Length = 598

Score = 187 bits (476), Expect = 2e-46, Method: Compositional matrix adjust.
Identities = 87/116 (75%), Positives = 99/116 (85%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLAFGVRFDDRVTK+EAFASR+KIVHIDID AEIGKNK P
Sbjct: 276 MLGMHGT VYANYAVDKADLLAFGVRFDDRVTKIEAFASRSKIVHIDIDPAEIGKNKQP 335

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLDGFWVRNENLVQKQKFPLSFKTFGEAVP 116
HVS+C DVKLALQG+N +L + LDFG W EL+ QK++FPL FKTFGEA+P
Sbjct: 336 HVSICADVKLALQGLNDLLNGSKAQQLDFGPWHKELDQKREFPLGFKTFGEAIP 391

>gb|AA053550.1| acetohydroxyacid synthase [Triticum aestivum]
Length = 598

Score = 187 bits (476), Expect = 2e-46, Method: Compositional matrix adjust.
Identities = 87/116 (75%), Positives = 99/116 (85%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLAFGVRFDDRVTK+EAFASR+KIVHIDID AEIGKNK P
Sbjct: 276 MLGMHGT VYANYAVDKADLLAFGVRFDDRVTKIEAFASRSKIVHIDIDPAEIGKNKQP 335

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLDGFWVRNENLVQKQKFPLSFKTFGEAVP 116

HVS+C DVKLALQG+N +L + LDFG W EL+ QK++FPL FKTFGEA+P
Sbjct: 336 HVSICADVKLALQGLNDLLNGSKAQQLDFGFWHKELDQQKREFPLGFKTFGEAIP 391

>gb|ABM92357.2| acetolactate synthase [Cyperus difformis]
Length = 569

Score = 187 bits (474), Expect = 3e-46, Method: Compositional matrix adjust.
Identities = 87/116 (75%), Positives = 100/116 (86%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
+LGMHGT VYANYAV+ +DLLAFGVRFDDRTVGKLEAFASR+KIVHIDID AEIGKNK P
Sbjct: 248 LLGMHGT VYANYAVDKADLLAFGVRFDDRTVGKLEAFASRSKIVHIDIDPAEIGKNKQP 307

Query: 61 HVSVC GDVKLALQGMNVLENRAEELKLDGFWVRNENLVQKQKFPLSFKTFGEAVP 116
HVS+C DVK ALQGMN++LE+ KLDF WR EL+ QK+ +PLS+KTFGE +P
Sbjct: 308 HVSICADV KPALQGMNQILESSGVHKKLDFSSWRAELDEQKKTYPLSYKTFGEEIP 363

>gb|AAC14572.1| acetohydroxyacid synthase [Hordeum vulgare]
Length = 541

Score = 187 bits (474), Expect = 3e-46, Method: Compositional matrix adjust.
Identities = 86/116 (74%), Positives = 99/116 (85%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLAFGVRFDDRTVGK+EAFASR+KIVHIDID AEIGKNK P
Sbjct: 219 MLGMHGT VYANYAVDKADLLAFGVRFDDRTVGKIEAFASRSKIVHIDIDPAEIGKNKQP 278

Query: 61 HVSVC GDVKLALQGMNVLENRAEELKLDGFWVRNENLVQKQKFPLSFKTFGEAVP 116
HVS+C DVKLALQG+N +L + LDFG W EL+ QK++FPL +KTFGEA+P
Sbjct: 279 HVSICADV KLALQGLNGLLSGSKAQQLDFGFWHKELDQQKREFPLGYKTFGEAIP 334

>dbj|BAE53610.1| acetolactate synthase [Monochoria vaginalis]
Length = 614

Score = 187 bits (474), Expect = 3e-46, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 99/116 (85%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYA++ +DLLAFGVRFDDRTVGKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 297 MLGMHGT VYANYAIDKADLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDPAEIGKNKQP 356

Query: 61 HVSVC GDVKLALQGMNVLENRAEELKLDGFWVRNENLVQKQKFPLSFKTFGEAVP 116
HVS+CGD+KLALQ MN+++E KLDF WR EL+ QK+ +PL +KTFG+ +P
Sbjct: 357 HVSICGDIKLALQEMNEMIEENGIHSLDFSAWREELDQQKKNYPLKYKTFGDLIP 412

>dbj|BAE53611.1| acetolactate synthase [Monochoria vaginalis]
Length = 614

Score = 187 bits (474), Expect = 3e-46, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 99/116 (85%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYA++ +DLLAFGVRFDDRTVGKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 297 MLGMHGT VYANYAIDKADLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDPAEIGKNKQP 356

Query: 61 HVSVC GDVKLALQGMNVLENRAEELKLDGFWVRNENLVQKQKFPLSFKTFGEAVP 116
HVS+CGD+KLALQ MN+++E KLDF WR EL+ QK+ +PL +KTFG+ +P
Sbjct: 357 HVSICGDIKLALQEMNEMIEENGIHSLDFSAWREELDQQKKNYPLKYKTFGDLIP 412

>dbj|BAE53612.1| acetolactate synthase [Monochoria vaginalis]
Length = 614

Score = 187 bits (474), Expect = 3e-46, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 99/116 (85%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYA++ +DLLLLAFGVRFDDRTVGKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 297 MLGMHGT VYANYAIDKADLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDPAEIGKNKQP 356

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+CGD+KLALQ MN+++E KLDF WR EL+ QK+ +PL +KTFG+ +P
Sbjct: 357 HVSICGDIKLALQEMNEMIEENGIH SKLDFS AWREELDQQKKNYPLKYKTFGDLIP 412

>dbj|BAE53605.1| acetolactate synthase [Monochoria vaginalis]
dbj|BAE53606.1| acetolactate synthase [Monochoria vaginalis]
dbj|BAE53607.1| acetolactate synthase [Monochoria vaginalis]
dbj|BAE53609.1| acetolactate synthase [Monochoria vaginalis]
Length = 614

Score = 187 bits (474), Expect = 3e-46, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 99/116 (85%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYA++ +DLLLLAFGVRFDDRTVGKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 297 MLGMHGT VYANYAIDKADLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDPAEIGKNKQP 356

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+CGD+KLALQ MN+++E KLDF WR EL+ QK+ +PL +KTFG+ +P
Sbjct: 357 HVSICGDIKLALQEMNEMIEENGIH SKLDFS AWREELDQQKKNYPLKYKTFGDLIP 412

>dbj|BAF37288.1| acetolactate synthase [Monochoria vaginalis]
Length = 614

Score = 187 bits (474), Expect = 3e-46, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 99/116 (85%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYA++ +DLLLLAFGVRFDDRTVGKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 297 MLGMHGT VYANYAIDKADLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDPAEIGKNKQP 356

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+CGD+KLALQ MN+++E KLDF WR EL+ QK+ +PL +KTFG+ +P
Sbjct: 357 HVSICGDIKLALQEMNEMIEENGIH SKLDFS AWREELDQQKKNYPLKYKTFGDLIP 412

>dbj|BAE53608.1| acetolactate synthase [Monochoria vaginalis]
Length = 614

Score = 187 bits (474), Expect = 3e-46, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 99/116 (85%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYA++ +DLLLLAFGVRFDDRTVGKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 297 MLGMHGT VYANYAIDKADLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDPAEIGKNKQP 356

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+CGD+KLALQ MN+++E KLDF WR EL+ QK+ +PL +KTFG+ +P
Sbjct: 357 HVSICGDIKLALQEMNEMIEENGIH SKLDFS AWREELDQQKKNYPLKYKTFGDLIP 412

>gb|AAM03119.1|AF488771_1 acetolactate synthase [Bromus tectorum]
Length = 583

Score = 186 bits (473), Expect = 4e-46, Method: Compositional matrix adjust.
Identities = 86/116 (74%), Positives = 98/116 (84%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLLLAFGVRFDDRTVGKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 261 MLGMHGT VYANYAVDKADLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDPAEIGKNKQP 320

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+C DVKLALQG+N +L + LDFG W+ EL QK+ FPL +KTFGEA+P
Sbjct: 321 HVSICADV KLALQGLNDLLNGSKAQKSLDFGPWQEELDQQKRTFPLGYKTFGEAIP 376

>gb|AAL93207.1|AF487459_1 acetolactate synthase [Bromus tectorum]
Length = 583

Score = 186 bits (473), Expect = 4e-46, Method: Compositional matrix adjust.
Identities = 86/116 (74%), Positives = 98/116 (84%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLLLAFGVRFD DRVTGK+EAFASR+KIVHIDID AEIGKNK P
Sbjct: 261 MLGMHGT VYANYAVDKADLLLA FGVRFD DRVTGKIEAFASRSKIVHIDIDPAEIGKNKQP 320

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+C DVKLALQG+N +L + LDFG W+ EL QK+ FPL +KTFGEA+P
Sbjct: 321 HVSICADVKLALQGLNDLLNGSKAQKSLDFGPWQEELEQQKRTFPLGYKTFGEAIP 376

>dbj|BAE53591.1| acetolactate synthase [Monochoria vaginalis]
dbj|BAE53592.1| acetolactate synthase [Monochoria vaginalis]
Length = 642

Score = 186 bits (473), Expect = 4e-46, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 99/116 (85%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYA++ +DLLLLAFGVRFD DRVTGKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 320 MLGMHGT VYANYAIDKADLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDPAEIGKNKQP 379

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+CGD+KLALQ MN+++E KLDF WR EL+ QK+ +PL +KTFG+ +P
Sbjct: 380 HVSICGDIKLALQEMNEMIEESGIHNKLD FSAWREELDQQKKNYPLEYKTFGDLIP 435

>dbj|BAE53593.1| acetolactate synthase [Monochoria vaginalis]
Length = 642

Score = 186 bits (473), Expect = 4e-46, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 99/116 (85%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYA++ +DLLLLAFGVRFD DRVTGKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 320 MLGMHGT VYANYAIDKADLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDPAEIGKNKQP 379

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+CGD+KLALQ MN+++E KLDF WR EL+ QK+ +PL +KTFG+ +P
Sbjct: 380 HVSICGDIKLALQEMNEMIEESGIHNKLD FSAWREELDQQKKNYPLEYKTFGDLIP 435

>dbj|BAE53587.1| acetolactate synthase [Monochoria vaginalis]
dbj|BAE53589.1| acetolactate synthase [Monochoria vaginalis]
dbj|BAE53590.1| acetolactate synthase [Monochoria vaginalis]
dbj|BAE53594.1| acetolactate synthase [Monochoria vaginalis]
Length = 642

Score = 186 bits (473), Expect = 4e-46, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 99/116 (85%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYA++ +DLLLLAFGVRFD DRVTGKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 320 MLGMHGT VYANYAIDKADLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDPAEIGKNKQP 379

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+CGD+KLALQ MN+++E KLDF WR EL+ QK+ +PL +KTFG+ +P
Sbjct: 380 HVSICGDIKLALQEMNEMIEESGIHNKLD FSAWREELDQQKKNYPLEYKTFGDLIP 435

>dbj|BAE53588.1| acetolactate synthase [Monochoria vaginalis]
Length = 642

Score = 186 bits (472), Expect = 5e-46, Method: Compositional matrix adjust.

Identities = 85/116 (73%), Positives = 99/116 (85%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYA++ +DLLAFGVRFDDRVTKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 320 MLGMHGT VYANYAIDKADLLAFGVRFDDRVTKLEAFASRAKIVHIDIDPAEIGKNKQP 379

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLDG VWRNELNVQKQKFP LSFKTFGEAVP 116
HVS+CGD+KLALQ MN+++E KLDF WR EL+ QK+ +PL +KTFG+ +P
Sbjct: 380 HVSICGDIKLALQEMNEMIEESGIHNKLD FSAWREELDQ QKKNYPLEYKTFGD LIP 435

>gb|EAZ23185.1| hypothetical protein OsJ_006668 [Oryza sativa (japonica cultivar-group)]
Length = 628

Score = 186 bits (471), Expect = 6e-46, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 98/116 (84%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLAFGVRFDDRVTK+EAFASRAKIVHIDID AEIGKNK P
Sbjct: 306 MLGMHGT VYANYAVDKADLLAFGVRFDDRVTKIEAFASRAKIVHIDIDPAEIGKNKQP 365

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLDG VWRNELNVQKQKFP LSFKTFGEAVP 116
HVS+C DVKLALQG+N +L+ + DF W NEL+ QK++FPL +KTFGE +P
Sbjct: 366 HVSICADVKLALQGLNALDQSTTKTSSDFS AWHNELDQ QKREFPLGYKTFGE EIP 421

>gb|EAY86003.1| hypothetical protein OsI_007236 [Oryza sativa (indica cultivar-group)]
Length = 644

Score = 186 bits (471), Expect = 7e-46, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 98/116 (84%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLAFGVRFDDRVTK+EAFASRAKIVHIDID AEIGKNK P
Sbjct: 322 MLGMHGT VYANYAVDKADLLAFGVRFDDRVTKIEAFASRAKIVHIDIDPAEIGKNKQP 381

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLDG VWRNELNVQKQKFP LSFKTFGEAVP 116
HVS+C DVKLALQG+N +L+ + DF W NEL+ QK++FPL +KTFGE +P
Sbjct: 382 HVSICADVKLALQGLNALDQSTTKTSSDFS AWHNELDQ QKREFPLGYKTFGE EIP 437

>dbj|BAB20813.1| acetolactate synthase [Oryza sativa (japonica cultivar-group)]
Length = 644

Score = 186 bits (471), Expect = 7e-46, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 98/116 (84%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLAFGVRFDDRVTK+EAFASRAKIVHIDID AEIGKNK P
Sbjct: 322 MLGMHGT VYANYAVDKADLLAFGVRFDDRVTKIEAFASRAKIVHIDIDPAEIGKNKQP 381

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLDG VWRNELNVQKQKFP LSFKTFGEAVP 116
HVS+C DVKLALQG+N +L+ + DF W NEL+ QK++FPL +KTFGE +P
Sbjct: 382 HVSICADVKLALQGLNALDQSTTKTSSDFS AWHNELDQ QKREFPLGYKTFGE EIP 437

>gb|AAx14282.1| acetolactate synthase [Oryza sativa (japonica cultivar-group)]
Length = 644

Score = 186 bits (471), Expect = 7e-46, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 98/116 (84%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLAFGVRFDDRVTK+EAFASRAKIVHIDID AEIGKNK P
Sbjct: 322 MLGMHGT VYANYAVDKADLLAFGVRFDDRVTKIEAFASRAKIVHIDIDPAEIGKNKQP 381

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLDG VWRNELNVQKQKFP LSFKTFGEAVP 116
HVS+C DVKLALQG+N +L+ + DF W NEL+ QK++FPL +KTFGE +P

Sbjct: 382 HVSICADVKKLALQGLNALLDQSTTKTSSDFSAAWHNELDQQKREFPLGYKTFGEEIP 437

```
>ref|NP_001046931.1| Os02g0510200 [Oryza sativa (japonica cultivar-group)]
sp|Q6K2E8|ILV1_ORYSJ Acetolactate synthase 1, chloroplast precursor (Acetohydroxy-acid
synthase 1)
dbj|BAB20812.1| acetolactate synthase [Oryza sativa (japonica cultivar-group)]
dbj|BAD23668.1| acetolactate synthase [Oryza sativa Japonica Group]
dbj|BAF08845.1| Os02g0510200 [Oryza sativa (japonica cultivar-group)]
Length = 644
```

Score = 185 bits (470), Expect = 8e-46, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 98/116 (84%)

```
Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV+ +DLLAFGVRFDDRTGK+EAFASRAKIVHIDID AEIGKNK P
Sbjct: 322 MLGMHGTIVYANYAVDKADLLAFGVRFDDRTGKIEAFASRAKIVHIDIDPAEIGKNKQP 381
```

```
Query: 61 HVSVCADVKKLALQGMNKVLENRAEELKLDGFWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+C DVKKLALQG+N +L+ + DF W NEL+ QK++FPL +KTFGE +P
Sbjct: 382 HVSICADVKKLALQGLNALLDQSTTKTSSDFSAAWHNELDQQKREFPLGYKTFGEEIP 437
```

```
>gb|ABF66052.1| acetolactate synthase [Oryza sativa (indica cultivar-group)]
Length = 644
```

Score = 185 bits (470), Expect = 8e-46, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 98/116 (84%)

```
Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV+ +DLLAFGVRFDDRTGK+EAFASRAKIVHIDID AEIGKNK P
Sbjct: 322 MLGMHGTIVYANYAVDKADLLAFGVRFDDRTGKIEAFASRAKIVHIDIDPAEIGKNKQP 381
```

```
Query: 61 HVSVCADVKKLALQGMNKVLENRAEELKLDGFWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+C DVKKLALQG+N +L+ + DF W NEL+ QK++FPL +KTFGE +P
Sbjct: 382 HVSICADVKKLALQGLNALLDQSTTKTSSDFSAAWHNELDQQKREFPLGYKTFGEEIP 437
```

```
>gb|AAX14281.1| acetolactate synthase [Oryza sativa]
gb|ABF66050.1| acetolactate synthase [Oryza sativa (indica cultivar-group)]
gb|ACD74787.1| acetolactate synthase [Oryza sativa Indica Group]
gb|ACD74788.1| acetolactate synthase [Oryza sativa Indica Group]
Length = 644
```

Score = 185 bits (470), Expect = 8e-46, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 98/116 (84%)

```
Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV+ +DLLAFGVRFDDRTGK+EAFASRAKIVHIDID AEIGKNK P
Sbjct: 322 MLGMHGTIVYANYAVDKADLLAFGVRFDDRTGKIEAFASRAKIVHIDIDPAEIGKNKQP 381
```

```
Query: 61 HVSVCADVKKLALQGMNKVLENRAEELKLDGFWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+C DVKKLALQG+N +L+ + DF W NEL+ QK++FPL +KTFGE +P
Sbjct: 382 HVSICADVKKLALQGLNALLDQSTTKTSSDFSAAWHNELDQQKREFPLGYKTFGEEIP 437
```

```
>gb|ACD74789.1| acetolactate synthase [Oryza sativa Indica Group]
gb|ACD74790.1| acetolactate synthase [Oryza sativa Indica Group]
Length = 644
```

Score = 185 bits (470), Expect = 8e-46, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 98/116 (84%)

```
Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV+ +DLLAFGVRFDDRTGK+EAFASRAKIVHIDID AEIGKNK P
Sbjct: 322 MLGMHGTIVYANYAVDKADLLAFGVRFDDRTGKIEAFASRAKIVHIDIDPAEIGKNKQP 381
```

```
Query: 61 HVSVCADVKKLALQGMNKVLENRAEELKLDGFWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+C DVKKLALQG+N +L+ + DF W NEL+ QK++FPL +KTFGE +P
Sbjct: 382 HVSICADVKKLALQGLNALLDQSTTKTSSDFSAAWHNELDQQKREFPLGYKTFGEEIP 437
```

>gb|ABF66048.1| acetolactate synthase [Oryza sativa (indica cultivar-group)]
Length = 644

Score = 185 bits (470), Expect = 8e-46, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 98/116 (84%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLLAFGVRFD DRVTGK+EAFASRAKIVHIDID AEIGKNK P
Sbjct: 322 MLGMHGT VYANYAVDKADLLLA FGVRFD DRVTGKIEAFASRAKIVHIDIDPAEIGKNKQP 381

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FG VWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+C DVKLALQG+N +L+ + DF W NEL+ QK++FPL +KTFGE +P
Sbjct: 382 HVSICADVKLALQGLNALLDQSTTKTSSDFS AWHNELDQKREFPLGYKTFGE EIP 437

>gb|ABF66051.1| acetolactate synthase [Oryza sativa (indica cultivar-group)]
Length = 644

