



**Event SYHT0H2 Soybean: Allergenicity and Toxicity Assessment of  
Start to Stop T-DNA ORFs with a Minimum Size of 30 Amino Acids**

**Final Report**

**DATA REQUIREMENT(S):** Not Applicable

**AUTHOR(S):**



**STUDY COMPLETION DATE:** June 12, 2012

**PERFORMING LABORATORY:** Syngenta Crop Protection, LLC  
3054 East Cornwallis Road  
Research Triangle Park, NC 27709-2257 USA

**LABORATORY PROJECT ID:** Report Number: SSB-165-12  
Task Number: TK0066954

**SUBMITTER:**  
Syngenta Seeds, Inc.  
3054 East Cornwallis Road  
Post Office Box 12257  
Research Triangle Park, NC 27709-2257 USA

**SPONSOR:**  
Syngenta Crop Protection, LLC  
410 Swing Road  
Post Office Box 18300  
Greensboro, NC 27419-8300 USA

## STATEMENT OF DATA CONFIDENTIALITY CLAIMS

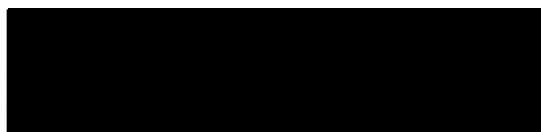
*The following statement applies to submissions to the United States Environmental Protection Agency (US EPA).*

### **No Claim of Confidentiality**

No claim of confidentiality is made for any information contained in this report on the basis of its falling within the scope of Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) Section 10 (d) (1) (A), (B), or (C).

**Company:** *Syngenta Seeds, Inc.*

**Company Representative:**



June 12, 2012

Date

*Manager, Regulatory Affairs*

These data are the property of Syngenta Seeds, Inc. and, as such, are considered to be confidential for all purposes other than compliance with the regulations implementing FIFRA Section 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 provision of common law or statute or in any other country.

*The following statement applies to submissions to regulatory agencies and other competent authorities other than the US EPA and all other viewers.*

### **This Document Contains Confidential Business Information**

This document contains information that is proprietary to Syngenta and, as such, is considered to be confidential for all purposes other than compliance with the relevant registration procedures.

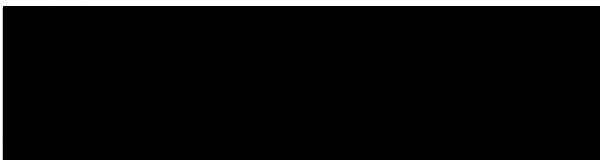
Without the prior written consent of Syngenta, this information may (i) not be used by any third party including, but not limited to, any other regulatory authority for the support of regulatory approval of this product or any other product, and (ii) not be published or disclosed to any third party including, but not limited to, any authority for the support of regulatory approval of any products.

Its submission does not constitute a waiver of any right to confidentiality that may exist in any other country.

## GOOD LABORATORY PRACTICE COMPLIANCE STATEMENT

This is not a study as defined by 40 CFR Part 160.3 and is therefore not subject to Federal Insecticide, Fungicide, and Rodenticide Act Good Laboratory Practice Standards (GLPS). However, all components of this analysis were performed according to accepted scientific practices, and relevant records have been retained.

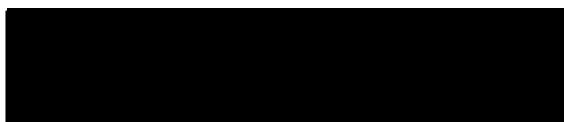
### Study Director:



*Technical Expert, Product Safety*  
Syngenta Crop Protection, LLC

June 12, 2012  
Date

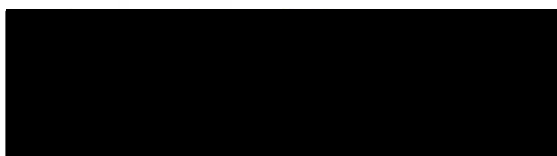
### Submitted by:



*Manager, Regulatory Affairs*  
Syngenta Seeds, Inc.  
3054 East Cornwallis Road  
PO Box 12257  
Research Triangle Park, NC 27709-2257 USA

June 12, 2012  
Date

### Sponsor:



*Technical Leader, Product Safety*  
Syngenta Crop Protection, LLC  
410 Swing Road  
Post Office Box 18300  
Greensboro, NC 27419-8300 USA

6/12/12  
Date

## GENERAL INFORMATION

### Contributors

The following contributed to this report in the capacities indicated:

Name	Title
[REDACTED]	Study Director, Syngenta Crop Protection, LLC
[REDACTED]	Technical Expert, Molecular Characterization

### Records Retention

Raw data, the original copy of this report, and other relevant records are archived at Syngenta, 3054 East Cornwallis Road, Research Triangle Park, NC 27709-2257, USA

## TABLE OF CONTENTS

<b>STATEMENT OF DATA CONFIDENTIALITY CLAIMS</b>	<b>2</b>
<b>GOOD LABORATORY PRACTICE COMPLIANCE STATEMENT</b>	<b>3</b>
<b>GENERAL INFORMATION</b>	<b>4</b>
<b>TABLE OF CONTENTS</b>	<b>5</b>
<b>LIST OF TABLES</b>	<b>6</b>
<b>LIST OF FIGURES</b>	<b>6</b>
<b>LIST OF ACRONYMS AND ABBREVIATIONS</b>	<b>7</b>
<b>1.0 EXECUTIVE SUMMARY</b>	<b>8</b>
<b>2.0 INTRODUCTION</b>	<b>9</b>
<b>3.0 MATERIALS AND METHODS</b>	<b>9</b>
3.1 Sequence Determination .....	9
3.2 Allergen Similarity Search.....	9
3.3 Toxin Similarity Search .....	10
3.4 Assessment of Alignment Significance .....	11
3.4.1 Allergen Similarity .....	11
3.4.2 Toxin Similarity .....	11
<b>4.0 RESULTS AND DISCUSSION</b>	<b>11</b>
4.1 Allergen Similarity Search.....	12
4.2 Toxin Similarity Search .....	12
<b>5.0 CONCLUSIONS</b>	<b>12</b>
<b>6.0 REFERENCES</b>	<b>19</b>
<b>APPENDICES SECTION</b>	<b>20</b>
APPENDIX A Protein sequences included in the FARRP AllergenOnline database, version 12 .....	21
APPENDIX B FASTA Allergen Search Alignments .....	109
APPENDIX C BLAST Toxin Search Alignments .....	326

## LIST OF TABLES

TABLE 1	Event SYHT0H2 T-DNA ORF sequences.....	13
---------	--	----

## LIST OF FIGURES

FIGURE 1	Location of the putative SYHT0H2 ORF sequences. ....	17
FIGURE 2	Location of the SYHT0H2 ORF sequences containing <i>avhppd-03</i> and <i>pat</i> .....	18

## LIST OF ACRONYMS AND ABBREVIATIONS

aa	amino acid
<i>avhppd-03</i>	<i>p</i> -hydroxyphenylpyruvate dioxygenase gene derived from oat
AvHPPD-03	<i>p</i> -hydroxyphenylpyruvate dioxygenase enzyme derived from oat
BLOSUM50	Blocks Substitution Matrix50
FARRP	Food Allergy Research and Resource Program
FIFRA	Federal Insecticide, Fungicide, and Rodenticide Act
GLPS	Good Laboratory Practice Standards
NCBI	National Center for Biotechnology Information
PAT	phosphinothricin acetyltransferase protein
T-DNA	Transfer deoxyribonucleic acid
US EPA	United States Environmental Protection Agency
®	registered trademark

## 1.0 EXECUTIVE SUMMARY

Soybean (*Glycine max* [L.] Merrill) has been genetically modified to express the novel gene *avhppd-03* derived from oat (*Avena sativa* L.). The gene *avhppd-03* encodes a *p*-hydroxyphenylpyruvate dioxygenase (HPPD) enzyme, designated AvHPPD-03, that catalyzes the formation of homogentisic acid, the aromatic precursor in plastoquinone and vitamin E biosynthesis. In comparison with the native soybean HPPD, the AvHPPD-03 isozyme from oat has lower binding affinity for mesotrione, an herbicide that inhibits HPPD. Expression of *avhppd-03* in transgenic Event SYHT0H2 soybean plants confers a mesotrione-tolerance phenotype.

Bioinformatic analysis of the deoxyribonucleic acid sequence within the Event SYHT0H2 T-DNA identified 52 sequences that are contained between known or putative start (ATG) and stop (TAG, TAA, or TGA) codons and have a minimum putative translation size of 30 amino acids (i.e., putative open reading frames ORFs). Translations include 5 DNA sequences encoding the AvHPPD and the PAT proteins in SYHT0H2. Each ORF DNA sequence was translated into its putative amino acid sequence and then systematically compared to the protein sequences of known or putative allergens or toxins in updated versions of 1) the Food Allergy Research and Resource Program (FARRP) AllergenOnline database, version 12 (FARRP 2012) and 2) a toxin database created from the National Center for Biotechnology Information (NCBI) Entrez® Protein database (NCBI 2011), respectively.

The allergen comparison consisted of two alignment searches. A full-length FASTA search evaluated alignments for each of the sequences with minimum criteria of at least 80 amino acids of alignment length with greater than 35% shared amino acid identity over the alignment length. An eight amino acid alignment match search was also performed to identify eight or more contiguous amino acids that may be shared with known or putative allergen sequences. To determine whether or not the sequences showed amino acid sequence similarity to toxins, a BLASTP search against the toxin database with an *E*-value of  $1 \times 10^{-5}$  as an initial threshold was used to identify potentially relevant alignments.

Results of the allergen comparisons revealed no significant sequence similarity between putative translated ORF sequences and any entry in the FARRP allergen database. Moreover, no matches were observed between any sequence of eight or more contiguous amino acids of any of the putative translated sequences and any entry in the FARRP AllergenOnline database (2012). No significant sequence similarity was observed between any of the SYHT0H2 translated T-DNA ORF sequences and any entry in the toxin database. Of the 47 putative, unintended ORFs that were identified, the potential for expression could not be ruled out bioinformatically for one. However, this ORF sequence does not show relevant biological similarity to known or putative allergens and toxins. In summary, this study supports the conclusion that ORF sequences derived from the T-DNA of SYHT0H2 soybean show no relevant biological similarity to known or putative allergens or toxins.



## 2.0 INTRODUCTION

Soybean (*Glycine max* [L.] Merrill) has been genetically modified to express the novel gene *avhppd-03* derived from oat (*Avena sativa* L.). The gene *avhppd-03* encodes a *p*-hydroxyphenylpyruvate dioxygenase (HPPD) enzyme, designated AvHPPD-03, that catalyzes the formation of homogentisic acid, the aromatic precursor in plastoquinone and vitamin E biosynthesis. In comparison with the native soybean HPPD, the AvHPPD-03 isozyme from oat has lower binding affinity for mesotrione, an herbicide that inhibits HPPD. Expression of *avhppd-03* in transgenic Event SYHT0H2 soybean plants confers a mesotrione-tolerance phenotype.

Bioinformatic analysis was used to determine deoxyribonucleic acid (DNA) sequences within the SYHT0H2 T-DNA that are contained between putative start (ATG) and stop (TAG, TAA, or TGA) codons and have a minimum putative translation size of 30 amino acids. Sequences were identified on all six reading frames of the SYHT0H2 T-DNA. Each DNA sequence was translated into its putative amino acid sequence and then systematically compared to the protein sequences of known or putative allergens or toxins in updated versions of 1) the Food Allergy Research and Resource Program (FARRP) AllergenOnline database, version 12 (FARRP 2012), located online at [www.allergenonline.org](http://www.allergenonline.org) as a source of known or putative allergens, and 2) a toxin database created from the National Center for Biotechnology Information (NCBI) Entrez® Protein database (NCBI 2012).

The purpose of these analyses was to determine whether the amino acid sequences of these *in silico* translations might have biologically relevant amino acid sequence similarity to known or putative protein allergens and toxins should they be expressed *in planta*.

## 3.0 MATERIALS AND METHODS

### 3.1 Sequence Determination

The Vector NTI Advance™ program, version 10.3.0, was used to analyze the DNA sequences within the SYHT0H2 T-DNA. DNA sequences that are contained between start (ATG) and stop codons (TAG, TAA, or TGA) and have a minimum putative translation size of 30 amino acids were determined and translated *in silico* into amino acid sequences. These translated sequences were compared using bioinformatics algorithms to protein sequences contained in either an allergen database or a database containing toxins to identify pair-wise alignments with potentially significant similarity to known or putative protein allergens and toxins.

### 3.2 Allergen Similarity Search

The FARRP AllergenOnline database (2012) contains the amino acid sequences of known and putative protein allergens. It is a curated, peer-reviewed database containing proteins identified as food allergens, respiratory allergens, allergenic venom proteins, contact allergens, gliadins, and glutenins. Entries were compiled primarily from searches of publicly available protein databases using the National Center for Biotechnology Information (NCBI) Entrez® search and retrieval system. The NCBI dataset was screened by searches for entries

associated with allergy or celiac disease; duplicate entries were removed, and additional entries were identified from publications. The list of candidate entries was then reviewed by an international panel of allergy experts who reviewed published clinical and laboratory evidence to support the candidate sequences as allergens. Proteins are classified as known or putative allergens according to predetermined criteria set by the FARRP expert review panel. The latest version of the FARRP AllergenOnline database (2012) contains 1,603 nonredundant entries (Appendix A).

Similarity searches were performed using an exact copy of the entire list of sequences in the current version of the FARRP AllergenOnline database (2012) (maintained at Syngenta Crop Protection, LLC, Research Triangle Park, NC, USA). Two sequence alignment searches were performed to compare the translated sequences with sequences in the FARRP AllergenOnline database (2012); 1) a FASTA search and 2) a separate, eight amino acid match search.

In the first search, the FASTA search algorithm, version 3.45 (Pearson and Lipman 1988), was used to assess overall sequence similarity by conducting a search of the each entire amino acid sequence with the sequences in the FARRP AllergenOnline database (2012). The default FASTA settings used include an extension penalty of two and gap creation penalty of 12. The scoring matrix for FASTA was the Blocks Substitution Matrix50 (BLOSUM50). The BLOSUM50 matrix is weighted to favor identical amino acids likely to impact protein structure.

In the second search, the each amino acid sequence was screened for matches of eight or more contiguous amino acids (Hileman *et al.* 2002) using a program developed by Syngenta; this program compared every possible peptide of eight contiguous amino acids in the query sequence with the sequences in the FARRP AllergenOnline database (2012).

### **3.3 Toxin Similarity Search**

The BLASTP software program (Altschul *et al.* 1997) was used to search a putative toxin database that was created as a subset of the National Center for Biotechnology Information Entrez® Protein database (NCBI 2011) on February 22, 2012. This putative toxin database was created with the search and retrieval system from the NCBI Entrez® Protein database using the keyword “toxin” and other toxin-related keywords (e.g., “ricin”) to broadly include toxin-associated proteins, while specific keywords (e.g., “toxin resistant” or “toxin receptor”) were also used to exclude proteins that are not toxins. This database was further refined by removal of *Bacillus thuringiensis* Cry protein entries, as these have been demonstrated to be non-toxic to mammals, including humans. This putative toxin database contains 23,302 non-redundant sequences, and while it has been refined to reduce the number of irrelevant entries (*i.e.*, proteins that are not actually toxins, but associated with that term such as toxin receptors or proteins that confer resistance to toxins, as well as Cry, Cry-like or related proteins), this putative toxin database has not been reviewed for clinical and laboratory evidence of toxicity and likely includes many non-toxic proteins. It is therefore considered a conservative and broadly inclusive database against which to conduct this bioinformatic analysis.

The BLASTP is a program used to search a database of protein sequences for identifying sequence similarities between a query sequence and a sequence from a database. The BLASTP algorithm is optimized to identify localized domains or other shared sequence similarities present within the full length query sequence; as a result, this approach detects more similarities than would a search started by aligning two sequences over their entire length from end-to-end. The BLASTP program, version 2.2.19, was used to perform the similarity searches with the following default parameters:

Low complexity filter

$E$ -value = 10

Word size = 3

Gap costs: existence = 11 and extension = 1

Similarity matrix: Blocks Substitution Matrix62 (BLOSUM62)

### **3.4 Assessment of Alignment Significance**

#### **3.4.1 Allergen Similarity**

The FASTA search produces localized alignments between each amino acid sequence and the sequences in the allergen database. The evaluation of each resulting alignment utilizes the minimum criterion of at least 80 amino acids of alignment length with greater than 35% shared amino acid identity over the alignment length. Any alignments exceeding these criteria for shared sequence similarity indicate the potential for immunologically relevant sequence similarity (Codex Alimentarius Commission 2009).

Additionally, any match of eight (or more) identical contiguous amino acids between each amino acid sequence, and any sequence in the allergen database, could indicate the potential for immunologically relevant sequence similarity. However, it should be noted that this specific search has been challenged due to high level of false positives and therefore results should be carefully considered for irrelevant alignments that may have no biological relevance for allergy risk (Goodman *et al.* 2008, Silvanovich *et al.* 2006).

#### **3.4.2 Toxin Similarity**

The expectation value ( $E$ -value) is a measure of the probability that matches between sequences occurred by chance. Search results involving comparisons between proteins with highly similar sequences yield  $E$ -values approaching zero, while the probability that sequence similarities occurred by chance increases with higher  $E$ -values (Ponting 2001). An  $E$ -value of  $1 \times 10^{-5}$  was set as an initial threshold for alignment significance; if an alignment met this criterion, it was then evaluated for potential biological relevance. All results under an  $E$ -value of 10 are shown in the results.

## **RESULTS AND DISCUSSION**

Analysis determined that there are 52 DNA sequences within the SYHT0H2 T-DNA that are contained between known or putative start (ATG) and stop (TAG, TAA, or TGA) codons and have a minimum putative translation size of 30 amino acids on the six DNA reading frames,

including the 5 ORFs that encode the AvHPPD-03 and PAT sequences. The putative amino acid sequences and location within the Event SYHT0H2 T-DNA are shown in Table 1 and Figure 1, respectively. Figure 2 illustrates the ORFs corresponding to the AvHPPD-03 and PAT proteins. Of the 47 putative, unintended ORFs that were identified, the potential for expression could not be ruled out bioinformatically for one. However, none of the ORF sequences show relevant biological similarity to known or putative allergens and toxins.

### **3.5 Allergen Similarity Search**

No significant sequence similarity was observed between any of the SYHT0H2 translated T-DNA ORF sequences and any entry in the FARRP AllergenOnline database (2012) that exceeded the minimum 35% shared identity over a minimum of 80 amino acids. Moreover, no matches of eight (or more) identical contiguous amino acids were observed between any sequence and any entry in the FARRP AllergenOnline database (2012). Alignment summary information and raw alignment data from the FASTA allergen search are shown in Appendix B.

### **3.6 Toxin Similarity Search**

No significant sequence similarity ( $E$ -value less than  $1 \times 10^{-5}$ ) was observed between any of the SYHT0H2 translated T-DNA ORF sequences and any entry in the toxin database. The most similar alignment with a protein from the database had an  $E$ -value of 0.089, greater than the significance threshold  $E$ -value of  $1 \times 10^{-5}$ . The SYHT0H2\_I\_22 sequence (corresponding to the AvHPPD-03 sequence) aligned with a protein described as putative RTX toxin (also described as an autotransporter adhesion) however, SYHT0H2\_I\_22 only aligned at 35% identity over 53 amino acids with the much larger 3,350 amino acid RTX protein. Alignment summary information and raw alignment data from the BLASTP toxin search are shown in Appendix C.

## **4.0 CONCLUSIONS**

Bioinformatic analysis of translations derived from putative open reading frames of DNA sequences within the SYHT0H2 T-DNA were evaluated for their alignment to and similarity with allergens and toxins. ORFs were defined as those that are contained between putative start (ATG) and stop (TAG, TAA, or TGA) codons and have a minimum putative translation size of 30 amino acids. An evaluation of these sequences using the FARRP AllergenOnline database (2012) show no alignments beyond minimum thresholds to any known or putative protein allergens. No significant sequence similarity was observed between any of the SYHT0H2 translated T-DNA ORF sequences and any entry in the toxin database. In summary, evaluation of these sequences using an allergen and toxin database supports the conclusion that these sequences show no biologically relevant similarity to any known or putative allergens or toxins.

**TABLE 1      Event SYHT0H2 T-DNA ORF sequences**

Sequence name (Figure 1)	Size of putative translation (base pair)	Size of putative translation (amino acids)	Putative translation
SYHT0H2_I_1	345	115	MIIARPATGFNLKKLYCQMFERSGKFG DLELDSYLGWPNWPWRSWQLKIPLP KTNIMPSTMLVSSCAQCTPGCVSQSLM QPNRWIVWKAYNSNHRLKTLRLHRLK QMCVQCRS
SYHT0H2_I_2	93	31	MVNDITNRGHISCCSWPNLNWSPLRRH GGIL
SYHT0H2_I_3	93	31	MPSHGGPPPVGHESYLFVRKSQTDCTV FPSW
SYHT0H2_I_4	363	121	MIIARPATGFNLKKLYCQMFERSAGRP MAPISLVQIWVTGLTGLGGAGNSKSLC QKPTSCHPPCLYPAARNVPRAVYPKAS CNLTDGSFGRPITATDLKPCASIDLSKC VYNVDPRPNL
SYHT0H2_I_5	213	71	MLQGHSLELFALTGLTDSLFCPSIMWS FGVLMQLVLLVVSALHLEHLLQDLIC LQGTQHMLLSYFDLEH
SYHT0H2_I_6	246	82	MLHLHKKLQLLQLLPFPFQQMLQEPL LLHMDLLSDLELGLLMQLKHFALVLL EEQDLLLLQQILVTDLDLLKWSMEM WF
SYHT0H2_I_7	111	37	MEPRDDLRSKPTSNTMVDQEFNTSHWL LTMCFELSGK
SYHT0H2_I_8	156	52	MKESVESLEMSCQRNRSRVKNWVFSL IETIKVCYSRSPNQLVIVPLFSSK
SYHT0H2_I_9	237	79	MTRRKSSSTWWSTTRLSTPKISKIQSQK TKGQLRFLNKDDFSTKGNIRKPPRIPLPS YLSLYCEDSGKGRWLLQMPSLR
SYHT0H2_I_10**	549	183	MSPERRPVEIRPATAADMAAVCDIVNH YIETSTVNFRTEPQTPQEWIDDLERLQD RYPWLVAEVEGVVAGIAYAGPWKARN AYDWTVESTVYVSHRHQRLGLGSTLYT HLLKSMEAQGFKSVVAVIGLPNDPSVR LHEALGYTARGTLRAAGYKHGGWHDV GFWQRDFELPAPPRPVRPVTQI
SYHT0H2_I_11	237	79	MFLQTTPTKYQVLFKQNDTDLNIGFYT NHMIFINSVILYGRWRQKSSRIIFEMRKL SSYLFSSSWLFFFLATCLNTSPP
SYHT0H2_I_12	105	35	MVDPHLSATNLICSEKAKPTAQFFHRG RDHFGID
SYHT0H2_I_13	120	40	MPLPTVVPKMDPHPRGASWKKKTFQP RLQSKWIDVISPLT
SYHT0H2_I_14	153	51	MWWVWKGELFRVVQVHRGLREVTGS QAVCRCSEELRALHMTIVQTFGNKVS

Sequence name (Figure 1)	Size of putative translation (base pair)	Size of putative translation (amino acids)	Putative translation
SYHT0H2_I_15	243	81	MTFQQRVISGNLLGFHCPAICHFIVKIVE KEGGSYKCHHCDKGKAIVEDASADSGP KDGPPPTRSIVEKEDVPTTSSKQVD
SYHT0H2_I_16**	549	183	MSPERRPVEIRPATAADMAAVCDIVNH YIETSTVNFRTEPQTPQEWIDDLERLQD RYPWLVAEVEGVVAGIAYAGPWKARN AYDWTVESTVYVSHRHQRLGLGSTLYT HLLKSMEAQGFKSVAVIGLPNDPSVR LHEALGYTARGTLRAAGYKHGGWHDV GFWQRDFELPAPPRPVRPVTQI
SYHT0H2_I_17	105	35	MRLTKVRSAAFNQYQSGPYHDGKTVQ SVWLFLTNN
SYHT0H2_I_18	192	64	MRLWDTQPGVHCAQLDTSMDVGMML VFGKGILSCQLLQGGQLGQLPRYESSRS PNFPDRSNIWQ
SYHT0H2_I_19	171	57	MPSTMLVSSCAQCTPGCVSQSLMQPNR WIVWKAYNSNHRLKTLRLHRLKQMCV QCGS
SYHT0H2_I_20	120	40	MVNDITNRGHISCCSWPNLNWSPLRRH GGLVISACPLQMK
SYHT0H2_I_21	222	74	MMAFVGATFLFHYLHNKVTDSWAMEP LLLHMDLLSDLELNIRKPPRIPLPSYLS LYCEDSGKGRWLLQMPSLR
SYHT0H2_I_22*	1317	439	MPPTPATATGAAAAAVTPEHAARSFPR VVRVNPRSDRFPVLSFHHVELWCADAA SAAGRFSFALGAPLAARSDLSTGNSAH ASLLLRSGALAFLLTAPYAPPPQEAATA ATASIPSFSADAARTFAAAHGLAVRSVG VRVADAAEAFRVSVAGGARPAFAPAD LGHGFGLALEVELYGDVVLRVFSYPDET DLPFLPGFERVSSPGAVDYGLTRFDHVV GNVPEMAPVIDYMKGFLGFHEFAEFTA EDVGTTESGLNSVVLANNSEAVLLPLN EPVHGTKRRSQIQTYLEYHGGPGVQHI ALASNDVLRTLREMRARTPMGGFEFM APPQAKYYEGVRRIAGDVLSEEQIKECQ ELGVLVDRDDQGVLLQIFTKPVGDRPT FFLEMIQRIGCMKDEVGQEYQKGGCG GFGKGNFSELFKSIEDYEKSLEVKQSVV AQKS
SYHT0H2_I_23	90	30	MVEHDTLVYSKNIKDTVSEDQRAIETF QQR
SYHT0H2_I_24	120	40	MPLPTVVPKMDPHPRGASWKKKTFQP RLQSKWIDVISPLT

Sequence name (Figure 1)	Size of putative translation (base pair)	Size of putative translation (amino acids)	Putative translation
SYHT0H2_I_25	195	65	MTHNPTILRKTLPLYKEVHFIWRGHAEI TSPPCLRRGDQLRLGQLQQLIWPRFVIS LTITLRRQLQ
SYHT0H2_I_26	207	69	MRLWDTQPGVHCAQLDTSMDGMML VFGKGILSCQLLQGQLGQLPRSELVISA PWVDLQIVQTFGNKVS
SYHT0H2_I_27	105	35	MSTHPSVYKYLAPPSLLREQNLREIVLE REREQVA
SYHT0H2_I_28**	549	183	MSPERRPVEIRPATAADMAAVCDIVNH YIETSTVNFRTPEQTPQEWIDDLERLQD RYPWLVAEVEGVVAGIAYAGPWKARN AYDWTVESTVYVSHRHQRLGLGSTLYT HLLKSMEAQGFKSVAVIGLPNDPSVR LHEALGYTARGTLRAAGYKHGGWHDV GFWQRDFELPAPPRPVRPVTQI
SYHT0H2_I_29	96	32	MMAFVGATFLFHYLHNKVTDSWAMES EEVSGY
SYHT0H2_I_30	132	44	MVGQFRQDSSRISEPHLHTAPLQQVQIR DQDLLEQKQVLLLQQH
SYHT0H2_I_31	171	57	MRCVELLVHHGIRGRFGSEIVSWFHEQ ARSMEEQLHCCWLEPHCSVQILWFQH PLL
SYHT0H2_I_32	363	121	MIIIARPATGFNLKKLYCQMFERSAGRP MAPISLVQIWVTGLTGLGGAGNSKSLC QKPTSCHPPCLYPAARNVPRAVYPKAS CNLTDGSFGRPITATDLKPCASIDLSKC VYNVDPRPNL
SYHT0H2_I_33	93	31	MPSHGGPPPVGHESYLFVRKSQTDCTV FPSW
SYHT0H2_I_34	93	31	MVNDITNRGHISCCSWPNLNWSPLRRH GGIL
SYHT0H2_I_35	345	115	MIIIARPATGFNLKKLYCQMFERSGKFG DLELDSYLGWPNWPWRSWQLKIPLP KTNIMPSTMLVSSCAQCTPGCVSQSLM QPNRWIVWKAYNSNHRLKTLRLHRLK QMCVQCRS
SYHT0H2_I_36	192	64	MRLWDTQPGVHCAQLDTSMDGMML VFGKGILSCQLLQGQLGQLPRYESSRS PNFPDRSNIWQ
SYHT0H2_I_37	105	35	MRLTKVRSAAFNQYQSGPYHDGKTVQ SVWLFLTNN

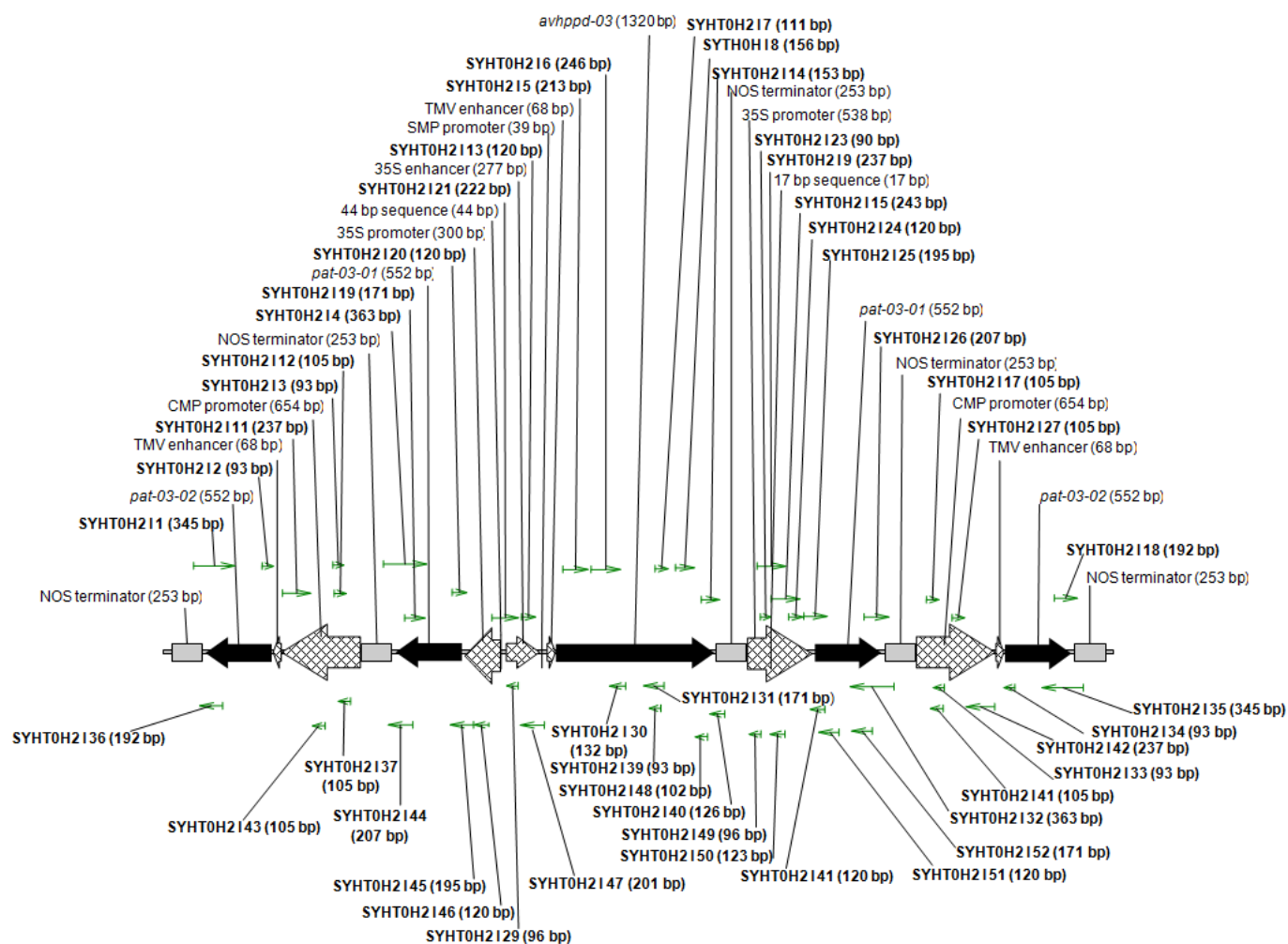
Sequence name (Figure 1)	Size of putative translation (base pair)	Size of putative translation (amino acids)	Putative translation
SYHT0H2_I_38**	549	183	MSPERRPVEIRPATAADMAAVCDIVNH YIETSTVNFRTEPQTPQEWIDDLERLQD RYPWLVAEVEGVVAGIAYAGPWKARN AYDWTVESTVYVSHRHQRLGLGSTLYT HLLKSMEAQGFKSVAVIGLPNDPSVR LHEALGYTARGTLRAAGYKHGGWHDV GFWQRDFELPAPPRPVRPVTQI
SYHT0H2_I_39	93	31	MVFEVGLDLRSSLGSMNRLVQWQKNS FTVVG
SYHT0H2_I_40	126	42	MIIIARPATGFNLKKLYCQMFERSYEE LLALLSNDRLLDQ
SYHT0H2_I_41	105	35	MVDPHLSATNLICSEKAKPTAQFFHRG RDHFGID
SYHT0H2_I_42	237	79	MFLQTTPTKYQVLFKQNDTDLNIGFYT NHMIFINSVILYGRWRQKSSRIIFEMRKL SSYLFSSSWLFFFLATCLNTSPP
SYHT0H2_I_43	105	35	MSTHPSVYKYLAPPSLLREQNLREIVLE REREQVA
SYHT0H2_I_44	207	69	MRLWDTQPGVHCAQLDTSMVDGMML VFGKGILSCQLLQGLGQLPRSELVISA PWVDLQIVQTFGNKVS
SYHT0H2_I_45	195	65	MTHNPTILRKTLPLYKEVHFIWRGHAIE TSPPCLRRGDQLRLGQLQLIWPRFVIS LTITLRLRLQ
SYHT0H2_I_46	120	40	MPLPTVVPKMDPHPRGASWKKKTFQP RLQSKWIDVISPLT
SYHT0H2_I_47	201	67	MNFLFIGSSRVDLQKDSGIVRHPLRQW RYHINPLALKTWLERLLFPRCSSLWGV HLWDHCRQRHLQR
SYHT0H2_I_48	102	34	MDLNNSEKFPFPNPPHPPFWYSCPTSSF SMHPIR
SYHT0H2_I_49	96	32	MLTKIFFLSLSRKRLCMNCSPVFTASSV RSSI
SYHT0H2_I_50	123	41	MMAFVGATFLFHYLHNKVTDSWAMES EEVSGYYPLLKSHLC
SYHT0H2_I_51	120	40	MVNDITNRGHISCCSWPNLNWSPLRRH GGLVISACPLQMK
SYHT0H2_I_52	171	57	MPSTMLVSSCAQCTPGCVSQSLMQPNR WIVWKAYNSNHRLKTLRLHRLKQMCV QCGS

\* The translation of SYHT0H2\_I\_22 contains the AvHPPD-03 sequence

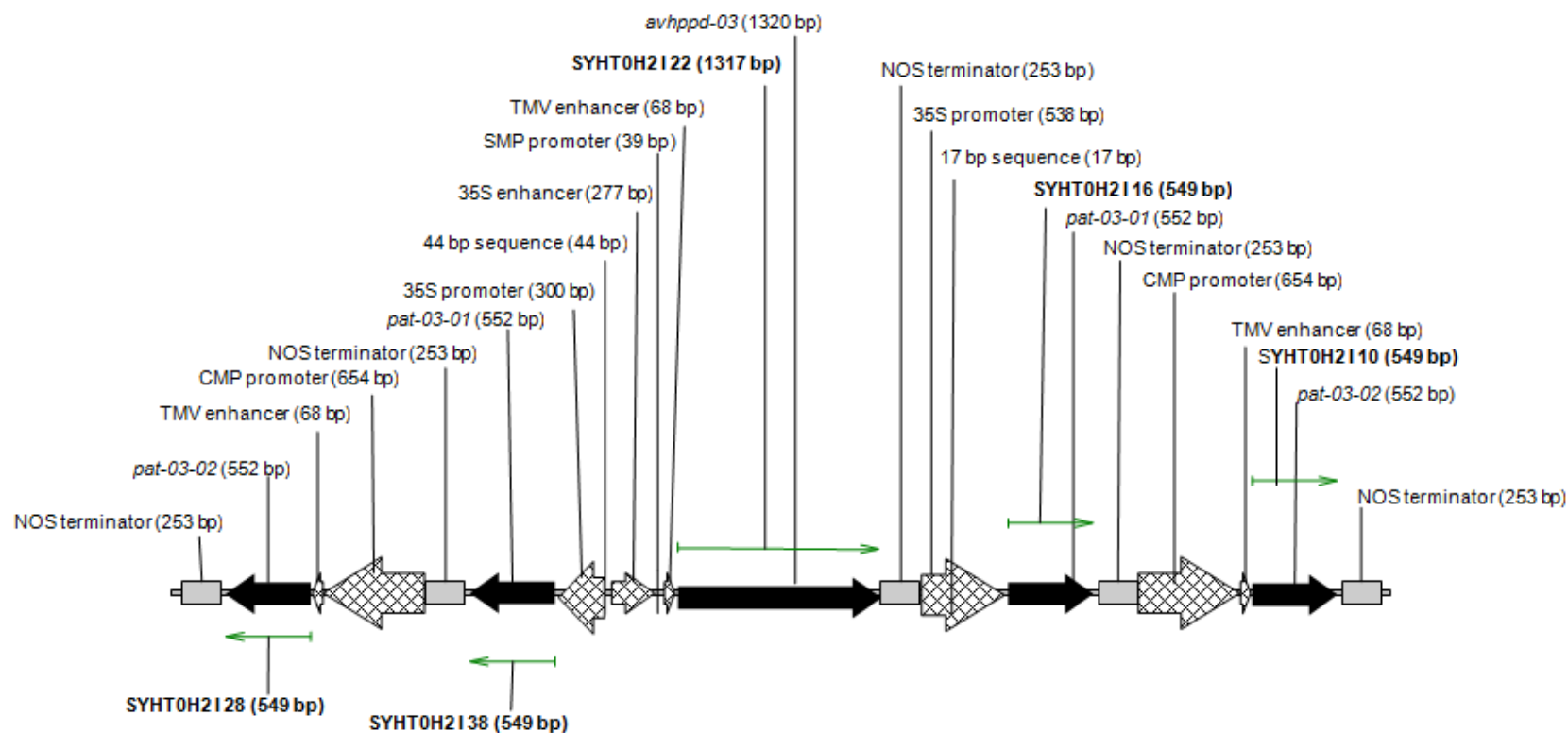
\*\* The translation of SYHT0H2\_I\_10, 16, 28, and 38 contains the PAT sequences



**FIGURE 1**      **Location of the putative SYHT0H2 ORF sequences.**



**FIGURE 2**      **Location of the SYHT0H2 ORF sequences containing *avhppd-03* and *pat*.**



## 5.0 REFERENCES

- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402.
- Codex Alimentarius Commission. 2009. Foods derived from modern biotechnology. FAO/WHO Rome, Italy: 1-85.
- FARRP. 2012. FARRP Protein AllergenOnline Database, version 12, updated February 7, 2012. Lincoln, NB: Food Allergy Research and Resource Program, University of Nebraska in Lincoln. <http://www.allergenonline.org>. (Accessed February 7, 2012).
- Goodman RE, Vieths S, Sampson HA, Hill D, Ebisawa M, Taylor SL, van Ree R. 2008. Allergenicity assessment of genetically modified crops--what makes sense? *Nat Biotechnol* 26: 73-81.
- Hileman RE, Silvanovich A, Goodman RE, Rice EA, Holleschak G, Astwood JD, Hefle SL. 2002. Bioinformatic methods for allergenicity assessment using a comprehensive allergen database. *Int Arch Allergy Imm* 128:280–291.
- NCBI. 2012. Entrez® Protein database. Bethesda, MD: National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health. <http://www.ncbi.nlm.nih.gov/sites/entrez?db=Protein>.
- Pearson WR, Lipman DJ. 1988. Improved tools for biological sequence comparison. *P Natl Acad Sci USA* 85:2444–2448.
- Ponting C. 2001. Issues in predicting protein function from sequence. *Brief Bioinform* 2:19–29.
- Silvanovich A, Nemeth MA, Song P, Herman R, Tagliani L, Bannon GA. 2006. The value of short amino acid sequence matches for prediction of protein allergenicity. *Toxicol Sci* 90: 252-8.
- US EPA. 1989. Good Laboratory Practice Standards. 40 CFR Part 160.3.

## **APPENDICES SECTION**

## APPENDIX A Protein sequences included in the FARRP AllergenOnline database, version 12

The GenInfo identifier is a number assigned to a nucleotide or protein sequence and used to differentiate reported sequence variations within an entry.

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1	<i>Acarus siro</i>	Mite	Unassigned	allergen Aca s 13 [Acarus siro]	118638268
2	<i>Actinidia chinensis</i>	Kiwi	Unassigned	RecName: Full=Actinidain; Short=Actinidin; AltName: Allergen=Act c 1; Flags: Precursor	190358935
3	<i>Actinidia chinensis</i>	Kiwi	Unassigned	bet v 1 related allergen [Actinidia chinensis]	281552896
4	<i>Actinidia chinensis</i>	Kiwi	Unassigned	Kiwellin (Allergen Act c ?)	85701136
5	<i>Actinidia chinensis</i>	Kiwi	Unassigned	Thaumatococcus-like protein (Allergen Act c 2)	68064399
6	<i>Actinidia deliciosa</i>	Kiwi	Unassigned	RecName: Full=Actinidain; Short=Actinidin; AltName: Full=Allergen Act d 1; AltName: Allergen=Act d 1; Flags: Precursor	193806686
7	<i>Actinidia deliciosa</i>	Kiwi	Unassigned	actinidin	166317
8	<i>Actinidia deliciosa</i>	Kiwi	Unassigned	unnamed protein product [Actinidia deliciosa]	15984
9	<i>Actinidia deliciosa</i>	Kiwi	Unassigned	bet v 1 related allergen [Actinidia deliciosa]	281552898
10	<i>Actinidia deliciosa</i>	Kiwi	Unassigned	RecName: Full=Kirola; AltName: Allergen=Act d 11	332319679
11	<i>Actinidia deliciosa</i>	Kiwi	Unassigned	phytoalexin [Actinidia deliciosa]	40807635
12	<i>Actinidia deliciosa</i>	Kiwi	Unassigned	thaumatococcus-like protein [Actinidia deliciosa]	146737976
13	<i>Actinidia deliciosa</i>	Kiwi	Act c 2	thaumatococcus-like protein [Actinidia deliciosa]	71057064

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
14	<i>Aedes aegypti</i>	Yellow fever mosquito	Aed a 2	RecName: Full=37 kDa salivary gland allergen Aed a 2; AltName: Full=Protein D7; AltName: Allergen=Aed a 2; Flags: Precursor	205525919
15	<i>Aedes aegypti</i>	Yellow fever mosquito	Aed a 2	D7 protein precursor (Allergen Aed a 2)	118216
16	<i>Aedes aegypti</i>	Yellow fever mosquito	Unassigned	30 kDa salivary gland allergen variant 3 [ <i>Aedes aegypti</i> ]	94468552
17	<i>Aedes aegypti</i>	Yellow fever mosquito	Unassigned	30 kDa salivary gland allergen variant 2 [ <i>Aedes aegypti</i> ]	94468546
18	<i>Aedes aegypti</i>	Yellow fever mosquito	Unassigned	30 kDa salivary gland allergen Aed a 3 [ <i>Aedes aegypti</i> ]	2114497
19	<i>Aedes aegypti</i>	Yellow fever mosquito	Unassigned	RecName: Full=Apyrase; AltName: Full=Adenosine diphosphatase; AltName: Full=ATP-diphosphohydrolase; AltName: Full=ATP-diphosphatase; Short=ADPase; AltName: Allergen=Aed a 1; Flags: Precursor	193806340
20	<i>Aedes aegypti</i>	Yellow fever mosquito	Aed a 1	apyrase	556272
21	<i>Agrostis alba</i>	Bent grass	Unassigned	Group I allergen Agr a I (Form 2), pollen	75139987
22	<i>Agrostis alba</i>	Bent grass	Unassigned	Group I allergen Agr a I (Form 1), pollen	75139989
23	<i>Agrostis alba</i>	Bent grass	Unassigned	pollen allergen Agr a I - bent grass (fragment)	320606
24	<i>Alnus glutinosa</i>	Alder	Aln g 1	Aln g I [ <i>Alnus glutinosa</i> ]	261407
25	<i>Alnus glutinosa</i>	Alder	Unassigned	pollen allergen Aln g 4 [ <i>Alnus glutinosa</i> ]	3319651

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
26	<i>Alternaria alternata</i>	Fungus	Unassigned	aldehyde dehydrogenase (NAD+) [ <i>Alternaria alternata</i> ]	76666767
27	<i>Alternaria alternata</i>	Fungus	Unassigned	RecName: Full=Glutathione-S-transferase; AltName: Allergen=Alt a 13	74611808
28	<i>Alternaria alternata</i>	Fungus	Unassigned	major allergen Alt a 1 subunit [ <i>Alternaria alternata</i> ]	45680856
29	<i>Alternaria alternata</i>	Fungus	Alt a 1	major allergen alt a1 [ <i>Alternaria alternata</i> ]	21913174
30	<i>Alternaria alternata</i>	Fungus	Alt a 1	major allergen Alt a 1 subunit [ <i>Alternaria alternata</i> ]	1842045
31	<i>Alternaria alternata</i>	Fungus	Alt a 6	Enolase (2-phosphoglycerate dehydratase) (2-phospho-D-glycerate hydro-lyase) (Major allergen Alt a 11) (Alt a XI)	14423684
32	<i>Alternaria alternata</i>	Fungus	Unassigned	Minor allergen Alt a 7 (Alt a VII)	1168402
33	<i>Alternaria alternata</i>	Fungus	Alt a 3	Heat shock 70 kDa protein (Allergen Alt a 3)	14423730
34	<i>Alternaria alternata</i>	Fungus	Unassigned	RecName: Full=Superoxide dismutase [Mn], mitochondrial; AltName: Full=Manganese-dependent superoxide dismutase; Short=MnSOD; AltName: Allergen=Alt a MnSOD	292630881
35	<i>Alternaria alternata</i>	Fungus	Unassigned	putative nuclear transport factor 2 [ <i>Alternaria alternata</i> ]	21748153
36	<i>Alternaria alternata</i>	Fungus	Alt a 12	60S acidic ribosomal protein P1 (Allergen Alt a 12) (Alt a XII)	1350779

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
37	<i>Alternaria alternata</i>	Fungus	Unassigned	RecName: Full=60S acidic ribosomal protein P2; AltName: Full=Minor allergen Alt a 5; AltName: Full=Allergen Alt a 6; AltName: Full=Allergen Alt a VI; AltName: Allergen=Alt a 5	1173071
38	<i>Alternaria alternata</i>	Fungus	Alt a 5	ribosomal P2 phosphoprotein [Alternaria alternata]	1850540
39	<i>Alternaria alternata</i>	Fungus	Unassigned	TCTP [Alternaria alternata]	112824341
40	<i>Alternaria alternata</i>	Fungus	Alt a 4	Protein disulfide-isomerase (PDI) (Allergen Alt a 4)	85701160
41	<i>Amaranthus retroflexus</i>	Common Amaranth	Unassigned	Ama r 2 pollen allergen [Amaranthus retroflexus]	227937304
42	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	putative pectate lyase precursor [Ambrosia artemisiifolia]	302127828
43	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	putative pectate lyase precursor [Ambrosia artemisiifolia]	302127826
44	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	putative pectate lyase precursor [Ambrosia artemisiifolia]	302127824
45	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	putative pectate lyase precursor [Ambrosia artemisiifolia]	302127822
46	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	putative pectate lyase precursor [Ambrosia artemisiifolia]	302127820
47	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	putative pectate lyase precursor [Ambrosia artemisiifolia]	302127818
48	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	putative pectate lyase precursor [Ambrosia artemisiifolia]	302127816
49	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	putative pectate lyase precursor [Ambrosia artemisiifolia]	302127814
50	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	putative pectate lyase precursor [Ambrosia artemisiifolia]	302127812



Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
51	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	putative pectate lyase precursor [ <i>Ambrosia artemisiifolia</i> ]	302127810
52	<i>Ambrosia artemisiifolia</i>	Short ragweed	Amb a 1.3	antigen E	166443
53	<i>Ambrosia artemisiifolia</i>	Short ragweed	Amb a 1.4	Pollen allergen Amb a 1.4 precursor (Antigen E) (Antigen Amb a I)	113478
54	<i>Ambrosia artemisiifolia</i>	Short ragweed	Amb a 1.3	Pollen allergen Amb a 1.3 precursor (Antigen E) (Antigen Amb a I)	113477
55	<i>Ambrosia artemisiifolia</i>	Short ragweed	Amb a 1.2	Pollen allergen Amb a 1.2 precursor (Antigen E) (Antigen Amb a I) (AaBA protein)	113476
56	<i>Ambrosia artemisiifolia</i>	Short ragweed	Amb a 1.1	Pollen allergen Amb a 1.1 precursor (Antigen E) (AgE) (Antigen Amb a I)	113475
57	<i>Ambrosia artemisiifolia</i>	Short ragweed	Amb a 2	Pollen allergen Amb a 2 precursor (Antigen K) (Antigen Amb a II)	113479
58	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	ragweed homologue of Art v 1 precursor [ <i>Ambrosia artemisiifolia</i> ]	291482318
59	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	ragweed homologue of Art v 1 precursor [ <i>Ambrosia artemisiifolia</i> ]	291482316
60	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	ragweed homologue of Art v 1 precursor [ <i>Ambrosia artemisiifolia</i> ]	291482314
61	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	ragweed homologue of Art v 1 precursor [ <i>Ambrosia artemisiifolia</i> ]	291482310
62	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	ragweed homologue of Art v 1 precursor [ <i>Ambrosia artemisiifolia</i> ]	291482308
63	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	ragweed homologue of Art v 1 precursor [ <i>Ambrosia artemisiifolia</i> ]	291482306
64	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	ragweed homologue of Art v 1 precursor [ <i>Ambrosia artemisiifolia</i> ]	291197394
65	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	ragweed homologue of Art v 1 precursor [ <i>Ambrosia artemisiifolia</i> ]	285005079

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
66	<i>Ambrosia artemisiifolia</i>	Short ragweed	Amb a 6	Nonspecific lipid-transfer protein precursor (LTP) (Pollen allergen Amb a 6) (Amb a VI) (Allergen Ra6)	14285595
67	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	profilin isoallergen 2 [Ambrosia artemisiifolia]	62249512
68	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	profilin isoallergen 1 [Ambrosia artemisiifolia]	62249502
69	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	profilin-like protein [Ambrosia artemisiifolia]	34851178
70	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	profilin-like protein [Ambrosia artemisiifolia]	34851180
71	<i>Ambrosia artemisiifolia</i>	Short ragweed	Unassigned	profilin-like protein [Ambrosia artemisiifolia]	34851182
72	<i>Ambrosia artemisiifolia (elator)</i>	Short ragweed	Amb a 3	Pollen allergen Amb a 3 (Amb a III) (Allergen Ra3)	416636
73	<i>Ambrosia artemisiifolia (elator)</i>	Short ragweed	Amb a 5	Pollen allergen Amb a 5 (Amb a V) (Allergen Ra5)	114090
74	<i>Ambrosia psilostachya</i>	Western ragweed	Unassigned	Amb p V allergen	515957
75	<i>Ambrosia psilostachya</i>	Western ragweed	Unassigned	Amb p V allergen	515956
76	<i>Ambrosia psilostachya</i>	Western ragweed	Unassigned	Amb p V allergen	515955
77	<i>Ambrosia psilostachya</i>	Western ragweed	Unassigned	Amb p V allergen	515954
78	<i>Ambrosia psilostachya</i>	Western ragweed	Unassigned	Amb p V allergen	515953
79	<i>Ambrosia trifida</i>	Giant ragweed	Amb t 5	Pollen allergen Amb t 5 precursor (Amb t V) (Allergen Ra5G)	114091
80	<i>Anacardium occidentale</i>	Cashew	Ana o 1	vicilin-like protein [Anacardium occidentale]	21914823

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
81	<i>Anacardium occidentale</i>	Cashew	Ana o 1	vicilin-like protein [Anacardium occidentale]	21666498
82	<i>Anacardium occidentale</i>	Cashew	Ana o 2	allergen Ana 0 2 [Anacardium occidentale]	25991543
83	<i>Anacardium occidentale</i>	Cashew	Ana o 3	2s albumin [Anacardium occidentale]	24473800
84	<i>Ananas comosus</i>	Pineapple	Unassigned	Bromelain precursor (Allergen Ana c 2)	75277440
85	<i>Ananas comosus</i>	Pineapple	Unassigned	RecName: Full=Profilin; AltName: Full=Minor food allergen Ana c 1; AltName: Allergen=Ana c 1	75306610
86	<i>Anisakis simplex</i>	Parasitic fish worm	Ani s 1	Major allergen Ani s 1 precursor (Excretory gland allergen Ans1) (21 kDa allergen)	47605452
87	<i>Anisakis simplex</i>	Parasitic fish worm	Unassigned	Ani s 11-like protein 2 precursor [Anisakis simplex]	323575365
88	<i>Anisakis simplex</i>	Parasitic fish worm	Unassigned	Ani s 11-like protein precursor [Anisakis simplex]	323575363
89	<i>Anisakis simplex</i>	Parasitic fish worm	Unassigned	Ani s 11 allergen precursor [Anisakis simplex]	323575361
90	<i>Anisakis simplex</i>	Parasitic fish worm	Unassigned	Ani s 12 allergen precursor [Anisakis simplex]	323575367
91	<i>Anisakis simplex</i>	Parasitic fish worm	Unassigned	Paramyosin (Allergen Ani s 2)	42559536
92	<i>Anisakis simplex</i>	Parasitic fish worm	Ani s 2	paramyosin isoform [Anisakis simplex]	8453086
93	<i>Anisakis simplex</i>	Parasitic fish worm	Unassigned	ani s 4 allergen [Anisakis simplex]	110346534
94	<i>Anisakis simplex</i>	Parasitic fish worm	Ani s 4	Allergen Ani s 4	47605398

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
95	<i>Anisakis simplex</i>	Parasitic fish worm	Unassigned	SXP/RAL-2 family protein [Anisakis simplex]	121308878
96	<i>Anisakis simplex</i>	Parasitic fish worm	Unassigned	UA3-recognized allergen [Anisakis simplex]	119524036
97	<i>Anisakis simplex</i>	Parasitic fish worm	Unassigned	SXP/RAL-2 family protein 2 isoform 10 [Anisakis simplex]	155676698
98	<i>Anisakis simplex</i>	Parasitic fish worm	Unassigned	SXP/RAL-2 family protein 2 isoform 9 [Anisakis simplex]	155676696
99	<i>Anisakis simplex</i>	Parasitic fish worm	Unassigned	SXP/RAL-2 family protein 2 isoform 8 [Anisakis simplex]	155676694
100	<i>Anisakis simplex</i>	Parasitic fish worm	Unassigned	SXP/RAL-2 family protein 2 isoform 7 [Anisakis simplex]	155676692
101	<i>Anisakis simplex</i>	Parasitic fish worm	Unassigned	SXP/RAL-2 family protein 2 isoform 6 [Anisakis simplex]	155676690
102	<i>Anisakis simplex</i>	Parasitic fish worm	Unassigned	SXP/RAL-2 family protein 2 isoform 5 [Anisakis simplex]	155676688
103	<i>Anisakis simplex</i>	Parasitic fish worm	Unassigned	SXP/RAL-2 family protein 2 isoform 4 [Anisakis simplex]	155676686
104	<i>Anisakis simplex</i>	Parasitic fish worm	Unassigned	SXP/RAL-2 family protein 2 isoform 3 [Anisakis simplex]	155676684
105	<i>Anisakis simplex</i>	Parasitic fish worm	Unassigned	SXP/RAL-2 family protein 2 isoform 2 [Anisakis simplex]	155676682
106	<i>Anisakis simplex</i>	Parasitic fish worm	Unassigned	SXP/RAL-2 family protein 2 isoform 1 [Anisakis simplex]	155676636
107	<i>Anisakis simplex</i>	Parasitic fish worm	Unassigned	Ani s 9 allergen precursor [Anisakis simplex]	157418806
108	<i>Anisakis simplex</i>	Parasitic fish worm	Unassigned	troponin-like protein [Anisakis simplex]	6065738
109	<i>Anthoxanthum odoratum</i>	Sweet vernal grass	Unassigned	Group I allergen Ant o I (Form 2), pollen	75139990

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
110	<i>Anthoxanthum odoratum</i>	Sweet vernal grass	Unassigned	Group I allergen Ant o I (Form 1), pollen	75139986
111	<i>Anthoxanthum odoratum</i>	Sweet vernal grass	Unassigned	pollen allergen Ant o I - sweet vernal grass (fragment)	320607
112	<i>Apis cerana</i>	Indian honeybee	Unassigned	phospholipase A2 (EC 3.1.1.4), venom - Indian honeybee	7435005
113	<i>Apis cerana cerana</i>	Indian honeybee	Unassigned	Phospholipase A2 (Phosphatidylcholine 2-acylhydrolase)	24638082
114	<i>Apis dorsata</i>	Giant honeybee	Unassigned	Phospholipase A2 (Phosphatidylcholine 2-acylhydrolase)	47117012
115	<i>Apis dorsata</i>	Giant honeybee	Unassigned	Melittin	126955
116	<i>Apis mellifera</i>	Honeybee	Api m 1	Phospholipase A2 precursor (Phosphatidylcholine 2-acylhydrolase) (Allergen Api m 1) (Api m I)	24418862
117	<i>Apis mellifera</i>	Honeybee	Api m 2	Hyaluronoglucosaminidase precursor (Hyaluronidase) (Hya) (Allergen Api m 2) (Api m II)	585279
118	<i>Apis mellifera</i>	Honeybee	Unassigned	venom acid phosphatase precursor [Apis mellifera]	60652325
119	<i>Apis mellifera</i>	Honeybee	Unassigned	venom acid phosphatase [Apis mellifera]	208342441
120	<i>Apis mellifera</i>	Honeybee	Unassigned	Melittin precursor (Allergen Api m 3) (Api m III)	126949
121	<i>Apis mellifera</i>	Honeybee	Unassigned	melittin, minor - honeybee	69552
122	<i>Apis mellifera</i>	Honeybee	Unassigned	RecName: Full=Venom dipeptidyl peptidase 4; AltName: Full=Allergen C; AltName: Full=Venom dipeptidyl peptidase IV; AltName: Allergen=Api m 5; Flags: Precursor	313471719
123	<i>Apis mellifera</i>	Honeybee	Unassigned	allergen Api m 6 variant 2 precursor [Apis mellifera]	88770352

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
124	<i>Apis mellifera</i>	Honeybee	Unassigned	allergen Api m 6 variant 1 [Apis mellifera]	94400907
125	<i>Apis mellifera</i>	Honeybee	Unassigned	icarapin variant 2 precursor [Apis mellifera]	94471624
126	<i>Apis mellifera</i>	Honeybee	Unassigned	icarapin variant 1 precursor [Apis mellifera]	94471622
127	<i>Apium graveolens</i>	Celery	Api g 1.0201	Major allergen Api g 2 (Api g 1.0201)	14423646
128	<i>Apium graveolens</i>	Celery	Api g 1.0101	Major allergen Api g 1 (Api g 1.0101) (Api g I)	1346568
129	<i>Apium graveolens</i>	Celery	Api g 2.0101	stalk non-specific lipid-transfer protein 1 [Apium graveolens]	256600126
130	<i>Apium graveolens</i>	Celery	Api g 4	profilin [Apium graveolens]	4761578
131	<i>Apium graveolens</i>	Celery	Unassigned	[Segment 4 of 4] Allergen Api g 5	32363126
132	<i>Apium graveolens</i>	Celery	Unassigned	[Segment 3 of 4] Allergen Api g 5	32363125
133	<i>Apium graveolens</i>	Celery	Unassigned	[Segment 2 of 4] Allergen Api g 5	32363124
134	<i>Apium graveolens</i>	Celery	Unassigned	[Segment 1 of 4] Allergen Api g 5	33300921
135	<i>Arachis hypogaea</i>	Peanut	Unassigned	peanut agglutinin precursor; prePNA [Arachis hypogaea]	253289
136	<i>Arachis hypogaea</i>	Peanut	Unassigned	Ara h 1 allergen [Arachis hypogaea]	312233063
137	<i>Arachis hypogaea</i>	Peanut	Unassigned	conarachin [Arachis hypogaea]	46560476
138	<i>Arachis hypogaea</i>	Peanut	Unassigned	conarachin [Arachis hypogaea]	46560472
139	<i>Arachis hypogaea</i>	Peanut	Unassigned	conarachin [Arachis hypogaea]	46560474
140	<i>Arachis hypogaea</i>	Peanut	Ara h 1	Allergen Ara h 1, clone P41B precursor (Ara h I)	1168391
141	<i>Arachis hypogaea</i>	Peanut	Ara h 1	Allergen Ara h 1, clone P17 precursor (Ara h I)	1168390
142	<i>Arachis hypogaea</i>	Peanut	Unassigned	Ara h 2.01 allergen [Arachis hypogaea]	224747150
143	<i>Arachis hypogaea</i>	Peanut	Unassigned	allergen II [Arachis hypogaea]	15418705

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
144	<i>Arachis hypogaea</i>	Peanut	Ara h 2	allergen Ara h 2 isoform [Arachis hypogaea]	31322017
145	<i>Arachis hypogaea</i>	Peanut	Ara h 2.02	allergen Ara h 2.02 [Arachis hypogaea]	26245447
146	<i>Arachis hypogaea</i>	Peanut	Unassigned	profilin [Arachis hypogaea]	284810529
147	<i>Arachis hypogaea</i>	Peanut	Ara h 5	profilin [Arachis hypogaea]	5902968
148	<i>Arachis hypogaea</i>	Peanut	Unassigned	RecName: Full=Conglutin; AltName: Allergen=Ara h 6; Flags: Precursor	75114094
149	<i>Arachis hypogaea</i>	Peanut	Unassigned	Chain A, Allergen Arah6 From Peanut (Arachis Hypogaea)	159163254
150	<i>Arachis hypogaea</i>	Peanut	Unassigned	conglutin [Arachis hypogaea]	17225991
151	<i>Arachis hypogaea</i>	Peanut	Ara h 6	allergen Arah6 [Arachis hypogaea]	5923742
152	<i>Arachis hypogaea</i>	Peanut	Unassigned	Ara h 7 allergen precursor [Arachis hypogaea]	158121995
153	<i>Arachis hypogaea</i>	Peanut	Ara h 7	allergen [Arachis hypogaea]	5931948
154	<i>Arachis hypogaea</i>	Peanut	Unassigned	pathogenesis-related protein 10 [Arachis hypogaea]	110676574
155	<i>Arachis hypogaea</i>	Peanut	Unassigned	Ara h 8 allergen isoform 3 [Arachis hypogaea]	169786740
156	<i>Arachis hypogaea</i>	Peanut	Unassigned	Ara h 8 allergen isoform [Arachis hypogaea]	145904610
157	<i>Arachis hypogaea</i>	Peanut	Ara h 8	Ara h 8 allergen [Arachis hypogaea]	37499626
158	<i>Arachis hypogaea</i>	Peanut	Unassigned	LTP isoallergen 2 [Arachis hypogaea]	161610580
159	<i>Arachis hypogaea</i>	Peanut	Unassigned	LTP isoallergen 1 precursor [Arachis hypogaea]	161087230
160	<i>Arachis hypogaea</i>	Peanut	Unassigned	Ara h 3 allergen [Arachis hypogaea]	312233065
161	<i>Arachis hypogaea</i>	Peanut	Unassigned	Chain A, Crystal Structure Of Peanut Major Allergen Ara H 3	224036293
162	<i>Arachis hypogaea</i>	Peanut	Unassigned	arachin Arah3 isoform [Arachis hypogaea]	199732457
163	<i>Arachis hypogaea</i>	Peanut	Unassigned	iso-Ara h3 [Arachis hypogaea]	112380623
164	<i>Arachis hypogaea</i>	Peanut	Unassigned	trypsin inhibitor [Arachis hypogaea]	22135348

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
165	<i>Arachis hypogaea</i>	Peanut	Unassigned	allergen Arah3/Arah4 [Arachis hypogaea]	21314465
166	<i>Arachis hypogaea</i>	Peanut	Ara h 4	glycinin [Arachis hypogaea]	5712199
167	<i>Arachis hypogaea</i>	Peanut	Ara h 3	glycinin [Arachis hypogaea]	3703107
168	<i>Argas reflexus</i>	European pigeon tick	Unassigned	Chain A, Crystal Structure Of Arg R 1 In Complex With Histamine	322812205
169	<i>Argas reflexus</i>	European pigeon tick	Arg r 1	Arg r 1 precursor [Argas reflexus]	58371884
170	<i>Artemisia vulgaris</i>	Mugwort	Unassigned	Amb a 1-like protein [Artemisia vulgaris]	62530263
171	<i>Artemisia vulgaris</i>	Mugwort	Art v 1	major pollen allergen Art v 1 precursor [Artemisia vulgaris]	27818335
172	<i>Artemisia vulgaris</i>	Mugwort	Unassigned	art v 2 allergen [Artemisia vulgaris]	148887203
173	<i>Artemisia vulgaris</i>	Mugwort	Unassigned	Art v 3 allergen precursor [Artemisia vulgaris]	189544595
174	<i>Artemisia vulgaris</i>	Mugwort	Unassigned	Art v 3.0301 allergen precursor [Artemisia vulgaris]	189544590
175	<i>Artemisia vulgaris</i>	Mugwort	Unassigned	Art v 3 allergen precursor [Artemisia vulgaris]	189544584
176	<i>Artemisia vulgaris</i>	Mugwort	Unassigned	Art v 3.0201 allergen precursor [Artemisia vulgaris]	189544578
177	<i>Artemisia vulgaris</i>	Mugwort	Unassigned	Nonspecific lipid-transfer protein (LTP) (Pollen allergen Art v 3)	73621307
178	<i>Artemisia vulgaris</i>	Mugwort	Unassigned	Profilin-2 (Pollen allergen Art v 4.02)	73621416
179	<i>Artemisia vulgaris</i>	Mugwort	Unassigned	Profilin-1 (Pollen allergen Art v 4.01)	73621415
180	<i>Arthroderma benhamiae</i>	Fungus	Unassigned	tri m 2 allergen [Arthroderma benhamiae]	23894244
181	<i>Arthroderma benhamiae</i>	Fungus	Unassigned	tri m 2 allergen [Arthroderma benhamiae]	23894240



Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
182	<i>Arthroderma benhamiae</i>	Fungus	Unassigned	tri m 4 allergen [Arthroderma benhamiae]	23894232
183	<i>Arthroderma vanbreuseghemii</i>	Fungus	Unassigned	allergen [Arthroderma vanbreuseghemii]	219687753
184	<i>Ascaris lumbricoides</i>	Parasitic roundworm	Unassigned	ABA-1 allergen [Ascaris lumbricoides]	2735104
185	<i>Ascaris lumbricoides</i>	Parasitic roundworm	Unassigned	ABA-1 allergen [Ascaris lumbricoides]	2735100
186	<i>Ascaris lumbricoides</i>	Parasitic roundworm	Unassigned	ABA-1 allergen [Ascaris lumbricoides]	2735118
187	<i>Ascaris lumbricoides</i>	Parasitic roundworm	Unassigned	ABA-1 allergen [Ascaris lumbricoides]	2735114
188	<i>Ascaris lumbricoides</i>	Parasitic roundworm	Unassigned	ABA-1 allergen [Ascaris lumbricoides]	2735112
189	<i>Ascaris lumbricoides</i>	Parasitic roundworm	Unassigned	ABA-1 allergen [Ascaris lumbricoides]	2735110
190	<i>Ascaris lumbricoides</i>	Parasitic roundworm	Unassigned	ABA-1 allergen [Ascaris lumbricoides]	2735108
191	<i>Ascaris lumbricoides</i>	Parasitic roundworm	Unassigned	ABA-1 allergen [Ascaris lumbricoides]	2735106
192	<i>Ascaris lumbricoides</i>	Parasitic roundworm	Unassigned	ABA-1 allergen [Ascaris lumbricoides]	2735102
193	<i>Ascaris lumbricoides</i>	Parasitic roundworm	Unassigned	ABA-1 allergen [Ascaris lumbricoides]	2735098
194	<i>Ascaris lumbricoides</i>	Parasitic roundworm	Unassigned	ABA-1 allergen [Ascaris lumbricoides]	2735096
195	<i>Ascaris lumbricoides</i>	Parasitic roundworm	Unassigned	tropomyosin [Ascaris lumbricoides]	224016002
196	<i>Ascaris suum</i>	Parasitic roundworm	Unassigned	Chain A, The Solution Structure Of ABA-1a Saturated With Oleic Acid	343197079

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
197	<i>Ascaris suum</i>	Parasitic roundworm	Asc s 1	ABA-1 polypeptide precursor (Body fluid allergen 1) (Allergen Asc s 1) (Nematode polypeptide allergen ABA-1) (NPA ABA-1) [Contains: ABA-1B1 repeat unit; ABA-1A4 repeat unit; ABA-1A3 repeat unit; ABA-1A2 repeat unit; ABA-1A1 repeat unit; C-terminal extensi	77416849
198	<i>Ascaris suum</i>	Parasitic roundworm	Asc s 1	major allergen ABA-1=TBA-1 allergen homolog {N-terminal} [Ascaris suum, pseudocoelomic body fluid, Peptide Partial, 68 aa]	299550
199	<i>Aspergillus flavus</i>	Fungus	Unassigned	Allergen Asp fl 1	74665726
200	<i>Aspergillus fumigatus</i>	Fungus	Unassigned	Ribonuclease mitogillin precursor (Major allergen Asp f 1) (Asp f I) (Allergen I/a) (IgE-binding ribotoxin)	54039254
201	<i>Aspergillus fumigatus</i>	Fungus	Asp f 1	major allergen I 18kDa antigen [Aspergillus fumigatus]	9280360
202	<i>Aspergillus fumigatus</i>	Fungus	Asp f 1	Aspf1 allergen [Aspergillus fumigatus]	3021324
203	<i>Aspergillus fumigatus</i>	Fungus	Asp f 10	aspergillopepsin i [Aspergillus fumigatus]	963013
204	<i>Aspergillus fumigatus</i>	Fungus	Asp f 11	PPIase [Aspergillus fumigatus]	5019414
205	<i>Aspergillus fumigatus</i>	Fungus	Asp f 12	Heat shock protein 90 (Heat shock protein hsp1) (65 kDa IgE-binding protein) (Allergen Asp f 12)	83303658
206	<i>Aspergillus fumigatus</i>	Fungus	Asp f 2	Major allergen Asp f 2 precursor (Asp f II)	83300352
207	<i>Aspergillus fumigatus</i>	Fungus	Asp f 2	Asp FII [Aspergillus fumigatus]	664852
208	<i>Aspergillus fumigatus</i>	Fungus	Unassigned	Enolase (2-phosphoglycerate dehydratase) (2-phospho-D-glycerate hydro-lyase) (Allergen Asp f 22)	83288046

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
209	<i>Aspergillus fumigatus</i>	Fungus	Asp f 22	enolase [Aspergillus fumigatus]	13925873
210	<i>Aspergillus fumigatus</i>	Fungus	Asp f 3	peroxisomal-like protein [Aspergillus fumigatus]	2769700
211	<i>Aspergillus fumigatus</i>	Fungus	Unassigned	Allergen Asp f 4 precursor	83300369
212	<i>Aspergillus fumigatus</i>	Fungus	Asp f 4	rAsp f 4 [Aspergillus fumigatus]	3005839
213	<i>Aspergillus fumigatus</i>	Fungus	Unassigned	Superoxide dismutase [Mn], mitochondrial precursor (Allergen Asp f 6)	83305645
214	<i>Aspergillus fumigatus</i>	Fungus	Asp f 6	manganese superoxide dismutase [Aspergillus fumigatus]	1648970
215	<i>Aspergillus fumigatus</i>	Fungus	Asp f 7	Allergen Asp f 7 precursor	83300389
216	<i>Aspergillus fumigatus</i>	Fungus	Unassigned	60S acidic ribosomal protein P2 (Allergen Asp f 8) (AfP2)	83305635
217	<i>Aspergillus fumigatus</i>	Fungus	Asp f 8	rAsp f 8 [Aspergillus fumigatus]	6686524
218	<i>Aspergillus fumigatus</i>	Fungus	Unassigned	Probable glycosidase crf1 precursor (Crh-like protein) (Allergen Asp f 9)	85540942
219	<i>Aspergillus fumigatus</i>	Fungus	Unassigned	60S ribosomal protein L3 (Allergen Asp f 23)	83305621
220	<i>Aspergillus fumigatus</i>	Fungus	Unassigned	large subunit ribosomal protein L3 [Aspergillus fumigatus]	21215170
221	<i>Aspergillus fumigatus</i>	Fungus	Unassigned	cellular serine proteinase [Aspergillus fumigatus]	2143220
222	<i>Aspergillus fumigatus</i> Af293	Fungus	Unassigned	major allergen Asp F2 [Aspergillus fumigatus Af293]	66849502
223	<i>Aspergillus fumigatus</i> Af293	Fungus	Unassigned	allergen Asp F3 [Aspergillus fumigatus Af293]	66845476
224	<i>Aspergillus fumigatus</i> Af293	Fungus	Unassigned	allergen Asp F4 [Aspergillus fumigatus Af293]	66847146
225	<i>Aspergillus niger</i>	Fungus	Asp n 14	beta-xylosidase [Aspergillus niger]	4235093
226	<i>Aspergillus niger</i>	Fungus	Asp n 14	xylosidase [Aspergillus niger]	2181180

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
227	<i>Aspergillus niger</i>	Fungus	Unassigned	serine protease	289172
228	<i>Aspergillus oryzae</i>	Fungus	Asp o 21	Alpha-amylase A type-1/2 precursor (Taka-amylase A) (TAA) (1,4-alpha-D-glucan glucanohydrolase)	94706935
229	<i>Aspergillus oryzae</i>	Fungus	Asp o 13	Oryzin precursor (Alkaline proteinase) (ALP) (Aspergillus proteinase B) (Aspergillopeptidase B)	129235
230	<i>Bacillus lentus</i>		Unassigned	RecName: Full=Subtilisin Savinase; AltName: Full=Alkaline protease	267048
231	<i>Bacillus licheniformis</i>		Unassigned	subtilisin precursor [Bacillus licheniformis]	11127680
232	<i>Bacillus licheniformis</i>		Unassigned	RecName: Full=Subtilisin Carlsberg; Flags: Precursor	135016
233	<i>Bacillus sp.</i>		Unassigned	prepro AprM [Bacillus sp.]	1225905
234	<i>Balanus rostratus</i>		Unassigned	tropomyosin [Balanus rostratus]	125659386
235	<i>Batillus cornutus</i>	Japanese turban shell	Unassigned	tropomyosin [Turbo cornutus]	219806588
236	<i>Batillus cornutus</i>	Japanese turban shell	Unassigned	[Segment 2 of 6] Tropomyosin (Major allergen Tur c 1)	47117351
237	<i>Batillus cornutus</i>	Japanese turban shell	Unassigned	[Segment 1 of 6] Tropomyosin (Major allergen Tur c 1)	47117350
238	<i>Bertholletia excelsa</i>	Brazil nut	Ber e 2	11S globulin [Bertholletia excelsa]	30313867
239	<i>Bertholletia excelsa</i>	Brazil nut	Ber e 1	2S sulfur-rich seed storage protein precursor (Allergen Ber e 1)	112754
240	<i>Bertholletia excelsa</i>	Brazil nut	Unassigned	2S albumin [Bertholletia excelsa]	17713
241	<i>Betula pendula</i>	European white birch	Unassigned	Chain A, Birch Pollen Profilin	157830684
242	<i>Betula pendula</i>	European white birch	Bet v 2	Profilin (Pollen allergen Bet v 2) (Bet v II)	130975

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
243	<i>Betula pendula</i>	European white birch	Bet v 1	Chain A, Birch Pollen Allergen Bet V 1	159162097
244	<i>Betula pendula</i>	European white birch	Bet v 1	Major pollen allergen Bet v 1-A (Allergen Bet v I-A)	114922
245	<i>Betula pendula</i>	European white birch	Bet v 1f/I	Major pollen allergen Bet v 1-F/I (Bet v I-F/I)	1168705
246	<i>Betula pendula</i>	European white birch	Bet v 1.0501	Major pollen allergen Bet v 1-E (Bet v I-E)	1168704
247	<i>Betula pendula</i>	European white birch	Bet v 1	Major pollen allergen Bet v 1-D/H (Bet v I-D/H)	1168703
248	<i>Betula pendula</i>	European white birch	Bet v 1.0301	Major pollen allergen Bet v 1-C (Bet v I-C)	1168702
249	<i>Betula pendula</i>	European white birch	Bet v 1.0201	Major pollen allergen Bet v 1-B (Bet v I-B)	1168701
250	<i>Betula pendula</i>	European white birch	Bet v 1m/n	Major pollen allergen Bet v 1-M/N (Bet v I-M/N)	1168710
251	<i>Betula pendula</i>	European white birch	Bet v 1.0901	Major pollen allergen Bet v 1-K (Bet v I-K)	1168708
252	<i>Betula pendula</i>	European white birch	Bet v 1.0801	Major pollen allergen Bet v 1-J (Bet v I-J)	1168707
253	<i>Betula pendula</i>	European white birch	Bet v 1.0701	Major pollen allergen Bet v 1-G (Bet v I-G)	1168706
254	<i>Betula pendula</i>	European white birch	Bet v 1.1001	Major pollen allergen Bet v 1-L (Bet v I-L)	1168709
255	<i>Betula pendula</i>	European white birch	Unassigned	pollen allergen Betv1, isoform at14 [Betula pendula]	4006947
256	<i>Betula pendula</i>	European white birch	Unassigned	pollen allergen Betv1, isoform at87 [Betula pendula]	4006963
257	<i>Betula pendula</i>	European white birch	Unassigned	Bet v I=major allergen [Betula verrucosa=birch trees, pollen, Peptide Partial, 43 aa]	239734

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
258	<i>Betula pendula</i>	European white birch	Bet v 1	Chain A, Birch Pollen Allergen Bet V 1 Mutant E45s	38492423
259	<i>Betula pendula</i>	European white birch	Bet v 1x	allergen Bet v 1x [Betula pendula]	30908931
260	<i>Betula pendula</i>	European white birch	Bet v 1	Chain A, Birch Pollen Allergen Bet V 1 Mutant N28t, K32q, E45s, P108g	11514622
261	<i>Betula pendula</i>	European white birch	bet v 1 b3	isoallergen bet v 1 b3 [Betula pendula]	4590396
262	<i>Betula pendula</i>	European white birch	Bet v 1 b2	isoallergen Bet v 1 b2 [Betula pendula]	4590394
263	<i>Betula pendula</i>	European white birch	Bet v 1 b1	isoallergen bet v 1 b1 [Betula pendula]	4590392
264	<i>Betula pendula</i>	European white birch	Bet v 1	pollen allergen, Betv1 [Betula pendula]	4376222
265	<i>Betula pendula</i>	European white birch	Bet v 1	pollen allergen, Betv1 [Betula pendula]	4376221
266	<i>Betula pendula</i>	European white birch	Bet v 1	pollen allergen, Betv1 [Betula pendula]	4376220
267	<i>Betula pendula</i>	European white birch	Bet v 1	pollen allergen, Betv1 [Betula pendula]	4376219
268	<i>Betula pendula</i>	European white birch	Bet v 1	pollen allergen, Betv1 [Betula pendula]	4376216
269	<i>Betula pendula</i>	European white birch	Bet v 1	pollen allergen Betv1, isoform at7 [Betula pendula]	4006967
270	<i>Betula pendula</i>	European white birch	Bet v 1	pollen allergen Betv1, isoform at5 [Betula pendula]	4006965
271	<i>Betula pendula</i>	European white birch	Bet v 1	pollen allergen Betv1, isoform at59 [Betula pendula]	4006961
272	<i>Betula pendula</i>	European white birch	Bet v 1	pollen allergen Betv1, isoform at50 [Betula pendula]	4006959

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
273	<i>Betula pendula</i>	European white birch	Bet v 1	pollen allergen Betv1, isoform at45 [Betula pendula]	4006957
274	<i>Betula pendula</i>	European white birch	Bet v 1	pollen allergen Betv1, isoform at42 [Betula pendula]	4006955
275	<i>Betula pendula</i>	European white birch	Bet v 1	pollen allergen Betv1, isoform at37 [Betula pendula]	4006953
276	<i>Betula pendula</i>	European white birch	Bet v 1	pollen allergen Betv1, isoform at10 [Betula pendula]	4006945
277	<i>Betula pendula</i>	European white birch	Bet v 1	pollen allergen Betv1, isoform at8 [Betula pendula]	4006928
278	<i>Betula pendula</i>	European white birch	Bet v 1	pollen allergen Betv1 [Betula pendula]	2564228
279	<i>Betula pendula</i>	European white birch	Bet v 1	pollen allergen Betv1 [Betula pendula]	2564224
280	<i>Betula pendula</i>	European white birch	Bet v 1	pollen allergen Betv1 [Betula pendula]	2564222
281	<i>Betula pendula</i>	European white birch	Bet v 1	pollen allergen Betv1 [Betula pendula]	2564220
282	<i>Betula pendula</i>	European white birch	Bet v 1.2301	major allergen Bet v 1 [Betula pendula]	2414158
283	<i>Betula pendula</i>	European white birch	Bet v 1.3001	pollen allergen Bet v 1 [Betula pendula]	1542873
284	<i>Betula pendula</i>	European white birch	Bet v 1.2901	pollen allergen Bet v 1 [Betula pendula]	1542871
285	<i>Betula pendula</i>	European white birch	Bet v 1.2801	pollen allergen Bet v 1 [Betula pendula]	1542869
286	<i>Betula pendula</i>	European white birch	Bet v 1.2701	pollen allergen Bet v 1 [Betula pendula]	1542867
287	<i>Betula pendula</i>	European white birch	Bet v 1.2601	pollen allergen Bet v 1 [Betula pendula]	1542865

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
288	<i>Betula pendula</i>	European white birch	Bet v 1.2501	pollen allergen Bet v 1 [Betula pendula]	1542863
289	<i>Betula pendula</i>	European white birch	Bet v 1.2401	pollen allergen Bet v 1 [Betula pendula]	1542861
290	<i>Betula pendula</i>	European white birch	Bet v 1.2201	major allergen Bet v 1 [Betula pendula]	1321728
291	<i>Betula pendula</i>	European white birch	Bet v 1.2101	major allergen Bet v 1 [Betula pendula]	1321726
292	<i>Betula pendula</i>	European white birch	Bet v 1.2001	major allergen Bet v 1 [Betula pendula]	1321724
293	<i>Betula pendula</i>	European white birch	Bet v 1.1901	major allergen Bet v 1 [Betula pendula]	1321722
294	<i>Betula pendula</i>	European white birch	Bet v 1.1502	major allergen Bet v 1 [Betula pendula]	1321720
295	<i>Betula pendula</i>	European white birch	Bet v 1.1801	major allergen Bet v 1 [Betula pendula]	1321718
296	<i>Betula pendula</i>	European white birch	Bet v 1.1701	major allergen Bet v 1 [Betula pendula]	1321716
297	<i>Betula pendula</i>	European white birch	Bet v 1.1601	major allergen Bet v 1 [Betula pendula]	1321714
298	<i>Betula pendula</i>	European white birch	Bet v 1	1-Sc1 [Betula pendula]	534910
299	<i>Betula pendula</i>	European white birch	Bet v 1	1 Sc2 [Betula pendula]	534900
300	<i>Betula pendula</i>	European white birch	Bet v 1	1 Sc-3 [Betula pendula]	534898
301	<i>Betula pendula</i>	European white birch	Bet v 1	major pollen allergen Bet v I - European white birch (fragment)	320545
302	<i>Betula pendula</i>	European white birch	Unassigned	major pollen allergen Bet v II - European white birch (fragment)	320546



Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
303	<i>Betula pendula</i>	European white birch	Bet v 3	Calcium-binding allergen Bet v 3 (Bet v III)	1168696
304	<i>Betula pendula</i>	European white birch	Bet v 4	Polcalcin Bet v 4 (Calcium-binding pollen allergen Bet v 4)	14423850
305	<i>Betula pendula</i>	European white birch	Bet v 6.0102	allergenic isoflavone reductase-like protein Bet v 6.0102 [Betula pendula]	10764491
306	<i>Betula pendula</i>	European white birch	Bet v 7	peptidylprolyl isomerase (cyclophilin) [Betula pendula]	21886603
307	<i>Betula platyphylla</i>	Japanese white birch	Unassigned	Bet vI jap3 [Betula platyphylla]	12583685
308	<i>Betula platyphylla</i>	Japanese white birch	Unassigned	Bet vI jap2 [Betula platyphylla]	12583683
309	<i>Betula platyphylla</i>	Japanese white birch	Unassigned	Bet vI jap1 [Betula platyphylla]	12583681
310	<i>Betula sp.</i>	Birch	Unassigned	isoallergen {N-terminal} [birch, pollen, Peptide Partial, 51 aa]	298736
311	<i>Betula sp.</i>	Birch	Unassigned	isoallergen {N-terminal} [birch, pollen, Peptide Partial, 51 aa]	298737
312	<i>Blattella germanica</i>	German cockroach	Bla g 1.0101	major allergen Bla g 1.0101 [Blattella germanica]	4572592
313	<i>Blattella germanica</i>	German cockroach	Bla g 1.02	major allergen Bla g 1.02 [Blattella germanica]	4240395
314	<i>Blattella germanica</i>	German cockroach	Unassigned	Chain A, Crystal Structure Of Bla G 2 Complexed With Fab 4c3	315113421
315	<i>Blattella germanica</i>	German cockroach	Unassigned	Bla g 2 allergen variant [Blattella germanica]	145105726
316	<i>Blattella germanica</i>	German cockroach	Bla g 2	Aspartic protease Bla g 2 precursor (Allergen Bla g II)	1703445
317	<i>Blattella germanica</i>	German cockroach	Unassigned	Chain A, The Structure Of Mutant (N93q) Of Bla G 2	62738637

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
318	<i>Blattella germanica</i>	German cockroach	Unassigned	Bla g 4 isoallergen 2 [Blattella germanica]	194350817
319	<i>Blattella germanica</i>	German cockroach	Unassigned	Bla g 4 isoallergen 1 [Blattella germanica]	194350815
320	<i>Blattella germanica</i>	German cockroach	Unassigned	Bla g 4 allergen [Blattella germanica]	212675308
321	<i>Blattella germanica</i>	German cockroach	Unassigned	Bla g 4 allergen [Blattella germanica]	144952778
322	<i>Blattella germanica</i>	German cockroach	Bla g 4	allergen Bla g 4	1166573
323	<i>Blattella germanica</i>	German cockroach	Unassigned	Bla g 5 variant allergen [Blattella germanica]	144952780
324	<i>Blattella germanica</i>	German cockroach	Unassigned	Glutathione S-transferase (GST class-sigma) (Major allergen Bla g 5)	6225491
325	<i>Blattella germanica</i>	German cockroach	Unassigned	allergen Bla g 6.0301 [Blattella germanica]	82704036
326	<i>Blattella germanica</i>	German cockroach	Unassigned	allergen Bla g 6.0201 [Blattella germanica]	82704034
327	<i>Blattella germanica</i>	German cockroach	Unassigned	allergen Bla g 6.0101 [Blattella germanica]	82704032
328	<i>Blattella germanica</i>	German cockroach	Unassigned	delta class glutathione S-transferase [Blattella germanica]	161137518
329	<i>Blattella germanica</i>	German cockroach	Unassigned	tropomyosin [Blattella germanica]	8101069
330	<i>Blattella germanica</i>	German cockroach	Unassigned	36 kda allergen {peptide 143-115} [Blattella germanica=German cockroaches, Peptide Partial, 25 aa]	544619
331	<i>Blattella germanica</i>	German cockroach	Unassigned	36 kda allergen {peptide 143-111} [Blattella germanica=German cockroaches, Peptide Partial, 20 aa]	544618

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
332	<i>Blomia tropicalis</i>	Mite	Blo t 1	cysteine protease precursor [Blomia tropicalis]	14276828
333	<i>Blomia tropicalis</i>	Mite	Unassigned	Blo t 1.02 [Blomia tropicalis] Manual Entry	2
334	<i>Blomia tropicalis</i>	Mite	Unassigned	Blo t 1 allergen [Blomia tropicalis]	33667928
335	<i>Blomia tropicalis</i>	Mite	Unassigned	group 10 allergen Blo t 10 [Blomia tropicalis]	156938889
336	<i>Blomia tropicalis</i>	Mite	Blo t 11	paramyosin allergen [Blomia tropicalis]	21954740
337	<i>Blomia tropicalis</i>	Mite	Unassigned	Fatty acid-binding protein (Allergen Blo t 13) (Bt6)	14423698
338	<i>Blomia tropicalis</i>	Mite	Unassigned	Blo t 13 allergen [Blomia tropicalis]	37958153
339	<i>Blomia tropicalis</i>	Mite	Unassigned	Blo t 21 allergen [Blomia tropicalis]	111494253
340	<i>Blomia tropicalis</i>	Mite	Unassigned	Blo t 21 allergen [Blomia tropicalis]	111120432
341	<i>Blomia tropicalis</i>	Mite	Unassigned	Blo t 21 allergen [Blomia tropicalis]	111120428
342	<i>Blomia tropicalis</i>	Mite	Unassigned	Blo t 21 allergen [Blomia tropicalis]	111120424
343	<i>Blomia tropicalis</i>	Mite	Unassigned	Blo t 21 allergen [Blomia tropicalis]	111120420
344	<i>Blomia tropicalis</i>	Mite	Unassigned	Blo t 3 allergen [Blomia tropicalis]	33667930
345	<i>Blomia tropicalis</i>	Mite	Unassigned	trypsin [Blomia tropicalis]	25989482
346	<i>Blomia tropicalis</i>	Mite	Unassigned	Chain A, Nmr Solution Structure Of Blo T 5, A Major Mite Allergen From Blomia Tropicalis	160285626
347	<i>Blomia tropicalis</i>	Mite	Unassigned	Blo t 5 allergen [Blomia tropicalis]	111120450
348	<i>Blomia tropicalis</i>	Mite	Unassigned	Blo t 5 allergen [Blomia tropicalis]	111120436
349	<i>Blomia tropicalis</i>	Mite	Blo t 5	major IgE-binding protein Blo t 5 [Blomia tropicalis]	4204917
350	<i>Bombus pennsylvanicus</i>	Bumblebee	Unassigned	RecName: Full=Phospholipase A2; AltName: Full=Phosphatidylcholine 2-acylhydrolase; AltName: Allergen=Bom p 1	47117013

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
351	<i>Bombus pennsylvanicus</i>	Bumblebee	Unassigned	RecName: Full=Venom protease; AltName: Allergen=Bom p 4	75009997
352	<i>Bombus terrestris</i>	Bumblebee	Unassigned	Phospholipase A2 (Phosphatidylcholine 2-acylhydrolase) (Allergen Bom t 1)	14423832
353	<i>Bombus terrestris</i>	Bumblebee	Unassigned	RecName: Full=Venom protease; AltName: Allergen=Bom t 4	313471465
354	<i>Bombyx mori</i>	Silkworm	Unassigned	arginine kinase [Bombyx mori]	204324083
355	<i>Bos taurus</i>	Bovine	Unassigned	alpha S1 casein [Bos taurus]	159793217
356	<i>Bos taurus</i>	Bovine	Unassigned	alpha S1 casein [Bos taurus]	159793201
357	<i>Bos taurus</i>	Bovine	Unassigned	alpha S1 casein [Bos taurus]	159793197
358	<i>Bos taurus</i>	Bovine	Unassigned	alpha-s1-casein	162927
359	<i>Bos taurus</i>	Bovine	Unassigned	alpha-S1-casein	162794
360	<i>Bos taurus</i>	Bovine	Unassigned	alpha-s1-casein precursor	162792
361	<i>Bos taurus</i>	Bovine	Unassigned	alpha-s1-casein	162650
362	<i>Bos taurus</i>	Bovine	Unassigned	alpha-s2-like casein precursor	162929
363	<i>Bos taurus</i>	Bovine	Unassigned	beta-casein A3 [Bos taurus]	459292
364	<i>Bos taurus</i>	Bovine	Unassigned	beta-casein precursor	162931
365	<i>Bos taurus</i>	Bovine	Unassigned	beta-casein	162805
366	<i>Bos taurus</i>	Bovine	Unassigned	beta-casein precursor	162797
367	<i>Bos taurus</i>	Bovine	Unassigned	Allergen Bos d 2 precursor (Dander major allergen BDA20) (Dermal allergen BDA20)	2497701
368	<i>Bos taurus</i>	Bovine	Bos d 3	S100 calcium-binding protein A7 (Allergen Bos d 3) (Dander minor allergen BDA11) (Dermal allergen BDA11) (Calcium-binding protein in amniotic fluid 2) (CAAF2)	2493414

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
369	<i>Bos taurus</i>	Bovine	Unassigned	Alpha-lactalbumin precursor (Lactose synthase B protein) (Allergen Bos d 4)	125996
370	<i>Bos taurus</i>	Bovine	Bos d 4	alpha-lactalbumin [Bos taurus]	295774
371	<i>Bos taurus</i>	Bovine	Unassigned	major allergen beta-lactoglobulin [Bos taurus]	195957138
372	<i>Bos taurus</i>	Bovine	Unassigned	Beta-lactoglobulin precursor (Beta-LG) (Allergen Bos d 5)	125910
373	<i>Bos taurus</i>	Bovine	Unassigned	beta-lactoglobulin	162750
374	<i>Bos taurus</i>	Bovine	Bos d 5	beta-lactoglobulin [Bos taurus]	520
375	<i>Bos taurus</i>	Bovine	Unassigned	RecName: Full=Serum albumin; AltName: Full=BSA; AltName: Allergen=Bos d 6; Flags: Precursor	1351907
376	<i>Bos taurus</i>	Bovine	Unassigned	bovine serum albumin [Bos taurus]	3336842
377	<i>Bos taurus</i>	Bovine	Unassigned	collagen alpha-2(I) chain precursor [Bos taurus]	27806257
378	<i>Bos taurus</i>	Bovine	Unassigned	kappa-casein precursor	162811
379	<i>Bos taurus</i>	Bovine	Unassigned	lactotransferrin [Bos taurus]	30794292
380	<i>Brassica juncea</i>	Mustard	Bra j 1	Allergen Bra j 1-E (Bra j I) [Contains: Allergen Bra j 1-E small chain; Allergen Bra j 1-E large chain]	32363444
381	<i>Brassica napus</i>	Rape	Bra n 1	Napin-3 (Napin BnIII) (Napin nIII) (1.7S seed storage protein) [Contains: Napin-3 small chain; Napin-3 large chain]	75107016
382	<i>Brassica napus</i>	Rape	Unassigned	Polcalcin Bra n 1 (Calcium-binding pollen allergen Bra n 1)	59800143
383	<i>Brassica napus</i>	Rape	Unassigned	recombinant Ib pronapin precursor [Brassica napus]	26985163
384	<i>Brassica napus</i>	Rape	Unassigned	Polcalcin Bra n 2 (Calcium-binding pollen allergen Bra n 2)	59800145

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
385	<i>Brassica napus</i>	Rape	Unassigned	pollen allergen group II (clone 44) - rape	2129802
386	<i>Brassica napus</i>	Rape	Unassigned	pollen allergen group II (clone 42) - rape	2129801
387	<i>Brassica oleracea</i>	Cabbage	Unassigned	LTP Bra o 3 [Brassica oleracea] Manual Entry	1
388	<i>Brassica rapa</i>	Turnip	Unassigned	pollen allergen group II (clone 4) - turnip (fragment)	2129805
389	<i>Brassica rapa subsp. rapa</i>	Turnip	Unassigned	Chitin-binding allergen Bra r 2	32363456
390	<i>Brassica rapa subsp. rapa</i>	Turnip	Unassigned	Polcalcin Bra r 1 (Calcium-binding pollen allergen Bra r 1)	59800144
391	<i>Brassica rapa subsp. rapa</i>	Turnip	Unassigned	Polcalcin Bra r 2 (Calcium-binding pollen allergen Bra r 2)	59800146
392	<i>Candida albicans</i>	Yeast	Cand a 3	29 kDa IgE-binding protein [Candida albicans]	37548637
393	<i>Candida albicans</i>	Yeast	Unassigned	Enolase 1 (2-phosphoglycerate dehydratase) (2-phospho-D-glycerate hydro-lyase)	232054
394	<i>Canis familiaris</i>	Dog	Can f 1	Major allergen Can f 1 precursor (Allergen Dog 1)	3121745
395	<i>Canis familiaris</i>	Dog	Can f 2	precursor Can f II [Canis familiaris]	29292274
396	<i>Canis familiaris</i>	Dog	Can f 2	precursor Can f II [Canis familiaris]	29292272
397	<i>Canis familiaris</i>	Dog	Can f 2	Minor allergen Can f 2 precursor (Allergen Dog 2)	3121746
398	<i>Canis familiaris</i>	Dog	Can f 3	albumin [Canis familiaris]	6687188
399	<i>Canis familiaris</i>	Dog	Can f 3	albumin [Canis familiaris]	3319897
400	<i>Canis familiaris</i>	Dog	Can f 3	albumin [Canis familiaris]	633938
401	<i>Canis familiaris</i>	Dog	Unassigned	allergen Can f 4 precursor [Canis lupus familiaris]	262232390

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
402	<i>Capsicum annuum</i>	Bell pepper	Cap a 2	profilin [Capsicum annuum]	16555785
403	<i>Carica papaya</i>	Papaya	Unassigned	Papain precursor (Papaya proteinase I) (PPI) (Allergen Car p 1)	129614
404	<i>Carpinus betulus</i>	Hornbeam	Unassigned	group 2 Car b 1=isoallergenic variant [Carpinus betulus=hornbeams, pollen, Peptide Recombinant Partial, 80 aa]	1008580
405	<i>Carpinus betulus</i>	Hornbeam	Unassigned	group 1 Car b 1=isoallergenic variant [Carpinus betulus=hornbeams, pollen, Peptide Recombinant Partial, 80 aa]	1008579
406	<i>Carpinus betulus</i>	Hornbeam	Unassigned	group 1 Car b 1=isoallergenic variant [Carpinus betulus=hornbeams, pollen, Peptide Recombinant Partial, 80 aa]	1008578
407	<i>Carpinus betulus</i>	Hornbeam	Unassigned	pollen allergen Car b 1 isoform [Carpinus betulus]	167472839
408	<i>Carpinus betulus</i>	Hornbeam	Unassigned	pollen allergen Car b 1 isoform [Carpinus betulus]	167472841
409	<i>Carpinus betulus</i>	Hornbeam	Unassigned	pollen allergen Car b 1 isoform [Carpinus betulus]	167472843
410	<i>Carpinus betulus</i>	Hornbeam	Unassigned	pollen allergen Car b 1 isoform [Carpinus betulus]	167472837
411	<i>Carpinus betulus</i>	Hornbeam	Unassigned	pollen allergen Car b 1 isoform [Carpinus betulus]	167472845
412	<i>Carpinus betulus</i>	Hornbeam	Unassigned	Car b I=major allergen [Carpinus betulus=hornbeam trees, pollen, Peptide Partial, 40 aa]	239735
413	<i>Carpinus betulus</i>	Hornbeam	Car b 1.0302	pollen allergen Car b 1 [Carpinus betulus]	1545897
414	<i>Carpinus betulus</i>	Hornbeam	Car b 1.0301	pollen allergen Car b 1 [Carpinus betulus]	1545895
415	<i>Carpinus betulus</i>	Hornbeam	Car b 1.0108	pollen allergen Car b 1 [Carpinus betulus]	1545893

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
416	<i>Carpinus betulus</i>	Hornbeam	Car b 1	pollen allergen Car b 1 [ <i>Carpinus betulus</i> ]	1545891
417	<i>Carpinus betulus</i>	Hornbeam	Car b 1.0105	pollen allergen Car b 1 [ <i>Carpinus betulus</i> ]	1545887
418	<i>Carpinus betulus</i>	Hornbeam	Car b 1.0104	pollen allergen Car b 1 [ <i>Carpinus betulus</i> ]	1545879
419	<i>Carpinus betulus</i>	Hornbeam	Car b 1.0104	pollen allergen Car b 1 [ <i>Carpinus betulus</i> ]	1545877
420	<i>Carpinus betulus</i>	Hornbeam	Car b 1.0103	pollen allergen Car b 1 [ <i>Carpinus betulus</i> ]	1545875
421	<i>Carpinus betulus</i>	Hornbeam	Car b 1	Major pollen allergen Car b 1 isoform 2 (Car b I)	730049
422	<i>Carpinus betulus</i>	Hornbeam	Car b 1	Major pollen allergen Car b 1 isoforms 1A and 1B (Car B I)	730048
423	<i>Carpinus betulus</i>	Hornbeam	Car b 1	Car b I [ <i>Carpinus betulus</i> ]	402745
424	<i>Castanea sativa</i>	European chestnut	Unassigned	Cas s 1 pollen allergen [ <i>Castanea sativa</i> ]	212291468
425	<i>Castanea sativa</i>	European chestnut	Unassigned	Cas s 1 pollen allergen [ <i>Castanea sativa</i> ]	212291464
426	<i>Castanea sativa</i>	European chestnut	Unassigned	Cas s 1 pollen allergen [ <i>Castanea sativa</i> ]	212291466
427	<i>Castanea sativa</i>	European chestnut	Cas s 1	ypr10 [ <i>Castanea sativa</i> ]	16555781
428	<i>Castanea sativa</i>	European chestnut	Unassigned	class I chitinase isoform 2 [ <i>Castanea sativa</i> ]	307159110
429	<i>Castanea sativa</i>	European chestnut	Cas s 5	chitinase Ib [ <i>Castanea sativa</i> ]	1359600
430	<i>Cavia porcellus</i>	Domestic guinea pig	Cav p 1	Major urinary protein (MUP) (Allergen Cav p 1)	32469617



Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
431	<i>Cavia porcellus</i>	Domestic guinea pig	Unassigned	allergen lipocalin Cav p 2.0101 precursor [Cavia porcellus]	325910590
432	<i>Cavia porcellus</i>	Domestic guinea pig	Unassigned	allergen lipocalin Cav p 3.0101 precursor [Cavia porcellus]	325910592
433	<i>Chamaecyparis obtusa</i>	Japanese cypress	Unassigned	Major pollen allergen Cha o 1 precursor	9087163
434	<i>Chamaecyparis obtusa</i>	Japanese cypress	Unassigned	pollen allergen [Chamaecyparis obtusa]	114841683
435	<i>Chamaecyparis obtusa</i>	Japanese cypress	Unassigned	Polygalacturonase precursor (PG) (Pectinase) (Major pollen allergen Cha o 2)	47606004
436	<i>Charybdis feriatus</i>	Crab	Cha f 1	Tropomyosin (Allergen Cha f 1) (Cha f I)	14285800
437	<i>Chenopodium album</i>	Pigweed	Unassigned	Pollen allergen Che a 1 precursor	47605504
438	<i>Chenopodium album</i>	Pigweed	Unassigned	Che a 2 pollen allergen [Chenopodium album]	238886048
439	<i>Chenopodium album</i>	Pigweed	Che a 2	pollen allergen Che a 2 [Chenopodium album]	29465666
440	<i>Chenopodium album</i>	Pigweed	Che a 3	pollen allergen Che a 3 [Chenopodium album]	29465668
441	<i>Chionoecetes opilio</i>	Snow Crab	Unassigned	RecName: Full=Tropomyosin; AltName: Full=Tropomyosin, slow-tonic isoform; Short=Tm-Chio-tonic	308191588
442	<i>Chironomus kiiensis</i>	Midge	Unassigned	Tropomyosin (Allergen Chi k 10)	42559556
443	<i>Chironomus thummi thummi</i>	Midge	Chi t 1.02	Globin CTT-IV precursor	121227
444	<i>Chironomus thummi thummi</i>	Midge	Chi t 1.01	Globin CTT-III precursor (Erythrocrutorin III)	121219
445	<i>Chironomus thummi thummi</i>	Midge	Chi t 2	Globin CTT-I/CTT-IA precursor (Erythrocrutorin)	2506460

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
446	<i>Chironomus thummi thummi</i>	Midge	Chi t 3	Globin CTT-II beta precursor	1707908
447	<i>Chironomus thummi thummi</i>	Midge	Chi t 4	Globin CTT-III A	121256
448	<i>Chironomus thummi thummi</i>	Midge	Chi t 5	Globin CTT-VI precursor	2506461
449	<i>Chironomus thummi thummi</i>	Midge	Chi t 7	Globin CTT-VIIB-5/CTT-VIIB-9 precursor	56405054
450	<i>Chironomus thummi thummi</i>	Midge	Chi t 7	Globin CTT-VIIB-4 precursor (Erythrocrurin)	56405052
451	<i>Chironomus thummi thummi</i>	Midge	Chi t 7	Globin CTT-VIIB-7 precursor	121249
452	<i>Chironomus thummi thummi</i>	Midge	Chi t 7	Globin CTT-VIIB-6 precursor	121248
453	<i>Chironomus thummi thummi</i>	Midge	Chi t 7	Globin CTT-VIIB-3 precursor	121244
454	<i>Chironomus thummi thummi</i>	Midge	Chi t 8	Globin CTT-VIII	121237
455	<i>Chironomus thummi thummi</i>	Midge	Chi t 9	Globin CTT-X	121259
456	<i>Citrus limon</i>	Lemon	Unassigned	Nonspecific lipid-transfer protein (LTP) (Allergen Cit 1 3)	52783176
457	<i>Citrus sinensis</i>	Navel orange	Unassigned	Germin-like protein (Allergen Cit s 1)	52782810
458	<i>Citrus sinensis</i>	Navel orange	Unassigned	RecName: Full=Profilin; AltName: Allergen=Cit s 2	261260074
459	<i>Citrus sinensis</i>	Navel orange	Cit s 3	lipid transfer protein [Citrus sinensis]	50199132
460	<i>Citrus sinensis</i>	Navel orange	Unassigned	Nonspecific lipid-transfer protein (LTP) (Allergen Cit s 3.0101)	52783177
461	<i>Cladosporium cladosporioides</i>		Unassigned	vacuolar serine protease [Cladosporium cladosporioides]	148361511

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
462	<i>Cochliobolus lunatus</i>		Unassigned	Cytochrome c	20137645
463	<i>Cochliobolus lunatus</i>		Cur l 2.01	enolase [Curvularia lunata]	14585753
464	<i>Coprinus comatus</i>	Shaggy mane	Cop c 1	Cop c1 allergen [Coprinus comatus]	4538529
465	<i>Corylus avellana</i>	European hazelnut	Cor a I	Major pollen allergen Cor a 1 isoforms 5, 6, 11 and 16 (Cor a I)	584968
466	<i>Corylus avellana</i>	European hazelnut	Cor a 1.0404	major allergen variant Cor a 1.0404 [Corylus avellana]	11762106
467	<i>Corylus avellana</i>	European hazelnut	Cor a 1.0403	major allergen variant Cor a 1.0403 [Corylus avellana]	11762104
468	<i>Corylus avellana</i>	European hazelnut	Cor a 1.0402	major allergen variant Cor a 1.0402 [Corylus avellana]	11762102
469	<i>Corylus avellana</i>	European hazelnut	Cor a 1.0401	major allergen Cor a 1.0401 [Corylus avellana]	5726304
470	<i>Corylus avellana</i>	European hazelnut	Cor a 1.0301	major allergen Cor a 1 [Corylus avellana]	1321733
471	<i>Corylus avellana</i>	European hazelnut	Cor a 1.0201	major allergen Cor a 1 [Corylus avellana]	1321731
472	<i>Corylus avellana</i>	European hazelnut	Cor a 1.0102	major allergen [Corylus avellana]	22690
473	<i>Corylus avellana</i>	European hazelnut	Cor a 1.0104	major allergen [Corylus avellana]	22686
474	<i>Corylus avellana</i>	European hazelnut	Cor a 1.0103	major allergen [Corylus avellana]	22684
475	<i>Corylus avellana</i>	European hazelnut	Cor a 10	putative luminal binding protein [Corylus avellana]	10944737
476	<i>Corylus avellana</i>	European hazelnut	Cor a 11	48-kDa glycoprotein precursor [Corylus avellana]	19338630
477	<i>Corylus avellana</i>	European hazelnut	Unassigned	2S albumin [Corylus avellana]	226437844

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
478	<i>Corylus avellana</i>	European hazelnut	Cor a 2	minor allergen hazelnut profilin [ <i>Corylus avellana</i> ]	12659208
479	<i>Corylus avellana</i>	European hazelnut	Cor a 2	minor allergen hazelnut profilin [ <i>Corylus avellana</i> ]	12659206
480	<i>Corylus avellana</i>	European hazelnut	Cor a 8	lipid transfer protein precursor [ <i>Corylus avellana</i> ]	13507262
481	<i>Corylus avellana</i>	European hazelnut	Cor a 9	11S globulin-like protein [ <i>Corylus avellana</i> ]	18479082
482	<i>Corylus avellana</i>	European hazelnut	Unassigned	oleosin [ <i>Corylus avellana</i> ]	29170509
483	<i>Crangon crangon</i>		Unassigned	tropomyosin [ <i>Crangon crangon</i> ]	238477263
484	<i>Crangon crangon</i>		Unassigned	arginine kinase [ <i>Crangon crangon</i> ]	238477265
485	<i>Crangon crangon</i>		Unassigned	sarcoplasmic calcium-binding protein [ <i>Crangon crangon</i> ]	238477327
486	<i>Crangon crangon</i>		Unassigned	myosin light chain [ <i>Crangon crangon</i> ]	238477331
487	<i>Crangon crangon</i>		Unassigned	troponin C [ <i>Crangon crangon</i> ]	238477333
488	<i>Crangon crangon</i>		Unassigned	triosephosphate isomerase [ <i>Crangon crangon</i> ]	238477329
489	<i>Crassostrea gigas</i>	American oyster	Unassigned	tropomyosin [ <i>Crassostrea gigas</i> ]	219806594
490	<i>Crassostrea gigas</i>	American oyster	Unassigned	tropomyosin [ <i>Crassostrea gigas</i> ]	15419048
491	<i>Crassostrea virginica</i>	Eastern oyster	Unassigned	tropomyosin [ <i>Crassostrea virginica</i> ]	3668408
492	<i>Crocus sativus</i>	Saffron crocus	Unassigned	profilin [ <i>Crocus sativus</i> ]	58700651
493	<i>Cryptomeria japonica</i>	Japanese cedar	Unassigned	class IV chitinase [ <i>Cryptomeria japonica</i> ]	56550550
494	<i>Cryptomeria japonica</i>	Japanese cedar	Unassigned	Cry j IB precursor [ <i>Cryptomeria japonica</i> ]	493634
495	<i>Cryptomeria japonica</i>	Japanese cedar	Cry j 1	Sugi basic protein precursor (SBP) (Major allergen Cry j 1) (Cry j I)	1173367

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
496	<i>Cryptomeria japonica</i>	Japanese cedar	Cry j 1	Cry j 1 precursor [Cryptomeria japonica]	19570315
497	<i>Cryptomeria japonica</i>	Japanese cedar	Unassigned	allergen Cry j 2 [Cryptomeria japonica]	123299282
498	<i>Cryptomeria japonica</i>	Japanese cedar	Unassigned	pollen allergen [Cryptomeria japonica]	114841671
499	<i>Cryptomeria japonica</i>	Japanese cedar	Unassigned	pollen allergen [Cryptomeria japonica]	114841665
500	<i>Cryptomeria japonica</i>	Japanese cedar	Unassigned	pollen allergen [Cryptomeria japonica]	114841663
501	<i>Cryptomeria japonica</i>	Japanese cedar	Unassigned	pollen allergen [Cryptomeria japonica]	114841657
502	<i>Cryptomeria japonica</i>	Japanese cedar	Unassigned	pollen allergen [Cryptomeria japonica]	114841653
503	<i>Cryptomeria japonica</i>	Japanese cedar	Unassigned	pollen allergen [Cryptomeria japonica]	114841641
504	<i>Cryptomeria japonica</i>	Japanese cedar	Unassigned	pollen allergen [Cryptomeria japonica]	114841635
505	<i>Cryptomeria japonica</i>	Japanese cedar	Unassigned	pollen allergen [Cryptomeria japonica]	114841629
506	<i>Cryptomeria japonica</i>	Japanese cedar	Unassigned	pollen allergen [Cryptomeria japonica]	114841617
507	<i>Cryptomeria japonica</i>	Japanese cedar	Unassigned	pollen allergen [Cryptomeria japonica]	114841607
508	<i>Cryptomeria japonica</i>	Japanese cedar	Cry j 2	allergen Cry j 2 [Cryptomeria japonica]	24898908
509	<i>Cryptomeria japonica</i>	Japanese cedar	Cry j 2	allergen Cry j 2 [Cryptomeria japonica]	24898906
510	<i>Cryptomeria japonica</i>	Japanese cedar	Cry j 2	allergen Cry j 2 [Cryptomeria japonica]	24898904
511	<i>Cryptomeria japonica</i>	Japanese cedar	Cry j 2	Polygalacturonase precursor (PG) (Pectinase) (Major pollen allergen Cry j 2) (Cry j II)	1171004
512	<i>Cryptomeria japonica</i>	Japanese cedar	Unassigned	isoflavone reductase-like protein CJP-6 [Cryptomeria japonica]	19847822
513	<i>Cryptomeria japonica</i>	Japanese cedar	Unassigned	pollen allergen CJP-8 [Cryptomeria japonica]	291621332
514	<i>Cryptomeria japonica</i>	Japanese cedar	Unassigned	pollen allergen CPA63 [Cryptomeria japonica]	293329689
515	<i>Cryptomeria japonica</i>	Japanese cedar	Unassigned	thaumatin-like protein [Cryptomeria japonica]	139002766

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
516	<i>Cucumis melo</i>	Muskmelon	Unassigned	Cucumisin precursor (Allergen Cuc m 1)	71153243
517	<i>Cucumis melo</i>	Muskmelon	Cuc m 2	profilin [ <i>Cucumis melo</i> ]	58263793
518	<i>Cucumis melo</i>	Muskmelon	Cuc m 2	profilin [ <i>Cucumis melo</i> ]	31559374
519	<i>Cucumis melo</i>	Muskmelon	Cuc m 3	[Segment 3 of 3] Pathogenesis-related protein (PR-1) (Allergen Cuc m 3)	46396598
520	<i>Cucumis melo</i>	Muskmelon	Cuc m 3	[Segment 2 of 3] Pathogenesis-related protein (PR-1) (Allergen Cuc m 3)	46396597
521	<i>Cucumis melo</i>	Muskmelon	Cuc m 3	[Segment 1 of 3] Pathogenesis-related protein (PR-1) (Allergen Cuc m 3)	46396596
522	<i>Cucumis melo</i> var. <i>inodorus</i>	Muskmelon	Unassigned	pathogen-related protein 1 [ <i>Cucumis melo</i> var. <i>inodorus</i> ]	171464770
523	<i>Cucumis melo</i> var. <i>reticulatus</i>	Netted muskmelon	Cuc m 2	profilin [ <i>Cucumis melo</i> var. <i>reticulatus</i> ]	57021110
524	<i>Cupressus arizonica</i>	Arizona Cypress	Unassigned	Major pollen allergen Cup a 1	9087167
525	<i>Cupressus arizonica</i>	Arizona Cypress	Unassigned	major allergen Cup a 1 [ <i>Cupressus arizonica</i> ]	118197955
526	<i>Cupressus arizonica</i>	Arizona Cypress	Cup a 1	putative allergen Cup a 1 [ <i>Cupressus arizonica</i> ]	19069497
527	<i>Cupressus arizonica</i>	Arizona Cypress	Unassigned	Cup a 3 protein [ <i>Cupressus arizonica</i> ]	9929163
528	<i>Cupressus arizonica</i>	Arizona Cypress	Unassigned	putative Cup a 4 allergen [ <i>Hesperocyparis arizonica</i> ]	261865475
529	<i>Cupressus sempervirens</i>	Mediterranean Cypress	Cup s 1.0105	Cup s 1 pollen allergen precursor [ <i>Cupressus sempervirens</i> ]	8101719
530	<i>Cupressus sempervirens</i>	Mediterranean Cypress	Cup s 1.0104	Cup s 1 pollen allergen precursor [ <i>Cupressus sempervirens</i> ]	8101717
531	<i>Cupressus sempervirens</i>	Mediterranean Cypress	Cup s 1.0103	Cup s 1 pollen allergen precursor [ <i>Cupressus sempervirens</i> ]	8101715
532	<i>Cupressus sempervirens</i>	Mediterranean Cypress	Cup s 1.0102	Cup s 1 pollen allergen precursor [ <i>Cupressus sempervirens</i> ]	8101713

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
533	<i>Cupressus sempervirens</i>	Mediterranean Cypress	Cup s 1.0101	Cup s 1 pollen allergen precursor [Cupressus sempervirens]	8101711
534	<i>Cupressus sempervirens</i>	Mediterranean Cypress	Unassigned	PR5 allergen Cup s 3.2 precursor [Cupressus sempervirens]	38456228
535	<i>Cupressus sempervirens</i>	Mediterranean Cypress	Unassigned	PR5 allergen Cup s 3.3 precursor [Cupressus sempervirens]	38456230
536	<i>Cynodon dactylon</i>	Bermuda grass	Cyn d 1.0203	acidic Cyn d 1 isoallergen isoform 4 precursor [Cynodon dactylon]	16076697
537	<i>Cynodon dactylon</i>	Bermuda grass	Cyn d 1	acidic Cyn d 1 isoallergen isoform 3 precursor [Cynodon dactylon]	16076695
538	<i>Cynodon dactylon</i>	Bermuda grass	Cyn d 1.0202	acidic Cyn d 1 isoallergen isoform 2 precursor [Cynodon dactylon]	16076693
539	<i>Cynodon dactylon</i>	Bermuda grass	Cyn d 1.0201	acidic allergen Cyn d 1 precursor [Cynodon dactylon]	15384338
540	<i>Cynodon dactylon</i>	Bermuda grass	Cyn d 1	Major pollen allergen Cyn d 1	14423757
541	<i>Cynodon dactylon</i>	Bermuda grass	Cyn d 1.0204	acidic Cyn d 1 isoallergen isoform 1 precursor [Cynodon dactylon]	10314021
542	<i>Cynodon dactylon</i>	Bermuda grass	Cyn d 1	Cyn d 1b isoallergen {N-terminal} [Cynodon dactylon=Bermuda grass, pollen, Peptide Partial, 34 aa]	691726
543	<i>Cynodon dactylon</i>	Bermuda grass	Cyn d 1	major allergen Cyn d I=29 kda polypeptide {N-terminal} [Cynodon dactylon=Bermuda grass, pollen, Peptide Partial, 38 aa]	451275
544	<i>Cynodon dactylon</i>	Bermuda grass	Cyn d 1	major allergen Cyn d I=34 kda polypeptide {N-terminal} [Cynodon dactylon=Bermuda grass, pollen, Peptide Partial, 25 aa]	451274
545	<i>Cynodon dactylon</i>	Bermuda grass	Cyn d 12	profilin 1 [Cynodon dactylon]	2154730

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
546	<i>Cynodon dactylon</i>	Bermuda grass	Cyn d 7	calcium-binding pollen allergen [Cynodon dactylon]	1871507
547	<i>Cynodon dactylon</i>	Bermuda grass	Unassigned	B4 protein allergen [Cynodon dactylon]	1247375
548	<i>Cynodon dactylon</i>	Bermuda grass	Unassigned	B1 protein allergen [Cynodon dactylon]	1247373
549	<i>Cyprinus carpio</i>	Carp	Unassigned	parvalbumin [Cyprinus carpio]	17977827
550	<i>Cyprinus carpio</i>	Carp	Unassigned	parvalbumin [Cyprinus carpio]	17977825
551	<i>Dactylis glomerata</i>	Orchard grass	Unassigned	group 1 allergen Dac g 1.01 precursor [Dactylis glomerata]	33149333
552	<i>Dactylis glomerata</i>	Orchard grass	Dac g 1	unnamed protein product [Dactylis glomerata]	18093991
553	<i>Dactylis glomerata</i>	Orchard grass	Dac g 2	pollen allergen (group II) [Dactylis glomerata]	4007040
554	<i>Dactylis glomerata</i>	Orchard grass	Dac g 2	allergen Dac g II	1093120
555	<i>Dactylis glomerata</i>	Orchard grass	Unassigned	Pollen allergen Dac g 3 (Dac g III)	14423759
556	<i>Dactylis glomerata</i>	Orchard grass	Unassigned	[Segment 4 of 4] Major pollen allergen Dac g 4	32363467
557	<i>Dactylis glomerata</i>	Orchard grass	Unassigned	[Segment 3 of 4] Major pollen allergen Dac g 4	32363466
558	<i>Dactylis glomerata</i>	Orchard grass	Unassigned	[Segment 2 of 4] Major pollen allergen Dac g 4	32363465
559	<i>Dactylis glomerata</i>	Orchard grass	Unassigned	[Segment 1 of 4] Major pollen allergen Dac g 4	32363464
560	<i>Dactylis glomerata</i>	Orchard grass	Dac g 5	unnamed protein product [Dactylis glomerata]	18093971
561	<i>Dactylis glomerata</i>	Orchard grass	Dac g 5	group 5 allergen precursor [Dactylis glomerata]	14423124
562	<i>Daucus carota</i>	Carrot	Unassigned	PRP-like protein [Daucus carota]	302379159



Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
563	<i>Daucus carota</i>	Carrot	Unassigned	PRP-like protein [Daucus carota]	302379157
564	<i>Daucus carota</i>	Carrot	Unassigned	PRP-like protein [Daucus carota]	302379155
565	<i>Daucus carota</i>	Carrot	Unassigned	PRP-like protein [Daucus carota]	302379153
566	<i>Daucus carota</i>	Carrot	Unassigned	PRP-like protein [Daucus carota]	302379151
567	<i>Daucus carota</i>	Carrot	Unassigned	PRP-like protein [Daucus carota]	302379149
568	<i>Daucus carota</i>	Carrot	Dau c 1.0301	PRP-like protein [Daucus carota]	302379147
569	<i>Daucus carota</i>	Carrot	Dau c 1.0105	Major allergen Dau c 1 (CR16) (Pathogenesis-related protein Gea20)	8928058
570	<i>Daucus carota</i>	Carrot	Unassigned	pathogenesis-related protein-like protein 1 [Daucus carota]	19912791
571	<i>Daucus carota</i>	Carrot	Dau c 1.0201	major allergen isoform Dau c 1.0201 [Daucus carota]	18652047
572	<i>Daucus carota</i>	Carrot	Dau c 1.0104	major allergen [Daucus carota]	2154734
573	<i>Daucus carota</i>	Carrot	Dau c 1.0103	major allergen [Daucus carota]	2154732
574	<i>Daucus carota</i>	Carrot	Dau c 1.0102	cr16 [Daucus carota]	1663522
575	<i>Daucus carota</i>	Carrot	Dau c 1.0101	pathogenesis-related protein	1335877
576	<i>Daucus carota</i>	Carrot	Unassigned	RecName: Full=Profilin; AltName: Full=Minor pollen allergen Dau c 4; AltName: Allergen=Dau c 4	47606043
577	<i>Davidiella tassiana</i>	Fungus	Unassigned	Aldehyde dehydrogenase (ALDDH) (Allergen Cla h 10) (Cla h 3) (Cla h III)	108935817
578	<i>Davidiella tassiana</i>	Fungus	Cla h 5	60S acidic ribosomal protein P2 (Minor allergen Cla h 4) (Cla h IV)	21542440
579	<i>Davidiella tassiana</i>	Fungus	Cla h 5	60S acidic ribosomal protein P2 (Allergen Cla h 3) (Cla h III)	1173074

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
580	<i>Davidiella tassiana</i>	Fungus	Cla h 6	Enolase (2-phosphoglycerate dehydratase) (2-phospho-D-glycerate hydro-lyase) (Allergen Cla h 6) (Cla h VI)	6015094
581	<i>Davidiella tassiana</i>	Fungus	Cla h 6	enolase; phosphopyruvate hydratase [Davidiella tassiana]	467660
582	<i>Davidiella tassiana</i>	Fungus	Cla h 7	Minor allergen Cla h 5 (Cla h V)	1168970
583	<i>Davidiella tassiana</i>	Fungus	Unassigned	Probable NADP-dependent mannitol dehydrogenase (MtDH) (Mannitol 2-dehydrogenase [NADP+]) (Allergen Cla h 8)	85701146
584	<i>Davidiella tassiana</i>	Fungus	Unassigned	vacuolar serine protease [Davidiella tassiana]	60116876
585	<i>Davidiella tassiana</i>	Fungus	Unassigned	Heat shock 70 kDa protein (Allergen Cla h 4) (Cla h IV)	729764
586	<i>Davidiella tassiana</i>	Fungus	Unassigned	hydrophobin [Davidiella tassiana]	22796153
587	<i>Davidiella tassiana</i>	Fungus	Unassigned	putative nuclear transport factor 2 [Davidiella tassiana]	21748151
588	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Chain A, Solution Structure Of Der F 13, Group 13 Allergen From House Dust Mites	99031759
589	<i>Dermatophagoides farinae</i>	House dust mite	Der f 16	gelsolin-like allergen Der f 16 [Dermatophagoides farinae]	21591547
590	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Der f 1 allergen [Dermatophagoides farinae]	37958161
591	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Der f 1 allergen [Dermatophagoides farinae]	156106765
592	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Der f 1 allergen precursor [Dermatophagoides farinae]	76097507

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
593	<i>Dermatophagoides farinae</i>	House dust mite	Der f 1	Der f 1 allergen preproenzyme [Dermatophagoides farinae]	27530349
594	<i>Dermatophagoides farinae</i>	House dust mite	Der f 1	Major mite fecal allergen Der f 1 precursor (Der f I)	730035
595	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Tropomyosin (Allergen Der f 10) (Mag44)	42559584
596	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Paramyosin (Allergen Der f 11) (Antigen Df642)	42559514
597	<i>Dermatophagoides farinae</i>	House dust mite	Der f 14	Mag3 [Dermatophagoides farinae]	1545803
598	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Allergen Mag	729979
599	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Der f 2 allergen [Dermatophagoides farinae]	37958157
600	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	group 2 allergen [Dermatophagoides farinae]	256631558
601	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Der f 2 allergen [Dermatophagoides farinae]	218203834
602	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Der f 2 allergen [Dermatophagoides farinae]	156480837
603	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Der f 2 allergen precursor [Dermatophagoides farinae]	76097511
604	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	mite allergen Der f 2 [Dermatophagoides farinae]	55859466
605	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Der f 2 [Dermatophagoides farinae]	55859468
606	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	mite allergen Der f 2 [Dermatophagoides farinae]	55859470
607	<i>Dermatophagoides farinae</i>	House dust mite	Der f 2	major Der f 2 isoform [Dermatophagoides farinae]	17978844

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
608	<i>Dermatophagoides farinae</i>	House dust mite	Der f 2	Der f II [Dermatophagoides farinae]	546852
609	<i>Dermatophagoides farinae</i>	House dust mite	Der f 2	mite allergen Der f II precursor [Dermatophagoides farinae]	217308
610	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Der f 3 allergen [Dermatophagoides farinae]	218203818
611	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Der f 3 allergen [Dermatophagoides farinae]	218203816
612	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Der f 3 allergen precursor [Dermatophagoides farinae]	163638970
613	<i>Dermatophagoides farinae</i>	House dust mite	Der f 3	Mite allergen Der f 3 precursor (Der f III)	2507248
614	<i>Dermatophagoides farinae</i>	House dust mite	Der f 3	Der f 3 mite allergen	1314736
615	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Der f 6 allergen [Dermatophagoides farinae]	218203828
616	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Der f 6 allergen [Dermatophagoides farinae]	218203826
617	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	DF5=allergen {N-terminal} [Dermatophagoides farinae=mites, Peptide Partial, 20 aa]	404371
618	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Mite allergen Der f 6 precursor (Der f VI) (DF5)	14424450
619	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Der f 7 allergen [Dermatophagoides farinae]	218203832
620	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Der f 7 allergen [Dermatophagoides farinae]	37958165
621	<i>Dermatophagoides farinae</i>	House dust mite	Der f 7	Mite allergen Der f 7 precursor (Der f VII)	2498299
622	<i>Dermatophagoides farinae</i>	House dust mite	Der f 18	60 kDa allergen Der f 18p [Dermatophagoides farinae]	27550039

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
623	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	group 21 allergen [Dermatophagoides farinae]	140089326
624	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	group 21 allergen [Dermatophagoides farinae]	140089324
625	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	group 21 allergen [Dermatophagoides farinae]	140089322
626	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	group 21 allergen [Dermatophagoides farinae]	140089320
627	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	group 21 allergen [Dermatophagoides farinae]	140089316
628	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	group 21 allergen [Dermatophagoides farinae]	140089314
629	<i>Dermatophagoides farinae</i>	House dust mite	Unassigned	Der f 5.02 allergen [Dermatophagoides farinae]	60679572
630	<i>Dermatophagoides microceras</i>	House dust mite	Der m 1	Major mite fecal allergen Der m 1 (Der m I)	127205
631	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	major house dust allergen [Dermatophagoides pteronyssinus]	387592
632	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	Der p 1 allergen precursor [Dermatophagoides pteronyssinus]	256095986
633	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	Der p 1 allergen precursor [Dermatophagoides pteronyssinus]	195933901
634	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	Chain B, C2 Crystal Form Of Mite Allergen Der P 1	223365887
635	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	Der p 1 allergen [Dermatophagoides pteronyssinus]	157696052
636	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	cysteine protease [Dermatophagoides pteronyssinus]	1460058
637	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	Chain B, Crystal Structure Of Mature And Fully Active Der P 1 Allergen	83754033

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
638	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 1	Major mite fecal allergen Der p 1 precursor (Der p I)	730036
639	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	Der p 1 allergen [Dermatophagoides pteronyssinus]	61608445
640	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 1	unnamed protein product [Dermatophagoides pteronyssinus]	21725580
641	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 1	unnamed protein product [Dermatophagoides pteronyssinus]	21725578
642	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 1	unnamed protein product [Dermatophagoides pteronyssinus]	21725576
643	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 1	unnamed protein product [Dermatophagoides pteronyssinus]	21725574
644	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 1	unnamed protein product [Dermatophagoides pteronyssinus]	21725572
645	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 1	unnamed protein product [Dermatophagoides pteronyssinus]	21725570
646	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 1	unnamed protein product [Dermatophagoides pteronyssinus]	21725568
647	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 1	unnamed protein product [Dermatophagoides pteronyssinus]	21725566
648	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 1	unnamed protein product [Dermatophagoides pteronyssinus]	21725564
649	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 1	unnamed protein product [Dermatophagoides pteronyssinus]	21725562
650	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 1	unnamed protein product [Dermatophagoides pteronyssinus]	21725560
651	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	tropomyosin [Dermatophagoides pteronyssinus]	208970286
652	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	tropomyosin [Dermatophagoides pteronyssinus]	80553470

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
653	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 10	tropomyosin [Dermatophagoides pteronyssinus]	2440053
654	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 10	tropomyosin [Dermatophagoides pteronyssinus]	2353266
655	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 11	HDM allergen [Dermatophagoides pteronyssinus]	37778944
656	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	Der p 13 allergen [Dermatophagoides pteronyssinus]	302035350
657	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 14	group 14 allergen protein [Dermatophagoides pteronyssinus]	20385544
658	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	Der p 2 allergen precursor [Dermatophagoides pteronyssinus]	256095984
659	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	Der p 2 allergen [Dermatophagoides pteronyssinus]	164415595
660	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	Chain A, Tertiary Structure Of The Major House Dust Mite Allergen Der P 2, Nmr, 10 Structures	157829757
661	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	group 2 allergen Der p 2 [Dermatophagoides pteronyssinus]	110560872
662	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	Der p 2 allergen precursor [Dermatophagoides pteronyssinus]	99644635
663	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	Der p 2 allergen precursor [Dermatophagoides pteronyssinus]	76097509
664	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 2	unnamed protein product [Dermatophagoides pteronyssinus]	21725604
665	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 2	unnamed protein product [Dermatophagoides pteronyssinus]	21725602
666	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 2	unnamed protein product [Dermatophagoides pteronyssinus]	21725600
667	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 2	unnamed protein product [Dermatophagoides pteronyssinus]	21725596

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
668	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 2	unnamed protein product [Dermatophagoides pteronyssinus]	21725594
669	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 2	unnamed protein product [Dermatophagoides pteronyssinus]	21725592
670	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 2	unnamed protein product [Dermatophagoides pteronyssinus]	21725590
671	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 2	unnamed protein product [Dermatophagoides pteronyssinus]	21725588
672	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 2	unnamed protein product [Dermatophagoides pteronyssinus]	21725586
673	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 2	unnamed protein product [Dermatophagoides pteronyssinus]	21725584
674	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 2	unnamed protein product [Dermatophagoides pteronyssinus]	21725582
675	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 2	Chain A, X-Ray Structure Of Der P 2, The Major House Dust Mite Allergen	21465915
676	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 2	Mite group 2 allergen Der p 2 precursor (Der p II) (DPX)	1352237
677	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	allergen precursor [Dermatophagoides pteronyssinus]	85687540
678	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 3	Der p 3 allergen	511476
679	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	Alpha-amylase (Allergen Der p 4) (Der p IV)	1351935
680	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 4	alpha-amylase [Dermatophagoides pteronyssinus]	5059162
681	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 5	Mite allergen Der p 5 (Der P V) (IgE-binding allergen)	1352238
682	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 5	unnamed protein product [Dermatophagoides pteronyssinus]	28798085



Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
683	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 5	Der p V allergen [Dermatophagoides pteronyssinus]	913285
684	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 6	Mite allergen Der p 6 (Der p VI) (DP5)	1352239
685	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	Mite allergen Der p 7 precursor (Der p VII)	1352240
686	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Der p 7	unnamed protein product [Dermatophagoides pteronyssinus]	10189811
687	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	Glutathione S-transferase (GST class-mu) (Major allergen Der p 8) (P dp 15)	1170095
688	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	glutathione transferase mu class Dp7019C10 [Dermatophagoides pteronyssinus]	60920878
689	<i>Dermatophagoides pteronyssinus</i>	House dust mite	Unassigned	group 18 allergen protein [Dermatophagoides pteronyssinus]	67975085
690	<i>Dermatophagoides siboney</i>	House dust mite	Unassigned	Der s 2 a allergen [Dermatophagoides siboney]	86450747
691	<i>Dolichovespula arenaria</i>	Yellow jacket	Dol a 5	Venom allergen 5 (Antigen 5) (Ag5) (Allergen Dol a 5) (Dol a V)	465052
692	<i>Dolichovespula maculata</i>	Whiteface hornet	Dol m 1	Phospholipase A1 2 (Allergen Dol m 1.02) (Dol m I)	1709542
693	<i>Dolichovespula maculata</i>	Whiteface hornet	Dol m 1	Phospholipase A1 1 precursor (Allergen Dol m 1.01) (Dol m I)	548449
694	<i>Dolichovespula maculata</i>	Whiteface hornet	Dol m 2	Hyaluronoglucosaminidase (Hyaluronidase) (Allergen Dol m 2) (Dol m II)	1346322
695	<i>Dolichovespula maculata</i>	Whiteface hornet	Dol m 5	Venom allergen 5.02 precursor (Antigen 5 form 3) (Ag5-3) (Allergen Dol m 5.02) (Dol m V-B)	549186

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
696	<i>Dolichovespula maculata</i>	Whiteface hornet	Dol m 5	Venom allergen 5.01 precursor (Antigen 5 form 2) (Ag5-2) (Allergen Dol m 5.01) (Dol m V-A)	137395
697	<i>Epicoccum nigrum</i>	Fungus	Unassigned	Major allergen Epi p 1 (EpI n I4625*)	24636820
698	<i>Equus caballus</i>	Horse	Equ c 1	Major allergen Equ c 1 precursor	3121758
699	<i>Equus caballus</i>	Horse	Equ c 2.0102	Dander allergen Equ c 2.0102	3121756
700	<i>Equus caballus</i>	Horse	Equ c 2.0101	Dander allergen Equ c 2.0101	3121755
701	<i>Equus caballus</i>	Horse	Unassigned	Serum albumin precursor (Allergen Equ c 3)	543794
702	<i>Equus caballus</i>	Horse	Unassigned	Latherin precursor (Dander allergen Equ c 4/Equ c 5)	152031631
703	<i>Equus caballus</i>	Horse	Unassigned	Latherin precursor (Dander allergen Equ c 4/Equ c 5)	38258932
704	<i>Erimacrus isenbeckii</i>		Unassigned	tropomyosin slow-tonic isoform [Erimacrus isenbeckii]	125995171
705	<i>Erimacrus isenbeckii</i>		Unassigned	tropomyosin slow-twitch isoform [Erimacrus isenbeckii]	125995169
706	<i>Euphausia pacifica</i>		Unassigned	tropomyosin [Euphausia pacifica]	156712754
707	<i>Euphausia superba</i>		Unassigned	tropomyosin [Euphausia superba]	156712752
708	<i>Euroglyphus maynei</i>	House dust mite	Eur m 2	Mite group 2 allergen Eur m 2 precursor	14423649
709	<i>Euroglyphus maynei</i>	House dust mite	Eur m 2.0102	group 2 allergen Eur m 2 0102 [Euroglyphus maynei]	3941386
710	<i>Fagopyrum esculentum</i>	Buckwheat	Unassigned	16 kDa allergen [Fagopyrum esculentum]	83416591
711	<i>Fagopyrum esculentum</i>	Buckwheat	Unassigned	BW 16kDa allergen [Fagopyrum esculentum]	61970231
712	<i>Fagopyrum esculentum</i>	Buckwheat	Unassigned	BW8KD allergen protein [Fagopyrum esculentum]	17907758

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
713	<i>Fagopyrum esculentum</i>	Buckwheat	Unassigned	13S globulin seed storage protein 3 precursor (Legumin-like protein 3) (Allergen Fag e 1) [Contains: 13S globulin seed storage protein 3 acidic chain; 13S globulin seed storage protein 3 basic chain]	29839419
714	<i>Fagopyrum esculentum</i>	Buckwheat	Unassigned	13S globulin seed storage protein 2 precursor (Legumin-like protein 2) [Contains: 13S globulin seed storage protein 2 acidic chain; 13S globulin seed storage protein 2 basic chain]	29839255
715	<i>Fagopyrum esculentum</i>	Buckwheat	Unassigned	13S globulin seed storage protein 1 precursor (Legumin-like protein 1) [Contains: 13S globulin seed storage protein 1 acidic chain; 13S globulin seed storage protein 1 basic chain]	29839254
716	<i>Fagopyrum esculentum</i>	Buckwheat	Unassigned	vicilin-like protein [Fagopyrum esculentum]	146217148
717	<i>Fagopyrum gracilipes</i>	Buckwheat	Unassigned	22kDa storage protein [Fagopyrum gracilipes]	6979766
718	<i>Fagopyrum tataricum</i>	Buckwheat	Unassigned	16 kDa major allergen [Fagopyrum tataricum]	320445237
719	<i>Fagopyrum tataricum</i>	Buckwheat	Unassigned	BW10KD allergen protein [Fagopyrum tataricum]	144228127
720	<i>Fagopyrum tataricum</i>	Buckwheat	Unassigned	allergenic protein [Fagopyrum tataricum]	113200131
721	<i>Fagus sylvatica</i>	European Beech	Unassigned	Fag s 1 pollen allergen [Fagus sylvatica]	212291474
722	<i>Fagus sylvatica</i>	European Beech	Fag s 1	Fag s 1 pollen allergen [Fagus sylvatica]	212291470

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
723	<i>Fagus sylvatica</i>	European Beech	Unassigned	Fag s 1 pollen allergen [ <i>Fagus sylvatica</i> ]	212291472
724	<i>Farfantepenaeus aztecus</i>	Brown shrimp	Pen a 1	Pen a 1 allergen [ <i>Farfantepenaeus aztecus</i> ]	73532979
725	<i>Felis catus</i>	Cat	Unassigned	major allergen I, polypeptide chain 1 [ <i>Felis catus</i> ]	114326420
726	<i>Felis catus</i>	Cat	Fel d 1	major allergen I	163825
727	<i>Felis catus</i>	Cat	Fel d 1	Major allergen I polypeptide chain 1 major form precursor (Allergen Fel d 1-A) (Fel d I-A) (Allergen Cat-1) (Fel dI) (AG4)	1169665
728	<i>Felis catus</i>	Cat	Fel d 1	fel d I chain 1 precursor with leader A [ <i>Felis catus</i> ]	1364213
729	<i>Felis catus</i>	Cat	Fel d 1	fel d I chain 1 precursor with leader B [ <i>Felis catus</i> ]	1364212
730	<i>Felis catus</i>	Cat	Unassigned	fel d I chain 2 precursor [ <i>Felis catus</i> ]	395407
731	<i>Felis catus</i>	Cat	Fel d 1	Major allergen I polypeptide chain 2 precursor (Allergen Fel d 1-B) (Fel d I-B) (Allergen Cat-1) (AG4) (FdI)	232086
732	<i>Felis catus</i>	Cat	Unassigned	Serum albumin precursor (Allergen Fel d 2)	1351908
733	<i>Felis catus</i>	Cat	Unassigned	Cystatin-A (Allergen Fel d 3)	47605720
734	<i>Felis catus</i>	Cat	Unassigned	Allergen Fel d 4 precursor	75062228
735	<i>Felis catus</i>	Cat	Unassigned	Fel d 7 allergen precursor [ <i>Felis catus</i> ]	301072397
736	<i>Felis catus</i>	Cat	Unassigned	lipid binding protein [ <i>Felis catus</i> ]	303387468
737	<i>Forcipomyia taiwana</i>	biting midges	Unassigned	serine/threonine protein kinase [ <i>Forcipomyia taiwana</i> ]	188572341
738	<i>Forcipomyia taiwana</i>	biting midges	Unassigned	eukaryotic translation initiation factor [ <i>Forcipomyia taiwana</i> ]	188572343

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
739	<i>Fragaria x ananassa</i>	Strawberry	Fra a 1	Fra a 1-A allergen [Fragaria x ananassa]	88082485
740	<i>Fragaria x ananassa</i>	Strawberry	Fra a 1	Major strawberry allergen Fra a 1-B [Fragaria x ananassa]	90185682
741	<i>Fragaria x ananassa</i>	Strawberry	Fra a 1	Major strawberry allergen Fra a 1-D [Fragaria x ananassa]	90185684
742	<i>Fragaria x ananassa</i>	Strawberry	Fra a 1	Major strawberry allergen Fra a 1-C [Fragaria x ananassa]	90185688
743	<i>Fragaria x ananassa</i>	Strawberry	Fra a 1	Major strawberry allergen Fra a 1-E [Fragaria x ananassa]	90185692
744	<i>Fragaria x ananassa</i>	Strawberry	Fra a 1	[Segment 3 of 3] Allergen Fra a 18 kDa	60389905
745	<i>Fragaria x ananassa</i>	Strawberry	Fra a 1	[Segment 2 of 3] Allergen Fra a 18 kDa	60389904
746	<i>Fraxinus excelsior</i>	European ash	Fra e 1	allergen Fra e 1.0101 [Fraxinus excelsior]	33327133
747	<i>Fraxinus excelsior</i>	European ash	Fra e 1	Fra e 1.0102 major allergen [Fraxinus excelsior]	56122438
748	<i>Fraxinus excelsior</i>	European ash	Fra e 1	allergen Fra e 1 [Fraxinus excelsior]	34978692
749	<i>Fulvia mutica</i>		Unassigned	tropomyosin [Fulvia mutica]	219806596
750	<i>Fusarium culmorum</i>	Fungus	Unassigned	helix-loop-helix protein [Fusarium culmorum]	25361513
751	<i>Fusarium culmorum</i>	Fungus	Unassigned	RecName: Full=60S acidic ribosomal protein P2; AltName: Allergen=Fus c 1	41688715
752	<i>Fusarium culmorum</i>	Fungus	Unassigned	Thioredoxin-like protein (Allergen Fus c 2)	52783462
753	<i>Gadus callarias</i>	Baltic cod	Gad c 1	Parvalbumin beta (Allergen Gad c 1) (Gad c I) (Allergen M)	131112
754	<i>Gadus morhua</i>	Atlantic cod	Unassigned	parvalbumin beta [Gadus morhua]	148356693
755	<i>Gadus morhua</i>	Atlantic cod	Unassigned	parvalbumin beta [Gadus morhua]	148356691
756	<i>Gadus morhua</i>	Atlantic cod	Unassigned	parvalbumin beta [Gadus morhua]	14531016

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
757	<i>Gadus morhua</i>	Atlantic cod	Unassigned	parvalbumin beta [ <i>Gadus morhua</i> ]	14531014
758	<i>Gallus gallus</i>	Chicken	Unassigned	ovomucoid [ <i>Gallus gallus</i> ]	209979542
759	<i>Gallus gallus</i>	Chicken	Unassigned	ovomucoid [ <i>Gallus gallus</i> ]	162952006
760	<i>Gallus gallus</i>	Chicken	Gal d 1	Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	124757
761	<i>Gallus gallus</i>	Chicken	Unassigned	Chain D, Crystal Structure Of S-Ovalbumin At 1.9 Angstrom Resolution	34811333
762	<i>Gallus gallus</i>	Chicken	Gal d 2	Chain A, Loop-Inserted Structure Of P1-P1' Cleaved Ovalbumin Mutant R339t	15826578
763	<i>Gallus gallus</i>	Chicken	Gal d 2	unnamed protein product [ <i>Gallus gallus</i> ]	808969
764	<i>Gallus gallus</i>	Chicken	Gal d 2	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	129293
765	<i>Gallus gallus</i>	Chicken	Gal d 2	unnamed protein product [ <i>Gallus gallus</i> ]	63052
766	<i>Gallus gallus</i>	Chicken	Gal d 3	Ovotransferrin precursor (Conalbumin) (Allergen Gal d 3) (Gal d III) (Serum transferrin)	1351295
767	<i>Gallus gallus</i>	Chicken	Gal d 3	ovotransferrin [ <i>Gallus gallus</i> ]	757851
768	<i>Gallus gallus</i>	Chicken	Gal d 4	lysozyme protein	212279
769	<i>Gallus gallus</i>	Chicken	Gal d 4	Lysozyme C precursor (1,4-beta-N-acetylmuramidase C) (Allergen Gal d 4) (Gal d IV)	126608
770	<i>Gallus gallus</i>	Chicken	Unassigned	Serum albumin precursor (Alpha-livetin) (Allergen Gal d 5)	113575
771	<i>Gallus gallus</i>	Chicken	Unassigned	parvalbumin [ <i>Gallus gallus</i> ]	225877920

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
772	<i>Gibberella zeae PH-1</i>	Fungus	Unassigned	RLA2_ALTAL 60S acidic ribosomal protein P2 (Minor allergen Alt a 6) (Alt a VI) [Gibberella zeae PH-1]	46122455
773	<i>Glossina morsitans morsitans</i>	Tsetse fly	Unassigned	antigen 5 precursor [Glossina morsitans morsitans]	8927462
774	<i>Glossina morsitans morsitans</i>	Tsetse fly	Unassigned	salivary antigen 5 precursor variant [Glossina morsitans morsitans]	289742483
775	<i>Glossina morsitans morsitans</i>	Tsetse fly	Unassigned	antigen 5 precursor [Glossina morsitans morsitans]	289742475
776	<i>Glossina morsitans morsitans</i>	Tsetse fly	Unassigned	salivary antigen 5 precursor [Glossina morsitans morsitans]	289740263
777	<i>Glycine max</i>	Soybean	Unassigned	CG4 beta-conglycinin [Glycine max]	256427
778	<i>Glycine max</i>	Soybean	Unassigned	beta-conglycinin storage protein [Glycine max]	169929
779	<i>Glycine max</i>	Soybean	Unassigned	beta-conglycinin-alpha subunit	169927
780	<i>Glycine max</i>	Soybean	Unassigned	unnamed protein product [Glycine max]	18536
781	<i>Glycine max</i>	Soybean	Unassigned	hydrophobic seed protein precursor [Glycine max]	76782249
782	<i>Glycine max</i>	Soybean	Unassigned	hydrophobic seed protein precursor-like [Glycine max]	76782247
783	<i>Glycine max</i>	Soybean	Gly m 1.0101	Gly m IA allergen, HPS=hydrophobic seed protein {N-terminal} [Glycine max=soybeans, hulls, Peptide Partial, 42 aa]	999355
784	<i>Glycine max</i>	Soybean	Gly m 2	hull allergen Gly m 2 - soybean (fragment)	1362049
785	<i>Glycine max</i>	Soybean	Unassigned	profilin [Glycine max]	156938901
786	<i>Glycine max</i>	Soybean	Gly m 3	Profilin-1 (GmPRO1) (Allergen Gly m 3)	3914435
787	<i>Glycine max</i>	Soybean	Gly m 3	profilin [Glycine max]	3021373

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
788	<i>Glycine max</i>	Soybean	Unassigned	Stress-induced protein SAM22 (Starvation-associated message 22) (Allergen Gly m 4)	134194
789	<i>Glycine max</i>	Soybean	Unassigned	Gly m Bd 28K allergen [Glycine max]	187766755
790	<i>Glycine max</i>	Soybean	Unassigned	Gly m Bd 28K allergen [Glycine max]	187766747
791	<i>Glycine max</i>	Soybean	Unassigned	Gly m Bd 28K allergen [Glycine max]	187766749
792	<i>Glycine max</i>	Soybean	Unassigned	Gly m Bd 28K allergen [Glycine max]	187766751
793	<i>Glycine max</i>	Soybean	Unassigned	allergen Gly m Bd 28K [Glycine max]	12697782
794	<i>Glycine max</i>	Soybean	Unassigned	Bd 30K [Glycine max]	3097321
795	<i>Glycine max</i>	Soybean	Unassigned	34 kDa maturing seed vacuolar thiol protease precursor [Glycine max]	1199563
796	<i>Glycine max</i>	Soybean	Unassigned	P34 probable thiol protease precursor	129353
797	<i>Glycine max</i>	Soybean	Unassigned	glycinin subunit G1 [Glycine max]	18635
798	<i>Glycine max</i>	Soybean	Unassigned	unnamed protein product [Glycine max]	18615
799	<i>Glycine max</i>	Soybean	Unassigned	glycinin subunit G2 [Glycine max]	18637
800	<i>Glycine max</i>	Soybean	Unassigned	unnamed protein product [Glycine max]	18609
801	<i>Glycine max</i>	Soybean	Unassigned	glycinin subunit G3 [Glycine max]	18639
802	<i>Glycine max</i>	Soybean	Unassigned	unnamed protein product [Glycine max]	732706
803	<i>Glycine max</i>	Soybean	Unassigned	glycinin [Glycine max]	18641
804	<i>Glycine max</i>	Soybean	Unassigned	glycinin precursor	169971
805	<i>Glycine max</i>	Soybean	Unassigned	glycinin	169969
806	<i>Glycine max</i>	Soybean	Unassigned	Major Gly 50 kDa allergen	85681057
807	<i>Glycine max</i>	Soybean	Unassigned	Kunitz trypsin inhibitor [Glycine max]	510515
808	<i>Glycine max</i>	Soybean	Unassigned	Kunitz trypsin inhibitor KTi2 [Glycine max]	256636



Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
809	<i>Glycine max</i>	Soybean	Unassigned	Kunitz trypsin inhibitor KTi1 [Glycine max]	256635
810	<i>Glycine max</i>	Soybean	Unassigned	Kunitz trypsin inhibitor; KTi [Glycine max]	256429
811	<i>Glycine max</i>	Soybean	Unassigned	trypsin inhibitor subtype B [Glycine max]	18772
812	<i>Glycine max</i>	Soybean	Unassigned	trypsin inhibitor subtype A [Glycine max]	18770
813	<i>Glycine soja</i>	Soybean	Unassigned	A5A4B3 subunit [Glycine soja]	806556
814	<i>Glycine soja</i>	Soybean	Unassigned	Gy5 [Glycine soja]	736002
815	<i>Glycyphagus domesticus</i>	Storage mite	Unassigned	Mite group 2 allergen Gly d 2.01	48428178
816	<i>Glycyphagus domesticus</i>	Storage mite	Unassigned	Mite group 2 allergen Gly d 2.02	48428170
817	<i>Glycyphagus domesticus</i>	Storage mite	Unassigned	Gly d 2.03 [Glycyphagus domesticus]	33772588
818	<i>Haliotis discus discus</i>	Disk abalone	Unassigned	tropomyosin [Haliotis discus discus]	219806586
819	<i>Haliotis discus discus</i>	Disk abalone	Unassigned	paramyosin [Haliotis discus discus]	318609972
820	<i>Haliotis diversicolor</i>	Abalone	Unassigned	tropomyosin [Haliotis diversicolor]	9954249
821	<i>Helianthus annuus</i>	Sunflower	Hel a 2	profilin [Helianthus annuus]	3581965
822	<i>Helianthus annuus</i>	Sunflower	Unassigned	Albumin-8 precursor (Methionine-rich 2S protein) (SFA8)	112745
823	<i>Helix aspersa</i>	Brown garden snail	Unassigned	Tropomyosin (Allergen Hel as 1)	42559558
824	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 1	Rubber elongation factor protein (REF) (Allergen Hev b 1)	132270
825	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 10.0103	IgE-binding protein MnSOD [Hevea brasiliensis]	10862818
826	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 10.0102	MnSOD [Hevea brasiliensis]	5777414

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
827	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 10.0101	superoxide dismutase (manganese)	348137
828	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 11	putative class I chitinase [Hevea brasiliensis]	14575525
829	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 12	lipid transfer precursor protein [Hevea brasiliensis]	20135538
830	<i>Hevea brasiliensis</i>	Para rubber tree	Unassigned	Esterase precursor (Early nodule-specific protein homolog) (Latex allergen Hev b 13)	51315784
831	<i>Hevea brasiliensis</i>	Para rubber tree	Unassigned	hevamine [Hevea brasiliensis]	313870530
832	<i>Hevea brasiliensis</i>	Para rubber tree	Unassigned	beta-1,3-glucanase [Hevea brasiliensis]	270315180
833	<i>Hevea brasiliensis</i>	Para rubber tree	Unassigned	beta-1,3-glucanase form 'RRII Gln 2' [Hevea brasiliensis]	268037674
834	<i>Hevea brasiliensis</i>	Para rubber tree	Unassigned	Chain D, Crystal Structure Of The Native Endo Beta-1,3-Glucanase (Hev B 2), A Major Allergen From Hevea Brasiliensis (Space Group P41)	261824817
835	<i>Hevea brasiliensis</i>	Para rubber tree	Unassigned	beta-1,3-glucanase [Hevea brasiliensis]	124365253
836	<i>Hevea brasiliensis</i>	Para rubber tree	Unassigned	beta-1,3-glucanase [Hevea brasiliensis]	124365251
837	<i>Hevea brasiliensis</i>	Para rubber tree	Unassigned	beta-1,3-glucanase [Hevea brasiliensis]	124365249
838	<i>Hevea brasiliensis</i>	Para rubber tree	Unassigned	beta-1,3-glucanase [Hevea brasiliensis]	124294785
839	<i>Hevea brasiliensis</i>	Para rubber tree	Unassigned	beta-1,3-glucanase [Hevea brasiliensis]	124294783
840	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 2	beta-1,3-glucanase [Hevea brasiliensis]	32765543
841	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 2	beta-1,3-glucanase	1184668
842	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 3	Small rubber particle protein (SRPP) (22 kDa rubber particle protein) (22 kDa RPP) (Latex allergen Hev b 3) (27 kDa natural rubber allergen)	14423933

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
843	<i>Hevea brasiliensis</i>	Para rubber tree	Unassigned	major latex allergen Hev b 4 [Hevea brasiliensis]	46410859
844	<i>Hevea brasiliensis</i>	Para rubber tree	Unassigned	Major latex allergen Hev b 5	7387766
845	<i>Hevea brasiliensis</i>	Para rubber tree	Unassigned	hevein [Hevea brasiliensis]	158342650
846	<i>Hevea brasiliensis</i>	Para rubber tree	Unassigned	Chain A, Crystal Structure Of A Hev B 6.02 Isoallergen	73535415
847	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 6	prohevein [Hevea brasiliensis]	2832430
848	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 6	Pro-hevein precursor (Major hevein) [Contains: Hevein (Allergen Hev b 6); Win-like protein]	123062
849	<i>Hevea brasiliensis</i>	Para rubber tree	Unassigned	putative latex allergen hev b 7.02 [Hevea brasiliensis]	41581137
850	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 7	latex protein allergen Hev b 7 [Hevea brasiliensis]	6707018
851	<i>Hevea brasiliensis</i>	Para rubber tree	Unassigned	latex allergen [Hevea brasiliensis]	3288200
852	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 7.02	latex allergen [Hevea brasiliensis]	3087805
853	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 7.01	latex patatin homolog [Hevea brasiliensis]	1916805
854	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 8.0102	Profilin-2 (Pollen allergen Hev b 8.0102)	14423868
855	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 8.0201	Profilin-3 (Pollen allergen Hev b 8.0201)	14423860
856	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 8.0202	Profilin-4 (Pollen allergen Hev b 8.0202)	14423859
857	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 8.0203	Profilin-5 (Pollen allergen Hev b 8.0203)	14423858
858	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 8.0204	Profilin-6 (Pollen allergen Hev b 8.0204)	14423856
859	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 8	Chain A, Latex Profilin Hevb8	11513601
860	<i>Hevea brasiliensis</i>	Para rubber tree	Hev b 8	profilin [Hevea brasiliensis]	3183706
861	<i>Hevea brasiliensis</i>	Para rubber tree	Unassigned	Enolase 1 (2-phosphoglycerate dehydratase 1) (2-phospho-D-glycerate hydro-lyase 1) (Allergen Hev b 9)	14423688

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
862	<i>Hevea brasiliensis</i>	Para rubber tree	Unassigned	Enolase 2 (2-phosphoglycerate dehydratase 2) (2-phospho-D-glycerate hydro-lyase 2) (Allergen Hev b 9)	14423687
863	<i>Hevea brasiliensis subsp. brasiliensis</i>	Para rubber tree	Hev b 11	class I chitinase [Hevea brasiliensis subsp. brasiliensis]	27526732
864	<i>Holcus lanatus</i>	Velvet grass	Unassigned	pollen allergen Hol l 5b [Holcus lanatus]	11991229
865	<i>Holcus lanatus</i>	Velvet grass	Unassigned	group V allergen [Holcus lanatus]	2266625
866	<i>Holcus lanatus</i>	Velvet grass	Unassigned	group V grass pollen allergen [Holcus lanatus]	2266623
867	<i>Holcus lanatus</i>	Velvet grass	Unassigned	Major pollen allergen Hol l 1 precursor (Hol l I) (Hol l 1.0101 and 1.0102)	1171005
868	<i>Holcus lanatus</i>	Velvet grass	Unassigned	major group I allergen Hol l 1 [Holcus lanatus]	3860384
869	<i>Holcus lanatus</i>	Velvet grass	Hol l 1.0102	protein with incomplete signal sequence [Holcus lanatus]	1167836
870	<i>Homarus americanus</i>	American lobster	Unassigned	Tropomyosin (Allergen Hom a 1)	14285796
871	<i>Homarus americanus</i>	American lobster	Unassigned	fast tropomyosin isoform [Homarus americanus]	2660868
872	<i>Hordeum vulgare</i>	Barley	Unassigned	Alpha-amylase/trypsin inhibitor CMb precursor (Chloroform/methanol-soluble protein CMb)	585290
873	<i>Hordeum vulgare</i>	Barley	Hor v 15	Alpha-amylase inhibitor BMAI-1 precursor (Allergen Hor v 1) (Alpha-amylase flour inhibitor)	2506771
874	<i>Hordeum vulgare</i>	Barley	Unassigned	amylase/protease inhibitor	167077
875	<i>Hordeum vulgare</i>	Barley	Unassigned	LTP 1 [Hordeum vulgare]	19039
876	<i>Hordeum vulgare</i>	Barley	Unassigned	trypsin inhibitor cme precursor [Hordeum vulgare]	1405736
877	<i>Hordeum vulgare subsp. vulgare</i>	Barley	Unassigned	BDAI-1; Barley dimeric alpha-amylase inhibitor [Hordeum vulgare subsp. vulgare]	3367714
878	<i>Hordeum vulgare subsp. vulgare</i>	Barley	Unassigned	CMA, component of tetrameric alpha-amylase inhibitor [Hordeum vulgare subsp. vulgare]	439275
879	<i>Hordeum vulgare subsp. vulgare</i>	Barley	Unassigned	alpha-amylase inhibitor [Hordeum vulgare subsp. vulgare]	18955

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
880	<i>Hordeum vulgare subsp. vulgare</i>	Barley	Unassigned	CMe [Hordeum vulgare subsp. vulgare]	19009
881	<i>Humulus japonicus</i>	Japanese hop	Hum j 1	Humj1 [Humulus japonicus]	33113263
882	<i>Humulus scandens</i>	Japanese hop	Unassigned	profilin-like protein [Humulus scandens]	34851174
883	<i>Humulus scandens</i>	Japanese hop	Unassigned	profilin-like protein [Humulus scandens]	34851176
884	<i>Juglans nigra</i>	Black walnut	Jug n 1	2S albumin seed storage protein [Juglans nigra]	31321942
885	<i>Juglans nigra</i>	Black walnut	Jug n 2	vicilin seed storage protein [Juglans nigra]	31321944
886	<i>Juglans regia</i>	English walnut	Jug r 1	albumin seed storage protein precursor [Juglans regia]	1794252
887	<i>Juglans regia</i>	English walnut	Jug r 2	vicilin-like protein precursor [Juglans regia]	6580762
888	<i>Juglans regia</i>	English walnut	Unassigned	nonspecific lipid transfer protein [Juglans regia]	209484145
889	<i>Juglans regia</i>	English walnut	Unassigned	seed storage protein [Juglans regia]	56788031
890	<i>Juniperus ashei</i>	Mountain cedar	Unassigned	Polygalacturonase precursor (PG) (Pectinase) (Major pollen allergen Jun a 2)	47606048
891	<i>Juniperus ashei</i>	Mountain cedar	Unassigned	Pathogenesis-related protein precursor (Pollen allergen Jun a 3)	9087177
892	<i>Juniperus ashei</i>	Mountain cedar	Unassigned	Major pollen allergen Jun a 1 precursor	9087152
893	<i>Juniperus oxycedrus</i>	Juniper	Unassigned	putative allergen jun o 1 [Juniperus oxycedrus]	15139849
894	<i>Juniperus oxycedrus</i>	Juniper	Unassigned	Polcalcin Jun o 2 (Calcium-binding pollen allergen Jun o 2)	14423843
895	<i>Juniperus rigida</i>	Cedar	Unassigned	PR5 allergen Jun r 3.1 precursor [Juniperus rigida]	38456222
896	<i>Juniperus rigida</i>	Cedar	Unassigned	PR5 allergen Jun r 3.2 precursor [Juniperus rigida]	38456224
897	<i>Juniperus virginiana</i>	Red cedar	Unassigned	Pathogenesis-related protein precursor (Potential major pollen allergen Jun v 3)	51316532
898	<i>Juniperus virginiana</i>	Red cedar	Jun v 1	pollen major allergen 1-1 [Juniperus virginiana]	8843921
899	<i>Juniperus virginiana</i>	Red cedar	Jun v 1	pollen major allergen 1-2 [Juniperus virginiana]	8843917
900	<i>Lens culinaris</i>	Lentil	Len c 1.0102	allergen Len c 1.0102 [Lens culinaris]	29539111
901	<i>Lens culinaris</i>	Lentil	Len c 1.0101	allergen Len c 1.0101 [Lens culinaris]	29539109
902	<i>Lepidoglyphus destructor</i>	Storage mite	Lep d 10	Tropomyosin (Allergen Lep d 10)	14423956

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
903	<i>Lepidoglyphus destructor</i>	Storage mite	Lep d 13	Fatty acid-binding protein (Allergen Lep d 13)	14423714
904	<i>Lepidoglyphus destructor</i>	Storage mite	Unassigned	Mite group 2 allergen Lep d 2 precursor (Lep d 1) (Lep d I)	1708793
905	<i>Lepidoglyphus destructor</i>	Storage mite	Unassigned	type 2 allergen Lep d 2.042 [Lepidoglyphus destructor]	34495290
906	<i>Lepidoglyphus destructor</i>	Storage mite	Unassigned	type 2 allergen Lep d 2.039 [Lepidoglyphus destructor]	34495288
907	<i>Lepidoglyphus destructor</i>	Storage mite	Unassigned	type 2 allergen Lep d 2.035 [Lepidoglyphus destructor]	34495286
908	<i>Lepidoglyphus destructor</i>	Storage mite	Unassigned	type 2 allergen Lep d 2.031 [Lepidoglyphus destructor]	34495284
909	<i>Lepidoglyphus destructor</i>	Storage mite	Unassigned	type 2 allergen Lep d 2.025 [Lepidoglyphus destructor]	34495282
910	<i>Lepidoglyphus destructor</i>	Storage mite	Unassigned	type 2 allergen Lep d 2.024 [Lepidoglyphus destructor]	34495280
911	<i>Lepidoglyphus destructor</i>	Storage mite	Unassigned	type 2 allergen Lep d 2.023 [Lepidoglyphus destructor]	34495278
912	<i>Lepidoglyphus destructor</i>	Storage mite	Unassigned	type 2 allergen Lep d 2.013 [Lepidoglyphus destructor]	34495274
913	<i>Lepidoglyphus destructor</i>	Storage mite	Lep d 2	allergen Lep d 1.01	1582222
914	<i>Lepidoglyphus destructor</i>	Storage mite	Lep d 2	allergen Lep d 1.02	1582223
915	<i>Lepidoglyphus destructor</i>	Storage mite	Lep d 2	Lep D 2 precursor [Lepidoglyphus destructor]	21213900
916	<i>Lepidoglyphus destructor</i>	Storage mite	Lep d 2	Lep D 2 precursor [Lepidoglyphus destructor]	21213898
917	<i>Lepidoglyphus destructor</i>	Storage mite	Lep d 2	allergen Lep d 1.02 precursor - Lepidoglyphus destructor	2147108
918	<i>Lepidoglyphus destructor</i>	Storage mite	Unassigned	type 2 allergen Lep d 5.04 [Lepidoglyphus destructor]	34495294
919	<i>Lepidoglyphus destructor</i>	Storage mite	Unassigned	type 2 allergen Lep d 5.02 [Lepidoglyphus destructor]	34495292
920	<i>Lepidoglyphus destructor</i>	Storage mite	Lep d 5	Mite allergen Lep d 5	14423651
921	<i>Lepidoglyphus destructor</i>	Storage mite	Lep d 7	Mite allergen Lep d 7 precursor	14423650
922	<i>Lepidorhombus whiffiagonis</i>		Unassigned	parvalbumin [Lepidorhombus whiffiagonis]	208608078
923	<i>Lepisma saccharina</i>	Silverfish	Unassigned	tropomyosin [Lepisma saccharina]	20387029
924	<i>Lepisma saccharina</i>	Silverfish	Lep s 1	tropomyosin [Lepisma saccharina]	20387027
925	<i>Ligustrum vulgare</i>	Privet	Unassigned	Major pollen allergen Lig v 1	14423737
926	<i>Ligustrum vulgare</i>	Privet	Lig v 1.0102	major allergen [Ligustrum vulgare]	3256212
927	<i>Lilium longiflorum</i>	Trumpet lily	Unassigned	polygalacturonase [Lilium longiflorum]	73913442
928	<i>Litchi chinensis</i>	Lychee nut	Unassigned	profilin [Litchi chinensis]	83317152

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
929	<i>Litchi chinensis</i>	Lychee nut	Lit c 1	profilin [ <i>Litchi chinensis</i> ]	15809696
930	<i>Litopenaeus vannamei</i>		Unassigned	sarcoplasmic calcium-binding protein [ <i>Litopenaeus vannamei</i> ]	223403273
931	<i>Litopenaeus vannamei</i>		Unassigned	arginine kinase [ <i>Litopenaeus vannamei</i> ]	115492980
932	<i>Litopenaeus vannamei</i>		Unassigned	Lit v 3 allergen myosin light chain [ <i>Litopenaeus vannamei</i> ]	184198734
933	<i>Lolium perenne</i>	Perennial ryegrass	Unassigned	pollen allergen	168316
934	<i>Lolium perenne</i>	Perennial ryegrass	Lol p 1	Pollen allergen	75274600
935	<i>Lolium perenne</i>	Perennial ryegrass	Lol p 1	pollen allergen	168314
936	<i>Lolium perenne</i>	Perennial ryegrass	Lol p 1	Pollen allergen Lol p 1 precursor (Lol p I) (Allergen R7)	126385
937	<i>Lolium perenne</i>	Perennial ryegrass	Lol p 11	Major pollen allergen Lol p 11 (Lol p XI)	47605808
938	<i>Lolium perenne</i>	Perennial ryegrass	Lol p 2	allergen Lol p II [ <i>Lolium perenne</i> ]	939932
939	<i>Lolium perenne</i>	Perennial ryegrass	Lol p 2	Pollen allergen Lol p 2-A (Lol p II-A)	126386
940	<i>Lolium perenne</i>	Perennial ryegrass	Lol p 3	Pollen allergen Lol p 3 (Lol p III)	126387
941	<i>Lolium perenne</i>	Perennial ryegrass	Unassigned	pollen allergen Lol p 4 [ <i>Lolium perenne</i> ]	55859464
942	<i>Lolium perenne</i>	Perennial ryegrass		RecName: Full=Major pollen allergen Lol p 5a; AltName: Full=Allergen Lol p Ib; AltName: Full=Allergen Lol p Va; AltName: Allergen=Lol p 5a; Flags: Precursor	332278195
943	<i>Lolium perenne</i>	Perennial ryegrass	Lol p 5	pollen allergen [ <i>Lolium perenne</i> ]	6634467
944	<i>Lolium perenne</i>	Perennial ryegrass	Lol p 5	pollen allergen Lol p VA precursor; major allergen [ <i>Lolium perenne</i> ]	4416516
945	<i>Lolium perenne</i>	Perennial ryegrass	Lol p 5	Major pollen allergen Lol p 5b precursor (Lol p Vb)	2498582
946	<i>Lupinus angustifolius</i>		Unassigned	conglutin beta [ <i>Lupinus angustifolius</i> ]	169950562
947	<i>Lupinus angustifolius</i>		Unassigned	conglutin beta [ <i>Lupinus angustifolius</i> ]	149208403
948	<i>Lupinus angustifolius</i>		Unassigned	conglutin beta [ <i>Lupinus angustifolius</i> ]	149208401
949	<i>Lycopersicon esculentum</i>	Tomato	Lyc e 1	profilin [ <i>Lycopersicon esculentum</i> ]	17224229
950	<i>Lycopersicon esculentum</i>	Tomato	Lyc e 1	profilin [ <i>Lycopersicon esculentum</i> ]	16555787
951	<i>Lycopersicon esculentum</i>	Tomato	Lyc e 2.0102	minor allergen beta-fructofuranosidase precursor [ <i>Lycopersicon esculentum</i> ]	18542115

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
952	<i>Lycopersicon esculentum</i>	Tomato	Lyc e 2.0101	minor allergen beta-fructofuranosidase precursor [Lycopersicon esculentum]	18542113
953	<i>Lycopersicon esculentum</i>	Tomato	Unassigned	non-specific lipid transfer protein [Lycopersicon esculentum]	71360930
954	<i>Lycopersicon esculentum</i>	Tomato	Unassigned	non-specific lipid transfer protein [Lycopersicon esculentum]	71360928
955	<i>Macrobrachium rosenbergii</i>		Unassigned	tropomyosin [Macrobrachium rosenbergii]	288819271
956	<i>Malassezia furfur</i>	Yeast	Unassigned	Putative peroxiredoxin (Thioredoxin reductase) (Allergen Mal f 2) (MF1)	3914386
957	<i>Malassezia furfur</i>	Yeast	Unassigned	Putative peroxiredoxin (Thioredoxin reductase) (Allergen Mal f 3) (MF2)	3914387
958	<i>Malassezia furfur</i>	Yeast	Mala f 4	major allergenic protein Mal f4 [Malassezia furfur]	4587985
959	<i>Malassezia furfur</i>	Yeast	Unassigned	Major allergen Mal f 1 precursor (Pit o 1)	13959403
960	<i>Malassezia sympodialis</i>	Yeast	Mala s 11	manganese superoxide dismutase [Malassezia sympodialis]	28569698
961	<i>Malassezia sympodialis</i>	Yeast	Unassigned	mala s 12 allergen precursor [Malassezia sympodialis]	78038796
962	<i>Malassezia sympodialis</i>	Yeast	Unassigned	Chain B, Cross-Reactivity And Crystal Structure Of Malassezia Sympodialis Thioredoxin (Mala S 13), A Member Of A New Pan- Allergen Family	119390336
963	<i>Malassezia sympodialis</i>	Yeast	Mala s 5	allergen [Malassezia sympodialis]	4138171
964	<i>Malassezia sympodialis</i>	Yeast	Mala s 6	allergen [Malassezia sympodialis]	4138173
965	<i>Malassezia sympodialis</i>	Yeast	Mala s 7	allergen [Malassezia sympodialis]	4138175
966	<i>Malassezia sympodialis</i>	Yeast	Mala s 8	allergen [Malassezia sympodialis]	7271239
967	<i>Malassezia sympodialis</i>	Yeast	Mala s 9	allergen [Malassezia sympodialis]	19069920
968	<i>Malus x domestica</i>	Apple	Unassigned	major allergen [Malus x domestica]	886683
969	<i>Malus x domestica</i>	Apple	Unassigned	Ribonuclease-like PR-10a	75306007
970	<i>Malus x domestica</i>	Apple	Unassigned	Ribonuclease-like PR-10c (Mal d 1.0109)	75306008
971	<i>Malus x domestica</i>	Apple	Unassigned	Major allergen Mal d 1 (Allergen Mal d I) (AP15)	42558971
972	<i>Malus x domestica</i>	Apple	Unassigned	major allergen Mal d 1.03D [Malus x domestica]	60280851
973	<i>Malus x domestica</i>	Apple	Unassigned	major allergen Mal d 1.07 [Malus x domestica]	60280829
974	<i>Malus x domestica</i>	Apple	Mal d 1	Major allergen Mal d 1 (Mal d I)	1346478



Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
975	<i>Malus x domestica</i>	Apple	Mal d 1	major allergen Mal d 1 [Malus x domestica]	27922941
976	<i>Malus x domestica</i>	Apple	Mal d 1	ypr10 [Malus x domestica]	16555783
977	<i>Malus x domestica</i>	Apple	Mal d 1	major allergen mal d 1 [Malus x domestica]	4590388
978	<i>Malus x domestica</i>	Apple	Mal d 1	major allergen mal d 1 [Malus x domestica]	4590382
979	<i>Malus x domestica</i>	Apple	Mal d 1	major allergen mal d 1 [Malus x domestica]	4590380
980	<i>Malus x domestica</i>	Apple	Mal d 1	major allergen mal d 1 [Malus x domestica]	4590378
981	<i>Malus x domestica</i>	Apple	Mal d 1	major allergen mal d 1 [Malus x domestica]	4590376
982	<i>Malus x domestica</i>	Apple	Mal d 1	major allergen mal d 1 [Malus x domestica]	4590368
983	<i>Malus x domestica</i>	Apple	Mal d 1	major allergen mal d 1 [Malus x domestica]	4590366
984	<i>Malus x domestica</i>	Apple	Mal d 1	major allergen mal d 1 [Malus x domestica]	4590364
985	<i>Malus x domestica</i>	Apple	Mal d 1	major allergen Mal d 1 [Malus x domestica]	1313966
986	<i>Malus x domestica</i>	Apple	Unassigned	thaumatin-like protein [Malus x domestica]	218059715
987	<i>Malus x domestica</i>	Apple	Unassigned	thaumatin-like protein [Malus x domestica]	218059718
988	<i>Malus x domestica</i>	Apple	Unassigned	Thaumatin-like protein 1a precursor (Allergen Mal d 2) (Mdtl1) (Pathogenesis-related protein 5a) (PR-5a)	30316292
989	<i>Malus x domestica</i>	Apple	Unassigned	thaumatin-like protein precursor [Malus x domestica]	60418848
990	<i>Malus x domestica</i>	Apple	Unassigned	thaumatin-like protein precursor [Malus x domestica]	60418842
991	<i>Malus x domestica</i>	Apple	Unassigned	31 kda major allergen/disease resistance protein homolog {N-terminal} [Malus domestica=apple trees, Golden Delicious, fruit, Peptide Partial, 26 aa]	1478293
992	<i>Malus x domestica</i>	Apple	Unassigned	Non-specific lipid-transfer protein precursor (LTP) (Allergen Mal d 3)	14423814
993	<i>Malus x domestica</i>	Apple	Unassigned	allergen Mal d 3 [Malus x domestica]	38492338
994	<i>Malus x domestica</i>	Apple	Unassigned	lipid transfer protein precursor [Malus x domestica]	50659859
995	<i>Malus x domestica</i>	Apple	Unassigned	lipid transfer protein precursor [Malus x domestica]	50659879
996	<i>Malus x domestica</i>	Apple	Unassigned	lipid transfer protein precursor [Malus x domestica]	50659885
997	<i>Malus x domestica</i>	Apple	Unassigned	lipid transfer protein precursor [Malus x domestica]	50659889
998	<i>Malus x domestica</i>	Apple	Unassigned	lipid transfer protein precursor [Malus x domestica]	50659891

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
999	<i>Malus x domestica</i>	Apple	Unassigned	profilin [Malus x domestica]	218059728
1000	<i>Malus x domestica</i>	Apple	Unassigned	profilin [Malus x domestica]	218059733
1001	<i>Malus x domestica</i>	Apple	Unassigned	profilin [Malus x domestica]	218059730
1002	<i>Malus x domestica</i>	Apple	Unassigned	profilin [Malus x domestica]	164510860
1003	<i>Malus x domestica</i>	Apple	Unassigned	profilin [Malus x domestica]	164510858
1004	<i>Malus x domestica</i>	Apple	Unassigned	profilin [Malus x domestica]	164510842
1005	<i>Malus x domestica</i>	Apple	Unassigned	profilin 2 [Malus x domestica]	60418866
1006	<i>Malus x domestica</i>	Apple	Unassigned	profilin 2 [Malus x domestica]	60418862
1007	<i>Malus x domestica</i>	Apple	Unassigned	profilin 1 [Malus x domestica]	60418858
1008	<i>Malus x domestica</i>	Apple	Unassigned	profilin 1 [Malus x domestica]	60418854
1009	<i>Malus x domestica</i>	Apple	Mal d 4	profilin [Malus x domestica]	28881455
1010	<i>Malus x domestica</i>	Apple	Mal d 4	profilin [Malus x domestica]	28881457
1011	<i>Malus x domestica</i>	Apple	Mal d 4	profilin [Malus x domestica]	28881453
1012	<i>Malus x domestica</i>	Apple	Mal d 4	Profilin-3 (GD4-5) (Pollen allergen Mal d 4)	14423875
1013	<i>Malus x domestica</i>	Apple	Mal d 4	Profilin-2 (GD4-2) (Pollen allergen Mal d 4)	14423874
1014	<i>Malus x domestica</i>	Apple	Mal d 4	Profilin-1 (GD4-1) (Pollen allergen Mal d 4)	14423873
1015	<i>Marsupenaeus japonicus</i>		Unassigned	tropomyosin fast isoform [Marsupenaeus japonicus]	125995159
1016	<i>Mercurialis annua</i>	Annual mercury grass	Mer a 1	Profilin [Mercurialis annua]	2959898
1017	<i>Metapenaeus ensis</i>	Greasyback shrimp	Unassigned	Tropomyosin (Allergen Met e 1) (Met e I)	6094504
1018	<i>Mimachlamys nobilis</i>	Noble scallop	Unassigned	tropomyosin [Chlamys nobilis]	9954253
1019	<i>Morus nigra</i>	Black mulberry	Unassigned	RecName: Full=Non-specific lipid-transfer protein 1; Short=LTP 1; AltName: Allergen=Mor n 3	288561913
1020	<i>Mus musculus</i>	Mouse	Mus m 1	Major urinary protein 6 precursor (MUP 6) (Alpha-2U-globulin) (Group 1, BS6) (Allergen Mus m 1)	20178291
1021	<i>Musa acuminata</i>	Banana	Mus xp 1	profilin [Musa acuminata]	14161635

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1022	<i>Myrmecia pilosula</i>	Jumper ant	Unassigned	Myr p I=allergenic polypeptide {N-terminal} [Myrmecia pilosula=jumper ants, venom, Peptide Partial, 112 aa]	1911819
1023	<i>Myrmecia pilosula</i>	Jumper ant	Myr p 1	Pilosulin-1 precursor (Major allergen Myr p 1) (Myr p I) [Contains: Pilosulin-1; Pilosulin-1 65->112; Pilosulin-1 68->112; Pilosulin-1 71->112; Pilosulin-1 86->112]	730091
1024	<i>Myrmecia pilosula</i>	Jumper ant	Myr p 2	Allergen Myr p 2 precursor (Myr p II)	2498604
1025	<i>Myrmecia pilosula</i>	Jumper ant	Myr p 2	major allergen Myr p II	1587177
1026	<i>Neptunea polycostata</i>		Unassigned	tropomyosin [Neptunea polycostata]	219806590
1027	<i>Nicotiana tabacum</i>	Tobacco	Unassigned	villin 1 [Nicotiana tabacum]	57283137
1028	<i>Nicotiana tabacum</i>	Tobacco	Unassigned	villin 2 [Nicotiana tabacum]	57283139
1029	<i>Octopus vulgaris</i>		Unassigned	tropomyosin [Octopus vulgaris]	83715936
1030	<i>Olea europaea</i>	Olive tree	Unassigned	major pollen allergen Ole e 1 [Olea europaea]	145313992
1031	<i>Olea europaea</i>	Olive tree	Unassigned	major pollen allergen Ole e 1 [Olea europaea]	145313990
1032	<i>Olea europaea</i>	Olive tree	Unassigned	major pollen allergen Ole e 1 [Olea europaea]	145313988
1033	<i>Olea europaea</i>	Olive tree	Unassigned	major pollen allergen Ole e 1 [Olea europaea]	145313984
1034	<i>Olea europaea</i>	Olive tree	Unassigned	major pollen allergen Ole e 1 [Olea europaea]	145313982
1035	<i>Olea europaea</i>	Olive tree	Ole e 1	Major pollen allergen (Allergen Ole e 1) (Ole e I)	14424429
1036	<i>Olea europaea</i>	Olive tree	Unassigned	major pollen allergen Ole e 1 [Olea europaea]	33325115
1037	<i>Olea europaea</i>	Olive tree	Unassigned	major pollen allergen Ole e 1 [Olea europaea]	33329732
1038	<i>Olea europaea</i>	Olive tree	Unassigned	major pollen allergen Ole e 1 [Olea europaea]	33329738
1039	<i>Olea europaea</i>	Olive tree	Unassigned	major pollen allergen Ole e 1 [Olea europaea]	33329744
1040	<i>Olea europaea</i>	Olive tree	Unassigned	major pollen allergen Ole e 1 [Olea europaea]	33329748
1041	<i>Olea europaea</i>	Olive tree	Unassigned	major pollen allergen Ole e 1 [Olea europaea]	33329750
1042	<i>Olea europaea</i>	Olive tree	Unassigned	major pollen allergen Ole e 1 [Olea europaea]	33329752
1043	<i>Olea europaea</i>	Olive tree	Unassigned	major pollen allergen Ole e 1 [Olea europaea]	33329754
1044	<i>Olea europaea</i>	Olive tree	Unassigned	major pollen allergen Ole e 1 [Olea europaea]	33329756
1045	<i>Olea europaea</i>	Olive tree	Unassigned	major pollen allergen Ole e 1 [Olea europaea]	33329758
1046	<i>Olea europaea</i>	Olive tree	Unassigned	major pollen allergen Ole e 1 [Olea europaea]	37548753

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1047	<i>Olea europaea</i>	Olive tree	Unassigned	major pollen allergen Ole e 1 [Olea europaea]	37724593
1048	<i>Olea europaea</i>	Olive tree	Unassigned	major pollen allergen Ole e 1 [Olea europaea]	37724597
1049	<i>Olea europaea</i>	Olive tree	Ole e 1.0101	main olive allergen [Olea europaea]	13195753
1050	<i>Olea europaea</i>	Olive tree	Ole e 1.0107	Ole e 1 protein [Olea europaea]	2465131
1051	<i>Olea europaea</i>	Olive tree	Ole e 1.0106	Ole e 1.0103 protein [Olea europaea]	2465129
1052	<i>Olea europaea</i>	Olive tree	Ole e 1.0105	Ole e 1.0102 protein [Olea europaea]	2465127
1053	<i>Olea europaea</i>	Olive tree	Unassigned	major allergen OLE6 - common olive (fragment)	1362137
1054	<i>Olea europaea</i>	Olive tree	Ole e 1.0103	major allergen OLE5c - common olive	1362136
1055	<i>Olea europaea</i>	Olive tree	Ole e 1.0102	major allergen OLE3c - common olive	1362135
1056	<i>Olea europaea</i>	Olive tree	Unassigned	major allergen OLE33/OLE37 - common olive (fragment)	1362134
1057	<i>Olea europaea</i>	Olive tree	Unassigned	major allergen OLE26 - common olive (fragment)	1362133
1058	<i>Olea europaea</i>	Olive tree	Ole e 1	major allergen OLE20 - common olive (fragment)	1362132
1059	<i>Olea europaea</i>	Olive tree	Ole e 1.0104	major allergen OLE1c - common olive (fragment)	1362131
1060	<i>Olea europaea</i>	Olive tree	Unassigned	major allergen OLE19 - common olive (fragment)	1362130
1061	<i>Olea europaea</i>	Olive tree	Unassigned	major allergen OLE17 - common olive (fragment)	1362129
1062	<i>Olea europaea</i>	Olive tree	Unassigned	major allergen OLE16 - common olive (fragment)	1362128
1063	<i>Olea europaea</i>	Olive tree	Ole e 10	allergen Ole e 10 [Olea europaea]	29465664
1064	<i>Olea europaea</i>	Olive tree	Ole e 11.0101	Ole e 11.0101 allergen precursor [Olea europaea]	269996495
1065	<i>Olea europaea</i>	Olive tree	Ole e 11.0102	Ole e 11.01 allergen precursor [Olea europaea]	68270856
1066	<i>Olea europaea</i>	Olive tree	Unassigned	allergenic thaumatin [Olea europaea]	269996497
1067	<i>Olea europaea</i>	Olive tree	Ole e 2	Profilin-3 (Pollen allergen Ole e 2)	3914428
1068	<i>Olea europaea</i>	Olive tree	Ole e 2	Profilin-2 (Pollen allergen Ole e 2)	3914427
1069	<i>Olea europaea</i>	Olive tree	Ole e 2	Profilin-1 (Pollen allergen Ole e 2)	3914426
1070	<i>Olea europaea</i>	Olive tree	Ole e 3	Ole e 3 allergen [Olea europaea]	37725377
1071	<i>Olea europaea</i>	Olive tree	Ole e 3	calcium-binding pollen allergen [Olea europaea]	3337403
1072	<i>Olea europaea</i>	Olive tree	Unassigned	Ole e 5 olive pollen allergen [Olea europaea]	160962611

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1073	<i>Olea europaea</i>	Olive tree	Unassigned	Ole e 5 olive pollen allergen [ <i>Olea europaea</i> ]	160962597
1074	<i>Olea europaea</i>	Olive tree	Unassigned	Ole e 5 olive pollen allergen [ <i>Olea europaea</i> ]	160962591
1075	<i>Olea europaea</i>	Olive tree	Unassigned	Ole e 5 olive pollen allergen [ <i>Olea europaea</i> ]	160962587
1076	<i>Olea europaea</i>	Olive tree	Unassigned	Ole e 5 olive pollen allergen [ <i>Olea europaea</i> ]	160962583
1077	<i>Olea europaea</i>	Olive tree	Unassigned	Ole e 5 olive pollen allergen [ <i>Olea europaea</i> ]	160962577
1078	<i>Olea europaea</i>	Olive tree	Unassigned	Ole e 5 olive pollen allergen [ <i>Olea europaea</i> ]	160962557
1079	<i>Olea europaea</i>	Olive tree	Unassigned	Ole e 5 olive pollen allergen [ <i>Olea europaea</i> ]	160962547
1080	<i>Olea europaea</i>	Olive tree	Unassigned	Ole e 5 olive pollen allergen [ <i>Olea europaea</i> ]	160962543
1081	<i>Olea europaea</i>	Olive tree	Unassigned	allergen Ole e 5 [ <i>Olea europaea</i> ]	160347138
1082	<i>Olea europaea</i>	Olive tree	Unassigned	allergen Ole e 5 [ <i>Olea europaea</i> ]	160347134
1083	<i>Olea europaea</i>	Olive tree	Unassigned	allergen Ole e 5 [ <i>Olea europaea</i> ]	160347130
1084	<i>Olea europaea</i>	Olive tree	Unassigned	allergen Ole e 5 [ <i>Olea europaea</i> ]	160347126
1085	<i>Olea europaea</i>	Olive tree	Unassigned	allergen Ole e 5 [ <i>Olea europaea</i> ]	160347124
1086	<i>Olea europaea</i>	Olive tree	Unassigned	allergen Ole e 5 [ <i>Olea europaea</i> ]	160347122
1087	<i>Olea europaea</i>	Olive tree	Unassigned	allergen Ole e 5 [ <i>Olea europaea</i> ]	160347120
1088	<i>Olea europaea</i>	Olive tree	Unassigned	allergen Ole e 5 [ <i>Olea europaea</i> ]	160347112
1089	<i>Olea europaea</i>	Olive tree	Unassigned	allergen Ole e 5 [ <i>Olea europaea</i> ]	160347108
1090	<i>Olea europaea</i>	Olive tree	Unassigned	allergen Ole e 5 [ <i>Olea europaea</i> ]	160347106
1091	<i>Olea europaea</i>	Olive tree	Unassigned	pollen allergen Ole e 5 [ <i>Olea europaea</i> ]	145313972
1092	<i>Olea europaea</i>	Olive tree	Ole e 5	Superoxide dismutase [Cu-Zn] (Allergen Ole e 5) (Ole e V)	122064581
1093	<i>Olea europaea</i>	Olive tree	Ole e 6	Pollen allergen Ole e 6	14423643
1094	<i>Olea europaea</i>	Olive tree	Ole e 7	Pollen allergen Ole e 7 (Ole e VII)	22002032
1095	<i>Olea europaea</i>	Olive tree	Ole e 8	Calcium-binding allergen Ole e 8 (PCA18/PCA23)	14423648
1096	<i>Olea europaea</i>	Olive tree	Ole e 8	calcium-binding protein [ <i>Olea europaea</i> ]	6901654
1097	<i>Olea europaea</i>	Olive tree	Unassigned	Chain A, Solution Structure Of The C-Terminal Domain Ole E 9	166235350

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1098	<i>Olea europaea</i>	Olive tree	Ole e 9	beta-1,3-glucanase-like protein [ <i>Olea europaea</i> ]	14279169
1099	<i>Ommastrephes bartramii</i>		Unassigned	tropomyosin [ <i>Ommastrephes bartramii</i> ]	83715934
1100	<i>Onchocerca volvulus</i>	Parasitic nematode	Unassigned	RecName: Full=Tropomyosin; AltName: Full=MOv-14; AltName: Full=Ov-tmy-1	42559586
1101	<i>Oncorhynchus mykiss</i>		Unassigned	RecName: Full=Parvalbumin beta 2	288559140
1102	<i>Oncorhynchus mykiss</i>		Unassigned	RecName: Full=Parvalbumin beta 1	288559139
1103	<i>Oratosquilla oratoria</i>		Unassigned	tropomyosin [ <i>Oratosquilla oratoria</i> ]	162286975
1104	<i>Oryza sativa</i>	Rice	Unassigned	Lactoylglutathione lyase (Methylglyoxalase) (Aldoketomutase) (Glyoxalase I) (Glx I) (Ketone-aldehyde mutase) (S-D-lactoylglutathione methylglyoxal lyase) (Allergen Ory s ?) (Allergen Glb33) (PP33)	84029333
1105	<i>Oryza sativa</i>	Rice	Ory s 1	Ory s 1	1173557
1106	<i>Oryza sativa</i>	Rice	Unassigned	beta-expansin [ <i>Oryza sativa</i> ]	8118439
1107	<i>Oryza sativa</i> ( <i>japonica</i> cultivar-group)	Rice	Unassigned	glyoxalase I [ <i>Oryza sativa</i> ( <i>japonica</i> cultivar-group)]	16580747
1108	<i>Oryza sativa</i> ( <i>japonica</i> cultivar-group)	Rice	Ory s 1	Expansin-B1 precursor (OsEXPB1) (Beta-expansin-1) (OsaEXPb1.2) (OsaEXPb1.3) (Major pollen allergen Ory s 1) (Ory s I)	109913547
1109	<i>Oryza sativa</i> ( <i>japonica</i> cultivar-group)	Rice	Unassigned	allergen RA5B precursor [ <i>Oryza sativa</i> ( <i>japonica</i> cultivar-group)]	23495787
1110	<i>Oryza sativa</i> ( <i>japonica</i> cultivar-group)	Rice	Unassigned	allergen RA16 [ <i>Oryza sativa</i> ( <i>japonica</i> cultivar-group)]	23616954
1111	<i>Oryza sativa</i> ( <i>japonica</i> cultivar-group)	Rice	Unassigned	Seed allergenic protein RAG1 precursor (Seed allergenic protein RA17)	114152864
1112	<i>Oryza sativa</i> ( <i>japonica</i> cultivar-group)	Rice	Unassigned	Seed allergenic protein RAG2 precursor (Seed allergenic protein RA14)	114152865
1113	<i>Oryza sativa</i> ( <i>japonica</i> cultivar-group)	Rice	Unassigned	putative allergenic protein [ <i>Oryza sativa</i> ( <i>japonica</i> cultivar-group)]	23616947
1114	<i>Oryza sativa</i> ( <i>japonica</i> cultivar-group)	Rice	Unassigned	allergenic protein [ <i>Oryza sativa</i> ]	2827316

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1115	<i>Oryza sativa</i> (japonica cultivar-group)	Rice	Unassigned	allergenic protein [Oryza sativa (japonica cultivar-group)]	1398918
1116	<i>Oryza sativa</i> (japonica cultivar-group)	Rice	Unassigned	allergenic protein [Oryza sativa (japonica cultivar-group)]	1398916
1117	<i>Oryza sativa</i> (japonica cultivar-group)	Rice	Unassigned	allergenic protein [Oryza sativa (japonica cultivar-group)]	1398915
1118	<i>Oryza sativa</i> (japonica cultivar-group)	Rice	Unassigned	allergenic protein [Oryza sativa (japonica cultivar-group)]	1398913
1119	<i>Oryza sativa</i> (japonica cultivar-group)	Rice	Unassigned	allergenic protein [Oryza sativa (japonica cultivar-group)]	1304218
1120	<i>Oryza sativa</i> (japonica cultivar-group)	Rice	Unassigned	allergenic protein [Oryza sativa (japonica cultivar-group)]	1304217
1121	<i>Oryza sativa</i> (japonica cultivar-group)	Rice	Unassigned	allergenic protein [Oryza sativa (japonica cultivar-group)]	1304216
1122	<i>Oryza sativa</i> (japonica cultivar-group)	Rice	Unassigned	allergenic protein [Oryza sativa (japonica cultivar-group)]	218197
1123	<i>Oryza sativa</i> (japonica cultivar-group)	Rice	Unassigned	allergenic protein [Oryza sativa (japonica cultivar-group)]	218193
1124	<i>Ostrya carpinifolia</i>		Unassigned	pollen allergen [Ostrya carpinifolia]	300872535
1125	<i>Pachycondyla chinensis</i>		Unassigned	Pac c 3 allergen [Pachycondyla chinensis]	169822894
1126	<i>Pandalus borealis</i>		Unassigned	tropomyosin, allergen Pan b 1 [Pandalus borealis]	312831088
1127	<i>Pandalus eous</i>		Unassigned	tropomyosin fast isoform [Pandalus eous]	125995161
1128	<i>Panulirus stimpsoni</i>	Lobster	Unassigned	Tropomyosin (Allergen Pan s 1) (Pan s I)	14285797
1129	<i>Paralithodes camtschaticus</i>		Unassigned	tropomyosin slow-tonic isoform [Paralithodes camtschaticus]	125995165
1130	<i>Paralithodes camtschaticus</i>		Unassigned	tropomyosin fast isoform [Paralithodes camtschaticus]	125995163
1131	<i>Parietaria judaica</i>	Weed	Par j 1.0101	Probable nonspecific lipid-transfer protein (LTP) (Major pollen allergen Par j 1.0101) (Par j I) (P5 protein)	3915783
1132	<i>Parietaria judaica</i>	Weed	Par j 1.0201	Probable nonspecific lipid-transfer protein 1 precursor (LTP) (Major pollen allergen Par j 1.0201) (Par j I) (P1 protein)	2497749

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1133	<i>Parietaria judaica</i>	Weed	Par j 1.0102	P9 protein [Parietaria judaica]	1532058
1134	<i>Parietaria judaica</i>	Weed	Par j 1	major allergen Par j I	741844
1135	<i>Parietaria judaica</i>	Weed	Par j 2.0101	Probable nonspecific lipid-transfer protein 2 precursor (LTP 2) (Major pollen allergen Par j 2.0101) (Par j II) (P2 protein)	2497750
1136	<i>Parietaria judaica</i>	Weed	Par j 2.0102	P8 protein [Parietaria judaica]	1532056
1137	<i>Parietaria judaica</i>	Weed	Par j 3	Profilin-1 (Pollen allergen Par j 3)	14423876
1138	<i>Parietaria judaica</i>	Weed	Par j 3	Profilin-2 (Pollen allergen Par j 3)	14423869
1139	<i>Parietaria officinalis</i>	Weed	Par o 1	Pollen major allergen Par o I	75139847
1140	<i>Parietaria officinalis</i>	Weed	Unassigned	Par o 1a=acidic allergen isoform {N-terminal} [Parietaria officinalis=pellitory, pollen, Peptide Partial, 25 aa]	1836010
1141	<i>Parietaria officinalis</i>	Weed	Par o 1	Par o 1b=basic allergen isoform {N-terminal} [Parietaria officinalis=pellitory, pollen, Peptide Partial, 24 aa]	1836011
1142	<i>Parietaria officinalis</i>	Weed	Par o 1	mAb 8B6-reactive major allergen {N-terminal} [Parietaria officinalis, pollen, Peptide Partial, 30 aa]	1311513
1143	<i>Parietaria officinalis</i>	Weed	Par o 1	mAb 3F8-reactive major allergen {N-terminal} [Parietaria officinalis, pollen, Peptide Partial, 15 aa]	1311512
1144	<i>Parietaria officinalis</i>	Weed	Par o 1	mAb 8C7-reactive major allergen {N-terminal, band 2} [Parietaria officinalis, pollen, Peptide Partial, 15 aa]	1311511
1145	<i>Parietaria officinalis</i>	Weed	Par o 1	mAb 8C7-reactive major allergen {N-terminal, band 1} [Parietaria officinalis, pollen, Peptide Partial, 15 aa]	1311510
1146	<i>Parietaria officinalis</i>	Weed	Par o 1	mAb 2F9-reactive major allergen {N-terminal} [Parietaria officinalis, pollen, Peptide Partial, 17 aa]	1311509
1147	<i>Paspalum notatum</i>	Bahia grass	Unassigned	group 13 grass pollen allergen [Paspalum notatum]	338930672
1148	<i>Paspalum notatum</i>	Bahia grass	Unassigned	group 13 grass pollen allergen [Paspalum notatum]	338930674
1149	<i>Paspalum notatum</i>	Bahia grass	Unassigned	group 13 grass pollen allergen [Paspalum notatum]	338930676
1150	<i>Paspalum notatum</i>	Bahia grass	Unassigned	group 13 grass pollen allergen [Paspalum notatum]	338930678
1151	<i>Paspalum notatum</i>	Bahia grass	Unassigned	group 13 grass pollen allergen [Paspalum notatum]	338930680
1152	<i>Paspalum notatum</i>	Bahia grass	Unassigned	group 13 grass pollen allergen [Paspalum notatum]	338930682



Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1153	<i>Paspalum notatum</i>	Bahia grass	Unassigned	group 13 grass pollen allergen [Paspalum notatum]	338930684
1154	<i>Paspalum notatum</i>	Bahia grass	Unassigned	group 13 grass pollen allergen [Paspalum notatum]	338930686
1155	<i>Paspalum notatum</i>	Bahia grass	Unassigned	Pas n 1 allergen precursor [Paspalum notatum]	168419914
1156	<i>Penaeus monodon</i>	Black tiger shrimp	Unassigned	tropomyosin fast isoform [Penaeus monodon]	125995157
1157	<i>Penaeus monodon</i>	Black tiger shrimp	Unassigned	RecName: Full=Arginine kinase; Short=AK; AltName: Allergen=Pen m 2	308154236
1158	<i>Penaeus monodon</i>	Black tiger shrimp	Pen m 2	allergen Pen m 2 [Penaeus monodon]	27463265
1159	<i>Penaeus monodon</i>	Black tiger shrimp	Unassigned	myosin light chain [Penaeus monodon]	317383196
1160	<i>Penaeus monodon</i>	Black tiger shrimp	Unassigned	sarcoplasmic calcium binding protein [Penaeus monodon]	317383198
1161	<i>Penicillium brevicompactum</i>	Fungus	Unassigned	60S acidic ribosomal P1 phosphoprotein Pen b 26 [Penicillium brevicompactum]	59894749
1162	<i>Penicillium chrysogenum</i>	Fungus	Pen ch 18	vacuolar serine protease [Penicillium chrysogenum]	14215732
1163	<i>Penicillium chrysogenum</i>	Fungus	Pen ch 18	allergen Pen n 18 [Penicillium chrysogenum]	7963902
1164	<i>Penicillium chrysogenum</i>	Fungus	Pen ch 13	alkaline serine protease [Penicillium chrysogenum]	21069093
1165	<i>Penicillium chrysogenum</i>	Fungus	Pen ch 13	allergen Pen n 13 [Penicillium chrysogenum]	6684758
1166	<i>Penicillium chrysogenum</i>	Fungus	Pen ch 20	68 kDa allergen [Penicillium chrysogenum]	999009
1167	<i>Penicillium citrinum</i>	Fungus	Unassigned	vacuolar serine protease [Penicillium citrinum]	12005501
1168	<i>Penicillium citrinum</i>	Fungus	Unassigned	alkaline serine protease Pen c2 [Penicillium citrinum]	4588118
1169	<i>Penicillium citrinum</i>	Fungus	Pen c 19	Heat shock 70 kDa protein (Allergen Pen c 19)	14423733
1170	<i>Penicillium citrinum</i>	Fungus	Unassigned	Enolase (2-phosphoglycerate dehydratase) (2-phospho-D-glycerate hydro-lyase) (Allergen Pen c 22)	74664773
1171	<i>Penicillium citrinum</i>	Fungus	Pen c 24	unknown [Penicillium citrinum]	38326693
1172	<i>Penicillium citrinum</i>	Fungus	Pen c 3	peroxisomal membrane protein [Penicillium citrinum]	5326864
1173	<i>Penicillium citrinum</i>	Fungus	Unassigned	Pen c 1; alkaline serine protease [Penicillium citrinum]	4587983
1174	<i>Penicillium oxalicum</i>	Fungus	Pen o 18	vacuolar serine protease [Penicillium oxalicum]	12005497
1175	<i>Periplaneta americana</i>	American cockroach	Unassigned	tropomyosin [Periplaneta americana]	239740599
1176	<i>Periplaneta americana</i>	American cockroach	Unassigned	Tropomyosin (Major allergen Per a 7)	14423957

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1177	<i>Periplaneta americana</i>	American cockroach	Per a 7.0101	tropomyosin [Periplaneta americana]	4378573
1178	<i>Periplaneta americana</i>	American cockroach	Unassigned	major allergen Cr-P II [Periplaneta americana]	284518361
1179	<i>Periplaneta americana</i>	American cockroach	Unassigned	major allergen Per a 1.0105 [Periplaneta americana]	30144660
1180	<i>Periplaneta americana</i>	American cockroach	Per a 1.0101	major allergen Per a 1.0101 [Periplaneta americana]	4240399
1181	<i>Periplaneta americana</i>	American cockroach	Per a 1.0102	allergen [Periplaneta americana]	2897849
1182	<i>Periplaneta americana</i>	American cockroach	Per a 1	Cr-P II [Periplaneta americana]	2580504
1183	<i>Periplaneta americana</i>	American cockroach	Per a 1.0104	Cr-P II allergen [Periplaneta americana]	2253610
1184	<i>Periplaneta americana</i>	American cockroach	Per a 1	Cr-P II allergen [Periplaneta americana]	2231297
1185	<i>Periplaneta americana</i>	American cockroach	Unassigned	Per a 3 allergen [Periplaneta americana]	289721058
1186	<i>Periplaneta americana</i>	American cockroach	Unassigned	major allergen Cr-P I [Periplaneta americana]	284518363
1187	<i>Periplaneta americana</i>	American cockroach	Per a 3.0101	Allergen Cr-P I precursor (Allergen Per a 3)	2833325
1188	<i>Periplaneta americana</i>	American cockroach	Per a 3.0203	allergen [Periplaneta americana]	1580797
1189	<i>Periplaneta americana</i>	American cockroach	Per a 3.0202	allergen [Periplaneta americana]	1580794
1190	<i>Periplaneta americana</i>	American cockroach	Per a 3.0201	allergen [Periplaneta americana]	1531589
1191	<i>Periplaneta americana</i>	American cockroach	Unassigned	Per a 4 allergen variant 1 [Periplaneta americana]	212675312
1192	<i>Periplaneta americana</i>	American cockroach	Unassigned	Chain A, Crystal Structure Of Major Allergens, Per A 4 From Cockroaches	215794707
1193	<i>Periplaneta americana</i>	American cockroach	Unassigned	Per a 4 allergen [Periplaneta americana]	60678787
1194	<i>Periplaneta fuliginosa</i>	Smokybrown cockroach	Unassigned	tropomyosin [Periplaneta fuliginosa]	19310971
1195	<i>Perna viridis</i>	Asian green mussel	Unassigned	tropomyosin [Perna viridis]	9954251
1196	<i>Persea americana</i>	Avocado	Pers a 1	endochitinase [Persea americana]	3201547
1197	<i>Phalaris aquatica</i>	Canary grass	Pha a 1	Major pollen allergen Pha a 1 precursor (Pha a I)	2498576
1198	<i>Phalaris aquatica</i>	Canary grass	Unassigned	Pha a I=34 kda pollen allergen {N-terminal} [Phalaris aquatica=canary grass, Peptide Partial, 20 aa]	409328
1199	<i>Phalaris aquatica</i>	Canary grass	Unassigned	Major pollen allergen Pha a 5.4 (Pha a 5) (Clone 5)	2498580
1200	<i>Phalaris aquatica</i>	Canary grass	Unassigned	Major pollen allergen Pha a 5.3 precursor (Pha a 5) (Clone 29)	2498579

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1201	<i>Phalaris aquatica</i>	Canary grass	Unassigned	Major pollen allergen Pha a 5.2 precursor (Pha a 5) (Clone 14)	2498578
1202	<i>Phalaris aquatica</i>	Canary grass	Unassigned	Major pollen allergen Pha a 5.1 precursor (Pha A 5) (Clone 28)	2498577
1203	<i>Phaseolus vulgaris</i>	Kidney bean	Unassigned	non-specific lipid transfer protein 1b precursor [Phaseolus vulgaris]	289064179
1204	<i>Phaseolus vulgaris</i>	Kidney bean	Unassigned	non-specific lipid transfer protein 1a precursor [Phaseolus vulgaris]	289064177
1205	<i>Phleum pratense</i>	Common timothy	Unassigned	allergen Phl p I	1582250
1206	<i>Phleum pratense</i>	Common timothy	Unassigned	Pollen allergen Phl p 1 precursor (Phl p I)	1171008
1207	<i>Phleum pratense</i>	Common timothy	Unassigned	unnamed protein product [Phleum pratense]	45823012
1208	<i>Phleum pratense</i>	Common timothy	Phl p 1	Chain A, Crystal Structure Of Phl P 1, A Major Timothy Grass Pollen Allergen	28373838
1209	<i>Phleum pratense</i>	Common timothy	Phl p 1.0101	pollen allergen Phl pI [Phleum pratense]	3901094
1210	<i>Phleum pratense</i>	Common timothy	Unassigned	Pollen allergen Phl p 11	47606039
1211	<i>Phleum pratense</i>	Common timothy	Phl p 12	profilin 4 [Phleum pratense]	2415702
1212	<i>Phleum pratense</i>	Common timothy	Phl p 12	profilin 3 [Phleum pratense]	2415700
1213	<i>Phleum pratense</i>	Common timothy	Phl p 12	Profilin-1 (Pollen allergen Phl p 12) (Phl p 11)	464471
1214	<i>Phleum pratense</i>	Common timothy	Phl p 13	polygalacturonase [Phleum pratense]	4826572
1215	<i>Phleum pratense</i>	Common timothy	Unassigned	Pollen allergen Phl p 2 precursor (Phl p II)	1171009
1216	<i>Phleum pratense</i>	Common timothy	Unassigned	pollen allergen Phl p 4.0204 [Phleum pratense]	189014272
1217	<i>Phleum pratense</i>	Common timothy	Unassigned	pollen allergen Phl p 4.0203 [Phleum pratense]	189014270
1218	<i>Phleum pratense</i>	Common timothy	Unassigned	pollen allergen Phl p 4.0202 [Phleum pratense]	189014268
1219	<i>Phleum pratense</i>	Common timothy	Unassigned	pollen allergen Phl p 4.0102 [Phleum pratense]	189014266
1220	<i>Phleum pratense</i>	Common timothy	Unassigned	major pollen allergen Phl p 4 precursor [Phleum pratense]	82492267
1221	<i>Phleum pratense</i>	Common timothy	Unassigned	unnamed protein product [Phleum pratense]	45108967
1222	<i>Phleum pratense</i>	Common timothy	Unassigned	unnamed protein product [Phleum pratense]	45108973

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1223	<i>Phleum pratense</i>	Common timothy	Unassigned	pollen allergen Phl p 4 [Phleum pratense]	54144332
1224	<i>Phleum pratense</i>	Common timothy	Phl p 5	Pollen allergen Phl p V	75139900
1225	<i>Phleum pratense</i>	Common timothy	Unassigned	phl p5a allergen precursor [Phleum pratense]	29500897
1226	<i>Phleum pratense</i>	Common timothy	Phl p 5	Chain A, Crystal Structure Of The Functional Domain Of The Major Grass Pollen Allergen Phl P 5b	28948464
1227	<i>Phleum pratense</i>	Common timothy	Unassigned	unnamed protein product [Phleum pratense]	21725632
1228	<i>Phleum pratense</i>	Common timothy	Unassigned	unnamed protein product [Phleum pratense]	21725630
1229	<i>Phleum pratense</i>	Common timothy	Unassigned	unnamed protein product [Phleum pratense]	21725628
1230	<i>Phleum pratense</i>	Common timothy	Unassigned	unnamed protein product [Phleum pratense]	21725626
1231	<i>Phleum pratense</i>	Common timothy	Unassigned	unnamed protein product [Phleum pratense]	21725624
1232	<i>Phleum pratense</i>	Common timothy	Unassigned	unnamed protein product [Phleum pratense]	21725622
1233	<i>Phleum pratense</i>	Common timothy	Unassigned	unnamed protein product [Phleum pratense]	21725620
1234	<i>Phleum pratense</i>	Common timothy	Unassigned	unnamed protein product [Phleum pratense]	21725618
1235	<i>Phleum pratense</i>	Common timothy	Unassigned	unnamed protein product [Phleum pratense]	21725616
1236	<i>Phleum pratense</i>	Common timothy	Unassigned	unnamed protein product [Phleum pratense]	21725614
1237	<i>Phleum pratense</i>	Common timothy	Unassigned	unnamed protein product [Phleum pratense]	21725612
1238	<i>Phleum pratense</i>	Common timothy	Unassigned	unnamed protein product [Phleum pratense]	21725610
1239	<i>Phleum pratense</i>	Common timothy	Unassigned	unnamed protein product [Phleum pratense]	21725608
1240	<i>Phleum pratense</i>	Common timothy	Unassigned	unnamed protein product [Phleum pratense]	21725606
1241	<i>Phleum pratense</i>	Common timothy	Phl p 5	group V allergen Phl p 5 precursor [Phleum pratense]	13430402
1242	<i>Phleum pratense</i>	Common timothy	Unassigned	group V allergen Phl p 5.0207 precursor [Phleum pratense]	3309047
1243	<i>Phleum pratense</i>	Common timothy	Unassigned	group V allergen Phl p 5.0206 precursor [Phleum pratense]	3309045
1244	<i>Phleum pratense</i>	Common timothy	Unassigned	group V allergen Phl p 5.0203 precursor [Phleum pratense]	3309041
1245	<i>Phleum pratense</i>	Common timothy	Phl p 5.0103	group V allergen Phl p 5.0103 precursor [Phleum pratense]	3309039
1246	<i>Phleum pratense</i>	Common timothy	Phl p 5.0108	major allergen Phl p 5 [Phleum pratense]	3135503
1247	<i>Phleum pratense</i>	Common timothy	Phl p 5.0107	major allergen Phl p 5 [Phleum pratense]	3135501
1248	<i>Phleum pratense</i>	Common timothy	Phl p 5.0106	major allergen Phl p 5 [Phleum pratense]	3135499

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1249	<i>Phleum pratense</i>	Common timothy	Phl p 5.0105	major allergen Phl p 5 [Phleum pratense]	3135497
1250	<i>Phleum pratense</i>	Common timothy	Phl p 5	Pollen allergen Phl p 5b precursor (Phl p Vb)	2851457
1251	<i>Phleum pratense</i>	Common timothy	Phl p 5	Major Pollen Allergen Phl p Va [Phleum pratense]	2398757
1252	<i>Phleum pratense</i>	Common timothy	Phl p 5.0104	major allergen Phl p 5 [Phleum pratense]	1684720
1253	<i>Phleum pratense</i>	Common timothy	Phl p 5.0202	major allergen Phl p 5 [Phleum pratense]	1684718
1254	<i>Phleum pratense</i>	Common timothy	Unassigned	major allergen Phl p Va	1092249
1255	<i>Phleum pratense</i>	Common timothy	Phl p 5	allergen Phl p Vb - common timothy	481397
1256	<i>Phleum pratense</i>	Common timothy	Phl p 5	PHLP5A protein - common timothy (fragment)	422005
1257	<i>Phleum pratense</i>	Common timothy	Phl p 5.0101	Phlp5 [Phleum pratense]	398830
1258	<i>Phleum pratense</i>	Common timothy	Unassigned	Chain N, Crystal Structure Of Phl P 6, A Major Timothy Grass Pollen Allergen Co-Crystallized With Zinc	28374072
1259	<i>Phleum pratense</i>	Common timothy	Unassigned	Phl p6 IgE binding fragment [Phleum pratense]	3004469
1260	<i>Phleum pratense</i>	Common timothy	Phl p 6	Phl p6 allergen [Phleum pratense]	3004467
1261	<i>Phleum pratense</i>	Common timothy	Phl p 6	Phl p6 allergen [Phleum pratense]	3004465
1262	<i>Phleum pratense</i>	Common timothy	Unassigned	Polcalcin Phl p 7 (Calcium-binding pollen allergen Phl p 7) (P7)	14423846
1263	<i>Phoenix dactylifera</i>	Date palm	Pho d 2	profilin [Phoenix dactylifera]	21322677
1264	<i>Pistacia vera</i>		Unassigned	Pis v 2.0201 allergen 11S globulin precursor [Pistacia vera]	110349085
1265	<i>Pistacia vera</i>		Unassigned	Pis v 2.0101 allergen 11S globulin precursor [Pistacia vera]	110349083
1266	<i>Pistacia vera</i>		Unassigned	11S globulin [Pistacia vera]	156001070
1267	<i>Pistacia vera</i>		Unassigned	Pis v 1 allergen 2S albumin [Pistacia vera]	110349081
1268	<i>Pistacia vera</i>		Unassigned	vicilin [Pistacia vera]	133711974
1269	<i>Pisum sativum</i>	Pea	Pis s 1	Vicilin [Pisum sativum]	42414627
1270	<i>Pisum sativum</i>	Pea	Pis s 1	Vicilin [Pisum sativum]	42414629
1271	<i>Plantago lanceolata</i>	Narrow-leaved plantain	Unassigned	unnamed protein product [Plantago lanceolata]	29163773
1272	<i>Plantago lanceolata</i>	Narrow-leaved plantain	Pla l 1	plantain pollen major allergen, Pla l 1.0103 [Plantago lanceolata]	14422363

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1273	<i>Plantago lanceolata</i>	Narrow-leaved plantain	Pla l 1	plantain pollen major allergen, Pla l 1.0102 [Plantago lanceolata]	14422361
1274	<i>Plantago lanceolata</i>	Narrow-leaved plantain	Pla l 1	plantain pollen major allergen, Pla l 1.0101 [Plantago lanceolata]	14422359
1275	<i>Platanus orientalis</i>		Unassigned	pollen allergen Pla o 1 [Platanus orientalis]	162949336
1276	<i>Platanus x acerifolia</i>	London plane tree	Unassigned	Putative invertase inhibitor precursor (Pollen allergen Pla a 1)	29839547
1277	<i>Platanus x acerifolia</i>	London plane tree	Pla a 2	polygalacturonase [Platanus x acerifolia]	49523394
1278	<i>Plodia interpunctella</i>	Indian meal moth	Unassigned	arginine kinase [Plodia interpunctella]	15886861
1279	<i>Poa pratensis</i>	Kentucky bluegrass	Unassigned	pollen allergen (group II) [Poa pratensis]	4007655
1280	<i>Poa pratensis</i>	Kentucky bluegrass	Poa p 1	group I pollen allergen [Poa pratensis]	4090265
1281	<i>Poa pratensis</i>	Kentucky bluegrass	Poa p 1	pollen allergen Poa p I - Kentucky bluegrass (fragment)	320620
1282	<i>Poa pratensis</i>	Kentucky bluegrass	Poa p 1	pollen allergen Poa-pI - Kentucky bluegrass (fragment)	280414
1283	<i>Poa pratensis</i>	Kentucky bluegrass	Poa p 5	pollen allergen Poa p 5 [Poa pratensis]	11991227
1284	<i>Poa pratensis</i>	Kentucky bluegrass	Unassigned	Pollen allergen KBG 31 precursor (Pollen allergen Poa p 9) (Poa p IX)	113560
1285	<i>Poa pratensis</i>	Kentucky bluegrass	Unassigned	Pollen allergen KBG 41 precursor (Pollen allergen Poa p 9) (Poa p IX)	113561
1286	<i>Poa pratensis</i>	Kentucky bluegrass	Unassigned	pollen allergen (clone 7.2) - Kentucky bluegrass (fragment)	539056
1287	<i>Poa pratensis</i>	Kentucky bluegrass	Unassigned	Pollen allergen KBG 60 precursor (Pollen allergen Poa p 9) (Poa p IX)	113562
1288	<i>Polistes annularis</i>	Paper wasp	Pol a 1	Phospholipase A1 (Allergen Pol a 1)	14423833
1289	<i>Polistes annularis</i>	Paper wasp	Pol a 2	Hyaluronoglucosaminidase precursor (Hyaluronidase) (Allergen Pol a 2)	14423735
1290	<i>Polistes annularis</i>	Paper wasp	Pol a 5	allergen 5	160780
1291	<i>Polistes dominulus</i>	Paper wasp	Unassigned	venom phospholipase A1 1 precursor [Polistes dominulus]	45510887
1292	<i>Polistes dominulus</i>	Paper wasp	Unassigned	venom phospholipase A1 2 precursor [Polistes dominulus]	45510889
1293	<i>Polistes dominulus</i>	Paper wasp	Unassigned	venom phospholipase A1 3 precursor [Polistes dominulus]	45510891

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1294	<i>Polistes dominulus</i>	Paper wasp	Unassigned	venom phospholipase A1 4 precursor [Polistes dominulus]	45510893
1295	<i>Polistes dominulus</i>	Paper wasp	Pol d 5	allergen Pol d 5 precursor [Polistes dominulus]	51093377
1296	<i>Polistes dominulus</i>	Paper wasp	Unassigned	venom serine protease precursor [Polistes dominulus]	30909091
1297	<i>Polistes exclamans</i>	Paper wasp	Unassigned	allergen Pol e 5 precursor [Polistes exclamans]	51093375
1298	<i>Polistes exclamans</i>	Paper wasp	Pol e 5	Venom allergen 5 (Antigen 5) (Ag5) (Allergen Pol e 5) (Pol e V)	549187
1299	<i>Polistes fuscatus</i>	Paper wasp	Pol f 5	Venom allergen 5 (Antigen 5) (Ag5) (Allergen Pol f 5) (Pol f V)	549188
1300	<i>Polistes gallicus</i>	Paper wasp	Unassigned	Phospholipase A1 (Allergen Pol g 1)	41017429
1301	<i>Polistes gallicus</i>	Paper wasp	Pol g 5	Venom allergen 5 (Antigen 5) (Ag5) (Allergen Pol g 5)	25091511
1302	<i>Polybia paulista</i>	wasp	Unassigned	RecName: Full=Hyaluronidase; Short=Hya; AltName: Full=Hyaluronoglucosaminidase	302425085
1303	<i>Polybia paulista</i>	wasp	Unassigned	hyaluronidase precursor [Polybia paulista]	302201583
1304	<i>Polybia paulista</i>	wasp	Unassigned	RecName: Full=Venom allergen 5; AltName: Full=Antigen 5; Short=Ag5	302595972
1305	<i>Polybia paulista</i>	wasp	Unassigned	venom allergen 5 [Polybia paulista]	290792375
1306	<i>Polybia paulista</i>	wasp	Unassigned	phospholipase A1 [Polybia paulista]	315190620
1307	<i>Polybia paulista</i>	wasp	Pol p 1.0101	Venom phospholipase A1 precursor (Allergen Poly p 1)	166216292
1308	<i>Protortonia cacti</i>		Unassigned	cochineal major allergen [Protortonia cacti]	237769615
1309	<i>Prunus armeniaca</i>	Apricot	Unassigned	non-specific lipid transfer protein [Prunus armeniaca]	313575736
1310	<i>Prunus armeniaca</i>	Apricot	Unassigned	non-specific lipid transfer protein [Prunus armeniaca]	313575734
1311	<i>Prunus armeniaca</i>	Apricot	Unassigned	non-specific lipid transfer protein [Prunus armeniaca]	313575732
1312	<i>Prunus armeniaca</i>	Apricot	Unassigned	non-specific lipid transfer protein [Prunus armeniaca]	313575730
1313	<i>Prunus avium</i>	Cherry	Unassigned	Chain A, Solution Structure Of The Major Cherry Allergen Pru Av 1 Mutant E45w	159162378
1314	<i>Prunus avium</i>	Cherry	Pru av 1	major cherry allergen Pru av 1.0201 [Prunus avium]	44409451
1315	<i>Prunus avium</i>	Cherry	Pru av 1	major cherry allergen Pru av 1.0202 [Prunus avium]	44409474
1316	<i>Prunus avium</i>	Cherry	Pru av 1	major cherry allergen Pru av 1.0203 [Prunus avium]	44409496

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1317	<i>Prunus avium</i>	Cherry	Pru av 1	cherry-allergen PRUA1	1513216
1318	<i>Prunus avium</i>	Cherry	Unassigned	non-specific lipid transfer protein [Prunus avium]	313575728
1319	<i>Prunus avium</i>	Cherry	Unassigned	non-specific lipid transfer protein [Prunus avium]	313575726
1320	<i>Prunus avium</i>	Cherry	Pru av 3	lipid transfer protein precursor [Prunus avium]	6715520
1321	<i>Prunus avium</i>	Cherry	Pru av 4	profilin [Prunus avium]	4761582
1322	<i>Prunus avium</i>	Cherry	Pru av 2	thaumatin-like protein precursor	1144346
1323	<i>Prunus domestica</i>	Plum	Pru d 3	Nonspecific lipid-transfer protein 1 (LTP 1) (Major allergen Pru d 3)	9297015
1324	<i>Prunus dulcis</i>	Almond	Pru du 4	profilin [Prunus dulcis]	24473794
1325	<i>Prunus dulcis</i>	Almond	Unassigned	Chain A, Crystal Structure Of Pru Du Amandin, An Allergenic Protein From Prunus Dulcis	258588247
1326	<i>Prunus dulcis</i>	Almond	Unassigned	Seed allergenic protein 2 (Conglutin gamma) (Allergen Pru du ?)	75107131
1327	<i>Prunus dulcis x Prunus persica</i>		Unassigned	putative allergen Pru du 4.02 [Prunus dulcis x Prunus persica]	190613937
1328	<i>Prunus dulcis x Prunus persica</i>		Unassigned	putative allergen Pru du 4.01 [Prunus dulcis x Prunus persica]	190613933
1329	<i>Prunus dulcis x Prunus persica</i>		Pru p 2.0301	putative allergen Pru p 2.02 [Prunus dulcis x Prunus persica]	190613903
1330	<i>Prunus dulcis x Prunus persica</i>		Pru p 2.0101	putative allergen Pru p 2.01B [Prunus dulcis x Prunus persica]	190613911
1331	<i>Prunus dulcis x Prunus persica</i>		Pru p 2.0201	putative allergen Pru p 2.01A [Prunus dulcis x Prunus persica]	190613907
1332	<i>Prunus persica</i>	Peach	Unassigned	major allergen Pru p 1 [Prunus persica]	82492265
1333	<i>Prunus persica</i>	Peach	Unassigned	non-specific lipid transfer protein [Prunus persica]	313575718
1334	<i>Prunus persica</i>	Peach	Pru p 3	Chain B, Crystal Structure Of Peach Pru P3, The Prototypic Member Of The Family Of Plant Non-Specific Lipid Transfer Protein Pan-Allergens	83754241
1335	<i>Prunus persica</i>	Peach	Unassigned	lipid transfer protein 1 precursor [Prunus persica]	54793477
1336	<i>Prunus persica</i>	Peach	Pru p 3	Nonspecific lipid-transfer protein 1 (LTP 1) (Major allergen Pru p 3) (Pru p 1)	3287877



Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1337	<i>Prunus persica</i>	Peach	Pru p 4.02	profilin [Prunus persica]	27528312
1338	<i>Prunus persica</i>	Peach	Pru p 4.01	profilin [Prunus persica]	27528310
1339	<i>Pseudocardium sachalinensis</i>		Unassigned	tropomyosin [Pseudocardium sachalinensis]	219806598
1340	<i>Pyrus communis</i>	Pear	Pyr c 1	major allergen Pyrc1 [Pyrus communis]	3044216
1341	<i>Pyrus communis</i>	Pear	Pyr c 4	profilin [Pyrus communis]	4761580
1342	<i>Pyrus communis</i>	Pear	Pyr c 5	isoflavone reductase related protein [Pyrus communis]	3243234
1343	<i>Quercus alba</i>	Oak	Unassigned	pollen allergen Que a 1 isoform [Quercus alba]	167472849
1344	<i>Quercus alba</i>	Oak	Unassigned	pollen allergen Que a 1 isoform [Quercus alba]	167472847
1345	<i>Quercus alba</i>	Oak	Que a 1	major pollen allergen Que a I - white oak (fragment)	543675
1346	<i>Rana esculenta</i>	Frog	Ran e 1	parvalbumin alpha [Rana esculenta]	20796729
1347	<i>Rana esculenta</i>	Frog	Ran e 2	parvalbumin beta protein [Rana esculenta]	20797081
1348	<i>Rana</i> sp. CH-2001	Frog	Unassigned	parvalbumin alpha [Rana sp. CH-2001]	20796733
1349	<i>Rana</i> sp. CH-2001	Frog	Unassigned	parvalbumin beta protein [Rana sp. CH-2001]	20797085
1350	<i>Rattus norvegicus</i>	Rat	Unassigned	PREDICTED: similar to Major urinary protein precursor (MUP) (Alpha-2u-globulin) (Alpha(2)-euglobulin) (Allergen Rat n 1) (Rat n I) [Rattus norvegicus]	109474987
1351	<i>Rattus norvegicus</i>	Rat	Rat n 1	Alpha-2u globulin	81890324
1352	<i>Rattus norvegicus</i>	Rat	Rat n 1	Major urinary protein precursor (MUP) (Alpha-2u-globulin) (15.5 kDa fatty acid binding protein) (15.5 kDa FABP) (Alpha(2)-euglobulin) (Allergen Rat n 1) (Rat n I)	127533
1353	<i>Rhodotorula mucilaginosa</i>	Fungus	Unassigned	Enolase (2-phosphoglycerate dehydratase) (2-phospho-D-glycerate hydro-lyase) (Allergen Rho m 1)	37078092
1354	<i>Rhodotorula mucilaginosa</i>	Fungus	Unassigned	vacuolar serine protease [Rhodotorula mucilaginosa]	54654335
1355	<i>Ricinus communis</i>	Castor bean	Ric c 1	2S albumin precursor (Allergen Ric c 1)	112762
1356	<i>Rubus idaeus</i>		Unassigned	putative allergen Rub i 1 [Rubus idaeus]	110180525
1357	<i>Rubus idaeus</i>		Unassigned	putative allergen Rub i 3 [Rubus idaeus]	110180523
1358	<i>Salmo salar</i>	Salmon	Unassigned	Parvalbumin beta 1 [Salmo salar]	209734468
1359	<i>Salmo salar</i>	Salmon	Sal s 1	Parvalbumin beta 2 (Major allergen Sal s 1)	18281421
1360	<i>Salmo salar</i>	Salmon	Sal s 1	Parvalbumin beta 1 (Major allergen Sal s 1)	2493445

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1361	<i>Salsola kali</i>	Thistle	Unassigned	Sal k 1 pollen allergen [Salsola kali]	225810597
1362	<i>Salsola kali</i>	Thistle	Unassigned	pectin methylesterase allergenic protein [Salsola kali]	59895730
1363	<i>Salsola kali</i>	Thistle	Unassigned	pectin methylesterase allergenic protein [Salsola kali]	59895728
1364	<i>Salsola kali</i>	Thistle	Unassigned	pectin-methyltransferase precursor [Salsola kali]	51242679
1365	<i>Salsola kali</i>	Thistle	Sal k 1	[Segment 4 of 4] Pollen allergen Sal k 1	25090951
1366	<i>Salsola kali</i>	Thistle	Sal k 1	[Segment 3 of 4] Pollen allergen Sal k 1	25090950
1367	<i>Salsola kali</i>	Thistle	Sal k 1	[Segment 2 of 4] Pollen allergen Sal k 1	25090949
1368	<i>Salsola kali</i>	Thistle	Sal k 1	[Segment 1 of 4] Pollen allergen Sal k 1	25090948
1369	<i>Salsola kali</i>	Thistle	Unassigned	Sal k 3 pollen allergen [Salsola kali]	225810599
1370	<i>Salsola kali</i>	Thistle	Unassigned	Sal k 4 pollen allergen [Salsola kali]	239916566
1371	<i>Salvelinus fontinalis</i>	Brook trout	Unassigned	parvalbumin [Salvelinus fontinalis]	288557440
1372	<i>Salvelinus fontinalis</i>	Brook trout	Unassigned	parvalbumin [Salvelinus fontinalis]	288557438
1373	<i>Sarcoptes scabiei type hominis</i>	Scabies mite	Unassigned	major allergen 1 [Sarcoptes scabiei type hominis]	27462848
1374	<i>Sarcoptes scabiei type hominis</i>	Scabies mite	Unassigned	Sar s 1 allergen Yv5032C08 [Sarcoptes scabiei type hominis]	46406002
1375	<i>Sarcoptes scabiei type hominis</i>	Scabies mite	Unassigned	Sar s 1 allergen SMIPP-C Yv6018B11 [Sarcoptes scabiei type hominis]	46406016
1376	<i>Sarcoptes scabiei type hominis</i>	Scabies mite	Unassigned	Sar s 1 allergen SMIPP-C Yv5009F04 [Sarcoptes scabiei type hominis]	46406014
1377	<i>Sarcoptes scabiei type hominis</i>	Scabies mite	Unassigned	Sar s 1 allergen SMIPP-C Yv4028C12 [Sarcoptes scabiei type hominis]	46406012
1378	<i>Sarcoptes scabiei type hominis</i>	Scabies mite	Unassigned	glutathione transferase mu class Yv5004H11 [Sarcoptes scabiei type hominis]	60920770
1379	<i>Sarcoptes scabiei type hominis</i>	Scabies mite	Unassigned	glutathione S-transferase [Sarcoptes scabiei type hominis]	27462836
1380	<i>Sardinops sagax</i>		Unassigned	parvalbumin [Sardinops sagax]	193247972
1381	<i>Scapharca broughtonii</i>		Unassigned	tropomyosin [Scapharca broughtonii]	219806592
1382	<i>Schedonorus arundinaceus</i>	Tall fescue	Unassigned	Group I allergen FeS e I, pollen	75139991
1383	<i>Schedonorus arundinaceus</i>	Tall fescue	Unassigned	pollen allergen Fes e I type B - reed fescue (fragment)	320611
1384	<i>Schedonorus arundinaceus</i>	Tall fescue	Unassigned	pollen allergen Fes e I type A - reed fescue (fragment)	320610

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1385	<i>Schistosoma japonicum</i>	Schistosoma	Unassigned	hypothetical protein, putative Profilin/allergen [Schistosoma japonicum]	29841461
1386	<i>Schistosoma japonicum</i>	Schistosoma	Unassigned	22.6 kDa tegumental antigen [Schistosoma japonicum]	2739154
1387	<i>Scomber japonicus</i>	Chub mackerel	Unassigned	parvalbumin [Scomber japonicus]	29420793
1388	<i>Scomber scombrus</i>	Atlantic mackerel	Unassigned	parvalbumin [Scomber scombrus]	288557436
1389	<i>Secale cereale</i>	Rye	Unassigned	pollen allergen Sec c 5 [Secale cereale]	332205751
1390	<i>Secale cereale</i>	Rye	Unassigned	30K allergen	75140047
1391	<i>Secale cereale</i>	Rye	Sec c 1	Major BAKER'S ASTHMA allergen SEC C 1	75198875
1392	<i>Secale cereale</i>	Rye	Unassigned	pollen allergen Sec c 4 [Secale cereale]	55859454
1393	<i>Secale cereale</i>	Rye	Unassigned	pollen allergen Sec c 4 [Secale cereale]	55859456
1394	<i>Sepia esculenta</i>		Unassigned	tropomyosin [Sepia esculenta]	83715928
1395	<i>Sepioteuthis lessoniana</i>		Unassigned	tropomyosin [Sepioteuthis lessoniana]	83715930
1396	<i>Sesamum indicum</i>	Sesame	Unassigned	2S albumin [Sesamum indicum]	209165427
1397	<i>Sesamum indicum</i>	Sesame	Ses i 1	2S albumin [Sesamum indicum]	13183175
1398	<i>Sesamum indicum</i>	Sesame	Ses i 2	2S albumin precursor [Sesamum indicum]	5381323
1399	<i>Sesamum indicum</i>	Sesame	Ses i 3	7S globulin [Sesamum indicum]	13183177
1400	<i>Sesamum indicum</i>	Sesame	Unassigned	15 kDa oleosin	75315271
1401	<i>Sesamum indicum</i>	Sesame	Unassigned	main allergen 15 kDa oleosin [Sesamum indicum]	198250343
1402	<i>Sinapis alba</i>	White mustard	Sin a 1	Allergen Sin a 1 precursor (Sin a I) [Contains: Allergen Sin a 1 small chain; Allergen Sin a 1 large chain]	51338758
1403	<i>Sinapis alba</i>	White mustard	Sin a 1	allergen sin a 1.0108 [Sinapis alba]	1009442
1404	<i>Sinapis alba</i>	White mustard	Sin a 1	allergen sin a 1.0107 [Sinapis alba]	1009440
1405	<i>Sinapis alba</i>	White mustard	Sin a 1	allergen sin a 1.0106 [Sinapis alba]	1009438
1406	<i>Sinapis alba</i>	White mustard	Sin a 1	allergen sin a 1.0105 [Sinapis alba]	1009436
1407	<i>Sinapis alba</i>	White mustard	Sin a 1	allergen sin a 1.0104 [Sinapis alba]	1009434
1408	<i>Sinapis alba</i>	White mustard	Unassigned	11S globulin precursor [Sinapis alba]	62240392
1409	<i>Sinapis alba</i>	White mustard	Sin a 2.0101	11S globulin precursor [Sinapis alba]	62240390
1410	<i>Sinapis alba</i>	White mustard	Sin a 3.0101	Sin a 3 allergen [Sinapis alba]	156778059

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1411	<i>Sinapis alba</i>	White mustard	Sin a 4.0101	Sin a 4 allergen [ <i>Sinapis alba</i> ]	156778061
1412	<i>Solanum tuberosum</i>	Potato	Unassigned	profilin-like protein [ <i>Solanum tuberosum</i> ]	77999277
1413	<i>Solanum tuberosum</i>	Potato	Unassigned	profilin-like [ <i>Solanum tuberosum</i> ]	77416979
1414	<i>Solanum tuberosum</i>	Potato	Sola t 1	RecName: Full=Patatin-B1; Flags: Precursor	158517845
1415	<i>Solanum tuberosum</i>	Potato	Unassigned	patatin	169500
1416	<i>Solanum tuberosum</i>	Potato	Unassigned	patatin [ <i>Solanum tuberosum</i> ]	21514
1417	<i>Solanum tuberosum</i>	Potato	Unassigned	patatin [ <i>Solanum tuberosum</i> ]	21512
1418	<i>Solanum tuberosum</i>	Potato	Unassigned	patatin B2 (AA 1 - 386) [ <i>Solanum tuberosum</i> ]	21510
1419	<i>Solanum tuberosum</i>	Potato	Sola t 2	Aspartic protease inhibitor 11 (Cathepsin D inhibitor PDI) (Allergen Sola t 2)	124148
1420	<i>Solanum tuberosum</i>	Potato	Sola t 3	Cysteine protease inhibitor 1 precursor (PCPI 8.3) (P340) (P34021)	20141344
1421	<i>Solanum tuberosum</i>	Potato	Sola t 4	Serine protease inhibitor 7 precursor (PIG) (PIGEN1) (Allergen Sola t 4) (STPIB) (STPIA) (pKEN14-28) (pF4)	20141714
1422	<i>Solanum tuberosum</i>	Potato	Unassigned	aspartic proteinase inhibitor [ <i>Solanum tuberosum</i> ]	21413
1423	<i>Solen strictus</i>		Unassigned	tropomyosin [ <i>Solen strictus</i> ]	219806602
1424	<i>Solenopsis geminata</i>	Tropical Fire Ant	Sol g 4	venom allergen Sol g 4.02 precursor [ <i>Solenopsis geminata</i> ]	7638030
1425	<i>Solenopsis geminata</i>	Tropical Fire Ant	Sol g 4	venom allergen Sol g 4.01 precursor [ <i>Solenopsis geminata</i> ]	7638028
1426	<i>Solenopsis invicta</i>	Red fire ant	Unassigned	allergen Sol i 1 precursor [ <i>Solenopsis invicta</i> ]	51093373
1427	<i>Solenopsis invicta</i>	Red fire ant	Unassigned	Sol i 1=antigen {N-terminal} [ <i>Solenopsis invicta</i> =imported fire ants, venom, Peptide Partial, 26 aa, segment 5 of 5]	1336813
1428	<i>Solenopsis invicta</i>	Red fire ant	Unassigned	Sol i 1=antigen {N-terminal} [ <i>Solenopsis invicta</i> =imported fire ants, venom, Peptide Partial, 26 aa, segment 4 of 5]	1336812
1429	<i>Solenopsis invicta</i>	Red fire ant	Unassigned	Sol i 1=antigen {N-terminal} [ <i>Solenopsis invicta</i> =imported fire ants, venom, Peptide Partial, 25 aa, segment 3 of 5]	1336811
1430	<i>Solenopsis invicta</i>	Red fire ant	Unassigned	Sol i 1=antigen {N-terminal} [ <i>Solenopsis invicta</i> =imported fire ants, venom, Peptide Partial, 58 aa, segment 1 of 5]	1336809

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1431	<i>Solenopsis invicta</i>	Red fire ant	Sol i 2	Venom allergen 2 precursor (Venom allergen II) (Allergen Sol i 2) (Sol i II)	549179
1432	<i>Solenopsis invicta</i>	Red fire ant	Sol i 3	Venom allergen III precursor (Allergen Sol i 3) (Sol i III)	14424466
1433	<i>Solenopsis invicta</i>	Red fire ant	Sol i 4	Venom allergen IV precursor (Allergen Sol i 4) (Sol i IV)	14424465
1434	<i>Solenopsis invicta</i>	Red fire ant	Sol i 4	venom allergen Sol i 4.02 precursor [Solenopsis invicta]	4038411
1435	<i>Solenopsis richteri</i>	Black fire ant	Unassigned	Venom allergen 2 (Venom allergen II) (Allergen Sol r 2) (Sol r II)	6136162
1436	<i>Solenopsis richteri</i>	Black fire ant	Unassigned	Venom allergen III (Allergen Sol r 3) (Sol r III)	6136163
1437	<i>Solenopsis saevissima</i>	Brazilian fire ant	Unassigned	Sol s 2 allergen [Solenopsis saevissima]	291092710
1438	<i>Staphylococcus aureus</i>		Unassigned	Chain B, Staphylococcal Enterotoxin A	1633233
1439	<i>Staphylococcus aureus</i>		Unassigned	staphylococcal enterotoxin B precursor (SEB) [Staphylococcus aureus]	83308249
1440	<i>Staphylococcus aureus</i>		Unassigned	RecName: Full=Enterotoxin type C-2; AltName: Full=SEC2; Flags: Precursor	462026
1441	<i>Staphylococcus aureus</i>		Unassigned	RecName: Full=Enterotoxin type D; AltName: Full=SED; Flags: Precursor	119654
1442	<i>Staphylococcus aureus</i>		Unassigned	RecName: Full=Toxic shock syndrome toxin-1; Short=TSST-1; Flags: Precursor	136457
1443	<i>Suidasia medanensis</i>		Unassigned	group 2 allergen Sui m 2 [Suidasia medanensis]	45738062
1444	<i>Sus scrofa</i>	Pig	Unassigned	RecName: Full=Pepsin A; Flags: Precursor	118572685
1445	<i>Syringa vulgaris</i>	Lilac	Syr v 1.0103	allergen-like protein Syr v I isoform 3 - Syringa vulgaris	631913
1446	<i>Syringa vulgaris</i>	Lilac	Syr v 1.0102	allergen-like protein Syr v I isoform 2 - Syringa vulgaris	631912
1447	<i>Syringa vulgaris</i>	Lilac	Syr v 1.0101	allergen-like protein Syr v I isoform 1 - Syringa vulgaris	631911
1448	<i>Tabanus yao</i>	Horse Fly	Tab y 1.0101	Tab y 3 allergen [Tabanus yao]	323473390
1449	<i>Tabanus yao</i>	Horse Fly	Tab y 2.0101	putative Tab y 2 allergen [Tabanus yao]	304273371
1450	<i>Tabanus yao</i>	Horse Fly	Tab y 5.0101	putative Tab y 1 allergen [Tabanus yao]	304273369
1451	<i>Thaumetopoea pityocampa</i>	Pine moth	Unassigned	Tha p 1 allergen [Thaumetopoea pityocampa]	301030229
1452	<i>Theragra chalcogramma</i>	Alaska pollock	Unassigned	parvalbumin [Theragra chalcogramma]	14531018

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1453	<i>Theragra chalcogramma</i>	Alaska pollock	Unassigned	parvalbumin [Theragra chalcogramma]	14531020
1454	<i>Todarodes pacificus</i>	Japanese flying squid	Unassigned	tropomyosin [Todarodes pacificus]	83715932
1455	<i>Trachurus japonicus</i>		Unassigned	dark muscle parvalbumin [Trachurus japonicus]	77799800
1456	<i>Tresus keenae</i>		Unassigned	tropomyosin [Tresus keenae]	219806600
1457	<i>Triatoma protracta</i>	Western conenose	Tria p 1	procalin [Triatoma protracta]	15426413
1458	<i>Trichophyton rubrum</i>	Fungus	Tri r 2	Tri r 2 allergen [Trichophyton rubrum]	5813790
1459	<i>Trichophyton rubrum</i>	Fungus	Tri r 4	Tri r 4 allergen [Trichophyton rubrum]	5813788
1460	<i>Trichophyton schoenleinii</i>	Fungus	Unassigned	RecName: Full=Subtilisin-like protease 6; AltName: Allergen=Tri m 2; Flags: Precursor	74663809
1461	<i>Trichophyton schoenleinii</i>	Fungus	Unassigned	tri s 4 allergen [Trichophyton schoenleinii]	23894227
1462	<i>Triticum aestivum</i>	Wheat	Unassigned	RecName: Full=Non-specific lipid-transfer protein; Short=LTP; AltName: Full=Phospholipid transfer protein; Short=PLTP; AltName: Full=ns-LTP1; Flags: Precursor	417370
1463	<i>Triticum aestivum</i>	Wheat	Unassigned	5a2 protein [Triticum aestivum]	66840998
1464	<i>Triticum aestivum</i>	Wheat	Unassigned	major allergen CM16 [Triticum aestivum]	195957140
1465	<i>Triticum aestivum</i>	Wheat	Unassigned	CM 17 protein precursor [Triticum aestivum]	21711
1466	<i>Triticum aestivum</i>	Wheat	Unassigned	unnamed protein product [Triticum aestivum]	21713
1467	<i>Triticum aestivum</i>	Wheat	Unassigned	alpha-gliadin [Triticum aestivum]	1304264
1468	<i>Triticum aestivum</i>	Wheat	Unassigned	alpha-gliadin	473876
1469	<i>Triticum aestivum</i>	Wheat	Unassigned	alpha-type gliadin	170728
1470	<i>Triticum aestivum</i>	Wheat	Unassigned	pre-alpha-/beta-gliadin A-III	170726
1471	<i>Triticum aestivum</i>	Wheat	Unassigned	pre-alpha-/beta-gliadin A-IV	170724
1472	<i>Triticum aestivum</i>	Wheat	Unassigned	pre-alpha-/beta-gliadin A-I	170722
1473	<i>Triticum aestivum</i>	Wheat	Unassigned	alpha/beta-gliadin precursor [Triticum aestivum]	170720
1474	<i>Triticum aestivum</i>	Wheat	Unassigned	alpha/beta-gliadin precursor	170718
1475	<i>Triticum aestivum</i>	Wheat	Unassigned	pre-alpha-/beta-gliadin A-II	170712
1476	<i>Triticum aestivum</i>	Wheat	Unassigned	alpha-type gliadin precursor protein	170710
1477	<i>Triticum aestivum</i>	Wheat	Unassigned	unnamed protein product [Triticum aestivum]	21765

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1478	<i>Triticum aestivum</i>	Wheat	Unassigned	unnamed protein product [Triticum aestivum]	21761
1479	<i>Triticum aestivum</i>	Wheat	Unassigned	unnamed protein product [Triticum aestivum]	21757
1480	<i>Triticum aestivum</i>	Wheat	Unassigned	unnamed protein product [Triticum aestivum]	21755
1481	<i>Triticum aestivum</i>	Wheat	Unassigned	unnamed protein product [Triticum aestivum]	21673
1482	<i>Triticum aestivum</i>	Wheat	Unassigned	Allergen C-C	3913017
1483	<i>Triticum aestivum</i>	Wheat	Unassigned	putative gamma-gliadin [Triticum aestivum]	62484809
1484	<i>Triticum aestivum</i>	Wheat	Unassigned	gamma-gliadin precursor [Triticum aestivum]	1063270
1485	<i>Triticum aestivum</i>	Wheat	Unassigned	gamma-gliadin	170738
1486	<i>Triticum aestivum</i>	Wheat	Unassigned	gamma-gliadin	170736
1487	<i>Triticum aestivum</i>	Wheat	Unassigned	gamma gliadin B-III	170734
1488	<i>Triticum aestivum</i>	Wheat	Unassigned	gamma-gliadin	170732
1489	<i>Triticum aestivum</i>	Wheat	Unassigned	pre-gamma-gliadin B-I	170730
1490	<i>Triticum aestivum</i>	Wheat	Unassigned	gamma-gliadin B precursor	170708
1491	<i>Triticum aestivum</i>	Wheat	Unassigned	gamma gliadin precursor	170702
1492	<i>Triticum aestivum</i>	Wheat	Unassigned	unnamed protein product [Triticum aestivum]	897811
1493	<i>Triticum aestivum</i>	Wheat	Unassigned	glutenin [Triticum aestivum]	736319
1494	<i>Triticum aestivum</i>	Wheat	Unassigned	HMW glutenin subunit Ax2* [Triticum aestivum]	170743
1495	<i>Triticum aestivum</i>	Wheat	Unassigned	HMW glutenin subunit 1By9 [Triticum aestivum]	22090
1496	<i>Triticum aestivum</i>	Wheat	Unassigned	unnamed protein product [Triticum aestivum]	21793
1497	<i>Triticum aestivum</i>	Wheat	Unassigned	unnamed protein product [Triticum aestivum]	21779
1498	<i>Triticum aestivum</i>	Wheat	Unassigned	high molecular weight glutenin subunit 10 [Triticum aestivum]	21751
1499	<i>Triticum aestivum</i>	Wheat	Unassigned	high molecular weight glutenin subunit 1Ax1 [Triticum aestivum]	21743
1500	<i>Triticum aestivum</i>	Wheat	Unassigned	low molecular weight glutenin [Triticum aestivum]	335331566
1501	<i>Triticum aestivum</i>	Wheat	Unassigned	LMM glutenin 1	75219081
1502	<i>Triticum aestivum</i>	Wheat	Unassigned	LMM glutenin 3	75317968

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1503	<i>Triticum aestivum</i>	Wheat	Unassigned	putative LMW-glutenin subunit [Triticum aestivum]	62550933
1504	<i>Triticum aestivum</i>	Wheat	Unassigned	low molecular weight glutenin [Triticum aestivum]	886967
1505	<i>Triticum aestivum</i>	Wheat	Unassigned	low molecular weight glutenin [Triticum aestivum]	886965
1506	<i>Triticum aestivum</i>	Wheat	Unassigned	low molecular weight glutenin [Triticum aestivum]	886963
1507	<i>Triticum aestivum</i>	Wheat	Unassigned	unnamed protein product [Triticum aestivum]	21783
1508	<i>Triticum aestivum</i>	Wheat	Unassigned	unnamed protein product [Triticum aestivum]	21773
1509	<i>Triticum aestivum</i>	Wheat	Unassigned	D-type LMW glutenin subunit [Triticum aestivum]	208605348
1510	<i>Triticum aestivum</i>	Wheat	Unassigned	D-type LMW glutenin subunit [Triticum aestivum]	208605346
1511	<i>Triticum aestivum</i>	Wheat	Unassigned	D-type LMW glutenin subunit [Triticum aestivum]	208605344
1512	<i>Triticum aestivum</i>	Wheat	Unassigned	omega-5 gliadin [Triticum aestivum]	73912496
1513	<i>Triticum aestivum</i>	Wheat	Unassigned	profilin [Triticum aestivum]	190684061
1514	<i>Triticum aestivum</i>	Wheat	Unassigned	profilin [Triticum aestivum]	1052817
1515	<i>Triticum aestivum</i>	Wheat	Unassigned	profilin [Triticum aestivum]	1008445
1516	<i>Triticum aestivum</i>	Wheat	Unassigned	profilin [Triticum aestivum]	1008443
1517	<i>Triticum aestivum</i>	Wheat	Unassigned	dehydrin [Triticum aestivum]	190684063
1518	<i>Triticum aestivum</i>	Wheat	Unassigned	peroxiredoxin [Triticum aestivum]	190684059
1519	<i>Triticum aestivum</i>	Wheat	Unassigned	glutathione transferase [Triticum aestivum]	190684057
1520	<i>Triticum aestivum</i>	Wheat	Unassigned	thioredoxin H [Triticum aestivum]	190684055
1521	<i>Triticum aestivum</i>	Wheat	Unassigned	putative leucine-rich repeat protein [Triticum aestivum]	66840996
1522	<i>Triticum aestivum</i>	Wheat	Unassigned	RecName: Full=Serine carboxypeptidase 2; AltName: Full=Serine carboxypeptidase II; AltName: Full=Carboxypeptidase D; AltName: Full=CPDW-II; Short=CP-WII; Contains: RecName: Full=Serine carboxypeptidase 2 chain A; AltName: Full=Serine carboxypeptidase II c	125987805
1523	<i>Triticum aestivum</i>	Wheat	Unassigned	serine carboxypeptidase II [Triticum aestivum]	66840994
1524	<i>Triticum aestivum</i>	Wheat	Unassigned	serpin [Triticum aestivum]	1885350



Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1525	<i>Triticum aestivum</i>	Wheat	Unassigned	RecName: Full=Subtilisin-chymotrypsin inhibitor WSCI; Flags: Precursor	122065237
1526	<i>Triticum aestivum</i>	Wheat	Unassigned	serine proteinase inhibitor-like allergen [Triticum aestivum]	154101366
1527	<i>Triticum aestivum</i>	Wheat	Unassigned	RecName: Full=Thaumatococcus-like protein PWIR2; Flags: Precursor	135917
1528	<i>Triticum aestivum</i>	Wheat	Unassigned	putative alpha-amylase inhibitor CM2 [Triticum aestivum]	283465827
1529	<i>Triticum aestivum</i>	Wheat	Unassigned	alpha amylase inhibitor CM1 [Triticum aestivum]	253783731
1530	<i>Triticum aestivum</i>	Wheat	Unassigned	unnamed protein product [Triticum aestivum]	21701
1531	<i>Triticum aestivum</i>	Wheat	Unassigned	triosephosphat-isomerase [Triticum aestivum]	11124572
1532	<i>Triticum turgidum</i>	Wheat	Unassigned	tetrameric alpha-amylase inhibitor 16 kDa subunit, CM16* [Triticum turgidum L.=pasta wheat, cv. Senatore Capelli, Peptide Partial, 18 aa]	244610
1533	<i>Triticum turgidum subsp. durum</i>	Wheat	Unassigned	precursor (AA -24 to 119) [Triticum turgidum subsp. durum]	21916
1534	<i>Triticum turgidum subsp. durum</i>	Wheat	Unassigned	alpha-amylase inhibitor, tetrameric, chain CM3 precursor - durum wheat	100834
1535	<i>Triticum turgidum subsp. durum</i>	Wheat	Unassigned	LMW glutenin [Triticum turgidum subsp. durum]	21930
1536	<i>Triticum turgidum subsp. durum</i>	Wheat	Unassigned	unnamed protein product [Triticum turgidum subsp. durum]	21926
1537	<i>Triticum turgidum subsp. durum</i>	Wheat	Unassigned	CM2 protein [Triticum turgidum subsp. durum]	21920
1538	<i>Triticum urartu</i>	Wheat	Unassigned	gliadin	170740
1539	<i>Tyrophagus putrescentiae</i>	Dust mite	Unassigned	tropomyosin [Tyrophagus putrescentiae]	48249227
1540	<i>Tyrophagus putrescentiae</i>	Dust mite	Unassigned	tropomyosin [Tyrophagus putrescentiae]	156938915
1541	<i>Tyrophagus putrescentiae</i>	Dust mite	Unassigned	tropomyosin [Tyrophagus putrescentiae]	148615631
1542	<i>Tyrophagus putrescentiae</i>	Dust mite	Unassigned	allergen Tyr p 13 [Tyrophagus putrescentiae]	156938917
1543	<i>Tyrophagus putrescentiae</i>	Dust mite	Unassigned	mite allergen Tyr p 13 [Tyrophagus putrescentiae]	121296500
1544	<i>Tyrophagus putrescentiae</i>	Dust mite	Tyr p 13	fatty acid-binding protein [Tyrophagus putrescentiae]	51860756

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1545	<i>Tyrophagus putrescentiae</i>	Dust mite	Unassigned	Mite group 2 allergen Tyr p 2 precursor	3182907
1546	<i>Tyrophagus putrescentiae</i>	Dust mite	Unassigned	troponin C [Tyrophagus putrescentiae]	219815476
1547	<i>Tyrophagus putrescentiae</i>	Dust mite	Unassigned	Tyr p 3 [Tyrophagus putrescentiae]	167540622
1548	<i>Vespa crabro</i>	European hornet	Vesp c 5	Venom allergen 5.01 (Antigen 5-1) (Ag5-1) (Allergen Vesp c 5.01) (Vesp c V.01)	549184
1549	<i>Vespa crabro</i>	European hornet	Vesp c 5	Venom allergen 5.02 (Antigen 5-2) (Ag5-2) (Allergen Vesp c 5.02) (Vesp c V.02)	549185
1550	<i>Vespa crabro</i>	European hornet	Unassigned	RecName: Full=Venom phospholipase A1; AltName: Allergen=Vesp c 1	313471397
1551	<i>Vespa mandarinia</i>	Wasp	Vesp m 5	Venom allergen 5 (Antigen 5) (Ag5) (Allergen Vesp m 5)	6136165
1552	<i>Vespula flavopilosa</i>	Wasp	Ves f 5	Venom allergen 5 (Antigen 5) (Ag5) (Allergen Ves f 5) (Ves f V)	549189
1553	<i>Vespula germanica</i>	Wasp	Unassigned	Ves g 1 allergen precursor [Vespula germanica]	74035843
1554	<i>Vespula germanica</i>	Wasp	Unassigned	Ves g 5 allergen precursor [Vespula germanica]	74035841
1555	<i>Vespula germanica</i>	Wasp	Ves g 5	Venom allergen 5 (Antigen 5) (Ag5) (Allergen Ves g 5) (Ves g V)	549190
1556	<i>Vespula germanica</i>	Wasp	Unassigned	hyaluronidase homologue [Vespula germanica]	116174182
1557	<i>Vespula germanica</i>	Wasp	Unassigned	hyaluronidase [Vespula germanica]	116174180
1558	<i>Vespula maculifrons</i>	Wasp	Ves m 1	Phospholipase A1 (Allergen Ves m 1) (Ves m I)	1709545
1559	<i>Vespula maculifrons</i>	Wasp	Unassigned	venom allergen 5 [Vespula maculifrons]	85681830
1560	<i>Vespula maculifrons</i>	Wasp	Ves m 5	Venom allergen 5 (Antigen 5) (Ag5) (Allergen Ves m 5) (Ves m V)	549191
1561	<i>Vespula maculifrons</i>	Wasp	Unassigned	RecName: Full=Hyaluronidase; Short=Hya; AltName: Full=Hyaluronoglucosaminidase; AltName: Allergen=Ves m 2	313118253
1562	<i>Vespula pensylvanica</i>	Wasp	Ves p 5	Venom allergen 5 (Antigen 5) (Ag5) (Allergen Ves p 5) (Ves p V)	549192
1563	<i>Vespula squamosa</i>	Wasp	Ves s 5	Venom allergen 5 (Antigen 5) (Ag5) (Allergen Ves s 5) (Ves s V)	549193

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1564	<i>Vespula squamosa</i>	Wasp	Unassigned	RecName: Full=Venom phospholipase A1; AltName: Allergen=Ves s 1	313471398
1565	<i>Vespula vidua</i>	Wasp	Ves vi 5	Venom allergen 5 (Antigen 5) (Ag5) (Allergen Ves vi 5) (Ves vi V)	549194
1566	<i>Vespula vulgaris</i>	Wasp	Ves v 1	allergen and phospholipase A1	897647
1567	<i>Vespula vulgaris</i>	Wasp	Ves v 5	Chain A, Ves V 5, An Allergen From Vespula Vulgaris Venom	11514279
1568	<i>Vespula vulgaris</i>	Wasp	Ves v 5	allergen 5; antigen 5 [Vespula vulgaris]	4826574
1569	<i>Vespula vulgaris</i>	Wasp	Ves v 5	allergen 5	162551
1570	<i>Vespula vulgaris</i>	Wasp	Unassigned	Chain A, Crystal Structure Of The Recombinant Allergen Ves V 2	109157163
1571	<i>Vespula vulgaris</i>	Wasp	Unassigned	hyaluronidase b [Vespula vulgaris]	62147665
1572	<i>Vespula vulgaris</i>	Wasp	Ves v 2	Hyaluronoglucosaminidase (Hyaluronidase) (Allergen Ves v 2) (Ves v II)	1346323
1573	<i>Vespula vulgaris</i>	Wasp	Unassigned	RecName: Full=Venom dipeptidyl peptidase 4; AltName: Full=Venom dipeptidyl peptidase IV; AltName: Allergen=Ves v 3; Flags: Precursor	313471718
1574	<i>Vigna radiata</i>		Unassigned	pathogenesis-related protein 10 [Vigna radiata]	60418924
1575	<i>Vitis sp.</i>	Grape	Unassigned	Nonspecific lipid-transfer protein P3 (LTP P3)	145559502
1576	<i>Vitis sp.</i>	Grape	Unassigned	Nonspecific lipid-transfer protein P2 (LTP P2)	462717
1577	<i>Vitis sp.</i>	Grape	Vit v 1	Nonspecific lipid-transfer protein P4 (LTP P4)	462719
1578	<i>Xiphias gladius</i>		Unassigned	beta-parvalbumin [Xiphias gladius]	222352960
1579	<i>Zea mays</i>	Corn	Unassigned	Expansin-B11 precursor (ZmEXPB11) (Beta-expansin-11) (Pollen allergen Zea m 1a) (Pollen allergen Zea m 1b)	115502168
1580	<i>Zea mays</i>	Corn	Unassigned	Expansin-B10 precursor (ZmEXPB10) (Beta-expansin-10) (Pollen allergen Zea m 1c)	115502167
1581	<i>Zea mays</i>	Corn	Unassigned	EXPB10 [Zea mays]	105969545
1582	<i>Zea mays</i>	Corn	Unassigned	EXPB10 [Zea mays]	105969543
1583	<i>Zea mays</i>	Corn	Unassigned	Zea m 1 allergen [Zea mays]	89892723

Entry	Organism	Common name	Allergen name	Description	GenInfo identifier
1584	<i>Zea mays</i>	Corn	Unassigned	Zea m 1 allergen [Zea mays]	89892721
1585	<i>Zea mays</i>	Corn	Unassigned	pollen profilin variant 7 [Zea mays]	110644964
1586	<i>Zea mays</i>	Corn	Unassigned	pollen profilin variant 6 [Zea mays]	110644962
1587	<i>Zea mays</i>	Corn	Unassigned	pollen profilin variant 5 [Zea mays]	110644960
1588	<i>Zea mays</i>	Corn	Unassigned	pollen profilin variant 4 [Zea mays]	110644958
1589	<i>Zea mays</i>	Corn	Unassigned	pollen profilin variant 3 [Zea mays]	110644956
1590	<i>Zea mays</i>	Corn	Unassigned	pollen profilin variant 2 [Zea mays]	110644954
1591	<i>Zea mays</i>	Corn	Unassigned	pollen profilin variant 1 [Zea mays]	110644952
1592	<i>Zea mays</i>	Corn	Unassigned	profilin [Zea mays]	2642324
1593	<i>Zea mays</i>	Corn	Unassigned	Zea m 13 allergen [Zea mays]	89892729
1594	<i>Zea mays</i>	Corn	Unassigned	Zea m 13 allergen [Zea mays]	89892727
1595	<i>Zea mays</i>	Corn	Unassigned	Zea m 13 allergen [Zea mays]	89892725
1596	<i>Zea mays</i>	Corn	Unassigned	Zm13	1588669
1597	<i>Zea mays</i>	Corn	Zea m 14	Nonspecific lipid-transfer protein precursor (LTP) (Phospholipid transfer protein) (PLTP) (Allergen Zea m 14)	128388
1598	<i>Zea mays</i>	Corn	Unassigned	thioredoxin h1 protein [Zea mays]	66841002
1599	<i>Zea mays</i>	Corn	Unassigned	Chain X, Crystal Structure Of Expb1 (Zea M 1), A Beta-Expansin And Group-1 Pollen Allergen From Maize	114794319
1600	<i>Zea mays</i>	Corn	Unassigned	beta-expansin 1 [Zea mays]	14193761
1601	<i>Zea mays</i>	Corn	Unassigned	beta-expansin 1 protein [Zea mays]	28630923
1602	<i>Zea mays</i>	Corn	Unassigned	beta-expansin 9 protein [Zea mays]	28630919
1603	<i>Ziziphus mauritiana</i>	Chinese-date	Unassigned	allergen Ziz m 1 [Ziziphus mauritiana]	61225281

## APPENDIX B FASTA Allergen Search Alignments

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_7	>>ALLERGEN_2012_2 gi 664852 gb AAB07620.1  Asp FII [Aspe (250 aa)	0.019	42.9	21
>SYHT0H2_I_7	>>ALLERGEN_2012_2 gi 66849502 gb EAL89830.1  major aller (304 aa)	0.023	42.9	21
>SYHT0H2_I_7	>>ALLERGEN_2012_2 gi 83300352 sp P79017.2  ALL2_ASPFU Rec (310 aa)	0.023	42.9	21
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 62550933 emb CAI79052.1  putative L (326 aa)	0.21	30.1	73
>SYHT0H2_I_10	>>ALLERGEN_2012_2 gi 5059162 gb AAD38942.1 AF144060_1 al (496 aa)	0.25	26.4	110
>SYHT0H2_I_16	>>ALLERGEN_2012_2 gi 5059162 gb AAD38942.1 AF144060_1 al (496 aa)	0.25	26.4	110
>SYHT0H2_I_28	>>ALLERGEN_2012_2 gi 5059162 gb AAD38942.1 AF144060_1 al (496 aa)	0.25	26.4	110
>SYHT0H2_I_38	>>ALLERGEN_2012_2 gi 5059162 gb AAD38942.1 AF144060_1 al (496 aa)	0.25	26.4	110
>SYHT0H2_I_50	>>ALLERGEN_2012_2 gi 22090 emb CAA43361.1  HMW glutenin (705 aa)	0.27	47.1	17
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 170710 gb AAA34275.1  alpha-type gl (318 aa)	0.29	45.8	24
>SYHT0H2_I_50	>>ALLERGEN_2012_2 gi 21779 emb CAA26847.1  unnamed prote (660 aa)	0.33	47.1	17
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 83300389 sp O42799.2  ALL7_ASPFU Rec (270 aa)	0.36	41.7	24
>SYHT0H2_I_50	>>ALLERGEN_2012_2 gi 21751 emb CAA31396.1  high molecula (648 aa)	0.41	47.1	17
>SYHT0H2_I_47	>>ALLERGEN_2012_2 gi 32363456 sp P81729.1 C HAL_BRARA Rec (91 aa)	0.44	40.7	27
>SYHT0H2_I_47	>>ALLERGEN_2012_2 gi 14423650 sp Q9U1G2.1  ALL7_LEPDS Rec (216 aa)	0.63	31.3	32

