

## **Appendix 4**

### **Bioinformatics Analysis of Deduced Amino Acid Sequences of Open Reading Frames Created by the Junctions of the Insert with Genomic DNA, in Herbicide-tolerant BPS-CV127-9 Soybean**

**BIOINFORMATICS ANALYSIS OF DEDUCED AMINO ACID SEQUENCES OF  
OPEN READING FRAMES CREATED BY THE JUNCTIONS OF THE INSERT  
WITH GENOMIC DNA, IN HERBICIDE-TOLERANT BPS-CV127-9 SOYBEAN**

**AUTHOR:**



**STUDY COMPLETED ON: MAY 15, 2009**

**PERFORMING LABORATORY:**

**SUNGENE GMBH  
CORRENSSTRASSE 3  
D-06466 GATERSLEBEN  
GERMANY**

**REPORT NUMBER: BPS-006-09**

**OECD UNIQUE IDENTIFIER: BPS-CV127-9**

**SUBMITTED BY:**

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



The Chemical Company  
Plant Science L. L. C.

BASF Reg. Doc. No. 2009/7003069

## STATEMENT OF NO DATA CONFIDENTIALITY CLAIMS

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

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

Company Agent: [REDACTED] Date: 15 May 2009

Title: Regulatory Affairs Manager

Signature: [REDACTED]


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

## STATEMENT OF COMPLIANCE

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

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

### PRINCIPAL INVESTIGATOR:

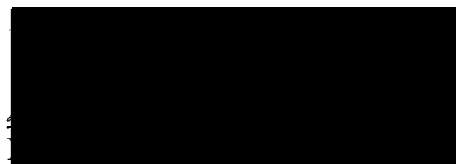


Scientist  
SunGene GmbH

15 MAY 2009

Date

### SPONSOR REPRESENTATIVE:

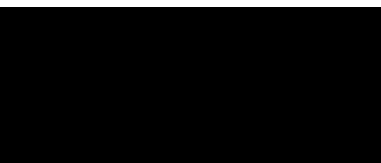


Senior Manager, Regulatory Science  
BASF Plant Science, L. L. C.

15 May 09

Date

### SUBMITTED BY:



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

15 May 2009

Date

## TABLE OF CONTENTS

STATEMENT OF NO DATA CONFIDENTIALITY CLAIMS.....	2
STATEMENT OF COMPLIANCE.....	3
TABLE OF CONTENTS.....	4
LIST OF TABLES .....	4
LIST OF FIGURES .....	4
ABBREVIATIONS AND DEFINITIONS.....	5
SUMMARY .....	6
INTRODUCTION .....	6
MATERIALS AND METHODS.....	7
RESULTS AND DISCUSSION .....	10
CONCLUSIONS.....	10
RECORDS RETENTION.....	12
STUDY PERSONNEL .....	12
REFERENCES .....	12
APPENDIX 1. 80-AMINO ACID ALLERGEN HOMOLOGY SEARCH RESULTS..	19
APPENDIX 2. TOXIN HOMOLOGY AND GENERAL PROTEIN HOMOLOGY SEARCH RESULTS .....	161

## LIST OF TABLES

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

## LIST OF FIGURES

Figure 1. Graphic Representation of the Insertion in Soybean BPS-CV127-9.....	15
Figure 2. Deduced amino Acid Sequences of Open Reading Frames from the Transgene Insert of BPS-CV127-9.....	15

## ABBREVIATIONS AND DEFINITIONS

AHAS	acetohydroxyacid synthase enzyme
<i>ahasl</i>	imidazolinone-tolerant alleles of <i>Arabidopsis thaliana</i> acetohydroxyacid synthase large subunit gene; includes the <i>ahasl</i> (S653N) allele found in plasmid pAC321 and the <i>ahasl</i> (R272K, S653N) allele found in BPS-CV127-9
AHASL	wild-type <i>Arabidopsis thaliana</i> acetohydroxyacid synthase large subunit gene; also referred to in the literature as <i>CSR1</i>
ahasl	imidazolinone-tolerant <i>Arabidopsis thaliana</i> acetohydroxyacid synthase large subunit protein bearing two mutations (R272K, S653N)
BLASTP	protein-protein Basic Local Alignment Search Tool
BLOSUM	<u>b</u> locks <u>s</u> ubstitution <u>m</u> atrix
<i>csr1-2</i>	Gene from <i>Arabidopsis thaliana</i> encoding an imidazolinone-tolerant AHAS enzyme <sup>1</sup>
FARRP	Food Allergy Research and Resource Program
E value	expect value
GI number	GenInfo identifier, or GI number, is a sequence identification number assigned by the NCBI
IgE	immunoglobulin E
NCBI	National Center for Biotechnology Information (United States of America)
ORF	open reading frame
R272K	arginine residue at position 272 of <i>Arabidopsis thaliana</i> acetohydroxyacid synthase large subunit replaced with lysine
RT-PCR	reverse-transcription polymerase chain reaction
S653N	serine residue at position 653 of <i>Arabidopsis thaliana</i> acetohydroxyacid synthase large subunit replaced with asparagine
<i>SEC61γ</i>	wild-type <i>Arabidopsis thaliana</i> <i>SEC61γ</i> subunit gene
SEC61γ	wild-type <i>Arabidopsis thaliana</i> SEC61γ subunit protein

<sup>1</sup> It should be noted that throughout this document, the gene that resides within the genome of BPS-CV127-9 soybean that encodes the imidazolinone-tolerance trait is referred to as the *csr1-2* gene which is derived from *A. thaliana* and it is recognized that the *csr1-2* gene in CV127 soybean differs from the *csr1-2* gene in *A. thaliana* by a single nucleotide change that results in the R272K replacement.

## **BIOINFORMATICS ANALYSIS OF DEDUCED AMINO ACID SEQUENCES OF OPEN READING FRAMES CREATED BY THE JUNCTIONS OF THE INSERT WITH GENOMIC DNA IN HERBICIDE-TOLERANT BPS-CV127-9 SOYBEAN**

### **SUMMARY**

Soybean plants have been developed that are tolerant to imidazolinone herbicides. These plants, referred to as BPS-CV127-9 soybean, were produced by the introduction of an imidazolinone-tolerant acetohydroxyacid synthase large subunit (*ahasl*) allele from *Arabidopsis* (referred to as the *csr1-2* gene) into the soybean plant genome via biolistics transformation.

The junctions between the BPS-CV127-9 transgene insert and the soybean genome were analyzed to determine if new open reading frames (ORFs) were created due to the insertion of the transgene into the soybean genome. At the 5' junction 24 ORFs were identified and at the 3' junction six putative ORFs were identified. In this study, the deduced amino acid sequences of these 30 ORFs were subjected to bioinformatics analyses as part of a weight-of-evidence assessment of the potential allergenicity and toxicity of any proteins potentially encoded by the junctions between the BPS-CV127-9 soybean insert and the soybean genome. The bioinformatics analyses reported herein did not indicate any potential allergenicity or toxicity concerns related to the transgene insert of BPS-CV127-9 soybean.

### **INTRODUCTION**

Soybean [*Glycine max* (L.) Merr.] plants have been developed that are tolerant to the imidazolinone class of agricultural herbicides. These herbicide-tolerant plants, referred to as BPS-CV127-9 soybean, are derived from a single transformation event. BPS-CV127-9 soybean was produced by the introduction of an imidazolinone-tolerant acetohydroxyacid synthase large subunit (*ahasl*) gene (*csr1-2*) from *Arabidopsis thaliana* (L.) Heynh. into the soybean plant genome via biolistics using the PvuII fragment of transformation vector pAC321. The PvuII fragment includes what, at the time of transformation, was defined as the wild-type *Arabidopsis AHASL* promoter, the herbicide-tolerant *Arabidopsis csr1-2* coding sequence, and the wild-type *Arabidopsis AHASL* terminator. In addition to the S653N mutation, a second mutation was discovered in the *csr1-2* coding sequence integrated in the BPS-CV127-9 soybean genome. This mutation, in which arginine at position 272 of the AHAS1 protein is replaced by lysine, does not impact the enzymatic function of the AHAS enzyme or its herbicide-tolerance properties (Stevenson Paulik, 2007). Also included in the region originally annotated as the *AHASL* promoter was the majority of the *Arabidopsis* putative *SEC61 $\gamma$*  subunit gene (*Arabidopsis* Genome Initiative locus code At3g48570), including the entire *Arabidopsis SEC61 $\gamma$*  coding sequence.

*SEC61 $\gamma$*  subunit gene (Arabidopsis Genome Initiative locus code At3g48570), including the entire Arabidopsis *SEC61 $\gamma$*  coding sequence.

Acetohydroxyacid synthase (AHAS) is found ubiquitously in the plant kingdom. The AHAS enzyme catalyzes the first step in the biosynthesis of branched-chain amino acids in plants. In conventional imidazolinone-sensitive plants, inhibition of the AHAS enzyme by imidazolinone herbicides leads to a deficiency in branched-chain amino acids and other compounds derived from this pathway that are needed for plant survival. The *csr1-2* gene from Arabidopsis confers tolerance to imidazolinone herbicides by encoding an AHAS catalytic subunit with altered herbicide-binding properties, while retaining its normal biosynthetic function in the soybean plant.

SEC61 is a heterotrimeric complex which forms a protein-conducting channel involved in transport across the endoplasmic reticulum membrane in eukaryotes. Two of the three subunits of SEC61, the  $\alpha$  and  $\gamma$  subunits, are found in all organisms (Osborne *et al.*, 2005). Although the AtSEC61 $\gamma$  subunit transcript is detectable in BPS-CV127-9 by reverse-transcription polymerase chain reaction (RT-PCR) analysis (Shen, 2007), the AtSEC61 $\gamma$  subunit protein product is not detectable in BPS-CV127-9 soybean leaves or grain (Jiang, 2008).

The transgene insert of BPS-CV127-9 contains the coding sequences of two proteins from *A. thaliana*: AHASL (R272K, S653N) and SEC61 $\gamma$ . The purpose of this current study was to determine if additional open reading frames (ORFs) were generated by the junctions between the transgene and the soybean genome, and to conduct a bioinformatics analysis of these ORFs as part of a weight-of-evidence assessment of the potential allergenicity and toxicity of any protein potentially encoded by the transgene-genome junctions.

## MATERIALS AND METHODS

**Databases.** The Food Allergy Research and Resource Program (FARRP) Allergen Protein Database (version 9.00; [allergenonline.com](http://allergenonline.com)) containing 1389 entries was utilized for all bioinformatics assessments of potential allergenicity. These 1389 entries are comprised of known or putative food, respiratory, venom/salivary, or contact allergenic proteins. Glutenins and gliadins from *Triticum aestivum* L. (bread wheat) were included in the database if there was published evidence of immunoglobulin E (IgE) from sera of allergic individuals binding to the glutenin or gliadin in question. Some of these same glutenins and gliadins are involved in Celiac disease, a non-IgE-mediated disorder. All allergen database entries have been vetted by a panel of seven academic allergy experts based on published evidence of allergenicity.

The GenBank non-redundant peptide sequence database was utilized for all bioinformatics assessments of potential toxicity in addition to the general assessment of protein homology. The GenBank non-redundant peptide sequence data (a total of



7,946,514 sequences) was downloaded March 2, 2009 from the National Center for Biotechnology Information (NCBI) website. This database is comprised of all non-redundant GenBank coding sequence translations, protein sequences from NCBI's Reference Sequence Project, sequences derived from the Protein Data Bank, the last major release of the SWISS-PROT protein sequence database, the Protein Information Resource (PIR) Protein Sequence Database, and the Protein Research Foundation (PRF) Protein Sequence Database.

Identification of open reading frames. The junctions between the transgene insert of BPS-CV127-9 soybean and the soybean genome were analyzed for open reading frames (ORFs) using Vector NTI, version 10.3.0 (Invitrogen; Carlsbad, CA). An ORF is defined as a sequence of nucleotides which has the potential to encode a protein; whether an ORF is transcribed and ultimately translated is dependent upon having the requisite regulatory elements in the proper context. For the purposes of this study, an ORF was defined as a contiguous nucleic acid sequence that contains a string of codons that is uninterrupted until the first termination codon (i.e. TAA, TAG, or TGA) in the same reading frame is encountered. Any ORF from any of the six potential reading frames with a deduced amino acid sequence of one amino acid or greater which is located in the junctions between the insert and the soybean genome was selected for further analysis.

Based on bioinformatic analyses of the soybean genomic DNA flanking the transgene insert in BPS-CV127-9, it was determined that the transgene insert is most likely integrated into soybean chromosome 2, with 141 bp in the 5' junction having sequence identity to soybean chromosome 18 (see Sánchez-Fernández, 2009). These 141 bp were included in the ORF search on the 5' junction, which contained a total of 24 putative ORFs of one amino acid or more. The 3' junction encoded six putative ORFs. The location of the 30 unique ORFs identified within the junctions between the transgene insert and the soybean genome are schematically shown in Figure 1.

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

This eight-amino acid search was originally suggested based on the concept that eight or more amino acids is a representative minimal size for an IgE-binding epitope (Metcalf *et al.*, 1996). Bannon and Ogawa (2006) compiled a list of characterized linear IgE-binding epitopes from major allergens and, although one epitope from a wheat  $\omega$ -5 gliadin was only four amino acids long, the majority of characterized epitopes were indeed eight amino acids or longer. However, this search does not detect conformational epitopes which are formed when non-linear amino acids are brought together by the higher-order folding of the protein. Moreover, the utility of such an eight-amino acid analysis has

been questioned due to the high rate of false positives identified by this search (Silvanovich *et al.*, 2006; Hileman *et al.*, 2002).

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

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

Database search for homology to known protein toxin sequences as well as general homology to known proteins. Each ORF deduced amino acid sequence was submitted to a protein-protein Basic Local Alignment Search Tool (BLASTP), (NCBI version 2.2.18) analysis. All parameters were left at the default setting for this BLASTP analysis with the exception of the expect (E) value cut-off which was set to one. Thus, either the top 250 unique protein alignments (the default value) or all alignments with an E value less than one, whichever is fewer, were analyzed. This BLASTP analysis will identify regions of local similarity (i.e. similarity over a portion of the protein sequence rather than the entire protein sequence) between the query sequence and proteins in the downloaded March 2, 2009 GenBank non-redundant peptide sequence database.

Descriptions of the proteins from the GenBank non-redundant database which show the most significant local homology to the translation of ORF CV127\_5319\_5915\_frame2 were manually compared to known toxins which act on humans as listed in the United States Code of Federal Regulations (40 CFR Part 725.421).

## RESULTS AND DISCUSSION

Identification of open reading frames. The region at the 5' junction contained a total of 24 putative ORFs of one amino acid or more. The region at the 3' junction encoded six putative ORFs (one from each possible reading frame). The deduced amino acid sequences of the 30 unique ORFs identified within the junctions between the transgene insert and the soybean genome are shown in Figure 1.

Database search for eight-amino acid regions of identity. None of the 30 ORF deduced amino acid sequences were found to share a sequence of eight or more consecutive identical amino acids with any potential allergen found in the FARRP Allergen Protein Database. Because the eight-amino acid word-match search did not identify any common identical sequences of eight or more amino acids, there are no search results to report.

Database search with eighty-amino acid segments of query sequence. None of the deduced amino acid sequences from the 30 ORFs identified showed more than 35% identity over 80 amino acids to any potential allergens found in the FARRP Allergen Protein Database. The allergen homology search results for all putative peptides encoded by the 30 identified ORFs are shown in Appendix 1. For ORFs with deduced amino acid sequences greater than 80, only the allergen homology search results for the 80-amino acid segment with the highest homology to a known or potential allergen are shown.

Database search for homology to known protein toxin sequences as well as general homology to known proteins. The ORF deduced amino acid sequences did not show significant homology to proteins that are toxic to humans as listed in the United States Code of Federal Regulations ([40 CFR Part 725.421](#)). The toxin homology search results are shown in Appendix 2. Of the 30 ORFs analyzed, 29 ORFs did not show homology to any known protein sequence using the search parameters defined in this report. The remaining ORF had homology to AHAS. The summary of the ORF deduced amino acid sequence protein homology search is shown in Table 1.

## CONCLUSIONS

The transgene insert of BPS-CV127-9 soybean contains the coding sequences of two proteins from *A. thaliana*: AHASL (R272K, S653N) and SEC61 $\gamma$ . In this analysis, the junctions between the transgene insert of BPS-CV127-9 soybean and the soybean genome were analyzed for potential additional newly created open reading frames (ORFs) of at least one amino acid. The search covering the 5' junction included 141 bp of DNA with identity with soybean chromosome 18 and produced 24 putative ORFs. The 3' junction produced six putative ORFs. The deduced amino acid sequences of these 30 ORFs were subjected to bioinformatics analyses as part of a weight-of-evidence assessment of the potential allergenicity and toxicity of any proteins potentially created

by the BPS-CV127-9 transgene insertion into the soybean genome. No ORF deduced amino acid sequence showed significant homology to a known protein toxin as defined in the United States Code of Federal Regulations ([40 CFR Part 725.421](#)). Finally, no ORF deduced amino acid sequence shared a sequence of eight or more consecutive identical amino acids with a potential allergen. Thus, the bioinformatics analyses of the putative ORFs newly created by the BPS-CV127-9 transgene insertion into the contiguous soybean genomic DNA do not provide any indication of a potential allergenicity or toxicity concern.

## RECORDS RETENTION

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

## STUDY PERSONNEL

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

## REFERENCES

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

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

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

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

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

Jiang, X. (2008) AtSEC61 $\gamma$  subunit protein expression in Cultivance Soybean Event 127. BASF Plant Science Report Number BPS-010-07A.

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

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

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

Sánchez-Fernández, R. (2009) Bioinformatics Analysis of the genomic area surrounding the transgene in herbicide-tolerant BPS-CV127-9 Soybean. BASF Plant Science Report Number BPS-005-09

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

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

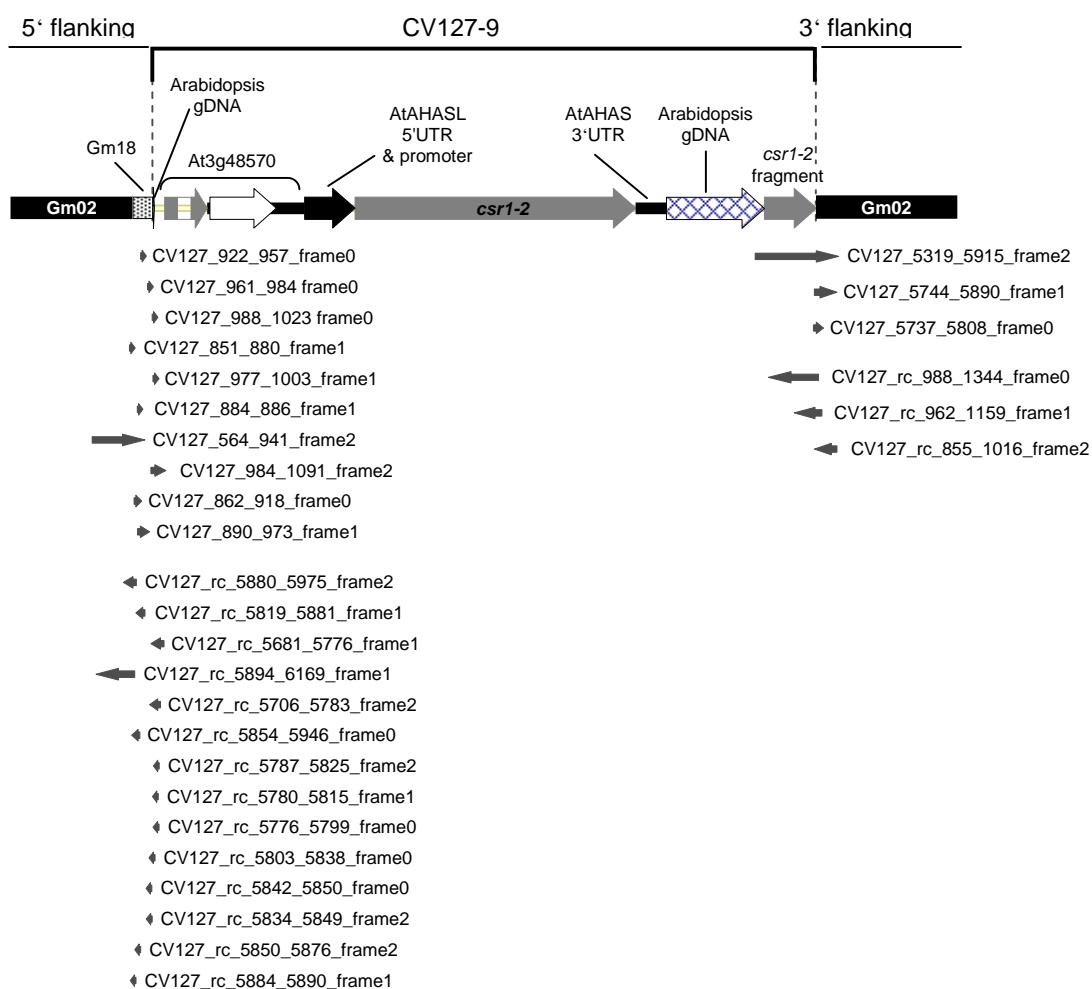
Stevenson Paulik, J. (2007) Comparison of inhibition of AtAHAS (S653N) and AtAHAS (S653N R272K) by imidazolinones and sulfonylurea. BASF Plant Science Amended Report Number BPS-004-06.

**Table 1. Summary of open reading frames present at the 5' and 3' junctions of the BPS-CV127-9 transgene insertion with the flanking soybean genomic sequence.**

Query Sequence	Sequence Length (aa)	Top BLASTP Result (accession number)
CV127_988_1023_frame0	12	None
CV127_862_918_frame0	19	None
CV127_922_957_frame0	12	None
CV127_961_984_frame0	8	None
CV127_977_1003_frame1	9	None
CV127_851_880_frame1	10	None
CV127_884_886_frame1	1	None
CV127_890_973_frame1	28	None
CV127_984_1091_frame2	36	None
CV127_564_941_frame2	126	None
CV127_rc_5776_5799_frame0	8	None
CV127_rc_5803_5838_frame0	12	None
CV127_rc_5842_5850_frame0	3	None
CV127_rc_5854_5946_frame0	31	None
CV127_rc_5681_5776_frame1	32	None
CV127_rc_5780_5815_frame1	12	None
CV127_rc_5819_5881_frame1	21	None
CV127_rc_5884_5890_frame1	2	None
CV127_rc_5894_6169_frame1	92	None
CV127_rc_5706_5783_frame2	26	None
CV127_rc_5735_5849_frame2	4	None
CV127_rc_5850_5876_frame2	9	None
CV127_rc_5880_5975_frame2	32	None
CV127_rc_5787_5825_frame2	13	None
CV127_5737_5808_frame0	24	None
CV127_5744_5890_frame1	49	None
CV127_5319_5915_frame2	199	Acetolactate synthase, <i>Arabidopsis thaliana</i> (AAM92569)
CV127_rc_988_1344_frame0	119	None
CV127_rc_962_1159_frame1	66	None
CV127_rc_855_1016_frame2	54	None

**Figure 1. Graphic representation of the insertion in soybean BPS-CV127-9.**

The insert in soybean BPS-CV127-9 (4758 bp) seems to have integrated into soybean chromosome 2 (Gm02). The 5' flanking sequence contains 141 bp of DNA that is identical to DNA in soybean chromosome 18 (see Sánchez-Fernández, 2009). Analysis of the junctions between the insert and the genomic DNA identified 30 putative ORFs, 24 of them in the 5' junction and six in the 3' junction.





**Figure 2. Deduced amino acid sequences of open reading frames from the transgene insert of BPS-CV127-9 and 1 kb of flanking sequence on each side**

The deduced amino acid sequences of open reading frames that are contained either fully or partially within the transgene insert from BPS-CV127-9 and 1 kb of flanking sequence on each side are shown below.

CV127\_988\_1023\_frame0  
IYIPQIKPNFHP

CV127\_977\_1003\_frame1  
RLIKFIFLK

CV127\_984\_1091\_frame2  
LNLYSSNKTQFSSLNEPAETLISITNSDLKRSHGSH

CV127\_rc\_5681\_5776\_frame1  
LLFRSELVIEIRVSAGSFKDENWVLFEEYKFN

CV127\_rc\_5706\_5783\_frame2  
SKLGFQQVRLRMKIGFYLRNINLINL

CV127\_5737\_5808\_frame0  
ALRRLGKL SHAHQRRQYTIRATNG

CV127\_5744\_5890\_frame1  
DVWGSCPMP IKEDSTRSELRMGRPKNKARRATQSN EGRQTNTKYPSNKDL

CV127\_5319\_5915\_frame2  
SCKFNKQKKFFIFFHFFFQHEGSYP CDEL SLHMLGMHGT VYANYAVEHSDLLLA FGVRF  
DDRVTGKLEAFASRAKIVHIDIDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELK  
LDFGVWRNELNVQKQKFPLSFKTFGEAVPCPSKKT VHDPSYEWVGPIRREGPPSPMREDK  
LTQNTHLIRTYKFVFFKCI

CV127\_rc\_988\_1344\_frame0  
WAWDSFPKRLKAQRKLLFLYVQLIPPNSKIKLKLRLSVLKNLIHPLQSQLNITTHRMRS  
LILPNLSRVNINMNNLSPTSKSLKLTRDTIIKPYSKRQQITMLHSVICIHSPMHSKHM

CV127\_rc\_962\_1159\_frame1  
LGSCTVFFDGHGTASPNVLKLN GNFCFCTFNSFLQTPKSSLSSSARFSRTLFI PCKASLT  
SPHTDT

CV127\_rc\_855\_1016\_frame2  
KIQTYRSLLDGYFVLVCLPSLDWVALLALLGLPIRSSDRVLSSLMGMGQLPQTS

CV127\_862\_918\_frame0  
NIKIFILKLNKRGPPLYLSP

CV127\_922\_957\_frame0  
FLFSFLSNKGAN

CV127\_961\_984\_frame0  
SRNILEVN

CV127\_851\_880\_frame1  
NGNKISKSLY

CV127\_884\_886\_frame1  
S

CV127\_890\_973\_frame1  
TKGALLIYPLSFYFHFLLIKGQTS�VIY

CV127\_564\_941\_frame2  
LWLYCLIDYLCCKLTCTNDFSTMCQLSLSFLKNKWLVSSSLKKIGSLGLKIEIAIGKINL  
GRVWEKVASNILFNSVQELLRRLVYRKAQTDVWKLETAIKYONLYIKAEQKGPSLFIPLV  
FIFISF

CV127\_rc\_5776\_5799\_frame0  
LTSNILRD

CV127\_rc\_5803\_5838\_frame0  
FAPLLERNENKN

CV127\_rc\_5842\_5850\_frame0  
GNK

CV127\_rc\_5854\_5946\_frame0  
GGPLLFSFNIKILIFYCRFKLPNISLRFPIH

CV127\_rc\_5780\_5815\_frame1  
PLIYYETSLPLY

CV127\_rc\_5819\_5881\_frame1  
KEMKIKTKGINKEGPFCSALI

CV127\_rc\_5884\_5890\_frame1  
RF

CV127\_rc\_5894\_6169\_frame1  
YFIAVSSFQTSVCAFLYTKRRRSSCTELKRMLDATFSQTRPKLILPIAISIFKPKLPIFL  
RDEDTSHLFFKKDKLNWHIVEKSLVFVRHLQR

CV127\_rc\_5787\_5825\_frame2  
YITRLVCPFIRKK

CV127\_rc\_5729\_5831\_frame2  
K

CV127\_rc\_5735\_5849\_frame2  
KLRG

CV127\_rc\_5850\_5876\_frame2  
IRRAPFVQL

CV127\_rc\_5880\_5975\_frame2  
YKDFDILLPFQASKHQFALSYTLNGEGALVLN

## APPENDIX 1. 80-AMINO ACID ALLERGEN HOMOLOGY SEARCH RESULTS

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

### CV127\_5737\_5808\_frame0

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

Please cite:

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

CV127\_5737\_5808\_frame0, 24 aa

vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences

Expectation\_n fit: rho(ln(x))= 3.4551+/-0.0031; mu= 4.1464+/- 0.162

mean\_var=23.3437+/- 7.656, 0's: 26 Z-trim: 31 B-trim: 9 in 1/41

Lambda= 0.265454

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

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

The best scores are:

					opt bits	E(1386)	%_id	%_sim	alen	
gi 2129805	gi 2129805 pir	S65143	p ( 80)		47	22.4	0.5	0.333	0.708	24
gi 2129801	gi 2129801 pir	S65144	p ( 83)		47	22.4	0.51	0.333	0.708	24
gi 2129802	gi 2129802 pir	S65145	p ( 83)		47	22.4	0.51	0.333	0.708	24
gi 59800145	gi 59800145 sp	P69198.1	( 83)		47	22.4	0.51	0.333	0.708	24
gi 59800146	gi 59800146 sp	P69199.1	( 83)		47	22.4	0.51	0.333	0.708	24
gi 29465668	gi 29465668 gb	AAL92871	( 86)		40	19.7	3.4	0.333	0.708	24
gi 20387027	gi 20387027 emb	CAC8459	( 284)		44	21.4	3.5	0.636	0.818	11
gi 1247373	gi 1247373 emb	CAA01909.	( 71)		39	19.3	3.7	0.375	0.667	24
gi 1247375	gi 1247375 emb	CAA01910.	( 73)		39	19.3	3.8	0.375	0.667	24
gi 14423846	gi 14423846 sp	O82040.1	( 78)		39	19.3	4.1	0.375	0.667	24
gi 1871507	gi 1871507 emb	CAA62634.	( 82)		39	19.3	4.2	0.375	0.667	24
gi 1311511	gi 1311511 gb	AAB36010.1	( 15)		31	16.0	7.5	0.500	0.750	8
gi 114841683	gi 114841683 dbj	BAF32	( 419)		42	20.6	8.6	0.389	0.611	18
gi 59800143	gi 59800143 sp	P69196.1	( 79)		36	18.1	9.1	0.292	0.667	24
gi 59800144	gi 59800144 sp	P69197.1	( 79)		36	18.1	9.1	0.292	0.667	24
gi 21913174	gi 21913174 gb	AAM77471	( 115)		37	18.6	9.9	0.625	1.000	8

>>>CV127\_5737\_5808\_frame0, 24 aa vs

/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

>>gi|2129805 gi|2129805|pir|S65143 pollen allergen grou (80 aa)

initn: 38 initl: 38 opt: 47 Z-score: 107.4 bits: 22.4 E(): 0.5

Smith-Waterman score: 47; 33.333% identity (70.833% similar) in 24 aa overlap (1-24:30-53)

			10		20	
CV127_			ALRRLGKLSHAHQRRQYTIRATNG			
			::: ::::: : ::::: :::			
gi 212	ETERAEHDRIFFKFDANGDGKISASELGDALKNLGSVTHDDIKRMMAEIDTDGDGYISYQ					
	10 20 30 40 50 60					
gi 212	EFSDFFASANRGLMKDVAKIF					

70 80

>>gi|2129801 gi|2129801|pir||S65144 pollen allergen grou (83 aa)  
 initn: 38 initl: 38 opt: 47 Z-score: 107.1 bits: 22.4 E(): 0.51  
 Smith-Waterman score: 47; 33.333% identity (70.833% similar) in 24 aa overlap  
 (1-24:33-56)

CV127\_ ALRRLGKLSHAHQRRQYTIRATNG  
 :: ::...: .:... ::  
 gi|212 MADATEKTEHDRFFKKFDANGDGTISSTELGDALKNLGVSVDHDDIKRMMAEIDTDGDGFI  
 10 20 30 40 50 60

gi|212 SYQEFSDFAKANRGLMKDVAKIF  
 70 80

>>gi|2129802 gi|2129802|pir||S65145 pollen allergen grou (83 aa)  
 initn: 38 initl: 38 opt: 47 Z-score: 107.1 bits: 22.4 E(): 0.51  
 Smith-Waterman score: 47; 33.333% identity (70.833% similar) in 24 aa overlap  
 (1-24:33-56)

CV127\_ ALRRLGKLSHAHQRRQYTIRATNG  
 :: ::...: .:... ::  
 gi|212 MADATEKAEHDRFFKKFDANGDGTISSTELGDALKNLGVSVDHDDIKRMMAEIDTDGDGFI  
 10 20 30 40 50 60

gi|212 SYQEFSDFAKANRGLMKDVAKIF  
 70 80

>>gi|59800145 gi|59800145|sp|P69198.1|POLC2\_BRANA RecNam (83 aa)  
 initn: 38 initl: 38 opt: 47 Z-score: 107.1 bits: 22.4 E(): 0.51  
 Smith-Waterman score: 47; 33.333% identity (70.833% similar) in 24 aa overlap  
 (1-24:33-56)

CV127\_ ALRRLGKLSHAHQRRQYTIRATNG  
 :: ::...: .:... ::  
 gi|598 MADATEKAEHDRIFKKFDANGDGKISASELGDALKNLGVSVDHDDIKRMMAEIDTDGDGYI  
 10 20 30 40 50 60

gi|598 SYQEFSDFAKANRGLMKDVAKIF  
 70 80

>>gi|59800146 gi|59800146|sp|P69199.1|POLC2\_BRARA RecNam (83 aa)  
 initn: 38 initl: 38 opt: 47 Z-score: 107.1 bits: 22.4 E(): 0.51  
 Smith-Waterman score: 47; 33.333% identity (70.833% similar) in 24 aa overlap  
 (1-24:33-56)

CV127\_ ALRRLGKLSHAHQRRQYTIRATNG  
 :: ::...: .:... ::  
 gi|598 MADATEKAEHDRIFKKFDANGDGKISASELGDALKNLGVSVDHDDIKRMMAEIDTDGDGYI  
 10 20 30 40 50 60

gi|598 SYQEFSDFAKANRGLMKDVAKIF  
 70 80

>>gi|29465668 gi|29465668|gb|AAL92871.1| pollen allergen (86 aa)  
 initn: 35 initl: 35 opt: 40 Z-score: 92.4 bits: 19.7 E(): 3.4  
 Smith-Waterman score: 40; 33.333% identity (70.833% similar) in 24 aa overlap  
 (1-24:36-59)

```

                                10      20
CV127_                ALRRLGKLSHAHQRRQYTIRATNG
                        :: :... : :... ::
gi|294 MAAEDTPQDIADRERIFKRFD TNGDGKISSSELGDALKTLG SVTPDEVRRMMAEIDTDGD
                        10      20      30      40      50      60

gi|294 GFISFDEFTDFARANRGLVKDVSKIF
                        70      80

>>gi|20387027 gi|20387027|emb|CAC84590.2| tropomyosin [L (284 aa)
  initn: 44 initl: 44 opt: 44 Z-score: 92.1 bits: 21.4 E(): 3.5
Smith-Waterman score: 44; 63.636% identity (81.818% similar) in 11 aa overlap
(1-11:109-119)

gi|203 MEAIKKKMQAMKLEKDNAMDKADALEAQARDANRKADKILEEVQDLKKKPSQVETDFTTT
                        10      20      30      40      50      60

CV127_                ALRRLGKLSHAH
                        :: :... ::
gi|203 KENLATANKNLEDKEKTLTNTSEVASLNRKVQMIEENLERSEERLGTALTKLGEASHAA
                        70      80      90      100     110     120

                        20
CV127_ QRRQYTIRATNG

gi|203 DEASRMCKVLENRSQQDEERMDQLTNQLKEARMLAEDADGKSDEVSRKMAQVEDDLEVAE
                        130     140     150     160     170     180

gi|203 DRVKS GDSKIMELEEELKVVGNSLKSLEVSEEKANQRVEEYKRQIKTLTVKLKEAEARAE
                        190     200     210     220     230     240

gi|203 YAEKYVKKLQKEVDRLEDELGINKDRYRALADEMDQTFAELSGY
                        250     260     270     280

>>gi|1247373 gi|1247373|emb|CAA01909.1| B1 protein aller (71 aa)
  initn: 33 initl: 33 opt: 39 Z-score: 91.7 bits: 19.3 E(): 3.7
Smith-Waterman score: 39; 37.500% identity (66.667% similar) in 24 aa overlap
(1-24:21-44)

                                10      20
CV127_                ALRRLGKLSHAHQRRQYTIRATNG
                        :: :... : :... ::
gi|124 GTRRFDTNGDGKISLAELTDALRTL GSTSADEVQRMMAEIDTDG DGFIDFDEFISFCNAN
                        10      20      30      40      50      60

gi|124 PGLMKDVAKVF
                        70

>>gi|1247375 gi|1247375|emb|CAA01910.1| B4 protein aller (73 aa)
  initn: 33 initl: 33 opt: 39 Z-score: 91.5 bits: 19.3 E(): 3.8
Smith-Waterman score: 39; 37.500% identity (66.667% similar) in 24 aa overlap
(1-24:23-46)

                                10      20
CV127_                ALRRLGKLSHAHQRRQYTIRATNG
                        :: :... : :... ::
gi|124 GTSFKRFD TNGDGKISLAELTDALRTL GSTSADEVQRMMAEIDTDG DGFIDFDEFISFCN
                        10      20      30      40      50      60

gi|124 ANPGLMKDVAKVF

```

70

```
>>gi|14423846 gi|14423846|sp|O82040.1|POLC7_PHLPR RecNam (78 aa)
  initn: 33 initl: 33 opt: 39 Z-score: 91.0 bits: 19.3 E(): 4.1
Smith-Waterman score: 39; 37.500% identity (66.667% similar) in 24 aa overlap
(1-24:28-51)
```

CV127\_ 10 20  
ALRRLGKLSHAHQRRQYTIRATNG  
::: ::: : . . . . . :::  
gi|144 MADDMERIFKRFDTNGDGKISLSELTDLALRTLTGSTSADDEVQRMMAEIDTDGDGFIDFNEF  
10 20 30 40 50 60

gi | 144 ISFCNANPGLMKDVAKVF  
70

```
>>gi|1871507 gi|1871507|emb|CAA62634.1| calcium-binding (82 aa)
  initn: 33 initl: 33 opt: 39 Z-score: 90.6 bits: 19.3 E(): 4.2
Smith-Waterman score: 39; 37.500% identity (66.667% similar) in 24 aa overlap
(1-24:32-55)
```

CV127\_ 10 20  
ALRRLGKLSHAHQRRQYTIRATNG  
::: :: : . : : : : : :  
gi|187 KTMADTGDMEHIFKRFDTNGDGKISLAELTDALRTLGSTSADEVQRMMAEIDTDGDFID  
10 20 30 40 50 60

gi | 187 FDEFISFCNANPGLMKDVAKV  
70 80

```
>>gi|1311511 gi|1311511|gb|AAB36010.1| mAb 8C7-reactive (15 aa)
  initn: 31 initl: 31 opt: 31 Z-score: 86.2 bits: 16.0 E(): 7.5
Smith-Waterman score: 31; 50.000% identity (75.000% similar) in 8 aa overlap
(5-12:3-10)
```

```

              10      20
CV127_  ALRRLGKLSHAHQRRQYTIRATNG
              .: . :::
gi|131   GPVGGVVHAHMMPLL
              10

```

```
>>gi|114841683 gi|114841683|dbj|BAF32143.1| pollen aller (419 aa)
  initn: 42 init1: 42 opt: 42 Z-score: 85.2 bits: 20.6 E(): 8.6
Smith-Waterman score: 42; 38.889% identity (61.111% similar) in 18 aa overlap
(7-24:263-280)
```

gi | 114 LDSDIEQYLRNRSLLKKLVHSRHDAAATVFNVEQYGAVGDGKHDSTAFETAWNAACKKAS

gi | 114 AVLVVPANKKFFVNNLVFRGPCQPHLSFKVDGTIVAQPDPARWKNSKIWLQFAQLTDFNL  
70 80 90 100 110 120gi | 114 MGTGVIDGGQQWWAGQCKVNVNGRTVCNDRNRPTAIKIDYSKSVTVKELTLMNSPEFHLV  
130 140 150 160 170 180gi|114 FGECEGVKIQGLKIKAPRDSNTDGDIDIFASKRFHIEKCVIGTGDDCIAIGTGSSNITIK  
190 200 210 220 230 240

```

                                10      20
CV127_      ALRRLGKLSHAHQRRQYTIRATNG
              ..:::  :  :  .  ::
qi|114 DLICPGHGISIGSLGRDNSRAEVSHVHVNRAKFIDTONGLRIKTWOGGSLASYITYEN

```

```

                250          260          270          280          290          300
gi|114 VEMINSENPIILINQFYCTTSASACQNRSAVQIQGVTYKNIHGTSATAAAIQLMCSDSVPC
                310          320          330          340          350          360

gi|114 TGIQLSNVSLKLTSGKPASCVDKNARGFYSGRLIPTCKNLRPGPSPKEFELQQPTTVM
                370          380          390          400          410

>>gi|59800143 gi|59800143|sp|P69196.1|POLC1_BRANA RecNam (79 aa)
  initn: 32 initl: 32 opt: 36 Z-score: 84.7 bits: 18.1 E(): 9.1
Smith-Waterman score: 36; 29.167% identity (66.667% similar) in 24 aa overlap
(1-24:29-52)

                10          20
CV127_          ALRRLGKLSHAHQRRQYTIRATNG
                :::::::::::....:::
gi|598 MADAHERIFKKFDTGDGKISAAELEEALKKLGSVTPDDVTRMMAKIDTDGDGNISFQE
                10          20          30          40          50          60

gi|598 FTEFASANPGLMKDVAKVF
                70

>>gi|59800144 gi|59800144|sp|P69197.1|POLC1_BRARA RecNam (79 aa)
  initn: 32 initl: 32 opt: 36 Z-score: 84.7 bits: 18.1 E(): 9.1
Smith-Waterman score: 36; 29.167% identity (66.667% similar) in 24 aa overlap
(1-24:29-52)

                10          20
CV127_          ALRRLGKLSHAHQRRQYTIRATNG
                :::::::::::....:::
gi|598 MADAHERIFKKFDTGDGKISAAELEEALKKLGSVTPDDVTRMMAKIDTDGDGNISFQE
                10          20          30          40          50          60

gi|598 FTEFASANPGLMKDVAKVF
                70

>>gi|21913174 gi|21913174|gb|AAM77471.1| major allergen (115 aa)
  initn: 37 initl: 37 opt: 37 Z-score: 84.1 bits: 18.6 E(): 9.9
Smith-Waterman score: 37; 62.500% identity (100.000% similar) in 8 aa overlap
(17-24:60-67)

                10
CV127_          ALRRLGKLSHAHQRRQY
                .
gi|219 MQFTTIASLFAAAGLAAAAPLESRQDNASCPVTTKGDYVWKISEFYGRKPEGTYNSLGF
                10          20          30          40          50          60

                20
CV127_ TIRATNG
                .....
gi|219 NIKATNGGTLDFTCSAQADKLEDHKWYSCGENSFMDFSFSDRSGLLLKQKVSDE
                70          80          90          100          110

24 residues in 1 query sequences
307888 residues in 1386 library sequences
Scomplib [34.26]
start: Wed Mar 11 05:56:24 2009 done: Wed Mar 11 05:56:24 2009
Total Scan time: 0.060 Total Display time: 0.010

```



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

## CV127\_5744\_5890\_frame1

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

Please cite:

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

CV127\_5744\_5890\_frame1, 49 aa

vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences

Expectation\_n fit:  $\rho(\ln(x)) = 5.8900 \pm 0.004$ ;  $\mu = -4.8365 \pm 0.206$

mean\_var=40.6831 $\pm$ -10.906, 0's: 12 Z-trim: 12 B-trim: 106 in 1/41

Lambda= 0.201079

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

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

The best scores are:

	opt	bits	E(1386)	%_id	%_sim	alen	
gi 162797 gi 162797 gb AAA30430.1	( 224)	57	22.9	1.9	0.394	0.576	33
gi 162805 gi 162805 gb AAA30431.1	( 224)	57	22.9	1.9	0.394	0.576	33
gi 459292 gi 459292 gb AAB29137.1	( 224)	57	22.9	1.9	0.394	0.576	33
gi 162931 gi 162931 gb AAA30480.1	( 224)	57	22.9	1.9	0.394	0.576	33
gi 47606039 gi 47606039 sp Q8H6L7.1	( 143)	51	21.3	3.7	0.375	0.656	32
gi 111120432 gi 111120432 gb ABH063	( 129)	50	21.1	4	0.333	0.800	15
gi 111120424 gi 111120424 gb ABH063	( 129)	50	21.1	4	0.333	0.800	15
gi 111494253 gi 111494253 gb ABH063	( 129)	50	21.1	4	0.333	0.800	15
gi 111120420 gi 111120420 gb ABH063	( 129)	50	21.1	4	0.333	0.800	15
gi 111120428 gi 111120428 gb ABH063	( 129)	50	21.1	4	0.333	0.800	15
gi 3703107 gi 3703107 gb AAC63045.1	( 507)	57	22.7	5	0.300	0.667	30
gi 21779 gi 21779 emb CAA26847.1	u ( 660)	58	22.9	5.5	0.256	0.605	43
gi 13183177 gi 13183177 gb AAK15089	( 585)	57	22.7	5.9	0.364	0.667	33
gi 20387029 gi 20387029 emb CAC8459	( 243)	51	21.2	6.9	0.293	0.585	41
gi 6094504 gi 6094504 sp Q25456.1	T ( 274)	51	21.2	8	0.293	0.585	41
gi 14285797 gi 14285797 sp O61379.1	( 274)	51	21.2	8	0.293	0.585	41
gi 125995159 gi 125995159 dbj BAF47	( 284)	51	21.1	8.3	0.293	0.585	41
gi 2660868 gi 2660868 gb AAC48288.1	( 284)	51	21.1	8.3	0.293	0.585	41
gi 125995167 gi 125995167 dbj BAF47	( 284)	51	21.1	8.3	0.293	0.585	41
gi 148615631 gi 148615631 gb ABQ966	( 284)	51	21.1	8.3	0.293	0.585	41
gi 125995163 gi 125995163 dbj BAF47	( 284)	51	21.1	8.3	0.293	0.585	41
gi 125995157 gi 125995157 dbj BAF47	( 284)	51	21.1	8.3	0.293	0.585	41
gi 73532979 gi 73532979 gb AAZ76743	( 284)	51	21.1	8.3	0.293	0.585	41
gi 14285796 gi 14285796 sp O44119.1	( 284)	51	21.1	8.3	0.293	0.585	41
gi 125995165 gi 125995165 dbj BAF47	( 284)	51	21.1	8.3	0.293	0.585	41
gi 162286975 gi 162286975 dbj BAF95	( 284)	51	21.1	8.3	0.293	0.585	41
gi 156712754 gi 156712754 dbj BAF76	( 284)	51	21.1	8.3	0.293	0.585	41
gi 125995161 gi 125995161 dbj BAF47	( 284)	51	21.1	8.3	0.293	0.585	41
gi 14423757 gi 14423757 sp O04701.1	( 246)	50	20.9	8.6	0.333	0.704	27
gi 5931948 gi 5931948 gb AAD56719.1	( 160)	47	20.1	9.4	0.312	0.625	32
gi 25991543 gi 25991543 gb AAN76862	( 457)	53	21.6	9.8	0.268	0.683	41

>>>CV127\_5744\_5890\_frame1, 49 aa vs

/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

>>gi|162797 gi|162797|gb|AAA30430.1| beta-casein precurs (224 aa)

initn: 57 init1: 57 opt: 57 Z-score: 97.0 bits: 22.9 E(): 1.9

Smith-Waterman score: 57; 39.394% identity (57.576% similar) in 33 aa overlap (7-39:24-56)

```

              10      20      30      40
CV127_          DVWGSCPMPIKEDSTRSELRMGRPNKARRATQSNAGRQNTKY
              :  :  .  .  :  :  .  :  :  .  :  :  :  :
gi|162 MKVLILACLVALALARELEELNVPGEIVESLSSEESITRINKKIEKFQSEEQQQTEDEL
              10      20      30      40      50      60

```

```

CV127_ PSNKDL

gi|162 QDKIHPPAQTQSLVYPFGPIPNLQNIPLTQTPVVVPPFLQPEVLGVSKVKEAMAPK
              70      80      90      100     110     120

gi|162 HKEMPFPPKYPVEPFTESSQSLTLTDVENLHLPPLLQSWMHQPHQLPPTVMFPPQSVLSL
              130     140     150     160     170     180

gi|162 SQSKVLPVPQKAVPYPQRDMPIQAFLLYQQPVLGPVGRGPFPIIV
              190     200     210     220

```

>>gi|162805 gi|162805|gb|AAA30431.1| beta-casein (224 aa)  
 initn: 57 initl: 57 opt: 57 Z-score: 97.0 bits: 22.9 E(): 1.9  
 Smith-Waterman score: 57; 39.394% identity (57.576% similar) in 33 aa overlap  
 (7-39:24-56)

```

              10      20      30      40
CV127_          DVWGSCPMPIKEDSTRSELRMGRPNKARRATQSNAGRQNTKY
              :  :  .  .  :  :  .  :  :  .  :  :  :  :
gi|162 MKVLILACLVALALARELEELNVPGEIVESLSSEESITRINKKIEKFQSEEQQQTEDEL
              10      20      30      40      50      60

```

```

CV127_ PSNKDL

gi|162 QDKIHPPAQTQSLVYPFGPIHNSLPQNIPLTQTPVVVPPFLQPEVMGVSKVKEAMAPK
              70      80      90      100     110     120

gi|162 HKEMPFPPKYPVEPFTESSQSLTLTDVENLHLPPLLQSWMHQPHQLPPTVMFPPQSVLSL
              130     140     150     160     170     180

gi|162 SQSKVLPVPQKAVPYPQRDMPIQAFLLYQEPVLGPVGRGPFPIIV
              190     200     210     220

```

>>gi|459292 gi|459292|gb|AAB29137.1| beta-casein A3 [Bos (224 aa)  
 initn: 57 initl: 57 opt: 57 Z-score: 97.0 bits: 22.9 E(): 1.9  
 Smith-Waterman score: 57; 39.394% identity (57.576% similar) in 33 aa overlap  
 (7-39:24-56)

```

              10      20      30      40
CV127_          DVWGSCPMPIKEDSTRSELRMGRPNKARRATQSNAGRQNTKY
              :  :  .  .  :  :  .  :  :  .  :  :  :  :
gi|459 MKVLILACLVALALARELEELNVPGEIVESLSSEESITRINKKIEKFQSEEQQQTEDEL
              10      20      30      40      50      60

```

```

CV127_ PSNKDL

gi|459 QDKIHPPAQTQSLVYPFGPIPNLQNIPLTQTPVVVPPFLQPEVMGVSKVKEAMAPK
              70      80      90      100     110     120

gi|459 QKEMPFPPKYPVEPFTESSQSLTLTDVENLHLPPLLQSWMHQPHQLPPTVMFPPQSVLSL
              130     140     150     160     170     180

```

gi|459 SQSKVLPVPQKAVPYPQRDMPPIQAFLLYQEPVLGPVRGPFPIIV  
190 200 210 220

>>gi|162931 gi|162931|gb|AAA30480.1| beta-casein precurs (224 aa)  
initn: 57 initl: 57 opt: 57 Z-score: 97.0 bits: 22.9 E(): 1.9  
Smith-Waterman score: 57; 39.394% identity (57.576% similar) in 33 aa overlap  
(7-39:24-56)

CV127\_ DVWGSCMPIKEDSTRSELRMGPRNKARRATQSNEGRQTNTKY  
: : . . : : . : : : . : : : . : :  
gi|162 MKVLILACLVALALARELEELNVPGEIVESLSSEESITRINKKIEKFQSEEQQQTEDEL  
10 20 30 40 50 60