Score = 185 bits (470), Expect = 8e-46, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 98/116 (84%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLLAFGVRFD DRVTGK+EAFASRAKIVHIDID AEIGKNK P
Sbjct: 322 MLGMHGT VYANYAVDKADLLLA FGVRFD DRVTGKIEAFASRAKIVHIDIDPAEIGKNKQP 381

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FG VWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+C DVKLALQG+N +L+ + DF W NEL+ QK++FPL +KTFGE +P
Sbjct: 382 HVSICADVKLALQGLNALLDQSTTKTSSDFS AWHNELDQKREFPLGYKTFGE EIP 437

>dbj|BAF37287.1| acetolactate synthase [Monochoria vaginalis]
Length = 642

Score = 184 bits (467), Expect = 2e-45, Method: Compositional matrix adjust.
Identities = 84/116 (72%), Positives = 99/116 (85%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYA++ +DLLLAFGVRFD +RVTGKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 320 MLGMHGT VYANYAIDKADLLLA FGVRFD ERVTGKLEAFASRAKIVHIDIDPAEIGKNKQP 379

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FG VWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+CGD+KLALQ MN+++E KLDF WR EL+ QK+ +PL +KTFG+ +P
Sbjct: 380 HVSICGDIKLALQEMNEMIEESGIHNKLDFAWREELDQKKNYPLEYKTFGDLIP 435

>gb|AAx14283.1| acetolactate synthase [Oryza sativa]
Length = 644

Score = 184 bits (466), Expect = 2e-45, Method: Compositional matrix adjust.
Identities = 84/116 (72%), Positives = 98/116 (84%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLLAFGVRFD DRVTGK+EAFASRAKIVHIDID AEIGKNK P
Sbjct: 322 MLGMHGT VYANYAVDKADLLLA FGVRFD DRVTGKIEAFASRAKIVHIDIDPAEIGKNKQP 381

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FG VWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+C DV+LALQG+N +L+ + DF W NEL+ QK++FPL +KTFGE +P
Sbjct: 382 HVSICADVELALQGLNALLDQSTTKTSSDFS AWHNELDQKREFPLGYKTFGE EIP 437

>gb|AAT07329.1| acetohydroxyacid synthase 3 [Helianthus annuus]
Length = 646

Score = 184 bits (466), Expect = 2e-45, Method: Compositional matrix adjust.
Identities = 86/116 (74%), Positives = 99/116 (85%), Gaps = 1/116 (0%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYA++ SDLLLA FGVRFD DRVTGK+EAFASRAKIVHIDID AEIGKNK P

Sbjct: 326 MLGMHGT VYANYAIDKSDLLLA FGVRFD DRVTGKIEAFASRAKIVHIDIDPAEIGKNKQP 385

Query: 61 HVS VCGDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF KTFGEAVP 116
H S+CGD+K ALQG+NK+LE R E+L DF W+ E+ QK PLS+KT FG+A+P

Sbjct: 386 HFSICGDIKAALQGLNKILE-RGEDLDFDFSPWKEEVMNQKASNPLSYKTFGDAIP 440

>gb|ABQ85869.1| acetolactate synthase [*Lolium rigidum*]
Length = 516

Score = 184 bits (466), Expect = 2e-45, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 96/116 (82%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLLLA FGVRFD DRVTGK+EAFASR+KIVHIDID AEIGKNK P

Sbjct: 213 MLGMHGT VYANYAVDKADLLLA FGVRFD DRVTGKIEAFASRSKIVHIDIDPAEIGKNKQP 272

Query: 61 HVS VCGDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF KTFGEAVP 116
HVS+C DVKLALQG+N +L DFG W EL QK++FPL +KTFGEA+P

Sbjct: 273 HVSICADV KLALQGLNALLTGTKAHKSFD FGSWHEELEQQKREFPLGYKTFGEAIP 328

>gb|AAG30931.1|AF310684_1 acetolactate synthase precursor [*Lolium multiflorum*]
Length = 640

Score = 183 bits (465), Expect = 3e-45, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 97/116 (83%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLLLA FGVRFD DRVTGK+EAFASR+KIVHIDID AEIGKNK P

Sbjct: 318 MLGMHGT VYANYAVDKADLLLA FGVRFD DRVTGKIEAFASRSKIVHIDIDPAEIGKNKQP 377

Query: 61 HVS VCGDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF KTFGEAVP 116
HVS+C DVKLALQG+N VL + DF W +EL QK++FPL +KTFGEA+P

Sbjct: 378 HVSICADV KLALQGLNAVLTGSKCDKSFD FASWHDELEQQKREFPLGYKTFGEAIP 433

>gb|ABQ85870.1| acetolactate synthase [*Lolium rigidum*]
Length = 516

Score = 183 bits (465), Expect = 3e-45, Method: Compositional matrix adjust.
Identities = 85/116 (73%), Positives = 96/116 (82%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLLLA FGVRFD DRVTGK+EAFASR+KIVHIDID AEIGKNK P

Sbjct: 213 MLGMHGT VYANYAVDKADLLLA FGVRFD DRVTGKIEAFASRSKIVHIDIDPAEIGKNKQP 272

Query: 61 HVS VCGDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF KTFGEAVP 116
HVS+C DVKLALQG+N +L DFG W EL QK++FPL +KTFGEA+P

Sbjct: 273 HVSICADV KLALQGLNALLTGTKAHKSFD FGSWHEELEQQKREFPLGYKTFGEAIP 328

>dbj|BAE53604.1| acetolactate synthase [*Monochoria vaginalis*]
Length = 614

Score = 182 bits (461), Expect = 1e-44, Method: Compositional matrix adjust.
Identities = 84/116 (72%), Positives = 97/116 (83%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
M GMHGT VYANYA++ +DLLLLA FGVRFD DRVTGKLEAFASRAKIVHIDID AEIGKNK P

Sbjct: 297 MWGMHGT VYANYAIDKADLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDPAEIGKNKQP 356

Query: 61 HVS VCGDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF KTFGEAVP 116
HVS+CGD+KLALQ MN+++E KLD F WR EL+ QK+ PL +KTFG+ +P

Sbjct: 357 HVSICGDIKLALQEMNEMIEENGIH SKLD FSAWREELDQQKKNPLKYKTFGD LIP 412

>emb|CAD24801.2| acetolactate synthase [*Alopecurus myosuroides*]
Length = 640

Score = 181 bits (460), Expect = 1e-44, Method: Compositional matrix adjust.
Identities = 85/117 (72%), Positives = 100/117 (85%), Gaps = 1/117 (0%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLLAFGVRFDDRVTKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 317 MLGMHGT VYANYAVDRADLLAFGVRFDDRVTKLEAFASRAKIVHIDIDPAEIGKNKQP 376

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLDFGVWRNELNVQKQKFP LSFKTFGEAVP 116
HVS+C DVK+ALQG+N +L N ++ K DF W EL+ QK++FPL +KTFGEA+P
Sbjct: 377 HVSICADVKIALQGLNTLLNLSKTRKCSDFSSWHEELDQKREFPLGYKTFGEAIP 433

>dbj|BAE53597.1| acetolactate synthase [Monochoria vaginalis]
dbj|BAE53599.1| acetolactate synthase [Monochoria vaginalis]
dbj|BAE53600.1| acetolactate synthase [Monochoria vaginalis]
dbj|BAE53603.1| acetolactate synthase [Monochoria vaginalis]
Length = 604

Score = 181 bits (458), Expect = 2e-44, Method: Compositional matrix adjust.
Identities = 82/116 (70%), Positives = 98/116 (84%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANY+++ +DLLLAFGVRFDDRVTKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 282 MLGMHGT VYANYSIDKADLLAFGVRFDDRVTKLEAFASRAKIVHIDIDPAEIGKNKQP 341

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLDFGVWRNELNVQKQKFP LSFKTFGEAVP 116
HVS+C D+KLALQ MNK++E KL+F WR EL+ K+ +PL++KTFG+ +P
Sbjct: 342 HVSICADIKLALQEMNKIIIESGIYNKLNFSAWREELDQHKKNYPLNYKTFGDLIP 397

>dbj|BAE53596.1| acetolactate synthase [Monochoria vaginalis]
Length = 604

Score = 181 bits (458), Expect = 2e-44, Method: Compositional matrix adjust.
Identities = 82/116 (70%), Positives = 98/116 (84%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANY+++ +DLLLAFGVRFDDRVTKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 282 MLGMHGT VYANYSIDKADLLAFGVRFDDRVTKLEAFASRAKIVHIDIDPAEIGKNKQP 341

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLDFGVWRNELNVQKQKFP LSFKTFGEAVP 116
HVS+C D+KLALQ MNK++E KL+F WR EL+ K+ +PL++KTFG+ +P
Sbjct: 342 HVSICADIKLALQEMNKIIIESGIYNKLNFSAWREELDQHKKNYPLNYKTFGDLIP 397

>dbj|BAE53598.1| acetolactate synthase [Monochoria vaginalis]
dbj|BAE53601.1| acetolactate synthase [Monochoria vaginalis]
Length = 604

Score = 181 bits (458), Expect = 2e-44, Method: Compositional matrix adjust.
Identities = 82/116 (70%), Positives = 98/116 (84%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANY+++ +DLLLAFGVRFDDRVTKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 282 MLGMHGT VYANYSIDKADLLAFGVRFDDRVTKLEAFASRAKIVHIDIDPAEIGKNKQP 341

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLDFGVWRNELNVQKQKFP LSFKTFGEAVP 116
HVS+C D+KLALQ MNK++E KL+F WR EL+ K+ +PL++KTFG+ +P
Sbjct: 342 HVSICADIKLALQEMNKIIIESGIYNKLNFSAWREELDQHKKNYPLNYKTFGDLIP 397

>dbj|BAE53602.1| acetolactate synthase [Monochoria vaginalis]
Length = 604

Score = 181 bits (458), Expect = 2e-44, Method: Compositional matrix adjust.
Identities = 82/116 (70%), Positives = 98/116 (84%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANY+++ +DLLLAFGVRFDDRVTKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 282 MLGMHGT VYANYSIDKADLLAFGVRFDDRVTKLEAFASRAKIVHIDIDPAEIGKNKQP 341

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+C D+KLALQ MNK++E KL+F WR EL+ K+ +PL++KTFG+ +P
Sbjct: 342 HVSICADIKLALQEMNKIIEESGIYNKLNFSAWREELDQHKKNYPLNYKTFGDLIP 397

>sp|Q41769|ILV2_MAIZE Acetolactate synthase 2, chloroplast precursor (Acetohydroxy-acid
synthase 2)
emb|CAA45117.1| acetohydroxyacid synthase [Zea mays]
Length = 638

Score = 179 bits (455), Expect = 5e-44, Method: Compositional matrix adjust.
Identities = 83/115 (72%), Positives = 95/115 (82%)

Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV+ +DLLLLAFGVRFDRTGK+EAF RAKIVHIDID AEIGKNK P
Sbjct: 316 MLGMHGTIVYANYAVDKADLLAFGVRFDRTGKIEAFAGRAKIVHIDIDPAEIGKNKQP 375

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAV 115
HVS+C DVKLALQGMN +LE + DFG W +EL+ QK++FPL +K F E +
Sbjct: 376 HVSICADVKLALQGMNTLLEGSTSKKSFDGFWHDELDQKREFPLGYKIFNEEI 430

>gb|ACF87819.1| unknown [Zea mays]
Length = 638

Score = 179 bits (455), Expect = 5e-44, Method: Compositional matrix adjust.
Identities = 83/115 (72%), Positives = 95/115 (82%)

Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV+ +DLLLLAFGVRFDRTGK+EAF RAKIVHIDID AEIGKNK P
Sbjct: 316 MLGMHGTIVYANYAVDKADLLAFGVRFDRTGKIEAFAGRAKIVHIDIDPAEIGKNKQP 375

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAV 115
HVS+C DVKLALQGMN +LE + DFG W +EL+ QK++FPL +K F E +
Sbjct: 376 HVSICADVKLALQGMNTLLEGSTSKKSFDGFWHDELDQKREFPLGYKIFNEEI 430

>gb|EAY93888.1| hypothetical protein OsI_015121 [Oryza sativa (indica
cultivar-group)]
Length = 662

Score = 178 bits (452), Expect = 1e-43, Method: Compositional matrix adjust.
Identities = 82/116 (70%), Positives = 98/116 (84%), Gaps = 1/116 (0%)

Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV+++DLLLL GVRFDRTGK+EAFASRAKIVH+DID +E+GKNK P
Sbjct: 340 MLGMHGTIVYANYAVDNADLLALGVRFDRTGKVEAFASRAKIVHVDIDPSELGKNKQP 399

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+C DVKLALQGMN LE + + LDF WR+EL +K +FPL ++TFGE +P
Sbjct: 400 HVSICADVKLALQGMNATLEQQQRK-NLDFSARWSELEKKKAEFPLGYRTFGEEIP 454

>sp|Q41768|ILV1_MAIZE Acetolactate synthase 1, chloroplast precursor (Acetohydroxy-acid
synthase 1)
emb|CAA45116.1| acetohydroxyacid synthase [Zea mays]
Length = 638

Score = 178 bits (451), Expect = 1e-43, Method: Compositional matrix adjust.
Identities = 82/115 (71%), Positives = 95/115 (82%)

Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV+ +DLLLL GVRFDRTGK+EAFASRAKIVH+DID AEIGKNK P
Sbjct: 316 MLGMHGTIVYANYAVDKADLLALGVRFDRTGKIEAFASRAKIVHVDIDPAEIGKNKQP 375

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFKTFGEAV 115
HVS+C DVKLALQGMN +LE + DFG W +EL+ QK++FPL +KT E +
Sbjct: 376 HVSICADVKLALQGMNALLEGSTSKKSFDGFWNDELDQKREFPLGYKTSNEEI 430

```
>gb|EAZ30535.1| hypothetical protein OsJ_014018 [Oryza sativa (japonica
cultivar-group)]
Length = 610

Score = 178 bits (451), Expect = 1e-43, Method: Compositional matrix adjust.
Identities = 83/119 (69%), Positives = 99/119 (83%), Gaps = 3/119 (2%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
          MLGMHGT VYANYAV++D LLLA GVRFD DRVTGK+EAFASRAKIVH+DID +E+GKNK P
Sbjct: 305 MLGMHGT VYANYAVDNADLLLA LGVRFD DRVTGKVEAFASRAKIVHVDIDPSELGKNKQP 364

Query: 61 HVSVC G DVKLALQGMNKVLENR---AEELK LDFGVWRNELNVQKQKFPLSFKTFGEAVP 116
          HVS+C DVKLALQGMN +LE + A LDF WR+EL +K +FPL ++TFGE +P
Sbjct: 365 HVSICADVKLALQGMNAMLEEQSAAAARKNLDFS AWRSELEKKKVEFPLGYRTFGEEIP 423

>gb|EAZ30530.1| hypothetical protein OsJ_014013 [Oryza sativa (japonica
cultivar-group)]
Length = 620

Score = 178 bits (451), Expect = 1e-43, Method: Compositional matrix adjust.
Identities = 82/116 (70%), Positives = 98/116 (84%), Gaps = 1/116 (0%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
          MLGMHGT VYANYAV++D LLLA GVRFD DRVTGK+EAFASRAKIVH+DID +E+GKNK P
Sbjct: 298 MLGMHGT VYANYAVDNADLLLA LGVRFD DRVTGKVEAFASRAKIVHVDIDPSELGKNKQP 357

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELK LDFGVWRNELNVQKQKFPLSFKTFGEAVP 116
          HVS+C DVKLALQGMN LE + + LDF WR+EL +K +FPL ++TFGE +P
Sbjct: 358 HVSICADVKLALQGMNATLEQQQRK-NLDFS AWRSELEKKKAEFPLGYRTFGEEIP 412

>sp|Q7XKQ8| ILV2_ORYSJ Probable acetolactate synthase 2, chloroplast precursor
(Aceto-hydroxy-acid synthase 2)
emb|CAE05539.2| OSJNBa0053B21.13 [Oryza sativa (japonica cultivar-group)]
gb|EAY93893.1| hypothetical protein OsI_015126 [Oryza sativa (indica
cultivar-group)]
Length = 663

Score = 178 bits (451), Expect = 2e-43, Method: Compositional matrix adjust.
Identities = 83/119 (69%), Positives = 99/119 (83%), Gaps = 3/119 (2%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
          MLGMHGT VYANYAV++D LLLA GVRFD DRVTGK+EAFASRAKIVH+DID +E+GKNK P
Sbjct: 336 MLGMHGT VYANYAVDNADLLLA LGVRFD DRVTGKVEAFASRAKIVHVDIDPSELGKNKQP 395

Query: 61 HVSVC G DVKLALQGMNKVLENR---AEELK LDFGVWRNELNVQKQKFPLSFKTFGEAVP 116
          HVS+C DVKLALQGMN +LE + A LDF WR+EL +K +FPL ++TFGE +P
Sbjct: 396 HVSICADVKLALQGMNAMLEEQSAAAARKNLDFS AWRSELEKKKVEFPLGYRTFGEEIP 454

>emb|CAH66432.1| OSIGBa0096P03.6 [Oryza sativa (indica cultivar-group)]
Length = 663

Score = 177 bits (450), Expect = 2e-43, Method: Compositional matrix adjust.
Identities = 83/119 (69%), Positives = 99/119 (83%), Gaps = 3/119 (2%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
          MLGMHGT VYANYAV++D LLLA GVRFD DRVTGK+EAFASRAKIVH+DID +E+GKNK P
Sbjct: 336 MLGMHGT VYANYAVDNADLLLA LGVRFD DRVTGKVEAFASRAKIVHVDIDPSELGKNKQP 395

Query: 61 HVSVC G DVKLALQGMNKVLENR---AEELK LDFGVWRNELNVQKQKFPLSFKTFGEAVP 116
          HVS+C DVKLALQGMN +LE + A LDF WR+EL +K +FPL ++TFGE +P
Sbjct: 396 HVSICADVKLALQGMNAMLEEQSAAAARKNLDFS AWRSELEKKKVEFPLGYRTFGEEIP 454

>gb|ABF66049.1| acetolactate synthase [Oryza sativa (indica cultivar-group)]
Length = 644
```

Score = 176 bits (445), Expect = 7e-43, Method: Compositional matrix adjust.
Identities = 81/116 (69%), Positives = 95/116 (81%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+ +DLLAFGVRFDDRTVGK+EAFASRAK +DID AEIGKNK P
Sbjct: 322 MLGMHGT VYANYAVDKADLLAFGVRFDDRTVGKIEAFASRAKSLVDIDPAEIGKNKQP 381

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS+C DVKLALQG+N +L+ + DF W NEL+ QK++FPL +KTFGE +P
Sbjct: 382 HVSICADV KLALQGLNALLDQSTTKTSSDFS AWHNELDQKREFPLGYKTFGE EIP 437

>emb|CAO49020.1| unnamed protein product [Vitis vinifera]
Length = 657

Score = 175 bits (444), Expect = 8e-43, Method: Compositional matrix adjust.
Identities = 86/116 (74%), Positives = 94/116 (81%), Gaps = 7/116 (6%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT+YANYAV+ SDLLAFGVRFDDRTVGK+EAFASRAKIVHIDID AEIGKNK P
Sbjct: 342 MLGMHGTMYANYAVDRSDLLAFGVRFDDRTVGKIEAFASRAKIVHIDIDPAEIGKNKQP 401

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
HVS AL GMN +LE + K DF WR+ELN QK K+PLSFKTFGEA+P
Sbjct: 402 HVS-----ALTGMNSLLEETGAKSKYDFSSWRDELNEQAKYPLSFKTFGEAIP 450