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 170722 gb AAA34281.1 pre-alpha-be (262 aa)	0.66	28.2	39
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 473876 gb AAA17741.1 alpha-gliadin (287 aa)	0.73	28.2	39
>SYHT0H2_I_47	>>ALLERGEN_2012_2 gi 4538529 emb CAB39376.1 Cop c1 alle (81 aa)	0.73	38.2	34
>SYHT0H2_I_15	>>ALLERGEN_2012_2 gi 22684 emb CAA50325.1 major allerge (160 aa)	0.76	40.6	32
>SYHT0H2_I_15	>>ALLERGEN_2012_2 gi 584968 sp Q08407.3 MPAC1_CORAV RecN (160 aa)	0.76	40.6	32
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 170728 gb AAA34284.1 alpha-type gl (186 aa)	0.79	30.6	49
>SYHT0H2_I_19	>>ALLERGEN_2012_2 gi 6136162 sp P35776.2 VA2_SOLRI RecNa (119 aa)	0.91	36.0	25
>SYHT0H2_I_52	>>ALLERGEN_2012_2 gi 6136162 sp P35776.2 VA2_SOLRI RecNa (119 aa)	0.91	36.0	25
>SYHT0H2_I_27	>>ALLERGEN_2012_2 gi 21783 emb CAA30570.1 unnamed prote (356 aa)	0.92	30.8	26
>SYHT0H2_I_43	>>ALLERGEN_2012_2 gi 21783 emb CAA30570.1 unnamed prote (356 aa)	0.92	30.8	26
>SYHT0H2_I_10	>>ALLERGEN_2012_2 gi 21748153 emb CAD38167.1 putative n (124 aa)	0.99	33.3	42
>SYHT0H2_I_16	>>ALLERGEN_2012_2 gi 21748153 emb CAD38167.1 putative n (124 aa)	0.99	33.3	42
>SYHT0H2_I_28	>>ALLERGEN_2012_2 gi 21748153 emb CAD38167.1 putative n (124 aa)	0.99	33.3	42
>SYHT0H2_I_38	>>ALLERGEN_2012_2 gi 21748153 emb CAD38167.1 putative n (124 aa)	0.99	33.3	42
>SYHT0H2_I_40	>>ALLERGEN_2012_2 gi 38326693 gb AAR17475.1 unknown [Pe (228 aa)	0.99	37.5	32

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 75219081 sp O22108 O22108_WHEAT LMM (285 aa)	1	36.4	44
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 75219081 sp O22108 O22108_WHEAT LMM (285 aa)	1	36.4	44
>SYHT0H2_I_5	>>ALLERGEN_2012_2 gi 1364213 emb CAA44344.1 fel d I cha (92 aa)	1	45.0	20
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 46410859 gb AAR98518.1 major latex (366 aa)	1.1	22.9	236
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 886965 emb CAA59339.1 low molecula (261 aa)	1.1	29.4	68
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 21773 emb CAA31685.1 unnamed prote (307 aa)	1.1	36.4	44
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 21773 emb CAA31685.1 unnamed prote (307 aa)	1.1	36.4	44
>SYHT0H2_I_39	>>ALLERGEN_2012_2 gi 1364213 emb CAA44344.1 fel d I cha (92 aa)	1.2	40.0	20
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 18536 emb CAA35691.1 unnamed prote (605 aa)	1.2	58.3	12
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 170732 gb AAA34286.1 gamma-gliadin (323 aa)	1.3	25.9	58
>SYHT0H2_I_15	>>ALLERGEN_2012_2 gi 1346568 sp P49372.1 ALL1_APIGR RecN (154 aa)	1.4	35.2	54
>SYHT0H2_I_47	>>ALLERGEN_2012_2 gi 14422361 emb CAC41634.1 plantain p (131 aa)	1.4	31.3	32
>SYHT0H2_I_47	>>ALLERGEN_2012_2 gi 14422359 emb CAC41633.1 plantain p (131 aa)	1.4	31.3	32
>SYHT0H2_I_47	>>ALLERGEN_2012_2 gi 14422363 emb CAC41635.1 plantain p (131 aa)	1.4	31.3	32
>SYHT0H2_I_21	>>ALLERGEN_2012_2 gi 208605344 emb CAR82265.1 D-type LM (359 aa)	1.4	37.0	27

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 21926 emb CAA36063.1  unnamed prote (295 aa)	1.4	41.7	24
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 21755 emb CAA25593.1  unnamed prote (286 aa)	1.5	30.6	49
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 21761 emb CAA26384.1  unnamed prote (286 aa)	1.5	30.6	49
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 170720 gb AAA34280.1  alpha/beta-gl (286 aa)	1.5	30.6	49
>SYHT0H2_I_27	>>ALLERGEN_2012_2 gi 75317968 sp O22116 O22116_WHEAT LMM (373 aa)	1.5	30.8	26
>SYHT0H2_I_43	>>ALLERGEN_2012_2 gi 75317968 sp O22116 O22116_WHEAT LMM (373 aa)	1.5	30.8	26
>SYHT0H2_I_21	>>ALLERGEN_2012_2 gi 303387468 gb ADM15668.1  lipid bind (228 aa)	1.6	21.6	74
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 1304264 dbj BAA12318.1  alpha-gliad (259 aa)	1.6	34.9	43
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 170738 gb AAA34289.1  gamma-gliadin (327 aa)	1.8	37.5	24
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 75219081 sp O22108 O22108_WHEAT LMM (285 aa)	1.9	39.3	28
>SYHT0H2_I_3	>>ALLERGEN_2012_2 gi 291482314 emb CBK62697.1  ragweed h (96 aa)	1.9	69.2	13
>SYHT0H2_I_33	>>ALLERGEN_2012_2 gi 291482314 emb CBK62697.1  ragweed h (96 aa)	1.9	69.2	13
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 886965 emb CAA59339.1  low molecula (261 aa)	2	25.4	71
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 886965 emb CAA59339.1  low molecula (261 aa)	2	25.4	71
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 21773 emb CAA31685.1  unnamed prote (307 aa)	2	39.3	28



Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_3	>>ALLERGEN_2012_2 gi 291482306 emb CBK62693.1  ragweed h (111 aa)	2	69.2	13
>SYHT0H2_I_33	>>ALLERGEN_2012_2 gi 291482306 emb CBK62693.1  ragweed h (111 aa)	2	69.2	13
>SYHT0H2_I_26	>>ALLERGEN_2012_2 gi 71360928 emb CAJ19705.1  non-specif (114 aa)	2.1	26.5	34
>SYHT0H2_I_44	>>ALLERGEN_2012_2 gi 71360928 emb CAJ19705.1  non-specif (114 aa)	2.1	26.5	34
>SYHT0H2_I_17	>>ALLERGEN_2012_2 gi 3021324 emb CAA06305.1  Aspfl aller (125 aa)	2.1	33.3	33
>SYHT0H2_I_37	>>ALLERGEN_2012_2 gi 3021324 emb CAA06305.1  Aspfl aller (125 aa)	2.1	33.3	33
>SYHT0H2_I_10	>>ALLERGEN_2012_2 gi 89892723 gb ABD79095.1  Zea m 1 all (252 aa)	2.2	30.2	53
>SYHT0H2_I_16	>>ALLERGEN_2012_2 gi 89892723 gb ABD79095.1  Zea m 1 all (252 aa)	2.2	30.2	53
>SYHT0H2_I_28	>>ALLERGEN_2012_2 gi 89892723 gb ABD79095.1  Zea m 1 all (252 aa)	2.2	30.2	53
>SYHT0H2_I_38	>>ALLERGEN_2012_2 gi 89892723 gb ABD79095.1  Zea m 1 all (252 aa)	2.2	30.2	53
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 170740 gb AAA34290.1  gliadin [Trit (296 aa)	2.2	32.6	43
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 21757 emb CAA26383.1  unnamed prote (296 aa)	2.2	32.6	43
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 170712 gb AAA34276.1  pre-alpha-/be (291 aa)	2.2	31.0	42
>SYHT0H2_I_25	>>ALLERGEN_2012_2 gi 23616947 dbj BAC20650.1  putative a (160 aa)	2.2	30.8	39
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 155676690 dbj BAF75708.1  SXP/RAL-2 (150 aa)	2.2	30.8	39

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_45	>>ALLERGEN_2012_2 gi 23616947 dbj BAC20650.1 putative a (160 aa)	2.2	30.8	39
>SYHT0H2_I_31	>>ALLERGEN_2012_2 gi 417370 sp P24296.2 NLTP1_WHEAT RecN (113 aa)	2.2	45.0	20
>SYHT0H2_I_3	>>ALLERGEN_2012_2 gi 291482310 emb CBK62695.1 ragweed h (134 aa)	2.2	69.2	13
>SYHT0H2_I_3	>>ALLERGEN_2012_2 gi 291482308 emb CBK62694.1 ragweed h (140 aa)	2.2	69.2	13
>SYHT0H2_I_33	>>ALLERGEN_2012_2 gi 291482310 emb CBK62695.1 ragweed h (134 aa)	2.2	69.2	13
>SYHT0H2_I_33	>>ALLERGEN_2012_2 gi 291482308 emb CBK62694.1 ragweed h (140 aa)	2.2	69.2	13
>SYHT0H2_I_10	>>ALLERGEN_2012_2 gi 89892721 gb ABD79094.1 Zea m 1 all (263 aa)	2.3	30.2	53
>SYHT0H2_I_10	>>ALLERGEN_2012_2 gi 115502167 sp Q1ZYQ8.2 EXB10_MAIZE R (270 aa)	2.3	30.2	53
>SYHT0H2_I_16	>>ALLERGEN_2012_2 gi 89892721 gb ABD79094.1 Zea m 1 all (263 aa)	2.3	30.2	53
>SYHT0H2_I_16	>>ALLERGEN_2012_2 gi 115502167 sp Q1ZYQ8.2 EXB10_MAIZE R (270 aa)	2.3	30.2	53
>SYHT0H2_I_28	>>ALLERGEN_2012_2 gi 89892721 gb ABD79094.1 Zea m 1 all (263 aa)	2.3	30.2	53
>SYHT0H2_I_28	>>ALLERGEN_2012_2 gi 115502167 sp Q1ZYQ8.2 EXB10_MAIZE R (270 aa)	2.3	30.2	53
>SYHT0H2_I_38	>>ALLERGEN_2012_2 gi 89892721 gb ABD79094.1 Zea m 1 all (263 aa)	2.3	30.2	53
>SYHT0H2_I_38	>>ALLERGEN_2012_2 gi 115502167 sp Q1ZYQ8.2 EXB10_MAIZE R (270 aa)	2.3	30.2	53
>SYHT0H2_I_4	>>ALLERGEN_2012_2 gi 7638030 gb AAF65313.1 AF230384_1 ve (137 aa)	2.3	30.4	46

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_4	>>ALLERGEN_2012_2 gi 7638028 gb AAF65312.1 AF230383_1 ve (137 aa)	2.3	30.4	46
>SYHT0H2_I_32	>>ALLERGEN_2012_2 gi 7638030 gb AAF65313.1 AF230384_1 ve (137 aa)	2.3	30.4	46
>SYHT0H2_I_32	>>ALLERGEN_2012_2 gi 7638028 gb AAF65312.1 AF230383_1 ve (137 aa)	2.3	30.4	46
>SYHT0H2_I_17	>>ALLERGEN_2012_2 gi 9280360 gb AAF86369.1 major allerg (150 aa)	2.4	33.3	33
>SYHT0H2_I_37	>>ALLERGEN_2012_2 gi 9280360 gb AAF86369.1 major allerg (150 aa)	2.4	33.3	33
>SYHT0H2_I_3	>>ALLERGEN_2012_2 gi 285005079 emb CBJ24286.1 ragweed h (164 aa)	2.4	69.2	13
>SYHT0H2_I_3	>>ALLERGEN_2012_2 gi 291197394 emb CBK52317.1 ragweed h (164 aa)	2.4	69.2	13
>SYHT0H2_I_33	>>ALLERGEN_2012_2 gi 285005079 emb CBJ24286.1 ragweed h (164 aa)	2.4	69.2	13
>SYHT0H2_I_33	>>ALLERGEN_2012_2 gi 291197394 emb CBK52317.1 ragweed h (164 aa)	2.4	69.2	13
>SYHT0H2_I_40	>>ALLERGEN_2012_2 gi 160285626 pdb 2JMH A Chain A, Nmr S (119 aa)	2.5	37.5	32
>SYHT0H2_I_29	>>ALLERGEN_2012_2 gi 25361513 gb AAN73248.1 helix-loop- (450 aa)	2.5	63.6	11
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 422005 pir  S32101 PHL5A protein - (257 aa)	2.5	66.7	9
>SYHT0H2_I_17	>>ALLERGEN_2012_2 gi 10189811 emb CAC09234.1 unnamed pr (215 aa)	2.6	28.6	28
>SYHT0H2_I_17	>>ALLERGEN_2012_2 gi 1352240 sp P49273.1 ALL7_DERPT RecN (215 aa)	2.6	28.6	28
>SYHT0H2_I_37	>>ALLERGEN_2012_2 gi 10189811 emb CAC09234.1 unnamed pr (215 aa)	2.6	28.6	28

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_37	>>ALLERGEN_2012_2 gi 1352240 sp P49273.1 ALL7_DERPT RecN (215 aa)	2.6	28.6	28
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 2266625 emb CAB10765.1  group V all (264 aa)	2.7	34.3	70
>SYHT0H2_I_15	>>ALLERGEN_2012_2 gi 291482314 emb CBK62697.1  ragweed h (96 aa)	2.7	28.9	38
>SYHT0H2_I_14	>>ALLERGEN_2012_2 gi 5381323 gb AAD42943.1 AF091841_1 2S (148 aa)	2.7	37.9	29
>SYHT0H2_I_21	>>ALLERGEN_2012_2 gi 25361513 gb AAN73248.1  helix-loop- (450 aa)	2.7	45.5	22
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 1092249 prf 2023228A major allerge (285 aa)	2.7	66.7	9
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 21673 emb CAA35238.1  unnamed prote (307 aa)	2.8	30.0	40
>SYHT0H2_I_50	>>ALLERGEN_2012_2 gi 14423933 sp O82803.1 SRPP_HEVBR Rec (204 aa)	2.8	29.4	34
>SYHT0H2_I_17	>>ALLERGEN_2012_2 gi 54039254 sp P67875.1 R NMG_ASPFU Rec (176 aa)	2.8	33.3	33
>SYHT0H2_I_37	>>ALLERGEN_2012_2 gi 54039254 sp P67875.1 R NMG_ASPFU Rec (176 aa)	2.8	33.3	33
>SYHT0H2_I_40	>>ALLERGEN_2012_2 gi 111120450 gb ABH06359.1  Blo t 5 al (134 aa)	2.8	37.5	32
>SYHT0H2_I_40	>>ALLERGEN_2012_2 gi 4204917 gb AAD10850.1  major IgE-bi (134 aa)	2.8	37.5	32
>SYHT0H2_I_40	>>ALLERGEN_2012_2 gi 111120436 gb ABH06352.1  Blo t 5 al (134 aa)	2.8	37.5	32
>SYHT0H2_I_25	>>ALLERGEN_2012_2 gi 114152864 sp Q01883.2 RAG1_ORYSJ Re (163 aa)	2.8	43.3	30
>SYHT0H2_I_45	>>ALLERGEN_2012_2 gi 114152864 sp Q01883.2 RAG1_ORYSJ Re (163 aa)	2.8	43.3	30

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 94400907 ref NP_001035360.1  allerg (92 aa)	2.8	40.0	20
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 88770352 gb ABD51779.1  allergen Ap (94 aa)	2.8	40.0	20
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 94400907 ref NP_001035360.1  allerg (92 aa)	2.8	40.0	20
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 88770352 gb ABD51779.1  allergen Ap (94 aa)	2.8	40.0	20
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 633938 gb AAB30434.1  albumin [Cani (265 aa)	2.8	47.1	17
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 633938 gb AAB30434.1  albumin [Cani (265 aa)	2.8	47.1	17
>SYHT0H2_I_21	>>ALLERGEN_2012_2 gi 4138175 emb CAA09885.1  allergen [M (187 aa)	2.8	53.3	15
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 886963 emb CAA59338.1  low molecula (229 aa)	2.9	26.0	73
>SYHT0H2_I_18	>>ALLERGEN_2012_2 gi 71360928 emb CAJ19705.1  non-specif (114 aa)	2.9	26.5	34
>SYHT0H2_I_36	>>ALLERGEN_2012_2 gi 71360928 emb CAJ19705.1  non-specif (114 aa)	2.9	26.5	34
>SYHT0H2_I_23	>>ALLERGEN_2012_2 gi 22135348 gb AAM93157.1  trypsin inh (219 aa)	2.9	32.1	28
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 62484809 emb CAI78902.1  putative g (285 aa)	2.9	60.0	15
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 160285626 pdb 2JMH A Chain A, Nmr S (119 aa)	3	31.5	54
>SYHT0H2_I_15	>>ALLERGEN_2012_2 gi 291482316 emb CBK62698.1  ragweed h (110 aa)	3	28.9	38
>SYHT0H2_I_4	>>ALLERGEN_2012_2 gi 4038411 gb AAC97370.1  venom allerg (137 aa)	3.1	34.8	46

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_32	>>ALLERGEN_2012_2 gi 4038411 gb AAC97370.1  venom allerg (137 aa)	3.1	34.8	46
>SYHT0H2_I_15	>>ALLERGEN_2012_2 gi 291482318 emb CBK62699.1  ragweed h (116 aa)	3.1	28.9	38
>SYHT0H2_I_8	>>ALLERGEN_2012_2 gi 154101366 gb ABS58503.1  serine pro (84 aa)	3.1	34.8	23
>SYHT0H2_I_19	>>ALLERGEN_2012_2 gi 2497701 sp Q28133.1 ALL2_BOVIN RecN (172 aa)	3.2	27.8	54
>SYHT0H2_I_52	>>ALLERGEN_2012_2 gi 2497701 sp Q28133.1 ALL2_BOVIN RecN (172 aa)	3.2	27.8	54
>SYHT0H2_I_10	>>ALLERGEN_2012_2 gi 105969545 gb ABF81662.1  EXPB10 [Ze (269 aa)	3.2	30.2	53
>SYHT0H2_I_10	>>ALLERGEN_2012_2 gi 115502168 sp P0C1Y5.1 EXB11_MAIZE R (269 aa)	3.2	30.2	53
>SYHT0H2_I_16	>>ALLERGEN_2012_2 gi 105969545 gb ABF81662.1  EXPB10 [Ze (269 aa)	3.2	30.2	53
>SYHT0H2_I_16	>>ALLERGEN_2012_2 gi 115502168 sp P0C1Y5.1 EXB11_MAIZE R (269 aa)	3.2	30.2	53
>SYHT0H2_I_28	>>ALLERGEN_2012_2 gi 105969545 gb ABF81662.1  EXPB10 [Ze (269 aa)	3.2	30.2	53
>SYHT0H2_I_28	>>ALLERGEN_2012_2 gi 115502168 sp P0C1Y5.1 EXB11_MAIZE R (269 aa)	3.2	30.2	53
>SYHT0H2_I_38	>>ALLERGEN_2012_2 gi 105969545 gb ABF81662.1  EXPB10 [Ze (269 aa)	3.2	30.2	53
>SYHT0H2_I_38	>>ALLERGEN_2012_2 gi 115502168 sp P0C1Y5.1 EXB11_MAIZE R (269 aa)	3.2	30.2	53
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 155676692 dbj BAF75709.1  SXP/RAL-2 (150 aa)	3.2	28.2	39
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 4538529 emb CAB39376.1  Cop c1 alle (81 aa)	3.2	38.9	36

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_14	>>ALLERGEN_2012_2 gi 112754 sp P04403.2 SS1_BEREX RecNa (146 aa)	3.2	37.5	24
>SYHT0H2_I_14	>>ALLERGEN_2012_2 gi 1405736 emb CAA35188.1  trypsin inh (144 aa)	3.2	39.1	23
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 111120450 gb ABH06359.1  Blo t 5 al (134 aa)	3.3	31.5	54
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 4204917 gb AAD10850.1  major IgE-bi (134 aa)	3.3	31.5	54
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 111120436 gb ABH06352.1  Blo t 5 al (134 aa)	3.3	31.5	54
>SYHT0H2_I_14	>>ALLERGEN_2012_2 gi 19009 emb CAA46705.1  CMe [Hordeum (148 aa)	3.3	39.1	23
>SYHT0H2_I_12	>>ALLERGEN_2012_2 gi 22135348 gb AAM93157.1  trypsin inh (219 aa)	3.3	40.0	15
>SYHT0H2_I_41	>>ALLERGEN_2012_2 gi 22135348 gb AAM93157.1  trypsin inh (219 aa)	3.3	40.0	15
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 886967 emb CAA59340.1  low molecula (276 aa)	3.4	25.0	84
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 886967 emb CAA59340.1  low molecula (276 aa)	3.4	25.0	84
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 3021324 emb CAA06305.1  Aspf1 aller (125 aa)	3.4	23.1	65
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 3021324 emb CAA06305.1  Aspf1 aller (125 aa)	3.4	23.1	65
>SYHT0H2_I_15	>>ALLERGEN_2012_2 gi 291482310 emb CBK62695.1  ragweed h (134 aa)	3.4	28.9	38
>SYHT0H2_I_25	>>ALLERGEN_2012_2 gi 1304217 dbj BAA07773.1  allergenic (109 aa)	3.4	43.3	30
>SYHT0H2_I_45	>>ALLERGEN_2012_2 gi 1304217 dbj BAA07773.1  allergenic (109 aa)	3.4	43.3	30

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_10	>>ALLERGEN_2012_2 gi 18615 emb CAA26723.1  unnamed prote (495 aa)	3.5	25.8	89
>SYHT0H2_I_16	>>ALLERGEN_2012_2 gi 18615 emb CAA26723.1  unnamed prote (495 aa)	3.5	25.8	89
>SYHT0H2_I_28	>>ALLERGEN_2012_2 gi 18615 emb CAA26723.1  unnamed prote (495 aa)	3.5	25.8	89
>SYHT0H2_I_38	>>ALLERGEN_2012_2 gi 18615 emb CAA26723.1  unnamed prote (495 aa)	3.5	25.8	89
>SYHT0H2_I_19	>>ALLERGEN_2012_2 gi 886965 emb CAA59339.1  low molecula (261 aa)	3.5	27.5	40
>SYHT0H2_I_52	>>ALLERGEN_2012_2 gi 886965 emb CAA59339.1  low molecula (261 aa)	3.5	27.5	40
>SYHT0H2_I_15	>>ALLERGEN_2012_2 gi 291482306 emb CBK62693.1  ragweed h (111 aa)	3.5	28.9	38
>SYHT0H2_I_25	>>ALLERGEN_2012_2 gi 1304216 dbj BAA07772.1  allergenic (111 aa)	3.5	43.3	30
>SYHT0H2_I_25	>>ALLERGEN_2012_2 gi 1304218 dbj BAA07774.1  allergenic (113 aa)	3.5	43.3	30
>SYHT0H2_I_45	>>ALLERGEN_2012_2 gi 1304216 dbj BAA07772.1  allergenic (111 aa)	3.5	43.3	30
>SYHT0H2_I_45	>>ALLERGEN_2012_2 gi 1304218 dbj BAA07774.1  allergenic (113 aa)	3.5	43.3	30
>SYHTOH2_I_27	>>ALLERGEN_2012_2 gi 21314465 gb AAM46958.1 AF510854_1 a (538 aa)	3.5	44.4	18
>SYHT0H2_I_43	>>ALLERGEN_2012_2 gi 21314465 gb AAM46958.1 AF510854_1 a (538 aa)	3.5	44.4	18
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 335331566 gb AEH31546.1  low molecu (369 aa)	3.6	25.0	64
>SYHT0H2_I_31	>>ALLERGEN_2012_2 gi 148887203 emb CAK50834.1  art v 2 a (162 aa)	3.6	26.0	50



Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_23	>>ALLERGEN_2012_2 gi 162927 gb AAA30478.1 alpha-s1-case (76 aa)	3.6	45.0	20
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 170734 gb AAA34287.1 gamma gliadin (244 aa)	3.7	25.8	66
>SYHT0H2_I_14	>>ALLERGEN_2012_2 gi 549186 sp P10737.3 VA53_DOLMA RecNa (215 aa)	3.7	28.3	46
>SYHT0H2_I_27	>>ALLERGEN_2012_2 gi 112754 sp P04403.2 2SS1_BEREX RecNa (146 aa)	3.7	19.4	31
>SYHT0H2_I_43	>>ALLERGEN_2012_2 gi 112754 sp P04403.2 2SS1_BEREX RecNa (146 aa)	3.7	19.4	31
>SYHT0H2_I_19	>>ALLERGEN_2012_2 gi 94400907 ref NP_001035360.1 allerg (92 aa)	3.7	38.9	18
>SYHT0H2_I_52	>>ALLERGEN_2012_2 gi 94400907 ref NP_001035360.1 allerg (92 aa)	3.7	38.9	18
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 21926 emb CAA36063.1 unnamed prote (295 aa)	3.8	32.6	46
>SYHT0H2_I_13	>>ALLERGEN_2012_2 gi 170722 gb AAA34281.1 pre-alpha-/be (262 aa)	3.8	30.0	30
>SYHT0H2_I_24	>>ALLERGEN_2012_2 gi 170722 gb AAA34281.1 pre-alpha-/be (262 aa)	3.8	30.0	30
>SYHT0H2_I_46	>>ALLERGEN_2012_2 gi 170722 gb AAA34281.1 pre-alpha-/be (262 aa)	3.8	30.0	30
>SYHT0H2_I_19	>>ALLERGEN_2012_2 gi 88770352 gb ABD51779.1 allergen Ap (94 aa)	3.8	38.9	18
>SYHT0H2_I_52	>>ALLERGEN_2012_2 gi 88770352 gb ABD51779.1 allergen Ap (94 aa)	3.8	38.9	18
>SYHT0H2_I_47	>>ALLERGEN_2012_2 gi 18772 emb CAA45778.1 trypsin inhib (217 aa)	3.9	38.9	36
>SYHT0H2_I_25	>>ALLERGEN_2012_2 gi 1703445 sp P54958.1 ASP2_BLAG RecN (352 aa)	3.9	46.7	30

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_25	>>ALLERGEN_2012_2 gi 145105726 gb ABP35603.1 Bla g 2 al (352 aa)	3.9	46.7	30
>SYHT0H2_I_45	>>ALLERGEN_2012_2 gi 1703445 sp P54958.1 ASP2_BLAG RecN (352 aa)	3.9	46.7	30
>SYHT0H2_I_45	>>ALLERGEN_2012_2 gi 145105726 gb ABP35603.1 Bla g 2 al (352 aa)	3.9	46.7	30
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 6136162 sp P35776.2 VA2_SOLRI RecNa (119 aa)	3.9	36.0	25
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 6136162 sp P35776.2 VA2_SOLRI RecNa (119 aa)	3.9	36.0	25
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 168419914 gb ACA23876.1 Pas n 1 al (265 aa)	4	24.6	69
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 168419914 gb ACA23876.1 Pas n 1 al (265 aa)	4	24.6	69
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 335331566 gb AEH31546.1 low molecu (369 aa)	4	27.6	29
>SYHT0H2_I_7	>>ALLERGEN_2012_2 gi 20387027 emb CAC84590.2 tropomyosi (284 aa)	4	33.3	24
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 170718 gb AAA34279.1 alpha/beta-gl (313 aa)	4.1	32.6	43
>SYHT0H2_I_15	>>ALLERGEN_2012_2 gi 291482308 emb CBK62694.1 ragweed h (140 aa)	4.1	28.9	38
>SYHT0H2_I_13	>>ALLERGEN_2012_2 gi 473876 gb AAA17741.1 alpha-gliadin (287 aa)	4.1	30.0	30
>SYHT0H2_I_24	>>ALLERGEN_2012_2 gi 473876 gb AAA17741.1 alpha-gliadin (287 aa)	4.1	30.0	30
>SYHT0H2_I_46	>>ALLERGEN_2012_2 gi 473876 gb AAA17741.1 alpha-gliadin (287 aa)	4.1	30.0	30
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 170702 gb AAA34272.1 gamma gliadin (302 aa)	4.1	50.0	12

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_14	>>ALLERGEN_2012_2 gi 6136165 sp P81657.1 V A5_VESMA RecNa (202 aa)	4.2	37.0	27
>SYHT0H2_I_27	>>ALLERGEN_2012_2 gi 62550933 emb CAI7905 2.1 putative L (326 aa)	4.2	25.0	24
>SYHT0H2_I_43	>>ALLERGEN_2012_2 gi 62550933 emb CAI7905 2.1 putative L (326 aa)	4.2	25.0	24
>SYHT0H2_I_27	>>ALLERGEN_2012_2 gi 56788031 gb AAW2981 0.1 seed storag (507 aa)	4.2	50.0	20
>SYHT0H2_I_43	>>ALLERGEN_2012_2 gi 56788031 gb AAW2981 0.1 seed storag (507 aa)	4.2	50.0	20
>SYHT0H2_I_15	>>ALLERGEN_2012_2 gi 8928058 sp O04298.1 D AU1_DAUCA RecN (154 aa)	4.4	33.3	54
>SYHT0H2_I_15	>>ALLERGEN_2012_2 gi 2154734 emb CAB0371 6.1 major aller (154 aa)	4.4	33.3	54
>SYHT0H2_I_21	>>ALLERGEN_2012_2 gi 7638028 gb AAF65312. 1 AF230383_1 ve (137 aa)	4.5	30.0	40
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 155676684 dbj BAF757 05.1 SXP/RAL-2 (150 aa)	4.6	28.2	39
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 155676696 dbj BAF757 11.1 SXP/RAL-2 (150 aa)	4.6	28.2	39
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 155676688 dbj BAF757 07.1 SXP/RAL-2 (150 aa)	4.6	28.2	39
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 155676636 dbj BAF756 81.1 SXP/RAL-2 (150 aa)	4.6	28.2	39
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 155676682 dbj BAF757 04.1 SXP/RAL-2 (150 aa)	4.6	28.2	39
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 155676698 dbj BAF757 12.1 SXP/RAL-2 (150 aa)	4.6	28.2	39
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 155676694 dbj BAF757 10.1 SXP/RAL-2 (150 aa)	4.6	28.2	39

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 155676686 dbj BAF75706.1 SXP/RAL-2 (150 aa)	4.6	28.2	39
>SYHT0H2_I_15	>>ALLERGEN_2012_2 gi 291197394 emb CBK52317.1 ragweed h (164 aa)	4.6	28.9	38
>SYHT0H2_I_15	>>ALLERGEN_2012_2 gi 285005079 emb CBJ24286.1 ragweed h (164 aa)	4.6	28.9	38
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 30794292 ref NP_851341.1 lactotran (708 aa)	4.6	36.8	38
>SYHT0H2_I_23	>>ALLERGEN_2012_2 gi 291482314 emb CBK62697.1 ragweed h (96 aa)	4.6	34.6	26
>SYHT0H2_I_27	>>ALLERGEN_2012_2 gi 149208403 gb ABR21772.1 conglutin (455 aa)	4.7	50.0	20
>SYHT0H2_I_43	>>ALLERGEN_2012_2 gi 149208403 gb ABR21772.1 conglutin (455 aa)	4.7	50.0	20
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 3319897 emb CAA76841.1 albumin [Ca (585 aa)	4.7	47.1	17
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 3319897 emb CAA76841.1 albumin [Ca (585 aa)	4.7	47.1	17
>SYHT0H2_I_50	>>ALLERGEN_2012_2 gi 4590366 gb AAD26547.1 AF124824_1 ma (159 aa)	4.7	42.9	14
>SYHT0H2_I_50	>>ALLERGEN_2012_2 gi 25361513 gb AAN73248.1 helix-loop- (450 aa)	4.7	63.6	11
>SYHT0H2_I_19	>>ALLERGEN_2012_2 gi 170730 gb AAA34285.1 gamma-gliadin (304 aa)	4.8	25.0	40
>SYHT0H2_I_52	>>ALLERGEN_2012_2 gi 170730 gb AAA34285.1 gamma-gliadin (304 aa)	4.8	25.0	40
>SYHT0H2_I_31	>>ALLERGEN_2012_2 gi 1709545 sp P51528.1 P A1_VESMC RecNa (300 aa)	4.8	50.0	28
>SYHT0H2_I_27	>>ALLERGEN_2012_2 gi 335331566 gb AEH31546.1 low molecu (369 aa)	4.8	25.0	24

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_43	>>ALLERGEN_2012_2 gi 335331566 gb AEH31546.1  low molecu (369 aa)	4.8	25.0	24
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 6687188 emb CAB64867.1  albumin [Ca (608 aa)	4.8	47.1	17
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 1351908 sp P49064.1 A LBU_FELCA RecN (608 aa)	4.8	47.1	17
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 6687188 emb CAB64867.1  albumin [Ca (608 aa)	4.8	47.1	17
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 1351908 sp P49064.1 A LBU_FELCA RecN (608 aa)	4.8	47.1	17
>SYHT0H2_I_39	>>ALLERGEN_2012_2 gi 5777414 emb CAB53458.1  MnSOD [Heve (205 aa)	4.8	29.4	17
>SYHT0H2_I_39	>>ALLERGEN_2012_2 gi 10862818 emb CAC13961.1  IgE-bindin (205 aa)	4.8	29.4	17
>SYHT0H2_I_10	>>ALLERGEN_2012_2 gi 18635 emb CAA33215.1  glycinin subu (495 aa)	4.9	25.8	89
>SYHT0H2_I_16	>>ALLERGEN_2012_2 gi 18635 emb CAA33215.1  glycinin subu (495 aa)	4.9	25.8	89
>SYHT0H2_I_28	>>ALLERGEN_2012_2 gi 18635 emb CAA33215.1  glycinin subu (495 aa)	4.9	25.8	89
>SYHT0H2_I_38	>>ALLERGEN_2012_2 gi 18635 emb CAA33215.1  glycinin subu (495 aa)	4.9	25.8	89
>SYHT0H2_I_25	>>ALLERGEN_2012_2 gi 1398915 dbj BAA07711.1  allergenic (160 aa)	4.9	43.3	30
>SYHT0H2_I_45	>>ALLERGEN_2012_2 gi 1398915 dbj BAA07711.1  allergenic (160 aa)	4.9	43.3	30
>SYHT0H2_I_50	>>ALLERGEN_2012_2 gi 218203828 gb ACK76297.1  Der f 6 al (279 aa)	4.9	55.6	9
>SYHT0H2_I_50	>>ALLERGEN_2012_2 gi 218203826 gb ACK76296.1  Der f 6 al (279 aa)	4.9	55.6	9

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_31	>>ALLERGEN_2012_2 gi 50659885 gb AAT80662.1  lipid trans (115 aa)	5	31.6	38
>SYHT0H2_I_31	>>ALLERGEN_2012_2 gi 50659891 gb AAT80665.1  lipid trans (115 aa)	5	31.6	38
>SYHT0H2_I_31	>>ALLERGEN_2012_2 gi 50659879 gb AAT80659.1  lipid trans (115 aa)	5	31.6	38
>SYHT0H2_I_10	>>ALLERGEN_2012_2 gi 18652047 gb AAL76932.1 AF456481_1 m (154 aa)	5.1	22.7	110
>SYHT0H2_I_16	>>ALLERGEN_2012_2 gi 18652047 gb AAL76932.1 AF456481_1 m (154 aa)	5.1	22.7	110
>SYHT0H2_I_28	>>ALLERGEN_2012_2 gi 18652047 gb AAL76932.1 AF456481_1 m (154 aa)	5.1	22.7	110
>SYHT0H2_I_38	>>ALLERGEN_2012_2 gi 18652047 gb AAL76932.1 AF456481_1 m (154 aa)	5.1	22.7	110
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 112754 sp P04403.2 SS1_BEREX RecNa (146 aa)	5.1	25.7	35
>SYHT0H2_I_25	>>ALLERGEN_2012_2 gi 218193 dbj BAA01998.1  allergenic p (165 aa)	5.1	43.3	30
>SYHT0H2_I_25	>>ALLERGEN_2012_2 gi 1398913 dbj BAA07710.1  allergenic (166 aa)	5.1	43.3	30
>SYHT0H2_I_25	>>ALLERGEN_2012_2 gi 114152865 sp Q01882.2 RAG2_ORYSJ Re (166 aa)	5.1	43.3	30
>SYHT0H2_I_45	>>ALLERGEN_2012_2 gi 218193 dbj BAA01998.1  allergenic p (165 aa)	5.1	43.3	30
>SYHT0H2_I_45	>>ALLERGEN_2012_2 gi 1398913 dbj BAA07710.1  allergenic (166 aa)	5.1	43.3	30
>SYHT0H2_I_45	>>ALLERGEN_2012_2 gi 114152865 sp Q01882.2 RAG2_ORYSJ Re (166 aa)	5.1	43.3	30
>SYHT0H2_I_50	>>ALLERGEN_2012_2 gi 170743 gb AAB02788.1  HMW glutenin (815 aa)	5.1	33.3	24

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_25	>>ALLERGEN_2012_2 gi 47117012 sp Q7M4I5.1 PA2_APIDO RecN (134 aa)	5.1	41.7	12
>SYHT0H2_I_45	>>ALLERGEN_2012_2 gi 47117012 sp Q7M4I5.1 PA2_APIDO RecN (134 aa)	5.1	41.7	12
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 21913174 gb AAM77471.1 major aller (115 aa)	5.1	66.7	9
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 1092249 prf 2023228A major allerge (285 aa)	5.2	28.7	101
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 2398757 emb CAA50281.1 Major Polle (286 aa)	5.2	28.7	101
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 21725630 emb CAD38396.1 unnamed pr (287 aa)	5.2	28.7	101
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 21725612 emb CAD38387.1 unnamed pr (287 aa)	5.2	28.7	101
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 21725614 emb CAD38388.1 unnamed pr (287 aa)	5.2	28.7	101
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 21725616 emb CAD38389.1 unnamed pr (287 aa)	5.2	28.7	101
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 21725618 emb CAD38390.1 unnamed pr (287 aa)	5.2	28.7	101
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 21725620 emb CAD38391.1 unnamed pr (287 aa)	5.2	28.7	101
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 21725622 emb CAD38392.1 unnamed pr (287 aa)	5.2	28.7	101
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 21725624 emb CAD38393.1 unnamed pr (287 aa)	5.2	28.7	101
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 21725626 emb CAD38394.1 unnamed pr (287 aa)	5.2	28.7	101
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 21725632 emb CAD38397.1 unnamed pr (287 aa)	5.2	28.7	101

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 21725606 emb CAD38384.1 unnamed pr (287 aa)	5.2	28.7	101
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 21725610 emb CAD38386.1 unnamed pr (287 aa)	5.2	28.7	101
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 21725608 emb CAD38385.1 unnamed pr (287 aa)	5.2	28.7	101
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 21725628 emb CAD38395.1 unnamed pr (287 aa)	5.2	28.7	101
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 21926 emb CAA36063.1 unnamed prote (295 aa)	5.2	26.1	69
>SYHT0H2_I_20	>>ALLERGEN_2012_2 gi 3287877 sp P81402.1 N LTP1_PRUPE Rec (91 aa)	5.2	37.0	27
>SYHT0H2_I_51	>>ALLERGEN_2012_2 gi 3287877 sp P81402.1 N LTP1_PRUPE Rec (91 aa)	5.2	37.0	27
>SYHT0H2_I_31	>>ALLERGEN_2012_2 gi 897647 gb AAB48072.1 allergen and (336 aa)	5.3	50.0	28
>SYHT0H2_I_20	>>ALLERGEN_2012_2 gi 83754241 pdb 2B5S B Chain B, Crysta (92 aa)	5.3	37.0	27
>SYHT0H2_I_51	>>ALLERGEN_2012_2 gi 83754241 pdb 2B5S B Chain B, Crysta (92 aa)	5.3	37.0	27
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 1063270 dbj BAA11251.1 gamma-gliad (279 aa)	5.4	34.5	29
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 398830 emb CAA52753.1 Phlp5 [Phleu (312 aa)	5.5	27.3	110
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 3309039 gb AAC25994.1 group V alle (312 aa)	5.5	28.7	101
>SYHT0H2_I_4	>>ALLERGEN_2012_2 gi 14424465 sp P35777.2 VA4_SOLIN RecN (137 aa)	5.6	34.8	46
>SYHT0H2_I_4	>>ALLERGEN_2012_2 gi 291092710 gb ADD74392.1 Sol s 2 al (137 aa)	5.6	34.8	46



Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_32	>>ALLERGEN_2012_2 gi 14424465 sp P35777.2 VA4_SOLIN RecN (137 aa)	5.6	34.8	46
>SYHT0H2_I_32	>>ALLERGEN_2012_2 gi 291092710 gb ADD74392.1 Sol s 2 al (137 aa)	5.6	34.8	46
>SYHT0H2_I_14	>>ALLERGEN_2012_2 gi 13183177 gb AAK15089.1 AF240006_1 7 (585 aa)	5.6	34.3	35
>SYHT0H2_I_8	>>ALLERGEN_2012_2 gi 29163773 emb CAD80019.1 unnamed pr (65 aa)	5.6	54.5	11
>SYHT0H2_I_25	>>ALLERGEN_2012_2 gi 897811 emb CAA24933.1 unnamed prot (101 aa)	5.7	47.8	23
>SYHT0H2_I_45	>>ALLERGEN_2012_2 gi 897811 emb CAA24933.1 unnamed prot (101 aa)	5.7	47.8	23
>SYHT0H2_I_17	>>ALLERGEN_2012_2 gi 171464770 gb ACB45874.1 pathogen-r (151 aa)	5.7	50.0	18
>SYHT0H2_I_37	>>ALLERGEN_2012_2 gi 171464770 gb ACB45874.1 pathogen-r (151 aa)	5.7	50.0	18
>SYHT0H2_I_5	>>ALLERGEN_2012_2 gi 194350815 gb ACF53836.1 Bla g 4 is (191 aa)	5.8	30.2	53
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 55859456 emb CAH92630.1 pollen all (520 aa)	5.8	36.7	30
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 21765 emb CAA26385.1 unnamed prote (313 aa)	5.8	40.7	27
>SYHT0H2_I_13	>>ALLERGEN_2012_2 gi 5381323 gb AAD42943.1 AF091841_1 2S (148 aa)	5.8	26.1	23
>SYHT0H2_I_24	>>ALLERGEN_2012_2 gi 5381323 gb AAD42943.1 AF091841_1 2S (148 aa)	5.8	26.1	23
>SYHT0H2_I_46	>>ALLERGEN_2012_2 gi 5381323 gb AAD42943.1 AF091841_1 2S (148 aa)	5.8	26.1	23
>SYHT0H2_I_14	>>ALLERGEN_2012_2 gi 3287877 sp P81402.1 NLTP1_PRUPE Rec (91 aa)	5.8	38.9	18

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_14	>>ALLERGEN_2012_2 gi 83754241 pdb 2B5S B Chain B, Crysta (92 aa)	5.8	38.9	18
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 100834 pir  S16031 alpha-amylase in (168 aa)	5.9	24.3	70
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 21713 emb CAA35597.1  unnamed prote (168 aa)	5.9	24.3	70
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 100834 pir  S16031 alpha-amylase in (168 aa)	5.9	24.3	70
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 21713 emb CAA35597.1  unnamed prote (168 aa)	5.9	24.3	70
>SYHT0H2_I_18	>>ALLERGEN_2012_2 gi 21743 emb CAA43331.1  high molecula (830 aa)	5.9	35.1	37
>SYHT0H2_I_36	>>ALLERGEN_2012_2 gi 21743 emb CAA43331.1  high molecula (830 aa)	5.9	35.1	37
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 112745 sp P23110.1 2SS 8_HELAN RecNa (141 aa)	5.9	33.3	33
>SYHT0H2_I_2	>>ALLERGEN_2012_2 gi 23894244 emb CAD236 14.1  tri m 2 al (404 aa)	5.9	30.0	30
>SYHT0H2_I_2	>>ALLERGEN_2012_2 gi 74663809 sp Q8J077.1 S UB6_TRISH Rec (405 aa)	5.9	30.0	30
>SYHT0H2_I_34	>>ALLERGEN_2012_2 gi 23894244 emb CAD236 14.1  tri m 2 al (404 aa)	5.9	30.0	30
>SYHT0H2_I_34	>>ALLERGEN_2012_2 gi 74663809 sp Q8J077.1 S UB6_TRISH Rec (405 aa)	5.9	30.0	30
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 75219081 sp O22108 O2 2108_WHEAT LMM (285 aa)	6	31.9	72
>SYHT0H2_I_49	>>ALLERGEN_2012_2 gi 1362131 pir  C53806 major allergen (145 aa)	6	34.5	29
>SYHT0H2_I_12	>>ALLERGEN_2012_2 gi 3703107 gb AAC63045.1  glycinin [Ar (507 aa)	6	40.0	15

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_12	>>ALLERGEN_2012_2 gi 224036293 pdb 3C3V A Chain A, Cryst (510 aa)	6	40.0	15
>SYHT0H2_I_41	>>ALLERGEN_2012_2 gi 3703107 gb AAC63045.1  glycinin [Ar (507 aa)	6	40.0	15
>SYHT0H2_I_41	>>ALLERGEN_2012_2 gi 224036293 pdb 3C3V A Chain A, Cryst (510 aa)	6	40.0	15
>SYHT0H2_I_15	>>ALLERGEN_2012_2 gi 30316292 sp Q9FSG7.1  TP1A_MALDO Rec (246 aa)	6.1	21.1	71
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 886967 emb CAA59340.1  low molecula (276 aa)	6.1	31.6	38
>SYHT0H2_I_25	>>ALLERGEN_2012_2 gi 7435005 pir  A59055 phospholipase A (134 aa)	6.1	41.7	12
>SYHT0H2_I_25	>>ALLERGEN_2012_2 gi 24638082 sp Q9BMK4.1 PA2_APICC RecN (134 aa)	6.1	41.7	12
>SYHT0H2_I_45	>>ALLERGEN_2012_2 gi 7435005 pir  A59055 phospholipase A (134 aa)	6.1	41.7	12
>SYHT0H2_I_45	>>ALLERGEN_2012_2 gi 24638082 sp Q9BMK4.1 PA2_APICC RecN (134 aa)	6.1	41.7	12
>SYHT0H2_I_31	>>ALLERGEN_2012_2 gi 50659889 gb AAT80664.1  lipid trans (115 aa)	6.2	31.6	38
>SYHT0H2_I_15	>>ALLERGEN_2012_2 gi 1545895 emb CAB0221 6.1  pollen alle (161 aa)	6.2	31.3	32
>SYHT0H2_I_11	>>ALLERGEN_2012_2 gi 16580747 dbj BAB7174 1.1  glyoxalase (291 aa)	6.2	35.7	28
>SYHT0H2_I_11	>>ALLERGEN_2012_2 gi 84029333 sp Q948T6.2  LGUL_ORYSJ Rec (291 aa)	6.2	35.7	28
>SYHT0H2_I_42	>>ALLERGEN_2012_2 gi 16580747 dbj BAB7174 1.1  glyoxalase (291 aa)	6.2	35.7	28
>SYHT0H2_I_42	>>ALLERGEN_2012_2 gi 84029333 sp Q948T6.2  LGUL_ORYSJ Rec (291 aa)	6.2	35.7	28

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_50	>>ALLERGEN_2012_2 gi 238477265 gb ACR43474.1  arginine k (356 aa)	6.2	30.8	26
>SYHT0H2_I_20	>>ALLERGEN_2012_2 gi 2506771 sp P16968.2 IA A1_HORVU RecN (146 aa)	6.2	37.5	24
>SYHT0H2_I_51	>>ALLERGEN_2012_2 gi 2506771 sp P16968.2 IA A1_HORVU RecN (146 aa)	6.2	37.5	24
>SYHT0H2_I_2	>>ALLERGEN_2012_2 gi 6136163 sp P35779.2 V A3_SOLRI RecNa (211 aa)	6.2	38.9	18
>SYHT0H2_I_34	>>ALLERGEN_2012_2 gi 6136163 sp P35779.2 V A3_SOLRI RecNa (211 aa)	6.2	38.9	18
>SYHT0H2_I_12	>>ALLERGEN_2012_2 gi 5712199 gb AAD47382.1  glycinin [Ar (530 aa)	6.2	40.0	15
>SYHT0H2_I_12	>>ALLERGEN_2012_2 gi 199732457 gb ACH91862.1  arachin Ar (530 aa)	6.2	40.0	15
>SYHT0H2_I_12	>>ALLERGEN_2012_2 gi 21314465 gb AAM46958.1 AF510854_1 a (538 aa)	6.2	40.0	15
>SYHT0H2_I_17	>>ALLERGEN_2012_2 gi 46410859 gb AAR98518.1  major latex (366 aa)	6.2	46.7	15
>SYHT0H2_I_37	>>ALLERGEN_2012_2 gi 46410859 gb AAR98518.1  major latex (366 aa)	6.2	46.7	15
>SYHT0H2_I_41	>>ALLERGEN_2012_2 gi 5712199 gb AAD47382.1  glycinin [Ar (530 aa)	6.2	40.0	15
>SYHT0H2_I_41	>>ALLERGEN_2012_2 gi 199732457 gb ACH91862.1  arachin Ar (530 aa)	6.2	40.0	15
>SYHT0H2_I_41	>>ALLERGEN_2012_2 gi 21314465 gb AAM46958.1 AF510854_1 a (538 aa)	6.2	40.0	15
>SYHT0H2_I_39	>>ALLERGEN_2012_2 gi 212279 gb AAA48944.1  lysozyme prot (24 aa)	6.2	54.5	11
>SYHT0H2_I_20	>>ALLERGEN_2012_2 gi 1311510 gb AAB36009.1  mAb 8C7-reac (15 aa)	6.2	55.6	9

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_51	>>ALLERGEN_2012_2 gi 1311510 gb AAB36009.1  mAb 8C7-reac (15 aa)	6.2	55.6	9
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 170730 gb AAA34285.1  gamma-gliadin (304 aa)	6.3	27.7	65
>SYHT0H2_I_27	>>ALLERGEN_2012_2 gi 169950562 gb ACB05815.1  conglutin (611 aa)	6.3	50.0	20
>SYHT0H2_I_43	>>ALLERGEN_2012_2 gi 169950562 gb ACB05815.1  conglutin (611 aa)	6.3	50.0	20
>SYHT0H2_I_9	>>ALLERGEN_2012_2 gi 162929 gb AAA30479.1  alpha-s2-like (222 aa)	6.4	20.5	73
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 21773 emb CAA31685.1  unnamed prote (307 aa)	6.4	26.1	69
>SYHT0H2_I_20	>>ALLERGEN_2012_2 gi 313575718 gb ADR66939.1  non-specif (117 aa)	6.4	37.0	27
>SYHT0H2_I_20	>>ALLERGEN_2012_2 gi 288561913 sp P85894.1  LTP1_MORNI Re (91 aa)	6.4	33.3	27
>SYHT0H2_I_51	>>ALLERGEN_2012_2 gi 313575718 gb ADR66939.1  non-specif (117 aa)	6.4	37.0	27
>SYHT0H2_I_51	>>ALLERGEN_2012_2 gi 288561913 sp P85894.1  LTP1_MORNI Re (91 aa)	6.4	33.3	27
>SYHT0H2_I_49	>>ALLERGEN_2012_2 gi 261824817 pdb 3F55 D Chain D, Cryst (316 aa)	6.4	50.0	20
>SYHT0H2_I_11	>>ALLERGEN_2012_2 gi 74664773 sp Q96X46.3  ENO_PENCI RecN (438 aa)	6.4	42.1	19
>SYHT0H2_I_42	>>ALLERGEN_2012_2 gi 74664773 sp Q96X46.3  ENO_PENCI RecN (438 aa)	6.4	42.1	19
>SYHT0H2_I_50	>>ALLERGEN_2012_2 gi 897811 emb CAA24933.1  unnamed prot (101 aa)	6.4	38.5	13
>SYHT0H2_I_9	>>ALLERGEN_2012_2 gi 4138175 emb CAA09885.1  allergen [M (187 aa)	6.5	53.3	15

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 62484809 emb CAI78902.1  putative g (285 aa)	6.6	34.5	29
>SYHT0H2_I_20	>>ALLERGEN_2012_2 gi 5777414 emb CAB53458.1  MnSOD [Heve (205 aa)	6.6	40.9	22
>SYHT0H2_I_20	>>ALLERGEN_2012_2 gi 10862818 emb CAC13961.1  IgE-bindin (205 aa)	6.6	40.9	22
>SYHT0H2_I_51	>>ALLERGEN_2012_2 gi 5777414 emb CAB53458.1  MnSOD [Heve (205 aa)	6.6	40.9	22
>SYHT0H2_I_51	>>ALLERGEN_2012_2 gi 10862818 emb CAC13961.1  IgE-bindin (205 aa)	6.6	40.9	22
>SYHT0H2_I_2	>>ALLERGEN_2012_2 gi 291482306 emb CBK62693.1  ragweed h (111 aa)	6.6	55.6	9
>SYHT0H2_I_34	>>ALLERGEN_2012_2 gi 291482306 emb CBK62693.1  ragweed h (111 aa)	6.6	55.6	9
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 208605344 emb CAR82265.1  D-type LM (359 aa)	6.7	39.1	23
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 1842045 gb AAB47552.1  major allerg (157 aa)	6.7	66.7	9
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 45680856 gb AAS75297.1  major aller (157 aa)	6.7	66.7	9
>SYHT0H2_I_40	>>ALLERGEN_2012_2 gi 9954251 gb AAG08988.1 AF216519_1 tr (284 aa)	6.8	42.9	35
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 169971 gb AAA33965.1  glycinin prec (240 aa)	6.8	39.1	23
>SYHT0H2_I_2	>>ALLERGEN_2012_2 gi 14424466 sp P35778.2 VA3_SOLIN RecN (234 aa)	6.8	38.9	18
>SYHT0H2_I_34	>>ALLERGEN_2012_2 gi 14424466 sp P35778.2 VA3_SOLIN RecN (234 aa)	6.8	38.9	18
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 13430402 gb AAK25823.1  group V all (275 aa)	6.9	25.3	166

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_23	>>ALLERGEN_2012_2 gi 3703107 gb AAC63045.1  glycinin [Ar (507 aa)	6.9	32.1	28
>SYHT0H2_I_23	>>ALLERGEN_2012_2 gi 224036293 pdb 3C3V A Chain A, Cryst (510 aa)	6.9	32.1	28
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 76782247 gb ABA54897.1  hydrophobic (134 aa)	7	40.9	22
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 2497701 sp Q28133.1 A LL2_BOVIN RecN (172 aa)	7.1	27.8	54
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 2497701 sp Q28133.1 A LL2_BOVIN RecN (172 aa)	7.1	27.8	54
>SYHT0H2_I_49	>>ALLERGEN_2012_2 gi 1063270 dbj BAA11251.1  gamma-gliad (279 aa)	7.1	36.0	25
>SYHT0H2_I_14	>>ALLERGEN_2012_2 gi 54793477 gb AAV40850.1  lipid trans (117 aa)	7.1	38.9	18
>SYHT0H2_I_14	>>ALLERGEN_2012_2 gi 313575718 gb ADR66939.1  non-specif (117 aa)	7.1	38.9	18
>SYHT0H2_I_11	>>ALLERGEN_2012_2 gi 127533 sp P02761.1 MU P_RAT RecName: (181 aa)	7.2	34.0	47
>SYHT0H2_I_42	>>ALLERGEN_2012_2 gi 127533 sp P02761.1 MU P_RAT RecName: (181 aa)	7.2	34.0	47
>SYHT0H2_I_23	>>ALLERGEN_2012_2 gi 5712199 gb AAD47382.1  glycinin [Ar (530 aa)	7.2	32.1	28
>SYHT0H2_I_23	>>ALLERGEN_2012_2 gi 199732457 gb ACH91862.1  arachin Ar (530 aa)	7.2	32.1	28
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 21755 emb CAA25593.1  unnamed prote (286 aa)	7.2	45.0	20
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 21761 emb CAA26384.1  unnamed prote (286 aa)	7.2	45.0	20
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 170720 gb AAA34280.1  alpha/beta-gl (286 aa)	7.2	45.0	20

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_47	>>ALLERGEN_2012_2 gi 146217148 gb ABQ10638.1  vicilin-li (136 aa)	7.2	60.0	10
>SYHT0H2_I_30	>>ALLERGEN_2012_2 gi 170738 gb AAA34289.1  gamma-gliadin (327 aa)	7.3	37.5	32
>SYHT0H2_I_23	>>ALLERGEN_2012_2 gi 21314465 gb AAM46958.1 AF510854_1 a (538 aa)	7.3	32.1	28
>SYHT0H2_I_20	>>ALLERGEN_2012_2 gi 348137 gb AAA16792.1  superoxide di (233 aa)	7.3	40.9	22
>SYHT0H2_I_51	>>ALLERGEN_2012_2 gi 348137 gb AAA16792.1  superoxide di (233 aa)	7.3	40.9	22
>SYHT0H2_I_29	>>ALLERGEN_2012_2 gi 322812205 pdb 2X45 A Chain A, Cryst (144 aa)	7.3	33.3	18
>SYHT0H2_I_18	>>ALLERGEN_2012_2 gi 13183175 gb AAK15088.1 AF240005_1 2 (153 aa)	7.4	52.2	23
>SYHT0H2_I_18	>>ALLERGEN_2012_2 gi 209165427 gb ACI41244.1  2S albumin (153 aa)	7.4	52.2	23
>SYHT0H2_I_36	>>ALLERGEN_2012_2 gi 13183175 gb AAK15088.1 AF240005_1 2 (153 aa)	7.4	52.2	23
>SYHT0H2_I_36	>>ALLERGEN_2012_2 gi 209165427 gb ACI41244.1  2S albumin (153 aa)	7.4	52.2	23
>SYHT0H2_I_5	>>ALLERGEN_2012_2 gi 105969545 gb ABF81662.1  EXPB10 [Ze (269 aa)	7.4	55.0	20
>SYHT0H2_I_5	>>ALLERGEN_2012_2 gi 115502168 sp P0C1Y5.1 EXB11_MAIZE R (269 aa)	7.4	55.0	20
>SYHT0H2_I_14	>>ALLERGEN_2012_2 gi 1398916 dbj BAA07712.1  allergenic (157 aa)	7.4	45.0	20
>SYHT0H2_I_18	>>ALLERGEN_2012_2 gi 166235350 pdb 2JON A Chain A, Solut (101 aa)	7.5	27.9	61
>SYHT0H2_I_36	>>ALLERGEN_2012_2 gi 166235350 pdb 2JON A Chain A, Solut (101 aa)	7.5	27.9	61



Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_13	>>ALLERGEN_2012_2 gi 1170095 sp P46419.1 GSTM1_DERPT Rec (219 aa)	7.5	35.3	34
>SYHT0H2_I_13	>>ALLERGEN_2012_2 gi 60920878 gb AAX37326.1  glutathione (219 aa)	7.5	35.3	34
>SYHT0H2_I_24	>>ALLERGEN_2012_2 gi 1170095 sp P46419.1 GSTM1_DERPT Rec (219 aa)	7.5	35.3	34
>SYHT0H2_I_24	>>ALLERGEN_2012_2 gi 60920878 gb AAX37326.1  glutathione (219 aa)	7.5	35.3	34
>SYHT0H2_I_46	>>ALLERGEN_2012_2 gi 1170095 sp P46419.1 GSTM1_DERPT Rec (219 aa)	7.5	35.3	34
>SYHT0H2_I_46	>>ALLERGEN_2012_2 gi 60920878 gb AAX37326.1  glutathione (219 aa)	7.5	35.3	34
>SYHT0H2_I_49	>>ALLERGEN_2012_2 gi 33327133 gb AAQ08947.1  allergen Fr (145 aa)	7.5	34.5	29
>SYHT0H2_I_29	>>ALLERGEN_2012_2 gi 462026 sp P34071.1 EN TC2_STAAU RecN (266 aa)	7.5	33.3	24
>SYHT0H2_I_49	>>ALLERGEN_2012_2 gi 124365253 gb ABN09655.1  beta-1,3-g (374 aa)	7.5	50.0	20
>SYHT0H2_I_49	>>ALLERGEN_2012_2 gi 124294785 gb ABN03966.1  beta-1,3-g (374 aa)	7.5	50.0	20
>SYHT0H2_I_49	>>ALLERGEN_2012_2 gi 268037674 gb ACY91851.1  beta-1,3-g (374 aa)	7.5	50.0	20
>SYHT0H2_I_47	>>ALLERGEN_2012_2 gi 47606039 sp Q8H6L7.1 PHLB_PHLPR Rec (143 aa)	7.6	31.3	32
>SYHT0H2_I_5	>>ALLERGEN_2012_2 gi 14285595 sp O04004.1 NLTP6_AMBAR Re (118 aa)	7.6	46.7	15
>SYHT0H2_I_11	>>ALLERGEN_2012_2 gi 323575367 dbj BAJ78223.1  Ani s 12 (295 aa)	7.7	37.0	27
>SYHT0H2_I_26	>>ALLERGEN_2012_2 gi 21926 emb CAA36063.1  unnamed prote (295 aa)	7.7	37.0	27

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_42	>>ALLERGEN_2012_2 gi 323575367 dbj BAJ78223.1  Ani s 12 (295 aa)	7.7	37.0	27
>SYHT0H2_I_44	>>ALLERGEN_2012_2 gi 21926 emb CAA36063.1  unnamed prote (295 aa)	7.7	37.0	27
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 244610 gb AAB21323.1  tetrameric al (18 aa)	7.7	47.4	19
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 244610 gb AAB21323.1  tetrameric al (18 aa)	7.7	47.4	19
>SYHT0H2_I_25	>>ALLERGEN_2012_2 gi 163825 gb AAC37318.1  major allerge (92 aa)	7.8	34.5	29
>SYHT0H2_I_45	>>ALLERGEN_2012_2 gi 163825 gb AAC37318.1  major allerge (92 aa)	7.8	34.5	29
>SYHT0H2_I_20	>>ALLERGEN_2012_2 gi 313575726 gb ADR66943.1  non-specif (117 aa)	7.8	37.0	27
>SYHT0H2_I_51	>>ALLERGEN_2012_2 gi 313575726 gb ADR66943.1  non-specif (117 aa)	7.8	37.0	27
>SYHT0H2_I_11	>>ALLERGEN_2012_2 gi 11991229 gb AAG42255.1 AF306708_1 p (296 aa)	7.8	40.9	22
>SYHT0H2_I_42	>>ALLERGEN_2012_2 gi 11991229 gb AAG42255.1 AF306708_1 p (296 aa)	7.8	40.9	22
>SYHT0H2_I_11	>>ALLERGEN_2012_2 gi 510515 emb CAA56343.1  Kunitz tryps (208 aa)	7.8	41.2	17
>SYHT0H2_I_42	>>ALLERGEN_2012_2 gi 510515 emb CAA56343.1  Kunitz tryps (208 aa)	7.8	41.2	17
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 121308878 dbj BAF43534.1  SXP/RAL-2 (152 aa)	7.8	54.5	11
>SYHT0H2_I_11	>>ALLERGEN_2012_2 gi 11991227 gb AAG42254.1 AF306707_1 p (303 aa)	7.9	40.9	22
>SYHT0H2_I_42	>>ALLERGEN_2012_2 gi 11991227 gb AAG42254.1 AF306707_1 p (303 aa)	7.9	40.9	22

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_21	>>ALLERGEN_2012_2 gi 4038411 gb AAC97370.1  venom allerg (137 aa)	8	27.5	40
>SYHT0H2_I_31	>>ALLERGEN_2012_2 gi 129614 sp P00784.1 PA PA1_CARPA RecN (345 aa)	8	33.3	24
>SYHT0H2_I_29	>>ALLERGEN_2012_2 gi 58371884 emb CAG26895.1  Arg r 1 pr (159 aa)	8	33.3	18
>SYHT0H2_I_40	>>ALLERGEN_2012_2 gi 633938 gb AAB30434.1  albumin [Cani (265 aa)	8	50.0	12
>SYHT0H2_I_9	>>ALLERGEN_2012_2 gi 18479082 gb AAL73404.1 AF449424_1 1 (515 aa)	8.1	24.0	50
>SYHT0H2_I_49	>>ALLERGEN_2012_2 gi 170736 gb AAA34288.1  gamma-gliadin (251 aa)	8.1	44.4	18
>SYHT0H2_I_2	>>ALLERGEN_2012_2 gi 291482308 emb CBK62694.1  ragweed h (140 aa)	8.1	55.6	9
>SYHT0H2_I_34	>>ALLERGEN_2012_2 gi 291482308 emb CBK62694.1  ragweed h (140 aa)	8.1	55.6	9
>SYHT0H2_I_39	>>ALLERGEN_2012_2 gi 1168402 sp P42058.1 A LTA7_ALTAL Rec (204 aa)	8.1	57.1	7
>SYHT0H2_I_22	>>ALLERGEN_2012_2 gi 29500897 emb CAD87529.1  phl p5a al (284 aa)	8.2	32.0	75
>SYHT0H2_I_29	>>ALLERGEN_2012_2 gi 757851 emb CAA26040.1  ovotransferr (705 aa)	8.2	41.2	17
>SYHT0H2_I_29	>>ALLERGEN_2012_2 gi 1351295 sp P02789.2 T RFE_CHICK RecN (705 aa)	8.2	41.2	17
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 335331566 gb AEH31546.1  low molecu (369 aa)	8.3	24.1	79
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 335331566 gb AEH31546.1  low molecu (369 aa)	8.3	24.1	79
>SYHT0H2_I_18	>>ALLERGEN_2012_2 gi 886967 emb CAA59340.1  low molecula (276 aa)	8.3	47.6	21

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_36	>>ALLERGEN_2012_2 gi 886967 emb CAA59340.1  low molecula (276 aa)	8.3	47.6	21
>SYHT0H2_I_23	>>ALLERGEN_2012_2 gi 159793201 gb ABW98945.1  alpha S1 c (172 aa)	8.3	45.0	20
>SYHT0H2_I_18	>>ALLERGEN_2012_2 gi 736319 emb CAA27052.1  glutenin [Tr (838 aa)	8.4	31.7	60
>SYHT0H2_I_36	>>ALLERGEN_2012_2 gi 736319 emb CAA27052.1  glutenin [Tr (838 aa)	8.4	31.7	60
>SYHT0H2_I_9	>>ALLERGEN_2012_2 gi 21666498 gb AAM73729.1 AF395893_1 v (536 aa)	8.4	28.3	46
>SYHT0H2_I_7	>>ALLERGEN_2012_2 gi 62484809 emb CAI78902.1  putative g (285 aa)	8.4	29.6	27
>SYHT0H2_I_9	>>ALLERGEN_2012_2 gi 315190620 gb ADT89774.1  phospholip (302 aa)	8.5	28.3	60
>SYHT0H2_I_9	>>ALLERGEN_2012_2 gi 21914823 gb AAM73730.2 AF395894_1 v (538 aa)	8.5	28.3	46
>SYHT0H2_I_21	>>ALLERGEN_2012_2 gi 631911 pir  S43242 allergen-like pr (145 aa)	8.5	25.0	28
>SYHT0H2_I_14	>>ALLERGEN_2012_2 gi 156778059 gb ABU95411.1  Sin a 3 al (92 aa)	8.5	30.8	26
>SYHT0H2_I_5	>>ALLERGEN_2012_2 gi 190613907 gb ACE80957.1  putative a (246 aa)	8.5	43.5	23
>SYHT0H2_I_39	>>ALLERGEN_2012_2 gi 544619 gb AAB29345.1  36 kda allerg (25 aa)	8.5	50.0	8
>SYHT0H2_I_10	>>ALLERGEN_2012_2 gi 2266625 emb CAB10765.1  group V all (264 aa)	8.6	28.8	59
>SYHT0H2_I_16	>>ALLERGEN_2012_2 gi 2266625 emb CAB10765.1  group V all (264 aa)	8.6	28.8	59
>SYHT0H2_I_28	>>ALLERGEN_2012_2 gi 2266625 emb CAB10765.1  group V all (264 aa)	8.6	28.8	59

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_38	>>ALLERGEN_2012_2 gi 2266625 emb CAB10765.1  group V all (264 aa)	8.6	28.8	59
>SYHT0H2_I_3	>>ALLERGEN_2012_2 gi 14422361 emb CAC41634.1  plantain p (131 aa)	8.6	40.0	25
>SYHT0H2_I_3	>>ALLERGEN_2012_2 gi 14422359 emb CAC41633.1  plantain p (131 aa)	8.6	40.0	25
>SYHT0H2_I_33	>>ALLERGEN_2012_2 gi 14422361 emb CAC41634.1  plantain p (131 aa)	8.6	40.0	25
>SYHT0H2_I_33	>>ALLERGEN_2012_2 gi 14422359 emb CAC41633.1  plantain p (131 aa)	8.6	40.0	25
>SYHT0H2_I_3	>>ALLERGEN_2012_2 gi 27818335 gb AAO24900.1  major polle (132 aa)	8.6	66.7	9
>SYHT0H2_I_33	>>ALLERGEN_2012_2 gi 27818335 gb AAO24900.1  major polle (132 aa)	8.6	66.7	9
>SYHT0H2_I_26	>>ALLERGEN_2012_2 gi 262232390 gb ACY38525.1  allergen C (174 aa)	8.7	30.2	43
>SYHT0H2_I_44	>>ALLERGEN_2012_2 gi 262232390 gb ACY38525.1  allergen C (174 aa)	8.7	30.2	43
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 886965 emb CAA59339.1  low molecula (261 aa)	8.7	26.7	30
>SYHT0H2_I_39	>>ALLERGEN_2012_2 gi 21413 emb CAA45723.1  aspartic prot (217 aa)	8.7	46.7	15
>SYHT0H2_I_40	>>ALLERGEN_2012_2 gi 85687540 gb ABC73706.1  allergen pr (140 aa)	8.7	50.0	14
>SYHT0H2_I_49	>>ALLERGEN_2012_2 gi 83300389 sp O42799.2  ALL7_ASPFU Rec (270 aa)	8.7	66.7	6
>SYHT0H2_I_15	>>ALLERGEN_2012_2 gi 14422361 emb CAC41634.1  plantain p (131 aa)	8.8	25.0	48
>SYHT0H2_I_15	>>ALLERGEN_2012_2 gi 14422363 emb CAC41635.1  plantain p (131 aa)	8.8	25.0	48

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_31	>>ALLERGEN_2012_2 gi 74035843 emb CAJ2893.1  Ves g l al (300 aa)	8.8	46.4	28
>SYHT0H2_I_8	>>ALLERGEN_2012_2 gi 3668408 gb AAC61869.1  tropomyosin (160 aa)	8.9	23.4	47
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 21757 emb CAA26383.1  unnamed prote (296 aa)	8.9	52.6	19
>SYHT0H2_I_6	>>ALLERGEN_2012_2 gi 170740 gb AAA34290.1  gliadin [Trit (296 aa)	8.9	52.6	19
>SYHT0H2_I_5	>>ALLERGEN_2012_2 gi 19009 emb CAA46705.1  CMe [Hordeum (148 aa)	8.9	33.3	15
>SYHT0H2_I_39	>>ALLERGEN_2012_2 gi 20141714 sp P30941.2 S PI7_SOLTU Rec (221 aa)	8.9	46.7	15
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 2506771 sp P16968.2 IA A1_HORVU RecN (146 aa)	9	30.4	46
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 2506771 sp P16968.2 IA A1_HORVU RecN (146 aa)	9	30.4	46
>SYHT0H2_I_48	>>ALLERGEN_2012_2 gi 157418806 gb ABV55106.1  Ani s 9 al (147 aa)	9	54.5	11
>SYHT0H2_I_31	>>ALLERGEN_2012_2 gi 170718 gb AAA34279.1  alpha/beta-gl (313 aa)	9.1	34.5	29
>SYHT0H2_I_31	>>ALLERGEN_2012_2 gi 21765 emb CAA26385.1  unnamed prote (313 aa)	9.1	34.5	29
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 18641 emb CAA37044.1  glycinin [Gly (562 aa)	9.2	26.5	34
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 18641 emb CAA37044.1  glycinin [Gly (562 aa)	9.2	26.5	34
>SYHT0H2_I_14	>>ALLERGEN_2012_2 gi 119524036 gb ABL77410.1  UA3-recogn (1096 aa)	9.2	37.9	29
>SYHT0H2_I_31	>>ALLERGEN_2012_2 gi 170710 gb AAA34275.1  alpha-type gl (318 aa)	9.2	34.5	29

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_19	>>ALLERGEN_2012_2 gi 194350817 gb ACF53837.1 Bla g 4 is (190 aa)	9.2	41.2	17
>SYHT0H2_I_52	>>ALLERGEN_2012_2 gi 194350817 gb ACF53837.1 Bla g 4 is (190 aa)	9.2	41.2	17
>SYHT0H2_I_20	>>ALLERGEN_2012_2 gi 23894244 emb CAD23614.1 tri m 2 al (404 aa)	9.3	30.0	30
>SYHT0H2_I_20	>>ALLERGEN_2012_2 gi 74663809 sp Q8J077.1 S UB6_TRISH Rec (405 aa)	9.3	30.0	30
>SYHT0H2_I_51	>>ALLERGEN_2012_2 gi 23894244 emb CAD23614.1 tri m 2 al (404 aa)	9.3	30.0	30
>SYHT0H2_I_51	>>ALLERGEN_2012_2 gi 74663809 sp Q8J077.1 S UB6_TRISH Rec (405 aa)	9.3	30.0	30
>SYHT0H2_I_49	>>ALLERGEN_2012_2 gi 170708 gb AAA34274.1 gamma-gliadin (291 aa)	9.3	36.0	25
>SYHT0H2_I_2	>>ALLERGEN_2012_2 gi 11514279 pdb 1QNX A Chain A, Ves V (209 aa)	9.3	37.5	16
>SYHT0H2_I_34	>>ALLERGEN_2012_2 gi 11514279 pdb 1QNX A Chain A, Ves V (209 aa)	9.3	37.5	16
>SYHT0H2_I_2	>>ALLERGEN_2012_2 gi 285005079 emb CBJ24286.1 ragweed h (164 aa)	9.3	55.6	9
>SYHT0H2_I_2	>>ALLERGEN_2012_2 gi 313118253 sp POCH89.1 HUGA_VESMC Re (31 aa)	9.3	44.4	9
>SYHT0H2_I_34	>>ALLERGEN_2012_2 gi 285005079 emb CBJ24286.1 ragweed h (164 aa)	9.3	55.6	9
>SYHT0H2_I_34	>>ALLERGEN_2012_2 gi 313118253 sp POCH89.1 HUGA_VESMC Re (31 aa)	9.3	44.4	9
>SYHT0H2_I_8	>>ALLERGEN_2012_2 gi 42559558 sp O97192.1 TPM_HELAS RecN (284 aa)	9.4	25.5	47
>SYHT0H2_I_49	>>ALLERGEN_2012_2 gi 56122438 gb AAV74343.1 Fra e 1.010 (145 aa)	9.4	34.5	29

Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_49	>>ALLERGEN_2012_2 gi 1184668 gb AAA87456.1  beta-1,3-glu (374 aa)	9.4	50.0	20
>SYHT0H2_I_49	>>ALLERGEN_2012_2 gi 124365249 gb ABN09653.1  beta-1,3-g (374 aa)	9.4	50.0	20
>SYHT0H2_I_49	>>ALLERGEN_2012_2 gi 124294783 gb ABN03965.1  beta-1,3-g (374 aa)	9.4	50.0	20
>SYHT0H2_I_49	>>ALLERGEN_2012_2 gi 32765543 gb AAP87281.1  beta-1,3-gl (374 aa)	9.4	50.0	20
>SYHT0H2_I_49	>>ALLERGEN_2012_2 gi 270315180 gb ACZ74626.1  beta-1,3-g (374 aa)	9.4	50.0	20
>SYHT0H2_I_49	>>ALLERGEN_2012_2 gi 124365251 gb ABN09654.1  beta-1,3-g (374 aa)	9.4	50.0	20
>SYHT0H2_I_31	>>ALLERGEN_2012_2 gi 1336813 gb AAB36121.1  Sol i 1=anti (26 aa)	9.4	41.7	12
>SYHT0H2_I_5	>>ALLERGEN_2012_2 gi 170712 gb AAA34276.1  pre-alpha-/be (291 aa)	9.6	29.7	64
>SYHT0H2_I_20	>>ALLERGEN_2012_2 gi 9929163 emb CAC05258.1  Cup a 3 pro (199 aa)	9.6	25.0	24
>SYHT0H2_I_51	>>ALLERGEN_2012_2 gi 9929163 emb CAC05258.1  Cup a 3 pro (199 aa)	9.6	25.0	24
>SYHT0H2_I_18	>>ALLERGEN_2012_2 gi 170743 gb AAB02788.1  HMW glutenin (815 aa)	9.7	26.6	64
>SYHT0H2_I_36	>>ALLERGEN_2012_2 gi 170743 gb AAB02788.1  HMW glutenin (815 aa)	9.7	26.6	64
>SYHT0H2_I_15	>>ALLERGEN_2012_2 gi 190613903 gb ACE80955.1  putative a (242 aa)	9.7	37.5	32
>SYHT0H2_I_50	>>ALLERGEN_2012_2 gi 219815476 gb ACL36923.1  troponin C (153 aa)	9.7	23.8	21
>SYHT0H2_I_13	>>ALLERGEN_2012_2 gi 105969543 gb ABF81661.1  EXPB10 [Ze (99 aa)	9.7	42.9	14



Query	Hit Name	E value	Percent Identity	Alignment Length
>SYHT0H2_I_24	>>ALLERGEN_2012_2 gi 105969543 gb ABF8166.1  EXPB10 [Ze (99 aa)	9.7	42.9	14
>SYHT0H2_I_46	>>ALLERGEN_2012_2 gi 105969543 gb ABF8166.1  EXPB10 [Ze (99 aa)	9.7	42.9	14
>SYHT0H2_I_19	>>ALLERGEN_2012_2 gi 886967 emb CAA59340.1  low molecula (276 aa)	9.8	25.6	43
>SYHT0H2_I_52	>>ALLERGEN_2012_2 gi 886967 emb CAA59340.1  low molecula (276 aa)	9.8	25.6	43
>SYHT0H2_I_31	>>ALLERGEN_2012_2 gi 48428170 sp Q9NFQ4.1  ALL22_GLYDO Re (125 aa)	9.8	44.4	18
>SYHT0H2_I_1	>>ALLERGEN_2012_2 gi 1170095 sp P46419.1 G STM1_DERPT Rec (219 aa)	9.9	29.5	61
>SYHT0H2_I_35	>>ALLERGEN_2012_2 gi 1170095 sp P46419.1 G STM1_DERPT Rec (219 aa)	9.9	29.5	61
>SYHT0H2_I_31	>>ALLERGEN_2012_2 gi 61970231 gb AAX5757.8.1  BW 16kDa al (127 aa)	9.9	27.3	33
>SYHT0H2_I_23	>>ALLERGEN_2012_2 gi 159793197 gb ABW989.43.1  alpha S1 c (205 aa)	10	45.0	20
>SYHT0H2_I_2	>>ALLERGEN_2012_2 gi 162551 gb AAA30333.1  allergen 5 [V (227 aa)	10	37.5	16
>SYHT0H2_I_34	>>ALLERGEN_2012_2 gi 162551 gb AAA30333.1  allergen 5 [V (227 aa)	10	37.5	16

>SYHT0H2\_I\_1  
vs /data/fasta/ALLERGEN\_2012\_2 library  
searching /data/fasta/ALLERGEN\_2012\_2 library

367796 residues in 1603 sequences  
Expectation\_n fit:  $\rho(\ln(x)) = 3.7938 \pm 0.00334$ ;  $\mu = 10.5821 \pm 0.172$   
mean\_var=54.5589 $\pm$ 14.161, 0's: 3 Z-trim: 3 B-trim: 20 in 2/41  
Lambda= 0.1736

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2  
join: 36, opt: 24, gap-pen: -12/-2, width: 16  
Scan time: 0.040

The best scores are: opt bits E(1603)  
ALLERGEN\_2012\_2|gi|75219081|sp|O22108|O22108\_WHEA ( 285) 71 26 1

ALLERGEN_2012_2	gi	emb	CAA31685.1	unnamed	( 307)	71	26	1.1
ALLERGEN_2012_2	gi	886965	emb	CAA59339.1	low mol	( 261)	67	2
ALLERGEN_2012_2	gi	94400907	ref	NP_001035360.1	a	( 92)	61	2.8
ALLERGEN_2012_2	gi	633938	gb	AAB30434.1	albumin	( 265)	65	2.8
ALLERGEN_2012_2	gi	88770352	gb	ABD51779.1	allerg	( 94)	61	2.8
ALLERGEN_2012_2	gi	3021324	emb	CAA06305.1	Aspf1	( 125)	61	3.4
ALLERGEN_2012_2	gi	886967	emb	CAA59340.1	low mol	( 276)	64	3.4
ALLERGEN_2012_2	gi	6136162	sp	P35776.2	VA2_SOLRI	( 119)	60	3.9
ALLERGEN_2012_2	gi	168419914	gb	ACA23876.1	Pas n	( 265)	63	4
ALLERGEN_2012_2	gi	3319897	emb	CAA76841.1	albumi	( 585)	65	4.7
ALLERGEN_2012_2	gi	6687188	emb	CAB64867.1	albumi	( 608)	65	4.8
ALLERGEN_2012_2	gi	1351908	sp	P49064.1	ALBU_FELCA	( 608)	65	4.8
ALLERGEN_2012_2	gi	100834	pir	S16031	alpha-amyla	( 168)	59	5.9
ALLERGEN_2012_2	gi	21713	emb	CAA35597.1	unnamed	( 168)	59	5.9
ALLERGEN_2012_2	gi	2497701	sp	Q28133.1	ALL2_BOVIN	( 172)	58	7.1
ALLERGEN_2012_2	gi	244610	gb	AAB21323.1	tetramer	( 18)	49	7.7
ALLERGEN_2012_2	gi	335331566	gb	AEH31546.1	low m	( 369)	60	8.3
ALLERGEN_2012_2	gi	2506771	sp	P16968.2	IAA1_HORVU	( 146)	56	9
ALLERGEN_2012_2	gi	18641	emb	CAA37044.1	glycinin	( 562)	61	9.2
ALLERGEN_2012_2	gi	1170095	sp	P46419.1	GSTM1 DERP	( 219)	57	9.9

>>ALLERGEN\_2012\_2|gi|75219081|sp|O22108|O22108\_WHEAT LMM (285 aa)  
initn: 40 initl: 40 opt: 71 Z-score: 102.8 bits: 25.6 E(): 1  
Smith-Waterman score: 71; 36.364% identity (39.024% ungapped) in 44 aa overlap  
(45-86:8-50)

	20	30	40	50	60	70
SYHT0H	KLYCQMFERSGKFGDLELDSYLG	NWPNW	PSWQLKIPLPKTNIMPSTMLVSSCA--QCT			
				:	:	:
ALLERG			RCIPGLERPWQQQ-PLPPQQTFFPQQPLFSQQQQQQLF			
				10	20	30

	80	90	100	110
SYHT0H	PGCVSQSLMQPNRWIVWKAYNSNHRLKTLRLHRLKQMCVQCRS			
	:	:	:	:
ALLERG	PQQPSFSQQQPPFWQQQPPFSQQQPILPQQPPFSQQQQLVLPQQSPFSQQQQLILPPQQQ			
	40	50	60	70

>>ALLERGEN\_2012\_2|gi|21773|emb|CAA31685.1| unnamed prote (307 aa)  
initn: 40 initl: 40 opt: 71 Z-score: 102.4 bits: 25.6 E(): 1.1  
Smith-Waterman score: 71; 36.364% identity (39.024% ungapped) in 44 aa overlap  
(45-86:31-73)

	20	30	40	50	60	70
SYHT0H	KLYCQMFERSGKFGDLELDSYLG	NWPNW	PSWQLKIPLPKTNIMPSTMLVSSCA--QCT			
				:	:	:
ALLERG	MKTFLV FALLAVAATS AIAQMETRCIPGLERPWQQQ-PLPPQQTFFPQQPLFSQQQQQQLF					
	10	20	30	40	50	

	80	90	100	110
SYHT0H	PGCVSQSLMQPNRWIVWKAYNSNHRLKTLRLHRLKQMCVQCRS			
	:	:	:	:
ALLERG	PQQPSFSQQQPPFWQQQPPFSQQQPILPQQPPFSQQQQLVLPQQPPFSQQQPPVLPPQQS			
	60	70	80	90

>>ALLERGEN\_2012\_2|gi|886965|emb|CAA59339.1| low molecula (261 aa)  
initn: 44 initl: 44 opt: 67 Z-score: 97.8 bits: 24.6 E(): 2  
Smith-Waterman score: 67; 25.352% identity (29.508% ungapped) in 71 aa overlap  
(48-114:140-204)

```

      20      30      40      50      60      70
SYHT0H CQMFERSGKFGDLELD SYLGNWPNWPWRSWQLKIPLPKTNIM---PSTMLVSSCAQCTP
      : .:: . .:: : .::: .:::
ALLERG QQQPPFSQQQQQPILLQQPPFSQHQQPVLPQQQIPSVQPSILQQLNPKVFLQQ--QCSP
      110      120      130      140      150      160

```

```

      80      90      100      110
SYHT0H GCVSQSLMQPNRWIVWKAYNSNHRLKTLRLHRLKQMCVQCRS
      . ::: . . .::: : : . . .::: . : :
ALLERG VAMPQSLARSQ--MLWQS--SCHVMQQCCRQLPQIPEQSRYDAIRAIYSIVLQEQQHG
      170      180      190      200      210      220

```

```

ALLERG QGLNQPPQQPPQQSVQGVSPQQQQKQLGQCSFQQPQQ
      230      240      250      260

```

>>ALLERGEN\_2012\_2|gi|94400907|ref|NP\_001035360.1| allerg (92 aa)  
 initn: 49 initl: 49 opt: 61 Z-score: 95.0 bits: 22.5 E(): 2.8  
 Smith-Waterman score: 61; 40.000% identity (47.059% ungapped) in 20 aa overlap  
 (57-76:56-72)

```

      30      40      50      60      70      80
SYHT0H FGDLELD SYLGNWPNWPWRSWQLKIPLPKTNIMPSTMLVSSCAQCTPGCVSQSLMQPNRW
      : .::: . . .::: ::::
ALLERG GGFGLGGRGKCP SNEIFSRCDGRCQRF CPNVVPKPLCIKICA---PGCVCR LGYL RNKK
      30      40      50      60      70      80

```

```

      90      100      110
SYHT0H IVWKAYNSNHRLKTLRLHRLKQMCVQCRS

```

```

ALLERG KVCVPRSKCG
      90

```

>>ALLERGEN\_2012\_2|gi|633938|gb|AAB30434.1| albumin [Cani (265 aa)  
 initn: 65 initl: 65 opt: 65 Z-score: 95.0 bits: 24.1 E(): 2.8  
 Smith-Waterman score: 65; 47.059% identity (47.059% ungapped) in 17 aa overlap  
 (12-28:196-212)

```

      10      20      30      40
SYHT0H MIIIRPATGFNLKKLYCQMFERSGKFGDLELD SYLGNWPN
      : : : : .::: .:::
ALLERG TLEKCCATDDPPTCYAKVLDEFKPLVDEPQNLVKTNCELF EKLGEYGFQ NALLVRYTKKA
      170      180      190      200      210      220

```

```

      50      60      70      80      90      100
SYHT0H WPWRSWQLKIPLPKTNIMPSTMLVSSCAQCTPGCVSQSLMQPNRWIVWKAYNSNHRLKTL
ALLERG PQVSTPTLVVEVSRKLGKVGTKCCKKPESERMSCADDFLS
      230      240      250      260

```

>>ALLERGEN\_2012\_2|gi|88770352|gb|ABD51779.1| allergen Ap (94 aa)  
 initn: 49 initl: 49 opt: 61 Z-score: 94.9 bits: 22.5 E(): 2.8  
 Smith-Waterman score: 61; 40.000% identity (47.059% ungapped) in 20 aa overlap  
 (57-76:56-72)

```

      30      40      50      60      70      80
SYHT0H FGDLELD SYLGNWPNWPWRSWQLKIPLPKTNIMPSTMLVSSCAQCTPGCVSQSLMQPNRW
      : .::: . . .::: ::::
ALLERG GGFGLGGRGKCP SNEIFSRCDGRCQRF CPNVVPKPLCIKICA---PGCVCR LGYL RNKK
      30      40      50      60      70      80

```

90 100 110  
SYHT0H IVWKAYNSNHLKTLRLHRLKQMCVQCRS

ALLERG KVCVPRSKCLPG  
90

>>ALLERGEN\_2012\_2|gi|3021324|emb|CAA06305.1| Aspf1 aller (125 aa)  
initn: 44 initl: 44 opt: 61 Z-score: 93.5 bits: 22.7 E(): 3.4  
Smith-Waterman score: 61; 23.077% identity (23.438% ungapped) in 65 aa overlap  
(15-78:52-116)

10 20 30 40  
SYHT0H MIIARPATGFNLKKLYCQMFERSGKFGDLELDSYLGWPNWP-  
: : . : : : : : : : : :  
ALLERG KTGSSYPHWFTNGYDGDGKLIKGRMPIKFGKADCDRPPKHGKDGMGKDDHYLLEFPTFPD  
30 40 50 60 70 80  
50 60 70 80 90 100  
SYHT0H WRSWQLKIPLPKTNIMPSTMLVSSCAQCTPGCVSQSLMQPNRWIVWKAYNSNHLKTLRL  
..... : : . : . . . : : :  
ALLERG GHDKYFDSKKPKEDPGPARVIYTYPNKVFCGIVAHERNQGDLR  
90 100 110 120

110  
SYHT0H HRLKQMCVQCRS

>>ALLERGEN\_2012\_2|gi|886967|emb|CAA59340.1| low molecula (276 aa)  
initn: 44 initl: 44 opt: 64 Z-score: 93.5 bits: 23.8 E(): 3.4  
Smith-Waterman score: 68; 25.000% identity (29.577% ungapped) in 84 aa overlap  
(40-114:140-219)

10 20 30 40 50 60  
SYHT0H GFNLKKLYCQMFERSGKFGDLELDSYLGWPNWPWRSWQLKIPLPKTNI---MPSTML-V  
: : : . . : : : : :  
ALLERG QRPPFSQQQQQPVLPPQPPFSQQQQQPILPQQPPFSLHQQPVLPQQQIPYVQPSILQQL  
110 120 130 140 150 160  
70 80 90 100 110  
SYHT0H SSCA-----QCTPGCVSQSLMQPNRWIVWKAYNSNHLKTLRLHRLKQMCVQCRS  
. : : : : . : : . . : : : : : : : : : :  
ALLERG NPCKVFLQQQCSPVAMPQSLARSQ--MLWQS--SCHVMQQQCCQQLPRIPEQSRIDAIRA  
170 180 190 200 210 220  
ALLERG IIYSIVLQEQQHGGQFNQPQQQQPQQSVQGVSPQQQQKQLGQCSFQQPQQ  
230 240 250 260 270

>>ALLERGEN\_2012\_2|gi|6136162|sp|P35776.2|VA2\_SOLRI RecNa (119 aa)  
initn: 60 initl: 60 opt: 60 Z-score: 92.4 bits: 22.4 E(): 3.9  
Smith-Waterman score: 60; 36.000% identity (36.000% ungapped) in 25 aa overlap  
(65-89:12-36)

40 50 60 70 80 90  
SYHT0H YLGWPNWPWRSWQLKIPLPKTNIMPSTMLVSSCAQCTPGCVSQSLMQPNRWIVWKAYNS  
.. : : : : : : : : : :  
ALLERG DIEAQRVLRKDIAECARTLPKCVNQPDPLARVDVWHCAMS  
10 20 30 40  
100 110  
SYHT0H NHRLKTLRLHRLKQMCVQCRS  
ALLERG KRGVYDNPDPVAVVEKNSKMCPKIITDPADVENCKKVSRCVDRETQRPNSNRQKAINIT  
50 60 70 80 90 100

>>ALLERGEN\_2012\_2|gi|168419914|gb|ACA23876.1| Pas n 1 al (265 aa)  
 initn: 27 initl: 27 opt: 63 Z-score: 92.3 bits: 23.6 E(): 4  
 Smith-Waterman score: 63; 24.638% identity (27.869% ungapped) in 69 aa overlap  
 (30-95:199-262)

```

              10      20      30      40      50
SYHT0H  MIIIIARPATGFNLKKLYCQMFERSGKFGDLELDSYLGWPNWPNW--SWQ--LKIPLPKT
              .::  ..:  :.  ::  ..  ::.
ALLERG  KIVFHVEKGSNPNYLAMLVKFVADDGDIVLMELKEKSSDWK--PMKLSWGAIWRMDTPKA
      170      180      190      200      210      220

              60      70      80      90      100      110
SYHT0H  NIMPSTMLVSSCAQCTPGCVSQSLMQPNRWIVWKAYNSNHLKTLRLHRLKQMCVQCRS
      . : .. .: .  ..... : :  .....
ALLERG  LVPPFSIRLTS--ESGKKVIAQDVI-PVNWKPDVTYNSNVQF
      230      240      250      260

```

>>ALLERGEN\_2012\_2|gi|3319897|emb|CAA76841.1| albumin [Ca (585 aa)  
 initn: 65 initl: 65 opt: 65 Z-score: 90.9 bits: 24.4 E(): 4.7  
 Smith-Waterman score: 65; 47.059% identity (47.059% ungapped) in 17 aa overlap  
 (12-28:387-403)

```

              10      20      30      40
SYHT0H                      MIIIIARPATGFNLKKLYCQMFERSGKFGDLELDSYLGWPN
              :: :  :::::  :::
ALLERG  TLEKCCATDDPPTCYAKVLDEFKPLVDEPQNLVKTNCLEFEKLGEYGFQNALLVRYTKKA
      360      370      380      390      400      410

              50      60      70      80      90      100
SYHT0H  WPWRSQLKIPLPKTNIMPSTMLVSSCAQCTPGCVSQSLMQPNRWIVWKAYNSNHLKTL
ALLERG  PQVSTPTLVEVSRKLGKVGTKCCKKPESERMSCADDFLSVVLNRLCVLHEKTPVSEKVTK
      420      430      440      450      460      470

```

>>ALLERGEN\_2012\_2|gi|6687188|emb|CAB64867.1| albumin [Ca (608 aa)  
 initn: 65 initl: 65 opt: 65 Z-score: 90.7 bits: 24.5 E(): 4.8  
 Smith-Waterman score: 65; 47.059% identity (47.059% ungapped) in 17 aa overlap  
 (12-28:410-426)

```

              10      20      30      40
SYHT0H                      MIIIIARPATGFNLKKLYCQMFERSGKFGDLELDSYLGWPN
              :: :  :::::  :::
ALLERG  TLEKCCATDDPPTCYAKVLDEFKPLVDEPQNLVKTNCLEFEKLGEYGFQNALLVRYTKKA
      380      390      400      410      420      430

              50      60      70      80      90      100
SYHT0H  WPWRSQLKIPLPKTNIMPSTMLVSSCAQCTPGCVSQSLMQPNRWIVWKAYNSNHLKTL
ALLERG  PQVSTPTLVEVSRKLGKVGTKCCKKPESERMSCAEDFLSVVLNRLCVLHEKTPVSEKVTK
      440      450      460      470      480      490

```

>>ALLERGEN\_2012\_2|gi|1351908|sp|P49064.1|ALBU\_FELCA RecN (608 aa)  
 initn: 65 initl: 65 opt: 65 Z-score: 90.7 bits: 24.5 E(): 4.8  
 Smith-Waterman score: 65; 47.059% identity (47.059% ungapped) in 17 aa overlap  
 (12-28:410-426)

```

              10      20      30      40
SYHT0H                      MIIIIARPATGFNLKKLYCQMFERSGKFGDLELDSYLGWPN
              :: :  :::::  :::
ALLERG  TLEKCCATDDPPACYAHVFDEFKPLVEEPNLVKTNCLEFEKLGEYGFQNALLVRYTKKV
      380      390      400      410      420      430

```

```

          50          60          70          80          90         100
SYHT0H WPWRSWQLKIPLPKTNIMPSTMLVSSCAQCTPGCVSQSLMQPNRWIVWKAYNSNHRLKTL

ALLERG PQVSTPTLVEVSRSLGKVGSKCCTHPEAERLSAEDYLSVVLNRLCVLHEKTPVSESVTK
      440      450      460      470      480      490

```

```

>>ALLERGEN_2012_2|gi|100834|pir||S16031 alpha-amylase in (168 aa)
  initn: 38 initl: 38 opt: 59 Z-score: 89.2 bits: 22.3 E(): 5.9
Smith-Waterman score: 59; 24.286% identity (31.481% ungapped) in 70 aa overlap
(55-115:36-98)

```

```

          30          40          50          60          70
SYHT0H GKFGDLELDSYLGWPNWPWRSWQLKIPLPKTNIMPST---MLVSSCAQCTPGC-----
                                     .:.:.:. .: .:. .:.
ALLERG SCSLLLLAAVLLSVLAAASASGSCVPGVAFRTNLLPHCRDYVLQQTCTFTPGSKLPEWM
      10      20      30      40      50      60

          80          90         100         110
SYHT0H VSQSLMQPNRWIVWKAYNSNHRLKTLRLHRLKQMCVQCRS
      .: .:.:.:. .: .: .: .: .:.:.:. .:.
ALLERG TSASIYSPGKPYLAKLYCCQE-----LAEISQQC-RCEALRYFIALPVPSQPVDPRSGN
      70      80      90      100      110

```

```

ALLERG VGESGLIDLPGCPREMOWDFVRLLVAPGQCENLATIHNVRYPYCPAVEQPLWI
      120      130      140      150      160

```

```

>>ALLERGEN_2012_2|gi|21713|emb|CAA35597.1| unnamed prote (168 aa)
  initn: 38 initl: 38 opt: 59 Z-score: 89.2 bits: 22.3 E(): 5.9
Smith-Waterman score: 59; 24.286% identity (31.481% ungapped) in 70 aa overlap
(55-115:36-98)

```

```

          30          40          50          60          70
SYHT0H GKFGDLELDSYLGWPNWPWRSWQLKIPLPKTNIMPST---MLVSSCAQCTPGC-----
                                     .:.:.:. .: .:. .:.
ALLERG SCSLLLLAAVLLSVLAAASASGSCVPGVAFRTNLLPHCRDYVLQQTCTFTPGSKLPEWM
      10      20      30      40      50      60

          80          90         100         110
SYHT0H VSQSLMQPNRWIVWKAYNSNHRLKTLRLHRLKQMCVQCRS
      .: .:.:.:. .: .: .: .: .:.:.:. .:.
ALLERG TSASIYSPGKPYLAKLYCCQE-----LAEISQQC-RCEALRYFIALPVPSQPVDPRSGN
      70      80      90      100      110

```

```

ALLERG VGESGLIDLPGCPREMOWDFVRLLVAPGQCENLATIHNVRYPYCPAVEQPLWI
      120      130      140      150      160

```

```

>>ALLERGEN_2012_2|gi|2497701|sp|Q28133.1|ALL2_BOVIN RecN (172 aa)
  initn: 29 initl: 29 opt: 58 Z-score: 87.8 bits: 22.1 E(): 7.1
Smith-Waterman score: 58; 27.778% identity (31.915% ungapped) in 54 aa overlap
(68-115:14-66)

```

```

          40          50          60          70          80          90
SYHT0H NWPWPNWRSWQLKIPLPKTNIMPSTMLVSSC-AQCTPGCVSQSLMQPNRW-IVWKAYNSN
                                     : : : . .: .: .: .: .: .:
ALLERG                                     MKAVFLTLLFLVCTAQETPAEIDPSKI-PGEWRRIYAAADNK
                                     10      20      30      40

          100         110
SYHT0H HRL---KTLRLHRLKQMCVQ-CRS
      ..      : : . .: .: .:
ALLERG DKIVEGGPLRNYRRIECINDCESLSITFYLDQGTCLLLTEVAKRQEGYVYVLEFYGTN
      50      60      70      80      90      100

```

>>ALLERGEN\_2012\_2|gi|244610|gb|AAB21323.1| tetrameric al (18 aa)  
 initn: 34 initl: 34 opt: 49 Z-score: 87.2 bits: 18.7 E(): 7.7  
 Smith-Waterman score: 49; 47.368% identity (50.000% ungapped) in 19 aa overlap  
 (36-54:1-18)

```

      10      20      30      40      50      60
SYHT0H RPATGFNLKKLYCQMFERSGKFGDLELDSYLGWPNWPNWRSWQLKIPLPKTNIMPSTMLV
      .::  ::  :  :  ::
ALLERG          IGNEDCTPWMS-TLITPLP
                        10

```

```

      70      80      90     100     110
SYHT0H SSACQCTPGCVSQSLMQPNRWIVWKAYNSNHRLKTLRLHRLKQMCVQCRS

```

>>ALLERGEN\_2012\_2|gi|335331566|gb|AEH31546.1| low molecu (369 aa)  
 initn: 44 initl: 44 opt: 60 Z-score: 86.5 bits: 23.0 E(): 8.3  
 Smith-Waterman score: 63; 24.051% identity (32.759% ungapped) in 79 aa overlap  
 (43-112:183-249)

```

      20      30      40      50      60
SYHT0H LKKLYCQMFERSGKFGDLELDSYLGWPNWPNWRSWQLKIPLPKTNI---MPSTML-VSSC
      . . : . : . : . : . : . :
ALLERG QSPFSQQQIVLQQQPPFLQQQPPSLPQQPPFSQQQQQLVLPQQQIPFVHPSILQQLNPC
      160      170      180      190      200      210

```

```

      70      80      90     100     110
SYHT0H A-----QCTPGCVSQSLMQPNRWIVWKAYNSNHRLKTLRLHRLKQMCVQCRS
      ::: . ::: . . . . : . : . :
ALLERG KVFLQQQCSPVAMPQSLAR-----SQMLQQSSCHVMQQQCCQQLPQIPQQSRY
      220      230                        240      250      260

```

```

ALLERG EAIRAIYSIILQEQQQVQGSIQTPQQQPQQLGQCVSQPQQQSQQQLGQQPQQQQLAQGT
      270      280      290      300      310      320

```

>>ALLERGEN\_2012\_2|gi|2506771|sp|P16968.2|IAA1\_HORVU RecN (146 aa)  
 initn: 37 initl: 37 opt: 56 Z-score: 85.9 bits: 21.5 E(): 9  
 Smith-Waterman score: 57; 30.435% identity (37.838% ungapped) in 46 aa overlap  
 (37-82:17-53)

```

      10      20      30      40      50      60
SYHT0H PATGFNLKKLYCQMFERSGKFGDLELDSYLGWPNWPNWRSWQLKIPLPKTNIMPSTMLVS
      .::  ::  . . . : . : . :
ALLERG          PTSVAVDQGSMSVNSPGEW-CWPGMGYPV-YPFPRCRA-----LVK
                        10      20      30

```

```

      70      80      90     100     110
SYHT0H SCAQCTPGCVSQSLMQPNRWIVWKAYNSNHRLKTLRLHRLKQMCVQCRS
      :  ::  :  :  : . : . .
ALLERG S--QCAGGQVVESIQKDCCRQIAAIGDEWCICGALGSMRGS MYKELGVALADDKATVAEV
      40      50      60      70      80      90

```

>>ALLERGEN\_2012\_2|gi|18641|emb|CAA37044.1| glycinin [Gly (562 aa)  
 initn: 50 initl: 50 opt: 61 Z-score: 85.7 bits: 23.4 E(): 9.2  
 Smith-Waterman score: 61; 26.471% identity (30.000% ungapped) in 34 aa overlap  
 (53-83:82-114)

```

      30      40      50      60      70
SYHT0H RSGKFGDLELDSYLGWPNWPNWRSWQLKIPLPKTNIMPSTMLVSS---CAQCTPGCVSQS
      . . . : . . . . : . : . :
ALLERG LIQTWNSQHPCLKCAGVTVSKLTLNRNGLHLPSYSPYPRMIIIAQKGKALQCKPGC-PET
      60      70      80      90     100     110

```

```

      80      90      100      110
SYHT0H LMQPNRWIVWKAYNSNHRLKTLRLHRLKQMCVQCRS
      . . .
ALLERG FEFPQEQSNRRGSRSQKQQLQDSHQKIRHFNEGDVLVIPPVGYWYNTGTDEPVVAISLL
      120      130      140      150      160      170

>>ALLERGEN_2012_2|gi|1170095|sp|P46419.1|GSTM1_DERPT Rec (219 aa)
  initn: 32 initl: 32 opt: 57 Z-score: 85.2 bits: 22.0 E(): 9.9
Smith-Waterman score: 57; 29.508% identity (30.508% ungapped) in 61 aa overlap
(11-69:159-219)

```

```

      10      20      30
SYHT0H MIIIPATGFNLKKLYCQMFERSGK-FGDLE-LDSYLG
      : : : . : : . . : : : : : :
ALLERG LKSLPDCLKLMSKFVGEHAFIAGANISYVDFNLYEYLCHVKVMVPEVFGQFENLKRYVER
      130      140      150      160      170      180

      40      50      60      70      80      90
SYHT0H WPNWPWRSWQLKIPLPKTNIMPSTMLVSSCAQCTPGCVSQSLMQPNRWIVWKAYNSNHRL
      . : : . : : : : : : : : :
ALLERG MESLPRVSDYIKKQPKTFNAPT SKWNASYA
      190      200      210

```

```

      100      110
SYHT0H KTLRLHRLKQMCVQCRS

```

```

115 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.040 Display time: 0.010

```

```

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 31 aa
>SYHT0H2_I_2
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

```

```

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 4.1921+/-0.00334; mu= 1.0916+/- 0.174
mean_var=39.8651+/- 8.917, 0's: 22 Z-trim: 22 B-trim: 0 in 0/41
Lambda= 0.2031

```

```

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.010

```

```

The best scores are:
                                opt bits E(1603)
ALLERGEN_2012_2|gi|23894244|emb|CAD23614.1| tri m ( 404) 51 22 5.9
ALLERGEN_2012_2|gi|74663809|sp|Q8J077.1|SUB6_TRIS ( 405) 51 22 5.9
ALLERGEN_2012_2|gi|6136163|sp|P35779.2|VA3_SOLRI ( 211) 48 21 6.2
ALLERGEN_2012_2|gi|291482306|emb|CBK62693.1| ragw ( 111) 45 20 6.6
ALLERGEN_2012_2|gi|14424466|sp|P35778.2|VA3_SOLIN ( 234) 48 21 6.8
ALLERGEN_2012_2|gi|291482308|emb|CBK62694.1| ragw ( 140) 45 20 8.1
ALLERGEN_2012_2|gi|285005079|emb|CBJ24286.1| ragw ( 164) 45 20 9.3
ALLERGEN_2012_2|gi|313118253|sp|POCH89.1|HUGA_VES ( 31) 38 17 9.3
ALLERGEN_2012_2|gi|11514279|pdb|1QNX|A Chain A, V ( 209) 46 20 9.3

```



ALLERGEN\_2012\_2|gi|162551|gb|AAA30333.1| allergen ( 227) 46 20 10

>>ALLERGEN\_2012\_2|gi|23894244|emb|CAD23614.1| tri m 2 al (404 aa)  
 initn: 35 initl: 35 opt: 51 Z-score: 89.2 bits: 21.7 E(): 5.9  
 Smith-Waterman score: 51; 30.000% identity (32.143% ungapped) in 30 aa overlap (1-28:39-68)

			10		20
SYHT0H			MVNDITNRGHISCCSW--PNLNWSPLRRHG		
			: : : : : . . : : : : : . : : :		
ALLERG	VLAALSAVNGAKILEAGPHAETIPNKYIVVMKKDVSDEAFSTHTTWLSQNLNRRRLMRRSG				
	10 20 30 40 50 60				

	30
SYHT0H	GIL
	.
ALLERG	SSKAMAGMQNKYSLGGIFRAYSGEFDDAMIKDISNHDDVDYIEPDFVVRTSTNGTNLTRQ
	70 80 90 100 110 120

>>ALLERGEN\_2012\_2|gi|74663809|sp|Q8J077.1|SUB6\_TRISH Rec (405 aa)  
 initn: 35 initl: 35 opt: 51 Z-score: 89.2 bits: 21.7 E(): 5.9  
 Smith-Waterman score: 51; 30.000% identity (32.143% ungapped) in 30 aa overlap (1-28:39-68)

			10		20
SYHT0H			MVNDITNRGHISCCSW--PNLNWSPLRRHG		
			: : : : : . . : : : : : . : : :		
ALLERG	VLAALSAVNGAKILEAGPHAETIPNKYIVVMKKDVSDEAFSTHTTWLSQNLNRRRLMRRSG				
	10 20 30 40 50 60				

	30
SYHT0H	GIL
	.
ALLERG	SSKAMAGMQNKYSLGGIFRAYSGEFDDAMIKDISNHDDVDYIEPDFVVRTSTNGTNLTRQ
	70 80 90 100 110 120

>>ALLERGEN\_2012\_2|gi|6136163|sp|P35779.2|VA3\_SOLRI RecNa (211 aa)  
 initn: 48 initl: 48 opt: 48 Z-score: 88.8 bits: 20.7 E(): 6.2  
 Smith-Waterman score: 48; 38.889% identity (38.889% ungapped) in 18 aa overlap (6-23:66-83)

			10		20		30
SYHT0H			MVNDITNRGHISCCSWPNLNWSPLRRHGGIL				
			: : . . : : : : :				
ALLERG	VGFTDAEKDAIVNKHNLQRQVASGKEMRGTNQPPPAVKMPNLTWDPELATIAQRWANQ						
	40 50 60 70 80 90						

ALLERG	CTFEHDACRNVERFAVGQNIATSSSGKNKSTLSDMILLWYNEVKDFDNRWISSFPDGN
	100 110 120 130 140 150

>>ALLERGEN\_2012\_2|gi|291482306|emb|CBK62693.1| ragweed h (111 aa)  
 initn: 45 initl: 45 opt: 45 Z-score: 88.3 bits: 19.7 E(): 6.6  
 Smith-Waterman score: 45; 55.556% identity (55.556% ungapped) in 9 aa overlap (14-22:23-31)

		10		20		30
SYHT0H		MVNDITNRGHISCCSWPNLNWSPLRRHGGIL				
		: : : : :				
ALLERG	MAAGLLVFLAISEIASVKGLCEKPSLTWSGKCKVKQTDKCDKRCIEWEGAKHGACHKR					
	10 20 30 40 50 60					

ALLERG	DSKATCFCYFDCDPTKNPGPPGAPKGKAPAPSPPSGGGAPPPSGGEGGER
	70 80 90 100 110

>>ALLERGEN\_2012\_2|gi|14424466|sp|P35778.2|VA3\_SOLIN RecN (234 aa)  
 initn: 48 initl: 48 opt: 48 Z-score: 88.1 bits: 20.7 E(): 6.8  
 Smith-Waterman score: 48; 38.889% identity (38.889% ungapped) in 18 aa overlap (6-23:88-105)

```

                                10      20      30
SYHT0H                      MVNDITNRGHISCCSWPNLNWSPLRRHGGIL
                                :: . . . :::::
ALLERG VGFTDAEKDAIVNKHNELRQRVASGKEMRGTNQPQPPAVKMPNLTWDPPEATIAQRWANQ
      60      70      80      90      100     110

ALLERG CTFEHDACRNVERFAVGQNIATSSSGKNKSTPNEMILLWYNEVKDFDNRWISSFPSSDDN
      120     130     140     150     160     170

```

>>ALLERGEN\_2012\_2|gi|291482308|emb|CBK62694.1| ragweed h (140 aa)  
 initn: 45 initl: 45 opt: 45 Z-score: 86.7 bits: 19.7 E(): 8.1  
 Smith-Waterman score: 45; 55.556% identity (55.556% ungapped) in 9 aa overlap (14-22:3-11)

```

                                10      20      30
SYHT0H MVNDITNRGHISCCSWPNLNWSPLRRHGGIL
                                : ::::
ALLERG          KLCEKPSLTWSGKCKVKQTDKCDKRCIEWEGAKHGACHKRDSKATCFY
                                10      20      30      40

```

>>ALLERGEN\_2012\_2|gi|285005079|emb|CBJ24286.1| ragweed h (164 aa)  
 initn: 45 initl: 45 opt: 45 Z-score: 85.7 bits: 19.7 E(): 9.3  
 Smith-Waterman score: 45; 55.556% identity (55.556% ungapped) in 9 aa overlap (14-22:23-31)

```

                                10      20      30
SYHT0H          MVNDITNRGHISCCSWPNLNWSPLRRHGGIL
                                : ::::
ALLERG MAAGLLVFVLAFSEIASVKGKLCEKPSLTWSGKCKVKQTDKCDKRCIEWEGAKHGACHKR
      10      20      30      40      50      60

ALLERG DSKATCFYFDCDPTKNPGPPGAPKGAAPAPSPPSGGGAPPPSGGEGGDGPPPPPEGEG
      70      80      90      100     110     120

```

>>ALLERGEN\_2012\_2|gi|313118253|sp|P0CH89.1|HUGA\_VESMC Re (31 aa)  
 initn: 28 initl: 28 opt: 38 Z-score: 85.7 bits: 17.3 E(): 9.3  
 Smith-Waterman score: 38; 44.444% identity (44.444% ungapped) in 9 aa overlap (14-22:3-11)

```

                                10      20      30
SYHT0H MVNDITNRGHISCCSWPNLNWSPLRRHGGIL
                                : ::. .:
ALLERG          DRCIWPKEGFSIYWNIPTHFCHNFGVYFKEL
                                10      20      30

```

>>ALLERGEN\_2012\_2|gi|11514279|pdb|1QNX|A Chain A, Ves V (209 aa)  
 initn: 32 initl: 32 opt: 46 Z-score: 85.7 bits: 20.1 E(): 9.3  
 Smith-Waterman score: 46; 37.500% identity (37.500% ungapped) in 16 aa overlap (3-18:154-169)

```

                                10      20      30
SYHT0H                      MVNDITNRGHISCCSWPNLNWSPLRRHGGIL
                                ::. . ::. .:
ALLERG TAAKYDDPVKLVKMWEDVKDYNPKKKFSGNDFLKTGHYTMVWANTKEVGCGSIKIYQE
      130     140     150     160     170     180

```

ALLERG KWHKHLYVCNYGPSGNFKNEELYQTK

190 200

>>ALLERGEN\_2012\_2|gi|162551|gb|AAA30333.1| allergen 5 [V (227 aa)  
initn: 32 initl: 32 opt: 46 Z-score: 85.1 bits: 20.1 E(): 10  
Smith-Waterman score: 46; 37.500% identity (37.500% ungapped) in 16 aa overlap (3-  
18:172-187)

```

                                10      20      30
SYHTOH                        MVNDITNRGHISCCSWPNLNWSPLRRHGGIL
                                ::. . ::. . : :
ALLERG TAAKYDDPVKLVKMWEDEVKDYNPKKKFSGNDFLKTGHYTQMVWANTKEVGCGSIKIYIQE
                                150      160      170      180      190      200

ALLERG KWHKHLYLCNYGPSGNFMNEELYQTK
                                210      220
```

31 residues in 1 query sequences  
367796 residues in 1603 library sequences  
Scomplib [34t11]  
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012  
Scan time: 0.010 Display time: 0.000

Function used was FASTA [version 3.4t11 Apr 17, 2002]  
FASTA searches a protein or DNA sequence data bank  
version 3.4t11 Apr 17, 2002  
Please cite:  
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

/grid/.qsub\_tmp/1330373600kde\_agent/grid\_NbCuMBoIO: 31 aa  
>SYHT0H2\_I\_3  
vs /data/fasta/ALLERGEN\_2012\_2 library  
searching /data/fasta/ALLERGEN\_2012\_2 library

367796 residues in 1603 sequences  
Expectation\_n fit: rho(ln(x))= 2.3220+/-0.00343; mu= 14.9746+/- 0.181  
mean\_var=42.6403+/-11.992, 0's: 18 Z-trim: 18 B-trim: 0 in 0/42  
Lambda= 0.1964

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2  
join: 36, opt: 24, gap-pen: -12/-2, width: 16  
Scan time: 0.010

The best scores are:

						opt	bits	E(1603)
ALLERGEN_2012_2	gi 291482314 emb CBK62697.1	ragw	( 96)	57	21	1.9		
ALLERGEN_2012_2	gi 291482306 emb CBK62693.1	ragw	( 111)	57	21	2		
ALLERGEN_2012_2	gi 291482310 emb CBK62695.1	ragw	( 134)	57	22	2.2		
ALLERGEN_2012_2	gi 291482308 emb CBK62694.1	ragw	( 140)	57	22	2.2		
ALLERGEN_2012_2	gi 285005079 emb CBJ24286.1	ragw	( 164)	57	22	2.4		
ALLERGEN_2012_2	gi 291197394 emb CBK52317.1	ragw	( 164)	57	22	2.4		
ALLERGEN_2012_2	gi 14422361 emb CAC41634.1	plant	( 131)	50	20	8.6		
ALLERGEN_2012_2	gi 14422359 emb CAC41633.1	plant	( 131)	50	20	8.6		
ALLERGEN_2012_2	gi 27818335 gb AAO24900.1	major	( 132)	50	20	8.6		

>>ALLERGEN\_2012\_2|gi|291482314|emb|CBK62697.1| ragweed h (96 aa)  
initn: 74 initl: 48 opt: 57 Z-score: 98.1 bits: 21.3 E(): 1.9  
Smith-Waterman score: 57; 69.231% identity (75.000% ungapped) in 13 aa overlap (2-  
13:74-86)

```

                                10      20      30
SYHT0H                      MPSHGG-PPPVGHESYLFVRKSQTDCTVFPS
                                :: :: ::: : :.
ALLERG SCFCYFDCDPTKNPGPPPGAPKGKAPAPSPPSGGGAPPPSGGEGGGDGGGGRR
                                50      60      70      80      90

SYHT0H W

>>ALLERGEN_2012_2|gi|291482306|emb|CBK62693.1| ragweed h (111 aa)
  initn: 74 initl: 48 opt: 57 Z-score: 97.6 bits: 21.4 E(): 2
Smith-Waterman score: 57; 69.231% identity (75.000% ungapped) in 13 aa overlap (2-
13:95-107)

                                10      20      30
SYHT0H                      MPSHGG-PPPVGHESYLFVRKSQTDCTVFPS
                                :: :: ::: : :.
ALLERG TCFCYFDCDPTKNPGPPPGAPKGKAPAPSPPSGGGAPPPSGGEGGER
                                70      80      90      100     110

SYHT0H W

>>ALLERGEN_2012_2|gi|291482310|emb|CBK62695.1| ragweed h (134 aa)
  initn: 74 initl: 48 opt: 57 Z-score: 96.9 bits: 21.5 E(): 2.2
Smith-Waterman score: 57; 69.231% identity (75.000% ungapped) in 13 aa overlap (2-
13:75-87)

                                10      20      30
SYHT0H                      MPSHGG-PPPVGHESYLFVRKSQTDCTVFPS
                                :: :: ::: : :.
ALLERG SCFCYFDCDPTKNPGPPPGAPKGKAPAPSPPSGGGAPPPSGGEGGGDGGGEGGGGGGGG
                                50      60      70      80      90      100

SYHT0H W

ALLERG EGGGEGGGGGEGGGEGGGGGEGGGEGGGER
                                110     120     130

>>ALLERGEN_2012_2|gi|291482308|emb|CBK62694.1| ragweed h (140 aa)
  initn: 99 initl: 48 opt: 57 Z-score: 96.8 bits: 21.6 E(): 2.2
Smith-Waterman score: 57; 69.231% identity (75.000% ungapped) in 13 aa overlap (2-
13:75-87)

                                10      20      30
SYHT0H                      MPSHGG-PPPVGHESYLFVRKSQTDCTVFPS
                                :: :: ::: : :.
ALLERG TCFCYFDCDPTKNPGPPPGAPKGKAPAPSPPSGGGAPPPSGGEGGDGPPPEGGEGGGDG
                                50      60      70      80      90      100

SYHT0H W

ALLERG GEGGGGDGGGEGGGEGGGDGGGDGGGEGGGEGGGER
                                110     120     130     140

>>ALLERGEN_2012_2|gi|285005079|emb|CBJ24286.1| ragweed h (164 aa)
  initn: 99 initl: 48 opt: 57 Z-score: 96.2 bits: 21.7 E(): 2.4
Smith-Waterman score: 57; 69.231% identity (75.000% ungapped) in 13 aa overlap (2-
13:95-107)

```

```

                                10      20      30
SYHT0H                      MP SHGG-PPPVGHESYLFVRKSQTDCTVFPS
                                :: :: ::: : :.
ALLERG TCFCYFDCDPTKNPGPPPGAPKGKAPAPSPPSGGGAPPPSGGEGGDGPPPPPEGGEGGGDG
        70      80      90      100     110     120

```

SYHT0H W

```

ALLERG GGE GGGEGGGDGGGEGGGEGGGDGGGEGGGEGGGER
        130     140     150     160

```

>>ALLERGEN\_2012\_2|gi|291197394|emb|CBK52317.1| ragweed h (164 aa)  
 initn: 99 initl: 48 opt: 57 Z-score: 96.2 bits: 21.7 E(): 2.4  
 Smith-Waterman score: 57; 69.231% identity (75.000% ungapped) in 13 aa overlap (2-13:95-107)

```

                                10      20      30
SYHT0H                      MP SHGG-PPPVGHESYLFVRKSQTDCTVFPS
                                :: :: ::: : :.
ALLERG TCFCYFDCDPTKNPGPPPGAPKGKAPAPSPPSGGGAPPPSGGEGGDGPPPPPEGGEGGGDG
        70      80      90      100     110     120

```

SYHT0H W

```

ALLERG GGE GGGEGGGDGGGEGGGEGGGDGGGEGGGEGGGER
        130     140     150     160

```

>>ALLERGEN\_2012\_2|gi|14422361|emb|CAC41634.1| plantain p (131 aa)  
 initn: 30 initl: 30 opt: 50 Z-score: 86.3 bits: 19.5 E(): 8.6  
 Smith-Waterman score: 50; 40.000% identity (45.455% ungapped) in 25 aa overlap (8-29:66-90)

```

                                10      20      30
SYHT0H                      MP SHGGPPPVG-HES--YLFVRKSQTDCTVFPSW
                                : :: ::. .... ::. ...
ALLERG AQVQLDCKDDSKKVIYSIGGETGQDGVYRLPVVGYHEDCEIKLVKSSRPDCSEIPKLAKG
        40      50      60      70      80      90

```

```

ALLERG TIQTSKVDLSKNTTITEKTRHVKPLSFRAKTDAPGC
        100     110     120     130

```

>>ALLERGEN\_2012\_2|gi|14422359|emb|CAC41633.1| plantain p (131 aa)  
 initn: 30 initl: 30 opt: 50 Z-score: 86.3 bits: 19.5 E(): 8.6  
 Smith-Waterman score: 50; 40.000% identity (45.455% ungapped) in 25 aa overlap (8-29:66-90)

```

                                10      20      30
SYHT0H                      MP SHGGPPPVG-HES--YLFVRKSQTDCTVFPSW
                                : :: ::. .... ::. ...
ALLERG AQVQLDCKDDSKKVIYSIGGETDQDGVYRLPVVGYHEDCEIKLVKSSRPDCSEIPKLAKG
        40      50      60      70      80      90

```

```

ALLERG TIQTSKVDLSKNTTITEKTRHVKPLSFRAKTDAPGC
        100     110     120     130

```

>>ALLERGEN\_2012\_2|gi|27818335|gb|AA024900.1| major polle (132 aa)  
 initn: 72 initl: 50 opt: 50 Z-score: 86.3 bits: 19.5 E(): 8.6  
 Smith-Waterman score: 50; 66.667% identity (66.667% ungapped) in 9 aa overlap (2-10:113-121)

```

              10      20      30
SYHT0H      MPSHGGPPPVGHESYLFVRKSQTDCTVFPWS
              :. :: :::
ALLERG GATPAPPGAAPPPAAGGSPSPADGGSPPPADGGSPFVDGGSPPPSTH
              90      100      110      120      130

```

```

31 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.010 Display time: 0.010

```

```

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 121 aa
>SYHT0H2_I_4
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

```

```

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 4.0029+/-0.00408; mu= 13.2344+/- 0.212
mean_var=74.1704+/-20.031, 0's: 2 Z-trim: 2 B-trim: 0 in 0/44
Lambda= 0.1489

```

```

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.040

```

```

The best scores are:
                                opt bits E(1603)
ALLERGEN_2012_2|gi|7638030|gb|AAF65313.1|AF230384 ( 137) 73 23 2.3
ALLERGEN_2012_2|gi|7638028|gb|AAF65312.1|AF230383 ( 137) 73 23 2.3
ALLERGEN_2012_2|gi|4038411|gb|AAC97370.1| venom a ( 137) 71 23 3.1
ALLERGEN_2012_2|gi|14424465|sp|P35777.2|VA4_SOLIN ( 137) 67 22 5.6
ALLERGEN_2012_2|gi|291092710|gb|ADD74392.1| Sol s ( 137) 67 22 5.6

```

```

>>ALLERGEN_2012_2|gi|7638030|gb|AAF65313.1|AF230384_1 ve (137 aa)
initn: 46 init1: 46 opt: 73 Z-score: 96.5 bits: 23.5 E(): 2.3
Smith-Waterman score: 73; 30.435% identity (34.146% ungapped) in 46 aa overlap
(74-119:68-108)

```

```

              50      60      70      80      90      100
SYHT0H GLGGAGNSKSLCQKPTSCHPPCLYPARNVPRAVYPKASCNLTDGSFGRPITATDLKPC
              :..: : : : : : :
ALLERG TVPKGENDPINPLKNVNVLYCAFSKRGIFTPKGVNTKQYINYCEKTIINP---ADIKQC
              40      50      60      70      80      90

```

```

              110      120
SYHT0H ASIDLKCVYNVDPRPNL
              .. :... : :
ALLERG KKL-ISKCIKKVYDRPGPIIERSKNLLSCVRKKGVLTLTVYGKKK
              100      110      120      130

```

```

>>ALLERGEN_2012_2|gi|7638028|gb|AAF65312.1|AF230383_1 ve (137 aa)
initn: 46 init1: 46 opt: 73 Z-score: 96.5 bits: 23.5 E(): 2.3
Smith-Waterman score: 73; 30.435% identity (34.146% ungapped) in 46 aa overlap
(74-119:68-108)

```

```

      50      60      70      80      90     100
SYHT0H GLGGAGNSKSLCQKPTSCHPPCLYPAARNVPRAVYPKASCNLTDGSFGRPITATTDLKPC
      :..: : : . . : . : : :
ALLERG TVPKGENDPINPLKNVNVLYCAFSKRGIFTPKGVNTKQYINYCEKTIINP----ADIKQC
      40      50      60      70      80      90

      110     120
SYHT0H ASIDLSKCVYNVDPRPNL
      .. : : : . : : :
ALLERG KKL-ISKCIKKVYDRPGPIIERSKNLLSCVLKKGVLLELTVYGKKK
      100     110     120     130

>>ALLERGEN_2012_2|gi|4038411|gb|AAC97370.1| venom allerg (137 aa)
  initn: 45 init1: 45 opt: 71 Z-score: 94.2 bits: 23.0 E(): 3.1
Smith-Waterman score: 71; 34.783% identity (39.024% ungapped) in 46 aa overlap
(74-119:68-108)

      50      60      70      80      90     100
SYHT0H GLGGAGNSKSLCQKPTSCHPPCLYPAARNVPRAVYPKASCNLTDGSFGRPITATTDLKPC
      :..: : : . . : . : : :
ALLERG TVPKRENDPINPLKNVNVLYCAFTKRGIFTPKGVNTKQYINYCE----KTIISPADIKLC
      40      50      60      70      80      90

      110     120
SYHT0H ASIDLSKCVYNVDPRPNL
      .: : : : . : : :
ALLERG KKIA-SKCVKKVYDRPGPIIERSKNLLSCVLKKGLELTVYGKKK
      100     110     120     130

>>ALLERGEN_2012_2|gi|14424465|sp|P35777.2|VA4_SOLIN RecN (137 aa)
  initn: 45 init1: 45 opt: 67 Z-score: 89.6 bits: 22.2 E(): 5.6
Smith-Waterman score: 67; 34.783% identity (39.024% ungapped) in 46 aa overlap
(74-119:68-108)

      50      60      70      80      90     100
SYHT0H GLGGAGNSKSLCQKPTSCHPPCLYPAARNVPRAVYPKASCNLTDGSFGRPITATTDLKPC
      :..: : : . . : . : : :
ALLERG TVPKRENDPINPLRNVNWYCAFTKRGIFTPKGVNTKQYINYCEKTAISP----ADIKLC
      40      50      60      70      80      90

      110     120
SYHT0H ASIDLSKCVYNVDPRPNL
      .: : : : . : : :
ALLERG KKIA-SKCVKKVYDRPGPIIERSKNLLSCVLKKGLELTVYGKKK
      100     110     120     130

>>ALLERGEN_2012_2|gi|291092710|gb|ADD74392.1| Sol s 2 al (137 aa)
  initn: 45 init1: 45 opt: 67 Z-score: 89.6 bits: 22.2 E(): 5.6
Smith-Waterman score: 67; 34.783% identity (39.024% ungapped) in 46 aa overlap
(74-119:68-108)

      50      60      70      80      90     100
SYHT0H GLGGAGNSKSLCQKPTSCHPPCLYPAARNVPRAVYPKASCNLTDGSFGRPITATTDLKPC
      :..: : : . . : . : : :
ALLERG TVPKHENDPINPLRNVNWYCAFTKRGIFTPKGVNTKQYINYCEKTAISP----ADIKLC
      40      50      60      70      80      90

      110     120
SYHT0H ASIDLSKCVYNVDPRPNL
      .: : : : . : : :
ALLERG KKIA-SKCVKKVYDRPGPIIERSKNLLSCVLKKGLELTVYGKKK
      100     110     120     130

```

121 residues in 1 query sequences  
 367796 residues in 1603 library sequences  
 Scomplib [34t11]  
 start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012  
 Scan time: 0.040 Display time: 0.000

Function used was FASTA [version 3.4t11 Apr 17, 2002]  
 FASTA searches a protein or DNA sequence data bank  
 version 3.4t11 Apr 17, 2002  
 Please cite:  
 W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

/grid/.qsub\_tmp/1330373600kde\_agent/grid\_NbCuMBoIO: 71 aa  
 >SYHT0H2\_I\_5  
 vs /data/fasta/ALLERGEN\_2012\_2 library  
 searching /data/fasta/ALLERGEN\_2012\_2 library

367796 residues in 1603 sequences  
 Expectation\_n fit:  $\rho(\ln(x)) = 3.4962 \pm 0.0031$ ;  $\mu = 7.8021 \pm 0.162$   
 mean\_var=39.9531  $\pm 10.614$ , 0's: 13 Z-trim: 13 B-trim: 0 in 0/42  
 Lambda= 0.2029

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2  
 join: 36, opt: 24, gap-pen: -12/-2, width: 16  
 Scan time: 0.020

The best scores are:

					opt	bits	E(1603)
ALLERGEN_2012_2	gi 1364213 emb CAA44344.1	fel d	( 92)	57	23	1	
ALLERGEN_2012_2	gi 194350815 gb ACF53836.1	Bla g	( 191)	51	22	5.8	
ALLERGEN_2012_2	gi 105969545 gb ABF81662.1	EXPB1	( 269)	51	22	7.4	
ALLERGEN_2012_2	gi 115502168 sp P0C1Y5.1	EXB11_MA	( 269)	51	22	7.4	
ALLERGEN_2012_2	gi 14285595 sp O04004.1	NLTP6_AMB	( 118)	48	21	7.6	
ALLERGEN_2012_2	gi 190613907 gb ACE80957.1	putat	( 246)	50	22	8.5	
ALLERGEN_2012_2	gi 19009 emb CAA46705.1	CME [Hor	( 148)	48	21	8.9	
ALLERGEN_2012_2	gi 170712 gb AAA34276.1	pre-alph	( 291)	50	22	9.6	

>>ALLERGEN\_2012\_2|gi|1364213|emb|CAA44344.1| fel d I cha (92 aa)  
 initn: 57 init1: 57 opt: 57 Z-score: 102.8 bits: 23.3 E(): 1  
 Smith-Waterman score: 57; 45.000% identity (45.000% ungapped) in 20 aa overlap (6-25:9-28)

	10	20	30	40	50
SYHT0H	MLQGHSL	ELFALT	LGLTDS	LCPSIM	WSFGVLM
	:	:	:	:	:
ALLERG	EGGLCSR	ASLGLA	LDLWKD	CEICPA	VKRDVD
	10	20	30	40	50
					60

>>ALLERGEN\_2012\_2|gi|194350815|gb|ACF53836.1| Bla g 4 is (191 aa)  
 initn: 33 init1: 33 opt: 51 Z-score: 89.3 bits: 21.8 E(): 5.8  
 Smith-Waterman score: 51; 30.189% identity (32.653% ungapped) in 53 aa overlap (22-71:127-178)

	10	20	30	40
SYHT0H	MLQGHSL	ELFALT	LGLTDS	LCPSIM
	:	:	:	:
ALLERG	TIDYDDE	GKAFS	APYSVL	ATDYD
	100	110	120	130
				140
				150