CV127\_ PSNKDL

gi|162 QDKIHPFAQTQSLVYPFGPIPNSLPQNIPPLTQTPVVPFPLQPEVMGVSKVKEAMAPK  
70 80 90 100 110 120

gi|162 HKEMPFKYPVEPFTESSQLTLTDVENLHLPLPLLQSWMHQPHQLPPTVMFPPQSVLSL  
130 140 150 160 170 180

gi|162 SQSKVLPVPQKAVPYPQRDMPPIQAFLLYQEPVLGPVRGPFPIIV  
190 200 210 220

>>gi|47606039 gi|47606039|sp|Q8H6L7.1|PHLB\_PHLPR RecName (143 aa)  
initn: 26 initl: 26 opt: 51 Z-score: 91.7 bits: 21.3 E(): 3.7  
Smith-Waterman score: 51; 37.500% identity (65.625% similar) in 32 aa overlap  
(13-43:82-112)

gi|476 DKGPFGVVTGRVYCDPCRAGFETNVSHNVQGATVAVDCRPFNGGESKLAETDGLGWY  
10 20 30 40 50 60

CV127\_ DVWGSCMPIKEDSTRSELRMGPRNKARRATQSNEG-RQTNTKYPSNKDL  
: : : : : : : : : : : : : : : :  
gi|476 KIEIDQDHQEEICEVVLAKSPDTCSEIEEFR-DRARVPLTSNNGIKQQGIRYANPIAFF  
70 80 90 100 110

gi|476 RKEPLKECGGILQAYDLRDAPETP  
120 130 140

>>gi|111120432 gi|111120432|gb|ABH06350.1| Blo t 21 alle (129 aa)  
initn: 21 initl: 21 opt: 50 Z-score: 91.1 bits: 21.1 E(): 4  
Smith-Waterman score: 50; 33.333% identity (80.000% similar) in 15 aa overlap  
(5-19:14-28)

CV127\_ DVWGSCMPIKEDSTRSELRMGPRNKARRATQSNEGRQTNTKYPSNKDL  
: : : : : : : : : : : : : : : :  
gi|111 MKFIIALAALIIVACALPVSDNFRHEFDHMIVNTATQRFHEIEKFLHITHEVDDLEET  
10 20 30 40 50 60

gi|111 GNKDEKARLLRELTVSEAFIEGSRGYFQRELKRTDLDLLEKFNFEALATGDLKDLKA  
70 80 90 100 110 120

gi|111 LQKRVQDSE

>>gi|111120424 gi|111120424|gb|ABH06346.1| Blo t 21 alle (129 aa)  
initn: 21 initl: 21 opt: 50 Z-score: 91.1 bits: 21.1 E(): 4

Smith-Waterman score: 50; 33.333% identity (80.000% similar) in 15 aa overlap (5-19:14-28)

```

              10      20      30      40
CV127_      DVWGSCMPIKEDSTRSELRMGRPKNKARRATQSNAGRQTNTKYPSNKDL
              .: .: .: .: .: .:
gi|111 MKFIIALAALIAVACALPVSNDNFRHEFDHMIVNTATQRFHEIEKFLHITHEVDDLEKT
              10      20      30      40      50      60
gi|111 GNKDEKARPLRELTVSEAFIEGSRGYFQRELKRTDLDLLEKFNFEAALATGDLLLKDLKA
              70      80      90      100     110     120
gi|111 LQKRVQDSE

```

>>gi|111494253 gi|111494253|gb|ABH06347.1| Blo t 21 alle (129 aa)  
initn: 21 init1: 21 opt: 50 Z-score: 91.1 bits: 21.1 E(): 4  
Smith-Waterman score: 50; 33.333% identity (80.000% similar) in 15 aa overlap (5-19:14-28)

```

              10      20      30      40
CV127_      DVWGSCMPIKEDSTRSELRMGRPKNKARRATQSNAGRQTNTKYPSNKDL
              .: .: .: .: .: .:
gi|111 MKFIIALAALIAVACALPVSNDNFRHEFDHMIVNAATQRFHEIEKFLHITHEVDDLEKT
              10      20      30      40      50      60
gi|111 GNKDEKARLLRELTVSEAFIEGSRGYFQRELKRTDLDLLEKFNFEAALATGDLLLKDLKA
              70      80      90      100     110     120
gi|111 LQKRVQDSE

```

>>gi|111120420 gi|111120420|gb|ABH06344.1| Blo t 21 alle (129 aa)  
initn: 21 init1: 21 opt: 50 Z-score: 91.1 bits: 21.1 E(): 4  
Smith-Waterman score: 50; 33.333% identity (80.000% similar) in 15 aa overlap (5-19:14-28)

```

              10      20      30      40
CV127_      DVWGSCMPIKEDSTRSELRMGRPKNKARRATQSNAGRQTNTKYPSNKDL
              .: .: .: .: .: .:
gi|111 MKFIIALAALIAVACALPVSNDNFRHEFDHMIVNTATQRFHEIEKFLHITHEVDDLEKT
              10      20      30      40      50      60
gi|111 GNKDEKARLLRELTVSEAFIEGSRGYFQRELKRTDLDLLEKFNFEAALATGDLLLKDLKA
              70      80      90      100     110     120
gi|111 LQKRVQDSE

```

>>gi|111120428 gi|111120428|gb|ABH06348.1| Blo t 21 alle (129 aa)  
initn: 21 init1: 21 opt: 50 Z-score: 91.1 bits: 21.1 E(): 4  
Smith-Waterman score: 50; 33.333% identity (80.000% similar) in 15 aa overlap (5-19:14-28)

```

              10      20      30      40
CV127_      DVWGSCMPIKEDSTRSELRMGRPKNKARRATQSNAGRQTNTKYPSNKDL
              .: .: .: .: .: .:
gi|111 MKFIIALAALIAVACALPVSNDNFRHEFDHMIINTATQRFHEIEKFLHITHEVDDLEKT
              10      20      30      40      50      60
gi|111 GNKDEKARLLRELTVSEAFIEGSRGYFQRELKRTDLDLLEKFNFEAALATGDLLLKDLKA
              70      80      90      100     110     120

```

gi|111 LQKRVQDSE

>>gi|3703107 gi|3703107|gb|AAC63045.1| glycinin [Arachis (507 aa)  
 initn: 48 initl: 48 opt: 57 Z-score: 89.4 bits: 22.7 E(): 5  
 Smith-Waterman score: 57; 30.000% identity (66.667% similar) in 30 aa overlap  
 (14-43:277-306)

gi|370 RQQPEENACQFQRLNAQRPDNRIESEGGYIETWNPNNQEFECAGVALSRLVLRNALRRP  
 10 20 30 40 50 60

gi|370 FYSNAPQEIFIQQGRGYFGLIFPGCPRHYYEPTQGRRSQSQRPPRRLQGEDQSQQQRDS  
 70 80 90 100 110 120

gi|370 HQKVHRFDEGDLIAVPTGVAFWLYNDHDTDVVAVSLTDTNNNDNQLDQFPFRFNLAGNTE  
 130 140 150 160 170 180

gi|370 QEFLRYQQQSRQSRRLSPYSPYSPQSPRQEEREFSPRGQHSRRERAGQEEENEGGNIF  
 190 200 210 220 230 240

CV127\_ DVWGSCPMPIKEDSTRSELRMGRPKNKARRATQSNRGR  
 10 20 30  
 . . . :  
 gi|370 SGFTPEFLEQAFQVDDRQIVQNLRGETESEEEGAIVTVRGGLRILSPDRKRRADEEEEEYD  
 250 260 270 280 290 300

CV127\_ QTNTKYPSNKDL  
 . . . :  
 gi|370 EDEYEEYDEEDRRRGRGSRGRNGIEETICTASAKKNIGNRNSPDIYNPQAGSLKTANDLN  
 310 320 330 340 350 360

gi|370 LLILRWLGPSAEYGNLYRNALFVAHYNTNAHSIIYRLRGRAHVQVDSNGNRVYDEELQE  
 370 380 390 400 410 420

gi|370 GHVLVVPQNFVAVAGKSQSENFEYVAFKTDSPSIANLAGENSVIDNLPEEVVANSYGLQR  
 430 440 450 460 470 480

gi|370 EQARQLKNNNPFKFFVPPSQSPRAVA  
 490 500

>>gi|21779 gi|21779|emb|CAA26847.1| unnamed protein prod (660 aa)  
 initn: 39 initl: 39 opt: 58 Z-score: 88.6 bits: 22.9 E(): 5.5  
 Smith-Waterman score: 58; 25.581% identity (60.465% similar) in 43 aa overlap  
 (4-46:544-582)

gi|217 MAKRLVLFAAVVIALVALTTAEGEASRLQCERELQESSLEACRQVVDQQLAGRLPWSTG  
 10 20 30 40 50 60

gi|217 LQMRCCQQLRDVSAKCRSVAVSQVARQYEQTVPVPPKGGSFYPGETTPLLQQLQGIFWGTS  
 70 80 90 100 110 120

gi|217 SQTQGGYYPSTSPRQGSYYPGQASPQQPGQGQPGKWQEPGQGGQWYYPTSLQQPGQGQ  
 130 140 150 160 170 180

gi|217 QIGKGKQGGYPTSLQQPGQGQIGQGQGGYPTSPQHTGQRQQPVGQGGQIGQGQGPQEQGQ  
 190 200 210 220 230 240

gi|217 QPGQWQGGYPTSPQQLGQGQPGQWQSGQGQGGHYPTSLQQPGQGQGGHYLASQQQPA  
 250 260 270 280 290 300

```
gi|217 QGQGHYPASQQPGQGQGHYPASQQPGQGQGHYPASQQEPGQGGQIPASQQQPG
      310      320      330      340      350      360
```

```
gi|217 QGQGHYPASLQQPGQGQHYPTSLQQLGQGGQIGQPGQKQQPGQGGQTGGQGPQEQEQP
      370      380      390      400      410      420
```

```
gi|217 GQGQGGYPTSLQQPGQGQGGQGGQGGYPTSLQQPGQGQGHYPASLQQPGQGQGPQG
      430      440      450      460      470      480
```

```
gi|217 RQPGQGQHPEQGQPGQGQGGYPTSPQQPGQGQGLGQGGYPTSPQQPGQGQGPQG
      490      500      510      520      530      540
```

```

      10      20      30      40
CV127_ DVWGSCPMPIKEDSTRSELRMGRPNKARRATQSNEGRQTNTKYPSNKDL
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi|217 GQGQHCPMSPQQTGQAQQL--GQGQIGQVQPGQGQGGY--YPTSLQQPGQGQSGQGQ
      550      560      570      580      590
```

```
gi|217 QSGQGHQPGQGQSGQEKQGYDSPYHVSAAEQQAASPMVAKAQQPATQLPTVCRMEGGDAL
      600      610      620      630      640      650
```

```
gi|217 SASQ
      660
```

```
>>gi|13183177 gi|13183177|gb|AAK15089.1|AF240006_1 7S gl (585 aa)
initn: 48 initl: 48 opt: 57 Z-score: 88.1 bits: 22.7 E(): 5.9
Smith-Waterman score: 57; 36.364% identity (66.667% similar) in 33 aa overlap
(5-36:464-495)
```

```
gi|131 MSCGGRLLCLVLFALLLASAVVASESKDPELKQCKHQCKAQQQISKEQKEACIQACKEYIR
      10      20      30      40      50      60
```

```
gi|131 QKHQGEHGRGGDILEEEVWNRKSPIERLRECSRGCEQQHGEQREECLRRCEEEYQREKG
      70      80      90      100      110      120
```

```
gi|131 RQDDDNPTDPEKQYQQCRLQCRRGEGGGFSREHCERRREEKYREQQGREGGRGEMYEGR
      130      140      150      160      170      180
```

```
gi|131 EREEEQEEQGRGRIPYVFEDQHFITGFRTHGRMRVLQKFTDRSELLRGIENYRVAILEA
      190      200      210      220      230      240
```

```
gi|131 EPQTFIVPNHWAESVVFVAKGRGTISLVRQDRRESLNKQGDILKINAGTTAYLINRDN
      250      260      270      280      290      300
```

```
gi|131 NERLVLAKLLQPVSTPGFELFFGAGGENPESFFKSFSDEILEAAFNTRRDLQRIFGQQ
      310      320      330      340      350      360
```

```
gi|131 RQGVIVKASEEQVRAMSRHEEGGIWPFGGESKGTINIYQQRPTHSNQYQLHEVDASQYR
      370      380      390      400      410      420
```

```

      10      20
CV127_ DVWGSCP-MPIKEDSTRSELR
      : : : : : : : : : : : : : : : : : : : : : :
gi|131 QLRDLTLTVSLANITQGAMTAPHYNSKATKIALVVDGEGYFEMACPHMSRSRGSYQGETR
      430      440      450      460      470      480
```

```

      30      40
CV127_ MGRPNKARRATQSNEGRQTNTKYPSNKDL
      : : : : : : : : : : : : : : : : : : : : : :
gi|131 -GRPSYQVRVASRLTRGTVVIIIPAGHPFVAVASSNQLQVLCFEVNANNNEKFPLAGRRNV
      490      500      510      520      530
```

```

gi|131 MNQLEREAKELAFGMPAREVEEVSRSQEEFFFKGPRQQQGRADA
      540      550      560      570      580

>>gi|20387029 gi|20387029|emb|CAC84593.2| troposmyosin [ (243 aa)
  initn: 35 initl: 35 opt: 51 Z-score: 86.8 bits: 21.2 E(): 6.9
Smith-Waterman score: 51; 29.268% identity (58.537% similar) in 41 aa overlap
(10-49:95-135)

gi|203 MEAIKKKMQAMKLEKDNVDRARQNEQEAKDANLRAEKAEEEEARSLQKKIQTIENELDQT
      10      20      30      40      50      60

CV127_
      10      20      30
      DVWGSCPMPIKEDSTRSELRMGRPN-KARRATQSN
      ...: ...: ...: .: .: .: .:

gi|203 QEQLMQVNAKLEEKDKALQNAESEVAALNRRIQLLEEDLERSEERLATATQKLAEASQAA
      70      80      90      100      110      120

CV127_
      40
      EGRQTNTKYP SNKDL
      .: .: .: .: .:

gi|203 DESERM RKVLEN RSLADEERMDALENQLKEARFLAEEADKKYDEVARKLAMVEADLERAE
      130      140      150      160      170      180

gi|203 ERAESGESKIVELEEEELRVVGNLKSLEVSEEKANQREEEYKQIKTLTTRLKEAEARAC
      190      200      210      220      230      240

gi|203 WLL

>>gi|6094504 gi|6094504|sp|Q25456.1|TPM_METEN RecName: F (274 aa)
  initn: 35 initl: 35 opt: 51 Z-score: 85.7 bits: 21.2 E(): 8
Smith-Waterman score: 51; 29.268% identity (58.537% similar) in 41 aa overlap
(10-49:85-125)

gi|609 MKLEKDNAMD RADTLEQQNKEANNRAEKSEEEVHNLQKRMQQLENDLDQVQESLLKANNQ
      10      20      30      40      50      60

CV127_
      10      20      30      40
      DVWGSCPMPIKEDSTRSELRMGRPN-KARRATQSN EGRQTNTKYP
      ...: ...: ...: .: .: .: .: .: .:

gi|609 LVEKDKALSNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAADSEERM RKVL
      70      80      90      100      110      120

CV127_
      SNKDL
      ...:

gi|609 ENRSL SDEERMDALENQLKEARFLAEEADRKYDEVARKLAMVEADLERAEERAETGESKI
      130      140      150      160      170      180

gi|609 VELEEEELRVVGNLKSLEVSEEKANQREEAYKEQIKTLTNKLKAAEARAFAERSVQKLQ
      190      200      210      220      230      240

gi|609 KEVDRLEDELVNEKEKYKSITDELDTQTFSEL SGY
      250      260      270

>>gi|14285797 gi|14285797|sp|O61379.1|TPM_PANST RecName: (274 aa)
  initn: 35 initl: 35 opt: 51 Z-score: 85.7 bits: 21.2 E(): 8
Smith-Waterman score: 51; 29.268% identity (58.537% similar) in 41 aa overlap
(10-49:85-125)

gi|142 MKLEKDNAMD RADTLEQQNKEANIRAEKAEEEEVHNLQKRMQQLENDLDQVQESLLKANTQ
      10      20      30      40      50      60

```

```

              10      20      30      40
CV127_      DVWGSCPMPIKEDSTRSELRMGRPN-KARRATQSNRQTNTKYP
              . . . . . : : : : : : : : : : : : : :
gi|142 LEEKDKALSNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAADSESRMRKVL
              70      80      90      100     110     120

```

```

CV127_ SNKDL
      . . . .
gi|142 ENRSLSDDEERMDALENQLKEARFLAEEADRKYDEVARKLAMVEADLERAEERAETGESKF
              130     140     150     160     170     180

gi|142 VELEEEELRVVGNLKSLEVSEEKANQREEAYKEQIKTLTNKLKAAEAREFAERSVQKLQ
              190     200     210     220     230     240

```

```

gi|142 KEVDRLEDELVNEKEKYKSITDELDTQTFSELGY
              250     260     270

```

>>gi|125995159 gi|125995159|dbj|BAF47263.1| tropomyosin (284 aa)  
 initn: 35 initl: 35 opt: 51 Z-score: 85.4 bits: 21.1 E(): 8.3  
 Smith-Waterman score: 51; 29.268% identity (58.537% similar) in 41 aa overlap  
 (10-49:95-135)

```

gi|125 MDAIKKKMQAMKLEKDNAMDRADTLEQQNKEANNRAEKSEEEVHNLQKRMQQLENDLDQV
              10      20      30      40      50      60

```

```

              10      20      30
CV127_      DVWGSCPMPIKEDSTRSELRMGRPN-KARRATQSN
              . . . . . : : : : : : : : : :
gi|125 QESLLKANIQLVEKDKALSNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA
              70      80      90      100     110     120

```

```

              40
CV127_ EGRQTNTKYPSNKDL
      . . . : : :
gi|125 DESERMRKVLNENRSLSDDEERMDALENQLKEARFLAEEADRKYDEVARKLAMVEADLERAE
              130     140     150     160     170     180

gi|125 ERAETGESKIVELEEEELRVVGNLKSLEVSEEKANQREEAYKEQIKTLTNKLKAAEAREAE
              190     200     210     220     230     240

gi|125 FAERSVQKLQKEVDRLEDELVNEKEKYKSITDELDTQTFSELGY
              250     260     270     280

```

>>gi|2660868 gi|2660868|gb|AAC48288.1| fast tropomyosin (284 aa)  
 initn: 35 initl: 35 opt: 51 Z-score: 85.4 bits: 21.1 E(): 8.3  
 Smith-Waterman score: 51; 29.268% identity (58.537% similar) in 41 aa overlap  
 (10-49:95-135)

```

gi|266 MDAIKKKMQAMKLEKDNAMDRADTLEQQNKEANIRAEKSEEEVHNLQKRMQQLENDLDQV
              10      20      30      40      50      60

```

```

              10      20      30
CV127_      DVWGSCPMPIKEDSTRSELRMGRPN-KARRATQSN
              . . . . . : : : : : : : : :
gi|266 QESLLKANTQLEEKDKALSNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA
              70      80      90      100     110     120

```

```

              40
CV127_ EGRQTNTKYPSNKDL
      . . . : : :

```



gi|266 DESERMRKVLENRSLSDERMDALENQLKEARFLAEEADRKYDEVARKLAMVEADLERAE  
130 140 150 160 170 180

gi|266 ERAETGESKIVELEEEELRVVGNNLKSLEVSEEKANQREEAYKEQIKTLANKLKAAEAREAE  
190 200 210 220 230 240

gi|266 FAERSVQKLQKEVDRLEDELVNEKEKYKSITDELDQTFSELSGY  
250 260 270 280

>>gi|125995167 gi|125995167|dbj|BAF47267.1| tropomyosin (284 aa)  
initn: 35 initl: 35 opt: 51 Z-score: 85.4 bits: 21.1 E(): 8.3  
Smith-Waterman score: 51; 29.268% identity (58.537% similar) in 41 aa overlap  
(10-49:95-135)

gi|125 MDAIKKKMQAMKLEKDNAMDKADTLEQQNKEANLRAEKTEEEIRANQKKSQVLVENELDHA  
10 20 30 40 50 60

CV127\_ DVWGSCPMPKEDSTRSELRMGRPN-KARRATQSN  
10 20 30  
...: ...: ...: ...: ...: ...:  
gi|125 QEQLSAATHKLVEKEKAFANAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA  
70 80 90 100 110 120

40  
CV127\_ EGRQTNTKYPSNKDL  
...: ...: ...: ...: ...: ...:  
gi|125 DESERMRKVLENRSLSDERMDALENQLKEARFLAEEADRKYDEVARKLAMVEADLERAE  
130 140 150 160 170 180

gi|125 ERAESGESKIVELEEEELRVVGNNLKSLEVSEEKANQREETYKEQIKTLANKLKAAEAREAE  
190 200 210 220 230 240

gi|125 FAERSVQKLQKEVDRLEDELVNEKEKYKNIADMDQAFSELSGF  
250 260 270 280

>>gi|148615631 gi|148615631|gb|ABQ96644.1| tropomyosin [ (284 aa)  
initn: 35 initl: 35 opt: 51 Z-score: 85.4 bits: 21.1 E(): 8.3  
Smith-Waterman score: 51; 29.268% identity (58.537% similar) in 41 aa overlap  
(10-49:95-135)

gi|148 MEAIKKKMQAMKLEKDNAIDRADTLEQQNKEANNRAEKSEEEVHNLQKRMQQLGNDLDQV  
10 20 30 40 50 60

CV127\_ DVWGSCPMPKEDSTRSELRMGRPN-KARRATQSN  
10 20 30  
...: ...: ...: ...: ...: ...:  
gi|148 QESLLKANIQLVEKDRLSNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA  
70 80 90 100 110 120

40  
CV127\_ EGRQTNTKYPSNKDL  
...: ...: ...: ...: ...: ...:  
gi|148 DESERMRKVLENRSLSDERMDALENQLKEARFLAEEADRKYDEVARKLAMVEADLERAE  
130 140 150 160 170 180

gi|148 ERAETGESKIVELEEEELRVVGNNLKSLEVSEEKANQREEAYKEQIKTLTNKLKAAEAREAE  
190 200 210 220 230 240

gi|148 FAERSVQKLQKEVDRLEDELVNEKEKYKSITDELDQTFSELSTGY  
250 260 270 280

>>gi|125995163 gi|125995163|dbj|BAF47265.1| tropomyosin (284 aa)

initn: 35 initl: 35 opt: 51 Z-score: 85.4 bits: 21.1 E(): 8.3  
Smith-Waterman score: 51; 29.268% identity (58.537% similar) in 41 aa overlap  
(10-49:95-135)

gi|125 MDAIKKKMQAMKLEKDNAMDKADTLEQQNKEANNRAEKAEVEVHGLQKRMQQLENDLDQV  
10 20 30 40 50 60

CV127\_ DVWGSCPMPIKEDSTRSELRMGRPN-KARRATQSN  
10 20 30  
...: ...: ...: ...: ...: ...:

gi|125 QESLLKANTQLEDKDKALSNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA  
70 80 90 100 110 120

40  
CV127\_ EGRQTNTKYPSNKDL  
...: ...: ...: ...: ...: ...:  
gi|125 DESERMRKVLENRSLSDERMDALENQLKEARFLAEADRKYDEVARKLAMVEADLERAE  
130 140 150 160 170 180

gi|125 ERAETGESKIVELEEEELRVVGNNLKSLEVSEEKANQREEAYKEQIKTLANKLKAAEARE  
190 200 210 220 230 240

gi|125 FAERSVQKLQKEVDRLEDELVNEKEKYKSITDELDTFSELSGY  
250 260 270 280

>>gi|125995157 gi|125995157|dbj|BAF47262.1| tropomyosin (284 aa)  
initn: 35 initl: 35 opt: 51 Z-score: 85.4 bits: 21.1 E(): 8.3  
Smith-Waterman score: 51; 29.268% identity (58.537% similar) in 41 aa overlap  
(10-49:95-135)

gi|125 MDAIKKKMQAMKLEKDNAMDRADTLEQQNKEANNRAEKSEEEVHNLQKRMQQLENDLDQV  
10 20 30 40 50 60

CV127\_ DVWGSCPMPIKEDSTRSELRMGRPN-KARRATQSN  
10 20 30  
...: ...: ...: ...: ...: ...:

gi|125 QESLLKANIQLVEKDKALSNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA  
70 80 90 100 110 120

40  
CV127\_ EGRQTNTKYPSNKDL  
...: ...: ...: ...: ...: ...:  
gi|125 DESERMRKVLENRSLSDERMDALENQLKEARFLAEADRKYDEVARKLAMVEADLERAE  
130 140 150 160 170 180

gi|125 ERAETGESKIVELEEEELRVVGNNLKSLEVSEEKANQREEAYKEQIKTLTNKLKAAEARE  
190 200 210 220 230 240

gi|125 FAERSVQKLQKEVDRLEDELVNEKEKYKSITDELDTFSELSGY  
250 260 270 280

>>gi|73532979 gi|73532979|gb|AAZ76743.1| Pen a 1 allerge (284 aa)  
initn: 35 initl: 35 opt: 51 Z-score: 85.4 bits: 21.1 E(): 8.3  
Smith-Waterman score: 51; 29.268% identity (58.537% similar) in 41 aa overlap  
(10-49:95-135)

gi|735 MDAIKKKMQAMKLEKDNAMDRADTLEQQNKEANNRAEKSEEEVHNLQKRMQQLENDLDQV  
10 20 30 40 50 60

CV127\_ DVWGSCPMPIKEDSTRSELRMGRPN-KARRATQSN  
10 20 30  
...: ...: ...: ...: ...: ...:

gi|735 QESLLKANIQLVEKDKALSNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA  
70 80 90 100 110 120

40  
CV127\_ EGRQTNTKYPSNKDL  
. . . : : . . .

gi|735 DESERMKVLNRSLSDEERMDALENQLKEARFLAEEADRKYDEVARKLAMVEADLERAE  
130 140 150 160 170 180

gi|735 ERAETGESKIVELEEEELRVVGNNLKSLEVSEEKANQREEAYKEQIKTLTNKLKAAEARAE  
190 200 210 220 230 240

gi|735 FAERSVQKLQKEVDRLEDELVNEKEKYKSITDELDQTFSELGY  
250 260 270 280

>>gi|14285796 gi|14285796|sp|O44119.1|TPM\_HOMAM Tropomyo (284 aa)  
initn: 35 initl: 35 opt: 51 Z-score: 85.4 bits: 21.1 E(): 8.3  
Smith-Waterman score: 51; 29.268% identity (58.537% similar) in 41 aa overlap  
(10-49:95-135)

gi|142 MDAIKKKMQAMKLEKDNAMDRADTLEQQNKEANIRAETKEEIRITHKKMQQVENELDQV  
10 20 30 40 50 60

10 20 30  
CV127\_ DVWGSCPMPIKEDSTRSELRMGRPN-KARRATQSN  
. . . : : . . . : . . .

gi|142 QEQLSLANTKLEEKEKALQNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA  
70 80 90 100 110 120

40  
CV127\_ EGRQTNTKYPSNKDL  
. . . : : . . .

gi|142 DESERMKVLNRSLSDEERMDALENQLKEARFLAEEADRKYDEVARKLAMVEADLERAE  
130 140 150 160 170 180

gi|142 ERAETGESKIVELEEEELRVVGNNLKSLEVSEEKANQREEAYKEQIKTLANKLKAAEARAE  
190 200 210 220 230 240

gi|142 FAERSVQKLQKEVDRLEDELVNEKEKYKSITDELDQTFSELGY  
250 260 270 280

>>gi|125995165 gi|125995165|dbj|BAF47266.1| tropomyosin (284 aa)  
initn: 35 initl: 35 opt: 51 Z-score: 85.4 bits: 21.1 E(): 8.3  
Smith-Waterman score: 51; 29.268% identity (58.537% similar) in 41 aa overlap  
(10-49:95-135)

gi|125 MDAIKKKMQAMKLEKDNAMDKADTLEQQNKEANNRAETKEEIRLTQKKMQQVENELDVA  
10 20 30 40 50 60

10 20 30  
CV127\_ DVWGSCPMPIKEDSTRSELRMGRPN-KARRATQSN  
. . . : : . . . : . . .

gi|125 QEQLSLANTKLEEKEKALQNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA  
70 80 90 100 110 120

40  
CV127\_ EGRQTNTKYPSNKDL  
. . . : : . . .

gi|125 DESERMKVLNRSLSDEERMDALENQLKEARFLAEEADRKYDEVARKLAMVEADLERAE  
130 140 150 160 170 180

gi|125 ERAETGESKIVELEEEELRVVGNNLKSLEVSEEKANQREEAYKEQIKTLANKLKAAEARAE

```

                190          200          210          220          230          240
gi|125 FAERSVQKLQKEVDRLEDELVNEKEKYKNIADMDQAFSELSGF
                250          260          270          280

>>gi|162286975 gi|162286975|dbj|BAF95206.1| tropomyosin (284 aa)
  initn: 35 initl: 35 opt: 51 Z-score: 85.4 bits: 21.1 E(): 8.3
Smith-Waterman score: 51; 29.268% identity (58.537% similar) in 41 aa overlap
(10-49:95-135)

gi|162 MDAIKKKMQAMKLEKDNAMDRADTLEQQNKEANNRAEKAEEVHNLQKRMQQLENDLDQV
                10          20          30          40          50          60

CV127_
                10          20          30
                DVWGSCPMPIKEDSTRSELRMGRPN-KARRATQSN
                . . . . . : : : . . : . : . .
gi|162 QEQLLKANTQLEEKDKALQNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA
                70          80          90          100          110          120

CV127_
                40
                EGRQTNTKYPSNKDL
                . . . : : . .
gi|162 DESERMRKVLENRSLSDERMDALENQLKEARFLAEEADRKYDEVARKLAMVEADLERAE
                130          140          150          160          170          180

gi|162 ERAETGESKIVELEEEELRVVGNNLKSLEVSEEKANQREEAYKEQIKQLTNKLKAAEARAE
                190          200          210          220          230          240

gi|162 FAERSVQKLQKEVDRLEDELVNEKEKYKSITDELDTQTFSELSGY
                250          260          270          280

>>gi|156712754 gi|156712754|dbj|BAF76431.1| tropomyosin (284 aa)
  initn: 35 initl: 35 opt: 51 Z-score: 85.4 bits: 21.1 E(): 8.3
Smith-Waterman score: 51; 29.268% identity (58.537% similar) in 41 aa overlap
(10-49:95-135)

gi|156 MDAIKKKMQAMKMEKDSAMDRADALEAQNKETNARAEKADDEVHNLQKRLQTLLENDLDQV
                10          20          30          40          50          60

CV127_
                10          20          30
                DVWGSCPMPIKEDSTRSELRMGRPN-KARRATQSN
                . . . . . : : : . . : . : . .
gi|156 SEALLKANTQLVEKDKALQNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA
                70          80          90          100          110          120

CV127_
                40
                EGRQTNTKYPSNKDL
                . . . : : . .
gi|156 DESERMRKVLENRSLLEERMDALENQLKEARLLAEEADRKYDEVARKLAMVEADLERAE
                130          140          150          160          170          180

gi|156 ERAEAGENKIVELEEEELRVVGNNLKSLEVSEEKANQREEAYKEQIKHLTHKLKTAEARAE
                190          200          210          220          230          240

gi|156 FAERSVQKLQKEVDRLEDELVNEKEKYKGITDELDTQTFNELSGY
                250          260          270          280

>>gi|125995161 gi|125995161|dbj|BAF47264.1| tropomyosin (284 aa)
  initn: 35 initl: 35 opt: 51 Z-score: 85.4 bits: 21.1 E(): 8.3
Smith-Waterman score: 51; 29.268% identity (58.537% similar) in 41 aa overlap
(10-49:95-135)

```

gi | 125 MDAIKKKMQAMKLEKDNAMDRADTLEQQNKEANNRAEKSEEEVFGLQKKLQQLENDLDSV  
10 20 30 40 50 60

CV127\_ DVWGSCPMPIKEDSTRSELRMGRPN-KARRATQSN  
          10      20      30  
            .: :.: .: .: .:

gi|125 QEALLKANQHLEEKDKALSNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA  
70 80 90 100 110 120

CV127\_ EGRQTNTKYPSNKDL  
          . . . : : : :

gi|125 DESERMKRVLENRSLDEERMDALENQLKEARFLAEEDRKYDEVARKLAMVEADLERAE  
130 140 150 160 170 180gi | 125 ERAETGESKIVELEEEELRVVGNNLKSLEVSEEKANQREEAYKEQIKTLTNKLKAAEARAE  
190 200 210 220 230 240

gi|125 FAERSVQKLQKEVDRLLEDELVNEKEKYKSITDELDQTFSELSGY  
250 260 270 280

```
>>gi|14423757 gi|14423757|sp|O04701.1|MPAC1_CYNDA RecNam (246 aa)
  initn: 36 initl: 36 opt: 50 Z-score: 85.1 bits: 20.9 E(): 8.6
Smith-Waterman score: 50; 33.333% identity (70.370% similar) in 27 aa overlap
(20-45:118-144)
```

gi|144 AIGDKPGPNITATYGSKWLEARATFYGSNPRGAAPDDHGGACGYKDVDKPPFDGMTACGN  
10 20 30 40 50 60

CV127\_ 10 20  
DVWGSCPMPPIKEDSTRSELRMG  
:

gi | 144 EPIFKDGLGCRACYEIKCKEPVECSGEPVLVKITDKNYEHIAAYHFDLSGKAFGAMAKKG  
70 80 90 100 110 120

CV127\_ RPNKARRATQ-SNEGRQTNTKYPSNKDL

gi|144 QEDKLRKAGELTLQFRRVKCKYPSGTKITFHIEKGSNDHYLALLVKYAAGDGNIVAVDIK  
130 140 150 160 170 180gi | 144 PRDSDEFIPMKSSWGAIWRIDPKKPLKGPFSIRLTSEGG AHLVQDDVIPANWKPD TVYTS  
190 200 210 220 230 240

gi | 144 KLQFGA

```
>>gi|5931948 gi|5931948|gb|AAD56719.1|AF091737_1 allerge (160 aa)
  initn: 36 initl: 36 opt: 47 Z-score: 84.4 bits: 20.1 E(): 9.4
Smith-Waterman score: 47; 31.250% identity (62.500% similar) in 32 aa overlap
(16-47:51-80)
```

CV127 10 20  
DVWGSCPMPRIKEDSTRSELRMGRPN

```
gi|593 MMVKLSILVALLGALLVVASATRWDPDRGSRGRWDAPSRGDDQCQRQLQRANLRPCEEH
      10      20      30      40      50      60
```

CV127\_ KARRATQSNEGRQTNTKYPSNKDL

gi|593 MRRRVEQEQE--QEODEYPYSRRGSRGRQPGESDENQEQRCNELNRFQNNQRCMCQALO

```

70          80          90          100          110
gi|593 QILQNQSFVWPAGQEPVASDGEQAQELAPLRVQVTKPLRPL
    120          130          140          150          160

>>gi|25991543 gi|25991543|gb|AAN76862.1|AF453947_1 aller (457 aa)
  initn: 38 initl: 38 opt: 53 Z-score: 84.1 bits: 21.6 E(): 9.8
Smith-Waterman score: 53; 26.829% identity (68.293% similar) in 41 aa overlap
(10-48:234-270)

gi|259 LSVCFLLILFHGCLASRQEWQQQDECQIDRLDALEPDNRVEYEAGTVEAWDPNHEQFRCAG
    10          20          30          40          50          60

gi|259 VALVRHTIQPNGLLLPQYSNAPQLIYVVGEGMTGISYPGCPETYQAPQQGRQQGQSGRF
    70          80          90          100          110          120

gi|259 QDRHQKIRRFRRGDIIAIPAGVAHWCYNEGNSPVVTVTLDDVSNSQNQLDRTPRKFHLAG
    130          140          150          160          170          180

CV127_
                                10
                                DVWGSCPMPIKEDSTR
                                .:.:
gi|259 NPKDVFQQQQQHQSRRNLFSGFDTLLAEAFQVDERLIKQLKSEDNRGGIVKVKDD---
    190          200          210          220          230

    20          30          40
CV127_ SELRMGRPNKAR--RATQSNRQNTNTKYPSNKDL
    :.:. :.:.:. :.:.:.: . . . . .
gi|259 -ELRVIRPSRSQSERGSESEEESEDEKRRWGQRDNGIEETICTMRLKENINDPARADIYT
    240          250          260          270          280          290

gi|259 PEVGRLTTLNSLNLPLKWLQLSVEKGVLYKNALVLPHWNLNSHSIIYGCKGKGQVQVVD
    300          310          320          330          340          350

gi|259 NFGNRVFDGEVREGQMLVVPQNFAVVKRAREERFEWISFKTNDRAMTSPLAGRTSVLGGM
    360          370          380          390          400          410

gi|259 PEEVLANAFQISREDARKIKFNNQQTTLTSGESSHHMRDDA
    420          430          440          450

```

49 residues in 1 query sequences  
 307888 residues in 1386 library sequences  
 Scomplib [34.26]  
 start: Wed Mar 11 05:58:41 2009 done: Wed Mar 11 05:58:41 2009  
 Total Scan time: 0.090 Total Display time: 0.050

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

## CV127\_851\_880\_frame1

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

Please cite:

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

CV127\_851\_880\_frame1, 10 aa

vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

```

307888 residues in 1386 sequences
Expectation_n fit: rho(ln(x))= 2.1237+/-0.00246; mu= 6.3300+/- 0.131
mean_var=15.7866+/- 4.315, 0's: 91 Z-trim: 91 B-trim: 0 in 0/36
Lambda= 0.322798

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 20, open/ext: -10/-2, width: 16
The best scores are:
gi|33300921 gi|33300921|sp|P81943_ ( 22) 31 17.0 2.3 0.625 1.000 8
gi|118216 gi|118216|sp|P18153.1|D7_ ( 321) 33 19.2 7.4 0.545 0.909 11
gi|205525919 gi|205525919|sp|P18153 ( 321) 33 19.2 7.4 0.545 0.909 11
gi|83308249 gi|83308249|emb|CAJ4356 ( 254) 32 18.6 8.7 0.571 1.000 7

>>>CV127_851_880_frame1, 10 aa vs
/n/na4/bioinfo/refsets/PUB_AllergenOnlineJan2009_V9.fasta library

>>gi|33300921 gi|33300921|sp|P81943_1 [Segment 1 of 4] (22 aa)
initn: 30 init1: 30 opt: 31 Z-score: 95.6 bits: 17.0 E(): 2.3
Smith-Waterman score: 31; 62.500% identity (100.000% similar) in 8 aa overlap
(3-10:12-19)

          10
CV127_      NGNKISKSLY
          .....:
gi|333 LPNPSGFVTCLSSISKSVYTPA
          10      20

>>gi|118216 gi|118216|sp|P18153.1|D7_AEDAE D7 protein pr (321 aa)
initn: 24 init1: 24 opt: 33 Z-score: 86.3 bits: 19.2 E(): 7.4
Smith-Waterman score: 33; 54.545% identity (90.909% similar) in 11 aa overlap
(1-10:147-157)

gi|118 MKLLLLLLAIFTTFSVVASMGFPDPEEMLFIFTRCMEDNLEDGANRLPMLAKWKEWINEPV
          10      20      30      40      50      60

gi|118 DSPATQCFGKCVLVRTGLYDPVAQKFDASVIQEQFKAYPSLGEKSKVEAYANAVKQLPST
          70      80      90      100      110      120

          10
CV127_      NGNK-ISKSLY
          ....: .....:
gi|118 NNDCAAVFKAYDPVHKAHKDTSKNLFHGNKELTKGLYEKLGKDIRKKQSYFEFCENKYY
          130      140      150      160      170      180

gi|118 PAGSDKRQQLCQIRQYTVLDDALFKEHTDCVMKGIRYITKDNQLDVEEVKRDVFLVKNKDT
          190      200      210      220      230      240

gi|118 KALEKVLNDCKSKEPSNAKEKSWHYKCLVESSVKDDFKEAFDYREVRSQIYAFNLPKKQ
          250      260      270      280      290      300

gi|118 AYSKPAVQSQVMEIDGKQCPQ
          310      320

>>gi|205525919 gi|205525919|sp|P18153.2|ALL2_AEDAE RecNa (321 aa)
initn: 24 init1: 24 opt: 33 Z-score: 86.3 bits: 19.2 E(): 7.4
Smith-Waterman score: 33; 54.545% identity (90.909% similar) in 11 aa overlap
(1-10:147-157)

gi|205 MKLPLLLAIVTTFSSVASTGFPDPEEMLFITFTRCMEDNLEDGPNRLPMLAKWKEWINEPV
          10      20      30      40      50      60

```

gi|205 DSPATQCFGKCVLVRTGLYDPVAQKFDASVIOEQFKAYPSLGEKSKVEAYANAVQQLPST  
70 80 90 100 110 120

CV127\_ NGNK-ISKSLY  
10  
::: :::  
gi|205 NNDCAAVFKAYDPVHKAHKDTSKNLFHGNKELTKGLYEKLGKDIRQKKQSYFEFCENKYY  
130 140 150 160 170 180

gi|205 PAGSDKRQQLCKIRQYTVLDDALFKEHTDCVMKGIRYITKNNELDAEEVKRDFMQVNKDT  
190 200 210 220 230 240

gi|205 KALEKVLNDCKSKEPSNAGEKSWHYKCLVESSVKDDFKEAFDYREVRSQIYAFNLPKKQ  
250 260 270 280 290 300

gi|205 VYSKPAVQSQVMEIDGKQCPQ  
310 320

>>gi|83308249 gi|83308249|emb|CAJ43561.1| staphylococcal (254 aa)  
initn: 32 initl: 32 opt: 32 Z-score: 85.0 bits: 18.6 E(): 8.7  
Smith-Waterman score: 32; 57.143% identity (100.000% similar) in 7 aa overlap  
(1-7:137-143)

gi|833 IFVLILVISTPNVLAESQDPKPKDELHKASKFTGLMENMKVLYDDNHVSAINVKSIDQFL  
10 20 30 40 50 60

gi|833 YFDLIYSIKDTKLGNYDNVRVEFKNKDLADKYKDKYVDVFGANYYYQCYFSKKTNDINSH  
70 80 90 100 110 120

CV127\_ NGNKISKSLY  
10  
:::  
gi|833 QTDKRKTCMYGGVTEHNGNQLDKYRSITVRVFEDGKNLLSFDVQTNKKKVTAQELDYLTR  
130 140 150 160 170 180

gi|833 HYLKVNKKLYEFNNSPYETGYIKFIENENSFWYDMMPAPGDKFDQSKYLMMYNDNKMVDS  
190 200 210 220 230 240

gi|833 KDVKIEVYLTTKK  
250

10 residues in 1 query sequences  
307888 residues in 1386 library sequences  
Scomplib [34.26]  
start: Wed Mar 11 06:22:11 2009 done: Wed Mar 11 06:22:11 2009  
Total Scan time: 0.050 Total Display time: 0.010

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

## CV127\_862\_918\_frame0

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007  
Please cite:  
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448



```

CV127_862_918_frame0, 19 aa
vs /n/na4/bioinfo/refsets/PUB_AllergenOnlineJan2009_V9.fasta library

307888 residues in 1386 sequences
Expectation_n fit: rho(ln(x))= 3.6040+/-0.00325; mu= 3.3033+/- 0.169
mean_var=27.3285+/- 7.417, 0's: 35 Z-trim: 36 B-trim: 0 in 0/42
Lambda= 0.245339

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 20, open/ext: -10/-2, width: 16
The best scores are:
      opt bits E(1386) %_id %_sim alen
gi|157418806 gi|157418806|gb|ABV551 ( 147) 60 26.0 0.058 0.333 0.722 18
gi|114841683 gi|114841683|dbj|BAF32 ( 419) 49 22.3 2.2 0.500 0.700 20
gi|47606004 gi|47606004|sp|Q7M1E7.1 ( 514) 49 22.3 2.6 0.500 0.700 20
gi|19338630 gi|19338630|gb|AAL86739 ( 448) 47 21.6 3.8 0.412 0.706 17
gi|585290 gi|585290|sp|P32936.2|IAA ( 149) 43 20.0 3.8 0.455 0.818 11
gi|2493414 gi|2493414|sp|Q28050.1|S ( 101) 41 19.2 4.4 0.444 0.611 18
gi|21711 gi|21711|emb|CAA42453.1| C ( 143) 42 19.6 4.7 0.833 1.000 6
gi|256427 gi|256427|gb|AAB23463.1| ( 439) 45 20.9 6.1 0.467 0.867 15
gi|115502167 gi|115502167|sp|Q1ZYQ8 ( 270) 42 19.7 8.3 0.312 0.625 16

>>>CV127_862_918_frame0, 19 aa vs
/n/na4/bioinfo/refsets/PUB_AllergenOnlineJan2009_V9.fasta library

>>gi|157418806 gi|157418806|gb|ABV55106.1| Ani s 9 aller (147 aa)
initn: 39 initl: 39 opt: 60 Z-score: 124.1 bits: 26.0 E(): 0.058
Smith-Waterman score: 60; 33.333% identity (72.222% similar) in 18 aa overlap
(2-19:8-25)

      10
CV127_ NIKIFILKLNKRGPPYLSP
      . . . . . : : :
gi|157 MKLCILAVAVFVAVSAQGPPPLPPFVANAPPAVQAEFRQLANGAPDKTEAEIEAQIEQW
      10 20 30 40 50 60

gi|157 VASKGGAVQAEFNKFKQMLEQGKARAEAAHQASLTRLSPAAKAADARLSAIASNRALKVG
      70 80 90 100 110 120

gi|157 EKQRQLAAAFQALDPAVKAELQKEMQG
      130 140

>>gi|114841683 gi|114841683|dbj|BAF32143.1| pollen aller (419 aa)
initn: 25 initl: 25 opt: 49 Z-score: 95.8 bits: 22.3 E(): 2.2
Smith-Waterman score: 49; 50.000% identity (70.000% similar) in 20 aa overlap
(1-18:68-87)

gi|114 LDSDIEQYLRSLNRSLKKLVHSRHDAAATVFNVEQYGAVGDKHDSFEAFETAWNAACKKAS
      10 20 30 40 50 60

      10
CV127_ NIKIFILKLNKRGPP--PYLSP
      : . . . . : : : :
gi|114 AVLVPANKKFFVNNLVFRGPCQPHLSFKVDGTIVAQPDPARWKNKSKIWLQFAQLTDFNL
      70 80 90 100 110 120

gi|114 MGTGVIDGQGQWQWAGQCKVNVNRTVCNDRNRPTAIKIDYSKSVTVKELTLMNSPEFHLV
      130 140 150 160 170 180

gi|114 FGECEGVKIQGLKIKAPRDSPTNDGIDIFASKRFHIEKCVIGTGDDCIAIGTGSSNITIK
      190 200 210 220 230 240

```

gi|114 DLICGPGHGISIGSLGRDNSRAEVSHVHVNRKAFIDTQNGLRIKTWQGGSGLASIYITYEN  
250 260 270 280 290 300gi | 114 VEMINSENPIILINQFYCTASACQNQRSASVQIQGVTYKNIHGTSATAAAIQLMCSDSVPC  
310 320 330 340 350 360gi|114 TGIQLSNVSLKLTSGKPASCVDKNARGFYSGRLIPTCKNLRPGPSPKEFELQQQPTTVM  
370 380 390 400 410

```
>>gi|47606004 gi|47606004|sp|Q7M1E7.1|PGLR2_CHAOB RecNam (514 aa)
  initn: 25 initl: 25 opt: 49 Z-score: 94.4 bits: 22.3 E(): 2.6
Smith-Waterman score: 49; 50.000% identity (70.000% similar) in 20 aa overlap
(1-18:98-117)
```

gi|476 MGMKFMAAVAF<sup>10</sup>LALQLIVMAAAEDQSAQIM<sup>20</sup>LDSDIEQYLR<sup>30</sup>SNRSLKKLVHSR<sup>40</sup>HDAA<sup>50</sup>TVFN<sup>60</sup>

CV127\_ 10  
NIKIFILKLNKRG--PYLSP  
: : : : : : : : : :  
gi|476 VEQYGAVGDGKHDSTEAFFATTWNAACKKASAVLLVPANKKFFVNNLVFRGPCQPPLSFKV  
70 80 90 100 110 120

gi|476 DGTIVAQPDPARWKNSKIWLQFAQLTDFNLMGTGVIDGQGQWWAGQCKVVNGRTVCNDR  
130 140 150 160 170 180gi|476 NRPTAIKIDYSKSVTVKELTLMNSPEFHLVFGEGVGIQGLKIKAPRDSPTDGDIDIFA  
190 200 210 220 230 240gi|476 SKRFHIEKCVIGTGDDCIAIGTGSSNITIKDLICGPGHGISIGSLGRDNSRAEVSHVHVN  
250 260 270 280 290 300gi|476 RAKFIDTQNGLRIKTWQGGSGLASYYITYENVEMINSENPIILINQFYCTSSASACQNQSAV  
310 320 330 340 350 360gi|476 QIQGVTYKNIHGTSATAAAIQLMCSDSVPCTGIQLSNVSLKLTSGKPA SCVDKNARGFYS  
370 380 390 400 410 420gi|476 GRLIPTCKNLRPGSPKEFELQQQPTTVMDENKGACAKGDSTCISLSSPPNCKNKCKGC  
430 440 450 460 470 480

gi|476 QPCKPKLIIVHPNKPQDYYPQKWVCSCHNKIYNP  
490 500 510

```
>>gi|19338630 gi|19338630|gb|AAL86739.1|AF441864_1 48-kD (448 aa)
  initn: 47 initl: 47 opt: 47 Z-score: 91.5 bits: 21.6 E(): 3.8
Smith-Waterman score: 47; 41.176% identity (70.588% similar) in 17 aa overlap
(3-19:238-254)
```

gi|193 MLPKEDPELKCKHKCRDERQFDEQQRRDGKQICEEKARERQQEEGNSSEESYGKEQEEN  
10 20 30 40 50 60gi|193 PYVFQDEHFESRVKTEEGRVQVLENFTKRSRLLSGIENFRLAILEANPHTFISPAHFDAE  
70 80 90 100 110 120gi|193 LVL FVAKGRATITMVREEKRESFNVEHGDIIRIPAGTPVYMINRDENEKLFIVKILQPVS  
130 140 150 160 170 180

CV127\_ NIKIF  
::  
qi | 193 APGHFEAFYGAGGEDPESFYRAFSWEVLEAALKVRREOLEKVFGEQSKGSIVKASREKIR

```

190      200      210      220      230      240
10
CV127_ ILKLNKRGPPYLSP
      .. ..::: . :
gi|193 ALSQHEEGPPRIWPFGGESSGPINLLHKHPSQSNQFGRLYEAHPDDHKQLQDLDMVSFA
      250      260      270      280      290      300

gi|193 NITKGMAGPYNSRATKISVVVEGEGFFEMACPHLSSSSGSYQKISARLRRGVVFVAPA
      310      320      330      340      350      360

gi|193 GHPVAVIASQNNNLQVLCFEVNAHGNSRFPLAGKGNIVNEFERDAKELAFNLPSREVERI
      370      380      390      400      410      420

gi|193 FKNQDQAFFFPGNKQQEEGGRGGRAFE
      430      440

>>gi|585290 gi|585290|sp|P32936.2|IAAB_HORVU RecName: Fu (149 aa)
  initn: 42 initl: 42 opt: 43 Z-score: 91.4 bits: 20.0 E(): 3.8
Smith-Waterman score: 43; 45.455% identity (81.818% similar) in 11 aa overlap
(8-18:54-64)

```

```

10
CV127_ NIKIFILKLNKRG
      ... ::
gi|585 MASKSSCDLLLA AVLVSIFA AVA AVG SEDCTPWTATPITPLPSCR DYVEQQACRIETPGP
      10      20      30      40      50      60

```

```

CV127_ PYLSP
      :::
gi|585 PYLAKQQCCGELANIPQQCRQCALRFFMGRKSRPDQSGLMELPGCPREVQMD FVRILVTP
      70      80      90      100      110      120

gi|585 GFCNLTTVHNTPYCLAMDEWQWNRQFCSS
      130      140

```

```

>>gi|2493414 gi|2493414|sp|Q28050.1|S10A7_BOVIN RecName: (101 aa)
  initn: 35 initl: 35 opt: 41 Z-score: 90.3 bits: 19.2 E(): 4.4
Smith-Waterman score: 41; 44.444% identity (61.111% similar) in 18 aa overlap
(1-18:39-56)