>gb|ABR26179.1| acetolactate synthase III [Oryza sativa (indica cultivar-group)]
Length = 151

Score = 173 bits (438), Expect = 4e-42, Method: Compositional matrix adjust.
Identities = 79/110 (71%), Positives = 92/110 (83%)

Query: 7 TVYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTPHVSVC 66
TVYANYAV+ +DLLAFGVRFDDRTVGK+EAFASRAKIVHIDID AEIGKNK PHVS+C
Sbjct: 6 TVYANYAVDKADLLAFGVRFDDRTVGKIEAFASRAKIVHIDIDPAEIGKNKQPHVSICA 65

Query: 67 DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVP 116
DVKLALQG+N +L+ + DF W NEL+ QK++FPL +KTFGE +P
Sbjct: 66 DVKLALQGLNALLDQSTTKTSSDFS AWHNELDQKREFPLGYKTFGE EIP 115

>emb|CAP09635.1| acetolactate-synthase-N-DnaE intein-N fusion protein
[Transformation Vector pICH13688]
Length = 558

Score = 171 bits (434), Expect = 1e-41, Method: Compositional matrix adjust.
Identities = 91/119 (76%), Positives = 95/119 (79%), Gaps = 5/119 (4%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP
Sbjct: 348 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 407

Query: 61 HVSVC GDVKLALQGMNKVLENRAE--ELKLD FGVWRNELNVQKQKFPLSFKTFGEAVPC 117
HVSVC GDVKLALQGMNKVLEN + E L FG L V+ P+ K E + C
Sbjct: 408 HVSVC GDVKLALQGMNKVLENDVKFAEYCLSGT--EILTVEYGPLPIG-KIVSEEINC 463

>ref|XP_001759950.1| predicted protein [Physcomitrella patens subsp. patens]
gb|EDQ75075.1| predicted protein [Physcomitrella patens subsp. patens]
Length = 583

Score = 162 bits (411), Expect = 7e-39, Method: Compositional matrix adjust.
Identities = 77/118 (65%), Positives = 94/118 (79%), Gaps = 1/118 (0%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV+++D+LLAFGVRFDDRTVGKLE+FASRA IVHIDID AEIGKNK P
Sbjct: 261 MLGMHGT VYANYAVDNADMLLAFGVRFDDRTVGKLESFASRASIVHIDIDPAEIGKNKQP 320

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVPCP 118

H+S+C DV+LAL G+NK+++ E + DF WR EL+ K K+P+ F F +V P
Sbjct: 321 HISICADVQLALAGLNKLIKEGPAE-RPDFSAWRAELDGVKLKWPMPKFPKFDNSVIVP 377

>emb|CAN73760.1| hypothetical protein [Vitis vinifera]
Length = 586

Score = 160 bits (405), Expect = 3e-38, Method: Compositional matrix adjust.
Identities = 78/129 (60%), Positives = 94/129 (72%), Gaps = 13/129 (10%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT+ ANYAV+ SDLLAFGVRFDDRVTK+EAF A IVHIDID AEIGKNK P
Sbjct: 256 MLGMHGTIQANYAVDRSDLLAFGVRFDDRVTKVEAFARNATIVHIDIDPAEIGKNKKP 315

Query: 61 HVSVC G DVKLALQGMNKVLENRAE-----ELKLD FGVWRNELNVQKQKFPLS 107
H+S+C DVKLAL+G+N +LE A + + W E++ QK+K+P S
Sbjct: 316 HLSICTDVKLALLEGINTILEKNAAKQPTAENKRGKGTKFNDNVSAWIEEIDEQKEKYPAS 375

Query: 108 FKTFGEAVP 116
+KTFGEA+P
Sbjct: 376 YKTFGEAIP 384

>ref|XP_001758473.1| predicted protein [Physcomitrella patens subsp. patens]
gb|EDQ76875.1| predicted protein [Physcomitrella patens subsp. patens]
Length = 618

Score = 159 bits (401), Expect = 9e-38, Method: Compositional matrix adjust.
Identities = 74/118 (62%), Positives = 91/118 (77%), Gaps = 1/118 (0%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYA+++SD+LLAFGVRFDDRVTKLE+FA SRA IVHIDID AEIGKNK P
Sbjct: 296 MLGMHGT VYANYAIDNSDMLLAFGVRFDDRVTKLESFASRASIVHIDIDPAEIGKNKQP 355

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVPCP 118
H+S+C DV+LAL G+NK+++ + F WR EL+ K+K+P+ F V P
Sbjct: 356 HISICADVQLALAGLNKLIKE-GPATRPSFSAWRKELDGVKEKWPMPKFPKLD SQVIVP 412

>dbj|BAE53595.1| acetolactate synthase [Monochoria vaginalis]
Length = 566

Score = 150 bits (379), Expect = 4e-35, Method: Compositional matrix adjust.
Identities = 71/89 (79%), Positives = 79/89 (88%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYA++ +DLLAFGVRFDDRVTKLEAFASRAKIVHIDID AEIGKNK P
Sbjct: 244 MLGMHGT VYANYAIDKADLLAFGVRFDDRVTKLEAFASRAKIVHIDIDPAEIGKNKQP 303

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD 89
HVS+CGD+KLALQ MN+++E KLD
Sbjct: 304 HVSICGDIKLALQEMNEMIEESGIHNKLD 332

>gb|AAC04854.1| acetolactate synthase [Vulvox carteri]
Length = 681

Score = 149 bits (375), Expect = 1e-34, Method: Compositional matrix adjust.
Identities = 72/115 (62%), Positives = 90/115 (78%), Gaps = 2/115 (1%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT V ANYAV+ +DLLA GVRFDDRVTG+L+AFASRA+IVH+DID+AEI KNKT
Sbjct: 348 MLGMHGTVAANYAVDQADLLVALGVRFDDRV TGR L DAFASRARIVHVIDIDAAEISKNKTA 407

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAV 115
HV VCGDVK AL+ +N++LE AE L F WR EL ++ +FPL + +A+
Sbjct: 408 HVPVCGDVKQALRHLNRMLE--AEPLSDRFVAWRAELA AKRAEFPLRYPQRDDAI 460

>gb|AAB88296.1| acetolactate synthase [Vulvox carteri]

Length = 681

Score = 148 bits (374), Expect = 1e-34, Method: Compositional matrix adjust.
Identities = 72/115 (62%), Positives = 90/115 (78%), Gaps = 2/115 (1%)Query: 1 MLGMHGTVYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTV ANYAV+ +DLL+A GVRFDDRVTKL+AFASRA+IVH+DID+AEI KNKT
Sbjct: 348 MLGMHGTVAANYAVDQADLLVALGVRFDDRVTKRLDAFASRARIVHVDIDAAEISKNKTA 407Query: 61 HVSVCADVKLALQGMNKVLENRAEELKLD-FGVWRNELNVQKQKFPLSFKTFGEAV 115
HV VCGDVK AL+ +N++LE AE L F WR EL ++ +FPL + +A+
Sbjct: 408 HVPVCGDVKQALRHLNRMLE--AEPLSDRFVAVRAELAAKRAEFPLRYPQRDDAI 460>ref|XP_001695168.1| acetolactate synthase, large subunit [Chlamydomonas reinhardtii]
gb|EDP01876.1| acetolactate synthase, large subunit [Chlamydomonas reinhardtii]
Length = 640Score = 145 bits (367), Expect = 9e-34, Method: Compositional matrix adjust.
Identities = 71/116 (61%), Positives = 90/116 (77%), Gaps = 3/116 (2%)Query: 1 MLGMHGTVYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTV+ANYAV+ +DLL+A GVRFDDRVTKL+AF+RA+IVHIDID+AEI KNKT
Sbjct: 306 MLGMHGTVFANYAVDQADLLVALGVRFDDRVTKLDAFAARARIVHIDIDAAEISKNKTA 365Query: 61 HVSVCADVKLALQGMNKVLENRAEELKLD-FGVWRNELNVQKQKFPLSFKTFGEAV 115
HV VCGDVK AL +N++L AE L D + WR EL ++ +FP+ + +A+
Sbjct: 366 HVPVCGDVKQALSHLNRLA--AEPLADKWAGWRAELAAKRAEFPMRYPQRDDAI 419>gb|AAC03784.1| acetolactate synthase [Chlamydomonas reinhardtii]
Length = 683Score = 145 bits (366), Expect = 9e-34, Method: Compositional matrix adjust.
Identities = 71/116 (61%), Positives = 90/116 (77%), Gaps = 3/116 (2%)Query: 1 MLGMHGTVYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTV+ANYAV+ +DLL+A GVRFDDRVTKL+AF+RA+IVHIDID+AEI KNKT
Sbjct: 349 MLGMHGTVFANYAVDQADLLVALGVRFDDRVTKLDAFAARARIVHIDIDAAEISKNKTA 408Query: 61 HVSVCADVKLALQGMNKVLENRAEELKLD-FGVWRNELNVQKQKFPLSFKTFGEAV 115
HV VCGDVK AL +N++L AE L D + WR EL ++ +FP+ + +A+
Sbjct: 409 HVPVCGDVKQALSHLNRLA--AEPLADKWAGWRAELAAKRAEFPMRYPQRDDAI 462>gb|AAB88292.1| acetolactate synthase [Chlamydomonas reinhardtii]
Length = 683Score = 145 bits (366), Expect = 1e-33, Method: Compositional matrix adjust.
Identities = 71/116 (61%), Positives = 90/116 (77%), Gaps = 3/116 (2%)Query: 1 MLGMHGTVYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTV+ANYAV+ +DLL+A GVRFDDRVTKL+AF+RA+IVHIDID+AEI KNKT
Sbjct: 349 MLGMHGTVFANYAVDQADLLVALGVRFDDRVTKLDAFAARARIVHIDIDAAEISKNKTA 408Query: 61 HVSVCADVKLALQGMNKVLENRAEELKLD-FGVWRNELNVQKQKFPLSFKTFGEAV 115
HV VCGDVK AL +N++L AE L D + WR EL ++ +FP+ + +A+
Sbjct: 409 HVPVCGDVKQALSHLNRLA--AEPLADKWAGWRAELAAKRAEFPMRYPQRDDAI 462>emb|CA045071.1| unnamed protein product [Vitis vinifera]
Length = 463Score = 144 bits (364), Expect = 2e-33, Method: Compositional matrix adjust.
Identities = 73/116 (62%), Positives = 86/116 (74%), Gaps = 2/116 (1%)Query: 1 MLGMHGTVYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT+ ANYAV+ SDLLAFGVRFDDRVTK+EAF A IVHIDID AEIGKNK P
Sbjct: 148 MLGMHGTIQANYAVDRSDLLAFGVRFDDRVTKVEAFARNATIVHIDIDPAEIGKNKKP 207

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPPLSFKTFGEAVP 116
H+S+C DVKLAL+G+N +LE A K + + P S++TFGEA+P
Sbjct: 208 HLSICTDVKLALLEGINTILEKNAA--KQPTAENKRGKGTGFNDNPPSYRTFGEAIP 261

>emb|CA017032.1| unnamed protein product [Vitis vinifera]
Length = 866

Score = 144 bits (364), Expect = 2e-33, Method: Compositional matrix adjust.
Identities = 73/116 (62%), Positives = 86/116 (74%), Gaps = 2/116 (1%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT+ ANYAV+ SDLLAFGVRFDRTGK+EAFA A IVHIDID AEIGKNK P
Sbjct: 551 MLGMHGTIQANYAVDRSDLLAFGVRFDRTGKVEAFARNATIVHIDIDPAEIGKNKKP 610

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPPLSFKTFGEAVP 116
H+S+C DVKLAL+G+N +LE A K + + P S++TFGEA+P
Sbjct: 611 HLSICTDVKLALLEGINTILEKNAA--KQPTAENKRGKGTGFNDNPPSYRTFGEAIP 664

>emb|CAL58226.1| acetolactate synthase 1 (ISS) [Ostreococcus tauri]
Length = 679

Score = 134 bits (336), Expect = 3e-30, Method: Compositional matrix adjust.
Identities = 64/106 (60%), Positives = 80/106 (75%), Gaps = 1/106 (0%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHG VYANYA++ +DLLAFGVRFDRTGKL FA RA IVHIDID AE+GKNK P
Sbjct: 343 MLGMHGA VYANYAIDSADLLAFGVRFDRTGKLAFAKRAAIVHIDIDPAELGKNKKP 402

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPPL 106
S+ ++K AL+ +NK++ A L DF WR E+ V+++FP
Sbjct: 403 FCSIFSNIKPALRSLNKLIALDAANLP-DFSDWRAEIEVKRKEFPF 447

>ref|XP_001421626.1| predicted protein [Ostreococcus lucimarinus CCE9901]
gb|AB099919.1| predicted protein [Ostreococcus lucimarinus CCE9901]
Length = 610

Score = 132 bits (331), Expect = 1e-29, Method: Compositional matrix adjust.
Identities = 64/106 (60%), Positives = 80/106 (75%), Gaps = 2/106 (1%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHG VYANYA++ +DLLAFGVRFDRTGKL FA RA IVHIDID AE+GKNK P
Sbjct: 275 MLGMHGA VYANYAIDSADLLAFGVRFDRTGKLAFAKRAAIVHIDIDPAELGKNKKP 334

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPPL 106
S+ ++K AL+ +NK++ E+ DF WR ++VQ++KFP
Sbjct: 335 FCSIFSNIKPALKTLNKLIA--LEKDMPDFSEWRAAIDVQRKKFPF 378

>ref|YP_001941018.1| Acetolactate synthase large subunit or other thiamine
pyrophosphate-requiring enzyme [Methylokorus inferorum
V4]
gb|ACD84421.1| Acetolactate synthase large subunit or other thiamine
pyrophosphate-requiring enzyme [Methylacidiphilum
inferorum V4]
Length = 589

Score = 125 bits (314), Expect = 1e-27, Method: Compositional matrix adjust.
Identities = 60/117 (51%), Positives = 86/117 (73%), Gaps = 2/117 (1%)

Query: 2 LGMHGT VYANYAVEHSDLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTPH 61
LGMHG+VY+N+AV+ SDLLAFGVRFDRTGK++AFA +A+IVHIDID++E+ KNK
Sbjct: 274 LGMHGSVYSNFAVDQSDLLAFGVRFDRTGNVQAFQAKARIVHIDIDNSELNKNKRV 333

Query: 62 VSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPPLSFKTFGEAVPCP 118
+ + GD+K AL+ +N+++E R + + +W+ E+ K K+P +K F E V P
Sbjct: 334 LPILGDIKEALRLNRLIEERKFQ-SPGYSLWQEEIQAWKIKYPFRYKPF-EGVIMP 388

>ref|NP_682086.1| acetolactate synthase 3 catalytic subunit [Thermosynechococcus
elongatus BP-1]
dbj|BAC08848.1| acetohydroxy acid synthase [Thermosynechococcus elongatus BP-1]
Length = 602

Score = 125 bits (313), Expect = 2e-27, Method: Compositional matrix adjust.
Identities = 61/119 (51%), Positives = 78/119 (65%), Gaps = 7/119 (5%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+ FASRAK++HIDID AE+GKN+ P
Sbjct: 259 MLGMHGTAYANFAVSEC DLLIAVGARFDDRVTKGLDEFASRAKVIHIDIDPAEVGKNRVP 318

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVPCPS 119
V + G VK LQ + K +E+ + ++ W + + KQ +PL VP PS
Sbjct: 319 DVPIVGSVKPVLQQLLKHIEDSGQAVEPQTQAWLERITIWKQDYPL-----IVPQPS 370

>gb|EDX84416.1| acetolactate synthase, large subunit, biosynthetic type
[Synechococcus sp. PCC 7335]
Length = 589

Score = 122 bits (305), Expect = 1e-26, Method: Compositional matrix adjust.
Identities = 59/109 (54%), Positives = 75/109 (68%), Gaps = 6/109 (5%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+ FAS AK++HIDID AE+GKN+ P
Sbjct: 274 MLGMHGTAYANFAVSEC DLLVAIGARFDDRVTKGLDEFASSAKVIHIDIDPAEVGKNRVP 333

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPL 106
V + GDVK L ++L+ + + LD WR+++ KQ FPL
Sbjct: 334 EVPIVGDVKAFLA---RMLDRLSGQAALDVNRSQSWRDQIETWKQDFPL 379

>ref|YP_502706.1| acetolactate synthase, large subunit, biosynthetic type
[Methanospirillum hungatei JF-1]
gb|ABD40987.1| acetolactate synthase, large subunit [Methanospirillum hungatei
JF-1]
Length = 563

Score = 119 bits (298), Expect = 8e-26, Method: Compositional matrix adjust.
Identities = 62/125 (49%), Positives = 80/125 (64%), Gaps = 5/125 (4%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YANYA+ SDLL + GVRFDDRVTKG+EAFA +AKI+HIDID AEIGKNK P
Sbjct: 254 MLGMHGTAYANYAITESD LLSIGVRFDDRVTKGIEAFAPQAKIIHIDIDPAEIGKNKQP 313

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVPCPSK 120
V + GD K L + K + + W ++ + K+ PL +++ G+ P
Sbjct: 314 DVPIVGDAKSVLSDILKKIPESCPHPE-----WSEKVAMWKKNHPLRYRSDGKLYPQFVI 68

Query: 121 KTVHD 125
KT+ D
Sbjct: 369 KTLAD 373

>ref|ZP_02963213.1| acetolactate synthase 1 catalytic subunit [Bifidobacterium animalis
subsp. lactis HN019]
gb|EDT89371.1| acetolactate synthase 1 catalytic subunit [Bifidobacterium animalis
subsp. lactis HN019]
Length = 642

Score = 119 bits (297), Expect = 1e-25, Method: Compositional matrix adjust.
Identities = 58/110 (52%), Positives = 75/110 (68%), Gaps = 1/110 (0%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT+ A AV+ SDLL+A G RFDDRVTKLEAFA A+++HIDID AEIGKN+TP
Sbjct: 283 MLGMHGTIAATGAVQRS DLLVAIGARFDDRVTKLEAFAPGARVIHIDIDPAEIGKNRTP 342

Query: 61 HVSVC G DVKLALQG-MNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFK 109

V + GDVK L M ++ A K + W +N +K+PL+++
Sbjct: 343 DVPIVGDVKTIVLSALMPRIERQHAIHGKPNLTTWWGLINSWVEKYPLTYQ 392

>ref|YP_001293618.1| acetohydroxyacid synthetase large subunit [Rhodomonas salina]
gb|AB070722.1| acetohydroxyacid synthetase large subunit [Rhodomonas salina]
Length = 574

Score = 118 bits (296), Expect = 1e-25, Method: Compositional matrix adjust.
Identities = 56/106 (52%), Positives = 73/106 (68%)

Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV DLL+A G RFDDRVTKGL+ FA A++VHIDID AE+GKN+ P
Sbjct: 267 MLGMHGTIVYANYAVSECDLLIALGARFDDRVTKGLDEFACHAQVHIDIDPAEVGKNRMP 326

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGFGVWRNENLVQKQKFPL 106
V + GDVK +++ + VL+ WR L+ ++++PL
Sbjct: 327 QVGIVGDVKESVKAILDVLLKKEDYLEPEQTQAWRQRLDRWRKEYPL 372

>ref|ZP_01731805.1| acetolactate synthase III large subunit [Cyanothec sp. CCY0110]
gb|EAS88774.1| acetolactate synthase III large subunit [Cyanothec sp. CCY0110]
Length = 619

Score = 118 bits (295), Expect = 2e-25, Method: Compositional matrix adjust.
Identities = 57/120 (47%), Positives = 79/120 (65%), Gaps = 9/120 (7%)

Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKGLE FASRA+++HIDID AE+GKN+ P
Sbjct: 275 MLGMHGTAYANFAVSECDLLIAGSRFDDRVTKGLEEFASRARVIHIDIDPAEVGKNRPL 334

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGFGV-----WRNENLVQKQKFPLSFKTFGEAV 115
V + GDV+ L+ L R+ E+ + W N +N ++++PL+ + ++
Sbjct: 335 EVPIVGDVRQVLEQ---LLQRSREIDVPLDAEKTPWLNRLNRWREQYPLTAPHYDHSI 390

>ref|YP_001936332.1| acetolactate synthase large subunit [Heterosigma akashiwo]
gb|ABV65938.1| acetolactate synthase large subunit [Heterosigma akashiwo]
gb|ABV70079.1| acetolactate synthase large subunit [Heterosigma akashiwo]
Length = 608

Score = 118 bits (295), Expect = 2e-25, Method: Compositional matrix adjust.
Identities = 67/166 (40%), Positives = 95/166 (57%), Gaps = 15/166 (9%)

Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT +AN+AV DLL+A G RFDDRVTKGL+ FA+ AKI+HIDID AEI KN P
Sbjct: 267 MLGMHGTAFANFAVNECDLLIAGARFDDRVTKGLDEFAGAKIIHIDIDPAEIAKNCIP 326

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGFGV-WRNENLVQKQKFPLSFKTFGEAVPCPS 119
V++ GDVK+ LQ + ++ N ++ + + WR+ +N ++ +PL +P P
Sbjct: 327 QVALIGDVKIILQELIEIYTNNIDQYEEKSTLPWRHRINAWQEAYPL-----LIPAPE 379

Query: 120 KKTVDHPSYEWVGPPIRREGPPSPMREDKLTQNTHLIRTYKFVFFKC 165
+ +Y I G +P RE T + + + F KC
Sbjct: 380 E-----AYSPQVINEIGKVAP-RETIFTTVDVGHQMWAAQFLKC 418

>ref|ZP_01857674.1| acetolactate synthase III (Precursor) [Planctomyces maris DSM 8797]
gb|EDL56438.1| acetolactate synthase III (Precursor) [Planctomyces maris DSM 8797]
Length = 593

Score = 117 bits (294), Expect = 2e-25, Method: Compositional matrix adjust.
Identities = 59/115 (51%), Positives = 77/115 (66%), Gaps = 3/115 (2%)

Query: 1 MLGMHGTIVYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGTIVYANYAV +DLLAFGVRFDDRVTKGLE FA KIVH+DID +E+ KNK
Sbjct: 269 MLGMHGTIVYANYAVNEADLLAFGVRFDDRVTKGLEEFAKHGKIVHVIDIDPSELQKNKEA 328

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGFGVWRNENLVQKQKFPLSFKTFGEAV 115
H+ + D+K L +N+ + + ++L W + K+KFPL + G+ +

Sbjct: 329 HIPINADLKHVLTTELNEAITD--DDL P-QVDSWLAQCKEWEKFP LKYPELG DVM 380

>ref|NP_849033.1| acetohydroxyacid synthase large subunit [Cyanidioschyzon merolae strain 10D]
dbj|BAC76195.1| acetolactate synthase large subunit [Cyanidioschyzon merolae strain 10D]
Length = 569

Score = 117 bits (293), Expect = 3e-25, Method: Composition-based stats.
Identities = 50/73 (68%), Positives = 61/73 (83%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRV T GKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRV T GK L+ FAS A+++HIDID+AEIGKN+ P
Sbjct: 255 MLGMHGTAYANFAVSECDLLIALGARFDDRV T GK LDEFASSAQVIHIDIDAAEIGKNRIP 314

Query: 61 HVSVC G DVKLALQ 73
+++C DVKL LQ
Sbjct: 315 QLAICSDV KLV LQ 327

>ref|YP_399158.1| acetolactate synthase 3 catalytic subunit [Synechococcus elongatus PCC 7942]
gb|ABB56171.1| acetolactate synthase, large subunit [Synechococcus elongatus PCC 7942]
prf||1611501A acetolactate synthase
Length = 612

Score = 117 bits (292), Expect = 4e-25, Method: Compositional matrix adjust.
Identities = 54/119 (45%), Positives = 79/119 (66%), Gaps = 7/119 (5%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRV T GKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRV T GK L+ FAS+A+++H+DID AE+GKN+ P
Sbjct: 268 MLGMHGTAYANFAVSECDLLIAGARFDDRV T GK LDEFASKAQVIHVIDIDPAEVGKNRVP 327

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELK LDFGV----WRNELNVQKQKFPLSFKTFGEAV 115
V + GDV+ Q +N++L E++ D W + + K+++PL + + +
Sbjct: 328 EVPIVG D VR---QVLNELLARAE EQISAD DATRTQPWLDRIAYWKREYPLQIPYYADVL 383

>ref|YP_172076.1| acetolactate synthase 3 catalytic subunit [Synechococcus elongatus PCC 6301]
dbj|BAD79556.1| acetolactate synthase [Synechococcus elongatus PCC 6301]
Length = 612

Score = 117 bits (292), Expect = 4e-25, Method: Compositional matrix adjust.
Identities = 54/119 (45%), Positives = 79/119 (66%), Gaps = 7/119 (5%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRV T GKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRV T GK L+ FAS+A+++H+DID AE+GKN+ P
Sbjct: 268 MLGMHGTAYANFAVSECDLLIAGARFDDRV T GK LDEFASKAQVIHVIDIDPAEVGKNRVP 327

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELK LDFGV----WRNELNVQKQKFPLSFKTFGEAV 115
V + GDV+ Q +N++L E++ D W + + K+++PL + + +
Sbjct: 328 EVPIVG D VR---QVLNELLARAE EQISAD DATRTQPWLDRIAYWKREYPLQIPYYADVL 383

>ref|NP_926225.1| acetolactate synthase 3 catalytic subunit [Gloeobacter violaceus PCC 7421]
dbj|BAC91220.1| acetohydroxyacid synthetase large subunit [Gloeobacter violaceus PCC 7421]
Length = 596

Score = 117 bits (292), Expect = 4e-25, Method: Compositional matrix adjust.
Identities = 58/112 (51%), Positives = 74/112 (66%), Gaps = 4/112 (3%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRV T GKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRV T GK L+ FA+RA+++HIDID AE+GKN+ P
Sbjct: 258 MLGMHGTAYANFAVTECDLLIAGARFDDRV T GK LDEFAARARVIHIDIDPAEVGKNRGP 317

Query: 61 HVSVCQDVKLALQGMNKLLENRAEELKLDLFGVWRNELNVQKQKPLSFKTFG 112
V + GDVKL LQ + K L E + W + K ++PL + G
Sbjct: 318 EVPIVGDVQLVLQDLIKKL----IEFQPRTRAWLERIERWKAEPLEVPDSDG 365

>ref|ZP_01621589.1| acetolactate synthase III large subunit [Lyngbya sp. PCC 8106]
gb|EAW36367.1| acetolactate synthase III large subunit [Lyngbya sp. PCC 8106]
Length = 612

Score = 116 bits (291), Expect = 5e-25, Method: Compositional matrix adjust.
Identities = 55/116 (47%), Positives = 75/116 (64%), Gaps = 1/116 (0%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKLE FASRA+++HIDID AE+GKN+ P
Sbjct: 268 MLGMHGTAYANFAVSECDLLIAGGARFDDRVTKLEEFASRAEVIHIDIDPAEVGKNRAP 327

Query: 61 HVSVCQDVKLALQGMNKLLENRAEELKLDLFGVWRNELNVQKQKPLSFKTFGEAV 115
V + GDV+ L + + E + W N +N ++ +PL + +++
Sbjct: 328 QVPVIGDVRQVLIDLLRRCRETEEPAPQTQTEAWLNRIWREDYPLEVPEYPSL 383

>ref|YP_001805892.1| acetolactate synthase [Cyanothecae sp. ATCC 51142]
gb|ACB53826.1| acetolactate synthase [Cyanothecae sp. ATCC 51142]
Length = 619

Score = 116 bits (291), Expect = 5e-25, Method: Compositional matrix adjust.
Identities = 56/111 (50%), Positives = 75/111 (67%), Gaps = 9/111 (8%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKL+ FASRA+++HIDID AE+GKN+ P
Sbjct: 275 MLGMHGTAYANFAVSECDLLIAGGARFDDRVTKLDEFASRARVIHIDIDPAEVGKNRPL 334

Query: 61 HVSVCQDVKLALQGMNKLLENRAEELKLDLFGV-----WRNELNVQKQKPL 106
V + GDV+ L+ L R+ E+ + W N +N ++++PL
Sbjct: 335 EVPIVGDVRQVLEQ----LLQRSREIDVPLDAEKTTPWLNRIWREQYPL 381

>ref|ZP_01079763.1| acetolactate synthase [Synechococcus sp. RS9917]
gb|EAQ69739.1| acetolactate synthase [Synechococcus sp. RS9917]
Length = 590

Score = 116 bits (291), Expect = 5e-25, Method: Composition-based stats.
Identities = 58/118 (49%), Positives = 75/118 (63%), Gaps = 9/118 (7%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT ANYAV+ SDLLAFGVRFDDRVTKL+ FA RA+++H +ID AEIGKN+
Sbjct: 250 MLGMHGTAYANFAVTECDLLIAGGARFDDRVTKLDTFAPRARVIHFEIDPAEIGKNRQA 309

Query: 61 HVSVCQDVKLALQGMNKLLENRAEELKLDLFGVWRNELNVQKQKPLSFKTFGEAVPCP 118
V+V GD+ L+L M ++ R EE + W + KQ++PL VP P
Sbjct: 310 DVAVLGDLGLSLARMVEMSLQRHEEPT--AAWLQRIEAWKQRYPL-----VVPSP 358

>emb|CAO42659.1| unnamed protein product [Vitis vinifera]
Length = 616

Score = 116 bits (290), Expect = 6e-25, Method: Compositional matrix adjust.
Identities = 63/116 (54%), Positives = 74/116 (63%), Gaps = 27/116 (23%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT ANYAV+ SDLLAFGVRFDDRVTKL+ FA A IVHIDID AEIGKNK P
Sbjct: 326 MLGMHGTIANYAVDRSDLLAFGVRFDDRVTKLVESFARNATIVHIDIDPAEIGKNKKP 385

Query: 61 HVSVCQDVKLALQGMNKLLENRAEELKLDLFGVWRNELNVQKQKPLSFKTFGEAVP 116
H+S+C D A EN K+ +++TFGEA+P
Sbjct: 386 HLSICTDQPTA-----EN-----KRGKGTNYRTFGEAIP 414

>ref|NP_897837.1| acetolactate synthase 3 catalytic subunit [Synechococcus sp. WH 8102]

sp|Q7U5G1|ILVB_SYNPX Acetolactate synthase large subunit (AHAS) (Acetohydroxy-acid synthase large subunit) (ALS)
emb|CAE08261.1| acetolactate synthase [Synechococcus sp. WH 8102]
Length = 617

Score = 116 bits (290), Expect = 6e-25, Method: Composition-based stats.
Identities = 55/107 (51%), Positives = 72/107 (67%), Gaps = 2/107 (1%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTGKL+ FA RA++VH +ID AEIGKN+
Sbjct: 277 MLGMHGTAYANFAVTECDLLIAVGARFDDRVTGKLDTFAPRARVVHFEIDPAEIGKNRKA 336

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLS 107
V+V GD+ L+L M ++ R E + W +N K ++PL+
Sbjct: 337 DVAVLGDLGLSLARMVEISLQRTAEPT--AAWLERINTWKDRYPLT 381

>ref|ZP_02971818.1| acetolactate synthase, large subunit, biosynthetic type [Cyanothecae sp. PCC 7424]
gb|EDU18743.1| acetolactate synthase, large subunit, biosynthetic type [Cyanothecae sp. PCC 7424]
Length = 619

Score = 116 bits (290), Expect = 6e-25, Method: Compositional matrix adjust.
Identities = 58/116 (50%), Positives = 76/116 (65%), Gaps = 1/116 (0%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTGKLE FASRAK+VHIDID AE+GK + P
Sbjct: 275 MLGMHGTAYANFAVTECDLLIAVGARFDDRVTGKLEEFASRAKVVHIDIDPAEVGKVRAP 334

Query: 61 HVSVC G DVKLALQGM-NKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAV 115
V + GDV+ L+ M +V E + W +N ++ +PL T+ +++
Sbjct: 335 DVPIVGDVQVLEQMLQVRVREMDYPTVPERTKDWLERINRWREDYPLMVPTYPDSI 390

>ref|ZP_01085428.1| acetolactate synthase III large subunit [Synechococcus sp. WH 5701]
gb|EAQ74912.1| acetolactate synthase III large subunit [Synechococcus sp. WH 5701]
Length = 590

Score = 115 bits (288), Expect = 1e-24, Method: Composition-based stats.
Identities = 53/110 (48%), Positives = 76/110 (69%), Gaps = 2/110 (1%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTG+L++FA RA+++HIDID+AE+GKN+ P
Sbjct: 250 MLGMHGTAYANFAVTECDLLIAAGARFDDRVTGRLD SFAPRAQVIHIDIDAAEVGKNRLP 309

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKT 110
V + DV+LAL+ ++ ++ + G W ++ K+ +PL T
Sbjct: 310 EVPIVSDVRLALEALLEASVGEDNGRT--GPWLERISRWRHYPLVIPT 357

>ref|ZP_01469460.1| acetolactate synthase III large subunit [Synechococcus sp. BL107]
gb|EAU70474.1| acetolactate synthase III large subunit [Synechococcus sp. BL107]
Length = 587

Score = 115 bits (288), Expect = 1e-24, Method: Composition-based stats.
Identities = 53/107 (49%), Positives = 74/107 (69%), Gaps = 2/107 (1%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTGKL+ FA RAK++H +ID AE+GKN+T
Sbjct: 278 MLGMHGTAYANFAVTECDLLIAVGARFDDRVTGKLDTFAPRAKVIHFEIDPAEVGKNRTA 337

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLS 107
V+V GD+ L+++ ++ R + W ++ KQ++PLS
Sbjct: 338 EVAVLGDLSLSMERLVQLSMER--QANPSTAAWLEQIQDWKQRYPLS 382

>ref|NP_488653.1| acetolactate synthase 3 catalytic subunit [Nostoc sp. PCC 7120]
dbj|BAB76312.1| acetohydroxy acid synthase [Nostoc sp. PCC 7120]
Length = 632

Score = 115 bits (288), Expect = 1e-24, Method: Compositional matrix adjust.
Identities = 56/117 (47%), Positives = 76/117 (64%), Gaps = 1/117 (0%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+ G RFDDRTGKL+ FASRAK++HIDID AE+GKN+ P
Sbjct: 289 MLGMHGTAYANFAVTD CDLLICVGARFD DRTGKLDEFASRAKVIHIDIDPAEVGKNRIP 348

Query: 61 HVSVC GDV-KLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFLSFKTFGEAVP 116
V + GDV K+ L + + + A+ W N +N ++ +PL +++P
Sbjct: 349 EVPIVG DVRKVL LLDLLRCKQTS AKTTPNQNEWLNLINRWREDYPLVVPQHPDSIP 405

>ref|YP_323265.1| acetolactate synthase 3 catalytic subunit [Anabaena variabilis ATCC
29413]
gb|ABA22370.1| acetolactate synthase, large subunit [Anabaena variabilis ATCC
29413]
Length = 632

Score = 115 bits (288), Expect = 1e-24, Method: Compositional matrix adjust.
Identities = 56/117 (47%), Positives = 76/117 (64%), Gaps = 1/117 (0%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+ G RFDDRTGKL+ FASRAK++HIDID AE+GKN+ P
Sbjct: 289 MLGMHGTAYANFAVTD CDLLICVGARFD DRTGKLDEFASRAKVIHIDIDPAEVGKNRIP 348

Query: 61 HVSVC GDV-KLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFLSFKTFGEAVP 116
V + GDV K+ L + + + A+ W N +N ++ +PL +++P
Sbjct: 349 EVPIVG DVRKVL LLDLLRCKQTG AKNTPNQNEWLNLINRWREDYPLVIPQHPDSIP 405

>ref|ZP_03272318.1| acetolactate synthase, large subunit, biosynthetic type
[Arthrospira maxima CS-328]
gb|EDZ96099.1| acetolactate synthase, large subunit, biosynthetic type
[Arthrospira maxima CS-328]
Length = 612

Score = 115 bits (287), Expect = 1e-24, Method: Compositional matrix adjust.
Identities = 58/121 (47%), Positives = 79/121 (65%), Gaps = 11/121 (9%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRTGKL+ FASRAK++HIDID AE+GKN+TP
Sbjct: 268 MLGMHGTAYANFAVSEC DLLIAGARFD DRTGKLDEFASRAKVIHIDIDPAEVGKNRTP 327

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD F-----VWRNELNVQKQKFLSFKTFGEA 114
V + GDV+ L + + R E+ D G W +N ++ +PL ++ ++
Sbjct: 328 EVPIVG DVRQVLIDL L R----RCREIG-DVGNDNQTSWLERINRWREDYPLVVP SYSDS 382

Query: 115 V 115
+
Sbjct: 383 L 383

>ref|YP_731202.1| acetolactate synthase 3 catalytic subunit [Synechococcus sp.
CC9311]
gb|ABI45115.1| acetolactate synthase, large subunit, biosynthetic type
[Synechococcus sp. CC9311]
Length = 581

Score = 115 bits (287), Expect = 1e-24, Method: Compositional matrix adjust.
Identities = 55/108 (50%), Positives = 74/108 (68%), Gaps = 2/108 (1%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRTGKL+ FA +AK++H +ID AE+GKN+ P
Sbjct: 265 MLGMHGTAYANFAVTD CDLLIAGARFD DRTGKLDTFAPKAKVIHFEIDPAEVGKNRCP 324