	50	60	70
SYHT0H	LLQDLIC	LQGTQH	MLLSYF
	:	:	:
ALLERG	ALQHYTV	HQVNQ	HKKAIE



```

160          170          180          190

>>ALLERGEN_2012_2|gi|105969545|gb|ABF81662.1| EXPB10 [Ze (269 aa)
  initn: 46 initl: 46 opt: 51 Z-score: 87.4 bits: 22.0 E(): 7.4
Smith-Waterman score: 51; 55.000% identity (64.706% ungapped) in 20 aa overlap
(24-43:5-21)

          10          20          30          40          50          60
SYHT0H MLQGHSLLEFALTGLGLTDSLFCPSIMWSFGVLMQLVLLVVSALHLEHHLLQDLICLQGTQ
          :::::  :... :... ::
ALLERG          MTVVSIMWS---LVQVQVLVAVALSFLVGGAWCGPPKVPVPG
                      10          20          30

          70
SYHT0H HMLLSYFDLEH

ALLERG KNITANYGSDWLDKATWYGKPTGAGPDDNGGGCGYKDVNKAPFNSMGACGNVPIFKDGL
          40          50          60          70          80          90

>>ALLERGEN_2012_2|gi|115502168|sp|P0C1Y5.1|EXB11_MAIZE R (269 aa)
  initn: 42 initl: 42 opt: 51 Z-score: 87.4 bits: 22.0 E(): 7.4
Smith-Waterman score: 51; 55.000% identity (64.706% ungapped) in 20 aa overlap
(24-43:5-21)

          10          20          30          40          50          60
SYHT0H MLQGHSLLEFALTGLGLTDSLFCPSIMWSFGVLMQLVLLVVSALHLEHHLLQDLICLQGTQ
          :::::  :... :... ::
ALLERG          MTVVSIMWS---LVQVQVLVAVALAFLVGGAWCGPPKVPVPG
                      10          20          30

          70
SYHT0H HMLLSYFDLEH

ALLERG KNITAKYGSDWLDKATWYGKPTGAGPDDNGGGCGYKDVNKAPFNSMGACGNVPIFKDGL
          40          50          60          70          80          90

>>ALLERGEN_2012_2|gi|14285595|sp|O04004.1|NLTP6_AMBAR Re (118 aa)
  initn: 48 initl: 48 opt: 48 Z-score: 87.2 bits: 20.7 E(): 7.6
Smith-Waterman score: 48; 46.667% identity (46.667% ungapped) in 15 aa overlap
(22-36:3-17)

          10          20          30          40          50          60
SYHT0H MLQGHSLLEFALTGLGLTDSLFCPSIMWSFGVLMQLVLLVVSALHLEHHLLQDLICLQGTQ
          :  :... :... ::
ALLERG          MDCIRILWSVAVGLLLVSWRPTMFAASPTCDTVQNILAPCA
                      10          20          30          40

          70
SYHT0H HMLLSYFDLEH

ALLERG GFLTGQEPSKACCTGVNNLNNRKTADRVAVCNCIKELTKSIAYDPKRMPLLSTKCGVK
          50          60          70          80          90          100

>>ALLERGEN_2012_2|gi|190613907|gb|ACE80957.1| putative a (246 aa)
  initn: 45 initl: 45 opt: 50 Z-score: 86.3 bits: 21.6 E(): 8.5
Smith-Waterman score: 50; 43.478% identity (45.455% ungapped) in 23 aa overlap (6-
28:51-72)

          10          20          30
SYHT0H          MLQGHSLLEFALTGLGLTDSLFCPSIMWSFGVLMQL
          ::  :  :  :... :  :  :
ALLERG GAHAAKITFTNKCSYTVWPGTLTGDKPQLSLTGFEATGISNSVDAPSP-WSGRFFGRT
          30          40          50          60          70

```

```

          40          50          60          70
SYHT0H VLLVVSALHLEHLLQDLICLQGTQHMLLSYFDLEH

```

```

ALLERG RCSTDASGKFTCATADCGSGQVSCNGNGAVPPATLVEITIAENGGQDFYDVSLVDGFNLP
          80          90         100         110         120         130

```

```

>>ALLERGEN_2012_2|gi|19009|emb|CAA46705.1| CMe [Hordeum (148 aa)
  initn: 41 initl: 41 opt: 48 Z-score: 86.0 bits: 20.8 E(): 8.9
Smith-Waterman score: 48; 33.333% identity (33.333% ungapped) in 15 aa overlap
(19-33:134-148)

```

```

                  10          20          30          40
SYHT0H          MLQGHSLFLFALTTLGLTDSLFCPSIMWSFGVLMQLVLLVVSALHLEHH
                  : .:: . . .::..
ALLERG DSPNCPREERQTSYAANLVTPQECNLGTIHGSAYCPQLPGYGVVL
          110         120         130         140

```

```

          50          60          70
SYHT0H LLQDLICLQGTQHMLLSYFDLEH

```

```

>>ALLERGEN_2012_2|gi|170712|gb|AAA34276.1| pre-alpha-/be (291 aa)
  initn: 35 initl: 35 opt: 50 Z-score: 85.4 bits: 21.7 E(): 9.6
Smith-Waterman score: 50; 29.688% identity (32.203% ungapped) in 64 aa overlap (3-
62:155-217)

```

```

                  10          20          30
SYHT0H          MLQGHSLFLFALTTLGLTDSLFCPSIMWSFGVL
                  : . : : : : .:: : . .::
ALLERG QQQQQQILQQILQQQLIPCRDVLQQHNIAGSSQVLQESTYQLVQQQLCCQQ-LWQIPEQ
          130         140         150         160         170         180

```

```

          40          50          60          70
SYHT0H --MQLVLLVVSAA--LHLEHLLQDLICLQGTQHMLLSYFDLEH
          : . : : : : : : : : :
ALLERG SRCQAIHNVVHAILHQQHHHQQQQQQQQQPLSQVFSFQQPQQQYPSGQGFQPSQQNP
          190         200         210         220         230         240

```

```

ALLERG QAQGSFQPQQLPQFEEIRNLALQTLPMCNVYIPPYCTIAPFGIFGTN
          250         260         270         280         290

```

```

71 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.020 Display time: 0.000

```

```

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 82 aa
>SYHT0H2_I_6
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

```

```

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 4.3795+/-0.00353; mu= 4.7410+/- 0.183
mean_var=49.6123+/-14.903, 0's: 13 Z-trim: 14 B-trim: 79 in 1/42

```

Lambda= 0.1821

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2  
join: 36, opt: 24, gap-pen: -12/-2, width: 16  
Scan time: 0.030

The best scores are:

					opt	bits	E(1603)
ALLERGEN_2012_2	gi 62550933 emb CAI79052.1	putat	( 326)	76	28	0.21	
ALLERGEN_2012_2	gi 886965 emb CAA59339.1	low mol	( 261)	66	25	1.1	
ALLERGEN_2012_2	gi 170732 gb AAA34286.1	gamma-gl	( 323)	66	25	1.3	
ALLERGEN_2012_2	gi 62484809 emb CAI78902.1	putat	( 285)	61	24	2.9	
ALLERGEN_2012_2	gi 886963 emb CAA59338.1	low mol	( 229)	60	23	2.9	
ALLERGEN_2012_2	gi 160285626 pdb 2JMH A Chain A,		( 119)	57	22	3	
ALLERGEN_2012_2	gi 4538529 emb CAB39376.1	Cop c1	( 81)	55	22	3.2	
ALLERGEN_2012_2	gi 111120450 gb ABH06359.1	Blo t	( 134)	57	22	3.3	
ALLERGEN_2012_2	gi 4204917 gb AAD10850.1	major I	( 134)	57	22	3.3	
ALLERGEN_2012_2	gi 111120436 gb ABH06352.1	Blo t	( 134)	57	22	3.3	
ALLERGEN_2012_2	gi 335331566 gb AEH31546.1	low m	( 369)	61	24	3.6	
ALLERGEN_2012_2	gi 170734 gb AAA34287.1	gamma gl	( 244)	59	23	3.7	
ALLERGEN_2012_2	gi 112754 sp P04403.2 2SS1_BEREX		( 146)	55	22	5.1	
ALLERGEN_2012_2	gi 21926 emb CAA36063.1	unnamed	( 295)	58	23	5.2	
ALLERGEN_2012_2	gi 112745 sp P23110.1 2SS8_HELAN		( 141)	54	22	5.9	
ALLERGEN_2012_2	gi 75219081 sp O22108 O22108_WHEA		( 285)	57	23	6	
ALLERGEN_2012_2	gi 170730 gb AAA34285.1	gamma-gl	( 304)	57	23	6.3	
ALLERGEN_2012_2	gi 21773 emb CAA31685.1	unnamed	( 307)	57	23	6.4	
ALLERGEN_2012_2	gi 21755 emb CAA25593.1	unnamed	( 286)	56	22	7.2	
ALLERGEN_2012_2	gi 21761 emb CAA26384.1	unnamed	( 286)	56	22	7.2	
ALLERGEN_2012_2	gi 170720 gb AAA34280.1	alpha/be	( 286)	56	22	7.2	
ALLERGEN_2012_2	gi 21757 emb CAA26383.1	unnamed	( 296)	55	22	8.9	
ALLERGEN_2012_2	gi 170740 gb AAA34290.1	gliadin	( 296)	55	22	8.9	

>>ALLERGEN\_2012\_2|gi|62550933|emb|CAI79052.1| putative L (326 aa)  
initn: 75 initl: 51 opt: 76 Z-score: 115.2 bits: 27.6 E(): 0.21  
Smith-Waterman score: 86; 30.137% identity (33.333% ungapped) in 73 aa overlap  
(12-80:117-186)

	10	20	30
SYHT0H	MLHLHKKLQLLQLLPF----	PPFQQMLQEPLLLHMDLLSDL	
	:: ::	::::: ::::: ..	
ALLERG	QLPPFSQQQQPPFSQQQQPVLPQQPSFSQQQLPPFSQQQLPPFSQQ-QQPVLLQQQI--PF		
	90	100	110
	120	130	140
	40	50	60
SYHT0H	LELGLLMQLKHFALVLEEQDLLLLQILVTDLDLLKWSCMEMWF		
	.. ::::: . : . . . . : . . : . :		
ALLERG	VHPSILQQINPCKVFLQQQCSPVAMPQSLARSQMLQQSSCHVMQQQCCQQLPQIPQQSRY		
	150	160	170
	180	190	200
ALLERG	EAIRAIVYSIILQEQQQVQGSIQTQQQQPQQLGQCVSQPQQQSQQQLGQCSFQQPQQLQQ		
	210	220	230
	240	250	260

>>ALLERGEN\_2012\_2|gi|886965|emb|CAA59339.1| low molecula (261 aa)  
initn: 95 initl: 65 opt: 66 Z-score: 102.4 bits: 24.9 E(): 1.1  
Smith-Waterman score: 76; 29.412% identity (30.769% ungapped) in 68 aa overlap  
(13-80:124-188)

	10	20	30	40
SYHT0H	MLHLHKKLQLLQLLPFPPFQQMLQEPLLLHMDLLSDLLGL			
	:: ::::: ::::: ..			
ALLERG	PQPPFSQQQQPPFSQQQQPPFSQQQQPILLQPPFSQH-QQPVLPQQQIPS--VQPSI			
	100	110	120	130
	140	150		

```

          50          60          70          80
SYHT0H LMQLKHFALVLLLEEQDLLLLQQILVTDLDLLKWSCMEMWF
      :.:. . : . . . . : :. . : . :. :
ALLERG LQQLNPKVFLQQQCSPVAMPQSLARSQMLWQSSCHVMQQQCCRQLPQIPEQSRDAIRA
          160          170          180          190          200          210

ALLERG IIYSIVLQEQQHGGQGLNQPPQQQPPQSVQGVSQPPQQQKQLGQCSFQQPPQQ
          220          230          240          250          260

>>ALLERGEN_2012_2|gi|170732|gb|AAA34286.1| gamma-gliadin (323 aa)
  initn: 100 initl: 50 opt: 66 Z-score: 101.0 bits: 25.0 E(): 1.3
Smith-Waterman score: 66; 25.862% identity (26.786% ungapped) in 58 aa overlap
(17-74:93-148)

          10          20          30          40
SYHT0H MLHLHKKLQLLQLLPFPFPFQQMLQEPLLLHMDLLSDLLELGLLMQL
      :.:. . : :. . . . : . . . :
ALLERG NNSPNNNFQYHTHPSNHNLPHTNNIQQQPPFSQQQPPFSQQQPPV--LPQQSPFSQQ
          70          80          90          100          110          120

          50          60          70          80
SYHT0H KHFAVLLLEEQDLLLLQQILVTDLDLLKWSCMEMWF
      ....: .:. . :. :. . . . :
ALLERG QQLVLPQQQQQQLVQQQIPVQPSVLQQLNPKVFLQQQCSPVAMPQRLARSQMWQQSS
          130          140          150          160          170          180

>>ALLERGEN_2012_2|gi|62484809|emb|CAI78902.1| putative g (285 aa)
  initn: 121 initl: 61 opt: 61 Z-score: 94.7 bits: 23.6 E(): 2.9
Smith-Waterman score: 61; 60.000% identity (60.000% ungapped) in 15 aa overlap
(12-26:87-101)

          10          20          30          40
SYHT0H MLHLHKKLQLLQLLPFPFPFQQMLQEPLLLHMDLLSDLLELG
      :.:. :. :. :. :
ALLERG QPQHQPFPPTQQFPQRPLLPFTTHPFLTTFPDQLLPQPPHQSFPQPPQSYQPPLQFPFPFP
          60          70          80          90          100          110

          50          60          70          80
SYHT0H LLMQLKHFAVLLLEEQDLLLLQQILVTDLDLLKWSCMEMWF

ALLERG QQKYEPQPPFPWQQPTIQLYLQQQLNPCKEFLQCCRQPVSLLSYLWSKIVQQSSCRVM
          120          130          140          150          160          170

>>ALLERGEN_2012_2|gi|886963|emb|CAA59338.1| low molecula (229 aa)
  initn: 70 initl: 44 opt: 60 Z-score: 94.7 bits: 23.3 E(): 2.9
Smith-Waterman score: 60; 26.027% identity (30.159% ungapped) in 73 aa overlap (9-
77:68-134)

          10          20          30
SYHT0H MLHLHKKLQLLQLLPFPFPFQQMLQEPLLLHMDLLSDLL
      : . :. :. . . :. :. :
ALLERG LFPQQQPPFPQQPPFSQQQPSFSQQQPPFSQQQPILPEPPFS-LQQQPVLPPQSPFSQQ-
          40          50          60          70          80          90

          40          50          60          70          80
SYHT0H ELGLLMQLKHFAVLLLEEQDLLLLQQILVTDLD---LLKWSCMEMWF
      . : : . . : :. . . :. . . :
ALLERG QLVLPQQQQ---QLPQQQISIVQPSVLQQLNPKVFLQQQCSPVAMPQRLARSQMWQQ
          100          110          120          130          140          150

ALLERG SRCHVMQQCCQQLSQIPEQSRDAICAITYSIILQEQQGQFVQAQQQPPQSGQGVQS
          160          170          180          190          200          210

```

```
>>ALLERGEN_2012_2|gi|160285626|pdb|2JMH|A Chain A, Nmr S (119 aa)
  initn: 39 initl: 39 opt: 57 Z-score: 94.5 bits: 22.3 E(): 3
Smith-Waterman score: 57; 31.481% identity (35.417% ungapped) in 54 aa overlap
(23-70:53-106)
```

```

              10      20      30      40
SYHT0H      MLHLHKKLQLLQLLPFPFQQMLQEPLLLHMDLLSDLLE--LGLLM-QLKHF
              ::: .. ..... ..: : : : :
ALLERG QANHAIEKGEHQLLYLQHQLDENENKSKELQEKIIRELDVVCAMIEGAQGALERELKRT
              30      40      50      60      70      80

              50      60      70      80
SYHT0H ALVLLLE---QDLLLLQQILVTDLDLLKWSCMEMWF
              : : : : .. : : : : : :
ALLERG DLNILERFNYEEAQTLSKILLKDLKETEQKVVDIQTQ
              90      100     110
```

```
>>ALLERGEN_2012_2|gi|4538529|emb|CAB39376.1| Cop c1 alle (81 aa)
  initn: 50 initl: 50 opt: 55 Z-score: 94.0 bits: 21.7 E(): 3.2
Smith-Waterman score: 55; 38.889% identity (41.176% ungapped) in 36 aa overlap
(13-47:25-59)
```

```

              10      20      30      40
SYHT0H      MLHLHKKLQLLQLLPFPFQQMLQEPLLLHMDLLSDLLELG-LLMQLK
              ::: : . . : : : : . : . : : : :
ALLERG RFLPSSSHLNPQHLPWLVPVLLPVLQKPVVHPLLL-LPLDTTTLHMPPLLLQLQ
              10      20      30      40      50

              50      60      70      80
SYHT0H HFALVLLLEEQDLLLLQQILVTDLDLLKWSCMEMWF

ALLERG LPPLLSQGNPACSPKWLQLLVP
              60      70      80
```

```
>>ALLERGEN_2012_2|gi|111120450|gb|ABH06359.1| Blo t 5 al (134 aa)
  initn: 39 initl: 39 opt: 57 Z-score: 93.7 bits: 22.3 E(): 3.3
Smith-Waterman score: 57; 31.481% identity (35.417% ungapped) in 54 aa overlap
(23-70:68-121)
```

```

              10      20      30      40
SYHT0H      MLHLHKKLQLLQLLPFPFQQMLQEPLLLHMDLLSDLLE--LGLLM-QLKHF
              ::: .. ..... ..: : : : :
ALLERG QANHAIEKGEHQLLYLQHQLDENENKSKELQEKIIRELDVVCAMIEGAQGALERELKRT
              40      50      60      70      80      90

              50      60      70      80
SYHT0H ALVLLLE---QDLLLLQQILVTDLDLLKWSCMEMWF
              : : : : .. : : : : : :
ALLERG DLNILERFNYEEAQTLSKILLKDLKETEQKVVDIQTQ
              100     110     120     130
```

```
>>ALLERGEN_2012_2|gi|4204917|gb|AAD10850.1| major IgE-bi (134 aa)
  initn: 39 initl: 39 opt: 57 Z-score: 93.7 bits: 22.3 E(): 3.3
Smith-Waterman score: 57; 31.481% identity (35.417% ungapped) in 54 aa overlap
(23-70:68-121)
```

```

              10      20      30      40
SYHT0H      MLHLHKKLQLLQLLPFPFQQMLQEPLLLHMDLLSDLLE--LGLLM-QLKHF
              ::: .. ..... ..: : : : :
ALLERG QANHAIEKGEHQLLYLQHQLDENENKSKELQEKIIRELDVVCAMIEGAQGALERELKRT
              40      50      60      70      80      90

              50      60      70      80
```

SYHT0H ALVLLEE---QDLLLLQQILVTDLDLLKWSMEMWF  
 : .:. . . :.:. . :  
 ALLERG DLNILERFNYEEAQTLSKILLKDLKETEQKVVDIQTQ  
 100 110 120 130

>>ALLERGEN\_2012\_2|gi|111120436|gb|ABH06352.1| Blo t 5 al (134 aa)  
 initn: 39 init1: 39 opt: 57 Z-score: 93.7 bits: 22.3 E(): 3.3  
 Smith-Waterman score: 57; 31.481% identity (35.417% ungapped) in 54 aa overlap  
 (23-70:68-121)

10 20 30 40  
 SYHT0H MLHLHKKLQLLQLLPFPFQQMLQEPLLLHMDLLSDLLE--LGLLM-QLKHF  
 :.:. . . :.:. . :  
 ALLERG QANHAIEKGEHQLLYLQHQDELNENKSKELQEKIIRELDVVCAMIEGAQGALERELKRT  
 40 50 60 70 80 90

50 60 70 80  
 SYHT0H ALVLLEE---QDLLLLQQILVTDLDLLKWSMEMWF  
 : .:. . . :.:. . :  
 ALLERG DLNILERFNYEEAQTLSKILLKDLKETEQKVVDIQTQ  
 100 110 120 130

>>ALLERGEN\_2012\_2|gi|335331566|gb|AEH31546.1| low molecu (369 aa)  
 initn: 57 init1: 57 opt: 61 Z-score: 93.1 bits: 23.7 E(): 3.6  
 Smith-Waterman score: 65; 25.000% identity (25.806% ungapped) in 64 aa overlap  
 (17-80:182-243)

10 20 30 40  
 SYHT0H MLHLHKKLQLLQLLPFPFQQMLQEPLLLHMDLLSDLLELGLLMQ  
 :.:. . . :.:. . :  
 ALLERG QQSPFSQQQIIVLQQPPFLQQQPSLPQQPPFSQQQQQLVLPQQQI--PFVHPSILQQL  
 160 170 180 190 200

50 60 70 80  
 SYHT0H KHFAVLLEEQDLLLLQQILVTDLDLLKWSMEMWF  
 . . : . . . : . . : . :  
 ALLERG NPCKVFLQQQCSPVAMPQSLARSQMLQQSSCHVMQQQCCQLPQIPQQSRYEAIIRAIYS  
 210 220 230 240 250 260

ALLERG IILQEQQQVQGSIQTPQQQPQQLGQCVSQPQQQSQQQLGQQPQQQQLAQGTFLQPHQIAQ  
 270 280 290 300 310 320

>>ALLERGEN\_2012\_2|gi|170734|gb|AAA34287.1| gamma gliadin (244 aa)  
 initn: 43 init1: 43 opt: 59 Z-score: 92.9 bits: 23.1 E(): 3.7  
 Smith-Waterman score: 59; 25.758% identity (27.419% ungapped) in 66 aa overlap  
 (15-80:57-118)

10 20 30 40  
 SYHT0H MLHLHKKLQLLQLLPFPFQQMLQEPLLLHMDLLSDLLELGLLM  
 :.:. . . :.:. . :  
 ALLERG QPSPQPQQVQIISPATPTTIPSAGKPTSAPFP--QQQQQHQLAQQQI--PVVQPSILQ  
 30 40 50 60 70 80

50 60 70 80  
 SYHT0H QLKHFAVLLEEQDLLLLQQILVTDLDLLKWSMEMWF  
 :. . : . . . : . . : . :  
 ALLERG QLNPKVFLQQQCSPVAMPQRLARSQMLQQSSCHVMQQQCCQLPQIPQQSRYQAIRAI  
 90 100 110 120 130 140

ALLERG YSIILQEQQQVQGSIQSQQQQPQQLGQCVSQPQQQSQQQLGQQPQQQQLAQGTFLQPHQI  
 150 160 170 180 190 200

>>ALLERGEN\_2012\_2|gi|112754|sp|P04403.2|2SS1\_BEREX RecNa (146 aa)

initn: 29 init1: 29 opt: 55 Z-score: 90.4 bits: 21.8 E(): 5.1  
 Smith-Waterman score: 55; 25.714% identity (25.714% ungapped) in 35 aa overlap (5-39:52-86)

```

                                10      20      30
SYHT0H                      MLHLHKKLQLLQLLPFPFQQLQEPLLLHMDLL
                                : . . . : . :. : . . . :.
ALLERG AFRATVTTTVVEEENQEECREQMQRQQLSHCRMYMRQQMEESPYQTMPRRGMPEHSEC
                                30      40      50      60      70      80

                                40      50      60      70      80
SYHT0H SDLLELGLLMQKHFAVLLEEQDLLLLQQILVTDLDLLKWSCMEMWF
                                . :.
ALLERG CEQLEGMDDESCRCEGLRMMMMRMQQEEMQPRGEQMRMMRLAENIPSRCNLSPMRCPMGG
                                90      100     110     120     130     140

```

>>ALLERGEN\_2012\_2|gi|21926|emb|CAA36063.1| unnamed prote (295 aa)  
 initn: 70 init1: 47 opt: 58 Z-score: 90.3 bits: 22.8 E(): 5.2  
 Smith-Waterman score: 61; 26.087% identity (28.125% ungapped) in 69 aa overlap (12-80:95-158)

```

                                10      20      30      40
SYHT0H                      MLHLHKKLQLLQLLPFPFQQLQEPLLLHMDLLSDLLELG
                                : : .: :. :. : . . . .
ALLERG SFLQQQPILPQLPFSQQQQPVLPQQSPFSQQQLVLPP-QQQYQQVLQQQIPIVQP----S
                                70      80      90      100     110

                                50      60      70      80
SYHT0H LLMQLKHFAVLLEEQDLLLLQQILVTDLDLLKWSCMEMWF
                                .: :. . : . . . : :. : :. :
ALLERG VLQQLNPCKVFLQQQCNPVAMPQRLARSQMLQQSSCHVMQQQCCQQLPQIPEQSRVDVIR
                                120     130     140     150     160     170

ALLERG AITYSIILQEQQQGFVQAQQQQPQQLGQGVVSQSQQQSQQQLGQCSFQQPQQQLGQQPQQQ
                                180     190     200     210     220     230

```

>>ALLERGEN\_2012\_2|gi|112745|sp|P23110.1|2SS8\_HELAN RecNa (141 aa)  
 initn: 38 init1: 38 opt: 54 Z-score: 89.2 bits: 21.6 E(): 5.9  
 Smith-Waterman score: 54; 33.333% identity (40.741% ungapped) in 33 aa overlap (5-31:85-117)

```

                                10      20
SYHT0H                      MLHLHKKLQLLQLLPF-----PPFQQLQEPLL
                                :. :. :. :. :. :. :. :. :. :.
ALLERG EAEMLNHCMYLMKNLGERSQVSPRMREEDHKQLCCMQLKNLDEKCMCPAIMMMLNPEPMW
                                60      70      80      90      100     110

                                30      40      50      60      70      80
SYHT0H LHMDLLSDLLELGLLMQKHFAVLLEEQDLLLLQQILVTDLDLLKWSCMEMWF
                                ...
ALLERG IRMRDQVMSMAHNLPIECNLMSQPCQM
                                120     130     140

```

>>ALLERGEN\_2012\_2|gi|75219081|sp|O22108|O22108\_WHEAT LMM (285 aa)  
 initn: 43 init1: 43 opt: 57 Z-score: 89.0 bits: 22.6 E(): 6  
 Smith-Waterman score: 64; 31.944% identity (38.333% ungapped) in 72 aa overlap (12-74:45-113)

```

                                10      20      30
SYHT0H                      MLHLHKKLQLLQLLPF----PPFQQLQEPLLLHMDLLSDL
                                : : : : :. :. :. :. :. :.
ALLERG LPPQQTFFPQQPLFSQQQQQQLFPPQPPFSQQQPPFWQQQPPFSQ--QQPILPQQPPFSQQ
                                20      30      40      50      60      70

```

```

      40      50      60      70      80
SYHT0H LELGLLMQLKHFA----LVLL-EEQDLLLLQQILVTDLDLLKWSMEMWF
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
ALLERG QQL-VLPQQSPFSQQQQLILPPQQQQQLPQQQISIVQPSVLQQLNPCKVFLQQQCSPVAM
      80      90      100      110      120      130

ALLERG PQLARSQMWQQSSCHVMQQQCCQQLSQIPEQSRDYDAIRAITYSIILQEQQQGFVQAQQQ
      140      150      160      170      180      190

```

>>ALLERGEN\_2012\_2|gi|170730|gb|AAA34285.1| gamma-gliadin (304 aa)  
 initn: 57 init1: 57 opt: 57 Z-score: 88.6 bits: 22.6 E(): 6.3  
 Smith-Waterman score: 64; 27.692% identity (31.579% ungapped) in 65 aa overlap  
 (17-77:82-142)

```

      10      20      30      40
SYHT0H      MLHLHKKLQLLQLLPFPFPFQQMLQEPLLLHMDLLSDLLELGLLMQL
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
ALLERG QQQQQPLPQQPSFSQQQPPFSQQQPILSQPPFSQQ-QQPVLPPQQSPFSQQQQLVLPQQQ
      60      70      80      90      100      110

      50      60      70      80
SYHT0H KHFAVLLEEQDLLLLQQILVTDLD---LLKWSMEMWF
      .. .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
ALLERG QQQQLV---QQQIPVQPSVLQQLNPCKVFLQQQCSPVAMPQRLARSQMWQQSSCHVMQQ
      120      130      140      150      160

ALLERG QCCQQLQQIPEQSRYEAIRAIYSIILQEQQQGFVQPQQQQPQQSGQGVSSQSQSQSQQL
      170      180      190      200      210      220

```

>>ALLERGEN\_2012\_2|gi|21773|emb|CAA31685.1| unnamed prote (307 aa)  
 initn: 43 init1: 43 opt: 57 Z-score: 88.6 bits: 22.6 E(): 6.4  
 Smith-Waterman score: 64; 26.087% identity (27.692% ungapped) in 69 aa overlap  
 (12-80:117-181)

```

      10      20      30      40
SYHT0H      MLHLHKKLQLLQLLPFPFPFQQMLQEPLLLHMDLLSDLLELG
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
ALLERG PQQPPFSQQQQLVLPQQPPFSQQQQPVLPQQSPFP--QQQQHQQLVQQQI--PVVQPS
      90      100      110      120      130      140

      50      60      70      80
SYHT0H LLMQLKHFAVLLEEQDLLLLQQILVTDLDLLKWSMEMWF
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
ALLERG ILQQLNPCKVFLQQQCSPVAMPQRLARSQMLQQSSCHVMQQQCCQQLPQIPQQSRYEAIR
      150      160      170      180      190      200

ALLERG AIIYSIILQEQQQVQGSIQSQQQPQQLGQCVSQPQQQSQQQLGQQPQQQQLAQGTFLQP
      210      220      230      240      250      260

```

>>ALLERGEN\_2012\_2|gi|21755|emb|CAA25593.1| unnamed prote (286 aa)  
 initn: 88 init1: 56 opt: 56 Z-score: 87.6 bits: 22.3 E(): 7.2  
 Smith-Waterman score: 56; 45.000% identity (45.000% ungapped) in 20 aa overlap (7-  
 26:72-91)

```

      10      20      30
SYHT0H      MLHLHKKLQLLQLLPFPFPFQQMLQEPLLLHMDLLSD
      .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
ALLERG VPLVQQQQFLGQQQPFPFPQQPYPQPFPFSQLPYLQLQFPFPQPLPYSPQPPFRPQQPYP
      50      60      70      80      90      100

```



```

      40      50      60      70      80
SYHT0H LLELGLLMQLKHFALVLEEQDLLLLQQILVTDLDLLKWSCMEMWF

ALLERG QPQPQYSQPQQPISQQQQQQQQQQQQQQQQQQQQILQQILQQQLIPCMDVVLQQHNIAGHRS
      110      120      130      140      150      160

>>ALLERGEN_2012_2|gi|21761|emb|CAA26384.1| unnamed prote (286 aa)
  initn: 88 initl: 56 opt: 56 Z-score: 87.6 bits: 22.3 E(): 7.2
Smith-Waterman score: 56; 45.000% identity (45.000% ungapped) in 20 aa overlap (7-
26:72-91)

      10      20      30
SYHT0H      MLHLHKKLQLLQLLPFPFPFQQMLQEPLLLHMDLLSD
      .: ::: ::: : ..:
ALLERG VPLVQQQQFLGQQQPFPPQQPYQPQPFPSQLPYLQLQFPQPQLPYSQPQFRPQQPY
      50      60      70      80      90      100

      40      50      60      70      80
SYHT0H LLELGLLMQLKHFALVLEEQDLLLLQQILVTDLDLLKWSCMEMWF

ALLERG QPQPQYSQPQQPISQQQQQQQQQQQQQQQQQQQQILQQILQQQLIPCMDVVLQQHNIAGHRS
      110      120      130      140      150      160

>>ALLERGEN_2012_2|gi|170720|gb|AAA34280.1| alpha/beta-gl (286 aa)
  initn: 88 initl: 56 opt: 56 Z-score: 87.6 bits: 22.3 E(): 7.2
Smith-Waterman score: 56; 45.000% identity (45.000% ungapped) in 20 aa overlap (7-
26:72-91)

      10      20      30
SYHT0H      MLHLHKKLQLLQLLPFPFPFQQMLQEPLLLHMDLLSD
      .: ::: ::: : ..:
ALLERG VPLVQQQQFLGQQQPFPPQQPYQPQPFPSQLPYLQLQFPQPQLPYSQPQFRPQQPY
      50      60      70      80      90      100

      40      50      60      70      80
SYHT0H LLELGLLMQLKHFALVLEEQDLLLLQQILVTDLDLLKWSCMEMWF

ALLERG QPQPQYSQPQQPISQQQQQQQQQQQQQQQQQQQQILQQILQQQLIPCMDVVLQQHNIAGHRS
      110      120      130      140      150      160

>>ALLERGEN_2012_2|gi|21757|emb|CAA26383.1| unnamed prote (296 aa)
  initn: 45 initl: 45 opt: 55 Z-score: 86.0 bits: 22.1 E(): 8.9
Smith-Waterman score: 55; 52.632% identity (62.500% ungapped) in 19 aa overlap
(11-26:76-94)

      10      20      30
SYHT0H      MLHLHKKLQLLQLLPFP---PFQQMLQEPLLLHMDLLSD
      ::: ::: : : :
ALLERG QQQQFPGQQQFPPQQPYQPQPFPSQQPYLQLQFPQPQPFPPQLPYQPFPFSPQQPY
      50      60      70      80      90      100

      40      50      60      70      80
SYHT0H LELGLLMQLKHFALVLEEQDLLLLQQILVTDLDLLKWSCMEMWF

ALLERG PQQPQYPQPPQISQQQAQQQQQQQQQQQQQQQQQQQQILQQILQQQLIPCRDVVLQQHNI
      110      120      130      140      150      160

>>ALLERGEN_2012_2|gi|170740|gb|AAA34290.1| gliadin [Trit (296 aa)
  initn: 45 initl: 45 opt: 55 Z-score: 86.0 bits: 22.1 E(): 8.9
Smith-Waterman score: 55; 52.632% identity (62.500% ungapped) in 19 aa overlap
(11-26:76-94)

```

```

              10          20          30
SYHT0H      MLHLHKKLQLLQLLPFP---PFQQMLQEPLLLHMDLLSDL
              ::: :::  ::  .:  :
ALLERG QQQQFPGQQQQFPPQPYQPQPFPSQQPYLQLQFPQPFPPQLPYQPFPFSPQQPY
          50          60          70          80          90          100

          40          50          60          70          80
SYHT0H LELGLLMQLKHFALVLLLEEQDLLLLQQILVTDLDLLKWSCMEMWF

ALLERG PQPQPYPQPQQPISQQQAQQQQQQQQQQQQQQQQQILPQILQQQLIPCRDVLVQQHNI
          110          120          130          140          150          160

```

82 residues in 1 query sequences  
367796 residues in 1603 library sequences  
Scomplib [34t11]  
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012  
Scan time: 0.030 Display time: 0.010

Function used was FASTA [version 3.4t11 Apr 17, 2002]  
FASTA searches a protein or DNA sequence data bank  
version 3.4t11 Apr 17, 2002  
Please cite:  
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 37 aa
>SYHT0H2_I_7
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

```

367796 residues in 1603 sequences  
Expectation\_n fit:  $\rho(\ln(x)) = 4.1229 \pm 0.00258$ ;  $\mu = 0.9804 \pm 0.134$   
mean\_var=26.4253 $\pm$  6.997, 0's: 19 Z-trim: 22 B-trim: 0 in 0/43  
Lambda= 0.2495

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2  
join: 36, opt: 24, gap-pen: -12/-2, width: 16  
Scan time: 0.010

The best scores are:

						opt	bits	E(1603)
ALLERGEN_2012_2	gi 664852 gb AAB07620.1	Asp FII	( 250)	67	30	0.019		
ALLERGEN_2012_2	gi 66849502 gb EAL89830.1	major	( 304)	67	30	0.023		
ALLERGEN_2012_2	gi 83300352 sp P79017.2	ALL2_ASPF	( 310)	67	30	0.023		
ALLERGEN_2012_2	gi 20387027 emb CAC84590.2	tropo	( 284)	46	22	4		
ALLERGEN_2012_2	gi 62484809 emb CAI78902.1	putat	( 285)	43	21	8.4		

```

>>ALLERGEN_2012_2|gi|664852|gb|AAB07620.1| Asp FII [Aspe (250 aa)
initn: 61 init1: 61 opt: 67 Z-score: 134.1 bits: 29.6 E(): 0.019
Smith-Waterman score: 67; 42.857% identity (45.000% ungapped) in 21 aa overlap
(12-31:85-105)

```

```

              10          20          30
SYHT0H      MEPRDDLRSKPTSNTMV-DQEFNTSHWLLTMCFELSGK
              :.:... :. ... :.:...:
ALLERG DKANVLFRCNDPNDGNCALGEGWHWRGANATSETVICDRSYTTRRWLVSMCSQGYTVAGS
          60          70          80          90          100          110

ALLERG ETNTFWASDLMHRLYHVPVAVGQGWVDHFDGYDEVIALAKSNGTESTHDSEAFEYFALEA
          120          130          140          150          160          170

```

```

>>ALLERGEN_2012_2|gi|66849502|gb|EAL89830.1| major aller (304 aa)
initn: 61 init1: 61 opt: 67 Z-score: 132.6 bits: 29.6 E(): 0.023

```

Smith-Waterman score: 67; 42.857% identity (45.000% ungapped) in 21 aa overlap  
(12-31:145-165)

```

              10      20      30
SYHT0H      MEPRDDLRSKPTSNTMV-DQEFNTSHWLLTMCFELSGK
              ::::: :. .: ::::
ALLERG DKANVLFRCNDPDGNCALEGWGGHWRGANATSETVICDRSYTTRRWLVSMCSQGYTVAGS
              120      130      140      150      160      170

ALLERG ETNTFWASDLMHRLYHVPVAVGGQGWVDHFADGYDEVIALAKSNGTESTHDSEALQYFALEA
              180      190      200      210      220      230
```

>>ALLERGEN\_2012\_2|gi|83300352|sp|P79017.2|ALL2\_ASPFU Rec (310 aa)  
initn: 61 initl: 61 opt: 67 Z-score: 132.4 bits: 29.6 E(): 0.023  
Smith-Waterman score: 67; 42.857% identity (45.000% ungapped) in 21 aa overlap  
(12-31:145-165)

```

              10      20      30
SYHT0H      MEPRDDLRSKPTSNTMV-DQEFNTSHWLLTMCFELSGK
              ::::: :. .: ::::
ALLERG DKANVLFRCNDPDGNCALEGWGGHWRGANATSETVICDRSYTTRRWLVSMCSQGYTVAGS
              120      130      140      150      160      170

ALLERG ETNTFWASDLMHRLYHVPVAVGGQGWVDHFADGYDEVIALAKSNGTESTHDSEALQYFALEA
              180      190      200      210      220      230
```

>>ALLERGEN\_2012\_2|gi|20387027|emb|CAC84590.2| tropomyosi (284 aa)  
initn: 34 initl: 34 opt: 46 Z-score: 92.3 bits: 22.0 E(): 4  
Smith-Waterman score: 46; 33.333% identity (38.095% ungapped) in 24 aa overlap (1-  
24:40-60)

```

              10      20      30
SYHT0H      MEPRDDLRSKPTSNTMVDQEFNTSHWLLTM
              .: ::::~::~ :. ::::~:
ALLERG AMKLEKDNAMDKADALEAQARDANRKADKILEEVQDLKKKPSQ--VETDFTTTKENLAT
              10      20      30      40      50      60
```

SYHT0H CFELSGK

```

ALLERG ANKNLEDKEKTLTNTSEVASLNRKVQMIEENLERSEERLGTALTKLGEASHAADEASRM
              70      80      90      100      110      120
```

>>ALLERGEN\_2012\_2|gi|62484809|emb|CAI78902.1| putative g (285 aa)  
initn: 29 initl: 29 opt: 43 Z-score: 86.4 bits: 20.9 E(): 8.4  
Smith-Waterman score: 43; 29.630% identity (30.769% ungapped) in 27 aa overlap  
(10-35:132-158)

```

              10      20      30
SYHT0H      MEPRDDLRSKPTSNTMVDQEFNT-SHWLLTMCFELSGK
              .:: . :~::~ :~::~ :. :
ALLERG QSYQPPLQPFPQPQQKYPEQPQQPFPWQQPTIQLYLQQQLNPCKEFLLQQCRPVSLLS
              110      120      130      140      150      160

ALLERG YLWSKIVQQSSCRMVQQQCCLQLAQIPEQYKCTAIDSIVHAIFMQQGQRQGVQIVQQQPQ
              170      180      190      200      210      220
```

37 residues in 1 query sequences  
367796 residues in 1603 library sequences  
Scomplib [34t11]  
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012

```

Scan time: 0.010 Display time: 0.000

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 52 aa
>SYHT0H2_I_8
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 5.6255+/-0.00284; mu= -4.7927+/- 0.147
mean_var=31.6374+/- 8.245, 0's: 9 Z-trim: 9 B-trim: 132 in 1/42
Lambda= 0.2280

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.020
The best scores are:
ALLERGEN_2012_2|gi|154101366|gb|ABS58503.1| serin ( 84) 45 21 3.1
ALLERGEN_2012_2|gi|29163773|emb|CAD80019.1| unnam ( 65) 41 20 5.6
ALLERGEN_2012_2|gi|3668408|gb|AAC61869.1| tropomy ( 160) 44 21 8.9
ALLERGEN_2012_2|gi|42559558|sp|O97192.1|TPM_HELAS ( 284) 47 21 9.4

>>ALLERGEN_2012_2|gi|154101366|gb|ABS58503.1| serine pro (84 aa)
initn: 40 initl: 40 opt: 45 Z-score: 94.2 bits: 21.1 E(): 3.1
Smith-Waterman score: 45; 34.783% identity (34.783% ungapped) in 23 aa overlap
(26-48:32-54)

          10          20          30          40          50
SYHT0H      MKESVESLEMSCQRNRSRSVKNWVFSLIETIKVCYSRSSPNQLVIVPLFSSK
              .. .. :: .. .: :.....
ALLERG SPVVKKPEGGNTDTGDHNNQKTEWPELVGKSVEEAKKVMQDKSEAQIVVLPVGTIVTME
          10          20          30          40          50          60
ALLERG YRIDRVRLFVDSLDKIAQVPRVG
          70          80

>>ALLERGEN_2012_2|gi|29163773|emb|CAD80019.1| unnamed pr (65 aa)
initn: 33 initl: 33 opt: 41 Z-score: 89.7 bits: 19.9 E(): 5.6
Smith-Waterman score: 41; 54.545% identity (54.545% ungapped) in 11 aa overlap
(33-43:44-54)

          10          20          30          40          50
SYHT0H ESVESLEMSCQRNRSRSVKNWVFSLIETIKVCYSRSSPNQLVIVPLFSSK
              ::::: :.:
ALLERG GAFFLIASITATQTSHPAKFHVEGEVYCNVCHSRNLINELSERMAGAQQVL
          20          30          40          50          60

>>ALLERGEN_2012_2|gi|3668408|gb|AAC61869.1| tropomyosin (160 aa)
initn: 32 initl: 32 opt: 44 Z-score: 86.0 bits: 20.5 E(): 8.9
Smith-Waterman score: 44; 23.404% identity (24.444% ungapped) in 47 aa overlap (4-
48:79-125)

          10          20          30
SYHT0H      MKESVESLEMSCQRNRSR--SVKNWVFSLIETI
              .....: :. : .. .: .
ALLERG EVDLERAEARLEAAEAKVLELEEEELKVVGNMKSLEISEQEASQREDSYEETIRDLTQRL
          50          60          70          80          90          100

```

```

          40          50
SYHT0H KVCYSRSSPNQLVIVPLFSSK
      :   . . . .   :
ALLERG KDAENRATEAERTVSKLQKEVDRLEDELLAEKERYKAISDELDQTFaelAGY
      110          120          130          140          150          160

>>ALLERGEN_2012_2|gi|42559558|sp|O97192.1|TPM_HELAS RecN (284 aa)
  initn: 32 initl: 32 opt: 47 Z-score: 85.6 bits: 21.3 E(): 9.4
Smith-Waterman score: 47; 25.532% identity (26.667% ungapped) in 47 aa overlap (4-
48:203-249)

```

```

                                10          20          30
SYHT0H                                MKESVESLEMSCQRNRSR--SVKNWVFSLIETI
                                . . . . . : . : : . . . . : . .
ALLERG EVDLERAEARLEAAEAKILELEELKVVGNMKSLEISEQEASQREDSYEETIRDLTQRL
      180          190          200          210          220          230

```

```

          40          50
SYHT0H KVCYSRSSPNQLVIVPLFSSK
      :   . . . :   :
ALLERG KDAENRASEAERTVSKLQKEVDRLEDELLAEKERYKATSDELDSTFAELAGY
      240          250          260          270          280

```

```

52 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.020 Display time: 0.000

```

```

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 79 aa
>SYHT0H2_I_9
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

```

```

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 5.1484+/-0.00358; mu= 1.8745+/- 0.184
mean_var=50.0852+/-12.821, 0's: 6 Z-trim: 6 B-trim: 0 in 0/43
Lambda= 0.1812

```

```

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.030

```

```

The best scores are:
ALLERGEN_2012_2|gi|162929|gb|AAA30479.1| alpha-s2 ( 222) 57 22 6.4
ALLERGEN_2012_2|gi|4138175|emb|CAA09885.1| allerg ( 187) 56 22 6.5
ALLERGEN_2012_2|gi|18479082|gb|AAL73404.1|AF44942 ( 515) 60 23 8.1
ALLERGEN_2012_2|gi|21666498|gb|AAM73729.1|AF39589 ( 536) 60 23 8.4
ALLERGEN_2012_2|gi|21914823|gb|AAM73730.2|AF39589 ( 538) 60 23 8.5
ALLERGEN_2012_2|gi|315190620|gb|ADT89774.1| phosp ( 302) 57 22 8.5

```

```

>>ALLERGEN_2012_2|gi|162929|gb|AAA30479.1| alpha-s2-like (222 aa)
  initn: 30 initl: 30 opt: 57 Z-score: 88.6 bits: 22.1 E(): 6.4
Smith-Waterman score: 57; 20.548% identity (21.127% ungapped) in 73 aa overlap (1-
72:138-209)

```

```

                                10      20      30
SYHT0H                      MTRRKSSSTWWSTTRLSTPKISKIQSQTK
                                ... .. . . . . . . . . . .
ALLERG POYLQYLYQGPIVLNPWDQVKRNAVPITPTLNREQLSTSEENSKKTVDMESTEFTKKTK
      110      120      130      140      150      160

                                40      50      60      70
SYHT0H GQLRFLFNKDDFSTKGNIRKPPRIPLPSYL-SLYCEDSGKGRWLLQMPSLR
      . . . . . . . . . . . . . . . . . . . . . . . . . .
ALLERG LTEEEKNRLNFKKISQRYQ-KFALPQYLKTVYQHQAAMKPWIQPKTKVIPYVRYL
      170      180      190      200      210      220

>>ALLERGEN_2012_2|gi|4138175|emb|CAA09885.1| allergen [M (187 aa)
  initn: 56 init1: 56 opt: 56 Z-score: 88.4 bits: 21.8 E(): 6.5
Smith-Waterman score: 56; 53.333% identity (53.333% ungapped) in 15 aa overlap
(54-68:6-20)

                                30      40      50      60      70
SYHT0H IQSQTKGQLRFLFNKDDFSTKGNIRKPPRIPLPSYLSLYCEDSGKGRWLLQMPSLR
                                : : : . : : : . : :
ALLERG                                AVSASPTPSKHNLVYCYAQGKDLFEFHINDTVTKDV
                                10      20      30

ALLERG CKSLNSGKYHNMNNEKYCSVADYDVKWFKERCQSHPTDVKTTKWIAGTDLKIEMDPKEPY
      40      50      60      70      80      90

>>ALLERGEN_2012_2|gi|18479082|gb|AAL73404.1|AF449424_1 1 (515 aa)
  initn: 48 init1: 48 opt: 60 Z-score: 86.7 bits: 22.9 E(): 8.1
Smith-Waterman score: 60; 24.000% identity (24.490% ungapped) in 50 aa overlap
(19-67:121-170)

                                10      20      30      40
SYHT0H                      MTRRKSSSTWWSTTRLSTPKISKIQSQTK-KGQLRFLFNKDDFSTKGN
                                : . . . : . . . : . . . : . . .
ALLERG YSNAPELIYIERGRGITGVLFPGCPETFEDPQQSQSQGQRQGGQSQQRSEQDRHQKIRHF
      100      110      120      130      140      150

                                50      60      70
SYHT0H RKPPRIPLPSYLSLYCEDSGKGRWLLQMPSLR
      . . . : . . . : . . . : . . . : . . . : . . . : . . .
ALLERG REGDIIALPAGVAHWYNDGDSPPVTVSLHTNNYANQLDENPRHFYLAGNPDEHQRRQG
      160      170      180      190      200      210

>>ALLERGEN_2012_2|gi|21666498|gb|AAM73729.1|AF395893_1 v (536 aa)
  initn: 45 init1: 45 opt: 60 Z-score: 86.4 bits: 22.9 E(): 8.4
Smith-Waterman score: 60; 28.261% identity (30.233% ungapped) in 46 aa overlap
(13-56:332-376)

                                10      20      30      40
SYHT0H                      MTRRKSSSTWWSTTRLSTPKISKIQSQTKGQLRFLFNKDDFS
                                . : : : . . . . : . . . . . . .
ALLERG KTSKDTLEKLFQDQGTIMKASKEQVRAMSRREGPKIWPFF-TEESTGSFKLFKKDPSQ
      310      320      330      340      350      360

                                50      60      70
SYHT0H TK--GNIRKPPRIPLPSYLSLYCEDSGKGRWLLQMPSLR
      . . . . . : : :
ALLERG SNKYGQLFEAERIDYPPLEKLDMMVVSYANITKGGMSVPFYNSRATKIAIVVSGEGCVEIA
      370      380      390      400      410      420

>>ALLERGEN_2012_2|gi|21914823|gb|AAM73730.2|AF395894_1 v (538 aa)
  initn: 45 init1: 45 opt: 60 Z-score: 86.4 bits: 22.9 E(): 8.5

```

Smith-Waterman score: 60; 28.261% identity (30.233% ungapped) in 46 aa overlap (13-56:334-378)

```

              10      20      30      40
SYHT0H          MTRRKSSSTWSTTRLSTPKISKIQSKTKGQLRLFNKDDFS
              . :   ::: . .... :.....: .
ALLERG KTSKDTLEKLFQKQDQGTIMKASKEQIRAMSRRGEGPKIWPF-TEESTGSFKLFKKDPSQ
              310      320      330      340      350      360
```

```

              50      60      70
SYHT0H TK--GNIRKPPRIPLPSYLSLYCEDSGKGRWLLQMPSLR
              .. :... . :: :
ALLERG SNKYGQLFEAERIDYPPLEKLDVVSYANITKGGMSVPFYNRATKIAIVVSGEGCVEIA
              370      380      390      400      410      420
```

>>ALLERGEN\_2012\_2|gi|315190620|gb|ADT89774.1| phospholip (302 aa)  
initn: 30 init1: 30 opt: 57 Z-score: 86.4 bits: 22.1 E(): 8.5  
Smith-Waterman score: 57; 28.333% identity (32.075% ungapped) in 60 aa overlap (9-67:246-299)

```

              10      20      30
SYHT0H          MTRRKSSSTWSTTRLSTPKISKIQSKTKGQLRL-FN
              : :::: :::: . :. : . . .:
ALLERG QPGCLVRTCSHTKAVLYMTECIKRECCLIGHTPWSSTG-SSPKPKPISACKRDICVCVGLN
              220      230      240      250      260      270
```

```

              40      50      60      70
SYHT0H KDDFSTKGNIRKPPRIPLPSYLSLYCEDSGKGRWLLQMPSLR
              .. :... : . : : : : :
ALLERG AKSYPAKGSFYVPVEKNSP-----YCHNEGIKL
              280      290      300
```

79 residues in 1 query sequences  
367796 residues in 1603 library sequences  
Scomplib [34t11]  
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012  
Scan time: 0.030 Display time: 0.000

Function used was FASTA [version 3.4t11 Apr 17, 2002]  
FASTA searches a protein or DNA sequence data bank  
version 3.4t11 Apr 17, 2002  
Please cite:  
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

/grid/.qsub\_tmp/1330373600kde\_agent/grid\_NbCuMBoIO: 183 aa  
>SYHT0H2\_I\_10  
vs /data/fasta/ALLERGEN\_2012\_2 library  
searching /data/fasta/ALLERGEN\_2012\_2 library

367796 residues in 1603 sequences  
Expectation\_n fit:  $\rho(\ln(x)) = 4.0748 \pm 0.00364$ ;  $\mu = 12.2148 \pm 0.190$   
mean\_var=60.0372 $\pm$ 15.922, 0's: 3 Z-trim: 4 B-trim: 0 in 0/44  
Lambda= 0.1655

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2  
join: 36, opt: 24, gap-pen: -12/-2, width: 16  
Scan time: 0.050

The best scores are:

			opt	bits	E(1603)
ALLERGEN_2012_2	gi 5059162 gb AAD38942.1 AF144060	( 496)	87	29	0.25
ALLERGEN_2012_2	gi 21748153 emb CAD38167.1  putat	( 124)	73	25	0.99
ALLERGEN_2012_2	gi 89892723 gb ABD79095.1  Zea m	( 252)	71	25	2.2

ALLERGEN_2012_2	gi 89892721 gb ABD79094.1	Zea m	( 263)	71	25	2.3
ALLERGEN_2012_2	gi 115502167 sp Q1ZYQ8.2	EXB10_MA	( 270)	71	25	2.3
ALLERGEN_2012_2	gi 105969545 gb ABF81662.1	EXPB1	( 269)	69	25	3.2
ALLERGEN_2012_2	gi 115502168 sp P0C1Y5.1	EXB11_MA	( 269)	69	25	3.2
ALLERGEN_2012_2	gi 18615 emb CAA26723.1	unnamed	( 495)	71	25	3.5
ALLERGEN_2012_2	gi 18635 emb CAA33215.1	glycinin	( 495)	69	25	4.9
ALLERGEN_2012_2	gi 18652047 gb AAL76932.1	AF45648	( 154)	64	23	5.1
ALLERGEN_2012_2	gi 2266625 emb CAB10765.1	group	( 264)	63	23	8.6

>>ALLERGEN\_2012\_2|gi|5059162|gb|AAD38942.1|AF144060\_1 al (496 aa)  
 initn: 50 initl: 50 opt: 87 Z-score: 113.9 bits: 29.1 E(): 0.25  
 Smith-Waterman score: 92; 26.364% identity (31.868% ungapped) in 110 aa overlap  
 (23-132:178-268)

	10	20	30	40	50
SYHT0H	MSPERRPVEIRPATAADMAAVCDIVNHYIETSTVNFRTPEQTPQEWIDDLER				
	: . . . . : . . . . . : . . . . .				
ALLERG	YTNPK	EARNCR	LSGLRDL	KQQSEYVR	QKQVDFLNHLIDIGVAGFRSDASTHQ-WPDDLRS
	150	160	170	180	190 200

	60	70	80	90	100	110
SYHT0H	LQDRYPWLVAEVEGVVAGIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLK					
	. . : : : : : : . . . . . : . . . . .					
ALLERG	IYSRLHNLNKEFF-----P---ENSQPFIIYHETIYYGGN----	GINSNEYTSLGR				
	210	220	230	240		

	120	130	140	150	160	170
SYHT0H	SMEAQGFKSVVAVIGLPNDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFQWQDFELPA					
	. . . : . . . : .					
ALLERG	IIIEFRFYKEITNVFR-GNNPLHLWLKNFGTEWGLVPSGDALVMIDSHDLRVGHTGKLGFN					
	250	260	270	280	290	300

>>ALLERGEN\_2012\_2|gi|21748153|emb|CAD38167.1| putative n (124 aa)  
 initn: 70 initl: 70 opt: 73 Z-score: 103.1 bits: 25.1 E(): 0.99  
 Smith-Waterman score: 73; 33.333% identity (34.146% ungapped) in 42 aa overlap  
 (92-132:16-57)

	70	80	90	100	110	120
SYHT0H	AEVEGVVAGIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKS					
	: . . : . . : . . . . .					
ALLERG	MSDFNAIAQQFVEFYKTFDGNRAGLGALYKEHSMLTFEAQGTQG					
		10	20	30	40	

	130	140	150	160	170	180
SYHT0H	VVAVI-GLPNDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRPV					
	. . . : : : .					
ALLERG	SAAIVEKLQNLPFQEIQHRTDTVDAQPSADDGILVLVTGALLLGESKPMSTQAFQLKN					
	50	60	70	80	90	100

>>ALLERGEN\_2012\_2|gi|89892723|gb|ABD79095.1| Zea m 1 all (252 aa)  
 initn: 71 initl: 71 opt: 71 Z-score: 96.8 bits: 25.0 E(): 2.2  
 Smith-Waterman score: 71; 30.189% identity (30.189% ungapped) in 53 aa overlap  
 (100-152:81-133)

	70	80	90	100	110	120
SYHT0H	GIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKSVVAVIGLP					
	: : : . . : . : : .					
ALLERG	NGGGCGYKDVNKPFFNSMGACGNIPFKDGLGCGSCFEIKCDKPVCECSGKPVVVHITDMN					
	60	70	80	90	100	110



```

      130      140      150      160      170      180
SYHT0H NDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRPVTQI
      .: . : :. :: :. :
ALLERG YEPIAAYHFDFLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKYDSKVTFFHLEKGC GPN
      120      130      140      150      160      170

>>ALLERGEN_2012_2|gi|89892721|gb|ABD79094.1| Zea m 1 all (263 aa)
  initn: 71 initl: 71 opt: 71 Z-score: 96.6 bits: 25.0 E(): 2.3
Smith-Waterman score: 71; 30.189% identity (30.189% ungapped) in 53 aa overlap
(100-152:92-144)

      70      80      90      100      110      120
SYHT0H GIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKSVVAVIGLP
      :: :: . . : .: .: :: . .
ALLERG NGGGCGYKDVNKPFFNSMGACGNIPFIDGLGCGSCFEIKCDKPVCECSGKPVVVHITDMN
      70      80      90      100      110      120

      130      140      150      160      170      180
SYHT0H NDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRPVTQI
      .: . : :. :: :. :
ALLERG YEPIAAYHFDFLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKFGSKVFSFHLEKGC GPN
      130      140      150      160      170      180

>>ALLERGEN_2012_2|gi|115502167|sp|Q1ZYQ8.2|EXB10_MAIZE R (270 aa)
  initn: 71 initl: 71 opt: 71 Z-score: 96.4 bits: 25.0 E(): 2.3
Smith-Waterman score: 71; 30.189% identity (30.189% ungapped) in 53 aa overlap
(100-152:99-151)

      70      80      90      100      110      120
SYHT0H GIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKSVVAVIGLP
      :: :: . . : .: .: :: . .
ALLERG NGGGCGYKDVNKPFFNSMGACGNIPFIDGLGCGSCFEIKCDKPVCECSGKPVVVHITDMN
      70      80      90      100      110      120

      130      140      150      160      170      180
SYHT0H NDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRPVTQI
      .: . : :. :: :. :
ALLERG YEPIAAYHFDFLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKYDSKVTFFHLEKGC GPN
      130      140      150      160      170      180

>>ALLERGEN_2012_2|gi|105969545|gb|ABF81662.1| EXPB10 [Ze (269 aa)
  initn: 94 initl: 69 opt: 69 Z-score: 93.9 bits: 24.5 E(): 3.2
Smith-Waterman score: 69; 30.189% identity (30.189% ungapped) in 53 aa overlap
(100-152:98-150)

      70      80      90      100      110      120
SYHT0H GIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKSVVAVIGLP
      :: :: . . : : : : :: . .
ALLERG NGGGCGYKDVNKPFFNSMGACGNVPIFIDGLGCGSCFEIKCDKPAECSGKPVVVYITDMN
      70      80      90      100      110      120

      130      140      150      160      170      180
SYHT0H NDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRPVTQI
      .: . : :. :: :. :
ALLERG YEPIAAYHFDFLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKYGSKVTFFHLEKGC NPN
      130      140      150      160      170      180

>>ALLERGEN_2012_2|gi|115502168|sp|P0C1Y5.1|EXB11_MAIZE R (269 aa)
  initn: 94 initl: 69 opt: 69 Z-score: 93.9 bits: 24.5 E(): 3.2
Smith-Waterman score: 69; 30.189% identity (30.189% ungapped) in 53 aa overlap
(100-152:98-150)

```

```

      70      80      90      100      110      120
SYHT0H GIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKSVVAVIGLP
      :: :: . . : : .: :: . .
ALLERG NGGGCGYKDVNKAPFNSMGACGNVPIFKDGLGCGSCFEIKCDKPAECSGKPVVVYITDMN
      70      80      90      100      110      120

      130      140      150      160      170      180
SYHT0H NDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRPVTQI
      .: . : :. :: :. :
ALLERG YEPiAAyHFDLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKYGSKVTFHLEKGCNPN
      130      140      150      160      170      180

```

>>ALLERGEN\_2012\_2|gi|18615|emb|CAA26723.1| unnamed prote (495 aa)  
 initn: 41 init1: 41 opt: 71 Z-score: 93.2 bits: 25.3 E(): 3.5  
 Smith-Waterman score: 71; 25.843% identity (29.114% ungapped) in 89 aa overlap  
 (39-118:113-200)

```

      10      20      30      40      50      60
SYHT0H EIRPATAADMAAVCDIVNHYIETSTVNFRTEPQTPQEWIDDLERLQDRYPWLVAEEVEGVV
      ::: ::. .. : ::. . :: .
ALLERG SYTNGPQEIYIQQGKGIFGMIYPGCSSTFEFPQQPQQRGQS-SRPQDRHQKIYNSREGDL
      90      100      110      120      130      140

```

```

      70      80      90      100      110
SYHT0H AGIAYAGPWKARNAYDWTVESTVYV---SHRHQ-----RLGLGSTLYTHLLKSMEAQGF
      .. . : : : :. . : ..: .: :... ::: . :
ALLERG IAVPTGVAWWMYNNEDTPVVAVSIIDTNSLENQLDQMPRRFYLAGNQEQEFLKYQQEQGG
      150      160      170      180      190      200

```

```

      120      130      140      150      160      170
SYHT0H KSVVAVIGLPNDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRP
ALLERG HQSQKGKHQQEEENEKGSILSGFTLEFLEHAFSVDKQIAKNLQGENEGEDKGAIVTVKGG
      210      220      230      240      250      260

```

>>ALLERGEN\_2012\_2|gi|18635|emb|CAA33215.1| glycinin subu (495 aa)  
 initn: 41 init1: 41 opt: 69 Z-score: 90.7 bits: 24.8 E(): 4.9  
 Smith-Waterman score: 69; 25.843% identity (29.114% ungapped) in 89 aa overlap  
 (39-118:113-200)

```

      10      20      30      40      50      60
SYHT0H EIRPATAADMAAVCDIVNHYIETSTVNFRTEPQTPQEWIDDLERLQDRYPWLVAEEVEGVV
      ::: ::. .. : ::. . :: .
ALLERG SYTNGPQEIYIQQGKGIFGMIYPGCPSTFEFPQQPQQRGQS-SRPQDRHQKIYNFREGDL
      90      100      110      120      130      140

```

```

      70      80      90      100      110
SYHT0H AGIAYAGPWKARNAYDWTVESTVYV---SHRHQ-----RLGLGSTLYTHLLKSMEAQGF
      .. . : : : :. . : ..: .: :... ::: . :
ALLERG IAVPTGVAWWMYNNEDTPVVAVSIIDTNSLENQLDQMPRRFYLAGNQEQEFLKYQQEQGG
      150      160      170      180      190      200

```

```

      120      130      140      150      160      170
SYHT0H KSVVAVIGLPNDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRP
ALLERG HQSQKGKHQQEEENEKGSILSGFTLEFLEHAFSVDKQIAKNLQGENEGEDKGAIVTVKGG
      210      220      230      240      250      260

```

>>ALLERGEN\_2012\_2|gi|18652047|gb|AAL76932.1|AF456481\_1 m (154 aa)  
 initn: 33 init1: 33 opt: 64 Z-score: 90.3 bits: 23.1 E(): 5.1  
 Smith-Waterman score: 64; 22.727% identity (25.773% ungapped) in 110 aa overlap  
 (44-146:21-124)

```

      20      30      40      50      60      70
SYHT0H TAADMAAVCDIVNHYIETSTVNFRTEPQTPQEWIDDLERLQDRY-PWLVAEVEGVVAGIA
      : .. :.. . . : :. :. :. :.
ALLERG      MGVSQKTEVEAPSTVSAEKMYQGFLDMDTVFPKVLPLQLIKSVE-ILEGDG
      10      20      30      40

      80      90      100      110      120
SYHT0H YAGPWKARNAYDWTVESTVYVSHRHQ-----RLGLGSTLYTHLLKSMEAQGFKSVVA-VI
      .: . . . :. :.. . . . :. :. :. . . . :. :. :. :. :. :. :. :.
ALLERG GVGTVRLVHLG---EATEYTTMKQKVDVIDKAGLGYT-YTTIGGDILVEGLESVVNQFV
      50      60      70      80      90      100

      130      140      150      160      170      180
SYHT0H GLPNDPSVRLHEALGYTARGLRAAGYKHGGWHDVGFWRQDFELPAPPRPVRPVTQI
      .: :. . . . :. :. :. :. :. :. :. :. :. :. :. :. :.
ALLERG VVPTDGGCIVKNTTIYNTKGDVLPEDKVKTEATEKSALAFKAVEAYLLAN
      110      120      130      140      150

```

```

>>ALLERGEN_2012_2|gi|2266625|emb|CAB10765.1| group V all (264 aa)
  initn: 34 initl: 34 opt: 63 Z-score: 86.2 bits: 23.1 E(): 8.6
Smith-Waterman score: 63; 28.814% identity (30.357% ungapped) in 59 aa overlap
(101-157:17-74)

```

```

      80      90      100      110      120
SYHT0H IAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKSVVAVIGL--
      : ..: .:..... : :. :. :. :.
ALLERG      ADAGYTPAAPAAAGAGGKATTDEQKLEDVNA-GFKTAVAAAAANVP
      10      20      30      40

      130      140      150      160      170      180
SYHT0H PNDPSVRLHEALGYTARGLRAAGYKHGGWHDVGFWRQDFELPAPPRPVRPVTQI
      : : . . . :. :. :. :. :. :. :. :. :. :. :. :. :.
ALLERG PADKYKTFEAAFTASSKASIAAAATKAPGLIPQLNAATNTAYAAAQGATPEAKYDAFVTT
      50      60      70      80      90      100

```

```

183 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.050 Display time: 0.010

```

```

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 79 aa
>SYHT0H2_I_11
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

```

```

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 2.8002+/-0.00316; mu= 11.3231+/- 0.164
mean_var=37.6476+/- 8.777, 0's: 4 Z-trim: 4 B-trim: 0 in 0/43
Lambda= 0.2090

```

```

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.020
The best scores are:
opt bits E(1603)

```

ALLERGEN_2012_2	gi 16580747 dbj BAB71741.1  glyox	( 291)	51	22	6.2
ALLERGEN_2012_2	gi 84029333 sp Q948T6.2 LGUL_ORYS	( 291)	51	22	6.2
ALLERGEN_2012_2	gi 74664773 sp Q96X46.3 ENO_PENCI	( 438)	52	23	6.4
ALLERGEN_2012_2	gi 127533 sp P02761.1 MUP_RAT Rec	( 181)	49	22	7.2
ALLERGEN_2012_2	gi 323575367 dbj BAJ78223.1  Ani	( 295)	50	22	7.7
ALLERGEN_2012_2	gi 11991229 gb AAG42255.1 AF30670	( 296)	50	22	7.8
ALLERGEN_2012_2	gi 510515 emb CAA56343.1  Kunitz	( 208)	49	22	7.8
ALLERGEN_2012_2	gi 11991227 gb AAG42254.1 AF30670	( 303)	50	22	7.9

>>ALLERGEN\_2012\_2|gi|16580747|dbj|BAB71741.1| glyoxalase (291 aa)  
 initn: 48 initl: 48 opt: 51 Z-score: 88.8 bits: 22.5 E(): 6.2  
 Smith-Waterman score: 51; 35.714% identity (40.000% ungapped) in 28 aa overlap (2-26:146-173)

			10		20	
SYHT0H			MFLQTTPTKY	---	QVLFKQNDTDLNIGFYTN	
			..: ::		::... .: : .: ::	
ALLERG	CKKITREPGVPVKG	STVIAFAQDPDGYMFELI	QRGPTPEPLCQV	MLRVGDLDRS	IKFYEK	
	120	130	140	150	160	170

	30	40	50	60	70	
SYHT0H	HMIFINSVILYGRWRQ	KSSRIIFEMRKLSSYL	FSSSWLFFFLATCLNTSPP			
ALLERG	ALGMKLLRKKDVPDYKYTIAM	LGYADEDKTTVIELTYNYGVTEYTKGN	NAYAQVAIGTEDV			
	180	190	200	210	220	230

>>ALLERGEN\_2012\_2|gi|84029333|sp|Q948T6.2|LGUL\_ORYSJ Rec (291 aa)  
 initn: 48 initl: 48 opt: 51 Z-score: 88.8 bits: 22.5 E(): 6.2  
 Smith-Waterman score: 51; 35.714% identity (40.000% ungapped) in 28 aa overlap (2-26:146-173)

			10		20	
SYHT0H			MFLQTTPTKY	---	QVLFKQNDTDLNIGFYTN	
			..: ::		::... .: : .: ::	
ALLERG	CKKITREPGVPVKG	STVIAFAQDPDGYMFELI	QRGPTPEPLCQV	MLRVGDLDRS	IKFYEK	
	120	130	140	150	160	170

	30	40	50	60	70	
SYHT0H	HMIFINSVILYGRWRQ	KSSRIIFEMRKLSSYL	FSSSWLFFFLATCLNTSPP			
ALLERG	ALGMKLLRKKDVPDYKYTIAM	LGYADEDKTTVIELTYNYGVTEYTKGN	NAYAQVAIGTEDV			
	180	190	200	210	220	230

>>ALLERGEN\_2012\_2|gi|74664773|sp|Q96X46.3|ENO\_PENCI RecN (438 aa)  
 initn: 41 initl: 41 opt: 52 Z-score: 88.5 bits: 23.0 E(): 6.4  
 Smith-Waterman score: 52; 42.105% identity (42.105% ungapped) in 19 aa overlap (2-20:252-270)

			10		20		30
SYHT0H			MFLQTTPTKYQV	LFKQNDTDLNIGFYTNHMI			
			: .:	::... ::. .:			
ALLERG	EEALDLITEAIEQAGYTGKISIAM	DVASSEFYKTD	AKYDLDFKNPDS	DPTKWLT	YEQLA		
	230	240	250	260	270	280	

	40	50	60	70		
SYHT0H	FINSVILYGRWRQ	KSSRIIFEMRKLSSYL	FSSSWLFFFLATCLNTSPP			
ALLERG	DLYKS	LAAKYPIVSIEDPFAEDDWE	AWSYFYKTSDFQIVGDDLTVTNPLRIKKAIELKSC			
	290	300	310	320	330	340

>>ALLERGEN\_2012\_2|gi|127533|sp|P02761.1|MUP\_RAT RecName: (181 aa)  
 initn: 47 initl: 47 opt: 49 Z-score: 87.7 bits: 21.6 E(): 7.2

Smith-Waterman score: 49; 34.043% identity (42.105% ungapped) in 47 aa overlap (16-56:113-156)

```

              10      20      30
SYHT0H      MFLQTTPTKYQVLFKQNDTLNIGFYTNHMI-FINS----VILY
              ..: : . : :.: : :. ....
ALLERG CRELYLVAYKTPEDGEYFVEYDGGNTFTILKTDYDRYVMF--HLINFKNGETFQLMVLY
              90      100      110      120      130
```

```

              40      50      60      70
SYHT0H GRWRQKSSRIIFEMRKLSSYLFSSSWLFFFLATCLNTSPP
              :: .. :: : .. ::
ALLERG GRTKDLSSDIKEKFAKLCEAHGITRDNIIDLTKTDRCLQARG
              140      150      160      170      180
```

>>ALLERGEN\_2012\_2|gi|323575367|dbj|BAJ78223.1| Ani s 12 (295 aa)  
initn: 44 initl: 44 opt: 50 Z-score: 87.1 bits: 22.2 E(): 7.7  
Smith-Waterman score: 50; 37.037% identity (38.462% ungapped) in 27 aa overlap (22-48:198-223)

```

              10      20      30      40      50
SYHT0H      MFLQTTPTKYQVLFKQNDTLNIGFYTNHMIFINSVILYGRWRQKSSRIIF
              : :.. :.:.: :.:.: ::
ALLERG ECVDALGTPPVTTAANGAYQMAAPLHRCIENGGMKMCSTWINATICE-RWKQEC SRDKD
              170      180      190      200      210      220
```

```

              60      70
SYHT0H EMRKLSSYLFSSSWLFFFLATCLNTSPP
ALLERG AEPPTNFSQCIQTQTVMLQCKLEFGDKCKALQEECVAATYAPTAYVDANPPIFTSETIRC
              230      240      250      260      270      280
```

>>ALLERGEN\_2012\_2|gi|11991229|gb|AAG42255.1|AF306708\_1 p (296 aa)  
initn: 37 initl: 37 opt: 50 Z-score: 87.1 bits: 22.2 E(): 7.8  
Smith-Waterman score: 50; 40.909% identity (50.000% ungapped) in 22 aa overlap (4-25:44-61)

```

              10      20      30
SYHT0H      MFLQTTPTKYQVLFKQNDTLNIGFYTNHMIFI
              . :.: : :.. :.:.:
ALLERG VALVAGPAASYAADAGYAPTPPAAAGAAAGKITPTQEQLME---DINVGFKA AVAAAA
              20      30      40      50      60
```

```

              40      50      60      70
SYHT0H NSVILYGRWRQKSSRIIFEMRKLSSYLFSSSWLFFFLATCLNTSPP
ALLERG GAPPADKFKTFQAAFSASVEASAAKLNAAPGFVSHVAATSDATYKAAVGATPEAKFDS
              70      80      90      100      110      120
```

>>ALLERGEN\_2012\_2|gi|510515|emb|CAA56343.1| Kunitz tryps (208 aa)  
initn: 49 initl: 49 opt: 49 Z-score: 87.0 bits: 21.7 E(): 7.8  
Smith-Waterman score: 49; 41.176% identity (41.176% ungapped) in 17 aa overlap (42-58:134-150)

```

              20      30      40      50      60      70
SYHT0H VLFKQNDTLNIGFYTNHMIFINSVILYGRWRQKSSRIIFEMRKLSSYLFSSSWLFFFLA
              : :: .: :.... ::
ALLERG FHLCTPLSLNSFSVDRYSQGSARRTPCQTHWLQKHNRWCWFRIQRASSESNYKLVFCTSN
              110      120      130      140      150      160
```

ALLERG DDSSCGDIVAPIDREGNRPLIVTHDQNHPLL<sup>VQ</sup>FQKVEAYESSTA  
170 180 190 200

SYHT0H

10 20 30

MFLQTTPTKYQVLFKQNDTDLNIGFYTNHMIFI

. :. . : . . :. . .

ALLERG VALVAGPAASYAADAGYAPATPAAAGAAAGKITPTQEQKLME----DINVGFKA AVAAAAA

40 50 60 70  
SYHT0H NSVILYGRWRQKSSRIIFEMRKLSSYLFSSSWLFFFLATCLNTSP

ALLERG GAPPADKFKTFQAAFSASVEASAAKLNAAQAPGVSHVAATSDATYKAAVGATPEAKFDS

Function used was FASTA [version 3.4.t11 Apr 17, 2002]  
FASTA searches a protein or DNA sequence data bank  
version 3.4.t11 Apr 17, 2002  
Please cite:  
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```
367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 2.9169+/-0.00319; mu= 9.8082+/- 0.166
mean_var=28.8941+/- 7.269, 0's: 14 Z-trim: 14 B-trim: 23 in 1/41
Lambda= 0.2386
```

```
FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.020
```

The best scores are:					opt	bits	E(1603)
ALLERGEN_2012_2	gi	22135348	gb AAM93157.1	trypsi (219)	49	22	3.3
ALLERGEN_2012_2	gi	3703107	gb AAC63045.1	glycini (507)	49	22	6
ALLERGEN_2012_2	gi	224036293	pdb 3C3V A Chain A,	(510)	49	22	6
ALLERGEN_2012_2	gi	5712199	gb AAD47382.1	glycini (530)	49	22	6.2
ALLERGEN_2012_2	gi	199732457	gb ACH91862.1	arach (530)	49	22	6.2
ALLERGEN_2012_2	gi	21314465	gb AAM46958.1	AF51085 (538)	49	22	6.2

**Report Number: SSB-165-12**

```

              10      20      30
SYHT0H      MVDPHLSATNLCSSSEKAKPTAQFFHRGRDHFID
              :      :...: :...
ALLERG NNQEFECAGVALSRLVLRNALRRPFYSNAPQEIFIQQGRGYFGLIFPGCPSTYEPAQQ
              70      80      90      100      110      120

ALLERG GRRHQSQRP RR FQGQDQSQQQDSHQKVHRFDEGDLIAVPTGVAFWV MYNDHDTDVVAVS
              130      140      150      160      170      180

>>ALLERGEN_2012_2|gi|3703107|gb|AAC63045.1| glycinin [Ar (507 aa)
  initn: 34 initl: 34 opt: 49 Z-score: 89.1 bits: 22.2 E(): 6
Smith-Waterman score: 49; 40.000% identity (40.000% ungapped) in 15 aa overlap
(20-34:66-80)

```

```

              10      20      30
SYHT0H      MVDPHLSATNLCSSSEKAKPTAQFFHRGRDHFID
              :      :...: :...
ALLERG NNQEFECAGVALSRLVLRNALRRPFYSNAPQEIFIQQGRGYFGLIFPGCPRH YEEPHTQ
              40      50      60      70      80      90

ALLERG GRRSQSQRP RR LQGEDQSQQQDSHQKVHRFDEGDLIAVPTGVAFWLYNDHDTDVVAVS
              100      110      120      130      140      150

>>ALLERGEN_2012_2|gi|224036293|pdb|3C3V|A Chain A, Cryst (510 aa)
  initn: 34 initl: 34 opt: 49 Z-score: 89.1 bits: 22.2 E(): 6
Smith-Waterman score: 49; 40.000% identity (40.000% ungapped) in 15 aa overlap
(20-34:69-83)

```

```

              10      20      30
SYHT0H      MVDPHLSATNLCSSSEKAKPTAQFFHRGRDHFID
              :      :...: :...
ALLERG NNQEFECAGVALSRLVLRNALRRPFYSNAPQEIFIQQGRGYFGLIFPGCPSTYEPAQQ
              40      50      60      70      80      90

ALLERG GRRYQSQRP RR LQEEDQSQQQDSHQKVHRFNEGDLIAVPTGVAFWLYNDHDTDVVAVS
              100      110      120      130      140      150

>>ALLERGEN_2012_2|gi|5712199|gb|AAD47382.1| glycinin [Ar (530 aa)
  initn: 34 initl: 34 opt: 49 Z-score: 88.9 bits: 22.2 E(): 6.2
Smith-Waterman score: 49; 40.000% identity (40.000% ungapped) in 15 aa overlap
(20-34:89-103)

```

```

              10      20      30
SYHT0H      MVDPHLSATNLCSSSEKAKPTAQFFHRGRDHFID
              :      :...: :...
ALLERG NNQEFECAGVALSRLVLRNALRRPFYSNAPQEIFIQQGRGYFGLIFPGCPSTYEPAQQ
              60      70      80      90      100      110

ALLERG GRRYQSQRP RR LQEEDQSQQQDSHQKVHRFNEGDLIAVPTGVAFWLYNDHDTDVVAVS
              120      130      140      150      160      170

>>ALLERGEN_2012_2|gi|199732457|gb|ACH91862.1| arachin Ar (530 aa)
  initn: 34 initl: 34 opt: 49 Z-score: 88.9 bits: 22.2 E(): 6.2
Smith-Waterman score: 49; 40.000% identity (40.000% ungapped) in 15 aa overlap
(20-34:89-103)

```

```

              10      20      30
SYHT0H      MVDPHLSATNLCSSSEKAKPTAQFFHRGRDHFID
              :      :...: :...
ALLERG NNQEFECAGVALSRLVLRNALRRPFYSNAPQEIFIQQGRGYFGLIFPGCPSTYEPAQQ
              60      70      80      90      100      110

```

```

ALLERG GRRYQSQRPRLQEEDQSQQQQDSHQVHRFNEGDIAVPTGVAFWLYNDHDTDVVAVS
      120      130      140      150      160      170

```

```

>>ALLERGEN_2012_2|gi|21314465|gb|AAM46958.1|AF510854_1 a (538 aa)
  initn: 34 initl: 34 opt: 49 Z-score: 88.8 bits: 22.2 E(): 6.2
Smith-Waterman score: 49; 40.000% identity (40.000% ungapped) in 15 aa overlap
(20-34:89-103)

```

```

      10      20      30
SYHT0H      MVDPHLSATNLICSSEKAKPTAQFFHRGRDHFID
              :  . . . . :  . . . .
ALLERG NNQEFECAGVALSRLVLRNALRRPFYSNAPQEIFIQQGRGYFGLIFGCPSTYEPAQQ
      60      70      80      90      100      110

```

```

ALLERG GRRHQSQRPRLRFQGDQSQQQQDSHQVHRFDEGDIAVPTGVAFWMYNDHDTDVVAVS
      120      130      140      150      160      170

```

```

35 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.020 Display time: 0.000

```

```

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 40 aa
>SYHT0H2_I_13
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

```

```

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 3.4130+/-0.00286; mu= 9.5190+/- 0.150
mean_var=41.7198+/-10.742, 0's: 9 Z-trim: 9 B-trim: 0 in 0/43
Lambda= 0.1986

```

```

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.020

```

```

The best scores are:
ALLERGEN_2012_2|gi|170722|gb|AAA34281.1| pre-alpha ( 262) 56 22 3.8
ALLERGEN_2012_2|gi|473876|gb|AAA17741.1| alpha-gl ( 287) 56 22 4.1
ALLERGEN_2012_2|gi|5381323|gb|AAD42943.1|AF091841 ( 148) 52 21 5.8
ALLERGEN_2012_2|gi|1170095|sp|P46419.1|GSTM1 DERP ( 219) 52 21 7.5
ALLERGEN_2012_2|gi|60920878|gb|AAX37326.1| glutat ( 219) 52 21 7.5
ALLERGEN_2012_2|gi|105969543|gb|ABF81661.1| EXPB1 ( 99) 48 19 9.7

```

```

>>ALLERGEN_2012_2|gi|170722|gb|AAA34281.1| pre-alpha-/be (262 aa)
  initn: 28 initl: 28 opt: 56 Z-score: 92.5 bits: 22.1 E(): 3.8
Smith-Waterman score: 56; 30.000% identity (30.000% ungapped) in 30 aa overlap (2-
31:56-85)

```

```

      10      20      30
SYHT0H      MPLPTVVPKMDPHPRGASWKKKTFQPRLQSK
              :  . :  . . . . :  . . . . :  . . . . :
ALLERG PQLQPQNPSQQQPQEQVPLVQQQQFLGQQQFPFPQPYPQPQPFPSQQPYLQLQPFLQPQ
      30      40      50      60      70      80

```



```

40
SYHT0H WIDVISPLT

ALLERG LPYSQPQPFRRPQQPYQPQPQYSQPQQPISQQQQQQQQQQQQQQQQQQIIQQILQQQLI
      90      100      110      120      130      140

>>ALLERGEN_2012_2|gi|473876|gb|AA17741.1| alpha-gliadin (287 aa)
  initn: 28 initl: 28 opt: 56 Z-score: 92.1 bits: 22.1 E(): 4.1
Smith-Waterman score: 56; 30.000% identity (30.000% ungapped) in 30 aa overlap (2-
31:56-85)

      10      20      30
SYHT0H      MPLPTVVPKMDPHPRGASWKKKTFQPRLQSK
      : : : : : : : : : : : : : : :
ALLERG PQLQPQNPSQQQPQEQVPLVQQQQFLGQQQPFPPQQPYQPQPFPSPQQPYLQLQPFLQPQ
      30      40      50      60      70      80

40
SYHT0H WIDVISPLT

ALLERG LPYSQPQPFRRPQQPYQPQPQYSQPQQPISQQQQQQQQQQQQQQQQQQIIQQILQQQLI
      90      100      110      120      130      140

>>ALLERGEN_2012_2|gi|5381323|gb|AAD42943.1|AF091841_1 2S (148 aa)
  initn: 40 initl: 40 opt: 52 Z-score: 89.4 bits: 20.6 E(): 5.8
Smith-Waterman score: 52; 26.087% identity (26.087% ungapped) in 23 aa overlap
(11-33:36-58)

      10      20      30      40
SYHT0H      MPLPTVVPKMDPHPRGASWKKKTFQPRLQSKWIDVISPLT
      . . : : : : : : : : : : : :
ALLERG IVLAVLFAAALVSASAHKTVVTTSTVAEEGEEENQRGCEWESRQCQMRHCMQWMRSMRGQY
      10      20      30      40      50      60

ALLERG EESFLRSAEANQGQFEHFRECCNELRDVKSHCRCEALRCMMRQMQQEYGMEQEMQQMQM
      70      80      90      100      110      120

>>ALLERGEN_2012_2|gi|1170095|sp|P46419.1|GSTM1_DERPT Rec (219 aa)
  initn: 39 initl: 39 opt: 52 Z-score: 87.3 bits: 20.8 E(): 7.5
Smith-Waterman score: 52; 35.294% identity (41.379% ungapped) in 34 aa overlap (3-
32:182-214)

      10      20
SYHT0H      MPLPTVVPKMDPHPRGASWKKKTFQPRL----
      : : : : : : : : : : : : : :
ALLERG ANISYVDFNLYEYLCHVKVMVPEVFGQFENLKRYVERMESLPRVSDYIKKQ-QPKTFNAP
      160      170      180      190      200      210

      30      40
SYHT0H QSKWIDVISPLT
      : : :
ALLERG TSKWNASYA

>>ALLERGEN_2012_2|gi|60920878|gb|AAX37326.1| glutathione (219 aa)
  initn: 39 initl: 39 opt: 52 Z-score: 87.3 bits: 20.8 E(): 7.5
Smith-Waterman score: 52; 35.294% identity (41.379% ungapped) in 34 aa overlap (3-
32:182-214)

```

```

              10      20
SYHT0H      MPLPTVVPKMDPHPRGASWKKKTFQPRL----
              :  :  :.  :.  :.  :.  :.  :.
ALLERG ANISYVDFYLYEYLCRVKVMVPEVFGQFENLKRYVERMESLPRVSDYIKKQ-QPKTFNAP
              160      170      180      190      200      210

```

```

              30      40
SYHT0H QSKWIDVISPLT
              :::
ALLERG TSKWNASYA

```

```

>>ALLERGEN_2012_2|gi|105969543|gb|ABF81661.1| EXPB10 [Ze (99 aa)
  initn: 35 initl: 35 opt: 48 Z-score: 85.3 bits: 19.3 E(): 9.7
Smith-Waterman score: 48; 42.857% identity (42.857% ungapped) in 14 aa overlap
(13-26:51-64)

```

```

              10      20      30      40
SYHT0H      MPLPTVVPKMDPHPRGASWKKKTFQPRLQSKWIDVISPLT
              :  :.  :.  :.  :.  :.  :.
ALLERG ALLVKYVDGDGDIVAVDIKEKGSPTYEPLKHSWGAIWRKDSDKPIKGPITVRLTTEGGTK
              30      40      50      60      70      80

```

```

ALLERG TVYDDVIPAEWKPNAYTT
              90

```

```

40 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.020 Display time: 0.000

```

```

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 51 aa
>SYHT0H2_I_14
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

```

```

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 4.2069+/-0.00333; mu= 3.5577+/- 0.174
mean_var=45.1697+/-12.375, 0's: 9 Z-trim: 9 B-trim: 49 in 1/42
Lambda= 0.1908

```

```

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.020

```

```

The best scores are:
ALLERGEN_2012_2|gi|5381323|gb|AAD42943.1|AF091841 ( 148) 55 22 2.7
ALLERGEN_2012_2|gi|1405736|emb|CAA35188.1| trypsi ( 144) 54 22 3.2
ALLERGEN_2012_2|gi|112754|sp|P04403.2|2SS1_BEREX ( 146) 54 22 3.2
ALLERGEN_2012_2|gi|19009|emb|CAA46705.1| CMe [Hor ( 148) 54 22 3.3
ALLERGEN_2012_2|gi|549186|sp|P10737.3|VA53_DOLMA ( 215) 55 22 3.7
ALLERGEN_2012_2|gi|6136165|sp|P81657.1|VA5_VESMA ( 202) 54 22 4.2
ALLERGEN_2012_2|gi|13183177|gb|AAK15089.1|AF24000 ( 585) 57 23 5.6
ALLERGEN_2012_2|gi|3287877|sp|P81402.1|NLTP1_PRUP ( 91) 49 20 5.8
ALLERGEN_2012_2|gi|83754241|pdb|2B5S|B Chain B, C ( 92) 49 20 5.8

```

```

ALLERGEN_2012_2|gi|54793477|gb|AAV40850.1| lipid ( 117) 49 20 7.1
ALLERGEN_2012_2|gi|313575718|gb|ADR66939.1| non-s ( 117) 49 20 7.1
ALLERGEN_2012_2|gi|1398916|dbj|BAA07712.1| allerg ( 157) 50 21 7.4
ALLERGEN_2012_2|gi|156778059|gb|ABU95411.1| Sin a ( 92) 47 20 8.5
ALLERGEN_2012_2|gi|119524036|gb|ABL77410.1| UA3-r (1096) 57 23 9.2

>>ALLERGEN_2012_2|gi|5381323|gb|AAD42943.1|AF091841_1 2S (148 aa)
  initn: 33 initl: 33 opt: 55 Z-score: 95.3 bits: 22.1 E(): 2.7
Smith-Waterman score: 55; 37.931% identity (44.000% ungapped) in 29 aa overlap
(19-47:90-114)

              10      20      30      40
SYHT0H      MWWVWKGELFRVVQVHRGLREVTGSQAVCRCSEELRALHMTIVQTFGN
              :::  .. ::: : :: . . : :
ALLERG SMRGQYEEFSLRSAEANQGQFEHFRECCNELRDV---KSHCRC-EALRCMMRMQMQQEYGM
        60      70      80      90      100     110

        50
SYHT0H KVS

ALLERG EQEMQQMMQMMQYLPRMCGMSYPTECRMRIPIFA
        120     130     140

>>ALLERGEN_2012_2|gi|1405736|emb|CAA35188.1| trypsin inh (144 aa)
  initn: 42 initl: 42 opt: 54 Z-score: 93.9 bits: 21.8 E(): 3.2
Smith-Waterman score: 54; 39.130% identity (40.909% ungapped) in 23 aa overlap
(21-43:70-91)

              10      20      30      40      50
SYHT0H      MWWVWKGELFRVVQVHRGLREVTGSQAVCRCSEELRALHMTIVQTFGNKV
              .... : ::: : :: . . : :
ALLERG LRACRTYVVSQICHQGPRLLTSDMKRRCCDELSAIPAYCRC-EALRIIMQGVVTWQGAFF
        40      50      60      70      80      90

SYHT0H S

ALLERG GAYFKDSPNCPRERQTSYAANLVTPQECNLGTIHGSAYCPQLPAY
        100     110     120     130     140

>>ALLERGEN_2012_2|gi|112754|sp|P04403.2|2SS1_BEREX RecNa (146 aa)
  initn: 33 initl: 33 opt: 54 Z-score: 93.9 bits: 21.8 E(): 3.2
Smith-Waterman score: 54; 37.500% identity (39.130% ungapped) in 24 aa overlap
(21-44:84-106)

              10      20      30      40      50
SYHT0H      MWWVWKGELFRVVQVHRGLREVTGSQAVCRCSEELRALHMTIVQTFGNKV
              .. : . ::: : :: . . : :
ALLERG RMYMRQQMEESPYQTMPPRGMEPHMSECCQLEGMDDESCRC-EGLRMMMRMQQEEMQPR
        60      70      80      90      100     110

SYHT0H S

ALLERG GEQMRMMRLAENIPSRCNLSPMRCMPMGSIAGF
        120     130     140

>>ALLERGEN_2012_2|gi|19009|emb|CAA46705.1| CMe [Hordeum (148 aa)
  initn: 42 initl: 42 opt: 54 Z-score: 93.8 bits: 21.8 E(): 3.3
Smith-Waterman score: 54; 39.130% identity (40.909% ungapped) in 23 aa overlap
(21-43:70-91)

```



>>ALLERGEN\_2012\_2|gi|3287877|sp|P81402.1|NLTP1\_PRUPE Rec (91 aa)  
 initn: 46 initl: 46 opt: 49 Z-score: 89.4 bits: 20.3 E(): 5.8  
 Smith-Waterman score: 49; 38.889% identity (38.889% ungapped) in 18 aa overlap  
 (20-37:39-56)

```

              10      20      30      40
SYHT0H      MWWVWKGELFRVVQVHRGLREVTGSQAVCRCSEELRALHMTIVQTFGNK
              : .   ::. : . : .
ALLERG ALAPCIPYVRGGGAVPPACCNGIRNVNNLARTTPDRQAACNCLKQLSASVPGVNPNNAAA
              10      20      30      40      50      60

50
SYHT0H VS

ALLERG LPGKCGVHIPYKISASTNCATVK
              70      80      90

```

>>ALLERGEN\_2012\_2|gi|83754241|pdb|2B5S|B Chain B, Crysta (92 aa)  
 initn: 46 initl: 46 opt: 49 Z-score: 89.3 bits: 20.3 E(): 5.8  
 Smith-Waterman score: 49; 38.889% identity (38.889% ungapped) in 18 aa overlap  
 (20-37:40-57)

```

              10      20      30      40
SYHT0H      MWWVWKGELFRVVQVHRGLREVTGSQAVCRCSEELRALHMTIVQTFGNK
              : .   ::. : . : .
ALLERG SLAPCIPYVRGGGAVPPACCNGIRNVNNLARTTPDRQAACNCLKQLSASVPGVNPNNAAA
              10      20      30      40      50      60

50
SYHT0H VS

ALLERG LPGKCGVSIPYKISASTNCATVK
              70      80      90

```

>>ALLERGEN\_2012\_2|gi|54793477|gb|AAV40850.1| lipid trans (117 aa)  
 initn: 46 initl: 46 opt: 49 Z-score: 87.8 bits: 20.4 E(): 7.1  
 Smith-Waterman score: 49; 38.889% identity (38.889% ungapped) in 18 aa overlap  
 (20-37:65-82)

```

              10      20      30      40
SYHT0H      MWWVWKGELFRVVQVHRGLREVTGSQAVCRCSEELRALHMTIVQTFGNK
              : .   ::. : . : .
ALLERG SLAPCIPYVRGGGAVPPACCNGIRNVNNLARTTPDRQAACNCLKQLSASVPGVNPNNAAA
              40      50      60      70      80      90

50
SYHT0H VS

ALLERG LPGKCGVSIPYKISASTNCATVK
              100     110

```

>>ALLERGEN\_2012\_2|gi|313575718|gb|ADR66939.1| non-specif (117 aa)  
 initn: 46 initl: 46 opt: 49 Z-score: 87.8 bits: 20.4 E(): 7.1  
 Smith-Waterman score: 49; 38.889% identity (38.889% ungapped) in 18 aa overlap  
 (20-37:65-82)

```

              10      20      30      40
SYHT0H      MWWVWKGELFRVVQVHRGLREVTGSQAVCRCSEELRALHMTIVQTFGNK
              : .   ::. : . : .
ALLERG SLAPCIPYVRGGGAVPPACCNGIRNVNNLARTTPDRQAACNCLKQLSASVPGVNPNNAAA
              40      50      60      70      80      90

```

```

50
SYHT0H VS

ALLERG LPGKCGVSIPYKISASTNCATVK
      100      110

>>ALLERGEN_2012_2|gi|1398916|dbj|BAA07712.1| allergenic (157 aa)
  initn: 33 initl: 33 opt: 50 Z-score: 87.5 bits: 20.7 E(): 7.4
Smith-Waterman score: 50; 45.000% identity (50.000% ungapped) in 20 aa overlap
(14-31:72-91)

      10      20      30      40
SYHT0H      MWWVWKGELFRVVQVHRGL-REVTG-SQAVCRCSEELRALHMT
              :::: :... :... :::.
ALLERG QPGIGYPTYPLPRCRAFKRQCVAPGTLDEQVRRGCCRQLAGIDSSWCRCDALNHMLRII
      50      60      70      80      90      100

50
SYHT0H IVQTFGNKVS

ALLERG YREERAADAGHPMAEVFRGCRRGDIERAAASLP AFCNV DIPNGVGGVCYWLP GTGY
      110      120      130      140      150

>>ALLERGEN_2012_2|gi|156778059|gb|ABU95411.1| Sin a 3 al (92 aa)
  initn: 47 initl: 47 opt: 47 Z-score: 86.3 bits: 19.8 E(): 8.5
Smith-Waterman score: 47; 30.769% identity (30.769% ungapped) in 26 aa overlap (6-
31:26-51)

      10      20      30      40
SYHT0H      MWWVWKGELFRVVQVHRGLREVTGSQAVCRCSEELRALHM
              :: :... :. :...
ALLERG ALSCGTVNSNLAACIGYLTQNAPLPKGCCTGVTNLNNMARTTPDRQQACRCLVGAANSFP
      10      20      30      40      50      60

50
SYHT0H TIVQTFGNKVS

ALLERG SLNAARAAALPKACGVNIPYKISKSTNCNSVR
      70      80      90

>>ALLERGEN_2012_2|gi|119524036|gb|ABL77410.1| UA3-recogn (1096 aa)
  initn: 53 initl: 53 opt: 57 Z-score: 85.7 bits: 23.2 E(): 9.2
Smith-Waterman score: 57; 37.931% identity (37.931% ungapped) in 29 aa overlap
(13-41:1019-1047)

      10      20      30      40
SYHT0H      MWWVWKGELFRVVQVHRGLREVTGSQAVCRCSEELRALHMTI
              : : :... :: :... :
ALLERG VIGSCVLYHPPSQCDQWVQQCATALQTSAGVTVAGGYRQLSPPMVAVCVASQDLMTRCMT
      990      1000      1010      1020      1030      1040

50
SYHT0H VQTFGNKVS

ALLERG LGQGTCQQAVKNCKRRFNTFPSSRLPGRLWSLSSELINCLYRPVNRASN
      1050      1060      1070      1080      1090

51 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012

```

```

Scan time: 0.020 Display time: 0.010

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 81 aa
>SYHT0H2_I_15
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 4.2668+/-0.00397; mu= 8.6816+/- 0.211
mean_var=62.5117+/-17.766, 0's: 3 Z-trim: 3 B-trim: 200 in 1/43
Lambda= 0.1622

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.040
The best scores are:

```

					opt	bits	E(1603)
ALLERGEN_2012_2	gi 22684 emb CAA50325.1	major al	( 160)	74	25	0.76	
ALLERGEN_2012_2	gi 584968 sp Q08407.3 MPAC1_CORAV	( 160)	74	25	0.76		
ALLERGEN_2012_2	gi 1346568 sp P49372.1 ALL1_APIGR	( 154)	70	24	1.4		
ALLERGEN_2012_2	gi 291482314 emb CBK62697.1	ragw ( 96)	64	22	2.7		
ALLERGEN_2012_2	gi 291482316 emb CBK62698.1	ragw ( 110)	64	22	3		
ALLERGEN_2012_2	gi 291482318 emb CBK62699.1	ragw ( 116)	64	22	3.1		
ALLERGEN_2012_2	gi 291482310 emb CBK62695.1	ragw ( 134)	64	22	3.4		
ALLERGEN_2012_2	gi 291482306 emb CBK62693.1	ragw ( 111)	63	22	3.5		
ALLERGEN_2012_2	gi 291482308 emb CBK62694.1	ragw ( 140)	63	22	4.1		
ALLERGEN_2012_2	gi 8928058 sp O04298.1 DAU1_DAUCA	( 154)	63	22	4.4		
ALLERGEN_2012_2	gi 2154734 emb CAB03716.1	major ( 154)	63	22	4.4		
ALLERGEN_2012_2	gi 291197394 emb CBK52317.1	ragw ( 164)	63	22	4.6		
ALLERGEN_2012_2	gi 285005079 emb CBJ24286.1	ragw ( 164)	63	22	4.6		
ALLERGEN_2012_2	gi 30316292 sp Q9FSG7.1 TP1A_MALD	( 246)	63	22	6.1		
ALLERGEN_2012_2	gi 1545895 emb CAB02216.1	pollen ( 161)	61	22	6.2		
ALLERGEN_2012_2	gi 14422361 emb CAC41634.1	plant ( 131)	58	21	8.8		
ALLERGEN_2012_2	gi 14422363 emb CAC41635.1	plant ( 131)	58	21	8.8		
ALLERGEN_2012_2	gi 190613903 gb ACE80955.1	putat ( 242)	60	22	9.7		

```

>>ALLERGEN_2012_2|gi|22684|emb|CAA50325.1| major allerge (160 aa)
initn: 71 initl: 71 opt: 74 Z-score: 105.2 bits: 24.7 E(): 0.76
Smith-Waterman score: 74; 40.625% identity (41.935% ungapped) in 32 aa overlap (2-33:81-111)

              10          20          30
SYHT0H          MTFQQRVISGNLLGFHCPAICHFIVKIVEKE
              ...  ::  ::::  .  .::  . ::
ALLERG PGTIKNITFGEGRYKYVKERVDEVNTNFTYSYTVIEGDVLGDKLEKVCHEL-KIVAAP
              60          70          80          90          100

              40          50          60          70          80
SYHT0H GGSYKCHHCDCGKAIVEDASADSGPKDGPPTRSIVEKEDVPTTSSKQVD
              ::.
ALLERG GGSILKISSKFHAKGDHEINAEIKGAKEMAELLRAVETYLLAHSAEYN
              110          120          130          140          150          160

>>ALLERGEN_2012_2|gi|584968|sp|Q08407.3|MPAC1_CORAV RecN (160 aa)
initn: 71 initl: 71 opt: 74 Z-score: 105.2 bits: 24.7 E(): 0.76
Smith-Waterman score: 74; 40.625% identity (41.935% ungapped) in 32 aa overlap (2-33:81-111)

```

```

              10      20      30
SYHT0H      MTFQQRVISGNLLGFHCPAICHFIVKIVEKE
              :.. :: :...: . .:: . :::
ALLERG PGTIKNITFGEGRYKYVKERVDEVNTNFTYSYTVIEGDVLGDKLEKVCHEL-KIVAAP
              60      70      80      90      100

              40      50      60      70      80
SYHT0H GGSYKCHHCDKGKAIVEDASADSGPKDGPPTRSIVEKEDVPTTSSKQVD
              :..
ALLERG GGSILKISSKFHAKGDHEINAEEMKGAKEMAELLRVETVYLLAHSAEYN
              110     120     130     140     150     160

>>ALLERGEN_2012_2|gi|1346568|sp|P49372.1|ALL1_APIGR RecN (154 aa)
  initn: 60 initl: 60 opt: 70 Z-score: 100.4 bits: 23.8 E(): 1.4
Smith-Waterman score: 70; 35.185% identity (40.426% ungapped) in 54 aa overlap (1-
49:79-130)

```

```

              10      20
SYHT0H      MTFQQRVISGN-LLGFHCPAICHFIVKIVE
              :... :...: :...: .: . :.
ALLERG GPGLTKIITLPDGGPITMTLRIDGVNKEALTFDYSVIDGDILLGF-IESIENHVVLVPT
              50      60      70      80      90      100

              30      40      50      60      70      80
SYHT0H KEGGSYKCHHC---DKGKAIVEDASADSGPKDGPPTRSIVEKEDVPTTSSKQVD
              :... :. :. :...:
ALLERG ADGGS-CKTTAIFHTKGDVAVPEENIKYANEQNTALFKALEAYLIAN
              110     120     130     140     150

>>ALLERGEN_2012_2|gi|291482314|emb|CBK62697.1| ragweed h (96 aa)
  initn: 53 initl: 53 opt: 64 Z-score: 95.3 bits: 22.1 E(): 2.7
Smith-Waterman score: 64; 28.947% identity (29.730% ungapped) in 38 aa overlap
(26-62:24-61)

```

```

              10      20      30      40      50
SYHT0H MTFQQRVISGNLLGFHCPAICHFIVKIVEKEGGSY-KCHHCDKGKAIVEDASADSGPKDG
              . : :...: :...: . : . :
ALLERG KLCEKPSVTWSGNKVKQTDKCDKRCIEWEGAKHGACHKRDSKASCFCYFDCDPTKNPG
              10      20      30      40      50

              60      70      80
SYHT0H PPPTRRSIVEKEDVPTTSSKQVD
              :..
ALLERG PPPGAPKGKAPAPSPPSGGGAPPPSGGEGGGDGGGGRR
              60      70      80      90

>>ALLERGEN_2012_2|gi|291482316|emb|CBK62698.1| ragweed h (110 aa)
  initn: 53 initl: 53 opt: 64 Z-score: 94.6 bits: 22.2 E(): 3
Smith-Waterman score: 64; 28.947% identity (29.730% ungapped) in 38 aa overlap
(26-62:45-82)

```

```

              10      20      30      40      50
SYHT0H      MTFQQRVISGNLLGFHCPAICHFIVKIVEKEGGSY-KCHHCDKGKAIVEDASADS
              . : :...: :...: . : . :
ALLERG IASVKGKLCCKPSVTWSGKCKVKQTDKCDKRCIEWEGAKHGACHKRDSKASCFCYFDCDP
              20      30      40      50      60      70

              60      70      80
SYHT0H GPKDGPPTRSIVEKEDVPTTSSKQVD
              . :...:
ALLERG TKNPGPPGAPKGKAPAPSPPSGGGGEGGGEGGGER
              80      90      100     110

```





```

        60          70          80
SYHT0H PPPTRSIVEKEDVPTTSSKQVD
      :::
ALLERG PPPGAPKKGAPAPSPPPSGGGAPPPSGGEGGDGPPPPPEGEGGGDGGGEGGGDGGGEGGGG
        60          70          80          90         100         110

>>ALLERGEN_2012_2|gi|8928058|sp|O04298.1|DAU1_DAUCA RecN (154 aa)
  initn: 26 initl: 26 opt: 63 Z-score: 91.5 bits: 22.1 E(): 4.4
Smith-Waterman score: 63; 33.333% identity (38.298% ungapped) in 54 aa overlap (1-
49:79-130)

```

```

                                10          20
SYHT0H                        MTFQQRVISGN-LLGFHCPAICHFIVKIVE
                                ..... :...: :...: :...:
ALLERG GAGTVRIITLPEGSPITSMTVVRTDAVNKEALTYDSTVIDGDILLGFIESIETHLVV-VPT
        50          60          70          80          90         100

        30          40          50          60          70          80
SYHT0H KEGGSY----KCHHCDKKGKAIVEDASADSGPKDGPPTTRSIVEKEDVPTTSSKQVD
      ..... : : : : :
ALLERG ADGGSITKTTAIFHT-KGDAVVPEENIKFADAQNTALFKAIEAYLIAN
        110         120         130         140         150

```

```

>>ALLERGEN_2012_2|gi|2154734|emb|CAB03716.1| major aller (154 aa)
  initn: 26 initl: 26 opt: 63 Z-score: 91.5 bits: 22.1 E(): 4.4
Smith-Waterman score: 63; 33.333% identity (38.298% ungapped) in 54 aa overlap (1-
49:79-130)

```

```

                                10          20
SYHT0H                        MTFQQRVISGN-LLGFHCPAICHFIVKIVE
                                ..... :...: :...: :...:
ALLERG GAGTVRIITLPEGSPITSMTVVRTDAVNKEALTYDSTVIDGDILLGFIESIETHLVV-VPT
        50          60          70          80          90         100

        30          40          50          60          70          80
SYHT0H KEGGSY----KCHHCDKKGKAIVEDASADSGPKDGPPTTRSIVEKEDVPTTSSKQVD
      ..... : : : : :
ALLERG ADGGSITKTTAIFHT-KGDAVVPEENIKFADEQNTALFKAIEAYLIAN
        110         120         130         140         150

```

```

>>ALLERGEN_2012_2|gi|291197394|emb|CBK52317.1| ragweed h (164 aa)
  initn: 52 initl: 52 opt: 63 Z-score: 91.2 bits: 22.1 E(): 4.6
Smith-Waterman score: 63; 28.947% identity (29.730% ungapped) in 38 aa overlap
(26-62:45-82)

```

```

                10          20          30          40          50
SYHT0H      MTFQQRVISGNLLGFHCPAICHFIVKIVEKEGGSY-KCHHCDKKGKAIVEDASADS
                . : : : : : : : : : :
ALLERG IASVKGKLCEKPSVTWSGKCKVKQTDKCDKRCIEWEGAKHGACHKRDSKATCFYFDCDP
        20          30          40          50          60          70

        60          70          80
SYHT0H GPKDGPPTTRSIVEKEDVPTTSSKQVD
      . :...:
ALLERG TKNPGPPPGAPKKGAPAPSPPPSGGGAPPPSGGEGGDGPPPPPEGEGGGDGGGEGGGEGGGG
        80          90         100         110         120         130

```

```

>>ALLERGEN_2012_2|gi|285005079|emb|CBJ24286.1| ragweed h (164 aa)
  initn: 52 initl: 52 opt: 63 Z-score: 91.2 bits: 22.1 E(): 4.6
Smith-Waterman score: 63; 28.947% identity (29.730% ungapped) in 38 aa overlap
(26-62:45-82)

```

```

                10          20          30          40          50

```

SYHT0H MTFQQRVISGNLLGFHCPAICHFIVKIVEKEGGSY-KCHHCDKGKAIVEDASADS  
ALLERG IASVKGKLCCKPSLTWSGKCKVKQTDKCDKRCIEWEGAKHGACHKRDSKATCFYFDCDP

```
>>ALLERGEN_2012_2|gi|30316292|sp|Q9FSG7.1|TP1A_MALDO Rec (246 aa)
  initn: 36 init1: 36 opt: 63 Z-score: 89.0 bits: 22.3 E(): 6.1
Smith-Waterman score: 63; 21.127% identity (23.077% ungapped) in 71 aa overlap
(17-81:156-226)
```

ALLERG P

SYHT0H MTFQQRVISGNLLGFHCPAICHFIVKIVEKEG  
.. :: :...: . :.: : .. :.  
ALLERG GTIKKITFSEGSVPKYVKERVEEVDHTNFKYSYTVIEGGFVGDKVEIKNEIKIVAAPDG

60 70 80 90 100 110

```
>>ALLERGEN_2012_2|gi|14422361|emb|CAC41634.1| plantain p (131 aa)
  initn: 27 initl: 27 opt: 58 Z-score: 86.1 bits: 20.9 E(): 8.8
Smith-Waterman score: 58; 25.000% identity (28.571% ungapped) in 48 aa overlap
(16-59:16-61)
```



```

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.070
The best scores are:
ALLERGEN_2012_2|gi|5059162|gb|AAD38942.1|AF144060 ( 496) 87 29 0.25
ALLERGEN_2012_2|gi|21748153|emb|CAD38167.1| putat ( 124) 73 25 0.99
ALLERGEN_2012_2|gi|89892723|gb|ABD79095.1| Zea m ( 252) 71 25 2.2
ALLERGEN_2012_2|gi|89892721|gb|ABD79094.1| Zea m ( 263) 71 25 2.3
ALLERGEN_2012_2|gi|115502167|sp|Q1ZYQ8.2|EXB10_MA ( 270) 71 25 2.3
ALLERGEN_2012_2|gi|105969545|gb|ABF81662.1| EXPB1 ( 269) 69 25 3.2
ALLERGEN_2012_2|gi|115502168|sp|POC1Y5.1|EXB11_MA ( 269) 69 25 3.2
ALLERGEN_2012_2|gi|18615|emb|CAA26723.1| unnamed ( 495) 71 25 3.5
ALLERGEN_2012_2|gi|18635|emb|CAA33215.1| glycinin ( 495) 69 25 4.9
ALLERGEN_2012_2|gi|18652047|gb|AAL76932.1|AF45648 ( 154) 64 23 5.1
ALLERGEN_2012_2|gi|2266625|emb|CAB10765.1| group ( 264) 63 23 8.6

>>ALLERGEN_2012_2|gi|5059162|gb|AAD38942.1|AF144060_1 al (496 aa)
initn: 50 initl: 50 opt: 87 Z-score: 113.9 bits: 29.1 E(): 0.25
Smith-Waterman score: 92; 26.364% identity (31.868% ungapped) in 110 aa overlap
(23-132:178-268)

          10      20      30      40      50
SYHT0H      MSPERRPVEIRPATAADMAAVCDIVNHYIETSTVNFRTPEQTPQEWIDDLER
              : : : : : : : : : : : : : : : : : : : : : : : : : :
ALLERG YTNPKEARNCRLSGLRDLKQQSEYVRQKQVDFLNHLIDIGVAGFRSDASTHQ-WPDDLRS
          150      160      170      180      190      200

          60      70      80      90      100      110
SYHT0H LQDRYPWLVAEVEGVVAGIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLK
          . . : : : : : : : : : : : : : : : : : : : : : : : :
ALLERG IYSRLHNLNKEFF-----P---ENSQPFYIYHETIYYGGN----GINSNEYTSLGR
          210      220      230      240

          120      130      140      150      160      170
SYHT0H SMEAQGFKSVVAVIGLPNDPSVRLHEALGYTARGLRAAGYKHGGWHDVGFWRQDFELPA
          . : . : . : . : : : : : : : : : : : : : : : : : : :
ALLERG IIEFRFYKEITNVFR-GNNPLHWLKNFGTEWGLVPSGDALVMIDSHDLRVGHTGKLGFN
          250      260      270      280      290      300

>>ALLERGEN_2012_2|gi|21748153|emb|CAD38167.1| putative n (124 aa)
initn: 70 initl: 70 opt: 73 Z-score: 103.1 bits: 25.1 E(): 0.99
Smith-Waterman score: 73; 33.333% identity (34.146% ungapped) in 42 aa overlap
(92-132:16-57)

          70      80      90      100      110      120
SYHT0H AEVEGVVAGIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKS
              : . : : : : : : : : : : : : : : : : : : :
ALLERG      MSDFNAIAQQFVEFYKTFDGNRAGLGALYKEHSMLTFEAQGTQG
              10      20      30      40

          130      140      150      160      170      180
SYHT0H VVAVI-GLPNDPSVRLHEALGYTARGLRAAGYKHGGWHDVGFWRQDFELPAPPRPVRPV
          . : . : : : : : : : : : : : : : : : : : : :
ALLERG SAAIVEKLQNLFPQEIQHRTDTVDAQPSADDGILVLVTGALLLGGESKPMSFTQAFQLKN
          50      60      70      80      90      100

>>ALLERGEN_2012_2|gi|89892723|gb|ABD79095.1| Zea m 1 all (252 aa)
initn: 71 initl: 71 opt: 71 Z-score: 96.8 bits: 25.0 E(): 2.2
Smith-Waterman score: 71; 30.189% identity (30.189% ungapped) in 53 aa overlap
(100-152:81-133)

```

```

      70      80      90      100      110      120
SYHT0H GIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKSVVAVIGLP
      :: :: . . : .: .: :: . .
ALLERG NGGGCGYKDVNKPFFNSMGACGNIPFKDGLGCGSCFEIKCDKPVVECSGKPVVVHITDMN
      60      70      80      90      100      110

      130      140      150      160      170      180
SYHT0H NDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFWQ RDFELPAPPRPVRPVTQI
      .: . : :. :: :. :
ALLERG YEPIAAYHFDLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKYDSKVTTFHLEKGCGPN
      120      130      140      150      160      170

>>ALLERGEN_2012_2|gi|89892721|gb|ABD79094.1| Zea m 1 all (263 aa)
  initn: 71 init1: 71 opt: 71 Z-score: 96.6 bits: 25.0 E(): 2.3
Smith-Waterman score: 71; 30.189% identity (30.189% ungapped) in 53 aa overlap
(100-152:92-144)

```

```

      70      80      90      100      110      120
SYHT0H GIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKSVVAVIGLP
      :: :: . . : .: .: :: . .
ALLERG NGGGCGYKDVNKPFFNSMGACGNIPFKDGLGCGSCFEIKCDKPVVECSGKPVVVHITDMN
      70      80      90      100      110      120

      130      140      150      160      170      180
SYHT0H NDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFWQ RDFELPAPPRPVRPVTQI
      .: . : :. :: :. :
ALLERG YEPIAAYHFDLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKFGSKVSFHFLEKGCGPN
      130      140      150      160      170      180

>>ALLERGEN_2012_2|gi|115502167|sp|Q1ZYQ8.2|EXB10_MAIZE R (270 aa)
  initn: 71 init1: 71 opt: 71 Z-score: 96.4 bits: 25.0 E(): 2.3
Smith-Waterman score: 71; 30.189% identity (30.189% ungapped) in 53 aa overlap
(100-152:99-151)

```

```

      70      80      90      100      110      120
SYHT0H GIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKSVVAVIGLP
      :: :: . . : .: .: :: . .
ALLERG NGGGCGYKDVNKPFFNSMGACGNIPFKDGLGCGSCFEIKCDKPVVECSGKPVVVHITDMN
      70      80      90      100      110      120

      130      140      150      160      170      180
SYHT0H NDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFWQ RDFELPAPPRPVRPVTQI
      .: . : :. :: :. :
ALLERG YEPIAAYHFDLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKYDSKVTTFHLEKGCGPN
      130      140      150      160      170      180

>>ALLERGEN_2012_2|gi|105969545|gb|ABF81662.1| EXPB10 [Ze (269 aa)
  initn: 94 init1: 69 opt: 69 Z-score: 93.9 bits: 24.5 E(): 3.2
Smith-Waterman score: 69; 30.189% identity (30.189% ungapped) in 53 aa overlap
(100-152:98-150)

```

```

      70      80      90      100      110      120
SYHT0H GIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKSVVAVIGLP
      :: :: . . : .: .: :: . .
ALLERG NGGGCGYKDVNKAPFNSMGACGNVPIFKDGLGCGSCFEIKCDKPAECSGKPVVVYITDMN
      70      80      90      100      110      120

      130      140      150      160      170      180
SYHT0H NDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFWQ RDFELPAPPRPVRPVTQI
      .: . : :. :: :. :
ALLERG YEPIAAYHFDLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKYGSKVTTFHLEKGCPNP
      130      140      150      160      170      180

```

>>ALLERGEN\_2012\_2|gi|115502168|sp|P0C1Y5.1|EXB11\_MAIZE R (269 aa)  
 initn: 94 initl: 69 opt: 69 Z-score: 93.9 bits: 24.5 E(): 3.2  
 Smith-Waterman score: 69; 30.189% identity (30.189% ungapped) in 53 aa overlap  
 (100-152:98-150)

```

      70      80      90      100      110      120
SYHT0H GIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKSVVAVIGLP
      :: :: . . : : .: :: . .
ALLERG NGGGCGYKDVNKAPFNSMGACGNVPIFKDGLGCGSCFEIKCDKPAECSGKPVVYITDMN
      70      80      90      100      110      120

      130      140      150      160      170      180
SYHT0H NDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRPVTQI
      .: . : :. :: :. . :
ALLERG YEPIAAYHFDLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKYGSKVTFHLEKGCNPN
      130      140      150      160      170      180

```

>>ALLERGEN\_2012\_2|gi|18615|emb|CAA26723.1| unnamed prote (495 aa)  
 initn: 41 initl: 41 opt: 71 Z-score: 93.2 bits: 25.3 E(): 3.5  
 Smith-Waterman score: 71; 25.843% identity (29.114% ungapped) in 89 aa overlap  
 (39-118:113-200)

```

      10      20      30      40      50      60
SYHT0H EIRPATAADMAAVCDIVNHYIETSTVNFRTEPQTPQEWIDDLERLQDRYPWLVAEVEGVV
      ::: :: . . : :: . . : :: .
ALLERG SYTNGPQEIYIQQKGIFGMIYPGCSSTFEEPQQPQQRGQS-SRPQDRHQKIYNSREGDL
      90      100      110      120      130      140

      70      80      90      100      110
SYHT0H AGIAYAGPWKARNAYDWTVESTVYV---SHRHQ-----RLGLGSTLYTHLLKSMEAQGF
      .. . : : : :. . : :. : :. : :. : :. : :. : :. : :. : :
ALLERG IAVPTGVAWWMYNNEDTPVVAVSIIDTNSLENQLDQMPRRFYLAGNQEQEFLKYQQEQGG
      150      160      170      180      190      200

      120      130      140      150      160      170
SYHT0H KSVVAVIGLPNDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRP
ALLERG HQSQKGKHQQEEENEGGSILSGFTLEFLEHAFSVDKQIAKNLQGENEGEDKGAIVTVKGG
      210      220      230      240      250      260

```

>>ALLERGEN\_2012\_2|gi|18635|emb|CAA33215.1| glycinin subu (495 aa)  
 initn: 41 initl: 41 opt: 69 Z-score: 90.7 bits: 24.8 E(): 4.9  
 Smith-Waterman score: 69; 25.843% identity (29.114% ungapped) in 89 aa overlap  
 (39-118:113-200)

```

      10      20      30      40      50      60
SYHT0H EIRPATAADMAAVCDIVNHYIETSTVNFRTEPQTPQEWIDDLERLQDRYPWLVAEVEGVV
      ::: :: . . : :: . . : :: .
ALLERG SYTNGPQEIYIQQKGIFGMIYPGCPSTFEEPQQPQQRGQS-SRPQDRHQKIYNFREGDL
      90      100      110      120      130      140

      70      80      90      100      110
SYHT0H AGIAYAGPWKARNAYDWTVESTVYV---SHRHQ-----RLGLGSTLYTHLLKSMEAQGF
      .. . : : : :. . : :. : :. : :. : :. : :. : :. : :. : :
ALLERG IAVPTGVAWWMYNNEDTPVVAVSIIDTNSLENQLDQMPRRFYLAGNQEQEFLKYQQEQGG
      150      160      170      180      190      200

      120      130      140      150      160      170
SYHT0H KSVVAVIGLPNDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRP
ALLERG HQSQKGKHQQEEENEGGSILSGFTLEFLEHAFSVDKQIAKNLQGENEGEDKGAIVTVKGG
      210      220      230      240      250      260

```

```
>>ALLERGEN_2012_2|gi|18652047|gb|AAL76932.1|AF456481_1 m (154 aa)
  initn: 33 initl: 33 opt: 64 Z-score: 90.3 bits: 23.1 E(): 5.1
Smith-Waterman score: 64; 22.727% identity (25.773% ungapped) in 110 aa overlap
(44-146:21-124)
```

```

                20          30          40          50          60          70
SYHT0H TAADMAAVCDIVNHYIETSTVNFRTPEQTPQEWIDDLERLQDRY-PWLVAEVEGVVAGIA
                : .. :.. . . : :. :: .. : .
ALLERG          MGVMQKTEVEAPSTVSAEKMYQGFLDMDTVFPKVLPLQLIKSVE-ILEGDG
                10          20          30          40
```

```

                80          90          100          110          120
SYHT0H YAGPWKARNAYDWTVESTVYVSHRHQ-----RLGLGSTLYTHLLKSMEAQGFKSVVA-VI
                .: . . . :.: :.. ... . :.: : : : . .. :.:.:.: : .
ALLERG GVGTVRLVHLG---EATEYTTMKQKVDVIDKAGLGYT-YTTIGGDILVEGLESVVNQFV
                50          60          70          80          90          100
```

```

                130          140          150          160          170          180
SYHT0H GLPNDPVRLHEALGYTARGTLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRPVTQI
                .:.: . . .... :.:.:
ALLERG VVPTDGGCIVKNNTTIYNTKGDAVLPEDKVKEATEKSALAFKAVEAYLLAN
                110          120          130          140          150
```

```
>>ALLERGEN_2012_2|gi|2266625|emb|CAB10765.1| group V all (264 aa)
  initn: 34 initl: 34 opt: 63 Z-score: 86.2 bits: 23.1 E(): 8.6
Smith-Waterman score: 63; 28.814% identity (30.357% ungapped) in 59 aa overlap
(101-157:17-74)
```

```

                80          90          100          110          120
SYHT0H IAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKSVVAVIGL--
                : ..: .:~::~~: :~::~~: .
ALLERG          ADAGYTPAAPAAAGAGGKATTDEQKLLLEDVNA-GFKTAVAAAAANVP
                10          20          30          40
```

```

                130          140          150          160          170          180
SYHT0H PNDPVRLHEALGYTARGTLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRPVTQI
                : : .. :. .... :.: : :
ALLERG PADKYKTFEAAFTASSKASIAAAATKAPGLIPQLNAATNTAYAAAQGATPEAKYDAFVTT
                50          60          70          80          90          100
```

```
183 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.070 Display time: 0.010
```

```
Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448
```

```
/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 35 aa
>SYHT0H2_I_17
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library
```

```
367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 3.8679+/-0.00302; mu= 3.9405+/- 0.158
mean_var=35.9280+/-10.523, 0's: 13 Z-trim: 13 B-trim: 0 in 0/43
Lambda= 0.2140
```



```

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
  join: 36, opt: 24, gap-pen: -12/-2, width: 16
  Scan time: 0.010
The best scores are:
ALLERGEN_2012_2|gi|3021324|emb|CAA06305.1| Aspf1 ( 125) 51 22 2.1
ALLERGEN_2012_2|gi|9280360|gb|AAF86369.1| major a ( 150) 51 22 2.4
ALLERGEN_2012_2|gi|10189811|emb|CAC09234.1| unnam ( 215) 52 22 2.6
ALLERGEN_2012_2|gi|1352240|sp|P49273.1|ALL7_DERPT ( 215) 52 22 2.6
ALLERGEN_2012_2|gi|54039254|sp|P67875.1|RNMG_ASPF ( 176) 51 22 2.8
ALLERGEN_2012_2|gi|171464770|gb|ACB45874.1| patho ( 151) 47 21 5.7
ALLERGEN_2012_2|gi|46410859|gb|AAR98518.1| major ( 366) 50 22 6.2

>>ALLERGEN_2012_2|gi|3021324|emb|CAA06305.1| Aspf1 aller (125 aa)
  initn: 51 init1: 51 opt: 51 Z-score: 97.4 bits: 21.7 E(): 2.1
Smith-Waterman score: 51; 33.333% identity (33.333% ungapped) in 33 aa overlap (2-34:1-33)

          10          20          30
SYHT0H MRLTKVRSAAFNQYQSGPYHDGKTVQSVWLFLTNNK
      :: . . : . . . . : :::: . : .:::
ALLERG RLVYNQAKAESNSHHAPLSDGKTGSSYPHWFTNGYDGDGKLIKGRMPIKFGKADCDRPP
          10          20          30          40          50

ALLERG KHGKDGMMGKDDHYLLEFPTFPDGHDKYKFDSSKPKEDPGPARVIYTPNKKVFCGIVAHERG
          60          70          80          90         100         110

>>ALLERGEN_2012_2|gi|9280360|gb|AAF86369.1| major allerg (150 aa)
  initn: 51 init1: 51 opt: 51 Z-score: 96.2 bits: 21.7 E(): 2.4
Smith-Waterman score: 51; 33.333% identity (33.333% ungapped) in 33 aa overlap (2-34:21-53)

          10          20          30
SYHT0H MRLTKVRSAAFNQYQSGPYHDGKTVQSVWLFLTNNK
      :: . . : . . . . : :::: . : .:::
ALLERG MTWTCINQQLNPKNKWKEDKRLLYNQAKAESNSHHAPLSDGKTGSSYAHWFTNGYDGNKG
          10          20          30          40          50          60

ALLERG LIKGRTPIKFGKADCDRPPKHSQNGMGKDDHYLLEFPTFPDGHDKYKFDSSKNKPKEDPGPA
          70          80          90         100         110         120

>>ALLERGEN_2012_2|gi|10189811|emb|CAC09234.1| unnamed pr (215 aa)
  initn: 32 init1: 32 opt: 52 Z-score: 95.5 bits: 22.1 E(): 2.6
Smith-Waterman score: 52; 28.571% identity (28.571% ungapped) in 28 aa overlap (1-28:2-29)

          10          20          30
SYHT0H MRLTKVRSAAFNQYQSGPYHDGKTVQSVWLFLTNNK
      :: . .::: . . : : : . . .
ALLERG MMKLLLIAAAFAVVSADPIHYDKITEEINKAVDEAVAAIEKSETFDPMKVPDHSKDKFER
          10          20          30          40          50          60

>>ALLERGEN_2012_2|gi|1352240|sp|P49273.1|ALL7_DERPT RecN (215 aa)
  initn: 32 init1: 32 opt: 52 Z-score: 95.5 bits: 22.1 E(): 2.6
Smith-Waterman score: 52; 28.571% identity (28.571% ungapped) in 28 aa overlap (1-28:2-29)

          10          20          30
SYHT0H MRLTKVRSAAFNQYQSGPYHDGKTVQSVWLFLTNNK
      :: . .::: . . : : : . . .
ALLERG MMKLLLIAAAFAVVSADPIHYDKITEEINKAVDEAVAAIEKSETFDPMKVPDHSKDKFER
          10          20          30          40          50          60

```

```
>>ALLERGEN_2012_2|gi|54039254|sp|P67875.1|RNMG_ASPFU Rec (176 aa)
  initn: 51 initl: 51 opt: 51 Z-score: 95.1 bits: 21.8 E(): 2.8
Smith-Waterman score: 51; 33.333% identity (33.333% ungapped) in 33 aa overlap (2-34:48-80)
```

```

                                10      20      30
SYHT0H                        MRLTKVRSAAFNQYQSGPYHDGKTVQSVWLF
                                ::  .. :  . . . :  :::: :  .
ALLERG LAAPSPLDARATWTCINQQLNPKTNKWEDKRLLYSQAKAESNSHHAPLSDGKTGSSYPHW
                                20      30      40      50      60      70
```

```

SYHT0H LTNK
      .::
ALLERG FTNGYDGNGKLIKGRTPIKFGKADCDRPPKHSQNGMGKDDHYLLEFPTFPDGHDKYKFDSK
      80      90     100     110     120     130
```

```
>>ALLERGEN_2012_2|gi|171464770|gb|ACB45874.1| pathogen-r (151 aa)
  initn: 47 initl: 47 opt: 47 Z-score: 89.5 bits: 20.5 E(): 5.7
Smith-Waterman score: 47; 50.000% identity (50.000% ungapped) in 18 aa overlap (8-25:19-36)
```

```

                                10      20      30
SYHT0H                        MRLTKVRSAAFNQYQSGPYHDGKTVQSVWLF LTNK
                                .:: :  :: :  .::
ALLERG MLPFSFAQDSIKDFVDAHNAARAQVGVPVHWNKTVADYAHQYANKRIKDCNLVHSGGPY
                                10      20      30      40      50      60
ALLERG GENIAWGSRNLAGTVAVRMWVSEKQFYNYDTNSCVRGKMCGHYTQVVWRNSVRIGCAKVR
                                70      80      90     100     110     120
```

```
>>ALLERGEN_2012_2|gi|46410859|gb|AAR98518.1| major latex (366 aa)
  initn: 50 initl: 50 opt: 50 Z-score: 88.8 bits: 21.6 E(): 6.2
Smith-Waterman score: 50; 46.667% identity (46.667% ungapped) in 15 aa overlap (11-25:63-77)
```

```

                                10      20      30
SYHT0H                        MRLTKVRSAAFNQYQSGPYHDGKTVQSVWLF LTNK
                                :  .:  .:  . ::::
ALLERG FGDGLYDAGNAKFIYPDKYLPSYHHYPGTTFDFDYPTGRFSDGRTTVDFVAENVSLPRIPP
                                40      50      60      70      80      90
ALLERG FKNKEANFTYGANFASEGATASDSNPLIDFRSQIRDFGELKLEWAVQLVNVTELARRLKK
                                100     110     120     130     140     150
```

```
35 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.010 Display time: 0.010
```

```
Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448
```

```
/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 64 aa
>SYHT0H2_I_18
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library
```

```

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 4.5451+/-0.00322; mu= 4.1974+/- 0.167
mean_var=55.6035+/-16.017, 0's: 7 Z-trim: 7 B-trim: 7 in 1/41
Lambda= 0.1720

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.020
The best scores are:
ALLERGEN_2012_2|gi|71360928|emb|CAJ19705.1| non-s ( 114) 59 22 2.9
ALLERGEN_2012_2|gi|21743|emb|CAA43331.1| high mol ( 830) 64 24 5.9
ALLERGEN_2012_2|gi|13183175|gb|AAK15088.1|AF24000 ( 153) 55 21 7.4
ALLERGEN_2012_2|gi|209165427|gb|ACI41244.1| 2S al ( 153) 55 21 7.4
ALLERGEN_2012_2|gi|166235350|pdb|2JON|A Chain A, ( 101) 53 20 7.5
ALLERGEN_2012_2|gi|886967|emb|CAA59340.1| low mol ( 276) 57 22 8.3
ALLERGEN_2012_2|gi|736319|emb|CAA27052.1| gluteni ( 838) 62 23 8.4
ALLERGEN_2012_2|gi|170743|gb|AAB02788.1| HMW glut ( 815) 61 23 9.7

>>ALLERGEN_2012_2|gi|71360928|emb|CAJ19705.1| non-specif (114 aa)
initn: 38 initl: 38 opt: 59 Z-score: 94.6 bits: 21.9 E(): 2.9
Smith-Waterman score: 59; 26.471% identity (28.125% ungapped) in 34 aa overlap
(12-43:27-60)

          10          20          30          40
SYHT0H      MRLWDTQPGVHCAQLDTSMDGMMLVFGKGIL--SCQLLQGQLGQ
          ::. .... . ....: : :. ... :::
ALLERG MEMVNKIACFVLLCMVVVAPHAEALTCGQVTSTLAPCLPYLMNRGPLRNCCDGVKGLLGQ
          10          20          30          40          50          60

          50          60
SYHT0H LPRYESSRSPNFPDRSNIWQ

ALLERG AKTTVDRQAACCTCLKSAASSFTGLNLGKAAALPNTCSVNIPYKISPSTDCSKVQ
          70          80          90          100          110

>>ALLERGEN_2012_2|gi|21743|emb|CAA43331.1| high molecula (830 aa)
initn: 261 initl: 57 opt: 64 Z-score: 89.2 bits: 23.8 E(): 5.9
Smith-Waterman score: 64; 35.135% identity (35.135% ungapped) in 37 aa overlap
(28-64:435-471)

          10          20          30          40          50
SYHT0H      MRLWDTQPGVHCAQLDTSMDGMMLVFGKGILSCQLLQGQLGQLPRYESSRSPNFP
          ::. . : ::: :: : . . : :
ALLERG GQEQQPRQLQQPEQQGQQGQQPEQQGQQGQQPGQEQGQQPGQGGQQGQPGQPGYYPTSP
          410          420          430          440          450          460

          60
SYHT0H DRSNIWQ
          ... :
ALLERG QQSGQQPGYYPTSPQQSGQLQQPAQGGQPGQEQGQQPGQGGQQPGQGGQQPGQGGQPG
          470          480          490          500          510          520

>>ALLERGEN_2012_2|gi|13183175|gb|AAK15088.1|AF240005_1 2 (153 aa)
initn: 46 initl: 46 opt: 55 Z-score: 87.5 bits: 21.0 E(): 7.4
Smith-Waterman score: 55; 52.174% identity (57.143% ungapped) in 23 aa overlap
(34-54:44-66)

          10          20          30          40          50          60
SYHT0H WDTQPGVHCAQLDTSMDGMMLVFGKGILSC-QLLQG-QLGQLPRYESSRSPNFPDRSN
          : : :: : . . : : :: :
ALLERG AMVALASATYTTTTVTTTAIDDEANQQSQCRQLQGRQFRSCQRYLSQGRSPYGGEED
          20          30          40          50          60          70