```

```

10
CV127_ NIKIFILKLNKRGPPYLSP
      :. :. .::: :::
gi|249 MSSSQLEQAITDLINLFHKYSGSDDTIEKEDLLRLMKDNFPNFLGACEKRGRDYLSNIFE
      10      20      30      40      50      60

gi|249 KQDKNKDRKIDFSEFLSLLADIATDYHNHSHGAQLCSGGNQ
      70      80      90      100

```

```

>>gi|21711 gi|21711|emb|CAA42453.1| CM 17 protein precur (143 aa)
  initn: 42 initl: 42 opt: 42 Z-score: 89.8 bits: 19.6 E(): 4.7
Smith-Waterman score: 42; 83.333% identity (100.000% similar) in 6 aa overlap
(13-18:59-64)

```

```

10
CV127_ NIKIFILKLNKRG
      ::
gi|217 MASKSNYNLLFTALLVFIFA AVA AVG NEDCTPWTSTLITPLPSCR NYVEEQACRIEMP GP
      10      20      30      40      50      60

```

CV127\_ PYLSP

:::

gi|217 PYLAKQECCEQLANIPQQCRQCQALRYFMGPKSRPDQSGLMELPGCPREVQMNFPILVTP  
70 80 90 100 110 120

gi|217 GYCNLTTHNTTPYCLGMEESQWS  
130 140

>>gi|256427 gi|256427|gb|AAB23463.1| CG4 beta-conglycini (439 aa)  
initn: 30 initl: 30 opt: 45 Z-score: 87.8 bits: 20.9 E(): 6.1  
Smith-Waterman score: 45; 46.667% identity (86.667% similar) in 15 aa overlap  
(1-14:50-64)

CV127\_ NIKIFIL-KLN  
10  
gi|256 MMRVRFPLLVLVLTGTVFLASVCVSLKVREDDNNPFYFRSSNSFQTLFENQNVRIIRLLQRFN  
10 20 30 40 50 60

CV127\_ KRGPYPYLSP  
:::  
gi|256 KRSPQLENLRDYRIVQFQSKPNTILLPHHADADFLFVLSGRAILTLVNNDDRDSYNLHP  
70 80 90 100 110 120

gi|256 GDAQRIAGTYYLVNPHDHQNLKIIKLAIPVNKPGRYDDFFLSSTQAQQSYLQGFHSHNI  
130 140 150 160 170 180

gi|256 LETSFHSEFEEINRVLFGEERQEQGVIVELSKEQIRQLSRRAKSSSRKTISSEDEPF  
190 200 210 220 230 240

gi|256 NLRSRNPIYSNNFGKFFEITPEKNPQLRDLDFLSSVDINEGALLPHFNSKAIVILVIN  
250 260 270 280 290 300

gi|256 EGDANIELVGIKEQQQKQKQEEPLEVQRYRAELSEDDVVFIPAAYPFVFNATSNLFLA  
310 320 330 340 350 360

gi|256 FGINAENNQRNFLAGKDNVVRQIERQVQELAFPGSAQDVERLLKKQRESYFVDAQPPQK  
370 380 390 400 410 420

gi|256 EEGSKGRKGPFPFPSILGALY  
430

>>gi|115502167 gi|115502167|sp|Q1ZYQ8.2|EXB10\_MAIZE RecN (270 aa)  
initn: 36 initl: 36 opt: 42 Z-score: 85.4 bits: 19.7 E(): 8.3  
Smith-Waterman score: 42; 31.250% identity (62.500% similar) in 16 aa overlap  
(4-19:23-38)

CV127\_ NIKIFILKLNKRGPYPYLSP  
10  
gi|115 MAVNVRTMSSMRAQVAMVVALVFLVRGAWCGPPKVPKGNITATYGKDWDALAKATWYGK  
10 20 30 40 50 60

gi|115 PTGAGPDDNGGCGYKDVNKPFPNSMGACGNIPFIDKGLGCGSCFEIKCDKPVECSGKPV  
70 80 90 100 110 120

gi|115 VVHITDMNYEPIAAYHFDLAGTAFGAMAKKGEEELRKAGIIDMQFRRVKCKYDSKVTFH  
130 140 150 160 170 180

gi|115 LEKGCGPNYLALLVKYVDGDIIVAVDVKEKGSPTYEPLKHSWGAIWRKDSKPLKGLPT

```

          190          200          210          220          230          240
gi|115 VRLTTEGGTKSVYDDVIPANWKANTAYTAK
          250          260          270
```

19 residues in 1 query sequences  
307888 residues in 1386 library sequences  
Scomplib [34.26]  
start: Wed Mar 11 06:15:46 2009 done: Wed Mar 11 06:15:46 2009  
Total Scan time: 0.050 Total Display time: 0.000

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

### CV127\_884\_886\_frame1

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

Please cite:

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

CV127\_884\_886\_frame1, 1 aa

vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences  
unscaled statistics: mu= 50.0000 var=10.0000; Lambda= 0.4056

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2  
join: 15, opt: 20, open/ext: -10/-2, width: 16  
!! No sequences with E() < 10.000000

1 residues in 1 query sequences  
307888 residues in 1386 library sequences  
Scomplib [34.26]  
start: Wed Mar 11 06:24:18 2009 done: Wed Mar 11 06:24:19 2009  
Total Scan time: 0.040 Total Display time: 0.000

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

### CV127\_890\_973\_frame1

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

Please cite:

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

CV127\_890\_973\_frame1, 28 aa

vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences  
Expectation\_n fit: rho(ln(x))= 0.4456+/-0.00317; mu= 20.3626+/- 0.166  
mean\_var=30.5286+/- 8.779, 0's: 28 Z-trim: 28 B-trim: 10 in 1/41  
Lambda= 0.232124

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

The best scores are:

	gi	5813788	gi	5813788	gb	AAD52012.1	(	726)	opt	bits	E(1386)	%_id	%_sim	alen	
	gi	5813788	gi	5813788	gb	AAD52012.1	(	726)	45	22.4		5	0.556	0.889	9
	gi	23894227	gi	23894227	emb	CAD2337	(	726)	45	22.4		5	0.556	0.889	9
	gi	23894232	gi	23894232	emb	CAD2361	(	726)	45	22.4		5	0.556	0.889	9
	gi	89892727	gi	89892727	gb	ABD79097	(	404)	43	21.0	7.6	0.545	0.818	11	
	gi	89892725	gi	89892725	gb	ABD79096	(	410)	43	21.0	7.6	0.545	0.818	11	

>>>CV127\_890\_973\_frame1, 28 aa vs

/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

>>gi|5813788 gi|5813788|gb|AAD52012.1|AF082514\_1 Tri r 4 (726 aa)  
initn: 41 initl: 41 opt: 45 Z-score: 89.3 bits: 22.4 E(): 5  
Smith-Waterman score: 45; 55.556% identity (88.889% similar) in 9 aa overlap  
(7-15:467-475)

gi|581 MAAKWLIASLAFASSGLAFTPEDFISAPRRGEAIPDPKGELAVFHVSKYNFDKKDRPSG  
10 20 30 40 50 60

gi|581 WNLLNLKNGDINVLTTSDVSEITWLGDKTKVVYINGTDSVKGKGVGIWISDAKNFGNAYK  
70 80 90 100 110 120

gi|581 AGSVNGAFSGLKLAKSGDKINFGYQSTTKGDLNEAAAKEAVSSARIYDSLFRHWD  
130 140 150 160 170 180

gi|581 YVGTQFNAVFSGTLTKSGDKYSFDGKLKLVQPVKYAESPPFGSGDYDLSSDGKTVA  
190 200 210 220 230 240

gi|581 FMSKAPELPAKANLTTSYIFLVPDGSRAEPINKRNGPRTPQIEGASSSPVFS  
250 260 270 280 290 300

gi|581 AYLQMAAKNYESDRRVIHIAEVGTNKPVQRIASNWDSPKAVKSSDGRITLYVTAEDHAT  
310 320 330 340 350 360

gi|581 GKLFITLPADARDNHKPAVVKHDSVSSFFYFIGSSKSVLISGNSLWSNALYQVATPDRPNR  
370 380 390 400 410 420

CV127\_ 10 20  
TKGALLIYPLSFYFHFLIK  
.: :.

gi|581 KLFYANEHDPKLGPNIDIEPLWVDGARTKIHSWIVKPTGFDKNKVYPLAFILHGGPQG  
430 440 450 460 470 480

CV127\_ GQTSLVIY

gi|581 SWGDNWSTRWNPRVWADQGYVVPAPNPTGSTGFGQKLTDDITNDWGGAPYKDLVKIWEHV  
490 500 510 520 530 540

gi|581 HDHIKYIDTDNGIAAGASFGGMVNWIIQGDGRKFKALVSHDGTTFVGSSKIGTDELFFI  
550 560 570 580 590 600

gi|581 EHDFNGTFFEARQNYDRWDCSKPELVAKWSTPQLVVHNDFFRLSVAEGVGLFNVLQKEG  
610 620 630 640 650 660

gi|581 VPSRFLNFPDETHWVTKPENSLVWHQQVLGWVKNKWSGINKSNPKSIKLSDCPIEVVDHEA  
670 680 690 700 710 720

gi|581 HSYFDY

>>gi|23894227 gi|23894227|emb|CAD23374.1| tri s 4 allerg (726 aa)

initn: 41 initl: 41 opt: 45 Z-score: 89.3 bits: 22.4 E(): 5  
Smith-Waterman score: 45; 55.556% identity (88.889% similar) in 9 aa overlap  
(7-15:467-475)

gi|238 MAAAKWLIASLAFASSGLAFTPEDFISAPRRGEAIPDPKGELAVFHVSKYNFDKKDRPSG  
10 20 30 40 50 60

gi|238 WNLLNLKNGDISVLTTDSVSEITWLGDGTVVYVNGTDSVEGGVGIWISDAKNFGNAYK  
70 80 90 100 110 120

gi|238 AGSVNGAFSGLKLAKAGDKINFVGYGQSTTKGDLYNEAAAKEAVSSARIYDGLFVRHWDT  
130 140 150 160 170 180

gi|238 YVGTQFNAVFSGSLTKNGDKYSFDGKLKLVQPVKYAESPPYPPFGGSGDYDLSSDGKTVA  
190 200 210 220 230 240

gi|238 FMSKAPELPKANLTTSYIFLVPDGRVAEPINKRNGPRTPOGIEGASSSPVFSPOGKRI  
250 260 270 280 290 300

gi|238 AYLQMATKNYESDRRVIHIAEVGSNKPVQRIASSWDRSPEAVKWSSDGRITLYVTAEDHAT  
310 320 330 340 350 360

gi|238 GKLFITLPADARDNHKPSVVKHDGVSFFYFIGSSKSVLISGNSLWSNALLYQVATPGRPNR  
370 380 390 400 410 420

CV127\_ 10 20  
TKGALLIYPLSFYFHFFLIK  
.: :.:

gi|238 KLFYANEHDPKGLGPKDIEPLWVDGARTKIHSWIVKPTGFDKNKVYPLAFLIHGGPQG  
430 440 450 460 470 480

CV127\_ GQTSLVIIY

gi|238 SWGDSWSTRWNPRVWADQGYVVVAPNPTGSTGFGQKLTDDITNDWGGAPYKDLVKIWEHV  
490 500 510 520 530 540

gi|238 RDHIKYIDTDNGIAAGASFGGFMVNWIIQGDGRKFKALVSHDGTGFGSSKIGTDELFFI  
550 560 570 580 590 600

gi|238 EHDENGTFEFARQNYDRWDCSKPELVAKWSTPQLVIHNDSDFRLSVAEGVGLFNVLQKEG  
610 620 630 640 650 660

gi|238 IPSRFLNFPDETHWVTKPENSLVWHQQVLGWINKWSGINKSNPKSIKLSDCPIEVIDHEA  
670 680 690 700 710 720

gi|238 HSYFDY

>>gi|23894232 gi|23894232|emb|CAD23611.1| tri m 4 allerg (726 aa)  
initn: 41 initl: 41 opt: 45 Z-score: 89.3 bits: 22.4 E(): 5  
Smith-Waterman score: 45; 55.556% identity (88.889% similar) in 9 aa overlap  
(7-15:467-475)

gi|238 MAAAKWLIASLAFASSGLAFTPEDFISAPRRGEAIPDPKGELAVFHVSKYNFDKKDRPSG  
10 20 30 40 50 60

gi|238 WNLLNLKNGDISVLTTDSVSEITWLGDGTVVYVNGTDSVEGGVGIWISDAKNFGNAYK  
70 80 90 100 110 120

gi|238 AGSVNGAFSGLKLAKAGDKINFVGYGQSTTKGDLYNEAAAKEAVSSARIYDGLFVRHWDT  
130 140 150 160 170 180

gi|238 YVGTQFNAVFSGLTKNGDKYSFDGKLKLNLPVKYAESPYPPFGGSGDYDLSSDGKTVA  
190 200 210 220 230 240

gi|238 FMSKAPELPAKANTTSYIFLVPHDGSRVAEPINKRNGPRTPOGIEGASSSPVFSPODGRI  
250 260 270 280 290 300

gi|238 AYLMATKNYESDRRVIHIAEVGSNKPVQRIASSWDRSPEAVKWSSDGRITLYVTAEDHAT  
310 320 330 340 350 360

gi|238 GKLFTLPADARDNHKPSVVKHDGVSFFYFIGSSKSVLISGNSLWSNLYQVATPGRPNR  
370 380 390 400 410 420

CV127\_ TKGALLIYPLSFYFHFFLIK  
10 20  
.:...: .:

gi|238 KLFYANEHDPKLGKLPKDIEPLWVDGARTKIHSWIVKPTGFDKNKVYPLAFLIHGGPQG  
430 440 450 460 470 480

CV127\_ GQTSLSVIY

gi|238 SWGDSWSTRWNPRVWADQGYVVVAPNPTGSTGFGQKLTDDITNDWGGAPYKDLVKIWEHV  
490 500 510 520 530 540

gi|238 RDHIKYIDTDNGIAAGASFGGFMVNWIIQGDGRKFKALVSHDGTGSSKIGTDELFFI  
550 560 570 580 590 600

gi|238 EHDFTNGTFFEARQNYDRWDCSKPELVAKWSTPQLVIHNDDFRLSVAEGVGLFNVLQEKG  
610 620 630 640 650 660

gi|238 IPSRFLNFPDETHWVTKPENSLVWHQQVLGWINKWSGINKSNPKSIKLSDCPIEVIDHEA  
670 680 690 700 710 720

gi|238 HSYFDY

>>gi|89892727 gi|89892727|gb|ABD79097.1| Zea m 13 allerg (404 aa)  
initn: 43 initl: 43 opt: 43 Z-score: 86.1 bits: 21.0 E(): 7.6  
Smith-Waterman score: 43; 54.545% identity (81.818% similar) in 11 aa overlap  
(2-12:74-84)

gi|898 ARGALFLLALFCVVHGEKAKSKDNDKASGPGGSFDITKLGASGNGKTDSTKAVQEAWAS  
10 20 30 40 50 60

CV127\_ TKGALLIYPLSFYFHFFLIKGQTSLSVIY  
10 20  
::...: .:

gi|898 ACGGTGKQTILIPKGDFLVGPLNFTGPCKGDVTIQVNGNLLATTDLSQYKDHGNWIEILR  
70 80 90 100 110 120

gi|898 VDNLVITGKGKLDGQGPVWWSKNSCVKKYDCKILPNSLVMDFVNNGEVSGITLLNSKFFH  
130 140 150 160 170 180

gi|898 MNMYKCKDMLIKDVNVTAPGDSPTNDGIHMGDSSGVITNTVIGVGDDCISIGPGTFKVN  
190 200 210 220 230 240

gi|898 ITGVTGPGHGISIGSLGRYKDEKDVTDINVKDCTLKKTANGVRIKAYEDAASVLTASKI  
250 260 270 280 290 300

gi|898 HYENIKMEDSGYPPIIDMKYCPNKLCTANGASKVTVKDVTFKNIPGTSSTPEAVNLLCSA  
310 320 330 340 350 360



```
gi|898 KIPCTGVTMDDVNIKYSGTNNKTMVCKNAKGSAGCLKELACF
      370      380      390      400

>>gi|89892725 gi|89892725|gb|ABD79096.1| Zea m 13 allerg (410 aa)
  initn: 43 initl: 43 opt: 43 Z-score: 86.1 bits: 21.0 E(): 7.6
Smith-Waterman score: 43; 54.545% identity (81.818% similar) in 11 aa overlap
(2-12:80-90)

gi|898 MACIDNAMRALFLLALFCVVHGEKAKSKDNDASGPGGSFDITKLGASGNGKTDSTKAV
      10      20      30      40      50      60

CV127_      10      20
      TKGALLIYPLSFYFHFFLIKGQTS�VIY
      :: .:: ::::
gi|898 QEAWASACGGTGKQTILIPKGDPLVGPLNFTGPCKGDVTIQVNGNLLATTDLSQYKDHN
      70      80      90      100      110      120

gi|898 WIEILRVDNLVITGKGKLDGQGPVWSKNSCVKKYDCKILPNSLVMDFVNNGEVSGITLL
      130      140      150      160      170      180

gi|898 NSKFFHMNMYKCKDMLIKDVNVTAPGDSPTDGIHMGDSSGVTITNTVIGVGDDCISIGP
      190      200      210      220      230      240

gi|898 GTSKVNITGVTCGPGHGISIGSLGRYKDEKDVTDINVKDCTLKKTANGVRIKAYEDAASV
      250      260      270      280      290      300

gi|898 LTASKIHYENIKMEDSGYPPIIDMKYCPNKLCTANGASKVTVKDVTFKNITGTSSTPEAV
      310      320      330      340      350      360

gi|898 NLLCSAKIPCTGVTMDDVNIKYSGTNNKTMVCKNAKGSAGCLKELACF
      370      380      390      400      410
```

28 residues in 1 query sequences  
 307888 residues in 1386 library sequences  
 Scomplib [34.26]  
 start: Wed Mar 11 06:26:24 2009 done: Wed Mar 11 06:26:24 2009  
 Total Scan time: 0.060 Total Display time: 0.010

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

## CV127\_922\_957\_frame0

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

Please cite:

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

CV127\_922\_957\_frame0, 12 aa

vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences

Expectation\_n fit:  $\rho(\ln(x)) = 3.5465 \pm 0.00228$ ;  $\mu = -0.7063 \pm 0.119$   
 mean\_var=14.2628 $\pm$  3.705, 0's: 59 Z-trim: 65 B-trim: 43 in 1/39  
 Lambda= 0.339604

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

```

join: 36, opt: 20, open/ext: -10/-2, width: 16
The best scores are:
gi|939932 gi|939932|emb|CAA51775.1| ( 88) 39 22.6 0.24 0.600 0.800 10
gi|4007655 gi|4007655|emb|CAA10348.1| ( 122) 40 22.9 0.25 0.600 0.800 10
gi|4007636 gi|4007636|emb|CAA10350.1| ( 122) 40 22.9 0.25 0.600 0.800 10
gi|1171009 gi|1171009|sp|P43214.1|M ( 122) 40 22.9 0.25 0.600 0.800 10
gi|4007040 gi|4007040|emb|CAA10345.1| ( 122) 40 22.9 0.25 0.600 0.800 10
gi|126386 gi|126386|sp|P14947.1|MPA ( 97) 39 22.5 0.27 0.600 0.800 10
gi|1093120 gi|1093120|prf|2103117A ( 196) 39 22.3 0.62 0.600 0.800 10
gi|94706935 gi|94706935|sp|P0C1B3.1 ( 499) 40 22.5 1.4 0.545 0.909 11
gi|14423843 gi|14423843|sp|O64943.2 ( 165) 34 19.9 2.8 0.600 0.800 10
gi|45510889 gi|45510889|gb|AAS67042 ( 316) 36 20.7 3.1 0.667 0.778 9
gi|45510887 gi|45510887|gb|AAS67041 ( 337) 36 20.7 3.3 0.667 0.778 9
gi|462717 gi|462717|sp|P33556.1|NLT ( 38) 28 17.4 3.6 0.444 0.889 9
gi|2506460 gi|2506460|sp|P02221.2|G ( 158) 33 19.4 3.7 0.556 0.889 9
gi|126387 gi|126387|sp|P14948.1|MPA ( 97) 31 18.6 4.1 0.625 0.750 8
gi|1545803 gi|1545803|dbj|BAA04558.1 ( 349) 35 20.2 4.9 0.625 0.875 8

```

```

>>>CV127_922_957_frame0, 12 aa vs
/n/na4/bioinfo/refsets/PUB_AllergenOnlineJan2009_V9.fasta library

>>gi|939932 gi|939932|emb|CAA51775.1| allergen Lol p II (88 aa)
initn: 39 initl: 39 opt: 39 Z-score: 113.1 bits: 22.6 E(): 0.24
Smith-Waterman score: 39; 60.000% identity (80.000% similar) in 10 aa overlap
(1-10:62-71)

```

```

gi|939 EFTVEKGSDEKNLALSIKYSKEGDAMAEVELKEHGSNEWLALKKNGDGVWEIKSDKPLKG
      10          20          30          40          50          60

      10
CV127_ FLFSFSLNKGAN
      : : : : :
gi|939 PFNFRFVSEKGMWNVFDVVPADFKVGT
      70          80

```

```

>>gi|4007655 gi|4007655|emb|CAA10348.1| pollen allergen (122 aa)
initn: 40 initl: 40 opt: 40 Z-score: 112.7 bits: 22.9 E(): 0.25
Smith-Waterman score: 40; 60.000% identity (80.000% similar) in 10 aa overlap
(1-10:90-99)

```

```

gi|400 MSMASSSSSSLLAMAVLAALFAGAWCPKVTFTVEKGSNEKHLAVLVKYEGLTMAEVELR
      10          20          30          40          50          60

      10
CV127_ FLFSFSLNKGAN
      : : : : :
gi|400 EHGSDEWVAMTKGEGGVWTFDSEEPLQGPFNFRFLTEKGMKNVFDVVPPEKYTIGATYAP
      70          80          90          100          110          120

```

gi|400 EE

```

>>gi|4007636 gi|4007636|emb|CAA10350.1| pollen allergen (122 aa)
initn: 40 initl: 40 opt: 40 Z-score: 112.7 bits: 22.9 E(): 0.25
Smith-Waterman score: 40; 60.000% identity (80.000% similar) in 10 aa overlap
(1-10:90-99)

```

```

gi|400 MSMASSSSSSLLAMAVLAALFAGAWCPKVTFTVEKGSNEKHLAVLVKYEGLTMAEVELR
      10          20          30          40          50          60

```

10

```

CV127_                                FLFSFSLNKGAN
                                     : : : : :
gi|400 EHGSDEWVAMTKGEGGVWTFDSEEPLQGPFNFRFLTEKGMKNVFDVVPPEKYTIGATYAP
      70          80          90          100          110          120

```

```

gi|400 EE

```

```

>>gi|1171009 gi|1171009|sp|P43214.1|MPAP2_PHLPR RecName: (122 aa)
  initn: 40 initl: 40 opt: 40 Z-score: 112.7 bits: 22.9 E(): 0.25
Smith-Waterman score: 40; 60.000% identity (80.000% similar) in 10 aa overlap
(1-10:90-99)

```

```

gi|117 MSMASSSSSLLAMAVLAALFAGAWCVPKVTFTVEKGSNEKHLAVLVKYEGLTMAEVELR
      10          20          30          40          50          60

```

```

                                     10
CV127_                                FLFSFSLNKGAN
                                     : : : : :
gi|117 EHGSDEWVAMTKGEGGVWTFDSEEPLQGPFNFRFLTEKGMKNVFDVVPPEKYTIGATYAP
      70          80          90          100          110          120

```

```

gi|117 EE

```

```

>>gi|4007040 gi|4007040|emb|CAA10345.1| pollen allergen (122 aa)
  initn: 40 initl: 40 opt: 40 Z-score: 112.7 bits: 22.9 E(): 0.25
Smith-Waterman score: 40; 60.000% identity (80.000% similar) in 10 aa overlap
(1-10:90-99)

```

```

gi|400 MSMASSSSSGLLAMAVLAALFAGAWCVPKVTFTVEKGSNEKHLAVLVKYEGLTMAEVELR
      10          20          30          40          50          60

```

```

                                     10
CV127_                                FLFSFSLNKGAN
                                     : : : : :
gi|400 EHGSDEWVAMTKGEGGVWTFDSEEPLQGPFNFRFLTEKGMKNVFDVVPPEKYTIGATYAP
      70          80          90          100          110          120

```

```

gi|400 EE

```

```

>>gi|126386 gi|126386|sp|P14947.1|MPAL2_LOLPR RecName: F (97 aa)
  initn: 39 initl: 39 opt: 39 Z-score: 112.2 bits: 22.5 E(): 0.27
Smith-Waterman score: 39; 60.000% identity (80.000% similar) in 10 aa overlap
(1-10:66-75)

```

```

gi|126 AAPVEFTVEKGSDEKNLALSIKYNKEGDSMAEVELKEHGSNEWLALKKNGDGVWEIKSDK
      10          20          30          40          50          60

```

```

                                     10
CV127_                                FLFSFSLNKGAN
                                     : : : : :
gi|126 PLKGPFFNFRFVSEKGMKNVFDVVPADFKVGTTYKPE
      70          80          90

```

```

>>gi|1093120 gi|1093120|prf||2103117A allergen Dac g II (196 aa)
  initn: 39 initl: 39 opt: 39 Z-score: 105.6 bits: 22.3 E(): 0.62
Smith-Waterman score: 39; 60.000% identity (80.000% similar) in 10 aa overlap
(1-10:66-75)

```

```

gi|109 EAPVTFTVEKGSDEKNLALSIKYNKEGDSMAEVELKEHGSNEWLALKKNGDGVWEIKSDK

```

```

10      20      30      40      50      60

CV127_      10
      FLFSFLSNKGAN
      : : : : :
gi|109 PLKGPFNFRFVSEKGMNRNVFVDVVPADFKVGTITYKPEEAAASASRRRSSEVFQFLILSCQ
      70      80      90      100      110      120

gi|109 GRIVNNCEVLICVMRRGNAMCLIASISMHHILTLDRFFFDGLEIIYKIFKMMFQKPRTRT
      130      140      150      160      170      180

gi|109 CIEKDFPRSSSSSIPT
      190

>>gi|94706935 gi|94706935|sp|P0C1B3.1|AMYA1_ASPOR RecNam (499 aa)
  initn: 37 initl: 37 opt: 40 Z-score: 99.4 bits: 22.5 E(): 1.4
Smith-Waterman score: 40; 54.545% identity (90.909% similar) in 11 aa overlap
(2-12:426-436)

gi|947 MMVAWWSLFLYGLQVAAPALAATPADWRSQSIYFLLTDRFARTDGGSTTATCNTADQKYCG
      10      20      30      40      50      60

gi|947 GTWQGIIDKLDYIQGMGFTAIWITPVTALPQTAYGDAYHGYWQQDIYSLNENYGTADD
      70      80      90      100      110      120

gi|947 LKALSSALHERGMYLMVDVVANHMGYDGAGSSVDYSVFKPFSSQDYFHPFCFIQNYEDQT
      130      140      150      160      170      180

gi|947 QVEDCWLGDNTVSLPDLDTTKDVVKNEWYDWVGSLSVSNYSIDGLRIDTVKHVQKDFWPGY
      190      200      210      220      230      240

gi|947 NKAAGVYCIGEVLDGDPAYTCPYQNVMDGVLNYPYIYPLLNAFKSTSGSMDDLYNMINTV
      250      260      270      280      290      300

gi|947 KSDCPDSTLLGTFFVENHDNPRFASYTNDIALAKNVAAFIILNDGIPIIYAGQEYHYAGGN
      310      320      330      340      350      360

gi|947 DPANREATWLSGYPTDSELYKLIASANAIRNYAISKDTGFVITYKNWPIYKDDTTIAMRK
      370      380      390      400      410      420

CV127_      10
      FLFSFLSNKGAN
      . . . . .
gi|947 TDGSQIVTILSNKGASGDSYTLISLGSAGYTAGQQLTEVIGCTTVTVGSDGNVPVPMAGGL
      430      440      450      460      470      480

gi|947 PRVLYPTEKLAGSKICSSS
      490

>>gi|14423843 gi|14423843|sp|O64943.2|POLC2_JUNOX RecNam (165 aa)
  initn: 27 initl: 27 opt: 34 Z-score: 93.9 bits: 19.9 E(): 2.8
Smith-Waterman score: 34; 60.000% identity (80.000% similar) in 10 aa overlap
(3-12:83-92)

gi|144 MDEVPSDGSKSACSGEVVMEQSVHELEEVEFKKFDANGDGKISGSELADILRSLGSDVGE
      10      20      30      40      50      60

CV127_      10
      FLFSFLSNKGAN
      : : : : :
gi|144 AEVKAMMEEADADGDGYVSLQEFVDLNNKGASVKDLKNAFKVFDRDCNGSISAAELCHTL
      70      80      90      100      110      120

```

gi|144 ESVGEPCTIEESKNIIHNVDKNGDGLISVEEFQTMMTSEMTDKSK  
130 140 150 160

>>gi|45510889 gi|45510889|gb|AAS67042.1| venom phospholi (316 aa)  
initn: 36 initl: 36 opt: 36 Z-score: 93.1 bits: 20.7 E(): 3.1  
Smith-Waterman score: 36; 66.667% identity (77.778% similar) in 9 aa overlap  
(4-12:186-194)

gi|455 ADDLTTLRNGTLDRGITPDCTFNEKDIELHVYSRDKRNGIILKKEILKNYDLFQKSQISH  
10 20 30 40 50 60

gi|455 QIAILIHGFLSTGNNENFDAMAKALIEIDNFLVISVDWKKGACNAFASTNDVLGYSQAVG  
70 80 90 100 110 120

gi|455 NTRHVGKYVADFTKLLVEQYKVPMSNIRLIGHSLGAHTSGFAGKEVQRLKLGKYKEIIGL  
130 140 150 160 170 180

10  
CV127\_ FLFSFSLNKGAN  
::: : :  
gi|455 DPAGPSFLTNKCPNRLCETDAEYVQAIHTSAILGVYYNVGSDVDFYVNYGKSQPGCSEPC  
190 200 210 220 230 240

gi|455 SHTKAVKYLTECIKRECCLIGTPWKSYPSTPKPISQCKRDTVCVGLNAQSYPAKGSFYV  
250 260 270 280 290 300

gi|455 PVEKDAPYCHNEGIKL  
310

>>gi|45510887 gi|45510887|gb|AAS67041.1| venom phospholi (337 aa)  
initn: 36 initl: 36 opt: 36 Z-score: 92.5 bits: 20.7 E(): 3.3  
Smith-Waterman score: 36; 66.667% identity (77.778% similar) in 9 aa overlap  
(4-12:207-215)

gi|455 MNFKYSILFICFVKVLDNCYAADDLTTLRNGTLDRGITPDCTFNEKDIELHVYSRDKRNG  
10 20 30 40 50 60

gi|455 IILKKEILKNYDLFQKSQISHQIAILIHGFLSTGNNENFDAMAKALIEIDNFLVISVDWK  
70 80 90 100 110 120

gi|455 KGACNAFASTNDVLGYSQAVGNTRHVGKYVADFTKLLVEQYKVPMSNIRLIGHSLGAHTS  
130 140 150 160 170 180

10  
CV127\_ FLFSFSLNKGAN  
::: : :  
gi|455 GFAGKEVQRLKLGKYKEIIGLDPAGPSFLTNKCPNRLCETDAEYVQAIHTSAILGVYYNV  
190 200 210 220 230 240

gi|455 GSDVDFYVNYGKSQPGCSEPCSHTKAVKYLTECIKRECCLIGTPWKSYPSTPKPISQCKR  
250 260 270 280 290 300

gi|455 DTCVCVGLNAQSYPAKGSFYVPVDKAPYCHNEGIKL  
310 320 330

>>gi|462717 gi|462717|sp|P33556.1|NLTP2\_VITSX RecName: F (38 aa)  
initn: 28 initl: 28 opt: 28 Z-score: 91.9 bits: 17.4 E(): 3.6  
Smith-Waterman score: 28; 44.444% identity (88.889% similar) in 9 aa overlap  
(3-11:15-23)

```

CV127_          FLFSFLSNKGAN
                  ...:. :
gi|462 AITCGQVSSALSSCLGYLKNNGGAVPPGSSCGIKNLNSA
                  10          20          30

>>gi|2506460 gi|2506460|sp|P02221.2|GLB1_CHITH RecName: (158 aa)
  initn: 24 initl: 24 opt: 33 Z-score: 91.7 bits: 19.4 E(): 3.7
Smith-Waterman score: 33; 55.556% identity (88.889% similar) in 9 aa overlap
(1-9:149-157)

gi|250 MKFLILALCVAAAMAGPSGDQIAAAKASWNTVKNNQVDILYAVFKANPDIQTAFSQFAGK
                  10          20          30          40          50          60

gi|250 DLDSIKGTPDFSKHAGRVVGLFSEVMDLLGNDANTPTILAKAKDFGKSHKSRASPAQLDN
                  70          80          90          100          110          120

                  10
CV127_          FLFSFLSNKGAN
                  :.:. :.:.
gi|250 FRKSLVVYLKGATKWDSAVESSWAPVLDVFVSTLKNEL
                  130          140          150

>>gi|126387 gi|126387|sp|P14948.1|MPAL3_LOLPR RecName: F (97 aa)
  initn: 24 initl: 24 opt: 31 Z-score: 91.0 bits: 18.6 E(): 4.1
Smith-Waterman score: 31; 62.500% identity (75.000% similar) in 8 aa overlap
(3-10:66-73)

gi|126 TKVDLTVEKGSDAKTLVLNIKYTRPGDTLAEVELRQHGSEEWEPMTKKGNLWEVKSAPKL
                  10          20          30          40          50          60

                  10
CV127_          FLFSFLSNKGAN
                  : :.:. :
gi|126 TGPMMNFRFLSKGGMKNVFDEVIPTAFTVGKTYTPEYN
                  70          80          90

>>gi|1545803 gi|1545803|dbj|BAA04558.1| Mag3 [Dermatopha (349 aa)
  initn: 35 initl: 35 opt: 35 Z-score: 89.6 bits: 20.2 E(): 4.9
Smith-Waterman score: 35; 62.500% identity (87.500% similar) in 8 aa overlap
(3-10:224-231)

gi|154 VTALELLLLKGETEDKTRRYVAELTAVGSPSNKQAKAQIEVTKGEEYKITLKSPEHEFNTE
                  10          20          30          40          50          60

gi|154 FTIHADKNNLKMHMDFPNVFAQDLTGTFQHDKENNVKRNQNLQYKFAGDEKPHTVDYEN
                  70          80          90          100          110          120

gi|154 EFSFNLKRSSKDKNSGVYRAKYMSSHFPILNHKVNIQFKYRPFKVNELNLEGEFGRELQ
                  130          140          150          160          170          180

                  10
CV127_          FLFSFLSNKGAN
                  :.. :.:.
gi|154 HKFQLMRNSQIEVEEVRPFKMHGNSDIKLMANDLDIDYDLKSEFKYESNKGTPIELQYKI
                  190          200          210          220          230          240

gi|154 SGKDRSKRAADLGAEDVEGVIDYKNNGSPIDSKMHAHLKMKGNNGYDSELKQTQPQQYE
                  250          260          270          280          290          300

gi|154 GKITLSKNDKKIFINHKSEMTKPTNTFHLKTDADVSYSDSMKKHYQME
                  310          320          330          340

```

12 residues in 1 query sequences  
307888 residues in 1386 library sequences  
Scomplib [34.26]  
start: Wed Mar 11 06:17:59 2009 done: Wed Mar 11 06:17:59 2009  
Total Scan time: 0.050 Total Display time: 0.010

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

## CV127\_961\_984\_frame0

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

Please cite:

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

CV127\_961\_984\_frame0, 8 aa

vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences

Expectation\_n fit:  $\rho(\ln(x)) = 3.2953 \pm 0.00231$ ;  $\mu = -1.8580 \pm 0.122$   
mean\_var=10.5314  $\pm$  2.688, 0's: 109 Z-trim: 109 B-trim: 0 in 0/36  
Lambda= 0.395213

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

The best scores are:

	gi 18536 gi 18536 emb CAA35691.1  u ( 605)	gi 42414629 gi 42414629 emb CAF2523 ( 415)	gi 42414627 gi 42414627 emb CAF2523 ( 415)	gi 29539109 gi 29539109 emb CAD8773 ( 418)	gi 32363467 gi 32363467 sp  P82946_ ( 15)	gi 169929 gi 169929 gb AAB01374.1  ( 639)	gi 46396598 gi 46396598 sp  P83834_ ( 10)	gi 29163773 gi 29163773 emb CAD8001 ( 65)	opt bits	E(1386)	%_id	%_sim	alen	
									36	22.6	1	0.750	1.000	8
									31	19.9	4.6	0.833	1.000	6
									31	19.9	4.6	0.833	1.000	6
									31	19.9	4.6	0.833	1.000	6
									20	15.1	4.7	0.667	0.667	6
									32	20.3	5.4	0.625	1.000	8
									17	13.6	9.1	0.500	1.000	4
									23	16.2	9.7	0.600	1.000	5

>>>CV127\_961\_984\_frame0, 8 aa vs

/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

>>gi|18536 gi|18536|emb|CAA35691.1| unnamed protein prod (605 aa)

initn: 36 initl: 36 opt: 36 Z-score: 101.6 bits: 22.6 E(): 1

Smith-Waterman score: 36; 75.000% identity (100.000% similar) in 8 aa overlap (1-8:338-345)

gi|185 MMRARFPLLLLGLVFLASVSVSFGIAYWEKENPKHNKCLQSCNSERDSYRNQACHARCNL  
10 20 30 40 50 60

gi|185 LKVEKEECEEGEIPRPRPQHPEREPQQPGEKEEDEDEQPRPIPFPRPQPRQEEHEQR  
70 80 90 100 110 120

gi|185 EEQEWPRKEEKRGKEKGSEEEDEDEDEEQDERQFPFPRPPHQKEERNEEEDEDEEQQRESE  
130 140 150 160 170 180

gi|185 ESEDSELRRHKKNPFLFGSNRFETLFKNQYGRIRVLQRFNQSPQLQNLRDYRILEFNS  
190 200 210 220 230 240

gi|185 KPNTLLLPNHADADYLIVILNGTAILSLVNDDRDYSYRLQSGDALRVPSGTTYVVPNDN

```

                250      260      270      280      290      300

CV127_                               SRNILEVN
                                   :::::..
gi|185 NENLRLITLAIPVKNKGRFESFFLSSTEAQQSYLQGFSRNILEASYDTKFEEINKVLFSSR
                310      320      330      340      350      360

gi|185 EEGQQQGEQRLQESVIVEISKEQIRALSKRAKSSSRKTISSEDKPFNLRSDPIYSNKLK
                370      380      390      400      410      420

gi|185 KFFEITPEKNPQLRDLDFLSIVDMNEGALLLPHFNKAIVILVINEGDANIELVGLKEQ
                430      440      450      460      470      480

gi|185 QQEQQQEEQPLEVRKYRAELSEQDIFVIPAGYPVVVNATSNLNFFAIGINAENNQRNFLA
                490      500      510      520      530      540

gi|185 GSQDNVISQIPSQVQELAFPGSAQAVEKLLKNQRESYFVDAQPKKKEEGNKGRKGPLSSI
                550      560      570      580      590      600

gi|185 LRAFY

```

```

>>gi|42414629 gi|42414629|emb|CAF25233.1| Vicilin [Pisum (415 aa)
  initn: 31 initl: 31 opt: 31 Z-score: 90.0 bits: 19.9 E(): 4.6
Smith-Waterman score: 31; 83.333% identity (100.000% similar) in 6 aa overlap
(1-6:151-156)

```

```

gi|424 SRSDQENPFIFKSNRFQTLYENENGHIRLLQKFDKRSKIFENLQNYRLLEYKSKPRTLFL
                10      20      30      40      50      60

gi|424 PQCTDADFILVVLGKATLTVLKSNDNRNSFNLERGDTIKLPAGTIAYLANRDDNEDLRVL
                70      80      90      100      110      120

```

```

CV127_                               SRNILEVN
                                   :::::
gi|424 DLTIPTVKNKPGQLQSFLSGTQNQPSLLSGFSKNILEAAFNTNYEEIEKVLLEQQEQEPQH
                130      140      150      160      170      180

gi|424 RRSCLKDRRQEINEENVIVKVSREQIEELSKNAKSSSKSVSSESGPFNLRSRNPISYKSF
                190      200      210      220      230      240

gi|424 GKFFEITPEKNQQLQDLDFVNSVDIKEGSLLLPNYNSRAIVIVTVTEGKGFELVQQRN
                250      260      270      280      290      300

gi|424 ENQKGKENDKEEQEEETSKQVQLYRAKLSPGDVFVIPAGHPVAINASSDLNLIGFGINAE
                310      320      330      340      350      360

gi|424 NNERNFLAGEEDNVISQVERPVKELAFPGSSHEVDRLKQKQSYFANAQPLQRE
                370      380      390      400      410

```

```

>>gi|42414627 gi|42414627|emb|CAF25232.1| Vicilin [Pisum (415 aa)
  initn: 31 initl: 31 opt: 31 Z-score: 90.0 bits: 19.9 E(): 4.6
Smith-Waterman score: 31; 83.333% identity (100.000% similar) in 6 aa overlap
(1-6:151-156)

```

```

gi|424 SRSDQENPFIFKSNRFQTLYENENGHIRLLQKFDKRSKIFENLQNYRLLEYKSKPHTLFL
                10      20      30      40      50      60

gi|424 PQYTDADFILVVLGKATLTVLKSNDNRNSFNLERGDAIKLPAGTIAYLANRDDNEDLRVL
                70      80      90      100      110      120

```



```

CV127_                               SRNILEVN
                                   :.::::
gi|424 DLAIPVKNPGQLQSFLSGTQNQPSLLSGFSKNILEAAFNTNYEEIEKVLLEQQEQEPQH
               130       140       150       160       170       180

gi|424 RRSLKDRRQEINEENVIVKVSREQIEELSKNAKSSSKSVSSESGPFNLRSRNPIYSNKF
               190       200       210       220       230       240

gi|424 GKFFEITPEKNQQLQDLDFVNSVDIKEGSLLLPNYNSRAIVIVTVTEGKGFELVQQRN
               250       260       270       280       290       300

gi|424 ENQKGENDKEEQEEETSKQVQLYRAKLSPGDVFVIPAGHPVAINASSDLNLIGFGINAE
               310       320       330       340       350       360

gi|424 NNERNFLAGEEDNVISQVERPVKELAFPGSSHEVDRLLNQKQSYFANAQPLQRE
               370       380       390       400       410

>>gi|29539109 gi|29539109|emb|CAD87730.1| allergen Len c (418 aa)
  initn: 31 initl: 31 opt: 31 Z-score: 90.0 bits: 19.9 E(): 4.6
Smith-Waterman score: 31; 83.333% identity (100.000% similar) in 6 aa overlap
(1-6:151-156)

```

```

gi|295 SRSDQENPFIFKSNRFQTIYENENGHIRLLQRFDKRSKIFENLQNYRLLEYKSKPHTIFL
               10       20       30       40       50       60

gi|295 PQFTDADFILVVLGKAILTVLNSNDRNSFNLERGDTIKLPAGTIAYLANRDDNEDLRLV
               70       80       90       100       110       120

```

```

CV127_                               SRNILEVN
                                   :.::::
gi|295 DLAIPVNRPGQLQSFLSGTQNQPSFLSGFSKNILEAAFNTYEEIEKVLLEEQEQKSQH
               130       140       150       160       170       180

gi|295 RRSLRDKRQEITNEDVIVKVSREQIEELSKNAKSSSKSVSSESEPFNLRSRNPIYSNKF
               190       200       210       220       230       240

gi|295 GKFFEITPEKNPQLQDLDFVNSVEIKEGSLLLPNYNSRAIVIVTVNEGKGFELVQQRN
               250       260       270       280       290       300

gi|295 ENQQEQREENDEEEGQEEETTKQVQRYRARLSPGDVLVIPAGHPVAINASSDLNLIGFI
               310       320       330       340       350       360

gi|295 NAKNNQRNFLAGEEDNVISQIQRPVKELAFPGSSREVDRLLTNQKQSHFANAQPLQIE
               370       380       390       400       410

>>gi|32363467 gi|32363467|sp||P82946_4 [Segment 4 of 4] (15 aa)
  initn: 20 initl: 20 opt: 20 Z-score: 89.9 bits: 15.1 E(): 4.7
Smith-Waterman score: 20; 66.667% identity (66.667% similar) in 6 aa overlap
(3-8:5-10)

```

```

CV127_   SRNILEVN
         ::  ::
gi|323 GVLFNIQYVNYWFAP
         10

```

```

>>gi|169929 gi|169929|gb|AAB01374.1| beta-conglycinin st (639 aa)
  initn: 24 initl: 24 opt: 32 Z-score: 88.7 bits: 20.3 E(): 5.4

```

Smith-Waterman score: 32; 62.500% identity (100.000% similar) in 8 aa overlap (1-8:373-380)

```
gi|169 MMRARFPLLLGVVFLASVSVSFGIAYWEKQNP SHNKCLRSCNSEKDSYRNQACHARNL
      10      20      30      40      50      60
gi|169 LKVEEEEECEEGQIPRPRPQHPERERQQHGEKEEDEGEQPRPFPFPRPRQPHQEEHEQK
      70      80      90     100     110     120
gi|169 EEHEWHRKEEKHGGKGSEEEQDEREHPRPHQKKEEKHEWQHKQEKHQGKESEEEEEED
     130     140     150     160     170     180
gi|169 QDEDEEQDKESQESEGESQREPRRHKNKNPFHFN SKRFQTLFKNQYGHVRLQRFNKRS
     190     200     210     220     230     240
gi|169 QQLQNLRDYRILEFNSKPNLTLPHHADADYLIVILNGTAILTLVNDDDRDSYNLQSGDA
     250     260     270     280     290     300
gi|169 LRVPA GTTFYVVPNDNENLRMIAGTTFYVVPNDNENLRMITLAIPVNKPGRFESFFLS
     310     320     330     340     350     360

CV127_          SRNILEVN
          : : : : :
gi|169 STQAQQSYLQGF SKNILEASYDTKFEEINKVLFGREEGQQQGEERLQESVIVEISKQIR
     370     380     390     400     410     420
gi|169 ELSKHAKSSSRKTISSEDKPFLNLSRDP IYSNKLGLFEITQRNPQLRDL DVFLSVVDMN
     430     440     450     460     470     480
gi|169 EGALFLPHFN SKAIVVLVINEGEANIELVGIKEQQQRQQQEEQPLEVRKYRAELSEQDIF
     490     500     510     520     530     540
gi|169 VIPAGYPVMVNATSDLNFFAFGINAENNQRNFLAGSKDNVISQIP SQVQELAFPRSAKDI
     550     560     570     580     590     600
gi|169 ENLIK SQSESYFVDAQPPQKKEEGNKGRKGPLSSILRAF Y
     610     620     630
```

>>gi|46396598 gi|46396598|sp||P83834\_3 [Segment 3 of 3] (10 aa)  
 initn: 17 initl: 17 opt: 17 Z-score: 84.7 bits: 13.6 E(): 9.1  
 Smith-Waterman score: 17; 50.000% identity (100.000% similar) in 4 aa overlap (2-5:2-5)

```
CV127_ SRNILEVN
          : : .
gi|463 DRNLVHSATR
      10
```

>>gi|29163773 gi|29163773|emb|CAD80019.1| unnamed protei (65 aa)  
 initn: 23 initl: 23 opt: 23 Z-score: 84.2 bits: 16.2 E(): 9.7  
 Smith-Waterman score: 23; 60.000% identity (100.000% similar) in 5 aa overlap (1-5:47-51)

```
CV127_          SRNILEVN
          : : : .
gi|291 MVKLTQVAAILLIGAFFLIASTSIATQTSHPAKFHVEGEVYCNVCHSRNLINELSERMAG
      10      20      30      40      50      60
gi|291 AQVQL
```

8 residues in 1 query sequences  
307888 residues in 1386 library sequences  
Scomplib [34.26]  
start: Wed Mar 11 06:20:06 2009 done: Wed Mar 11 06:20:06 2009  
Total Scan time: 0.050 Total Display time: 0.000

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

## CV127\_977\_1003\_frame1

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

Please cite:

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

CV127\_977\_1003\_frame1, 9 aa

vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences

Expectation\_n fit:  $\rho(\ln(x)) = 1.3298 \pm 0.00265$ ;  $\mu = 9.6085 \pm 0.144$   
mean\_var=11.1184  $\pm$  3.194, 0's: 96 Z-trim: 96 B-trim: 214 in 1/34  
Lambda= 0.384639

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

The best scores are:

	opt	bits	E(1386)	%_id	%_sim	alen
gi 113575 gi 113575 sp P19121.2 ALB ( 615)	36	23.2	0.81	0.857	1.000	7
gi 34495286 gi 34495286 gb AAQ73490 ( 141)	28	17.7	8.3	0.571	0.857	7
gi 34495290 gi 34495290 gb AAQ73492 ( 141)	28	17.7	8.3	0.571	0.857	7
gi 34495278 gi 34495278 gb AAQ73486 ( 141)	28	17.7	8.3	0.571	0.857	7
gi 34495282 gi 34495282 gb AAQ73488 ( 141)	28	17.7	8.3	0.571	0.857	7
gi 34495284 gi 34495284 gb AAQ73489 ( 141)	28	17.7	8.3	0.571	0.857	7
gi 34495288 gi 34495288 gb AAQ73491 ( 141)	28	17.7	8.3	0.571	0.857	7
gi 34495274 gi 34495274 gb AAQ73484 ( 141)	28	17.7	8.3	0.571	0.857	7
gi 33772588 gi 33772588 gb AAQ54603 ( 141)	28	17.7	8.3	0.571	0.857	7
gi 2414158 gi 2414158 emb CAA96545. ( 160)	28	17.8	8.8	0.625	0.875	8

>>>CV127\_977\_1003\_frame1, 9 aa vs

/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

>gi|113575 gi|113575|sp|P19121.2|ALBU\_CHICK RecName: Fu (615 aa)  
initn: 36 initl: 36 opt: 36 Z-score: 103.5 bits: 23.2 E(): 0.81  
Smith-Waterman score: 36; 85.714% identity (100.000% similar) in 7 aa overlap  
(2-8:6-12)

CV127\_ RLIKFI FLK  
::: :::

gi|113 MKWVT LISFIF LFSSATSRNLQRFARDAEHKSEIAHRYNDLKEETFKAVAMITFAQYLQR  
10 20 30 40 50 60

gi|113 CSYEGLSKLVKDVVDLAQKCVANEDAPECSKPLPSIILDEICQVEKL RDSYGAMADCCSK  
70 80 90 100 110 120

gi|113 ADPERNECFLSFKVSQPDFVQPYQRPASDVICQEYQDNRVSFLGHFIYSVARRHPFLYAP  
130 140 150 160 170 180

gi|113 AILSFAVDFEHALQSCCKESDVGACLDTKEIVMREKAKGVSVKQYFCGILKQFGDRVFQ  
190 200 210 220 230 240

gi|113 ARQLIYLSQKYKAPFSEVSKFVHDSIGVHKECCEGDMVECMDDMARMSNLCSQQDVFS  
250 260 270 280 290 300

gi|113 GKI KCCEKPIVERSQCIMEAEFDEKPADLPSLVEKYIEDKEVCKSF EAGHDAFMAEFVY  
310 320 330 340 350 360

gi|113 EYSRRHPEFSIQLIMRIAKGYESLLEKCKCTDNPAECYANAQEQNLQH IKETQDVVKTNC  
370 380 390 400 410 420

gi|113 DLLHDHGEADFLKSILIRYTKMPQVPTDLLLETGKKMTTIGTKCCQLGEDRRMACSEGY  
430 440 450 460 470 480

gi|113 LSIVIHDTCRKQETTPINDNVSQCCS QLYANRRPCFTAMGVDTKYVPPFPNPMFSFDEK  
490 500 510 520 530 540

gi|113 LCSAPAEEREVGQMKLLINLIKRPQMTEEQIKTIADGFTAMVDKCKQSDINTCFGEEG  
550 560 570 580 590 600

gi|113 ANLIVQSRATLGIGA  
610

>>gi|34495286 gi|34495286|gb|AAQ73490.1| type 2 allergen (141 aa)  
initn: 23 init1: 23 opt: 28 Z-score: 85.4 bits: 17.7 E(): 8.3  
Smith-Waterman score: 28; 57.143% identity (85.714% similar) in 7 aa overlap  
(2-8:1-7)