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFLSF 108
V V GDV L+L + V ++R +L W ++N K+ +PL+
Sbjct: 325 DVVVLGDVGLSLAQL--VDQSRPHSAELTTSSWLEQINSWKELYPLTI 370


```
>ref|YP_001111655.1| acetolactate synthase, large subunit, biosynthetic type
[Desulfotomaculum reducens MI-1]
gb|AB048830.1| acetolactate synthase, large subunit [Desulfotomaculum reducens
MI-1]
Length = 554

Score = 115 bits (287), Expect = 1e-24, Method: Compositional matrix adjust.
Identities = 58/115 (50%), Positives = 79/115 (68%), Gaps = 6/115 (5%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+A+ DLL+A GVRFDDRVTK+E+FA AKI+HIDID AE+GKN
Sbjct: 254 MLGMHGT KYANFAICDCDLLIAVGVRFDDRVTKVESFAPNAKIIHIDIDPAELGKNVRV 313

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAV 115
V + G+VKL LQ + + ++ R E W++++ K+++PL F GE +
Sbjct: 314 DVPIAGNVKLV LQQLLERVQTRIPE-----AWQHKVAQWKKEYPLDFDENGEG 362

>ref|ZP_00514506.1| Acetolactate synthase, large subunit, biosynthetic type
[Crocospaera watsonii WH 8501]
gb|EAM52344.1| Acetolactate synthase, large subunit, biosynthetic type
[Crocospaera watsonii WH 8501]
Length = 619

Score = 114 bits (286), Expect = 2e-24, Method: Compositional matrix adjust.
Identities = 56/109 (51%), Positives = 76/109 (69%), Gaps = 5/109 (4%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+ FASRA+++HIDID AE+GKN+ P
Sbjct: 275 MLGMHGT AYANFAVSECDLLIAVGARFDDRVTKGLDEFASRARVIHIDIDPAEVGKNRRP 334

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD---FGVWRNELNVQKQKFPL 106
V + GDV L+ + + +R ++ LD W + +N +Q++PL
Sbjct: 335 DVPIVGDAQVLEQL--LQRSREIDVPLDPQKTEAWLHRINRWRQQYPL 381

>ref|ZP_02737021.1| acetolactate synthase III [Gemmata obscuriglobus UQM 2246]
Length = 590

Score = 114 bits (286), Expect = 2e-24, Method: Compositional matrix adjust.
Identities = 60/111 (54%), Positives = 75/111 (67%), Gaps = 3/111 (2%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VY+NYA+ +DLLAFGVRFDDRVTKGL FA KIVH+DID +EI KNK
Sbjct: 268 MLGMHGT VYSNYAINDADLLAFGVRFDDRVTKGLAEFAKHGKIVHVIDIKSEIHKNKFA 327

Query: 61 HVSVC GDVKLALQGMNKVL-ENRAEELKLD--FGVWRNELNVQKQKFPLSF 108
HV V D+K AL G+N +L E + +L + W +++ + PL F
Sbjct: 328 HVPVHSDLKHALHGLNALLKEKNADLTAGGRYTDWWRQVD AWRDAEPLKF 378

>ref|YP_001046985.1| acetolactate synthase, large subunit, biosynthetic type
[Methanoculleus marisnigri JR1]
gb|ABN57003.1| acetolactate synthase, large subunit [Methanoculleus marisnigri
JR1]
Length = 558

Score = 114 bits (285), Expect = 2e-24, Method: Compositional matrix adjust.
Identities = 56/109 (51%), Positives = 69/109 (63%), Gaps = 5/109 (4%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT ANYAV DLL+A GVRFDDRVTK+E FA A ++HIDID AEIGKNK
Sbjct: 254 MLGMHGT QSANYAVTECDLLVAVGVRFDDRVTKIETFAPNAAVIHIDIDPAEIGKNKKV 313

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFK 109
V + GDVK LQ ++ R D W + + K ++PL ++
Sbjct: 314 DVPIVGDKAVLQAFLSRMQKRG-----DTAAWVSRI GAWKAQYPLGYR 357
```

>ref|ZP_03142166.1| acetolactate synthase, large subunit, biosynthetic type [Cyanothecae
sp. PCC 8802]
gb|EDY03144.1| acetolactate synthase, large subunit, biosynthetic type [Cyanothecae
sp. PCC 8802]
Length = 619

Score = 114 bits (285), Expect = 2e-24, Method: Compositional matrix adjust.
Identities = 57/120 (47%), Positives = 77/120 (64%), Gaps = 9/120 (7%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+ FAS AK++HIDID AE+GKN+TP
Sbjct: 275 MLGMHGTAYANFAVSECDDLLIAGGARFDDRVTKGLDEFASHAKVIHIDIDPAEVGKNRTP 334

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGFGV-----WRNELNVQKQKFPLSFKTFGEAV 115
V + GDV+ L+ L RA E+ W + +N ++++PL + ++
Sbjct: 335 DVPIVGDVRRVLE----QLLQRRAREIDCPHTAEKTPWLSRINRWREQYPLVIPHYDHSI 390

>ref|ZP_02942745.1| acetolactate synthase, large subunit, biosynthetic type [Cyanothecae
sp. PCC 8801]
gb|EDT59304.1| acetolactate synthase, large subunit, biosynthetic type [Cyanothecae
sp. PCC 8801]
Length = 594

Score = 114 bits (285), Expect = 3e-24, Method: Compositional matrix adjust.
Identities = 57/120 (47%), Positives = 77/120 (64%), Gaps = 9/120 (7%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+ FAS AK++HIDID AE+GKN+TP
Sbjct: 250 MLGMHGTAYANFAVSECDDLLIAGGARFDDRVTKGLDEFASHAKVIHIDIDPAEVGKNRTP 309

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKLDGFGV-----WRNELNVQKQKFPLSFKTFGEAV 115
V + GDV+ L+ L RA E+ W + +N ++++PL + ++
Sbjct: 310 DVPIVGDVRRVLE----QLLQRRAREIDCPHTAEKTPWLSRINRWREQYPLVIPHYDHSI 365

>ref|NP_441297.1| acetolactate synthase 3 catalytic subunit [Synechocystis sp. PCC
6803]
dbj|BAA17977.1| acetohydroxy acid synthase [Synechocystis sp. PCC 6803]
Length = 621

Score = 114 bits (285), Expect = 3e-24, Method: Compositional matrix adjust.
Identities = 60/130 (46%), Positives = 80/130 (61%), Gaps = 11/130 (8%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+ FASRAK++HIDID AE+GKN+ P
Sbjct: 277 MLGMHGTAYANFAVSECDDLLIAGGARFDDRVTKGLDEFASRAKVIHIDIDPAEVGKNRAP 336

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKL-----DFGVWRNELNVQKQKFPLSFKTFGEAV 115
V + GDV+ L+ L RA EL W N ++ + +PL + + +
Sbjct: 337 DVPIVGDVRRHVLE----QLLQRRARELDYPTHPHTTQAWLNRIHWRTDYPLQVPHYEDTI 392

Query: 116 PCPSKKTVDH 125
++ VH+
Sbjct: 393 --APQEVVHE 400

>ref|YP_001517397.1| acetolactate synthase 3 catalytic subunit [Acaryochloris marina
MBIC11017]
gb|ABW28081.1| acetolactate synthase, large subunit, biosynthetic type
[Acaryochloris marina MBIC11017]
Length = 602

Score = 114 bits (285), Expect = 3e-24, Method: Compositional matrix adjust.
Identities = 54/106 (50%), Positives = 71/106 (66%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+ FASRAK++HIDID AE+GKN+ P
Sbjct: 259 MLGMHGTAYANFAVSECDDLLIAGGARFDDRVTKGLDEFASRAKVIHIDIDPAEVGKNRAP 318

Query: 61 HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPL 106
V + G VK L + K+ + + WR+ + ++ +PL
Sbjct: 319 TVPIVGSVKPVLVELLKI AKQSSPPDTAQTQEWRSRVERWQRDYPL 364

>ref|NP_053940.1| acetohydroxyacid synthase large subunit [Porphyra purpurea]
sp|P69683|ILVB_PORPU Acetolactate synthase large subunit (AHAS) (Acetohydroxy-acid
synthase large subunit) (ALS)
sp|P69684|ILVB_PORUM Acetolactate synthase large subunit (AHAS) (Acetohydroxy-acid
synthase large subunit) (ALS)
gb|AAA03052.1| acetolactate synthase
gb|AAC08216.1| acetohydroxyacid synthase large subunit [Porphyra purpurea]
Length = 590

Score = 114 bits (285), Expect = 3e-24, Method: Compositional matrix adjust.
Identities = 50/106 (47%), Positives = 72/106 (67%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRTVGKL+ FA A+++H+DID AE+GKN+ P
Sbjct: 267 MLGMHGTAYANFAVSECDLLIALGARFDDRTVGKLEDEFACNAQVIHVDIDPAEVGKNRIP 326

Query: 61 HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPL 106
V++ GDV + + +L+N + W+ ++ +Q++PL
Sbjct: 327 QVAIVGDVTEVVTSLNLLKNNFKPYPEQIISWQERIHRWRQYPL 372

>ref|NP_050806.1| acetohydroxyacid synthase large subunit [Guillardia theta]
sp|O78518|ILVB_GUIITH Acetolactate synthase large subunit (AHAS) (Acetohydroxy-acid
synthase large subunit) (ALS)
gb|AAC35740.1| acetohydroxyacid synthetase large subunit [Guillardia theta]
Length = 575

Score = 114 bits (285), Expect = 3e-24, Method: Compositional matrix adjust.
Identities = 52/106 (49%), Positives = 74/106 (69%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV DLL+A G RFDDRTVGKL+ FA A+++H+DID AEIGKN+TP
Sbjct: 266 MLGMHGT VYANYAVSECDLLIALGARFDDRTVGKLEDEFACHAQVIHVDIDPAEIGKNRTP 325

Query: 61 HVSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPL 106
+ + G++K ++ + + L+N WR+ + ++++PL
Sbjct: 326 QIGIVGEIKDFVRDLIECLKNDINFDSQSAWRSRIIRWRKEYPL 371

>ref|NP_870771.1| acetolactate synthase III [precursor] [Rhodopirellula baltica SH 1]
emb|CAD77848.1| acetolactate synthase III [Precursor] [Rhodopirellula baltica SH 1]
Length = 619

Score = 114 bits (285), Expect = 3e-24, Method: Compositional matrix adjust.
Identities = 57/107 (53%), Positives = 71/107 (66%), Gaps = 6/107 (5%)

Query: 2 LGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTPH 61
LGMHG YANYAV DLL+A GVRFDDRTVGK+EAFA AKI+H+DIDS+E+ KNK H
Sbjct: 296 LGMHGAAYANYAVRDCDLLIALGVRFDDRTVGKVEAFKDAKIIHVDIDDSSELNKNKQAH 355

Query: 62 VSVCGDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSF 108
+ V GDVK L +NK+++ E W+ K K+PL +
Sbjct: 356 IPVRGDKDVLVQLNKIVQPPEIE-----AWQKTCTDLKAKYPLKY 396

>ref|YP_063568.1| acetohydroxyacid synthase large subunit [Gracilaria tenuistipitata
var. liui]
gb|AAT79643.1| acetolactate synthase large subunit [Gracilaria tenuistipitata var.
liui]
Length = 596

Score = 113 bits (283), Expect = 4e-24, Method: Compositional matrix adjust.
Identities = 53/108 (49%), Positives = 74/108 (68%), Gaps = 2/108 (1%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60

MLGMHGT YAN+AV DLL+A G RFDDRVTGKL+ FA AK++HIDID AEIGKN+ P
Sbjct: 268 MLGMHGTAYANFAVSECDLLIALGARFDDRVTGKLDEFACNAKVIHIDIDPAEIGKNRVP 327

Query: 61 HVSVCQDVKLALQGMNKVLENRAEEL--KLDFGVWRNELNVQKQKFPL 106
V++ GDVK + + L+N + + W++ +++ + ++PL
Sbjct: 328 QVAILGDVKQVMVQILSYLDNSSFVFNSSQQTYSWKDRISMWRDQYPL 375

>ref|ZP_03137434.1| acetolactate synthase, large subunit, biosynthetic type [Cyanothecce
sp. PCC 7425]
gb|EDY12146.1| acetolactate synthase, large subunit, biosynthetic type [Cyanothecce
sp. PCC 7425]
Length = 602

Score = 113 bits (283), Expect = 5e-24, Method: Compositional matrix adjust.
Identities = 60/122 (49%), Positives = 76/122 (62%), Gaps = 9/122 (7%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTGKL+ FASRAK++HIDID AE+GKN+ P
Sbjct: 259 MLGMHGTAYANFAVSECDLLIAVGARFDDRVTGKLDEFASRAKVIHIDIDPAEIGKNRVP 318

Query: 61 HVSVCQDVKLALQGMNKVLENRAEELKL-DFGVWRNELNVQKQKFPLSFKTFGEAVPCPS 119
V + G V+ L + + L E KL W ++ KQ +PL VP P+
Sbjct: 319 EVPIVGSQVPVLAELLR-LAREGEPALGQTRAWLERIDRWKQDYPL-----VVPYPA 370

Query: 120 KK 121
+
Sbjct: 371 DR 372

>ref|YP_001736021.1| acetolactate synthase, large subunit, biosynthetic type
[Synechococcus sp. PCC 7002]
gb|ACB00766.1| acetolactate synthase, large subunit, biosynthetic type
[Synechococcus sp. PCC 7002]
Length = 591

Score = 113 bits (282), Expect = 5e-24, Method: Compositional matrix adjust.
Identities = 57/116 (49%), Positives = 76/116 (65%), Gaps = 1/116 (0%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHG+ YAN+AV DLL+A G RFDDRVTGKL+ FAS AK++HIDID AE+GKN+ P
Sbjct: 275 MLGMHGSAYANFAVSECDLLIAVGARFDDRVTGKLDEFASEAKVIHIDIDPAEIGKNRVP 334

Query: 61 HVSVCQDVKLALQGMNKVLENRAEEL-KLDFGVWRNELNVQKQKFPLSFKTFGEAV 115
V++ GDV+ L M + L + L W N++ K+K PL F +++
Sbjct: 335 EVAIVGDVHRVLIEMLQQLGGKGLPLGNKPTTEEWLNKIQGWKEKHPLVAPRFADSI 390

>ref|NP_618663.1| acetolactate synthase 3 catalytic subunit [Methanosarcina
acetivorans C2A]
gb|AAM07143.1| acetolactate synthase, large subunit [Methanosarcina acetivorans
C2A]
Length = 564

Score = 113 bits (282), Expect = 5e-24, Method: Compositional matrix adjust.
Identities = 55/111 (49%), Positives = 76/111 (68%), Gaps = 10/111 (9%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHG YANYAV+ SDL++A G RFDDRVTGKLE+FA A+++HID+D AEI KN
Sbjct: 259 MLGMHGCKYANYAVQESDLIIAVGARFDDRVTGKLESFAPNARVIHIDVDPAEISKNVKV 318

Query: 61 HVSVCQDVKLALQGMNKVLE--NRAEELKLDFGVWRNELNVQKQKFPLSFK 109
HV + GD K L+ + K ++ N AE W ++N K+++PL+++
Sbjct: 319 HVPVIGDAKQILKSLIKYIQRCSAE-----WIEKINQWKKEYPLAYR 361

>ref|YP_377644.1| acetolactate synthase 3 catalytic subunit [Synechococcus sp.
CC9902]
gb|ABB26600.1| acetolactate synthase, large subunit [Synechococcus sp. CC9902]
Length = 618

Score = 113 bits (282), Expect = 5e-24, Method: Compositional matrix adjust.
Identities = 53/107 (49%), Positives = 72/107 (67%), Gaps = 2/107 (1%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRTVGKL+ FA RAK++H +ID AE+GKN++P
Sbjct: 278 MLGMHGTAYANFAVTECDLLIAGGARFDDRTVGKLDTFAPRAKVIHFEIDPAEVGKNRSP 337

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLS 107
V+V GD+ L++ + ++ R E W + KQ +PL+
Sbjct: 338 EVAVLGDLSLSMARLVQLSLGR--EADPHTAAWLKRIQSWKQHYPLT 382

>ref|YP_722490.1| acetolactate synthase 3 catalytic subunit [Trichodesmium erythraeum
IMS101]
gb|ABG52017.1| acetolactate synthase, large subunit [Trichodesmium erythraeum
IMS101]
Length = 587

Score = 112 bits (281), Expect = 7e-24, Method: Compositional matrix adjust.
Identities = 58/120 (48%), Positives = 77/120 (64%), Gaps = 9/120 (7%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YANYAV DLL+A G RFDDRTVGKL+ FAS AK++HIDID AE+GKN++P
Sbjct: 268 MLGMHGTAYANYAVSECDLLIAGGARFDDRTVGKLDQFASYAKVIHIDIDPAEVGKNRSP 327

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLDGFG----VWRNELNVQKQKFPLSFKTFGEAV 115
V + GDV+ L M LE R +++ + W + KQ +PL + +++
Sbjct: 328 EVPIVG DVVRQVLIDM---LE-RCQKIDTNVSSNQTTWLERIKRWKQDYPLVVPHYSDSM 383

>ref|YP_537012.1| acetohydroxyacid synthase large subunit [Porphyra yezoensis]
sp|Q1XDF6|ILVB_PORYE Acetolactate synthase large subunit (AHAS) (Acetohydroxy-acid
synthase large subunit) (ALS)
dbj|BAE92455.1| acetohydroxyacid synthase large subunit [Porphyra yezoensis]
Length = 590

Score = 112 bits (281), Expect = 8e-24, Method: Compositional matrix adjust.
Identities = 50/106 (47%), Positives = 70/106 (66%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRTVGKL+ FA A+++H+DID AE+GKN+ P
Sbjct: 267 MLGMHGTAYANFAVSECDLLIALGARFDDRTVGKLDEFACNAQVIHVDIDPAEVGKNRIP 326

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPL 106
V++ GDV + + +L+ W+ +N +Q++PL
Sbjct: 327 QVAIVGDVAEVVSEILNLLKTSFPPYPEQIISWQERINRWRQYPL 372

>ref|YP_001930430.1| acetolactate synthase, large subunit, biosynthetic type
[Sulfurihydrogenibium sp. Y03AOP1]
gb|ACD65876.1| acetolactate synthase, large subunit, biosynthetic type
[Sulfurihydrogenibium sp. Y03AOP1]
Length = 582

Score = 112 bits (280), Expect = 8e-24, Method: Compositional matrix adjust.
Identities = 58/109 (53%), Positives = 71/109 (65%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN AV HSDLL+A G RFDDRTVGK+ FA AKI+HIDID A I K T
Sbjct: 259 MLGMHGTYYANMAVYHSDLLIAGGARFDDRTVGKINEFAPEAKIIHIDIDPASISKITITV 318

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPLSFK 109
V + GDVK L+ + K LE + E W ++N K+K PL+++
Sbjct: 319 DVPIVG DVKNVLRKLIKELEEKPIEWIAAREQWLKQINNEWKEKHPLNYR 367

>ref|YP_359378.1| acetolactate synthase, large subunit, biosynthetic type
[Carboxydotherrmus hydrogenoformans Z-2901]
gb|ABB14028.1| acetolactate synthase, large subunit, biosynthetic type

[Carboxydotherrmus hydrogenoformans Z-2901]
Length = 554

Score = 112 bits (280), Expect = 9e-24, Method: Compositional matrix adjust.
Identities = 58/111 (52%), Positives = 73/111 (65%), Gaps = 10/111 (9%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLG+HGT YAN AV D+L+ GVRF DRVTG+L FA +AKI+HID+D AEIGKN
Sbjct: 253 MLGLHGTRYANLAVTECDVLIGLGVRFADRVTEGELSGFAPKAKIIHIDVDPAEIGKNVRA 312

Query: 61 HVSVC GDVKLALQGMNKVLE--NRAEELKLD FGVWRNELNVQKQKFPLSFK 109
V + GDVK LQ M K +E NR E W +++N KQ+FPL ++
Sbjct: 313 DVPIVG DVKNVLQEMLKQIEPQNRQE-----WLSQINTWKQEFPLKYE 355

>ref|YP_001227109.1| acetolactate synthase 3 catalytic subunit [Synechococcus sp.
RCC307]
emb|CAK27756.1| Acetolactate synthase large subunit [Synechococcus sp. RCC307]
Length = 615

Score = 112 bits (280), Expect = 1e-23, Method: Compositional matrix adjust.
Identities = 53/107 (49%), Positives = 74/107 (69%), Gaps = 2/107 (1%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTGKL+ FA RAK++H +ID AE+GKN+ P
Sbjct: 275 MLGMHGTAYANFAVTD CDLLI AVGARFDDRVTGKLDTFAPRAKVIHFIDPAEVGKNRKP 334

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLS 107
V+V GDV+ +L+ + + + R ++ W ++ KQ +PL+
Sbjct: 335 DVAVLGDVRRSLEELLQSQQR--HVEPHTAAWLARIDSWKQNYPLT 379

>ref|YP_002249089.1| acetolactate synthase, large subunit, biosynthetic type
[Thermodesulfovibrio yellowstonii DSM 11347]
gb|ACI21111.1| acetolactate synthase, large subunit, biosynthetic type
[Thermodesulfovibrio yellowstonii DSM 11347]
Length = 576