```

ALLERG VLEMSTGNQQSEQSLRDCCQQLRNVDRCRCEAIRQAVRQQQQEGGYEGQSQQVYQRR

SYHT0H WDTQPGVHCAQLDTSMDGMMLVFGKGILSC-QLLQG-QLGQLPRYESSRSPNFPDRSN  
ALLERG AMVALASATYTTTTVTTTAIDDEANQQSQQCROQLGRQFRSCQRYLSQLGRSPYGGEED

ALLERG VLEMSTGNQQSEQSLRDCCQQLGNVDERCRCEAIRQAVRQQQEGGYEGQSQQVYQRR

SYHT0H MRLWDTQPGVHCAQLDTSMDVG--MMLVF---GKG  
::: : . : . : : : :  
ALLERG WCVPKPGVSDDQLTGNINYACSGIDCGPIQPGGACFEPNTVKAHAAAYVMNLYYQHAGRN  
20 30 40 50 60 70

```
>>ALLERGEN_2012_2|gi|886967|emb|CAA59340.1| low molecula (276 aa)
  initn: 38 initl: 38 opt: 57 Z-score: 86.6 bits: 21.7 E(): 8.3
Smith-Waterman score: 57; 47.619% identity (50.000% ungapped) in 21 aa overlap
(33-52:199-219)
```

SYHT0H LWDTPQGVHCAQLDTSMDGMLLVFGKGILSCQLLQGQL-GQLPRYESSRSPNFPDRSN  
ALLERG LNPKCVFLQQCSPVAMPQSLARSQMLWQSSCHVMQQCCQQLPRIPEQSRYDAIRAIYY

10 20 30 40 50 60  
170 180 190 200 210 220

ALLERG SIVLQEQQHGGFNQPQQQPPQQSVQGVSQPQQQKQLGQCSFQQPQQ  
230 240 250 260 270

```
>>ALLERGEN_2012_2|gi|736319|emb|CAA27052.1| glutenin [Tr (838 aa)
  initn: 245 init1: 53 opt: 62 Z-score: 86.5 bits: 23.3 E(): 8.4
Smith-Waterman score: 72; 31.667% identity (32.759% ungapped) in 60 aa overlap (7-
64:454-513)
```

```

              10      20      30
SYHT0H      MRLWDTQPG--VHCAQLDTSMDGMMLVFGKILSC
              ::: . :: . . . :. :. .
ALLERG PGYYPTSPLQSGQGPGYYLTSPQQSGQGQPGQLQSSAQGQKGGQPGQGQPGQGQGGQ
      430      440      450      460      470      480

```

```

      40      50      60
SYHT0H QLLQGQLGQLPRYESSRSPNFPDRSNIWQ
      : : : : : . . :. :. : :
ALLERG QPGQGQGGQPGQGQPGYYPTSPLQSGQGQPGQWQPGQGQPGYYPTSPLQPGQGQPGY
      490      500      510      520      530      540

```

```

>>ALLERGEN_2012_2|gi|170743|gb|AAB02788.1| HMW glutenin (815 aa)
  initn: 241 initl: 57 opt: 61 Z-score: 85.3 bits: 23.0 E(): 9.7
Smith-Waterman score: 61; 26.562% identity (29.310% ungapped) in 64 aa overlap (7-
64:609-672)

```

```

              10      20      30
SYHT0H      MRLWDTQPGVHCAQLDTS---MVDGMMLVFGKG---
              ::: . . :. : . :. :. :
ALLERG QPGQGQGGQPGQGQPGQGQPGQGQPGYYPTSPLQSGQGQPGQWQPGQGQPGQ
      580      590      600      610      620      630

```

```

      40      50      60
SYHT0H ILSCQLLQGQLGQLPRYESSRSPNFPDRSNIWQ
      . : : . : : : :. :. :. :. :
ALLERG YYPTSSLQPEQGQGGYYPTSQQQPGQPGQWQSGQGQGGYYPTSPLQSGQGQPGQW
      640      650      660      670      680      690

```

```

ALLERG LQPGQWLQSGYYLTSPQQLGQGQPPQWLQPRQGQGGYYPTSPLQSGQGQGLQGQGGQY
      700      710      720      730      740      750

```

```

64 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.020 Display time: 0.010

```

```

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 57 aa
>SYHT0H2_I_19
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

```

```

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 3.5776+/-0.00299; mu= 8.1422+/- 0.157
mean_var=41.7344+/-11.088, 0's: 12 Z-trim: 13 B-trim: 0 in 0/43
Lambda= 0.1985

```

```

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.020

```

```

The best scores are:
ALLERGEN_2012_2|gi|6136162|sp|P35776.2|VA2_SOLRI ( 119) 60 24 0.91
ALLERGEN_2012_2|gi|2497701|sp|Q28133.1|ALL2_BOVIN ( 172) 55 22 3.2
ALLERGEN_2012_2|gi|886965|emb|CAA59339.1| low mol ( 261) 56 23 3.5

```

```

ALLERGEN_2012_2|gi|94400907|ref|NP_001035360.1| a ( 92) 52 21 3.7
ALLERGEN_2012_2|gi|88770352|gb|ABD51779.1| allerg ( 94) 52 21 3.8
ALLERGEN_2012_2|gi|170730|gb|AAA34285.1| gamma-gl ( 304) 55 22 4.8
ALLERGEN_2012_2|gi|194350817|gb|ACF53837.1| Bla g ( 190) 50 21 9.2
ALLERGEN_2012_2|gi|886967|emb|CAA59340.1| low mol ( 276) 51 21 9.8

>>ALLERGEN_2012_2|gi|6136162|sp|P35776.2|VA2_SOLRI RecNa (119 aa)
  initn: 60 initl: 60 opt: 60 Z-score: 103.8 bits: 23.5 E(): 0.91
Smith-Waterman score: 60; 36.000% identity (36.000% ungapped) in 25 aa overlap (7-
31:12-36)

          10          20          30          40          50
SYHT0H      MPSTMLVSSCAQCTPGCVSQSLMQPNRWIVWKAYNSNHRLKTLRLHRLKQMCVQC
          .. :. : :. : : :. : :. : :. :
ALLERG DIEAQRVLRKDIACARTLPKCVNQPDPLARVDVWHCAMS KRGVYDNPDP AVVKEKNSK
          10          20          30          40          50          60

SYHT0H GS

ALLERG MCPKIITDPADVENCKKVSRCDRETQRPRSNRQKAINITGCILRAGVVEATVLAREK
          70          80          90         100         110

>>ALLERGEN_2012_2|gi|2497701|sp|Q28133.1|ALL2_BOVIN RecN (172 aa)
  initn: 29 initl: 29 opt: 55 Z-score: 94.0 bits: 22.2 E(): 3.2
Smith-Waterman score: 55; 27.778% identity (31.915% ungapped) in 54 aa overlap
(10-57:14-66)

          10          20          30          40          50
SYHT0H      MPSTMLVSSC-AQCTPGCVSQSLMQPNRW-IVWKAYNSNHRL---KTLRLHRLKQM
          : : :. :. :. :. :. :. :. :. :. :. :. :. :. :. :. :. :. :. :. :
ALLERG MKAVFLTLLFGLVCTAQETPAEIDPSKI-PGEWRRIYAAADNKDKIVEGGPLRNYRRIE
          10          20          30          40          50

SYHT0H CVQ-CGS
          :. : :
ALLERG CINDCESLSITFYLKDQGTCLLLTEVAKRQEGYVYVLEFYGTNTLEVIHVS ENMLVTYVE
          60          70          80          90         100         110

>>ALLERGEN_2012_2|gi|886965|emb|CAA59339.1| low molecula (261 aa)
  initn: 44 initl: 44 opt: 56 Z-score: 93.3 bits: 22.7 E(): 3.5
Smith-Waterman score: 56; 27.500% identity (30.556% ungapped) in 40 aa overlap
(12-51:164-199)

          10          20          30          40
SYHT0H      MPSTMLVSSCAQCTPGCVSQSLMQPNRWIVWKAYNSNHRLK
          :. :. :. :. :. :. :. :. :. :. :. :. :. :. :. :. :. :. :. :
ALLERG QQPVLPPQQIPSVQPSILQQLNPKCVFLQQQCSVPAMPQSLARSQ--MLWQS--SCHVMQ
          140          150          160          170          180

          50
SYHT0H TLRLHRLKQMCVQCGS
          :. : :
ALLERG QQCCRQLPQIQEQRYSRYDAIRAIYISIVLQEQQHGGQLNQPPQQQPPQSVQGVSPQPPQQK
          190          200          210          220          230          240

>>ALLERGEN_2012_2|gi|94400907|ref|NP_001035360.1| allerg (92 aa)
  initn: 49 initl: 49 opt: 52 Z-score: 92.8 bits: 21.1 E(): 3.7
Smith-Waterman score: 52; 38.889% identity (41.176% ungapped) in 18 aa overlap (2-
18:55-72)

```

```

                10      20      30
SYHT0H          MPSTMLVSSCAQ-CTPGCVSQSLMQPNRWIV
                :...  : . :...:
ALLERG FGGFGGLGGRGKCPSNEIFSRCDGRCQRFPCNVVVKPLCIKICAPGCVCR LGYLRNKKKV
                30      40      50      60      70      80

```

```

                40      50
SYHT0H WKAYNSNHRLKTLRLHRLKQMCVQCGS

ALLERG CVPRSKCG
                90

```

```

>>ALLERGEN_2012_2|gi|88770352|gb|ABD51779.1| allergen Ap (94 aa)
  initn: 49 initl: 49 opt: 52 Z-score: 92.7 bits: 21.1 E(): 3.8
Smith-Waterman score: 52; 38.889% identity (41.176% ungapped) in 18 aa overlap (2-18:55-72)

```

```

                10      20      30
SYHT0H          MPSTMLVSSCAQ-CTPGCVSQSLMQPNRWIV
                :...  : . :...:
ALLERG FGGFGGLGGRGKCPSNEIFSRCDGRCQRFPCNVVVKPLCIKICAPGCVCR LGYLRNKKKV
                30      40      50      60      70      80

```

```

                40      50
SYHT0H WKAYNSNHRLKTLRLHRLKQMCVQCGS

ALLERG CVPRSKCLPG
                90

```

```

>>ALLERGEN_2012_2|gi|170730|gb|AAA34285.1| gamma-gliadin (304 aa)
  initn: 32 initl: 32 opt: 55 Z-score: 90.9 bits: 22.5 E(): 4.8
Smith-Waterman score: 55; 25.000% identity (27.778% ungapped) in 40 aa overlap (12-51:141-176)

```

```

                10      20      30      40
SYHT0H          MPSTMLVSSCAQCTPGCVSQSLMQPNRWIVWKAYNSNHRLK
                :...  : . :...:
ALLERG QQQQLVQQQIPIVQPSVLQQLNPCKVFLQQQCSPVAMPQRLARSQMW----QQSSCHVMQ
                120     130     140     150     160

```

```

                50
SYHT0H TLRLHRLKQMCVQCGS
                ..:..
ALLERG QQCCQLQQIPEQSRYEAIRAIYSIILQEQQQGFVQPQQQQPQQSGQGVSSQSSQQSSQQ
                170     180     190     200     210     220

```

```

>>ALLERGEN_2012_2|gi|194350817|gb|ACF53837.1| Bla g 4 is (190 aa)
  initn: 38 initl: 38 opt: 50 Z-score: 85.7 bits: 20.8 E(): 9.2
Smith-Waterman score: 50; 41.176% identity (41.176% ungapped) in 17 aa overlap (19-35:51-67)

```

```

                10      20      30      40
SYHT0H          MPSTMLVSSCAQCTPGCVSQSLMQPNRWIVWKAYNSNHRLKTLRLHRL
                :...  : . :...:
ALLERG ADEDCFRHESLVPNLDYERFRGMWVIVAGTSEALTQYKCWIDWFSYDDALVSKYTDSKQG
                30      40      50      60      70      80

```

```

                50
SYHT0H KQMCVQCGS

ALLERG NKILIGKIKFEGNKFTIDYDEGKAFSAPYSVLATDYDNYAIVEGCPAAANGHVIVQLR
                90      100     110     120     130     140

```

```
>>ALLERGEN_2012_2|gi|886967|emb|CAA59340.1| low molecula (276 aa)
  initn: 44 initl: 44 opt: 51 Z-score: 85.2 bits: 21.3 E(): 9.8
Smith-Waterman score: 55; 25.581% identity (35.484% ungapped) in 43 aa overlap
(12-54:179-209)
```

```

                        10      20      30      40
SYHT0H                MPSTMLVSSCAQCTPGCVSQSLMQPNRWIVWKAYNSNHLK
                        ::: . ::: . . :::..
ALLERG QQPVLPPQQIPYVQPSILQQLNPCKVFLQQQCSPVAMPQSLARSQ--MLWQS-----
      150      160      170      180      190

                        50
SYHT0H TLRHLRLKQMCVQCGS
      : .::: :
ALLERG --SCHVMQQCCQQLPRIPEQSRDYAIRAIYSIVLQEQQHGQGFNQPPQQPPQSVQGV
      200      210      220      230      240      250
```

```
57 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.020 Display time: 0.000
```

```
Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448
```

```
/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuBoIO: 40 aa
>SYHT0H2_I_20
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library
```

```
367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 3.9573+/-0.00339; mu= 4.1111+/- 0.177
mean_var=42.0783+/-10.208, 0's: 12 Z-trim: 12 B-trim: 68 in 2/40
Lambda= 0.1977
```

```
FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.010
```

```
The best scores are:
                                opt bits E(1603)
ALLERGEN_2012_2|gi|3287877|sp|P81402.1|NLTP1_PRUP ( 91) 48 20 5.2
ALLERGEN_2012_2|gi|83754241|pdb|2B5S|B Chain B, C ( 92) 48 20 5.3
ALLERGEN_2012_2|gi|1311510|gb|AAB36009.1| mAb 8C7 ( 15) 40 17 6.2
ALLERGEN_2012_2|gi|2506771|sp|P16968.2|IAA1_HORVU ( 146) 49 21 6.2
ALLERGEN_2012_2|gi|313575718|gb|ADR66939.1| non-s ( 117) 48 20 6.4
ALLERGEN_2012_2|gi|288561913|sp|P85894.1|LTP1_MOR ( 91) 47 20 6.4
ALLERGEN_2012_2|gi|5777414|emb|CAB53458.1| MnSOD ( 205) 50 21 6.6
ALLERGEN_2012_2|gi|10862818|emb|CAC13961.1| IgE-b ( 205) 50 21 6.6
ALLERGEN_2012_2|gi|348137|gb|AAA16792.1| superoxi ( 233) 50 21 7.3
ALLERGEN_2012_2|gi|313575726|gb|ADR66943.1| non-s ( 117) 47 20 7.8
ALLERGEN_2012_2|gi|23894244|emb|CAD23614.1| tri m ( 404) 51 21 9.3
ALLERGEN_2012_2|gi|74663809|sp|Q8J077.1|SUB6_TRIS ( 405) 51 21 9.3
ALLERGEN_2012_2|gi|9929163|emb|CAC05258.1| Cup a ( 199) 48 20 9.6
```

```
>>ALLERGEN_2012_2|gi|3287877|sp|P81402.1|NLTP1_PRUPE Rec (91 aa)
  initn: 44 initl: 44 opt: 48 Z-score: 90.1 bits: 20.1 E(): 5.2
Smith-Waterman score: 48; 37.037% identity (41.667% ungapped) in 27 aa overlap (9-35:4-27)
```



```

          10          20          30          40
SYHT0H MVNDITNRGHISCCSWPNLNWSPLRRHGGLVISACPLQMK
      :..:      : .      : : : : : :
ALLERG      ITCGQVSSALAPCI---PYVRGGGAVPPACNGIRNVNNLARTTPDRQAACNCLK
          10          20          30          40          50

ALLERG QLSASVPGVNPNNAAALPGKCGVHIPYKISASTNCATVK
          60          70          80          90

```

>>ALLERGEN\_2012\_2|gi|83754241|pdb|2B5S|B Chain B, Crysta (92 aa)  
 initn: 44 initl: 44 opt: 48 Z-score: 90.1 bits: 20.1 E(): 5.3  
 Smith-Waterman score: 48; 37.037% identity (41.667% ungapped) in 27 aa overlap (9-35:5-28)

```

          10          20          30          40
SYHT0H MVNDITNRGHISCCSWPNLNWSPLRRHGGLVISACPLQMK
      :..:      : .      : : : : : :
ALLERG      MITCGQVSSSLAPCI---PYVRGGGAVPPACNGIRNVNNLARTTPDRQAACNCLK
          10          20          30          40          50

ALLERG QLSASVPGVNPNNAAALPGKCGVSIPIYKISASTNCATVK
          60          70          80          90

```

>>ALLERGEN\_2012\_2|gi|1311510|gb|AAB36009.1| mAb 8C7-reac (15 aa)  
 initn: 40 initl: 40 opt: 40 Z-score: 88.8 bits: 17.2 E(): 6.2  
 Smith-Waterman score: 40; 55.556% identity (55.556% ungapped) in 9 aa overlap (28-36:4-12)

```

          10          20          30          40
SYHT0H MVNDITNRGHISCCSWPNLNWSPLRRHGGLVISACPLQMK
      : : : : : :
ALLERG      APAGGVVVAAMPPPL
          10

```

>>ALLERGEN\_2012\_2|gi|2506771|sp|P16968.2|IAA1\_HORVU RecN (146 aa)  
 initn: 30 initl: 30 opt: 49 Z-score: 88.8 bits: 20.5 E(): 6.2  
 Smith-Waterman score: 49; 37.500% identity (42.857% ungapped) in 24 aa overlap (14-35:20-42)

```

          10          20          30          40
SYHT0H      MVNDITNRGHISCCSWPNLNWS--PLRRHGGLVISACPLQMK
      : : : : : :
ALLERG      PTSVAVDQGSMSVNSPGEWC-WPGMGYPVYFPFRCRALVKSQCAGGQVVESIQKDCCRQI
          10          20          30          40          50

ALLERG AAIGDEWCICGALGSMRGS MYKELGVALADDKATVAEVFPGCRTEVMDRAVASLPVCNQ
          60          70          80          90          100          110

```

>>ALLERGEN\_2012\_2|gi|313575718|gb|ADR66939.1| non-specif (117 aa)  
 initn: 44 initl: 44 opt: 48 Z-score: 88.6 bits: 20.2 E(): 6.4  
 Smith-Waterman score: 48; 37.037% identity (41.667% ungapped) in 27 aa overlap (9-35:30-53)

```

          10          20          30
SYHT0H      MVNDITNRGHISCCSWPNLNWSPLRRHGGLVISACPLQM
      :..:      : .      : : : : : :
ALLERG      MAYSAMTKLALVVALCMVSVPIAQAITCGQVSSSLAPCI---PYVRGGGAVPPACNGI
          10          20          30          40          50

```

```

40
SYHT0H K

ALLERG RNVNNLARTTTPDRQAACNCLKQLSASVPGVNPNNAAALPGKCGVSIPYKISASTNCATVK
      60      70      80      90     100     110

>>ALLERGEN_2012_2|gi|288561913|sp|P85894.1|LTP1_MORNI Re (91 aa)
  initn: 32 initl: 32 opt: 47 Z-score: 88.6 bits: 19.8 E(): 6.4
Smith-Waterman score: 47; 33.333% identity (37.500% ungapped) in 27 aa overlap (9-
35:4-27)

      10      20      30      40
SYHT0H MVNDITNRGHISCCSWPNLNWSPLRRHGGLVISACPLQMK
      :.. : :. : :.: :
ALLERG ITCGQVSSSLAPCINYL--RAGGVVPANCCNGVRSLNNAAKTTADRQAACNCLK
      10      20      30      40      50

ALLERG SAFNSIKGLNLNLAAGLPGKCGVSPYKISPSTDCKSVK
      60      70      80      90

>>ALLERGEN_2012_2|gi|5777414|emb|CAB53458.1| MnSOD [Heve (205 aa)
  initn: 38 initl: 38 opt: 50 Z-score: 88.3 bits: 20.9 E(): 6.6
Smith-Waterman score: 50; 40.909% identity (47.368% ungapped) in 22 aa overlap (9-
29:71-90)

      10      20      30
SYHT0H MVNDITNRGHIS-CCSWPNLNWSPLRRHGGLVISACPL
      :.. : :. : :.: :
ALLERG KALEQLNDAIEKGDSAADVVKLQSAIKFNGGGHVNHSIFWKNL--APVREGGGELPHGSLG
      50      60      70      80      90

40
SYHT0H QMK

ALLERG WAIDADFGSLEKLIQLMNAEGVALQGSGVWVLALDKELKKLVVETTANQDPLVTKGPTLV
      100     110     120     130     140     150

>>ALLERGEN_2012_2|gi|10862818|emb|CAC13961.1| IgE-bindin (205 aa)
  initn: 38 initl: 38 opt: 50 Z-score: 88.3 bits: 20.9 E(): 6.6
Smith-Waterman score: 50; 40.909% identity (47.368% ungapped) in 22 aa overlap (9-
29:71-90)

      10      20      30
SYHT0H MVNDITNRGHIS-CCSWPNLNWSPLRRHGGLVISACPL
      :.. : :. : :.: :
ALLERG KALEQLNDAIEKGDSAADVVKLQSAIKFNGGGHVNHSIFWKNL--APVREGGGELPHGSLG
      50      60      70      80      90

40
SYHT0H QMK

ALLERG WAIDADFGSLEKLIQLMNAEGAALRGSGVWVLALDKELKKLVVETTANQDPLVTKGPTLV
      100     110     120     130     140     150

>>ALLERGEN_2012_2|gi|348137|gb|AAA16792.1| superoxide di (233 aa)
  initn: 38 initl: 38 opt: 50 Z-score: 87.5 bits: 21.0 E(): 7.3
Smith-Waterman score: 50; 40.909% identity (47.368% ungapped) in 22 aa overlap (9-
29:99-118)

```

```

                                10      20      30
SYHT0H                      MVNDITNRGHIS-CCSWPNLNWSPLRRHGGGLVISACPL
                                :...  :  :  :...  :
ALLERG KALEQLNDAIEKGDSA AVVKLQSAIKFNGGGHVNHSIFWKNL--APVREGGGELPHGSLG
      70      80      90      100      110      120

      40
SYHT0H QMK

ALLERG WAIDADFGSLEKLIQLMNAEGAALQGSGVWVLALDKELKKLVVETTANQDPLVTKGPTLV
      130      140      150      160      170      180

>>ALLERGEN_2012_2|gi|313575726|gb|ADR66943.1| non-specif (117 aa)
  initn: 44 init1: 44 opt: 47 Z-score: 87.1 bits: 19.9 E(): 7.8
Smith-Waterman score: 47; 37.037% identity (41.667% ungapped) in 27 aa overlap (9-
35:30-53)

                                10      20      30
SYHT0H                      MVNDITNRGHISCCSWPNLNWSPLRRHGGGLVISACPLQM
                                :...  :  :  :...  :
ALLERG MACSAMTKLALVVALCMVVSVPPIAQALTGQVSSNLAPCI---PYVRGGGAVPPACCNGI
      10      20      30      40      50

      40
SYHT0H K

ALLERG RNINNLAKT TADRQTACNCLKQLSASVPGVNNANNAALPGKCGVNVVPYKISPSTNCATVK
      60      70      80      90      100      110

>>ALLERGEN_2012_2|gi|23894244|emb|CAD23614.1| tri m 2 al (404 aa)
  initn: 35 init1: 35 opt: 51 Z-score: 85.7 bits: 21.4 E(): 9.3
Smith-Waterman score: 51; 30.000% identity (32.143% ungapped) in 30 aa overlap (1-
28:39-68)

                                10      20
SYHT0H                      MVNDITNRGHISCCSW--PNLNWSPLRRHG
                                : .....  .  :  :...  :...  :
ALLERG VLAALSAVNGAKILEAGPHAETIPNKYIVVMKKDVSDEAFSTHTTWLSQNLNRRMLMRRSG
      10      20      30      40      50      60

      30      40
SYHT0H GLVISACPLQMK
      .
ALLERG SSKAMAGMQNKYSLGGIFRAYSGEFDDAMIKDISNHDDVDYIEPDFVVRTSTNGTNLTRQ
      70      80      90      100      110      120

>>ALLERGEN_2012_2|gi|74663809|sp|Q8J077.1|SUB6_TRISH Rec (405 aa)
  initn: 35 init1: 35 opt: 51 Z-score: 85.7 bits: 21.4 E(): 9.3
Smith-Waterman score: 51; 30.000% identity (32.143% ungapped) in 30 aa overlap (1-
28:39-68)

                                10      20
SYHT0H                      MVNDITNRGHISCCSW--PNLNWSPLRRHG
                                : .....  .  :  :...  :...  :
ALLERG VLAALSAVNGAKILEAGPHAETIPNKYIVVMKKDVSDEAFSTHTTWLSQNLNRRMLMRRSG
      10      20      30      40      50      60

      30      40
SYHT0H GLVISACPLQMK
      .
ALLERG SSKAMAGMQNKYSLGGIFRAYSGEFDDAMIKDISNHDDVDYIEPDFVVRTSTNGTNLTRQ
      70      80      90      100      110      120

```



```

        60          70
SYHT0H SLYCEDSGKGRWLLQMPSLR

ALLERG HQPQQPYPQQQPYGSSLTSIDGQ
        340          350

>>ALLERGEN_2012_2|gi|303387468|gb|ADM15668.1| lipid bind (228 aa)
  initn: 39 initl: 39 opt: 60 Z-score: 99.5 bits: 24.0 E(): 1.6
Smith-Waterman score: 60; 21.622% identity (22.535% ungapped) in 74 aa overlap (1-
74:14-84)

        10          20          30          40
SYHT0H          MMAFVGATFLFHYLHNKVTDSWAMEPLLLHMDLLSDLLELNIRKPPR
        ..: .. .. . :...: :. : : : : : : : : : : : : :
ALLERG MLKVSSLFILLCGLLASSSTQEVLSRVSSHITD--ALTQGLLG MNFLPTLQTIDFQGPKL
        10          20          30          40          50

        50          60          70
SYHT0H IPLPSYLSLYCEDSGKGRWLLQMPSLR
        . : . . . :... :...: :..
ALLERG -DIFSLVLGHQLTNGEANFMVQMKDLRLFQVFIETSPDFKGIDLRMPLAFSIQIKFPALN
        60          70          80          90          100          110

>>ALLERGEN_2012_2|gi|25361513|gb|AAN73248.1| helix-loop- (450 aa)
  initn: 48 initl: 48 opt: 61 Z-score: 95.4 bits: 24.3 E(): 2.7
Smith-Waterman score: 61; 45.455% identity (47.619% ungapped) in 22 aa overlap
(11-32:201-221)

        10          20          30          40
SYHT0H          MMAFVGATFLFHYLHNKVTDSWAMEPLLLHMDLLSDLLEL
        : : : : : : : : : : : : : : : : : :
ALLERG ANQFNFEVPGHMNVSGHPHLEHSSTNPNSFHYEHNIVSPS-SIHPSTAHDGEVPSQWD
        180          190          200          210          220

        50          60          70
SYHT0H NIRKPPRIPLPSYLSLYCEDSGKGRWLLQMPSLR

ALLERG DSLGHGASTPKVTRTPSHHVSSNPWAEINEPTGGDNDNLAPVTRPRKPARARRQKKEPRKL
        230          240          250          260          270          280

>>ALLERGEN_2012_2|gi|4138175|emb|CAA09885.1| allergen [M (187 aa)
  initn: 56 initl: 56 opt: 56 Z-score: 95.0 bits: 22.9 E(): 2.8
Smith-Waterman score: 56; 53.333% identity (53.333% ungapped) in 15 aa overlap
(49-63:6-20)

        20          30          40          50          60          70
SYHT0H TDSWAMEPLLLHMDLLSDLLELNIRKPPRIPLPSYLSLYCEDSGKGRWLLQMPSLR
        : : : : : : : : : : : : : : : : : :
ALLERG          AVSASPTPSKHNLYCYAQGKDLFEFHINDTVTKDV
        10          20          30

ALLERG CKSLNSGKYHNMNNEKYCSVADYDVKWKERCQSHPTDVKTTKWIAGTDLKIEMDPKEPY
        40          50          60          70          80          90

>>ALLERGEN_2012_2|gi|7638028|gb|AAF65312.1|AF230383_1 ve (137 aa)
  initn: 34 initl: 34 opt: 52 Z-score: 91.4 bits: 21.8 E(): 4.5
Smith-Waterman score: 52; 30.000% identity (35.294% ungapped) in 40 aa overlap
(31-64:25-64)

```

```

      10      20      30      40      50
SYHT0H MMAFVGATFLFHYLHNKVTDSWAMEPLLLHMDLLSDLLELNIRKPPR-----IPLPSYL
      . . . . .:: :: :. :: .
ALLERG      MKTFVLVSCLLVFTQIIYAVDIKELKIMNRILEKCIRTVPKGENDPINPLKNVN
      10      20      30      40      50

      60      70
SYHT0H SLYCEDSGKGRWLLQMPSLR
      ::: : .:
ALLERG VLYCAFSKRGIFTPKGVNTKQYINYCEKTIINPADIKQCKKLISKCIKKVYDRPGPIIER
      60      70      80      90      100      110

>>ALLERGEN_2012_2|gi|4038411|gb|AAC97370.1| venom allerg (137 aa)
  initn: 34 init1: 34 opt: 49 Z-score: 86.8 bits: 20.9 E(): 8
Smith-Waterman score: 49; 27.500% identity (32.353% ungapped) in 40 aa overlap
(31-64:25-64)

```

```

      10      20      30      40      50
SYHT0H MMAFVGATFLFHYLHNKVTDSWAMEPLLLHMDLLSDLLELNIRKPPR-----IPLPSYL
      . . . . .:: :: :. :: .
ALLERG      MKTFVLVSCLLVFTQIIYALDIKEISIMNRILEKCIRTVPKRENDPINPLKNVN
      10      20      30      40      50

      60      70
SYHT0H SLYCEDSGKGRWLLQMPSLR
      ::: . .:
ALLERG VLYCAFTKRGIFTPKGVNTKQYINYCEKTIISPADIKLCCKKIASKCVKKVYDRPGPIIER
      60      70      80      90      100      110

>>ALLERGEN_2012_2|gi|631911|pir||S43242 allergen-like pr (145 aa)
  initn: 27 init1: 27 opt: 49 Z-score: 86.3 bits: 20.9 E(): 8.5
Smith-Waterman score: 49; 25.000% identity (29.167% ungapped) in 28 aa overlap
(46-70:4-30)

```

```

      20      30      40      50      60      70
SYHT0H NKVTDSWAMEPLLLHMDLLSDLLELNIRKPPRIPLPSYL---SLYCEDSGKGRWLLQMPS
      :. :.:. .:: :. :.:. .
ALLERG      EDVPQPPPIQFHIQGQVYC-DTCRARFITEELSE
      10      20      30

SYHT0H LR

ALLERG FIPGASIRLQCKDRENGKITFTEIGYTRAEGLYSMLVEGDHKNFCEITLISSGREDCDE
      40      50      60      70      80      90

```

74 residues in 1 query sequences  
 367796 residues in 1603 library sequences  
 Scomplib [34t11]  
 start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012  
 Scan time: 0.020 Display time: 0.000

Function used was FASTA [version 3.4t11 Apr 17, 2002]  
 FASTA searches a protein or DNA sequence data bank  
 version 3.4t11 Apr 17, 2002  
 Please cite:  
 W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

/grid/.qsub\_tmp/1330373600kde\_agent/grid\_NbCuMBoIO: 439 aa  
 >SYHT0H2\_I\_22  
 vs /data/fasta/ALLERGEN\_2012\_2 library

searching /data/fasta/ALLERGEN\_2012\_2 library

367796 residues in 1603 sequences

Expectation\_n fit: rho(ln(x))= 4.1238+/-0.00439; mu= 19.1602+/- 0.227

mean\_var=69.4357+/-17.631, 0's: 2 Z-trim: 2 B-trim: 79 in 1/43

Lambda= 0.1539

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2

join: 37, opt: 25, gap-pen: -12/-2, width: 16

Scan time: 0.080

The best scores are:

					opt	bits	E(1603)
ALLERGEN_2012_2	gi 46410859 gb AAR98518.1	major	( 366)	87	28	1.1	
ALLERGEN_2012_2	gi 2266625 emb CAB10765.1	group	( 264)	80	26	2.7	
ALLERGEN_2012_2	gi 1092249 prf 2023228A	major al	( 285)	76	25	5.2	
ALLERGEN_2012_2	gi 2398757 emb CAA50281.1	Major	( 286)	76	25	5.2	
ALLERGEN_2012_2	gi 21725630 emb CAD38396.1	unnam	( 287)	76	25	5.2	
ALLERGEN_2012_2	gi 21725612 emb CAD38387.1	unnam	( 287)	76	25	5.2	
ALLERGEN_2012_2	gi 21725614 emb CAD38388.1	unnam	( 287)	76	25	5.2	
ALLERGEN_2012_2	gi 21725616 emb CAD38389.1	unnam	( 287)	76	25	5.2	
ALLERGEN_2012_2	gi 21725618 emb CAD38390.1	unnam	( 287)	76	25	5.2	
ALLERGEN_2012_2	gi 21725620 emb CAD38391.1	unnam	( 287)	76	25	5.2	
ALLERGEN_2012_2	gi 21725622 emb CAD38392.1	unnam	( 287)	76	25	5.2	
ALLERGEN_2012_2	gi 21725624 emb CAD38393.1	unnam	( 287)	76	25	5.2	
ALLERGEN_2012_2	gi 21725626 emb CAD38394.1	unnam	( 287)	76	25	5.2	
ALLERGEN_2012_2	gi 21725632 emb CAD38397.1	unnam	( 287)	76	25	5.2	
ALLERGEN_2012_2	gi 21725606 emb CAD38384.1	unnam	( 287)	76	25	5.2	
ALLERGEN_2012_2	gi 21725610 emb CAD38386.1	unnam	( 287)	76	25	5.2	
ALLERGEN_2012_2	gi 21725608 emb CAD38385.1	unnam	( 287)	76	25	5.2	
ALLERGEN_2012_2	gi 21725628 emb CAD38395.1	unnam	( 287)	76	25	5.2	
ALLERGEN_2012_2	gi 3309039 gb AAC25994.1	group V	( 312)	76	25	5.5	
ALLERGEN_2012_2	gi 398830 emb CAA52753.1	Phlp5 [	( 312)	76	25	5.5	
ALLERGEN_2012_2	gi 13430402 gb AAK25823.1	group	( 275)	74	25	6.9	
ALLERGEN_2012_2	gi 29500897 emb CAD87529.1	phl p	( 284)	73	25	8.2	

>>ALLERGEN\_2012\_2|gi|46410859|gb|AAR98518.1| major latex (366 aa)

initn: 40 init1: 40 opt: 87 Z-score: 102.2 bits: 27.8 E(): 1.1

Smith-Waterman score: 117; 22.881% identity (26.087% ungapped) in 236 aa overlap  
(175-404:37-249)

```

      150      160      170      180      190      200
SYHT0H EAFRVSVAGGARPAFAPADLGHGFGFLAEVELYGDVVLRFVSYPDETDLPFLPGFERVSSP
      ::      ..      ::::      :.....      :
ALLERG SLFILSLFTFTLLNPVCTELDEYLFSGDGLYDAGNAKFI-YPDK----YLPSYHH---P
      10       20       30       40       50

      210      220      230      240      250      260
SYHT0H GAVDYGLTRFDHVVGNVPPEMAPVIDYMKGFLGFHEFAEFTAEDVGTTESGLNSVVLANN
      :: : :: . : . : :: . . . . : . . . : : : . :
ALLERG ---YGTTFDFDYPTGRFSDGRTTVVDFVAENVSLPRIPPFKNKEANFTY-GANFA-----S
      60       70       80       90      100

      270      280      290      300      310      320
SYHT0H EAVLLPLNEPVHGTKRRSQIQTYLEYHGGPGVQHIALASNDVLRRTLREMRARTPMGGFEF
      :..      .:.      :::: . : . : :: . . . . : : : . :
ALLERG EGATASDSNPL--IDFRSQIRDFGELKLEWAVQLVNV--ELARRLKKAVYLISFGADDY
      110      120      130      140      150      160

      330      340      350      360      370
SYHT0H MA--PPQAKYYEGVRRAGDVLSE--EQIKECQELGV--LVDRDDQGVLLQIFTKPVGDR
      . : . : . : . : :: . : . : : : : : : : : :
ALLERG LNYEIPSEASREQLESIVDVVLGNISDRIKELYDFGARKFVVENVAPLGLIPFIKQTSN
      170      180      190      200      210      220
```

```

      380      390      400      410      420      430
SYHT0H PTFFLEMIQRIGCMEKDEVGQEYQKGGCGGFGKGNFSELFKSIEDYEKSLEVKQSVVAQK
      :: :: :: :: :: :: :: :: :: ::
ALLERG STLFEYELAS-LHAMKLPQILEKIQDGYLFPEFNITVFNYFGIIEKIIDAPGEHGFYKYGDI
      230      240      250      260      270      280

>>ALLERGEN_2012_2|gi|2266625|emb|CAB10765.1| group V all (264 aa)
  initn: 68 initl: 68 opt: 80 Z-score: 95.4 bits: 26.1 E(): 2.7
Smith-Waterman score: 80; 34.286% identity (35.821% ungapped) in 70 aa overlap
(95-162:57-125)

```

```

      70      80      90      100      110      120
SYHT0H GAPLAARSDLSTGNSAHASLLLRSGALAFLLTAPYAPPPQEAATAATASIPSFSADAART
      :::: :::: :::: :::: :::: ::::
ALLERG LEDVNAGFKTAVAAAANVPPADKYKTFEAAFTASSKASIAAAATKAPGLIPQLNAATNTA
      30      40      50      60      70      80

      130      140      150      160      170      180
SYHT0H FAAAHGLAVRS-VGVRVADAAEAFRVSVAGGAR-PAFAPADLGHGFGGLAEVELYGDVVLRL
      :::: :::: :::: :::: :::: ::::
ALLERG YAAAQGATPEAKYDAFVTTLTEALRV-IAGALEVHAVKPATEEVGAAKIPAGELQIVDKI
      90      100      110      120      130      140

```

```

      190      200      210      220      230      240
SYHT0H FVSYPDETDLPLPGFERVSSPGAVDYLGLTRFDHVVGNVPEMAPVIDYMKGFLGFHEFAE

ALLERG DAAFRIAATAANAAPVNDKFTVFEGAFNKAIKESTGGAYEAYKFIPSLETAVKQAYAATV
      150      160      170      180      190      200

```

```

>>ALLERGEN_2012_2|gi|1092249|prf||2023228A major allergen (285 aa)
  initn: 122 initl: 57 opt: 76 Z-score: 90.2 bits: 25.2 E(): 5.2
Smith-Waterman score: 77; 28.713% identity (31.522% ungapped) in 101 aa overlap
(72-172:1-92)

```

```

      50      60      70      80      90      100
SYHT0H SFHHVELWCADAASAAGRFSFALGAPLAARSDLSTGNSAHASLLLRSGALAFLLTAPYAP
      :::: :::: :::: :::: :::: ::::
ALLERG ADLGYGPATPAA-----PAAGYTPATPAAP
      10      20

      110      120      130      140      150      160
SYHT0H PPQEAATAATASIPSFSADAARTFAAAHGLAVRSVGVRVADAAEAFRVSVAGGARPAFAP
      :::: :::: :::: :::: :::: ::::
ALLERG AGADAAGKATTEEQKLIKINAGFKAA--LA--GAGVQPADKYRTFVATFGPASNKAFAE
      30      40      50      60      70      80

```

```

      170      180      190      200      210      220
SYHT0H ADLGHGFGGLAEVELYGDVVLRFVSYYPDETDLPLPGFERVSSPGAVDYLGLTRFDHVVGNV
      . :: ::::
ALLERG GLSGEPKGAAESSSKAALTSKLDAAAYKLAYKTAEGATPEAKYDAYVATLSEALRIIAGTL
      90      100      110      120      130      140

```

```

>>ALLERGEN_2012_2|gi|2398757|emb|CAA50281.1| Major Polle (286 aa)
  initn: 111 initl: 57 opt: 76 Z-score: 90.2 bits: 25.2 E(): 5.2
Smith-Waterman score: 77; 28.713% identity (31.522% ungapped) in 101 aa overlap
(72-172:1-92)

```

```

      50      60      70      80      90      100
SYHT0H SFHHVELWCADAASAAGRFSFALGAPLAARSDLSTGNSAHASLLLRSGALAFLLTAPYAP
      :::: :::: :::: :::: :::: ::::
ALLERG ADLGYGPATPAA-----PAAGYTPATPAAP
      10      20

```



```

      110      120      130      140      150      160
SYHT0H PPQEAATAATASIPFSADAARTFAAAHGLAVRSVGVRVADAAEAFRVSVAGGARPAFAP
      .:: ::. .. : :: :: :... :: :.. .. . . :::
ALLERG AGADAAGKATTEEQKLIKINAGFKAA--LA--GAGVQPADKYRTFVATFGPASNKAFAE
      30      40      50      60      70      80

```

```

      170      180      190      200      210      220
SYHT0H ADLGHGFGLAEVELYGDVVLRFVSYPDETDLPFLPGFERVSSPGAVDYGLTRFDHVGNV
      . :. : ::
ALLERG GLSGEPKGAAESSSKAALTSKLDAAAYKLAYKTAEGATPEAKYDAYVATLSEALRIIAGTL
      90      100      110      120      130      140

```

```

>>ALLERGEN_2012_2|gi|21725630|emb|CAD38396.1| unnamed pr (287 aa)
  initn: 52 init1: 52 opt: 76 Z-score: 90.2 bits: 25.2 E(): 5.2
Smith-Waterman score: 76; 28.713% identity (31.183% ungapped) in 101 aa overlap
(72-172:1-93)

```

```

      50      60      70      80      90      100
SYHT0H SFHHVELWCADAASAAGRFSFALGAPLAARSDLSTGNSAHASLLLRSGALAFLLFTAPYAP
      :... : .. :. : .. :.. ::
ALLERG                                ADLGYGPATPAA-----PAAGYTPATPAAP
                                10                                20

```

```

      110      120      130      140      150      160
SYHT0H PPQEAATAATASIPFSADAARTFAAAHGLAVRSVGVRVADAAEAFRVSVAGGARPAFAP
      : : ::. .. : :: :: :... :: :.. .. . . :::
ALLERG AGAEPAGKATTEEQKLIKKNAGFKAA--LAA-AAGVPPADKYRTFVATFGAASNKAFAE
      30      40      50      60      70      80

```

```

      170      180      190      200      210      220
SYHT0H ADLGHGFGLAEVELYGDVVLRFVSYPDETDLPFLPGFERVSSPGAVDYGLTRFDHVGNV
      . :. : ::
ALLERG GLSGEPKGAAESSSKAALTSKLDAAAYKLAYKTAEGATPEAKYDAYVATVSSALRIIAGTL
      90      100      110      120      130      140

```

```

>>ALLERGEN_2012_2|gi|21725612|emb|CAD38387.1| unnamed pr (287 aa)
  initn: 52 init1: 52 opt: 76 Z-score: 90.2 bits: 25.2 E(): 5.2
Smith-Waterman score: 76; 28.713% identity (31.183% ungapped) in 101 aa overlap
(72-172:1-93)

```

```

      50      60      70      80      90      100
SYHT0H SFHHVELWCADAASAAGRFSFALGAPLAARSDLSTGNSAHASLLLRSGALAFLLFTAPYAP
      :... : .. :. : .. :.. ::
ALLERG                                ADLGYGPATPAA-----PAAGYTPATPAAP
                                10                                20

```

```

      110      120      130      140      150      160
SYHT0H PPQEAATAATASIPFSADAARTFAAAHGLAVRSVGVRVADAAEAFRVSVAGGARPAFAP
      : : ::. .. : :: :: :... :: :.. .. . . :::
ALLERG AGAEPAGKATTEEQKLIKKNAGFKAA--LAA-AAGVPPADKYRTFVATFGAASNKAFAE
      30      40      50      60      70      80

```

```

      170      180      190      200      210      220
SYHT0H ADLGHGFGLAEVELYGDVVLRFVSYPDETDLPFLPGFERVSSPGAVDYGLTRFDHVGNV
      . :. : ::
ALLERG GLSGEPKGAAESSSKAALTSKLDAAAYKLAYKTAEGATPEAKYDAYVATVSEALRKIAGTL
      90      100      110      120      130      140

```

```

>>ALLERGEN_2012_2|gi|21725614|emb|CAD38388.1| unnamed pr (287 aa)
  initn: 52 init1: 52 opt: 76 Z-score: 90.2 bits: 25.2 E(): 5.2
Smith-Waterman score: 76; 28.713% identity (31.183% ungapped) in 101 aa overlap
(72-172:1-93)

```

```

          50          60          70          80          90          100
SYHT0H SFHHVELWCADAASAAGRFSFALGAPLAARSDLSTGNSAHASLLLRSGALAFLLFTAPYAP
          . . . . : . . . : . . . . : . . . . :
ALLERG                                ADLGYGPATPAA-----PAAGYTPATPAAP
                                   10              20

```

```

          110          120          130          140          150          160
SYHT0H PPQEAATAATASIPSFSADAARTFAAAHGLAVRSVGVRVADAAEAFRVSVAGGARPAFAP
          : : : . . . : : : . . . : : : . . . : . . . . : . . . : . . . :
ALLERG AGAEPAGKATTEEQKLIEKKNAGFKAA--LAA-AAGVPPADKYRTFVATFGAASNKAFAE
          30          40          50          60          70          80

```

```

          170          180          190          200          210          220
SYHT0H ADLGHGFGGLAEVELYGDVVLRFVSYPDETDLPFLPGFERVSSPGAVDYLTRFDHVVGNV
          . . . : :
ALLERG GLSGEPKGAAESSSKAALTSKLDAAAYKLAYKTAEGATPEAKYDAYVATVSSALRIIAGTL
          90          100          110          120          130          140

```

>>ALLERGEN\_2012\_2|gi|21725616|emb|CAD38389.1| unnamed pr (287 aa)  
 initn: 52 initl: 52 opt: 76 Z-score: 90.2 bits: 25.2 E(): 5.2  
 Smith-Waterman score: 76; 28.713% identity (31.183% ungapped) in 101 aa overlap  
 (72-172:1-93)

```

          50          60          70          80          90          100
SYHT0H SFHHVELWCADAASAAGRFSFALGAPLAARSDLSTGNSAHASLLLRSGALAFLLFTAPYAP
          . . . . : . . . : . . . . : . . . . :
ALLERG                                ADLGYGPATPAA-----PAAGYTPATPAAP
                                   10              20

```

```

          110          120          130          140          150          160
SYHT0H PPQEAATAATASIPSFSADAARTFAAAHGLAVRSVGVRVADAAEAFRVSVAGGARPAFAP
          : : : . . . : : : . . . : : : . . . : . . . . : . . . : . . . :
ALLERG AGAEPAGKATTEEQKLIEKKNAGFKAA--LAA-AAGVPPADKYRTFVATFGAASNKAFAE
          30          40          50          60          70          80

```

```

          170          180          190          200          210          220
SYHT0H ADLGHGFGGLAEVELYGDVVLRFVSYPDETDLPFLPGFERVSSPGAVDYLTRFDHVVGNV
          . . . : :
ALLERG GLSGEPKGAAESSSKAALTSKLDAAAYKLAYKTAEGATPEAKYDAYVATVSSALRIIAGTL
          90          100          110          120          130          140

```

>>ALLERGEN\_2012\_2|gi|21725618|emb|CAD38390.1| unnamed pr (287 aa)  
 initn: 52 initl: 52 opt: 76 Z-score: 90.2 bits: 25.2 E(): 5.2  
 Smith-Waterman score: 76; 28.713% identity (31.183% ungapped) in 101 aa overlap  
 (72-172:1-93)

```

          50          60          70          80          90          100
SYHT0H SFHHVELWCADAASAAGRFSFALGAPLAARSDLSTGNSAHASLLLRSGALAFLLFTAPYAP
          . . . . : . . . : . . . . : . . . . :
ALLERG                                ADLGYGPATPAA-----PAAGYTPATPAAP
                                   10              20

```

```

          110          120          130          140          150          160
SYHT0H PPQEAATAATASIPSFSADAARTFAAAHGLAVRSVGVRVADAAEAFRVSVAGGARPAFAP
          : : : . . . : : : . . . : : : . . . : . . . . : . . . : . . . :
ALLERG AGAEPAGKATTEEQKLIEKKNAGFKAA--LAA-AAGVPPADKYRTFVATFGAASNKAFAE
          30          40          50          60          70          80

```

```

          170          180          190          200          210          220
SYHT0H ADLGHGFGGLAEVELYGDVVLRFVSYPDETDLPFLPGFERVSSPGAVDYLTRFDHVVGNV
          . . . : :
ALLERG GLSGEPKGAAESSSKAALTSKLDAAAYKLAYKTAEGATPEAKYDAYVATVSSALRIIAGTL
          90          100          110          120          130          140

```

>>ALLERGEN\_2012\_2|gi|21725620|emb|CAD38391.1| unnamed pr (287 aa)  
 initn: 52 initl: 52 opt: 76 Z-score: 90.2 bits: 25.2 E(): 5.2  
 Smith-Waterman score: 76; 28.713% identity (31.183% ungapped) in 101 aa overlap  
 (72-172:1-93)

```

      50      60      70      80      90     100
SYHT0H SFHHVELWCADAASAAGRFSFALGAPLAARSDLSTGNSAHASLLLRSGALAFLLFTAPYAP
      . . . . : . . . : . . . . : . . . . :
ALLERG                                ADLGYGPATPAA-----PAAGYTPATPAAP
                                   10      20

      110     120     130     140     150     160
SYHT0H PPQEAATAATASIPSFSADAARTFAAAHGLAVRSVGVRVADAAEAFRVSVAGGARPAFAP
      : : : . . . : : : : . . . : : . . . . . : :
ALLERG AGAEPAGKATTEEQKLIEKKNAGFKAA--LAA-AAGVPPADKYRTFVATFGAASNKAFAE
      30      40      50      60      70      80

      170     180     190     200     210     220
SYHT0H ADLGHGFGGLAEVELYGDVVLRFVSYPDETDLPFLPGFERVSSPGAVDYLTRFDHVVGNV
      . . : : :
ALLERG GLSGEPKGAAESSSKAALTSKLDAAAYKLAYKTAEGATPEAKYDAYVATVSSALRIIAGTL
      90      100     110     120     130     140

```

>>ALLERGEN\_2012\_2|gi|21725622|emb|CAD38392.1| unnamed pr (287 aa)  
 initn: 52 initl: 52 opt: 76 Z-score: 90.2 bits: 25.2 E(): 5.2  
 Smith-Waterman score: 76; 28.713% identity (31.183% ungapped) in 101 aa overlap  
 (72-172:1-93)

```

      50      60      70      80      90     100
SYHT0H SFHHVELWCADAASAAGRFSFALGAPLAARSDLSTGNSAHASLLLRSGALAFLLFTAPYAP
      . . . . : . . . : . . . . : . . . . :
ALLERG                                ADLGYGPATPAA-----PAAGYTPATPAAP
                                   10      20

      110     120     130     140     150     160
SYHT0H PPQEAATAATASIPSFSADAARTFAAAHGLAVRSVGVRVADAAEAFRVSVAGGARPAFAP
      : : : . . . : : : : . . . : : . . . . . : :
ALLERG AGAEPAGKATTEEQKLIEKINAGFKAA--LAA-AAGVPPADKYNTFVATFGAASNKAFAE
      30      40      50      60      70      80

      170     180     190     200     210     220
SYHT0H ADLGHGFGGLAEVELYGDVVLRFVSYPDETDLPFLPGFERVSSPGAVDYLTRFDHVVGNV
      . . : : :
ALLERG GLSGEPKGAAESSSKAALTSKLDAAAYKLAYKTAEGATPEAKYDAYVATVSEALSIIAGTL
      90      100     110     120     130     140

```

>>ALLERGEN\_2012\_2|gi|21725624|emb|CAD38393.1| unnamed pr (287 aa)  
 initn: 52 initl: 52 opt: 76 Z-score: 90.2 bits: 25.2 E(): 5.2  
 Smith-Waterman score: 76; 28.713% identity (31.183% ungapped) in 101 aa overlap  
 (72-172:1-93)

```

      50      60      70      80      90     100
SYHT0H SFHHVELWCADAASAAGRFSFALGAPLAARSDLSTGNSAHASLLLRSGALAFLLFTAPYAP
      . . . . : . . . : . . . . : . . . . :
ALLERG                                ADLGYGPATPAA-----PAAGYTPATPAAP
                                   10      20

      110     120     130     140     150     160
SYHT0H PPQEAATAATASIPSFSADAARTFAAAHGLAVRSVGVRVADAAEAFRVSVAGGARPAFAP
      : : : . . . : : : : . . . : : . . . . . : :
ALLERG AGAEPAGKATTEEQKLIEKINAGFKAA--LAA-AAGVPPADKYNTFVATFGAASNKAFAE
      30      40      50      60      70      80

```

```

      170      180      190      200      210      220
SYHT0H ADLGHGFGLAEEVELYGDVVLRFVSYPDETDLPFLPGFERVSSPGAVDYGLTRFDHVGVN
.  . . : : :
ALLERG GLSGEPKGAAESSSKAALTSKLDAAAYKLAYKTAEGATPEAKYDAYVATVSEALRKIAGTL
      90      100      110      120      130      140

```

```

>>ALLERGEN_2012_2|gi|21725626|emb|CAD38394.1| unnamed pr (287 aa)
  initn: 52 init1: 52 opt: 76 Z-score: 90.2 bits: 25.2 E(): 5.2
Smith-Waterman score: 76; 28.713% identity (31.183% ungapped) in 101 aa overlap
(72-172:1-93)

```

```

      50      60      70      80      90      100
SYHT0H SFHHVELWCADAASAAGRFSFALGAPLAARSDLSTGNSAHASLLLRSGALAFLLFTAPYAP
      . . . : . . . : . . . : :
ALLERG                                ADLGYGPATPAA-----PAAGYTPATPAAP
                                10      20

```

```

      110      120      130      140      150      160
SYHT0H PPQEAATAATASIPSFSADAARTFAAAHGLAVRSVGVVADAAEAFRVSVAGGARPAFAP
      : : : . . : : : : : : : : : . . . . . : :
ALLERG AGAEPAGKATTEEQKLEKKNAGFKAA--LAA-AAGVPPADKYRTFVATFGAASNKAFAE
      30      40      50      60      70      80

```

```

      170      180      190      200      210      220
SYHT0H ADLGHGFGLAEEVELYGDVVLRFVSYPDETDLPFLPGFERVSSPGAVDYGLTRFDHVGVN
.  . . : : :
ALLERG GLSGEPKGAAESSSKAALTSKLDAAAYKLAYKTAEGATPEAKYDAYVATVSSALRIIAGTL
      90      100      110      120      130      140

```

```

>>ALLERGEN_2012_2|gi|21725632|emb|CAD38397.1| unnamed pr (287 aa)
  initn: 52 init1: 52 opt: 76 Z-score: 90.2 bits: 25.2 E(): 5.2
Smith-Waterman score: 76; 28.713% identity (31.183% ungapped) in 101 aa overlap
(72-172:1-93)

```

```

      50      60      70      80      90      100
SYHT0H SFHHVELWCADAASAAGRFSFALGAPLAARSDLSTGNSAHASLLLRSGALAFLLFTAPYAP
      . . . : . . . : . . . : :
ALLERG                                ADLGYGPATPAA-----PAAGYTPATPAAP
                                10      20

```

```

      110      120      130      140      150      160
SYHT0H PPQEAATAATASIPSFSADAARTFAAAHGLAVRSVGVVADAAEAFRVSVAGGARPAFAP
      : : : . . : : : : : : : : . . . . . : :
ALLERG AGAEPAGKATTEEQKLEKKNAGFKAA--LAA-AAGVPPADKYRTFVATFGAASNKAFAE
      30      40      50      60      70      80

```

```

      170      180      190      200      210      220
SYHT0H ADLGHGFGLAEEVELYGDVVLRFVSYPDETDLPFLPGFERVSSPGAVDYGLTRFDHVGVN
.  . . : : :
ALLERG GLSGEPKGAAESSSKAALTSKLDAAAYKLAYKTAEGATPEAKYDAYVATVSSALRIIAGTL
      90      100      110      120      130      140

```

```

>>ALLERGEN_2012_2|gi|21725606|emb|CAD38384.1| unnamed pr (287 aa)
  initn: 52 init1: 52 opt: 76 Z-score: 90.2 bits: 25.2 E(): 5.2
Smith-Waterman score: 76; 28.713% identity (31.183% ungapped) in 101 aa overlap
(72-172:1-93)

```

```

      50      60      70      80      90      100
SYHT0H SFHHVELWCADAASAAGRFSFALGAPLAARSDLSTGNSAHASLLLRSGALAFLLFTAPYAP
      . . . : . . . : . . . : :
ALLERG                                ADLGYGPATPAA-----PAAGYTPATPAAP
                                10      20

```

```

      110      120      130      140      150      160
SYHT0H PPQEAATAATASIPSFSADAARTFAAAHGLAVRSVGVVRVADAAEAFRVSVAGGARPAFAP
      : :  ::.  ..      : ::  ::.  :::  ::  ::.  ..  ....  :::
ALLERG AGAEPAGKATTEEQKLIEKKNAGFKAA--LAA-AAGVPPADKYRTFVATFGAASNKAFAE
      30      40      50      60      70      80

```

```

      170      180      190      200      210      220
SYHT0H ADLGHGFGLAEVELYGDVVLRFVSYPDETDLPFLPGFERVSSPGAVDYGGLTRFDHVGNV
      .  ..  :  ::
ALLERG GLSGEPKGAAESSSKAALTSKLDAAAYKLAYKTAEGATPEAKYDAYVATVSSALRIIAGTL
      90      100      110      120      130      140

```

```

>>ALLERGEN_2012_2|gi|21725610|emb|CAD38386.1| unnamed pr (287 aa)
  initn: 52 initl: 52 opt: 76 Z-score: 90.2 bits: 25.2 E(): 5.2
Smith-Waterman score: 76; 28.713% identity (31.183% ungapped) in 101 aa overlap
(72-172:1-93)

```

```

      50      60      70      80      90      100
SYHT0H SFHHVELWCADAASAAGRFSFALGAPLAARSDLSTGNSAHASLLLRSGALAFLLFTAPYAP
      .:::  :  ..  :.      :  ..  ..:  ::
ALLERG                                ADLGYGPATPAA-----PAAGYTPATPAAP
                                10                                20

```

```

      110      120      130      140      150      160
SYHT0H PPQEAATAATASIPSFSADAARTFAAAHGLAVRSVGVVRVADAAEAFRVSVAGGARPAFAP
      : :  ::.  ..      : ::  ::.  :::  ::  ::.  ..  ....  :::
ALLERG AGAEPAGKATTEEQKLIEKKNAGFKAA--LAA-AAGVPPADKYRTFVATFGAASNKAFAE
      30      40      50      60      70      80

```

```

      170      180      190      200      210      220
SYHT0H ADLGHGFGLAEVELYGDVVLRFVSYPDETDLPFLPGFERVSSPGAVDYGGLTRFDHVGNV
      .  ..  :  ::
ALLERG GLSGEPKGAAESSSKAALTSKLDAAAYKLAYKTAEGATPEAKYDAYVATVSEALSIIAGTL
      90      100      110      120      130      140

```

```

>>ALLERGEN_2012_2|gi|21725608|emb|CAD38385.1| unnamed pr (287 aa)
  initn: 52 initl: 52 opt: 76 Z-score: 90.2 bits: 25.2 E(): 5.2
Smith-Waterman score: 76; 28.713% identity (31.183% ungapped) in 101 aa overlap
(72-172:1-93)

```

```

      50      60      70      80      90      100
SYHT0H SFHHVELWCADAASAAGRFSFALGAPLAARSDLSTGNSAHASLLLRSGALAFLLFTAPYAP
      .:::  :  ..  :.      :  ..  ..:  ::
ALLERG                                ADLGYGPATPAA-----PAAGYTPATPAAP
                                10                                20

```

```

      110      120      130      140      150      160
SYHT0H PPQEAATAATASIPSFSADAARTFAAAHGLAVRSVGVVRVADAAEAFRVSVAGGARPAFAP
      : :  ::.  ..      : ::  ::.  :::  ::  ::.  ..  ....  :::
ALLERG AGAEPAGKATTEEQKLIEKINAGFKAA--LAA-AAGVPPADKYNTFVATFGAASNKAFAE
      30      40      50      60      70      80

```

```

      170      180      190      200      210      220
SYHT0H ADLGHGFGLAEVELYGDVVLRFVSYPDETDLPFLPGFERVSSPGAVDYGGLTRFDHVGNV
      .  ..  :  ::
ALLERG GLSGEPKGAAESSSKAALTSKLDAAAYKLAYKTAEGATPEAKYDAYVATVSSALRIIAGTL
      90      100      110      120      130      140

```

```

>>ALLERGEN_2012_2|gi|21725628|emb|CAD38395.1| unnamed pr (287 aa)
  initn: 52 initl: 52 opt: 76 Z-score: 90.2 bits: 25.2 E(): 5.2
Smith-Waterman score: 76; 28.713% identity (31.183% ungapped) in 101 aa overlap
(72-172:1-93)

```

```

              50              60              70              80              90              100
SYHTOH SFHHVELWCADAASAAGRFSFALGAPLAARSDLSTGNSAHASLLLRSGALAFLEFTAPYAP
              .:. .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .:
ALLERG ADLGYGPATPAA-----PAAGYTPATPAAP
                      10                      20

```

	170	180	190	200	210	220
SYHT0H	ADLGHGFGGLAEVELYGDVVLRFVSYPDETDLFPFLPGFERVSSPGAVDYGLTRFDHVVGNV					
	. : . : . :					
ALLERG	GLSGEPKGAAESSKAAALTSKLDAAAYKLAYKTAEGATPEAKYDAYVATVSSALRIIAGTL					
	90	100	110	120	130	140

```

              50          60          70          80          90          100
SYHT0H  SFHHVELWCADAASAAGRFSFALGAPLAARSDLSTGNSAHASLLLRSGALAFLLFTAPYAP
              . . . . . : . . . . : . . . . :
ALLERG  MAVHQYTVALFLAVALVAGPAGSYAADLGYGPATPAA-----PAAGYTPATPAAP
              10          20          30          40          50

```

SYHT0H ADLGHGFGGLAEVELYGDVVLRFVSYSPDETDLFPFLPGFERVSSPGAVDYGLTRFDHVVGNV  
 . . . : :  
 ALLERG GLSGEPKGAEESSSKAALTSKLDAAAYKLAYKTAEGATPEAKYDAYVATVSEALRIIAGTL  
 110 120 130 140 150 160

SYHT0H RSDRFPVLSFHHVELWCADAASAAGRFSFALGAPLAARS-DLSTGNSAHASLLLRSGALA  
ALLERG MAVHQYTVALFLAVALVAGPAASYAADLGYPATPAA-----PAAG

40 50 60 70 80 90  
... : . . : . : . :  
10 20 30 40

```

          160          170          180          190          200          210
SYHT0H  GGARPAFAPADLGHGFGGLAEVELYGDVVLRFVSYPDETDLPFLPGFERVSSPGAVDYGLT
...   ::: .  . :  ::
ALLERG  AASNKAFAEGLSGEPKGAEESSSKAALTSKLDAAAYKLAYKTAEGATPEAKYDAYVATLSE
          100          110          120          130          140          150

```

```

>>ALLERGEN_2012_2|gi|13430402|gb|AAK25823.1| group V all (275 aa)
  initn: 70 initl: 70 opt: 74 Z-score: 88.0 bits: 24.7 E(): 6.9
Smith-Waterman score: 77; 25.301% identity (28.966% ungapped) in 166 aa overlap
(2-158:19-172)

```

```

              10          20          30          40
SYHT0H      MPPTPATATGAAAAAVTPEHAARSF-PRVVRVNPNSDRFPVLS
              :: :   :: . :   :: :   : . :   : . .
ALLERG  PISVTAPPPQLPRPPATPPPPPPQLGASPYKLGPSPKARSERPAIV---PPADKY----
              10          20          30          40          50

```

```

          50          60          70          80          90          100
SYHT0H  FHHVELWCADAASAAGRFSFALGAPLAARSDL-STGNSAHASLLLRSGALAFLLFTAPYAP
. . :   . . . . . : :   . .   . . . . :   :   .   : . .   .   :
ALLERG  ---RTFVATFGAASNK-AFAEGLSGEPKGAEESSSKAALTSKLDAAAYKLAYKTAEGATP
          60          70          80          90          100

```

```

          110          120          130          140          150
SYHT0H  PPQEAATAATAS----IPSFSAADAARTFAAAHGLAVRSVG-VRVADAAEA-FRVSV-AGG
. :   : : :   :   . . . . . : :   . :   . :   . :   . :   . :
ALLERG  EAKYDAYVATLSEALRIIAGTLEVHAVKPAAEEVKVIPAGELQVIEKVDAAFKVAATAAN
          110          120          130          140          150          160

```

```

          160          170          180          190          200          210
SYHT0H  ARPAFAPADLGHGFGGLAEVELYGDVVLRFVSYPDETDLPFLPGFERVSSPGAVDYGLTRF
: :
ALLERG  AAPANDKFTVFEEAFNNAIKAGTGGAYESYKFIPALEAAVKQAYAATVATAPEVKYTVFE
          170          180          190          200          210          220

```

```

>>ALLERGEN_2012_2|gi|29500897|emb|CAD87529.1| phl p5a al (284 aa)
  initn: 52 initl: 52 opt: 73 Z-score: 86.7 bits: 24.5 E(): 8.2
Smith-Waterman score: 73; 32.000% identity (33.333% ungapped) in 75 aa overlap
(98-172:19-90)

```

```

          70          80          90          100          110          120
SYHT0H  LAARSDLSTGNSAHASLLLRSGALAFLLFTAPYAPPPQEAATAATASIPSFSAADAARTFAA
: : :   : :   : :   : :   : :
ALLERG      ADLGYGPATPAAPAAGYTPAAPAGAEPAGKATTEEQKLIKINAGFKA
              10          20          30          40

```

```

          130          140          150          160          170          180
SYHT0H  AHGLAVRSVGVRVADAAEAFRVSVAGGARPAFAPADLGHGFGGLAEVELYGDVVLRFVSY
:   : :   . : :   : :   . :   . . . . . : :   . :   : :
ALLERG  A--LAA-AAGVPPADKYRTFVATFGAASNKAFAEGLSGEPKGAEESSSKAALTSKLDAA
          50          60          70          80          90          100

```

```

          190          200          210          220          230          240
SYHT0H  DETDLPFLPGFERVSSPGAVDYGLTRFDHVVGNVPEMAPVIDYMKGFLGFHEFAEFTAED
ALLERG  KLAYKTAEGATPEAKYDAYVATLSEALRIIAGTLEVHAVKPAAEEVKVIPAGELQVIEK
          110          120          130          140          150          160

```

```

439 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]

```

```

start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.080 Display time: 0.050

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 30 aa
>SYHT0H2_I_23
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 4.7147+/-0.00291; mu= -0.9595+/- 0.152
mean_var=35.3312+/- 9.607, 0's: 27 Z-trim: 27 B-trim: 23 in 1/40
Lambda= 0.2158

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.010
The best scores are:

```

					opt	bits	E(1603)
ALLERGEN_2012_2	gi 22135348 gb AAM93157.1	trypsi	( 219)		51	22	2.9
ALLERGEN_2012_2	gi 162927 gb AAA30478.1	alpha-s1	( 76)		45	20	3.6
ALLERGEN_2012_2	gi 291482314 emb CBK62697.1	ragw	( 96)		45	20	4.6
ALLERGEN_2012_2	gi 3703107 gb AAC63045.1	glycini	( 507)		51	22	6.9
ALLERGEN_2012_2	gi 224036293 pdb 3C3V A Chain A,		( 510)		51	22	6.9
ALLERGEN_2012_2	gi 5712199 gb AAD47382.1	glycini	( 530)		51	22	7.2
ALLERGEN_2012_2	gi 199732457 gb ACH91862.1	arach	( 530)		51	22	7.2
ALLERGEN_2012_2	gi 21314465 gb AAM46958.1	AF51085	( 538)		51	22	7.3
ALLERGEN_2012_2	gi 159793201 gb ABW98945.1	alpha	( 172)		45	20	8.3
ALLERGEN_2012_2	gi 159793197 gb ABW98943.1	alpha	( 205)		45	20	10

```

>>ALLERGEN_2012_2|gi|22135348|gb|AAM93157.1| trypsin inh (219 aa)
initn: 43 initl: 43 opt: 51 Z-score: 94.7 bits: 21.8 E(): 2.9
Smith-Waterman score: 51; 32.143% identity (33.333% ungapped) in 28 aa overlap (3-30:171-197)

SYHT0H
MVEHDTLVYSKNIKDTVSEDQRAIETFQQR
::: : . . . : : . . . . . : :
ALLERG QQQDSHQKVHRFDEGDLIAVPTGVAFWMYNDHDTDVAVSLTDTNNNDNQ-LDQFPRRFN
150 160 170 180 190

ALLERG LAGNHEQEFLRYQQQSRRRS
200 210

>>ALLERGEN_2012_2|gi|162927|gb|AAA30478.1| alpha-s1-case (76 aa)
initn: 36 initl: 36 opt: 45 Z-score: 93.0 bits: 19.9 E(): 3.6
Smith-Waterman score: 45; 45.000% identity (47.368% ungapped) in 20 aa overlap
(10-29:2-20)

SYHT0H MVEHDTLVYSKNIKDTVSEDQRAIETFQQR
::: . . : : : : :
ALLERG LSKDIGSESTEDQ-AMEDIKQMEAESISSSEEIVPNSVEQKQIQKEDVPSEK
10 20 30 40 50

>>ALLERGEN_2012_2|gi|291482314|emb|CBK62697.1| ragweed h (96 aa)
initn: 27 initl: 27 opt: 45 Z-score: 91.1 bits: 19.9 E(): 4.6
Smith-Waterman score: 45; 34.615% identity (36.000% ungapped) in 26 aa overlap (1-25:2-27)

```



```

          10          20          30
SYHT0H  MVEHDTLVYSKN-IKDTVSEDQRAIETFQQR
          . . . . . : . . . : . . . : .
ALLERG  KLCEKPSVTWSGNKVKQTDKCDKRCIEWEGAKHGACHKRDSKASCFCYFDCDPTKNPGPP
          10          20          30          40          50          60

```

>>ALLERGEN\_2012\_2|gi|3703107|gb|AAC63045.1| glycinin [Ar (507 aa)  
 initn: 43 initl: 43 opt: 51 Z-score: 88.0 bits: 21.8 E(): 6.9  
 Smith-Waterman score: 51; 32.143% identity (33.333% ungapped) in 28 aa overlap (3-30:146-172)

```

          10          20          30
SYHT0H                                MVEHDTLVYSKNIKDTVSEDQRAIETFQQR
          . . . : . . . : . . . . . : . .
ALLERG  QQRDASHQKVHRFDEGDLIAVPTGVAFWLYNDHDTDVAVVSLTDTNNNDNQ-LDQFPRRFN
          120          130          140          150          160          170
ALLERG  LAGNTEQEFLRYQQQSRQSRRRSLPYPSPQSQPRQEEREFSPRGQHSRRERAGQEEEN
          180          190          200          210          220          230

```

>>ALLERGEN\_2012\_2|gi|224036293|pdb|3C3V|A Chain A, Cryst (510 aa)  
 initn: 43 initl: 43 opt: 51 Z-score: 88.0 bits: 21.8 E(): 6.9  
 Smith-Waterman score: 51; 32.143% identity (33.333% ungapped) in 28 aa overlap (3-30:149-175)

```

          10          20          30
SYHT0H                                MVEHDTLVYSKNIKDTVSEDQRAIETFQQR
          . . . : . . . : . . . . . : . .
ALLERG  QQQDSHQKVHRFNEGDLIAVPTGVAFWLYNDHDTDVAVVSLTDTNNNDNQ-LDQFPRRFN
          120          130          140          150          160          170
ALLERG  LAGNHEQEFLRYQQQSRQSRRRSLPYPSPQSQPRQEEREFSPRGQHSRRERAGQEEEH
          180          190          200          210          220          230

```

>>ALLERGEN\_2012\_2|gi|5712199|gb|AAD47382.1| glycinin [Ar (530 aa)  
 initn: 43 initl: 43 opt: 51 Z-score: 87.7 bits: 21.8 E(): 7.2  
 Smith-Waterman score: 51; 32.143% identity (33.333% ungapped) in 28 aa overlap (3-30:169-195)

```

          10          20          30
SYHT0H                                MVEHDTLVYSKNIKDTVSEDQRAIETFQQR
          . . . : . . . : . . . . . : . .
ALLERG  QQQDSHQKVHRFNEGDLIAVPTGVAFWLYNDHDTDVAVVSLTDTNNNDNQ-LDQFPRRFN
          140          150          160          170          180          190
ALLERG  LAGNHEQEFLRYQQQSRQSRRRSLPYPSPHSPRPREEREFRPRGQHSRRERAGQEEED
          200          210          220          230          240          250

```

>>ALLERGEN\_2012\_2|gi|199732457|gb|ACH91862.1| arachin Ar (530 aa)  
 initn: 43 initl: 43 opt: 51 Z-score: 87.7 bits: 21.8 E(): 7.2  
 Smith-Waterman score: 51; 32.143% identity (33.333% ungapped) in 28 aa overlap (3-30:169-195)

```

          10          20          30
SYHT0H                                MVEHDTLVYSKNIKDTVSEDQRAIETFQQR
          . . . : . . . : . . . . . : . .
ALLERG  QQQDSHQKVHRFNEGDLIAVPTGVAFWLYNDHDTDVAVVSLTDTNNNDNQ-LDQFPRRFN
          140          150          160          170          180          190
ALLERG  LAGNHEQEFLRYQQQSRQSRRRSLPYPSPQSQPRQEEREFSPRGQHSRRERAGQEEEN
          200          210          220          230          240          250

```

```
>>ALLERGEN_2012_2|gi|21314465|gb|AAM46958.1|AF510854_1 a (538 aa)
  initn: 43 initl: 43 opt: 51 Z-score: 87.5 bits: 21.8 E(): 7.3
Smith-Waterman score: 51; 32.143% identity (33.333% ungapped) in 28 aa overlap (3-30:169-195)
```

```

                                10      20      30
SYHT0H                        MVEHDTLVYSKNIKDTVSEDQRAIETFQQR
                                .::: . . . .:: .::.. . . : .:
ALLERG QQQDSHQKVHRFDEGDLIAVPTGVAFWMYNDHDTDVVAVSLTDTNNNDNQ-LDQFPRRFN
      140      150      160      170      180      190

ALLERG LAGNHEQEFLRYQQSRRRSLPYSPYSPQTQPKQEDREFSPRGQHGRERERAGQEENEGG
      200      210      220      230      240      250
```

```
>>ALLERGEN_2012_2|gi|159793201|gb|ABW98945.1| alpha S1 c (172 aa)
  initn: 36 initl: 36 opt: 45 Z-score: 86.5 bits: 19.9 E(): 8.3
Smith-Waterman score: 45; 45.000% identity (47.368% ungapped) in 20 aa overlap
(10-29:14-32)
```

```

                                10      20      30
SYHT0H      MVEHDTLVYSKNIKDTVSEDQRAIETFQQR
                                .::: . . . .:: .::.. . . : .:
ALLERG FSEVFGKEKVNELSKDIGSESTEDQ-AMEDIKQMEAESISSSEEIVPNSVEQKHIQKEDV
      10      20      30      40      50

ALLERG PSERYLGYLEQLRLKKYKVPQLEIVPNSAEERLHSMKEGIHAQQKEPMIGVNQELAYFY
      60      70      80      90      100      110
```

```
>>ALLERGEN_2012_2|gi|159793197|gb|ABW98943.1| alpha S1 c (205 aa)
  initn: 36 initl: 36 opt: 45 Z-score: 85.1 bits: 19.9 E(): 10
Smith-Waterman score: 45; 45.000% identity (47.368% ungapped) in 20 aa overlap
(10-29:47-65)
```

```

                                10      20      30
SYHT0H                        MVEHDTLVYSKNIKDTVSEDQRAIETFQQR
                                .::: . . . .:: .::.. . . : .:
ALLERG LPQEVNLNENLLRFFVAPFPEVFGKEKVNELSKDIGSESTEDQ-AMEDIKQMEAESISSSE
      20      30      40      50      60      70

ALLERG EIVPNSVEQKHIQKEDVPSERYLGYLEQLRLKKYKVPQLEIVPNSAEERLHSMKEGIHA
      80      90      100      110      120      130
```

```
30 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.010 Display time: 0.000
```

```
Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448
```

```
/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 40 aa
>SYHT0H2_I_24
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library
```

```
367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 3.4130+/-0.00286; mu= 9.5190+/- 0.150
```

mean\_var=41.7198+/-10.742, 0's: 9 Z-trim: 9 B-trim: 0 in 0/43  
Lambda= 0.1986

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2  
join: 36, opt: 24, gap-pen: -12/-2, width: 16  
Scan time: 0.020

The best scores are:

					opt	bits	E(1603)
ALLERGEN_2012_2	gi 170722 gb AAA34281.1	pre-alpha	( 262)	56	22	3.8	
ALLERGEN_2012_2	gi 473876 gb AAA17741.1	alpha-gl	( 287)	56	22	4.1	
ALLERGEN_2012_2	gi 5381323 gb AAD42943.1	AF091841	( 148)	52	21	5.8	
ALLERGEN_2012_2	gi 1170095 sp P46419.1	GSTM1_DERP	( 219)	52	21	7.5	
ALLERGEN_2012_2	gi 60920878 gb AA37326.1	glutat	( 219)	52	21	7.5	
ALLERGEN_2012_2	gi 105969543 gb ABF81661.1	EXPB1	( 99)	48	19	9.7	

>>ALLERGEN\_2012\_2|gi|170722|gb|AAA34281.1| pre-alpha-/be (262 aa)  
initn: 28 initl: 28 opt: 56 Z-score: 92.5 bits: 22.1 E(): 3.8  
Smith-Waterman score: 56; 30.000% identity (30.000% ungapped) in 30 aa overlap (2-31:56-85)

		10	20	30
SYHT0H		MPLPTVVPKMDPHPRGASWKKKTFQPR	LQSK	
		:: : :: :	.. :: :	
ALLERG	PQLQPQNPSQQQPQEQVPLVQQQQLGQQQFPFPQQPYQPQPFPSQQPYLQLQPF	LQPF	LQPF	LQPF
	30 40 50 60 70 80			
	40			
SYHT0H	WIDVISPLT			
ALLERG	LPYSQPQPFPRPQQPYQPQPYSPQQPISQQQQQQQQQQQQQQQQQQIIQQILQQQLI			
	90 100 110 120 130 140			

>>ALLERGEN\_2012\_2|gi|473876|gb|AAA17741.1| alpha-gliadin (287 aa)  
initn: 28 initl: 28 opt: 56 Z-score: 92.1 bits: 22.1 E(): 4.1  
Smith-Waterman score: 56; 30.000% identity (30.000% ungapped) in 30 aa overlap (2-31:56-85)

		10	20	30
SYHT0H		MPLPTVVPKMDPHPRGASWKKKTFQPR	LQSK	
		:: : :: :	.. :: :	
ALLERG	PQLQPQNPSQQQPQEQVPLVQQQQLGQQQFPFPQQPYQPQPFPSQQPYLQLQPF	LQPF	LQPF	LQPF
	30 40 50 60 70 80			
	40			
SYHT0H	WIDVISPLT			
ALLERG	LPYSQPQPFPRPQQPYQPQPYSPQQPISQQQQQQQQQQQQQQQQQQIIQQILQQQLI			
	90 100 110 120 130 140			

>>ALLERGEN\_2012\_2|gi|5381323|gb|AAD42943.1|AF091841\_1 2S (148 aa)  
initn: 40 initl: 40 opt: 52 Z-score: 89.4 bits: 20.6 E(): 5.8  
Smith-Waterman score: 52; 26.087% identity (26.087% ungapped) in 23 aa overlap (11-33:36-58)

		10	20	30	40
SYHT0H		MPLPTVVPKMDPHPRGASWKKKTFQPR	LQSKWIDVISPLT		
		. . :: :...	: : ..		
ALLERG	IVLAVLFAAALVSASAHKTVVTTSVAAEEGEEENQRCGEWESRQCQMRHCMQWMRSMRGQY				
	10 20 30 40 50 60				
ALLERG	EESFLRSAEANQGQFEHFRECCNELRDVKSHCRCEALRCMMRQMQQEYGMQEQEMQMQM				
	70 80 90 100 110 120				

>>ALLERGEN\_2012\_2|gi|1170095|sp|P46419.1|GSTM1\_DERPT Rec (219 aa)

initn: 39 initl: 39 opt: 52 Z-score: 87.3 bits: 20.8 E(): 7.5  
 Smith-Waterman score: 52; 35.294% identity (41.379% ungapped) in 34 aa overlap (3-32:182-214)

```

              10      20
SYHT0H      MPLPTVVPKMDPHPRGASWKKKTFQPRL----
              :  :  .:.  ::  ...  ::  :..
ALLERG ANISYVDFNLYEYLCHVKVMVPEVFGQFENLKRYVERMESLPRVSDYIKKQ-QPKTFNAP
              160      170      180      190      200      210

```

```

      30      40
SYHT0H QSKWIDVISPLT
      :::
ALLERG TSKWNASYA

```

>>ALLERGEN\_2012\_2|gi|60920878|gb|AAX37326.1| glutathione (219 aa)  
 initn: 39 initl: 39 opt: 52 Z-score: 87.3 bits: 20.8 E(): 7.5  
 Smith-Waterman score: 52; 35.294% identity (41.379% ungapped) in 34 aa overlap (3-32:182-214)

```

              10      20
SYHT0H      MPLPTVVPKMDPHPRGASWKKKTFQPRL----
              :  :  .:.  ::  ...  ::  :..
ALLERG ANISYVDFYLYEYLCRVKVMVPEVFGQFENLKRYVERMESLPRVSDYIKKQ-QPKTFNAP
              160      170      180      190      200      210

```

```

      30      40
SYHT0H QSKWIDVISPLT
      :::
ALLERG TSKWNASYA

```

>>ALLERGEN\_2012\_2|gi|105969543|gb|ABF81661.1| EXPB10 [Ze (99 aa)  
 initn: 35 initl: 35 opt: 48 Z-score: 85.3 bits: 19.3 E(): 9.7  
 Smith-Waterman score: 48; 42.857% identity (42.857% ungapped) in 14 aa overlap (13-26:51-64)

```

              10      20      30      40
SYHT0H      MPLPTVVPKMDPHPRGASWKKKTFQPRLQSKWIDVISPLT
              :  :  :.:.  .  .  :
ALLERG ALLVKYVDGDGDIVAVDIKEKGSPTYEPLKHSWGAIWKDSDKPIKGPITVRLTTEGGTK
              30      40      50      60      70      80

```

```

ALLERG TVYDDVIPAEWKPNTAYTT
      90

```

40 residues in 1 query sequences  
 367796 residues in 1603 library sequences  
 Scomplib [34t11]  
 start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012  
 Scan time: 0.020 Display time: 0.000

Function used was FASTA [version 3.4t11 Apr 17, 2002]  
 FASTA searches a protein or DNA sequence data bank  
 version 3.4t11 Apr 17, 2002  
 Please cite:  
 W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

/grid/.qsub\_tmp/1330373600kde\_agent/grid\_NbCuMBoIO: 65 aa  
 >SYHT0H2\_I\_25

vs /data/fasta/ALLERGEN\_2012\_2 library  
searching /data/fasta/ALLERGEN\_2012\_2 library

367796 residues in 1603 sequences

Expectation\_n fit:  $\rho(\ln(x)) = 4.8252 \pm 0.00351$ ;  $\mu = 0.8977 \pm 0.184$   
mean\_var=42.9082 $\pm$ 11.183, 0's: 7 Z-trim: 7 B-trim: 0 in 0/42  
Lambda= 0.1958

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2  
join: 36, opt: 24, gap-pen: -12/-2, width: 16  
Scan time: 0.020

The best scores are:

			opt	bits	E(1603)
ALLERGEN_2012_2	gi 23616947 dbj BAC20650.1  putat	( 160)	56	23	2.2
ALLERGEN_2012_2	gi 114152864 sp Q01883.2 RAG1_ORY	( 163)	55	23	2.8
ALLERGEN_2012_2	gi 1304217 dbj BAA07773.1  allerg	( 109)	52	22	3.4
ALLERGEN_2012_2	gi 1304216 dbj BAA07772.1  allerg	( 111)	52	22	3.5
ALLERGEN_2012_2	gi 1304218 dbj BAA07774.1  allerg	( 113)	52	22	3.5
ALLERGEN_2012_2	gi 1703445 sp P54958.1 ASP2_BLAG	( 352)	57	23	3.9
ALLERGEN_2012_2	gi 145105726 gb ABP35603.1  Bla g	( 352)	57	23	3.9
ALLERGEN_2012_2	gi 1398915 dbj BAA07711.1  allerg	( 160)	52	22	4.9
ALLERGEN_2012_2	gi 47117012 sp Q7M4I5.1 PA2_APIDO	( 134)	51	21	5.1
ALLERGEN_2012_2	gi 218193 dbj BAA01998.1  allerge	( 165)	52	22	5.1
ALLERGEN_2012_2	gi 1398913 dbj BAA07710.1  allerg	( 166)	52	22	5.1
ALLERGEN_2012_2	gi 114152865 sp Q01882.2 RAG2_ORY	( 166)	52	22	5.1
ALLERGEN_2012_2	gi 897811 emb CAA24933.1  unnamed	( 101)	49	21	5.7
ALLERGEN_2012_2	gi 7435005 pir A59055 phospholip	( 134)	50	21	6.1
ALLERGEN_2012_2	gi 24638082 sp Q9BMK4.1 PA2_APICC	( 134)	50	21	6.1
ALLERGEN_2012_2	gi 163825 gb AAC37318.1  major al	( 92)	47	20	7.8

>>ALLERGEN\_2012\_2|gi|23616947|dbj|BAC20650.1| putative a (160 aa)  
initn: 48 init1: 48 opt: 56 Z-score: 96.7 bits: 22.8 E(): 2.2  
Smith-Waterman score: 56; 30.769% identity (32.432% ungapped) in 39 aa overlap (5-42:90-127)

```

                                10      20      30
SYHT0H                      MTHNPTILRKTLPLYKEVHFIWRGHA-EITSPPC
                                :. . . . :. . . . :
ALLERG RQCAGGAVDEQVRQDCCRQLAAIDDSFCRCPALSHMLVGMKELGAPAKGQPMDEVFPGC
      60      70      80      90     100     110

      40      50      60
SYHT0H LRRGDQLRLGQLQQLIWPRFVISLTITLRRLLQ
      :. . . :. .
ALLERG -RRGDMKRVAASLP AFCNV D I P I G I G G V C Y W L S Y P M N P A T G H
      120     130     140     150     160
```

>>ALLERGEN\_2012\_2|gi|114152864|sp|Q01883.2|RAG1\_ORYSJ Re (163 aa)  
initn: 42 init1: 42 opt: 55 Z-score: 95.1 bits: 22.5 E(): 2.8  
Smith-Waterman score: 55; 43.333% identity (50.000% ungapped) in 30 aa overlap (14-41:102-129)

```

                                10      20      30      40
SYHT0H                      MTHNPTILRKTLPLYKEVHFIWRGH--AEITSPPC LRRGDQLR
                                :. . . . :. :. :. :. :. :. :
ALLERG EQVWQDCCRQLAAVDDGWCRGALDHMLSGIYRELGATEAGHPMAEVF-PGC-RRGDLER
      80      90     100     110     120

      50      60
SYHT0H LGQLQQLIWPRFVISLTITLRRLLQ

ALLERG AAASLP AFCNV D I P N G P G G V C Y W L G Y P R T P R T G H
      130     140     150     160
```

```
>>ALLERGEN_2012_2|gi|1304217|dbj|BAA07773.1| allergenic (109 aa)
  initn: 42 initl: 42 opt: 52 Z-score: 93.5 bits: 21.7 E(): 3.4
Smith-Waterman score: 52; 43.333% identity (50.000% ungapped) in 30 aa overlap
(14-41:48-75)
```

```

                        10      20      30      40
SYHT0H                MTHNPTILRKTLPLYKEVHFIWRGH--AEITSPPCLRRGDQLR
                        .:.:.   ::  ::  : : : : : :
ALLERG EQLRQDCCRQLAAVDDSWCRCSALNHMVGGIYRELGATDVGHMAEVF-PGC-RRGDLER
                        20      30      40      50      60      70

                        50      60
SYHT0H LGQLQQLIWPRFVISLTITLRLRQ

ALLERG ARATLPAFCNVDIPNGTGGVCYWLGYPRTPRTGH
                        80      90     100
```

```
>>ALLERGEN_2012_2|gi|1304216|dbj|BAA07772.1| allergenic (111 aa)
  initn: 42 initl: 42 opt: 52 Z-score: 93.3 bits: 21.7 E(): 3.5
Smith-Waterman score: 52; 43.333% identity (50.000% ungapped) in 30 aa overlap
(14-41:50-77)
```

```

                        10      20      30      40
SYHT0H                MTHNPTILRKTLPLYKEVHFIWRGH--AEITSPPCLRRGDQLR
                        .:.:.   ::  ::  : : : : : :
ALLERG EQLRQDCCRQLAAVDDSWCRCSALNHMVGGIYRELGATDVGHMAEVF-PGC-RRGDLER
                        20      30      40      50      60      70

                        50      60
SYHT0H LGQLQQLIWPRFVISLTITLRLRQ

ALLERG RRRSLPAFCNVDIPNGTGGVCYWLGYPRTPRTGH
                        80      90     100     110
```

```
>>ALLERGEN_2012_2|gi|1304218|dbj|BAA07774.1| allergenic (113 aa)
  initn: 42 initl: 42 opt: 52 Z-score: 93.2 bits: 21.7 E(): 3.5
Smith-Waterman score: 52; 43.333% identity (50.000% ungapped) in 30 aa overlap
(14-41:52-79)
```

```

                        10      20      30      40
SYHT0H                MTHNPTILRKTLPLYKEVHFIWRGH--AEITSPPCLRRGDQLR
                        .:.:.   ::  ::  : : : : : :
ALLERG QQLRQDCCRQLAAVDDSWCRCSALNHMVGGIYRELGATDVGHMAEVF-PGC-RRGDLER
                        30      40      50      60      70

                        50      60
SYHT0H LGQLQQLIWPRFVISLTITLRLRQ

ALLERG AAASLPAFCNVDIPNGTGGVCYWLGYPRTPRTGH
                        80      90     100     110
```

```
>>ALLERGEN_2012_2|gi|1703445|sp|P54958.1|ASP2_BLAG RecN (352 aa)
  initn: 43 initl: 43 opt: 57 Z-score: 92.5 bits: 23.2 E(): 3.9
Smith-Waterman score: 57; 46.667% identity (51.852% ungapped) in 30 aa overlap (1-
29:16-43)
```

```

                        10      20      30      40
SYHT0H                MTHNPTILRKTLPLYKEVH-FIWRGHAEITSPPCLRRGDQLRLGQ
                        .: : . : .: : : : : : : : : : :
ALLERG MIGLKLVTVLFVAVATITHAAELQR--VPLYKLHVHFINTQYAGITKIGNQNFLTVDST
                        10      20      30      40      50
```

```

      50      60
SYHT0H LQQLIWPRFVISLTITLRLRQ

ALLERG CNVVVASQECVGGACVCPNLQKYEKLKPKYISDGNVQVKFFDTGSAVGRGIEDSLTISNL
      60      70      80      90     100     110

>>ALLERGEN_2012_2|gi|145105726|gb|ABP35603.1| Bla g 2 al (352 aa)
  initn: 43 initl: 43 opt: 57 Z-score: 92.5 bits: 23.2 E(): 3.9
Smith-Waterman score: 57; 46.667% identity (51.852% ungapped) in 30 aa overlap (1-
29:16-43)

      10      20      30      40
SYHT0H      MTHNPTILRKTLPLYKEVH-FIWRGHAEITSPPCLLRRGDQLRLGQ
      .:: . : .::: :: :: .: ::
ALLERG MIGLKIVTVLFAVATITHAAELQR--VPLYKLVHVFINTQYAGITKIGNQNFLTTFVDSTS
      10      20      30      40      50

      50      60
SYHT0H LQQLIWPRFVISLTITLRLRQ

ALLERG CNVVVASQECVGGACVCPNLQKYEKLKPRYISDGNVQVKFFDTGSAVGRGIEDSLTIFNL
      60      70      80      90     100     110

>>ALLERGEN_2012_2|gi|1398915|dbj|BAA07711.1| allergenic (160 aa)
  initn: 42 initl: 42 opt: 52 Z-score: 90.6 bits: 21.7 E(): 4.9
Smith-Waterman score: 52; 43.333% identity (50.000% ungapped) in 30 aa overlap
(14-41:98-125)

      10      20      30      40
SYHT0H      MTHNPTILRKTLPLYKEVHFIWRGH--AEITSPPCLLRRGDQLR
      .::: .: .: .: :
ALLERG EQLRQDCCRQLAAVDDSWCRCSALNHMVGGIYRELGATDVGHMAEVF-PGC-RRGDLER
      70      80      90     100     110     120

      50      60
SYHT0H LGQLQLIWPRFVISLTITLRLRQ

ALLERG AAASLP AFCNV DIPNGTGGVCYWLGYTPRTPRTGH
      130     140     150     160

>>ALLERGEN_2012_2|gi|47117012|sp|Q7M4I5.1|PA2_APIDO RecN (134 aa)
  initn: 38 initl: 38 opt: 51 Z-score: 90.4 bits: 21.4 E(): 5.1
Smith-Waterman score: 51; 41.667% identity (41.667% ungapped) in 12 aa overlap
(21-32:7-18)

      10      20      30      40      50      60
SYHT0H MTHNPTILRKTLPLYKEVHFIWRGHAEITSPPCLLRRGDQLRLGQLQLIWPRFVISLTIT
      .: ::::: :
ALLERG      IIYPGTLWCCHGNVSSSPDELGRFKHTDSCCRSHDMCPDVMSAGES
      10      20      30      40

SYHT0H LRLRQ

ALLERG KHGLTNTASHTRLSCDCDDKFYDCLKNSSDTISSYFVGEMYFNILDTKCYKLEHPVTGCG
      50      60      70      80      90     100

>>ALLERGEN_2012_2|gi|218193|dbj|BAA01998.1| allergenic p (165 aa)
  initn: 42 initl: 42 opt: 52 Z-score: 90.4 bits: 21.7 E(): 5.1
Smith-Waterman score: 52; 43.333% identity (50.000% ungapped) in 30 aa overlap
(14-41:104-131)

```

```

              10      20      30      40
SYHT0H      MTHNPTILRKTLPLYKEVHFIWRGH--AEITSPPCLRRGDQLR
              .:.:.      ::  :.  : : :.:.  :
ALLERG  EQLAQDCCRELA AVDDSWCRCSALNHMVGGIYRELGATDVGHMAEVF-PGC-RRGDLER
              80      90      100      110      120      130

              50      60
SYHT0H  LGQLQQLIWPRFVISLTITLRRQLQ

ALLERG  AAASLP AFCNV DIPNGTGGVCYWLGYPRTPRTGH
              140      150      160

>>ALLERGEN_2012_2|gi|1398913|dbj|BAA07710.1| allergenic (166 aa)
  initn: 42 initl: 42 opt: 52 Z-score: 90.4 bits: 21.7 E(): 5.1
Smith-Waterman score: 52; 43.333% identity (50.000% ungapped) in 30 aa overlap
(14-41:105-132)

              10      20      30      40
SYHT0H      MTHNPTILRKTLPLYKEVHFIWRGH--AEITSPPCLRRGDQLR
              .:.:.      ::  :.  : : :.:.  :
ALLERG  EQLRQDCCRQLAAVDDSWCRCSALNHMVGGIYRELGATDVGHMAEVF-PGC-RRGDLER
              80      90      100      110      120      130

              50      60
SYHT0H  LGQLQQLIWPRFVISLTITLRRQLQ

ALLERG  AAASLP AFCNV DIPNGTGGVCYWLGYPRTPRTGH
              140      150      160

>>ALLERGEN_2012_2|gi|114152865|sp|Q01882.2|RAG2_ORYSJ Re (166 aa)
  initn: 42 initl: 42 opt: 52 Z-score: 90.4 bits: 21.7 E(): 5.1
Smith-Waterman score: 52; 43.333% identity (50.000% ungapped) in 30 aa overlap
(14-41:105-132)

              10      20      30      40
SYHT0H      MTHNPTILRKTLPLYKEVHFIWRGH--AEITSPPCLRRGDQLR
              .:.:.      ::  :.  : : :.:.  :
ALLERG  EQLRQDCCRQLAAVDDSWCRCSALNHMVGGIYRELGATDVGHMAEVF-PGC-RRGDLER
              80      90      100      110      120      130

              50      60
SYHT0H  LGQLQQLIWPRFVISLTITLRRQLQ

ALLERG  AAASLP AFCNV DIPNGTGGVCYWLGYPRTPRTGH
              140      150      160

>>ALLERGEN_2012_2|gi|897811|emb|CAA24933.1| unnamed prot (101 aa)
  initn: 41 initl: 41 opt: 49 Z-score: 89.4 bits: 20.8 E(): 5.7
Smith-Waterman score: 49; 47.826% identity (52.381% ungapped) in 23 aa overlap
(29-51:25-45)

              10      20      30      40      50      60
SYHT0H  MTHNPTILRKTLPLYKEVHFIWRGHAEITSPPCLRRGDQLRLGQLQQLIWPRFVISLTIT
              ::      .:.:.  ::  :.  :.  :.
ALLERG      EKLGGQQQPRQLQPRQGQQGYPTSPQQSGQQQL--GQQQQGYPTSPQQSGQG
              10      20      30      40      50

SYHT0H  LRRQLQ

ALLERG  QQQYDSPYHVS AEHQ AASLKVAKAQQQLAAQLPAMCRLEGGDALLASQ
              60      70      80      90      100

```



>>ALLERGEN\_2012\_2|gi|7435005|pir||A59055 phospholipase A (134 aa)  
 initn: 37 initl: 37 opt: 50 Z-score: 88.9 bits: 21.1 E(): 6.1  
 Smith-Waterman score: 50; 41.667% identity (41.667% ungapped) in 12 aa overlap  
 (21-32:7-18)

```

      10      20      30      40      50      60
SYHT0H MTHNPTILRKTLPLYKEVHFIWRGHAETSPPCLLRRGDLRLGQLQLIWPFRFVISLTIT
      .: :.....:
ALLERG      IIYPGTLWCGHGNVSSGPNELGRFKHTDACCRTDHMCPPDVMSAGES
              10      20      30      40

```

SYHT0H LRRLQ

```

ALLERG KHGLTNTASHTRLSCDDTFYDCLKNSEKISSYFVGKMYFNLIIDTKCYKLEHPVTGCG
      50      60      70      80      90     100

```

>>ALLERGEN\_2012\_2|gi|24638082|sp|Q9BMK4.1|PA2\_APICC RecN (134 aa)  
 initn: 37 initl: 37 opt: 50 Z-score: 88.9 bits: 21.1 E(): 6.1  
 Smith-Waterman score: 50; 41.667% identity (41.667% ungapped) in 12 aa overlap  
 (21-32:7-18)

```

      10      20      30      40      50      60
SYHT0H MTHNPTILRKTLPLYKEVHFIWRGHAETSPPCLLRRGDLRLGQLQLIWPFRFVISLTIT
      .: :.....:
ALLERG      IIYPGTLWCGHGNVSSGPNELGRFKHTDACCRTDHMCPPDVMSAGES
              10      20      30      40

```

SYHT0H LRRLQ

```

ALLERG KHGLTNTASHTRLSCDDTFYDCLKNSEKISSYFVGKMYFNLIIDTKCYKLEHPVTGCG
      50      60      70      80      90     100

```

>>ALLERGEN\_2012\_2|gi|163825|gb|AAC37318.1| major allergen (92 aa)  
 initn: 40 initl: 40 opt: 47 Z-score: 87.1 bits: 20.2 E(): 7.8  
 Smith-Waterman score: 47; 34.483% identity (37.037% ungapped) in 29 aa overlap  
 (14-42:10-36)

```

      10      20      30      40      50      60
SYHT0H MTHNPTILRKTLPLYKEVHFIWRGHAETSPPCLLRRGDLRLGQLQLIWPFRFVISLTIT
      .: . .: .: .: .: .: .: .: .:
ALLERG      MKGARVLVLLWAALLLIWGGNCEIC--PAVKRDVDLFLTGTPEYVEQVAQYKALP
              10      20      30      40      50

```

SYHT0H LRRLQ

```

ALLERG VVLENARILKNCVDAKMTEEDKENALSLLDKIYTSPLC
      60      70      80      90

```

65 residues in 1 query sequences  
 367796 residues in 1603 library sequences  
 Scomplib [34t11]  
 start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012  
 Scan time: 0.020 Display time: 0.000

Function used was FASTA [version 3.4t11 Apr 17, 2002]  
 FASTA searches a protein or DNA sequence data bank  
 version 3.4t11 Apr 17, 2002  
 Please cite:

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

```
/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 69 aa
>SYHT0H2_I_26
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library
```

```
367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 4.3276+/-0.00324; mu= 5.1316+/- 0.168
mean_var=49.9800+/-13.628, 0's: 3 Z-trim: 3 B-trim: 0 in 0/43
Lambda= 0.1814
```

```
FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.020
```

```
The best scores are:
                                opt bits E(1603)
ALLERGEN_2012_2|gi|71360928|emb|CAJ19705.1| non-s ( 114)  59  23  2.1
ALLERGEN_2012_2|gi|21926|emb|CAA36063.1| unnamed ( 295)  56  22  7.7
ALLERGEN_2012_2|gi|262232390|gb|ACY38525.1| aller ( 174)  53  21  8.7
```

```
>>ALLERGEN_2012_2|gi|71360928|emb|CAJ19705.1| non-specif (114 aa)
initn: 38 initl: 38 opt: 59 Z-score: 97.2 bits: 22.5 E(): 2.1
Smith-Waterman score: 59; 26.471% identity (28.125% ungapped) in 34 aa overlap
(12-43:27-60)
```

```

              10      20      30      40
SYHT0H      MRLWDTQPGVHCAQLDTSMDGMMLVFGKGIL--SCQLLQGQLGQ
              :.:. .... . ....: : :. .: :.
ALLERG MEMVNKIACFVLLCMVVVAPHAEALTCGQVTSTLAPCLPYLMNRGPLRNCCDGVKGLLGQ
              10      20      30      40      50      60
              50      60
SYHT0H LPRSELVISAPWVDLQIVQTFGNKVS
```

```
ALLERG AKTTVDRQAACTCLKSAASSFTGLNLGKAAALPNTCSVNIPYKISPSTDCSKVQ
              70      80      90      100     110
```

```
>>ALLERGEN_2012_2|gi|21926|emb|CAA36063.1| unnamed prote (295 aa)
initn: 69 initl: 44 opt: 56 Z-score: 87.1 bits: 22.0 E(): 7.7
Smith-Waterman score: 56; 37.037% identity (37.037% ungapped) in 27 aa overlap
(30-56:221-247)
```

```

              10      20      30      40      50
SYHT0H MRLWDTQPGVHCAQLDTSMDGMMLVFGKGILSCQLLQGQLGQLPRSELVISAPWVDLQ
              : : : : : : : : : : : : : : : :
ALLERG QQGFBVQAQQQQPQQQLGQGVSSQSSQQLGQCSFQQPQQQLGQQPQQQQVLQGTFLQPH
              200     210     220     230     240     250
              60
SYHT0H IVQTFGNKVS
```

```
ALLERG QIAHLEVMTSIALRTLPTMCSNVNPLYSSSTTSVPFSGTGVGAYL
              260     270     280     290
```

```
>>ALLERGEN_2012_2|gi|262232390|gb|ACY38525.1| allergen C (174 aa)
initn: 44 initl: 44 opt: 53 Z-score: 86.1 bits: 21.1 E(): 8.7
Smith-Waterman score: 53; 30.233% identity (32.500% ungapped) in 43 aa overlap
(29-68:2-44)
```

```

          10          20          30          40          50
SYHTOH MRLWDTQPGVHCAQLDTSMDGMMLVFGKGILSC-QLLQGQLGQLPRSELV--ISAPWVD
          : : : : . . . : : . . . : :
ALLERG          MKILLCLALVLASDAQPLPNVLTQVSGPWKT
                      10          20          30

          60
SYHTOH LQIVQTFGNKVS
          : : . . : :
ALLERG LYISSNNLDKIGDNGPFRIYMRGINVDIPRLKMSFNFYVKVDGECVENSVGASIGRDNLI
          40          50          60          70          80          90

```

69 residues in 1 query sequences  
367796 residues in 1603 library sequences  
Scomplib [34t11]  
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012  
Scan time: 0.020 Display time: 0.000

Function used was FASTA [version 3.4t11 Apr 17, 2002]  
FASTA searches a protein or DNA sequence data bank  
version 3.4t11 Apr 17, 2002  
Please cite:  
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 35 aa
>SYHTOH2_I_27
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

```

367796 residues in 1603 sequences  
Expectation\_n fit:  $\rho(\ln(x)) = 4.4141 \pm 0.0033$ ;  $\mu = 1.1420 \pm 0.172$   
mean\_var=31.1013  $\pm$  7.648, 0's: 21 Z-trim: 22 B-trim: 0 in 0/40  
Lambda= 0.2300

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2  
join: 36, opt: 24, gap-pen: -12/-2, width: 16  
Scan time: 0.010

The best scores are:

					opt	bits	E(1603)
ALLERGEN_2012_2	gi 21783 emb CAA30570.1	unnamed	( 356)	57	24	0.92	
ALLERGEN_2012_2	gi 75317968 sp O22116 O22116_WHEA	( 373)	55	24	1.5		
ALLERGEN_2012_2	gi 21314465 gb AAM46958.1 AF51085	( 538)	53	23	3.5		
ALLERGEN_2012_2	gi 112754 sp P04403.2 2SS1_BEREX	( 146)	47	21	3.7		
ALLERGEN_2012_2	gi 56788031 gb AAW29810.1 seed s	( 507)	52	23	4.2		
ALLERGEN_2012_2	gi 62550933 emb CAI79052.1 putat	( 326)	50	22	4.2		
ALLERGEN_2012_2	gi 149208403 gb ABR21772.1 congl	( 455)	51	22	4.7		
ALLERGEN_2012_2	gi 335331566 gb AEH31546.1 low m	( 369)	50	22	4.8		
ALLERGEN_2012_2	gi 169950562 gb ACB05815.1 congl	( 611)	51	22	6.3		

```

>>ALLERGEN_2012_2|gi|21783|emb|CAA30570.1| unnamed prote (356 aa)
  initn: 44 initl: 44 opt: 57 Z-score: 103.7 bits: 24.4 E(): 0.92
Smith-Waterman score: 57; 30.769% identity (30.769% ungapped) in 26 aa overlap (4-
29:187-212)

```

```

          10          20          30
SYHTOH          MSTHPSVYKYLAPPSLLREQNLREIVLEREREQ
          : : : : . . . : : . . . : :
ALLERG QQPPISQQQPPFLQQQRPPFSRQQQIPVIHPSVLQQNLNPKCVFLQQQCIPVAMQRCCLAR
          160          170          180          190          200          210

```

ALLERG SQMLQQSICHVMQQQCCQQLRQIPEQSRHESIRAIYISIILQQQQQQQQQQQQQGGQSII  
220 230 240 250 260 270

SYHTOH MSTHPSVYKYLAPPSLLREQNLREIVLEREREQ  
::: . : : .. :. ....  
ALLERG QQPPISQQQQPPFSQQQQPPFSQQQQIPVIHPSVLQQLNPCMVFLQQQCIPVAMQRCLAR  
170 180 190 200 210 220

ALLERG SQMLQQSICHVMQRQCCQQLRQIPEQSRHESIRAIITYSIILQQQQQQQQQQQQQQGQSII

230 240 250 260 270 280

```

                                10      20      30
SYHTOH                          MSTHPSVYKYLAPPSLLREQNLREIVLEREREQVA
                                .....  ::::  ..
ALLERG SFIDNLP EEVVANSYGLPREQARQLKNNNPFKFFVPPS---EQSLRAVA
                        500      510      520      530

```

SYHTOH MSTHPSVYKYLAPPSLLREQNLR EIVLEREREQVA  
:  
ALLERG MAKISVAAAALLVL MALGHATAFRATVT TTVVEENQEECREQMQRQQMLSHCRM YMRQQ

SYHTOH MSTHPSVYKYLAPPSLLREQNLRIVLEREREQVA  
:  
ALLERG DTETARRLQSENDHRHSIRVEGRQLQVIPRWSREEQEREERKERERERESSESERRQSR

250 260 270 280 290 300

```
>>ALLERGEN_2012_2|gi|62550933|emb|CAI79052.1| putative L (326 aa)
  initn: 43 initl: 43 opt: 50 Z-score: 91.8 bits: 22.0 E(): 4.2
```

Smith-Waterman score: 50; 25.000% identity (25.000% ungapped) in 24 aa overlap (4-27:145-168)

```

                                10      20      30
SYHTOH                        MSTHPSVYKYLAPPSLLREQNLREIVLEREREQ
                                : : : : : : : : : :
ALLERG QQQLPPFSQQLPFFSQQQQPVLLQQQIPFVHPSILQQLNPKVFLQQQCSPVAMPQSLAR
                                120      130      140      150      160      170
```

SYHTOH VA

```

                                180      190      200      210      220      230
ALLERG SQMLQQSSCHVMQQQCCQQLPQIPQQSRYEAIRAIVYSIILQEQQQVQGSIQTQQQQPQQ
```

>>ALLERGEN\_2012\_2|gi|149208403|gb|ABR21772.1| conglutin (455 aa)  
initn: 47 initl: 47 opt: 51 Z-score: 91.0 bits: 22.4 E(): 4.7  
Smith-Waterman score: 51; 50.000% identity (50.000% ungapped) in 20 aa overlap  
(14-33:123-142)

```

                                10      20      30
SYHTOH                        MSTHPSVYKYLAPPSLLREQNLREIVLEREREQVA
                                : : : : : : : : : :
ALLERG RREHHREREQEQQPRPQRRQEEEEEEEWQPRRQRPQSRREEREEREQEQQSSSGSQRGS
                                100      110      120      130      140      150
```

```

                                160      170      180      190      200      210
ALLERG GDERRQHRRRVHREEREQEQDSRSDSRQRNPNYHFSNRFQTYRNRNGQIRVLERFNQ
```

>>ALLERGEN\_2012\_2|gi|335331566|gb|AEH31546.1| low molecu (369 aa)  
initn: 43 initl: 43 opt: 50 Z-score: 90.8 bits: 22.0 E(): 4.8  
Smith-Waterman score: 50; 25.000% identity (25.000% ungapped) in 24 aa overlap (4-27:202-225)

```

                                10      20      30
SYHTOH                        MSTHPSVYKYLAPPSLLREQNLREIVLEREREQ
                                : : : : : : : : : :
ALLERG QQQQPSLPQQPPFSQQQQQLVLPQQQIPFVHPSILQQLNPKVFLQQQCSPVAMPQSLAR
                                180      190      200      210      220      230
```

SYHTOH VA

```

                                240      250      260      270      280      290
ALLERG SQMLQQSSCHVMQQQCCQQLPQIPQQSRYEAIRAIIYSIILQEQQQVQGSIQTPQQQPQQ
```

>>ALLERGEN\_2012\_2|gi|169950562|gb|ACB05815.1| conglutin (611 aa)  
initn: 47 initl: 47 opt: 51 Z-score: 88.6 bits: 22.4 E(): 6.3  
Smith-Waterman score: 51; 50.000% identity (50.000% ungapped) in 20 aa overlap  
(14-33:123-142)

```

                                10      20      30
SYHTOH                        MSTHPSVYKYLAPPSLLREQNLREIVLEREREQVA
                                : : : : : : : : : :
ALLERG RREHHREREQEQQPRPQRRQEEEEEEEWQPRRQRPQSRREEREEREQEQQSSSGSQRGG
                                100      110      120      130      140      150
```

```

                                160      170      180      190      200      210
ALLERG GDERRQHRRRVHREEREQEQDSRSDSRQRNPNYHFSNRFQTYRNRNGQIRVLERFNQ
```

35 residues in 1 query sequences



Smith-Waterman score: 73; 33.333% identity (34.146% ungapped) in 42 aa overlap  
(92-132:16-57)

```

              70          80          90          100          110          120
SYHT0H AEVEGVVAGIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKS
              : . . . : : : . . . : . . : : . .
ALLERG          MSDFNAIAQQFVEFYKTFDGNRAGLGALYKEHSMLETFEAQGTQG
              10          20          30          40
```

```

              130          140          150          160          170          180
SYHT0H VVAVI-GLPNDPSVRLHEALGYTARGLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRPV
              . . . . : : :
ALLERG SAAIVEKLQNLFPQEIQHRTDTVDAQPSADDGILVLVTGALLLGESKPMSTQAFQLKN
              50          60          70          80          90          100
```

>>ALLERGEN\_2012\_2|gi|89892723|gb|ABD79095.1| Zea m 1 all (252 aa)  
initn: 71 init1: 71 opt: 71 Z-score: 96.8 bits: 25.0 E(): 2.2  
Smith-Waterman score: 71; 30.189% identity (30.189% ungapped) in 53 aa overlap  
(100-152:81-133)

```

              70          80          90          100          110          120
SYHT0H GIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKS VVAVIGLP
              : : : . . : : : : : :
ALLERG NGGGCGYKDVNKPFFNSMGACGNIPFKDGLGCGSCFEIKCDKPVVECSGKPVVVHITDMN
              60          70          80          90          100          110
```

```

              130          140          150          160          170          180
SYHT0H NDPSVRLHEALGYTARGLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRPVTQI
              . : . : : : : : :
ALLERG YEPAAAYHFDLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKYDSKVTFHLEKGC GPN
              120          130          140          150          160          170
```

>>ALLERGEN\_2012\_2|gi|89892721|gb|ABD79094.1| Zea m 1 all (263 aa)  
initn: 71 init1: 71 opt: 71 Z-score: 96.6 bits: 25.0 E(): 2.3  
Smith-Waterman score: 71; 30.189% identity (30.189% ungapped) in 53 aa overlap  
(100-152:92-144)

```

              70          80          90          100          110          120
SYHT0H GIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKS VVAVIGLP
              : : : . . : : : : : :
ALLERG NGGGCGYKDVNKPFFNSMGACGNIPFKDGLGCGSCFEIKCDKPVVECSGKPVVVHITDMN
              70          80          90          100          110          120
```

```

              130          140          150          160          170          180
SYHT0H NDPSVRLHEALGYTARGLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRPVTQI
              . : . : : : : : :
ALLERG YEPAAAYHFDLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKFGSKVFSFHLEKGC GPN
              130          140          150          160          170          180
```

>>ALLERGEN\_2012\_2|gi|115502167|sp|Q1ZYQ8.2|EXB10\_MAIZE R (270 aa)  
initn: 71 init1: 71 opt: 71 Z-score: 96.4 bits: 25.0 E(): 2.3  
Smith-Waterman score: 71; 30.189% identity (30.189% ungapped) in 53 aa overlap  
(100-152:99-151)

```

              70          80          90          100          110          120
SYHT0H GIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKS VVAVIGLP
              : : : . . : : : : : :
ALLERG NGGGCGYKDVNKPFFNSMGACGNIPFKDGLGCGSCFEIKCDKPVVECSGKPVVVHITDMN
              70          80          90          100          110          120
```

```

      130      140      150      160      170      180
SYHT0H NDPSVRLHEALGYTARGLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRPVTQI
      .: . : :. :: :. :
ALLERG YEPIAAYHFDLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKYDSKVTFFHLEKGCNPN
      130      140      150      160      170      180

>>ALLERGEN_2012_2|gi|105969545|gb|ABF81662.1| EXPB10 [Ze (269 aa)
  initn: 94 initl: 69 opt: 69 Z-score: 93.9 bits: 24.5 E(): 3.2
Smith-Waterman score: 69; 30.189% identity (30.189% ungapped) in 53 aa overlap
(100-152:98-150)

      70      80      90      100      110      120
SYHT0H GIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKSVAVIGLP
      :: :: . . : : .: :: . .
ALLERG NGGGCGYKDVNKAPFNSMGACGNVPIFKDGLGCGSCFEIKCDKPAECSGKPVVYITDMN
      70      80      90      100      110      120

      130      140      150      160      170      180
SYHT0H NDPSVRLHEALGYTARGLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRPVTQI
      .: . : :. :: :. :
ALLERG YEPIAAYHFDLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKYGSKVTFFHLEKGCNPN
      130      140      150      160      170      180

>>ALLERGEN_2012_2|gi|115502168|sp|P0C1Y5.1|EXB11_MAIZE R (269 aa)
  initn: 94 initl: 69 opt: 69 Z-score: 93.9 bits: 24.5 E(): 3.2
Smith-Waterman score: 69; 30.189% identity (30.189% ungapped) in 53 aa overlap
(100-152:98-150)

      70      80      90      100      110      120
SYHT0H GIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKSVAVIGLP
      :: :: . . : : .: :: . .
ALLERG NGGGCGYKDVNKAPFNSMGACGNVPIFKDGLGCGSCFEIKCDKPAECSGKPVVYITDMN
      70      80      90      100      110      120

      130      140      150      160      170      180
SYHT0H NDPSVRLHEALGYTARGLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRPVTQI
      .: . : :. :: :. :
ALLERG YEPIAAYHFDLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKYGSKVTFFHLEKGCNPN
      130      140      150      160      170      180

>>ALLERGEN_2012_2|gi|18615|emb|CAA26723.1| unnamed prote (495 aa)
  initn: 41 initl: 41 opt: 71 Z-score: 93.2 bits: 25.3 E(): 3.5
Smith-Waterman score: 71; 25.843% identity (29.114% ungapped) in 89 aa overlap
(39-118:113-200)

      10      20      30      40      50      60
SYHT0H EIRPATAADMAAVCDIVNHYIETSTVNFRTEPQTPQEWIDDLERLQDRYPWLVAEEGVV
      ::: ::. . . : ::. . : : .
ALLERG SYTNGPQEIYIQQKGIFGMIYPGCSSTFEEPQQPQQRGQS-SRPQDRHQKIYNSREGDL
      90      100      110      120      130      140

      70      80      90      100      110
SYHT0H AGIAYAGPWKARNAYDWTVESTVYV---SHRHQ-----RLGLGSTLYTHLLKSMEAQGF
      .. . : : : : .. . : :: : . :... .::: .. ::
ALLERG IAVPTGVAWWMYNNEDTPVVAVSIIDTNSLENQLDQMPRRFYLAGNQEQEFLKYQQEQGG
      150      160      170      180      190      200

      120      130      140      150      160      170
SYHT0H KSVVAVIGLPNDPSVRLHEALGYTARGLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRP
ALLERG HQSQKGKHQQEEENEGGSILSGFTLEFLEHAFSVDKQIAKNLQGENEGEDKGAIVTVKGG
      210      220      230      240      250      260

```



>>ALLERGEN\_2012\_2|gi|18635|emb|CAA33215.1| glycinin subu (495 aa)  
 initn: 41 initl: 41 opt: 69 Z-score: 90.7 bits: 24.8 E(): 4.9  
 Smith-Waterman score: 69; 25.843% identity (29.114% ungapped) in 89 aa overlap  
 (39-118:113-200)

```

      10      20      30      40      50      60
SYHT0H EIRPATAADMAAVCDIVNHYIETSTVNFRTEPQTPQEWIDDLERLQDRYPWLVAEEGVV
      :: :: .. : :: . . :: .
ALLERG SYTNGPQEIIYIQQKGIFGMIYPGCPSTFEEPQQPQQRGQS-SRPQDRHQKIYNFREGDL
      90      100      110      120      130      140

      70      80      90      100      110
SYHT0H AGIAYAGPWKARNAYDWTVESTVYV---SHRHQ-----RLGLGSTLYTHLLKSMEAQGF
      .. . : : : .. . : :: :. :... .::: .. ::
ALLERG IAVPTGVAVWWMYNNEDTPVVAVSIIDTNSLENQLDQMPRRFYLAGNQEQEFLKYQQEQGG
      150      160      170      180      190      200

      120      130      140      150      160      170
SYHT0H KSVVAVIGLPNDPSVRLHEALGYTARGLRAAGYKHGGWHDVGFWRDFELPAPPRPVRP
ALLERG HQSQKGKHQQEEENEGGSILSGFTLEFLEHAFSVDKQIAKNLQGENEGEDKGAIVTVKGG
      210      220      230      240      250      260

```

>>ALLERGEN\_2012\_2|gi|18652047|gb|AAL76932.1|AF456481\_1 m (154 aa)  
 initn: 33 initl: 33 opt: 64 Z-score: 90.3 bits: 23.1 E(): 5.1  
 Smith-Waterman score: 64; 22.727% identity (25.773% ungapped) in 110 aa overlap  
 (44-146:21-124)

```

      20      30      40      50      60      70
SYHT0H TAADMAAVCDIVNHYIETSTVNFRTEPQTPQEWIDDLERLQDRY-PWLVAEEGVVAGIA
      : .. :... . . : :. :: .. ::
ALLERG      MGVMQKTEVEAPSTVSAEKMYQGFLDMDTVFPKVLPLIKSVE-ILEGDG
      10      20      30      40

      80      90      100      110      120
SYHT0H YAGPWKARNAYDWTVESTVYVSHRHQ-----RLGLGSTLYTHLLKSMEAQGFKSVVA-VI
      .: . . . :.: :.. ... . :.: : :. . . . :.: :
ALLERG GVGTVRLVHLG---EATEYTTMKQKVDVIDKAGLGYT-YTTIGGDILVEGLESVVNQFV
      50      60      70      80      90      100

      130      140      150      160      170      180
SYHT0H GLPNDPSVRLHEALGYTARGLRAAGYKHGGWHDVGFWRDFELPAPPRPVRPVTQI
      .:.: . .... :.:.:
ALLERG VVPTDGGCIVKNTTIYNTKGDVLPEDKVKEATEKSALAFKAVEAYLLAN
      110      120      130      140      150

```

>>ALLERGEN\_2012\_2|gi|2266625|emb|CAB10765.1| group V all (264 aa)  
 initn: 34 initl: 34 opt: 63 Z-score: 86.2 bits: 23.1 E(): 8.6  
 Smith-Waterman score: 63; 28.814% identity (30.357% ungapped) in 59 aa overlap  
 (101-157:17-74)

```

      80      90      100      110      120
SYHT0H IAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKSVVAVIGL--
      : .. : .:.....: :.:.....:
ALLERG      ADAGYTPAAPAAAGAGGKATTDEQKLLEDVNA-GFKTAVAAAAANVP
      10      20      30      40

      130      140      150      160      170      180
SYHT0H PNDPSVRLHEALGYTARGLRAAGYKHGGWHDVGFWRDFELPAPPRPVRPVTQI
      : : .. :. .... :.: : :
ALLERG PADKYKTFAAFTASSKASIAAAATKAPGLIPQLNAATNTAYAAAQGATPEAKYDAFVTT
      50      60      70      80      90      100

```

183 residues in 1 query sequences  
 367796 residues in 1603 library sequences  
 Scomplib [34t11]  
 start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012  
 Scan time: 0.050 Display time: 0.010

Function used was FASTA [version 3.4t11 Apr 17, 2002]  
 FASTA searches a protein or DNA sequence data bank  
 version 3.4t11 Apr 17, 2002  
 Please cite:  
 W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

/grid/.qsub\_tmp/1330373600kde\_agent/grid\_NbCuMBoIO: 32 aa  
 >SYHT0H2\_I\_29  
 vs /data/fasta/ALLERGEN\_2012\_2 library  
 searching /data/fasta/ALLERGEN\_2012\_2 library

367796 residues in 1603 sequences  
 Expectation\_n fit:  $\rho(\ln(x)) = 3.4303 \pm 0.00255$ ;  $\mu = 4.7585 \pm 0.134$   
 mean\_var=23.4858  $\pm$  6.194, 0's: 16 Z-trim: 16 B-trim: 0 in 0/42  
 Lambda= 0.2647

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2  
 join: 36, opt: 24, gap-pen: -12/-2, width: 16  
 Scan time: 0.010

The best scores are:

					opt	bits	E(1603)
ALLERGEN_2012_2	gi 25361513 gb AAN73248.1	helix-	( 450)	48	23	2.5	
ALLERGEN_2012_2	gi 322812205 pdb 2X45 A Chain A,	( 144)	40	20	7.3		
ALLERGEN_2012_2	gi 462026 sp P34071.1 ENTC2_STAAU	( 266)	42	21	7.5		
ALLERGEN_2012_2	gi 58371884 emb CAG26895.1  Arg r	( 159)	40	20	8		
ALLERGEN_2012_2	gi 757851 emb CAA26040.1  ovotran	( 705)	45	22	8.2		
ALLERGEN_2012_2	gi 1351295 sp P02789.2 TRFE_CHICK	( 705)	45	22	8.2		

>>ALLERGEN\_2012\_2|gi|25361513|gb|AAN73248.1| helix-loop- (450 aa)  
 initn: 48 initl: 48 opt: 48 Z-score: 96.0 bits: 23.2 E(): 2.5  
 Smith-Waterman score: 48; 63.636% identity (63.636% ungapped) in 11 aa overlap  
 (11-21:201-211)

	10	20	30	
SYHT0H	MMAFVGATFLFHYLHNKVTDSWAMESEEVSGY			
	::: :: :. :			
ALLERG	ANQFNFEVPGHMNVSGHP	THLEHSSTNPNSFHYEHNIVSPSSIH	PSTA	FDGEVPSQWDD
	180	190	200	210 220 230
ALLERG	SLGHGASTPKVRTPSHHVSSNPWAEINEPTGGDNDNLAPVTRPRKPARARRQKKEPRKLS			
	240	250	260	270 280 290

>>ALLERGEN\_2012\_2|gi|322812205|pdb|2X45|A Chain A, Cryst (144 aa)  
 initn: 36 initl: 36 opt: 40 Z-score: 87.5 bits: 19.9 E(): 7.3  
 Smith-Waterman score: 40; 33.333% identity (33.333% ungapped) in 18 aa overlap  
 (15-32:99-116)

	10	20	30	
SYHT0H	MMAFVGATFLFHYLHNKVTDSWAMESEEVSGY			
	:. :. :. :. :. :			
ALLERG	TATAKGS	DIVVGS	DTSTVIYTDGKTCDVVKHGHTELVWHSSKTS	GGYNNCCDKKFTETR
	70	80	90	100 110 120
ALLERG	GSTPANEVYKKCPGMP			
	130	140		

>>ALLERGEN\_2012\_2|gi|462026|sp|P34071.1|ENTC2\_STAAU RecN (266 aa)  
 initn: 29 initl: 29 opt: 42 Z-score: 87.3 bits: 20.8 E(): 7.5  
 Smith-Waterman score: 42; 33.333% identity (33.333% ungapped) in 24 aa overlap (4-27:235-258)

```

                                10      20      30
SYHT0H                      MMAFVGATFLFHYLHNKVTDSWAMESEEVSGY
                                :  .  .  .  :  :  :  :  .  :
ALLERG SSPYETGYIKFIENNGNTFWYDMMPAPGDKFDQSKYLMYNDNKTVDKSVKIEVHLTTK
                                210      220      230      240      250      260

```

ALLERG NG

>>ALLERGEN\_2012\_2|gi|58371884|emb|CAG26895.1| Arg r 1 pr (159 aa)  
 initn: 36 initl: 36 opt: 40 Z-score: 86.8 bits: 20.0 E(): 8  
 Smith-Waterman score: 40; 33.333% identity (33.333% ungapped) in 18 aa overlap (15-32:114-131)

```

                                10      20      30
SYHT0H                      MMAFVGATFLFHYLHNKVTDSWAMESEEVSGY
                                :  .  .  .  :  :  :  :  .  :
ALLERG TATAKGSDIVVGSdTSTVIYTDGKTCDVVKHGGHTELWVHSSKTSGGYNNCCDKKFTETR
                                90      100      110      120      130      140

```

ALLERG GSTPANEVYKKCPGMP  
 150

>>ALLERGEN\_2012\_2|gi|757851|emb|CAA26040.1| ovotransferr (705 aa)  
 initn: 45 initl: 45 opt: 45 Z-score: 86.6 bits: 22.1 E(): 8.2  
 Smith-Waterman score: 45; 41.176% identity (41.176% ungapped) in 17 aa overlap (16-32:277-293)

```

                                10      20      30
SYHT0H                      MMAFVGATFLFHYLHNKVTDSWAMESEEVSGY
                                :  :  :  :  .  .  :  .
ALLERG CLDGSRQPDNYKTCNWARVAAHAVVARDNKNVEDIWSFLSKAQSDFGVDTKSDFHFLFGP
                                250      260      270      280      290      300

```

ALLERG PGKKDPVLKDFLFKDSAIMLKRVP SLMDSQLYLGFEYSAIQSMRKDQLTPSPRENRIQW  
 310 320 330 340 350 360

>>ALLERGEN\_2012\_2|gi|1351295|sp|P02789.2|TRFE\_CHICK RecN (705 aa)  
 initn: 45 initl: 45 opt: 45 Z-score: 86.6 bits: 22.1 E(): 8.2  
 Smith-Waterman score: 45; 41.176% identity (41.176% ungapped) in 17 aa overlap (16-32:277-293)