CV127\_ RLIKFI FLK  
..::: :

gi|344 MMKFIVLFALNAVASAGKMKFKDCGKG EVTELDITDCSGDFCVLHRGKSVTLDAKFVAN  
10 20 30 40 50

gi|344 QDSAKATIKVLAKVAGTQIQVPGFDTDGCKIIKCP IKGDPIDFKYSGTIPAITPKIAE  
60 70 80 90 100 110

gi|344 VTAE LIGDHGILACGT VNGQVE  
120 130 140

>>gi|34495290 gi|34495290|gb|AAQ73492.1| type 2 allergen (141 aa)  
initn: 23 init1: 23 opt: 28 Z-score: 85.4 bits: 17.7 E(): 8.3  
Smith-Waterman score: 28; 57.143% identity (85.714% similar) in 7 aa overlap  
(2-8:1-7)

CV127\_ RLIKFI FLK  
..::: :

gi|344 MMKFIVLFALVAVASAGKMKFKDCGKG EVTELDITDCSGDFCVLHRGKSVTLDAKFVAN  
10 20 30 40 50

gi|344 QDSAKATIKVLAKVVGTQIQVPGFDTDGCKIIKCP IKGDPIDFKYSGTIPAITPKIAE  
60 70 80 90 100 110

gi|344 VTAE LIGDHGILACGT VNGTVE  
120 130 140

>>gi|34495278 gi|34495278|gb|AAQ73486.1| type 2 allergen (141 aa)

initn: 23 initl: 23 opt: 28 Z-score: 85.4 bits: 17.7 E(): 8.3  
Smith-Waterman score: 28; 57.143% identity (85.714% similar) in 7 aa overlap  
(2-8:1-7)

CV127\_ RLIKFI FLK

..::: :

gi|344 MMKFIVL FALDAVASAGKMKFKDCGKGEVTELDITDCSGDFCVLHRGKSVTLDAKFVAN  
10 20 30 40 50

gi|344 QDSAKATIKVLAKVAGTQIQVPGFDTDGCKIIKCPKKGDPIDFKYSGTIPAITPKIKAE  
60 70 80 90 100 110

gi|344 VTAE LIGDHGILACGTVNGQVE  
120 130 140

>>gi|34495282 gi|34495282|gb|AAQ73488.1| type 2 allergen (141 aa)  
initn: 23 initl: 23 opt: 28 Z-score: 85.4 bits: 17.7 E(): 8.3  
Smith-Waterman score: 28; 57.143% identity (85.714% similar) in 7 aa overlap  
(2-8:1-7)

CV127\_ RLIKFI FLK

..::: :

gi|344 MMKFIVL FALVAVASAGNMKFKDCGKGEVTELDITDCSGDFCVLHRGKSVTLDAKFVAN  
10 20 30 40 50

gi|344 QDSAKATIKVLAKVAGTQIQVPGFDTDGCKIIKCPKKGDPIDFKYSGTIPAITPKIKAE  
60 70 80 90 100 110

gi|344 VTAE LIGDHGILACGTVNGQVE  
120 130 140

>>gi|34495284 gi|34495284|gb|AAQ73489.1| type 2 allergen (141 aa)  
initn: 23 initl: 23 opt: 28 Z-score: 85.4 bits: 17.7 E(): 8.3  
Smith-Waterman score: 28; 57.143% identity (85.714% similar) in 7 aa overlap  
(2-8:1-7)

CV127\_ RLIKFI FLK

..::: :

gi|344 MMKFIVL FALVAVASAGKMKFKDCGKGEVTELDITDCSGGFCVLHRGKSVTLDAKFVAN  
10 20 30 40 50

gi|344 QDSAKATIKVLAKVAGTQIQVPGFDTDGCKIIKCPKKGDPIDFKYSGTIPAITPKIKAE  
60 70 80 90 100 110

gi|344 VTAE LIGDHGILACGTVNGQVE  
120 130 140

>>gi|34495288 gi|34495288|gb|AAQ73491.1| type 2 allergen (141 aa)  
initn: 23 initl: 23 opt: 28 Z-score: 85.4 bits: 17.7 E(): 8.3  
Smith-Waterman score: 28; 57.143% identity (85.714% similar) in 7 aa overlap  
(2-8:1-7)

CV127\_ RLIKFI FLK

..::: :

gi|344 MMKFIVL FALVAVASAGKMKFKDCGKGEVTELDITDCSGDFCVLHRGKSVTLDAKFVAN  
10 20 30 40 50

gi|344 QDSAKATIKVLAKVAGTQIQVPGFDTDGCKIIKCPKKGDPIDFKYSGTIPAITPKIKAE

```

        60          70          80          90          100          110

gi|344 VTAEIIGDHGILACGTVNGQVE
    120          130          140

>>gi|34495274 gi|34495274|gb|AAQ73484.1| type 2 allergen (141 aa)
    initn: 23 initl: 23 opt: 28 Z-score: 85.4 bits: 17.7 E(): 8.3
Smith-Waterman score: 28; 57.143% identity (85.714% similar) in 7 aa overlap
(2-8:1-7)

CV127_ RLIKFIPLK
    ..::: :
gi|344 MMKFIVLFALIAVASAGKMKFKDCGKGEVTELDITDCSGDFCVLHRGKSVTLDAKFVAN
    10          20          30          40          50

gi|344 QDSAKATIKVLAKVAGTQIQVPGFDTGCKIIKCPKIKGDPIDFKYSGTIPAITPKIAE
    60          70          80          90          100          110

gi|344 VTAEIIGDHGILACGTVNGQVE
    120          130          140

>>gi|33772588 gi|33772588|gb|AAQ54603.1| Gly d 2.03 [Gly (141 aa)
    initn: 23 initl: 23 opt: 28 Z-score: 85.4 bits: 17.7 E(): 8.3
Smith-Waterman score: 28; 57.143% identity (85.714% similar) in 7 aa overlap
(2-8:1-7)

CV127_ RLIKFIPLK
    ..::: :
gi|337 MMKFIVLFALVAVASAGKMKFKDCGKGEVTELDITDCSGDFCVLHRGKSVTLDAKFVAN
    10          20          30          40          50

gi|337 QDSAKATIKVLAKVAGTQIQVPGFDTGCKIIKCPKIKGDPIDFKYSGTIPAITPKIAE
    60          70          80          90          100          110

gi|337 VTAEIIGDHGILACGTVNGQVE
    120          130          140

>>gi|2414158 gi|2414158|emb|CAA96545.1| major allergen B (160 aa)
    initn: 28 initl: 28 opt: 28 Z-score: 84.9 bits: 17.8 E(): 8.8
Smith-Waterman score: 28; 62.500% identity (87.500% similar) in 8 aa overlap
(1-8:18-25)

CV127_ RLIKFIPLK
    ::: :
gi|241 MGVFNYESVETTSVIPARLFKAFFLDGDNLFKVPAPQAISSEVENIEGNGGPGTIKKISFP
    10          20          30          40          50          60

gi|241 EGFPFRYVKDRVDEVDHTNFKYSYSVIEGGPVGDTLEKISNEIKIVATPDGGSILKISNK
    70          80          90          100          110          120

gi|241 YHTKGDHEVKEEQIKASKEMGETLLRAVESYLLAHSDAYN
    130          140          150          160

9 residues in 1 query sequences
307888 residues in 1386 library sequences
Scomplib [34.26]
```

start: Wed Mar 11 05:47:21 2009 done: Wed Mar 11 05:47:21 2009  
Total Scan time: 0.050 Total Display time: 0.010

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

## CV127\_984\_1091\_frame2

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

Please cite:

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

CV127\_984\_1091\_frame2, 36 aa

vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences

Expectation\_n fit:  $\rho(\ln(x)) = 5.0357 \pm 0.00333$ ;  $\mu = -2.2226 \pm 0.171$

mean\_var=33.9305  $\pm$  8.567, 0's: 18 Z-trim: 18 B-trim: 14 in 1/42

Lambda= 0.220180

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

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

The best scores are:

			opt	bits	E(1386)	%_id	%_sim	alen
gi 2129805	gi 2129805 pir	S65143 p ( 80)	50	21.9	1	0.375	0.750	24
gi 59800145	gi 59800145 sp	P69198.1 ( 83)	50	21.9	1.1	0.375	0.750	24
gi 59800146	gi 59800146 sp	P69199.1 ( 83)	50	21.9	1.1	0.375	0.750	24
gi 25090951	gi 25090951 sp	P83181_ ( 14)	38	18.4	2.1	0.417	0.833	12
gi 41057621	gi 41057621 gb	AAR98813 ( 231)	51	22.1	2.6	0.421	0.684	19
gi 2129801	gi 2129801 pir	S65144 p ( 83)	45	20.3	3.2	0.350	0.800	20
gi 2129802	gi 2129802 pir	S65145 p ( 83)	45	20.3	3.2	0.350	0.800	20
gi 3337403	gi 3337403 gb	AAD05375.1 ( 84)	45	20.3	3.2	0.375	0.667	24
gi 729979	gi 729979 sp	P39673.1 MAG ( 341)	50	21.7	5.1	0.625	0.875	16
gi 29465668	gi 29465668 gb	AAL92871 ( 86)	42	19.4	6.4	0.333	0.708	24
gi 1916805	gi 1916805 gb	AAC27724.1 ( 388)	49	21.3	7.3	0.429	0.667	21

>>>CV127\_984\_1091\_frame2, 36 aa vs

/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

>>gi|2129805 gi|2129805|pir||S65143 pollen allergen grou (80 aa)

initn: 50 initl: 50 opt: 50 Z-score: 101.8 bits: 21.9 E(): 1

Smith-Waterman score: 50; 37.500% identity (75.000% similar) in 24 aa overlap (8-31:21-44)

```

              10      20      30
CV127_      LNLYSNKTQFSSLNEPAETLISITNSDLKRSHGSH
              : . : : . . . : : : : : : :
gi|212  ETERAEHDRIFKKFDANGDGKISASELGDALKNLGVTHTDDIKRMMAEIDTDGDGYISYQ
              10      20      30      40      50      60

gi|212  EFSDFASANRGLMKDVAKIF
              70      80

```

>>gi|59800145 gi|59800145|sp|P69198.1|POLC2\_BRANA RecNam (83 aa)

initn: 50 initl: 50 opt: 50 Z-score: 101.5 bits: 21.9 E(): 1.1

Smith-Waterman score: 50; 37.500% identity (75.000% similar) in 24 aa overlap (8-31:24-47)

```

              10      20      30
CV127_      LNLYSNKTQFSSLNEPAETLISITNSDLKRSHGSH

```

```

: . . : . . . . : . . . . :
gi|598 MADATEKAEHDRIFKKFDANGDGKISASELGDALKNLGSVTHDDIKRMMAEIDTDGDGYI
      10      20      30      40      50      60

gi|598 SYQEFSDFASANRGLMKDVAKIF
      70      80

>>gi|59800146 gi|59800146|sp|P69199.1|POLC2_BRARA RecNam (83 aa)
  initn: 50 init1: 50 opt: 50 Z-score: 101.5 bits: 21.9 E(): 1.1
Smith-Waterman score: 50; 37.500% identity (75.000% similar) in 24 aa overlap
(8-31:24-47)

      10      20      30
CV127_ LNLyssNKTQFSSLNEPAETLISITNSDLKRSHGSH
      : . . : . . . . : . . . . :
gi|598 MADATEKAEHDRIFKKFDANGDGKISASELGDALKNLGSVTHDDIKRMMAEIDTDGDGYI
      10      20      30      40      50      60

gi|598 SYQEFSDFASANRGLMKDVAKIF
      70      80

>>gi|25090951 gi|25090951|sp||P83181_4 [Segment 4 of 4] (14 aa)
  initn: 38 init1: 38 opt: 38 Z-score: 96.2 bits: 18.4 E(): 2.1
Smith-Waterman score: 38; 41.667% identity (83.333% similar) in 12 aa overlap
(16-27:1-12)

      10      20      30
CV127_ LNLyssNKTQFSSLNEPAETLISITNSDLKRSHGSH
      : : : . . : . . . .
gi|250 EPAEEFTTISDAVK
      10

>>gi|41057621 gi|41057621|gb|AAR98813.1| glutathione-S-t (231 aa)
  initn: 51 init1: 51 opt: 51 Z-score: 94.3 bits: 22.1 E(): 2.6
Smith-Waterman score: 51; 42.105% identity (68.421% similar) in 19 aa overlap
(14-32:162-180)

gi|410 MSDKPSELAVQKLVLFAVKGTATSTHNTVRPLILLDELGVPHEIYVDRVSAPWFTEINP
      10      20      30      40      50      60

gi|410 HRMPVPVILEKSPDGRDTRLAWESTSTLMYIADAYDKDGTFGGRNVQESSDINNWLTLHTA
      70      80      90      100      110      120

      10      20      30
CV127_ LNLyssNKTQFSSLNEPAETLISITNSDLKRS
      : : : . . . . : . . : . :
gi|410 ALGPTAKYWLYFYKLHPKLPKTIKLRSNITVQYDILERRLNEPGQQYLAWLNEKFKRS
      130      140      150      160      170      180

CV127_ HGSH

gi|410 SYNRRHCYASLCYEKYRRVVRAGVKVAQTARVVCYPYGGDTRRGVWPARKST
      190      200      210      220      230

>>gi|2129801 gi|2129801|pir||S65144 pollen allergen grou (83 aa)
  initn: 45 init1: 45 opt: 45 Z-score: 92.9 bits: 20.3 E(): 3.2
Smith-Waterman score: 45; 35.000% identity (80.000% similar) in 20 aa overlap
(12-31:28-47)

      10      20      30
CV127_ LNLyssNKTQFSSLNEPAETLISITNSDLKRSHGSH

```



```

      . . . . . : : : : : : : :
gi|212 MADATEKTEHDRFFKKFDANGDGTISSTELGDALKNLGSVTHDDIKRMMAEIDTDGDGFI
      10      20      30      40      50      60

gi|212 SYQEFSDFAKANRGLMKDVAKIF
      70      80

>>gi|2129802 gi|2129802|pir||S65145 pollen allergen grou (83 aa)
  initn: 45 initl: 45 opt: 45 Z-score: 92.9 bits: 20.3 E(): 3.2
Smith-Waterman score: 45; 35.000% identity (80.000% similar) in 20 aa overlap
(12-31:28-47)

      10      20      30
CV127_      LNLyssNKTQFSSSLNEPAETLISITNSDLKRSHGSH
      . . . . . : : : : : : : :
gi|212 MADATEKAEHDRFFKKFDANGDGTISSTELGDALKNLGSVTHDDIKRMMAEIDTDGDGFI
      10      20      30      40      50      60

gi|212 SYQEFSDFAKANRGLMKDVAKIF
      70      80

>>gi|3337403 gi|3337403|gb|AAD05375.1| calcium-binding p (84 aa)
  initn: 30 initl: 30 opt: 45 Z-score: 92.8 bits: 20.3 E(): 3.2
Smith-Waterman score: 45; 37.500% identity (66.667% similar) in 24 aa overlap
(8-31:25-48)

      10      20      30
CV127_      LNLyssNKTQFSSSLNEPAETLISITNSDLKRSHGSH
      : . : : : . : : : : . : :
gi|333 MADDPQEVAEHERIFKRFDANGDGKISSSELGETLKTGLSVTPPEIQRMMAEIDTDGDGF
      10      20      30      40      50      60

gi|333 ISFEEFTVFARANRGLVKDVAKIF
      70      80

>>gi|729979 gi|729979|sp|P39673.1|MAG_DERFA Allergen Mag (341 aa)
  initn: 38 initl: 38 opt: 50 Z-score: 89.2 bits: 21.7 E(): 5.1
Smith-Waterman score: 50; 62.500% identity (87.500% similar) in 16 aa overlap
(6-20:295-310)

gi|729 FVMKREPLRFRDITVEGNENAYIKNGKLHLSLMDPSTLSLVTKADGKIDMTVDLISPVTK
      10      20      30      40      50      60

gi|729 RASLKIDSKKYNLFHEGELSASIVNPRLSWHQYTKRDSREYKSDVELSLRSSDIALKITM
      70      80      90      100      110      120

gi|729 PDYNSKIHYSRQGDQINMDIDGTLIEGHAQGTIREGKIHIGRQTDFEIESNYRYEDGKL
      130      140      150      160      170      180

gi|729 IIEPVKSENGKLEGVLSRKVPShLTLETPrVKMNMKYDRYAPVKVFKLDYDGIHFekHTD
      190      200      210      220      230      240

      10
CV127_      LNLyssNKTQF
      : : : : :
gi|729 IEYEPGVRYKIIIGNGLKDDGRHYSIDVQGIPRKAFNLADLMDFKLVSKPEDSNKAQF
      250      260      270      280      290      300

      20      30
CV127_ S-SLNEPAETLISITNSDLKRSHGSH
      : : : : :
gi|729 SYTFNEYTETEEYEFDPHrayyVnWLSSIRKYIQNFIVEDN

```

```

310      320      330      340

>>gi|29465668 gi|29465668|gb|AAL92871.1| pollen allergen (86 aa)
  initn: 31 initl: 31 opt: 42 Z-score: 87.4 bits: 19.4 E(): 6.4
Smith-Waterman score: 42; 33.333% identity (70.833% similar) in 24 aa overlap
(8-31:27-50)

```

```

              10      20      30
CV127_          LNLSSNKTQFSSLNEPAETLISITNSDLKRSHGSH
              :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|294 MAAEDTPQDIADRERIFKRFDTNNGDKISSSELGDALKTLGSVTPDEVRRMMAEIDTDGD
              10      20      30      40      50      60

gi|294 GFISFDEFTDFARANRGLVKDVS KIF
              70      80

```

```

>>gi|1916805 gi|1916805|gb|AAC27724.1| latex patatin hom (388 aa)
  initn: 43 initl: 43 opt: 49 Z-score: 86.4 bits: 21.3 E(): 7.3
Smith-Waterman score: 49; 42.857% identity (66.667% similar) in 21 aa overlap
(15-35:278-298)

```

```

gi|191 MATGSTTLTQGKKITVLSIDGGGIRGIIPGIILASLESKLQDLDPDARIADYFDIIAGT
              10      20      30      40      50      60

gi|191 STGGLITTMLTAPNEDKKPMYQAKDIKDFYLENCPKIFPKESRDNYDPIHSIGPIYDGEY
              70      80      90      100     110     120

gi|191 LRELNNLLKDLTVKDTSTDVVIPTFDIKLLLPVIFPSDDAKCNALKNARLADVCI STSA
              130     140     150     160     170     180

gi|191 APVLLPAHSFTTEDDKNIHTFELIDGGVAATNPTLLALTHIRNEIIRQNPRFIGANLTES
              190     200     210     220     230     240

              10      20      30
CV127_          LNLSSNKTQFSSLNEPAETLISITNSDLKRSHGSH
              :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|191 KSRLVLSLGTGKSEYKEKYNADMTSKWRLYNWALYNGNSPAVDIFSNASSDMVDSHLSAL
              250     260     270     280     290     300

gi|191 FKSLDCEDYYLRIQDDTLTGEESSGHIATEENLQRLVEIGTELLEKQESRINLDTGRLES
              310     320     330     340     350     360

gi|191 IPGAPTNEAAIAKFAKLLSEERKLRQLK
              370     380

```

```

36 residues in 1 query sequences
307888 residues in 1386 library sequences
Scomplib [34.26]
start: Wed Mar 11 05:49:30 2009 done: Wed Mar 11 05:49:31 2009
Total Scan time: 0.080 Total Display time: 0.010

```

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

## CV127\_988\_1023\_frame0

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

Please cite:

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

CV127\_988\_1023\_frame0, 12 aa  
vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences  
Expectation\_n fit:  $\rho(\ln(x)) = 0.5017 \pm 0.003$ ;  $\mu = 20.6451 \pm 0.161$   
mean\_var=36.2052 $\pm$ -12.708, 0's: 100 Z-trim: 100 B-trim: 10 in 1/32  
Lambda= 0.213152

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2  
join: 36, opt: 20, open/ext: -10/-2, width: 16  
The best scores are: opt bits E(1386) %\_id %\_sim alen  
gi|4538529 gi|4538529|emb|CAB39376. ( 81) 51 19.4 1.9 0.600 0.800 10

>>>CV127\_988\_1023\_frame0, 12 aa vs  
/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

>>gi|4538529 gi|4538529|emb|CAB39376.1| Cop c1 allergen (81 aa)  
initn: 51 init1: 51 opt: 51 Z-score: 96.8 bits: 19.4 E(): 1.9  
Smith-Waterman score: 51; 60.000% identity (80.000% similar) in 10 aa overlap  
(3-12:29-38)

```

                                10
CV127_                          IYIPQIKPNFHP
                                .:.:.:  ::
gi|453 RFLPSSSHLNPQHLPWLVHPAPVLLLPVLPQLKPVAPHLLLLPLDTTTLHMPPLLLQLQL
                                10      20      30      40      50      60

gi|453 PPLLSQGNPACSPKWLQLLVP
                                70      80
```

12 residues in 1 query sequences  
307888 residues in 1386 library sequences  
Scomplib [34.26]  
start: Wed Mar 11 05:44:43 2009 done: Wed Mar 11 05:44:43 2009  
Total Scan time: 0.060 Total Display time: 0.000

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

## CV127\_rc\_5681\_5776\_frame1

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

Please cite:

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

CV127\_rc\_5681\_5776\_frame1, 32 aa  
vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences  
Expectation\_n fit:  $\rho(\ln(x)) = 4.6650 \pm 0.00299$ ;  $\mu = -1.0939 \pm 0.154$   
mean\_var=25.4290 $\pm$ -6.795, 0's: 21 Z-trim: 21 B-trim: 5 in 1/39  
Lambda= 0.254337

```

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 20, open/ext: -10/-2, width: 16
The best scores are:
      opt bits E(1386) %_id %_sim alen
gi|49523394 gi|49523394|emb|CAE5283 ( 377) 50 23.0      2 0.292 0.750 24
gi|1314736 gi|1314736|gb|AAA99805.1 ( 232) 47 22.0      2.4 0.412 0.588 17
gi|163638970 gi|163638970|gb|ABY281 ( 259) 47 22.0      2.8 0.412 0.588 17
gi|2507248 gi|2507248|sp|P49275.2|D ( 259) 47 22.0      2.8 0.412 0.588 17
gi|22595342 gi|22595342|gb|AAN02510 ( 244) 46 21.6      3.3 0.600 0.800 10
gi|4235093 gi|4235093|gb|AAD13106.1 ( 804) 51 23.1      3.8 0.409 0.682 22
gi|511476 gi|511476|gb|AAA19973.1| ( 261) 43 20.5      7.7 0.500 0.800 10

```

```

>>>CV127_rc_5681_5776_frame1, 32 aa vs
/n/na4/bioinfo/refsets/PUB_AllergenOnlineJan2009_V9.fasta library

>>gi|49523394 gi|49523394|emb|CAE5283.1| polygalacturon (377 aa)
initn: 41 initl: 41 opt: 50 Z-score: 96.4 bits: 23.0 E(): 2
Smith-Waterman score: 50; 29.167% identity (75.000% similar) in 24 aa overlap
(3-26:68-91)

```

```

gi|495 RGVQSSGSVFNVNDYGAKGAGDISQAVMKAWKAACASQGPSTVLIPKGNYNMGEVAMQGP
      10          20          30          40          50          60

      10          20          30
CV127_ LLFRSELVIEIRVSAGSFKDENWVLFEEYKFN
      :. . :.. .. :.:.:.:.: :
gi|495 CKGSKIGFQIDGVVKAPADPSKFKSDGWVSFYRIDGLTVSGTGTLDGQGQTAWAKNNCDK
      70          80          90          100          110          120

gi|495 NPNCKHAAMNLRFDLKHAMVRDITSLNSKMFHINVLECEDITFQHVTVTAPGTSINTDG
      130          140          150          160          170          180

gi|495 IHVGISKGVTTITNTKIATGDDCISIGPGSQNVTTITQVNCGPGHGIGISGLGRYNNEKEVR
      190          200          210          220          230          240

gi|495 GITVKGCTFSGTMNGVRVKTWPNSPPGAATDLTFQDLTMNNVQNPVILDQEYCPYGQCSR
      250          260          270          280          290          300

gi|495 QAPSRIKLSNINFNNIRGTSTGKVAVVIACSHGMPCSNMKIGEINLSYRGAGGPATSTCS
      310          320          330          340          350          360

gi|495 NVKPTFSGKQVPAIKCA
      370

```

```

>>gi|1314736 gi|1314736|gb|AAA99805.1| Der f 3 mite alle (232 aa)
initn: 46 initl: 46 opt: 47 Z-score: 95.0 bits: 22.0 E(): 2.4
Smith-Waterman score: 47; 41.176% identity (58.824% similar) in 17 aa overlap
(16-32:27-43)

```

```

      10          20          30
CV127_ LLFRSELVIEIRVSAGSFKDENWVLFEEYKFN
      :. . :.:.:.:.: :
gi|131 IVGGVKAKAGDCPYQISLQSSSHFCGGSILDEYWILTAACVNGQSAKKLSIRYNTLKHA
      10          20          30          40          50          60

gi|131 SGGEKIQVAEIIYQHENYDSMTIDNDVALIKLKTTPMTLDQTNAPVPLPPQGSVDKVKDKI
      70          80          90          100          110          120

gi|131 RVSGWGYLQEGSYSLPSELQRVDIDVVSREQCDQLYSKAGADVSENMICGGDVANGGVDS
      130          140          150          160          170          180

gi|131 CQGDSSGGPVVDIATKQIVGIVSWGYGCCARKGYPGVYTRVGNFVDWIESKRSQ

```

```

190      200      210      220      230
>>gi|163638970 gi|163638970|gb|ABY28115.1| Der f 3 aller (259 aa)
  initn: 46 initl: 46 opt: 47 Z-score: 94.0 bits: 22.0 E(): 2.8
Smith-Waterman score: 47; 41.176% identity (58.824% similar) in 17 aa overlap
(16-32:54-70)

```

```

                                10      20
CV127_                        LLFRSELVIEIRVSAGSFKDEN
                                ... ..
gi|163 MMILTIVVLLAANILATPILPSSPNATIVGGVKAQAGDCPYQISLQSSSHFCGGSILDEY
      10      20      30      40      50      60

      30
CV127_ WVLFEYYKFN
      ... ..
gi|163 WILTAAHCVNGQSAKKLSIRNNTLKHASGGEKIQVAEIIYQHENVDSMTIDNDVALIKLKT
      70      80      90     100     110     120

gi|163 PTTLDQTNAPVPLPAQGSQDVKGDKIRVSGWGYLQEGSYSLPSELQRVDDIDVVSREQCD
      130     140     150     160     170     180

gi|163 QLYSKAGADVSENMICGGDVANGGVDSCQGDSSGPVVDIATKQIVGIVSWGYGCAKRGYP
      190     200     210     220     230     240

gi|163 GVYTRVGNFVDWIESKRSQ
      250

```

```

>>gi|2507248 gi|2507248|sp|P49275.2|DERF3_DERFA RecName: (259 aa)
  initn: 46 initl: 46 opt: 47 Z-score: 94.0 bits: 22.0 E(): 2.8
Smith-Waterman score: 47; 41.176% identity (58.824% similar) in 17 aa overlap
(16-32:54-70)

```

```

                                10      20
CV127_                        LLFRSELVIEIRVSAGSFKDEN
                                ... ..
gi|250 MMILTIVVLLAANILATPILPSSPNATIVGGVKAQAGDCPYQISLQSSSHFCGGSILDEY
      10      20      30      40      50      60

      30
CV127_ WVLFEYYKFN
      ... ..
gi|250 WILTAAHCVNGQSAKKLSIRYNTLKHASGGEKIQVAEIIYQHENVDSMTIDNDVALIKLKT
      70      80      90     100     110     120

gi|250 PMTLDQTNAPVPLPAQGSQDVKGDKIRVSGWGYLQEGSYSLPSELQRVDDIDVVSREQCD
      130     140     150     160     170     180

gi|250 QLYSKAGADVSENMICGGDVANGGVDSCQGDSSGPVVDVATKQIVGIVSWGYGCAKRGYP
      190     200     210     220     230     240

gi|250 GVYTRVGNFVDWIESKRSQ
      250

```

```

>>gi|22595342 gi|22595342|gb|AAN02510.1|AF409110_1 serin (244 aa)
  initn: 46 initl: 46 opt: 46 Z-score: 92.5 bits: 21.6 E(): 3.3
Smith-Waterman score: 46; 60.000% identity (80.000% similar) in 10 aa overlap
(16-25:51-60)

```

```

                                10      20
CV127_                        LLFRSELVIEIRVSAGSFKDENWVL
                                ... ..

```

gi|225 MIILCTPLIILAIILNATRNIPLGIVGGSNASPGDAVYQIALQSASHFCGGSILDEYWIL  
10 20 30 40 50 60

30  
CV127\_ FEEYKFN

gi|225 TAAHCVDGQTVSKLIRSKVLGEKISVSKIFAHEKYDSRLLDNDIALIKLKSPRLNSKNAR  
70 80 90 100 110 120

gi|225 VLPGSDVVKDGQVQSVWGYLEEGSYSLPPELRRVDIGGASRKECNELYSKVNAEVTDNMI  
130 140 150 160 170 180

gi|225 CGGDVANGGKDCQGDSSGGPLVDVKNNQVVGNSWGYGCERKGYPGVYTRVGNFIDWIES  
190 200 210 220 230 240

gi|225 KRPQ

>>gi|4235093 gi|4235093|gb|AAD13106.1| beta-xylosidase [ (804 aa)  
initn: 41 initl: 41 opt: 51 Z-score: 91.4 bits: 23.1 E(): 3.8  
Smith-Waterman score: 51; 40.909% identity (68.182% similar) in 22 aa overlap  
(9-26:749-770)

gi|423 MAHSMRPPVAATAAALLALALPQALAQANTSIVDYNVEANPDLYPLCIETIPLSFPDCQN  
10 20 30 40 50 60

gi|423 GPLRSHLICDESATPYDRAASLISLFTLDELIAANTGNTGLGVSRLGLPAYQVWSEALHGL  
70 80 90 100 110 120

gi|423 DRANFSDSGSYNWATSFPPQILTTAALNRTLHQAISIISTQGRAFNAGRYGLDVYAPN  
130 140 150 160 170 180

gi|423 INTFRHPVWGRGQETPGEDVSLAAVYAYEYITGIQGPDPDSNLKLAATAKHYAGYDIENW  
190 200 210 220 230 240

gi|423 HNHSRLGNDMNITQQDLSEYYTPQFHVAARDAKVHVSVMCAYNVAVDGPACADSYFLQTLL  
250 260 270 280 290 300

gi|423 RDTFGFVDHGYVSSDCDAAYNIYNPHGYASSQAAAAAEAILAGTDIDCGTTYQWHLNESI  
310 320 330 340 350 360

gi|423 TAGDLSRDDIEKGVIRLYTTLVQAGYFDSNTTKANNPYRDLTWSVDLETDAWNISYQAAT  
370 380 390 400 410 420

gi|423 QGIVLLKNSNNVLPLTEKAYPPSNTTVALIGPWANATTQLLGNYGNAPYMISSPRAAFEE  
430 440 450 460 470 480

gi|423 AGYKVNFAEGTGISSTSTSGFAAALSAARSADVIIYAGGIDNTLEAEALDRESIAWPGNQ  
490 500 510 520 530 540

gi|423 LDLIQKLASSAGSKPLIVLQMGQQVDSSSLKNNTNVTALLWGGYPGQSGGFALRDIITG  
550 560 570 580 590 600

gi|423 KKNPAGRLVTTQYPASYAEFFPATDMNLRPEGDNPGQTYKWTGEAVYEFHGHLFYTTFA  
610 620 630 640 650 660

gi|423 ESSSNTTTKEVKLNIQDILSQTHEELASITQLPVLNFTANIKNTGKLESYDTAMVFANTS  
670 680 690 700 710 720

10 20 30  
CV127\_ LLFRSELVIEIRVSAGSF----KDENWVLFEEYKFN  
... : .::: .: .:::

```
gi|423 DAGPAPYPVKWLVGWDR LGDVKVGETREL RVPVEVGSFARVNEDGDWVLFPGTFELALNL
      730      740      750      760      770      780

gi|423 ERKVRVKVVLEGE EEEVVLKWP GKE
      790      800

>>gi|511476 gi|511476|gb|AAA19973.1| Der p 3 allergen (261 aa)
  initn: 36 initl: 36 opt: 43 Z-score: 86.0 bits: 20.5 E(): 7.7
Smith-Waterman score: 43; 50.000% identity (80.000% similar) in 10 aa overlap
(16-25:56-65)

CV127_
      10      20
      LLFRSELVIEIRVSAGSFKD
      :...:
gi|511 MIIYNILIVLLLAIN TLANPILPASPNATIVGGEKALAGECPYQISLQSSSHFCGGTILD
      10      20      30      40      50      60

      30
CV127_ ENWVLFEEYKFN
      :...:
gi|511 EYWILTAAHCVAGQ TASKLSIRYNSLKHSLGGEKISVAKIFAHEKYDSYQIDNDIALIKL
      70      80      90      100      110      120

gi|511 KSPMKLNQKNAKAVGLPAKGSDVKVG DQVRVSGWGYLEEGSYSLPSELRRVDIAVVS RKE
      130      140      150      160      170      180

gi|511 CNELYSKANA EVDNMICGGDVANGGKDSCQGD SGGPVVDVKNNQVVGIVSWGYGCARKG
      190      200      210      220      230      240

gi|511 YPGVYTRVGNFIDWIESKRSQ
      250      260
```

32 residues in 1 query sequences  
307888 residues in 1386 library sequences  
Scomplib [34.26]  
start: Wed Mar 11 05:51:51 2009 done: Wed Mar 11 05:51:51 2009  
Total Scan time: 0.080 Total Display time: 0.000

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

## CV127\_rc\_5706\_5783\_frame2

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

Please cite:

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

CV127\_rc\_5706\_5783\_frame2, 26 aa

vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences

Expectation\_n fit:  $\rho(\ln(x)) = 3.3435 \pm 0.00275$ ;  $\mu = 2.9372 \pm 0.142$

mean\_var=21.4432  $\pm$  5.656, 0's: 37 Z-trim: 37 B-trim: 0 in 0/41

Lambda= 0.276968

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

The best scores are:

	gi	emb	(	opt	bits	E(1386)	%_id	%_sim	alen
	gi 3201547	gi 3201547 emb CAB01591.	( 326)	44	22.6	1.9	0.348	0.739	23
	gi 27526732	gi 27526732 emb CAD2406	( 295)	42	21.7	3	0.333	0.778	18
	gi 256427	gi 256427 gb AAB23463.1	( 439)	43	22.2	3.3	0.429	0.667	21
	gi 14575525	gi 14575525 emb CAC4288	( 295)	41	21.3	4	0.400	0.733	15
	gi 1359600	gi 1359600 emb CAA64868.	( 316)	40	21.0	5.6	0.545	0.727	11
	gi 14423650	gi 14423650 sp Q9U1G2.1	( 216)	38	20.1	6.8	0.320	0.640	25
	gi 119390336	gi 119390336 pdb 2J23	( 121)	36	19.3	7	0.400	0.800	10
	gi 17907758	gi 17907758 dbj BAB7944	( 133)	36	19.3	7.6	0.400	0.733	15

>>>CV127\_rc\_5706\_5783\_frame2, 26 aa vs  
/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

>>gi|3201547 gi|3201547|emb|CAB01591.1| endochitinase [P (326 aa)  
initn: 39 initl: 39 opt: 44 Z-score: 96.9 bits: 22.6 E(): 1.9  
Smith-Waterman score: 44; 34.783% identity (73.913% similar) in 23 aa overlap  
(4-26:280-301)

```
gi|320 MVYCTASLPLLLLLLVGLLAGEAFAEQCGRQAGGALCPGGLCCSQFGWCGSTSDYCGPTC
      10          20          30          40          50          60

gi|320 QSQCGGVTPSPGGGVASLISQSVFNQMLKHRNDAAQAKGFYTYNAFIAAANSFNFGFASV
      70          80          90         100         110         120

gi|320 GDTATRKREIAAFLAQTSHETTGGWATAPDGPYAWGYCFLKEQGNPPDYCVPTAQWPCAP
      130         140         150         160         170         180

gi|320 GKYYGRGPIQISYNNYGPAGRAIGYDLINNPDAVATDPVISFKTALWFWMTQSPKPS
      190         200         210         220         230         240
```

```

                                10          20
CV127_                        SKLGFQQVRLRMKIGFYLRNINLI
                                :... .. :...
gi|320 CHNVITGRWTPSAADRAAGRLPGYGVITNIINGGIECGKGFND-KVADRIGFYKRYCDLL
      250         260         270         280         290
```

```
CV127_ NL
      ..
gi|320 GVSYGSNLDCYNQRSFGVSTNPLAASS
      300         310         320
```

>>gi|27526732 gi|27526732|emb|CAD24068.1| class I chitin (295 aa)  
initn: 35 initl: 35 opt: 42 Z-score: 93.3 bits: 21.7 E(): 3  
Smith-Waterman score: 42; 33.333% identity (77.778% similar) in 18 aa overlap  
(9-26:261-278)

```
gi|275 EQCGRQAGGALCPGGLCCSQYGCANTPEYCGSGCQSCDGGVGGEGGCVDLGSIIISRST
      10          20          30          40          50          60

gi|275 FEEMLKHRNNAACPAKGFYTYDAFISAAKAFPAFGTTGDVDTCKREIAAFFGQTSHATTG
      70          80          90         100         110         120

gi|275 GWPTAPDGPYAWGYCHKEELNQASSYCSPPAYPCAPGKYYGRGPIQLSWNNYNGQCQG
      130         140         150         160         170         180

gi|275 ALGLDLLNNPDLVATDRVISFKAIIWFWMTQFPKPSCHDVITGQWSPTGHDISAGRAPG
      190         200         210         220         230         240
```

```

                                10          20
CV127_                        SKLGFQQVRLRMKIGFYLRNINLINL
```



```

                                :.. .::: : .....
gi|275 YGVITNIINGGLECGSGWDARVEDRIGFYKRYCDMFGVGYGSNLDCYNQTPFGLG
                                250      260      270      280      290

>>gi|256427 gi|256427|gb|AAB23463.1| CG4 beta-conglycini (439 aa)
  initn: 43 initl: 43 opt: 43 Z-score: 92.6 bits: 22.2 E(): 3.3
Smith-Waterman score: 43; 42.857% identity (66.667% similar) in 21 aa overlap
(6-26:49-69)

                                10
CV127_ SKLGFQQVRLRMKIGFY
                                :.. .::: :
gi|256 MMRVRFPLLVLLGTVFLASVCVSLKVREDENNPFYFRSSNSFQTLFENQNVRIIRLLQRFN
                                10      20      30      40      50      60

                                20
CV127_ LRNINLINL
                                :.. .::: :
gi|256 KRSPQLENLRDYRIVQFQSKPNTILLPHHADADFLFVLSGRAILTLVNNDDRDSYNLHP
                                70      80      90      100     110     120

gi|256 GDAQRIAGTYYLVNPHDHQNLKIIKLAIPVKNKPGRYDDFFLSSTQAQQSYLQGFSHNI
                                130     140     150     160     170     180

gi|256 LETSFHSEFEEINRVLFGEERQEQGVIVELSKEQIRQLSRRAKSSSRKTISSEDEPF
                                190     200     210     220     230     240

gi|256 NLRSRNPIYSNNFGKFFEITPEKNPQLRDLDFLSSVDINEGALLLPHFNKAIVILVIN
                                250     260     270     280     290     300

gi|256 EGDANIELVGIKEQQQKQKEEPELVQRYRAELSEDDVFVIPAAYPFVVNATSNLNFLA
                                310     320     330     340     350     360

gi|256 FGINAENNQRNFLAGKDNVVRQIERQVQELAFPGSAQDVERLLKKQRESYFVDAQPQQK
                                370     380     390     400     410     420

gi|256 EEGSKGRKGPFPSILGALY
                                430

>>gi|14575525 gi|14575525|emb|CAC42881.1| putative class (295 aa)
  initn: 34 initl: 34 opt: 41 Z-score: 91.1 bits: 21.3 E(): 4
Smith-Waterman score: 41; 40.000% identity (73.333% similar) in 15 aa overlap
(9-23:261-275)

gi|145 EQCGRQAGGALCPGGLCCSYGWCANTPEYCGSGCQSQCDGGGGEDGGIDLGSIIIRST
                                10      20      30      40      50      60

gi|145 FEEMLKHRNDAACPAKGFYTYDAFISAAKAFPAFGTTGDVDTCKREIAAFFGQTSHATTG
                                70      80      90      100     110     120

gi|145 GWPTAPDGPYAWGYCYKEELNQASSYCSPPAYPCAPGKKYYGRGPIQLSWNYNYGQCQG
                                130     140     150     160     170     180

gi|145 ALGLDLLNNPDVATDRVISFKAAIWFWMTPQFPKPSCHDVITGQWSPTGHDISAGRAPG
                                190     200     210     220     230     240

                                10      20
CV127_ SKLGFQQVRLRMKIGFYLRNINLINL
                                :.. .::: : ..
gi|145 YGVITNIINGGLECGRGWDARVEDRIGFYKRYCDMFAVGYGSNLDCYNQTPFGLG
                                250     260     270     280     290

```

>>gi|1359600 gi|1359600|emb|CAA64868.1| chitinase Ib [Ca (316 aa)  
 initn: 33 initl: 33 opt: 40 Z-score: 88.5 bits: 21.0 E(): 5.6  
 Smith-Waterman score: 40; 54.545% identity (72.727% similar) in 11 aa overlap  
 (9-19:284-294)

gi|135 MKLFSLLLFLAFLLGTSAEQCGRQAGGAACANNLCCSQFGWCGNTAEYCGAGCQSQCSSP  
 10 20 30 40 50 60

gi|135 TTTTSSPTASSGGGGDVGSLISASLFDQMLKYRNDPRCKSNGFYTYNAFIAAARSFNFGF  
 70 80 90 100 110 120

gi|135 TTGDVTTTRKRELAFLAQTSHETTGGWATAPDGPYAWGYCFVMENNKQTYCTSKSWPCVF  
 130 140 150 160 170 180

gi|135 GKQYYGRGPIQLTHNNGYQAGKAIGADLINNPDLVATNPITISFKTAIWFWMTQANKPS  
 190 200 210 220 230 240

CV127\_ SKLGFQQVRLRMKIGFYLRNINLIN  
 :. :. :. :.

gi|135 SHDVIIGNWRPSAADTSAGRVPSYGVITNIINGGLECGHGSDDRANRIGFYKRYCDTLG  
 250 260 270 280 290 300

CV127\_ L

gi|135 VSYGNNLDCYNQKPFA  
 310

>>gi|14423650 gi|14423650|sp|Q9U1G2.1|ALL7\_LEPDS RecName (216 aa)  
 initn: 38 initl: 38 opt: 38 Z-score: 86.9 bits: 20.1 E(): 6.8  
 Smith-Waterman score: 38; 32.000% identity (64.000% similar) in 25 aa overlap  
 (2-26:105-129)

gi|144 MQYLAIAVIVALAGLSAAAHKPAYYDDNMANQMVDQIVKSLTTKKELDPFKIEQTKVPID  
 10 20 30 40 50 60

CV127\_ SKLGFQQVRLRMKIGFY  
 ::: :. :. :.

gi|144 KKIGLIHIKGSATIKNAVITGLSHISRRGDAKIDTDGGAFATLKLGDKNIRIKTDLHL  
 70 80 90 100 110 120

CV127\_ LRNINLINL  
 :. :. :.

gi|144 LGKIIHPNLKFEGHIGDIDMKLKLDAEGKPSLDQFEIDEFEQVELFIHGLGPLDPLVD  
 130 140 150 160 170 180

gi|144 VIADSFVKYFNPQARKLVTDMLKPILVEEIKKLKLN  
 190 200 210

>>gi|119390336 gi|119390336|pdb|2J23|B Chain B, Cross-Re (121 aa)  
 initn: 34 initl: 34 opt: 36 Z-score: 86.8 bits: 19.3 E(): 7  
 Smith-Waterman score: 36; 40.000% identity (80.000% similar) in 10 aa overlap  
 (13-22:66-75)

CV127\_ SKLGFQQ

gi|119 MRGSHHHHHHLVPRGSVQVISSYDQFKQVTGGDKVVVIDFWATWCGPCKMIGPVFEKISD  
 10 20 30 40 50 60

```

      10      20
CV127_ VRLRMKIGFYLRNINLINL
      .:.:. .
gi|119 TPAGDKVGFYKVDVDEQSQIAQEVGIRAMPTFVFFKNGQKIDTVVGADPSKLQAAITQHS
      70      80      90      100      110      120

```

gi|119 A

```

>>gi|17907758 gi|17907758|dbj|BAB79444.1| BW8KD allergen (133 aa)
  initn: 36 initl: 36 opt: 36 Z-score: 86.1 bits: 19.3 E(): 7.6
Smith-Waterman score: 36; 40.000% identity (73.333% similar) in 15 aa overlap
(1-15:114-128)

```

```

gi|179 MKLFLILAAASLLIVASHADSQMRSKCRKQMRMMEPQLEQCEGYMTMDMDDDSMRGREG
      10      20      30      40      50      60

```

```

CV127_ SKLGFQQ
      .:.:.
gi|179 RSEESCMRGCCCLAMKEMDDECMCEWMKMMVQQQRGEMGEEDMRMVMRKMKQLPNKCGMGH
      70      80      90      100      110      120

```

```

      10      20
CV127_ VRLRMKIGFYLRNINLINL
      .:.:.
gi|179 MRCHMGIGTRDYE
      130

```

```

26 residues in 1 query sequences
307888 residues in 1386 library sequences
Scomplib [34.26]
start: Wed Mar 11 05:54:10 2009 done: Wed Mar 11 05:54:10 2009
Total Scan time: 0.060 Total Display time: 0.010

```

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

## CV127\_rc\_5729\_5831\_frame2

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

Please cite:

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

```

CV127_rc_5729_5831_frame2, 1 aa
vs /n/na4/bioinfo/refsets/PUB_AllergenOnlineJan2009_V9.fasta library

```

```

307888 residues in 1386 sequences
unscaled statistics: mu= 50.0000 var=10.0000; Lambda= 0.4056

```

```

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 15, opt: 20, open/ext: -10/-2, width: 16
!! No sequences with E() < 10.000000

```

```

1 residues in 1 query sequences
307888 residues in 1386 library sequences
Scomplib [34.26]

```

start: Wed Mar 11 06:52:42 2009 done: Wed Mar 11 06:52:42 2009  
Total Scan time: 0.040 Total Display time: 0.000

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

## CV127\_rc\_5735\_5849\_frame2

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

Please cite:

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

CV127\_rc\_5735\_5849\_frame2, 4 aa

vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences

Expectation\_n fit:  $\rho(\ln(x)) = 1.5306 \pm 0.0021$ ;  $\mu = 6.0050 \pm 0.113$

mean\_var=6.4914  $\pm$  2.073, 0's: 224 Z-trim: 224 B-trim: 0 in 0/31

Lambda= 0.503391

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

join: 27, opt: 20, open/ext: -10/-2, width: 16

The best scores are:		opt bits	E(1386)	%_id	%_sim	alen
gi 29839255	gi 29839255 sp O23880.1 ( 504)	26	19.4	4	1.000	1.000 4
gi 113200131	gi 113200131 gb ABI321 ( 515)	26	19.4	4.1	1.000	1.000 4
gi 29839419	gi 29839419 sp Q9XFM4.1 ( 538)	26	19.4	4.2	1.000	1.000 4
gi 29839254	gi 29839254 sp O23878.1 ( 565)	26	19.5	4.4	1.000	1.000 4