Score = 112 bits (279), Expect = 1e-23, Method: Compositional matrix adjust.
Identities = 61/123 (49%), Positives = 82/123 (66%), Gaps = 8/123 (6%)

Query: 1 MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN AV++SDL++A G+RFDDRVTGK +AFA AKI+HIDID I KN
Sbjct: 255 MLGMHGTYANMAVQNSDLIVAIGMRFD DRVTGKTDAPAPNAKIIHIDIDPTSIRKNVRV 314

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD F---GVWRNELNVQKQKFPLSFKTFGEAVP 116
+ + GDV LQ +NK+L+ EE+K + W ++N K++ PL+++ F EAV
Sbjct: 315 DIPIVG DVSRVLQVLNKILK---EEVKPQWKEIRQAWLKQINQWKKERPLTYE-FDEAVI 370

Query: 117 CPS 119
P
Sbjct: 371 KPQ 373

>ref|YP_002250909.1| acetolactate synthase, large subunit, biosynthetic type
[Dictyoglomus thermophilum H-6-12]
gb|ACI19136.1| acetolactate synthase, large subunit, biosynthetic type
[Dictyoglomus thermophilum H-6-12]
Length = 557

Score = 111 bits (278), Expect = 1e-23, Method: Compositional matrix adjust.
Identities = 53/114 (46%), Positives = 73/114 (64%), Gaps = 6/114 (5%)

Query: 2 LGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHIDIDSAEIGKNKTPH 61
+GMHG YANYA+ SDL++A GVRF DR TGK+E FA AKI+HIDID AEIGKN P+
Sbjct: 255 IGMHGAAYANYAINDSDLI AIGVRFSDRSTGKVETFAPNAKIIHIDIDPAEIGKNVQPY 314

Query: 62 VSVCG DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAV 115
V + D K AL+ + +V+E + +W ++ K K+PL ++ E +
Sbjct: 315 VPIVADAKRALEKLI EVVEPKTNP-----IWWEKIKDWKNKYPLRYRMSNEVI 362

>ref|YP_001211077.1| thiamine pyrophosphate-requiring enzymes [Pelotomaculum
thermopropionicum SI]
dbj|BAF58708.1| thiamine pyrophosphate-requiring enzymes [Pelotomaculum
thermopropionicum SI]
Length = 555

Score = 111 bits (278), Expect = 2e-23, Method: Compositional matrix adjust.
Identities = 57/116 (49%), Positives = 73/116 (62%), Gaps = 6/116 (5%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKLE FA A++HIDID AEIGKN
Sbjct: 256 MLGMHGTYANFAVCECDLLIAGGARFDDRVTKLETFAPPEARIIHIDIDPAEIGKNVRV 315

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKPLSFKTFGEAVP 116
+ + GDVK L + ++L E WR ++ K+++PL++ G P
Sbjct: 316 DIPIVGDVKRVLSQLLEILRPGLRE-----AWREKIEAWKKEYPLTYCEQGR LKP 365

>ref|YP_303782.1| acetolactate synthase 3 catalytic subunit [Methanosarcina barkeri
str. Fusaro]
gb|AAZ69202.1| acetolactate synthase, large subunit [Methanosarcina barkeri str.
Fusaro]
Length = 588

Score = 111 bits (278), Expect = 2e-23, Method: Compositional matrix adjust.
Identities = 53/118 (44%), Positives = 75/118 (63%), Gaps = 6/118 (5%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT ANYA++ SDL++A G RFDDRVTKLE+FA A+++HID+D AEI KN
Sbjct: 259 MLGMHGTSANYAIQESDLIIAGGARFDDRVTKLESFAPNARVIHIDVDPAEISKNNV 318

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKPLSFKTFGEAVPCP 118
+ + GD K L+ +NK ++ E W ++ K+++PL++K V P
Sbjct: 319 QIPIVGDAKQVLKSLNKYIKCKSE-----AWLEKIKQWKKEYPLTYKQMSSDVIMP 370

>gb|EDX75556.1| acetolactate synthase, large subunit, biosynthetic type
[Microcoleus chthonoplastes PCC 7420]
Length = 619

Score = 111 bits (278), Expect = 2e-23, Method: Compositional matrix adjust.
Identities = 56/119 (47%), Positives = 75/119 (63%), Gaps = 7/119 (5%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKL+ FA AK++HIDID AE+GKN+ P
Sbjct: 275 MLGMHGTAYANFAVSECDLLIAGGARFDDRVTKLDQFACHAKVIHIDIDPAEVGKNRAP 334

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKL----DFGVWRNELNVQKQKPLSFKTFGEAV 115
V + GDV+ Q + K+L E+ L VW + +Q++PL + + +
Sbjct: 335 QVPIVGDVQ---QVLAKLLRYYEQDILADPNQTQVWLKRIQRWRQEYPLEAPQYPDTI 390

>ref|YP_754808.1| acetolactate synthase large subunit [Syntrophomonas wolfei subsp.
wolfei str. Goettingen]
gb|ABI69437.1| acetolactate synthase, large subunit [Syntrophomonas wolfei subsp.
wolfei str. Goettingen]
Length = 562

Score = 111 bits (277), Expect = 2e-23, Method: Compositional matrix adjust.
Identities = 56/118 (47%), Positives = 74/118 (62%), Gaps = 3/118 (2%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YANYA+ DLL+A GVRFDDRVTK++ FA A+++HIDID+AEIGKN
Sbjct: 254 MLGMHGTRYANYAIGECDLLIAGVRFDDRVTKIDTFAPHARVIHIDIDAAEIGKNVEV 313

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKPLSFKTFGEAVPCP 118
V + G VK L +N+ LE E + W ++ K+++PL + E P
Sbjct: 314 EVPIVGQVKEVLAAINQRLEAIESE---ELNEWHETIHRWKEEYPLRYGDSSEGRIMP 368

>ref|ZP_01472584.1| acetolactate synthase III large subunit [Synechococcus sp. RS9916]
gb|EAU73275.1| acetolactate synthase III large subunit [Synechococcus sp. RS9916]
Length = 624

Score = 110 bits (276), Expect = 3e-23, Method: Composition-based stats.
Identities = 54/118 (45%), Positives = 78/118 (66%), Gaps = 5/118 (4%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKL+ FA RA+++H +ID AE+GK +
Sbjct: 284 MLGMHGTAYANFAVTECDLLIAVGARFDDRVTKGLDTFAPRARVIHFEIDPAEVGKTRRA 343

Query: 61 HVSVC GDVKLALQGMNKVLE-NRAEELKLD FGVWRNELNVQKQKFP LSFKT-FGEAVP 116
V+V GD+ Q M +++E +R +++ W ++ K+++PL+ T GE P
Sbjct: 344 EVAVLGDLA---QSMQLVERSRQDIRPTTA AWLAKITTWKEQYPLTVPTPEGEIYP 398

>ref|NP_045101.1| acetohydroxyacid synthase large subunit [Cyanidium caldarium]
sp|O19929|ILVB_CYACA Acetolactate synthase large subunit (AHAS) (Acetohydroxy-acid
synthase large subunit) (ALS)
gb|AAB82660.1| unknown; acetohydroxyacid synthase large subunit [Cyanidium
caldarium]
Length = 585

Score = 110 bits (276), Expect = 3e-23, Method: Compositional matrix adjust.
Identities = 52/116 (44%), Positives = 76/116 (65%), Gaps = 15/116 (12%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKL+ FA A+++H+DID AEIGKN+ P
Sbjct: 268 MLGMHGTAYANFAVSECDLLIALGARFDDRVTKGLDEFACNAQVIHVDIDPAEIGKNRIP 327

Query: 61 HVSVC GDVKLAL-----QGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 108
+++ D+K+ L +G N + +N+ + W + ++ K+++PLS
Sbjct: 328 QLAIISDIKIVLKELLSSMKEGTNNMDKNQTQ-----AWLHRIHKWKKEYPLSI 376

>ref|ZP_01089063.1| acetolactate synthase III (Precursor) [Blastopirellula marina DSM
3645]
gb|EAQ82178.1| acetolactate synthase III (Precursor) [Blastopirellula marina DSM
3645]
Length = 590

Score = 110 bits (276), Expect = 3e-23, Method: Compositional matrix adjust.
Identities = 58/115 (50%), Positives = 73/115 (63%), Gaps = 5/115 (4%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT VYANYAV++ DLL++ GVRFDDRVTKL FA AKI+HIDID +EI KNK
Sbjct: 268 MLGMHGT VYANYAVQNCDDLISLGVRFDDRVTKLA EFAKHAKIIHIDIDPSEINKNKLA 327

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSFKT FGEAV 115
H V D+K AL+ + ++E D W + K+ PL+ T E +
Sbjct: 328 HFPVVSDLK PALRMLCDIVEKPE-----DISEWVAQCAKWKDEPLKYDTSFEGI 377

>ref|NP_632694.1| acetolactate synthase 3 catalytic subunit [Methanosarcina mazei
Gol]
gb|AAM30366.1| Acetolactate synthase large subunit [Methanosarcina mazei Gol]
Length = 564

Score = 110 bits (275), Expect = 4e-23, Method: Compositional matrix adjust.
Identities = 51/110 (46%), Positives = 77/110 (70%), Gaps = 8/110 (7%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHG YANYA++ SDL++A G RFDDRVTKLE+FA A+++HID+D AEI KN
Sbjct: 259 MLGMHGCKYANYAIQESDLIIAVGARFDDRVTKLESFAPNARVIHIDVDPAEISKNVKV 318

Query: 61 HVSVC GDVKLALQGMNKVLE-NRAEELKLD FGVWRNELNVQKQKFP LSF 109
H+ + GD K L+ + + ++ R+ E W ++N+ K+++PL+++
Sbjct: 319 HIPIVGDAKQVLKSLIRYVQCCRS AE-----WIEKVN LWKKEYPLNYR 361

>ref|YP_001660084.1| acetolactate synthase 3 catalytic subunit [Microcystis aeruginosa NIES-843]
dbj|BAG04892.1| acetohydroxy acid synthase [Microcystis aeruginosa NIES-843]
Length = 617

Score = 110 bits (275), Expect = 4e-23, Method: Compositional matrix adjust.
Identities = 54/107 (50%), Positives = 71/107 (66%), Gaps = 1/107 (0%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRV TGKL+ FAS+AK++H+DID AE+GKN+ P
Sbjct: 275 MLGMHGTAYANFAVTECDLLIAGGARFDDRV TGKLDEFASKAKVIHLIDIDPAEVGKNRAP 334

Query: 61 HVSVC G DVKLAL-QGMNKVLENRAEELKLD FGVWRNELNVQKQKFPL 106
V + GDV++ L Q + K E W ++ +Q +PL
Sbjct: 335 DVPIVG DV RVVLEQILQKAREFDYPTNSDRTQAWLAQIERWRQDYPL 381

>gb|AAA26594.1| acetohydroxy acid synthase (AHAS)
Length = 612

Score = 110 bits (274), Expect = 4e-23, Method: Compositional matrix adjust.
Identities = 56/121 (46%), Positives = 76/121 (62%), Gaps = 11/121 (9%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV D L+A G RFDDRV TGKL+ F SRAK++HIDID AE+GKN+TP
Sbjct: 268 MLGMHGTAYANFAVSECDFLIAGGARFDDRV TGKLDEF GSRKVIHIDIDPAEVGKNRTP 327

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD F-----VWRNELNVQKQKFPLSFKTFGEA 114
V + GDV+ L + + R E+ D G W +N + +PL ++ ++
Sbjct: 328 EVPIVG DV RVQLIHLR----RCREIG-DVGNDNQTSWLERINRWPEDYPLVVP SYSDS 382

Query: 115 V 115
+
Sbjct: 383 L 383

>ref|ZP_01123205.1| acetolactate synthase III large subunit [Synechococcus sp. WH 7805]
gb|EAR20062.1| acetolactate synthase III large subunit [Synechococcus sp. WH 7805]
Length = 597

Score = 110 bits (274), Expect = 4e-23, Method: Composition-based stats.
Identities = 52/107 (48%), Positives = 71/107 (66%), Gaps = 2/107 (1%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRV TGKL+ FA RA+++H +ID AEIGKN+
Sbjct: 257 MLGMHGTAYANFAVTDCLLIAGGARFDDRV TGKLDTFAPRARVIHFEIDPAEIGKNRRA 316

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLS 107
V+V GDV +L + ++ + E + W + KQ++PL+
Sbjct: 317 DVAVLGDVGASLAALVELSLQQGHEPRT--AAWLERIKDWKQRYPLT 361

>emb|CAA12081.1| acetohydroxy acid synthase [Porphyridium sp.]
Length = 587

Score = 110 bits (274), Expect = 4e-23, Method: Compositional matrix adjust.
Identities = 50/106 (47%), Positives = 68/106 (64%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+ G RFDDRV TGKL+ FA AK++H+DID AE+GKN+ P
Sbjct: 268 MLGMHGTAYANFAVSECDLLITL GARFDDRV TGKLDEFACNAKVIHVIDIDPAEVGKNRIP 327

Query: 61 HVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPL 106
V++ GD+ L L+ L+ + W + +Q++PL
Sbjct: 328 QVAIVGDISLVLEQWLLYLDRLQLDDSHLRSWHERIFRWRQEYPL 373

>ref|ZP_00206563.1| COG0028: Thiamine pyrophosphate-requiring enzymes [acetolactate

synthase, pyruvate dehydrogenase (cytochrome),
glyoxylate carboligase, phosphonopyruvate decarboxylase]
[Bifidobacterium longum DJO10A]
ref|YP_001954986.1| Acetolactate synthase large unit [Bifidobacterium longum DJO10A]
gb|ACD98488.1| Acetolactate synthase large unit [Bifidobacterium longum DJO10A]
Length = 649

Score = 110 bits (274), Expect = 5e-23, Method: Composition-based stats.
Identities = 58/137 (42%), Positives = 84/137 (61%), Gaps = 6/137 (4%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT+ A AV+ +DLL+A G RFDDRV TGKL+AFA A+++HIDID AEIGKN+ P
Sbjct: 276 MLGMHGTIAATGAVQRADLLVAIGARFDDRV TGKLD AFAPTARVIHIDIDPAEIGKNRQP 335

Query: 61 HVSVC G DVKLALQGM-NKVL ENRAEELKLD FGVWRNELNVQKQKFPLSFK--TFGEAVP- 116
V + GDV L + ++ +A + K + W ++ ++++P+++ T G P
Sbjct: 336 DVPIVGDVATVLDLPIEQRTQAIQKPNLAPWWKAIDGWREEYPM TWDEPTDGLAPQ 395

Query: 117 --CPSKKT VHDPSYEWV 131
+ DPS WV
Sbjct: 396 WVIKKLSEMAPSTI WV 412

>ref|ZP_02028034.1| hypothetical protein BIFADO_00444 [Bifidobacterium adolescentis
L2-32]
gb|EDN83535.1| hypothetical protein BIFADO_00444 [Bifidobacterium adolescentis
L2-32]
Length = 624

Score = 110 bits (274), Expect = 5e-23, Method: Compositional matrix adjust.
Identities = 60/137 (43%), Positives = 85/137 (62%), Gaps = 6/137 (4%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT+ A AV+ DLL+A G RFDDRV TGKL+AFA A+++HIDID AEIGKN+ P
Sbjct: 286 MLGMHGTIAATGAVQRCDLLVAIGARFDDRV TGKLD AFAPGARVIHIDIDPAEIGKNRQP 345

Query: 61 HVSVC G DVKLALQGM-NKVL ENRAEELKLD FGVWRNELNVQKQKFPLSFK--TFGEAVPC 117
V + GDV L + ++ +A K D W + +N ++++P+++ T G P
Sbjct: 346 DVPIVGDVATVLDLPIEKREQAVHGKPD TSHWVDVINKWREEYPITWDEPTDGLAPQ 405

Query: 118 PSKKT V H---DPSYEWV 131
K + DP+ WV
Sbjct: 406 WVKQLSDMAPNTV WV 422

>ref|ZP_03155476.1| acetolactate synthase, large subunit, biosynthetic type [Cyanothec
sp. PCC 7822]
gb|EDX97152.1| acetolactate synthase, large subunit, biosynthetic type [Cyanothec
sp. PCC 7822]
Length = 619

Score = 110 bits (274), Expect = 5e-23, Method: Compositional matrix adjust.
Identities = 54/116 (46%), Positives = 74/116 (63%), Gaps = 1/116 (0%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRV TGKLE FA AK++HIDID AE+GK + P
Sbjct: 275 MLGMHGTAYANFAVTECDLLIAVGARFDDRV TGKLEFARNAKVIHIDIDPAEIVGKVRGP 334

Query: 61 HVSVC G DVKLAL-QGMNKVL ENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAV 115
V + GDV+ L Q + +V E W + +N ++ +PL ++ ++
Sbjct: 335 DVPIVGDVVRQVLEQILQVRVIDYPTNSTRTKAWLDRINRWREDYPLMVPSYPSI 390

>ref|ZP_03286153.1| acetolactate synthase, large subunit, biosynthetic type
[Dictyoglomus turgidum DSM 6724]
gb|EEA17352.1| acetolactate synthase, large subunit, biosynthetic type
[Dictyoglomus turgidum DSM 6724]
Length = 557

Score = 110 bits (274), Expect = 5e-23, Method: Compositional matrix adjust.