```

                                10      20      30
SYHT0H                      MMAFVGATFLFHYLHNKVTDSWAMESEEVSGY
                                :  :  :  :  .  .  :  .
ALLERG CLDGSRQPDNYKTCNWARVAAHAVVARDNKNVEDIWSFLSKAQSDFGVDTKSDFHFLFGP
                                250      260      270      280      290      300

```

ALLERG PGKKDPVLKDLLFKDSAIMLKRVP SLMDSQLYLGFEYSAIQSMRKDQLTPSPRENRIQW  
 310 320 330 340 350 360

32 residues in 1 query sequences  
 367796 residues in 1603 library sequences  
 Scomplib [34t11]  
 start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012  
 Scan time: 0.010 Display time: 0.000

Function used was FASTA [version 3.4t11 Apr 17, 2002]

FASTA searches a protein or DNA sequence data bank

version 3.4t11 Apr 17, 2002

Please cite:

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

/grid/.qsub\_tmp/1330373600kde\_agent/grid\_NbCuMBoIO: 44 aa

>SYHT0H2\_I\_30

vs /data/fasta/ALLERGEN\_2012\_2 library

searching /data/fasta/ALLERGEN\_2012\_2 library

367796 residues in 1603 sequences

Expectation\_n fit: rho(ln(x))= 5.8225+/-0.00365; mu= -5.3535+/- 0.189

mean\_var=50.3779+/-14.454, 0's: 12 Z-trim: 12 B-trim: 79 in 1/42

Lambda= 0.1807

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2

join: 36, opt: 24, gap-pen: -12/-2, width: 16

Scan time: 0.010

The best scores are:

						opt	bits	E(1603)
ALLERGEN_2012_2	gi 170722 gb AAA34281.1	pre-alpha	( 262)	67	25		0.66	
ALLERGEN_2012_2	gi 473876 gb AAA17741.1	alpha-gl	( 287)	67	25		0.73	
ALLERGEN_2012_2	gi 170728 gb AAA34284.1	alpha-ty	( 186)	64	24		0.79	
ALLERGEN_2012_2	gi 21755 emb CAA25593.1	unnamed	( 286)	63	24		1.5	
ALLERGEN_2012_2	gi 21761 emb CAA26384.1	unnamed	( 286)	63	24		1.5	
ALLERGEN_2012_2	gi 170720 gb AAA34280.1	alpha/be	( 286)	63	24		1.5	
ALLERGEN_2012_2	gi 1304264 dbj BAA12318.1	alpha-	( 259)	62	23		1.6	
ALLERGEN_2012_2	gi 170712 gb AAA34276.1	pre-alpha	( 291)	61	23		2.2	
ALLERGEN_2012_2	gi 170740 gb AAA34290.1	gliadin	( 296)	61	23		2.2	
ALLERGEN_2012_2	gi 21757 emb CAA26383.1	unnamed	( 296)	61	23		2.2	
ALLERGEN_2012_2	gi 155676690 dbj BAF75708.1	SXP/	( 150)	57	22		2.2	
ALLERGEN_2012_2	gi 21673 emb CAA35238.1	unnamed	( 307)	60	23		2.8	
ALLERGEN_2012_2	gi 155676692 dbj BAF75709.1	SXP/	( 150)	55	22		3.2	
ALLERGEN_2012_2	gi 21926 emb CAA36063.1	unnamed	( 295)	58	22		3.8	
ALLERGEN_2012_2	gi 170718 gb AAA34279.1	alpha/be	( 313)	58	22		4.1	
ALLERGEN_2012_2	gi 155676684 dbj BAF75705.1	SXP/	( 150)	53	21		4.6	
ALLERGEN_2012_2	gi 155676696 dbj BAF75711.1	SXP/	( 150)	53	21		4.6	
ALLERGEN_2012_2	gi 155676688 dbj BAF75707.1	SXP/	( 150)	53	21		4.6	
ALLERGEN_2012_2	gi 155676636 dbj BAF75681.1	SXP/	( 150)	53	21		4.6	
ALLERGEN_2012_2	gi 155676682 dbj BAF75704.1	SXP/	( 150)	53	21		4.6	
ALLERGEN_2012_2	gi 155676698 dbj BAF75712.1	SXP/	( 150)	53	21		4.6	
ALLERGEN_2012_2	gi 155676694 dbj BAF75710.1	SXP/	( 150)	53	21		4.6	
ALLERGEN_2012_2	gi 155676686 dbj BAF75706.1	SXP/	( 150)	53	21		4.6	
ALLERGEN_2012_2	gi 30794292 ref NP_851341.1	lact	( 708)	62	23		4.6	
ALLERGEN_2012_2	gi 55859456 emb CAH92630.1	polle	( 520)	59	23		5.8	
ALLERGEN_2012_2	gi 21765 emb CAA26385.1	unnamed	( 313)	56	22		5.8	
ALLERGEN_2012_2	gi 886967 emb CAA59340.1	low mol	( 276)	55	22		6.1	
ALLERGEN_2012_2	gi 170738 gb AAA34289.1	gamma-gl	( 327)	55	22		7.3	

>>ALLERGEN\_2012\_2|gi|170722|gb|AAA34281.1| pre-alpha-/be (262 aa)

initn: 39 initl: 39 opt: 67 Z-score: 106.3 bits: 24.7 E(): 0.66

Smith-Waterman score: 67; 28.205% identity (28.205% ungapped) in 39 aa overlap (5-43:100-138)

```

                                10      20      30
SYHT0H                        MVGQFRQDSSRISEPHLHTAPLQQVQIRDQDLLE
                                . : . . : . . . : . : . : .
ALLERG PSQQPYLQLQPLQPLQPYSQPQPFRPQQPYQPQPQYSQPQQPISQQQQQQQQQQQQQ
                                70      80      90      100     110     120
```

```

40
SYHT0H QKQVLLLQQH
.: .:.
ALLERG QQQQIIQQILQQQLIPCMDVVLQQHNIVHGKSQVLQQSTYQLLQELCCQHLWQIPEQSQ
130      140      150      160      170      180

>>ALLERGEN_2012_2|gi|473876|gb|AAA17741.1| alpha-gliadin (287 aa)
initn: 39 initl: 39 opt: 67 Z-score: 105.5 bits: 24.7 E(): 0.73
Smith-Waterman score: 67; 28.205% identity (28.205% ungapped) in 39 aa overlap (5-43:100-138)

```

```

10      20      30
SYHT0H      MVGQFRQDSSRISEPHLHTAPLQQVQIRDQDLLE
.: . . :. . . :. . .
ALLERG PSQQPYLQLQPFLQPQLPYSQPQPFRPQQPYQPQPQYSQPQQPISQQQQQQQQQQQQQQ
70      80      90      100      110      120

```

```

40
SYHT0H QKQVLLLQQH
.: .:.
ALLERG QQQQIIQQILQQQLIPCMDVVLQQHNIVHGKSQVLQQSTYQLLQELCCQHLWQIPEQSQ
130      140      150      160      170      180

>>ALLERGEN_2012_2|gi|170728|gb|AAA34284.1| alpha-type gl (186 aa)
initn: 73 initl: 48 opt: 64 Z-score: 104.8 bits: 24.0 E(): 0.79
Smith-Waterman score: 66; 30.612% identity (36.585% ungapped) in 49 aa overlap (4-44:6-54)

```

```

10      20      30      40
SYHT0H      MVGQFRQDSSRISEPHLHTAPLQQVQIRDQDLLEQ---KQVL----LLQQH
.: . . :. . . :. . . :. . . :. . .
ALLERG PQPQPQYSQPQQPISQQQQQQQQQQQQQQEQQILQQILQQQLIPCMDVVLQQHNIAHGR
10      20      30      40      50      60

ALLERG SQVLQQSTYQLLQELCCQHLWQIPEQSQCQAIHNVVHAILHQQQKQQQQPSSQSFQQ
70      80      90      100      110      120

```

```

>>ALLERGEN_2012_2|gi|21755|emb|CAA25593.1| unnamed prote (286 aa)
initn: 79 initl: 53 opt: 63 Z-score: 99.9 bits: 23.7 E(): 1.5
Smith-Waterman score: 64; 30.612% identity (36.585% ungapped) in 49 aa overlap (4-44:106-154)

```

```

10      20      30
SYHT0H      MVGQFRQDSSRISEPHLHTAPLQQVQIRDQDL
.: . . :. . . :. . .
ALLERG LQLQFPQPQLPYSQPQPFRPQQPYQPQPQYSQPQQPISQQQQQQQQQQQQQQQQQIL
80      90      100      110      120      130

```

```

40
SYHT0H EQ---KQVL----LLQQH
.: .:.
ALLERG QQILQQQLIPCMDVVLQQHNIAHGRSQVLQQSTYQLLQELCCQHLWQIPEQSQCQAILKV
140      150      160      170      180      190

```

```

>>ALLERGEN_2012_2|gi|21761|emb|CAA26384.1| unnamed prote (286 aa)
initn: 82 initl: 56 opt: 63 Z-score: 99.9 bits: 23.7 E(): 1.5
Smith-Waterman score: 64; 30.612% identity (36.585% ungapped) in 49 aa overlap (4-44:106-154)

```

```

10      20      30
SYHT0H      MVGQFRQDSSRISEPHLHTAPLQQVQIRDQDL
.: . . :. . . :. . .
ALLERG LQLQFPQPQLPYSQPQQFRPQQPYQPQPQYSQPQQPISQQQQQQQQQQQQQQQQQIL

```

```

      80          90          100          110          120          130
      40
SYHT0H EQ---KQVL-----LLQQH
      .:  .:. .: .:
ALLERG QQILQQQLIPCMDVVLQQHNIAGRSQVLQQSTYQLLQELCCQHLWQIPEQSQCQAIHNV
      140          150          160          170          180          190

>>ALLERGEN_2012_2|gi|170720|gb|AAA34280.1| alpha/beta-gl (286 aa)
      initn: 79 initl: 53 opt: 63 Z-score: 99.9 bits: 23.7 E(): 1.5
Smith-Waterman score: 64; 30.612% identity (36.585% ungapped) in 49 aa overlap (4-
44:106-154)

```

```

      10          20          30
SYHT0H MVGQFRQDSSRISEPHLHTAPLQQVQIRDQDLL
      .: .: .: .: .: .: .: .:
ALLERG LQLQFPFPQPLPYSQPQFPRPQQPYQPQPQYSQPQQPISQQQQQQQQQQQQQQQQQQIL
      80          90          100          110          120          130

```

```

      40
SYHT0H EQ---KQVL-----LLQQH
      .:  .:. .: .:
ALLERG QQILQQQLIPCMDVVLQQHNIAGRSQVLQQSTYQLLQELCCQHLWQIPEQSQCQAIHNV
      140          150          160          170          180          190

```

```

>>ALLERGEN_2012_2|gi|1304264|dbj|BAA12318.1| alpha-gliad (259 aa)
      initn: 35 initl: 35 opt: 62 Z-score: 99.3 bits: 23.4 E(): 1.6
Smith-Waterman score: 62; 34.884% identity (37.500% ungapped) in 43 aa overlap (4-
44:86-127)

```

```

      10          20          30
SYHT0H MVGQFRQDSSRISEPHLHTAPLQQV--QIRDQD
      .: .: .: .: .: .: .: .:
ALLERG LQLQFPFPQPLPYSQPQFPRPQQPYQPQPQYSQPQEPISQQQQQQQQQQQQILQQILQQQ
      60          70          80          90          100          110

```

```

      40
SYHT0H LLEQKQVLLQLQH
      .:  .: .: .:
ALLERG LIPCMDVVL--QQHNIAGRSQVLQQSTYQLLQELCCQHLWQIPEQSQCQAIQNVVHAIL
      120          130          140          150          160          170

```

```

>>ALLERGEN_2012_2|gi|170712|gb|AAA34276.1| pre-alpha-/be (291 aa)
      initn: 40 initl: 40 opt: 61 Z-score: 96.9 bits: 23.2 E(): 2.2
Smith-Waterman score: 61; 30.952% identity (34.211% ungapped) in 42 aa overlap (5-
44:100-139)

```

```

      10          20          30
SYHT0H MVGQFRQDSSRISEPHLHTAPLQQVQIRDQDLLE
      .: .: .: .: .: .: .: .:
ALLERG PSQQPYLQLQFPFPQPLPYPQPQFPRPQQPYQPQPQYSQPQQPISQQQQQQQQQQQQ--Q
      70          80          90          100          110          120

```

```

      40
SYHT0H QKQVL--LLQQH
      .: .: .: .:
ALLERG QQQILQQILQQQLIPCRDVLQQHNIAGSSQVLQESTYQLVQQLCCQQLWQIPEQSRCQ
      130          140          150          160          170          180

```

```

>>ALLERGEN_2012_2|gi|170740|gb|AAA34290.1| gliadin [Trit (296 aa)
      initn: 46 initl: 46 opt: 61 Z-score: 96.8 bits: 23.2 E(): 2.2
Smith-Waterman score: 61; 32.558% identity (35.897% ungapped) in 43 aa overlap (4-
44:111-151)

```



>>ALLERGEN\_2012\_2|gi|155676692|dbj|BAF75709.1| SXP/RAL-2 (150 aa)  
 initn: 36 initl: 36 opt: 55 Z-score: 93.9 bits: 21.6 E(): 3.2  
 Smith-Waterman score: 55; 28.205% identity (28.205% ungapped) in 39 aa overlap (5-43:46-84)

```

                                10      20      30
SYHT0H                        MVGQFRQDSSRISEPHLHTAPLQQVQIRDQDLLE
                                :...: : .. :. :.: :
ALLERG AQDQGPPPLPKFLDGADQATKDAFAALAQTFKDDTDKQVEDAVQQFANDHPAIKDAYEAE
                                20      30      40      50      60      70

```

```

                                40
SYHT0H QKQVLLQQH
                                :.: :.:
ALLERG KKEVLAAQQAAEEHKKLVAAALPPDAQKADAELTATADDASLTAAKHDKIVQTFESLPP
                                80      90     100     110     120     130

```

>>ALLERGEN\_2012\_2|gi|21926|emb|CAA36063.1| unnamed prote (295 aa)  
 initn: 52 initl: 52 opt: 58 Z-score: 92.6 bits: 22.4 E(): 3.8  
 Smith-Waterman score: 58; 32.609% identity (34.884% ungapped) in 46 aa overlap (2-44:205-250)

```

                                10      20      30
SYHT0H                        MVGQFRQDSSRISEPHLHTAPLQQVQIRDQD
                                .: :. :. :. :. :. :. :
ALLERG YDVIRAITYSIILQEQQQGFVQAQQQQPQQLGQGVSSQQQSQQQLGQCSFQQPQQQLGQ
                                180     190     200     210     220     230

```

```

                                40
SYHT0H LLEQKQVL---LLQQH
                                :.: :.: :
ALLERG QPQQQQVLQGTFLQPHQIAHLEVMTSIALRTLPTMCSVNVPLYSSTTSVPFVSFGTGVGAY
                                240     250     260     270     280     290

```

>>ALLERGEN\_2012\_2|gi|170718|gb|AAA34279.1| alpha/beta-gl (313 aa)  
 initn: 79 initl: 53 opt: 58 Z-score: 92.1 bits: 22.4 E(): 4.1  
 Smith-Waterman score: 58; 32.558% identity (36.842% ungapped) in 43 aa overlap (4-44:108-147)

```

                                10      20      30
SYHT0H                        MVGQFRQDSSRISEPHLHTAPLQQVQIRDQDL
                                .. :. :. :. :. :. :. :
ALLERG FPQPQPFLPQLPYQPQSFPQPYPQRPKYLPQPQPISQQQ--AQQQQQQQQQQQQQ
                                80      90      100     110     120     130

```

```

                                40
SYHT0H EQKQVL--LLQQH
                                :.: :.: :
ALLERG QQQQILQQILQQQLIPCRDVVLQQHNIAHASSQVLQQSTYQLLQQLCCQQLLQIPEQSRC
                                140     150     160     170     180     190

```

>>ALLERGEN\_2012\_2|gi|155676684|dbj|BAF75705.1| SXP/RAL-2 (150 aa)  
 initn: 36 initl: 36 opt: 53 Z-score: 91.1 bits: 21.1 E(): 4.6  
 Smith-Waterman score: 53; 28.205% identity (28.205% ungapped) in 39 aa overlap (5-43:46-84)

```

                                10      20      30
SYHT0H                        MVGQFRQDSSRISEPHLHTAPLQQVQIRDQDLLE
                                :...: : .. :. :.: :
ALLERG AQDQGPPPLPKFLDGADQATKDAFAALAQTFKDDTDKQVEDAVQQFVNDHPAIKDAYEAE
                                20      30      40      50      60      70

```



```

      40
SYHT0H QKQVLLLQQH
      .:.:.:  ::
ALLERG KKEVLAAQQATEEEHKKLVAAALPPDAQKADAELTAIADDASLTLAAKHKIVQTFESLPP
      80      90      100      110      120      130

>>ALLERGEN_2012_2|gi|155676696|dbj|BAF75711.1| SXP/RAL-2 (150 aa)
  initn: 36 initl: 36 opt: 53 Z-score: 91.1 bits: 21.1 E(): 4.6
Smith-Waterman score: 53; 28.205% identity (28.205% ungapped) in 39 aa overlap (5-
43:46-84)

      10      20      30
SYHT0H          MVGQFRQDSSRISEPHLHTAPLQQVQIRDQDLLE
      .:.:.:  :  ..  ..  :.:  :
ALLERG AQDQGPPPLPKFLDGADQATKDFAAALAQTFKDDTDKQVEDAVQQFVNDHHPAIKDAYEAE
      20      30      40      50      60      70

      40
SYHT0H QKQVLLLQQH
      .:.:.:  ::
ALLERG KKEVLAAQQAAEEEHKKLVAAALPPDAQKADAESTAIADDASLTLAAKHKIVQTFESLPP
      80      90      100      110      120      130

>>ALLERGEN_2012_2|gi|155676688|dbj|BAF75707.1| SXP/RAL-2 (150 aa)
  initn: 36 initl: 36 opt: 53 Z-score: 91.1 bits: 21.1 E(): 4.6
Smith-Waterman score: 53; 28.205% identity (28.205% ungapped) in 39 aa overlap (5-
43:46-84)

      10      20      30
SYHT0H          MVGQFRQDSSRISEPHLHTAPLQQVQIRDQDLLE
      .:.:.:  :  ..  ..  :.:  :
ALLERG AQDQGPPPLPKFLDGADQATKDAYAALAQTFKDDTDKQVEDAVQQFVNDHHPAIKDAYEAE
      20      30      40      50      60      70

      40
SYHT0H QKQVLLLQQH
      .:.:.:  ::
ALLERG KKEVLAAQQAAEEEHKKLVAAALPPDAQKADAELTAIADDASLTLAAKHKIVQTFESLPP
      80      90      100      110      120      130

>>ALLERGEN_2012_2|gi|155676636|dbj|BAF75681.1| SXP/RAL-2 (150 aa)
  initn: 36 initl: 36 opt: 53 Z-score: 91.1 bits: 21.1 E(): 4.6
Smith-Waterman score: 53; 28.205% identity (28.205% ungapped) in 39 aa overlap (5-
43:46-84)

      10      20      30
SYHT0H          MVGQFRQDSSRISEPHLHTAPLQQVQIRDQDLLE
      .:.:.:  :  ..  ..  :.:  :
ALLERG AQDQGPPPLPKFLDGADQATKDFAAALAQTFKDDTDKQVEDAVQQFVNDHHPAIKDAYEAE
      20      30      40      50      60      70

      40
SYHT0H QKQVLLLQQH
      .:.:.:  ::
ALLERG KKEVLAAQQAAEEEHKKLVAAALPPDAQKADAELTAIADDASLTLAAKHKIVQTFESLPP
      80      90      100      110      120      130

>>ALLERGEN_2012_2|gi|155676682|dbj|BAF75704.1| SXP/RAL-2 (150 aa)
  initn: 36 initl: 36 opt: 53 Z-score: 91.1 bits: 21.1 E(): 4.6
Smith-Waterman score: 53; 28.205% identity (28.205% ungapped) in 39 aa overlap (5-
43:46-84)

```

```

              10      20      30
SYHT0H      MVGQFRQDSSRISEPHLHTAPLQQVQIRDQDLLE
              :...: : .. :.: :
ALLERG AQDQGPPPLPKFLDGADQATKDAFAALAQTFKDDTDKQVEDAVQQFVNDHHPAIKDAYEAE
              20      30      40      50      60      70

              40
SYHT0H QKQVLLLQQH
              :.: :.
ALLERG KKEVLAAQQAEEEEHKKLVAALPPDAQKADAELTAIADDASLTAAKHDKIVQTFESLPP
              80      90      100      110      120      130

```

>>ALLERGEN\_2012\_2|gi|155676698|dbj|BAF75712.1| SXP/RAL-2 (150 aa)  
 initn: 36 initl: 36 opt: 53 Z-score: 91.1 bits: 21.1 E(): 4.6  
 Smith-Waterman score: 53; 28.205% identity (28.205% ungapped) in 39 aa overlap (5-43:46-84)

```

              10      20      30
SYHT0H      MVGQFRQDSSRISEPHLHTAPLQQVQIRDQDLLE
              :...: : .. :.: :
ALLERG AQDQGPPPLPKFLDGADQATKDAFAALAQTFKDDTDKQVEDAVQQFVNDHHPAIKDAYEAE
              20      30      40      50      60      70

              40
SYHT0H QKQVLLLQQH
              :.: :.
ALLERG KKEVLAAQQAEEEEHKKLVAALPPDAQKADAELTAIADDASLTAAKHDKIVQTFESLPP
              80      90      100      110      120      130

```

>>ALLERGEN\_2012\_2|gi|155676694|dbj|BAF75710.1| SXP/RAL-2 (150 aa)  
 initn: 36 initl: 36 opt: 53 Z-score: 91.1 bits: 21.1 E(): 4.6  
 Smith-Waterman score: 53; 28.205% identity (28.205% ungapped) in 39 aa overlap (5-43:46-84)

```

              10      20      30
SYHT0H      MVGQFRQDSSRISEPHLHTAPLQQVQIRDQDLLE
              :...: : .. :.: :
ALLERG AQDQGPPPLPKFLDGADQATKDAFAALAQTFKDDTDKQVEDAVQQFVNNHPAIKDAYEAE
              20      30      40      50      60      70

              40
SYHT0H QKQVLLLQQH
              :.: :.
ALLERG KKEVLAAQQAEEEEHKKLVAALPPDAQKADAELTAIADDASLTAAKHDKIVQTFESLPP
              80      90      100      110      120      130

```

>>ALLERGEN\_2012\_2|gi|155676686|dbj|BAF75706.1| SXP/RAL-2 (150 aa)  
 initn: 36 initl: 36 opt: 53 Z-score: 91.1 bits: 21.1 E(): 4.6  
 Smith-Waterman score: 53; 28.205% identity (28.205% ungapped) in 39 aa overlap (5-43:46-84)

```

              10      20      30
SYHT0H      MVGQFRQDSSRISEPHLHTAPLQQVQIRDQDLLE
              :...: : .. :.: :
ALLERG AQDQGPPPLPKFLDGADQATKDAFAALAQTFKDDTDKQVEDAVQQFVNDHHPAIKDAYEAE
              20      30      40      50      60      70

              40
SYHT0H QKQVLLLQQH
              :.: :.
ALLERG KKEVLAAQQAEEEEHKKLVAALPPDAQKADAELTAIADDASLTAAKHDKIVQTFESLPP
              80      90      100      110      120      130

```

>>ALLERGEN\_2012\_2|gi|30794292|ref|NP\_851341.1| lactotran (708 aa)  
 initn: 60 initl: 60 opt: 62 Z-score: 91.1 bits: 23.4 E(): 4.6  
 Smith-Waterman score: 62; 36.842% identity (36.842% ungapped) in 38 aa overlap (6-43:597-634)

```

              10      20      30
SYHT0H          MVGQFRQDSSRISEPHLHTAPLQQVQIRDQDLLEQ
              :. . . . :. :. : :. . .
ALLERG VVENTNGESTADWAKNLNREDFRLLCCLDGTRKPVTEAQSCHLAVAPNHAVVSRSDRAAHV
      570      580      590      600      610      620

      40
SYHT0H KQVLLLQQH
      :. :. :. :. :.
ALLERG KQVLLHQQALFGKNGKNCSDKFCLFKSETKNLLFNDNTECLAKLGGRPTYEEYLGTEYVT
      630      640      650      660      670      680

```

>>ALLERGEN\_2012\_2|gi|55859456|emb|CAH92630.1| pollen all (520 aa)  
 initn: 30 initl: 30 opt: 59 Z-score: 89.4 bits: 22.6 E(): 5.8  
 Smith-Waterman score: 59; 36.667% identity (37.931% ungapped) in 30 aa overlap (4-32:436-465)

```

              10      20      30
SYHT0H          MVGQFRQDSSRISEPHLHTAPLQQ-VQIRDQDL
              :. . : :. :. :. :. :. :. :.
ALLERG ATPFPHRKDVLFNQYVNYWFDEAGGAAPLQWSKDMYRFMEPYVSKNPRQAYANYRDIDL
      410      420      430      440      450      460

      40
SYHT0H LEQKQVLLLQQH

ALLERG GRNEVVNDISTYASGKVGKEYFKGNFQRLAITKGKVDPPQDYFRNEQSIPPLL GK
      470      480      490      500      510      520

```

>>ALLERGEN\_2012\_2|gi|21765|emb|CAA26385.1| unnamed prote (313 aa)  
 initn: 79 initl: 53 opt: 56 Z-score: 89.3 bits: 21.9 E(): 5.8  
 Smith-Waterman score: 56; 40.741% identity (44.000% ungapped) in 27 aa overlap (20-44:121-147)

```

              10      20      30      40
SYHT0H          MVGQFRQDSSRISEPHLHTAPLQQVQIRDQDLLEQKQVL--LLQQH
              : :. : :. :. :. :. :. :. :.
ALLERG PQPQSFPQQPYQPQRPMYLPQQPISQQQAQQQQQQQQQQQQQQQQQQILQQILQQQLIP
      100      110      120      130      140      150

ALLERG CRDVVLQQHNIAHASSQVLQQSTYQLLQQLCCQQLLQIPEQSRCQAIHNVVHAIIMHQEQ
      160      170      180      190      200      210

```

>>ALLERGEN\_2012\_2|gi|886967|emb|CAA59340.1| low molecula (276 aa)  
 initn: 51 initl: 51 opt: 55 Z-score: 88.9 bits: 21.6 E(): 6.1  
 Smith-Waterman score: 57; 31.579% identity (32.432% ungapped) in 38 aa overlap (1-38:230-266)

```

              10      20      30
SYHT0H          MVGQFRQDSSRISEPHLHTAPLQQVQIRDQ
              :. :. :. :. :. :. :. :. :.
ALLERG CHVMQQQCCQQLPRIPEQSRDAIRAIISIVLQEQQHGGFNQPPQ-QQQPQQSVQGVSQ
      200      210      220      230      240      250

```

```

              40
SYHT0H DLLEQKQVLLLQQH
      .:.:.
ALLERG PQQQKQLGQCSFQQPQQ
      260      270

>>ALLERGEN_2012_2|gi|170738|gb|AAA34289.1| gamma-gliadin (327 aa)
  initn: 45 initl: 45 opt: 55 Z-score: 87.5 bits: 21.6 E(): 7.3
Smith-Waterman score: 55; 37.500% identity (37.500% ungapped) in 32 aa overlap
(12-43:16-47)

```

```

              10      20      30      40
SYHT0H      MVGQFRQDSSRISEPHLHTAPLQQVQIRDQDLLEQKQVLLLQQH
              . . . . . : : : : . : : : : : : :
ALLERG MKTLLILITILAMAITIGTANIQVDPGQVQWLQQQLVPLQLQQPLSQQPQQTFPQPQQTFP
              10      20      30      40      50      60

ALLERG HQPQQQVPQPQQPQQPFLQPQQPFPQQPQQPFPQTQQPQQPFPQQPQQPFPQTQQPQQPFP
              70      80      90      100      110      120

```

44 residues in 1 query sequences  
 367796 residues in 1603 library sequences  
 Scomplib [34t11]  
 start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012  
 Scan time: 0.010 Display time: 0.010

Function used was FASTA [version 3.4t11 Apr 17, 2002]  
 FASTA searches a protein or DNA sequence data bank  
 version 3.4t11 Apr 17, 2002  
 Please cite:  
 W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 57 aa
>SYHT0H2_I_31
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

```

367796 residues in 1603 sequences  
 Expectation\_n fit:  $\rho(\ln(x)) = 3.9568 \pm 0.00352$ ;  $\mu = 6.1512 \pm 0.183$   
 mean\_var =  $41.6044 \pm 10.677$ , 0's: 13 Z-trim: 13 B-trim: 69 in 1/42  
 Lambda = 0.1988

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2  
 join: 36, opt: 24, gap-pen: -12/-2, width: 16  
 Scan time: 0.020

The best scores are:

					opt	bits	E(1603)
ALLERGEN_2012_2	gi 417370 sp P24296.2 NLTP1_WHEAT	( 113)	55	22	2.2		
ALLERGEN_2012_2	gi 148887203 emb CAK50834.1  art	( 162)	54	22	3.6		
ALLERGEN_2012_2	gi 1709545 sp P51528.1 PA1_VESMC	( 300)	55	22	4.8		
ALLERGEN_2012_2	gi 50659885 gb AAT80662.1  lipid	( 115)	51	21	5		
ALLERGEN_2012_2	gi 50659891 gb AAT80665.1  lipid	( 115)	51	21	5		
ALLERGEN_2012_2	gi 50659879 gb AAT80659.1  lipid	( 115)	51	21	5		
ALLERGEN_2012_2	gi 897647 gb AAB48072.1  allergen	( 336)	55	22	5.3		
ALLERGEN_2012_2	gi 50659889 gb AAT80664.1  lipid	( 115)	50	21	6.2		
ALLERGEN_2012_2	gi 129614 sp P00784.1 PAPA1_CARPA	( 345)	53	22	8		
ALLERGEN_2012_2	gi 74035843 emb CAJ28931.1  Ves g	( 300)	52	22	8.8		
ALLERGEN_2012_2	gi 170718 gb AAA34279.1  alpha/be	( 313)	52	22	9.1		
ALLERGEN_2012_2	gi 21765 emb CAA26385.1  unnamed	( 313)	52	22	9.1		
ALLERGEN_2012_2	gi 170710 gb AAA34275.1  alpha-ty	( 318)	52	22	9.2		
ALLERGEN_2012_2	gi 1336813 gb AAB36121.1  Sol i 1	( 26)	42	18	9.4		
ALLERGEN_2012_2	gi 48428170 sp Q9NFQ4.1 ALL22_GLY	( 125)	48	20	9.8		

ALLERGEN\_2012\_2|gi|61970231|gb|AA57578.1| BW 16k ( 127) 48 20 9.9

>>ALLERGEN\_2012\_2|gi|417370|sp|P24296.2|NLTP1\_WHEAT RecN (113 aa)  
initn: 34 init1: 34 opt: 55 Z-score: 96.7 bits: 22.1 E(): 2.2  
Smith-Waterman score: 55; 45.000% identity (47.368% ungapped) in 20 aa overlap  
(21-39:54-73)

```

              10      20      30      40
SYHT0H      MRCVELLVHHGIRGRFGSEIVSWFHEQARSM AEEQLHC-CWLEPHCSVQI
              :. . . . . : : :
ALLERG IDCGHVDSLVRPCLSYVQGGPGPSGQCCDGVKNLHNQARSQSDRQSACNCLKGIARGIHN
              30      40      50      60      70      80
```

```

50
SYHT0H LWFQHPLL

ALLERG LNEEDNARSIPPKCGVNLPTISLNIIDCSRVR
              90      100      110
```

>>ALLERGEN\_2012\_2|gi|148887203|emb|CAK50834.1| art v 2 a (162 aa)  
initn: 33 init1: 33 opt: 54 Z-score: 93.0 bits: 22.0 E(): 3.6  
Smith-Waterman score: 54; 26.000% identity (28.261% ungapped) in 50 aa overlap (2-  
51:80-125)

```

              10      20      30
SYHT0H      MRCVELLVHHGIRGRFGSEIVSWFHEQARSM
              : : . . . : : . . . : . :
ALLERG MKPLCWDEIGKVAQAWAETRTPDCLSIHSDRCGENMAQGAINGSMVQL--WLDERLDYD
              50      60      70      80      90      100
```

```

              40      50
SYHT0H AEEQLHCCWLEPHCSVQILWLFQHPLL
              . . . : . : : :
ALLERG YNEN-KCIKMCGHYT-QIVWANSERVGCGRALCSNGWAYIIVCNYPDPGNVVGQKPY
              110      120      130      140      150      160
```

>>ALLERGEN\_2012\_2|gi|1709545|sp|P51528.1|PA1\_VESMC RecNa (300 aa)  
initn: 41 init1: 41 opt: 55 Z-score: 90.7 bits: 22.4 E(): 4.8  
Smith-Waterman score: 55; 50.000% identity (58.333% ungapped) in 28 aa overlap  
(14-39:220-245)

```

              10      20      30      40
SYHT0H      MRCVELLVHHGIRGRFGSEIVSWFHEQAR-SMAEEQLHCCWL
              : : : : : : : : : :
ALLERG IIHTSNYLGTETKILGTVDGYMNNGNKNNPGCGRFFSEVCS--HTRAVIYMAECIKHECCLI
              190      200      210      220      230      240
```

```

50
SYHT0H EPHCSVQILWLFQHPLL

ALLERG GIPRSKSSQPISRCTKQECVCVGLNAKKYPSRGSFYVPVESTAPFCNNKGKII
              250      260      270      280      290      300
```

>>ALLERGEN\_2012\_2|gi|50659885|gb|AAT80662.1| lipid trans (115 aa)  
initn: 26 init1: 26 opt: 51 Z-score: 90.4 bits: 21.0 E(): 5  
Smith-Waterman score: 51; 31.579% identity (33.333% ungapped) in 38 aa overlap (3-  
39:16-52)

```

              10      20      30      40
SYHT0H      MRCVELLVHHGIR-GRFGSEIVSWFHEQARSM AEEQLHCCWLEPHCS
              : . : : : : : . . . : : : : :
ALLERG MASSAVINLALVVALCMAVSAHAITCQVSSNLVPCF-DYVRSGGPVPFACNGIRTIN
              10      20      30      40      50
```

```

50
SYHT0H VQILWFQHPLL

ALLERG GLAKTTPDRQAACNCLKSLAGSVSGVNPNGNAESLPGKCGVNPYPKISTSTNCATVK
60      70      80      90      100     110

>>ALLERGEN_2012_2|gi|50659891|gb|AAT80665.1| lipid trans (115 aa)
  initn: 26 initl: 26 opt: 51 Z-score: 90.4 bits: 21.0 E(): 5
Smith-Waterman score: 51; 31.579% identity (33.333% ungapped) in 38 aa overlap (3-
39:16-52)

          10      20      30      40
SYHT0H      MRCVELLVHHGIR-GRFGSEIVSWFHEQARSMAEEQLHCCWLEPHCS
          :. . : :. : :. :. : :. :. : :. :. : :. :
ALLERG MASSAVINLALVVALCMAVSVAHAITCGQVSSNLVPCF-DYVRSGGPVPPACCNGIRTIN
          10      20      30      40      50

50
SYHT0H VQILWFQHPLL

ALLERG GLAKTTPDRQAACNCLKSLAGSVSGVNPNGNVESLPGKCGVNPYPKISTSTNCATVK
60      70      80      90      100     110

>>ALLERGEN_2012_2|gi|50659879|gb|AAT80659.1| lipid trans (115 aa)
  initn: 26 initl: 26 opt: 51 Z-score: 90.4 bits: 21.0 E(): 5
Smith-Waterman score: 51; 31.579% identity (33.333% ungapped) in 38 aa overlap (3-
39:16-52)

          10      20      30      40
SYHT0H      MRCVELLVHHGIR-GRFGSEIVSWFHEQARSMAEEQLHCCWLEPHCS
          :. . : :. : :. :. : :. :. : :. :. : :. :
ALLERG MASSAVIKLALVVALCMAVSVAHAITCGQVSSNLVPCF-DYVRSGGPVPPACCNGIRTIN
          10      20      30      40      50

50
SYHT0H VQILWFQHPLL

ALLERG GLAKTTPDRQAACNCLKSLAGSVSGVNPNGNAESLPGKCGVNPYPKISTSTNCATVK
60      70      80      90      100     110

>>ALLERGEN_2012_2|gi|897647|gb|AAB48072.1| allergen and (336 aa)
  initn: 40 initl: 40 opt: 55 Z-score: 90.0 bits: 22.5 E(): 5.3
Smith-Waterman score: 55; 50.000% identity (58.333% ungapped) in 28 aa overlap
(14-39:256-281)

          10      20      30      40
SYHT0H      MRCVELLVHHGIRGRFGSEIVSWFHEQAR-SMAEEQLH-CCWL
          :. : :. : :. :. : :. :. : :. :. : :. :
ALLERG IIHTSNYLGTEKTLGTVDGYMNGKNQPGCGRFFSEVCS--HSRAVIYMAECIKHECCLI
          230     240     250     260     270     280

50
SYHT0H EPHCSVQILWFQHPLL

ALLERG GIPKSKSSQPISSCTKQECVCVGLNAKKYPSRGSFYVPVESTAPFCNNKGKII
          290     300     310     320     330

>>ALLERGEN_2012_2|gi|50659889|gb|AAT80664.1| lipid trans (115 aa)
  initn: 26 initl: 26 opt: 50 Z-score: 88.9 bits: 20.7 E(): 6.2
Smith-Waterman score: 50; 31.579% identity (33.333% ungapped) in 38 aa overlap (3-
39:16-52)

```

```

              10      20      30      40
SYHT0H      MRCVELLVHHGIR-GRFGSEIVSWFHEQARSMEEQLHCCWLEPHCS
              :. . : :. :. . : :. . : :
ALLERG MASSAVINLALVVALCMAVSVAHAITCGQVSSNLVPCF-DYVRSDGPVPPACCNGIRTIN
              10      20      30      40      50

              50
SYHT0H VQILWFQHPLL

ALLERG GLAKTTPDRQAACNCLKSLAGSVSGVNPAGNAESLPGKCGVNPYPKISTSTNCATVK
              60      70      80      90      100     110

>>ALLERGEN_2012_2|gi|129614|sp|P00784.1|PAPA1_CARPA RecN (345 aa)
  initn: 28 initl: 28 opt: 53 Z-score: 86.8 bits: 21.9 E(): 8
Smith-Waterman score: 53; 33.333% identity (36.364% ungapped) in 24 aa overlap
(19-40:136-159)

              10      20      30      40
SYHT0H      MRCVELLVHHGIRGRFGSEIVSWFHEQARSMEEQLHC--CWLEPHCS
              : :. :. :. :. :. :. : :
ALLERG EKYTGSIAGNYTTTETLSYEEVLNDGDVNIPEYVDWRQKGAVTPVKNQGSCGSCWAFSAVV
              110     120     130     140     150     160

              50
SYHT0H VQILWFQHPLL

ALLERG TIEGIIKIRTGNLNEYSEQELLDCDRRSYGCNGGYPWSALQLVAQYGIHYRNTYPYEGVQ
              170     180     190     200     210     220

>>ALLERGEN_2012_2|gi|74035843|emb|CAJ28931.1| Ves g 1 al (300 aa)
  initn: 37 initl: 37 opt: 52 Z-score: 86.1 bits: 21.6 E(): 8.8
Smith-Waterman score: 52; 46.429% identity (54.167% ungapped) in 28 aa overlap
(14-39:220-245)

              10      20      30      40
SYHT0H      MRCVELLVHHGIRGRFGSEIVSWFHEQAR-SMAEEQLH-CWLEPHCS
              : :. :. :. :. :. :. : :
ALLERG ILHTSNHLGTERILGTVDGYMNNNGKNQPGCGRFFTEVCS--HSRAVIYMAECIKHECCLI
              190     200     210     220     230     240

              50
SYHT0H EPHCSVQILWFQHPLL

ALLERG GIPKSKSSQPISSCTKQECVCVGLNAKKYPSRGSFYVPVESTAPFCNNKGKII
              250     260     270     280     290     300

>>ALLERGEN_2012_2|gi|170718|gb|AAA34279.1| alpha/beta-g1 (313 aa)
  initn: 42 initl: 42 opt: 52 Z-score: 85.8 bits: 21.6 E(): 9.1
Smith-Waterman score: 52; 34.483% identity (38.462% ungapped) in 29 aa overlap
(27-52:279-307)

              10      20      30      40      50
SYHT0H      MRCVELLVHHGIRGRFGSEIVSWFHEQARSMEEQL--HC-CWLEPHCSVQILWFQ
              . :. :. :. :. :. :. : :
ALLERG YPSSQGSFQPSQONPQAQGSVQPQQLPQFAEIRNLALQTLPAMCNVYIPPHCSTTIAPFG
              250     260     270     280     290     300

SYHT0H HPLL

ALLERG IFGTN
              310

```

>>ALLERGEN\_2012\_2|gi|21765|emb|CAA26385.1| unnamed prote (313 aa)  
 initn: 42 initl: 42 opt: 52 Z-score: 85.8 bits: 21.6 E(): 9.1  
 Smith-Waterman score: 52; 34.483% identity (38.462% ungapped) in 29 aa overlap  
 (27-52:279-307)

```

              10      20      30      40      50
SYHT0H      MRCVELLVHHGIRGRFGSEIVSWFHEQARSMEEQL--HC-CWLEPHCSVQILWFG
              . :.:. : : : : .. :.:. : :
ALLERG YPSSQGSFQPSQQNPQAQGSVQPQQLPQFAEIRNLALQTL PAMCNVYIPPHCSTTIAPFG
          250      260      270      280      290      300

```

SYHT0H HPLL

ALLERG IFGTN  
 310

>>ALLERGEN\_2012\_2|gi|170710|gb|AAA34275.1| alpha-type g1 (318 aa)  
 initn: 42 initl: 42 opt: 52 Z-score: 85.7 bits: 21.6 E(): 9.2  
 Smith-Waterman score: 52; 34.483% identity (38.462% ungapped) in 29 aa overlap  
 (27-52:284-312)

```

              10      20      30      40      50
SYHT0H      MRCVELLVHHGIRGRFGSEIVSWFHEQARSMEEQL--HC-CWLEPHCSVQILWFG
              . :.:. : : : : .. :.:. : :
ALLERG YPSSQVSFQPSQLNPQAQGSVQPQQLPQFAEIRNLALQTL PAMCNVYIPPHCSTTIAPFG
          260      270      280      290      300      310

```

SYHT0H HPLL

ALLERG ISGTN

>>ALLERGEN\_2012\_2|gi|1336813|gb|AAB36121.1| Sol i l=anti (26 aa)  
 initn: 28 initl: 28 opt: 42 Z-score: 85.6 bits: 18.0 E(): 9.4  
 Smith-Waterman score: 42; 41.667% identity (41.667% ungapped) in 12 aa overlap  
 (35-46:2-13)

```

              10      20      30      40      50
SYHT0H ELLVHHGIRGRFGSEIVSWFHEQARSMEEQLHCCWLEPHCSVQILWFQHPLL
              : : : . :.
ALLERG                      SQPACSWYDVP CSHSEXSVYXTGVPS
                              10      20

```

>>ALLERGEN\_2012\_2|gi|48428170|sp|Q9NFQ4.1|ALL22\_GLYDO Re (125 aa)  
 initn: 36 initl: 36 opt: 48 Z-score: 85.3 bits: 20.2 E(): 9.8  
 Smith-Waterman score: 48; 44.444% identity (50.000% ungapped) in 18 aa overlap (5-  
 20:107-124)

```

              10      20      30
SYHT0H      MRCVELLVHHGIR--GRFGSEIVSWFHEQARSM
              :. :. :. :.
ALLERG CPIXKGDPIDFKYTTTVPAILPKVKAETAEVLVDHGV LACGRFGRQVE
          80      90      100      110      120

```

```

              40      50
SYHT0H EEQLHCCWLEPHCSVQILWFQHPLL

```

>>ALLERGEN\_2012\_2|gi|61970231|gb|AA57578.1| BW 16kDa al (127 aa)  
 initn: 28 initl: 28 opt: 48 Z-score: 85.2 bits: 20.2 E(): 9.9  
 Smith-Waterman score: 48; 27.273% identity (30.000% ungapped) in 33 aa overlap  
 (13-45:89-118)



```

              10      20      30      40
SYHT0H      MRCVELLVHHGIRGRFGSEIVSWFHEQARSMAEQLHCCWLE
              .:.:. .:. .:. .:. .:. .:.
ALLERG ESCMRGCCVAMKEMDDECVCCEWMKMMVENQKGRIGERLIK---EGVRDLKELPSKCGLSE
              60      70      80      90      100     110

              50
SYHT0H PHCSVQILWFQHPLL
              .:.
ALLERG MECGSRGNRYFV
              120

```

57 residues in 1 query sequences  
 367796 residues in 1603 library sequences  
 Scomplib [34t11]  
 start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012  
 Scan time: 0.020 Display time: 0.000

Function used was FASTA [version 3.4t11 Apr 17, 2002]  
 FASTA searches a protein or DNA sequence data bank  
 version 3.4t11 Apr 17, 2002  
 Please cite:  
 W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 121 aa
>SYHT0H2_I_32
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

```

367796 residues in 1603 sequences  
 Expectation\_n fit:  $\rho(\ln(x)) = 4.0029 \pm 0.00408$ ;  $\mu = 13.2344 \pm 0.212$   
 mean\_var=74.1704 $\pm$ 20.031, 0's: 2 Z-trim: 2 B-trim: 0 in 0/44  
 Lambda= 0.1489

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2  
 join: 36, opt: 24, gap-pen: -12/-2, width: 16  
 Scan time: 0.050

The best scores are:

	gi	gb	AAF65313.1	AF230384	( 137)	73	23	2.3
ALLERGEN_2012_2	gi	gb	AAF65312.1	AF230383	( 137)	73	23	2.3
ALLERGEN_2012_2	gi	gb	AAC97370.1	venom a	( 137)	71	23	3.1
ALLERGEN_2012_2	gi	sp	P35777.2	VA4_SOLIN	( 137)	67	22	5.6
ALLERGEN_2012_2	gi	gb	ADD74392.1	Sol s	( 137)	67	22	5.6

```

>>ALLERGEN_2012_2|gi|7638030|gb|AAF65313.1|AF230384_1 ve (137 aa)
  initn: 46 initl: 46 opt: 73 Z-score: 96.5 bits: 23.5 E(): 2.3
Smith-Waterman score: 73; 30.435% identity (34.146% ungapped) in 46 aa overlap
(74-119:68-108)

```

```

              50      60      70      80      90      100
SYHT0H GLGGAGNSKSLCQKPTSCHPPCLYPARNVPRAVYPKASCNLTGDSFGRPITATTDLKPC
              .:. .:. .:. .:. .:. .:.
ALLERG TVPKGENDPINPLKNVNVLYCAFSKRGIFTPKGVNTKQYINYCEKTIINP---ADIKQC
              40      50      60      70      80      90

              110      120
SYHT0H ASIDLKCVYNVDPRPNL
              .. .:. .:. .:.
ALLERG KKL-ISKCIKKVYDRPGPIIERSKNLLSCVRKKGVLTLTVYGKKK
              100      110      120      130

```

>>ALLERGEN\_2012\_2|gi|7638028|gb|AAF65312.1|AF230383\_1 ve (137 aa)  
 initn: 46 initl: 46 opt: 73 Z-score: 96.5 bits: 23.5 E(): 2.3  
 Smith-Waterman score: 73; 30.435% identity (34.146% ungapped) in 46 aa overlap  
 (74-119:68-108)

```

      50      60      70      80      90     100
SYHT0H GLGGAGNSKSLCQKPTSCHPPCLYPAARNVPRAVYPKASCNLTDGSFGRPITATTDLKPC
      :..: : : . . : : : :
ALLERG TVPKGENDPINPLKNVNVLYCAFSKRGIPTPKGVNTKQYINYCEKTIINP---ADIKQC
      40      50      60      70      80      90

      110     120
SYHT0H ASIDLSKCVYNVDPRPNL
      .: :..: .: :.
ALLERG KKL-ISKCIKKVYDRPGPIIERSKNLLSCVLKKGVLELTVYGKKK
      100     110     120     130

```

>>ALLERGEN\_2012\_2|gi|4038411|gb|AAC97370.1| venom allerg (137 aa)  
 initn: 45 initl: 45 opt: 71 Z-score: 94.2 bits: 23.0 E(): 3.1  
 Smith-Waterman score: 71; 34.783% identity (39.024% ungapped) in 46 aa overlap  
 (74-119:68-108)

```

      50      60      70      80      90     100
SYHT0H GLGGAGNSKSLCQKPTSCHPPCLYPAARNVPRAVYPKASCNLTDGSFGRPITATTDLKPC
      :..: : : . . : : : :
ALLERG TVPKRENDPINPLKNVNVLYCAFTKRGIFTPKGVNTKQYINYCE---KTIISPADIKLC
      40      50      60      70      80      90

      110     120
SYHT0H ASIDLSKCVYNVDPRPNL
      .: :..: .: :.
ALLERG KKIA-SKCVKKVYDRPGPIIERSKNLLSCVLKKGLLELTVYGKKK
      100     110     120     130

```

>>ALLERGEN\_2012\_2|gi|14424465|sp|P35777.2|VA4\_SOLIN RecN (137 aa)  
 initn: 45 initl: 45 opt: 67 Z-score: 89.6 bits: 22.2 E(): 5.6  
 Smith-Waterman score: 67; 34.783% identity (39.024% ungapped) in 46 aa overlap  
 (74-119:68-108)

```

      50      60      70      80      90     100
SYHT0H GLGGAGNSKSLCQKPTSCHPPCLYPAARNVPRAVYPKASCNLTDGSFGRPITATTDLKPC
      :..: : : . . : : : :
ALLERG TVPKRENDPINPLRNVNWYCAFTKRGIFTPKGVNTKQYINYCEKTAISP---ADIKLC
      40      50      60      70      80      90

      110     120
SYHT0H ASIDLSKCVYNVDPRPNL
      .: :..: .: :.
ALLERG KKIA-SKCVKKVYDRPGPIIERSKNLLSCVLKKGLLELTVYGKKK
      100     110     120     130

```

>>ALLERGEN\_2012\_2|gi|291092710|gb|ADD74392.1| Sol s 2 al (137 aa)  
 initn: 45 initl: 45 opt: 67 Z-score: 89.6 bits: 22.2 E(): 5.6  
 Smith-Waterman score: 67; 34.783% identity (39.024% ungapped) in 46 aa overlap  
 (74-119:68-108)

```

      50      60      70      80      90     100
SYHT0H GLGGAGNSKSLCQKPTSCHPPCLYPAARNVPRAVYPKASCNLTDGSFGRPITATTDLKPC
      :..: : : . . : : : :
ALLERG TVPKHENDPINPLRNVNWYCAFTKRGIFTPKGVNTKQYINYCEKTAISP---ADIKLC
      40      50      60      70      80      90

```

```

          110          120
SYHT0H ASIDLSKCVYNVDPRPNL
      .:  :::: .:  ::.
ALLERG KKIA-SKCVKKVYDRPGPIIERSKNLLSCVLKKGLLELTVYGKKK
          100          110          120          130

```

```

121 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.050 Display time: 0.010

```

```

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 31 aa
>SYHT0H2_I_33
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

```

```

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 2.3220+/-0.00343; mu= 14.9746+/- 0.181
mean_var=42.6403+/-11.992, 0's: 18 Z-trim: 18 B-trim: 0 in 0/42
Lambda= 0.1964

```

```

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.020

```

```

The best scores are:
                                opt bits E(1603)
ALLERGEN_2012_2|gi|291482314|emb|CBK62697.1| ragw ( 96) 57 21 1.9
ALLERGEN_2012_2|gi|291482306|emb|CBK62693.1| ragw ( 111) 57 21 2
ALLERGEN_2012_2|gi|291482310|emb|CBK62695.1| ragw ( 134) 57 22 2.2
ALLERGEN_2012_2|gi|291482308|emb|CBK62694.1| ragw ( 140) 57 22 2.2
ALLERGEN_2012_2|gi|285005079|emb|CBJ24286.1| ragw ( 164) 57 22 2.4
ALLERGEN_2012_2|gi|291197394|emb|CBK52317.1| ragw ( 164) 57 22 2.4
ALLERGEN_2012_2|gi|14422361|emb|CAC41634.1| plant ( 131) 50 20 8.6
ALLERGEN_2012_2|gi|14422359|emb|CAC41633.1| plant ( 131) 50 20 8.6
ALLERGEN_2012_2|gi|27818335|gb|AAO24900.1| major ( 132) 50 20 8.6

```

```

>>ALLERGEN_2012_2|gi|291482314|emb|CBK62697.1| ragweed h (96 aa)
initn: 74 initl: 48 opt: 57 Z-score: 98.1 bits: 21.3 E(): 1.9
Smith-Waterman score: 57; 69.231% identity (75.000% ungapped) in 13 aa overlap (2-13:74-86)

```

```

          10          20          30
SYHT0H          MPSHGG-PPPVGHESYLFVRKSQTDCTVFPS
      :: :: :::: : ::.
ALLERG SCFCYFDCDPTKNPGPPPGAPKGKAPAPSPPSGGGAPPPSGGEGGGDGGGGRR
          50          60          70          80          90

```

SYHT0H W

```

>>ALLERGEN_2012_2|gi|291482306|emb|CBK62693.1| ragweed h (111 aa)
initn: 74 initl: 48 opt: 57 Z-score: 97.6 bits: 21.4 E(): 2
Smith-Waterman score: 57; 69.231% identity (75.000% ungapped) in 13 aa overlap (2-13:95-107)

```

```

                                10      20      30
SYHT0H                      MPSHGG-PPPVGHESYLFVRKSQTDCTVFPS
                                :: :: ::: : :.
ALLERG TCFCYFDCDPTKNPGPPPGAPKGKAPAPSPPSGGGAPPPSGGEGGER
                                70      80      90      100     110

SYHT0H W

>>ALLERGEN_2012_2|gi|291482310|emb|CBK62695.1| ragweed h (134 aa)
  initn: 74 initl: 48 opt: 57 Z-score: 96.9 bits: 21.5 E(): 2.2
Smith-Waterman score: 57; 69.231% identity (75.000% ungapped) in 13 aa overlap (2-
13:75-87)

                                10      20      30
SYHT0H                      MPSHGG-PPPVGHESYLFVRKSQTDCTVFPS
                                :: :: ::: : :.
ALLERG SCFCYFDCDPTKNPGPPPGAPKGKAPAPSPPSGGGAPPPSGGEGGDGGGEGGGEGGGG
                                50      60      70      80      90     100

SYHT0H W

ALLERG EGGGEGGGGGEGGGEGGGGGEGGGEGGGER
                                110     120     130

>>ALLERGEN_2012_2|gi|291482308|emb|CBK62694.1| ragweed h (140 aa)
  initn: 99 initl: 48 opt: 57 Z-score: 96.8 bits: 21.6 E(): 2.2
Smith-Waterman score: 57; 69.231% identity (75.000% ungapped) in 13 aa overlap (2-
13:75-87)

                                10      20      30
SYHT0H                      MPSHGG-PPPVGHESYLFVRKSQTDCTVFPS
                                :: :: ::: : :.
ALLERG TCFCYFDCDPTKNPGPPPGAPKGKAPAPSPPSGGGAPPPSGGEGGDGPPPEGGEGGGDG
                                50      60      70      80      90     100

SYHT0H W

ALLERG GGEGGGDGGGEGGGEGGGDGGGDGGGEGGGEGGGER
                                110     120     130     140

>>ALLERGEN_2012_2|gi|285005079|emb|CBJ24286.1| ragweed h (164 aa)
  initn: 99 initl: 48 opt: 57 Z-score: 96.2 bits: 21.7 E(): 2.4
Smith-Waterman score: 57; 69.231% identity (75.000% ungapped) in 13 aa overlap (2-
13:95-107)

                                10      20      30
SYHT0H                      MPSHGG-PPPVGHESYLFVRKSQTDCTVFPS
                                :: :: ::: : :.
ALLERG TCFCYFDCDPTKNPGPPPGAPKGKAPAPSPPSGGGAPPPSGGEGGDGPPPEGGEGGGDG
                                70      80      90      100     110     120

SYHT0H W

ALLERG GGEGGGEGGGDGGGEGGGEGGGDGGGDGGGEGGGEGGGER
                                130     140     150     160

>>ALLERGEN_2012_2|gi|291197394|emb|CBK52317.1| ragweed h (164 aa)
  initn: 99 initl: 48 opt: 57 Z-score: 96.2 bits: 21.7 E(): 2.4

```

Smith-Waterman score: 57; 69.231% identity (75.000% ungapped) in 13 aa overlap (2-13:95-107)

```

                                10      20      30
SYHT0H                        MPSHGG-PPVVGHESYLFVRKSQTDCTVFPS
                                :: :: ::: : :.
ALLERG TCFCYFDCDPTKNPGPPPGAPKKGKAPAPSPPSGGGAPPPSGGEGGDGPPPPPEGGEGGGDG
                                70      80      90      100     110     120
```

SYHT0H W

```

ALLERG GEGEGGEGGGDGGGEGGEGGGDGGGEGGGEGGEGGER
                                130     140     150     160
```

>>ALLERGEN\_2012\_2|gi|14422361|emb|CAC41634.1| plantain p (131 aa)  
initn: 30 init1: 30 opt: 50 Z-score: 86.3 bits: 19.5 E(): 8.6  
Smith-Waterman score: 50; 40.000% identity (45.455% ungapped) in 25 aa overlap (8-29:66-90)

```

                                10      20      30
SYHT0H                        MPSHGGPPPVG-HES--YLFVRKSQTDCTVFPSW
                                : :: ::. ....: ::. ...
ALLERG AQVQLDCKDDSKKVIYSIGGETGQDGVYRLPVVGYHEDCEIKLVKSSRPDCSEIPKLAKG
                                40      50      60      70      80      90
```

```

ALLERG TIQTSKVDLSKNTTITEKTRHVKPLSFRAKTDAPGC
                                100     110     120     130
```

>>ALLERGEN\_2012\_2|gi|14422359|emb|CAC41633.1| plantain p (131 aa)  
initn: 30 init1: 30 opt: 50 Z-score: 86.3 bits: 19.5 E(): 8.6  
Smith-Waterman score: 50; 40.000% identity (45.455% ungapped) in 25 aa overlap (8-29:66-90)

```

                                10      20      30
SYHT0H                        MPSHGGPPPVG-HES--YLFVRKSQTDCTVFPSW
                                : :: ::. ....: ::. ...
ALLERG AQVQLDCKDDSKKVIYSIGGETDQDGVYRLPVVGYHEDCEIKLVKSSRPDCSEIPKLAKG
                                40      50      60      70      80      90
```

```

ALLERG TIQTSKVDLSKNTTITEKTRHVKPLSFRAKTDAPGC
                                100     110     120     130
```

>>ALLERGEN\_2012\_2|gi|27818335|gb|AAO24900.1| major polle (132 aa)  
initn: 72 init1: 50 opt: 50 Z-score: 86.3 bits: 19.5 E(): 8.6  
Smith-Waterman score: 50; 66.667% identity (66.667% ungapped) in 9 aa overlap (2-10:113-121)

```

                                10      20      30
SYHT0H                        MPSHGGPPPVGHESYLFVRKSQTDCTVFPSW
                                :. :: :::
ALLERG GATPAPPGAAPPPAAGGSPSPPADGGSPPPPADGGSPVPDGGSPPPSTH
                                90      100     110     120     130
```

31 residues in 1 query sequences  
367796 residues in 1603 library sequences  
Scomplib [34t11]  
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012  
Scan time: 0.020 Display time: 0.000

Function used was FASTA [version 3.4t11 Apr 17, 2002]

FASTA searches a protein or DNA sequence data bank  
version 3.4t11 Apr 17, 2002  
Please cite:  
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

/grid/.qsub\_tmp/1330373600kde\_agent/grid\_NbCuMBoIO: 31 aa  
>SYHT0H2\_I\_34  
vs /data/fasta/ALLERGEN\_2012\_2 library  
searching /data/fasta/ALLERGEN\_2012\_2 library

367796 residues in 1603 sequences  
Expectation\_n fit: rho(ln(x))= 4.1921+/-0.00334; mu= 1.0916+/- 0.174  
mean\_var=39.8651+/- 8.917, 0's: 22 Z-trim: 22 B-trim: 0 in 0/41  
Lambda= 0.2031

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2  
join: 36, opt: 24, gap-pen: -12/-2, width: 16  
Scan time: 0.020

The best scores are:

					opt	bits	E(1603)
ALLERGEN_2012_2	gi 23894244 emb CAD23614.1	tri m ( 404)	51	22		5.9	
ALLERGEN_2012_2	gi 74663809 sp Q8J077.1 SUB6_TRIS	( 405)	51	22		5.9	
ALLERGEN_2012_2	gi 6136163 sp P35779.2 VA3_SOLRI	( 211)	48	21		6.2	
ALLERGEN_2012_2	gi 291482306 emb CBK62693.1	ragw ( 111)	45	20		6.6	
ALLERGEN_2012_2	gi 14424466 sp P35778.2 VA3_SOLIN	( 234)	48	21		6.8	
ALLERGEN_2012_2	gi 291482308 emb CBK62694.1	ragw ( 140)	45	20		8.1	
ALLERGEN_2012_2	gi 285005079 emb CBJ24286.1	ragw ( 164)	45	20		9.3	
ALLERGEN_2012_2	gi 313118253 sp POCH89.1 HUGA_VES	( 31)	38	17		9.3	
ALLERGEN_2012_2	gi 11514279 pdb 1QNX A Chain A, V	( 209)	46	20		9.3	
ALLERGEN_2012_2	gi 162551 gb AAA30333.1	allergen ( 227)	46	20		10	

>>ALLERGEN\_2012\_2|gi|23894244|emb|CAD23614.1| tri m 2 al (404 aa)  
initn: 35 initl: 35 opt: 51 Z-score: 89.2 bits: 21.7 E(): 5.9  
Smith-Waterman score: 51; 30.000% identity (32.143% ungapped) in 30 aa overlap (1-28:39-68)

```

                                10      20
SYHT0H                      MVNDITNRGHISCCSW--PNLNWSPLRRHG
                                : :..... : : : : : : : : : :
ALLERG VLAALSAVNGAKILEAGPHAETIPNKYIVVMKKDVSDEAFSTHTTWLSQNLNRRRLMRRSG
      10      20      30      40      50      60

      30
SYHT0H GIL
      .
ALLERG SSKAMAGMQNKYSLGGIFRAYSGEFDDAMIKDISNHDDVDYIEPDFVVRTSTNGTNLTRQ
      70      80      90     100     110     120
```

>>ALLERGEN\_2012\_2|gi|74663809|sp|Q8J077.1|SUB6\_TRISH Rec (405 aa)  
initn: 35 initl: 35 opt: 51 Z-score: 89.2 bits: 21.7 E(): 5.9  
Smith-Waterman score: 51; 30.000% identity (32.143% ungapped) in 30 aa overlap (1-28:39-68)

```

                                10      20
SYHT0H                      MVNDITNRGHISCCSW--PNLNWSPLRRHG
                                : :..... : : : : : : : : : :
ALLERG VLAALSAVNGAKILEAGPHAETIPNKYIVVMKKDVSDEAFSTHTTWLSQNLNRRRLMRRSG
      10      20      30      40      50      60

      30
SYHT0H GIL
      .
ALLERG SSKAMAGMQNKYSLGGIFRAYSGEFDDAMIKDISNHDDVDYIEPDFVVRTSTNGTNLTRQ
      70      80      90     100     110     120
```

>>ALLERGEN\_2012\_2|gi|6136163|sp|P35779.2|VA3\_SOLRI RecNa (211 aa)  
 initn: 48 initl: 48 opt: 48 Z-score: 88.8 bits: 20.7 E(): 6.2  
 Smith-Waterman score: 48; 38.889% identity (38.889% ungapped) in 18 aa overlap (6-23:66-83)

		10	20	30	
SYHT0H		MVNDITNRGHISCCSWPNLNWSPLRRHGGIL			
		:: . . : : : : :			
ALLERG	VGFTDAEKDAIVNKHNELRQ	RVASGKEMRG	TNGPQPPAVKMPNLTWDPE	LATIAQRWANQ	
	40	50	60	70	80 90
ALLERG	CTFEHDACRNVERFAVGQNIAATSSSGKNKSTLSMILLWYNEVKDFDNRWISSFPDGN				
	100	110	120	130	140 150

>>ALLERGEN\_2012\_2|gi|291482306|emb|CBK62693.1|ragweed h (111 aa)  
 initn: 45 initl: 45 opt: 45 Z-score: 88.3 bits: 19.7 E(): 6.6  
 Smith-Waterman score: 45; 55.556% identity (55.556% ungapped) in 9 aa overlap (14-22:23-31)

		10	20	30	
SYHT0H		MVNDITNRGHISCCSWPNLNWSPLRRHGGIL			
		: : : : :			
ALLERG	MAAGLLVFLAISEIASVKGKLCEKPSLTW	SGKCKVKQTDKCDKRCIEWE	GAKHGACHKR		
	10	20	30	40	50 60
ALLERG	DSKATCFCYFDCDPTKNPGPPGAPKGA				
	70	80	90	100	110

>>ALLERGEN\_2012\_2|gi|14424466|sp|P35778.2|VA3\_SOLIN RecN (234 aa)  
 initn: 48 initl: 48 opt: 48 Z-score: 88.1 bits: 20.7 E(): 6.8  
 Smith-Waterman score: 48; 38.889% identity (38.889% ungapped) in 18 aa overlap (6-23:88-105)

		10	20	30	
SYHT0H		MVNDITNRGHISCCSWPNLNWSPLRRHGGIL			
		:: . . : : : : :			
ALLERG	VGFTDAEKDAIVNKHNELRQ	RVASGKEMRG	TNGPQPPAVKMPNLTWDPE	LATIAQRWANQ	
	60	70	80	90	100 110
ALLERG	CTFEHDACRNVERFAVGQNIAATSSSGKNKSTPNEMILLWYNEVKDFDNRWISSFPDGN				
	120	130	140	150	160 170

>>ALLERGEN\_2012\_2|gi|291482308|emb|CBK62694.1|ragweed h (140 aa)  
 initn: 45 initl: 45 opt: 45 Z-score: 86.7 bits: 19.7 E(): 8.1  
 Smith-Waterman score: 45; 55.556% identity (55.556% ungapped) in 9 aa overlap (14-22:3-11)

		10	20	30	
SYHT0H	MVNDITNRGHISCCSWPNLNWSPLRRHGGIL				
		: : : : :			
ALLERG	KLCEKPSLTW	SGKCKVKQTDKCDKRCIEWE	GAKHGACHKR	DSKATCFCY	
		10	20	30	40

>>ALLERGEN\_2012\_2|gi|285005079|emb|CBJ24286.1|ragweed h (164 aa)  
 initn: 45 initl: 45 opt: 45 Z-score: 85.7 bits: 19.7 E(): 9.3  
 Smith-Waterman score: 45; 55.556% identity (55.556% ungapped) in 9 aa overlap (14-22:23-31)

```

          10          20          30
SYHT0H      MVNDITNRGHISCCSWPNLNWSPLRRHGGIL
              :  ::::
ALLERG  MAAGLLVFVLA FSEIASVKGLCEKPSLTWSGKCKVKQTDKCDKRCIEWEGAKHGACHKR
          10          20          30          40          50          60

ALLERG  DSKATCFCYFDCDPTKNPGPPPGAPKGKAPAPSPPSGGGAPPPSGGEGGDGPPPPPEGGEG
          70          80          90          100         110         120

>>ALLERGEN_2012_2|gi|313118253|sp|P0CH89.1|HUGA_VESMC Re (31 aa)
  initn: 28 init1: 28 opt: 38 Z-score: 85.7 bits: 17.3 E(): 9.3
Smith-Waterman score: 38; 44.444% identity (44.444% ungapped) in 9 aa overlap (14-22:3-11)

```

```

          10          20          30
SYHT0H  MVNDITNRGHISCCSWPNLNWSPLRRHGGIL
              :  ::. .:
ALLERG      DRCIWPKEGFSIYWNIPTHFCHNFGVYFKEL
              10          20          30

>>ALLERGEN_2012_2|gi|11514279|pdb|1QNX|A Chain A, Ves V (209 aa)
  initn: 32 init1: 32 opt: 46 Z-score: 85.7 bits: 20.1 E(): 9.3
Smith-Waterman score: 46; 37.500% identity (37.500% ungapped) in 16 aa overlap (3-18:154-169)

```

```

          10          20          30
SYHT0H      MVNDITNRGHISCCSWPNLNWSPLRRHGGIL
              ::. . ::. . : :
ALLERG  TAAKYDDPVKLVKMWEDVKDYNPKKKFSGNDFLKTGHYTQMVWANTKEVGC GSIK YIQE
          130          140          150          160          170          180

ALLERG  KWHKH YLVCNYGPSGNFKNEELYQTK
          190          200

```

```

>>ALLERGEN_2012_2|gi|162551|gb|AAA30333.1| allergen 5 [V (227 aa)
  initn: 32 init1: 32 opt: 46 Z-score: 85.1 bits: 20.1 E(): 10
Smith-Waterman score: 46; 37.500% identity (37.500% ungapped) in 16 aa overlap (3-18:172-187)

```

```

          10          20          30
SYHT0H      MVNDITNRGHISCCSWPNLNWSPLRRHGGIL
              ::. . ::. . : :
ALLERG  TAAKYDDPVKLVKMWEDVKDYNPKKKFSGNDFLKTGHYTQMVWANTKEVGC GSIK YIQE
          150          160          170          180          190          200

ALLERG  KWHKH YLVCNYGPSGNFMNEELYQTK
          210          220

```

```

31 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.020 Display time: 0.000

```

```

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 115 aa

```



```

>SYHT0H2_I_35
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 3.7938+/-0.00334; mu= 10.5821+/- 0.172
mean_var=54.5589+/-14.161, 0's: 3 Z-trim: 3 B-trim: 20 in 2/41
Lambda= 0.1736

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.050
The best scores are:

```

					opt	bits	E(1603)
ALLERGEN_2012_2	gi 75219081 sp O22108 O22108_WHEA	( 285)	71	26		1	
ALLERGEN_2012_2	gi 21773 emb CAA31685.1  unnamed	( 307)	71	26		1.1	
ALLERGEN_2012_2	gi 886965 emb CAA59339.1  low mol	( 261)	67	25		2	
ALLERGEN_2012_2	gi 94400907 ref NP_001035360.1  a	( 92)	61	23		2.8	
ALLERGEN_2012_2	gi 633938 gb AAB30434.1  albumin	( 265)	65	24		2.8	
ALLERGEN_2012_2	gi 88770352 gb ABD51779.1  allerg	( 94)	61	23		2.8	
ALLERGEN_2012_2	gi 3021324 emb CAA06305.1  Aspfl	( 125)	61	23		3.4	
ALLERGEN_2012_2	gi 886967 emb CAA59340.1  low mol	( 276)	64	24		3.4	
ALLERGEN_2012_2	gi 6136162 sp P35776.2 VA2_SOLRI	( 119)	60	22		3.9	
ALLERGEN_2012_2	gi 168419914 gb ACA23876.1  Pas n	( 265)	63	24		4	
ALLERGEN_2012_2	gi 3319897 emb CAA76841.1  albumi	( 585)	65	24		4.7	
ALLERGEN_2012_2	gi 6687188 emb CAB64867.1  albumi	( 608)	65	24		4.8	
ALLERGEN_2012_2	gi 1351908 sp P49064.1 ALBU_FELCA	( 608)	65	24		4.8	
ALLERGEN_2012_2	gi 100834 pir  S16031 alpha-amyla	( 168)	59	22		5.9	
ALLERGEN_2012_2	gi 21713 emb CAA35597.1  unnamed	( 168)	59	22		5.9	
ALLERGEN_2012_2	gi 2497701 sp Q28133.1 ALL2_BOVIN	( 172)	58	22		7.1	
ALLERGEN_2012_2	gi 244610 gb AAB21323.1  tetramer	( 18)	49	19		7.7	
ALLERGEN_2012_2	gi 335331566 gb AEH31546.1  low m	( 369)	60	23		8.3	
ALLERGEN_2012_2	gi 2506771 sp P16968.2 IAA1_HORVU	( 146)	56	22		9	
ALLERGEN_2012_2	gi 18641 emb CAA37044.1  glycinin	( 562)	61	23		9.2	
ALLERGEN_2012_2	gi 1170095 sp P46419.1 GSTM1 DERP	( 219)	57	22		9.9	

```

>>ALLERGEN_2012_2|gi|75219081|sp|O22108|O22108_WHEAT LMM (285 aa)
initn: 40 initl: 40 opt: 71 Z-score: 102.8 bits: 25.6 E(): 1
Smith-Waterman score: 71; 36.364% identity (39.024% ungapped) in 44 aa overlap
(45-86:8-50)

```

20	30	40	50	60	70
SYHT0H	KLYCQMFERSGKFGDLELDSYLG	NWPNW	PSWQLKIPLPKTNIMPSTMLVSSCA--QCT		
			: :: . :::	. . . : ::	:
ALLERG		RCIPGLERP	WQQQ-PLPPQQT	FPQQPLFSQQQQQQLF	
		10	20	30	

```

80 90 100 110
SYHT0H PGCVSQSLMQPNRWIVWKAYNSNHLRLKTLRLHRLKQMCVQCRS
: : : .:::
ALLERG PQQPSFSQQQPPFWQQQPPFSQQQPILPQQPPFSQQQQLVLPQQSPFSQQQQLILPPQQQ
40 50 60 70 80 90

```

```

>>ALLERGEN_2012_2|gi|21773|emb|CAA31685.1| unnamed prote (307 aa)
initn: 40 initl: 40 opt: 71 Z-score: 102.4 bits: 25.6 E(): 1.1
Smith-Waterman score: 71; 36.364% identity (39.024% ungapped) in 44 aa overlap
(45-86:31-73)

```

20	30	40	50	60	70
SYHT0H	KLYCQMFERSGKFGDLELDSYLG	NWPNW	PSWQLKIPLPKTNIMPSTMLVSSCA--QCT		
			: :: . :::	. . . : ::	:
ALLERG	MKTFLVFALLAVAATSAIAQMETRCIPGLERP	WQQQ-PLPPQQT	FPQQPLFSQQQQQQLF		
	10	20	30	40	50

```

      80      90      100      110
SYHT0H PGCVSQSLMQPNRWIVWKAYNSNHLKTLRLHRLKQMCVQCRS
      :   :   :   :   :
ALLERG PQQPSFSQQQPPFWQQPPFSQQQPILPQQPPFSQQQQLVLPQQPPFSQQQQPVLPPQQS
      60      70      80      90      100      110

>>ALLERGEN_2012_2|gi|886965|emb|CAA59339.1| low molecula (261 aa)
  initn: 44 initl: 44 opt: 67 Z-score: 97.8 bits: 24.6 E(): 2
Smith-Waterman score: 67; 25.352% identity (29.508% ungapped) in 71 aa overlap
(48-114:140-204)

```

```

      20      30      40      50      60      70
SYHT0H CQMFERSGKFGDLELDSYLGWPNWPNWRSWQLKIPLPKTNIM---PSTMLVSSCAQCTP
      :   :   :   :   :   :
ALLERG QQQPPFSQQQQPILLQQPPFSQHQQPVLPPQQIPSVQPSILQQLNPKVFLQQ--QCSP
      110      120      130      140      150      160

```

```

      80      90      100      110
SYHT0H GCVSQSLMQPNRWIVWKAYNSNHLKTLRLHRLKQMCVQCRS
      .   :   :   :   :   :   :
ALLERG VAMPQSLARSQ--MLWQS--SCHVMQQQCCRQLPQIPEQSRDYAIRAIYSIVLQEQQH
      170      180      190      200      210      220

```

```

ALLERG QGLNQPQQQPPQQSVQGVSPQQQQKQLGQCSFQQPQQ
      230      240      250      260

```

```

>>ALLERGEN_2012_2|gi|94400907|ref|NP_001035360.1| allerg (92 aa)
  initn: 49 initl: 49 opt: 61 Z-score: 95.0 bits: 22.5 E(): 2.8
Smith-Waterman score: 61; 40.000% identity (47.059% ungapped) in 20 aa overlap
(57-76:56-72)

```

```

      30      40      50      60      70      80
SYHT0H FGDLELDSYLGWPNWPNWRSWQLKIPLPKTNIMPSTMLVSSCAQCTPGCVSQSLMQPNRW
      :   :   :   :   :   :
ALLERG GGFGGLGGRGKCPSEIFSRCDGRCQRFQPNVVPKPLCIKICA---PGCVCRLLGYLRNKK
      30      40      50      60      70      80

```

```

      90      100      110
SYHT0H IVWKAYNSNHLKTLRLHRLKQMCVQCRS

```

```

ALLERG KVCVPRSKCG
      90

```

```

>>ALLERGEN_2012_2|gi|633938|gb|AAB30434.1| albumin [Cani (265 aa)
  initn: 65 initl: 65 opt: 65 Z-score: 95.0 bits: 24.1 E(): 2.8
Smith-Waterman score: 65; 47.059% identity (47.059% ungapped) in 17 aa overlap
(12-28:196-212)

```

```

      10      20      30      40
SYHT0H MIIIPATGFNLKKLYCQMFERSGKFGDLELDSYLGWPNWPN
      :   :   :   :   :
ALLERG TLEKCCATDDPPTCYAKVLDEFKPLVDEPQNLVKTNCELFEKLGEGYGFQNALLVRYTKKA
      170      180      190      200      210      220

```

```

      50      60      70      80      90      100
SYHT0H WPWRSWQLKIPLPKTNIMPSTMLVSSCAQCTPGCVSQSLMQPNRWIVWKAYNSNHLKTL

```

```

ALLERG PQVSTPTLVVEVSRKLGKVGTKCCKKPESERMSCADDFLS
      230      240      250      260

```

```

>>ALLERGEN_2012_2|gi|88770352|gb|ABD51779.1| allergen Ap (94 aa)
  initn: 49 initl: 49 opt: 61 Z-score: 94.9 bits: 22.5 E(): 2.8

```

Smith-Waterman score: 61; 40.000% identity (47.059% ungapped) in 20 aa overlap  
(57-76:56-72)

```

      30      40      50      60      70      80
SYHT0H FGDLELDSYLGWPNWPWRSWQLKIPLPKNIMPSTMLVSSCAQCTPGCVSQSLMQPNRW
      :...: . . . : : : :
ALLERG GGFGLGGRGKCPSEIFSRCDGRCQRFPCNVVPKPLCIKICA--PGCVCLGYLRNKK
      30      40      50      60      70      80

      90     100     110
SYHT0H IVWKAYNSNHRLKTLRLHRLKQMCVQCRS

ALLERG KVCVPRSKCLPG
      90
```

>>ALLERGEN\_2012\_2|gi|3021324|emb|CAA06305.1| Aspfl aller (125 aa)  
initn: 44 initl: 44 opt: 61 Z-score: 93.5 bits: 22.7 E(): 3.4  
Smith-Waterman score: 61; 23.077% identity (23.438% ungapped) in 65 aa overlap  
(15-78:52-116)

```

      10      20      30      40
SYHT0H MIIARPATGFNLKKLYCQMFERSGKFGDLELDSYLGWPNWP-
      : : . . : : . : : : :
ALLERG KTGSSYPHWFTNGYDGDGKLIKGRMPIKFGKADCDRPPKHGKDGMDGDDHYLEFPTFPD
      30      40      50      60      70      80

      50      60      70      80      90     100
SYHT0H WRSWQLKIPLPKNIMPSTMLVSSCAQCTPGCVSQSLMQPNRWIVWKAYNSNHRLKTLRL
      :...: : . : . . . : : :
ALLERG GHDKYFDSKKPKEDPGPARVIYTPNKFVFCGIVAHERGNGQDLR
      90     100     110     120

      110
SYHT0H HRLKQMCVQCRS
```

>>ALLERGEN\_2012\_2|gi|886967|emb|CAA59340.1| low molecula (276 aa)  
initn: 44 initl: 44 opt: 64 Z-score: 93.5 bits: 23.8 E(): 3.4  
Smith-Waterman score: 68; 25.000% identity (29.577% ungapped) in 84 aa overlap  
(40-114:140-219)

```

      10      20      30      40      50      60
SYHT0H GFNLKKLYCQMFERSGKFGDLELDSYLGWPNWPWRSWQLKIPLPKNI---MPSTML-V
      : : : . . : : . : :
ALLERG QRPPFSQQQQQPVLPPQPPFSQQQQQPILPQQPPFSLHQQPVLPQQQIPYVQPSILQQL
      110     120     130     140     150     160

      70      80      90     100     110
SYHT0H SSCA-----QCTPGCVSQSLMQPNRWIVWKAYNSNHRLKTLRLHRLKQMCVQCRS
      . : : : : : : : : : : : : : : :
ALLERG NPCKVFLQQQCSPVAMPQSLARSQ--MLWQS--SCHVMQQQCCQQLPRIPEQSRDAIRA
      170     180     190     200     210     220

ALLERG IIYSIVLQEQHGQGFNQPPQQPPQQSVQGVSQPPQQQKQLGQCSFQQPPQQ
      230     240     250     260     270
```

>>ALLERGEN\_2012\_2|gi|6136162|sp|P35776.2|VA2\_SOLRI RecNa (119 aa)  
initn: 60 initl: 60 opt: 60 Z-score: 92.4 bits: 22.4 E(): 3.9  
Smith-Waterman score: 60; 36.000% identity (36.000% ungapped) in 25 aa overlap  
(65-89:12-36)



>>ALLERGEN\_2012\_2|gi|1351908|sp|P49064.1|ALBU\_FELCA RecN (608 aa)  
 initn: 65 initl: 65 opt: 65 Z-score: 90.7 bits: 24.5 E(): 4.8  
 Smith-Waterman score: 65; 47.059% identity (47.059% ungapped) in 17 aa overlap  
 (12-28:410-426)

```

                                10      20      30      40
SYHT0H      MIIIARPATGFNLKKLYCQMFERSGKFGDLELDSYLGWPN
                                :: : : ::::: :::
ALLERG TLEKCCATDDPPACYAHVFDEFKPLVEEPHNLVKTNCELFEKLGEYGFQNALLVRYTKKV
      380      390      400      410      420      430

                                50      60      70      80      90     100
SYHT0H WPWRSWQLKIPLPKTNIMPSTMLVSSCAQCTPGCVSQSLMQPNRWIVWKAYNSNHRLKTL
ALLERG PQVSTPTLVEVSRSLGKVGSKCCTHPEAERLSAEDYLSVVLNRLCVLHEKTPVSEVTK
      440      450      460      470      480      490

```

>>ALLERGEN\_2012\_2|gi|100834|pir||S16031 alpha-amylase in (168 aa)  
 initn: 38 initl: 38 opt: 59 Z-score: 89.2 bits: 22.3 E(): 5.9  
 Smith-Waterman score: 59; 24.286% identity (31.481% ungapped) in 70 aa overlap  
 (55-115:36-98)

```

      30      40      50      60      70
SYHT0H GKFGDLELDSYLGWPNWPWRSWQLKIPLPKTNIMPST---MLVSSCAQCTPGC-----
                                ::::: : : ::: : :::
ALLERG SCSLLLLAAVLLSVLAAASASGSCVPGVAFRTNLLPHCRDYVLQQTCTGFTPGSKLPEWM
      10      20      30      40      50      60

      80      90     100     110
SYHT0H VSQSLMQPNRWIVWKAYNSNHRLKTLRLHRLKQMCVQCERS
      .: :::::.. . : : .. : ::::: :...
ALLERG TSASIYSPGKPYLAKLYCCQE-----LAEISQQC-RCEALRYFIALPVSPQVDPVPSGN
      70      80      90      100     110

ALLERG VGESGLIDLPGCPREMWDVFRLLVAPGQCENLATIHNVRYPVEQPLWI
      120     130     140     150     160

```

>>ALLERGEN\_2012\_2|gi|21713|emb|CAA35597.1| unnamed prote (168 aa)  
 initn: 38 initl: 38 opt: 59 Z-score: 89.2 bits: 22.3 E(): 5.9  
 Smith-Waterman score: 59; 24.286% identity (31.481% ungapped) in 70 aa overlap  
 (55-115:36-98)

```

      30      40      50      60      70
SYHT0H GKFGDLELDSYLGWPNWPWRSWQLKIPLPKTNIMPST---MLVSSCAQCTPGC-----
                                ::::: : : ::: : :::
ALLERG SCSLLLLAAVLLSVLAAASASGSCVPGVAFRTNLLPHCRDYVLQQTCTGFTPGSKLPEWM
      10      20      30      40      50      60

      80      90     100     110
SYHT0H VSQSLMQPNRWIVWKAYNSNHRLKTLRLHRLKQMCVQCERS
      .: :::::.. . : : .. : ::::: :...
ALLERG TSASIYSPGKPYLAKLYCCQE-----LAEISQQC-RCEALRYFIALPVSPQVDPVPSGN
      70      80      90      100     110

ALLERG VGESGLIDLPGCPREMWDVFRLLVAPGQCENLATIHNVRYPVEQPLWI
      120     130     140     150     160

```

>>ALLERGEN\_2012\_2|gi|2497701|sp|Q28133.1|ALL2\_BOVIN RecN (172 aa)  
 initn: 29 initl: 29 opt: 58 Z-score: 87.8 bits: 22.1 E(): 7.1  
 Smith-Waterman score: 58; 27.778% identity (31.915% ungapped) in 54 aa overlap  
 (68-115:14-66)

```

      40      50      60      70      80      90
SYHT0H NWPNWPWRSWQLKIPLPKTNIMPSTMLVSSC-AQCTPGCVSQSLMQPNRW-IVWKAYNSN
      : :: ::. . . : . :. : . : .
ALLERG          MKAVFLTLLFGLVCTAQETPAEIDPSKI-PGEWRIIYAAADNK
              10      20      30      40

```

```

      100      110
SYHT0H HRL---KTLRLHRLKQMCVQ-CRS
      ..      :: . . :.. ::
ALLERG DKIVEGGPLRNYRRIECINDCESLSITFYLKDQGTCLLLTEVAKRQEGYVYVLEFYGTN
              50      60      70      80      90      100

```

>>ALLERGEN\_2012\_2|gi|244610|gb|AAB21323.1| tetrameric al (18 aa)  
 initn: 34 initl: 34 opt: 49 Z-score: 87.2 bits: 18.7 E(): 7.7  
 Smith-Waterman score: 49; 47.368% identity (50.000% ungapped) in 19 aa overlap  
 (36-54:1-18)

```

      10      20      30      40      50      60
SYHT0H RPATGFNLKKLYCQMFERSGKFGDLELDSYLGWPNWPWRSWQLKIPLPKTNIMPSTMLV
      .::      :: : : : ::
ALLERG          IGNEDCTPWMS-TLITPLP
              10

```

```

      70      80      90      100      110
SYHT0H SSCAQCTPGCVSQSLMQPNRWIVWKAYNSNHRLKTLRLHRLKQMCVQCRS

```

>>ALLERGEN\_2012\_2|gi|335331566|gb|AEH31546.1| low molecu (369 aa)  
 initn: 44 initl: 44 opt: 60 Z-score: 86.5 bits: 23.0 E(): 8.3  
 Smith-Waterman score: 63; 24.051% identity (32.759% ungapped) in 79 aa overlap  
 (43-112:183-249)

```

      20      30      40      50      60
SYHT0H LKKLYCQMFERSGKFGDLELDSYLGWPNWPWRSWQLKIPLPKTNI---MPSTML-VSSC
      :. . : .. :. :. :. :. :. :. :
ALLERG QSPFSQQQIVLQQQPFLQQQPPSLPQQPPFSQQQQQLVLPQQQIPFVHPSILQQLNPC
      160      170      180      190      200      210

```

```

      70      80      90      100      110
SYHT0H A-----QCTPGCVSQSLMQPNRWIVWKAYNSNHRLKTLRLHRLKQMCVQCRS
      ::. : . :. : . . . : :. :. :
ALLERG KVFLQQQCSPVAMPQSLAR-----SQMLQQSSCHVMQQQCCQLPQIPQQSRY
      220      230      240      250      260

```

```

ALLERG EAIRAIIYSIILQEQQVQGSIQTPQQQPQQLGQCVSQPQQQSQQQLGQQPQQQQLAQGT
      270      280      290      300      310      320

```

>>ALLERGEN\_2012\_2|gi|2506771|sp|P16968.2|IAA1\_HORVU RecN (146 aa)  
 initn: 37 initl: 37 opt: 56 Z-score: 85.9 bits: 21.5 E(): 9  
 Smith-Waterman score: 57; 30.435% identity (37.838% ungapped) in 46 aa overlap  
 (37-82:17-53)

```

      10      20      30      40      50      60
SYHT0H PATGFNLKKLYCQMFERSGKFGDLELDSYLGWPNWPWRSWQLKIPLPKTNIMPSTMLVS
      :. : : . . . :. :. :. :. :
ALLERG          PTSVAVDQGSMSVNSPGEW-CWPGMGYPV-YPFPRCRA-----LVK
              10      20      30

```

```

      70      80      90      100      110
SYHT0H SCAQCTPGCVSQSLMQPNRWIVWKAYNSNHRLKTLRLHRLKQMCVQCRS
      : :. : : :. :. :
ALLERG S--QCAGGQVVESIQKDCCRQIAAIGDEWCICGALGSMRGS MYKELGVALADDKATVAEV
      40      50      60      70      80      90

```

```
>>ALLERGEN_2012_2|gi|18641|emb|CAA37044.1| glycinin [Gly (562 aa)
  initn: 50 initl: 50 opt: 61 Z-score: 85.7 bits: 23.4 E(): 9.2
Smith-Waterman score: 61; 26.471% identity (30.000% ungapped) in 34 aa overlap
(53-83:82-114)
```

```

              30          40          50          60          70
SYHT0H RSGKFGDLELDSYLGWPNWPNWPSWQLKIPLPKNIMPSTMLVSS---CAQCTPGCVSQS
              :. . .: . . . . . : : : : .
ALLERG LIQTWNSQHPELKCAGVTVSKLTLNRNGLHLPSPYPRMIIIAQGKGALQCKPGC-PET
              60          70          80          90          100          110

              80          90          100          110
SYHT0H LMQPNRWIVWKAYNSNHRLKTLRLHRLKQMCVQCRS
              . . .
ALLERG FEEPQEQSNRRGRSRSQKQLQDSHQKIRHFNEGDVLVIPPGVVPYWTYNTGDEPVVAISLL
              120          130          140          150          160          170
```

```
>>ALLERGEN_2012_2|gi|1170095|sp|P46419.1|GSTM1_DERPT Rec (219 aa)
  initn: 32 initl: 32 opt: 57 Z-score: 85.2 bits: 22.0 E(): 9.9
Smith-Waterman score: 57; 29.508% identity (30.508% ungapped) in 61 aa overlap
(11-69:159-219)
```

```

              10          20          30
SYHT0H MIIIPATGFNLKKLYCQMFERSGK-FGDLE-LDSYLGW
              :. . :. . . : : : : :
ALLERG LKSLPDCCLKMSKFVGEHAFIAGANISYVDFNLYEYLCHVKVMVPEVFGQFENLKRYVER
              130          140          150          160          170          180

              40          50          60          70          80          90
SYHT0H WPNWPNWPSWQLKIPLPKNIMPSTMLVSSCAQCTPGCVSQSLMQPNRWIVWKAYNSNHRL
              . : : . : : :. . . : :
ALLERG MESLPRVSDYIKKQPKTFNAPTSKWNASYA
              190          200          210

              100          110
SYHT0H KTLRLHRLKQMCVQCRS
```

```
115 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.050 Display time: 0.010
```

```
Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448
```

```
/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 64 aa
>SYHT0H2_I_36
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library
```

```
367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 4.5451+/-0.00322; mu= 4.1974+/- 0.167
mean_var=55.6035+/-16.017, 0's: 7 Z-trim: 7 B-trim: 7 in 1/41
Lambda= 0.1720
```

```
FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
```

Scan time: 0.020  
The best scores are:

					opt	bits	E(1603)
ALLERGEN_2012_2	gi 71360928 emb CAJ19705.1  non-s	( 114)	59	22	2.9		
ALLERGEN_2012_2	gi 21743 emb CAA43331.1  high mol	( 830)	64	24	5.9		
ALLERGEN_2012_2	gi 13183175 gb AAK15088.1 AF24000	( 153)	55	21	7.4		
ALLERGEN_2012_2	gi 209165427 gb ACI41244.1  2S al	( 153)	55	21	7.4		
ALLERGEN_2012_2	gi 166235350 pdb 2JON A Chain A,	( 101)	53	20	7.5		
ALLERGEN_2012_2	gi 886967 emb CAA59340.1  low mol	( 276)	57	22	8.3		
ALLERGEN_2012_2	gi 736319 emb CAA27052.1  gluteni	( 838)	62	23	8.4		
ALLERGEN_2012_2	gi 170743 gb AAB02788.1  HMW glut	( 815)	61	23	9.7		

>>ALLERGEN\_2012\_2|gi|71360928|emb|CAJ19705.1| non-specif (114 aa)  
initn: 38 init1: 38 opt: 59 Z-score: 94.6 bits: 21.9 E(): 2.9  
Smith-Waterman score: 59; 26.471% identity (28.125% ungapped) in 34 aa overlap  
(12-43:27-60)

	10	20	30	40
SYHT0H	MRLWDTQPGVHCAQLDTSMDGMMLVFGKGIL--SCQLLQGQLGQ			
	:.. .... . ....: : .. ..::			
ALLERG	MEMVNK	IACFVLL	CMVVVAP	HAEALTCGQVTSTLAPCLPYLMNRGPLRNCCDGVKGLLGQ
	10	20	30	40 50 60

	50	60
SYHT0H	LPRYESSSRSPNFPDRSNIWQ	
ALLERG	AKTTVDRAAACLKSAASSFTGLNLGKAAALPNTCSVNIPYKISPSTDCSKVQ	
	70	80 90 100 110

>>ALLERGEN\_2012\_2|gi|21743|emb|CAA43331.1| high molecula (830 aa)  
initn: 261 init1: 57 opt: 64 Z-score: 89.2 bits: 23.8 E(): 5.9  
Smith-Waterman score: 64; 35.135% identity (35.135% ungapped) in 37 aa overlap  
(28-64:435-471)

	10	20	30	40	50
SYHT0H	MRLWDTQPGVHCAQLDTSMDGMMLVFGKGILSCQLLQGQLGQLPRYESSSRSPNFP				
	:: . : ::: :: : . . : .				
ALLERG	GQEQQPRQLQ	QPEQQGQ	QPEQQGQ	QPGQGEQQG	QPGQGGQPGQGP
	410	420	430	440	450 460

	60
SYHT0H	DRSNIWQ
	... :
ALLERG	QQSQQGQPGYYPTSPQQSGQLQPPAQGQQPGQEQQGQQPGQGGQQPGQGGQPGQGGQPG
	470 480 490 500 510 520

>>ALLERGEN\_2012\_2|gi|13183175|gb|AAK15088.1|AF240005\_1 2 (153 aa)  
initn: 46 init1: 46 opt: 55 Z-score: 87.5 bits: 21.0 E(): 7.4  
Smith-Waterman score: 55; 52.174% identity (57.143% ungapped) in 23 aa overlap  
(34-54:44-66)

	10	20	30	40	50	60
SYHT0H	WDTQPGVHCAQLDTSMDGMMLVFGKGILSC-QLLQG-QLGQLPRYESSSRSPNFPDRSN					
	: : ::: :. . :: :...::					
ALLERG	AMVALASAT	YTTT	TVT	TTAIDDEAN	QSQCRQLQGRQFRSCQRYLSQGRSPYGGEDE	
	20	30	40	50	60	70

SYHT0H	IWQ
ALLERG	VLEMSTGNQQSEQSLRDCCQQLRNVDERCCEAIRQAVRQQQEGGYQEGQSQQVYQRAR
	80 90 100 110 120 130

>>ALLERGEN\_2012\_2|gi|209165427|gb|ACI41244.1| 2S albumin (153 aa)



initn: 46 initl: 46 opt: 55 Z-score: 87.5 bits: 21.0 E(): 7.4  
 Smith-Waterman score: 55; 52.174% identity (57.143% ungapped) in 23 aa overlap  
 (34-54:44-66)

```

          10          20          30          40          50          60
SYHT0H WDTQPGVHCAQLDTSMDGMMLVFGKGILSC-QLLQG-QLGQLPRYESSRSPNFPDRSN
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ALLERG AMVALASATYTTTTVTTTAIDDEANQQSQQCRQQLQGRQFRSCQRYLSQGRSPYGGEEDE
          20          30          40          50          60          70

```

SYHT0H IWQ

```

ALLERG VLEMSTGNQQSEQSLRDCCQQLGNVDERCRCEAIRQAVRQQQEGGYQEGQSQQVYQRAR
          80          90          100          110          120          130

```

>>ALLERGEN\_2012\_2|gi|166235350|pdb|2JON|A Chain A, Solut (101 aa)  
 initn: 34 initl: 34 opt: 53 Z-score: 87.3 bits: 20.4 E(): 7.5  
 Smith-Waterman score: 53; 27.869% identity (32.075% ungapped) in 61 aa overlap (7-61:43-101)

```

          10          20          30
SYHT0H MRLWDTQPGVHCAQLDTSMDG---MMLVF---GKG
          : : : : : : : : : : : : : : : : : : : : : :
ALLERG WCVPKPGVSDDQLTGNINYACSQGIDCGPIQPGGACFEPNTVKAHAAYVMNLYYQHAGRN
          20          30          40          50          60          70

```

```

          40          50          60
SYHT0H ILSCQLLQGQLGQLPRYESSRSPNFPDRSNIWQ
          . . . . : : : : : : : : : : : : : : : :
ALLERG SWNCDF--SQTATLTNTNPSYGACNFPSPGSN
          80          90          100

```

>>ALLERGEN\_2012\_2|gi|886967|emb|CAA59340.1| low molecula (276 aa)  
 initn: 38 initl: 38 opt: 57 Z-score: 86.6 bits: 21.7 E(): 8.3  
 Smith-Waterman score: 57; 47.619% identity (50.000% ungapped) in 21 aa overlap  
 (33-52:199-219)

```

          10          20          30          40          50          60
SYHT0H LWDTQPGVHCAQLDTSMDGMMLVFGKGILSCQLLQGQL-GQLPRYESSRSPNFPDRSN
          : : : : : : : : : : : : : : : : : : : : : :
ALLERG LNPCKVFLQQCSPVAMPQSLARSQMLWQSSCHVMQQQCCQQLPRIPEQSRDAIRAIY
          170          180          190          200          210          220

```

SYHT0H IWQ

```

ALLERG SIVLQEQQHGGQFNQPQQQQPQQSVQGVSQPQQQKQLGQCSFQQPQQ
          230          240          250          260          270

```

>>ALLERGEN\_2012\_2|gi|736319|emb|CAA27052.1| glutenin [Tr (838 aa)  
 initn: 245 initl: 53 opt: 62 Z-score: 86.5 bits: 23.3 E(): 8.4  
 Smith-Waterman score: 72; 31.667% identity (32.759% ungapped) in 60 aa overlap (7-64:454-513)

```

          10          20          30
SYHT0H MRLWDTQPG--VHCAQLDTSMDGMMLVFGKGILSC
          : : : : : : : : : : : : : : : : : : : : : :
ALLERG PGYYPTSPLQSGQGQPGYYLTSPQQSGQGQPGQLQQSAQGQKGGQPGQGQPGQGQGGQ
          430          440          450          460          470          480

```

```

          40          50          60
SYHT0H QLLQGQLGQLPRYESSRSPNFPDRSNIWQ
      :  :::  ::  :   .  .  .  :  ::  :
ALLERG QPGQGQQGQQPGQGPGYYPTSPQQSGQGQQPGQWQQPGQGQPGYYPTSPLQPGQGQPGY
          490          500          510          520          530          540

```

```

>>ALLERGEN_2012_2|gi|170743|gb|AAB02788.1| HMW glutenin (815 aa)
  initn: 241 initl: 57 opt: 61 Z-score: 85.3 bits: 23.0 E(): 9.7
Smith-Waterman score: 61; 26.562% identity (29.310% ungapped) in 64 aa overlap (7-
64:609-672)

```

```

                      10          20          30
SYHT0H                      MRLWDTQPGVHCAQLDTS---MVDGMMLVFGKG---
                      :::  .  .  .  :   .  .  .  :  ::
ALLERG QQPGQGQQGQQPGEGQQGQQPGQGQQPGQGQPGYYPTSLQQSGQGQQPGQWQQPGQGQPGY
          580          590          600          610          620          630

```

```

          40          50          60
SYHT0H ILSCQLLQGQLGQLPRYESSRSPNFPDRSNIWQ
      .  ::  .  ::  :  .  ::  .  .  .  .  .  .  :
ALLERG YYPTSSLQPEQQGQQGYYPTSQQQPGQGPQPGQWQQSGQGQQGYYPTSPQQSGQGQQPGQW
          640          650          660          670          680          690

```

```

ALLERG LQPGQWLQSGYYLTSPQQLGQGQQPRQWLQPRQGQQGYYPTSPQQSGQGQQLGQGQQGY
          700          710          720          730          740          750

```

```

64 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.020 Display time: 0.010

```

```

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 35 aa
>SYHT0H2_I_37
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

```

```

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 3.8679+/-0.00302; mu= 3.9405+/- 0.158
mean_var=35.9280+/-10.523, 0's: 13 Z-trim: 13 B-trim: 0 in 0/43
Lambda= 0.2140

```

```

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.010

```

```

The best scores are:
ALLERGEN_2012_2|gi|3021324|emb|CAA06305.1| Aspfl ( 125) 51 22 2.1
ALLERGEN_2012_2|gi|9280360|gb|AAF86369.1| major a ( 150) 51 22 2.4
ALLERGEN_2012_2|gi|10189811|emb|CAC09234.1| unnam ( 215) 52 22 2.6
ALLERGEN_2012_2|gi|1352240|sp|P49273.1|ALL7_DERPT ( 215) 52 22 2.6
ALLERGEN_2012_2|gi|54039254|sp|P67875.1|RNMG_ASPF ( 176) 51 22 2.8
ALLERGEN_2012_2|gi|171464770|gb|ACB45874.1| patho ( 151) 47 21 5.7
ALLERGEN_2012_2|gi|46410859|gb|AAR98518.1| major ( 366) 50 22 6.2

```

```

>>ALLERGEN_2012_2|gi|3021324|emb|CAA06305.1| Aspfl aller (125 aa)

```

initn: 51 initl: 51 opt: 51 Z-score: 97.4 bits: 21.7 E(): 2.1  
 Smith-Waterman score: 51; 33.333% identity (33.333% ungapped) in 33 aa overlap (2-34:1-33)

```

          10          20          30
SYHT0H MRLTKVRSAAFNQYQSGPYHDGKTVQSVWLFLTnk
      ::  .. : .. . .:  ::::  .:  ....
ALLERG RLVYNQAKAESNSHHAPLSDGKTGSSYPHWFTNGYDGDGKLIKGRMPIKFGKADCDRPP
          10          20          30          40          50

ALLERG KHGKDGMGKDDHYLLEFPTFPDGHDKFDSSKKPKEDPGPARVIYTPNKKVFCGIVAHERG
      60          70          80          90         100         110

```

>>ALLERGEN\_2012\_2|gi|9280360|gb|AAF86369.1| major allerg (150 aa)  
 initn: 51 initl: 51 opt: 51 Z-score: 96.2 bits: 21.7 E(): 2.4  
 Smith-Waterman score: 51; 33.333% identity (33.333% ungapped) in 33 aa overlap (2-34:21-53)

```

          10          20          30
SYHT0H MRLTKVRSAAFNQYQSGPYHDGKTVQSVWLFLTnk
      ::  .. : .. . .:  ::::  .:  ....
ALLERG MTWTCINQQLNPKNKWKEDKRLLYNQAKAESNSHHAPLSDGKTGSSYAHWFTNGYDGNKG
          10          20          30          40          50          60

ALLERG LIKGRTPIKFGKADCDRPPKHSQNGMGKDDHYLLEFPTFPDGHDKFDSSKNKPKEDPGPA
      70          80          90         100         110         120

```

>>ALLERGEN\_2012\_2|gi|10189811|emb|CAC09234.1| unnamed pr (215 aa)  
 initn: 32 initl: 32 opt: 52 Z-score: 95.5 bits: 22.1 E(): 2.6  
 Smith-Waterman score: 52; 28.571% identity (28.571% ungapped) in 28 aa overlap (1-28:2-29)

```

          10          20          30
SYHT0H MRLTKVRSAAFNQYQSGPYHDGKTVQSVWLFLTnk
      ::  . .:::  .. : : : .. .
ALLERG MMKLLLIAAAAFVAVSADPIHYDKITEEINKAVDEAVAAIEKSETFDPMKVPDHSKDFER
          10          20          30          40          50          60

```

>>ALLERGEN\_2012\_2|gi|1352240|sp|P49273.1|ALL7\_DERPT RecN (215 aa)  
 initn: 32 initl: 32 opt: 52 Z-score: 95.5 bits: 22.1 E(): 2.6  
 Smith-Waterman score: 52; 28.571% identity (28.571% ungapped) in 28 aa overlap (1-28:2-29)

```

          10          20          30
SYHT0H MRLTKVRSAAFNQYQSGPYHDGKTVQSVWLFLTnk
      ::  . .:::  .. : : : .. .
ALLERG MMKLLLIAAAAFVAVSADPIHYDKITEEINKAVDEAVAAIEKSETFDPMKVPDHSKDFER
          10          20          30          40          50          60

```

>>ALLERGEN\_2012\_2|gi|54039254|sp|P67875.1|RNMG\_ASPFU Rec (176 aa)  
 initn: 51 initl: 51 opt: 51 Z-score: 95.1 bits: 21.8 E(): 2.8  
 Smith-Waterman score: 51; 33.333% identity (33.333% ungapped) in 33 aa overlap (2-34:48-80)

```

          10          20          30
SYHT0H MRLTKVRSAAFNQYQSGPYHDGKTVQSVWLf
      ::  .. : .. . .:  ::::  .:  .
ALLERG LAAPSPLDARATWTCINQQLNPKNKWKEDKRLLYSQAKAESNSHHAPLSDGKTGSSYPHW
      20          30          40          50          60          70

```

```

SYHT0H LTNK
.:
ALLERG FTNGYDGNKLIKGRTPIKFGKADCDRPPKHSQNGMGKDDHYLLEFPTFPDGHDKFDSK
      80      90      100      110      120      130

>>ALLERGEN_2012_2|gi|171464770|gb|ACB45874.1| pathogen-r (151 aa)
  initn: 47 initl: 47 opt: 47 Z-score: 89.5 bits: 20.5 E(): 5.7
Smith-Waterman score: 47; 50.000% identity (50.000% ungapped) in 18 aa overlap (8-
25:19-36)

```

```

      10      20      30
SYHT0H      MRLTKVRSAAFNQYQSGPYHDGKTVQSVWLFLTnk
      .:. : : : :
ALLERG MLPFSFAQDSIKDFVDAHNAARAQVGVGPVHWNKTVADYAHQYANKRIKDCNLVHSGKGPY
      10      20      30      40      50      60

ALLERG GENIAWGSRLAGTVAVRMWVSEKQFYNYDTNSCVRGKMGHGTQVVRNNSVRIGCAKVR
      70      80      90      100      110      120

```

```

>>ALLERGEN_2012_2|gi|46410859|gb|AAR98518.1| major latex (366 aa)
  initn: 50 initl: 50 opt: 50 Z-score: 88.8 bits: 21.6 E(): 6.2
Smith-Waterman score: 50; 46.667% identity (46.667% ungapped) in 15 aa overlap
(11-25:63-77)

```

```

      10      20      30
SYHT0H      MRLTKVRSAAFNQYQSGPYHDGKTVQSVWLFLTnk
      : .: .: . : :
ALLERG FGDGLYDAGNAKFIYPDKYLPSYHHPYGTTFDDYPTGRFSDGRTVVDFVAENVSLPRIPP
      40      50      60      70      80      90

ALLERG FKNKEANFTYGANFASEGATASDSNPLIDFRSQIRDFGELKLEWAVQLVNVTELARRLKK
      100      110      120      130      140      150

```

```

35 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.010 Display time: 0.010

```

```

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 183 aa
>SYHT0H2_I_38
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

```

```

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 4.0748+/-0.00364; mu= 12.2148+/- 0.190
mean_var=60.0372+/-15.922, 0's: 3 Z-trim: 4 B-trim: 0 in 0/44
Lambda= 0.1655

```

```

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.050
The best scores are:
ALLERGEN_2012_2|gi|5059162|gb|AAD38942.1|AF144060 ( 496) 87 29 0.25
ALLERGEN_2012_2|gi|21748153|emb|CAD38167.1| putat ( 124) 73 25 0.99

```

```
>>ALLERGEN_2012_2|gi|5059162|gb|AAD38942.1|AF144060_1 al (496 aa)
  initn: 50 init1: 50 opt: 87 Z-score: 113.9 bits: 29.1 E(): 0.25
Smith-Waterman score: 92; 26.364% identity (31.868% ungapped) in 110 aa overlap
(23-132:178-268)
```

```
>>ALLERGEN_2012_2|gi|21748153|emb|CAD38167.1| putative n (124 aa)
  initn: 70 initl: 70 opt: 73 Z-score: 103.1 bits: 25.1 E(): 0.99
Smith-Waterman score: 73; 33.33% identity (34.14% ungapped) in 42 aa overlap
(92-132:16-57)
```

```
>>ALLERGEN_2012_2|gi|89892723|gb|ABD79095.1| Zea m 1 all (252 aa)
  initn: 71 initl: 71 opt: 71 Z-score: 96.8 bits: 25.0 E(): 2.2
Smith-Waterman score: 71; 30.189% identity (30.189% ungapped) in 53 aa overlap
(100-152:81-133)
```

```

      130      140      150      160      170      180
SYHT0H NDPSVRLHEALGYTARGLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRPVTQI
      .: . : :. :: :. :
ALLERG YEPIAAYHFDLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKYDSKVTFFHLEKGC GPN
      120      130      140      150      160      170

>>ALLERGEN_2012_2|gi|89892721|gb|ABD79094.1| Zea m 1 all (263 aa)
  initn: 71 initl: 71 opt: 71 Z-score: 96.6 bits: 25.0 E(): 2.3
Smith-Waterman score: 71; 30.189% identity (30.189% ungapped) in 53 aa overlap
(100-152:92-144)

      70      80      90      100      110      120
SYHT0H GIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKS VVAVIGLP
      : : : . . : : . : : : .
ALLERG NGGGCGYKDVNKPFFNSMGACGNIPFIDGLGCGSCFEIKCDKPVCECSGKPVVVHITDMN
      70      80      90      100      110      120

      130      140      150      160      170      180
SYHT0H NDPSVRLHEALGYTARGLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRPVTQI
      .: . : :. :: :. :
ALLERG YEPIAAYHFDLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKFGSKVFSFHLEKGC GPN
      130      140      150      160      170      180

>>ALLERGEN_2012_2|gi|115502167|sp|Q1ZYQ8.2|EXB10_MAIZE R (270 aa)
  initn: 71 initl: 71 opt: 71 Z-score: 96.4 bits: 25.0 E(): 2.3
Smith-Waterman score: 71; 30.189% identity (30.189% ungapped) in 53 aa overlap
(100-152:99-151)

      70      80      90      100      110      120
SYHT0H GIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKS VVAVIGLP
      : : : . . : : . : : : .
ALLERG NGGGCGYKDVNKPFFNSMGACGNIPFIDGLGCGSCFEIKCDKPVCECSGKPVVVHITDMN
      70      80      90      100      110      120

      130      140      150      160      170      180
SYHT0H NDPSVRLHEALGYTARGLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRPVTQI
      .: . : :. :: :. :
ALLERG YEPIAAYHFDLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKYDSKVTFFHLEKGC GPN
      130      140      150      160      170      180

>>ALLERGEN_2012_2|gi|105969545|gb|ABF81662.1| EXPB10 [Ze (269 aa)
  initn: 94 initl: 69 opt: 69 Z-score: 93.9 bits: 24.5 E(): 3.2
Smith-Waterman score: 69; 30.189% identity (30.189% ungapped) in 53 aa overlap
(100-152:98-150)

      70      80      90      100      110      120
SYHT0H GIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKS VVAVIGLP
      : : : . . : : . : : : .
ALLERG NGGGCGYKDVNKPFFNSMGACGNVPIFIDGLGCGSCFEIKCDKPAECSGKPVVVYITDMN
      70      80      90      100      110      120

      130      140      150      160      170      180
SYHT0H NDPSVRLHEALGYTARGLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRPVTQI
      .: . : :. :: :. :
ALLERG YEPIAAYHFDLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKYGSKVTFFHLEKGC NPN
      130      140      150      160      170      180

>>ALLERGEN_2012_2|gi|115502168|sp|P0C1Y5.1|EXB11_MAIZE R (269 aa)
  initn: 94 initl: 69 opt: 69 Z-score: 93.9 bits: 24.5 E(): 3.2
Smith-Waterman score: 69; 30.189% identity (30.189% ungapped) in 53 aa overlap
(100-152:98-150)

```

```

      70      80      90      100      110      120
SYHT0H GIAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKSVVAVIGLP
      :: :: . . : : .: :: . .
ALLERG NGGGCGYKDVNKAPFNSMGACGNVPIFKDGLGCGSCFEIKCDKPAECSGKPVVVYITDMN
      70      80      90      100      110      120

      130      140      150      160      170      180
SYHT0H NDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRPVTQI
      .: . : :. :: :. :
ALLERG YEPiAAyHFDLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKYGSKVTFHLEKGCNPN
      130      140      150      160      170      180

>>ALLERGEN_2012_2|gi|18615|emb|CAA26723.1| unnamed prote (495 aa)
  initn: 41 init1: 41 opt: 71 Z-score: 93.2 bits: 25.3 E(): 3.5
Smith-Waterman score: 71; 25.843% identity (29.114% ungapped) in 89 aa overlap
(39-118:113-200)

```

```

      10      20      30      40      50      60
SYHT0H EIRPATAADMAAVCDIVNHYIETSTVNFRTPEQTPQEWIDDLERLQDRYPWLVAEEVEGVV
      ::: ::. . . : ::. . : : .
ALLERG SYTNGPQEIYIQQGKGIFGMIYPGCSSTFEFPQQPQQRGQS-SRPQDRHQKIYNSREGDL
      90      100      110      120      130      140

```

```

      70      80      90      100      110
SYHT0H AGIAYAGPWKARNAYDWTVESTVYV---SHRHQ-----RLGLGSTLYTHLLKSMEAQGF
      .. . : : : : .. . : :: : : : : : : : : : : : : : : : : : : : :
ALLERG IAVPTGVAWWMYNNEDTPVVAVSIIDTNSLENQLDQMPRRFYLAGNQEQEFLKYQQEQGG
      150      160      170      180      190      200

```

```

      120      130      140      150      160      170
SYHT0H KSVVAVIGLPNDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRP
ALLERG HQSQKGKHQQEEENEKGSILSGFTLEFLEHAFSVDKQIAKNLQGENEGEDKGAIVTVKGG
      210      220      230      240      250      260

```

```

>>ALLERGEN_2012_2|gi|18635|emb|CAA33215.1| glycinin subu (495 aa)
  initn: 41 init1: 41 opt: 69 Z-score: 90.7 bits: 24.8 E(): 4.9
Smith-Waterman score: 69; 25.843% identity (29.114% ungapped) in 89 aa overlap
(39-118:113-200)

```

```

      10      20      30      40      50      60
SYHT0H EIRPATAADMAAVCDIVNHYIETSTVNFRTPEQTPQEWIDDLERLQDRYPWLVAEEVEGVV
      ::: ::. . . : ::. . : : .
ALLERG SYTNGPQEIYIQQGKGIFGMIYPGCPSTFEFPQQPQQRGQS-SRPQDRHQKIYNFREGDL
      90      100      110      120      130      140

```

```

      70      80      90      100      110
SYHT0H AGIAYAGPWKARNAYDWTVESTVYV---SHRHQ-----RLGLGSTLYTHLLKSMEAQGF
      .. . : : : : .. . : :: : : : : : : : : : : : : : : : : : : : :
ALLERG IAVPTGVAWWMYNNEDTPVVAVSIIDTNSLENQLDQMPRRFYLAGNQEQEFLKYQQEQGG
      150      160      170      180      190      200

```

```

      120      130      140      150      160      170
SYHT0H KSVVAVIGLPNDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFQWQDFELPAPPRPVRP
ALLERG HQSQKGKHQQEEENEKGSILSGFTLEFLEHAFSVDKQIAKNLQGENEGEDKGAIVTVKGG
      210      220      230      240      250      260

```

```

>>ALLERGEN_2012_2|gi|18652047|gb|AAL76932.1|AF456481_1 m (154 aa)
  initn: 33 init1: 33 opt: 64 Z-score: 90.3 bits: 23.1 E(): 5.1
Smith-Waterman score: 64; 22.727% identity (25.773% ungapped) in 110 aa overlap
(44-146:21-124)

```

```

      20      30      40      50      60      70
SYHT0H TAADMAAVCDIVNHYIETSTVNFRTEPQTPQEWIDDLERLQDRY-PWLVAEVEGVVAGIA
      : .. :.. . . : :. :. :. :.
ALLERG      MGVSQKTEVEAPSTVSAEKMYQGFLDMDTVFPKVLPLQLIKSVE-ILEGDG
      10      20      30      40

      80      90      100      110      120
SYHT0H YAGPWKARNAYDWTVESTVYVSHRHQ-----RLGLGSTLYTHLLKSMEAQGFKSVVA-VI
      .: . . . :. :.. . . . :. : : :. . . :. :. :. :.
ALLERG GVGTVRLVHLG---EATEYTTMKQKVDVIDKAGLGYT-YTTIGGDILVEGLESVVNQFV
      50      60      70      80      90      100

      130      140      150      160      170      180
SYHT0H GLPNDPVSRVLHEALGYTARGLTRAAGYKHGGWHDVGFWRQDFELPAPPRPVRPVTQI
      .: : . . . . :. :. :
ALLERG VVPTDGGCIVKNTTIYNTKGDVLPEDKVKTEATEKSALAFKAVEAYLLAN
      110      120      130      140      150

```

```

>>ALLERGEN_2012_2|gi|2266625|emb|CAB10765.1| group V all (264 aa)
  initn: 34 initl: 34 opt: 63 Z-score: 86.2 bits: 23.1 E(): 8.6
Smith-Waterman score: 63; 28.814% identity (30.357% ungapped) in 59 aa overlap
(101-157:17-74)

```

```

      80      90      100      110      120
SYHT0H IAYAGPWKARNAYDWTVESTVYVSHRHQRLGLGSTLYTHLLKSMEAQGFKSVVAVIGL--
      : ..: .:..... : :. :. :.
ALLERG      ADAGYTPAAPAAAGAGGKATTDEQKLEDVNA-GFKTAVAAAAANVP
      10      20      30      40

      130      140      150      160      170      180
SYHT0H PNDPSVRLHEALGYTARGLTRAAGYKHGGWHDVGFWRQDFELPAPPRPVRPVTQI
      : : . . . . :. : : :
ALLERG PADKYKTFEAAFTASSKASIAAAATKAPGLIPQLNAATNTAYAAAQGATPEAKYDAFVTT
      50      60      70      80      90      100

```

```

183 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.050 Display time: 0.010

```

```

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 31 aa
>SYHT0H2_I_39
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

```

```

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 4.2101+/-0.00274; mu= -0.3038+/- 0.144
mean_var=23.8043+/- 5.508, 0's: 19 Z-trim: 19 B-trim: 0 in 0/42
Lambda= 0.2629

```

```

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.010
The best scores are:
opt bits E(1603)

```



ALLERGEN_2012_2	gi 1364213 emb CAA44344.1	fel d	( 92)	44	22	1.2
ALLERGEN_2012_2	gi 5777414 emb CAB53458.1	MnSOD	( 205)	42	21	4.8
ALLERGEN_2012_2	gi 10862818 emb CAC13961.1	IgE-b	( 205)	42	21	4.8
ALLERGEN_2012_2	gi 212279 gb AAA48944.1	lysozyme	( 24)	32	18	6.2
ALLERGEN_2012_2	gi 1168402 sp P42058.1	ALTA7_ALTA	( 204)	40	20	8.1
ALLERGEN_2012_2	gi 544619 gb AAB29345.1	36 kda a	( 25)	31	17	8.5
ALLERGEN_2012_2	gi 21413 emb CAA45723.1	aspartic	( 217)	40	20	8.7
ALLERGEN_2012_2	gi 20141714 sp P30941.2	SPI7_SOLT	( 221)	40	20	8.9

>>ALLERGEN\_2012\_2|gi|1364213|emb|CAA44344.1| fel d I cha (92 aa)  
 initn: 34 init1: 34 opt: 44 Z-score: 101.8 bits: 21.9 E(): 1.2  
 Smith-Waterman score: 44; 40.000% identity (40.000% ungapped) in 20 aa overlap (4-23:1-20)

	10	20	30
SYHT0H	MVFEVGLDLRSSLGSMNRLVQWQKNSFTVVG		
	: :: : : : . . . .		
ALLERG	EGGLCSRASLGCLALDLGWKDCEICPAVKRDVDLFLTGTTPDEYVEQVAQYNALPVVL		
	10	20	30
		40	50

>>ALLERGEN\_2012\_2|gi|5777414|emb|CAB53458.1| MnSOD [Heve (205 aa)  
 initn: 35 init1: 35 opt: 42 Z-score: 90.8 bits: 21.0 E(): 4.8  
 Smith-Waterman score: 42; 29.412% identity (29.412% ungapped) in 17 aa overlap (5-21:97-113)

	10	20	30
SYHT0H	MVFEVGLDLRSSLGSMNRLVQWQKNSFTVVG		
	. : . . . . : : : : .		
ALLERG	FNGGGHVNHISIFWKNLAPVREGGGELPHGSLGWAIDADFGSLEKLIQLMNAEGVALQGSG		
	70	80	90
	100	110	120
ALLERG	WWVLALDKELKKLVVETTANQDPLVTKGPTLVPLLGIDVWEHAYYLQYKNVRPDYLKNIW		
	130	140	150
	160	170	180

>>ALLERGEN\_2012\_2|gi|10862818|emb|CAC13961.1| IgE-bindin (205 aa)  
 initn: 35 init1: 35 opt: 42 Z-score: 90.8 bits: 21.0 E(): 4.8  
 Smith-Waterman score: 42; 29.412% identity (29.412% ungapped) in 17 aa overlap (5-21:97-113)

	10	20	30
SYHT0H	MVFEVGLDLRSSLGSMNRLVQWQKNSFTVVG		
	. : . . . . : : : : .		
ALLERG	FNGGGHVNHISIFWKNLAPVREGGGELPHGSLGWAIDADFGSLEKLIQLMNAEGAALRGSG		
	70	80	90
	100	110	120
ALLERG	WWVLALDKELKKLVVETTANQDPLVTKGPTLVPLLGIDVWEHAYYLQYKNVRPDYLKNIW		
	130	140	150
	160	170	180

>>ALLERGEN\_2012\_2|gi|212279|gb|AAA48944.1| lysozyme prot (24 aa)  
 initn: 31 init1: 31 opt: 32 Z-score: 88.8 bits: 17.5 E(): 6.2  
 Smith-Waterman score: 32; 54.545% identity (60.000% ungapped) in 11 aa overlap (16-26:1-10)

	10	20	30
SYHT0H	MVFEVGLDLRSSLGSMNRLVQWQKNSFTVVG		
	: : : : : : : :		
ALLERG	MNAWVAW-RNSCKGTDVQAWIRGCR		
	10	20	

>>ALLERGEN\_2012\_2|gi|1168402|sp|P42058.1|ALTA7\_ALTAL Rec (204 aa)  
 initn: 40 init1: 40 opt: 40 Z-score: 86.7 bits: 20.3 E(): 8.1  
 Smith-Waterman score: 40; 57.143% identity (57.143% ungapped) in 7 aa overlap (21-27:100-106)

```

              10      20      30
SYHT0H      MVFEVGLDLRSSLGSMNRLVQWQKNSFTTVVG
              :::::
ALLERG EEFDGILFGIPTRYGNFPAQFKTFWDKTGKQWQGAFWGKYAGVVFVSTGTLGGGQETTAI
      70      80      90      100      110      120

ALLERG TSMSTLVHDGFIYVPLGYKTAFSMLANLDEVHGGSPWGAGTFSAGDGSRQPSELELNIAQ
      130      140      150      160      170      180

```

```

>>ALLERGEN_2012_2|gi|544619|gb|AAB29345.1| 36 kda allerg (25 aa)
  initn: 28 init1: 28 opt: 31 Z-score: 86.4 bits: 17.2 E(): 8.5
Smith-Waterman score: 31; 50.000% identity (50.000% ungapped) in 8 aa overlap (1-
8:18-25)

```

```

              10      20      30
SYHT0H      MVFEVGLDLRSSLGSMNRLVQWQKNSFTTVVG
              .::.:
ALLERG GLNICQIDCNKIQSLPXLVFTIGGD
      10      20

```

```

>>ALLERGEN_2012_2|gi|21413|emb|CAA45723.1| aspartic prot (217 aa)
  initn: 27 init1: 27 opt: 40 Z-score: 86.2 bits: 20.2 E(): 8.7
Smith-Waterman score: 40; 46.667% identity (50.000% ungapped) in 15 aa overlap (2-
16:126-139)

```

```

              10      20      30
SYHT0H      MVFEVGLDLRSSLGSMNRLVQWQKNSFTTVVG
              ...::.:
ALLERG FSHFGQGIFENELLNIQFAISTSKLCVSYTIWKVG-DYDASLGTMLLETGGTIGQADSSW
      100      110      120      130      140      150

ALLERG FKIVKSSQFGYNLLYCPVTSTMSCPFSSDDQFCLKVGVVHQNGKRRRLALVKDNPLDVSKF
      160      170      180      190      200      210

```

```

>>ALLERGEN_2012_2|gi|20141714|sp|P30941.2|SPI7_SOLTU Rec (221 aa)
  initn: 27 init1: 27 opt: 40 Z-score: 86.0 bits: 20.2 E(): 8.9
Smith-Waterman score: 40; 46.667% identity (50.000% ungapped) in 15 aa overlap (2-
16:130-143)

```

```

              10      20      30
SYHT0H      MVFEVGLDLRSSLGSMNRLVQWQKNSFTTVVG
              ...::.:
ALLERG SSHFGQGIFENELLNIQFAISTSKLCVSYTIWKVG-DYDASLGTMLLETGGTIGQADSSW
      100      110      120      130      140      150

ALLERG FKIVKSSQFGYNLLYCPVTSTMSCPFSSDDQFCLKVGVVHQNGKRRRLALVKDNPLDVSKF
      160      170      180      190      200      210

```

```

31 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.010 Display time: 0.000

```

```

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 42 aa
>SYHT0H2_I_40
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 4.1005+/-0.00278; mu= 2.2805+/- 0.151
mean_var=35.1069+/- 9.868, 0's: 11 Z-trim: 11 B-trim: 236 in 1/41
Lambda= 0.2165

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.020
The best scores are:
                                opt bits E(1603)
ALLERGEN_2012_2|gi|38326693|gb|AAR17475.1| unknow ( 228) 56 24 0.99
ALLERGEN_2012_2|gi|160285626|pdb|2JMH|A Chain A, ( 119) 49 22 2.5
ALLERGEN_2012_2|gi|111120450|gb|ABH06359.1| Blo t ( 134) 49 22 2.8
ALLERGEN_2012_2|gi|4204917|gb|AAD10850.1| major I ( 134) 49 22 2.8
ALLERGEN_2012_2|gi|111120436|gb|ABH06352.1| Blo t ( 134) 49 22 2.8
ALLERGEN_2012_2|gi|9954251|gb|AAG08988.1|AF216519 ( 284) 48 21 6.8
ALLERGEN_2012_2|gi|633938|gb|AAB30434.1| albumin ( 265) 47 21 8
ALLERGEN_2012_2|gi|85687540|gb|ABC73706.1| allerg ( 140) 44 20 8.7

>>ALLERGEN_2012_2|gi|38326693|gb|AAR17475.1| unknown [Pe (228 aa)
initn: 40 initl: 40 opt: 56 Z-score: 103.1 bits: 23.9 E(): 0.99
Smith-Waterman score: 56; 37.500% identity (41.379% ungapped) in 32 aa overlap (8-
36:182-213)

                                10      20      30
SYHT0H                      MIIIRPATGFNLKKLYCQMF---ERSSYEELLALLS
                          :.....: : : : : : : : : :
ALLERG WDETNNLEELEANVRAIEMDGLVWGASKFVAVGFGIKKLQINLVVEDEKVDSTDELQAQIE
                        160      170      180      190      200      210

                        40
SYHT0H NDRLLDFQ
                          : :
ALLERG EDEDHVQSTDVAAMQKL
                        220

>>ALLERGEN_2012_2|gi|160285626|pdb|2JMH|A Chain A, Nmr S (119 aa)
initn: 49 initl: 49 opt: 49 Z-score: 95.8 bits: 21.6 E(): 2.5
Smith-Waterman score: 49; 37.500% identity (37.500% ungapped) in 32 aa overlap (1-
32:67-98)

                                10      20      30
SYHT0H                      MIIIRPATGFNLKKLYCQMFERSSEYELL
                          : : : : : : : : : : : :
ALLERG YLQHQDELNENKSKELQEKIIRELDVVCAMIEGAQGALERELKRTDLNILERFNYYEEAQ
                        40      50      60      70      80      90

                        40
SYHT0H ALLSNDRLDFQ
                          : :
ALLERG TLSKILLKDLKETEQKVVDIQTQ
                        100      110

>>ALLERGEN_2012_2|gi|111120450|gb|ABH06359.1| Blo t 5 al (134 aa)
initn: 49 initl: 49 opt: 49 Z-score: 95.0 bits: 21.6 E(): 2.8
Smith-Waterman score: 49; 37.500% identity (37.500% ungapped) in 32 aa overlap (1-
32:82-113)

                                10      20      30

```

```

SYHT0H                                     MIIARPATGFNLKKLYCQMFERS10SYEELL20
                                     ::  ::  :  .::  .:::  :::
ALLERG YLQHQLDELNENKSKELQEKIIRELDVVCAMIEGAQ60GALERELKRTDLNILERFN70YEEAQ80
                                     90      100      110

      40
SYHT0H ALLSNDRL40LD40FQ
      .:
ALLERG T120LSKILLKDLKETE130QKV130KDIQTQ

>>ALLERGEN_2012_2|gi|4204917|gb|AAD10850.1| major IgE-bi (134 aa)
  initn: 49 init1: 49 opt: 49 Z-score: 95.0 bits: 21.6 E(): 2.8
Smith-Waterman score: 49; 37.500% identity (37.500% ungapped) in 32 aa overlap (1-32:82-113)

      10      20      30
SYHT0H                                     MIIARPATGFNLKKLYCQMFERS10SYEELL20
                                     ::  ::  :  .::  .:::  :::
ALLERG YLQHQLDELNENKSKELQEKIIRELDVVCAMIEGAQ60GALERELKRTDLNILERFN70YEEAQ80
                                     90      100      110

      40
SYHT0H ALLSNDRL40LD40FQ
      .:
ALLERG T120LSKILLKDLKETE130QKV130KDIQTQ

>>ALLERGEN_2012_2|gi|111120436|gb|ABH06352.1| Blo t 5 al (134 aa)
  initn: 49 init1: 49 opt: 49 Z-score: 95.0 bits: 21.6 E(): 2.8
Smith-Waterman score: 49; 37.500% identity (37.500% ungapped) in 32 aa overlap (1-32:82-113)

      10      20      30
SYHT0H                                     MIIARPATGFNLKKLYCQMFERS10SYEELL20
                                     ::  ::  :  .::  .:::  :::
ALLERG YLQHQLDELNENKSKELQEKIIRELDVVCAMIEGAQ60GALERELKRTDLNILERFN70YEEAQ80
                                     90      100      110

      40
SYHT0H ALLSNDRL40LD40FQ
      .:
ALLERG T120LSKILLKDLKETE130QKV130KDIQTQ

>>ALLERGEN_2012_2|gi|9954251|gb|AAG08988.1|AF216519_1 tr (284 aa)
  initn: 26 init1: 26 opt: 48 Z-score: 88.1 bits: 21.4 E(): 6.8
Smith-Waterman score: 48; 42.857% identity (50.000% ungapped) in 35 aa overlap (10-40:201-234)

      10      20      30
SYHT0H                                     MIIARPATGFNLKKLYCQMFERS10----SYEELL20LALLSN30
                                     :  ::  :  :  .  :  :::  .  :::
ALLERG ITEVDLERAEARLEAAEAKVIDLEEQLTVVGANIKTLQVQNDQASQREDSYEETIRD180LTN190
                                     200      210      220      230

      40
SYHT0H DRL40LD40FQ
      :: :
ALLERG -RLKDAENRATEAERTVSKLQKEVDRL240EDEL250LTEKEKYKAISDEL260DATFAELAGY270
                                     280

>>ALLERGEN_2012_2|gi|633938|gb|AAB30434.1| albumin [Cani (265 aa)

```

initn: 47 initl: 47 opt: 47 Z-score: 86.9 bits: 21.1 E(): 8  
 Smith-Waterman score: 47; 50.000% identity (50.000% ungapped) in 12 aa overlap  
 (12-23:196-207)

```

              10      20      30      40
SYHT0H      MIIIRPATGFNLKKLYCQMFERSSEYELLALLSNDRLLD
              :: : :...:.
ALLERG TLEKCCATDDPPTCYAKVLDEFKPLVDEPQNLVKTNCELFEKLGEYGFQNALLVRYTKKA
              170      180      190      200      210      220

```

SYHT0H Q

```

ALLERG PQVSTPTLVVEVSRKLGKVGTKCCKKPESERMSCADDFLS
              230      240      250      260

```

>>ALLERGEN\_2012\_2|gi|85687540|gb|ABC73706.1| allergen pr (140 aa)  
 initn: 41 initl: 41 opt: 44 Z-score: 86.2 bits: 20.1 E(): 8.7  
 Smith-Waterman score: 44; 50.000% identity (50.000% ungapped) in 14 aa overlap  
 (14-27:98-111)

```

              10      20      30      40
SYHT0H      MIIIRPATGFNLKKLYCQMFERSSEYELLALLSNDRLLD
              :. .:....:..
ALLERG SKELKEQILRELTIGENFMKGALKFFEMEAKRTDLNMFERYNYEFALESIKLLIKKLDEL
              70      80      90      100      110      120

```

```

ALLERG AKKVKA VNPDEYY
              130      140

```

42 residues in 1 query sequences  
 367796 residues in 1603 library sequences  
 Scomplib [34t11]  
 start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012  
 Scan time: 0.020 Display time: 0.000

Function used was FASTA [version 3.4t11 Apr 17, 2002]  
 FASTA searches a protein or DNA sequence data bank  
 version 3.4t11 Apr 17, 2002  
 Please cite:  
 W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 35 aa
>SYHT0H2_I_41
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

```

367796 residues in 1603 sequences  
 Expectation\_n fit: rho(ln(x))= 2.9169+/-0.00319; mu= 9.8082+/- 0.166  
 mean\_var=28.8941+/- 7.269, 0's: 14 Z-trim: 14 B-trim: 23 in 1/41  
 Lambda= 0.2386

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2  
 join: 36, opt: 24, gap-pen: -12/-2, width: 16  
 Scan time: 0.010