>>>CV127\_rc\_5735\_5849\_frame2, 4 aa vs

/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

>>gi|29839255 gi|29839255|sp|O23880.1|13S2\_FAGES RecName (504 aa)

initn: 26 initl: 26 opt: 26 Z-score: 91.1 bits: 19.4 E(): 4

Smith-Waterman score: 26; 100.000% identity (100.000% similar) in 4 aa overlap (1-4:266-269)

gi|298 MSTKLILSFSLCLMVLSCSAQLWPWQKGQGSRPHHGRQQHQFQHQCDIQRLTASEPSRRV  
10 20 30 40 50 60

gi|298 RSEAGVTEIWDHDTPEFRCTGFVAVRVVIQPGLLLPSSNAPYITFVEQGRGVQGVVIP  
70 80 90 100 110 120

gi|298 GCPETFQSDSEFEYFQSQGRHSRQSESEESSRGDQHQKIFRIREGDVIPSPAGVVQWT  
130 140 150 160 170 180

gi|298 HNDGNDLISVTLLDANSYHKQLDENVRSFFLAGQSQRETREEGSDRQSRESDDDEALLG  
190 200 210 220 230 240

CV127\_ KLRG

::::

gi|298 ANILSGFQDEILHELFRDVTRETISKLRGENDQRGFIVQAQDLKLRVPQDFEEYERERG  
250 260 270 280 290 300

gi|298 DRRRGQGGSGRSNGVEQGFENLKFRRNFNTPTNTYVFNPAGRINTVNSNSLPILFLQL  
310 320 330 340 350 360

gi|298 SAQHVVLYKNAIIGPRWNLNAHSALYVTRGEGRVQVVGDEGKSVFDDKVQRGQILVVPQG  
370 380 390 400 410 420

gi|298 FAVVLKAGREGLEWVELKNSGNAITSPIGGRTSVLRAIPVEVLANSYDISTKEAYKLKNG  
430 440 450 460 470 480

gi|298 RQEVFVFRPFQSRDEKERERFSIV  
490 500

>>gi|113200131 gi|113200131|gb|ABI32184.1| allergenic pr (515 aa)  
initn: 26 initl: 26 opt: 26 Z-score: 91.0 bits: 19.4 E(): 4.1  
Smith-Waterman score: 26; 100.000% identity (100.000% similar) in 4 aa overlap  
(1-4:273-276)

gi|113 MSTKLILSFSLSCLMVLSQSAQAQLWPWRKGQDSRPHHGHQQFQQQCDIQLRTASEPSRR  
10 20 30 40 50 60

gi|113 VRSEAGVTEIWDHNTPEFRCTGFVAVRYVIQPGGLLLPSYSNAPYITFVEQGRGVQGVVI  
70 80 90 100 110 120

gi|113 PGCPETFQSDSEYPQSQRGQHSRESSESQESSRGDQHQKIFRVREGDVIPSPAGVVQWTHN  
130 140 150 160 170 180

gi|113 DGDQDLISVTLLDANSFHNQLDENVRSFFLAGQSQQGREERRSQQTREEGGDRQSRES  
190 200 210 220 230 240

CV127\_ KLRG  
:::  
gi|113 DVEALIGANILSGFQDEILHELFRDVEDRETISKLRGENDQRGFIVQAQDLKLRVPEDSEE  
250 260 270 280 290 300

gi|113 GYERQGRDKRDERGSGRSNGLEQAFCNLKFQNVNRP SHADVFNP RAGRINTVNSNNLP  
310 320 330 340 350 360

gi|113 ILEFLQLSAQHVVLYKNAIIGPRWNLNAHSALYVTRGEGRVQVVGDEGKSVFDDNVQRGQ  
370 380 390 400 410 420

gi|113 ILVVPQGFVAVVKAGRQGLEWVELKNNDNAITSPIAGRTSVLRAIPVEVLANSYDISTEE  
430 440 450 460 470 480

gi|113 AYKLKNGRQEVFVFRPFQSRYEKEEKEKERERFSIV  
490 500 510

>>gi|29839419 gi|29839419|sp|Q9XFM4.1|13S3\_FAGES RecName (538 aa)  
initn: 26 initl: 26 opt: 26 Z-score: 90.7 bits: 19.4 E(): 4.2  
Smith-Waterman score: 26; 100.000% identity (100.000% similar) in 4 aa overlap  
(1-4:302-305)

gi|298 MSTKLILSFSLSCLMVLSQSAQLLPWQKGQSRPHHGHQQFQHQCDIQLRTASEPSRRVRS  
10 20 30 40 50 60

gi|298 EAGVTEIWDHDTPEFRCAQFVAVRVVIQPGGLLLPSYSNAPYITFVEQGRGVQGVVPGC  
70 80 90 100 110 120

gi|298 PETFQSGSEFEYPRSQRDQRSRQSESGESSRGDQRSRQSESEESSRGDQRSRQSESEEF  
130 140 150 160 170 180

gi|298 RGDQHQKIFRIRDGDVIPSPAGVVQWTHNNGDNDLISITLYDANSFQNLDENVRNFFLA  
190 200 210 220 230 240

gi|298 GQSKQSREDRRSQRQTREEGSDRQSRESQDDEALLEANILSGFEDEILQEIFRNVDQETI  
250 260 270 280 290 300

CV127\_ KLRG

::::

gi|298 SKLRGENDQRGFIVQARDLKLRVPPEEYEEELQRRERGDGRKGGSGRSNGLEQAFCNLKFRQ  
310 320 330 340 350 360

gi|298 NVNRPSRADVFNPRAGRINTVDSNNLPILFIQLSAQHVVLYKNAILGPRWNLNAHSALY  
370 380 390 400 410 420

gi|298 VTRGEGRVQVVGDEGRSVFDDNVQRGQILVVPQGFVAVLKAGREGLEWVELKNDDNAITS  
430 440 450 460 470 480

gi|298 PIAGKTSVLRAIPVEVLANSYDISTKEAFRLKNGRQEVFVRPFQSRDEKERERFSIV  
490 500 510 520 530

>>gi|29839254 gi|29839254|sp|O23878.1|13S1\_FAGES RecName (565 aa)  
initn: 26 initl: 26 opt: 26 Z-score: 90.4 bits: 19.5 E(): 4.4  
Smith-Waterman score: 26; 100.000% identity (100.000% similar) in 4 aa overlap  
(1-4:332-335)

gi|298 MSTKLILSFSCLMVLSCSAQLLPWRKGQSRPHRGHQFHHQCDVQRLTASEPSRRVRS  
10 20 30 40 50 60

gi|298 EAGVTEIWDNDTPEFRFCAGFVAVRVVIQPGGLLLPSYSNAPYITFVEQGRGVQGVVPGC  
70 80 90 100 110 120

gi|298 PETFQSESEFEYPQSQRDQSRQSESEESSRGDQTRQSESEEFSSRGDQTRQSESEEFSS  
130 140 150 160 170 180

gi|298 RGDQTRQSESEEFSSRGDQTRQSESEEFSSRGDQHKKIFRIRDGDVIPSPAGVVQWTHND  
190 200 210 220 230 240

gi|298 GDNDLISITLYDANSFQNLQDGNVRNFFLAGQSKQSREDDRSQRQTRREEGSDRQSRESDD  
250 260 270 280 290 300

CV127\_

KLRG

::::

gi|298 DEALLEANILTGFDQDEILQEIFRNVDQETISKLRGDNDQRGFIVQARDLKLRVPPEEYEE  
310 320 330 340 350 360

gi|298 LQRRERGDGRKGGSGRSNGLEQAFCNLKFKQNVNRPSRADVFNPRAGRINTVNSNNLPIL  
370 380 390 400 410 420

gi|298 FIQLSAQHVVLYKNAILGPRWNLNAHSALYVTRGEGRVQVVGDEGRSVFDDNVQRGQILV  
430 440 450 460 470 480

gi|298 VPQGFVAVLKAGREGLEWVELKNDDNAITSPIAGKTSVLRAIPVEVLANSYDISTKEAFR  
490 500 510 520 530 540

gi|298 LKNGRQEVFVFLPFQSRDEKERERF  
550 560

4 residues in 1 query sequences

307888 residues in 1386 library sequences

Scomplib [34.26]

start: Wed Mar 11 06:54:46 2009 done: Wed Mar 11 06:54:46 2009

Total Scan time: 0.050 Total Display time: 0.000

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

## CV127\_rc\_5776\_5799\_frame0

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

Please cite:

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

CV127\_rc\_5776\_5799\_frame0, 8 aa

vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences

Expectation\_n fit:  $\rho(\ln(x)) = 2.8941 \pm 0.00236$ ;  $\mu = -0.0740 \pm 0.128$

mean\_var=8.7475 +/- 2.216, 0's: 134 Z-trim: 134 B-trim: 183 in 1/35

Lambda= 0.433644

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

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

The best scores are:		opt	bits	E(1386)	%_id	%_sim	alen
gi 110644964	gi 110644964 gb ABG813 ( 130)	28	19.6	1.8	0.500	1.000	6
gi 110644958	gi 110644958 gb ABG813 ( 131)	28	19.6	1.8	0.500	1.000	6
gi 110644954	gi 110644954 gb ABG813 ( 131)	28	19.6	1.8	0.500	1.000	6
gi 110644962	gi 110644962 gb ABG813 ( 131)	28	19.6	1.8	0.500	1.000	6
gi 110644960	gi 110644960 gb ABG813 ( 131)	28	19.6	1.8	0.500	1.000	6
gi 110644952	gi 110644952 gb ABG813 ( 131)	27	19.0	2.8	0.500	1.000	6
gi 110644956	gi 110644956 gb ABG813 ( 131)	27	19.0	2.8	0.500	1.000	6
gi 2154730	gi 2154730 emb CAA69670. ( 131)	27	19.0	2.8	0.500	1.000	6
gi 3367714	gi 3367714 emb CAA08836. ( 152)	26	18.3	5.2	0.600	1.000	5
gi 129353	gi 129353 sp P22895.1 P34 ( 379)	28	19.2	6.9	0.667	0.833	6
gi 1199563	gi 1199563 gb AAB09252.1 ( 379)	28	19.2	6.9	0.667	0.833	6
gi 1 AOID 2607 LTP Bra o 3 [Brassic ( 20)		19	14.7	8.5	0.600	1.000	5
gi 1335877	gi 1335877 gb AAB01092.1 ( 168)	25	17.6	9.2	0.667	1.000	6
gi 4240395	gi 4240395 gb AAD13531.1 ( 492)	28	19.1	9.6	0.800	1.000	5

>>>CV127\_rc\_5776\_5799\_frame0, 8 aa vs

/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

>>gi|110644964 gi|110644964|gb|ABG81318.1| pollen profil (130 aa)

initn: 26 init1: 26 opt: 28 Z-score: 97.3 bits: 19.6 E(): 1.8

Smith-Waterman score: 28; 50.000% identity (100.000% similar) in 6 aa overlap (3-8:48-53)

CV127\_

LTSNLRD

. . . . .

gi|110 MSWQAYVDEHLMCEIEGHHLTSAAIIVGHDGAVWAQSTAFPQFKTEEMTNIMKDFDEPGFL  
10 20 30 40 50 60

gi|110 APTGLFLGPTKYMVIQGEPGAIVIRGKKGSGGITVKKTGQAMVVGIIYDEPMTPGQCNMVVE  
70 80 90 100 110 120

gi|110 RLGDYLLNRA  
130

>>gi|110644958 gi|110644958|gb|ABG81315.1| pollen profil (131 aa)

initn: 26 init1: 26 opt: 28 Z-score: 97.2 bits: 19.6 E(): 1.8

Smith-Waterman score: 28; 50.000% identity (100.000% similar) in 6 aa overlap (3-8:48-53)

```
CV127_                                LTSNILRD
                                     .:..:
gi|110 MSWQAYVDEHLMCEIEGHHLTSAAIVGHDGAVWAQSTAFPQFKTEEMTNIMKDFDEPGFL
      10      20      30      40      50      60

gi|110 APTGLFLGPTKYMVIQGEPEGAVIRGKKGSGGITVKKTGQALVIGIYDEPMTPGQCNMVVE
      70      80      90     100     110     120

gi|110 RLGDYLLEQGM
      130

>>gi|110644954 gi|110644954|gb|ABG81313.1| pollen profil (131 aa)
  initn: 26 initl: 26 opt: 28 Z-score: 97.2 bits: 19.6 E(): 1.8
Smith-Waterman score: 28; 50.000% identity (100.000% similar) in 6 aa overlap
(3-8:48-53)
```

```
CV127_                                LTSNILRD
                                     .:..:
gi|110 MSWQAYVDEHLMCEIEGHHLTSAAIVGHDGAVWAQSTAFPQFKTEEMTNIMKDFDEPGFL
      10      20      30      40      50      60

gi|110 APTGLFLGPTKYMVIQGEPEGAVIRGKKGSGGITVKKTGQAMVVGIYDEPMTPGQCNMVVE
      70      80      90     100     110     120

gi|110 RLGDYLLEQGL
      130

>>gi|110644962 gi|110644962|gb|ABG81317.1| pollen profil (131 aa)
  initn: 26 initl: 26 opt: 28 Z-score: 97.2 bits: 19.6 E(): 1.8
Smith-Waterman score: 28; 50.000% identity (100.000% similar) in 6 aa overlap
(3-8:48-53)
```

```
CV127_                                LTSNILRD
                                     .:..:
gi|110 MSWQAYVDEHLMCEIEGHHLTSAAIVGHDGAVWAQSTAFPQSKTEEMTNIMKDFDEPGFL
      10      20      30      40      50      60

gi|110 APTGLFLGPTKYMVIQGEPEGAVIRGKKGSGGITVKKTGQAMVVGIYDEPMTPGQCNMVVE
      70      80      90     100     110     120

gi|110 RLGDYLLEQGL
      130

>>gi|110644960 gi|110644960|gb|ABG81316.1| pollen profil (131 aa)
  initn: 26 initl: 26 opt: 28 Z-score: 97.2 bits: 19.6 E(): 1.8
Smith-Waterman score: 28; 50.000% identity (100.000% similar) in 6 aa overlap
(3-8:48-53)
```

```
CV127_                                LTSNILRD
                                     .:..:
gi|110 MSWQAYVDEHLMCEIEGHHLTSAAIVGHDGAVWAQSTAFPQFKTEEMTNIMKDFDEPGFL
      10      20      30      40      50      60

gi|110 APTGLFLGPTKYMVIQGEPEGAVIRGKKGSGGITVKKTGQALVIGIYDEPMTPGQCNMVVE
      70      80      90     100     110     120

gi|110 RLGDYLLEQGL
```



130

```
>>gi|110644952 gi|110644952|gb|ABG81312.1| pollen profil (131 aa)
  initn: 25 initl: 25 opt: 27 Z-score: 93.8 bits: 19.0 E(): 2.8
Smith-Waterman score: 27; 50.000% identity (100.000% similar) in 6 aa overlap
(3-8:48-53)
```

```
CV127_                                LTSNILRD
                                     .:..:
gi|110 MSWQTYVDEHLMCEIEGHHLSSAAIVGHDGAVWAQSTAFPQFKPEEMTNIKDFDEPGFL
      10          20          30          40          50          60

gi|110 APIGLFLGPTKYMVIQGEPEGAVIRGKKGSGGITVKKTGQALVIGIYDEPMTPGQCNMVVE
      70          80          90         100         110         120

gi|110 RLGDYLVKQGL
      130
```

```
>>gi|110644956 gi|110644956|gb|ABG81314.1| pollen profil (131 aa)
  initn: 26 initl: 26 opt: 27 Z-score: 93.8 bits: 19.0 E(): 2.8
Smith-Waterman score: 27; 50.000% identity (100.000% similar) in 6 aa overlap
(3-8:48-53)
```

```
CV127_                                LTSNILRD
                                     .:..:
gi|110 MSWQAYVDEHLMCEIEGHHLTSAAIVGHDGAAWAQSTAFPEFKTEDMANIMKDFDEPGHL
      10          20          30          40          50          60

gi|110 APTGLFLGPTKYMVIQGEPEGAVIRGKKGSGGITVKKTGQALVVGIYDEPMTPGQCNMVVE
      70          80          90         100         110         120

gi|110 RLGDYLLKQGL
      130
```

```
>>gi|2154730 gi|2154730|emb|CAA69670.1| profilin 1 [Cyno (131 aa)
  initn: 26 initl: 26 opt: 27 Z-score: 93.8 bits: 19.0 E(): 2.8
Smith-Waterman score: 27; 50.000% identity (100.000% similar) in 6 aa overlap
(3-8:48-53)
```

```
CV127_                                LTSNILRD
                                     .:..:
gi|215 MSWQAYVDDHLMCEIEGHHLTSAAIIGHDGTVWAQSAAFPAPFKPEEMANIMKDFDEPGFL
      10          20          30          40          50          60

gi|215 APTGLFLGPTKYMVIQGEPEGAVIRGKKGSGGVTVKKTGQALVIGIYDEPMTPGQCNMVIE
      70          80          90         100         110         120

gi|215 KLGDYLIQGM
      130
```

```
>>gi|3367714 gi|3367714|emb|CAA08836.1| BDAI-1; Barley d (152 aa)
  initn: 26 initl: 26 opt: 26 Z-score: 89.0 bits: 18.3 E(): 5.2
Smith-Waterman score: 26; 60.000% identity (100.000% similar) in 5 aa overlap
(4-8:67-71)
```

```
gi|336 MGAMWMKSMLLVLLLCMLMVTPTMGARSDNSGPWMWCDEPMGHKVSPLTRCRLVKLECV
      10          20          30          40          50          60
```

```
CV127_      LTSNILRD
          ..:::
gi|336 GNRVPEDVLRDCCQEVAISNEWCRCDLGSMLRSVYAALGVGGGPPEEVFPGCQKDVMLK
          70          80          90          100          110          120

gi|336 LVAGVPALCNVPIPNEAAGTRGVCIYWSASTDT
          130          140          150

>>gi|129353 gi|129353|sp|P22895.1|P34_SOYBN RecName: Ful (379 aa)
  initn: 28 initl: 28 opt: 28 Z-score: 86.8 bits: 19.2 E(): 6.9
Smith-Waterman score: 28; 66.667% identity (83.333% similar) in 6 aa overlap
(3-8:72-77)

gi|129 MGFLVLLLFSLLGLSSSSSISTHRSILDLDLTKFTTQKQVSSLFQLWKSEHGRVYHNHEE
          10          20          30          40          50          60

CV127_      LTSNILRD
          ::.:
gi|129 EAKRLEIFKNNSNYIRDMNANRKSPHSHRLGLNKFADITPQEFSSKKYLQAPKDVSSQIKM
          70          80          90          100          110          120

gi|129 ANKKMKKEQYSCDHPPASWDWRKKGVITQVKYQGGCGRGWAFSATGAIEAAHAIATGDLV
          130          140          150          160          170          180

gi|129 SLSEQELVDCVEESEGSYNGWQYQSFVLEHGGIATDDDYPRAKEGRCKANKIQDKVT
          190          200          210          220          230          240

gi|129 IDGYETLIMSDESTESETEQAFLSAILEQPISVSIDAKDFHLYTGGIYDGENCTSPYGIN
          250          260          270          280          290          300

gi|129 HFVLLVGYGSADGVYWIAKNSWGFWDGEDGYIWIQRNTGNLLGVCGMNYFASYPTKEES
          310          320          330          340          350          360

gi|129 ETLVSARVKGHRRVDHSPL
          370

>>gi|1199563 gi|1199563|gb|AAB09252.1| 34 kDa maturing s (379 aa)
  initn: 28 initl: 28 opt: 28 Z-score: 86.8 bits: 19.2 E(): 6.9
Smith-Waterman score: 28; 66.667% identity (83.333% similar) in 6 aa overlap
(3-8:72-77)

gi|119 MGFLVLLLFSLLGLSSSSSISTHRSILDLDLTKFTTQKQVSSLFQLWKSEHGRVYHNHEE
          10          20          30          40          50          60

CV127_      LTSNILRD
          ::.:
gi|119 EAKRLEIFKNNSNYIRDMNANRKSPHSHRLGLNKFADITPQEFSSKKYLQAPKDVSSQIKM
          70          80          90          100          110          120

gi|119 ANKKMKKEQYSCDHPPASWDWRKKGVITQVKYQGGCGRGWAFSATGAIEAAHAIATGDLV
          130          140          150          160          170          180

gi|119 SLSEQELVDCVEESEGSYNGWQYQSFVLEHGGIATDDDYPRAKEGRCKANKIQDKVT
          190          200          210          220          230          240

gi|119 IDGYETLIMSDESTESETEQAFLSAILEQPISVSIDAKDFHLYTGGIYDGENCTSPYGIN
          250          260          270          280          290          300

gi|119 HFVLLVGYGSADGVYWIAKNSWGFWDGEDGYIWIQRNTGNLLGVCGMNYFASYPTKEES
          310          320          330          340          350          360
```

gi|119 ETLVSARVKGHRRVDHSPL  
370

>>gi|1 AOID|2607|LTP Bra o 3 [Brassica oleracea] Manual (20 aa)  
initn: 19 initl: 19 opt: 19 Z-score: 85.2 bits: 14.7 E(): 8.5  
Smith-Waterman score: 20; 60.000% identity (100.000% similar) in 5 aa overlap  
(1-5:7-11)

CV127\_ LTSNILRD  
.:...  
gi|1 AISC GTVT SNLAPCAVYLMK  
10 20

>>gi|1335877 gi|1335877|gb|AAB01092.1| pathogenesis-rela (168 aa)  
initn: 25 initl: 25 opt: 25 Z-score: 84.6 bits: 17.6 E(): 9.2  
Smith-Waterman score: 25; 66.667% identity (100.000% similar) in 6 aa overlap  
(1-6:7-12)

CV127\_ LTSNILRD  
:..:..  
gi|133 LIFLLILSNILNIMGAQSHSLEITSSVSAEKIFXXIVLDVDTVIPKAAPGAYKSVDVKG  
10 20 30 40 50 60

gi|133 DGGAGTVRIITLPEGSPITSMVTVDVAVNKEALTYDSTVIDGDILLEFIESIETHMVVVP  
70 80 90 100 110 120

gi|133 TADGGSITKTTAIFHTKGDVAVPEENIKFADAQNTALFKAIEAYLIAN  
130 140 150 160

>>gi|4240395 gi|4240395|gb|AAD13531.1| major allergen B1 (492 aa)  
initn: 28 initl: 28 opt: 28 Z-score: 84.3 bits: 19.1 E(): 9.6  
Smith-Waterman score: 28; 80.000% identity (100.000% similar) in 5 aa overlap  
(4-8:113-117)

gi|424 NAEFLNNIHDLLGIPHIPVTARKHRRGVGITGLIDDI IAILPVDDLYALFQEKLETSP  
10 20 30 40 50 60

CV127\_ LTSNILRD  
:..:..  
gi|424 EFKALYDAIRSPEFQSIVGTLEAMPEYQNLIQKLKDKGVDVDHIIELIHQIFNIVRDRG  
70 80 90 100 110 120

gi|424 LPEDLQDFLALIPDQVLAIAADYLANDAEVKA AVEY LKSDEFETIVVTVDLSLPEFKNFL  
130 140 150 160 170 180

gi|424 NFLQTNGLNIAEFLNNIHDLLGIPHIPVTGRKHLRRGVGITGLIDDI IAILPVDDLYALF  
190 200 210 220 230 240

gi|424 QEKLETSPFEFKALYDAIRSPEFQSIVETLKAMPEYQSLIQKLKDKGVDVDHIIELIHQIF  
250 260 270 280 290 300

gi|424 NIVRDRGLPEDLQDFLALIPIDQILAIADYLANDAEVQA AVEY LKSDEFETIVVTVDLS  
310 320 330 340 350 360

gi|424 LPEFKNFLNFLQTNGLNIAEFLNNIHDLLGIPHIPATGRKHVRRGVGINGLIDDVIAILP  
370 380 390 400 410 420

gi|424 VDELYALFQEKLESSPEFKALYDAIRSPEFQSIVQTLKAMPEYQDLIQRLKDKGVDVDHF

```

430      440      450      460      470      480
gi|424 IELIKKLFGLSH
490

```

8 residues in 1 query sequences  
307888 residues in 1386 library sequences  
Scomplib [34.26]  
start: Wed Mar 11 06:32:25 2009 done: Wed Mar 11 06:32:25 2009  
Total Scan time: 0.040 Total Display time: 0.000

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

## CV127\_rc\_5780\_5815\_frame1

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

Please cite:

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

CV127\_rc\_5780\_5815\_frame1, 12 aa  
vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences  
Expectation\_n fit: rho(ln(x))= 1.2029+/-0.0028; mu= 12.4721+/- 0.147  
mean\_var=20.3422+/- 6.029, 0's: 80 Z-trim: 80 B-trim: 105 in 1/36  
Lambda= 0.284365

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

The best scores are:

	opt	bits	E(1386)	%_id	%_sim	alen
gi 5059162 gi 5059162 gb AAD38942.1 ( 496)	37	20.4	6.1	0.714	1.000	7
gi 3201547 gi 3201547 emb CAB01591. ( 326)	36	19.6	7	0.500	0.800	10
gi 14423733 gi 14423733 sp Q92260.1 ( 503)	36	20.0	8.1	0.500	0.800	10
gi 21783 gi 21783 emb CAA30570.1  u ( 356)	35	19.2	9.6	0.545	0.636	11
gi 75317968 gi 75317968 sp O22116 O ( 373)	35	19.3	9.7	0.545	0.636	11

>>>CV127\_rc\_5780\_5815\_frame1, 12 aa vs  
/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

>>gi|5059162 gi|5059162|gb|AAD38942.1|AF144060\_1 alpha-a (496 aa)  
initn: 37 initl: 37 opt: 37 Z-score: 87.8 bits: 20.4 E(): 6.1  
Smith-Waterman score: 37; 71.429% identity (100.000% similar) in 7 aa overlap  
(1-7:225-231)

gi|505 KYHNPFIIGNRSVITHLMEWKYDDIGDECERFLGPYGYGGVQVSPVNEHAILDRRPWYER  
10 20 30 40 50 60

gi|505 YQPVSYDIRTRSGDEQQFRMRVKRCNKAGVRIYVDIVLNHMTGAQSGKGTNGHHYDGNTL  
70 80 90 100 110 120

gi|505 QYPGVFPFGPNDFHGHESCPTQDLEIHDYTNPKERNCRSLGLRDLKQQSEYVRQKQVDFL  
130 140 150 160 170 180

CV127\_ 10  
PLIYYETSLPLY

```

                                :.:.:.:.
gi|505 NHLIDIGVAGFRSDASTHQWPDLLRSIYSLHNLNKEFFPENSQPFYIYHETIYYGGNGIN
      190      200      210      220      230      240

gi|505 SNEYTSLGRIIEFRFYKEITNVFRGNNPLHLKNGFGTEWGLVPSGDALVMIDSHDLRVGH
      250      260      270      280      290      300

gi|505 TGKLGFNINCFEGRLLKAATAFMLAWNYGVPRVMSSYFWNQIIKDGKDVNDWVGPPSDKN
      310      320      330      340      350      360

gi|505 GNILSVHPNPDMTCNHEWICEHRWREIYNMVKFRMIAGQEPVHNWWDNGDYQIAFSRGNR
      370      380      390      400      410      420

gi|505 AFIAINLQKNQQNLQQLHTGLPAGTYCDIISGNLIDNKCTGKSIHVDKNGQADVYVGHD
      430      440      450      460      470      480

gi|505 EFDAFVAYHIGARIVS
      490

```

```

>>gi|3201547 gi|3201547|emb|CAB01591.1| endochitinase [P (326 aa)
  initn: 36 initl: 36 opt: 36 Z-score: 86.7 bits: 19.6 E(): 7
Smith-Waterman score: 36; 50.000% identity (80.000% similar) in 10 aa overlap
(2-11:1-10)

```

```

      10
CV127_ PLIYYETSLPLY
      ..: .:..:
gi|320 MUYCTASLPLLLLLLVGLLAGEAFAEQCGRQAGGALCPGGLCCSQFGWCGSTSDYCGPT
      10      20      30      40      50

gi|320 CQSQCQGVTPSPGGGVASLISQSVFNQMLKHRNDAACQAKGFYTYNAFIAAANSFNGFAS
      60      70      80      90      100      110

gi|320 VGDATATRKREIAAFLAQTSHETTGGWATAPDGPYAWGYCFLKEQGNPPDYCVPTAQWPCA
      120      130      140      150      160      170

gi|320 PGKKYYGRGPIQISYNNYGPAGRAIGYDLINNPDAVATDPVISFKTALWFWMTQSPKP
      180      190      200      210      220      230

gi|320 SCHNVITGRWTPSAADRAAGRLPGYGVITNIINGGIECGKGFNDKVADRIGFYKRYCDLL
      240      250      260      270      280      290

gi|320 GVSYGSNLDCYNQRSFGVSTNPLAASS
      300      310      320

```

```

>>gi|14423733 gi|14423733|sp|Q92260.1|HSP70_PENCI RecNam (503 aa)
  initn: 36 initl: 36 opt: 36 Z-score: 85.6 bits: 20.0 E(): 8.1
Smith-Waterman score: 36; 50.000% identity (80.000% similar) in 10 aa overlap
(1-10:280-289)

```

```

gi|144 AYLGGTVNNAVITVPAYFNDSQRQATKDAGLIAGLNLVRIINEPTAAAIAYGLDKKTEGE
      10      20      30      40      50      60

gi|144 RNVLIFDLGGGTFDVSLLTIEEGIFEVKSTAGDTHLGGEDFDNRLVNHVFNVEFKRKHKKD
      70      80      90      100      110      120

gi|144 LTTNARALRRLTACERAKRTLSSAAQTSIEIDSLFEGIDFYTSITRARFEELCQDLFRG
      130      140      150      160      170      180

gi|144 TMEPVERVLRDAKIDKSSVHEIVLVGGSTRIPKIQKLVSDFFNKDANKSINPDEAVAYGA
      190      200      210      220      230      240

```

```

                                10
CV127_                          PLIYYETSLPLY
                                :::  .:..:
gi|144 AVQAAILSGDTSSKSTNEILLLDVAPLSLGIETAGGVMTPLIKRNTTIPTKKSETFSTYS
                                250      260      270      280      290      300

gi|144 DNQPGVLIQVFEGERARTKDNLLGKFELTGIPPAPRGVPQIEVTFDL DANGIMNV SASE
                                310      320      330      340      350      360

gi|144 KGTGKSNKITITNDKGRLSKEEIERMLAEAEKYKAEDEAEASRIQAKNGLESYAYSLKNT
                                370      380      390      400      410      420

gi|144 ITEGKLQMSDDDKKKIEDKISEIISWLDNNQTAEKDEYESQQKELEAIANPIMQAAYGAA
                                430      440      450      460      470      480

gi|144 GGAPPQQRADGETEEKKDEEELD
                                490      500

>>gi|21783 gi|21783|emb|CAA30570.1| unnamed protein prod (356 aa)
  initn: 35 initl: 35 opt: 35 Z-score: 84.3 bits: 19.2 E(): 9.6
Smith-Waterman score: 35; 54.545% identity (63.636% similar) in 11 aa overlap
(1-11:337-347)

gi|217 MKTFLV FALLALAAASAVAQISQQQQAPPFSSQQQQPPFSSQQQQPPFSSQQQSPFSSQQQQQ
                                10      20      30      40      50      60

gi|217 PPFAQQQQPPFSSQQQPISQQQQPPFSSQQQQPPFSSQQQQPPFSSQQQQPPFSSQQ
                                70      80      90      100     110     120

gi|217 QQQPPFSSQQQQPPFTQQQQQQQQQQPPFTQQQQPPFSSQQQPISQQQQPPFLQQQRPPFSRQ
                                130     140     150     160     170     180

gi|217 QQIPVIHPSVLQQLNPCKVFLQQQCIPVAMQRCLARSQMLQQSICHVMQQQCCQQLRQIP
                                190     200     210     220     230     240

gi|217 EQSRHESIRAIYISIILQQQQQQQQQQQQQQGQSIIQYQQQPPQQLGQCVSQPLQQLQQQ
                                250     260     270     280     290     300

                                10
CV127_                          PLIYYETSLPLY
                                ::  ::..:
gi|217 LGQQPQQQQLAHQIAQLEVMTSIALRTLPTMCNVNPLYETTTSVPLGVGIGVGVY
                                310     320     330     340     350

>>gi|75317968 gi|75317968|sp|O22116|O22116_WHEAT LMM glu (373 aa)
  initn: 35 initl: 35 opt: 35 Z-score: 84.2 bits: 19.3 E(): 9.7
Smith-Waterman score: 35; 54.545% identity (63.636% similar) in 11 aa overlap
(1-11:354-364)

gi|753 ASAVAQISQQQQPPPFSSQQQQPPFSSQQQQPPFSSQQQQSPFSSQQQQPPFSSQQ
                                10      20      30      40      50      60

gi|753 PLISQQQQLPFSSQQQQPPFSSQQQQPPFSSQQQQPPFSSQQQQPPFSSQQQQPSFS
                                70      80      90      100     110     120

gi|753 QQQQPPFTQQQQPPFSSQQSPISQQQQQQQQQQPPFTQQQQPPFSSQQQPISQQQQPPFSS
                                130     140     150     160     170     180

gi|753 QQQQPPFSSQQQIPVIHPSVLQQLNPCMVFLQQQCIPVAMQRCLARSQMLQQSICHVMQR
                                190     200     210     220     230     240

gi|753 QCCQQLRQIPEQSRHESIRAIYISIILQQQQQQQQQQQQQQGQSIIQYQQQPPQQLGQCV

```

```

                250          260          270          280          290          300

CV127_                                PLIYYET
                                ::      :
gi|753 SQPQQQLQQQLGQQPQQQQLAHGTFLQPHKIAQLEVMTSIALRTLPRMCSVNVPLYETTT
                310          320          330          340          350          360

                10
CV127_ SLPLY
                ::::
gi|753 SVPLGVGIGVGVY
                370

```

12 residues in 1 query sequences  
 307888 residues in 1386 library sequences  
 Scomplib [34.26]  
 start: Wed Mar 11 06:41:06 2009 done: Wed Mar 11 06:41:06 2009  
 Total Scan time: 0.050 Total Display time: 0.000

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

## CV127\_rc\_5787\_5825\_frame2

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

Please cite:

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

CV127\_rc\_5787\_5825\_frame2, 13 aa  
 vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences  
 Expectation\_n fit:  $\rho(\ln(x)) = 1.5688 \pm 0.00306$ ;  $\mu = 11.0911 \pm 0.162$   
 mean\_var = 28.1764  $\pm$  8.526, 0's: 61 Z-trim: 61 B-trim: 102 in 1/38  
 Lambda = 0.241619

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

	opt	bits	E(1386)	%_id	%_sim	alen
gi 41057621 gi 41057621 gb AAR98813 ( 231)	40	19.5	5.7	0.667	1.000	6
gi 124148 gi 124148 sp P16348.1 API ( 188)	38	18.6	8.5	0.571	0.857	7

```

>>>CV127_rc_5787_5825_frame2, 13 aa vs
/n/na4/bioinfo/refsets/PUB_AllergenOnlineJan2009_V9.fasta library

>>gi|41057621 gi|41057621|gb|AAR98813.1| glutathione-S-t (231 aa)
  initn: 32 initl: 32 opt: 40 Z-score: 88.4 bits: 19.5 E(): 5.7
Smith-Waterman score: 40; 66.667% identity (100.000% similar) in 6 aa overlap
(4-9:211-216)

```

```

gi|410 MSDKPSELAVQKLVLFAVKGTATSTHNTVRPLILLDELGVPHIYVVDVRSAPWFTEINP
                10          20          30          40          50          60

gi|410 HRMPVPVILEKSPDGRDTRLRAWESTSTLMYIADAYDKDGTFGGRNVQESSDINNWLTLHTA
                70          80          90          100          110          120

```

```

gi|410 ALGPTAKYWLYFYKLHPEKLPKTIEKLRSNITVQYDILERRLNPEGQYYLAWLNEKFKRS
          130          140          150          160          170          180

                                10
CV127_                                YITRLVCPFIRKK
                                :.:.:.
gi|410 SYNRRHCYASLCYEKYRRVVRAGVKVAQTARVVCYPYGGDTRRGVWPARKST
          190          200          210          220          230

>>gi|124148 gi|124148|sp|P16348.1|API11_SOLTU RecName: F (188 aa)
  initn: 36 initl: 36 opt: 38 Z-score: 85.2 bits: 18.6 E(): 8.5
Smith-Waterman score: 38; 57.143% identity (85.714% similar) in 7 aa overlap
(5-11:148-154)

gi|124 ESPLPKPVLDITNGKELNPNSSYRIISIGRGALGGDVYLGKSPNSDAPCPDGVFRYNSDVG
          10          20          30          40          50          60

gi|124 PSGTPVRFIPLSGGIFEDQLLNQFNIAITVKLCVSYTIWKVGNLNAFYRTMLLETGGTIG
          70          80          90          100          110          120

                                10
CV127_                                YITRLVCPFIRKK
                                ..:.:.
gi|124 QADSSYFKIVKLSNFGYNLLYCPITPPFLCPFCRDDNFCAKVGVVIQNGKRRLLALVNENP
          130          140          150          160          170          180

gi|124 LDVLFQEV

```

13 residues in 1 query sequences  
 307888 residues in 1386 library sequences  
 Scomplib [34.26]  
 start: Wed Mar 11 06:50:34 2009 done: Wed Mar 11 06:50:34 2009  
 Total Scan time: 0.050 Total Display time: 0.000

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

## CV127\_rc\_5803\_5838\_frame0

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

Please cite:

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

CV127\_rc\_5803\_5838\_frame0, 12 aa

vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences

Expectation\_n fit:  $\rho(\ln(x)) = 2.9498 \pm 0.00257$ ;  $\mu = 3.4139 \pm 0.134$   
 mean\_var=16.8096 $\pm$  4.887, 0's: 64 Z-trim: 64 B-trim: 2 in 1/37  
 Lambda= 0.312821

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

The best scores are: opt bits E(1386) %\_id %\_sim alen  
 gi|47117351 gi|47117351|sp||Q7M3Y8\_ ( 27) 34 18.6 1.1 0.714 1.000 7



gi 7638030	gi 7638030 gb AAF65313.1 ( 137)	36	19.7	2.7	0.500	1.000	8
gi 14424465	gi 14424465 sp P35777.2 ( 137)	36	19.7	2.7	0.500	1.000	8
gi 7638028	gi 7638028 gb AAF65312.1 ( 137)	36	19.7	2.7	0.500	1.000	8
gi 4038411	gi 4038411 gb AAC97370.1 ( 137)	36	19.7	2.7	0.500	1.000	8
gi 21913174	gi 21913174 gb AAM77471 ( 115)	32	17.8	8.1	0.556	0.778	9

>>>CV127\_rc\_5803\_5838\_frame0, 12 aa vs

/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

>>gi|47117351 gi|47117351|sp||Q7M3Y8\_2 [Segment 2 of 6] (27 aa)  
 initn: 34 initl: 34 opt: 34 Z-score: 100.9 bits: 18.6 E(): 1.1  
 Smith-Waterman score: 34; 71.429% identity (100.000% similar) in 7 aa overlap  
 (5-11:9-15)

	10
CV127_	FAPLLERNENKN
	: : : : : . .
gi 471	KITLLEEDLERNEERLQTATERLEEAS
	10 20

>>gi|7638030 gi|7638030|gb|AAF65313.1|AF230384\_1 venom a (137 aa)  
 initn: 22 initl: 22 opt: 36 Z-score: 94.1 bits: 19.7 E(): 2.7  
 Smith-Waterman score: 36; 50.000% identity (100.000% similar) in 8 aa overlap  
 (3-10:110-117)

gi 763	MKTFVLVSCLLVFTQIIYAVDIKELKIVNRILKKCIRTVPKGENDPINPLKNVNVLYCAF
	10 20 30 40 50 60

	10
CV127_	FAPLLERNENKN
	: : : : : . .
gi 763	SKRGIFTPKGVNTKQYINYCEKTIINPADIKQCKKLISKCIKKVYDRPGPIIERSKNLLS
	70 80 90 100 110 120

gi 763	CVRKKGVLELTVYGKKK
	130

>>gi|14424465 gi|14424465|sp|P35777.2|VA4\_SOLIN RecName: (137 aa)  
 initn: 22 initl: 22 opt: 36 Z-score: 94.1 bits: 19.7 E(): 2.7  
 Smith-Waterman score: 36; 50.000% identity (100.000% similar) in 8 aa overlap  
 (3-10:110-117)

gi 144	MKTFVLVSCLLVFTQIIYALDIKEISIMNRILEKCIIRTVPKRENDPINPLRNVNVWYCAF
	10 20 30 40 50 60

	10
CV127_	FAPLLERNENKN
	: : : : : . .
gi 144	TKRGIFTPKGVNTKQYINYCEKTAISPADIKLCKKIASKCVKKVYDRPGPIIERSKNLLS
	70 80 90 100 110 120

gi 144	CVLKKGLLELTVYGKKK
	130

>>gi|7638028 gi|7638028|gb|AAF65312.1|AF230383\_1 venom a (137 aa)  
 initn: 22 initl: 22 opt: 36 Z-score: 94.1 bits: 19.7 E(): 2.7  
 Smith-Waterman score: 36; 50.000% identity (100.000% similar) in 8 aa overlap  
 (3-10:110-117)

gi 763	MKTFVLVSCLLVFTQIIYAVDIKELKIMNRILEKCIIRTVPKGENDPINPLKNVNVLYCAF
	10 20 30 40 50 60

```

                                10
CV127_                          FAPLLERNENKN
                                .....:
gi|763 SKRGIFTPKGVNTKQYINYCEKTIINPADIKQCKKLISKCIKKVYDRPGPIIERSKNLLS
                                70      80      90      100     110     120

gi|763 CVLKKGVLELTVYGKKK
                                130

>>gi|4038411 gi|4038411|gb|AAC97370.1| venom allergen So (137 aa)
  initn: 22 initl: 22 opt: 36 Z-score: 94.1 bits: 19.7 E(): 2.7
Smith-Waterman score: 36; 50.000% identity (100.000% similar) in 8 aa overlap
(3-10:110-117)

gi|403 MKTFVLVSCLLVFTQIIYALDIKEISIMNRILEKCIRTVPKRENDPINPLKNVNVLYCAF
                                10      20      30      40      50      60

CV127_                          FAPLLERNENKN
                                .....:
gi|403 TKGIFTPKGVNTKQYINYCEKTIISPADIKLCKKIASKCVKKVYDRPGPIIERSKNLLS
                                70      80      90      100     110     120

gi|403 CVLKKGLLELTVYGKKK
                                130

>>gi|21913174 gi|21913174|gb|AAM77471.1| major allergen (115 aa)
  initn: 32 initl: 32 opt: 32 Z-score: 85.6 bits: 17.8 E(): 8.1
Smith-Waterman score: 32; 55.556% identity (77.778% similar) in 9 aa overlap
(2-10:19-27)

                                10
CV127_                          FAPLLERNENKN
                                ::: :..:
gi|219 MQFTTIASLFAAAGLAAAAPLESRQDNASCPVTTKGDYVWKISEFYGRKPEGTYNSLGF
                                10      20      30      40      50      60

gi|219 NIKATNGGTLDFTCQAQADKLEDHKWYSCGENSFMDFSFSDRSGLLLKQKVSDE
                                70      80      90      100     110

12 residues in 1 query sequences
307888 residues in 1386 library sequences
Scomplib [34.26]
start: Wed Mar 11 06:34:31 2009 done: Wed Mar 11 06:34:31 2009
Total Scan time: 0.050 Total Display time: 0.000

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

```

## CV127\_rc\_5819\_5881\_frame1

FASTA searches a protein or DNA sequence data bank version 34.26.5 April 26, 2007  
Please cite:  
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448  
  
CV127\_rc\_5819\_5881\_frame1, 21 aa

```

vs /n/na4/bioinfo/refsets/PUB_AllergenOnlineJan2009_V9.fasta library

307888 residues in 1386 sequences
Expectation_n fit: rho(ln(x))= 2.3852+/-0.00363; mu= 10.8741+/- 0.189
mean_var=26.3961+/- 7.314, 0's: 38 Z-trim: 38 B-trim: 47 in 1/38
Lambda= 0.249634

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2
join: 36, opt: 20, open/ext: -10/-2, width: 16
The best scores are:
gi|19570315 gi|19570315|dbj|BAB8628 ( 374) 45 21.0 5.3 0.471 0.706 17
gi|493634 gi|493634|dbj|BAA05543.1 ( 374) 45 21.0 5.3 0.471 0.706 17
gi|1173367 gi|1173367|sp|P18632.2|S ( 374) 45 21.0 5.3 0.471 0.706 17

>>>CV127_rc_5819_5881_frame1, 21 aa vs
/n/na4/bioinfo/refsets/PUB_AllergenOnlineJan2009_V9.fasta library

>>gi|19570315 gi|19570315|dbj|BAB86286.1| Cry j 1 precur (374 aa)
initn: 27 init1: 27 opt: 45 Z-score: 88.9 bits: 21.0 E(): 5.3
Smith-Waterman score: 45; 47.059% identity (70.588% similar) in 17 aa overlap
(1-17:232-247)

gi|195 MDSPCLVALLVFSFVIGSCFSDNPIDSCWRGDSNWAQNRMKLADCAVGFSGSSTMGGKGGD
10 20 30 40 50 60
gi|195 LYTVTNSDDDPVNPAPGTLRYGATRDRLWIIFSGNMNIKLKMPMYIAGYKTFDGRGAQV
70 80 90 100 110 120
gi|195 YIGNGGPCVFIKRVSNVIIHGLYLYGCSTSVLGNVLINESFGVEPVHPQDGDALTLRTAT
130 140 150 160 170 180

CV127_ KEMKIKTKG
: : : :
gi|195 NIWIDHNSFSNSSDGLVDVTLTSTGVTISNNLFFNHHKVMLLGHDDAYSDDKSMKV-TVA
190 200 210 220 230

10 20
CV127_ INKEGPFCSALI
: : : :
gi|195 FNQFGPNCQGQRMPRARYGLVHVANNNDPWTIYAIGGSSNPTILSEGNSFTAPNESYKKQ
240 250 260 270 280 290
gi|195 VTIRIGCKTSSSCSNVWQSTQDVFYNGAYFVSSGKYEGGNIYTKKEAFNVENGATPQL
300 310 320 330 340 350
gi|195 TKNAGVLTCSLSKRC
360 370

>>gi|493634 gi|493634|dbj|BAA05543.1| Cry j IB precursor (374 aa)
initn: 27 init1: 27 opt: 45 Z-score: 88.9 bits: 21.0 E(): 5.3
Smith-Waterman score: 45; 47.059% identity (70.588% similar) in 17 aa overlap
(1-17:232-247)

gi|493 MDSPCLVALLVFSFVIGSCFSDNPIDSCWRGDSNWAQNRMKLADCAVGFSGSSTMGGKGGD
10 20 30 40 50 60
gi|493 LYTVTNSDDDPVNPAPGTLRYGATRDRLWIIFSGNMNIKLKMPMYIAGYKTFDGRGAQV
70 80 90 100 110 120
gi|493 YIGNGGPCVFIKRVSNVIIHGLYLYGCSTSVLGNVLINESFGVEPVHPQDGDALTLRTAT

```

```

130      140      150      160      170      180

CV127_                                     KEMKIKTKG
: : : . : .
gi|493 NIWIDHNSFSNSSDGLVDVTLTSTGVTISNNLFFNHHKVMSLGHDDAYSDDKSMKV-TVA
190      200      210      220      230

10      20
CV127_ INKEGPFCALI
. . . : : :
gi|493 FNQFGPNCGQRMPrARYGLVHVANNNDPWTIYAIGGSSNPTILSEGNSTAPNESYKKQ
240      250      260      270      280      290

gi|493 VTIRIGCKTSSSCSNWVWQSTQDVFYNGAYFVSSGKYEGGNIYTKKEAFNVENGATPHL
300      310      320      330      340      350

gi|493 TQNAGVLTCSLSKRC
360      370

>>gi|1173367 gi|1173367|sp|P18632.2|SBP_CRYJA RecName: F (374 aa)
initn: 27 initl: 27 opt: 45 Z-score: 88.9 bits: 21.0 E(): 5.3
Smith-Waterman score: 45; 47.059% identity (70.588% similar) in 17 aa overlap
(1-17:232-247)

```

```

gi|117 MDSPCLVALLVLSFVIGSCFSDNPIDSCWRGDSNWAQNRMKLADCAVGFSGSSTMGGKGGD
10      20      30      40      50      60

gi|117 LYTVTNSDDDPVNPAPGTLRYGATRDRPLWIIIFSGNMNIKLKMPMYIAGYKTFDGRGAQV
70      80      90      100      110      120

gi|117 YIGNGGPCVFIFKRVSNIHGLHLYGCSTSVLGNVLINESFGVEPVHPQDGDALTLRTAT
130      140      150      160      170      180

```

```

CV127_                                     KEMKIKTKG
: : : . : .
gi|117 NIWIDHNSFSNSSDGLVDVTLSSSTGVTISNNLFFNHHKVMMLLGHDDAYSDDKSMKV-TVA
190      200      210      220      230

10      20
CV127_ INKEGPFCALI
. . . : : :
gi|117 FNQFGPNCGQRMPrARYGLVHVANNNDPWTIYAIGGSSNPTILSEGNSTAPNESYKKQ
240      250      260      270      280      290

gi|117 VTIRIGCKTSSSCSNWVWQSTQDVFYNGAYFVSSGKYEGGNIYTKKEAFNVENGATPQL
300      310      320      330      340      350

gi|117 TKNAGVLTCSLSKRC
360      370

```

21 residues in 1 query sequences  
307888 residues in 1386 library sequences  
Scomplib [34.26]  
start: Wed Mar 11 06:43:16 2009 done: Wed Mar 11 06:43:16 2009  
Total Scan time: 0.060 Total Display time: 0.000

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

## CV127\_rc\_5842\_5850\_frame0

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

Please cite:

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

CV127\_rc\_5842\_5850\_frame0, 3 aa

vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences

Expectation\_n fit:  $\rho(\ln(x)) = 0.4803 \pm 0.00239$ ;  $\mu = 12.3563 \pm 0.131$

mean\_var=3.2831  $\pm$  0.913, 0's: 515 Z-trim: 515 B-trim: 0 in 0/28

Lambda= 0.707837

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

join: 24, opt: 20, open/ext: -10/-2, width: 16

The best scores are:

		opt	bits	E(1386)	%_id	%_sim	alen
gi 29465664	gi 29465664 gb AAL92578 ( 123)	21	15.8	8.8	1.000	1.000	3
gi 111494253	gi 111494253 gb ABH063 ( 129)	21	15.9	8.9	1.000	1.000	3
gi 111120424	gi 111120424 gb ABH063 ( 129)	21	15.9	8.9	1.000	1.000	3
gi 111120428	gi 111120428 gb ABH063 ( 129)	21	15.9	8.9	1.000	1.000	3
gi 111120420	gi 111120420 gb ABH063 ( 129)	21	15.9	8.9	1.000	1.000	3
gi 111120432	gi 111120432 gb ABH063 ( 129)	21	15.9	8.9	1.000	1.000	3
gi 730049	gi 730049 sp P38950.2 MPA ( 160)	21	16.1	9.6	1.000	1.000	3
gi 1545895	gi 1545895 emb CAB02216. ( 161)	21	16.1	9.6	1.000	1.000	3
gi 24418862	gi 24418862 sp P00630.3 ( 167)	21	16.1	9.8	1.000	1.000	3

>>>CV127\_rc\_5842\_5850\_frame0, 3 aa vs

/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

>>gi|29465664 gi|29465664|gb|AAL92578.1| allergen Ole e (123 aa)

initn: 21 initl: 21 opt: 21 Z-score: 84.9 bits: 15.8 E(): 8.8

Smith-Waterman score: 21; 100.000% identity (100.000% similar) in 3 aa overlap (1-3:33-35)

CV127\_

GNK

:::

gi|294 MRGTAGVPDQVPVPTPTSPVPTSSSPVPKPPTQGNKKWCVPKAEATDAQLSNIDYVCSQS  
10 20 30 40 50 60

gi|294 GMDCGPIQANGACFNPNPTVRAHASYAMNSWYQSKGRNDFDCDFSGTGAITSSDPSNGSCS  
70 80 90 100 110 120

gi|294 FLS

>>gi|111494253 gi|111494253|gb|ABH06347.1| Blo t 21 alle (129 aa)

initn: 21 initl: 21 opt: 21 Z-score: 84.8 bits: 15.9 E(): 8.9

Smith-Waterman score: 21; 100.000% identity (100.000% similar) in 3 aa overlap (1-3:61-63)

gi|111 MKFIIALAALIAVACALPVSNDFRHEFDHMIVNAATQRFHEIEKFLHLHITHEVDDLEKT  
10 20 30 40 50 60

CV127\_ GNK

```
gi|111 GNKDEKARLLRELTVSEAFIEGSRGYFQRELKRTDLDLLEKFNFEALATGDLLKDLKA
      70      80      90      100      110      120
```

gi|111 LQKRVQDSE

```
>>gi|111120424 gi|111120424|gb|ABH06346.1| Blo t 21 alle (129 aa)
  initn: 21 init1: 21 opt: 21 Z-score: 84.8 bits: 15.9 E(): 8.9
Smith-Waterman score: 21; 100.000% identity (100.000% similar) in 3 aa overlap
(1-3:61-63)
```

```
gi|111 MKFIIALAALIAVACALPVSNDNFRHEFDHMIVNTATQRFHEIEKFLHITHEVDDLEKT
      10      20      30      40      50      60
```

CV127\_ GNK

```
gi|111 GNKDEKARPLRELTVSEAFIEGSRGYFQRELKRTDLDLLEKFNFEALATGDLLKDLKA
      70      80      90      100      110      120
```

gi|111 LQKRVQDSE

```
>>gi|111120428 gi|111120428|gb|ABH06348.1| Blo t 21 alle (129 aa)
  initn: 21 init1: 21 opt: 21 Z-score: 84.8 bits: 15.9 E(): 8.9
Smith-Waterman score: 21; 100.000% identity (100.000% similar) in 3 aa overlap
(1-3:61-63)
```

```
gi|111 MKFIIALAALIAVACALPVSNDNFRHEFDHMIINTATQRFHEIEKFLHITHEVDDLEKT
      10      20      30      40      50      60
```

CV127\_ GNK

```
gi|111 GNKDEKARLLRELTVSEAFIEGSRGYFQRELKRTDLDLLEKFNFEALATGDLLKDLKA
      70      80      90      100      110      120
```

gi|111 LQKRVQDSE

```
>>gi|111120420 gi|111120420|gb|ABH06344.1| Blo t 21 alle (129 aa)
  initn: 21 init1: 21 opt: 21 Z-score: 84.8 bits: 15.9 E(): 8.9
Smith-Waterman score: 21; 100.000% identity (100.000% similar) in 3 aa overlap
(1-3:61-63)
```

```
gi|111 MKFIIALAALIAVACALPVSNDNFRHEFDHMIVNTATQRFHEIEKFLHITHEVDDLEKT
      10      20      30      40      50      60
```

CV127\_ GNK

```
gi|111 GNKDEKARLLRELTVSEAFIEGSRGYFQRELKRTDLDLLEKFNFEALATGDLLKDLKA
      70      80      90      100      110      120
```

gi|111 LQKRVQDSE

```
>>gi|111120432 gi|111120432|gb|ABH06350.1| Blo t 21 alle (129 aa)
  initn: 21 init1: 21 opt: 21 Z-score: 84.8 bits: 15.9 E(): 8.9
Smith-Waterman score: 21; 100.000% identity (100.000% similar) in 3 aa overlap
(1-3:61-63)
```

gi|111 MKFIIALAALIAVACALPVSNDFRHEFDHMIVNTATQRFHEIEKFLHITHEVDDLEET  
10 20 30 40 50 60

CV127\_ GNK  
:::

gi|111 GNKDEKARLLRELTVSEAFIEGSRGYFQRELKRTDLDLLEKFNFEAALATGDLLLKDLKA  
70 80 90 100 110 120

gi|111 LQKRVQDSE

>>gi|730049 gi|730049|sp|P38950.2|MPAC2\_CARBE RecName: F (160 aa)  
initn: 21 init1: 21 opt: 21 Z-score: 84.3 bits: 16.1 E(): 9.6  
Smith-Waterman score: 21; 100.000% identity (100.000% similar) in 3 aa overlap  
(1-3:27-29)