Identities = 52/108 (48%), Positives = 71/108 (65%), Gaps = 6/108 (5%)

Query: 2 LGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHIDIDSAEIGKNKTPH 61
+GMHG YANYA+ SDL++A GVRF DR TGK+E FA AKI+HIDID AEIGKN P+
Sbjct: 255 IGMHGAAYANYAINSDLIIGAIVRFSDRSTGKVETFAPNAKIIHIDIDPAEIGKNVNPY 314

Query: 62 VSVCGDVKLALQGMNKLLENRAEELKLDGFWRNELNVQKQKFPPLSFK 109
V + D K AL+ + +V+E + + W ++ K K+PL ++
Sbjct: 315 VPIVADAKRALERLVEVVEPKINAM-----WWEKIKEWKTKYPLRYR 356

>ref|NP_695500.1| acetolactate synthase 1 catalytic subunit [Bifidobacterium longum
NCC2705]
gb|AAN24136.1| IlvB [Bifidobacterium longum NCC2705]
Length = 655

Score = 109 bits (273), Expect = 5e-23, Method: Composition-based stats.
Identities = 58/137 (42%), Positives = 84/137 (61%), Gaps = 6/137 (4%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT+ A AV+ +DLL+A G RFDDRV TGKL+AFA A+++HIDID AEIGKN+ P
Sbjct: 282 MLGMHGTIAATGAVQRADLLVAIGARFDDRV TGKLD AFAPTARVIHIDIDPAEIGKNRQP 341

Query: 61 HVSVC GDVKLALQGM-NKVLNRAEELKLDGFWRNELNVQKQKFPPLSFK--TFGEAVP- 116
V + GDV L + ++ +A + K + W ++ +++P+++ T G P
Sbjct: 342 DVPIVGDVATVLDLIP EIQR TQAIQ GKPNLAPWWKAIDGWRE EYPMTWDEPTD GSLAPQ 401

Query: 117 --CPSKKT VHDPSYEWV 131
+ DPS WV
Sbjct: 402 WVIKKLSE MADPSTIWV 418

>ref|YP_909084.1| acetolactate synthase 1 catalytic subunit [Bifidobacterium
adolescentis ATCC 15703]
dbj|BAF39002.1| acetohydroxy-acid synthase [Bifidobacterium adolescentis ATCC
15703]
Length = 620

Score = 109 bits (273), Expect = 5e-23, Method: Compositional matrix adjust.
Identities = 60/137 (43%), Positives = 85/137 (62%), Gaps = 6/137 (4%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT+ A AV+ DLL+A G RFDDRV TGKL+AFA A+++HIDID AEIGKN+ P
Sbjct: 282 MLGMHGTIAATGAVQRCDLLVAIGARFDDRV TGKLD AFA PGARVIHIDIDPAEIGKNRQP 341

Query: 61 HVSVC GDVKLALQGM-NKVLNRAEELKLDGFWRNELNVQKQKFPPLSFK--TFGEAVPC 117
V + GDV L+ + ++ +A K D W + +N +++P+++ T G P
Sbjct: 342 DVPIVGDVATV LKDLIPEIKREQAVHGKPD TSHWWDVINKWRE EYPITWDEPTD GSLAPQ 401

Query: 118 PSKKT V H---DPSYEWV 131
K + DP+ WV
Sbjct: 402 WVVKQLSDMADPNTVWV 418

>ref|ZP_03129981.1| acetolactate synthase, large subunit, biosynthetic type
[Chthoniobacter flavus Ellin428]
gb|EDY19461.1| acetolactate synthase, large subunit, biosynthetic type
[Chthoniobacter flavus Ellin428]
Length = 591

Score = 109 bits (273), Expect = 5e-23, Method: Compositional matrix adjust.
Identities = 57/111 (51%), Positives = 72/111 (64%), Gaps = 4/111 (3%)

Query: 2 LGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHIDIDSAEIGKNKTPH 61
LGMHGT VYAN AV +DLLLA GVRFDDRV TGK+ FA IVHIDID++EI KNK
Sbjct: 254 LGMHGT VYANNAVNEADLLAIGVRFDDRV TGKISEFAKHGTIVHIDIDNSEINKNVVK 313

Query: 62 VSVCGDVKLALQGMNKLLEN----RAEELKLDGFWRNELNVQKQKFPPLSF 108
+ + DVK AL +N++LE R + F W ++N ++ FP +F
Sbjct: 314 LPILSDVKYALGR LNLQLEESGAKRVQSGFKAFP DWYVKINAWREAFPFTF 364

>ref|YP_001865343.1| acetolactate synthase, large subunit, biosynthetic type [Nostoc punctiforme PCC 73102]
gb|ACC80400.1| acetolactate synthase, large subunit, biosynthetic type [Nostoc punctiforme PCC 73102]
Length = 637

Score = 109 bits (273), Expect = 5e-23, Method: Compositional matrix adjust.
Identities = 54/107 (50%), Positives = 70/107 (65%), Gaps = 1/107 (0%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+ G RFDDRVTKGL+ FASRAK++HIDID AE+GKN+ P
Sbjct: 294 MLGMHGTAYANFAVSDCDLLICVGARFDDRVTKGLDEFASRAKVIHIDIDPAEVGKNRIP 353

Query: 61 HVSVC GDVKLALQG-MNKVLENRAEELKLD FGVWRNELNVQKQKFPL 106
V + GDV+ L + + E + W N +N + ++PL
Sbjct: 354 EVPIVGDVNRVLVDLLRRCKETGVKATPNQNQEWNLVNRWRDEYPL 400

>ref|YP_001379064.1| acetolactate synthase, large subunit, biosynthetic type [Anaeromyxobacter sp. Fw109-5]
gb|ABS26080.1| acetolactate synthase, large subunit, biosynthetic type [Anaeromyxobacter sp. Fw109-5]
Length = 570

Score = 109 bits (273), Expect = 6e-23, Method: Compositional matrix adjust.
Identities = 58/119 (48%), Positives = 75/119 (63%), Gaps = 1/119 (0%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
M GMHG+ YANYAV++SDL++A G RFDDRVTKGL FA AKIVH+DID AEI K T
Sbjct: 256 MFGMHGSR YANYAVQNSDLIIALGARFDDRVTKGLSTFAPEAKIVHLDIDPAEISKLVTA 315

Query: 61 HVSVC GDVKLAL-QGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVPCP 118
V + GD+KL L Q +V A K D W ++ ++K+PL ++ G+ P
Sbjct: 316 TVPLVGD LKLLLPQLTEEVRRAEAAHGKPDVSGWWKRIDDWRKKYPLKYEQEGQKYL LP 374

>ref|ZP_02850743.1| acetolactate synthase, large subunit, biosynthetic type [Paenibacillus sp. JDR-2]
gb|EDS49383.1| acetolactate synthase, large subunit, biosynthetic type [Paenibacillus sp. JDR-2]
Length = 581

Score = 109 bits (273), Expect = 7e-23, Method: Compositional matrix adjust.
Identities = 56/109 (51%), Positives = 71/109 (65%), Gaps = 5/109 (4%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
M GMHGT AN +++ +DLL+ G RFDDRVTKGL FA AKIVHIDID AEIGKN
Sbjct: 275 MPGMHGT YANKSIQSADLLINIGARFDDRVTKLAGFAPLAKIVHIDIDPAEIGKNVPT 334

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFK 109
+ + GDVK LQ +N+ L RA++ WR E+ K ++P S+K
Sbjct: 335 DIPIVGDVKT V LQQVNQ-LAKRADKAD----AWRQEIQESKAQYPPFSYK 378

>ref|NP_789495.1| acetolactate synthase 1 catalytic subunit [Tropheryma whipplei TW08/27]
emb|CAD67233.1| acetolactate synthase [Tropheryma whipplei TW08/27]
Length = 581

Score = 109 bits (272), Expect = 7e-23, Method: Compositional matrix adjust.
Identities = 56/108 (51%), Positives = 70/108 (64%), Gaps = 3/108 (2%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
M GMHGT V A A++ SDDL+A G RFDDRVTKG+L++FA A++VH+DID AEIGKN+ P
Sbjct: 258 MPGMHGT VAAAILALQKSDLLIALGTRFDDRVTKRLDSFAQNARVVHVDIDPAEIGKNRRP 317

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 108
+ V GDVK L +N ENR D W L +Q +PL F

Sbjct: 318 DLFVIGDVKEVLGYINSHSENRTTG---DLTAWWRSLEKMRQIYPLVF 362

>emb|CA090567.1| ilvG [Microcystis aeruginosa PCC 7806]
Length = 617

Score = 109 bits (272), Expect = 9e-23, Method: Compositional matrix adjust.
Identities = 57/122 (46%), Positives = 76/122 (62%), Gaps = 8/122 (6%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+ FAS+AK++H+DID AE+GKN+ P
Sbjct: 275 MLGMHGTAYANFAVTGCDLLIAVGARFDDRVTKGLDEFASKAKVIHVIDIDPAEVGKNRAP 334

Query: 61 HVSVC GDVKLAL-QGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVPCPS 119
V + GDV++ L Q ++K E W ++ +Q +PL VP P
Sbjct: 335 DVPIVGDVVRVLEQILHKAREFDYPTNCDRTQAWLAQIERWRQDYPLQ-----VPLPE 387

Query: 120 KK 121
+
Sbjct: 388 GR 389

>ref|YP_002049511.1| acetolactate synthase III large subunit [Paulinella chromatophora]
gb|ACB43301.1| acetolactate synthase III large subunit [Paulinella chromatophora]
Length = 590

Score = 108 bits (271), Expect = 1e-22, Method: Compositional matrix adjust.
Identities = 54/106 (50%), Positives = 71/106 (66%), Gaps = 2/106 (1%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+ FA RA+I+H++ID AE+GKN+ P
Sbjct: 250 MLGMHGTAYANFAVTECDLLIAAGARFDDRVTKGLDTFAPRARIHLEIDPAELGKNRPL 309

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPL 106
V+V GDV +L M V L+ + +W + K ++PL
Sbjct: 310 DVAVLGDVAESLNLN--VQAKGPTALETNTHLWLERIQEWKTRYPL 353

>ref|ZP_01628200.1| acetolactate synthase III large subunit [Nodularia spumigena
CCY9414]
gb|EAW47101.1| acetolactate synthase III large subunit [Nodularia spumigena
CCY9414]
Length = 632

Score = 108 bits (271), Expect = 1e-22, Method: Compositional matrix adjust.
Identities = 53/107 (49%), Positives = 70/107 (65%), Gaps = 1/107 (0%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+ G RFDDRVTKGL+ FAS AK++HIDID AE+GKN+ P
Sbjct: 289 MLGMHGTAYANFAVTDCDLLICVGARFDDRVTKGLDEFASHAKVIHIDIDPAEVGKNRVP 348

Query: 61 HVSVC GDVKLALQG-MNKVLENRAEELKLD FGVWRNELNVQKQKFPL 106
V + GDV+ L + + E + + W + +N +Q +PL
Sbjct: 349 EVPIVGDVQRVLDLLRRCQETGIKNIPHNQEWLHLINRWRQDYPL 395

>ref|YP_475476.1| acetolactate synthase 3 catalytic subunit [Synechococcus sp.
JA-3-3Ab]
gb|ABD00213.1| acetolactate synthase, large subunit, biosynthetic type
[Synechococcus sp. JA-3-3Ab]
Length = 598

Score = 108 bits (270), Expect = 1e-22, Method: Compositional matrix adjust.
Identities = 56/120 (46%), Positives = 73/120 (60%), Gaps = 2/120 (1%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+ FA AK++HIDID AE+GKN+TP
Sbjct: 256 MLGMHGTAYANFAVTECDLLIAVGARFDDRVTKGLDEFASHAKVIHIDIDPAEVGKNRTP 315

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVPCPSK 120

V + GDV+ L + L ++ W + K+ +PL + E V P +
Sbjct: 316 DVPIVGDVRTVLVELLAQL-SKMPHTPGQTQAWLERIERWKRDPYPLQVPRY-EGVIAPQQ 373

>ref|YP_476898.1| acetolactate synthase 3 catalytic subunit [Synechococcus sp.
JA-2-3B'a(2-13)]
gb|ABD01635.1| acetolactate synthase, large subunit, biosynthetic type
[Synechococcus sp. JA-2-3B'a(2-13)]
Length = 598

Score = 108 bits (270), Expect = 1e-22, Method: Compositional matrix adjust.
Identities = 56/120 (46%), Positives = 73/120 (60%), Gaps = 2/120 (1%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+ FA AK++HIDID AE+GKN+TP
Sbjct: 256 MLGMHGTAYANFAVTECDLLIAVGARFDDRVTKGLDEFKAKHAKVIHIDIDPAEVGKNRTP 315

Query: 61 HVSVC G DVK LALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAVPCPSK 120
V + GDV+ L + L ++ W + K+ +PL + E V P +
Sbjct: 316 DVPIVGDVRTVLVELLAQL-SKMPHTPGQTQAWLERIERWKRDPYPLQVPRY-EGVIAPQQ 373

>ref|YP_001243976.1| acetolactate synthase, large subunit, biosynthetic type [Thermotoga
petrophila RKU-1]
gb|ABQ46400.1| acetolactate synthase, large subunit [Thermotoga petrophila RKU-1]
Length = 581

Score = 108 bits (270), Expect = 1e-22, Method: Compositional matrix adjust.
Identities = 55/114 (48%), Positives = 72/114 (63%), Gaps = 6/114 (5%)

Query: 2 LGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTPH 61
+GMHGT Y NYAV ++DL++A G VRF DR+ G FA AKIVH+DID AEIGKN
Sbjct: 256 IGMHGTYYGNVAVANADLVIALGVRFSDRILGNPRTFAKNAKIVHVIDIDPAEIGKNVRVD 315

Query: 62 VSVCG DVK LALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAV 115
V + GD+K LQ E E++ DF W EL K+K+PL++K G+ +
Sbjct: 316 VPIVGLDLKSVLQ-----EFLKYEIETDFSDVVEELQEIKKKYPLTYKRDGKLI 363

>ref|NP_895067.1| acetolactate synthase 3 catalytic subunit [Prochlorococcus marinus
str. MIT 9313]
emb|CAE21414.1| acetolactate synthase [Prochlorococcus marinus str. MIT 9313]
Length = 573

Score = 108 bits (269), Expect = 2e-22, Method: Compositional matrix adjust.
Identities = 51/106 (48%), Positives = 71/106 (66%), Gaps = 2/106 (1%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+ FA RA+++H +ID AE+ KN+ P
Sbjct: 257 MLGMHGTAYANFAVTECDLLIAVGARFDDRVTKGLDTFAPRARVIHFEIDPAEVAKNRRP 316

Query: 61 HVSVC G DVK LALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPL 106
V+V GDV ++L + + + ++ EL+ W + K +PL
Sbjct: 317 DVAVLGDVGVS LKLFDSLKQKSVELRTS--AWLTRIESWKNLYPL 360

>ref|YP_001224368.1| acetolactate synthase 3 catalytic subunit [Synechococcus sp. WH
7803]
emb|CAK23071.1| Acetolactate synthase large subunit [Synechococcus sp. WH 7803]
Length = 617

Score = 108 bits (269), Expect = 2e-22, Method: Compositional matrix adjust.
Identities = 52/108 (48%), Positives = 70/108 (64%), Gaps = 2/108 (1%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+ FA RA+++H +ID AEIGKN+
Sbjct: 277 MLGMHGTAYANFAVTDCLLIAGARFDDRVTKGLDTFAPRARVIHFEIDPAEIGKNRRA 336

Query: 61 HVSVC G DVK LALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 108
V+V GDV +L + ++ E + W + KQ++PL+

Sbjct: 337 DVAVLGDVGASLALVELSLQDTPETRT--AAWLERIKDWKQRYPLTI 382

>ref|YP_001030046.1| acetolactate synthase, large subunit [Methanocorpusculum labreanum
Z]
gb|ABN06779.1| acetolactate synthase, large subunit [Methanocorpusculum labreanum
Z]
Length = 567

Score = 108 bits (269), Expect = 2e-22, Method: Compositional matrix adjust.
Identities = 57/102 (55%), Positives = 66/102 (64%), Gaps = 11/102 (10%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YANYAV +DLL+A G RFDDRVTKGL FA+ AK++HIDID AEIGKN P
Sbjct: 256 MLGMHGT EYANYAVSEADLLIAGGARFDDRVTKLSHFATHAKVIHIDIDPAEIGKNVNP 315

Query: 61 HVSVC GDVKLALQGMNKVLENRA-----EELKLD FGVWRN 95
V + GD K L M + EN E++KL WR
Sbjct: 316 DVPIVGD AKSVLADMICLAENNGCISEPWLEQVKL----WRT 353

>gb|ABD96281.1| acetolactate synthase [uncultured marine type-A Synechococcus GOM
3M9]
Length = 618

Score = 107 bits (268), Expect = 2e-22, Method: Composition-based stats.
Identities = 53/108 (49%), Positives = 74/108 (68%), Gaps = 4/108 (3%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+ FA RA++VH +ID AE+GKN+
Sbjct: 278 MLGMHGT AYANFAVTECDLLIAGGARFDDRVTKGLDTFAPRARVVHFEIDPAE V GKNNRQA 337

Query: 61 HVSVC GDVKLALQGMNKV-LENRAEELKLD FGVWRNELNVQKQKFPLS 107
V+V GD+ L+L + ++ L++ AE W + K+++PL+
Sbjct: 338 EVAVLGDLGLSLARLVELSLQHPAEPRTT---AWLARIASWKERYPLT 382

>ref|NP_787333.1| acetolactate synthase 1 catalytic subunit [Tropheryma whipplei str.
Twist]
gb|AA044302.1| acetolactate synthase large subunit [Tropheryma whipplei str.
Twist]
Length = 581

Score = 107 bits (268), Expect = 2e-22, Method: Compositional matrix adjust.
Identities = 55/108 (50%), Positives = 70/108 (64%), Gaps = 3/108 (2%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
M GMHGT V A A++ SDLL+A G RFDDRVTKG+L++FA A++VH+DID AEIGKN+ P
Sbjct: 258 MPGMHGT VAAIALQKSDLLIALGTRFDDRVTKRLDSFAQNARVVHVDIDPAEIGKNRRP 317

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 108
+ V GDVK L +N EN+ D W L +Q +PL F
Sbjct: 318 DLFVIGDVKEVLGYINSHSENKTTG---DLTAWWRSLEKMRQIYPLVF 362

>ref|ZP_01188450.1| Acetolactate synthase, large subunit, biosynthetic type
[Halothermothrix orenii H 168]
gb|EAR80052.1| Acetolactate synthase, large subunit, biosynthetic type
[Halothermothrix orenii H 168]
Length = 555

Score = 107 bits (268), Expect = 2e-22, Method: Compositional matrix adjust.
Identities = 53/109 (48%), Positives = 73/109 (66%), Gaps = 6/109 (5%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
M GMHGT YANYA+ +DLL+A GVRFDDRVTKG+E FA AKI+HIDID AEI KN
Sbjct: 254 MPGMHGT RYANYAISEADLLVAVGVRFDDRVTKGIEEFAPDAKIIHIDIDPAEICKNVEA 313

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFK 109
H+ + GD + L+ + ++E++ + G W ++ K+K PL ++

Sbjct: 314 HIPIVGDARNILKELLPLIEHK-----ERGAWLQQIKWEKKNPLKYE 356

>ref|ZP_02917198.1| hypothetical protein BIFDEN_00474 [Bifidobacterium dentium ATCC 27678]
gb|EDT44666.1| hypothetical protein BIFDEN_00474 [Bifidobacterium dentium ATCC 27678]
Length = 625

Score = 107 bits (268), Expect = 2e-22, Method: Compositional matrix adjust.
Identities = 59/137 (43%), Positives = 80/137 (58%), Gaps = 6/137 (4%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKGLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT+ A AV+ DLL+A G RFDDRVTKGL+AFA A+++HIDID AEIGKN+ P
Sbjct: 286 MLGMHGTIAATGAVQRCDLLVAIGARFDDRVTKGLDAFAPGARVIHIDIDPAEIGKNRQP 345

Query: 61 HVSVC GDVKLALQGM-NKVLNRAEELKLD FGVWRNELNVQKQKFLSFK--TFGEAVP- 116
V + GDV L+ + ++ +A K D W +N + +P+ ++ T G P
Sbjct: 346 DVPIVGDVATVLKDLIPEIKREQAMHGKPDISHWWEVINKWVEDYPIKWEQPTDGLAPQ 405

Query: 117 --CPSKKT VHDPSYEWV 131
+ DP WV
Sbjct: 406 WVKQLSVMADPDTIWV 422

>ref|ZP_00828201.1| COG0028: Thiamine pyrophosphate-requiring enzymes [acetolactate synthase, pyruvate dehydrogenase (cytochrome), glyoxylate carboligase, phosphonopyruvate decarboxylase] [Yersinia frederiksenii ATCC 33641]
Length = 548

Score = 107 bits (268), Expect = 2e-22, Method: Compositional matrix adjust.
Identities = 53/118 (44%), Positives = 76/118 (64%), Gaps = 9/118 (7%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKGLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT AN+AV+ DLL+A G RFDDRVTKGL FA++AK++H+DID AE+GK +
Sbjct: 247 MLGMHGTAAANFAVQDCDLLIAVGARFDDRVTKGLNTFAAKAKVIHMDIDPAELGKLRQA 306

Query: 61 HVSVC GDVKLALQGMNKVLNRAEELKLD FGVWRNELNVQKQKFLSFKTFGEAVPCP 118
HV++ GD+K+ L + + L+ WR+E+ KQ+ + G+A+ P
Sbjct: 307 HVALQGDLKVLPLALQQ-----SLNIQPWREVMTLKQQHSWRYDHPGQAIYAP 355

>ref|YP_001016784.1| acetolactate synthase 3 catalytic subunit [Prochlorococcus marinus str. MIT 9303]
gb|ABM77519.1| Acetolactate synthase large subunit [Prochlorococcus marinus str. MIT 9303]
Length = 593