The best scores are:

	gi	22135348	gb	AAM93157.1	trypsi	( 219)	49	22	3.3
ALLERGEN_2012_2	gi	3703107	gb	AAC63045.1	glycini	( 507)	49	22	6
ALLERGEN_2012_2	gi	224036293	pdb	3C3V A Chain A,		( 510)	49	22	6
ALLERGEN_2012_2	gi	5712199	gb	AAD47382.1	glycini	( 530)	49	22	6.2
ALLERGEN_2012_2	gi	199732457	gb	ACH91862.1	arach	( 530)	49	22	6.2

ALLERGEN\_2012\_2|gi|21314465|gb|AAM46958.1|AF51085 ( 538) 49 22 6.2

>>ALLERGEN\_2012\_2|gi|22135348|gb|AAM93157.1| trypsin inh (219 aa)  
initn: 34 init1: 34 opt: 49 Z-score: 93.7 bits: 21.8 E(): 3.3  
Smith-Waterman score: 49; 40.000% identity (40.000% ungapped) in 15 aa overlap  
(20-34:91-105)

```

              10      20      30
SYHT0H      MVDPHLSATNLICSSEKAKPTAQFFHRGRDHFID
              :      :...: :...
ALLERG NNQEFECAGVALSRLVLRNALRRPFYSNAPQEIFIQQGRGYFGLIFPGCPSTYEPAQQ
              70      80      90      100     110     120
```

```

ALLERG GRRHQSQRP RRFGQDQSQQQDSDHQVHRFDEGDLIAVPTGVAFWV MYNDHDTDVVAVS
              130     140     150     160     170     180
```

>>ALLERGEN\_2012\_2|gi|3703107|gb|AAC63045.1| glycinin [Ar (507 aa)  
initn: 34 init1: 34 opt: 49 Z-score: 89.1 bits: 22.2 E(): 6  
Smith-Waterman score: 49; 40.000% identity (40.000% ungapped) in 15 aa overlap  
(20-34:66-80)

```

              10      20      30
SYHT0H      MVDPHLSATNLICSSEKAKPTAQFFHRGRDHFID
              :      :...: :...
ALLERG NNQEFECAGVALSRLVLRNALRRPFYSNAPQEIFIQQGRGYFGLIFPGCPRH YEEPHTQ
              40      50      60      70      80      90
```

```

ALLERG GRRSQSQRP RRRLQGEDQSQQQDSDHQVHRFDEGDLIAVPTGVAFWLYNDHDTDVVAVS
              100     110     120     130     140     150
```

>>ALLERGEN\_2012\_2|gi|224036293|pdb|3C3V|A Chain A, Cryst (510 aa)  
initn: 34 init1: 34 opt: 49 Z-score: 89.1 bits: 22.2 E(): 6  
Smith-Waterman score: 49; 40.000% identity (40.000% ungapped) in 15 aa overlap  
(20-34:69-83)

```

              10      20      30
SYHT0H      MVDPHLSATNLICSSEKAKPTAQFFHRGRDHFID
              :      :...: :...
ALLERG NNQEFECAGVALSRLVLRNALRRPFYSNAPQEIFIQQGRGYFGLIFPGCPSTYEPAQQ
              40      50      60      70      80      90
```

```

ALLERG GRRYQSQRP RRRLQEEDQSQQQDSDHQVHRFNEGDLIAVPTGVAFWLYNDHDTDVVAVS
              100     110     120     130     140     150
```

>>ALLERGEN\_2012\_2|gi|5712199|gb|AAD47382.1| glycinin [Ar (530 aa)  
initn: 34 init1: 34 opt: 49 Z-score: 88.9 bits: 22.2 E(): 6.2  
Smith-Waterman score: 49; 40.000% identity (40.000% ungapped) in 15 aa overlap  
(20-34:89-103)

```

              10      20      30
SYHT0H      MVDPHLSATNLICSSEKAKPTAQFFHRGRDHFID
              :      :...: :...
ALLERG NNQEFECAGVALSRLVLRNALRRPFYSNAPQEIFIQQGRGYFGLIFPGCPSTYEPAQQ
              60      70      80      90      100     110
```

```

ALLERG GRRYQSQRP RRRLQEEDQSQQQDSDHQVHRFNEGDLIAVPTGVAFWLYNDHDTDVVAVS
              120     130     140     150     160     170
```

>>ALLERGEN\_2012\_2|gi|199732457|gb|ACH91862.1| arachin Ar (530 aa)  
initn: 34 init1: 34 opt: 49 Z-score: 88.9 bits: 22.2 E(): 6.2  
Smith-Waterman score: 49; 40.000% identity (40.000% ungapped) in 15 aa overlap  
(20-34:89-103)

```

              10      20      30
SYHT0H      MVDPHLSATNLICSSSEKAKPTAQFFHRGRDHFID
              :      :...: :...
ALLERG NNQEFECAGVALSRLVLRNALRRPFYSNAPQEIFIQQGRGYFGLIFPGCPSTYEPAQQ
        60      70      80      90      100     110

ALLERG GRRYQSQRPRLQEEDQSQQQDSHQKVRHFNEGDLIAVPTGVAFWLYNDHDTDVVAVS
        120     130     140     150     160     170

>>ALLERGEN_2012_2|gi|21314465|gb|AAM46958.1|AF510854_1 a (538 aa)
  initn: 34 initl: 34 opt: 49 Z-score: 88.8 bits: 22.2 E(): 6.2
Smith-Waterman score: 49; 40.000% identity (40.000% ungapped) in 15 aa overlap
(20-34:89-103)

```

```

              10      20      30
SYHT0H      MVDPHLSATNLICSSSEKAKPTAQFFHRGRDHFID
              :      :...: :...
ALLERG NNQEFECAGVALSRLVLRNALRRPFYSNAPQEIFIQQGRGYFGLIFPGCPSTYEPAQQ
        60      70      80      90      100     110

ALLERG GRRHQSQRPPRRFQGDQSQQQDSHQKVRHFDEGDLIAVPTGVAFWLYNDHDTDVVAVS
        120     130     140     150     160     170

```

```

35 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.010 Display time: 0.000

```

```

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 79 aa
>SYHT0H2_I_42
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

```

```

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 2.8002+/-0.00316; mu= 11.3231+/- 0.164
mean_var=37.6476+/- 8.777, 0's: 4 Z-trim: 4 B-trim: 0 in 0/43
Lambda= 0.2090

```

```

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.020

```

```

The best scores are:
ALLERGEN_2012_2|gi|16580747|dbj|BAB71741.1| glyox ( 291) 51 22 6.2
ALLERGEN_2012_2|gi|84029333|sp|Q948T6.2|LGUL_ORYS ( 291) 51 22 6.2
ALLERGEN_2012_2|gi|74664773|sp|Q96X46.3|ENO_PENCI ( 438) 52 23 6.4
ALLERGEN_2012_2|gi|127533|sp|P02761.1|MUP_RAT Rec ( 181) 49 22 7.2
ALLERGEN_2012_2|gi|323575367|dbj|BAJ78223.1| Ani ( 295) 50 22 7.7
ALLERGEN_2012_2|gi|11991229|gb|AAG42255.1|AF30670 ( 296) 50 22 7.8
ALLERGEN_2012_2|gi|510515|emb|CAA56343.1| Kunitz ( 208) 49 22 7.8
ALLERGEN_2012_2|gi|11991227|gb|AAG42254.1|AF30670 ( 303) 50 22 7.9

```

```

>>ALLERGEN_2012_2|gi|16580747|dbj|BAB71741.1| glyoxalase (291 aa)
  initn: 48 initl: 48 opt: 51 Z-score: 88.8 bits: 22.5 E(): 6.2

```

Smith-Waterman score: 51; 35.714% identity (40.000% ungapped) in 28 aa overlap (26:146-173)

```

                                10      20
SYHT0H                      MFLQTTPTKY--QVLFKQNDTDLNIGFYTN
                                ..:  ::  :...  .:  :  .:  ::
ALLERG CCKITREPGVPVKGGSSTVIAFAQDPDGYMFELIQRGPTPEPLCQVMLRVGDLDRSIKFYEK
                                120      130      140      150      160      170

                                30      40      50      60      70
SYHT0H HMIFINSVILYGRWRQKSSRIIFEMRKLSSYLFSSSWLFFFLATCLNTSPP

ALLERG ALGMKLLRKKDVPDYKYTIAMLGYADEDKTTVIELTYNYGVTEYTKGNAYAQVAIGTEDV
                                180      190      200      210      220      230
```

>>ALLERGEN\_2012\_2|gi|84029333|sp|Q948T6.2|LGUL\_ORYSJ Rec (291 aa)  
initn: 48 initl: 48 opt: 51 Z-score: 88.8 bits: 22.5 E(): 6.2  
Smith-Waterman score: 51; 35.714% identity (40.000% ungapped) in 28 aa overlap (26:146-173)

```

                                10      20
SYHT0H                      MFLQTTPTKY--QVLFKQNDTDLNIGFYTN
                                ..:  ::  :...  .:  :  .:  ::
ALLERG CCKITREPGVPVKGGSSTVIAFAQDPDGYMFELIQRGPTPEPLCQVMLRVGDLDRSIKFYEK
                                120      130      140      150      160      170

                                30      40      50      60      70
SYHT0H HMIFINSVILYGRWRQKSSRIIFEMRKLSSYLFSSSWLFFFLATCLNTSPP

ALLERG ALGMKLLRKKDVPDYKYTIAMLGYADEDKTTVIELTYNYGVTEYTKGNAYAQVAIGTEDV
                                180      190      200      210      220      230
```

>>ALLERGEN\_2012\_2|gi|74664773|sp|Q96X46.3|ENO\_PENCI RecN (438 aa)  
initn: 41 initl: 41 opt: 52 Z-score: 88.5 bits: 23.0 E(): 6.4  
Smith-Waterman score: 52; 42.105% identity (42.105% ungapped) in 19 aa overlap (20:252-270)

```

                                10      20      30
SYHT0H                      MFLQTTPTKYQVLFKQNDTDLNIGFYTNHMI
                                :  .:  :...  :...  .:  :
ALLERG EEALDLITEAIEQAGYTGKISIAMDVASSEFYKTDAKKYDLDFKNPDSPTKWLTYEQLA
                                230      240      250      260      270      280

                                40      50      60      70
SYHT0H FINSVILYGRWRQKSSRIIFEMRKLSSYLFSSSWLFFFLATCLNTSPP

ALLERG DLYKSAAKYPIVSIEDPFAEDDWEAWSYFYKTSDFQIVGDDLTVTNPLRIKKAIELKSC
                                290      300      310      320      330      340
```

>>ALLERGEN\_2012\_2|gi|127533|sp|P02761.1|MUP\_RAT RecName: (181 aa)  
initn: 47 initl: 47 opt: 49 Z-score: 87.7 bits: 21.6 E(): 7.2  
Smith-Waterman score: 49; 34.043% identity (42.105% ungapped) in 47 aa overlap (16-56:113-156)

```

                                10      20      30
SYHT0H                      MFLQTTPTKYQVLFKQNDTDLNIGFYTNHMI-FINS-----VILY
                                ..:  :  .:  :...  :...  :...  :
ALLERG CRELYLVAYKTPEDGEYFVEYDGGNTFTILKTDYDRYVMF---HLINFKNGETFQLMVLY
                                90      100      110      120      130
```



```

      40      50      60      70
SYHT0H GRWRQKSSRIIFEMRKLSSYLFSSSWLFFFLATCLNTSPP
      :: .. :: : .. ::
ALLERG GRTKDLSSDIKEKFAKLCEAHGITRDNIIDLTKTDRCLQARG
      140      150      160      170      180

>>ALLERGEN_2012_2|gi|323575367|dbj|BAJ78223.1| Ani s 12 (295 aa)
  initn: 44 initl: 44 opt: 50 Z-score: 87.1 bits: 22.2 E(): 7.7
Smith-Waterman score: 50; 37.037% identity (38.462% ungapped) in 27 aa overlap
(22-48:198-223)

      10      20      30      40      50
SYHT0H      MFLQTTPTKYQVLFKQNDTDLNIGFYTNHMIFINSVILYGRWRQKSSRIIF
      : :.. :.... :.... ::
ALLERG ECVDALGTPPVTTAANGAYQMAAPLHRCIENGWWMKMCSTWINATICE-RWKQECSDRKD
      170      180      190      200      210      220

      60      70
SYHT0H EMRKLSSYLFSSSWLFFFLATCLNTSPP

ALLERG AEPPTNFSQCIQTQTVMLQCKLEFGDKCKALQEECVAAATYAPTAYVDANPPIFTSETIRC
      230      240      250      260      270      280

>>ALLERGEN_2012_2|gi|11991229|gb|AAG42255.1|AF306708_1 p (296 aa)
  initn: 37 initl: 37 opt: 50 Z-score: 87.1 bits: 22.2 E(): 7.8
Smith-Waterman score: 50; 40.909% identity (50.000% ungapped) in 22 aa overlap (4-
25:44-61)

      10      20      30
SYHT0H      MFLQTTPTKYQVLFKQNDTDLNIGFYTNHMIFI
      . :.... : :.. :.... :
ALLERG VALVAGPAASYAADAGYAPTPAAAGAAAGKITPTQEQLME----DINVGFKAAVAAAA
      20      30      40      50      60

      40      50      60      70
SYHT0H NSVILYGRWRQKSSRIIFEMRKLSSYLFSSSWLFFFLATCLNTSPP

ALLERG GAPPADKFKTFQAAFSASVEASAAKLNAQAAPGFVSHVAATSDATYKAAVGATPEAKFDS
      70      80      90      100      110      120

>>ALLERGEN_2012_2|gi|510515|emb|CAA56343.1| Kunitz tryps (208 aa)
  initn: 49 initl: 49 opt: 49 Z-score: 87.0 bits: 21.7 E(): 7.8
Smith-Waterman score: 49; 41.176% identity (41.176% ungapped) in 17 aa overlap
(42-58:134-150)

      20      30      40      50      60      70
SYHT0H VLFKQNDTDLNIGFYTNHMIFINSVILYGRWRQKSSRIIFEMRKLSSYLFSSSWLFFFLA
      : :: . : :.... ::
ALLERG FHLCTPLSLNSFSVDRYSQGSARTPCQTHWLQKHNRWCFRIQRASSESNNYKLVFCTSN
      110      120      130      140      150      160

SYHT0H TCLNTSPP

ALLERG DDSSCGDIVAPIDREGNRPLIVTHDQNHPLLQVQFKVEAYESSTA
      170      180      190      200

>>ALLERGEN_2012_2|gi|11991227|gb|AAG42254.1|AF306707_1 p (303 aa)
  initn: 37 initl: 37 opt: 50 Z-score: 87.0 bits: 22.2 E(): 7.9
Smith-Waterman score: 50; 40.909% identity (50.000% ungapped) in 22 aa overlap (4-
25:44-61)

```

```

              10      20      30
SYHT0H      MFLQTTPTKYQVLFKQNDTDLNIGFYTNHMIFI
              . : : . : : . : : :
ALLERG VALVAGPAASYAADAGYAPATPAAAGAAAGKITPTQEQLME---DINVGFKA AVAAAA
              20      30      40      50      60

              40      50      60      70
SYHT0H NSVILYGRWRQKSSRIIFEMRKLSSYLFSSSWLFFFLATCLNTSPP

ALLERG GAPPADKFKTFQAAFSASVEASAAKLNAAPGFVSHVAATSDATYKAAVGATPEAKFDS
              70      80      90      100      110      120

```

79 residues in 1 query sequences  
 367796 residues in 1603 library sequences  
 Scomplib [34t11]  
 start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012  
 Scan time: 0.020 Display time: 0.000

Function used was FASTA [version 3.4t11 Apr 17, 2002]  
 FASTA searches a protein or DNA sequence data bank  
 version 3.4t11 Apr 17, 2002  
 Please cite:  
 W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

/grid/.qsub\_tmp/1330373600kde\_agent/grid\_NbCuMBoIO: 35 aa  
 >SYHT0H2\_I\_43  
 vs /data/fasta/ALLERGEN\_2012\_2 library  
 searching /data/fasta/ALLERGEN\_2012\_2 library

367796 residues in 1603 sequences  
 Expectation\_n fit: rho(ln(x))= 4.4141+/-0.0033; mu= 1.1420+/- 0.172  
 mean\_var=31.1013+/- 7.648, 0's: 21 Z-trim: 22 B-trim: 0 in 0/40  
 Lambda= 0.2300

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2  
 join: 36, opt: 24, gap-pen: -12/-2, width: 16  
 Scan time: 0.010

The best scores are:

					opt	bits	E(1603)
ALLERGEN_2012_2	gi 21783 emb CAA30570.1	unnamed	( 356)	57	24	0.92	
ALLERGEN_2012_2	gi 75317968 sp O22116 O22116_WHEA	( 373)	55	24	1.5		
ALLERGEN_2012_2	gi 21314465 gb AAM46958.1 AF51085	( 538)	53	23	3.5		
ALLERGEN_2012_2	gi 112754 sp P04403.2 2SS1_BEREX	( 146)	47	21	3.7		
ALLERGEN_2012_2	gi 56788031 gb AAW29810.1 seed s	( 507)	52	23	4.2		
ALLERGEN_2012_2	gi 62550933 emb CAI79052.1 putat	( 326)	50	22	4.2		
ALLERGEN_2012_2	gi 149208403 gb ABR21772.1 congl	( 455)	51	22	4.7		
ALLERGEN_2012_2	gi 335331566 gb AEH31546.1 low m	( 369)	50	22	4.8		
ALLERGEN_2012_2	gi 169950562 gb ACB05815.1 congl	( 611)	51	22	6.3		

>>ALLERGEN\_2012\_2|gi|21783|emb|CAA30570.1| unnamed prote (356 aa)  
 initn: 44 initl: 44 opt: 57 Z-score: 103.7 bits: 24.4 E(): 0.92  
 Smith-Waterman score: 57; 30.769% identity (30.769% ungapped) in 26 aa overlap (4-29:187-212)

```

              10      20      30
SYHT0H      MSTHPSVYKYLAPPSLLREQNLREIVLEREREQ
              : : : : : : : : : :
ALLERG QQPPIQQQQPPFLQQQRPPFSRQQQIPVIHPSVLQQQLNPCKVFLQQQCIPVAMQRCCLAR
              160      170      180      190      200      210

```



Smith-Waterman score: 50; 25.000% identity (25.000% ungapped) in 24 aa overlap (4-27:145-168)

```

                                10      20      30
SYHT0H                      MSTHPSVYKYLAPPSLLREQNLREIVLEREREQ
                                : : : : : : : : : :
ALLERG QQQLPPFSQQLPFFSQQQQPVLLQQQIPFVHPSILQQLNPKVFLQQQCSPVAMPQSLAR
                                120      130      140      150      160      170
```

SYHT0H VA

```

ALLERG SQMLQQSSCHVMQQQCCQQLPQIPQQSRYEAIRAIVYSIILQEQQQVQGSIQTQQQQPQQ
                                180      190      200      210      220      230
```

>>ALLERGEN\_2012\_2|gi|149208403|gb|ABR21772.1| conglutin (455 aa)  
initn: 47 init1: 47 opt: 51 Z-score: 91.0 bits: 22.4 E(): 4.7  
Smith-Waterman score: 51; 50.000% identity (50.000% ungapped) in 20 aa overlap  
(14-33:123-142)

```

                                10      20      30
SYHT0H                      MSTHPSVYKYLAPPSLLREQNLREIVLEREREQVA
                                : : : : : : : : : :
ALLERG RREHHREREQEQQPRPQRRQEEEEEEWQPRRQRPQSRREEREEREQEQQSSSGSQRGS
                                100      110      120      130      140      150

ALLERG GDERRQHRRRVHREEREQEQDSRSDSRQRNPNYHFSSNRFQTYRNRNGQIRVLERFNQ
                                160      170      180      190      200      210
```

>>ALLERGEN\_2012\_2|gi|335331566|gb|AEH31546.1| low molecu (369 aa)  
initn: 43 init1: 43 opt: 50 Z-score: 90.8 bits: 22.0 E(): 4.8  
Smith-Waterman score: 50; 25.000% identity (25.000% ungapped) in 24 aa overlap (4-27:202-225)

```

                                10      20      30
SYHT0H                      MSTHPSVYKYLAPPSLLREQNLREIVLEREREQ
                                : : : : : : : : : :
ALLERG QQQQPSLPQQPPFSQQQQQLVLPQQQIPFVHPSILQQLNPKVFLQQQCSPVAMPQSLAR
                                180      190      200      210      220      230
```

SYHT0H VA

```

ALLERG SQMLQQSSCHVMQQQCCQQLPQIPQQSRYEAIRAIIYSIILQEQQQVQGSIQTPQQQPQQ
                                240      250      260      270      280      290
```

>>ALLERGEN\_2012\_2|gi|169950562|gb|ACB05815.1| conglutin (611 aa)  
initn: 47 init1: 47 opt: 51 Z-score: 88.6 bits: 22.4 E(): 6.3  
Smith-Waterman score: 51; 50.000% identity (50.000% ungapped) in 20 aa overlap  
(14-33:123-142)

```

                                10      20      30
SYHT0H                      MSTHPSVYKYLAPPSLLREQNLREIVLEREREQVA
                                : : : : : : : : : :
ALLERG RREHHREREQEQQPRPQRRQEEEEEEWQPRRQRPQSRREEREEREQEQQSSSGSQRGG
                                100      110      120      130      140      150

ALLERG GDERRQHRRRVHREEREQEQDSRSDSRQRNPNYHFSSNRFQTYRNRNGQIRVLERFNQ
                                160      170      180      190      200      210
```

35 residues in 1 query sequences



```
>>ALLERGEN_2012_2|gi|262232390|gb|ACY38525.1| allergen C (174 aa)
  initn: 44 initl: 44 opt: 53 Z-score: 86.1 bits: 21.1 E(): 8.7
Smith-Waterman score: 53; 30.233% identity (32.500% ungapped) in 43 aa overlap
(29-68:2-44)
```

```

          10          20          30          40          50
SYHT0H MRLWDTQPGVHCAQLDTSMDGMLLVFGKGILSC-QLLQGQLGQLPRSELV--ISAPWVD
          : . : : . . . : : . . . : :
ALLERG          MKILLLLCLALVLASDAQPLPNVLTQVSGPWKT
                      10          20          30

          60
SYHT0H LQIVQTFGNKVS
          : : . . . : .
ALLERG LYISSNNLDKIGDNGPFRIYMRGINVDIPRLKMSFNFYVKVDGECVENSVGASIGRDNLI
          40          50          60          70          80          90
```

```
69 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.020 Display time: 0.000
```

```
Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448
```

```
/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuBoIO: 65 aa
>SYHT0H2_I_45
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library
```

```
367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 4.8252+/-0.00351; mu= 0.8977+/- 0.184
mean_var=42.9082+/-11.183, 0's: 7 Z-trim: 7 B-trim: 0 in 0/42
Lambda= 0.1958
```

```
FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.020
```

```
The best scores are:
                                opt bits E(1603)
ALLERGEN_2012_2|gi|23616947|dbj|BAC20650.1| putat ( 160) 56 23 2.2
ALLERGEN_2012_2|gi|114152864|sp|Q01883.2|RAG1_ORY ( 163) 55 23 2.8
ALLERGEN_2012_2|gi|1304217|dbj|BAA07773.1| allerg ( 109) 52 22 3.4
ALLERGEN_2012_2|gi|1304216|dbj|BAA07772.1| allerg ( 111) 52 22 3.5
ALLERGEN_2012_2|gi|1304218|dbj|BAA07774.1| allerg ( 113) 52 22 3.5
ALLERGEN_2012_2|gi|1703445|sp|P54958.1|ASP2_BLAG ( 352) 57 23 3.9
ALLERGEN_2012_2|gi|145105726|gb|ABP35603.1| Bla g ( 352) 57 23 3.9
ALLERGEN_2012_2|gi|1398915|dbj|BAA07711.1| allerg ( 160) 52 22 4.9
ALLERGEN_2012_2|gi|47117012|sp|Q7M4I5.1|PA2_APIDO ( 134) 51 21 5.1
ALLERGEN_2012_2|gi|218193|dbj|BAA01998.1| allerge ( 165) 52 22 5.1
ALLERGEN_2012_2|gi|1398913|dbj|BAA07710.1| allerg ( 166) 52 22 5.1
ALLERGEN_2012_2|gi|114152865|sp|Q01882.2|RAG2_ORY ( 166) 52 22 5.1
ALLERGEN_2012_2|gi|897811|emb|CAA24933.1| unnamed ( 101) 49 21 5.7
ALLERGEN_2012_2|gi|7435005|pir|A59055 phospholip ( 134) 50 21 6.1
ALLERGEN_2012_2|gi|24638082|sp|Q9BMK4.1|PA2_APICC ( 134) 50 21 6.1
ALLERGEN_2012_2|gi|163825|gb|AAC37318.1| major al ( 92) 47 20 7.8
```

```
>>ALLERGEN_2012_2|gi|23616947|dbj|BAC20650.1| putative a (160 aa)
```

initn: 48 initl: 48 opt: 56 Z-score: 96.7 bits: 22.8 E(): 2.2  
 Smith-Waterman score: 56; 30.769% identity (32.432% ungapped) in 39 aa overlap (5-42:90-127)

```

                                10      20      30
SYHT0H                      MTHNPTILRKTLPLYKEVHFIWRGHA-EITSPPC
                                :. . . . :. . . . :
ALLERG RQCAGGAVDEQVRQDCCRQLAAIDDSFCRCPALSHMLVGMVKELGAPAKGQPMDEVFPGC
        60      70      80      90      100     110

```

```

                40      50      60
SYHT0H LRRGDQLRLGQLQQLIWPRFVISLTITLRRRLQ
        :. . . :. .
ALLERG -RRGDMKRVAASLP AFCNV D I P I G I G V C Y W L S Y P M N P A T G H
        120     130     140     150     160

```

>>ALLERGEN\_2012\_2|gi|114152864|sp|Q01883.2|RAG1\_ORYSJ Re (163 aa)  
 initn: 42 initl: 42 opt: 55 Z-score: 95.1 bits: 22.5 E(): 2.8  
 Smith-Waterman score: 55; 43.333% identity (50.000% ungapped) in 30 aa overlap  
 (14-41:102-129)

```

                                10      20      30      40
SYHT0H                      MTHNPTILRKTLPLYKEVHFIWRGH--AEITSPPCLRRGDQLR
                                :. . . . :. . . . : : : : :
ALLERG EQVWQDCCRQLAAVDDGWCRGALDHMLSGIYRELGATEAGHPMAEVF-PGC-RRGDLER
        80      90      100     110     120

```

```

                50      60
SYHT0H LGQLQQLIWPRFVISLTITLRRRLQ

ALLERG AAASLP AFCNV D I P N G P G V C Y W L G Y P R T P R T G H
        130     140     150     160

```

>>ALLERGEN\_2012\_2|gi|1304217|dbj|BAA07773.1|allergenic (109 aa)  
 initn: 42 initl: 42 opt: 52 Z-score: 93.5 bits: 21.7 E(): 3.4  
 Smith-Waterman score: 52; 43.333% identity (50.000% ungapped) in 30 aa overlap  
 (14-41:48-75)

```

                                10      20      30      40
SYHT0H                      MTHNPTILRKTLPLYKEVHFIWRGH--AEITSPPCLRRGDQLR
                                :. . . . :. . . . : : : : :
ALLERG EQLRQDCCRQLAAVDDSWCRCSALNHMVGGIYRELGATDVGHHPMAEVF-PGC-RRGDLER
        20      30      40      50      60      70

```

```

                50      60
SYHT0H LGQLQQLIWPRFVISLTITLRRRLQ

ALLERG ARATLP AFCNV D I P N G T G G V C Y W L G Y P R T P R T G H
        80      90      100

```

>>ALLERGEN\_2012\_2|gi|1304216|dbj|BAA07772.1|allergenic (111 aa)  
 initn: 42 initl: 42 opt: 52 Z-score: 93.3 bits: 21.7 E(): 3.5  
 Smith-Waterman score: 52; 43.333% identity (50.000% ungapped) in 30 aa overlap  
 (14-41:50-77)

```

                                10      20      30      40
SYHT0H                      MTHNPTILRKTLPLYKEVHFIWRGH--AEITSPPCLRRGDQLR
                                :. . . . :. . . . : : : : :
ALLERG EQLRQDCCRQLAAVDDSWCRCSALNHMVGGIYRELGATDVGHHPMAEVF-PGC-RRGDLER
        20      30      40      50      60      70

```

```

          50          60
SYHT0H LGQLQQLIWPRFVISLTITLRRQLQ

ALLERG RRRSLPAFCNVNDIPNGTGGVCYWLGYPRTPRTGH
          80          90          100          110

>>ALLERGEN_2012_2|gi|1304218|dbj|BAA07774.1| allergenic (113 aa)
  initn: 42 initl: 42 opt: 52 Z-score: 93.2 bits: 21.7 E(): 3.5
Smith-Waterman score: 52; 43.333% identity (50.000% ungapped) in 30 aa overlap
(14-41:52-79)

          10          20          30          40
SYHT0H          MTHNPTILRKTLPLYKEVHFIWRGH--AEITSPPCLRRGDQLR
          .:.:.      ::  ::  : : ::::  :
ALLERG QQLRQDCCRQLAAVDDSWCRCALNHMVGGIYRELGATDVGHMAEVF-PGC-RRGDLER
          30          40          50          60          70

          50          60
SYHT0H LGQLQQLIWPRFVISLTITLRRQLQ

ALLERG AAASLPAFCNVNDIPNGTGGVCYWLGYPRTPRTGH
          80          90          100          110

>>ALLERGEN_2012_2|gi|1703445|sp|P54958.1|ASP2_BLAG RecN (352 aa)
  initn: 43 initl: 43 opt: 57 Z-score: 92.5 bits: 23.2 E(): 3.9
Smith-Waterman score: 57; 46.667% identity (51.852% ungapped) in 30 aa overlap (1-
29:16-43)

          10          20          30          40
SYHT0H          MTHNPTILRKTLPLYKEVH-FIWRGHAEITSPPCLRRGDQLRLGQ
          .::  . :  .::: :: ::  .: ::
ALLERG MIGLKLVTVLFAVATITHAAELQR--VPLYKLHVHFINTQYAGITKIGNQNFLTTFDSTS
          10          20          30          40          50

          50          60
SYHT0H LQQLIWPRFVISLTITLRRQLQ

ALLERG CNVVVASQECVGGACVCPNLQKYEKLKPKYISDGNVQVKFFDTGSAVGRGIEDSLTISNL
          60          70          80          90          100          110

>>ALLERGEN_2012_2|gi|145105726|gb|ABP35603.1| Bla g 2 al (352 aa)
  initn: 43 initl: 43 opt: 57 Z-score: 92.5 bits: 23.2 E(): 3.9
Smith-Waterman score: 57; 46.667% identity (51.852% ungapped) in 30 aa overlap (1-
29:16-43)

          10          20          30          40
SYHT0H          MTHNPTILRKTLPLYKEVH-FIWRGHAEITSPPCLRRGDQLRLGQ
          .::  . :  .::: :: ::  .: ::
ALLERG MIGLKIVTVLFAVATITHAAELQR--VPLYKLHVHFINTQYAGITKIGNQNFLTTFDSTS
          10          20          30          40          50

          50          60
SYHT0H LQQLIWPRFVISLTITLRRQLQ

ALLERG CNVVVASQECVGGACVCPNLQKYEKLKPRYISDGNVQVKFFDTGSAVGRGIEDSLTIFNL
          60          70          80          90          100          110

>>ALLERGEN_2012_2|gi|1398915|dbj|BAA07711.1| allergenic (160 aa)
  initn: 42 initl: 42 opt: 52 Z-score: 90.6 bits: 21.7 E(): 4.9
Smith-Waterman score: 52; 43.333% identity (50.000% ungapped) in 30 aa overlap
(14-41:98-125)

```



```

              10      20      30      40
SYHT0H      MTHNPTILRKTLPLYKEVHFIWRGH--AEITSPPCLRRGDQLR
              .:.:.      ::  ::  : : : : :
ALLERG EQLRQDCCRQLAAVDDSWCRCSALNHMVGGIYRELGATDVGHMPAEVF-PGC-RRGDLER
              70      80      90      100      110      120

```

```

              50      60
SYHT0H LGQLQQLIWPRFVISLTITLRRQLQ

ALLERG AAASLP AFCNV DIPNGTGGVCYWLGYTPRTPRTGH
              130      140      150      160

```

>>ALLERGEN\_2012\_2|gi|47117012|sp|Q7M4I5.1|PA2\_APIDO RecN (134 aa)  
 initn: 38 init1: 38 opt: 51 Z-score: 90.4 bits: 21.4 E(): 5.1  
 Smith-Waterman score: 51; 41.667% identity (41.667% ungapped) in 12 aa overlap  
 (21-32:7-18)

```

              10      20      30      40      50      60
SYHT0H MTHNPTILRKTLPLYKEVHFIWRGHAETSPPCLLRRGDLRLGQLQQLIWPRFVISLTIT
              .: : : : : : :
ALLERG      IIYPGTLWC GHNVS SSPDELGRFKHTDSCCRSHDMCPDVMSAGES
              10      20      30      40

```

```

SYHT0H LRRQLQ

ALLERG KHGLTNTASHTRLSCDCDDKFYDCLKNSSDTISSYFVGEMYFNILDTKCYKLEHPVTGCG
              50      60      70      80      90      100

```

>>ALLERGEN\_2012\_2|gi|218193|dbj|BAA01998.1| allergenic p (165 aa)  
 initn: 42 init1: 42 opt: 52 Z-score: 90.4 bits: 21.7 E(): 5.1  
 Smith-Waterman score: 52; 43.333% identity (50.000% ungapped) in 30 aa overlap  
 (14-41:104-131)

```

              10      20      30      40
SYHT0H      MTHNPTILRKTLPLYKEVHFIWRGH--AEITSPPCLRRGDQLR
              .:.:.      ::  ::  : : : : :
ALLERG EQLAQDCCRQLAAVDDSWCRCSALNHMVGGIYRELGATDVGHMPAEVF-PGC-RRGDLER
              80      90      100      110      120      130

```

```

              50      60
SYHT0H LGQLQQLIWPRFVISLTITLRRQLQ

ALLERG AAASLP AFCNV DIPNGTGGVCYWLGYTPRTPRTGH
              140      150      160

```

>>ALLERGEN\_2012\_2|gi|1398913|dbj|BAA07710.1| allergenic (166 aa)  
 initn: 42 init1: 42 opt: 52 Z-score: 90.4 bits: 21.7 E(): 5.1  
 Smith-Waterman score: 52; 43.333% identity (50.000% ungapped) in 30 aa overlap  
 (14-41:105-132)

```

              10      20      30      40
SYHT0H      MTHNPTILRKTLPLYKEVHFIWRGH--AEITSPPCLRRGDQLR
              .:.:.      ::  ::  : : : : :
ALLERG EQLRQDCCRQLAAVDDSWCRCSALNHMVGGIYRELGATDVGHMPAEVF-PGC-RRGDLER
              80      90      100      110      120      130

```

```

              50      60
SYHT0H LGQLQQLIWPRFVISLTITLRRQLQ

ALLERG AAASLP AFCNV DIPNGTGGVCYWLGYTPRTPRTGH
              140      150      160

```

>>ALLERGEN\_2012\_2|gi|114152865|sp|Q01882.2|RAG2\_ORYSJ Re (166 aa)  
 initn: 42 initl: 42 opt: 52 Z-score: 90.4 bits: 21.7 E(): 5.1  
 Smith-Waterman score: 52; 43.333% identity (50.000% ungapped) in 30 aa overlap  
 (14-41:105-132)

```

              10      20      30      40
SYHT0H      MTHNPTILRKTLPLYKEVHFIWRGH--AEITSPPCLRGDLR
              .:.:.      ::  ::  : : : : :
ALLERG  EQLRQDCCRQLAAVDDSWCRCSALNHMVGGIYRELGATDVGHMAEVF-PGC-RRGDLER
              80      90      100      110      120      130

              50      60
SYHT0H  LGQLQQLIWPRFVISLTLRLRQ

ALLERG  AAASLP AFCNV DIPNGTGGVCYWLGYPRTPRTGH
              140      150      160

```

>>ALLERGEN\_2012\_2|gi|897811|emb|CAA24933.1| unnamed prot (101 aa)  
 initn: 41 initl: 41 opt: 49 Z-score: 89.4 bits: 20.8 E(): 5.7  
 Smith-Waterman score: 49; 47.826% identity (52.381% ungapped) in 23 aa overlap  
 (29-51:25-45)

```

              10      20      30      40      50      60
SYHT0H  MTHNPTILRKTLPLYKEVHFIWRGHAETSPPCLRRGDLRLGQLQQLIWPRFVISLTLR
              :::  .:.:  ::  ::  .:
ALLERG      EKLGGQGGPRQWLQPRQGGQGYPTSPQQSGQGQL--GQGGQGYPTSPQQSGQG
              10      20      30      40      50

```

SYHT0H LRLRQ

```

ALLERG  QQGYDSPYHVSAEHQAASLKVAKAQQQLAAQLPAMCRLEGGDALLASQ
              60      70      80      90      100

```

>>ALLERGEN\_2012\_2|gi|7435005|pir|A59055 phospholipase A (134 aa)  
 initn: 37 initl: 37 opt: 50 Z-score: 88.9 bits: 21.1 E(): 6.1  
 Smith-Waterman score: 50; 41.667% identity (41.667% ungapped) in 12 aa overlap  
 (21-32:7-18)

```

              10      20      30      40      50      60
SYHT0H  MTHNPTILRKTLPLYKEVHFIWRGHAETSPPCLRRGDLRLGQLQQLIWPRFVISLTLR
              .: : : : : :
ALLERG      IIYPGTLWCCHGNVSSGPNELGRFKHTDACCRTDHMC PDVMSAGES
              10      20      30      40

```

SYHT0H LRLRQ

```

ALLERG  KHGLTNTASHTRLSCDCDDTFYDCLKNSGEKISSYFVGKMYFNLIDTKCYKLEHPVTGCG
              50      60      70      80      90      100

```

>>ALLERGEN\_2012\_2|gi|24638082|sp|Q9BMK4.1|PA2\_APICC RecN (134 aa)  
 initn: 37 initl: 37 opt: 50 Z-score: 88.9 bits: 21.1 E(): 6.1  
 Smith-Waterman score: 50; 41.667% identity (41.667% ungapped) in 12 aa overlap  
 (21-32:7-18)

```

              10      20      30      40      50      60
SYHT0H  MTHNPTILRKTLPLYKEVHFIWRGHAETSPPCLRRGDLRLGQLQQLIWPRFVISLTLR
              .: : : : : :
ALLERG      IIYPGTLWCCHGNVSSGPNELGRFKHTDACCRTDHMC PDVMSAGES
              10      20      30      40

```

SYHT0H LRRLQ

ALLERG KHGLTNTASHTRLSCDDTFYDCLKNSGDKISSYFVGKMYFNLIDTKCYKLEHPVTGCG  
50 60 70 80 90 100

>>ALLERGEN\_2012\_2|gi|163825|gb|AAC37318.1| major allerge (92 aa)  
initn: 40 initl: 40 opt: 47 Z-score: 87.1 bits: 20.2 E(): 7.8  
Smith-Waterman score: 47; 34.483% identity (37.037% ungapped) in 29 aa overlap  
(14-42:10-36)

10 20 30 40 50 60  
SYHT0H MTHNPTILRKTLPLYKEVHFIWRGHAEITSPCLRRGDQLRLGQLQLIWPFRFVISLTIT  
:. . .: .: .: : .: .: :  
ALLERG MKGARVLVLLWAALLLIWGGNCEIC--PAVKRDVDLFLTGTTPDEYVEQVAQYKALP  
10 20 30 40 50

SYHT0H LRRLQ

ALLERG VVLENARILKNCVDAKMTEEDKENALSLLDKIYTSPLC  
60 70 80 90

65 residues in 1 query sequences  
367796 residues in 1603 library sequences  
Scomplib [34t11]  
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012  
Scan time: 0.020 Display time: 0.000

Function used was FASTA [version 3.4t11 Apr 17, 2002]  
FASTA searches a protein or DNA sequence data bank  
version 3.4t11 Apr 17, 2002  
Please cite:  
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

/grid/.qsub\_tmp/1330373600kde\_agent/grid\_NbCuMBoIO: 40 aa  
>SYHT0H2\_I\_46  
vs /data/fasta/ALLERGEN\_2012\_2 library  
searching /data/fasta/ALLERGEN\_2012\_2 library

367796 residues in 1603 sequences  
Expectation\_n fit:  $\rho(\ln(x)) = 3.4130 \pm 0.00286$ ;  $\mu = 9.5190 \pm 0.150$   
mean\_var=41.7198 $\pm$ 10.742, 0's: 9 Z-trim: 9 B-trim: 0 in 0/43  
Lambda= 0.1986

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2  
join: 36, opt: 24, gap-pen: -12/-2, width: 16  
Scan time: 0.020

The best scores are:

			opt	bits	E(1603)
ALLERGEN_2012_2	gi 170722 gb AAA34281.1	pre-alpha ( 262)	56	22	3.8
ALLERGEN_2012_2	gi 473876 gb AAA17741.1	alpha-gl ( 287)	56	22	4.1
ALLERGEN_2012_2	gi 5381323 gb AAD42943.1	AF091841 ( 148)	52	21	5.8
ALLERGEN_2012_2	gi 1170095 sp P46419.1	GSTM1 DERP ( 219)	52	21	7.5
ALLERGEN_2012_2	gi 60920878 gb AAX37326.1	glutat ( 219)	52	21	7.5
ALLERGEN_2012_2	gi 105969543 gb ABF81661.1	EXPB1 ( 99)	48	19	9.7

>>ALLERGEN\_2012\_2|gi|170722|gb|AAA34281.1| pre-alpha-/be (262 aa)  
initn: 28 initl: 28 opt: 56 Z-score: 92.5 bits: 22.1 E(): 3.8  
Smith-Waterman score: 56; 30.000% identity (30.000% ungapped) in 30 aa overlap (2-31:56-85)



Smith-Waterman score: 52; 35.294% identity (41.379% ungapped) in 34 aa overlap (32:182-214)

```

                                10      20
SYHT0H                        MPLPTVVPKMDPHPRGASWKKKTFQPRL----
                                :  :  ::  ::  ::  ::  ::
ALLERG ANISYVDFYLYEYLCRVKVMVPEVFGQFENLKRYVERMESLPRVSDYIKKQ-QPKTFNAP
                                160      170      180      190      200      210

                                30      40
SYHT0H QSKWIDVISPLT
                                :::
ALLERG TSKWNASYA
```

>>ALLERGEN\_2012\_2|gi|105969543|gb|ABF81661.1| EXPB10 [Ze (99 aa)  
initn: 35 init1: 35 opt: 48 Z-score: 85.3 bits: 19.3 E(): 9.7  
Smith-Waterman score: 48; 42.857% identity (42.857% ungapped) in 14 aa overlap  
(13-26:51-64)

```

                                10      20      30      40
SYHT0H                        MPLPTVVPKMDPHPRGASWKKKTFQPRLQSKWIDVISPLT
                                :  ::  ::  .  .  :
ALLERG ALLVKYVDGDGDIVAVDIKEKGSDTYEPLKHSWGAIWRKDSDKPIKGPITVRLTTEGGTK
                                30      40      50      60      70      80

ALLERG TVYDDVIPAEWKPNAYTT
                                90
```

40 residues in 1 query sequences  
367796 residues in 1603 library sequences  
Scomplib [34t11]  
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012  
Scan time: 0.020 Display time: 0.000

Function used was FASTA [version 3.4t11 Apr 17, 2002]  
FASTA searches a protein or DNA sequence data bank  
version 3.4t11 Apr 17, 2002  
Please cite:  
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

/grid/.qsub\_tmp/1330373600kde\_agent/grid\_NbCuMBoIO: 67 aa  
>SYHT0H2\_I\_47  
vs /data/fasta/ALLERGEN\_2012\_2 library  
searching /data/fasta/ALLERGEN\_2012\_2 library

367796 residues in 1603 sequences  
Expectation\_n fit:  $\rho(\ln(x)) = 4.3391 \pm 0.00314$ ;  $\mu = 2.9756 \pm 0.165$   
mean\_var=39.7260 $\pm$ 11.294, 0's: 11 Z-trim: 14 B-trim: 0 in 0/43  
Lambda= 0.2035

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2  
join: 36, opt: 24, gap-pen: -12/-2, width: 16  
Scan time: 0.020

The best scores are:

			opt	bits	E(1603)
ALLERGEN_2012_2	gi 32363456 sp P81729.1 CHAL_BRAR	( 91)	60	24	0.44
ALLERGEN_2012_2	gi 14423650 sp Q9U1G2.1 ALL7_LEPD	( 216)	62	25	0.63
ALLERGEN_2012_2	gi 4538529 emb CAB39376.1  Cop c1	( 81)	57	24	0.73
ALLERGEN_2012_2	gi 14422361 emb CAC41634.1  plant	( 131)	56	23	1.4
ALLERGEN_2012_2	gi 14422359 emb CAC41633.1  plant	( 131)	56	23	1.4
ALLERGEN_2012_2	gi 14422363 emb CAC41635.1  plant	( 131)	56	23	1.4

ALLERGEN\_2012\_2|gi|18772|emb|CAA45778.1| trypsin ( 217) 53 22 3.9  
 ALLERGEN\_2012\_2|gi|146217148|gb|ABQ10638.1| vicil ( 136) 48 21 7.2  
 ALLERGEN\_2012\_2|gi|47606039|sp|Q8H6L7.1|PHLB\_PHLF ( 143) 48 21 7.6

>>ALLERGEN\_2012\_2|gi|32363456|sp|P81729.1|CHAL\_BRARA Rec (91 aa)  
 initn: 56 initl: 56 opt: 60 Z-score: 109.4 bits: 24.4 E(): 0.44  
 Smith-Waterman score: 60; 40.741% identity (44.000% ungapped) in 27 aa overlap  
 (26-51:33-58)

```

              10      20      30      40      50
SYHT0H      MNFLFIGSSRVDLQKDSGIVRHPLRQWRYHI-NPLALKTWLERLLFPRCSSWVG
              :  ::  ::  :  ::  :  :  :  :  :  :  :  :  :  :  :  :
ALLERG  GGQTCAGNICCSQYGYCGTTADYCSPDNNCQATYHYNYP-AQNNWDLRAVSAYCSTWDAD
              10      20      30      40      50      60

              60
SYHT0H  HLWDHCRQRHLQR

ALLERG  KPYSWRYGWTAFCGPAGPRCLRTNAAVTVR
              70      80      90

```

>>ALLERGEN\_2012\_2|gi|14423650|sp|Q9U1G2.1|ALL7\_LEPDS Rec (216 aa)  
 initn: 59 initl: 59 opt: 62 Z-score: 106.6 bits: 25.1 E(): 0.63  
 Smith-Waterman score: 62; 31.250% identity (31.250% ungapped) in 32 aa overlap  
 (10-41:114-145)

```

              10      20      30
SYHT0H      MNFLFIGSSRVDLQKDSGIVRHPLRQWRYHINPLALKTW
              ..::  :  :  :  :  :  :  :  :  :  :  :  :  :
ALLERG  HISRRGDAKIDTDGGAFAATLKLGDKNIRIKTDLHLDLGGKIIHPNLKFEGHIGDIDMKLK
              90      100      110      120      130      140

              40      50      60
SYHT0H  LERLLFPRCSSWVGHLWDHCRQRHLQR
              :.
ALLERG  LKLDAEGKPSLDQFEIDFEQVELFIHGLGPLDPLVDVIADSFVKYFNPQARKLVTDMKL
              150      160      170      180      190      200

```

>>ALLERGEN\_2012\_2|gi|4538529|emb|CAB39376.1| Cop c1 alle (81 aa)  
 initn: 46 initl: 46 opt: 57 Z-score: 105.5 bits: 23.5 E(): 0.73  
 Smith-Waterman score: 57; 38.235% identity (43.333% ungapped) in 34 aa overlap  
 (23-52:43-76)

```

              10      20      30      40
SYHT0H      MNFLFIGSSRVDLQKDSGIVRHPLRQWRYHINPLALKTWLERLLF---PRCS
              :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
ALLERG  HLPWLVPAPVLLLPVLPQLKPVVAHPLLLPLDTTTTLHMPPLLLQLQLPPLLSQGNPACS
              20      30      40      50      60      70

              50      60
SYHT0H  -SWVGVLWDHCRQRHLQR
              :..
ALLERG  PKWLQLLVP
              80

```

>>ALLERGEN\_2012\_2|gi|14422361|emb|CAC41634.1| plantain p (131 aa)  
 initn: 31 initl: 31 opt: 56 Z-score: 100.6 bits: 23.3 E(): 1.4  
 Smith-Waterman score: 56; 31.250% identity (37.037% ungapped) in 32 aa overlap (6-  
 37:97-123)

```

              10      20      30
SYHT0H      MNFLFIGSSRVDLQKDSGIVRHPLRQWRYHINPLA
              : . . . . . : . . . . . : . . . . .
ALLERG VVGYPHEDCEIKLVKSSRPDCSEIPKLAKGTIQTSKVDLSKNTTITEKT-----RHVKPLS
              70      80      90      100      110      120

```

```

              40      50      60
SYHT0H LKTWLERLLFPRCSSWVGVLWDHCRQRHLQR
...
ALLERG FRAKTDAPGC
              130

```

>>ALLERGEN\_2012\_2|gi|14422359|emb|CAC41633.1| plantain p (131 aa)  
 initn: 31 init1: 31 opt: 56 Z-score: 100.6 bits: 23.3 E(): 1.4  
 Smith-Waterman score: 56; 31.250% identity (37.037% ungapped) in 32 aa overlap (6-37:97-123)

```

              10      20      30
SYHT0H      MNFLFIGSSRVDLQKDSGIVRHPLRQWRYHINPLA
              : . . . . . : . . . . . : . . . . .
ALLERG VVGYPHEDCEIKLVKSSRPDCSEIPKLAKGTIQTSKVDLSKNTTITEKT-----RHVKPLS
              70      80      90      100      110      120

```

```

              40      50      60
SYHT0H LKTWLERLLFPRCSSWVGVLWDHCRQRHLQR
...
ALLERG FRAKTDAPGC
              130

```

>>ALLERGEN\_2012\_2|gi|14422363|emb|CAC41635.1| plantain p (131 aa)  
 initn: 31 init1: 31 opt: 56 Z-score: 100.6 bits: 23.3 E(): 1.4  
 Smith-Waterman score: 56; 31.250% identity (37.037% ungapped) in 32 aa overlap (6-37:97-123)

```

              10      20      30
SYHT0H      MNFLFIGSSRVDLQKDSGIVRHPLRQWRYHINPLA
              : . . . . . : . . . . . : . . . . .
ALLERG VVGYPHEDCEIKLVKSGRPDCSEIPKLAKGTIQTSKVDLSKNTTITEKT-----RHVKPLS
              70      80      90      100      110      120

```

```

              40      50      60
SYHT0H LKTWLERLLFPRCSSWVGVLWDHCRQRHLQR
...
ALLERG FRAKTDAPGC
              130

```

>>ALLERGEN\_2012\_2|gi|18772|emb|CAA45778.1| trypsin inhib (217 aa)  
 initn: 31 init1: 31 opt: 53 Z-score: 92.3 bits: 22.5 E(): 3.9  
 Smith-Waterman score: 53; 38.889% identity (41.176% ungapped) in 36 aa overlap (4-37:66-101)

```

              10      20      30
SYHT0H      MNFLFIGSSRVDLQKDSG-IVRHPLR-QWRYHI
              : . . . . : . . . : . . . . .
ALLERG LDSGGTYIILSDITAFGGIRAAPTGNERCPLTVVQSRNELDKGIGTIISPPFRIRFIAEG
              40      50      60      70      80      90

```

```

              40      50      60
SYHT0H NPLALKTWLERLLFPRCSSWVGVLWDHCRQRHLQR
...
ALLERG NPLRLKFDSFAVIMLCVGIPTEWSVVEDLPEGPAVKIGENKDAVDGWGFRIERVSDDEFNN
              100      110      120      130      140      150

```

```
>>ALLERGEN_2012_2|gi|146217148|gb|ABQ10638.1| vicilin-li (136 aa)
  initn: 46 initl: 46 opt: 48 Z-score: 87.6 bits: 20.9 E(): 7.2
Smith-Waterman score: 48; 60.000% identity (60.000% ungapped) in 10 aa overlap
(58-67:12-21)
```

```

      30      40      50      60
SYHT0H RYHINPLALKTWLERLLFPRCSSWVGVLWDHCRQRHLQR
      .: .: .: .: .: .:
ALLERG      GDYPFEPRRQIEHCRQRCQDRSQGFRETRECLETCERAGF
      10      20      30      40

ALLERG DDEPSKEVDSYTTDTRSRDPQQEFRQCRHRCQTQEEGGRQQRKCEQRCEKQLERQQGYDD
      50      60      70      80      90     100
```

```
>>ALLERGEN_2012_2|gi|47606039|sp|Q8H6L7.1|PHLB_PHLPR Rec (143 aa)
  initn: 37 initl: 37 opt: 48 Z-score: 87.3 bits: 21.0 E(): 7.6
Smith-Waterman score: 48; 31.250% identity (35.714% ungapped) in 32 aa overlap (5-
36:91-118)
```

```

      10      20      30
SYHT0H      MNFLFIGSSRVLDLQKDSGIVRHPLRQWRYHINPL
      : .: .: .: .: .: .: .: .: .: .: .:
ALLERG KIEIDQDHQEEICEVVLAKSPD TTCSEIEEFRDRARVPLTSNNGIKQQGIR---YANPI
      70      80      90     100     110
```

```

      40      50      60
SYHT0H ALKTLWLERLLFPRCSSWVGVLWDHCRQRHLQR
      : .
ALLERG AFFRKEPLKECGGILQAYDLRDAPETP
      120     130     140
```

```
67 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.020 Display time: 0.000
```

```
Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448
```

```
/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 34 aa
>SYHT0H2_I_48
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library
```

```
367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 4.8934+/-0.00396; mu= -0.2124+/- 0.204
mean_var=51.6225+/-12.815, 0's: 24 Z-trim: 26 B-trim: 79 in 1/41
Lambda= 0.1785
```

```
FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.010
```

```
The best scores are:
ALLERGEN_2012_2|gi|170710|gb|AAA34275.1| alpha-ty ( 318) 73 26 0.29
ALLERGEN_2012_2|gi|83300389|sp|O42799.2|ALL7_AS PF ( 270) 71 25 0.36
ALLERGEN_2012_2|gi|18536|emb|CAA35691.1| unnamed ( 605) 68 25 1.2
ALLERGEN_2012_2|gi|21926|emb|CAA36063.1| unnamed ( 295) 64 23 1.4
```





Smith-Waterman score: 64; 41.667% identity (43.478% ungapped) in 24 aa overlap (10-33:50-72)

```

                                10      20      30
SYHT0H                        MDLNNSEKFFPNPPHPPFWYSCPTSSFSMHPIR
                                :.:. .:.. . : : .....:
ALLERG QMETSCIPGLERPWQEQLPPQHHTLFPQQQPPFPQQQPPFSQQQP-SFLQQQPILPQLPF
      20      30      40      50      60      70

ALLERG SQQQQPVLPPQQSPFSQQQLVLPPQQQYQQVLQQQIPIVQPSVLQQLNPCKVFLQQQCNPV
      80      90     100     110     120     130
```

>>ALLERGEN\_2012\_2|gi|170738|gb|AAA34289.1| gamma-gliadin (327 aa)  
initn: 55 initl: 55 opt: 63 Z-score: 98.5 bits: 23.3 E(): 1.8  
Smith-Waterman score: 63; 37.500% identity (39.130% ungapped) in 24 aa overlap (10-32:83-106)

```

                                10      20      30
SYHT0H                        MDLNNSEKFFPNPPHPPF-WYSCPTSSFSMHPIR
                                :.:. .:.. . : : .....:
ALLERG PQPQQTFPHQPQQQVPQPQQPQQPFLQPQQPFPQPQQPFPQTQQPQQPFPQPQQPFPQ
      60      70      80      90     100     110

ALLERG TQQPQQPFPQPQQPFPQTQQPQQPFPQLQQPQQPFPQPQQQLPQPQQPQQSFPQQQRPF
      120     130     140     150     160     170
```

>>ALLERGEN\_2012\_2|gi|75219081|sp|O22108|O22108\_WHEAT LMM (285 aa)  
initn: 72 initl: 50 opt: 62 Z-score: 98.1 bits: 23.0 E(): 1.9  
Smith-Waterman score: 62; 39.286% identity (50.000% ungapped) in 28 aa overlap (11-33:36-62)

```

                                10      20      30
SYHT0H                        MDLNNSEKFFPNPP-----HPPFWYSCPTSSFSMHPIR
                                :. : . : : : : : : : : : : :
ALLERG LERPWQQQLPPLPPQQTFFPQQPLFSQQQQQLLFPQQPSFSQQQPPFWQQQPPFS-QQQPILP
      10      20      30      40      50      60

ALLERG QQPPFSQQQLVLPPQQSPFSQQQQQLLILPPQQQQQLPQQQISIVQPSVLQQLNPCKVFLQQ
      70      80      90     100     110     120
```

>>ALLERGEN\_2012\_2|gi|21773|emb|CAA31685.1| unnamed prote (307 aa)  
initn: 72 initl: 50 opt: 62 Z-score: 97.6 bits: 23.0 E(): 2  
Smith-Waterman score: 62; 39.286% identity (50.000% ungapped) in 28 aa overlap (11-33:59-85)

```

                                10      20      30
SYHT0H                        MDLNNSEKFFPNPP-----HPPFWYSCPTSSFSMHPIR
                                :. : . : : : : : : : : : : :
ALLERG LERPWQQQLPPLPPQQTFFPQQPLFSQQQQQLLFPQQPSFSQQQPPFWQQQPPFS-QQQPILP
      30      40      50      60      70      80

ALLERG QQPPFSQQQLVLPPQQPPFSQQQQPVLPPQQSPFPQQQQQHQQQLVQQQIPVVQPSILQQ
      90     100     110     120     130     140
```

>>ALLERGEN\_2012\_2|gi|422005|pir||S32101 PHLP5A protein - (257 aa)  
initn: 60 initl: 60 opt: 60 Z-score: 96.0 bits: 22.4 E(): 2.5  
Smith-Waterman score: 60; 66.667% identity (66.667% ungapped) in 9 aa overlap (10-18:232-240)

```

              10      20      30
SYHT0H      MDLNNSEKFPFPNPPHPPFWYSCPTSSFSMHPPIR
              :: ::::
ALLERG APEVKYTVFETALKKAITAMSEAQKAAPPLPPPPQPPLAATGAATAATGGYKV
              210      220      230      240      250

>>ALLERGEN_2012_2|gi|1092249|prf||2023228A major allergen (285 aa)
  initn: 60 initl: 60 opt: 60 Z-score: 95.3 bits: 22.5 E(): 2.7
Smith-Waterman score: 60; 66.667% identity (66.667% ungapped) in 9 aa overlap (10-
18:260-268)

```

```

              10      20      30
SYHT0H      MDLNNSEKFPFPNPPHPPFWYSCPTSSFSMHPPIR
              :: ::::
ALLERG APEVKYTVFETALKKAITAMSEAQKAAPPLPPPPQPPLAATGAATAATGGYKV
              230      240      250      260      270      280

>>ALLERGEN_2012_2|gi|335331566|gb|AEH31546.1| low molecu (369 aa)
  initn: 45 initl: 45 opt: 59 Z-score: 92.2 bits: 22.2 E(): 4
Smith-Waterman score: 59; 27.586% identity (27.586% ungapped) in 29 aa overlap (5-
33:106-134)

```

```

              10      20      30
SYHT0H      MDLNNSEKFPFPNPPHPPFWYSCPTSSFSMHPPIR
              .... :: . :::: . : : ....
ALLERG QQQQPPFSQQQPPFSQQQQPLSQQQQPPFSQQQPPFSQQQPPFSQQQPPFSQQQQPVL
              80      90      100      110      120      130

ALLERG PQQPSFSQQQLPPFSQQQSPFSQQQQIVLQQQPPFLQQQPSLPQQPPFSQQQQQLVLPQ
              140      150      160      170      180      190

>>ALLERGEN_2012_2|gi|170702|gb|AAA34272.1| gamma gliadin (302 aa)
  initn: 129 initl: 58 opt: 58 Z-score: 92.1 bits: 22.0 E(): 4.1
Smith-Waterman score: 58; 50.000% identity (50.000% ungapped) in 12 aa overlap (8-
19:104-115)

```

```

              10      20      30
SYHT0H      MDLNNSEKFPFPNPPHPPFWYSCPTSSFSMHPPIR
              ..... :: ::
ALLERG YPHQPQQQFPQTQQPQQPFPQPQQTFFPQQPQLPFPQQPQQPFPQPQQPQQPFPQSQQPQQ
              80      90      100      110      120      130

ALLERG PFPQPQQQFPQPQQPQQSFPQQQQPAIQSFLQQQMNPCKNFLQCCNHVSLVSSLVSIIL
              140      150      160      170      180      190

>>ALLERGEN_2012_2|gi|21913174|gb|AAM77471.1| major aller (115 aa)
  initn: 52 initl: 52 opt: 52 Z-score: 90.4 bits: 20.2 E(): 5.1
Smith-Waterman score: 52; 66.667% identity (66.667% ungapped) in 9 aa overlap (20-
28:86-94)

```

```

              10      20      30
SYHT0H      MDLNNSEKFPFPNPPHPPFWYSCPTSSFSMHPPIR
              :::: .::
ALLERG NSLGFNIKATNGGTLDFTCQAQADKLEDHKWYSCGENSFMDFSFDSRSGLLLKQKVSDE
              60      70      80      90      100      110

>>ALLERGEN_2012_2|gi|1063270|dbj|BAA11251.1| gamma-gliad (279 aa)
  initn: 119 initl: 47 opt: 56 Z-score: 89.9 bits: 21.4 E(): 5.4
Smith-Waterman score: 56; 34.483% identity (45.455% ungapped) in 29 aa overlap
(10-32:17-44)

```

```

          10          20          30
SYHT0H      MDLNNSEKFFPNPPHPPF-----WYSCPTSSFSMHPIR
              ::::: : ::      . : :: :
ALLERG NIQVDPGQVQWPQQQFPQP-HQPFSSQQPQOTFPQPQQTFFHQPQQQFSQPQQPQQQFI
          10          20          30          40          50

ALLERG QPQQPFPQQPQQTYPQRPQQPFPQTQQPQQPFPQSQQPQQPFPQPQQQFPQPQQPQQSFP
          60          70          80          90          100          110

>>ALLERGEN_2012_2|gi|62484809|emb|CAI78902.1| putative g (285 aa)
  initn: 55 init1: 55 opt: 55 Z-score: 88.3 bits: 21.2 E(): 6.6
Smith-Waterman score: 59; 34.483% identity (41.667% ungapped) in 29 aa overlap (9-
34:84-110)

```

```

          10          20          30
SYHT0H      MDLNNSEKFP--FPNPPHPPFWYSCPTSSFSMHPIR
              :: :.::: : : : :.:.:.
ALLERG CLQQPQHQPQPTQQPQRPLLPFTHPFLTFPDQLLPQPPHQSF--PQPPQSYPPPLQP
          60          70          80          90          100          110

ALLERG FPQPPQKYPEQPQQPFPWQQPTIQLYLQQQLNPKCKEFLQCCRPVSLLSYLWSKIVQQS
          120          130          140          150          160          170

>>ALLERGEN_2012_2|gi|1842045|gb|AAB47552.1| major allerg (157 aa)
  initn: 52 init1: 52 opt: 52 Z-score: 88.2 bits: 20.3 E(): 6.7
Smith-Waterman score: 52; 66.667% identity (66.667% ungapped) in 9 aa overlap (20-
28:86-94)

```

```

          10          20          30
SYHT0H      MDLNNSEKFFPNPPHPPFWYSCPTSSFSMHPIR
              :::: :.::
ALLERG NSLGFNIKATNGGTLDFTCQAQADKLEDHKWYSCGENSFMDFSFSDRSGLLLKQKVSDD
          60          70          80          90          100          110

ALLERG ITYVATATLPNYCRAGNGPKDFVCQGVADAYITLVTLPKSS
          120          130          140          150

>>ALLERGEN_2012_2|gi|45680856|gb|AAS75297.1| major aller (157 aa)
  initn: 52 init1: 52 opt: 52 Z-score: 88.2 bits: 20.3 E(): 6.7
Smith-Waterman score: 52; 66.667% identity (66.667% ungapped) in 9 aa overlap (20-
28:86-94)

```

```

          10          20          30
SYHT0H      MDLNNSEKFFPNPPHPPFWYSCPTSSFSMHPIR
              :::: :.::
ALLERG NSLGFNIKATNGGTLDFTCQAQADKLEDHKWYSCGENSFMNFSFGSDRSGLLLKQKVSDD
          60          70          80          90          100          110

ALLERG ITYVATATLPNYCRAGNGPKDFVCQGVADAYITLVTLPKSS
          120          130          140          150

>>ALLERGEN_2012_2|gi|208605344|emb|CAR82265.1| D-type LM (359 aa)
  initn: 56 init1: 56 opt: 56 Z-score: 88.2 bits: 21.5 E(): 6.7
Smith-Waterman score: 67; 39.130% identity (40.909% ungapped) in 23 aa overlap
(10-32:245-266)

```

```

          10          20          30
SYHT0H      MDLNNSEKFFPNPPHPPFWYSCPTSSFSMHPIR
              :::: :.:: :.::
ALLERG PLQPQQPFPQQSQPFPQQPQQPCPLQPQQPFPQQPQQPFPQQ-PQQPFPQPQQPFPPLR
          220          230          240          250          260          270

```

ALLERG PQQPFSSQQPQQSQSFPQPQPQQPQQPSILQPQQPFLLPQQQLSQQLEQTISQQPQQPFP  
280 290 300 310 320 330

>>ALLERGEN\_2012\_2|gi|169971|gb|AAA33965.1| glycinin prec (240 aa)  
initn: 48 initl: 48 opt: 54 Z-score: 88.1 bits: 20.9 E(): 6.8  
Smith-Waterman score: 54; 39.130% identity (39.130% ungapped) in 23 aa overlap (2-24:31-53)

SYHT0H MDLNNSEKFFFPNPPHPPFWYSCPTSSFSMHLV  
: . . . . : : : : : :  
ALLERG QEDEDEDEDEEYEQTPSYPPRRPSHGKHEDDEDEDEEHQPRPDHPPQRPSRPEQQEPRG  
10 20 30 40 50 60

SYHT0H PIR

ALLERG RGCQTRNGVEENICTMKLHENIARPSRADFYNNPKAGRISTLNSLTLPALRQFGLSAQYLV  
70 80 90 100 110 120

>>ALLERGEN\_2012\_2|gi|76782247|gb|ABA54897.1| hydrophobic (134 aa)  
initn: 74 initl: 46 opt: 51 Z-score: 87.9 bits: 20.0 E(): 7  
Smith-Waterman score: 54; 40.909% identity (42.857% ungapped) in 22 aa overlap  
(10-31:52-72)

SYHT0H MDLNNSEKFFFPNPPHPPFWYSCPTSSFSMHPIR  
: . : : : : : : : :  
ALLERG MVSSSSHYDPPPPPCYVPAPLTPPPSLSPPPSLSP-PPSGPSCPDLVCLNILDGSPAD  
30 40 50 60 70 80

ALLERG DCCALIADLVLEASVCLCIQLRVLGIVNLDLNLQLILNACGPSYSPSNATCPRT  
90 100 110 120 130

>>ALLERGEN\_2012\_2|gi|121308878|dbj|BAF43534.1| SXP/RAL-2 (152 aa)  
initn: 46 initl: 46 opt: 51 Z-score: 87.1 bits: 20.0 E(): 7.8  
Smith-Waterman score: 51; 54.545% identity (54.545% ungapped) in 11 aa overlap  
(14-24:22-32)

SYHT0H MDLNNSEKFFFPNPPHPPFWYSCPTSSFSMHPIR  
: : : : : : : : :  
ALLERG MKTLIVAALFCTIGMALADDTPPPPPFLLAGAPQDVVKAFFELLKKDETKTDPEIEKDLDA  
10 20 30 40 50 60

ALLERG WVDTLGGDYKAKFETFKKEMKAKEAELAKAHEEAVAKMTPEAKKADAELSKIAEDDSLNG  
70 80 90 100 110 120

>>ALLERGEN\_2012\_2|gi|886965|emb|CAA59339.1| low molecula (261 aa)  
initn: 61 initl: 53 opt: 53 Z-score: 86.2 bits: 20.6 E(): 8.7  
Smith-Waterman score: 53; 26.667% identity (26.667% ungapped) in 30 aa overlap (4-33:42-71)

SYHT0H MDLNNSEKFFFPNPPHPPFWYSCPTSSFSMHPI  
..... : . . : : : : : : : : : :  
ALLERG QPLPLQQILWYQQQQPIQQQPQFPFPQQPQSQQQQPPLSQQQQPPFSQQQPFFSQQQQPI  
20 30 40 50 60 70

SYHT0H R

ALLERG LPQQPPFSQQQQQFPQQQQPLLPQQPPFSQQQPPFSQQQQQPPFSQQQQQPILLQQPPFS  
80 90 100 110 120 130

>>ALLERGEN\_2012\_2|gi|157418806|gb|ABV55106.1| Ani s 9 al (147 aa)  
initn: 45 initl: 45 opt: 50 Z-score: 85.9 bits: 19.8 E(): 9  
Smith-Waterman score: 50; 54.545% identity (54.545% ungapped) in 11 aa overlap  
(14-24:21-31)

10 20 30  
SYHT0H MDLNNSEKFPFPNPPHPPFWYSCPTSSFSMHPIR  
:: :: :  
ALLERG MKLCILAVAVFVVAQGPPLPPFVANAPPAVQAEFRQLANGAPDKTEAEIEAQIEQW  
10 20 30 40 50 60

ALLERG VASKGGAVQAEFNKFKQMLEQGKARAEAAHQASLTRLSPAAKAADARLSAIASNRALKVG  
70 80 90 100 110 120

34 residues in 1 query sequences  
367796 residues in 1603 library sequences  
Scomplib [34t11]  
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012  
Scan time: 0.010 Display time: 0.010

Function used was FASTA [version 3.4t11 Apr 17, 2002]  
FASTA searches a protein or DNA sequence data bank  
version 3.4t11 Apr 17, 2002  
Please cite:  
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

/grid/.qsub\_tmp/1330373600kde\_agent/grid\_NbCuMBoIO: 32 aa  
>SYHT0H2\_I\_49  
vs /data/fasta/ALLERGEN\_2012\_2 library  
searching /data/fasta/ALLERGEN\_2012\_2 library

367796 residues in 1603 sequences  
Expectation\_n fit: rho(ln(x))= 4.2703+/-0.00298; mu= -1.5553+/- 0.155  
mean\_var=32.5097+/- 8.267, 0's: 21 Z-trim: 21 B-trim: 130 in 1/40  
Lambda= 0.2249

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2  
join: 36, opt: 24, gap-pen: -12/-2, width: 16  
Scan time: 0.010

The best scores are: opt bits E(1603)

ALLERGEN_2012_2	gi 1362131 pir	C53806 major alle	( 145)	42	20	6
ALLERGEN_2012_2	gi 261824817 pdb 3F55	D Chain D,	( 316)	45	21	6.4
ALLERGEN_2012_2	gi 1063270 dbj BAA11251.1	gamma-	( 279)	44	21	7.1
ALLERGEN_2012_2	gi 33327133 gb AAQ08947.1	allerg	( 145)	41	20	7.5
ALLERGEN_2012_2	gi 124365253 gb ABN09655.1	beta-	( 374)	45	21	7.5
ALLERGEN_2012_2	gi 124294785 gb ABN03966.1	beta-	( 374)	45	21	7.5
ALLERGEN_2012_2	gi 268037674 gb ACY91851.1	beta-	( 374)	45	21	7.5
ALLERGEN_2012_2	gi 170736 gb AAA34288.1	gamma-gl	( 251)	43	21	8.1
ALLERGEN_2012_2	gi 83300389 sp O42799.2	ALL7_ASPF	( 270)	43	21	8.7
ALLERGEN_2012_2	gi 170708 gb AAA34274.1	gamma-gl	( 291)	43	21	9.3
ALLERGEN_2012_2	gi 56122438 gb AAV74343.1	Fra e	( 145)	40	20	9.4
ALLERGEN_2012_2	gi 1184668 gb AAA87456.1	beta-1,	( 374)	44	21	9.4
ALLERGEN_2012_2	gi 124365249 gb ABN09653.1	beta-	( 374)	44	21	9.4
ALLERGEN_2012_2	gi 124294783 gb ABN03965.1	beta-	( 374)	44	21	9.4
ALLERGEN_2012_2	gi 32765543 gb AAP87281.1	beta-1	( 374)	44	21	9.4
ALLERGEN_2012_2	gi 270315180 gb ACZ74626.1	beta-	( 374)	44	21	9.4

ALLERGEN\_2012\_2|gi|124365251|gb|ABN09654.1| beta- ( 374) 44 21 9.4

>>ALLERGEN\_2012\_2|gi|1362131|pir|C53806 major allergen (145 aa)  
 initn: 30 initl: 30 opt: 42 Z-score: 89.1 bits: 20.3 E(): 6  
 Smith-Waterman score: 42; 34.483% identity (40.000% ungapped) in 29 aa overlap (4-32:79-103)

	10	20	30
SYHT0H	MLTKIFFLSLSRKRLCMNCSPVFTASSVRSSI		
	..	... :	... :
ALLERG	GKITFTEVG	YTRAEG	LYSMLIERDHKNEFCEITLLSSSRK---DCDEIPT
	50	60	70 80 90 100

ALLERG FILNTVNGTTRTINPLGFFKKEALPKCPQVFNKLGMYPNNM

	110	120	130	140
--	-----	-----	-----	-----

>>ALLERGEN\_2012\_2|gi|261824817|pdb|3F55|D Chain D, Cryst (316 aa)  
 initn: 45 initl: 45 opt: 45 Z-score: 88.5 bits: 21.3 E(): 6.4  
 Smith-Waterman score: 45; 50.000% identity (50.000% ungapped) in 20 aa overlap (5-24:163-182)

	10	20	30
SYHT0H	MLTKIFFLSLSRKRLCMNCSPVFTASSVRSSI		
	:	... :	:
ALLERG	TAIDLTLVGNSYPPSAGAFRDDVRSYLNPIIRFLSSIRSPILLANIYPYFTYAGNPRDISL		
	140	150	160 170 180 190

ALLERG PYALFTSPSVVWDGQRGYKNLFDATLDALYSALERASGGSLVVSESGWPSAGAFAT

	200	210	220	230	240	250
--	-----	-----	-----	-----	-----	-----

>>ALLERGEN\_2012\_2|gi|1063270|dbj|BAA11251.1| gamma-gliad (279 aa)  
 initn: 27 initl: 27 opt: 44 Z-score: 87.7 bits: 20.9 E(): 7.1  
 Smith-Waterman score: 44; 36.000% identity (36.000% ungapped) in 25 aa overlap (8-32:134-158)

	10	20	30
SYHT0H	MLTKIFFLSLSRKRLCMNCSPVFTASSVRSSI		
	..	.. :	... :
ALLERG	PQQQFPQPQQPQQSFPQQQPSLIQQQLQQQLNPCKNLLQCKPVS		
	110	120	130 140 150 160

ALLERG CQVMRQQCCQQLAQIPQQLQCAAIHSIVHSIIMQQEQQEQRQGVQILVPLSQQQVQGQT

	170	180	190	200	210	220
--	-----	-----	-----	-----	-----	-----

>>ALLERGEN\_2012\_2|gi|33327133|gb|AAQ08947.1| allergen Fr (145 aa)  
 initn: 31 initl: 31 opt: 41 Z-score: 87.4 bits: 19.9 E(): 7.5  
 Smith-Waterman score: 41; 34.483% identity (40.000% ungapped) in 29 aa overlap (4-32:79-103)

	10	20	30
SYHT0H	MLTKIFFLSLSRKRLCMNCSPVFTASSVRSSI		
	..	... :	... :
ALLERG	GKVTFTTEVG	YTRAEG	LYSMVIERDHKNEFCEIVLLSSSRK---DCDEIPT
	50	60	70 80 90 100

ALLERG FILNTVNGTTRTINPLGFFKKEVLPKCPQVYNKLGMYPNNM

	110	120	130	140
--	-----	-----	-----	-----

>>ALLERGEN\_2012\_2|gi|124365253|gb|ABN09655.1| beta-1,3-g (374 aa)  
 initn: 45 initl: 45 opt: 45 Z-score: 87.3 bits: 21.3 E(): 7.5  
 Smith-Waterman score: 45; 50.000% identity (50.000% ungapped) in 20 aa overlap (5-24:199-218)

```

              10      20      30
SYHT0H      MLTKIFFLSLSRKRLCMNCSPVFTASSVRSSI
              : ::: :. : : : ::
ALLERG TAIDLTLVGNSYPPSAGAFRDDVRSYLNPIIRFLSSIRSPLLANIYPYFTYAGNPRDISL
      170      180      190      200      210      220

ALLERG PYALFTSPSVVVWDGQRGYKNLFDATLDALYSALERASGGSLVVVSESGWPSAGAFAAAT
      230      240      250      260      270      280

>>ALLERGEN_2012_2|gi|124294785|gb|ABN03966.1| beta-1,3-g (374 aa)
  initn: 45 initl: 45 opt: 45 Z-score: 87.3 bits: 21.3 E(): 7.5
Smith-Waterman score: 45; 50.000% identity (50.000% ungapped) in 20 aa overlap (5-
24:199-218)

```

```

              10      20      30
SYHT0H      MLTKIFFLSLSRKRLCMNCSPVFTASSVRSSI
              : ::: :. : : : ::
ALLERG TAIDLTLVRNSYPPSAGAFRDDVRSYLNPIIRFLSSIRSPLLANIYPYFTYAGNPRDISL
      170      180      190      200      210      220

ALLERG PYALFTSPSVVVWDGQRGYKNLFDATLDALYSALERASGGSLVVVSESGWPSAGAFAAAT
      230      240      250      260      270      280

>>ALLERGEN_2012_2|gi|268037674|gb|ACY91851.1| beta-1,3-g (374 aa)
  initn: 45 initl: 45 opt: 45 Z-score: 87.3 bits: 21.3 E(): 7.5
Smith-Waterman score: 45; 50.000% identity (50.000% ungapped) in 20 aa overlap (5-
24:199-218)

```

```

              10      20      30
SYHT0H      MLTKIFFLSLSRKRLCMNCSPVFTASSVRSSI
              : ::: :. : : : ::
ALLERG TAIDLTLVGNSYPPSAGAFRDDVRSYLNPIIRFLSSIRSPLLANIYPYFTYAGNPRDISL
      170      180      190      200      210      220

ALLERG PYALFTSPSVVVWDGQRGYKNLFDATLDALYSALERASGGSLVVVSESGWPSAGAFAAAT
      230      240      250      260      270      280

>>ALLERGEN_2012_2|gi|170736|gb|AAA34288.1| gamma-gliadin (251 aa)
  initn: 27 initl: 27 opt: 43 Z-score: 86.8 bits: 20.6 E(): 8.1
Smith-Waterman score: 43; 44.444% identity (44.444% ungapped) in 18 aa overlap
(15-32:160-177)

```

```

              10      20      30
SYHT0H      MLTKIFFLSLSRKRLCMNCSPVFTASSVRSSI
              : ..... :... : :
ALLERG PQQPQQSFPQQQPPFIQPSLQQQVNPCKNPLLQCKPVSLVSSLWSMIWPQSDCQVMRQQ
      130      140      150      160      170      180

ALLERG CCQQLAQIPQQLQCAAIHTIIHSIIMQQEQQEQQQGMHILLPLYQQQQVGQGTLVQGQGI
      190      200      210      220      230      240

>>ALLERGEN_2012_2|gi|83300389|sp|O42799.2|ALL7_ASPFU Rec (270 aa)
  initn: 43 initl: 43 opt: 43 Z-score: 86.2 bits: 20.6 E(): 8.7
Smith-Waterman score: 43; 66.667% identity (66.667% ungapped) in 6 aa overlap (16-
21:236-241)

```

```

              10      20      30
SYHT0H      MLTKIFFLSLSRKRLCMNCSPVFTASSVRSSI
              : .....
ALLERG PVGIMTDADCGKTVTITYNGITKTATVVDKCMGCKPTDLASRHLFGELADFSAGRIDGM
      210      220      230      240      250      260

```

ALLERG SWYFN



270

>>ALLERGEN\_2012\_2|gi|170708|gb|AAA34274.1| gamma-gliadin (291 aa)  
initn: 26 initl: 26 opt: 43 Z-score: 85.7 bits: 20.6 E(): 9.3  
Smith-Waterman score: 43; 36.000% identity (36.000% ungapped) in 25 aa overlap (8-32:146-170)

```

                                10      20      30
SYHT0H                        MLTKIFFLSLSRKRLCMNCSPVFTASSVRSSI
                                :.  .. : ..... :... : :
ALLERG PQQPFPPQPQQPQQSFPQQQPSLIQQSLQQQLNPCKNFLQQCKPVSLSVSSLWSIILPPSD
      120      130      140      150      160      170

ALLERG CQVMRQQCCQQLAQIPQQLQCAAIHVSVVHSIIMQQEQEQQLQGVQILVPLSQQQQVQGQI
      180      190      200      210      220      230
```

>>ALLERGEN\_2012\_2|gi|56122438|gb|AAV74343.1| Fra e 1.010 (145 aa)  
initn: 31 initl: 31 opt: 40 Z-score: 85.6 bits: 19.6 E(): 9.4  
Smith-Waterman score: 40; 34.483% identity (40.000% ungapped) in 29 aa overlap (4-32:79-103)

```

                                10      20      30
SYHT0H                        MLTKIFFLSLSRKRLCMNCSPVFTASSVRSSI
                                .:  .: : : : .:  .: .: .: .:
ALLERG GKVTFTEVGYTRAEGLYSMVIERDHKNEFCEIVLLSSSRK---DCHEIPTEGWVKPSLK
      50      60      70      80      90      100

ALLERG FILNTVNGTTRTINPLGFFKKEALPKCPQVYNKLGMYPPNM
      110      120      130      140
```

>>ALLERGEN\_2012\_2|gi|1184668|gb|AAA87456.1| beta-1,3-glu (374 aa)  
initn: 44 initl: 44 opt: 44 Z-score: 85.5 bits: 21.0 E(): 9.4  
Smith-Waterman score: 44; 50.000% identity (50.000% ungapped) in 20 aa overlap (5-24:199-218)

```

                                10      20      30
SYHT0H                        MLTKIFFLSLSRKRLCMNCSPVFTASSVRSSI
                                : : : : .:  : : : :
ALLERG TAIDLTLVGNSYPPSAGAFRDDVRSYLDPIIGFLSSIRSPLLANIYPYFTYAYNPRDISL
      170      180      190      200      210      220

ALLERG PYALFTSPSVVVWDGQRGYKNLFDATLDALYSALERASGGSLVVVSESGWPSAGAFAAAT
      230      240      250      260      270      280
```

>>ALLERGEN\_2012\_2|gi|124365249|gb|ABN09653.1| beta-1,3-g (374 aa)  
initn: 44 initl: 44 opt: 44 Z-score: 85.5 bits: 21.0 E(): 9.4  
Smith-Waterman score: 44; 50.000% identity (50.000% ungapped) in 20 aa overlap (5-24:199-218)

```

                                10      20      30
SYHT0H                        MLTKIFFLSLSRKRLCMNCSPVFTASSVRSSI
                                : : : : .:  : : : :
ALLERG TAIDLTLVGNSYPPSAGAFRDDVRSYLDPIIGFLSSIRSPLLTNIYPYFTYAYNPRDISL
      170      180      190      200      210      220

ALLERG PYALFTSPSVVVWDGQRGYKNLFDATLDALYSALERASGGSLVVVSESGWPSAGAFAAAT
      230      240      250      260      270      280
```

>>ALLERGEN\_2012\_2|gi|124294783|gb|ABN03965.1| beta-1,3-g (374 aa)  
initn: 44 initl: 44 opt: 44 Z-score: 85.5 bits: 21.0 E(): 9.4  
Smith-Waterman score: 44; 50.000% identity (50.000% ungapped) in 20 aa overlap (5-24:199-218)

```

              10      20      30
SYHT0H      MLTKIFFLSLSRKRLCMNCSPVFTASSVRSSI
              : ::: :. : : : ::
ALLERG TAIDLTLVGNSYPPSAGAFRDDVRSYLDPIIGFLSSIRSPLLANIYPYFTYADNPRDISL
      170      180      190      200      210      220

ALLERG PYALFTSPSVVVWDGQRGYKNLFDATLDALYSALERASGGSLVVVSESGWPSAGAFAT
      230      240      250      260      270      280

>>ALLERGEN_2012_2|gi|32765543|gb|AAP87281.1| beta-1,3-gl (374 aa)
  initn: 44 initl: 44 opt: 44 Z-score: 85.5 bits: 21.0 E(): 9.4
Smith-Waterman score: 44; 50.000% identity (50.000% ungapped) in 20 aa overlap (5-
24:199-218)

```

```

              10      20      30
SYHT0H      MLTKIFFLSLSRKRLCMNCSPVFTASSVRSSI
              : ::: :. : : : ::
ALLERG TAIDLTLVGNSYPPSAGAFRDDVRSYLDPIIGFLSSIRSPLLANIYPYFTYAGNPRDISL
      170      180      190      200      210      220

ALLERG PYALFTSPSVVVWDGQRGYKNLFDATLDALYSALERASGGSLVVVSESGWPSAGAFAT
      230      240      250      260      270      280

>>ALLERGEN_2012_2|gi|270315180|gb|ACZ74626.1| beta-1,3-g (374 aa)
  initn: 44 initl: 44 opt: 44 Z-score: 85.5 bits: 21.0 E(): 9.4
Smith-Waterman score: 44; 50.000% identity (50.000% ungapped) in 20 aa overlap (5-
24:199-218)

```

```

              10      20      30
SYHT0H      MLTKIFFLSLSRKRLCMNCSPVFTASSVRSSI
              : ::: :. : : : ::
ALLERG TAIDLTLMGNTYPPSAGAFRDDVRSYLDPIIGFLSSIRSPLLANIYPYFTYAGNPRDISL
      170      180      190      200      210      220

ALLERG PYALFTSPSVVVWDGQRGYKNLFDATLDALYSALERASGGSLVVVSESGWPSAGAFAT
      230      240      250      260      270      280

>>ALLERGEN_2012_2|gi|124365251|gb|ABN09654.1| beta-1,3-g (374 aa)
  initn: 44 initl: 44 opt: 44 Z-score: 85.5 bits: 21.0 E(): 9.4
Smith-Waterman score: 44; 50.000% identity (50.000% ungapped) in 20 aa overlap (5-
24:199-218)

```

```

              10      20      30
SYHT0H      MLTKIFFLSLSRKRLCMNCSPVFTASSVRSSI
              : ::: :. : : : ::
ALLERG TAIDLTLVGNSYPPSAGAFRDDVRSYLDPIIGFLSSXSPLLANIYPYFTYAYNPRDISL
      170      180      190      200      210      220

ALLERG PYALFTSPSVVVWDGQRGYKNLFDATLDALYSALERASGGSLVVVSESGWPSAGAFAT
      230      240      250      260      270      280

```

```

32 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.010 Display time: 0.000

```

```

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:

```

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

```
/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 41 aa
>SYHT0H2_I_50
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library
```

367796 residues in 1603 sequences

Expectation\_n fit:  $\rho(\ln(x)) = 3.8720 \pm 0.00253$ ;  $\mu = 3.7082 \pm 0.132$   
mean\_var=25.4268  $\pm$  6.826, 0's: 15 Z-trim: 18 B-trim: 0 in 0/42  
Lambda= 0.2543

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2  
join: 36, opt: 24, gap-pen: -12/-2, width: 16  
Scan time: 0.010

The best scores are:

					opt	bits	E(1603)
ALLERGEN_2012_2	gi 22090 emb CAA43361.1	HMW glut	( 705)	61	27	0.27	
ALLERGEN_2012_2	gi 21779 emb CAA26847.1	unnamed	( 660)	60	27	0.33	
ALLERGEN_2012_2	gi 21751 emb CAA31396.1	high mol	( 648)	59	27	0.41	
ALLERGEN_2012_2	gi 14423933 sp O82803.1	SRPP_HEVB	( 204)	47	22	2.8	
ALLERGEN_2012_2	gi 4590366 gb AAD26547.1	AF124824	( 159)	44	21	4.7	
ALLERGEN_2012_2	gi 25361513 gb AAN73248.1	helix-	( 450)	48	23	4.7	
ALLERGEN_2012_2	gi 218203828 gb ACK76297.1	Der f	( 279)	46	22	4.9	
ALLERGEN_2012_2	gi 218203826 gb ACK76296.1	Der f	( 279)	46	22	4.9	
ALLERGEN_2012_2	gi 170743 gb AAB02788.1	HMW glut	( 815)	50	23	5.1	
ALLERGEN_2012_2	gi 238477265 gb ACR43474.1	argin	( 356)	46	22	6.2	
ALLERGEN_2012_2	gi 897811 emb CAA24933.1	unnamed	( 101)	41	20	6.4	
ALLERGEN_2012_2	gi 219815476 gb ACL36923.1	tropo	( 153)	41	20	9.7	

```
>>ALLERGEN_2012_2|gi|22090|emb|CAA43361.1| HMW glutenin (705 aa)
initn: 37 init1: 37 opt: 61 Z-score: 113.3 bits: 27.4 E(): 0.27
Smith-Waterman score: 61; 47.059% identity (47.059% ungapped) in 17 aa overlap
(22-38:117-133)
```

```

              10      20      30      40
SYHT0H      MMAFVGATFLFHYLHNKVTDSWAMESEEVSGYYPLLKSHLC
              . . . . .
ALLERG QYEQTVPVPKGGSFYPGETTPLQQLQQVIFWGTSSQTVQGYYPSSVSPQQGPYPGQASP
      90      100      110      120      130      140

ALLERG QQPGQGQQPGKWQELGQGQQGYPTSLHQSGQGQQGYYPSSLQQPGQGQQIGQGQQGYYP
      150      160      170      180      190      200
```

```
>>ALLERGEN_2012_2|gi|21779|emb|CAA26847.1| unnamed prote (660 aa)
initn: 36 init1: 36 opt: 60 Z-score: 111.8 bits: 27.0 E(): 0.33
Smith-Waterman score: 60; 47.059% identity (47.059% ungapped) in 17 aa overlap
(22-38:117-133)
```

```

              10      20      30      40
SYHT0H      MMAFVGATFLFHYLHNKVTDSWAMESEEVSGYYPLLKSHLC
              . . . . .
ALLERG QYEQTVPVPKGGSFYPGETTPLQQLQQGIFWGTSSQTVQGYYPSTSPRQGSYYPGQASP
      90      100      110      120      130      140

ALLERG QQPGQGQQPGKWQEPGQGQQWYYPTSLQQPGQGQQIGKGKQGYPTSLQQPGQGQQIGQG
      150      160      170      180      190      200
```

```
>>ALLERGEN_2012_2|gi|21751|emb|CAA31396.1| high molecula (648 aa)
initn: 35 init1: 35 opt: 59 Z-score: 109.9 bits: 26.6 E(): 0.41
Smith-Waterman score: 59; 47.059% identity (47.059% ungapped) in 17 aa overlap
(22-38:117-133)
```





```

      10      20      30      40
SYHT0H MMAFVGATFLFHYLHNKVTDSWAMESEEVSGYYPLLKSHLC
      : . . . : : :
ALLERG      EKLGGQQPRQWLQPRQGGQGYPTSPQQSGQGQQLGQGQQGYPTSPQQ
      10      20      30      40      50

ALLERG SGQGQQGYDSPYHVSAEHQAASLKVAKAQQLAAQLPAMCRLEGGDALLASQ
      60      70      80      90      100

>>ALLERGEN_2012_2|gi|219815476|gb|ACL36923.1| troponin C (153 aa)
  initn: 25 init1: 25 opt: 41 Z-score: 85.3 bits: 20.0 E(): 9.7
Smith-Waterman score: 41; 23.810% identity (23.810% ungapped) in 21 aa overlap
(14-34:86-106)

```

```

      10      20      30      40
SYHT0H      MMAFVGATFLFHYLHNKVTDSWAMESEEVSGYYPLLKSHLC
      ..... : : : : :
ALLERG EIDADGSGELEFDEFLTTLTARFLVEEDTEAMQEELREAFRMYDKENGXIPTSALREILR
      60      70      80      90      100      110

ALLERG ALDDKLTDELDemiaEIDTDGSGTVDFDEFMEMMTGD
      120      130      140      150

```

```

41 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.010 Display time: 0.010

```

```

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 40 aa
>SYHT0H2_I_51
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

```

```

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 3.9573+/-0.00339; mu= 4.1111+/- 0.177
mean_var=42.0783+/-10.208, 0's: 12 Z-trim: 12 B-trim: 68 in 2/40
Lambda= 0.1977

```

```

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.010

```

```

The best scores are:
                                opt bits E(1603)
ALLERGEN_2012_2|gi|3287877|sp|P81402.1|NLTP1_PRUP ( 91) 48 20 5.2
ALLERGEN_2012_2|gi|83754241|pdb|2B5S|B Chain B, C ( 92) 48 20 5.3
ALLERGEN_2012_2|gi|1311510|gb|AAB36009.1| mAb 8C7 ( 15) 40 17 6.2
ALLERGEN_2012_2|gi|2506771|sp|P16968.2|IAA1_HORVU ( 146) 49 21 6.2
ALLERGEN_2012_2|gi|313575718|gb|ADR66939.1| non-s ( 117) 48 20 6.4
ALLERGEN_2012_2|gi|288561913|sp|P85894.1|LTP1_MOR ( 91) 47 20 6.4
ALLERGEN_2012_2|gi|5777414|emb|CAB53458.1| MnSOD ( 205) 50 21 6.6
ALLERGEN_2012_2|gi|10862818|emb|CAC13961.1| IgE-b ( 205) 50 21 6.6
ALLERGEN_2012_2|gi|348137|gb|AAA16792.1| superoxi ( 233) 50 21 7.3
ALLERGEN_2012_2|gi|313575726|gb|ADR66943.1| non-s ( 117) 47 20 7.8
ALLERGEN_2012_2|gi|23894244|emb|CAD23614.1| tri m ( 404) 51 21 9.3
ALLERGEN_2012_2|gi|74663809|sp|Q8J077.1|SUB6_TRIS ( 405) 51 21 9.3

```

ALLERGEN\_2012\_2|gi|9929163|emb|CAC05258.1| Cup a ( 199) 48 20 9.6

>>ALLERGEN\_2012\_2|gi|3287877|sp|P81402.1|NLTP1\_PRUPE Rec (91 aa)  
 initn: 44 initl: 44 opt: 48 Z-score: 90.1 bits: 20.1 E(): 5.2  
 Smith-Waterman score: 48; 37.037% identity (41.667% ungapped) in 27 aa overlap (9-35:4-27)

```

          10          20          30          40
SYHT0H MVNDITNRGHISCCSWPNLNWSPLRRHGGLVISACPLQMK
      :..: : . : : : : :
ALLERG ITCGQVSSALAPCI---PYVRGGGAVPPACNGIRNVNNLARTTPDRQAACNCLK
          10          20          30          40          50

ALLERG QLSASVPGVNPNNAAALPGKCGVHIPYKISASTNCATVK
          60          70          80          90

```

>>ALLERGEN\_2012\_2|gi|83754241|pdb|2B5S|B Chain B, Crysta (92 aa)  
 initn: 44 initl: 44 opt: 48 Z-score: 90.1 bits: 20.1 E(): 5.3  
 Smith-Waterman score: 48; 37.037% identity (41.667% ungapped) in 27 aa overlap (9-35:5-28)

```

          10          20          30          40
SYHT0H MVNDITNRGHISCCSWPNLNWSPLRRHGGLVISACPLQMK
      :..: : . : : : : :
ALLERG MITCGQVSSSLAPCI---PYVRGGGAVPPACNGIRNVNNLARTTPDRQAACNCLK
          10          20          30          40          50

ALLERG QLSASVPGVNPNNAAALPGKCGVSIPYKISASTNCATVK
          60          70          80          90

```

>>ALLERGEN\_2012\_2|gi|1311510|gb|AAB36009.1| mAb 8C7-reac (15 aa)  
 initn: 40 initl: 40 opt: 40 Z-score: 88.8 bits: 17.2 E(): 6.2  
 Smith-Waterman score: 40; 55.556% identity (55.556% ungapped) in 9 aa overlap (28-36:4-12)

```

          10          20          30          40
SYHT0H MVNDITNRGHISCCSWPNLNWSPLRRHGGLVISACPLQMK
      : : : : : :
ALLERG APAGGVVVAAMPPL
          10

```

>>ALLERGEN\_2012\_2|gi|2506771|sp|P16968.2|IAA1\_HORVU RecN (146 aa)  
 initn: 30 initl: 30 opt: 49 Z-score: 88.8 bits: 20.5 E(): 6.2  
 Smith-Waterman score: 49; 37.500% identity (42.857% ungapped) in 24 aa overlap (14-35:20-42)

```

          10          20          30          40
SYHT0H MVNDITNRGHISCCSWPNLNWS--PLRRHGGLVISACPLQMK
      : : : : : :
ALLERG PTSVAVDQGSMSVNSPGEWC-WPGMGYPVYFPFRCRALVKSQCAGGQVVESIQKDCCRQI
          10          20          30          40          50

ALLERG AAIGDEWCICGALGSMRGS MYKELGVALADDKATVAEVFPGCRTEVMDRAVASLPVCNQ
          60          70          80          90          100          110

```

>>ALLERGEN\_2012\_2|gi|313575718|gb|ADR66939.1| non-specif (117 aa)  
 initn: 44 initl: 44 opt: 48 Z-score: 88.6 bits: 20.2 E(): 6.4  
 Smith-Waterman score: 48; 37.037% identity (41.667% ungapped) in 27 aa overlap (9-35:30-53)

```

              10      20      30
SYHT0H      MVNDITNRGHISCCSWPNLNWSPLRRHGGLVISACPLQM
              :..:   : .   : : : : : :
ALLERG MAYSAMTKLALVVALCMVSVPIAQAITCGQVSSSLAPCI---PYVRGGGAVPPACCNIGI
              10      20      30      40      50

              40
SYHT0H K

ALLERG RNVNNLARTTPDRQAACNCLKQLSASVPGVNPNNAAALPGKCGVSIPYKISASTNCATVK
              60      70      80      90      100     110

>>ALLERGEN_2012_2|gi|288561913|sp|P85894.1|LTP1_MORNI Re (91 aa)
  initn: 32 initl: 32 opt: 47 Z-score: 88.6 bits: 19.8 E(): 6.4
Smith-Waterman score: 47; 33.333% identity (37.500% ungapped) in 27 aa overlap (9-
35:4-27)

              10      20      30      40
SYHT0H MVNDITNRGHISCCSWPNLNWSPLRRHGGLVISACPLQMK
              :..:   : .:   : : : : . :
ALLERG      ITCGQVSSSLAPCINYL---RAGGVVPANCCNGVRSLNNAAKTTADRQAACNCLK
              10      20      30      40      50

ALLERG SAFNSIKGLNLAAGLPGKCGVSPYKISPSTDCKSVK
              60      70      80      90

>>ALLERGEN_2012_2|gi|5777414|emb|CAB53458.1| MnSOD [Heve (205 aa)
  initn: 38 initl: 38 opt: 50 Z-score: 88.3 bits: 20.9 E(): 6.6
Smith-Waterman score: 50; 40.909% identity (47.368% ungapped) in 22 aa overlap (9-
29:71-90)

              10      20      30
SYHT0H      MVNDITNRGHIS-CCSWPNLNWSPLRRHGGLVISACPL
              :..:   : :   : : : : : :
ALLERG KALEQLNDAIEKGDSAADVVKLQSAIKFNGGGHVNHSIFWKNL--APVREGGGELPHGSLG
              50      60      70      80      90

              40
SYHT0H QMK

ALLERG WAIDADFGSLEKLIQLMNAEGVALQSGGWVWLALDKELKKLVVETTANQDPLVTKGPTLV
              100     110     120     130     140     150

>>ALLERGEN_2012_2|gi|10862818|emb|CAC13961.1| IgE-bindin (205 aa)
  initn: 38 initl: 38 opt: 50 Z-score: 88.3 bits: 20.9 E(): 6.6
Smith-Waterman score: 50; 40.909% identity (47.368% ungapped) in 22 aa overlap (9-
29:71-90)

              10      20      30
SYHT0H      MVNDITNRGHIS-CCSWPNLNWSPLRRHGGLVISACPL
              :..:   : :   : : : : : :
ALLERG KALEQLNDAIEKGDSAADVVKLQSAIKFNGGGHVNHSIFWKNL--APVREGGGELPHGSLG
              50      60      70      80      90

              40
SYHT0H QMK

ALLERG WAIDADFGSLEKLIQLMNAEGAALRGSGGWVWLALDKELKKLVVETTANQDPLVTKGPTLV
              100     110     120     130     140     150

>>ALLERGEN_2012_2|gi|348137|gb|AAA16792.1| superoxide di (233 aa)
  initn: 38 initl: 38 opt: 50 Z-score: 87.5 bits: 21.0 E(): 7.3

```



Smith-Waterman score: 50; 40.909% identity (47.368% ungapped) in 22 aa overlap (9-29:99-118)

```

              10      20      30
SYHT0H          MVNDITNRGHIS-CCSWPNLNWSPLRRHGGLVISACPL
              :... : : :... :
ALLERG KALEQLNDAIEKGDAAVVKLQSAIKFNGGGHVNHSIFWKNL--APVREGGGELPHGSLG
              70      80      90      100      110      120
```

```

              40
SYHT0H QMK
ALLERG WAIDADFGSLEKLIQLMNAEGAALQSGWVWLALDKELKKLVVETTANQDPLVTKGPTLV
              130      140      150      160      170      180
```

>>ALLERGEN\_2012\_2|gi|313575726|gb|ADR66943.1| non-specif (117 aa)  
initn: 44 initl: 44 opt: 47 Z-score: 87.1 bits: 19.9 E(): 7.8  
Smith-Waterman score: 47; 37.037% identity (41.667% ungapped) in 27 aa overlap (9-35:30-53)

```

              10      20      30
SYHT0H          MVNDITNRGHISCCSWPNLNWSPLRRHGGLVISACPLQM
              :... : . : : : :
ALLERG MACSAMTKLALVVALCMVSVPIAQALTCGQVSSNLAPCI---PYVRGGGAVPPACNGI
              10      20      30      40      50
```

```

              40
SYHT0H K
ALLERG RNINNLAKTADTADRTACNCLKQLSASVPGVNNANNAALPGKCGVNVVPYKISPSTNCATVK
              60      70      80      90      100      110
```

>>ALLERGEN\_2012\_2|gi|23894244|emb|CAD23614.1| tri m 2 al (404 aa)  
initn: 35 initl: 35 opt: 51 Z-score: 85.7 bits: 21.4 E(): 9.3  
Smith-Waterman score: 51; 30.000% identity (32.143% ungapped) in 30 aa overlap (1-28:39-68)

```

              10      20
SYHT0H          MVNDITNRGHISCCSW--PNLNWSPLRRHG
              : ..... . . : : : :
ALLERG VLAALSAVNGAKILEAGPHAETIPNKYIVVMKKDVSDEAFSTHTTWLSQNLNRRMLMRRSG
              10      20      30      40      50      60
```

```

              30      40
SYHT0H GLVISACPLQMK
              .
ALLERG SSKAMAGMQNKYSLGGIFRAYSGEFDDAMIKDISNHDDVDYIEPDFVVRTSTNGTNLTRQ
              70      80      90      100      110      120
```

>>ALLERGEN\_2012\_2|gi|74663809|sp|Q8J077.1|SUB6\_TRISH Rec (405 aa)  
initn: 35 initl: 35 opt: 51 Z-score: 85.7 bits: 21.4 E(): 9.3  
Smith-Waterman score: 51; 30.000% identity (32.143% ungapped) in 30 aa overlap (1-28:39-68)

```

              10      20
SYHT0H          MVNDITNRGHISCCSW--PNLNWSPLRRHG
              : ..... . . : : : :
ALLERG VLAALSAVNGAKILEAGPHAETIPNKYIVVMKKDVSDEAFSTHTTWLSQNLNRRMLMRRSG
              10      20      30      40      50      60
```

```

      30      40
SYHT0H GLVISACPLQMK
.
ALLERG SSKAMAGMQNKYSLGGIFRAYSGEFDDAMIKDISNHDDVDYIEPDFVVRTSTNGTNLTRQ
      70      80      90     100     110     120

>>ALLERGEN_2012_2|gi|9929163|emb|CAC05258.1| Cup a 3 pro (199 aa)
  initn: 48 initl: 48 opt: 48 Z-score: 85.4 bits: 20.3 E(): 9.6
Smith-Waterman score: 48; 25.000% identity (25.000% ungapped) in 24 aa overlap
(13-36:149-172)

```

```

              10      20      30      40
SYHT0H          MVNDITNRGHISCCSWPNLNWSPLRRHGGGLVISACPLQMK
              ::  .:  :  ..  .  ::
ALLERG KADINAVCPSELKVDGGCNSACNVLQTDQYCCRNAYVNNCPATNYSKIFKNQCPQAYSYA
      120      130      140      150      160      170

ALLERG KDDTATFACASGTDYSIVFCP
      180      190

```

```

40 residues in 1 query sequences
367796 residues in 1603 library sequences
Scomplib [34t11]
start: Mon Feb 27 15:13:32 2012 done: Mon Feb 27 15:13:32 2012
Scan time: 0.010 Display time: 0.010

```

```

Function used was FASTA [version 3.4t11 Apr 17, 2002]
FASTA searches a protein or DNA sequence data bank
version 3.4t11 Apr 17, 2002
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

```

```

/grid/.qsub_tmp/1330373600kde_agent/grid_NbCuMBoIO: 57 aa
>SYHT0H2_I_52
vs /data/fasta/ALLERGEN_2012_2 library
searching /data/fasta/ALLERGEN_2012_2 library

```

```

367796 residues in 1603 sequences
Expectation_n fit: rho(ln(x))= 3.5776+/-0.00299; mu= 8.1422+/- 0.157
mean_var=41.7344+/-11.088, 0's: 12 Z-trim: 13 B-trim: 0 in 0/43
Lambda= 0.1985

```

```

FASTA (3.45 Mar 2002) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 24, gap-pen: -12/-2, width: 16
Scan time: 0.030

```

```

The best scores are:
                                opt bits E(1603)
ALLERGEN_2012_2|gi|6136162|sp|P35776.2|VA2_SOLRI ( 119) 60 24 0.91
ALLERGEN_2012_2|gi|2497701|sp|Q28133.1|ALL2_BOVIN ( 172) 55 22 3.2
ALLERGEN_2012_2|gi|886965|emb|CAA59339.1| low mol ( 261) 56 23 3.5
ALLERGEN_2012_2|gi|94400907|ref|NP_001035360.1| a ( 92) 52 21 3.7
ALLERGEN_2012_2|gi|88770352|gb|ABD51779.1| allerg ( 94) 52 21 3.8
ALLERGEN_2012_2|gi|170730|gb|AAA34285.1| gamma-gl ( 304) 55 22 4.8
ALLERGEN_2012_2|gi|194350817|gb|ACF53837.1| Bla g ( 190) 50 21 9.2
ALLERGEN_2012_2|gi|886967|emb|CAA59340.1| low mol ( 276) 51 21 9.8

```

```

>>ALLERGEN_2012_2|gi|6136162|sp|P35776.2|VA2_SOLRI RecNa (119 aa)
  initn: 60 initl: 60 opt: 60 Z-score: 103.8 bits: 23.5 E(): 0.91
Smith-Waterman score: 60; 36.000% identity (36.000% ungapped) in 25 aa overlap (7-
31:12-36)

```





50  
SYHT0H TLRHLRLKQMCVQCGS  
: . . . . :  
ALLERG --SCHVMQQQCCQQLPRIPEQSRYDAIRAIISIVLQEQQHGGFNQPQQQQPQQSVQGV  
200 210 220 230 240 250

## APPENDIX C BLAST Toxin Search Alignments

Query	Subject ID	Description
SYHT0H2_I_22	TOXIN_2012_2 gi 90406967 ref ZP_01215158.1	putative RTX toxin [Psychromonas sp. CNPT3] gi 90312009 gb EAS40103.1  putative RTX toxin [Psychromonas sp. CNPT3]
SYHT0H2_I_4	TOXIN_2012_2 gi 157951698 ref NP_033910.2	complement C4-B precursor [Mus musculus] gi 341940526 sp P01029.3 CO4B_MOUSE RecName: Full=Complement C4-B; Contains: RecName: Full=Complement C4 beta chain; Contains: RecName: Full=Complement C4 alpha chain; Contains: RecName: Full=C4a anaphylatoxin; Contains: RecName: Full=Complement C4 gamma chain; Flags: Precursor
SYHT0H2_I_32	TOXIN_2012_2 gi 157951698 ref NP_033910.2	complement C4-B precursor [Mus musculus] gi 341940526 sp P01029.3 CO4B_MOUSE RecName: Full=Complement C4-B; Contains: RecName: Full=Complement C4 beta chain; Contains: RecName: Full=Complement C4 alpha chain; Contains: RecName: Full=C4a anaphylatoxin; Contains: RecName: Full=Complement C4 gamma chain; Flags: Precursor
SYHT0H2_I_30	TOXIN_2012_2 gi 282856127 ref ZP_06265411.1	addiction module toxin, RelE/StbE [Pyramidobacter piscicola W5455] gi 282586054 gb EFB91338.1  addiction module toxin, RelE/StbE [Pyramidobacter piscicola W5455]
SYHT0H2_I_15	TOXIN_2012_2 gi 71281801 ref YP_268598.1	perfringolysin O regulator protein [Colwellia psychrerythraea 34H] gi 71147541 gb AAZ28014.1  perfringolysin O regulator protein [Colwellia psychrerythraea 34H]
SYHT0H2_I_10	TOXIN_2012_2 gi 333979976 ref YP_004517921.1	beta-lysine N-acetyltransferase [Desulfotomaculum kuznetsovii DSM 6115] gi 333823457 gb AEG16120.1  putative beta-lysine N-acetyltransferase [Desulfotomaculum kuznetsovii DSM 6115]
SYHT0H2_I_16	TOXIN_2012_2 gi 333979976 ref YP_004517921.1	beta-lysine N-acetyltransferase [Desulfotomaculum kuznetsovii DSM 6115] gi 333823457 gb AEG16120.1  putative beta-lysine N-acetyltransferase [Desulfotomaculum kuznetsovii DSM 6115]
SYHT0H2_I_28	TOXIN_2012_2 gi 333979976 ref YP_004517921.1	beta-lysine N-acetyltransferase [Desulfotomaculum kuznetsovii DSM 6115] gi 333823457 gb AEG16120.1  putative beta-lysine N-acetyltransferase [Desulfotomaculum kuznetsovii DSM 6115]
SYHT0H2_I_38	TOXIN_2012_2 gi 333979976 ref YP_004517921.1	beta-lysine N-acetyltransferase [Desulfotomaculum kuznetsovii DSM 6115] gi 333823457 gb AEG16120.1  putative beta-lysine N-acetyltransferase [Desulfotomaculum kuznetsovii DSM 6115]

Query	Subject ID	Description
SYHT0H2_I_10	TOXIN_2012_2 gi 357038678 ref ZP_09100475.1	putative beta-lysine N-acetyltransferase [Desulfotomaculum gibsoniae DSM 7213] gi 355359470 gb EHG07232.1  putative beta-lysine N-acetyltransferase [Desulfotomaculum gibsoniae DSM 7213]
SYHT0H2_I_11	TOXIN_2012_2 gi 302566051 pdb 3I5V A	Chain A, Crystal Structure Of Beta Toxin 275-280 From Staphylococcus Aureus gi 302566052 pdb 3I5V B Chain B, Crystal Structure Of Beta Toxin 275-280 From Staphylococcus Aureus gi 302566053 pdb 3I5V C Chain C, Crystal Structure Of Beta Toxin 275-280 From Staphylococcus Aureus gi 302566054 pdb 3I5V D Chain D, Crystal Structure Of Beta Toxin 275-280 From Staphylococcus Aureus
SYHT0H2_I_16	TOXIN_2012_2 gi 357038678 ref ZP_09100475.1	putative beta-lysine N-acetyltransferase [Desulfotomaculum gibsoniae DSM 7213] gi 355359470 gb EHG07232.1  putative beta-lysine N-acetyltransferase [Desulfotomaculum gibsoniae DSM 7213]
SYHT0H2_I_28	TOXIN_2012_2 gi 357038678 ref ZP_09100475.1	putative beta-lysine N-acetyltransferase [Desulfotomaculum gibsoniae DSM 7213] gi 355359470 gb EHG07232.1  putative beta-lysine N-acetyltransferase [Desulfotomaculum gibsoniae DSM 7213]
SYHT0H2_I_38	TOXIN_2012_2 gi 357038678 ref ZP_09100475.1	putative beta-lysine N-acetyltransferase [Desulfotomaculum gibsoniae DSM 7213] gi 355359470 gb EHG07232.1  putative beta-lysine N-acetyltransferase [Desulfotomaculum gibsoniae DSM 7213]
SYHT0H2_I_42	TOXIN_2012_2 gi 302566051 pdb 3I5V A	Chain A, Crystal Structure Of Beta Toxin 275-280 From Staphylococcus Aureus gi 302566052 pdb 3I5V B Chain B, Crystal Structure Of Beta Toxin 275-280 From Staphylococcus Aureus gi 302566053 pdb 3I5V C Chain C, Crystal Structure Of Beta Toxin 275-280 From Staphylococcus Aureus gi 302566054 pdb 3I5V D Chain D, Crystal Structure Of Beta Toxin 275-280 From Staphylococcus Aureus
SYHT0H2_I_11	TOXIN_2012_2 gi 300508421 pdb 3I41 A	Chain A, Crystal Structure Of Beta Toxin From Staphylococcus Aureus F277a, P278a Mutant gi 300508422 pdb 3I41 B Chain B, Crystal Structure Of Beta Toxin From Staphylococcus Aureus F277a, P278a Mutant gi 300508423 pdb 3I46 A Chain A, Crystal Structure Of Beta Toxin From Staphylococcus Aureus F277a, P278a Mutant With Bound Calcium Ions gi 300508424 pdb 3I46 B Chain B, Crystal Structure Of Beta Toxin From Staphylococcus Aureus F277a, P278a Mutant With Bound Calcium Ions

Query	Subject ID	Description
		gi 300508425 pdb 3I48 A Chain A, Crystal Structure Of Beta Toxin From Staphylococcus Aureus F277a, P278a Mutant With Bound Magnesium Ions gi 300508426 pdb 3I48 B Chain B, Crystal Structure Of Beta Toxin From Staphylococcus Aureus F277a, P278a Mutant With Bound Magnesium Ions
SYHT0H2_I_42	TOXIN_2012_2 gi 300508421 pdb 3I41 A	Chain A, Crystal Structure Of Beta Toxin From Staphylococcus Aureus F277a, P278a Mutant gi 300508422 pdb 3I41 B Chain B, Crystal Structure Of Beta Toxin From Staphylococcus Aureus F277a, P278a Mutant gi 300508423 pdb 3I46 A Chain A, Crystal Structure Of Beta Toxin From Staphylococcus Aureus F277a, P278a Mutant With Bound Calcium Ions gi 300508424 pdb 3I46 B Chain B, Crystal Structure Of Beta Toxin From Staphylococcus Aureus F277a, P278a Mutant With Bound Calcium Ions gi 300508425 pdb 3I48 A Chain A, Crystal Structure Of Beta Toxin From Staphylococcus Aureus F277a, P278a Mutant With Bound Magnesium Ions gi 300508426 pdb 3I48 B Chain B, Crystal Structure Of Beta Toxin From Staphylococcus Aureus F277a, P278a Mutant With Bound Magnesium Ions
SYHT0H2_I_11	TOXIN_2012_2 gi 320089760 pdb 3K55 A	Chain A, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089761 pdb 3K55 B Chain B, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089762 pdb 3K55 C Chain C, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089763 pdb 3K55 D Chain D, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089764 pdb 3K55 E Chain E, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089765 pdb 3K55 F Chain F, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089766 pdb 3K55 G Chain G, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089767 pdb 3K55 H Chain H, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089768 pdb 3K55 I Chain I, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089769 pdb 3K55 J Chain J, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus



Query	Subject ID	Description
		gi 320089770 pdb 3K55 K Chain K, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089771 pdb 3K55 L Chain L, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089772 pdb 3K55 M Chain M, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089773 pdb 3K55 N Chain N, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089774 pdb 3K55 O Chain O, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089775 pdb 3K55 P Chain P, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus
SYHT0H2_I_42	TOXIN_2012_2 gi 320089760 pdb 3K55 A	Chain A, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089761 pdb 3K55 B Chain B, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089762 pdb 3K55 C Chain C, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089763 pdb 3K55 D Chain D, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089764 pdb 3K55 E Chain E, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089765 pdb 3K55 F Chain F, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089766 pdb 3K55 G Chain G, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089767 pdb 3K55 H Chain H, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089768 pdb 3K55 I Chain I, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089769 pdb 3K55 J Chain J, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089770 pdb 3K55 K Chain K, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089771 pdb 3K55 L Chain L, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089772 pdb 3K55 M Chain M, Structure Of Beta Hairpin Deletion Mutant Of Beta

Query	Subject ID	Description
		Toxin From Staphylococcus Aureus gi 320089773 pdb 3K55 N Chain N, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089774 pdb 3K55 O Chain O, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus gi 320089775 pdb 3K55 P Chain P, Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin From Staphylococcus Aureus
SYHT0H2_I_11	TOXIN_2012_2 gi 345014741  ref YP_004817095.1	Ricin B lectin [Streptomyces violaceusniger Tu 4113] gi 344041090 gb AEM86815.1  Ricin B lectin [Streptomyces violaceusniger Tu 4113]
SYHT0H2_I_42	TOXIN_2012_2 gi 345014741  ref YP_004817095.1	Ricin B lectin [Streptomyces violaceusniger Tu 4113] gi 344041090 gb AEM86815.1  Ricin B lectin [Streptomyces violaceusniger Tu 4113]
SYHT0H2_I_11	TOXIN_2012_2 gi 57650692  ref YP_186826.1	phospholipase C [Staphylococcus aureus subsp. aureus COL] gi 282894608 ref ZP_06302836.1  sphingomyelin phosphodiesterase [Staphylococcus aureus A8117] gi 130085 sp P09978.1 PHLC_STAAU RecName: Full=Phospholipase C; AltName: Full=Beta-hemolysin; AltName: Full=Beta- toxin; AltName: Full=Sphingomyelinase; Short=SMase; Flags: Precursor gi 61213889 sp Q5HEI1.1 PHLC_STAAC RecName: Full=Phospholipase C; AltName: Full=Beta-hemolysin; AltName: Full=Beta- toxin; AltName: Full=Sphingomyelinase; Short=SMase; Flags: Precursor gi 110815915 sp Q2FWP1.2 PHLC_STAA8 RecName: Full=Phospholipase C; AltName: Full=Beta-hemolysin; AltName: Full=Beta- toxin; AltName: Full=Sphingomyelinase; Short=SMase; Flags: Precursor gi 46588 emb CAA31769.1  unnamed protein product [Staphylococcus aureus subsp. aureus COL] gi 57284878 gb AAW36972.1  phospholipase C [Staphylococcus aureus subsp. aureus COL] gi 269941409 emb CBI49806.1  phospholipase C precursor (pseudogene) [Staphylococcus aureus subsp. aureus TW20] gi 282763095 gb EFC03227.1  sphingomyelin phosphodiesterase [Staphylococcus aureus A8117] gi 298695279 gb ADI98501.1  beta- hemolysin [Staphylococcus aureus subsp. aureus ED133] gi 329731457 gb EGG67820.1  sphingomyelin phosphodiesterase [Staphylococcus aureus subsp. aureus 21193] gi 365168342 gb EHM59689.1  sphingomyelin phosphodiesterase

Query	Subject ID	Description
		[Staphylococcus aureus subsp. aureus 21178] gi 375025481 gb EHS18885.1  sphingomyelin phosphodiesterase [Staphylococcus aureus subsp. aureus IS-91]
SYHT0H2_I_42	TOXIN_2012_2 gi 57650692 ref YP_186826.1	phospholipase C [Staphylococcus aureus subsp. aureus COL] gi 282894608 ref ZP_06302836.1  sphingomyelin phosphodiesterase [Staphylococcus aureus A8117] gi 130085 sp P09978.1 PHLC_STAAU RecName: Full=Phospholipase C; AltName: Full=Beta-hemolysin; AltName: Full=Beta-toxin; AltName: Full=Sphingomyelinase; Short=SMase; Flags: Precursor gi 61213889 sp Q5HEI1.1 PHLC_STAAC RecName: Full=Phospholipase C; AltName: Full=Beta-hemolysin; AltName: Full=Beta-toxin; AltName: Full=Sphingomyelinase; Short=SMase; Flags: Precursor gi 110815915 sp Q2FWP1.2 PHLC_STAA8 RecName: Full=Phospholipase C; AltName: Full=Beta-hemolysin; AltName: Full=Beta-toxin; AltName: Full=Sphingomyelinase; Short=SMase; Flags: Precursor gi 46588 emb CAA31769.1  unnamed protein product [Staphylococcus aureus subsp. aureus COL] gi 57284878 gb AAW36972.1  phospholipase C [Staphylococcus aureus subsp. aureus COL] gi 269941409 emb CBI49806.1  phospholipase C precursor (pseudogene) [Staphylococcus aureus subsp. aureus TW20] gi 282763095 gb EFC03227.1  sphingomyelin phosphodiesterase [Staphylococcus aureus A8117] gi 298695279 gb ADI98501.1  beta-hemolysin [Staphylococcus aureus subsp. aureus ED133] gi 329731457 gb EGG67820.1  sphingomyelin phosphodiesterase [Staphylococcus aureus subsp. aureus 21193] gi 365168342 gb EHM59689.1  sphingomyelin phosphodiesterase [Staphylococcus aureus subsp. aureus 21178] gi 375025481 gb EHS18885.1  sphingomyelin phosphodiesterase [Staphylococcus aureus subsp. aureus IS-91]
SYHT0H2_I_10	TOXIN_2012_2 gi 269796843 ref YP_003316298.1	Pertussis toxin subunit 1 [Sanguibacter keddiei DSM 10542] gi 269099028 gb ACZ23464.1  Pertussis toxin, subunit 1 [Sanguibacter keddiei DSM 10542]
SYHT0H2_I_16	TOXIN_2012_2 gi 269796843 ref YP_003316298.1	Pertussis toxin subunit 1 [Sanguibacter keddiei DSM 10542] gi 269099028 gb ACZ23464.1  Pertussis toxin, subunit 1 [Sanguibacter keddiei DSM 10542]

Query	Subject ID	Description
SYHT0H2_I_28	TOXIN_2012_2 gi 269796843 ref YP_003316298.1	Pertussis toxin subunit 1 [Sanguibacter keddiei DSM 10542] gi 269099028 gb ACZ23464.1  Pertussis toxin, subunit 1 [Sanguibacter keddiei DSM 10542]
SYHT0H2_I_38	TOXIN_2012_2 gi 269796843 ref YP_003316298.1	Pertussis toxin subunit 1 [Sanguibacter keddiei DSM 10542] gi 269099028 gb ACZ23464.1  Pertussis toxin, subunit 1 [Sanguibacter keddiei DSM 10542]
SYHT0H2_I_18	TOXIN_2012_2 gi 23015941 ref ZP_00055704.1	COG2931: RTX toxins and related Ca <sup>2+</sup> -binding proteins [Magnetospirillum magnetotacticum MS-1]
SYHT0H2_I_36	TOXIN_2012_2 gi 23015941 ref ZP_00055704.1	COG2931: RTX toxins and related Ca <sup>2+</sup> -binding proteins [Magnetospirillum magnetotacticum MS-1]
SYHT0H2_I_47	TOXIN_2012_2 gi 170022729 ref YP_001719234.1	virulence plasmid 65kDa B protein [Yersinia pseudotuberculosis YPIII] gi 149211827 gb ABR22251.1  toxin complex subunit TcaC [Yersinia pseudotuberculosis] gi 169749263 gb ACA66781.1  virulence plasmid 65kDa B protein [Yersinia pseudotuberculosis YPIII]
SYHT0H2_I_14	TOXIN_2012_2 gi 374599467 ref ZP_09672469.1	toxic anion resistance family protein [Myroides odoratus DSM 2801] gi 373910937 gb EHQ42786.1  toxic anion resistance family protein [Myroides odoratus DSM 2801]
SYHT0H2_I_18	TOXIN_2012_2 gi 23015941 ref ZP_00055704.1	COG2931: RTX toxins and related Ca <sup>2+</sup> -binding proteins [Magnetospirillum magnetotacticum MS-1]
SYHT0H2_I_36	TOXIN_2012_2 gi 23015941 ref ZP_00055704.1	COG2931: RTX toxins and related Ca <sup>2+</sup> -binding proteins [Magnetospirillum magnetotacticum MS-1]
SYHT0H2_I_4	TOXIN_2012_2 gi 29337194 sp P08649.3 CO4_RAT	RecName: Full=Complement C4; Contains: RecName: Full=Complement C4 beta chain; Contains: RecName: Full=Complement C4 alpha chain; Contains: RecName: Full=C4a anaphylatoxin; Contains: RecName: Full=Complement C4 gamma chain; Flags: Precursor gi 25244410 gb AAN72415.1  complement component C4 [Rattus norvegicus]
SYHT0H2_I_32	TOXIN_2012_2 gi 29337194 sp P08649.3 CO4_RAT	RecName: Full=Complement C4; Contains: RecName: Full=Complement C4 beta chain; Contains: RecName: Full=Complement C4 alpha chain; Contains: RecName: Full=C4a anaphylatoxin; Contains: RecName: Full=Complement C4 gamma chain; Flags: Precursor gi 25244410 gb AAN72415.1  complement component C4 [Rattus norvegicus]
SYHT0H2_I_31	TOXIN_2012_2 gi 347602872 gb EGY27821.1	RTX toxin [Candidatus Regiella insecticola R5.15]

Query	Subject ID	Description
SYHT0H2_I_4	TOXIN_2012_2 gi 317025490 ref XP_001389187.2	toxin biosynthesis protein [Aspergillus niger CBS 513.88]
SYHT0H2_I_22	TOXIN_2012_2 gi 149190738 ref ZP_01869004.1	putative RTX toxin [Vibrio shilonii AK1] gi 148835394 gb EDL52365.1  putative RTX toxin [Vibrio shilonii AK1]
SYHT0H2_I_26	TOXIN_2012_2 gi 23015941 ref ZP_00055704.1	COG2931: RTX toxins and related Ca <sup>2+</sup> -binding proteins [Magnetospirillum magnetotacticum MS-1]
SYHT0H2_I_32	TOXIN_2012_2 gi 317025490 ref XP_001389187.2	toxin biosynthesis protein [Aspergillus niger CBS 513.88]
SYHT0H2_I_44	TOXIN_2012_2 gi 23015941 ref ZP_00055704.1	COG2931: RTX toxins and related Ca <sup>2+</sup> -binding proteins [Magnetospirillum magnetotacticum MS-1]
SYHT0H2_I_10	TOXIN_2012_2 gi 148273353 ref YP_001222914.1	putative RTX toxin [Clavibacter michiganensis subsp. michiganensis NCPPB 382] gi 147831283 emb CAN02239.1  putative RTX toxin [Clavibacter michiganensis subsp. michiganensis NCPPB 382]
SYHT0H2_I_16	TOXIN_2012_2 gi 148273353 ref YP_001222914.1	putative RTX toxin [Clavibacter michiganensis subsp. michiganensis NCPPB 382] gi 147831283 emb CAN02239.1  putative RTX toxin [Clavibacter michiganensis subsp. michiganensis NCPPB 382]
SYHT0H2_I_28	TOXIN_2012_2 gi 148273353 ref YP_001222914.1	putative RTX toxin [Clavibacter michiganensis subsp. michiganensis NCPPB 382] gi 147831283 emb CAN02239.1  putative RTX toxin [Clavibacter michiganensis subsp. michiganensis NCPPB 382]
SYHT0H2_I_38	TOXIN_2012_2 gi 148273353 ref YP_001222914.1	putative RTX toxin [Clavibacter michiganensis subsp. michiganensis NCPPB 382] gi 147831283 emb CAN02239.1  putative RTX toxin [Clavibacter michiganensis subsp. michiganensis NCPPB 382]
SYHT0H2_I_1	TOXIN_2012_2 gi 212637285 ref YP_002313810.1	unnamed protein product [Shewanella piezotolerans WP3] gi 212558769 gb ACJ31223.1  RTX toxin, putative [Shewanella piezotolerans WP3]
SYHT0H2_I_35	TOXIN_2012_2 gi 212637285 ref YP_002313810.1	unnamed protein product [Shewanella piezotolerans WP3] gi 212558769 gb ACJ31223.1  RTX toxin, putative [Shewanella piezotolerans WP3]
SYHT0H2_I_26	TOXIN_2012_2 gi 23015941 ref ZP_00055704.1	COG2931: RTX toxins and related Ca <sup>2+</sup> -binding proteins [Magnetospirillum magnetotacticum MS-1]
SYHT0H2_I_44	TOXIN_2012_2 gi 23015941 ref ZP_00055704.1	COG2931: RTX toxins and related Ca <sup>2+</sup> -binding proteins [Magnetospirillum magnetotacticum MS-1]
SYHT0H2_I_1	TOXIN_2012_2 gi 375262803 ref YP_005025033.1	unnamed protein product [Vibrio sp. EJY3] gi 369844458 gb AEX25286.1  RTX toxin [Vibrio sp. EJY3]

Query	Subject ID	Description
SYHT0H2_I_35	TOXIN_2012_2 gi 375262803 ref YP_005025033.1	unnamed protein product [Vibrio sp. EJY3] gi 369844458 gb AEX25286.1  RTX toxin [Vibrio sp. EJY3]
SYHT0H2_I_4	TOXIN_2012_2 gi 238492287 ref XP_002377380.1	MFS gliotoxin efflux transporter GliA [Aspergillus flavus NRRL3357] gi 220695874 gb EED52216.1  MFS gliotoxin efflux transporter GliA [Aspergillus flavus NRRL3357]
SYHT0H2_I_32	TOXIN_2012_2 gi 238492287 ref XP_002377380.1	MFS gliotoxin efflux transporter GliA [Aspergillus flavus NRRL3357] gi 220695874 gb EED52216.1  MFS gliotoxin efflux transporter GliA [Aspergillus flavus NRRL3357]
SYHT0H2_I_4	TOXIN_2012_2 gi 358367040 dbj GAA83660.1	toxin biosynthesis protein [Aspergillus kawachii IFO 4308]
SYHT0H2_I_32	TOXIN_2012_2 gi 358367040 dbj GAA83660.1	toxin biosynthesis protein [Aspergillus kawachii IFO 4308]
SYHT0H2_I_23	TOXIN_2012_2 gi 326664496 ref XP_001920822.3	PREDICTED: neoverrucotoxin subunit beta [Danio rerio]
SYHT0H2_I_15	TOXIN_2012_2 gi 52140196 ref YP_086635.1	membrane protein; regulatory protein [Bacillus cereus E33L] gi 196036525 ref ZP_03103920.1  membrane protein PfoR [Bacillus cereus W] gi 196040980 ref ZP_03108277.1  membrane protein PfoR [Bacillus cereus NVH0597-99] gi 196045481 ref ZP_03112712.1  membrane protein PfoR [Bacillus cereus 03BB108] gi 218906545 ref YP_002454379.1  membrane protein PfoR [Bacillus cereus AH820] gi 225867349 ref YP_002752727.1  membrane protein PfoR [Bacillus cereus 03BB102] gi 228917980 ref ZP_04081512.1  Transcriptional regulator pfoR [Bacillus thuringiensis serovar pulsiensis BGSC 4CC1] gi 228936653 ref ZP_04099446.1  Transcriptional regulator pfoR [Bacillus thuringiensis serovar andalousiensis BGSC 4AW1] gi 228949092 ref ZP_04111362.1  Transcriptional regulator pfoR [Bacillus thuringiensis serovar monterrey BGSC 4AJ1] gi 229187595 ref ZP_04314733.1  Transcriptional regulator pfoR [Bacillus cereus BGSC 6E1] gi 301056836 ref YP_003795047.1  unnamed protein product [Bacillus cereus biovar anthracis str. CI] gi 51973665 gb AAU15215.1  membrane protein; possible regulatory protein [Bacillus cereus E33L] gi 195990858 gb EDX54831.1  membrane protein PfoR [Bacillus cereus W] gi 196023688 gb EDX62364.1  membrane protein PfoR [Bacillus cereus 03BB108] gi 196028148 gb EDX66758.1  membrane protein PfoR [Bacillus cereus NVH0597-99]