CV127\_ GNK  
:::  
gi|730 MGVFNIEAETTSVIPAAARLFKAFILDGNKLIPKVSPQAVSSVENVEGNGGPGTIKKITFS  
10 20 30 40 50 60

gi|730 EGSPVKYVKERVEEIDHTNFKYNYTVIEGDVLGDKLEKVSHELKIVAAPGGGSIVKISSK  
70 80 90 100 110 120

gi|730 FHAKGYHEVNAAEMKGAKEMAEKLLRAVESYLLAHTAEYN  
130 140 150 160

>>gi|1545895 gi|1545895|emb|CAB02216.1| pollen allergen (161 aa)  
initn: 21 init1: 21 opt: 21 Z-score: 84.2 bits: 16.1 E(): 9.6  
Smith-Waterman score: 21; 100.000% identity (100.000% similar) in 3 aa overlap  
(1-3:27-29)

CV127\_ GNK  
:::  
gi|154 MGVFNIEAETTSVIPAAARLFKAFILDGNKLIPKVSPQAVSSVENVEGNGGPGTIKKITFS  
10 20 30 40 50 60

gi|154 EGSPVKYVKERVEEVDHTNFKYSYTVIEGGFVGDKVEKICNEIKIVAAPDGGGSILKITSK  
70 80 90 100 110 120

gi|154 YHTKGDHEVPAEHIKGGKERVEGLLKPVEAYLLAHTAEYNN  
130 140 150 160

>>gi|24418862 gi|24418862|sp|P00630.3|PA2\_APIME RecName: (167 aa)  
initn: 21 init1: 21 opt: 21 Z-score: 84.1 bits: 16.1 E(): 9.8  
Smith-Waterman score: 21; 100.000% identity (100.000% similar) in 3 aa overlap  
(1-3:45-47)

CV127\_ GNK  
:::  
gi|244 MQVVLGSLFLLLLSTSHGWQIRDRIKDNELEERIIYPGTLWCGHGNKSSGPNELGRFKHT  
10 20 30 40 50 60

gi|244 DACCRTHDMCPDVMMSAGESKHGLTNTASHTRLSCDCDDKFYDCLKNSADTISSYFVGKMY  
70 80 90 100 110 120

gi|244 FNLIDTKCYKLEHPVTGCGERTGRCLHYTVDKSKPKVYQWFDLRKY

130 140 150 160

3 residues in 1 query sequences  
307888 residues in 1386 library sequences  
Scomplib [34.26]  
start: Wed Mar 11 06:36:40 2009 done: Wed Mar 11 06:36:40 2009  
Total Scan time: 0.040 Total Display time: 0.010

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

## CV127\_rc\_5850\_5876\_frame2

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

Please cite:

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

CV127\_rc\_5850\_5876\_frame2, 9 aa  
vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences  
Expectation\_n fit:  $\rho(\ln(x)) = 3.1166 \pm 0.00267$ ;  $\mu = 1.4920 \pm 0.138$   
mean\_var=15.1669  $\pm 3.907$ , 0's: 125 Z-trim: 125 B-trim: 10 in 1/36  
Lambda= 0.329326

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

	opt	bits	E(1386)	%_id	%_sim	alen
gi 170738 gi 170738 gb AAA34289.1  ( 327)	35	19.7	4.8	0.556	0.889	9
gi 55859456 gi 55859456 emb CAH9263 ( 520)	35	19.7	7.7	0.444	0.889	9
gi 1364212 gi 1364212 emb CAA44343. ( 88)	29	16.9	9	0.600	1.000	5

>>>CV127\_rc\_5850\_5876\_frame2, 9 aa vs  
/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library  
  
>>gi|170738 gi|170738|gb|AAA34289.1| gamma-gliadin (327 aa)  
initn: 32 init1: 32 opt: 35 Z-score: 89.7 bits: 19.7 E(): 4.8  
Smith-Waterman score: 35; 55.556% identity (88.889% similar) in 9 aa overlap  
(1-9:312-320)

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

gi|170 HQPQQQVPQPQQPQQPFLQPQQPFPQQPQQPFPQTQQPQQPFPQQPQQPFPQTQQPQQPFP
      70          80          90         100         110         120

gi|170 PQQPQQPFPQTQQPQQPFPQLQQPQQPFPQPQQQLPQPQQPQQSFPQQQRPFIQPSLQQQ
     130         140         150         160         170         180

gi|170 LNPCKNILLQQSKPASLVSSLWSIIWPQSDCQVMRQQCCQQLAQIPQQLQCAAIHVVHS
     190         200         210         220         230         240

gi|170 IIMQQQQQQQQQQGIDIFLPLSQHEQVGQGSLSVQGGIIQPQQPAQLEAIRSLVLQTLPS
     250         260         270         280         290         300
```



```
CV127_          IRRAPFVQL
                : ::::...
gi|170 MCNVYVPPECSIMRAPFASIVAGIGGQ
                310          320
```

```
>>gi|55859456 gi|55859456|emb|CAH92630.1| pollen allerge (520 aa)
  initn: 28 initl: 28 opt: 35 Z-score: 86.0 bits: 19.7 E(): 7.7
Smith-Waterman score: 35; 44.444% identity (88.889% similar) in 9 aa overlap
(1-9:331-339)
```

```
gi|558 NSRAFALVPLLICVLSCHAAVSIAAPVPAKEDFFGCLVKEIPARLLYAKSSPAFPTVLA
                10          20          30          40          50          60

gi|558 QTIRNSRWSSPQSVKPLYIITPTNASHIQSAVVCGRRHGVRIVRSGGHDYEGLSYRSER
                70          80          90          100         110         120

gi|558 PEAFAVVDLNKMRVVVDGKARTAWVDSGAQLGELYAIAKNSPVLAFFPAGVCPTIGVGG
                130         140         150         160         170         180

gi|558 NFAGGGFGMLLRKYGIAAENVIDVKVVDANGTLLDKSSMSADHFWAVRGGGGESFGIVVS
                190         200         210         220         230         240

gi|558 WQVKLLPVPPTVTVFKIPKTVQEGAVELINKWQLVAPALPDDLIRIIAFGGTAKFEAMY
                250         260         270         280         290         300
```

```
CV127_          IRRAPFVQL
                :. :::::
gi|558 LGTCKALTPLMSSRFPELGMNASHCNEMPIKSVFPIHLGKQATLSDLLNRNNTFKPFAE
                310         320         330         340         350         360

gi|558 YKSDYVYQVPVKPVWAQIFVWLVKPGAGIMVMDPYGAAISATPEAATPFPHRKDVLFNIQ
                370         380         390         400         410         420

gi|558 YVNYWFDEAGGAAPLQWSKDMYRFMEPYVSKNPRQAYANYRDIDLGRNEVVNDISTYASG
                430         440         450         460         470         480

gi|558 KVGGEKYFKGNFQRLAITKGKVDPDQDYFRNEQSIPPLLK
                490         500         510         520
```

```
>>gi|1364212 gi|1364212|emb|CAA44343.1| fel d I chain 1 (88 aa)
  initn: 25 initl: 25 opt: 29 Z-score: 84.8 bits: 16.9 E(): 9
Smith-Waterman score: 29; 60.000% identity (100.000% similar) in 5 aa overlap
(1-5:1-5)
```

```
CV127_ IRRAPFVQL
        :. ::::
gi|136 VRRSPSTLPYCCGHSRDCEICPAVKRDVDLFLTGTDPDEYVEQVAQYNALPVLLENARILK
                10          20          30          40          50          60

gi|136 NCVDAKMTEEDKENALSVLDKIYTSPLC
                70          80
```

```
9 residues in 1 query sequences
307888 residues in 1386 library sequences
Scomplib [34.26]
start: Wed Mar 11 06:56:50 2009 done: Wed Mar 11 06:56:51 2009
Total Scan time: 0.060 Total Display time: 0.010
```

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

## CV127\_rc\_5854\_5946\_frame0

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

Please cite:

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

CV127\_rc\_5854\_5946\_frame0, 31 aa

vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences

Expectation\_n fit:  $\rho(\ln(x)) = 3.6296 \pm 0.00313$ ;  $\mu = 4.5307 \pm 0.161$

mean\_var=29.2688  $\pm$  8.212, 0's: 24 Z-trim: 24 B-trim: 64 in 1/40

Lambda= 0.237067

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

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

The best scores are:		opt	bits	E(1386)	%_id	%_sim	alen
gi 2506461	gi 2506461 sp P02224.2 G ( 162)	49	22.0	1.6	0.435	0.652	23
gi 121237	gi 121237 sp P02227.1 GLB ( 151)	47	21.3	2.5	0.400	0.700	20
gi 121259	gi 121259 sp P02228.1 GLB ( 151)	46	21.0	3.1	0.400	0.650	20
gi 132270	gi 132270 sp P15252.2 REF ( 138)	42	19.6	7.5	0.467	0.800	15
gi 1707908	gi 1707908 sp P02222.2 G ( 160)	42	19.6	8.5	0.400	0.650	20

>>>CV127\_rc\_5854\_5946\_frame0, 31 aa vs

/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

>>gi|2506461 gi|2506461|sp|P02224.2|GLB6\_CHITH RecName: (162 aa)

initn: 34 initl: 34 opt: 49 Z-score: 98.1 bits: 22.0 E(): 1.6

Smith-Waterman score: 49; 43.478% identity (65.217% similar) in 23 aa overlap (8-29:35-56)

		10	20	30
CV127_		GGPLLFSFNIKILIFYCRFK-LPNISLRFPIH		
		::	::	::
gi 250	MKFLVLALCIAAASAAVLTTEQADLVKKTWSTVKFN-EVDILYAVFKAYPDIMAKFPQFA			
		10	20	30

gi 250	GKDLDSIKDSAAFATHATRIVSFLSEVISLAGSDANIPAIONLAKELATSHKPRGVSKDQ
	60 70 80 90 100 110

gi 250	FTEFRTALFTYLKAHINFDGPTETAWTLALDTTYAMLFSAMDS
	120 130 140 150 160

>>gi|121237 gi|121237|sp|P02227.1|GLB8\_CHITH RecName: Fu (151 aa)

initn: 36 initl: 36 opt: 47 Z-score: 94.8 bits: 21.3 E(): 2.5

Smith-Waterman score: 47; 40.000% identity (70.000% similar) in 20 aa overlap (11-29:24-43)

		10	20	30
CV127_		GGPLLFSFNIKILIFYCRFKL-PNISLRFPIH		
		::	::	::
gi 121	AVTPMSADQLALFKSSWNTVKHNEVDILYAVFKANPDIQAKFPQFAGKDLDSIKDSADFA			
		10	20	30

gi 121	VHSGRIVGFFSEVIGLIGNPENRPALKTLIDGLASSHKARGIEKAQFEFFRASLVLDYLSH
--------	---

```

70      80      90      100     110     120

gi|121 HLDWNDTMKSTWDLALNNMFFYILHALEVAQ
      130      140      150

>>gi|121259 gi|121259|sp|P02228.1|GLB10_CHITH RecName: F (151 aa)
  initn: 34 initl: 34 opt: 46 Z-score: 93.0 bits: 21.0 E(): 3.1
Smith-Waterman score: 46; 40.000% identity (65.000% similar) in 20 aa overlap
(11-29:26-45)

      10      20      30
CV127_      GGPLLFSFNIKILIFYCRFKL-PNISLRFPIH
      .. :. :. :. :. :.
gi|121 DPEWHTLDAHEVEQVQATWKAVSHDEVEILYTVFKAHPDIMAKFPKFAGKDLEAIKDTAD
      10      20      30      40      50      60

gi|121 FAVHASRIIGFFGEYVTLLGSSGNQAARTLLHDLGVFHKTRGITKAQFGEFRETMAYL
      70      80      90      100     110     120

gi|121 KGHNKWNADISHSWDDAFDKAFSVIFEVLES
      130      140      150

>>gi|132270 gi|132270|sp|P15252.2|REF_HEVBR RecName: Ful (138 aa)
  initn: 30 initl: 30 opt: 42 Z-score: 86.2 bits: 19.6 E(): 7.5
Smith-Waterman score: 42; 46.667% identity (80.000% similar) in 15 aa overlap
(15-28:64-78)

      10
CV127_      GGPLLFSFNIK

gi|132 MAEDEDNQGGQGEGLKYLGFVQDAATYAVTTFSNVYLFADKSGPLQPGVDIIEGPVKNV
      10      20      30      40      50      60

      20      30
CV127_ ILIFYCRFK-LPNISLRFPIH
      .: :. :. :. :. :.
gi|132 AVPLYNRFSYIPNGALKFVDSTVVASVTIIDRSLPPIVKDASIQVVSIRAAPAEARSLSLA
      70      80      90      100     110     120

gi|132 SSLPGQTKILAKVIFYGEN
      130

>>gi|1707908 gi|1707908|sp|P02222.2|GLB2_CHITH RecName: (160 aa)
  initn: 34 initl: 34 opt: 42 Z-score: 85.2 bits: 19.6 E(): 8.5
Smith-Waterman score: 42; 40.000% identity (65.000% similar) in 20 aa overlap
(11-29:37-56)

      10      20      30
CV127_      GGPLLFSFNIKILIFYCRFKL-PNISLRFPIH
      .. :. :. :. :. :.
gi|170 MKFLVLALCIAAAVAAPLSADEASLVRGSWAQVKHSEVDILYYIFKANPDIMAKFPQFAG
      10      20      30      40      50      60

gi|170 KDLETLKGTGQFATHAGRIVGFVSEIVALMGNSANMPAMETLIKDMAANHKARGIPKAQF
      70      80      90      100     110     120

gi|170 NEFRASLVSYLQSKVSWNDSLGAAWTQGLDNVFNMMFSYL
      130      140      150      160

```

31 residues in 1 query sequences  
 307888 residues in 1386 library sequences  
 Scomplib [34.26]  
 start: Wed Mar 11 06:38:47 2009 done: Wed Mar 11 06:38:48 2009  
 Total Scan time: 0.070 Total Display time: 0.000

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

## CV127\_rc\_5880\_5975\_frame2

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

Please cite:

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

CV127\_rc\_5880\_5975\_frame2, 32 aa  
 vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences  
 Expectation\_n fit:  $\rho(\ln(x)) = 3.4429 \pm 0.0032$ ;  $\mu = 6.8338 \pm 0.167$   
 mean\_var=26.0573  $\pm$  6.975, 0's: 24 Z-trim: 25 B-trim: 0 in 0/42  
 Lambda= 0.251252

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2  
 join: 36, opt: 20, open/ext: -10/-2, width: 16  
 The best scores are: opt bits E(1386) %\_id %\_sim alen  
 gi|21748151 gi|21748151|emb|CAD3816 ( 125) 47 21.3 2.1 0.636 0.727 11

>>>CV127\_rc\_5880\_5975\_frame2, 32 aa vs  
 /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library  
 >>gi|21748151 gi|21748151|emb|CAD38166.1| putative nucle (125 aa)  
 initn: 47 initl: 47 opt: 47 Z-score: 96.1 bits: 21.3 E(): 2.1  
 Smith-Waterman score: 47; 63.636% identity (72.727% similar) in 11 aa overlap  
 (8-18:56-66)

CV127_						10
						YKDFDILLPFQA
						:::
gi 217	MSDFNAIAQQFT	EFYKTFD	TDRAQLAP	LYRENSML	TFEQSPFL	GTANIVGKLQELPFQR
	10	20	30	40	50	60
	20	30				
CV127_	SKHQFALS	YTLNGEGAL	VVLN			
	:::	:				
gi 217	IEHQVATVDAQ	PSNESGGIL	VVVGALLVEE	ERRPMSYT	QTFQLLPAD	GAYYVFNDVFRL
	70	80	90	100	110	120
gi 217	VYPAA					

32 residues in 1 query sequences  
 307888 residues in 1386 library sequences  
 Scomplib [34.26]  
 start: Wed Mar 11 06:59:00 2009 done: Wed Mar 11 06:59:00 2009  
 Total Scan time: 0.080 Total Display time: 0.010

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

## CV127\_rc\_5884\_5890\_frame1

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

Please cite:

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

CV127\_rc\_5884\_5890\_frame1, 2 aa

vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences

MLE\_cen statistics: Lambda= 0.3260; K=0.003271 (cen=69)

FASTA (3.5 Sept 2006) function [optimized, BL50 matrix (15:-5)] ktup: 2  
join: 21, opt: 20, open/ext: -10/-2, width: 16

The best scores are:

			opt	bits	E(1386)	%_id	%_sim	alen
gi 37725377	gi 37725377 gb AA033897	( 52)	15	15.3	3.5	1.000	1.000	2
gi 1247373	gi 1247373 emb CAA01909.	( 71)	15	15.3	4.8	1.000	1.000	2
gi 1247375	gi 1247375 emb CAA01910.	( 73)	15	15.3	5	1.000	1.000	2
gi 14423846	gi 14423846 sp O82040.1	( 78)	15	15.3	5.3	1.000	1.000	2
gi 4538529	gi 4538529 emb CAB39376.	( 81)	15	15.3	5.5	1.000	1.000	2
gi 1871507	gi 1871507 emb CAA62634.	( 82)	15	15.3	5.6	1.000	1.000	2
gi 2129801	gi 2129801 pir S65144 p	( 83)	15	15.3	5.7	1.000	1.000	2
gi 2129802	gi 2129802 pir S65145 p	( 83)	15	15.3	5.7	1.000	1.000	2
gi 121308880	gi 121308880 dbj BAF43	( 84)	15	15.3	5.7	1.000	1.000	2
gi 3337403	gi 3337403 gb AAD05375.1	( 84)	15	15.3	5.7	1.000	1.000	2
gi 14423850	gi 14423850 sp Q39419.1	( 85)	15	15.3	5.8	1.000	1.000	2
gi 29465668	gi 29465668 gb AAL92871	( 86)	15	15.3	5.9	1.000	1.000	2
gi 939932	gi 939932 emb CAA51775.1	( 88)	15	15.3	6	1.000	1.000	2
gi 94400907	gi 94400907 ref NP_0010	( 92)	15	15.3	6.3	1.000	1.000	2
gi 94400875	gi 94400875 ref NP_0010	( 94)	15	15.3	6.4	1.000	1.000	2
gi 14423759	gi 14423759 sp P93124.1	( 96)	15	15.3	6.6	1.000	1.000	2
gi 126387	gi 126387 sp P14948.1 MPA	( 97)	15	15.3	6.6	1.000	1.000	2
gi 126386	gi 126386 sp P14947.1 MPA	( 97)	15	15.3	6.6	1.000	1.000	2
gi 416636	gi 416636 sp P00304.2 MPA	( 101)	15	15.3	6.9	1.000	1.000	2
gi 51316532	gi 51316532 sp Q9LD79.2	( 110)	15	15.3	7.5	1.000	1.000	2
gi 1304218	gi 1304218 dbj BAA07774.	( 113)	15	15.3	7.7	1.000	1.000	2
gi 160285626	gi 160285626 pdb 2JMH	( 119)	15	15.3	8.1	1.000	1.000	2
gi 4007655	gi 4007655 emb CAA10348.	( 122)	15	15.3	8.3	1.000	1.000	2
gi 4007636	gi 4007636 emb CAA10350.	( 122)	15	15.3	8.3	1.000	1.000	2
gi 4007040	gi 4007040 emb CAA10345.	( 122)	15	15.3	8.3	1.000	1.000	2
gi 1171009	gi 1171009 sp P43214.1 M	( 122)	15	15.3	8.3	1.000	1.000	2
gi 48428170	gi 48428170 sp Q9NFQ4.1	( 125)	15	15.3	8.5	1.000	1.000	2
gi 111120432	gi 111120432 gb ABH063	( 129)	15	15.3	8.8	1.000	1.000	2
gi 111494253	gi 111494253 gb ABH063	( 129)	15	15.3	8.8	1.000	1.000	2
gi 32363444	gi 32363444 sp P80207.1	( 129)	15	15.3	8.8	1.000	1.000	2
gi 111120424	gi 111120424 gb ABH063	( 129)	15	15.3	8.8	1.000	1.000	2
gi 111120420	gi 111120420 gb ABH063	( 129)	15	15.3	8.8	1.000	1.000	2
gi 111120428	gi 111120428 gb ABH063	( 129)	15	15.3	8.8	1.000	1.000	2
gi 33329758	gi 33329758 gb AAQ10281	( 131)	15	15.3	8.9	1.000	1.000	2
gi 156938917	gi 156938917 gb ABU974	( 131)	15	15.3	8.9	1.000	1.000	2
gi 118638268	gi 118638268 gb ABL093	( 131)	15	15.3	8.9	1.000	1.000	2
gi 51860756	gi 51860756 gb AAU11502	( 131)	15	15.3	8.9	1.000	1.000	2
gi 33329754	gi 33329754 gb AAQ10279	( 132)	15	15.3	9	1.000	1.000	2
gi 33329756	gi 33329756 gb AAQ10280	( 132)	15	15.3	9	1.000	1.000	2
gi 47117012	gi 47117012 sp Q7M4I5.1	( 134)	15	15.3	9.1	1.000	1.000	2
gi 111120436	gi 111120436 gb ABH063	( 134)	15	15.3	9.1	1.000	1.000	2

gi 7435005	gi 7435005 pir A59055 p	( 134)	15	15.3	9.1	1.000	1.000	2
gi 111120450	gi 111120450 gb ABH063	( 134)	15	15.3	9.1	1.000	1.000	2
gi 4204917	gi 4204917 gb AAD10850.1	( 134)	15	15.3	9.1	1.000	1.000	2
gi 19039	gi 19039 emb CAA42832.1  L	( 134)	15	15.3	9.1	1.000	1.000	2
gi 24638082	gi 24638082 sp Q9BMK4.1	( 134)	15	15.3	9.1	1.000	1.000	2
gi 1362129	gi 1362129 pir E53806 m	( 136)	15	15.3	9.3	1.000	1.000	2
gi 1362137	gi 1362137 pir H53806 m	( 136)	15	15.3	9.3	1.000	1.000	2
gi 1362132	gi 1362132 pir A38968 m	( 137)	15	15.3	9.3	1.000	1.000	2
gi 1362128	gi 1362128 pir I53806 m	( 137)	15	15.3	9.3	1.000	1.000	2
gi 24473800	gi 24473800 gb AAL91665	( 138)	15	15.3	9.4	1.000	1.000	2
gi 132270	gi 132270 sp P15252.2 REF	( 138)	15	15.3	9.4	1.000	1.000	2
gi 145313982	gi 145313982 gb ABP586	( 140)	15	15.3	9.6	1.000	1.000	2
gi 145313992	gi 145313992 gb ABP586	( 140)	15	15.3	9.6	1.000	1.000	2
gi 145313988	gi 145313988 gb ABP586	( 140)	15	15.3	9.6	1.000	1.000	2
gi 112745	gi 112745 sp P23110.1 2SS	( 141)	15	15.3	9.6	1.000	1.000	2
gi 3256212	gi 3256212 emb CAA54819.	( 145)	15	15.3	9.9	1.000	1.000	2
gi 1362131	gi 1362131 pir C53806 m	( 145)	15	15.3	9.9	1.000	1.000	2
gi 631912	gi 631912 pir S43243 all	( 145)	15	15.3	9.9	1.000	1.000	2
gi 631911	gi 631911 pir S43242 all	( 145)	15	15.3	9.9	1.000	1.000	2
gi 631913	gi 631913 pir S43244 all	( 145)	15	15.3	9.9	1.000	1.000	2
gi 1362136	gi 1362136 pir B53806 m	( 145)	15	15.3	9.9	1.000	1.000	2
gi 14423737	gi 14423737 sp O82015.2	( 145)	15	15.3	9.9	1.000	1.000	2
gi 34978692	gi 34978692 gb AAQ83588	( 146)	15	15.3	10	1.000	1.000	2
gi 2465127	gi 2465127 emb CAA73038.	( 146)	15	15.3	10	1.000	1.000	2
gi 2465131	gi 2465131 emb CAA73036.	( 146)	15	15.3	10	1.000	1.000	2

&gt;&gt;&gt;CV127\_rc\_5884\_5890\_framel, 2 aa vs

/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

&gt;&gt;gi|37725377 gi|37725377|gb|AAO33897.1| Ole e 3 allerge (52 aa)

initn: 15 initl: 15 opt: 15 Z-score: 92.0 bits: 15.3 E(): 3.5

Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap (1-2:8-9)

CV127\_ RF

::

gi|377 EHERIFKRFDAKGDKKISSSELGETLKPLGSVTLEEIQRMMAEIDTDGDGFL

10 20 30 40 50

&gt;&gt;gi|1247373 gi|1247373|emb|CAA01909.1| B1 protein aller (71 aa)

initn: 15 initl: 15 opt: 15 Z-score: 89.6 bits: 15.3 E(): 4.8

Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap (1-2:4-5)

CV127\_ RF

::

gi|124 GTRRFDTNGDGKISLAELTDALRTLGSTSADEVQRMMAEIDTDGDGFIDFDEFISFCNAN

10 20 30 40 50 60

gi|124 PGLMKDVAKVF

70

&gt;&gt;gi|1247375 gi|1247375|emb|CAA01910.1| B4 protein aller (73 aa)

initn: 15 initl: 15 opt: 15 Z-score: 89.4 bits: 15.3 E(): 5

Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap (1-2:6-7)

CV127\_ RF

```

::
gi|124 GTSFKRFDTNNGDGKISLAELTDALRTLGGSTSADEVQRMMAEIDTDGDGFIDFDEFISFCN
      10      20      30      40      50      60

gi|124 ANPGLMKDVAKVF
      70

>>gi|14423846 gi|14423846|sp|O82040.1|POLC7_PHLPR RecNam (78 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 88.9 bits: 15.3 E(): 5.3
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:11-12)

CV127_ RF
      ::
gi|144 MADDMERIFKRFDTNNGDGKISLSELTALRTLGGSTSADEVQRMMAEIDTDGDGFIDFNEF
      10      20      30      40      50      60

gi|144 ISFCNANPGLMKDVAKVF
      70

>>gi|4538529 gi|4538529|emb|CAB39376.1| Cop c1 allergen (81 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 88.6 bits: 15.3 E(): 5.5
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:1-2)

CV127_ RF
      ::
gi|453 RFLPSSSHLNPQHLPWLVHPAPVLLLPVLPQLKPVVAHPLLLLPLDTTTLHMPPLLLQLQL
      10      20      30      40      50      60

gi|453 PPLLSQGNPACSPKWLQLLVP
      70      80

>>gi|1871507 gi|1871507|emb|CAA62634.1| calcium-binding (82 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 88.5 bits: 15.3 E(): 5.6
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:15-16)

CV127_ RF
      ::
gi|187 KTMADTGDMEHIFKRFDTNNGDGKISLAELTDALRTLGGSTSADEVQRMMAEIDTDGDGFID
      10      20      30      40      50      60

gi|187 FDEFISFCNANPGLMKDVAKVF
      70      80

>>gi|2129801 gi|2129801|pir||S65144 pollen allergen grou (83 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 88.4 bits: 15.3 E(): 5.7
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:12-13)

CV127_ RF
      ::
gi|212 MADATEKTEHDFRFFKKFDANGDGTISSTELGDALKNLGSVTHDDIKRMAEIDTDGDGFI
      10      20      30      40      50      60

gi|212 SYQEFSDFAKANRGLMKDVAKIF
      70      80

```

```
>>gi|2129802 gi|2129802|pir||S65145 pollen allergen grou (83 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 88.4 bits: 15.3 E(): 5.7
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:12-13)
```

```
CV127_          RF
                ::
gi|212 MADATEKAEHDFRFFKKFDANGDGTISSTELGDALKNLGVSVTHDDIKRMMAEIDTDGDGFI
      10          20          30          40          50          60

gi|212 SYQEFSDFAKANRGLMKDVAKIF
      70          80
```

```
>>gi|121308880 gi|121308880|dbj|BAF43535.1| protease inh (84 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 88.3 bits: 15.3 E(): 5.7
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:18-19)
```

```
CV127_          RF
                ::
gi|121 MFQSTFFLVLMVCVATARFANKDHCPPNEEYNECGNPCQEKCDNGEPVICTYQCEHRCFC
      10          20          30          40          50          60

gi|121 KQGYVRLTEDGECVPEEFCKPIHY
      70          80
```

```
>>gi|3337403 gi|3337403|gb|AAD05375.1| calcium-binding p (84 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 88.3 bits: 15.3 E(): 5.7
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:17-18)
```

```
CV127_          RF
                ::
gi|333 MADDPQEVAEHERIFKRFDANGDGKISSSELGETLKTLSVTPPEEIQRMMAEIDTDGDGF
      10          20          30          40          50          60

gi|333 ISFEEFTVFARANRGLVKDVAKIF
      70          80
```

```
>>gi|14423850 gi|14423850|sp|Q39419.1|POLC4 BETVE RecNam (85 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 88.2 bits: 15.3 E(): 5.8
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:18-19)
```

```
CV127_          RF
                ::
gi|144 MADDHPQDKAERERIFKRFDANGDGKISAAELGEALKTLGSITPDEVKHMMAEIDTDGDG
      10          20          30          40          50          60

gi|144 FISFQEFTDFGRANRGLLKDVAKIF
      70          80
```

```
>>gi|29465668 gi|29465668|gb|AAL92871.1| pollen allergen (86 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 88.1 bits: 15.3 E(): 5.9
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:19-20)
```



```

CV127_      RF
           ::
gi|294 MAAEDTPQDIADRERIFKRFDITNGDGKISSSELGDALKTLGSVTPDEVRRMMAEIDTDGD
           10          20          30          40          50          60

gi|294 GFISFDEFTDFARANRGLVKDVKIF
           70          80

>>gi|939932 gi|939932|emb|CAA51775.1| allergen Lol p II (88 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 87.9 bits: 15.3 E(): 6
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:65-66)

gi|939 EFTVEKGSDEKNLALSIIKYSKEGDAMAEVELKEHGSNEWLALKKNGDGVWEIKSDKPLKG
           10          20          30          40          50          60

CV127_      RF
           ::
gi|939 PFNFRFVSEKGMWNVFDVVPADFKVGT
           70          80

>>gi|94400907 gi|94400907|ref|NP_001035360.1| allergen A (92 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 87.6 bits: 15.3 E(): 6.3
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:52-53)

CV127_      RF
           ::
gi|944 MSRLVLASFLLLAIFSMVLVGGFGGFGGGLGGRGKCPSEIFSRCDGRCQRFPCNVVVK
           10          20          30          40          50          60

gi|944 PLCIKICAPGCVCRLLGYLRNKKKVCVPRSKCG
           70          80          90

>>gi|94400875 gi|94400875|ref|NP_001035358.1| allergen A (94 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 87.4 bits: 15.3 E(): 6.4
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:52-53)

CV127_      RF
           ::
gi|944 MSRLVLASFLLLAIFSMVLVGGFGGFGGGLGGRGKCPSEIFSRCDGRCQRFPCNVVVK
           10          20          30          40          50          60

gi|944 PLCIKICAPGCVCRLLGYLRNKKKVCVPRSKCLPG
           70          80          90

>>gi|14423759 gi|14423759|sp|P93124.1|MPAG3_DACGL RecNam (96 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 87.2 bits: 15.3 E(): 6.6
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:67-68)

gi|144 VKVTFKVEKGSDDPKKLVLVDIKYTRPGDTLAEVELRQHGSEWEPLTKKGNLWEVKSSKPL
           10          20          30          40          50          60

CV127_      RF
           ::

```

```
gi|144 TGPFNFRFMSKGGMRNVFDEVIPTAFKIGTTYTPEE
      70          80          90
```

```
>>gi|126387 gi|126387|sp|P14948.1|MPAL3_LOLPR RecName: F (97 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 87.2 bits: 15.3 E(): 6.6
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:67-68)
```

```
gi|126 TKVDLTVEKGSDAKTLVLNIKYTRPGDTLAEVELRQHGSEEWEPMTKKGNLWEVKSAPKL
      10          20          30          40          50          60
```

```
CV127_      RF
           ::
```

```
gi|126 TGPMNFRFLSKGGMKNVFDEVIPTAFTVGKTYTPEYN
      70          80          90
```

```
>>gi|126386 gi|126386|sp|P14947.1|MPAL2_LOLPR RecName: F (97 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 87.2 bits: 15.3 E(): 6.6
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:69-70)
```

```
gi|126 AAPVEFTVEKGSDEKNLALSIKYNKEGDSMAEVELKEHGSNEWLALKKNGDGVWEIKSDK
      10          20          30          40          50          60
```

```
CV127_      RF
           ::
```

```
gi|126 PLKGPFNFRFVSEKGMNRNVFDDVVPADFKVGTTYKPE
      70          80          90
```

```
>>gi|416636 gi|416636|sp|P00304.2|MPAA3_AMBEL RecName: F (101 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 86.9 bits: 15.3 E(): 6.9
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:76-77)
```

```
gi|416 GKVYLVGGPELGGWKLQSDPRAYALWSARQQFKTTDVLWFNFTTGEDSVAEVWREEAYHA
      10          20          30          40          50          60
```

```
CV127_      RF
           ::
```

```
gi|416 CDIKDPIRLEPGGPDRTLLTPGSHFICTKDQKFVACVPGR
      70          80          90         100
```

```
>>gi|51316532 gi|51316532|sp|Q9LD79.2|PRR3_JUNVI RecName (110 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 86.2 bits: 15.3 E(): 7.5
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:62-63)
```

```
gi|513 AFLLAATLTISSHMQEAGAVKFDIKNQCGYTVWAAGLPGGGKRLDQGQTWTVNLAAGTAS
      10          20          30          40          50          60
```

```
CV127_      RF
           ::
```

```
gi|513 ARFWGRTGCTFDASGKGSCQTGDCGRQLSCTVSGAVPATLAEYTSQSDQDY
      70          80          90         100         110
```

```
>>gi|1304218 gi|1304218|dbj|BAA07774.1| allergenic prote (113 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 86.0 bits: 15.3 E(): 7.7
```

Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap (1-2:4-5)

CV127\_ RF  
::

gi|130 FGTRFRQRQCVGPRRTRRRWTTQQLRQDCCRQLAAVDDSWCRCSALNHMVGGIYRELGATD  
10 20 30 40 50 60

gi|130 VGHPMAEVFPGCRRGDLERAAAASLPAFCNVDPNGTGGVCYWLGYPRTPRTGH  
70 80 90 100 110

>>gi|160285626 gi|160285626|pdb|2JMH|A Chain A, Nmr Solu (119 aa)  
initn: 15 init1: 15 opt: 15 Z-score: 85.6 bits: 15.3 E(): 8.1  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap (1-2:89-90)

gi|160 GSQEHKPKKDDFRNEFDHLLIEQANHAIEKGEHQLLYLQHQDELNENKSKELQEKIIRE  
10 20 30 40 50 60

CV127\_ RF  
::

gi|160 LDVVCAMIEGAQGALERELKRTDLNILERFNYEEAQTLSKILLKDLKETEQKVVDIQTQ  
70 80 90 100 110

>>gi|4007655 gi|4007655|emb|CAA10348.1| pollen allergen (122 aa)  
initn: 15 init1: 15 opt: 15 Z-score: 85.4 bits: 15.3 E(): 8.3  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap (1-2:93-94)

gi|400 MSMASSSSSSLLAMAVLAALFAGAWCVPKVTFTVEKGSNEKHLAVLVKYE GDTMAEVELR  
10 20 30 40 50 60

CV127\_ RF  
::

gi|400 EHGSDEWVAMTKGEGGVWTFDSEEPLQGPFNFRFLTEKGMKNVFDVVPKYTIGATYAP  
70 80 90 100 110 120

gi|400 EE

>>gi|4007636 gi|4007636|emb|CAA10350.1| pollen allergen (122 aa)  
initn: 15 init1: 15 opt: 15 Z-score: 85.4 bits: 15.3 E(): 8.3  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap (1-2:93-94)

gi|400 MSMASSSSSSLLAMAVLAALFAGAWCVPKVTFTVEKGSNEKHLAVLVKYE GDTMAEVELR  
10 20 30 40 50 60

CV127\_ RF  
::

gi|400 EHGSDEWVAMTKGEGGVWTFDSEEPLQGPFNFRFLTEKGMKNVFDVVPKYTIGATYAP  
70 80 90 100 110 120

gi|400 EE

>>gi|4007040 gi|4007040|emb|CAA10345.1| pollen allergen (122 aa)  
initn: 15 init1: 15 opt: 15 Z-score: 85.4 bits: 15.3 E(): 8.3

Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap (1-2:93-94)

```
gi|400 MSMASSSSSGLLAMAVLAALFAGAWCVPKVTFTVEKGSNEKHLAVLVKYE GDTMAEVELR
      10      20      30      40      50      60
```

```
CV127_
      RF
      ::
gi|400 EHGSDEWVAMTKGEGGVWTFDSEEPLQGPFNFRFLTEKGMKNVFDVVP EKYTIGATYAP
      70      80      90      100      110      120
```

gi|400 EE

>>gi|1171009 gi|1171009|sp|P43214.1|MPAP2\_PHLPR RecName: (122 aa)  
initn: 15 initl: 15 opt: 15 Z-score: 85.4 bits: 15.3 E(): 8.3  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap (1-2:93-94)

```
gi|117 MSMASSSSSLLAMAVLAALFAGAWCVPKVTFTVEKGSNEKHLAVLVKYE GDTMAEVELR
      10      20      30      40      50      60
```

```
CV127_
      RF
      ::
gi|117 EHGSDEWVAMTKGEGGVWTFDSEEPLQGPFNFRFLTEKGMKNVFDVVP EKYTIGATYAP
      70      80      90      100      110      120
```

gi|117 EE

>>gi|48428170 gi|48428170|sp|Q9NFQ4.1|ALL22\_GLYDO RecNam (125 aa)  
initn: 15 initl: 15 opt: 15 Z-score: 85.2 bits: 15.3 E(): 8.5  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap (1-2:119-120)

```
gi|484 GKMFKFKDCGKGEVTELDITDCSGDFCVIHRGKPLTLEAKFAANQDTTKATIKVLAKVAGT
      10      20      30      40      50      60
```

```
CV127_
      RF
      ::
gi|484 PIQVPGLETDGCKFVKCPIKKGDPIDFKYTTTVPAILPKVKAEVTAELVGDHGV LACGRF
      70      80      90      100      110      120
```

gi|484 GRQVE

>>gi|111120432 gi|111120432|gb|ABH06350.1| Blo t 21 alle (129 aa)  
initn: 15 initl: 15 opt: 15 Z-score: 84.9 bits: 15.3 E(): 8.8  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap (1-2:39-40)

```
CV127_
      RF
      ::
gi|111 MKFIIALAALIAVACALPVSNDFRHEFDHMIVNTATQRFHEIEKFL LHITHEVDDLEET
      10      20      30      40      50      60
```

```
gi|111 GNKDEKARLLRELTVSEAFIEGSRGYFQRELKRTDLDLLEKFNFEAALATGDL LLLKDLKA
      70      80      90      100      110      120
```

gi | 111 LQKRVQDSE

```
>>gi|111494253 gi|111494253|gb|ABH06347.1| Blo t 21 alle (129 aa)
  initn: 15 init1: 15 opt: 15 Z-score: 84.9 bits: 15.3 E(): 8.8
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:39-40)
```

CV127\_ RF  
gi|111 MKFIIALAALIAVACALPVSNDFRHEFDHMIVNAATQRFHEIEKFL LHITHEVDDLEKT  
10 20 30 40 50 60

gi|111 GNKDEKARLLRELTVSEAFIEGSRGYFQRELKRTDLDLLEKFNFEEAALATGDLCLKDLKA  
70 80 90 100 110 120

gi | 111 LQKRVQDSE

```
>>gi|32363444 gi|32363444|sp|P80207.1|ALL1_BRAJU RecName (129 aa)
initn: 15 initl: 15 opt: 15 Z-score: 84.9 bits: 15.3 E(): 8.8
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:5-6)
```

CV127\_ RF  
::  
gi | 323 AGPFRFP<sup>10</sup>RCRKEF<sup>20</sup>QQAQHLRAC<sup>30</sup>QQWLHKQAMQSGSGPQ<sup>40</sup>PGPQQR<sup>50</sup>PL<sup>60</sup>LQ<sup>70</sup>CCNELHQEE

gi|323 PLCVCPTLKGASKAVKQQIRQQGQQGQQGQQQLQHEISRIYQTATHLPRVCNIPRSICP  
70 80 90 100 110 120

gi | 323 FQKTMPGPS

```
>>gi|111120424|gb|ABH06346.1| Blo t 21 alle (129 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 84.9 bits: 15.3 E(): 8.8
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:39-40)
```

CV127\_ RF  
::  
gi|111 MKFIIALAALIAVACALPVSDNFRHEFDHMIVNTATQRFHEIEKFL LHITHEVDDLEKT  
10 20 30 40 50 60

gi|111 GNKDEKARPLRELTVSEAFIEGSRGYFQRELKRTDLDLLEKFNFEAALATGDLLLLKDLKA  
70 80 90 100 110 120

gi | 111 LQKRVQDSE

```
>>gi|111120420 gi|111120420|gb|ABH06344.1| Blo t 21 alle (129 aa)
  initn: 15 init1: 15 opt: 15 Z-score: 84.9 bits: 15.3 E(): 8.8
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:39-40)
```

CV127\_ RF

```

::
gi|111 MKFIIALAALIIVACALPVSNDFRHEFDHMIVNTATQRFHEIEKFLHITHEVDDLEKT
      10      20      30      40      50      60

gi|111 GNKDEKARLLRELTVSEAFIEGSRGYFQRELKRTDLDLLEKFNFEAALATGDLLLKDLKA
      70      80      90     100     110     120

gi|111 LQKRVQDSE

```

```

>>gi|111120428 gi|111120428|gb|ABH06348.1| Blo t 21 alle (129 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 84.9 bits: 15.3 E(): 8.8
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:39-40)

```

```

CV127_ RF
      ::
gi|111 MKFIIALAALIIVACALPVSNDFRHEFDHMIINTATQRFHEIEKFLHITHEVDDLEKT
      10      20      30      40      50      60

gi|111 GNKDEKARLLRELTVSEAFIEGSRGYFQRELKRTDLDLLEKFNFEAALATGDLLLKDLKA
      70      80      90     100     110     120

gi|111 LQKRVQDSE

```

```

>>gi|33329758 gi|33329758|gb|AAQ10281.1| major pollen al (131 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 84.8 bits: 15.3 E(): 8.9
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:11-12)

```

```

CV127_ RF
      ::
gi|333 QDSYSGHCRARFITESEFIPGAGVRLQCKDGENGNVTFTEVGYTRAEGLYSMLIERDHK
      10      20      30      40      50      60

gi|333 NEFCEITLLSSSRKDCDEIPVEGWVKPSLKFMLNTVNGTTRTINPLGFFKKEALPKCPQV
      70      80      90     100     110     120

gi|333 FNKLGMYPPDM
      130

```

```

>>gi|156938917 gi|156938917|gb|ABU97480.1| allergen Tyr (131 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 84.8 bits: 15.3 E(): 8.9
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:126-127)

```

```

gi|156 MVQLNGSYKLEKSDNFDAFLKELGVNFVTRNLAKSASPTVEVIVDGDSTYIKTSSTLKNS
      10      20      30      40      50      60

gi|156 EIKFKLGEEFEEDRADGKKVQTSVTKEGDNKLVQVQKGDKPVTIVREFSEEGTLVTATVN
      70      80      90     100     110     120

CV127_ RF
      ::
gi|156 GATSVRFYKRQ
      130

```

```

>>gi|118638268 gi|118638268|gb|ABL09307.1| allergen Aca (131 aa)

```

initn: 15 initl: 15 opt: 15 Z-score: 84.8 bits: 15.3 E(): 8.9  
 Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
 (1-2:126-127)

gi|118 MVQINGSYKLEKSDNFDALFKELGLNFVTRNLAKSATPTVEVSVNGDSYTIKTASTLKNT  
 10 20 30 40 50 60

gi|118 EISFKLGEEFEERADGKTVKTVVNKESDTKFVQVQQGDKEVTIVREFSDEGLTVTATVN  
 70 80 90 100 110 120

CV127\_ RF  
 ::

gi|118 GVTSVRFYKRQ  
 130

>>gi|51860756 gi|51860756|gb|AAU11502.1| fatty acid-bidi (131 aa)  
 initn: 15 initl: 15 opt: 15 Z-score: 84.8 bits: 15.3 E(): 8.9  
 Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
 (1-2:126-127)

gi|518 MVQLNGSYKLEKSDNFDALFKELGVNFVTRNLAKSASPTVEVIVDGDSTYTIKTSSTLKNS  
 10 20 30 40 50 60

gi|518 EIKFKLGEEFEEDRADGKKVQTSVTKEGDNKLVQVQKGDKPVTIVREFSEEGLTVTATVN  
 70 80 90 100 110 120

CV127\_ RF  
 ::

gi|518 GVTSVRFYKRQ  
 130

>>gi|33329754 gi|33329754|gb|AAQ10279.1| major pollen al (132 aa)  
 initn: 15 initl: 15 opt: 15 Z-score: 84.8 bits: 15.3 E(): 9  
 Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
 (1-2:12-13)

CV127\_ RF  
 ::

gi|333 IRTVSQDTCCARFITESEFIPGAGVRLQCKDGENGNVTFTEVGYTRAEGLYSMLIERDH  
 10 20 30 40 50 60

gi|333 KNEFCEITLLSSSRKDCDEIPVEGWVKPSLKFMLNTVNGTTRTINPLGFFKKEALPKCPQ  
 70 80 90 100 110 120

gi|333 VFNKLGMYPPDM  
 130

>>gi|33329756 gi|33329756|gb|AAQ10280.1| major pollen al (132 aa)  
 initn: 15 initl: 15 opt: 15 Z-score: 84.8 bits: 15.3 E(): 9  
 Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
 (1-2:12-13)

CV127\_ RF  
 ::

gi|333 IRTVFRVTCRARFITESEFIPGAGVRLQCKDGENGNVTFTEVGYTRAEGLYSMLIERDH  
 10 20 30 40 50 60

gi|333 KNEFCEITLLSSSRKDCDEIPVEGWVKPSLKFMLNTVNGTTRTINPLGFFKKEALPKCPQ

```

70      80      90      100      110      120

gi|333 VFNKLGMYPDDM
130

>>gi|47117012 gi|47117012|sp|Q7M4I5.1|PA2_APIDO RecName: (134 aa)
initn: 15 initl: 15 opt: 15 Z-score: 84.6 bits: 15.3 E(): 9.1
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:23-24)

CV127_ RF
::
gi|471 IIYPGTLWCGHGNVSSSPDELGRFKHTDSCCRSHDMCPDVMSAGESKHGLTNTASHTRLS
10      20      30      40      50      60

gi|471 CDCDDKFYDCLKNSSDTISSYFVGEMYFNILDTKCYKLEHPVTGCGKRTEGRCLNYTVDK
70      80      90      100      110      120

gi|471 SKPKVYQWFDLRKY
130

>>gi|111120436 gi|111120436|gb|ABH06352.1|Blo t 5 aller (134 aa)
initn: 15 initl: 15 opt: 15 Z-score: 84.6 bits: 15.3 E(): 9.1
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:104-105)

gi|111 MKFAIVLIACFAASVLAQGHKPKKDDFRNEFDHLLIEQANHAIEKGEHQLLYLQHQLDLDEL
10      20      30      40      50      60

CV127_ RF
::
gi|111 NENKSKELQEKIIRELDVVCAMIEGAQGALERELKRTDLNILERFNYEEAQTLSKILLKD
70      80      90      100      110      120

gi|111 LKETEQQVKDIQTQ
130

>>gi|7435005 gi|7435005|pir||A59055 phospholipase A2 (EC (134 aa)
initn: 15 initl: 15 opt: 15 Z-score: 84.6 bits: 15.3 E(): 9.1
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:23-24)

CV127_ RF
::
gi|743 IIYPGTLWCGHGNVSSGPNELGRFKHTDACCRTTHDMCPDVMSAGESKHGLTNTASHTRLS
10      20      30      40      50      60

gi|743 CDCDDTFYDCLKNSGEKISSYFVGKMYFNILDTKCYKLEHPVTGCGERTEGRCLRYTVDK
70      80      90      100      110      120

gi|743 SKPKVYQWFDLRKY
130

>>gi|111120450 gi|111120450|gb|ABH06359.1|Blo t 5 aller (134 aa)
initn: 15 initl: 15 opt: 15 Z-score: 84.6 bits: 15.3 E(): 9.1
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:104-105)

gi|111 MKFAIVLIACFAASVLAQEHKPEKDDFRNEFDHLLIEQANHAIEKGEHQLLYLQHQLDLDEL

```



	10	20	30	40	50	60
--	----	----	----	----	----	----

CV127\_ RF  
::

gi|111 NENKSKELQEKIIRELDVVCAMIEGAQGALERELKRTDLNILERFNYYEAAQTLSKILLKD  
70 80 90 100 110 120

gi|111 LKETEQQVKDIQTQ  
130

>>gi|4204917 gi|4204917|gb|AAD10850.1| major IgE-binding (134 aa)  
initn: 15 initl: 15 opt: 15 Z-score: 84.6 bits: 15.3 E(): 9.1  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
(1-2:104-105)

gi|420 MKFAIVLIACFAASVLAQEHKPKKDDFRNEFDHLLIEQANHAIEKGEHQLLYLQHQDEL  
10 20 30 40 50 60

CV127\_ RF  
::

gi|420 NENKSKELQEKIIRELDVVCAMIEGAQGALERELKRTDLNILERFNYYEAAQTLSKILLKD  
70 80 90 100 110 120

gi|420 LKETEQQVKDIQTQ  
130

>>gi|19039 gi|19039|emb|CAA42832.1| LTP 1 [Hordeum vulga (134 aa)  
initn: 15 initl: 15 opt: 15 Z-score: 84.6 bits: 15.3 E(): 9.1  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
(1-2:115-116)

gi|190 MARAQVLLMAAALVLMMLTAAPRAAVALNCGQVDSKMKPCLTYVQGGPGPSGECCNGVRDL  
10 20 30 40 50 60

CV127\_ RF  
::

gi|190 HNQAQSSGDRQTVCNCLKGIARGIHNLNLNNAASIPSKCNVNVPTISPDIIDCSRFTERR  
70 80 90 100 110 120

gi|190 SVKLVLSSSIHVEL  
130

>>gi|24638082 gi|24638082|sp|Q9BMK4.1|PA2\_APICC RecName: (134 aa)  
initn: 15 initl: 15 opt: 15 Z-score: 84.6 bits: 15.3 E(): 9.1  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
(1-2:23-24)

CV127\_ RF  
::

gi|246 IIYPGTLWCCHGNVSSGPNELGRFKHTDACCRTHDMCPDVMSAGESKHGLTNTASHTRLS  
10 20 30 40 50 60

gi|246 CDCDDTFYDCLKNSGDKISSYFVGKMYFNLIDTKCYKLEHPVTGCGERTEGRCRLRYTVDK  
70 80 90 100 110 120

gi|246 SKPKAYQWFDLRKY  
130

>>gi|1362129 gi|1362129|pir||E53806 major allergen OLE17 (136 aa)  
initn: 15 initl: 15 opt: 15 Z-score: 84.5 bits: 15.3 E(): 9.3  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
(1-2:16-17)