Score = 107 bits (267), Expect = 3e-22, Method: Compositional matrix adjust.
Identities = 51/106 (48%), Positives = 70/106 (66%), Gaps = 2/106 (1%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKGLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+ FA RA+++H +ID AE+ KN+ P
Sbjct: 277 MLGMHGTAYANFAVTECDLLIAVGARFDDRVTKGLDTFAPRARVIHFEIDPAEVAKNRRP 336

Query: 61 HVSVC GDVKLALQGMNKVLNRAEELKLD FGVWRNELNVQKQKFL 106
V+V GDV ++L + + + + EL+ W + K +PL
Sbjct: 337 DVAVLGDVGSLVKLLDLSKQKPVLELTS--AWLTRIESWKDLYPL 380

>ref|YP_431088.1| acetolactate synthase, large subunit [Moorella thermoacetica ATCC 39073]
gb|ABC20545.1| acetolactate synthase, large subunit [Moorella thermoacetica ATCC 39073]
Length = 569

Score = 107 bits (267), Expect = 3e-22, Method: Compositional matrix adjust.
Identities = 53/112 (47%), Positives = 75/112 (66%), Gaps = 5/112 (4%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
M+GMHGT ANYA+ +DL++ GVRFDDRVTK+EAFA +AKI+HIDID+AEIGK
Sbjct: 264 MVGMHGTAANYAMCETDLIIGVGVRFDDRVTKGVEAFAPKAKIIHIDIDAAEIGKVQA 323

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSFKT FG 112
H+ + D + QG+ +LE + K D WR ++ +++ PL ++ G
Sbjct: 324 HIPIVSDAR---QGLAAILEKLSG--KGDHQA WRQQIRRWQEENPLRYEKS G 370

>ref|YP_001716542.1| acetolactate synthase, large subunit, biosynthetic type [Candidatus
Desulforudis audaxviator MP104C]
gb|ACA58910.1| acetolactate synthase, large subunit, biosynthetic type [Candidatus
Desulforudis audaxviator MP104C]
Length = 556

Score = 107 bits (266), Expect = 4e-22, Method: Compositional matrix adjust.
Identities = 54/112 (48%), Positives = 71/112 (63%), Gaps = 6/112 (5%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHG Y NYAV D L L L A G V R F D D R V T G K + + + F A S A K + + H I D I D A E G K N
Sbjct: 257 MLGMHGARYTNYAVNECDLLAIGVRFDDRVTKGVDSFASHAKVIHIDIDPAEFGKNVPV 316

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSFKT FG 112
+S+ GDV+ L + K L + W+ ++ K+++PL + G
Sbjct: 317 DISIAGDVRTVLVDLIKNLTCNPDP-----AWQEKIARWKKEYPLDYSREG 362

>ref|ZP_01389342.1| acetolactate synthase, large subunit, biosynthetic type [Geobacter
sp. FRC-32]
gb|EAT61366.1| acetolactate synthase, large subunit, biosynthetic type [Geobacter
sp. FRC-32]
Length = 566

Score = 107 bits (266), Expect = 4e-22, Method: Compositional matrix adjust.
Identities = 53/118 (44%), Positives = 73/118 (61%), Gaps = 3/118 (2%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
+LGMHGT YAN AV H DLL+A G RFDDRVTK+ +FA A+I+HID+D I KN
Sbjct: 254 LLGMHGTYANMAVSHCDLLIAGVRFDDRVTKIASFAPHAQIIHIDVDPTSIKKNVRV 313

Query: 61 HVSVC GDVKLALQGMNKVLENRAEE---LKLD FGVWRNELNVQKQKFP LSFKT FG EAV 115
+ + GDVK L+ + KVL +AE+ + W E+ K K P+++K+ +
Sbjct: 314 DLPVIGDVKDVLKQLLKVLAGQA EKAATFQKQIVPWSQEIEAWKAKHPMTYKSSASVI 371

>sp|Q59498|ILVB_MYCAV Acetolactate synthase (Acetohydroxy-acid synthase) (ALS)
gb|AAB38426.1| acetolactate synthase
Length = 621

Score = 107 bits (266), Expect = 4e-22, Method: Compositional matrix adjust.
Identities = 52/108 (48%), Positives = 69/108 (63%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
M GMHGT V A A++ SDLL+A G RFDDRVTKL+ FA AK++H DID AEIGKN+
Sbjct: 292 MPGMHGTVA AVALQRS D L L I A L G T R F D D R V T G K L D T F A P E A K V I H A D I D P A E I G K N R H A 351

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 108
V + GDVK + + ++L + LD W L+ + +PLS+
Sbjct: 352 DVPIVGDVKAVIAELVEILRHDGAPGNLDIADWWAYLDDVQSTYPLSY 399

>ref|YP_001004540.1| acetolactate synthase 2 catalytic subunit [Yersinia enterocolitica
subsp. enterocolitica 8081]
emb|CAL10288.1| acetolactate synthase isozyme II large subunit [Yersinia
enterocolitica subsp. enterocolitica 8081]
Length = 548

Score = 106 bits (265), Expect = 5e-22, Method: Compositional matrix adjust.
Identities = 53/118 (44%), Positives = 76/118 (64%), Gaps = 9/118 (7%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT AN+AV+ DLL+A G RFDDRVTKGL FA++AK++H+DID AE+GK +
Sbjct: 247 MLGMHGT KAAFAVQDCDLLIAVGARFDDRVTKGLNTFAAKAKVIHMDIDPAELGKLRQA 306

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLDGFWRNELNVQKQKFPLSFKTGFEAVPCP 118
HV++ GD+K+ L + + L+ WR+E+ KQ+ + G+A+ P
Sbjct: 307 HVALQGDLKVLPLPALQQ-----PLNIQPWRDEVIALKQQHSWRYDHPGQAIYAP 355

>ref|YP_842552.1| acetolactate synthase 3 catalytic subunit [Methanosaeta thermophila PT]
gb|ABK13912.1| acetolactate synthase, large subunit [Methanosaeta thermophila PT]
Length = 557

Score = 106 bits (265), Expect = 5e-22, Method: Compositional matrix adjust.
Identities = 52/108 (48%), Positives = 74/108 (68%), Gaps = 6/108 (5%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN AV+ +D+LLA G RFDDRVTKG+ +FA +A+I+HID+D+AEIGKN
Sbjct: 254 MLGMHGT K YANLAVQEADVLLAVGARFDDRVTKIASFAPKARIHIDVDAAEIGKNVRV 313

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLDGFWRNELNVQKQKFPLSF 108
+ V GD K+AL + K ++ R + W ++ K+++PL +
Sbjct: 314 DIPVVGDAKIALTELLKHVQQR-----WTEWNEKILSWKREYPLRY 355

>ref|YP_381046.1| acetolactate synthase 3 catalytic subunit [Synechococcus sp. CC9605]
gb|ABB34491.1| acetolactate synthase, large subunit, biosynthetic type [Synechococcus sp. CC9605]
Length = 574

Score = 106 bits (265), Expect = 5e-22, Method: Composition-based stats.
Identities = 50/107 (46%), Positives = 73/107 (68%), Gaps = 2/107 (1%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+ FA RA+++H +ID AEIGK + P
Sbjct: 257 MLGMHGT AYANFAVTECDLLIAVGARFDDRVTKGLDTFAPRAQVIHFEIDPAEIGKTRRP 316

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLDGFWRNELNVQKQKFPLSF 107
V+V GD+ L++ + ++ + ++++ W + KQ +PL+
Sbjct: 317 DVAVLGDLGLSVARLVEL--SLQQVQVPRTS AWLARIAEWKQTYPLT 361

>ref|ZP_03185165.1| acetolactate synthase 1 catalytic subunit [Streptomyces clavuligerus ATCC 27064]
ref|YP_002190658.1| acetohydroxy acid synthase large subunit [Streptomyces clavuligerus ATCC 27064]
gb|EDY52149.1| acetohydroxy acid synthase large subunit [Streptomyces clavuligerus ATCC 27064]
Length = 618

Score = 106 bits (265), Expect = 5e-22, Method: Composition-based stats.
Identities = 49/110 (44%), Positives = 75/110 (68%), Gaps = 1/110 (0%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
M GMHG+V A A++ +DL++A G RFDDRVTKGL++FA AKIVH DID AEIGKN+
Sbjct: 280 MPGMHGSVTAVTALQKADLIVALGARFDDRVTKGLDSFAPHAKIVHADIDPAEIGKNRAA 339

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLDGFWRNELNVQKQKFPLSF 109
V + GD + + + + ++ E ++ D+G+W +LN ++ +PL ++
Sbjct: 340 DVPIVGDA REVIADLVQAVQAEHSEGRIGDYGLWWTDLNRWRETYPLGYE 389

>ref|YP_001232446.1| acetolactate synthase, large subunit, biosynthetic type [Geobacter uraniireducens Rf4]
gb|ABQ27873.1| acetolactate synthase, large subunit [Geobacter uraniireducens Rf4]
Length = 566

Score = 106 bits (265), Expect = 6e-22, Method: Compositional matrix adjust.

Identities = 56/112 (50%), Positives = 71/112 (63%), Gaps = 3/112 (2%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
+LGMHGT YAN AV + DLL+A G RFDDRVTKG+ +FA AKI+HID+D I KN
Sbjct: 254 LLGMHGTYYANMAVSNCDLLVAVGARFDDRVTKGIASFAPHAKIIHIDVDPTSIKKNVRV 313

Query: 61 HVSVC GDVKLALQGMNKVLENRAEEL---KLD FGVWRNELNVQKQKFPLSFK 109
+ + GDVK L M KVLE + ++L K W E+ K K P+S+K
Sbjct: 314 DLPVIGDVKDVL TQMLKVLEE QGDKLDDFKKSIVPWTEEIEGWKAKHPMSYK 365

>ref|ZP_03275274.1| acetolactate synthase, large subunit, biosynthetic type
[Arthrospira maxima CS-328]
gb|EDZ93129.1| acetolactate synthase, large subunit, biosynthetic type
[Arthrospira maxima CS-328]
Length = 577

Score = 106 bits (264), Expect = 6e-22, Method: Compositional matrix adjust.
Identities = 53/108 (49%), Positives = 70/108 (64%), Gaps = 1/108 (0%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A GVRFDDRV G + FA AK++HIDID AE+GKN+T
Sbjct: 267 MLGMHGTAYANFAVSECDLLIAGVRFDDRVAGTGDQFAHSAKVIHIDIDPAEVGKNRTT 326

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKD-FGVWRNELNVQKQKFPLS 107
V + GDV+ L M + E+K + W + LN + ++PL+
Sbjct: 327 DVPVIGDVRQVLGDM LQRSYQLGTEIKPEQTQEWHSHLNQLRNQYPLT 374

>ref|YP_002121866.1| acetolactate synthase, large subunit, biosynthetic type
[Hydrogenobaculum sp. Y04AAS1]
gb|ACG57888.1| acetolactate synthase, large subunit, biosynthetic type
[Hydrogenobaculum sp. Y04AAS1]
Length = 584

Score = 106 bits (264), Expect = 6e-22, Method: Compositional matrix adjust.
Identities = 57/125 (45%), Positives = 79/125 (63%), Gaps = 10/125 (8%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN AV HSDLL+A G RFDDRVTKG++ FA AKI+HIDID A I KN T
Sbjct: 259 MLGMHGTYYANMAVYHSDLLIAGVSRFDDRVTKGVDEFAPFAKIIHIDIDPASISKNTIV 318

Query: 61 HVSVC GDVKLALQGMNKVLENRAEEL--KLD FGVWRNELNVQKQKFPLSFKTFGEAVPCP 118
V + GD++ L + + ++ + ++ + + W +++ K+ +PLS+K P
Sbjct: 319 DVPVIGDIRSVLVR LIEEIKGKNSKIVHEEERKTWLEQIDAWKKAYPLSYK-----P 370

Query: 119 SKKTV 123
S K V
Sbjct: 371 SDKVV 375

>ref|NP_246567.1| acetolactate synthase 2 catalytic subunit [Pasteurella multocida
subsp. multocida str. Pm70]
gb|AAK03712.1| IlvG [Pasteurella multocida subsp. multocida str. Pm70]
Length = 549

Score = 106 bits (264), Expect = 7e-22, Method: Compositional matrix adjust.
Identities = 54/110 (49%), Positives = 69/110 (62%), Gaps = 9/110 (8%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTP 60
M+GMHGT ANYAV+ D LLLAFG RFDDRVTKL+ FA AK++H DID AEIGK + P
Sbjct: 247 MIGMHGTTAANYAVQEC D LLLAFGARFDDRVTKGLDTFAPHAKVIHADIDMAEIGKLRRP 306

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKT 110
V++CGD+ +Q N A + L+ W+ + KQ + T
Sbjct: 307 DVALCGDI---VQAFN-----ALSIPLNIAEWQAHIKQLKQTHDFCYST 347

>ref|YP_001738428.1| acetolactate synthase, large subunit, biosynthetic type [Thermotoga
sp. RQ2]

gb|ACB08745.1| acetolactate synthase, large subunit, biosynthetic type [Thermotoga
sp. RQ2]
Length = 581

Score = 106 bits (264), Expect = 7e-22, Method: Compositional matrix adjust.
Identities = 53/114 (46%), Positives = 72/114 (63%), Gaps = 6/114 (5%)

Query: 2 LGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHIDIDSAEIGKNKTPH 61
+GMHGT Y NYAV ++DL++A G VRF DR+ G FA A+IVH+DID AEIGKN
Sbjct: 256 IGMHGTYYGN YAVANADLI ALGVRFSDRILGNPRTFAKNARIVHVDIDPAEIGKNVRVD 315

Query: 62 VSVCGDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAV 115
V + GD+K L+ E E++ DF W EL K+K+PL++K G+ +
Sbjct: 316 VPIVGDLKSVLE-----EFLKYEIETDFSDWIEELQEIKKKYPLTYKRDGKLI 363

>ref|NP_228358.1| acetolactate synthase, large subunit [Thermotoga maritima MSB8]
gb|AAD35633.1|AE001730.1 acetolactate synthase, large subunit [Thermotoga maritima MSB8]
Length = 584

Score = 106 bits (264), Expect = 7e-22, Method: Compositional matrix adjust.
Identities = 53/114 (46%), Positives = 72/114 (63%), Gaps = 6/114 (5%)

Query: 2 LGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHIDIDSAEIGKNKTPH 61
+GMHGT Y NYAV ++DL++A G VRF DR+ G FA A+IVH+DID AEIGKN
Sbjct: 259 IGMHGTYYGN YAVANADLI ALGVRFSDRILGNPRTFAKNARIVHVDIDPAEIGKNVRVD 318

Query: 62 VSVCGDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSFKTFGEAV 115
V + GD+K L+ E E++ DF W EL K+K+PL++K G+ +
Sbjct: 319 VPIVGDLKSVLE-----EFLKYEIETDFSDWIEELQEIKKKYPLTYKRDGKLI 366

>ref|YP_001014410.1| acetolactate synthase 3 catalytic subunit [Prochlorococcus marinus
str. NATL1A]
gb|ABM75145.1| Acetolactate synthase large subunit [Prochlorococcus marinus str.
NATL1A]
Length = 587

Score = 106 bits (264), Expect = 7e-22, Method: Compositional matrix adjust.
Identities = 53/106 (50%), Positives = 68/106 (64%), Gaps = 2/106 (1%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRV TGKL+ FA +AK++H +ID AEI KN+
Sbjct: 277 MLGMHGTAYANFAVTECDLLIAIGARFDDRV TGKLDTFAPKAKVIHFEIDPAEINKNRVV 336

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPL 106
VSV GDV ++L + + R + W N++ K FPL
Sbjct: 337 EVSVLGDVGISLVKLLDLSNQKTNPRTS--QWLNKIKNWKNFPL 380

>ref|YP_684797.1| acetolactate synthase, large subunit [uncultured methanogenic
archaeon RC-I]
emb|CAJ35471.1| acetolactate synthase, large subunit [uncultured methanogenic
archaeon RC-I]
Length = 566

Score = 105 bits (263), Expect = 8e-22, Method: Compositional matrix adjust.
Identities = 56/109 (51%), Positives = 73/109 (66%), Gaps = 6/109 (5%)

Query: 1 MLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YANY++ DLL+A G RF DRV TGKLE+FA AKIVHIDID AE+GKN
Sbjct: 259 MLGMHGTTRYANYSIIDCDLLIAIGARFSDRV TGKLESFAPNAKIVHIDIDPAEMGKNVKV 318

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGV--WRNELNVQKQKFPLSF 108
V + GD K+ L+ + L +LKL W ++++ K+++PL F
Sbjct: 319 DVPIVGDAKVILKKIISTL-----KLKLPKSAEWLKKIDLWKREYPLDF 362

>ref|YP_293049.1| acetolactate synthase 3 catalytic subunit [Prochlorococcus marinus
str. NATL2A]



The Chemical Company
Plant Science L. L. C.

BASF Reg. Doc. No. 2008/7019371

gb|AAZ59346.1| acetolactate synthase, large subunit [Prochlorococcus marinus str.
NATL2A]
Length = 582

Score = 105 bits (263), Expect = 8e-22, Method: Compositional matrix adjust.
Identities = 53/106 (50%), Positives = 68/106 (64%), Gaps = 2/106 (1%)

Query: 1 MLGMHGT VYANYAVEHS DLLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
MLGMHGT YAN+AV DLL+A G RFDDRVTGKL+ FA +AK++H +ID AEI KN+
Sbjct: 277 MLGMHGTAYANFAVTECDLLIAIGARFDDRVTGKLDTFAPKAKVIHFEIDPAEINKNRVV 336

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPL 106
VSV GDV ++L + + R + W N++ K FPL
Sbjct: 337 EVSVLGDVGISLVKLLDLSNQKTNPRTS--QWLNKIKNWKNFPL 380

>ref|NP_961972.1| acetolactate synthase 1 catalytic subunit [Mycobacterium avium
subsp. paratuberculosis K-10]
gb|AAS05586.1| IlvB_1 [Mycobacterium avium subsp. paratuberculosis K-10]
Length = 622

Score = 105 bits (263), Expect = 8e-22, Method: Compositional matrix adjust.
Identities = 52/108 (48%), Positives = 69/108 (63%)

Query: 1 MLGMHGT VYANYAVEHS DLLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNKTP 60
M GMHGT V A A++ SDLL+A G RFDDRVTGKL+ FA AK++H DID AEIGKN+
Sbjct: 293 MPGMHGTVA AVALQRS DLLIALGTRFDDRVTGKLDTFAPKAKVIHADIDPAEIGKNRHA 352

Query: 61 HVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 108
V + GDVK + + ++L + LD W L+ + +PLS+
Sbjct: 353 DVPIVGDVKAVIAELVEMLRHDGAPGNLDIADWWAYLDDVQSTYPLSY 400

Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF excluding environmental samples
from WGS projects
Posted date: Oct 26, 2008 5:54 PM
Number of letters in database: 2,468,333,968
Number of sequences in database: 7,155,275

Lambda	K	H
0.321	0.138	0.420

Gapped Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 7155275
Number of Hits to DB: 808,283,659
Number of extensions: 31473911
Number of successful extensions: 70461
Number of sequences better than 1.0: 1926
Number of HSP's gapped: 70594
Number of HSP's successfully gapped: 1926
Length of query: 166
Length of database: 2,468,333,968
Length adjustment: 124
Effective length of query: 42
Effective length of database: 1,581,079,868
Effective search space: 66405354456
Effective search space used: 66405354456
Neighboring words threshold: 11
Window for multiple hits: 40
X1: 16 (7.4 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (21.8 bits)
S2: 82 (36.2 bits)