Query	Subject ID	Description
		gi 218535498 gb ACK87896.1  membrane protein PfoR [Bacillus cereus AH820] gi 225790834 gb ACO31051.1  membrane protein PfoR [Bacillus cereus 03BB102] gi 228595847 gb EEK53529.1  Transcriptional regulator pfoR [Bacillus cereus BGSC 6E1] gi 228810533 gb EEM56884.1  Transcriptional regulator pfoR [Bacillus thuringiensis serovar monterrey BGSC 4AJ1] gi 228822992 gb EEM68831.1  Transcriptional regulator pfoR [Bacillus thuringiensis serovar andalousiensis BGSC 4AW1] gi 228841645 gb EEM86757.1  Transcriptional regulator pfoR [Bacillus thuringiensis serovar pulsensis BGSC 4CC1] gi 300379005 gb ADK07909.1  putative membrane protein [Bacillus cereus biovar anthracis str. CI] gi 364515083 gb AEW58482.1  Perfringolysin O regulator protein PfoR [Bacillus cereus F837/76]
SYHT0H2_I_4	TOXIN_2012_2 gi 315081364 gb EFT53340.1	toxin secretion/phage lysis holin [Propionibacterium acnes HL078PA1]
SYHT0H2_I_32	TOXIN_2012_2 gi 315081364 gb EFT53340.1	toxin secretion/phage lysis holin [Propionibacterium acnes HL078PA1]
SYHT0H2_I_4	TOXIN_2012_2 gi 335050940 ref ZP_08543885.1	toxin secretion/phage lysis holin [Propionibacterium sp. 409-HC1] gi 342212266 ref ZP_08704991.1  transporter, SSS family [Propionibacterium sp. CC003-HC2] gi 313794187 gb EFS42207.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL110PA1] gi 313839161 gb EFS76875.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL086PA1] gi 314963843 gb EFT07943.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL082PA1] gi 315078326 gb EFT50365.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL053PA2] gi 327455936 gb EGF02591.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL092PA1] gi 333768512 gb EGL45694.1  toxin secretion/phage lysis holin [Propionibacterium sp. 409-HC1] gi 340767810 gb EGR90335.1  transporter, SSS family [Propionibacterium sp. CC003-HC2]
SYHT0H2_I_4	TOXIN_2012_2 gi 313812423 gb EFS50137.1	toxin secretion/phage lysis holin [Propionibacterium acnes HL025PA1]
SYHT0H2_I_32	TOXIN_2012_2 gi 335050940 ref ZP_08543885.1	toxin secretion/phage lysis holin [Propionibacterium sp. 409-HC1] gi 342212266 ref ZP_08704991.1  transporter,

Query	Subject ID	Description
		SSS family [Propionibacterium sp. CC003-HC2] gi 313794187 gb EFS42207.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL110PA1] gi 313839161 gb EFS76875.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL086PA1] gi 314963843 gb EFT07943.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL082PA1] gi 315078326 gb EFT50365.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL053PA2] gi 327455936 gb EGF02591.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL092PA1] gi 333768512 gb EGL45694.1  toxin secretion/phage lysis holin [Propionibacterium sp. 409-HC1] gi 340767810 gb EGR90335.1  transporter, SSS family [Propionibacterium sp. CC003-HC2]
SYHT0H2_I_32	TOXIN_2012_2 gi 313812423 gb EFS50137.1	toxin secretion/phage lysis holin [Propionibacterium acnes HL025PA1]
SYHT0H2_I_4	TOXIN_2012_2 gi 327450317 gb EGE96971.1	toxin secretion/phage lysis holin [Propionibacterium acnes HL013PA2]
SYHT0H2_I_4	TOXIN_2012_2 gi 289426084 ref ZP_06427831.1	toxin secretion/phage lysis holin [Propionibacterium acnes SK187] gi 289427915 ref ZP_06429619.1  toxin secretion/phage lysis holin [Propionibacterium acnes J165] gi 295131714 ref YP_003582377.1  toxin secretion/phage lysis holin [Propionibacterium acnes SK137] gi 335055081 ref ZP_08547872.1  toxin secretion/phage lysis holin [Propionibacterium sp. 434-HC2] gi 365963814 ref YP_004945380.1  unnamed protein product [Propionibacterium acnes TypeIA2 P.acn31] gi 365966054 ref YP_004947619.1  unnamed protein product [Propionibacterium acnes TypeIA2 P.acn17] gi 365974993 ref YP_004956552.1  unnamed protein product [Propionibacterium acnes TypeIA2 P.acn33] gi 289153627 gb EFD02341.1  toxin secretion/phage lysis holin [Propionibacterium acnes SK187] gi 289158798 gb EFD06998.1  toxin secretion/phage lysis holin [Propionibacterium acnes J165] gi 291375283 gb ADD99137.1  toxin secretion/phage lysis holin [Propionibacterium acnes SK137]



Query	Subject ID	Description
		gi 313765210 gb EFS36574.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL013PA1] gi 313771631 gb EFS37597.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL074PA1] gi 313808642 gb EFS47102.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL087PA2] gi 313810645 gb EFS48359.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL083PA1] gi 313817068 gb EFS54782.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL059PA1] gi 313819118 gb EFS56832.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL046PA2] gi 313821169 gb EFS58883.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL036PA1] gi 313823846 gb EFS61560.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL036PA2] gi 313826948 gb EFS64662.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL063PA1] gi 313829213 gb EFS66927.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL063PA2] gi 313831458 gb EFS69172.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL007PA1] gi 313834108 gb EFS71822.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL056PA1] gi 314916787 gb EFS80618.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL005PA4] gi 314919058 gb EFS82889.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL050PA1] gi 314921100 gb EFS84931.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL050PA3] gi 314926186 gb EFS90017.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL036PA3] gi 314931365 gb EFS95196.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL067PA1] gi 314956221 gb EFT00593.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL027PA1] gi 314958716 gb EFT02818.1  toxin

Query	Subject ID	Description
		<p>secretion/phage lysis holin  [Propionibacterium acnes HL002PA1]  gi 314961023 gb EFT05124.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL002PA2]  gi 314968848 gb EFT12946.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL037PA1]  gi 314974848 gb EFT18943.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL053PA1]  gi 314977825 gb EFT21919.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL045PA1]  gi 314979020 gb EFT23114.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL072PA2]  gi 314984529 gb EFT28621.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL005PA1]  gi 314986006 gb EFT30098.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL005PA2]  gi 314989309 gb EFT33400.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL005PA3]  gi 315084996 gb EFT56972.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL027PA2]  gi 315087526 gb EFT59502.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL002PA3]  gi 315089567 gb EFT61543.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL072PA1]  gi 315095673 gb EFT67649.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL038PA1]  gi 315100002 gb EFT71978.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL059PA2]  gi 315102838 gb EFT74814.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL046PA1]  gi 315106633 gb EFT78609.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL030PA1]  gi 315110565 gb EFT82541.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL030PA2]  gi 327325857 gb EGE67648.1  sodium:solute  symporter [Propionibacterium acnes  HL096PA3] gi 327332953 gb EGE74685.1   sodium:solute symporter [Propionibacterium  acnes HL096PA2]</p>

Query	Subject ID	Description
		gi 327448659 gb EGE95313.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL043PA1] gi 327451116 gb EGE97770.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL043PA2] gi 327455475 gb EGF02130.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL087PA3] gi 327457616 gb EGF04271.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL083PA2] gi 328756865 gb EGF70481.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL087PA1] gi 328757750 gb EGF71366.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL020PA1] gi 328759480 gb EGF73096.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL025PA2] gi 328761925 gb EGF75434.1  sodium:solute symporter [Propionibacterium acnes HL099PA1] gi 333762917 gb EGL40399.1  toxin secretion/phage lysis holin [Propionibacterium sp. 434-HC2] gi 335278595 gb AEH30500.1  toxin secretion/phage lysis holin [Propionibacterium acnes 6609] gi 340769716 gb EGR92238.1  transporter, SSS family [Propionibacterium acnes SK182] gi 365740495 gb AEW84697.1  toxin secretion/phage lysis holin [Propionibacterium acnes TypeIA2 P.acn31] gi 365742735 gb AEW82429.1  toxin secretion/phage lysis holin [Propionibacterium acnes TypeIA2 P.acn17] gi 365744992 gb AEW80189.1  toxin secretion/phage lysis holin [Propionibacterium acnes TypeIA2 P.acn33]
SYHT0H2_I_32	TOXIN_2012_2 gi 327450317 gb EGE96971.1	toxin secretion/phage lysis holin [Propionibacterium acnes HL013PA2]
SYHT0H2_I_32	TOXIN_2012_2 gi 289426084 ref ZP_06427831.1	toxin secretion/phage lysis holin [Propionibacterium acnes SK187] gi 289427915 ref ZP_06429619.1  toxin secretion/phage lysis holin [Propionibacterium acnes J165] gi 295131714 ref YP_003582377.1  toxin secretion/phage lysis holin [Propionibacterium acnes SK137] gi 335055081 ref ZP_08547872.1  toxin secretion/phage lysis holin [Propionibacterium sp. 434-HC2] gi 365963814 ref YP_004945380.1  unnamed protein product [Propionibacterium acnes

Query	Subject ID	Description
		TypeIA2 P.acn31] gi 365966054 ref YP_004947619.1  unnamed protein product [Propionibacterium acnes TypeIA2 P.acn17] gi 365974993 ref YP_004956552.1  unnamed protein product [Propionibacterium acnes TypeIA2 P.acn33] gi 289153627 gb EFD02341.1  toxin secretion/phage lysis holin [Propionibacterium acnes SK187] gi 289158798 gb EFD06998.1  toxin secretion/phage lysis holin [Propionibacterium acnes J165] gi 291375283 gb ADD99137.1  toxin secretion/phage lysis holin [Propionibacterium acnes SK137] gi 313765210 gb EFS36574.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL013PA1] gi 313771631 gb EFS37597.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL074PA1] gi 313808642 gb EFS47102.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL087PA2] gi 313810645 gb EFS48359.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL083PA1] gi 313817068 gb EFS54782.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL059PA1] gi 313819118 gb EFS56832.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL046PA2] gi 313821169 gb EFS58883.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL036PA1] gi 313823846 gb EFS61560.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL036PA2] gi 313826948 gb EFS64662.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL063PA1] gi 313829213 gb EFS66927.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL063PA2] gi 313831458 gb EFS69172.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL007PA1] gi 313834108 gb EFS71822.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL056PA1] gi 314916787 gb EFS80618.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL005PA4]

Query	Subject ID	Description
		gi 314919058 gb EFS82889.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL050PA1] gi 314921100 gb EFS84931.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL050PA3] gi 314926186 gb EFS90017.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL036PA3] gi 314931365 gb EFS95196.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL067PA1] gi 314956221 gb EFT00593.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL027PA1] gi 314958716 gb EFT02818.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL002PA1] gi 314961023 gb EFT05124.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL002PA2] gi 314968848 gb EFT12946.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL037PA1] gi 314974848 gb EFT18943.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL053PA1] gi 314977825 gb EFT21919.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL045PA1] gi 314979020 gb EFT23114.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL072PA2] gi 314984529 gb EFT28621.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL005PA1] gi 314986006 gb EFT30098.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL005PA2] gi 314989309 gb EFT33400.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL005PA3] gi 315084996 gb EFT56972.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL027PA2] gi 315087526 gb EFT59502.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL002PA3] gi 315089567 gb EFT61543.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL072PA1] gi 315095673 gb EFT67649.1  toxin secretion/phage lysis holin [Propionibacterium acnes HL038PA1] gi 315100002 gb EFT71978.1  toxin

Query	Subject ID	Description
		<p>secretion/phage lysis holin  [Propionibacterium acnes HL059PA2]  gi 315102838 gb EFT74814.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL046PA1]  gi 315106633 gb EFT78609.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL030PA1]  gi 315110565 gb EFT82541.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL030PA2]  gi 327325857 gb EGE67648.1  sodium:solute  symporter [Propionibacterium acnes  HL096PA3] gi 327332953 gb EGE74685.1   sodium:solute symporter [Propionibacterium  acnes HL096PA2]  gi 327448659 gb EGE95313.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL043PA1]  gi 327451116 gb EGE97770.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL043PA2]  gi 327455475 gb EGF02130.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL087PA3]  gi 327457616 gb EGF04271.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL083PA2]  gi 328756865 gb EGF70481.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL087PA1]  gi 328757750 gb EGF71366.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL020PA1]  gi 328759480 gb EGF73096.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes HL025PA2]  gi 328761925 gb EGF75434.1  sodium:solute  symporter [Propionibacterium acnes  HL099PA1] gi 333762917 gb EGL40399.1   toxin secretion/phage lysis holin  [Propionibacterium sp. 434-HC2]  gi 335278595 gb AEH30500.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes 6609]  gi 340769716 gb EGR92238.1  transporter,  SSS family [Propionibacterium acnes SK182]  gi 365740495 gb AEW84697.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes TypeIA2 P.acn31]  gi 365742735 gb AEW82429.1  toxin  secretion/phage lysis holin  [Propionibacterium acnes TypeIA2 P.acn17]  gi 365744992 gb AEW80189.1  toxin  secretion/phage lysis holin</p>

Query	Subject ID	Description
		[Propionibacterium acnes TypeIA2 P.acn33]

Query= SYHT0H2\_I\_1  
(115 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

Sequences producing significant alignments:

	Score (bits)	E Value
--	-----------------	------------

TOXIN_2012_2 gi 212637285 ref YP_002313810.1  unnamed protein pr...	25	5.6
TOXIN_2012_2 gi 375262803 ref YP_005025033.1  unnamed protein pr...	25	5.9

>TOXIN\_2012\_2|gi|212637285|ref|YP\_002313810.1| unnamed protein  
product [Shewanella piezotolerans WP3]  
gi|212558769|gb|ACJ31223.1| RTX toxin, putative  
[Shewanella piezotolerans WP3]  
Length = 744

Score = 25.0 bits (53), Expect = 5.6, Method: Composition-based stats.  
Identities = 9/32 (28%), Positives = 17/32 (53%)

Query: 8 ATGFNLKKLYCQMFERSGKFGDLELDSYLGW 39  
A G++ ++ + + G+FG L +D G W  
Sbjct: 315 AQGYHSAQINWSINQSQGQFGSLNIDPVTGQW 346

>TOXIN\_2012\_2|gi|375262803|ref|YP\_005025033.1| unnamed protein  
product [Vibrio sp. EJY3] gi|369844458|gb|AEX25286.1|  
RTX toxin [Vibrio sp. EJY3]  
Length = 3043

Score = 25.0 bits (53), Expect = 5.9, Method: Composition-based stats.  
Identities = 8/17 (47%), Positives = 12/17 (70%)

Query: 23 RSGKFGDLELDSYLGW 39  
+ G +G+L LDS+ G W  
Sbjct: 248 KGGHYGELTLDHSGEW 264

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.323	0.134	0.477

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
 Gap Penalties: Existence: 11, Extension: 1  
 Number of Sequences: 23302  
 Number of Hits to DB: 1,525,158  
 Number of extensions: 50998  
 Number of successful extensions: 90  
 Number of sequences better than 10.0: 2  
 Number of HSP's gapped: 90  
 Number of HSP's successfully gapped: 2  
 Length of query: 115  
 Length of database: 7,243,252  
 Length adjustment: 81  
 Effective length of query: 34  
 Effective length of database: 5,355,790  
 Effective search space: 182096860  
 Effective search space used: 182096860  
 Neighboring words threshold: 11  
 Window for multiple hits: 40  
 X1: 16 ( 7.5 bits)  
 X2: 38 (14.6 bits)  
 X3: 64 (24.7 bits)  
 S1: 40 (21.6 bits)  
 S2: 51 (24.3 bits)  
 BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_2  
 (31 letters)

Database: TOXIN\_2012\_2  
 23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
 Posted date: Feb 22, 2012 11:49 AM  
 Number of letters in database: 7,243,252  
 Number of sequences in database: 23,302

Lambda	K	H
0.325	0.141	0.534

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
 Gap Penalties: Existence: 11, Extension: 1  
 Number of Sequences: 23302



Number of Hits to DB: 539,802  
 Number of extensions: 10564  
 Number of successful extensions: 55  
 Number of sequences better than 10.0: 0  
 Number of HSP's gapped: 55  
 Number of HSP's successfully gapped: 0  
 Length of query: 31  
 Length of database: 7,243,252  
 Length adjustment: 6  
 Effective length of query: 25  
 Effective length of database: 7,103,440  
 Effective search space: 177586000  
 Effective search space used: 177586000  
 Neighboring words threshold: 11  
 Window for multiple hits: 40  
 X1: 15 ( 7.0 bits)  
 X2: 38 (14.6 bits)  
 X3: 64 (24.7 bits)  
 S1: 40 (21.6 bits)  
 S2: 51 (24.3 bits)  
 BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_3  
 (31 letters)

Database: TOXIN\_2012\_2  
 23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
 Posted date: Feb 22, 2012 11:49 AM  
 Number of letters in database: 7,243,252  
 Number of sequences in database: 23,302

Lambda	K	H
0.319	0.138	0.491

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
 Gap Penalties: Existence: 11, Extension: 1  
 Number of Sequences: 23302  
 Number of Hits to DB: 477,123  
 Number of extensions: 6978  
 Number of successful extensions: 5

Number of sequences better than 10.0: 0  
 Number of HSP's gapped: 5  
 Number of HSP's successfully gapped: 0  
 Length of query: 31  
 Length of database: 7,243,252  
 Length adjustment: 6  
 Effective length of query: 25  
 Effective length of database: 7,103,440  
 Effective search space: 177586000  
 Effective search space used: 177586000  
 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 (21.7 bits)  
 S2: 51 (24.3 bits)  
 BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_4  
 (121 letters)

Database: TOXIN\_2012\_2  
 23,302 sequences; 7,243,252 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
TOXIN_2012_2 gi 157951698 ref NP_033910.2  complement C4-B precu...	27	1.7
TOXIN_2012_2 gi 29337194 sp P08649.3 CO4_RAT RecName: Full=Compl...	25	4.5
TOXIN_2012_2 gi 317025490 ref XP_001389187.2  toxin biosynthesis...	25	5.4
TOXIN_2012_2 gi 238492287 ref XP_002377380.1  MFS gliotoxin effl...	25	6.8
TOXIN_2012_2 gi 358367040 dbj GAA83660.1  toxin biosynthesis pro...	25	7.5
TOXIN_2012_2 gi 315081364 gb EFT53340.1  toxin secretion/phage l...	24	9.7
TOXIN_2012_2 gi 335050940 ref ZP_08543885.1  toxin secretion/pha...	24	9.8
TOXIN_2012_2 gi 313812423 gb EFS50137.1  toxin secretion/phage l...	24	9.8
TOXIN_2012_2 gi 327450317 gb EGE96971.1  toxin secretion/phage l...	24	9.9
TOXIN_2012_2 gi 289426084 ref ZP_06427831.1  toxin secretion/pha...	24	9.9

>TOXIN\_2012\_2|gi|157951698|ref|NP\_033910.2| complement C4-B  
 precursor [Mus musculus]  
 gi|341940526|sp|P01029.3|CO4B\_MOUSE RecName:  
 Full=Complement C4-B; Contains: RecName: Full=Complement  
 C4 beta chain; Contains: RecName: Full=Complement C4  
 alpha chain; Contains: RecName: Full=C4a anaphylatoxin;  
 Contains: RecName: Full=Complement C4 gamma chain;  
 Flags: Precursor  
 Length = 1738

Score = 26.9 bits (58), Expect = 1.7, Method: Composition-based stats.  
Identities = 25/93 (26%), Positives = 37/93 (39%), Gaps = 11/93 (11%)

Query: 35 VQIWVTGLTGL--GGAGNSKSLCQKPTSCHPPCLYP---AARNVPRAVYPKASCNLTGD 88  
V + VT + GL G G P P + AA NVP V + +L D  
Sbjct: 863 VSVHVTPVEGLCLAGGGMAQQVTPAGSARPVAFSVVPTAAANVPLKVVARGVFDLGDA 922

Query: 89 SFGRPITATTDLKPCASIDLSKCVYNVDPRPNL 121  
++ ++ +I + VYN+DP NL  
Sbjct: 923 -----VSKILQIEKEGAIHREELVYNLDPLNNL 950

>TOXIN\_2012\_2|gi|29337194|sp|P08649.3|CO4\_RAT RecName:  
Full=Complement C4; Contains: RecName: Full=Complement  
C4 beta chain; Contains: RecName: Full=Complement C4  
alpha chain; Contains: RecName: Full=C4a anaphylatoxin;  
Contains: RecName: Full=Complement C4 gamma chain;  
Flags: Precursor gi|25244410|gb|AAN72415.1| complement  
component C4 [Rattus norvegicus]  
Length = 1737

Score = 25.4 bits (54), Expect = 4.5, Method: Composition-based stats.  
Identities = 18/60 (30%), Positives = 30/60 (50%), Gaps = 7/60 (11%)

Query: 69 AARNVPRAVYPKASCNLTGDSF--GRPITATTDLKPCASIDLSKCVYNVDPRPNL 121  
+AR V +V P A+ ++ GSF G ++ ++ +I + VYN+DP NL  
Sbjct: 890 SARPVAFSVVPTAAASIPLKVVARGSFTIGDAVSKILQIEKEGAIHREEIVYNLDPLNNL 949

>TOXIN\_2012\_2|gi|317025490|ref|XP\_001389187.2| toxin biosynthesis  
protein [Aspergillus niger CBS 513.88]  
Length = 446

Score = 25.0 bits (53), Expect = 5.4, Method: Composition-based stats.  
Identities = 11/30 (36%), Positives = 16/30 (53%)

Query: 71 RNVPRVYPKASCNLTGDSFGRPITATTDL 100  
R +P A++P +GS RP+T T L  
Sbjct: 241 RELPTAIHPLDESQQKEGSNDRPVTLRTTL 270

>TOXIN\_2012\_2|gi|238492287|ref|XP\_002377380.1| MFS gliotoxin efflux  
transporter GliA [Aspergillus flavus NRRL3357]  
gi|220695874|gb|EED52216.1| MFS gliotoxin efflux  
transporter GliA [Aspergillus flavus NRRL3357]  
Length = 523

Score = 25.0 bits (53), Expect = 6.8, Method: Composition-based stats.  
Identities = 23/69 (33%), Positives = 29/69 (42%), Gaps = 2/69 (2%)

Query: 29 MAPISLVQIWVTGLTGLGGAGNSKSLCQKPTSCHPPCLYPAARNVPRAVYPKASC--NLT 86  
+AP SL +TG+GGAG S PP PA + A Y A+ L  
Sbjct: 108 VAPNSLALCIGRAITGVGGAGISSGAFTIIALSAPPKQRPAYIGILGASYGVAAAIGPLV 167

Query: 87 DGSFGRPIT 95  
G+F IT  
Sbjct: 168 GGAFITNIT 176

>TOXIN\_2012\_2|gi|358367040|dbj|GAA83660.1| toxin biosynthesis  
protein [Aspergillus kawachii IFO 4308]  
Length = 447

Score = 24.6 bits (52), Expect = 7.5, Method: Composition-based stats.  
Identities = 11/30 (36%), Positives = 16/30 (53%)

Query: 71 RNVPRVYPKASCNLTGDSFGRPITATD 100  
R +P A++P +GS RP+T T L  
Sbjct: 241 RELPTAIHPLDESQQREGSDDRPVTLRTTL 270

>TOXIN\_2012\_2|gi|315081364|gb|EFT53340.1| toxin secretion/phage  
lysis holin [Propionibacterium acnes HL078PA1]  
Length = 569

Score = 24.3 bits (51), Expect = 9.7, Method: Composition-based stats.  
Identities = 12/41 (29%), Positives = 20/41 (48%)

Query: 33 SLVQIWVTGLTGLGGAGNSKSLCQKPTSCHPPCLYPAAARNV 73  
SL+ + + GL +GG K+ HP + PAA+ +  
Sbjct: 196 SLLPLVIIGLHRVGWWSGLKAKITAAAEHPDVVTPAAQQL 236

>TOXIN\_2012\_2|gi|335050940|ref|ZP\_08543885.1| toxin secretion/phage  
lysis holin [Propionibacterium sp. 409-HC1]  
gi|342212266|ref|ZP\_08704991.1| transporter, SSS family  
[Propionibacterium sp. CC003-HC2]  
gi|313794187|gb|EFS42207.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL110PA1]  
gi|313839161|gb|EFS76875.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL086PA1]  
gi|314963843|gb|EFT07943.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL082PA1]  
gi|315078326|gb|EFT50365.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL053PA2]  
gi|327455936|gb|EGF02591.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL092PA1]  
gi|333768512|gb|EGL45694.1| toxin secretion/phage lysis  
holin [Propionibacterium sp. 409-HC1]  
gi|340767810|gb|EGR90335.1| transporter, SSS family  
[Propionibacterium sp. CC003-HC2]  
Length = 569

Score = 24.3 bits (51), Expect = 9.8, Method: Composition-based stats.  
Identities = 12/41 (29%), Positives = 20/41 (48%)

Query: 33 SLVQIWVTGLTGLGGAGNSKSLCQKPTSCHPPCLYPAAARNV 73  
SL+ + + GL +GG K+ HP + PAA+ +  
Sbjct: 196 SLLPLVIIGLHRVGWWSGLKAKITAAAEHPDVVTPAAQQL 236

>TOXIN\_2012\_2|gi|313812423|gb|EFS50137.1| toxin secretion/phage  
lysis holin [Propionibacterium acnes HL025PA1]  
Length = 569

Score = 24.3 bits (51), Expect = 9.8, Method: Composition-based stats.  
Identities = 12/41 (29%), Positives = 20/41 (48%)

Query: 33 SLVQIWVTGLTGLGGAGNSKSLCQKPTSCHPPCLYPAAARNV 73  
SL+ + + GL +GG K+ HP + PAA+ +  
Sbjct: 196 SLLPLVIIGLHRVGWWSGLKAKITAAAEHPDVVTPAAQQL 236

>TOXIN\_2012\_2|gi|327450317|gb|EGE96971.1| toxin secretion/phage  
lysis holin [Propionibacterium acnes HL013PA2]

Length = 569

Score = 24.3 bits (51), Expect = 9.9, Method: Composition-based stats.  
Identities = 12/41 (29%), Positives = 20/41 (48%)

Query: 33 SLVQIWVTGLTGLGGAGNSKSLCQKPTSCHPPCLYPAARNV 73  
SL+ + + GL +GG K+ HP + PAA+ +  
Sbjct: 196 SLLPLVIIGLHRVGGWSGLKAKITAAAEHPDVVTPAAQQL 236

>TOXIN\_2012\_2|gi|289426084|ref|ZP\_06427831.1| toxin secretion/phage  
lysis holin [Propionibacterium acnes SK187]  
gi|289427915|ref|ZP\_06429619.1| toxin secretion/phage  
lysis holin [Propionibacterium acnes J165]  
gi|295131714|ref|YP\_003582377.1| toxin secretion/phage  
lysis holin [Propionibacterium acnes SK137]  
gi|335055081|ref|ZP\_08547872.1| toxin secretion/phage  
lysis holin [Propionibacterium sp. 434-HC2]  
gi|365963814|ref|YP\_004945380.1| unnamed protein product  
[Propionibacterium acnes TypeIA2 P.acn31]  
gi|365966054|ref|YP\_004947619.1| unnamed protein product  
[Propionibacterium acnes TypeIA2 P.acn17]  
gi|365974993|ref|YP\_004956552.1| unnamed protein product  
[Propionibacterium acnes TypeIA2 P.acn33]  
gi|289153627|gb|EFD02341.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes SK187]  
gi|289158798|gb|EFD06998.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes J165]  
gi|291375283|gb|ADD99137.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes SK137]  
gi|313765210|gb|EFS36574.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL013PA1]  
gi|313771631|gb|EFS37597.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL074PA1]  
gi|313808642|gb|EFS47102.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL087PA2]  
gi|313810645|gb|EFS48359.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL083PA1]  
gi|313817068|gb|EFS54782.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL059PA1]  
gi|313819118|gb|EFS56832.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL046PA2]  
gi|313821169|gb|EFS58883.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL036PA1]  
gi|313823846|gb|EFS61560.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL036PA2]  
gi|313826948|gb|EFS64662.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL063PA1]  
gi|313829213|gb|EFS66927.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL063PA2]  
gi|313831458|gb|EFS69172.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL007PA1]  
gi|313834108|gb|EFS71822.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL056PA1]  
gi|314916787|gb|EFS80618.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL005PA4]  
gi|314919058|gb|EFS82889.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL050PA1]  
gi|314921100|gb|EFS84931.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL050PA3]  
gi|314926186|gb|EFS90017.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL036PA3]  
gi|314931365|gb|EFS95196.1| toxin secretion/phage lysis

holin [Propionibacterium acnes HL067PA1]  
 gi|314956221|gb|EFT00593.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL027PA1]  
 gi|314958716|gb|EFT02818.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL002PA1]  
 gi|314961023|gb|EFT05124.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL002PA2]  
 gi|314968848|gb|EFT12946.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL037PA1]  
 gi|314974848|gb|EFT18943.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL053PA1]  
 gi|314977825|gb|EFT21919.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL045PA1]  
 gi|314979020|gb|EFT23114.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL072PA2]  
 gi|314984529|gb|EFT28621.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL005PA1]  
 gi|314986006|gb|EFT30098.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL005PA2]  
 gi|314989309|gb|EFT33400.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL005PA3]  
 gi|315084996|gb|EFT56972.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL027PA2]  
 gi|315087526|gb|EFT59502.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL002PA3]  
 gi|315089567|gb|EFT61543.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL072PA1]  
 gi|315095673|gb|EFT67649.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL038PA1]  
 gi|315100002|gb|EFT71978.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL059PA2]  
 gi|315102838|gb|EFT74814.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL046PA1]  
 gi|315106633|gb|EFT78609.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL030PA1]  
 gi|315110565|gb|EFT82541.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL030PA2]  
 gi|327325857|gb|EGE67648.1| sodium:solute symporter  
 [Propionibacterium acnes HL096PA3]  
 gi|327332953|gb|EGE74685.1| sodium:solute symporter  
 [Propionibacterium acnes HL096PA2]  
 gi|327448659|gb|EGE95313.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL043PA1]  
 gi|327451116|gb|EGE97770.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL043PA2]  
 gi|327455475|gb|EGF02130.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL087PA3]  
 gi|327457616|gb|EGF04271.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL083PA2]  
 gi|328756865|gb|EGF70481.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL087PA1]  
 gi|328757750|gb|EGF71366.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL020PA1]  
 gi|328759480|gb|EGF73096.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL025PA2]  
 gi|328761925|gb|EGF75434.1| sodium:solute symporter  
 [Propionibacterium acnes HL099PA1]  
 gi|333762917|gb|EGL40399.1| toxin secretion/phage lysis  
 holin [Propionibacterium sp. 434-HC2]  
 gi|335278595|gb|AEH30500.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes 6609]  
 gi|340769716|gb|EGR92238.1| transporter, SSS family  
 [Propionibacterium acnes SK182]

gi|365740495|gb|AEW84697.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes TypeIA2 P.acn31]  
gi|365742735|gb|AEW82429.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes TypeIA2 P.acn17]  
gi|365744992|gb|AEW80189.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes TypeIA2 P.acn33]  
Length = 569

Score = 24.3 bits (51), Expect = 9.9, Method: Composition-based stats.  
Identities = 12/41 (29%), Positives = 20/41 (48%)

Query: 33 SLVQIWVTGLTGLGGAGNSKSLCQKPTSCHPPCLYPAARNV 73  
          SL+ + + GL +GG K+ HP + PAA+ +  
Sbjct: 196 SLLPLVIIGLHRVGGWSGLKAKITAAAEHPDVVTPAAQQL 236

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.321	0.137	0.442

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 1,783,989  
Number of extensions: 65418  
Number of successful extensions: 122  
Number of sequences better than 10.0: 13  
Number of HSP's gapped: 122  
Number of HSP's successfully gapped: 13  
Length of query: 121  
Length of database: 7,243,252  
Length adjustment: 83  
Effective length of query: 38  
Effective length of database: 5,309,186  
Effective search space: 201749068  
Effective search space used: 201749068  
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: 52 (24.6 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_5  
(71 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.332	0.143	0.446

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 842,371  
Number of extensions: 22089  
Number of successful extensions: 67  
Number of sequences better than 10.0: 0  
Number of HSP's gapped: 67  
Number of HSP's successfully gapped: 0  
Length of query: 71  
Length of database: 7,243,252  
Length adjustment: 42  
Effective length of query: 29  
Effective length of database: 6,264,568  
Effective search space: 181672472  
Effective search space used: 181672472  
Neighboring words threshold: 11  
Window for multiple hits: 40  
X1: 15 ( 7.2 bits)  
X2: 38 (14.6 bits)  
X3: 64 (24.7 bits)  
S1: 40 (22.0 bits)  
S2: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_6



(82 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.336	0.133	0.634

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 150,352  
Number of extensions: 722  
Number of sequences better than 10.0: 0  
Number of HSP's gapped: 0  
Number of HSP's successfully gapped: 0  
Length of query: 82  
Length of database: 7,243,252  
Length adjustment: 52  
Effective length of query: 30  
Effective length of database: 6,031,548  
Effective search space: 180946440  
Effective search space used: 180946440  
Neighboring words threshold: 11  
Window for multiple hits: 40  
X1: 15 ( 7.3 bits)  
X2: 38 (14.6 bits)  
X3: 64 (24.7 bits)  
S1: 39 (21.8 bits)  
S2: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_7  
(37 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2

Posted date: Feb 22, 2012 11:49 AM

Number of letters in database: 7,243,252

Number of sequences in database: 23,302

Lambda	K	H
0.316	0.128	0.399

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62

Gap Penalties: Existence: 11, Extension: 1

Number of Sequences: 23302

Number of Hits to DB: 543,914

Number of extensions: 9272

Number of successful extensions: 12

Number of sequences better than 10.0: 0

Number of HSP's gapped: 12

Number of HSP's successfully gapped: 0

Length of query: 37

Length of database: 7,243,252

Length adjustment: 11

Effective length of query: 26

Effective length of database: 6,986,930

Effective search space: 181660180

Effective search space used: 181660180

Neighboring words threshold: 11

Window for multiple hits: 40

X1: 16 ( 7.3 bits)

X2: 38 (14.6 bits)

X3: 64 (24.7 bits)

S1: 41 (21.7 bits)

S2: 51 (24.3 bits)

BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_8  
(52 letters)

Database: TOXIN\_2012\_2

23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2

Posted date: Feb 22, 2012 11:49 AM

Number of letters in database: 7,243,252

Number of sequences in database: 23,302

Lambda	K	H
0.317	0.126	0.355

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62

Gap Penalties: Existence: 11, Extension: 1

Number of Sequences: 23302

Number of Hits to DB: 528,832

Number of extensions: 10508

Number of successful extensions: 25

Number of sequences better than 10.0: 0

Number of HSP's gapped: 25

Number of HSP's successfully gapped: 0

Length of query: 52

Length of database: 7,243,252

Length adjustment: 25

Effective length of query: 27

Effective length of database: 6,660,702

Effective search space: 179838954

Effective search space used: 179838954

Neighboring words threshold: 11

Window for multiple hits: 40

X1: 16 ( 7.3 bits)

X2: 38 (14.6 bits)

X3: 64 (24.7 bits)

S1: 41 (21.7 bits)

S2: 51 (24.3 bits)

BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_9  
(79 letters)

Database: TOXIN\_2012\_2

23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.318	0.137	0.416

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 904,108  
Number of extensions: 29320  
Number of successful extensions: 36  
Number of sequences better than 10.0: 1  
Number of HSP's gapped: 36  
Number of HSP's successfully gapped: 1  
Length of query: 79  
Length of database: 7,243,252  
Length adjustment: 49  
Effective length of query: 30  
Effective length of database: 6,101,454  
Effective search space: 183043620  
Effective search space used: 183043620  
Neighboring words threshold: 11  
Window for multiple hits: 40  
X1: 16 ( 7.3 bits)  
X2: 38 (14.6 bits)  
X3: 64 (24.7 bits)  
S1: 41 (21.7 bits)  
S2: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_10  
(183 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
TOXIN_2012_2 gi 333979976 ref YP_004517921.1  beta-lysine N-acet...	28	2.5

TOXIN\_2012\_2|gi|357038678|ref|ZP\_09100475.1| putative beta-lysin... 27 2.6  
 TOXIN\_2012\_2|gi|269796843|ref|YP\_003316298.1| Pertussis toxin su... 27 4.1  
 TOXIN\_2012\_2|gi|148273353|ref|YP\_001222914.1| putative RTX toxin... 27 5.5

>TOXIN\_2012\_2|gi|333979976|ref|YP\_004517921.1| beta-lysine  
 N-acetyltransferase [Desulfotomaculum kuznetsovii DSM  
 6115] gi|333823457|gb|AEG16120.1| putative beta-lysine  
 N-acetyltransferase [Desulfotomaculum kuznetsovii DSM  
 6115]  
 Length = 280

Score = 27.7 bits (60), Expect = 2.5, Method: Compositional matrix adjust.  
 Identities = 17/73 (23%), Positives = 31/73 (42%)

Query: 97 HQRLGLGSTLYTHLLKSMEAQGFKSVVAVIGLPNDPSVRLHEALGYTARGTLRAAGYKHG 156  
 ++ + L S L H+ K + + ++ + + LGY RGTL + G  
 Sbjct: 208 YRGMALNSLLLLHIEKECLKRNINCLYSLARASSYGMNLVLHRLGYVFRGTLINNCHIDG 267

Query: 157 GWHDVGFQWQDFE 169  
 G+ ++ W R E  
 Sbjct: 268 GFENMNIWVRPVE 280

>TOXIN\_2012\_2|gi|357038678|ref|ZP\_09100475.1| putative beta-lysine  
 N-acetyltransferase [Desulfotomaculum gibsoniae DSM  
 7213] gi|355359470|gb|EHG07232.1| putative beta-lysine  
 N-acetyltransferase [Desulfotomaculum gibsoniae DSM  
 7213]  
 Length = 284

Score = 27.3 bits (59), Expect = 2.6, Method: Compositional matrix adjust.  
 Identities = 16/66 (24%), Positives = 28/66 (42%)

Query: 101 GLGSTLYTHLLKSMEAQGFKSVVAVIGLPNDPSVRLHEALGYTARGTLRAAGYKHGGWHD 160  
 G + + L K +G K + ++ + + LGY RGTL + GG+ D  
 Sbjct: 216 GFMLGIISELEKECLGKGIKCLYSLARASSYGMNLVHRLGYRFRGTLINNCHISGGYED 275

Query: 161 VGFWQR 166  
 + W +  
 Sbjct: 276 MNIWVK 281

>TOXIN\_2012\_2|gi|269796843|ref|YP\_003316298.1| Pertussis toxin  
 subunit 1 [Sanguibacter keddiei DSM 10542]  
 gi|269099028|gb|ACZ23464.1| Pertussis toxin, subunit 1  
 [Sanguibacter keddiei DSM 10542]  
 Length = 810

Score = 26.9 bits (58), Expect = 4.1, Method: Compositional matrix adjust.  
 Identities = 21/69 (30%), Positives = 28/69 (40%), Gaps = 4/69 (5%)

Query: 99 RLGLGSTLYTHL--LKSMEAQG--FKSVVAVIGLPNDPSVRLHEALGYTARGTLRAAGYK 154  
 RL L ++ L L+++ A V+ +G D L LG A G  
 Sbjct: 300 RLALAESIPADLARLEAVTASDTLISKVMETVGRSVDVGSELIPYLGAATGYALREDVD 359

Query: 155 HGGWHDVGF 163  
 HG W DVGF  
 Sbjct: 360 HGNWADVGF 368

>TOXIN\_2012\_2|gi|148273353|ref|YP\_001222914.1| putative RTX toxin  
 [Clavibacter michiganensis subsp. michiganensis NCPPB  
 382] gi|147831283|emb|CAN02239.1| putative RTX toxin

[Clavibacter michiganensis subsp. michiganensis NCPPB  
382]  
Length = 2004

Score = 26.6 bits (57), Expect = 5.5, Method: Compositional matrix adjust.  
Identities = 14/37 (37%), Positives = 19/37 (51%), Gaps = 1/37 (2%)

Query: 130 NDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFWQR 166  
DP V + G AR +R +G ++ GW D WQR  
Sbjct: 304 GDPVVVTGDVRGDPAR-PVRVSGCEYAGWADGSGWQR 339

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.320	0.135	0.439

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 2,522,842  
Number of extensions: 93013  
Number of successful extensions: 176  
Number of sequences better than 10.0: 8  
Number of HSP's gapped: 176  
Number of HSP's successfully gapped: 8  
Length of query: 183  
Length of database: 7,243,252  
Length adjustment: 89  
Effective length of query: 94  
Effective length of database: 5,169,374  
Effective search space: 485921156  
Effective search space used: 485921156  
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: 55 (25.8 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_11  
(79 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
TOXIN_2012_2 gi 302566051 pdb 3I5V A Chain A, Crystal Structure ...	26	2.6
TOXIN_2012_2 gi 300508421 pdb 3I41 A Chain A, Crystal Structure ...	26	2.7
TOXIN_2012_2 gi 320089760 pdb 3K55 A Chain A, Structure Of Beta ...	26	2.8
TOXIN_2012_2 gi 345014741 ref YP_004817095.1  Ricin B lectin [St...	26	3.2
TOXIN_2012_2 gi 57650692 ref YP_186826.1  phospholipase C [Staph...	26	3.4

>TOXIN\_2012\_2|gi|302566051|pdb|3I5V|A Chain A, Crystal Structure  
Of Beta Toxin 275-280 From Staphylococcus Aureus  
gi|302566052|pdb|3I5V|B Chain B, Crystal Structure Of  
Beta Toxin 275-280 From Staphylococcus Aureus  
gi|302566053|pdb|3I5V|C Chain C, Crystal Structure Of  
Beta Toxin 275-280 From Staphylococcus Aureus  
gi|302566054|pdb|3I5V|D Chain D, Crystal Structure Of  
Beta Toxin 275-280 From Staphylococcus Aureus  
Length = 313

Score = 26.2 bits (56), Expect = 2.6, Method: Composition-based stats.  
Identities = 11/30 (36%), Positives = 19/30 (63%), Gaps = 3/30 (10%)

Query: 15 KQNDTDLNIGFYTNHMIFINSVILYGRWRQ 44  
K++DTDL + +H +++ S +LY W Q  
Sbjct: 24 KKDDTDLKLL--VSHNVYMLSTVLYPNWGQ 50

>TOXIN\_2012\_2|gi|300508421|pdb|3I41|A Chain A, Crystal Structure  
Of Beta Toxin From Staphylococcus Aureus F277a, P278a  
Mutant gi|300508422|pdb|3I41|B Chain B, Crystal  
Structure Of Beta Toxin From Staphylococcus Aureus  
F277a, P278a Mutant gi|300508423|pdb|3I46|A Chain A,  
Crystal Structure Of Beta Toxin From Staphylococcus  
Aureus F277a, P278a Mutant With Bound Calcium Ions  
gi|300508424|pdb|3I46|B Chain B, Crystal Structure Of  
Beta Toxin From Staphylococcus Aureus F277a, P278a  
Mutant With Bound Calcium Ions gi|300508425|pdb|3I48|A  
Chain A, Crystal Structure Of Beta Toxin From  
Staphylococcus Aureus F277a, P278a Mutant With Bound  
Magnesium Ions gi|300508426|pdb|3I48|B Chain B, Crystal  
Structure Of Beta Toxin From Staphylococcus Aureus  
F277a, P278a Mutant With Bound Magnesium Ions  
Length = 317

Score = 26.2 bits (56), Expect = 2.7, Method: Composition-based stats.  
Identities = 11/30 (36%), Positives = 19/30 (63%), Gaps = 3/30 (10%)

Query: 15 KQNDTDLNIGFYTNHMIFINSVILYGRWRQ 44  
K++DTDL + +H +++ S +LY W Q  
Sbjct: 24 KKDDTDLKLL--VSHNVYMLSTVLYPNWGQ 50

>TOXIN\_2012\_2|gi|320089760|pdb|3K55|A Chain A, Structure Of Beta  
Hairpin Deletion Mutant Of Beta Toxin From  
Staphylococcus Aureus gi|320089761|pdb|3K55|B Chain B,

Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin  
 From Staphylococcus Aureus gi|320089762|pdb|3K55|C  
 Chain C, Structure Of Beta Hairpin Deletion Mutant Of  
 Beta Toxin From Staphylococcus Aureus  
 gi|320089763|pdb|3K55|D Chain D, Structure Of Beta  
 Hairpin Deletion Mutant Of Beta Toxin From  
 Staphylococcus Aureus gi|320089764|pdb|3K55|E Chain E,  
 Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin  
 From Staphylococcus Aureus gi|320089765|pdb|3K55|F  
 Chain F, Structure Of Beta Hairpin Deletion Mutant Of  
 Beta Toxin From Staphylococcus Aureus  
 gi|320089766|pdb|3K55|G Chain G, Structure Of Beta  
 Hairpin Deletion Mutant Of Beta Toxin From  
 Staphylococcus Aureus gi|320089767|pdb|3K55|H Chain H,  
 Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin  
 From Staphylococcus Aureus gi|320089768|pdb|3K55|I  
 Chain I, Structure Of Beta Hairpin Deletion Mutant Of  
 Beta Toxin From Staphylococcus Aureus  
 gi|320089769|pdb|3K55|J Chain J, Structure Of Beta  
 Hairpin Deletion Mutant Of Beta Toxin From  
 Staphylococcus Aureus gi|320089770|pdb|3K55|K Chain K,  
 Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin  
 From Staphylococcus Aureus gi|320089771|pdb|3K55|L  
 Chain L, Structure Of Beta Hairpin Deletion Mutant Of  
 Beta Toxin From Staphylococcus Aureus  
 gi|320089772|pdb|3K55|M Chain M, Structure Of Beta  
 Hairpin Deletion Mutant Of Beta Toxin From  
 Staphylococcus Aureus gi|320089773|pdb|3K55|N Chain N,  
 Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin  
 From Staphylococcus Aureus gi|320089774|pdb|3K55|O  
 Chain O, Structure Of Beta Hairpin Deletion Mutant Of  
 Beta Toxin From Staphylococcus Aureus  
 gi|320089775|pdb|3K55|P Chain P, Structure Of Beta  
 Hairpin Deletion Mutant Of Beta Toxin From  
 Staphylococcus Aureus  
 Length = 306

Score = 25.8 bits (55), Expect = 2.8, Method: Composition-based stats.  
 Identities = 11/30 (36%), Positives = 19/30 (63%), Gaps = 3/30 (10%)

Query: 15 KQNDTDLNIGFYTNHMFINSVILYGRWRQ 44  
 K++DTDL + +H +++ S +LY W Q  
 Sbjct: 24 KKDDTDLKLL--VSHNVYMLSTVLYPNWGQ 50

>TOXIN\_2012\_2|gi|345014741|ref|YP\_004817095.1| Ricin B lectin  
 [Streptomyces violaceusniger Tu 4113]  
 gi|344041090|gb|AEM86815.1| Ricin B lectin [Streptomyces  
 violaceusniger Tu 4113]  
 Length = 714

Score = 25.8 bits (55), Expect = 3.2, Method: Composition-based stats.  
 Identities = 12/26 (46%), Positives = 15/26 (57%), Gaps = 2/26 (7%)

Query: 4 QTTPTKYQVLFKQN--DTDLNIGFYT 27  
 Q +Y VLFK D+N+GFYT  
 Sbjct: 205 QFGSNRYAVLFKPGTYSADVNVGFYT 230

>TOXIN\_2012\_2|gi|57650692|ref|YP\_186826.1| phospholipase C  
 [Staphylococcus aureus subsp. aureus COL]  
 gi|282894608|ref|ZP\_06302836.1| sphingomyelin  
 phosphodiesterase [Staphylococcus aureus A8117]



gi|130085|sp|P09978.1|PHLC\_STAAU RecName:  
 Full=Phospholipase C; AltName: Full=Beta-hemolysin;  
 AltName: Full=Beta-toxin; AltName:  
 Full=Sphingomyelinase; Short=SMase; Flags: Precursor  
 gi|61213889|sp|Q5HEI1.1|PHLC\_STAAC RecName:  
 Full=Phospholipase C; AltName: Full=Beta-hemolysin;  
 AltName: Full=Beta-toxin; AltName:  
 Full=Sphingomyelinase; Short=SMase; Flags: Precursor  
 gi|110815915|sp|Q2FWP1.2|PHLC\_STAA8 RecName:  
 Full=Phospholipase C; AltName: Full=Beta-hemolysin;  
 AltName: Full=Beta-toxin; AltName:  
 Full=Sphingomyelinase; Short=SMase; Flags: Precursor  
 gi|46588|emb|CAA31769.1| unnamed protein product  
 [Staphylococcus aureus subsp. aureus COL]  
 gi|57284878|gb|AAW36972.1| phospholipase C  
 [Staphylococcus aureus subsp. aureus COL]  
 gi|269941409|emb|CBI49806.1| phospholipase C precursor  
 (pseudogene) [Staphylococcus aureus subsp. aureus TW20]  
 gi|282763095|gb|EFC03227.1| sphingomyelin  
 phosphodiesterase [Staphylococcus aureus A8117]  
 gi|298695279|gb|ADI98501.1| beta-hemolysin  
 [Staphylococcus aureus subsp. aureus ED133]  
 gi|329731457|gb|EGG67820.1| sphingomyelin  
 phosphodiesterase [Staphylococcus aureus subsp. aureus  
 21193] gi|365168342|gb|EHM59689.1| sphingomyelin  
 phosphodiesterase [Staphylococcus aureus subsp. aureus  
 21178] gi|375025481|gb|EHS18885.1| sphingomyelin  
 phosphodiesterase [Staphylococcus aureus subsp. aureus  
 IS-91]  
 Length = 330

Score = 25.8 bits (55), Expect = 3.4, Method: Composition-based stats.  
 Identities = 11/30 (36%), Positives = 19/30 (63%), Gaps = 3/30 (10%)

Query: 15 KQNDTDLNIGFYTNHMIFINSVILYGRWRQ 44  
 K++DTDL + +H +++ S +LY W Q  
 Sbjct: 37 KKDDTDLKL--VSHNVYMLSTVLYPNWGQ 63

Database: TOXIN\_2012\_2  
 Posted date: Feb 22, 2012 11:49 AM  
 Number of letters in database: 7,243,252  
 Number of sequences in database: 23,302

Lambda	K	H
0.327	0.137	0.425

Gapped		
Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
 Gap Penalties: Existence: 11, Extension: 1  
 Number of Sequences: 23302  
 Number of Hits to DB: 885,535  
 Number of extensions: 26026  
 Number of successful extensions: 67  
 Number of sequences better than 10.0: 11  
 Number of HSP's gapped: 67  
 Number of HSP's successfully gapped: 11  
 Length of query: 79  
 Length of database: 7,243,252

Length adjustment: 49  
Effective length of query: 30  
Effective length of database: 6,101,454  
Effective search space: 183043620  
Effective search space used: 183043620  
Neighboring words threshold: 11  
Window for multiple hits: 40  
X1: 15 ( 7.1 bits)  
X2: 38 (14.6 bits)  
X3: 64 (24.7 bits)  
S1: 40 (21.7 bits)  
S2: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_12  
(35 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.324	0.136	0.429

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 460,698  
Number of extensions: 7244  
Number of successful extensions: 7  
Number of sequences better than 10.0: 0  
Number of HSP's gapped: 7  
Number of HSP's successfully gapped: 0  
Length of query: 35  
Length of database: 7,243,252  
Length adjustment: 9  
Effective length of query: 26  
Effective length of database: 7,033,534

Effective search space: 182871884  
Effective search space used: 182871884  
Neighboring words threshold: 11  
Window for multiple hits: 40  
X1: 15 ( 7.0 bits)  
X2: 38 (14.6 bits)  
X3: 64 (24.7 bits)  
S1: 40 (21.6 bits)  
S2: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_13  
(40 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.319	0.135	0.454

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 625,904  
Number of extensions: 14096  
Number of successful extensions: 4  
Number of sequences better than 10.0: 0  
Number of HSP's gapped: 4  
Number of HSP's successfully gapped: 0  
Length of query: 40  
Length of database: 7,243,252  
Length adjustment: 14  
Effective length of query: 26  
Effective length of database: 6,917,024  
Effective search space: 179842624  
Effective search space used: 179842624  
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 (21.7 bits)  
S2: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_14  
(51 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
TOXIN_2012_2 gi 374599467 ref ZP_09672469.1  toxic anion resista...	25	4.4
>TOXIN_2012_2 gi 374599467 ref ZP_09672469.1  toxic anion resistance family protein [Myroides odoratus DSM 2801] gi 373910937 gb EHQ42786.1  toxic anion resistance family protein [Myroides odoratus DSM 2801] Length = 357		

Score = 25.4 bits (54), Expect = 4.4, Method: Composition-based stats.  
Identities = 10/32 (31%), Positives = 17/32 (53%)

Query: 4 VWKGELEFRVQVHRGLREVTGSQAVCRCSEEL 35  
VWK +L V +HR + + Q + +EE+  
Sbjct: 251 VWKNQLTLAVAMHRQQQHIEVQQKISNTTEEI 282

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.327	0.135	0.458

Gapped  
Lambda

K	H
0.267	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1

Number of Sequences: 23302  
 Number of Hits to DB: 640,379  
 Number of extensions: 12902  
 Number of successful extensions: 18  
 Number of sequences better than 10.0: 1  
 Number of HSP's gapped: 18  
 Number of HSP's successfully gapped: 1  
 Length of query: 51  
 Length of database: 7,243,252  
 Length adjustment: 24  
 Effective length of query: 27  
 Effective length of database: 6,684,004  
 Effective search space: 180468108  
 Effective search space used: 180468108  
 Neighboring words threshold: 11  
 Window for multiple hits: 40  
 X1: 15 ( 7.1 bits)  
 X2: 38 (14.6 bits)  
 X3: 64 (24.7 bits)  
 S1: 40 (21.8 bits)  
 S2: 51 (24.3 bits)  
 BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_15  
 (81 letters)

Database: TOXIN\_2012\_2  
 23,302 sequences; 7,243,252 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
TOXIN_2012_2 gi 71281801 ref YP_268598.1  perfringolysin O regul...	27	2.1
TOXIN_2012_2 gi 52140196 ref YP_086635.1  membrane protein; regu...	24	9.6

>TOXIN\_2012\_2|gi|71281801|ref|YP\_268598.1| perfringolysin O  
 regulator protein [Colwellia psychrerythraea 34H]  
 gi|71147541|gb|AAZ28014.1| perfringolysin O regulator  
 protein [Colwellia psychrerythraea 34H]  
 Length = 343

Score = 26.6 bits (57), Expect = 2.1, Method: Composition-based stats.  
 Identities = 11/22 (50%), Positives = 15/22 (68%)

Query: 9 SGNLLGFHCPAICHFIVKIVEK 30  
 +G L GF IC+FI K++EK  
 Sbjct: 101 TGILAGFIAGYICYFISKVIEK 122

```

>TOXIN_2012_2|gi|52140196|ref|YP_086635.1|    membrane protein;
regulatory protein [Bacillus cereus E33L]
gi|196036525|ref|ZP_03103920.1| membrane protein Pfor
[Bacillus cereus W] gi|196040980|ref|ZP_03108277.1|
membrane protein Pfor [Bacillus cereus NVH0597-99]
gi|196045481|ref|ZP_03112712.1| membrane protein Pfor
[Bacillus cereus 03BB108]
gi|218906545|ref|YP_002454379.1| membrane protein Pfor
[Bacillus cereus AH820] gi|225867349|ref|YP_002752727.1|
membrane protein Pfor [Bacillus cereus 03BB102]
gi|228917980|ref|ZP_04081512.1| Transcriptional
regulator pfor [Bacillus thuringiensis serovar
pulsiensis BGSC 4CC1] gi|228936653|ref|ZP_04099446.1|
Transcriptional regulator pfor [Bacillus thuringiensis
serovar andalousiensis BGSC 4AW1]
gi|228949092|ref|ZP_04111362.1| Transcriptional
regulator pfor [Bacillus thuringiensis serovar monterrey
BGSC 4AJ1] gi|229187595|ref|ZP_04314733.1|
Transcriptional regulator pfor [Bacillus cereus BGSC
6E1] gi|301056836|ref|YP_003795047.1| unnamed protein
product [Bacillus cereus biovar anthracis str. CI]
gi|51973665|gb|AAU15215.1| membrane protein; possible
regulatory protein [Bacillus cereus E33L]
gi|195990858|gb|EDX54831.1| membrane protein Pfor
[Bacillus cereus W] gi|196023688|gb|EDX62364.1| membrane
protein Pfor [Bacillus cereus 03BB108]
gi|196028148|gb|EDX66758.1| membrane protein Pfor
[Bacillus cereus NVH0597-99] gi|218535498|gb|ACK87896.1|
membrane protein Pfor [Bacillus cereus AH820]
gi|225790834|gb|ACO31051.1| membrane protein Pfor
[Bacillus cereus 03BB102] gi|228595847|gb|EEK53529.1|
Transcriptional regulator pfor [Bacillus cereus BGSC
6E1] gi|228810533|gb|EEM56884.1| Transcriptional
regulator pfor [Bacillus thuringiensis serovar monterrey
BGSC 4AJ1] gi|228822992|gb|EEM68831.1| Transcriptional
regulator pfor [Bacillus thuringiensis serovar
andalousiensis BGSC 4AW1] gi|228841645|gb|EEM86757.1|
Transcriptional regulator pfor [Bacillus thuringiensis
serovar pulsiensis BGSC 4CC1]
gi|300379005|gb|ADK07909.1| putative membrane protein
[Bacillus cereus biovar anthracis str. CI]
gi|364515083|gb|AEW58482.1| Perfringolysin O regulator
protein Pfor [Bacillus cereus F837/76]
Length = 345

```

Score = 24.3 bits (51), Expect = 9.6, Method: Composition-based stats.  
Identities = 10/22 (45%), Positives = 14/22 (63%)

```

Query: 10  GNLLGFHCPAICHFIVKIVEKE 31
          G L GF      +C F+VK +EK+
Sbjct: 102 GILPGFFAGYVCAFVVKFLEKK 123

```

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.316	0.135	0.416

Gapped

Lambda      K      H  
0.267      0.0410      0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 1,185,906  
Number of extensions: 40954  
Number of successful extensions: 78  
Number of sequences better than 10.0: 2  
Number of HSP's gapped: 78  
Number of HSP's successfully gapped: 2  
Length of query: 81  
Length of database: 7,243,252  
Length adjustment: 51  
Effective length of query: 30  
Effective length of database: 6,054,850  
Effective search space: 181645500  
Effective search space used: 181645500  
Neighboring words threshold: 11  
Window for multiple hits: 40  
X1: 16 ( 7.3 bits)  
X2: 38 (14.6 bits)  
X3: 64 (24.7 bits)  
S1: 41 (21.6 bits)  
S2: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_16  
(183 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
TOXIN_2012_2 gi 333979976 ref YP_004517921.1  beta-lysine N-acet...	28	2.5
TOXIN_2012_2 gi 357038678 ref ZP_09100475.1  putative beta-lysin...	27	2.6
TOXIN_2012_2 gi 269796843 ref YP_003316298.1  Pertussis toxin su...	27	4.1
TOXIN_2012_2 gi 148273353 ref YP_001222914.1  putative RTX toxin...	27	5.5

>TOXIN\_2012\_2|gi|333979976|ref|YP\_004517921.1| beta-lysine  
N-acetyltransferase [Desulfotomaculum kuznetsovii DSM  
6115] gi|333823457|gb|AEG16120.1| putative beta-lysine  
N-acetyltransferase [Desulfotomaculum kuznetsovii DSM  
6115]

Length = 280

Score = 27.7 bits (60), Expect = 2.5, Method: Compositional matrix adjust.  
Identities = 17/73 (23%), Positives = 31/73 (42%)

Query: 97 HQRLGLGSTLYTHLLKSMEAQGFKSVVAVIGLPNDPSVRLHEALGYTARGTLRAAGYKHG 156  
++ + L S L H+ K + + ++ + + LGY RGT L + G  
Sbjct: 208 YRGMALNSLLLLHIEKECLKRNINCLYSLARASSYGMNLVLHRLGYVFRGT LINNCHIDG 267

Query: 157 GWHDVGFWRDVE 169  
G+ ++ W R E  
Sbjct: 268 GFENMNIWVRPVE 280

>TOXIN\_2012\_2|gi|357038678|ref|ZP\_09100475.1| putative beta-lysine  
N-acetyltransferase [Desulfotomaculum gibsoniae DSM  
7213] gi|355359470|gb|EHG07232.1| putative beta-lysine  
N-acetyltransferase [Desulfotomaculum gibsoniae DSM  
7213]  
Length = 284

Score = 27.3 bits (59), Expect = 2.6, Method: Compositional matrix adjust.  
Identities = 16/66 (24%), Positives = 28/66 (42%)

Query: 101 GLGSTLYTHLLKSMEAQGFKSVVAVIGLPNDPSVRLHEALGYTARGTLRAAGYKHGGWHD 160  
G + + L K +G K + ++ + + LGY RGT L + GG+ D  
Sbjct: 216 GFMLGIISELEKECLGKGIKCLYSLARASSYGMNLVHRLGYRFRGT LINNCHISGGYED 275

Query: 161 VGFWRQ 166  
+ W +  
Sbjct: 276 MNIWVK 281

>TOXIN\_2012\_2|gi|269796843|ref|YP\_003316298.1| Pertussis toxin  
subunit 1 [Sanguibacter keddieii DSM 10542]  
gi|269099028|gb|ACZ23464.1| Pertussis toxin, subunit 1  
[Sanguibacter keddieii DSM 10542]  
Length = 810

Score = 26.9 bits (58), Expect = 4.1, Method: Compositional matrix adjust.  
Identities = 21/69 (30%), Positives = 28/69 (40%), Gaps = 4/69 (5%)

Query: 99 RLGLGSTLYTHL--LKSMEAQG--FKSVVAVIGLPNDPSVRLHEALGYTARGTLRAAGYK 154  
RL L ++ L L+++ A V+ +G D L LG A G  
Sbjct: 300 RLALAESIPADLARLEAVTASDTLISKVMETVGRSVDVGSELIPYLGVAATGYALREDVD 359

Query: 155 HGGWHDVGF 163  
HG W DVGF  
Sbjct: 360 HGNWADVGF 368

>TOXIN\_2012\_2|gi|148273353|ref|YP\_001222914.1| putative RTX toxin  
[Clavibacter michiganensis subsp. michiganensis NCPPB  
382] gi|147831283|emb|CAN02239.1| putative RTX toxin  
[Clavibacter michiganensis subsp. michiganensis NCPPB  
382]  
Length = 2004

Score = 26.6 bits (57), Expect = 5.5, Method: Compositional matrix adjust.  
Identities = 14/37 (37%), Positives = 19/37 (51%), Gaps = 1/37 (2%)



Query: 130 NDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFWR 166  
DP V + G AR +R +G ++ GW D WQR  
Sbjct: 304 GDPVVVTGDVRGDPAR-PVRVSGCEYAGWADGSGWQR 339

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda K H  
0.320 0.135 0.439

Gapped  
Lambda K H  
0.267 0.0410 0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 2,522,842  
Number of extensions: 93013  
Number of successful extensions: 176  
Number of sequences better than 10.0: 8  
Number of HSP's gapped: 176  
Number of HSP's successfully gapped: 8  
Length of query: 183  
Length of database: 7,243,252  
Length adjustment: 89  
Effective length of query: 94  
Effective length of database: 5,169,374  
Effective search space: 485921156  
Effective search space used: 485921156  
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: 55 (25.8 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_17  
(35 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
 Posted date: Feb 22, 2012 11:49 AM  
 Number of letters in database: 7,243,252  
 Number of sequences in database: 23,302

Lambda	K	H
0.320	0.130	0.398

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
 Gap Penalties: Existence: 11, Extension: 1  
 Number of Sequences: 23302  
 Number of Hits to DB: 502,057  
 Number of extensions: 8645  
 Number of successful extensions: 7  
 Number of sequences better than 10.0: 0  
 Number of HSP's gapped: 7  
 Number of HSP's successfully gapped: 0  
 Length of query: 35  
 Length of database: 7,243,252  
 Length adjustment: 9  
 Effective length of query: 26  
 Effective length of database: 7,033,534  
 Effective search space: 182871884  
 Effective search space used: 182871884  
 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: 51 (24.3 bits)  
 BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_18  
 (64 letters)

Database: TOXIN\_2012\_2  
 23,302 sequences; 7,243,252 total letters

Searching.....done

Sequences producing significant alignments:	Score	E
	(bits)	Value

TOXIN\_2012\_2|gi|23015941|ref|ZP\_00055704.1| COG2931: RTX toxins ... 25 4.2

>TOXIN\_2012\_2|gi|23015941|ref|ZP\_00055704.1| COG2931: RTX toxins and  
related Ca2+-binding proteins [Magnetospirillum  
magnetotacticum MS-1]  
Length = 12295

Score = 25.4 bits (54), Expect = 4.2, Method: Composition-based stats.  
Identities = 10/21 (47%), Positives = 12/21 (57%)

Query: 2 RLWDTQPGVHCAQLDTSMVDG 22  
R WD G H A++DTS G  
Sbjct: 6416 RAWDQTSNGTHGAKVDTSTNGG 6436

Score = 25.4 bits (54), Expect = 4.4, Method: Composition-based stats.  
Identities = 10/21 (47%), Positives = 13/21 (61%)

Query: 2 RLWDTQPGVHCAQLDTSMVDG 22  
R WD G H A++DTS+ G  
Sbjct: 6901 RAWDQTSNGTHGAKVDTSVNGG 6921

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.322	0.135	0.452

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 740,630  
Number of extensions: 14869  
Number of successful extensions: 9  
Number of sequences better than 10.0: 1  
Number of HSP's gapped: 9  
Number of HSP's successfully gapped: 2  
Length of query: 64  
Length of database: 7,243,252  
Length adjustment: 36  
Effective length of query: 28  
Effective length of database: 6,404,380  
Effective search space: 179322640  
Effective search space used: 179322640  
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: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_19  
(57 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.322	0.125	0.446

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 572,235  
Number of extensions: 9841  
Number of successful extensions: 11  
Number of sequences better than 10.0: 0  
Number of HSP's gapped: 11  
Number of HSP's successfully gapped: 0  
Length of query: 57  
Length of database: 7,243,252  
Length adjustment: 30  
Effective length of query: 27  
Effective length of database: 6,544,192  
Effective search space: 176693184  
Effective search space used: 176693184  
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: 40 (21.6 bits)  
S2: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_20  
(40 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.325	0.138	0.504

Gapped Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 618,726  
Number of extensions: 14460  
Number of successful extensions: 55  
Number of sequences better than 10.0: 1  
Number of HSP's gapped: 55  
Number of HSP's successfully gapped: 1  
Length of query: 40  
Length of database: 7,243,252  
Length adjustment: 14  
Effective length of query: 26  
Effective length of database: 6,917,024  
Effective search space: 179842624  
Effective search space used: 179842624  
Neighboring words threshold: 11  
Window for multiple hits: 40  
X1: 15 ( 7.0 bits)  
X2: 38 (14.6 bits)  
X3: 64 (24.7 bits)  
S1: 40 (21.6 bits)  
S2: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_21  
(74 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.325	0.140	0.469

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 847,053  
Number of extensions: 19112  
Number of successful extensions: 21  
Number of sequences better than 10.0: 0  
Number of HSP's gapped: 21  
Number of HSP's successfully gapped: 0  
Length of query: 74  
Length of database: 7,243,252  
Length adjustment: 45  
Effective length of query: 29  
Effective length of database: 6,194,662  
Effective search space: 179645198  
Effective search space used: 179645198  
Neighboring words threshold: 11  
Window for multiple hits: 40  
X1: 15 ( 7.0 bits)  
X2: 38 (14.6 bits)  
X3: 64 (24.7 bits)  
S1: 40 (21.6 bits)  
S2: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_22  
(439 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
TOXIN_2012_2 gi 90406967 ref ZP_01215158.1  putative RTX toxin [...	34	0.089
TOXIN_2012_2 gi 149190738 ref ZP_01869004.1  putative RTX toxin ...	28	5.4
>TOXIN_2012_2 gi 90406967 ref ZP_01215158.1  putative RTX toxin		
[Psychromonas sp. CNPT3] gi 90312009 gb EAS40103.1		
putative RTX toxin [Psychromonas sp. CNPT3]		
Length = 3350		

Score = 34.3 bits (77), Expect = 0.089, Method: Composition-based stats.  
Identities = 19/53 (35%), Positives = 32/53 (60%), Gaps = 3/53 (5%)

Query: 168 FGLAEVELYGDVVLRFV--SYPDETDLPLPGFERVSSPGAVDYGLTRFDHVV 218  
+G+ EVE+ G+ V + YP++T++ F+P E + S D G+ FDHV+  
Sbjct: 2038 YGVVEVEIEGNWVAMQLGEKYPEDTNVRFVPDTEAIESI-TKDIGVGSFDHVL 2089

>TOXIN\_2012\_2|gi|149190738|ref|ZP\_01869004.1| putative RTX toxin  
[Vibrio shilonii AK1] gi|148835394|gb|EDL52365.1|  
putative RTX toxin [Vibrio shilonii AK1]  
Length = 2496

Score = 28.1 bits (61), Expect = 5.4, Method: Composition-based stats.  
Identities = 17/44 (38%), Positives = 23/44 (52%), Gaps = 5/44 (11%)

Query: 52 DAASAAGRFSFALGAPLAARSDLSTGNSAHASLLLRSGALAF LF 95  
D+AS AG+F ALGA DL T S +AS+ G ++  
Sbjct: 954 DSASGAGQFDVALGA----DDLDTNISNNASVYFEIGEQLPVY 992

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.321	0.137	0.402

Gapped Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 5,017,141  
Number of extensions: 197387  
Number of successful extensions: 560

Number of sequences better than 10.0: 3  
 Number of HSP's gapped: 560  
 Number of HSP's successfully gapped: 3  
 Length of query: 439  
 Length of database: 7,243,252  
 Length adjustment: 98  
 Effective length of query: 341  
 Effective length of database: 4,959,656  
 Effective search space: 1691242696  
 Effective search space used: 1691242696  
 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: 59 (27.3 bits)  
 BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_23  
 (30 letters)

Database: TOXIN\_2012\_2  
 23,302 sequences; 7,243,252 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
TOXIN_2012_2 gi 326664496 ref XP_001920822.3  PREDICTED: neoverr...	24	8.2
>TOXIN_2012_2 gi 326664496 ref XP_001920822.3  PREDICTED:		
neoverrucotoxin subunit beta [Danio rerio]		
Length = 347		

Score = 24.3 bits (51), Expect = 8.2, Method: Composition-based stats.  
 Identities = 12/27 (44%), Positives = 15/27 (55%)

Query: 4 HDTLVYSKNIKDTVSEDQRAIETFQQR 30  
 HD LV SKNIK T +D + +R  
 Sbjct: 185 HDHLVMSKNIKATYVKDPQPYTDHPPER 211

Database: TOXIN\_2012\_2  
 Posted date: Feb 22, 2012 11:49 AM  
 Number of letters in database: 7,243,252  
 Number of sequences in database: 23,302

Lambda K H



0.313      0.126      0.327

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62

Gap Penalties: Existence: 11, Extension: 1

Number of Sequences: 23302

Number of Hits to DB: 351,579

Number of extensions: 4464

Number of successful extensions: 11

Number of sequences better than 10.0: 1

Number of HSP's gapped: 11

Number of HSP's successfully gapped: 1

Length of query: 30

Length of database: 7,243,252

Length adjustment: 5

Effective length of query: 25

Effective length of database: 7,126,742

Effective search space: 178168550

Effective search space used: 178168550

Neighboring words threshold: 11

Window for multiple hits: 40

X1: 16 ( 7.2 bits)

X2: 38 (14.6 bits)

X3: 64 (24.7 bits)

S1: 42 (22.0 bits)

S2: 51 (24.3 bits)

BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_24  
(40 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2

Posted date: Feb 22, 2012 11:49 AM

Number of letters in database: 7,243,252

Number of sequences in database: 23,302

Lambda	K	H
0.319	0.135	0.454

Gapped

Lambda      K      H  
0.267      0.0410      0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 625,904  
Number of extensions: 14096  
Number of successful extensions: 4  
Number of sequences better than 10.0: 0  
Number of HSP's gapped: 4  
Number of HSP's successfully gapped: 0  
Length of query: 40  
Length of database: 7,243,252  
Length adjustment: 14  
Effective length of query: 26  
Effective length of database: 6,917,024  
Effective search space: 179842624  
Effective search space used: 179842624  
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 (21.7 bits)  
S2: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_25  
(65 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda      K      H  
0.330      0.142      0.487

Gapped  
Lambda      K      H  
0.267      0.0410      0.140

Matrix: BLOSUM62  
 Gap Penalties: Existence: 11, Extension: 1  
 Number of Sequences: 23302  
 Number of Hits to DB: 614,280  
 Number of extensions: 11107  
 Number of successful extensions: 36  
 Number of sequences better than 10.0: 0  
 Number of HSP's gapped: 36  
 Number of HSP's successfully gapped: 0  
 Length of query: 65  
 Length of database: 7,243,252  
 Length adjustment: 37  
 Effective length of query: 28  
 Effective length of database: 6,381,078  
 Effective search space: 178670184  
 Effective search space used: 178670184  
 Neighboring words threshold: 11  
 Window for multiple hits: 40  
 X1: 15 ( 7.1 bits)  
 X2: 38 (14.6 bits)  
 X3: 64 (24.7 bits)  
 S1: 40 (21.8 bits)  
 S2: 51 (24.3 bits)  
 BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_26  
 (69 letters)

Database: TOXIN\_2012\_2  
 23,302 sequences; 7,243,252 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
TOXIN_2012_2 gi 23015941 ref ZP_00055704.1  COG2931: RTX toxins ...	25	5.4
>TOXIN_2012_2 gi 23015941 ref ZP_00055704.1  COG2931: RTX toxins and related Ca2+-binding proteins [Magnetospirillum magnetotacticum MS-1] Length = 12295		

Score = 25.0 bits (53), Expect = 5.4, Method: Composition-based stats.  
 Identities = 10/21 (47%), Positives = 12/21 (57%)

Query: 2 RLWDTQPGVHCAQLDTSMVDG 22  
 R WD G H A++DTS G  
 Sbjct: 6416 RAWDQTSNGTHAKVDTSTNGG 6436

Score = 25.0 bits (53), Expect = 5.7, Method: Composition-based stats.  
Identities = 10/21 (47%), Positives = 13/21 (61%)

Query: 2 RLWDTQPGVHCAQLDTSMVDG 22  
R WD G H A++DTS+ G  
Sbjct: 6901 RAWDQTSGSHGAKVDTSVNGG 6921

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.324	0.137	0.443

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 767,387  
Number of extensions: 15429  
Number of successful extensions: 11  
Number of sequences better than 10.0: 1  
Number of HSP's gapped: 11  
Number of HSP's successfully gapped: 2  
Length of query: 69  
Length of database: 7,243,252  
Length adjustment: 40  
Effective length of query: 29  
Effective length of database: 6,311,172  
Effective search space: 183023988  
Effective search space used: 183023988  
Neighboring words threshold: 11  
Window for multiple hits: 40  
X1: 15 ( 7.0 bits)  
X2: 38 (14.6 bits)  
X3: 64 (24.7 bits)  
S1: 40 (21.6 bits)  
S2: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHTOH2\_I\_27  
(35 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.315	0.130	0.353

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 384,936  
Number of extensions: 6200  
Number of successful extensions: 6  
Number of sequences better than 10.0: 0  
Number of HSP's gapped: 6  
Number of HSP's successfully gapped: 0  
Length of query: 35  
Length of database: 7,243,252  
Length adjustment: 9  
Effective length of query: 26  
Effective length of database: 7,033,534  
Effective search space: 182871884  
Effective search space used: 182871884  
Neighboring words threshold: 11  
Window for multiple hits: 40  
X1: 16 ( 7.3 bits)  
X2: 38 (14.6 bits)  
X3: 64 (24.7 bits)  
S1: 41 (21.6 bits)  
S2: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_28  
(183 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
TOXIN_2012_2 gi 333979976 ref YP_004517921.1  beta-lysine N-acet...	28	2.5
TOXIN_2012_2 gi 357038678 ref ZP_09100475.1  putative beta-lysin...	27	2.6
TOXIN_2012_2 gi 269796843 ref YP_003316298.1  Pertussis toxin su...	27	4.1
TOXIN_2012_2 gi 148273353 ref YP_001222914.1  putative RTX toxin...	27	5.5

>TOXIN\_2012\_2|gi|333979976|ref|YP\_004517921.1| beta-lysine  
N-acetyltransferase [Desulfotomaculum kuznetsovii DSM  
6115] gi|333823457|gb|AEG16120.1| putative beta-lysine  
N-acetyltransferase [Desulfotomaculum kuznetsovii DSM  
6115]  
Length = 280

Score = 27.7 bits (60), Expect = 2.5, Method: Compositional matrix adjust.  
Identities = 17/73 (23%), Positives = 31/73 (42%)

Query: 97 HQRLGLGSTLYTHLLKSMEAQGFKSVVAVIGLPNDPSVRLHEALGYTARGTLRAAGYKHG 156  
++ + L S L H+ K + + ++ + + LGY RGT L + G  
Sbjct: 208 YRGMALNSLLLLHIEKECLKRNINCLYSLARASSYGMNLVLHRLGYVFRGT LINNCHIDG 267

Query: 157 GWHDVGFWRDQFE 169  
G+ ++ W R E  
Sbjct: 268 GFENMNIWVRPVE 280

>TOXIN\_2012\_2|gi|357038678|ref|ZP\_09100475.1| putative beta-lysine  
N-acetyltransferase [Desulfotomaculum gibsoniae DSM  
7213] gi|355359470|gb|EHG07232.1| putative beta-lysine  
N-acetyltransferase [Desulfotomaculum gibsoniae DSM  
7213]  
Length = 284

Score = 27.3 bits (59), Expect = 2.6, Method: Compositional matrix adjust.  
Identities = 16/66 (24%), Positives = 28/66 (42%)

Query: 101 GLGSTLYTHLLKSMEAQGFKSVVAVIGLPNDPSVRLHEALGYTARGTLRAAGYKHGGWHD 160  
G + + L K +G K + ++ + + LGY RGT L + GG+ D  
Sbjct: 216 GFMLGIISELEKECLGKGIKCLYSLARASSYGMNLVFRHRLGYRFRGT LINNCHISGGYED 275

Query: 161 VGFWQR 166  
+ W +  
Sbjct: 276 MNIWVK 281

>TOXIN\_2012\_2|gi|269796843|ref|YP\_003316298.1| Pertussis toxin  
subunit 1 [Sanguibacter keddieii DSM 10542]  
gi|269099028|gb|ACZ23464.1| Pertussis toxin, subunit 1  
[Sanguibacter keddieii DSM 10542]  
Length = 810

Score = 26.9 bits (58), Expect = 4.1, Method: Compositional matrix adjust.  
Identities = 21/69 (30%), Positives = 28/69 (40%), Gaps = 4/69 (5%)

Query: 99 RLGLGSTLYTHL--LKSMEAQG--FKSVVAVIGLPNDPSVRLHEALGYTARGTLRAAGYK 154  
RL L ++ L L+++ A V+ +G D L LG A G  
Sbjct: 300 RLALAESIPADLARLEAVTASDTLISKVMETVGRSVDVGSELIPYLGAATGYALREDVD 359

Query: 155 HGGWHDVGF 163  
HG W DVGF  
Sbjct: 360 HGNWADVGF 368

>TOXIN\_2012\_2|gi|148273353|ref|YP\_001222914.1| putative RTX toxin  
[Clavibacter michiganensis subsp. michiganensis NCPPB  
382] gi|147831283|emb|CAN02239.1| putative RTX toxin  
[Clavibacter michiganensis subsp. michiganensis NCPPB  
382]  
Length = 2004

Score = 26.6 bits (57), Expect = 5.5, Method: Compositional matrix adjust.  
Identities = 14/37 (37%), Positives = 19/37 (51%), Gaps = 1/37 (2%)

Query: 130 NDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFQWR 166  
DP V + G AR +R +G ++ GW D WQR  
Sbjct: 304 GDPVVVTGDRVGPDPAR-PVRVSGCEYAGWADGSGWQR 339

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.320	0.135	0.439

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 2,522,842  
Number of extensions: 93013  
Number of successful extensions: 176  
Number of sequences better than 10.0: 8  
Number of HSP's gapped: 176  
Number of HSP's successfully gapped: 8  
Length of query: 183  
Length of database: 7,243,252  
Length adjustment: 89  
Effective length of query: 94  
Effective length of database: 5,169,374  
Effective search space: 485921156  
Effective search space used: 485921156  
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: 55 (25.8 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_29  
(32 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.323	0.131	0.412

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 421,053  
Number of extensions: 6789  
Number of successful extensions: 9  
Number of sequences better than 10.0: 0  
Number of HSP's gapped: 9  
Number of HSP's successfully gapped: 0  
Length of query: 32  
Length of database: 7,243,252  
Length adjustment: 7  
Effective length of query: 25  
Effective length of database: 7,080,138  
Effective search space: 177003450  
Effective search space used: 177003450  
Neighboring words threshold: 11  
Window for multiple hits: 40  
X1: 16 ( 7.5 bits)  
X2: 38 (14.6 bits)  
X3: 64 (24.7 bits)  
S1: 40 (21.6 bits)  
S2: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_30  
(44 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
TOXIN_2012_2 gi 282856127 ref ZP_06265411.1  addiction module to...	27	2.0
>TOXIN_2012_2 gi 282856127 ref ZP_06265411.1  addiction module toxin, RelE/StbE [Pyramidobacter piscolens W5455] gi 282586054 gb EFB91338.1  addiction module toxin, RelE/StbE [Pyramidobacter piscolens W5455] Length = 90		

Score = 26.6 bits (57), Expect = 2.0, Method: Compositional matrix adjust.  
Identities = 11/22 (50%), Positives = 15/22 (68%)

Query: 4 QFRQDSSRISEPHLHTAPLQQV 25  
QFR+D RI + L TAPL++  
Sbjct: 9 QFRRDVKRIKKQGLDTAPLKEA 30

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.322	0.132	0.387

Gapped Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 322,500  
Number of extensions: 4034  
Number of successful extensions: 1  
Number of sequences better than 10.0: 1  
Number of HSP's gapped: 1  
Number of HSP's successfully gapped: 1  
Length of query: 44  
Length of database: 7,243,252  
Length adjustment: 18  
Effective length of query: 26  
Effective length of database: 6,823,816  
Effective search space: 177419216  
Effective search space used: 177419216  
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 (22.0 bits)  
S2: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_31  
(57 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
TOXIN_2012_2 gi 347602872 gb EGY27821.1  RTX toxin [Candidatus R...	25	4.6
>TOXIN_2012_2 gi 347602872 gb EGY27821.1  RTX toxin [Candidatus Regiella insecticola R5.15] Length = 1696		

Score = 25.0 bits (53), Expect = 4.6, Method: Composition-based stats.  
Identities = 10/19 (52%), Positives = 14/19 (73%)

Query: 12 IRGRFGSEIVSWFHEQARS 30  
+RG+ GS+I +HEQA S  
Sbjct: 1374 LRGKAGSDIYHIYHEQAGS 1392

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.332	0.141	0.524

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 685,479  
Number of extensions: 15776

Number of successful extensions: 28  
 Number of sequences better than 10.0: 2  
 Number of HSP's gapped: 28  
 Number of HSP's successfully gapped: 2  
 Length of query: 57  
 Length of database: 7,243,252  
 Length adjustment: 30  
 Effective length of query: 27  
 Effective length of database: 6,544,192  
 Effective search space: 176693184  
 Effective search space used: 176693184  
 Neighboring words threshold: 11  
 Window for multiple hits: 40  
 X1: 15 ( 7.2 bits)  
 X2: 38 (14.6 bits)  
 X3: 64 (24.7 bits)  
 S1: 40 (22.0 bits)  
 S2: 51 (24.3 bits)  
 BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_32  
 (121 letters)

Database: TOXIN\_2012\_2  
 23,302 sequences; 7,243,252 total letters

Searching.....done

			Score	E
Sequences producing significant alignments:			(bits)	Value
TOXIN_2012_2	gi 157951698 ref NP_033910.2	complement C4-B precu...	27	1.7
TOXIN_2012_2	gi 29337194 sp P08649.3 CO4_RAT	RecName: Full=Compl...	25	4.5
TOXIN_2012_2	gi 317025490 ref XP_001389187.2	toxin biosynthesis...	25	5.4
TOXIN_2012_2	gi 238492287 ref XP_002377380.1	MFS gliotoxin effl...	25	6.8
TOXIN_2012_2	gi 358367040 dbj GAA83660.1	toxin biosynthesis pro...	25	7.5
TOXIN_2012_2	gi 315081364 gb EFT53340.1	toxin secretion/phage l...	24	9.7
TOXIN_2012_2	gi 335050940 ref ZP_08543885.1	toxin secretion/pha...	24	9.8
TOXIN_2012_2	gi 313812423 gb EFS50137.1	toxin secretion/phage l...	24	9.8
TOXIN_2012_2	gi 327450317 gb EGE96971.1	toxin secretion/phage l...	24	9.9
TOXIN_2012_2	gi 289426084 ref ZP_06427831.1	toxin secretion/pha...	24	9.9

>TOXIN\_2012\_2|gi|157951698|ref|NP\_033910.2| complement C4-B  
 precursor [Mus musculus]  
 gi|341940526|sp|P01029.3|CO4B\_MOUSE RecName:  
 Full=Complement C4-B; Contains: RecName: Full=Complement  
 C4 beta chain; Contains: RecName: Full=Complement C4  
 alpha chain; Contains: RecName: Full=C4a anaphylatoxin;  
 Contains: RecName: Full=Complement C4 gamma chain;  
 Flags: Precursor

Length = 1738

Score = 26.9 bits (58), Expect = 1.7, Method: Composition-based stats.  
Identities = 25/93 (26%), Positives = 37/93 (39%), Gaps = 11/93 (11%)

Query: 35 VQIWVTGLTGL--GGAGNSKSLCQKPTSCHPPCLYP---AARNVPRAVYPKASCNLTDG 88  
V + VT + GL G G P P + AA NVP V + +L D  
Sbjct: 863 VSVHVTPEGLCLAGGGMAQQVTVPAGSARPVAFSVVPTAAANVPLKVVARGVFDLGD 922

Query: 89 SFGRPITATTDLKPCASIDLSKCVYNVDPRPNL 121  
++ ++ +I + VYN+DP NL  
Sbjct: 923 -----VSKILQIEKEGAIHREELVYNLDPLNNL 950

>TOXIN\_2012\_2|gi|29337194|sp|P08649.3|CO4\_RAT RecName:  
Full=Complement C4; Contains: RecName: Full=Complement  
C4 beta chain; Contains: RecName: Full=Complement C4  
alpha chain; Contains: RecName: Full=C4a anaphylatoxin;  
Contains: RecName: Full=Complement C4 gamma chain;  
Flags: Precursor gi|25244410|gb|AAN72415.1| complement  
component C4 [Rattus norvegicus]  
Length = 1737

Score = 25.4 bits (54), Expect = 4.5, Method: Composition-based stats.  
Identities = 18/60 (30%), Positives = 30/60 (50%), Gaps = 7/60 (11%)

Query: 69 AARNVPRAVYPKASCNL-----TDGSF--GRPITATTDLKPCASIDLSKCVYNVDPRPNL 121  
+AR V +V P A+ ++ GSF G ++ ++ +I + VYN+DP NL  
Sbjct: 890 SARPVAFSVVPTAAASIPLKVVARGSFITIGDAVSKILQIEKEGAIHREIIVYNLDPLNNL 949

>TOXIN\_2012\_2|gi|317025490|ref|XP\_001389187.2| toxin biosynthesis  
protein [Aspergillus niger CBS 513.88]  
Length = 446

Score = 25.0 bits (53), Expect = 5.4, Method: Composition-based stats.  
Identities = 11/30 (36%), Positives = 16/30 (53%)

Query: 71 RNVPRAYPKASCNLTDGSFGRPITATTDL 100  
R +P A++P +GS RP+T T L  
Sbjct: 241 RELPTAIHPLDESQQKEGSNDRPVTLRTTL 270

>TOXIN\_2012\_2|gi|238492287|ref|XP\_002377380.1| MFS gliotoxin efflux  
transporter GliA [Aspergillus flavus NRRL3357]  
gi|220695874|gb|EED52216.1| MFS gliotoxin efflux  
transporter GliA [Aspergillus flavus NRRL3357]  
Length = 523

Score = 25.0 bits (53), Expect = 6.8, Method: Composition-based stats.  
Identities = 23/69 (33%), Positives = 29/69 (42%), Gaps = 2/69 (2%)

Query: 29 MAPISLVQIWVTGLTGLGGAGNSKSLCQKPTSCHPPCLYPAAARNVPRAVYPKASC--NLT 86  
+AP SL +TG+GGAG S PP PA + A Y A+ L  
Sbjct: 108 VAPNSLALCIGRAITGVGGAGISSGAFTIIALSAPPKQRPAYIGILGASYGVAAAIGPLV 167

Query: 87 DGSFGRPIT 95  
G+F IT  
Sbjct: 168 GGAFTTNIT 176

>TOXIN\_2012\_2|gi|358367040|dbj|GAA83660.1| toxin biosynthesis  
protein [Aspergillus kawachii IFO 4308]

Length = 447

Score = 24.6 bits (52), Expect = 7.5, Method: Composition-based stats.  
Identities = 11/30 (36%), Positives = 16/30 (53%)

Query: 71 RNVPRVYPKASCNLTGDSFGRPITATDDL 100  
R +P A++P +GS RP+T T L  
Sbjct: 241 RELPTAIHPLDESQQREGSDDRPVTLRTTL 270

>TOXIN\_2012\_2|gi|315081364|gb|EFT53340.1| toxin secretion/phage  
lysis holin [Propionibacterium acnes HL078PA1]  
Length = 569

Score = 24.3 bits (51), Expect = 9.7, Method: Composition-based stats.  
Identities = 12/41 (29%), Positives = 20/41 (48%)

Query: 33 SLVQIWVTGLTGLGGAGNSKSLCQKPTSCHPPCLYPAARNV 73  
SL+ + + GL +GG K+ HP + PAA+ +  
Sbjct: 196 SLLPLVIIGLHRVGGWSGLKAKITAAAEHPDVVTPAAQQL 236

>TOXIN\_2012\_2|gi|335050940|ref|ZP\_08543885.1| toxin secretion/phage  
lysis holin [Propionibacterium sp. 409-HC1]  
gi|342212266|ref|ZP\_08704991.1| transporter, SSS family  
[Propionibacterium sp. CC003-HC2]  
gi|313794187|gb|EFS42207.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL110PA1]  
gi|313839161|gb|EFS76875.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL086PA1]  
gi|314963843|gb|EFT07943.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL082PA1]  
gi|315078326|gb|EFT50365.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL053PA2]  
gi|327455936|gb|EGF02591.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL092PA1]  
gi|333768512|gb|EGL45694.1| toxin secretion/phage lysis  
holin [Propionibacterium sp. 409-HC1]  
gi|340767810|gb|EGR90335.1| transporter, SSS family  
[Propionibacterium sp. CC003-HC2]  
Length = 569

Score = 24.3 bits (51), Expect = 9.8, Method: Composition-based stats.  
Identities = 12/41 (29%), Positives = 20/41 (48%)

Query: 33 SLVQIWVTGLTGLGGAGNSKSLCQKPTSCHPPCLYPAARNV 73  
SL+ + + GL +GG K+ HP + PAA+ +  
Sbjct: 196 SLLPLVIIGLHRVGGWSGLKAKITAAAEHPDVVTPAAQQL 236

>TOXIN\_2012\_2|gi|313812423|gb|EFS50137.1| toxin secretion/phage  
lysis holin [Propionibacterium acnes HL025PA1]  
Length = 569

Score = 24.3 bits (51), Expect = 9.8, Method: Composition-based stats.  
Identities = 12/41 (29%), Positives = 20/41 (48%)

Query: 33 SLVQIWVTGLTGLGGAGNSKSLCQKPTSCHPPCLYPAARNV 73  
SL+ + + GL +GG K+ HP + PAA+ +  
Sbjct: 196 SLLPLVIIGLHRVGGWSGLKAKITAAAEHPDVVTPAAQQL 236

>TOXIN\_2012\_2|gi|327450317|gb|EGE96971.1| toxin secretion/phage

lysis holin [Propionibacterium acnes HL013PA2]  
Length = 569

Score = 24.3 bits (51), Expect = 9.9, Method: Composition-based stats.  
Identities = 12/41 (29%), Positives = 20/41 (48%)

Query: 33 SLVQIWVTGLTGLGGAGNSKSLCQKPTSCHPPCLYPAARNV 73  
SL+ + + GL +GG K+ HP + PAA+ +  
Sbjct: 196 SLLPLVIIGLHRVGGWSGLKAKITAAAEHPDVVTPAAQQL 236

>TOXIN\_2012\_2|gi|289426084|ref|ZP\_06427831.1| toxin secretion/phage  
lysis holin [Propionibacterium acnes SK187]  
gi|289427915|ref|ZP\_06429619.1| toxin secretion/phage  
lysis holin [Propionibacterium acnes J165]  
gi|295131714|ref|YP\_003582377.1| toxin secretion/phage  
lysis holin [Propionibacterium acnes SK137]  
gi|335055081|ref|ZP\_08547872.1| toxin secretion/phage  
lysis holin [Propionibacterium sp. 434-HC2]  
gi|365963814|ref|YP\_004945380.1| unnamed protein product  
[Propionibacterium acnes TypeIA2 P.acn31]  
gi|365966054|ref|YP\_004947619.1| unnamed protein product  
[Propionibacterium acnes TypeIA2 P.acn17]  
gi|365974993|ref|YP\_004956552.1| unnamed protein product  
[Propionibacterium acnes TypeIA2 P.acn33]  
gi|289153627|gb|EFD02341.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes SK187]  
gi|289158798|gb|EFD06998.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes J165]  
gi|291375283|gb|ADD99137.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes SK137]  
gi|313765210|gb|EFS36574.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL013PA1]  
gi|313771631|gb|EFS37597.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL074PA1]  
gi|313808642|gb|EFS47102.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL087PA2]  
gi|313810645|gb|EFS48359.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL083PA1]  
gi|313817068|gb|EFS54782.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL059PA1]  
gi|313819118|gb|EFS56832.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL046PA2]  
gi|313821169|gb|EFS58883.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL036PA1]  
gi|313823846|gb|EFS61560.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL036PA2]  
gi|313826948|gb|EFS64662.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL063PA1]  
gi|313829213|gb|EFS66927.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL063PA2]  
gi|313831458|gb|EFS69172.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL007PA1]  
gi|313834108|gb|EFS71822.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL056PA1]  
gi|314916787|gb|EFS80618.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL005PA4]  
gi|314919058|gb|EFS82889.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL050PA1]  
gi|314921100|gb|EFS84931.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL050PA3]  
gi|314926186|gb|EFS90017.1| toxin secretion/phage lysis  
holin [Propionibacterium acnes HL036PA3]

gi|314931365|gb|EFS95196.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL067PA1]  
 gi|314956221|gb|EFT00593.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL027PA1]  
 gi|314958716|gb|EFT02818.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL002PA1]  
 gi|314961023|gb|EFT05124.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL002PA2]  
 gi|314968848|gb|EFT12946.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL037PA1]  
 gi|314974848|gb|EFT18943.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL053PA1]  
 gi|314977825|gb|EFT21919.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL045PA1]  
 gi|314979020|gb|EFT23114.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL072PA2]  
 gi|314984529|gb|EFT28621.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL005PA1]  
 gi|314986006|gb|EFT30098.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL005PA2]  
 gi|314989309|gb|EFT33400.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL005PA3]  
 gi|315084996|gb|EFT56972.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL027PA2]  
 gi|315087526|gb|EFT59502.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL002PA3]  
 gi|315089567|gb|EFT61543.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL072PA1]  
 gi|315095673|gb|EFT67649.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL038PA1]  
 gi|315100002|gb|EFT71978.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL059PA2]  
 gi|315102838|gb|EFT74814.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL046PA1]  
 gi|315106633|gb|EFT78609.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL030PA1]  
 gi|315110565|gb|EFT82541.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL030PA2]  
 gi|327325857|gb|EGE67648.1| sodium:solute symporter  
 [Propionibacterium acnes HL096PA3]  
 gi|327332953|gb|EGE74685.1| sodium:solute symporter  
 [Propionibacterium acnes HL096PA2]  
 gi|327448659|gb|EGE95313.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL043PA1]  
 gi|327451116|gb|EGE97770.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL043PA2]  
 gi|327455475|gb|EGF02130.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL087PA3]  
 gi|327457616|gb|EGF04271.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL083PA2]  
 gi|328756865|gb|EGF70481.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL087PA1]  
 gi|328757750|gb|EGF71366.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL020PA1]  
 gi|328759480|gb|EGF73096.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes HL025PA2]  
 gi|328761925|gb|EGF75434.1| sodium:solute symporter  
 [Propionibacterium acnes HL099PA1]  
 gi|333762917|gb|EGL40399.1| toxin secretion/phage lysis  
 holin [Propionibacterium sp. 434-HC2]  
 gi|335278595|gb|AEH30500.1| toxin secretion/phage lysis  
 holin [Propionibacterium acnes 6609]  
 gi|340769716|gb|EGR92238.1| transporter, SSS family

```

[Propionibacterium acnes SK182]
gi|365740495|gb|AEW84697.1| toxin secretion/phage lysis
holin [Propionibacterium acnes TypeIA2 P.acn31]
gi|365742735|gb|AEW82429.1| toxin secretion/phage lysis
holin [Propionibacterium acnes TypeIA2 P.acn17]
gi|365744992|gb|AEW80189.1| toxin secretion/phage lysis
holin [Propionibacterium acnes TypeIA2 P.acn33]
Length = 569

```

Score = 24.3 bits (51), Expect = 9.9, Method: Composition-based stats.  
Identities = 12/41 (29%), Positives = 20/41 (48%)

```

Query: 33  SLVQIWVTGLTGLGGAGNSKSLCQKPTSCHPPCLYPAARNV 73
          SL+ + + GL +GG K+ HP + PAA+ +
Sbjct: 196 SLLPLVIIGLHRVGGWSGLKAKITAAAEHPDVVTPAAQQL 236

```

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.321	0.137	0.442

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 1,783,989  
Number of extensions: 65418  
Number of successful extensions: 122  
Number of sequences better than 10.0: 13  
Number of HSP's gapped: 122  
Number of HSP's successfully gapped: 13  
Length of query: 121  
Length of database: 7,243,252  
Length adjustment: 83  
Effective length of query: 38  
Effective length of database: 5,309,186  
Effective search space: 201749068  
Effective search space used: 201749068  
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: 52 (24.6 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_33  
(31 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.319	0.138	0.491

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 477,123  
Number of extensions: 6978  
Number of successful extensions: 5  
Number of sequences better than 10.0: 0  
Number of HSP's gapped: 5  
Number of HSP's successfully gapped: 0  
Length of query: 31  
Length of database: 7,243,252  
Length adjustment: 6  
Effective length of query: 25  
Effective length of database: 7,103,440  
Effective search space: 177586000  
Effective search space used: 177586000  
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 (21.7 bits)  
S2: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_34  
(31 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.325	0.141	0.534

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 539,802  
Number of extensions: 10564  
Number of successful extensions: 55  
Number of sequences better than 10.0: 0  
Number of HSP's gapped: 55  
Number of HSP's successfully gapped: 0  
Length of query: 31  
Length of database: 7,243,252  
Length adjustment: 6  
Effective length of query: 25  
Effective length of database: 7,103,440  
Effective search space: 177586000  
Effective search space used: 177586000  
Neighboring words threshold: 11  
Window for multiple hits: 40  
X1: 15 ( 7.0 bits)  
X2: 38 (14.6 bits)  
X3: 64 (24.7 bits)  
S1: 40 (21.6 bits)  
S2: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_35  
(115 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
TOXIN_2012_2 gi 212637285 ref YP_002313810.1  unnamed protein pr...	25	5.6
TOXIN_2012_2 gi 375262803 ref YP_005025033.1  unnamed protein pr...	25	5.9

>TOXIN\_2012\_2|gi|212637285|ref|YP\_002313810.1| unnamed protein  
product [Shewanella piezotolerans WP3]  
gi|212558769|gb|ACJ31223.1| RTX toxin, putative  
[Shewanella piezotolerans WP3]  
Length = 744

Score = 25.0 bits (53), Expect = 5.6, Method: Composition-based stats.  
Identities = 9/32 (28%), Positives = 17/32 (53%)

Query: 8 ATGFNLKKLYCQMFERSGKFGDLELDSYLGW 39  
A G++ ++ + + G+FG L +D G W  
Sbjct: 315 AQGYHSAQINWSINQSQQQFGLNIDPVTGQW 346

>TOXIN\_2012\_2|gi|375262803|ref|YP\_005025033.1| unnamed protein  
product [Vibrio sp. EJY3] gi|369844458|gb|AEX25286.1|  
RTX toxin [Vibrio sp. EJY3]  
Length = 3043

Score = 25.0 bits (53), Expect = 5.9, Method: Composition-based stats.  
Identities = 8/17 (47%), Positives = 12/17 (70%)

Query: 23 RSGKFGDLELDSYLGW 39  
+ G +G+L LDS+ G W  
Sbjct: 248 KGGHYGELTLDSHSGEW 264

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.323	0.134	0.477

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 1,525,158  
Number of extensions: 50998  
Number of successful extensions: 90  
Number of sequences better than 10.0: 2  
Number of HSP's gapped: 90  
Number of HSP's successfully gapped: 2  
Length of query: 115

Length of database: 7,243,252  
 Length adjustment: 81  
 Effective length of query: 34  
 Effective length of database: 5,355,790  
 Effective search space: 182096860  
 Effective search space used: 182096860  
 Neighboring words threshold: 11  
 Window for multiple hits: 40  
 X1: 16 ( 7.5 bits)  
 X2: 38 (14.6 bits)  
 X3: 64 (24.7 bits)  
 S1: 40 (21.6 bits)  
 S2: 51 (24.3 bits)  
 BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_36  
 (64 letters)

Database: TOXIN\_2012\_2  
 23,302 sequences; 7,243,252 total letters

Searching.....done

	Score	E
	(bits)	Value
Sequences producing significant alignments:		
TOXIN_2012_2 gi 23015941 ref ZP_00055704.1  COG2931: RTX toxins ...	25	4.2
>TOXIN_2012_2 gi 23015941 ref ZP_00055704.1  COG2931: RTX toxins and related Ca2+-binding proteins [Magnetospirillum magnetotacticum MS-1] Length = 12295		

Score = 25.4 bits (54), Expect = 4.2, Method: Composition-based stats.  
 Identities = 10/21 (47%), Positives = 12/21 (57%)

Query: 2 RLWDTQPGVHCAQLDTSMVDG 22  
 R WD G H A++DTS G  
 Sbjct: 6416 RAWDQTSNGTHGAKVDTSTNGG 6436

Score = 25.4 bits (54), Expect = 4.4, Method: Composition-based stats.  
 Identities = 10/21 (47%), Positives = 13/21 (61%)

Query: 2 RLWDTQPGVHCAQLDTSMVDG 22  
 R WD G H A++DTS+ G  
 Sbjct: 6901 RAWDQTSNGTHGAKVDTSVNGG 6921

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.322	0.135	0.452

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 740,630  
Number of extensions: 14869  
Number of successful extensions: 9  
Number of sequences better than 10.0: 1  
Number of HSP's gapped: 9  
Number of HSP's successfully gapped: 2  
Length of query: 64  
Length of database: 7,243,252  
Length adjustment: 36  
Effective length of query: 28  
Effective length of database: 6,404,380  
Effective search space: 179322640  
Effective search space used: 179322640  
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: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_37  
(35 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252

Number of sequences in database: 23,302

Lambda	K	H
0.320	0.130	0.398

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 502,057  
Number of extensions: 8645  
Number of successful extensions: 7  
Number of sequences better than 10.0: 0  
Number of HSP's gapped: 7  
Number of HSP's successfully gapped: 0  
Length of query: 35  
Length of database: 7,243,252  
Length adjustment: 9  
Effective length of query: 26  
Effective length of database: 7,033,534  
Effective search space: 182871884  
Effective search space used: 182871884  
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: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_38  
(183 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
TOXIN_2012_2 gi 333979976 ref YP_004517921.1  beta-lysine N-acet...	28	2.5
TOXIN_2012_2 gi 357038678 ref ZP_09100475.1  putative beta-lysin...	27	2.6
TOXIN_2012_2 gi 269796843 ref YP_003316298.1  Pertussis toxin su...	27	4.1
TOXIN_2012_2 gi 148273353 ref YP_001222914.1  putative RTX toxin...	27	5.5

>TOXIN\_2012\_2|gi|333979976|ref|YP\_004517921.1| beta-lysine  
N-acetyltransferase [Desulfotomaculum kuznetsovii DSM  
6115] gi|333823457|gb|AEG16120.1| putative beta-lysine  
N-acetyltransferase [Desulfotomaculum kuznetsovii DSM  
6115]  
Length = 280

Score = 27.7 bits (60), Expect = 2.5, Method: Compositional matrix adjust.  
Identities = 17/73 (23%), Positives = 31/73 (42%)

Query: 97 HQRGLGSTLYTHLLKSMEAQGFKSVVAVIGLPNDPSVRLHEALGYTARGTLRAAGYKHG 156  
++ + L S L H+ K + + ++ + + LGY RGT L + G  
Sbjct: 208 YRGMALNSLLLLHIEKECLKRNINCLYSLARASSYGMNLVLHRLGYVFRGT LINNCHIDG 267

Query: 157 GWHDVGFQWQDFE 169  
G+ ++ W R E  
Sbjct: 268 GFENMNIWVRPVE 280

>TOXIN\_2012\_2|gi|357038678|ref|ZP\_09100475.1| putative beta-lysine  
N-acetyltransferase [Desulfotomaculum gibsoniae DSM  
7213] gi|355359470|gb|EHG07232.1| putative beta-lysine  
N-acetyltransferase [Desulfotomaculum gibsoniae DSM  
7213]  
Length = 284

Score = 27.3 bits (59), Expect = 2.6, Method: Compositional matrix adjust.  
Identities = 16/66 (24%), Positives = 28/66 (42%)

Query: 101 GLGSTLYTHLLKSMEAQGFKSVVAVIGLPNDPSVRLHEALGYTARGTLRAAGYKHGGWHD 160  
G + + L K +G K + ++ + + LGY RGT L + GG+ D  
Sbjct: 216 GFMLGIISELEKECLGKGIKCLYSLARASSYGMNLVHRLGYRFRGT LINNCHISGGYED 275

Query: 161 VGFWQR 166  
+ W +  
Sbjct: 276 MNIWVK 281

>TOXIN\_2012\_2|gi|269796843|ref|YP\_003316298.1| Pertussis toxin  
subunit 1 [Sanguibacter keddieii DSM 10542]  
gi|269099028|gb|ACZ23464.1| Pertussis toxin, subunit 1  
[Sanguibacter keddieii DSM 10542]  
Length = 810

Score = 26.9 bits (58), Expect = 4.1, Method: Compositional matrix adjust.  
Identities = 21/69 (30%), Positives = 28/69 (40%), Gaps = 4/69 (5%)

Query: 99 RLGLGSTLYTHL--LKSMEAQG--FKSVVAVIGLPNDPSVRLHEALGYTARGTLRAAGYK 154  
RL L ++ L L+++ A V+ +G D L LG A G  
Sbjct: 300 RLALAESIPADLARLEAVTASDTLISKVMETVGRSVDVGSELIPYLGVAATGYALREDVD 359

Query: 155 HGGWHDVGF 163  
HG W DVGF  
Sbjct: 360 HGNWADVGF 368

>TOXIN\_2012\_2|gi|148273353|ref|YP\_001222914.1| putative RTX toxin  
[Clavibacter michiganensis subsp. michiganensis NCPPB  
382] gi|147831283|emb|CAN02239.1| putative RTX toxin  
[Clavibacter michiganensis subsp. michiganensis NCPPB  
382]  
Length = 2004

Score = 26.6 bits (57), Expect = 5.5, Method: Compositional matrix adjust.  
Identities = 14/37 (37%), Positives = 19/37 (51%), Gaps = 1/37 (2%)

Query: 130 NDPSVRLHEALGYTARGTLRAAGYKHGGWHDVGFQWQR 166  
DP V + G AR +R +G ++ GW D WQR  
Sbjct: 304 GDPVVVTGDVRGDPAR-PVRVSGCEYAGWADGSGWQR 339

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.320	0.135	0.439

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 2,522,842  
Number of extensions: 93013  
Number of successful extensions: 176  
Number of sequences better than 10.0: 8  
Number of HSP's gapped: 176  
Number of HSP's successfully gapped: 8  
Length of query: 183  
Length of database: 7,243,252  
Length adjustment: 89  
Effective length of query: 94  
Effective length of database: 5,169,374  
Effective search space: 485921156  
Effective search space used: 485921156  
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: 55 (25.8 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_39  
(31 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters



Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2

Posted date: Feb 22, 2012 11:49 AM

Number of letters in database: 7,243,252

Number of sequences in database: 23,302

Lambda	K	H
0.323	0.135	0.393

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62

Gap Penalties: Existence: 11, Extension: 1

Number of Sequences: 23302

Number of Hits to DB: 396,061

Number of extensions: 5435

Number of successful extensions: 31

Number of sequences better than 10.0: 0

Number of HSP's gapped: 31

Number of HSP's successfully gapped: 0

Length of query: 31

Length of database: 7,243,252

Length adjustment: 6

Effective length of query: 25

Effective length of database: 7,103,440

Effective search space: 177586000

Effective search space used: 177586000

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 (22.0 bits)

S2: 51 (24.3 bits)

BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_40  
(42 letters)

Database: TOXIN\_2012\_2

23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2

Posted date: Feb 22, 2012 11:49 AM

Number of letters in database: 7,243,252

Number of sequences in database: 23,302

Lambda	K	H
0.328	0.139	0.412

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62

Gap Penalties: Existence: 11, Extension: 1

Number of Sequences: 23302

Number of Hits to DB: 360,840

Number of extensions: 5237

Number of successful extensions: 2

Number of sequences better than 10.0: 0

Number of HSP's gapped: 2

Number of HSP's successfully gapped: 0

Length of query: 42

Length of database: 7,243,252

Length adjustment: 16

Effective length of query: 26

Effective length of database: 6,870,420

Effective search space: 178630920

Effective search space used: 178630920

Neighboring words threshold: 11

Window for multiple hits: 40

X1: 15 ( 7.1 bits)

X2: 38 (14.6 bits)

X3: 64 (24.7 bits)

S1: 40 (21.8 bits)

S2: 51 (24.3 bits)

BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_41  
(35 letters)

Database: TOXIN\_2012\_2

23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.324	0.136	0.429

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 460,698  
Number of extensions: 7244  
Number of successful extensions: 7  
Number of sequences better than 10.0: 0  
Number of HSP's gapped: 7  
Number of HSP's successfully gapped: 0  
Length of query: 35  
Length of database: 7,243,252  
Length adjustment: 9  
Effective length of query: 26  
Effective length of database: 7,033,534  
Effective search space: 182871884  
Effective search space used: 182871884  
Neighboring words threshold: 11  
Window for multiple hits: 40  
X1: 15 ( 7.0 bits)  
X2: 38 (14.6 bits)  
X3: 64 (24.7 bits)  
S1: 40 (21.6 bits)  
S2: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_42  
(79 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
TOXIN_2012_2 gi 302566051 pdb 3I5V A Chain A, Crystal Structure ...	26	2.6

TOXIN_2012_2 gi 300508421 pdb 3I41 A Chain A, Crystal Structure ...	26	2.7
TOXIN_2012_2 gi 320089760 pdb 3K55 A Chain A, Structure Of Beta ...	26	2.8
TOXIN_2012_2 gi 345014741 ref YP_004817095.1  Ricin B lectin [St...	26	3.2
TOXIN_2012_2 gi 57650692 ref YP_186826.1  phospholipase C [Staph...	26	3.4

>TOXIN\_2012\_2|gi|302566051|pdb|3I5V|A Chain A, Crystal Structure  
 Of Beta Toxin 275-280 From Staphylococcus Aureus  
 gi|302566052|pdb|3I5V|B Chain B, Crystal Structure Of  
 Beta Toxin 275-280 From Staphylococcus Aureus  
 gi|302566053|pdb|3I5V|C Chain C, Crystal Structure Of  
 Beta Toxin 275-280 From Staphylococcus Aureus  
 gi|302566054|pdb|3I5V|D Chain D, Crystal Structure Of  
 Beta Toxin 275-280 From Staphylococcus Aureus  
 Length = 313

Score = 26.2 bits (56), Expect = 2.6, Method: Composition-based stats.  
 Identities = 11/30 (36%), Positives = 19/30 (63%), Gaps = 3/30 (10%)

Query: 15 KQNDTDLNIGFYTNHMIFINSVILYGRWRQ 44  
 K++DTD L + +H +++ S +LY W Q  
 Sbjct: 24 KKDDTDLKL--VSHNVYMLSTVLYPNWGQ 50

>TOXIN\_2012\_2|gi|300508421|pdb|3I41|A Chain A, Crystal Structure  
 Of Beta Toxin From Staphylococcus Aureus F277a, P278a  
 Mutant gi|300508422|pdb|3I41|B Chain B, Crystal  
 Structure Of Beta Toxin From Staphylococcus Aureus  
 F277a, P278a Mutant gi|300508423|pdb|3I46|A Chain A,  
 Crystal Structure Of Beta Toxin From Staphylococcus  
 Aureus F277a, P278a Mutant With Bound Calcium Ions  
 gi|300508424|pdb|3I46|B Chain B, Crystal Structure Of  
 Beta Toxin From Staphylococcus Aureus F277a, P278a  
 Mutant With Bound Calcium Ions gi|300508425|pdb|3I48|A  
 Chain A, Crystal Structure Of Beta Toxin From  
 Staphylococcus Aureus F277a, P278a Mutant With Bound  
 Magnesium Ions gi|300508426|pdb|3I48|B Chain B, Crystal  
 Structure Of Beta Toxin From Staphylococcus Aureus  
 F277a, P278a Mutant With Bound Magnesium Ions  
 Length = 317

Score = 26.2 bits (56), Expect = 2.7, Method: Composition-based stats.  
 Identities = 11/30 (36%), Positives = 19/30 (63%), Gaps = 3/30 (10%)

Query: 15 KQNDTDLNIGFYTNHMIFINSVILYGRWRQ 44  
 K++DTD L + +H +++ S +LY W Q  
 Sbjct: 24 KKDDTDLKL--VSHNVYMLSTVLYPNWGQ 50

>TOXIN\_2012\_2|gi|320089760|pdb|3K55|A Chain A, Structure Of Beta  
 Hairpin Deletion Mutant Of Beta Toxin From  
 Staphylococcus Aureus gi|320089761|pdb|3K55|B Chain B,  
 Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin  
 From Staphylococcus Aureus gi|320089762|pdb|3K55|C  
 Chain C, Structure Of Beta Hairpin Deletion Mutant Of  
 Beta Toxin From Staphylococcus Aureus  
 gi|320089763|pdb|3K55|D Chain D, Structure Of Beta  
 Hairpin Deletion Mutant Of Beta Toxin From  
 Staphylococcus Aureus gi|320089764|pdb|3K55|E Chain E,  
 Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin  
 From Staphylococcus Aureus gi|320089765|pdb|3K55|F  
 Chain F, Structure Of Beta Hairpin Deletion Mutant Of  
 Beta Toxin From Staphylococcus Aureus  
 gi|320089766|pdb|3K55|G Chain G, Structure Of Beta

Hairpin Deletion Mutant Of Beta Toxin From  
 Staphylococcus Aureus gi|320089767|pdb|3K55|H Chain H,  
 Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin  
 From Staphylococcus Aureus gi|320089768|pdb|3K55|I  
 Chain I, Structure Of Beta Hairpin Deletion Mutant Of  
 Beta Toxin From Staphylococcus Aureus  
 gi|320089769|pdb|3K55|J Chain J, Structure Of Beta  
 Hairpin Deletion Mutant Of Beta Toxin From  
 Staphylococcus Aureus gi|320089770|pdb|3K55|K Chain K,  
 Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin  
 From Staphylococcus Aureus gi|320089771|pdb|3K55|L  
 Chain L, Structure Of Beta Hairpin Deletion Mutant Of  
 Beta Toxin From Staphylococcus Aureus  
 gi|320089772|pdb|3K55|M Chain M, Structure Of Beta  
 Hairpin Deletion Mutant Of Beta Toxin From  
 Staphylococcus Aureus gi|320089773|pdb|3K55|N Chain N,  
 Structure Of Beta Hairpin Deletion Mutant Of Beta Toxin  
 From Staphylococcus Aureus gi|320089774|pdb|3K55|O  
 Chain O, Structure Of Beta Hairpin Deletion Mutant Of  
 Beta Toxin From Staphylococcus Aureus  
 gi|320089775|pdb|3K55|P Chain P, Structure Of Beta  
 Hairpin Deletion Mutant Of Beta Toxin From  
 Staphylococcus Aureus  
 Length = 306

Score = 25.8 bits (55), Expect = 2.8, Method: Composition-based stats.  
 Identities = 11/30 (36%), Positives = 19/30 (63%), Gaps = 3/30 (10%)

Query: 15 KQNDTDLNIGFYTNHMIFINSVILYGRWRQ 44  
 K++DTDL + +H +++ S +LY W Q  
 Sbjct: 24 KKDDTDLKL--VSHNVYMLSTVLYPNWGQ 50

>TOXIN\_2012\_2|gi|345014741|ref|YP\_004817095.1| Ricin B lectin  
 [Streptomyces violaceusniger Tu 4113]  
 gi|344041090|gb|AEM86815.1| Ricin B lectin [Streptomyces  
 violaceusniger Tu 4113]  
 Length = 714

Score = 25.8 bits (55), Expect = 3.2, Method: Composition-based stats.  
 Identities = 12/26 (46%), Positives = 15/26 (57%), Gaps = 2/26 (7%)

Query: 4 QTTPTKYQVLFKQN--DTDLNIGFYT 27  
 Q +Y VLFK D+N+GFYT  
 Sbjct: 205 QFGSNRYAVLFKPGTYSADVNVGFYT 230

>TOXIN\_2012\_2|gi|57650692|ref|YP\_186826.1| phospholipase C  
 [Staphylococcus aureus subsp. aureus COL]  
 gi|282894608|ref|ZP\_06302836.1| sphingomyelin  
 phosphodiesterase [Staphylococcus aureus A8117]  
 gi|130085|sp|P09978.1|PHLC\_STAAU RecName:  
 Full=Phospholipase C; AltName: Full=Beta-hemolysin;  
 AltName: Full=Beta-toxin; AltName:  
 Full=Sphingomyelinase; Short=SMase; Flags: Precursor  
 gi|61213889|sp|Q5HEI1.1|PHLC\_STAAC RecName:  
 Full=Phospholipase C; AltName: Full=Beta-hemolysin;  
 AltName: Full=Beta-toxin; AltName:  
 Full=Sphingomyelinase; Short=SMase; Flags: Precursor  
 gi|110815915|sp|Q2FWP1.2|PHLC\_STAA8 RecName:  
 Full=Phospholipase C; AltName: Full=Beta-hemolysin;  
 AltName: Full=Beta-toxin; AltName:  
 Full=Sphingomyelinase; Short=SMase; Flags: Precursor

gi|46588|emb|CAA31769.1| unnamed protein product  
 [Staphylococcus aureus subsp. aureus COL]  
 gi|57284878|gb|AAW36972.1| phospholipase C  
 [Staphylococcus aureus subsp. aureus COL]  
 gi|269941409|emb|CBI49806.1| phospholipase C precursor  
 (pseudogene) [Staphylococcus aureus subsp. aureus TW20]  
 gi|282763095|gb|EFC03227.1| sphingomyelin  
 phosphodiesterase [Staphylococcus aureus A8117]  
 gi|298695279|gb|ADI98501.1| beta-hemolysin  
 [Staphylococcus aureus subsp. aureus ED133]  
 gi|329731457|gb|EGG67820.1| sphingomyelin  
 phosphodiesterase [Staphylococcus aureus subsp. aureus  
 21193] gi|365168342|gb|EHM59689.1| sphingomyelin  
 phosphodiesterase [Staphylococcus aureus subsp. aureus  
 21178] gi|375025481|gb|EHS18885.1| sphingomyelin  
 phosphodiesterase [Staphylococcus aureus subsp. aureus  
 IS-91]  
 Length = 330

Score = 25.8 bits (55), Expect = 3.4, Method: Composition-based stats.  
 Identities = 11/30 (36%), Positives = 19/30 (63%), Gaps = 3/30 (10%)

Query: 15 KQNDTDLNIGFYTNHMIFINSVILYGRWRQ 44  
 K++DTD L + +H +++ S +LY W Q  
 Sbjct: 37 KKDDTDLKL--VSHNVYMLSTVLYPNWGQ 63

Database: TOXIN\_2012\_2  
 Posted date: Feb 22, 2012 11:49 AM  
 Number of letters in database: 7,243,252  
 Number of sequences in database: 23,302

Lambda	K	H
0.327	0.137	0.425

Gapped		
Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
 Gap Penalties: Existence: 11, Extension: 1  
 Number of Sequences: 23302  
 Number of Hits to DB: 885,535  
 Number of extensions: 26026  
 Number of successful extensions: 67  
 Number of sequences better than 10.0: 11  
 Number of HSP's gapped: 67  
 Number of HSP's successfully gapped: 11  
 Length of query: 79  
 Length of database: 7,243,252  
 Length adjustment: 49  
 Effective length of query: 30  
 Effective length of database: 6,101,454  
 Effective search space: 183043620  
 Effective search space used: 183043620  
 Neighboring words threshold: 11  
 Window for multiple hits: 40  
 X1: 15 ( 7.1 bits)  
 X2: 38 (14.6 bits)  
 X3: 64 (24.7 bits)  
 S1: 40 (21.7 bits)  
 S2: 51 (24.3 bits)

BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_43  
(35 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.315	0.130	0.353

Gapped		
Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 384,936  
Number of extensions: 6200  
Number of successful extensions: 6  
Number of sequences better than 10.0: 0  
Number of HSP's gapped: 6  
Number of HSP's successfully gapped: 0  
Length of query: 35  
Length of database: 7,243,252  
Length adjustment: 9  
Effective length of query: 26  
Effective length of database: 7,033,534  
Effective search space: 182871884  
Effective search space used: 182871884  
Neighboring words threshold: 11  
Window for multiple hits: 40  
X1: 16 ( 7.3 bits)  
X2: 38 (14.6 bits)  
X3: 64 (24.7 bits)  
S1: 41 (21.6 bits)  
S2: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_44  
(69 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
TOXIN_2012_2 gi 23015941 ref ZP_00055704.1  COG2931: RTX toxins ...	25	5.4
>TOXIN_2012_2 gi 23015941 ref ZP_00055704.1  COG2931: RTX toxins and related Ca2+-binding proteins [Magnetospirillum magnetotacticum MS-1] Length = 12295		

Score = 25.0 bits (53), Expect = 5.4, Method: Composition-based stats.  
Identities = 10/21 (47%), Positives = 12/21 (57%)

Query: 2 RLWDTQPGVHCAQLDTSMDVG 22  
R WD G H A++DTS G  
Sbjct: 6416 RAWDQTSGTHGAKVDTSTNGG 6436

Score = 25.0 bits (53), Expect = 5.7, Method: Composition-based stats.  
Identities = 10/21 (47%), Positives = 13/21 (61%)

Query: 2 RLWDTQPGVHCAQLDTSMDVG 22  
R WD G H A++DTS+ G  
Sbjct: 6901 RAWDQTSGSHGAKVDTSVNGG 6921

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.324	0.137	0.443

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302



Number of Hits to DB: 767,387  
 Number of extensions: 15429  
 Number of successful extensions: 11  
 Number of sequences better than 10.0: 1  
 Number of HSP's gapped: 11  
 Number of HSP's successfully gapped: 2  
 Length of query: 69  
 Length of database: 7,243,252  
 Length adjustment: 40  
 Effective length of query: 29  
 Effective length of database: 6,311,172  
 Effective search space: 183023988  
 Effective search space used: 183023988  
 Neighboring words threshold: 11  
 Window for multiple hits: 40  
 X1: 15 ( 7.0 bits)  
 X2: 38 (14.6 bits)  
 X3: 64 (24.7 bits)  
 S1: 40 (21.6 bits)  
 S2: 51 (24.3 bits)  
 BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_45  
 (65 letters)

Database: TOXIN\_2012\_2  
 23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
 Posted date: Feb 22, 2012 11:49 AM  
 Number of letters in database: 7,243,252  
 Number of sequences in database: 23,302

Lambda	K	H
0.330	0.142	0.487

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
 Gap Penalties: Existence: 11, Extension: 1  
 Number of Sequences: 23302  
 Number of Hits to DB: 614,280  
 Number of extensions: 11107  
 Number of successful extensions: 36

Number of sequences better than 10.0: 0  
 Number of HSP's gapped: 36  
 Number of HSP's successfully gapped: 0  
 Length of query: 65  
 Length of database: 7,243,252  
 Length adjustment: 37  
 Effective length of query: 28  
 Effective length of database: 6,381,078  
 Effective search space: 178670184  
 Effective search space used: 178670184  
 Neighboring words threshold: 11  
 Window for multiple hits: 40  
 X1: 15 ( 7.1 bits)  
 X2: 38 (14.6 bits)  
 X3: 64 (24.7 bits)  
 S1: 40 (21.8 bits)  
 S2: 51 (24.3 bits)  
 BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_46  
 (40 letters)

Database: TOXIN\_2012\_2  
 23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
 Posted date: Feb 22, 2012 11:49 AM  
 Number of letters in database: 7,243,252  
 Number of sequences in database: 23,302

Lambda	K	H
0.319	0.135	0.454

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
 Gap Penalties: Existence: 11, Extension: 1  
 Number of Sequences: 23302  
 Number of Hits to DB: 625,904  
 Number of extensions: 14096  
 Number of successful extensions: 4  
 Number of sequences better than 10.0: 0  
 Number of HSP's gapped: 4  
 Number of HSP's successfully gapped: 0

Length of query: 40  
 Length of database: 7,243,252  
 Length adjustment: 14  
 Effective length of query: 26  
 Effective length of database: 6,917,024  
 Effective search space: 179842624  
 Effective search space used: 179842624  
 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 (21.7 bits)  
 S2: 51 (24.3 bits)  
 BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_47  
 (67 letters)

Database: TOXIN\_2012\_2  
 23,302 sequences; 7,243,252 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
TOXIN_2012_2 gi 170022729 ref YP_001719234.1  virulence plasmid ...	25	4.2

>TOXIN\_2012\_2|gi|170022729|ref|YP\_001719234.1| virulence plasmid 65kDa  
 B protein [Yersinia pseudotuberculosis YPIII]  
 gi|149211827|gb|ABR22251.1| toxin complex subunit TcaC  
 [Yersinia pseudotuberculosis] gi|169749263|gb|ACA66781.1|  
 virulence plasmid 65kDa B protein [Yersinia  
 pseudotuberculosis YPIII]  
 Length = 1489

Score = 25.4 bits (54), Expect = 4.2, Method: Composition-based stats.  
 Identities = 13/34 (38%), Positives = 21/34 (61%), Gaps = 1/34 (2%)

Query: 29 YHINPLALKTWLERLLFPRCSSLVGVHLWDHCRQ 62  
 Y++ PLAL+ W+ R P + +VG+ L + RQ  
 Sbjct: 1298 YYLCPALRRWI-RQSNPLITEYVGLTLKNPVRQ 1330

Database: TOXIN\_2012\_2  
 Posted date: Feb 22, 2012 11:49 AM  
 Number of letters in database: 7,243,252  
 Number of sequences in database: 23,302

Lambda	K	H
0.332	0.142	0.519

Gapped Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 23302  
Number of Hits to DB: 1,002,363  
Number of extensions: 30839  
Number of successful extensions: 37  
Number of sequences better than 10.0: 1  
Number of HSP's gapped: 37  
Number of HSP's successfully gapped: 1  
Length of query: 67  
Length of database: 7,243,252  
Length adjustment: 39  
Effective length of query: 28  
Effective length of database: 6,334,474  
Effective search space: 177365272  
Effective search space used: 177365272  
Neighboring words threshold: 11  
Window for multiple hits: 40  
X1: 15 ( 7.2 bits)  
X2: 38 (14.6 bits)  
X3: 64 (24.7 bits)  
S1: 40 (21.9 bits)  
S2: 51 (24.3 bits)  
BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_48  
(34 letters)

Database: TOXIN\_2012\_2  
23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
Posted date: Feb 22, 2012 11:49 AM  
Number of letters in database: 7,243,252  
Number of sequences in database: 23,302

Lambda	K	H
0.317	0.125	0.427

Gapped			
Lambda	K	H	
0.267	0.0410	0.140	

Matrix: BLOSUM62  
 Gap Penalties: Existence: 11, Extension: 1  
 Number of Sequences: 23302  
 Number of Hits to DB: 297,319  
 Number of extensions: 3711  
 Number of successful extensions: 1  
 Number of sequences better than 10.0: 0  
 Number of HSP's gapped: 1  
 Number of HSP's successfully gapped: 0  
 Length of query: 34  
 Length of database: 7,243,252  
 Length adjustment: 8  
 Effective length of query: 26  
 Effective length of database: 7,056,836  
 Effective search space: 183477736  
 Effective search space used: 183477736  
 Neighboring words threshold: 11  
 Window for multiple hits: 40  
 X1: 16 ( 7.3 bits)  
 X2: 38 (14.6 bits)  
 X3: 64 (24.7 bits)  
 S1: 41 (21.7 bits)  
 S2: 51 (24.3 bits)  
 BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_49  
 (32 letters)

Database: TOXIN\_2012\_2  
 23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
 Posted date: Feb 22, 2012 11:49 AM  
 Number of letters in database: 7,243,252  
 Number of sequences in database: 23,302

Lambda	K	H
0.333	0.133	0.387

Gapped			
Lambda	K	H	
0.267	0.0410	0.140	

Matrix: BLOSUM62  
 Gap Penalties: Existence: 11, Extension: 1  
 Number of Sequences: 23302  
 Number of Hits to DB: 246,219  
 Number of extensions: 2273  
 Number of successful extensions: 20  
 Number of sequences better than 10.0: 0  
 Number of HSP's gapped: 20  
 Number of HSP's successfully gapped: 0  
 Length of query: 32  
 Length of database: 7,243,252  
 Length adjustment: 7  
 Effective length of query: 25  
 Effective length of database: 7,080,138  
 Effective search space: 177003450  
 Effective search space used: 177003450  
 Neighboring words threshold: 11  
 Window for multiple hits: 40  
 X1: 15 ( 7.2 bits)  
 X2: 38 (14.6 bits)  
 X3: 64 (24.7 bits)  
 S1: 39 (21.7 bits)  
 S2: 51 (24.3 bits)  
 BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_50  
 (41 letters)

Database: TOXIN\_2012\_2  
 23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
 Posted date: Feb 22, 2012 11:49 AM  
 Number of letters in database: 7,243,252  
 Number of sequences in database: 23,302

Lambda	K	H
0.325	0.135	0.438

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62

Gap Penalties: Existence: 11, Extension: 1  
 Number of Sequences: 23302  
 Number of Hits to DB: 583,756  
 Number of extensions: 12610  
 Number of successful extensions: 32  
 Number of sequences better than 10.0: 0  
 Number of HSP's gapped: 32  
 Number of HSP's successfully gapped: 0  
 Length of query: 41  
 Length of database: 7,243,252  
 Length adjustment: 15  
 Effective length of query: 26  
 Effective length of database: 6,893,722  
 Effective search space: 179236772  
 Effective search space used: 179236772  
 Neighboring words threshold: 11  
 Window for multiple hits: 40  
 X1: 15 ( 7.0 bits)  
 X2: 38 (14.6 bits)  
 X3: 64 (24.7 bits)  
 S1: 40 (21.7 bits)  
 S2: 51 (24.3 bits)  
 BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_51  
 (40 letters)

Database: TOXIN\_2012\_2  
 23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
 Posted date: Feb 22, 2012 11:49 AM  
 Number of letters in database: 7,243,252  
 Number of sequences in database: 23,302

Lambda	K	H
0.325	0.138	0.504

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
 Gap Penalties: Existence: 11, Extension: 1  
 Number of Sequences: 23302  
 Number of Hits to DB: 618,726

Number of extensions: 14460  
 Number of successful extensions: 55  
 Number of sequences better than 10.0: 1  
 Number of HSP's gapped: 55  
 Number of HSP's successfully gapped: 1  
 Length of query: 40  
 Length of database: 7,243,252  
 Length adjustment: 14  
 Effective length of query: 26  
 Effective length of database: 6,917,024  
 Effective search space: 179842624  
 Effective search space used: 179842624  
 Neighboring words threshold: 11  
 Window for multiple hits: 40  
 X1: 15 ( 7.0 bits)  
 X2: 38 (14.6 bits)  
 X3: 64 (24.7 bits)  
 S1: 40 (21.6 bits)  
 S2: 51 (24.3 bits)  
 BLASTP 2.2.19 [Nov-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= SYHT0H2\_I\_52  
 (57 letters)

Database: TOXIN\_2012\_2  
 23,302 sequences; 7,243,252 total letters

Searching.....done

\*\*\*\*\* No hits found \*\*\*\*\*

Database: TOXIN\_2012\_2  
 Posted date: Feb 22, 2012 11:49 AM  
 Number of letters in database: 7,243,252  
 Number of sequences in database: 23,302

Lambda	K	H
0.322	0.125	0.446

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
 Gap Penalties: Existence: 11, Extension: 1  
 Number of Sequences: 23302  
 Number of Hits to DB: 572,235  
 Number of extensions: 9841  
 Number of successful extensions: 11  
 Number of sequences better than 10.0: 0



Number of HSP's gapped: 11  
Number of HSP's successfully gapped: 0  
Length of query: 57  
Length of database: 7,243,252  
Length adjustment: 30  
Effective length of query: 27  
Effective length of database: 6,544,192  
Effective search space: 176693184  
Effective search space used: 176693184  
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: 40 (21.6 bits)  
S2: 51 (24.3 bits)