```
CV127_          RF
               ::
gi|136  QFHIQGQVYCDTCHARFITESEFIPGASVRLQCREKENGDITFTEIGYTRAEGLYSMLV
          10          20          30          40          50          60

gi|136  ERDHKNEFCEITLISSGRKDCDEIPIEGWAKPSLKFILNTVNGTTRTVNPLGFFKKEALP
          70          80          90         100         110         120

gi|136  KCAQVYNKLGMYPPNM
          130
```

>>gi|1362137 gi|1362137|pir||H53806 major allergen OLE6 (136 aa)  
initn: 15 initl: 15 opt: 15 Z-score: 84.5 bits: 15.3 E(): 9.3  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
(1-2:16-17)

```
CV127_          RF
               ::
gi|136  QFHIQGQVYCDTCRARFITESEFIPGAGVRLECKDGGKGSITFTEVGYTRAEGLYSMLI
          10          20          30          40          50          60

gi|136  ERDHKNEFCEITLASSSRKDCDEIPVEGWVKPSLKFKLNTVNGTTRTINPLGFFKKEVLP
          70          80          90         100         110         120

gi|136  KCPQVFNKLGMYPNNM
          130
```

>>gi|1362132 gi|1362132|pir||A38968 major allergen OLE20 (137 aa)  
initn: 15 initl: 15 opt: 15 Z-score: 84.5 bits: 15.3 E(): 9.3  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
(1-2:16-17)

```
CV127_          RF
               ::
gi|136  QFHIQGQVYCDTCRARFITESEFIPGAGVRLECRDGEKGSITFTEVGYTRAEGLYSMLI
          10          20          30          40          50          60

gi|136  ERDHKNEFCEITLASSSRKDCDEIPVEGWVKPSLKFKLNTVNGTTRTINPLGFFKKEVLP
          70          80          90         100         110         120

gi|136  KCAPQVFNKLGMYPNNM
          130
```

>>gi|1362128 gi|1362128|pir||I53806 major allergen OLE16 (137 aa)  
initn: 15 initl: 15 opt: 15 Z-score: 84.5 bits: 15.3 E(): 9.3  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
(1-2:17-18)

```
CV127_          RF
               ::
gi|136  SQFHIQGQVYCDTCRARFITESEFIPGAGVRLECRDGEKGSITFTEVGYTRAEGLYSML
          10          20          30          40          50          60
```

gi|136 IERDHNKNEFCEITLASSSRKDCDEIPVEGWVKPSLKFKLNTVNGTTRTINPLGFFKKEVL  
70 80 90 100 110 120

gi|136 PKCPQVFNKLGMYPNM  
130

>>gi|24473800 gi|24473800|gb|AAL91665.1| 2s albumin [Ana (138 aa)  
initn: 15 initl: 15 opt: 15 Z-score: 84.4 bits: 15.3 E(): 9.4  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
(1-2:48-49)

CV127\_ RF  
:  
gi|244 MAKFLLLLSAFVLLLVANASIYRAIVEVEEDSGREQSCQRQFEEQQRFRNCQRYVKQEV  
10 20 30 40 50 60

gi|244 QRGGRYNQRQESLRECCQELQEVDRRCRCQNLEQMVRQLQQEQIKGEEVRELYETASEL  
70 80 90 100 110 120

gi|244 PRICISIPSQGCQFQSSY  
130

>>gi|132270 gi|132270|sp|P15252.2|REF\_HEVBR RecName: Ful (138 aa)  
initn: 15 initl: 15 opt: 15 Z-score: 84.4 bits: 15.3 E(): 9.4  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
(1-2:67-68)

gi|132 MAEDEDNQGGQEGGLKYLGFVQDAATYAVTTFSNVYLFADKSGPLQPGVDIIEGPKVKNV  
10 20 30 40 50 60

CV127\_ RF  
:  
gi|132 AVPLYNRFYSYIPNGALKFVDSTVVASVTIIDRSLPPIVKDASIQVVSIRAAPAEARSLSA  
70 80 90 100 110 120

gi|132 SSLPGQTKILAKVfyGEN  
130

>>gi|145313982 gi|145313982|gb|ABP58632.1| major pollen (140 aa)  
initn: 15 initl: 15 opt: 15 Z-score: 84.3 bits: 15.3 E(): 9.6  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
(1-2:20-21)

CV127\_ RF  
:  
gi|145 PPVSQFHIQGGIYCDTCRSRFITELSEFIPGASVRLQCREKENGDTTFTEIGYTRAEGLY  
10 20 30 40 50 60

gi|145 SMLVERDHNKNEFCEITLISSGRKDCDEIPIEGWAKPSLKFILNTVNGTTRTVNPLGFFKK  
70 80 90 100 110 120

gi|145 EALPKCAQVYNKLGMYPNM  
130 140

>>gi|145313992 gi|145313992|gb|ABP58637.1| major pollen (140 aa)  
initn: 15 initl: 15 opt: 15 Z-score: 84.3 bits: 15.3 E(): 9.6  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
(1-2:99-100)

```
gi|145 PPVSQFHVQGQIYCDTCRAGFIAELSEFILGAGVRLQCRDSKKGNIITFTEVGYTRAEGLY
      10      20      30      40      50      60
```

```
CV127_ RF
      ::
gi|145 SMLIERDHKNEFCEVNSLSSSRKDCDEIPIEGWVRPSLRFLNLTVNGTCTIKPLGFFKN
      70      80      90      100      110      120
```

```
gi|145 KALPKCPQVFNKLGMYPDDL
      130      140
```

```
>>gi|145313988 gi|145313988|gb|ABP58635.1| major pollen (140 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 84.3 bits: 15.3 E(): 9.6
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:20-21)
```

```
CV127_ RF
      ::
gi|145 PPVSQFHIQGVYCDTCRARFITEELSEFIPGASVRLQCRDGENGSITFTEVGSTRAEGLY
      10      20      30      40      50      60

gi|145 SMLIERDHKDEFCEITLISSSRKDCDEIPVEGWVKPSLKFMLNTVNGTTRTVNPLGFFKK
      70      80      90      100      110      120
```

```
gi|145 EALPKCPQVFNKLGMYPDDM
      130      140
```

```
>>gi|112745 gi|112745|sp|P23110.1|2SS8_HELAN RecName: Fu (141 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 84.3 bits: 15.3 E(): 9.6
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:3-4)
```

```
CV127_ RF
      ::
gi|112 MARFSIVFAAAGVLLLVAMAPVSEASTTTIITTIIEENPYGRGRTESGCYQQMEEAEMLN
      10      20      30      40      50      60

gi|112 HCGMYLMKNLGERSQVSPRMREEDHKQLCCMQLKNLDEKCMCPAIMMMLNEPMWIRMRDQ
      70      80      90      100      110      120
```

```
gi|112 VMSMAHNLPIECNLMSQPCQM
      130      140
```

```
>>gi|3256212 gi|3256212|emb|CAA54819.1| major allergen [ (145 aa)
  initn: 15 initl: 15 opt: 15 Z-score: 84.0 bits: 15.3 E(): 9.9
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap
(1-2:25-26)
```

```
CV127_ RF
      ::
gi|325 EDVPQPPVSQFYIQGVYCDTCRARFITEELSEFIPGAGVRLQCKDGENGKITFTEVGYTR
      10      20      30      40      50      60

gi|325 AEGLYSMLIERDHKNEFCEITLLSSSRKDCDEIPIEGWVKPSLKFMLNTVNGTTRTINPL
      70      80      90      100      110      120
```

```
gi|325 GFFKKEALPKCPQVFNKLGMYPNM
      130      140
```

>>gi|1362131 gi|1362131|pir||C53806 major allergen OLE1c (145 aa)  
 initn: 15 initl: 15 opt: 15 Z-score: 84.0 bits: 15.3 E(): 9.9  
 Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
 (1-2:25-26)

```

CV127_                                RF
                                     ::
gi|136 EDVPQPPISQFYVQGQVYCDTCRTRFITEFSEFIPGAGVRLQCKDGENGKITFTEVGYTR
      10          20          30          40          50          60

gi|136 AEGLYSMLIERDHKNEFCEITLLSSSRKDCDEIPTEGWVKPSVKFILNTVNGTTRTINPL
      70          80          90         100         110         120

gi|136 GFFKKEALPKCPQVFNKLGMYPNNM
      130         140
  
```

>>gi|631912 gi|631912|pir||S43243 allergen-like protein (145 aa)  
 initn: 15 initl: 15 opt: 15 Z-score: 84.0 bits: 15.3 E(): 9.9  
 Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
 (1-2:25-26)

```

CV127_                                RF
                                     ::
gi|631 EDVPQPPVPQFHIQGQVYCDTCRARFITEFSEFIPGAGIRLQCKDGEHGKITFTEIGYTR
      10          20          30          40          50          60

gi|631 AEGLYSMLVEGDHKNEFCEITLLSSSRKDCDEIPIEGWVKPSLKFILNTVNGTTRTINPL
      70          80          90         100         110         120

gi|631 GFFKKEVLPKCPQVFNKLGMYPNNM
      130         140
  
```

>>gi|631911 gi|631911|pir||S43242 allergen-like protein (145 aa)  
 initn: 15 initl: 15 opt: 15 Z-score: 84.0 bits: 15.3 E(): 9.9  
 Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
 (1-2:25-26)

```

CV127_                                RF
                                     ::
gi|631 EDVPQPPIPQFHIQGQVYCDTCRARFITEFSEFIPGASIRLQCKDRENGKITFTEIGYTR
      10          20          30          40          50          60

gi|631 AEGLYSMLVEGDHKNEFCEITLISSGREDCDEIPVEGWAKPSLKFILNTVNGTTRTINPI
      70          80          90         100         110         120

gi|631 GFFKKEALPKCTQVYNKLGMYPNNM
      130         140
  
```

>>gi|631913 gi|631913|pir||S43244 allergen-like protein (145 aa)  
 initn: 15 initl: 15 opt: 15 Z-score: 84.0 bits: 15.3 E(): 9.9  
 Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
 (1-2:25-26)

```

CV127_                                RF
                                     ::
gi|631 EDVPQPPVPQFHIQGQVYCDTCRARFITEFSEFIPGASIRLQCKDGENGKITFTEIGYTR
      10          20          30          40          50          60
  
```

gi|631 AEGLYSMLVEGDHKNEFCEITLISSGRKDCDEIPVEGWVKPSLKFKLNTVNGTTRTINPI  
70 80 90 100 110 120

gi|631 GFFKKEALPKCTQVYNKLGMYPPNM  
130 140

>>gi|1362136 gi|1362136|pir||B53806 major allergen OLE5c (145 aa)  
initn: 15 init1: 15 opt: 15 Z-score: 84.0 bits: 15.3 E(): 9.9  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
(1-2:25-26)

CV127\_ RF  
::  
gi|136 EDVPQPPVSQFHIQGQVYCDTCRARFITESEFIPGASVRLQCKDGENGSITTFEVGYTR  
10 20 30 40 50 60

gi|136 AEGLYSMLIERDHKDEFCEITLISSSRKDCDEIPVEGWVKPSLKFMLNTVNGTTRTINPL  
70 80 90 100 110 120

gi|136 GFFKKEALPKCPQVFNKLGMYPNNM  
130 140

>>gi|14423737 gi|14423737|sp|O82015.2|LIGV1\_LIGVU RecNam (145 aa)  
initn: 15 init1: 15 opt: 15 Z-score: 84.0 bits: 15.3 E(): 9.9  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
(1-2:25-26)

CV127\_ RF  
::  
gi|144 EDVPQPPVSQFYIQGQVYCDTCRARFITESEFIPGAGVRLQCKDGENGKVTTFEVGYTK  
10 20 30 40 50 60

gi|144 AEGLYNMLIERDHKNEFCEITLISSSRKDCDEIPTEGWVKPSLKFVLNTVNGTTRTINPL  
70 80 90 100 110 120

gi|144 GFLKKEVLPKCPQVFNKLGMYPNNM  
130 140

>>gi|34978692 gi|34978692|gb|AAQ83588.1| allergen Fra e (146 aa)  
initn: 15 init1: 15 opt: 15 Z-score: 84.0 bits: 15.3 E(): 10  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
(1-2:26-27)

CV127\_ RF  
::  
gi|349 MEDVPQPPVSQFHIQGQVYCDTCRARFITKLSEFITGASVRLQCRDKENGDTVTFTEIGYT  
10 20 30 40 50 60

gi|349 RGEGLYSMFVERDHKNEFCEITLLSSGRKDCNEIPIEGWVKPSLKFILNTVNGTTRTINP  
70 80 90 100 110 120

gi|349 LGFFKKEALPQCAQVYNKLGMYPPNM  
130 140

>>gi|2465127 gi|2465127|emb|CAA73038.1| Ole e 1.0102 pro (146 aa)  
initn: 15 init1: 15 opt: 15 Z-score: 84.0 bits: 15.3 E(): 10  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
(1-2:122-123)

gi|246 MEDVPQPPVSQFHIQGQVYCDTCRAGFITELSEFIPGASVRLQCKDKENGDTVTFTEIGYT  
10 20 30 40 50 60

gi|246 RAEGLYSMLVERDHKNEFCEITLISSGRKDCNEIPTEGWAKPSLKFILNTVNGTTRTVNP  
70 80 90 100 110 120

CV127\_ RF  
::

gi|246 LRFYKKEALPKCAQVYNKLGMYPPNM  
130 140

>>gi|2465131 gi|2465131|emb|CAA73036.1| Ole e 1 protein (146 aa)  
initn: 15 initl: 15 opt: 15 Z-score: 84.0 bits: 15.3 E(): 10  
Smith-Waterman score: 15; 100.000% identity (100.000% similar) in 2 aa overlap  
(1-2:26-27)

CV127\_ RF  
::  
gi|246 MEDVPQPPVSQFHIQGQVYCDTCRSRFFITELSEFIPGASVRLQCREKENGDTTFTEIGYT  
10 20 30 40 50 60

gi|246 RAEGLYSMLVERDHKNEFCEITLISSGRKDCDEIPIEGWAKPSLKFILNTVNGTTRTVNP  
70 80 90 100 110 120

gi|246 LGFFKKEALPKCAQVYNKLGMYPPNM  
130 140

2 residues in 1 query sequences  
307888 residues in 1386 library sequences  
Scomplib [34.26]  
start: Wed Mar 11 06:45:27 2009 done: Wed Mar 11 06:45:27 2009  
Total Scan time: 0.060 Total Display time: 0.010

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

## CV127\_rc\_855\_1016\_frame2

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

Please cite:

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

CV127\_rc\_855\_1016\_frame2, 54 aa

vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences

Expectation\_n fit:  $\rho(\ln(x)) = 4.4681 \pm 0.00303$ ;  $\mu = 0.7675 \pm 0.157$   
mean\_var=29.0283 $\pm$  7.624, 0's: 16 Z-trim: 17 B-trim: 52 in 1/42  
Lambda= 0.238047

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

The best scores are: opt bits E(1386) %\_id %\_sim alen  
gi|4538529 gi|4538529|emb|CAB39376. ( 81) 52 23.8 0.42 0.359 0.590 39





CV127 S

gi | 777 FIEDELKLF LQNFCAGARALSDAETKAF LKAGSDGDGKIGVDEFAAMVKH  
60 70 80 90 100

```
>>gi|18281421 gi|18281421|sp|Q91483.3|PRVB2_SALSA RecNam (108 aa)
  initn: 33 initl: 33 opt: 48 Z-score: 98.8 bits: 22.4 E(): 1.5
Smith-Waterman score: 48; 33.333% identity (70.000% similar) in 30 aa overlap
(20-49:23-52)
```

CV127\_            KIQTYRSLLDGYFVLVCLPSLDWVALLALLGLPIRSSDRVLSSLMGMGQLPQTS  
                                ::   :::  ::  :::  :  ...  :  
gi|182 MSFAGLNDADVAALAATAADSFNHHKAFKAVGLASKSSDDVKKAFYVIDQDKSGFIEE  
                      10            20            30            40            50            60

gi|182 DELKLFLQNFSASARALTD AETKAFLADGDKDGDGMIGVDEFAAMIKG  
70 80 90 100

```
>>gi|47605720 gi|47605720|sp|Q8WNR9.1|CYTA_FELCA RecName (98 aa)
      initn: 37 initl: 37 opt: 47 Z-score: 97.8 bits: 22.0 E(): 1.7
Smith-Waterman score: 47; 22.581% identity (70.968% similar) in 31 aa overlap
(12-42:53-83)
```

CV127\_ 10  
KIQTYRSLLDGYFVLVCLP  
: . . : .  
gi|476 MIPGGLSEAKPATPEIQEIANEVKPKQLEEKTNETYQKFEAIEYKTQVVAGINYYIKVQVD  
10 20 30 40 50 60

CV127\_ SLDWVALLALLGLPIRSSDRVLSSLMGMGQLPQTS  
 . . . . : : : : . : .  
 gi|476 DNRYYIHKVFVKGLPVQDSSLTLTGYQTGKSEDELTF

```
>>gi|1478293 gi|1478293|gb|AAB35897.1| 31 kda major alle (26 aa)
initn: 41 initl: 41 opt: 41 Z-score: 97.7 bits: 20.1 E(): 1.7
Smith-Waterman score: 41; 66.667% identity (77.778% similar) in 9 aa overlap
(44-52:18-26)
```

CV127\_ KIQTYRSLLDGYFVLVCLPSLDWVALLALLGLPIRSSDRVLSSLMGMGQLPQTS  
gi|147 AKITFTNNXPNTVWPGILTFGQKPKQ

```
>>gi|17977825 gi|17977825|emb|CAC83658.1| parvalbumin [C (109 aa)
  initn: 25 initl: 25 opt: 46 Z-score: 95.0 bits: 21.7 E(): 2.4
Smith-Waterman score: 46; 23.333% identity (66.667% similar) in 30 aa overlap
(20-49:24-53)
```

CV127\_            10                 20                 30                 40                 50  
KIQTYRSLLDGYFVLVCLPSLDWVALLALLGLPIRSSDRVLSSLMGMGQLPQTS  
                        : . . : . . : . . : . . :  
gi|179 MAFAGILNDADITAALQGCAADSFDYKSFFAKVGLSAKTPDDIKKAFAVIDQDKSGFIE  
                        10                 20                 30                 40                 50                 60

gi|179 EDELKLFLQNFSAGARALTAETKAFKAGDSDGDGKIGVDEFAALVKA  
70 80 90 100

```
>>q1|29539109 q1|29539109|emb|CAD87730.1| allergen Len c (418 aa)
```

initn: 35 initl: 35 opt: 47 Z-score: 85.8 bits: 21.9 E(): 7.9  
Smith-Waterman score: 47; 31.373% identity (56.863% similar) in 51 aa overlap  
(1-46:323-372)

gi|295 SRSDQENPFIFKSNRFQTIYENENGHIRLLQRFDKRSKIFENLQNYRLLEYKSKPHTIFL  
10 20 30 40 50 60

gi|295 PQFTDADFILVVLGKAILTVLNSNDRNSFNLERGDTIKLPAGTIAYLANRDDNEDLRVL  
70 80 90 100 110 120

gi|295 DLAIIPVNRPGQLQSFLSGTQNPFSLSGFSKNILEAAFNTYEEIEKVLLEEQQEKSQH  
130 140 150 160 170 180

gi|295 RRLRDKRQEITNEDVIVKVSREQIEELSKNAKSSSKSVSSESEPFNLRSRNPIYSNKF  
190 200 210 220 230 240

gi|295 GKFFEITPEKNPQLQDLDFVNSVEIKEGSLLLPNYNSRAIVIVTVNEGKGFELVQQRN  
250 260 270 280 290 300

CV127\_ 10 20 30  
KIQTYRSLLDGYFVLVCLPSLDWVAL----LALLGLP  
..: :. :. :. :. :. :. :.  
gi|295 ENQQEQREENDEEEGQEEETTKQVQRYRRLSPGDVLV-IPAGHPVAINASSDLNLIGFG  
310 320 330 340 350

CV127\_ 40 50  
IRSSDRVLSSLMGMLPQTS  
: ... : : :  
gi|295 INAKNNQRNFLAGEDNVISQIQRPVKELAFPGSSREVDRLLTNQKQSHFANAQPLQIE  
360 370 380 390 400 410

>gi|1885350 gi|1885350|emb|CAA72273.1| serpin [Triticum (399 aa)  
initn: 46 initl: 46 opt: 46 Z-score: 84.3 bits: 21.6 E(): 9.6  
Smith-Waterman score: 46; 40.741% identity (59.259% similar) in 27 aa overlap  
(26-52:299-325)

gi|188 MATTLATDVRLSIAHQTRFALRLASTISSNPKSAASNAAFSPVSLHSALSLAAGAGSAT  
10 20 30 40 50 60

gi|188 RDQLVATLGTGEVEGGHALAEQVVQFVLADASSAGGPRVAFANGVFVDASLLKPSFQEL  
70 80 90 100 110 120

gi|188 AVCKYKAETQSVDFQTKAAEVTTQVNSWVEKVTSGRIKNILPSGSVDNTTKLVLANALYF  
130 140 150 160 170 180

gi|188 KGAWTDQFDSYGTKNDFYLLDGSSVQTPFMSSMDDQYISSSDGLKVLKLPYKQGGDNR  
190 200 210 220 230 240

CV127\_ 10 20  
KIQTYRSLLDGYFVLVCLPSLDWVALL  
:  
gi|188 QFSMYILLPEAPGGLSSLAEKLSAEPDFLERHIPRQVRVAIRQFKLPKFKISFGMEASDLL  
250 260 270 280 290 300

CV127\_ 30 40 50  
ALLGLPIRSSDRVLSSLMGMLPQTS  
::: . :. : : . :.  
gi|188 KCLGLQLPFSDEADFSEMVDSPMPQGLRVSSVFHQAFVEVNEQGTEAAASTAIKMVPQQA  
310 320 330 340 350 360

gi|188 RPPSVMDFIADHPFLFLLEDISGVVLFMGHVVNPLLS  
370 380 390

54 residues in 1 query sequences  
307888 residues in 1386 library sequences  
Scomplib [34.26]  
start: Wed Mar 11 06:13:14 2009 done: Wed Mar 11 06:13:15 2009  
Total Scan time: 0.080 Total Display time: 0.000

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

## CV127\_rc\_962\_1159\_frame1

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

Please cite:

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

CV127\_rc\_962\_1159\_frame1, 66 aa  
vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences  
Expectation\_n fit:  $\rho(\ln(x)) = 4.5225 \pm 0.00391$ ;  $\mu = 2.9286 \pm 0.208$   
mean\_var=36.1718  $\pm$  8.941, 0's: 12 Z-trim: 12 B-trim: 216 in 1/41  
Lambda= 0.213250

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

The best scores are:

		opt	bits	E(1386)	%_id	%_sim	alen
gi 113475	gi 113475 sp P27759.1 MPA ( 396)	57	23.8	2.4	0.467	0.867	15
gi 66840996	gi 66840996 emb CAI6439 ( 137)	51	21.9	3.2	0.333	0.767	30
gi 2832430	gi 2832430 emb CAA05978. ( 187)	49	21.3	6.5	0.343	0.629	35
gi 3183706	gi 3183706 emb CAA75312. ( 131)	47	20.7	7.1	0.500	0.750	16
gi 158342650	gi 158342650 gb ABW349 ( 204)	49	21.3	7.1	0.343	0.629	35
gi 123062	gi 123062 sp P02877.2 HEV ( 204)	48	21.0	8.8	0.314	0.629	35

>>>CV127\_rc\_962\_1159\_frame1, 66 aa vs  
/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

>>gi|113475 gi|113475|sp|P27759.1|MPA11\_AMBAR RecName: F (396 aa)  
initn: 57 initl: 57 opt: 57 Z-score: 94.9 bits: 23.8 E(): 2.4  
Smith-Waterman score: 57; 46.667% identity (86.667% similar) in 15 aa overlap  
(12-26:299-313)

gi|113 MGIKHCCYILYFTLALVTLLQPVRSAEDLQEILPVNETRRLTTSGAYNIIDGCWRGKADW  
10 20 30 40 50 60

gi|113 AENRKALADCAQGFQKGTGKDGDIYTVTSELDDDDVANPKEGTLRFQAAQNRPLWIIIF  
70 80 90 100 110 120

gi|113 RDMVIRLDKEMVNSDKTIDGRGAKVEIINAGFTLNGVKNVIIHNINMHDVKVNPGLLIK  
130 140 150 160 170 180

gi|113 SNDGPAAPRAGSDGDAISISGSSQIWIWDHCSLSKSVGLVDAKLGTTRLTVSNLSFTQHQ  
190 200 210 220 230 240

CV127\_ 10  
LGSCTVFFDGHGT

```

: .
gi|113 FVLLFGAGDENIEDRGLATVAFNTFTDNDQRMPCRHGFFQVVMNNYDKWGSYAIGGS
      250      260      270      280      290      300

      20      30      40      50      60
CV127_ ASPNVLKLNGNFCFCTFNSFLQTPKSSLSSSARFSRTLFIIPCKASLTSPHTDT
      : : : : : . . : :
gi|113 ASPTILSQGNRFCAPDERSKKNVLGRHGEAAAESMKWNWRTNKDVLNGAIFVASGVDPV
      310      320      330      340      350      360

gi|113 LTPEQSAGMIPAEPGESALSLTSSAGVLSQPGAPC
      370      380      390

>>gi|66840996 gi|66840996|emb|CAI64397.1| putative leuci (137 aa)
  initn: 24 initl: 24 opt: 51 Z-score: 92.9 bits: 21.9 E(): 3.2
Smith-Waterman score: 61; 33.333% identity (76.667% similar) in 30 aa overlap
(22-49:106-135)

gi|668 KPVAANQDVDALSALRRGLQDPNGELKNWDANLVDACTWSHITCDRDNNRVTRIDLNMN
      10      20      30      40      50      60

      10      20      30
CV127_ LGSCTVFFDGHGTASPNVLKLNGNFC--FCTFNSFL
      : : : : . : : :
gi|668 LSGPLAPELGKLDRLQYLEIDHNRLTGPIPRELAGLSNLKHADFSSNNLCGPIPTTGAFAQ
      70      80      90      100      110      120

      40      50      60
CV127_ QTPKSSLSSSARFSRTLFIIPCKASLTSPHTDT
      . : : : : . : :
gi|668 RIPRSSFANNPRLGRKC
      130

>>gi|2832430 gi|2832430|emb|CAA05978.1| prohevein [Hevea (187 aa)
  initn: 30 initl: 30 opt: 49 Z-score: 87.3 bits: 21.3 E(): 6.5
Smith-Waterman score: 49; 34.286% identity (62.857% similar) in 35 aa overlap
(9-41:152-185)

gi|283 EQCGRQAGGKLCPPNNLCCSQWGWCGSTDEYCSDPHNCQSNCKDSGEGVGGGSASNVLATY
      10      20      30      40      50      60

gi|283 HLYNSQDHGWDLNAASAYCSTWDANKPYSWRSKYGTAFCEGPGVGAHGQPSGKCLSVTNT
      70      80      90      100      110      120

      10      20      30
CV127_ LGSCTVFFDGHGTASPNVLKLNGNFCFC--TFNSFLQ
      : : : . : : : : : : : :
gi|283 GTGAKATVRIVDQCSNGGLDLVDNVFRQLDLDGKGYERGH-LTVNYQFVDCGDSFNPLFS
      130      140      150      160      170

      40      50      60
CV127_ TPKSSLSSSARFSRTLFIIPCKASLTSPHTDT
      . : : :
gi|283 VMKSSVIN
      180

>>gi|3183706 gi|3183706|emb|CAA75312.1| profilin [Hevea (131 aa)
  initn: 47 initl: 47 opt: 47 Z-score: 86.6 bits: 20.7 E(): 7.1
Smith-Waterman score: 47; 50.000% identity (75.000% similar) in 16 aa overlap
(8-23:54-69)

```

```

CV127_                                     LGSCTVFFDGHGTA
                                         ::  ::
gi|318 MSWQTYVDERLMCEIEGNHLTAAAIIGQDGSVWAQSSNFPQFKSEEITAIMSDFDEPGTL
      10      20      30      40      50      60

      20      30      40      50      60
CV127_ SPNVLKLNGNFCFCFTFNSFLQTPKSSLSSSARFSRTLFIIPCKASLTSPHTDT
      ..  ::::
gi|318 APTGLHLGGTKYMVIIQGEAGAVIRGKKGPGGVTVRKTNQALIIGIYDEPMTPGQCNMIVE
      70      80      90      100      110      120

gi|318 RLG DYLLLEQGM
      130

>>gi|158342650 gi|158342650|gb|ABW34946.1| hevein [Hevea (204 aa)
  initn: 30 initl: 30 opt: 49 Z-score: 86.6 bits: 21.3 E(): 7.1
Smith-Waterman score: 49; 34.286% identity (62.857% similar) in 35 aa overlap
(9-41:169-202)

gi|158 MNIFIVLLCLTGVAIAEQCGRQAGGKLCPPNNLCCSQWGWCGSTDEYCSPDHNCQSNCKD
      10      20      30      40      50      60

gi|158 SGEGVGGGSASNVLATYHLYNSQDHGWDLNAASAYCSTWDANKPYSWR SKYGTAF CGPV
      70      80      90      100      110      120

CV127_                                     10      20
                                         LGSCTVFFDGHGTASPNVLK
                                         ::::  . :
gi|158 GAHQPPSCGKCLSVTNTGTGAKTTVRIVDQCSNGGLDLVDNVFRQLD TDGKGYERGH-LT
      130      140      150      160      170

      30      40      50      60
CV127_ LNGNFCFC--TFNSFLQTPKSSLSSSARFSRTLFIIPCKASLTSPHTDT
      .: .: : .:: .... :::.
gi|158 VNYQFVDCGDSFNPLFSVMKSSVIN
      180      190      200

>>gi|123062 gi|123062|sp|P02877.2|HEVE_HEVBR RecName: Fu (204 aa)
  initn: 30 initl: 30 opt: 48 Z-score: 85.0 bits: 21.0 E(): 8.8
Smith-Waterman score: 48; 31.429% identity (62.857% similar) in 35 aa overlap
(9-41:169-202)

gi|123 MNIFIVLLCLTGVAIAEQCGRQAGGKLCPPNNLCCSQWGWCGSTDEYCSPDHNCQSNCKD
      10      20      30      40      50      60

gi|123 SGEGVGGGSASNVLATYHLYNSQDHGWDLNAASAYCSTWDANKPYSWR SKYGTAF CGPV
      70      80      90      100      110      120

CV127_                                     10      20
                                         LGSCTVFFDGHGTASPNVLK
                                         ::::  ..
gi|123 GAHQSSCGKCLSVTNTGTGAKTTVRIVDQCSNGGLDLVDNVFRQLD TDGKGYERGHIT-
      130      140      150      160      170

      30      40      50      60
CV127_ LNGNFCFC--TFNSFLQTPKSSLSSSARFSRTLFIIPCKASLTSPHTDT
      .: .: : .:: .... :::.
gi|123 VNYQFVDCGDSFNPLFSVMKSSVIN
      180      190      200

```

66 residues in 1 query sequences  
307888 residues in 1386 library sequences  
Scomplib [34.26]  
start: Wed Mar 11 06:10:34 2009 done: Wed Mar 11 06:10:35 2009  
Total Scan time: 0.100 Total Display time: 0.010

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

## CV127\_5319\_5915\_frame2, amino acids 38-117

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

Please cite:

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

CV127\_5319\_5915\_frame2\_38-117, 80 aa  
vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences  
Expectation\_n fit:  $\rho(\ln(x)) = 2.4164 \pm 0.00359$ ;  $\mu = 17.7277 \pm 0.186$   
mean\_var=43.7830 +/- 12.578, 0's: 5 Z-trim: 5 B-trim: 114 in 2/42  
Lambda= 0.193830

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

The best scores are:		opt	bits	E(1386)	%_id	%_sim	alen
gi 60280803	gi 60280803 gb AAX18294 ( 160)	59	22.6	2.8	0.228	0.667	57
gi 51860756	gi 51860756 gb AAU11502 ( 131)	56	21.6	4.6	0.346	0.551	78
gi 156938917	gi 156938917 gb ABU974 ( 131)	56	21.6	4.6	0.346	0.551	78
gi 60280823	gi 60280823 gb AAX18304 ( 159)	56	21.7	5	0.228	0.632	57
gi 60280825	gi 60280825 gb AAX18305 ( 159)	56	21.7	5	0.228	0.632	57
gi 164510828	gi 164510828 emb CAK93 ( 159)	56	21.7	5	0.246	0.614	57
gi 14285797	gi 14285797 sp O61379.1 ( 274)	57	22.4	5.3	0.269	0.615	52
gi 2660868	gi 2660868 gb AAC48288.1 ( 284)	57	22.5	5.4	0.269	0.615	52
gi 125995171	gi 125995171 dbj BAF47 ( 284)	57	22.5	5.4	0.269	0.615	52
gi 125995169	gi 125995169 dbj BAF47 ( 284)	57	22.5	5.4	0.269	0.615	52
gi 14285796	gi 14285796 sp O44119.1 ( 284)	57	22.5	5.4	0.269	0.615	52
gi 156712752	gi 156712752 dbj BAF76 ( 284)	57	22.5	5.4	0.269	0.615	52
gi 125995167	gi 125995167 dbj BAF47 ( 284)	57	22.5	5.4	0.269	0.615	52
gi 19338630	gi 19338630 gb AAL86739 ( 448)	58	23.1	5.5	0.230	0.557	61
gi 60280827	gi 60280827 gb AAX18306 ( 159)	54	21.2	7.4	0.214	0.643	70
gi 1531589	gi 1531589 gb AAB09632.1 ( 631)	57	23.1	7.9	0.588	0.765	17
gi 729764	gi 729764 sp P40918.1 HSP ( 643)	57	23.1	8	0.237	0.593	59
gi 2739154	gi 2739154 gb AAC67308.1 ( 191)	54	21.3	8.1	0.298	0.526	57
gi 29839547	gi 29839547 sp Q8GT41.1 ( 179)	53	21.0	9.5	0.292	0.492	65

>>>CV127\_5319\_5915\_frame2\_38-117, 80 aa vs  
/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

>>gi|60280803 gi|60280803|gb|AAX18294.1| major allergen (160 aa)  
initn: 26 initl: 26 opt: 59 Z-score: 93.8 bits: 22.6 E(): 2.8  
Smith-Waterman score: 59; 22.807% identity (66.667% similar) in 57 aa overlap  
(10-66:75-129)

gi|602 MGVTYETEFTSVIPAPRLFKAFILDGDNLPKIAPQAIKSTIEIGDGGVGTIKKVTFG  
10 20 30 40 50 60  
10 20 30 40 50

CV127\_ HGTVYANYAVEHSDLLLAFGVRFDDRVTKGLEAFASRAKIVHIDIDSAEIGKNKT  
 ..... : .. :: .. .. : .. : .. :  
 gi|602 EGSQYGYVKQRVNGIDKNFTYSYSMIEGDTLSDKLEKITYETKLI-ASPDGGSIIKT-T  
 70 80 90 100 110

```

          60          70          80
CV127_ PHVSVCGDVKLALQGMNKVLENRAE
      :  .  : : . .
gi|602 SHYHAKGDVEIKEEHVKAGKEKASGLFKLLEAYLVANPDAYN
      120      130      140      150      160

```

```
>>gi|51860756 gi|51860756|gb|AAU11502.1| fatty acid-bidi (131 aa)
  initn: 34 initl: 34 opt: 56 Z-score: 90.0 bits: 21.6 E(): 4.6
Smith-Waterman score: 56; 34.615% identity (55.128% similar) in 78 aa overlap
(7-80:7-76)
```

CV127\_ HGTVYANYAVEHSDLLLLAF---GVRFDDRVTGKLEAFASRAKIVHIDIDS AEIGKNKTP  
 . : . : . : . : . : . : . : . : . : . : . : . : . : . : . :  
 gi|518 MVQLNGSYKLEKSDNFDAFLKELGVNF---VTRNLAKSASPTVEIVDGD SYTI---KTS  
 10 20 30 40 50

CV127\_ HVSVCGDVKLALQGMNKVLENRAE  
 . . . . . : : . . : : .  
 gi|518 STLKNSEIKFKL-G-EEFEEDRADGKKVQTSVTKEGDNKLQVQVQKGDKPVTIVREFSEEG  
 60 70 80 90 100 110

gi|518 LTVTATVNGVTSVRFYKRQ  
120 130

```
>>gi|156938917 gi|156938917|gb|ABU97480.1| allergen Tyr (131 aa)
initn: 34 initl: 34 opt: 56 Z-score: 90.0 bits: 21.6 E(): 4.6
Smith-Waterman score: 56; 34.615% identity (55.128% similar) in 78 aa overlap
(7-80:7-76)
```

CV127\_ HGTVYANYAVEHSDLLLLAF---GVRFDDRVTGKLEAFASRAKIVHIDIDS AEIGKNKTP  
gi|156 MVQLNGSYKLEKSDNFDAFLKELGVNF---VTRNLAKSASPTVEIVDGD SYTI---KTS

CV127\_ HVSVCGDVKLALQGMNKVLENRAE  
 . . . . . : : : : :  
 gi|156 STLKNSEIKFKL-G-EEFEEDRADGKKVQTSVTKEGDNKLQVQVQKGDKPVTIVREFSEEG  
 60 70 80 90 100 110

gi|156 LTVTATVNGATSVRFYKRQ  
120 130

```
>>gi|60280823 gi|60280823|gb|AAX18304.1| major allergen (159 aa)
  initn: 24 initl: 24 opt: 56 Z-score: 89.3 bits: 21.7 E(): 5
Smith-Waterman score: 56; 22.807% identity (63.158% similar) in 57 aa overlap
(10-66:75-128)
```

gi|602 MGVLTYETEYASVIPPARYNALVLDADNLIPKIA PQAVKTEILEGDDGGVGTIKKVSFG  
10 20 30 40 50 60

CV127\_ HGT VYANYAVEHSDLLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKT

gi|602 EGSEYSYVKHKVEGIDKDNFVSYNIEGDAISDKIEKISYEIKLVASG--SGSIKKNIS  
70 80 90 100 110

CV127\_ PHVSVC GDVKLALQGMNKVLENRAE  
: . : : .

gi|602 -HYHTKGDVEIKEEHVKAGKERAHGLFKLIENHLVANPDAYN  
120 130 140 150

```
>>gi|60280825 gi|60280825|gb|AAX18305.1| major allergen (159 aa)
initn: 24 initl: 24 opt: 56 Z-score: 89.3 bits: 21.7 E(): 5
Smith-Waterman score: 56; 22.807% identity (63.158% similar) in 57 aa overlap
(10-66:75-128)
```

gi|602 MGVLTYET EYASVIPP ARLYNALVLDADNLIPK IAPQAVK TVEILEG DGGVG TIKKVSFG  
10 20 30 40 50 60

CV127\_ HGT VYANYAVEHSDLLAFGVRFDDRTGKLEAFASRAKIVHIDIDSAEIGKNKT

gi|602 EGSEYSYVKHKVEGIDKDNFVYSYSLIEGDAISDKIEKISYEIKLVASG--SGSIIKNIS  
70 80 90 100 110

CV127\_ PHVSVC GDVKLALQGMNKVLENRAE

gi|602 -HYHTKGDVEIKEEHVKAGKERAHGLFKLIENHLVANPDAYN  
120 130 140 150

```
>>gi|164510828 gi|164510828|emb|CAK93677.1| PR-10 protei (159 aa)
  initn: 24 initl: 24 opt: 56 Z-score: 89.3 bits: 21.7 E(): 5
Smith-Waterman score: 56; 24.561% identity (61.404% similar) in 57 aa overlap
(10-66:75-128)
```

gi|164 MGVLTYEPNTPQSSPLLGCTMPLFLMLTISFRRL LHKQSKLLKFSREMAVLEPSRKLLALV  
10 20 30 40 50 60

CV127\_ HGT VYANYAVEHS D L L A F G V R F D D R V T G K L E A F A S R A K I V H I D I D S A E I G K N K T

gi|164 KGSEYSYVKHKVEGIDKDNFDYSYSLIEGDAISDKIEKISYEIKLVASG--SGSIKN-T  
70 80 90 100 110

CV127\_ PHVSVCGDVKLALQGMNKVLENRAE  
: . : : .

gi|164 SHYHTKGDVEIKEEHVKAGKDKAHGLFKLIENYLVANPDAYN  
120 130 140 150

```
>>gi|14285797 gi|14285797|sp|O61379.1|TPM_PANST RecName: (274 aa)
  initn: 37 initl: 37 opt: 57 Z-score: 88.9 bits: 22.4 E(): 5.3
Smith-Waterman score: 57; 26.923% identity (61.538% similar) in 52 aa overlap
(29-79:73-124)
```

CV127\_10  
HGTVYANYAVEHSDLL

gi|142 MKLEKDNAMDRADTLEQQNKEANIRAEKAEFEVHNLQKRMQQLENDLDQVQESLLKANTQ  
10 20 30 40 50 60

CV127 LAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNK-TPHVSVCGDVKLALOGMKNKVL



```

      :.. :. : .... :.. : : : ... .. : :
gi|142 LEEKDKALSNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAADSESRMRKVL
      70      80      90      100      110      120

```

```

      80
CV127_ ENRAE
      :..
gi|142 ENRSLSDDEERMDALENQLKEARFLAEEADRKYDEVARKLAMVEADLERAEERAETGESKF
      130      140      150      160      170      180

```

```

gi|142 VELEEEELRVVGNLKSLEVSEEKANQREEAYKEQIKTLTNKLKAAEARAFAERSVQKLQ
      190      200      210      220      230      240

```

```

gi|142 KEVDRLEDELVNEKEKYKSITDELDTFSELSGY
      250      260      270

```

```

>>gi|2660868 gi|2660868|gb|AAC48288.1| fast tropomyosin (284 aa)
  initn: 37 initl: 37 opt: 57 Z-score: 88.7 bits: 22.5 E(): 5.4
Smith-Waterman score: 57; 26.923% identity (61.538% similar) in 52 aa overlap
(29-79:83-134)

```

```

CV127_ HGTVYA
gi|266 MDAIKKKMQAMKLEKDNAMDRADTLEQQNKEANIRAESSEEVHNLQKRMQQLENDLDQV
      10      20      30      40      50      60

```

```

      10      20      30      40      50      60
CV127_ NYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNK-TPHVSVCQDVK
      :.. :. : .... :.. : : : ... ..
gi|266 QESLLKANTQLEEKDKALSNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA
      70      80      90      100      110      120

```

```

      70      80
CV127_ LALQGMNKVLENRAE
      . : : : : :
gi|266 DESERMRKVLENRSLSDDEERMDALENQLKEARFLAEEADRKYDEVARKLAMVEADLERAE
      130      140      150      160      170      180

```

```

gi|266 ERAETGESKIVELEEEELRVVGNLKSLEVSEEKANQREEAYKEQIKTLANKLKAAEARA
      190      200      210      220      230      240

```

```

gi|266 FAERSVQKLQKEVDRLEDELVNEKEKYKSITDELDTFSELSGY
      250      260      270      280

```

```

>>gi|125995171 gi|125995171|dbj|BAF47269.1| tropomyosin (284 aa)
  initn: 37 initl: 37 opt: 57 Z-score: 88.7 bits: 22.5 E(): 5.4
Smith-Waterman score: 57; 26.923% identity (61.538% similar) in 52 aa overlap
(29-79:83-134)

```

```

CV127_ HGTVYA
gi|125 MDAIKKKMQAMKLEKDNAMDRAHTLEQQNKEANLRAEKTEEEIRATQKKMQQIENELDHA
      10      20      30      40      50      60

```

```

      10      20      30      40      50      60
CV127_ NYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNK-TPHVSVCQDVK
      :.. :. : .... :.. : : : ... ..
gi|125 QEQLSAANTKLDEKEKALQNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA
      70      80      90      100      110      120

```

```

      70      80
CV127_ LALQGMNKVLENRAE
      . : : : : .
gi|125 DESERMRKVLENRSLSD EERMDALENQLKEARFLAEEADRKYDEVARKLAMVEADLERAE
      130      140      150      160      170      180

gi|125 ERAESGESKIVELEEEELRVVGNNLKSLEVSEEKANQREETYKEQIKTLANKLKAAEAREAE
      190      200      210      220      230      240

gi|125 FAERSVQKLQKEVDRLEDELVNEKEKYKNIADMDQAFSELSGF
      250      260      270      280

>>gi|125995169 gi|125995169|dbj|BAF47268.1| tropomyosin (284 aa)
  initn: 37 initl: 37 opt: 57 Z-score: 88.7 bits: 22.5 E(): 5.4
Smith-Waterman score: 57; 26.923% identity (61.538% similar) in 52 aa overlap
(29-79:83-134)

```

```

CV127_ HGTVYA

gi|125 MDAIKKKMQAMKLEKDNAMDRAHTLEQQNKEANLRAEKTEEEIRATQKKMQQIENELDHA
      10      20      30      40      50      60

      10      20      30      40      50      60
CV127_ NYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNK-TPHVSVC GDVK
      : . . : . . : . . : . . : . . : . . : . . : . . : . . : . .
gi|125 QEQLSAANTKLEKEKALQNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA
      70      80      90      100      110      120

```

```

      70      80
CV127_ LALQGMNKVLENRAE
      . : : : : .
gi|125 DESERMRKVLENRSLSD EERMDALENQLKEARFLAEEADRKYDEVARKLAMVEADLERAE
      130      140      150      160      170      180

gi|125 ERAESGESKIVELEEEELRVVGNNLKSLEVSEEKANQREETYKEQIKTLANKLKAAEAREAE
      190      200      210      220      230      240

gi|125 FAERSVQKLQKEVDRLEDELVNEKEKYKSITDELDQTFSELSGY
      250      260      270      280

>>gi|14285796 gi|14285796|sp|O44119.1|TPM_HOMAM Tropomyo (284 aa)
  initn: 37 initl: 37 opt: 57 Z-score: 88.7 bits: 22.5 E(): 5.4
Smith-Waterman score: 57; 26.923% identity (61.538% similar) in 52 aa overlap
(29-79:83-134)

```

```

CV127_ HGTVYA

gi|142 MDAIKKKMQAMKLEKDNAMDRA DTLEQQNKEANIRA EKTEEEIRITHKKMQQVENELDQV
      10      20      30      40      50      60

      10      20      30      40      50      60
CV127_ NYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNK-TPHVSVC GDVK
      : . . : . . : . . : . . : . . : . . : . . : . . : . . : . .
gi|142 QEQLSLANTKLEKEKALQNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA
      70      80      90      100      110      120

```

```

      70      80
CV127_ LALQGMNKVLENRAE
      . : : : : .
gi|142 DESERMRKVLENRSLSD EERMDALENQLKEARFLAEEADRKYDEVARKLAMVEADLERAE

```

```

130      140      150      160      170      180
gi|142 ERAETGESKIVELEEEELRVVGNLKSLEVSEEKANQREEAYKEQIKTLANKLKAAEAREAE
190      200      210      220      230      240

gi|142 FAERSVQKLQKEVDRLEDELVNEKEKYKSITDELDQTFSELSGY
250      260      270      280

>>gi|156712752 gi|156712752|dbj|BAF76430.1| tropomyosin (284 aa)
  initn: 37 initl: 37 opt: 57 Z-score: 88.7 bits: 22.5 E(): 5.4
Smith-Waterman score: 57; 26.923% identity (61.538% similar) in 52 aa overlap
(29-79:83-134)

```

```

CV127_                                     HGTVYA

gi|156 MDAIKKKMQAMKMEKDSAMDRSDALEAQNKETNAKADKADDEVHNLQKRLQTLENDLDQV
10      20      30      40      50      60

CV127_ NYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNK-TPHVSVCQDVK
      ... .. : .... :... .. : : ... ...
gi|156 SEALLKANTQLVEKDKALQNAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA
70      80      90      100     110     120

70      80
CV127_ LALQGMNKVLENRAE
      . : : : : .
gi|156 DESERMKVLNRSMLDEERMDALESQKLEARLLAEEDRKYDEVARKLAMVEADLERAE
130     140     150     160     170     180

gi|156 ERAEAGENKIVELEEEELRVVGNLKSLEVSEEKANQREEAYKEQIKHLTHKLKTAEAREAE
190     200     210     220     230     240

gi|156 FAERSVQKLQKEVDRLEDELVNEKEKYKGITDELDQTFNELSGY
250     260     270     280

>>gi|125995167 gi|125995167|dbj|BAF47267.1| tropomyosin (284 aa)
  initn: 37 initl: 37 opt: 57 Z-score: 88.7 bits: 22.5 E(): 5.4
Smith-Waterman score: 57; 26.923% identity (61.538% similar) in 52 aa overlap
(29-79:83-134)

```

```

CV127_                                     HGTVYA

gi|125 MDAIKKKMQAMKLEKDNAMDKADTLEQQNKEANLRAEKTTEEIRANQKKSQVLVENELDHA
10      20      30      40      50      60

CV127_ NYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNK-TPHVSVCQDVK
      ... .. : .... :... .. : : ... ...
gi|125 QEQLSAATHKLVEKEKAFANAEGEVAALNRRIQLLEEDLERSEERLNTATTKLAEASQAA
70      80      90      100     110     120

70      80
CV127_ LALQGMNKVLENRAE
      . : : : : .
gi|125 DESERMKVLNRSLSDEERMDALENQLKEARFLAEEDRKYDEVARKLAMVEADLERAE
130     140     150     160     170     180

gi|125 ERAESGESKIVELEEEELRVVGNLKSLEVSEEKANQREETYKEQIKTLANKLKAAEAREAE
190     200     210     220     230     240

```

gi|125 FAERSVQKLQKEVDRLEDELVNEKEKYKNIADEMDQAFSELSGF  
250 260 270 280

```
>>gi|19338630 gi|19338630|gb|AAL86739.1|AF441864_1 48-kD (448 aa)
  initn: 63 initl: 40 opt: 58 Z-score: 88.6 bits: 23.1 E(): 5.5
Smith-Waterman score: 58; 22.951% identity (55.738% similar) in 61 aa overlap
(2-59:277-337)
```

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

CV127\_ HGT VYANYAVEHS DLL-LAFGVRFD  
: : . : : : : : :  
gi|193 ALSQHEEGPPRIWPFGGESSGPINLLHKHPSQSNQFGRLYEAHPDDHKQLQDL DLMVSFA  
250 260 270 280 290 300

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

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

```
>>gi|60280827 gi|60280827|gb|AAX18306.1| major allergen (159 aa)
  initn: 24 initl: 24 opt: 54 Z-score: 86.3 bits: 21.2 E(): 7.4
Smith-Waterman score: 54; 21.429% identity (64.286% similar) in 70 aa overlap
(10-79:75-140)
```

gi|602 MGVLTYETEYASVIPPAPRLYNALVLDADNLIPKIPQAVKTEILEGDDGGVGTIKKVSFG  
10 20 30 40 50 60

CV127\_                    10                 20                 30                 40                 50  
HGTVYANYAVEHSDDLALAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGKNKT  
..... : . : . : . :  
gi|602 EGSEYSYVKHKVEGIDKDNFVYSYSLIEGDAISDKIQIKISYEIKLVASG--SGSIKNIS  
                70                 80                 90                 100                 110

```

          60          70          80
CV127_ PHVSVC GDVKLALQGMNKVLENRAE
      :  .  ::..  . . .  .  .  .
gi | 602 -HYHTKGDVEIKEENV-KAGKERAHGLFKLIENHLVANPDAYN
      120      130      140      150

```

```
>>gi|1531589 gi|1531589|gb|AAB09632.1| allergen [Peripla (631 aa)
initn: 45 initl: 45 opt: 57 Z-score: 85.8 bits: 23.1 E(): 7.9
```

```

              10      20      30      40      50      60
CV127_  HGTVYANYAVEHSDLLLLAFGVRFDRTGKLEAFASRAKIVHIDIDSAEIGKNKTPHVSV
              :  ::  : ::  : ::
gi|153              DIGDHYDIE-ANIGHYKYPHVVK
                      10      20

```

	70	80
CV127_	CGDVKLALQGMNKVLENRAE	
gi   153	NFISYYKKGLLPRGEPFSVYVEKHREQAIKLFELFFAANDYDTFYKTACWARDRVNEGMF	
	30 40 50 60 70 80	
gi   153	MYALTVAAFHREDTKDLVLPPEYVNPYLFVEDDVIQQAYKYWTKESGTDKHVEHVIPVN	
	90 100 110 120 130 140	
gi   153	FTARSQEDLVAYFREDVDLNAFNMYFRIYIPSWFNTTLYGKSFDRRGEQFYITYHQIYAR	
	150 160 170 180 190 200	
gi   153	YFLERLSNSLPDVKPFQYSKPLKTGYNPHLRYNHGEEMPAPRPSNMYPNTFDLFYVSDIKN	
	210 220 230 240 250 260	
gi   153	YERRVEKAIDFGYAFDEHRTPYSLYHDQHGM DYLGQMIETRNSPHQYFYGSVFHFYRLL	
	270 280 290 300 310 320	
gi   153	VGHVDPYPYHKNGLAPSALEHPQTALRDPAFYQLWKRIDHIVQKYKNRLPRYTYDELSFPG	
	330 340 350 360 370 380	
gi   153	VKIENVDVGKLYTYFEHFEHSLGNAMYLGKLEDYMKASIRARHYRLNHKPFITYNIEVSSD	
	390 400 410 420 430 440	
gi   153	KAQDVYVRIFLGPKYDLSLGECELDERRHYFVEMDRFVHKVEAGKTVIERKSHDSSIISD	
	450 460 470 480 490 500	
gi   153	SHDSYRNLFKKVSDALQEKDQYYIDKSHKYCGYPENLLLPKGKKGGQTFTFYVIVTPYVK	
	510 520 530 540 550 560	
gi   153	QDEHDFEPYHYKAFSYCGVGHGRKYPDDKPLGFPPDRKIHDYDFYTPNMYFKDVVIFHKK	
	570 580 590 600 610 620	
gi   153	YDEVHDTVH	
	630	

```
>>gi|729764 gi|729764|sp|P40918.1|HSP70_CLAHE RecName: F (643 aa)
  initn: 31 initl: 31 opt: 57 Z-score: 85.7 bits: 23.1 E(): 8
Smith-Waterman score: 57; 23.729% identity (59.322% similar) in 59 aa overlap
(2-60:10-65)
```

```

              10      20      30      40      50
CV127_      HGTVYANYAVEHSDLLLAFGVRFDDRVTGKLEAFASRAKIVHIDIDSAEIGK
              :::. . . .: . . . :. . :. . . . :. . . :. . .
gi|729      MAPAIGIDLGTTYSCVGIYRDDRIEIIANDQGNRTTPSFVAFTDTERLIG--DSAKNQV
              10      20      30      40      50

              60      70      80
CV127_      NKTPHVSVCGDVKLALQGMNKNVLENRAE
              ::. :.
gi|729      AINPHNTVFDAKRLIGRKFDQAEVQADMKHFPFKVIEKAGKPVTVQVEFKGETKDFTPEEI
              60      70      80      90      100     110

```

gi|729 SSMILTKMRETAESYLGGTVNNAVITVPAYFNDSQRQATK DAGLIAGLNVLRIINEPTAA  
120 130 140 150 160 170gi | 729 AIAYGLDKKQEGEKNVLIFDLGGGTFDVSFLTIEEGIFEVKSTAGDTHLGGEDFDNRLVN  
180 190 200 210 220 230gi | 729 HFSNEFKRKHKDLSDNARALRRLRTACERAKRTLSSSAQTSIEIDSLFEGIDFFTSNTR  
240 250 260 270 280 290gi|729 ARFEEVGQDLFRGNMEPGERTLRDDKIDKSSVHEIVLGGGSTRI PKVQKLVSDFNGKEP  
300 310 320 330 340 350gi|729 CKSINPDEAVAYGAAVQAAILSGDTSSKSTKEILLLDVAPLSLGIETAGGVMTALIKRNT  
360 370 380 390 400 410gi|729 TIPTKKSETFSTFSDNQPGVLIQVFEGEARTKIDINLMGKFELSGIRPAPRGVPQIEVTF  
420 430 440 450 460 470

```
gi|729 DLDANGIMNVSALEKGTGKTNKIVITNDKGRLSKEEIERMLADAEEKYKEEDEAEAGRIQA
480          490          500          510          520          530
```

gi|729 KNGLESYAYSLKNTVSDPKVEEKL~~SAEDKET~~LTGAIDKTVAWIDENQTATKEEYEA~~EQKQ~~

gi|729 LESVANPVMKKIYGAEGGAPGGMPGQGAGAPPPGAGDDGPTVEEVD  
600 610 620 630 640

```
>>gi|2739154 gi|2739154|gb|AAC67308.1| 22.6 kDa tegument (191 aa)
initn: 31 initl: 31 opt: 54 Z-score: 85.6 bits: 21.3 E(): 8.1
Smith-Waterman score: 54; 29.825% identity (52.632% similar) in 57 aa overlap
(22-73:54-110)
```

CV127\_ 10 20  
HGT VYANYAVEHS D L L A F G V R F D D R V T  
:::  
gi|273 MATTEYRLSLMEQFIRAFIEIDKDNNELIDKQELTKYCQNQMDMKQIDPWIARFDTDKD  
10 20 30 40 50 60

```

          30          40          50          60          70          80
CV127_ GK--LEAFASR--AKIVHIDIDS AEIGKNKTPHVSVCG-DVKLALQGMNKVLENRAE
      ::  ::  ::      . . . . . . . . . . . . . . . . . . . . . .
gi|273  GKSLEEFCRGFGLKLVWEVRREKEELKKDKEGKVSTLPLDIQIIAATMSKAKQYNICCKF
          70          80          90          100         110         120

```

gi | 273 KELLDKTSRTGDEVRAVANDLKAFLDSEYGRVWQVIIITGSYWMNFSHEPFLSMQFKYSN  
130 140 150 160 170 180

gi|273 YVCLLW RTPSS  
190

```
>>gi|29839547 gi|29839547|sp|Q8GT41.1|PLA1_PLAAC RecName (179 aa)
  initn: 29 initl: 29 opt: 53 Z-score: 84.4 bits: 21.0 E(): 9.5
Smith-Waterman score: 53; 29.231% identity (49.231% similar) in 65 aa overlap
(15-79:15-72)
```

```

      10      20      30      40      50      60
CV127_ HGT VYANYAVEHSDLLLAFGVRFD RVTGKLEAFASRAKIVHIDIDS AEIGKNKTPHVS V
      :::  .:  : : :  .  :.:  :.  :.  :.:  :..
gi | 298 MKLSFSLC IFFFNLLLLQAVISADIVQG TCKKVAQRSPNVNYDFCVKSLGADPKSHTA-
      10      20      30      40      50
      70      80

```

```
CV127_ CGDVKLALQGMNKVLENRAE
      :  ::...  : :
gi|298 --D----LQGLGVISANLAIQHGSKIQTFIGRILKSKVDPALKKYLNDCVGLYADAKSSV
      60          70          80          90         100         110

gi|298 QEAIADFKSKDYASANVKMSAALDDSVTCEDGFKEKKGIVSPVTKENKDYVQLTAISLAI
      120         130         140         150         160         170

gi|298 TKLLGA
```

80 residues in 1 query sequences  
307888 residues in 1386 library sequences  
Scomplib [34.26]  
start: Wed Mar 11 06:01:46 2009 done: Wed Mar 11 06:01:46 2009  
Total Scan time: 0.100 Total Display time: 0.040

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

### CV127\_564\_941\_frame2, amino acids 38-117

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

Please cite:

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

CV127\_564\_941\_frame2\_38-117, 80 aa  
vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences  
Expectation\_n fit:  $\rho(\ln(x)) = 4.0968 \pm 0.00298$ ;  $\mu = 5.9948 \pm 0.154$   
mean\_var=31.1106 +/- 8.603, 0's: 8 Z-trim: 9 B-trim: 47 in 1/42  
Lambda= 0.229943

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

The best scores are:

		opt bits	E(1386)	%_id	%_sim	alen
gi 4240395	gi 4240395 gb AAD13531.1 ( 492)	64	26.9	0.43	0.284	0.554 74
gi 15984	gi 15984 emb CAA34486.1  u ( 380)	56	24.2	2.1	0.304	0.609 23
gi 166317	gi 166317 gb AAA32629.1  ( 380)	56	24.2	2.1	0.304	0.609 23
gi 190358935	gi 190358935 sp P00785 ( 380)	56	24.2	2.1	0.304	0.609 23
gi 2253610	gi 2253610 gb AAC34737.1 ( 274)	54	23.5	2.5	0.271	0.604 48
gi 18652047	gi 18652047 gb AAL76932 ( 154)	51	22.5	2.9	0.205	0.718 39
gi 1168696	gi 1168696 sp P43187.1 A ( 205)	52	22.9	3	0.293	0.655 58
gi 4572592	gi 4572592 gb AAD13530.2 ( 412)	53	23.3	4.6	0.417	0.667 24
gi 37548637	gi 37548637 gb AAN11300 ( 236)	49	21.9	6.8	0.269	0.692 26
gi 116174180	gi 116174180 emb CAL59 ( 331)	50	22.2	7.4	0.327	0.564 55
gi 89892727	gi 89892727 gb ABD79097 ( 404)	50	22.3	9	0.333	0.565 69
gi 1336809	gi 1336809 gb AAB36117.1 ( 58)	42	19.4	9.1	0.375	0.583 24
gi 89892725	gi 89892725 gb ABD79096 ( 410)	50	22.3	9.1	0.333	0.565 69
gi 109157163	gi 109157163 pdb 2ATM  ( 331)	49	21.9	9.3	0.327	0.564 55
gi 1346323	gi 1346323 sp P49370.1 H ( 331)	49	21.9	9.3	0.327	0.564 55

>>>CV127\_564\_941\_frame2\_38-117, 80 aa vs  
/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

>>gi|4240395 gi|4240395|gb|AAD13531.1| major allergen B1 (492 aa)  
 initn: 56 initl: 56 opt: 64 Z-score: 108.5 bits: 26.9 E(): 0.43  
 Smith-Waterman score: 64; 28.378% identity (55.405% similar) in 74 aa overlap  
 (5-75:28-98)

```

                                10      20      30
CV127_      SSSLKKIGSLGLKIEI-AIGKIN--LGRVWEKVASNI
              . .:  ::  .:  ::  ..  .  ::.  ..
gi|424  NAIEFLNNIHDLLGIPHIPVTARKHHRGVGITGLIDDIAILPVDDLYALFQEKLETSP
              10      20      30      40      50      60

              40      50      60      70      80
CV127_  LFNSVQELLRRLLVYRKAQTDVWKLETAIKYONLYIKAEQKGPSLFI
              :...  .  .  .  :  :  :  :  :  :  :  :  :  :  :  :
gi|424  EFKALYDAIRSPEF---QSIVGTLEAMPEYQNLIQKLKDKGVDVDHIELIHQIFNIVRD
              70      80      90      100     110

gi|424  TRGLPEDLQDFLALIPTDQVLAIAADYLANDAEVKAABVEYLKSDEFETIVVTVDLPEFK
              120     130     140     150     160     170

gi|424  NFLNFLQTNGLNAIEFLNNIHDLLGIPHIPVTGRKHLRRGVGITGLIDDIAILPVDDLY
              180     190     200     210     220     230

gi|424  ALFQEKLETSPFEKALYDAIRSPEFQSIVETLKAMPEYQSLIQKLKDKGVDVDHIELIH
              240     250     260     270     280     290

gi|424  QIFNIVRDTRGLPEDLQDFLALIPIDQILAIADYLANDAEVQAABVEYLKSDEFETIVVT
              300     310     320     330     340     350

gi|424  VDSLPEFKNFLNFLQTNGLNAIEFINNIHDLLGIPHIPATGRKHVRRGVGINGLIDDVIA
              360     370     380     390     400     410

gi|424  ILPVDELYALFQEKLESSPEFKALYDAIRSPEFQSIVQTLKAMPEYQDLIQRLKDKGVDV
              420     430     440     450     460     470

gi|424  DHFIELIKKLFGLSH
              480     490

```

>>gi|15984 gi|15984|emb|CAA34486.1| unnamed protein prod (380 aa)  
 initn: 35 initl: 35 opt: 56 Z-score: 96.0 bits: 24.2 E(): 2.1  
 Smith-Waterman score: 56; 30.435% identity (60.870% similar) in 23 aa overlap  
 (56-78:248-270)

```

              10      20      30      40      50      60
gi|159  MGLPKSFVSMSELLFFSTLLILSLAFNAKNLTQRTNDEVKAMYESWLIKYGKSYNSLGEWE
              70      80      90      100     110     120

gi|159  RRFEIFKETLRFIDEHNADTNRSYKVGLNQFADLTDEEFRSTYLGFSTSGSNKTKVSNRYE
              130     140     150     160     170     180

              10      20      30      40
CV127_      SSSLKKIGSLGLKIEIAIGKINLGRVWEKVASNILFNSVQELLRRLLVY
              190     200     210     220     230     240

              50      60      70      80
CV127_  RKAQTDVWKLETAIKYONLYIKAEQKGPSLFI
              :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
gi|159  NVPYNNEWALQTAVTYQPVSVDAAGDAFKHYSSGIFTGPCGTAIDHAVTIVGYGTEGG

```



```

                250          260          270          280          290          300
gi|159 IDYWIVKNSWDTTWGEEGYMRILRNVGGAGTCGIATMPSYPVKYNNQNHKPKPYSSLINPP
                310          320          330          340          350          360

gi|159 AFSMSKDGPGVGVDDGQRYSA
                370          380

>>gi|166317 gi|166317|gb|AAA32629.1| actinidin (380 aa)
  initn: 35 initl: 35 opt: 56 Z-score: 96.0 bits: 24.2 E(): 2.1
Smith-Waterman score: 56; 30.435% identity (60.870% similar) in 23 aa overlap
(56-78:248-270)

gi|166 MGLPKSFVMSLLFFSTLLILSLAFNAKNLTQRTNDEVKAMYESWLIKYGKSYNSLGEWE
                10          20          30          40          50          60

gi|166 RRFEIFKETLRFIDEHNADTNRSYKVGLNQFADLTDEEFRSTYLGFSTSGSNKTKVSNRYE
                70          80          90          100         110         120

gi|166 PRVGQVLPSYVDWRSAGAVVDIKSQGECGGCWAFAIATVEGINKIVTGVLIISLSEQELI
                130         140         150         160         170         180

                10          20          30          40
CV127_          SSSLKKIGSLGLKIEIAIGKINLGRVWEKVASNILFNSVQELLRRLVY

gi|166 DCGRTQNTRGCGGYITDGFQFIINNGGINTEENYPYTAQDGECNVELQNEKYVTIDTYE
                190         200         210         220         230         240

                50          60          70          80
CV127_ RKAQTDVWKLETAIKYQONLYIKAEQKGPSLFI
          : : : : : : : : : : : : : : : :
gi|166 NVPYNNEWALQTAVTYQPVSVLDAAGDAFKQYSSGIFTGPCGTDAIDHAVTIVGYGTEGG
                250         260         270         280         290         300

gi|166 IDYWIVKNSWDTTWGEEGYMRILRNVGGAGTCGIATMPSYPVKYNNQNYPEPYSSLINPP
                310         320         330         340         350         360

gi|166 AFSMSKDGPGVGVDDGQRYSA
                370         380

>>gi|190358935 gi|190358935|sp|P00785.4|ACTN_ACTCH RecNa (380 aa)
  initn: 35 initl: 35 opt: 56 Z-score: 96.0 bits: 24.2 E(): 2.1
Smith-Waterman score: 56; 30.435% identity (60.870% similar) in 23 aa overlap
(56-78:248-270)

gi|190 MGLPKSFVMSLLFFSTLLILSLAFNAKNLTQRTNDEVKAMYESWLIKYGKSYNSLGEWE
                10          20          30          40          50          60

gi|190 RRFEIFKETLRFIDEHNADTNRSYKVGLNQFADLTDEEFRSTYLRFSTSGSNKTKVSNRYE
                70          80          90          100         110         120

gi|190 PRVGQVLPSYVDWRSAGAVVDIKSQGECGGCWAFAIATVEGINKIVTGVLIISLSEQELI
                130         140         150         160         170         180

                10          20          30          40
CV127_          SSSLKKIGSLGLKIEIAIGKINLGRVWEKVASNILFNSVQELLRRLVY

gi|190 DCGRTQNTRGCGGYITDGFQFIINNGGINTEENYPYTAQDGECNVDLQNEKYVTIDTYE
                190         200         210         220         230         240

                50          60          70          80
CV127_ RKAQTDVWKLETAIKYQONLYIKAEQKGPSLFI

```

```

      : : : : : : : : : : : :
gi|190 NVPYNNEWALQTAVTYQPVSVVALDAAGDAFKQYSSGIFTGPCGTAVDHAHTIVGYGTEGG
      250      260      270      280      290      300

gi|190 IDYWIVKNSWDTTWGEEGYMRILRNVGGAAGCTCGIATMPSYPVKYNNQNHKPKPYSSLINPP
      310      320      330      340      350      360

gi|190 AFSMSKDGPGVGVDDGQRYSA
      370      380

```

>>gi|2253610 gi|2253610|gb|AAC34737.1| Cr-PII allergen [ (274 aa)  
 initn: 46 initl: 46 opt: 54 Z-score: 94.8 bits: 23.5 E(): 2.5  
 Smith-Waterman score: 54; 27.083% identity (60.417% similar) in 48 aa overlap  
 (28-75:25-69)

```

      10      20      30      40      50      60
CV127_ SSSLKKIGSLGLKIEIAIGKINLGRVWEKVASNILFNSVQELLRRLVYRKAQTDVWKLET
      : : : : : : : : : : : : : : : :
gi|225 VGV DGLIDDI IAILPIDDLKALFQEKLETSLDFKAFYDAVRSPEF--QSIVQTLNA
      10      20      30      40      50

      70      80
CV127_ AIKYQONLYIKAEQKGPSLFI
      : : : : : : : : : :
gi|225 MPEYQDLLQKLRDKGVDVDHYIELIRALFGLTREARNLQDDLNDLFLALIPTDQILAIAMD
      60      70      80      90      100      110

gi|225 YLANDAEVQELVAYLQSDDFHKIINTIEALPEFANFYNFLKGHGLDVANYINEIHSIIGL
      120      130      140      150      160      170

gi|225 PPFVPPSRRHARRGVGINGLIDDVIAILPVDELKTLFQEKLETSPDFKALYDAIRSPEFQ
      180      190      200      210      220      230

gi|225 SIISTLNAMPEYQELLQNLRDKGVDVDHFIELIRSWFGLP
      240      250      260      270

```

>>gi|18652047 gi|18652047|gb|AAL76932.1|AF456481\_1 major (154 aa)  
 initn: 37 initl: 37 opt: 51 Z-score: 93.7 bits: 22.5 E(): 2.9  
 Smith-Waterman score: 51; 20.513% identity (71.795% similar) in 39 aa overlap  
 (9-47:67-104)

```

CV127_ SS
gi|186 MG VQKTEVEAPSTVSAEKMYQGFLDMDTVFPKVLPLIKSVEILEGDDGGVGTVRLVHLG
      10      20      30      40      50      60

      10      20      30      40      50      60
CV127_ SLKKIGSLGLKIEIAIGKINLGRVWEKVASNILFNSVQELLRRLVYRKAQTDVWKLETAI
      .. : : : : : : : : : : : : : : : :
gi|186 EATEYTTMKQKVDV-IDKAGLGYTYTTIGGDILVEGLESVVNQFVVVPTDGGCIVKNTTI
      70      80      90      100      110

      70      80
CV127_ KYQONLYIKAEQKGPSLFI
gi|186 YNTKGDAVLPEDKVKEATEKSALAFKAVEAYLLAN
      120      130      140      150

```

>>gi|1168696 gi|1168696|sp|P43187.1|ALLB3\_BETVE RecName: (205 aa)  
 initn: 43 initl: 43 opt: 52 Z-score: 93.4 bits: 22.9 E(): 3

Smith-Waterman score: 52; 29.310% identity (65.517% similar) in 58 aa overlap (9-63:28-82)

```

              10      20      30
CV127_      SSSLKKIGSLGLKIEIAIGKINLGRVWE--KVASNILFNS
              :. :. : ... : : ... : ... :
gi|116 MPCSTEAMEKAGHGHASTPRKRSLSNSFRRLRSE-SLNTLRRLRRIFDLFDKNSDGII--T
              10      20      30      40      50

              40      50      60      70      80
CV127_ VQELLRRLVYRKAQTDVWKLETAIKYQONLYIKAEQKGPSLFI
              :. :. : : :. :. :. :. :
gi|116 VDELSRALNLLGLETDLSELESTVKSFRTREGNIGLQFEDFISLHQSINDSYFAYGGEDED
              60      70      80      90      100     110

gi|116 DNEEDMRKSILSQEEADSFGGFKVFDDEDGDGYISARELQMVLGKLGFESEIDRVEKMI
              120     130     140     150     160     170

gi|116 VSVDSNRDGRVDFFFEFKDMMRSVLVRS
              180     190     200

```

>>gi|4572592 gi|4572592|gb|AAD13530.2| major allergen B1 (412 aa)  
 initn: 53 initl: 53 opt: 53 Z-score: 90.0 bits: 23.3 E(): 4.6  
 Smith-Waterman score: 53; 41.667% identity (66.667% similar) in 24 aa overlap (52-75:179-202)

```

gi|457 NLEKLREKGVVDVKIIEELIRALFGLTLNAKASRNLDLQDFLALIPVDQIIAIATDYL
              10      20      30      40      50      60

gi|457 ANDAEVQAAYVAYLQSDEFETIVVALDALPELQNFLEANGLNIDFLNGIHDLLGIPH
              70      80      90      100     110     120

              10      20      30      40      50
CV127_      SSSLKKIGSLGLKIEIAIGKINLGRVWEKVASNILFNSVQELLRRLVYRKAQT
              :.
gi|457 IPVSGRKYHIRRGVGITGLIDDLAILPIEDLKALFNEKLETSPDFLALYNAIRSPEFQS
              130     140     150     160     170     180

              60      70      80
CV127_ DVWKLETAIKYQONLYIKAEQKGPSLFI
              : :. :. :. :. :. :
gi|457 IVQTLNAMPEYQNLQKLREKGVVDVKIIEELIRALFGLTLNGKASRNLDLQDFLALIP
              190     200     210     220     230     240

gi|457 VDQIIAIATDYLANDAEVQAAYVAYLQSDEFETIVVTLDALPELQNFLEANGLNIDF
              250     260     270     280     290     300

gi|457 LNGIHDLLGIPHIPVSGRKYHIRRGVGITGLIDDLAILPLDDLKALFNEKLETSPDFLA
              310     320     330     340     350     360

gi|457 LYNAIKSPEFQSIVQTLNAMPEYQNLLEKLREKGVVDVKIIEELIRALFGLTH
              370     380     390     400     410

```

>>gi|37548637 gi|37548637|gb|AAN11300.1| 29 kDa IgE-bing (236 aa)  
 initn: 29 initl: 29 opt: 49 Z-score: 87.0 bits: 21.9 E(): 6.8  
 Smith-Waterman score: 49; 26.923% identity (69.231% similar) in 26 aa overlap (54-79:117-142)

```

gi|375 MVKVLLALTSYNETFYSDGKKTGVFVVEALHPFEVFRKKGYEIQLASSETGTFGWDDHSVV
              10      20      30      40      50      60

              10      20      30      40      50

```

CV127\_ SSSLKKIGSLGLKIEIAIGKINLGRVWEKVASNILFNSVQELLRRLLVYRKAQTDVWK : :

gi|375 PDFLNGEDKEIFDQVNSEFNVALKNLKKASDLDPNQYDIFFGSAGHGTLFDYPHAKDLQK  
70 80 90 100 110 120

60                      70                      80

CV127 LETAIKYONLYIKAEOKGPSLFI

[illegible]gi|375 IATTVYNKGGVVS AVCHGPAIFENLNDPKTGEPLIKGKKITGFTDIGEDILGVTDIMKKG  
130 140 150 160 170 180gi|375 NLLTIKQVAEKEGATYIEPEGPWDNFTVTDDGRIVTGVNPPQSAVKTAEDVIAAFECN  
190 200 210 220 230

```
>>gi|116174180 gi|116174180|emb|CAL59818.1| hyaluronidas (331 aa)
initn: 31 initl: 31 opt: 50 Z-score: 86.3 bits: 22.2 E(): 7.4
Smith-Waterman score: 50; 32.727% identity (56.364% similar) in 55 aa overlap
(18-69:234-287)
```

gi|116 SERPKRVFNIYWNVPTFMCHQYDLYFDEVTFNFIKRNSKDDFQGDKIAIFYDPGEFPALL  
10 20 30 40 50 60gi|116 SLKDGKYKRRNGGVPQEGNITIHLLQKFIEENLDKTYPNRNFSGIGVIDFERWRPIFRQNWG  
70 80 90 100 110 120gi|116 NMKIHKNFSIDLVRNEHPTWNKKMIELEASKRFEKYARFFMEETLKLAKKTRKQADWGYY  
130 140 150 160 170 180

CV127 SSSLKKIGSLGLKIEIAIGKINLG

\_\_\_\_\_

gi|116 GYPYCFNMSPNLVP ECDVTAMHENDKMSWLFNNQNVL LPSVYVRQELTPDQRIGLVQ-G  
190 200 210 220 230

CV127 RVWEKV-ASNILFNSVOELLRRLVYRKAOPTDVKLETAIK--YONLYIKAEOKGPSLFI

[illegible]gi|116 RVKEAVRISNNLKHSPKVL<sup>240</sup>SYWWYVYQ<sup>250</sup>DETNTYLTETDVKKTFQEIVINGGDGIIIWGSS<sup>290</sup>

gi|116 SDVNSLSKCKRLQDYLLTVLGPIAVNVTEAVN  
300 310 320 330

```
>>gi|89892727 gi|89892727|gb|ABD79097.1| Zea m 13 allerg (404 aa)
initn: 45 initl: 45 opt: 50 Z-score: 84.8 bits: 22.3 E(): 9
Smith-Waterman score: 50; 33.333% identity (56.522% similar) in 69 aa overlap
(16-80:252-316)
```

gi|898 ARGALFLLALFCVVHGEKAKSKDNDAKASGPGGSFDITKLGASGNGKTDSTKAVQEAWAS

gi|898 ACGGTGKQTILIPKGDFLVGPLNFTGPCKGDVTIQVNGNLLATTDLSQYKDHGNWIEILR  
70 80 90 100 110 120gi|898 VDNLVITGKGKLDGQGPVWSKNSCVKKYDCKILPNSLVMDFVNNGEVSGITLLNSKFFH  
130 140 150 160 170 180

CV127 SSSL

gi|898 MNMYKCKDMLIKDVNVTAPGDSPTNDGIHMGDSSGVITITNTVIGVGDDCISIGPGTFKVN  
190 200 210 220 230 240

```

      10      20      30      40      50      60
CV127_ KKIGSLGLKIEIAIGKINLGRVW-EKVASNILFNSVQELLRRL---VYRKAQTDVWKLET
      .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:.
gi|898 ITGVTCGPGHGIGSIG--SLGRYKDEKDVTDI--NVKDCTLKKTANGVRIKAYEDAASVLT
      250      260      270      280      290

```

```

      70      80
CV127_ AIKYQNLVIKAEQKGPSLFI
      .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:.
gi|898 ASKIHENIKMEDSGYPIIIDMKYCPNKLCTANGASKVTVKDVTFKNIPGTSSTPEAVNL
      300      310      320      330      340      350

```

```

gi|898 LCSAKIPCTGVTMDDVNIKYSGTNNKTMVCKNAKGSAGCLKELACF
      360      370      380      390      400

```

```

>>gi|1336809 gi|1336809|gb|AAB36117.1| Sol i 1=antigen { (58 aa)
  initn: 42 initl: 42 opt: 42 Z-score: 84.7 bits: 19.4 E(): 9.1
Smith-Waterman score: 42; 37.500% identity (58.333% similar) in 24 aa overlap
(1-24:4-27)

```

```

      10      20      30      40      50
CV127_ SSSLKKIGSLGLKIEIAIGKINLGRVWEKVASNILFNSVQELLRRLVYRKAQTDVWK
      .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:.
gi|133 GAASNTYDIGQLMAKYTVDLMNKLNQSQCDINASLPVVFITHGFTSSAQVSTFK
      10      20      30      40      50

```

```

      60      70      80
CV127_ LETAIKYQNLVIKAEQKGPSLFI

```

```

>>gi|89892725 gi|89892725|gb|ABD79096.1| Zea m 13 allerg (410 aa)
  initn: 45 initl: 45 opt: 50 Z-score: 84.7 bits: 22.3 E(): 9.1
Smith-Waterman score: 50; 33.333% identity (56.522% similar) in 69 aa overlap
(16-80:258-322)

```

```

gi|898 MACIDNAMRALFLLALFCVVHGEKAKSKDNDKASGPGGSFDITKLGASGNGKTDSTKAV
      10      20      30      40      50      60

```

```

gi|898 QEAWASACGGTGKQITILIPKGDFLVGPLNFTGPCKGDVTIQVNGNLLATTDLSQYKDHGN
      70      80      90      100      110      120

```

```

gi|898 WIEILRVDNLVITGKGKLDGQGPVWVSKNSCVKKYDCKILPNSLVMDFVNNGEVSGITLL
      130      140      150      160      170      180

```

```

gi|898 NSKFFHMNMYKCKDMLIKDVNVTAPGDSPTDGIHMGDSSGVTITNTVIGVGDDCISIGP
      190      200      210      220      230      240

```

```

      10      20      30      40      50
CV127_ SSSLKKIGSLGLKIEIAIGKINLGRVW-EKVASNILFNSVQELLRRL---VYRKAQTD
      .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:.
gi|898 GTSKVNTITGVTCGPGHGIGSIG--SLGRYKDEKDVTDI--NVKDCTLKKTANGVRIKAYED
      250      260      270      280      290

```

```

      60      70      80
CV127_ VWKLETAIKYQNLVIKAEQKGPSLFI
      .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:. .:.
gi|898 AASVLTASKIHENIKMEDSGYPIIIDMKYCPNKLCTANGASKVTVKDVTFKNITGTSST
      300      310      320      330      340      350

```

```

gi|898 PEAVNLLCSAKIPCTGVTMDDVNIKYSGTNNKTMVCKNAKGSAGCLKELACF
      360      370      380      390      400      410

```

```
>>gi|109157163 gi|109157163|pdb|2ATM|A Chain A, Crystal (331 aa)
  initn: 30 initl: 30 opt: 49 Z-score: 84.5 bits: 21.9 E(): 9.3
Smith-Waterman score: 49; 32.727% identity (56.364% similar) in 55 aa overlap
(18-69:234-287)
```

gi|109 SERPKRVFNIYWNVPFTFMCHQYDLYFDEVTFNFIKRNSKDDFQGDKIAIFYDPGEFPALL  
10 20 30 40 50 60gi|109 SLKDGKYKRRGGVPQEGNITIH LQKF IENLDKIYPNRNFSGIGVIDFERWRPIFRQNWG  
70 80 90 100 110 120gi|109 NMKIHKNFSIDLVRNEHPTWNKKXIELEASKRFEKYARFFMEETLKLAKKTRKQADWGYY  
130 140 150 160 170 180

CV127\_ SSSLKKIGSLGLKIEIAIGKINLG  
10 20  
:: :: :

gi|109 GYPYCFNMSPNNLVPECDVTAMHENDKMSWLFNNQNVLLPSVYVRQELTPDQRIGLVQ-G  
190 200 210 220 230

CV127\_ RVWEKV-ASNILFNSVQELLRRLVYRKAQTDVWKLETAIK--YQONLYIKAEQKGPSLFI

gi|109 RVKEAVRISNNLKHSPKVL<sup>240</sup>SYWWYVYQ<sup>250</sup>DETNTFLT<sup>260</sup>ETDVKKTF<sup>270</sup>QEI<sup>280</sup>VINGGDGII<sup>290</sup>WGSS

gi|109 SDVNSLSKCKRLQDYLLTVLGPIAINVTEAVN  
300 310 320 330

```
>>gi|1346323 gi|1346323|sp|P49370.1|HUGAA_VESVU RecName: (331 aa)
  initn: 30 initl: 30 opt: 49 Z-score: 84.5 bits: 21.9 E(): 9.3
Smith-Waterman score: 49; 32.727% identity (56.364% similar) in 55 aa overlap
(18-69;234-287)
```

gi|134 SERPKRVFNIIYWNVPFTFMCHQYDLYFDEVTFNFIKRNSKDDFQGDKIAIFYDPGEFPALL  
10 20 30 40 50 60gi | 134 SLKDGKYKRNGGVPQEGNITIH LQKF IENLDKIYPN RNFSGIGVIDFERWRPIFRQNWG  
70 80 90 100 110 120

gi | 134 NMKIHKNFSIDLVRNEHPTWNKKMIELEASKRFEKYARFFMEETLKLAKKTRKQADWGY

CV127\_ SSSLKKIGSLGLKIEIAIGKINLG  
10 20  
::: :

gi | 134 GYPYCFNMSPNNLVPECDVTAMHENDKMSWLFNNQNVLLPSVYVRQELTPDQRIGLVQ-G  
190 200 210 220 230

CV127\_ RVWEKV-ASNILFNSVQELLRRLVYRKAQTDVWKLETAIK--YONLYYIKAEQKGPSLFI

gi|134 RVKEAVRISNNLKHSPKVLSYWWYVYQDETNTFLTETDVKKTFQEIVINGGDGIIIWGSS  
240 250 260 270 280 290

gi|134 SDVNSLSKCKRLQDYLLTVLGPIAINVTEAVN  
300 310 320 330

80 residues in 1 query sequences

307888 residues in 1386 library sequences  
Scomplib [34.26]  
start: Wed Mar 11 06:29:13 2009 done: Wed Mar 11 06:29:13 2009  
Total Scan time: 0.100 Total Display time: 0.030

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

## CV127\_rc\_988\_1344\_frame0, amino acids 1-80

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

Please cite:

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

CV127\_rc\_988\_1344\_frame0\_1-80, 80 aa  
vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences  
Expectation\_n fit:  $\rho(\ln(x)) = 3.7122 \pm 0.00348$ ;  $\mu = 8.6415 \pm 0.182$   
mean\_var =  $31.7298 \pm 8.372$ , 0's: 7 Z-trim: 11 B-trim: 61 in 1/43  
Lambda = 0.227688

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

				opt	bits	E(1386)	%_id	%_sim	alen
gi 1460058	gi 1460058 emb CAA46317.1	( 211)		66	27.2	0.15	0.296	0.493	71
gi 61608445	gi 61608445 gb AA47076.1	( 216)		66	27.2	0.16	0.296	0.493	71
gi 83754033	gi 83754033 pdb 2AS8 B	( 222)		66	27.2	0.16	0.296	0.493	71
gi 730036	gi 730036 sp P08176.2 PEP	( 320)		66	27.3	0.22	0.296	0.493	71
gi 21725562	gi 21725562 emb CAD3836.1	( 222)		59	24.9	0.78	0.387	0.645	31
gi 21725568	gi 21725568 emb CAD3836.1	( 222)		59	24.9	0.78	0.387	0.645	31
gi 21725566	gi 21725566 emb CAD3836.1	( 222)		59	24.9	0.78	0.387	0.645	31
gi 21725572	gi 21725572 emb CAD3836.1	( 222)		59	24.9	0.78	0.387	0.645	31
gi 21725560	gi 21725560 emb CAD3836.1	( 222)		59	24.9	0.78	0.387	0.645	31
gi 21725564	gi 21725564 emb CAD3836.1	( 222)		59	24.9	0.78	0.387	0.645	31
gi 21725570	gi 21725570 emb CAD3836.1	( 222)		59	24.9	0.78	0.387	0.645	31
gi 62550933	gi 62550933 emb CAI7905.1	( 326)		58	24.7	1.4	0.255	0.617	47
gi 157696052	gi 157696052 gb ABV662.1	( 223)		55	23.6	2	0.282	0.479	71
gi 4538529	gi 4538529 emb CAB39376.1	( 81)		51	22.1	2.1	0.366	0.512	41
gi 170734	gi 170734 gb AAA34287.1	( 244)		54	23.3	2.7	0.333	0.667	33
gi 21773	gi 21773 emb CAA31685.1	u ( 307)		54	23.3	3.2	0.333	0.667	33
gi 21725580	gi 21725580 emb CAD3837.1	( 222)		51	22.3	4.8	0.355	0.645	31
gi 21725576	gi 21725576 emb CAD3836.1	( 222)		51	22.3	4.8	0.355	0.645	31
gi 19847822	gi 19847822 gb AAK27264.1	( 306)		52	22.7	5.1	0.292	0.585	65
gi 21725578	gi 21725578 emb CAD3837.1	( 222)		50	21.9	6.1	0.355	0.645	31
gi 21725574	gi 21725574 emb CAD3836.1	( 222)		50	21.9	6.1	0.355	0.645	31
gi 49523394	gi 49523394 emb CAE5283.1	( 377)		51	22.4	7.6	0.214	0.518	56
gi 21783	gi 21783 emb CAA30570.1	u ( 356)		50	22.1	9.1	0.294	0.647	34
gi 51242679	gi 51242679 gb AAT99258.1	( 362)		50	22.1	9.2	0.357	0.643	28

>>>CV127\_rc\_988\_1344\_frame0\_1-80, 80 aa vs  
/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

>>gi|1460058 gi|1460058|emb|CAA46317.1| cysteine proteas (211 aa)  
initn: 49 initl: 49 opt: 66 Z-score: 116.6 bits: 27.2 E(): 0.15  
Smith-Waterman score: 66; 29.577% identity (49.296% similar) in 71 aa overlap  
(8-71:105-175)

gi|146 ETNACSINGNAPAEIDLRLQMRVTPIRMQGGCGSCWAFSGVAATESAYLAYRNQSLDLAE

Page 143 of 237





Smith-Waterman score: 59; 38.710% identity (64.516% similar) in 31 aa overlap (8-37:104-134)

```
gi|217 TNACSINGNAPAEIDLRQMRTVTTIRMQGGCGSCWAFSGVAATESAYLAVRNQSLDLAEQ
      10          20          30          40          50          60

CV127_                                10          20
                                WAWDSFPKRLKAQRKLLFLYVQLI
                                .: .: .: .: .: .: .: .: .: .: .: .:
gi|217 ELVDCANQHGCHGDTIPRGIEYIQHNGVVQESYYRYVAEEQSCRRPNAQRFGISNYCQIY
      70          80          90          100          110          120

      30          40          50          60          70          80
CV127_ PPN-SKIKLKLRLSVLKNLIHPLQSQLNITTHRHMRSLILPNLSRVNINMNNLSPTS
      .: .: .: .: .: .: .: .: .: .: .: .:
gi|217 PPNVVKIREALAQTHSAIAVIIGIKDLDAFRHYDGQTIIQQDNGYQTNHVNIVGYNSA
      130          140          150          160          170          180

gi|217 QGVDYWIVRNSWDTNWGDNGYGYFAANIDLMMIEEYPYVVIL
      190          200          210          220
```

>>gi|21725566 gi|21725566|emb|CAD38364.1| unnamed protei (222 aa)  
 initn: 54 initl: 54 opt: 59 Z-score: 103.8 bits: 24.9 E(): 0.78  
 Smith-Waterman score: 59; 38.710% identity (64.516% similar) in 31 aa overlap (8-37:104-134)

```
gi|217 TNACSINGNAPAEIDLRQMRTVTPIRMQGGCGSCWAFSGVAATESAYLAVRNQSLDLAEQ
      10          20          30          40          50          60

CV127_                                10          20
                                WAWDSFPKRLKAQRKLLFLYVQLI
                                .: .: .: .: .: .: .: .: .: .: .: .:
gi|217 ELVDCASQHGCHGDTIPEGIEYIQHNGVVQESYYRYVAEEQSCRRPNAQRFGISNYCQIY
      70          80          90          100          110          120

      30          40          50          60          70          80
CV127_ PPN-SKIKLKLRLSVLKNLIHPLQSQLNITTHRHMRSLILPNLSRVNINMNNLSPTS
      .: .: .: .: .: .: .: .: .: .: .: .:
gi|217 PPNVVKIREALAQTHSAIAVIIGIKDLDAFRHYDGQTIIQQDNGYQTNHVNIVGYNSA
      130          140          150          160          170          180

gi|217 QGVDYWIVRNSWDTNWGDNGYGYFAANIDLMMIEEYPYVVIL
      190          200          210          220
```

>>gi|21725572 gi|21725572|emb|CAD38367.1| unnamed protei (222 aa)  
 initn: 54 initl: 54 opt: 59 Z-score: 103.8 bits: 24.9 E(): 0.78  
 Smith-Waterman score: 59; 38.710% identity (64.516% similar) in 31 aa overlap (8-37:104-134)

```
gi|217 TNACSINGNAPAEIDLRQMRTVTPIRMQGGCGSCWAFSGVAATESAYLAVRNQSLDLAEQ
      10          20          30          40          50          60

CV127_                                10          20
                                WAWDSFPKRLKAQRKLLFLYVQLI
                                .: .: .: .: .: .: .: .: .: .: .: .:
gi|217 ELVDCANQHGCHGDTIPRGIEYIQHNGVVQESYYRYVAEEQSCRRPNAQRFGISNYCQIY
      70          80          90          100          110          120

      30          40          50          60          70          80
CV127_ PPN-SKIKLKLRLSVLKNLIHPLQSQLNITTHRHMRSLILPNLSRVNINMNNLSPTS
      .: .: .: .: .: .: .: .: .: .: .: .:
gi|217 PPNVVKIREALAQTHSAIAVIIGIKDLDAFRHYDGETIIQEDNGYQTNHVNIVGYNSA
```



gi|157 YPQNVNKIREALAQTHSAIAVIIGIKDLDAFRHYDGRTIIQRDNGYQPNYHAVNIVGYSN

```

130      140      150      160      170      180

      80
CV127_ SPTS

gi|157 AQGVYWIWRNSWDTNWGDNGYGYFAANIDLMMIEEYPYVVIL
      190      200      210      220

>>gi|4538529 gi|4538529|emb|CAB39376.1| Cop c1 allergen (81 aa)
  initn: 45 initl: 45 opt: 51 Z-score: 96.2 bits: 22.1 E(): 2.1
Smith-Waterman score: 51; 36.585% identity (51.220% similar) in 41 aa overlap
(23-63:17-56)

      10      20      30      40      50      60
CV127_ WAWDSFPKRLKAQRKLLFLYVQLIPPNSKIKLLRSVLKNLIHPLQSQLNITTHRHMS
      . . . . .
gi|453 RFLPSSSHLNPQHLPWLVPAPVLLLPVLPQ-LKPVAPLPLPLDPTTLHMPP
      10      20      30      40      50

      70      80
CV127_ LILPNLSRVNINMNNLSPTS
      . . .
gi|453 LLLQLQLPPLLSQGNPACSPKWLQLLVP
      60      70      80

>>gi|170734 gi|170734|gb|AAA34287.1| gamma gliadin B-III (244 aa)
  initn: 38 initl: 38 opt: 54 Z-score: 94.3 bits: 23.3 E(): 2.7
Smith-Waterman score: 54; 33.333% identity (66.667% similar) in 33 aa overlap
(42-73:196-226)

gi|170 PQQPFPLQPQQSFLWQSQQPFLQPPQPPSPQPQQVQIISPATPTTIPSAGKPTSAPFPQ
      10      20      30      40      50      60

gi|170 QQQQHQQLAQQQIPVVQPSILQQLNPCVKVFLQQQCSPVAMPQRLARSQMLQQSSCHVMQQ
      70      80      90      100      110      120

      10      20
CV127_ WAWDSFPKRLKAQRKLLFLYVQLIPP

gi|170 QCCQQLPQIPQQSRYQAIRAIYSIILQEQQQVQGSISQQQQPQQLGQCVSQPQQQSQQ
      130      140      150      160      170      180

      30      40      50      60      70      80
CV127_ NSKIKLLRSVLKNLIHPLQ-SQLNITTHRHMSLILPNLSRVNINMNNLSPTS
      . . . . .
gi|170 QLGGQPQQQQLAQGTFLQPHQIAQLEVMTSIALR--ILPTMCSVNVPLYRTTTSVPFGVG
      190      200      210      220      230

gi|170 TGVGAY
      240

>>gi|21773 gi|21773|emb|CAA31685.1| unnamed protein prod (307 aa)
  initn: 38 initl: 38 opt: 54 Z-score: 92.8 bits: 23.3 E(): 3.2
Smith-Waterman score: 54; 33.333% identity (66.667% similar) in 33 aa overlap
(42-73:259-289)

gi|217 MKTFLVFAALLAVAATSIAIOMETRCIPGLERPWQQQPLPPQQTFFPQQPLFSQQQQQQQLFP
      10      20      30      40      50      60

gi|217 QQPSFSQQQPPFWQQQPPFSQQQPILPQQPPFSQQQQLVLPQQPPFSQQQPPVLPQQSP
      70      80      90      100      110      120

```

gi|217 QGVDYWIVRNSWDTNWGDNGYGYFAANIDLMMIEEYPYVVIL

190 200 210 220

>>gi|19847822 gi|19847822|gb|AAK27264.1| isoflavone redu (306 aa)  
 initn: 51 initl: 51 opt: 52 Z-score: 89.3 bits: 22.7 E(): 5.1  
 Smith-Waterman score: 52; 29.231% identity (58.462% similar) in 65 aa overlap  
 (12-76:25-83)

10 20 30 40  
 CV127\_ WAWDSFPKRLKAQRKLLFLYVQLIPPNSKIKLKLRSVLKLNLIHPLQ  
 :: . :: . : : : : : : : : : : : :  
 gi|198 MGGSRVLIIGGTGYIGRHVTNASLAQGHPTFLLVREITPSNPEKAQLLESFTSKGATLVQ  
 10 20 30 40 50 60

50 60 70 80  
 CV127\_ SQLNITTHRHMRSLILPNLSRVNINMNNLSPTS  
 . . : : : : : : : : : : : :  
 gi|198 G--SIDDHA---SLV-AALKKVDVVISTLGAPQIADQFNLIKAIKEVGTIKRFFPSEFGN  
 70 80 90 100 110

gi|198 DVDKHHAVEPMKSMFDLKIKLRRRTIEAEGIPHTYVVPFCFAGYFLTNLALGLAAPPDRK  
 120 130 140 150 160 170

gi|198 IVIYGDGTTKAVYMKEEDIGTFTIKAVDDPRTLNTLYLKPPANTISTNDLVALWEAKIG  
 180 190 200 210 220 230

gi|198 KTLKQVYLSEEQVLKLLQDTPFPPTFMVSIFHTIYVKGDQTNFQIGPDGVEASALYPDVK  
 240 250 260 270 280 290

gi|198 YTTVEEYISAFV  
 300

>>gi|21725578 gi|21725578|emb|CAD38370.1| unnamed protei (222 aa)  
 initn: 42 initl: 42 opt: 50 Z-score: 87.8 bits: 21.9 E(): 6.1  
 Smith-Waterman score: 50; 35.484% identity (64.516% similar) in 31 aa overlap  
 (8-37:104-134)

gi|217 TNACSSINGNAPASIDLRQMRTVTTIRMQGGCGSCWAFSGVAATESAYLAVRNQSLDLAEQ  
 10 20 30 40 50 60

10 20  
 CV127\_ WAWDSFPKRLKAQRKLLFLYVQLI  
 . : : : : : : : : : : :  
 gi|217 ELVDCASQHGCHGDTIPEGIEYIQHNGVVQESYRYVAQEQSCRPNADRFGISNYCQIY  
 70 80 90 100 110 120

30 40 50 60 70 80  
 CV127\_ PPN-SKIKLKLRSVLKLNLIHPLQSQLNITTHRHMRSLILPNLSRVNINMNNLSPTS  
 : : : : : : : : : : : :  
 gi|217 PPNVNKIEEALAQTHSAIAVIIGIKDLDAFRHYDGQTIIQEDNGYQTNHVNIVGYSNA  
 130 140 150 160 170 180

gi|217 QGVYDWIVRNSFDTNWGDNGYGYFAANIDLMMIEEYPYVIL  
 190 200 210 220

>>gi|21725574 gi|21725574|emb|CAD38368.1| unnamed protei (222 aa)  
 initn: 42 initl: 42 opt: 50 Z-score: 87.8 bits: 21.9 E(): 6.1  
 Smith-Waterman score: 50; 35.484% identity (64.516% similar) in 31 aa overlap  
 (8-37:104-134)

gi|217 TNACSSINGNAPASIDLRQMRTVTTIRMQGGCGSCWAFSGVAATESAYLAVRNQSLDLAEQ  
 10 20 30 40 50 60

```

                                10      20
CV127_ WAWDSFPKRLKAQRKLLFLYVQLI
                                .: .:.: . : :.
gi|217 ELVDCASQHGCHGDTIPEGIEYIQHNGVVQESYYRYVAQEQSCRRPNADRFGISNYCQIY
                                70      80      90      100     110     120

                                30      40      50      60      70      80
CV127_ PPN-SKIKLKLRSVLKNLIHPLQSQLNITTHRHMRSLILPNLSRVNINMNNLSPTS
                                ::: .:.: : ..
gi|217 PPNVNKIEEALQTHSAIAVIGIKDLDAFRHYDGQTIIQEDNGYQTNHVNIVGYSNA
                                130     140     150     160     170     180

gi|217 QGVYDWIVRNSWDTNWGDNGYGYFAANIDLMMIEEYPYVVIL
                                190     200     210     220

>>gi|49523394 gi|49523394|emb|CAE52833.1| polygalacturon (377 aa)
  initn: 42 initl: 42 opt: 51 Z-score: 86.1 bits: 22.4 E(): 7.6
Smith-Waterman score: 51; 21.429% identity (51.786% similar) in 56 aa overlap
(25-80:265-320)

gi|495 RGVQSSGSVFNVNDYGAKGAGDISQAVMKAWKAACASQGPSTVLIPKGNYNMGEVAMQGP
                                10      20      30      40      50      60

gi|495 CKGSKIGFQIDGVVKAPADPSKFKSDGWVSFYRIDGLTVSGTGTLDGQGQTAWAKNNCDK
                                70      80      90      100     110     120

gi|495 NPNCXHAAMNLRFDLKHAMVRDITSLNSKMFHINVLECEDITFQHVTVTAPGTSINTDG
                                130     140     150     160     170     180

gi|495 IHVGISKGVTTITNTKIATGDDCISIGPGSQNVTTITQVNCGPGHGISIGSLGRYNNEKEVR
                                190     200     210     220     230     240

                                10      20      30      40      50      60
CV127_ WAWDSFPKRLKAQRKLLFLYVQLIPPNSKIKLKLRSVLKNLIHPLQSQLNITTHRHMR
                                :... : . :... :. : . : .
gi|495 GITVKGCTFSGTMNGVRVKTWPNSPPGAATDLTFQDLTMNNVQNPVILDQEYCPYQCSR
                                250     260     270     280     290     300

                                70      80
CV127_ LILPNLSRVNINMNNLSPTS
                                .. :...: :
gi|495 QAPSRIKLSNINFNNIRGTSTGKVAVVIACSHGMPCSNMKIGEINLSYRGAGGPATSTCS
                                310     320     330     340     350     360

gi|495 NVKPTFSGKQVPAIKCA
                                370

>>gi|21783 gi|21783|emb|CAA30570.1| unnamed protein prod (356 aa)
  initn: 33 initl: 33 opt: 50 Z-score: 84.7 bits: 22.1 E(): 9.1
Smith-Waterman score: 50; 29.412% identity (64.706% similar) in 34 aa overlap
(40-73:308-338)

gi|217 MKTFLV FALLALAAASAVAQISQQQQAPPFSQQQQPPFSQQQQPPFSQQQQSPFSQQQQQ
                                10      20      30      40      50      60

gi|217 PPFAQQQQPPFSQQQPISQQQQPPFSQQQQPPFSQQQQPPYSQQQQPPYSQQQQPPFSQQ
                                70      80      90      100     110     120

gi|217 QQQPPFSQQQQPPFTQQQQQQQQQQPPFTQQQQPPFSQQQPISQQQQPPFLQQQRPPFSRQ
                                130     140     150     160     170     180

gi|217 QQIPVIHPSVLQQLNPCKVFLQQQCIPVAMQRCLARSQMLQQSICHVMQQQCCQQLRQIP

```



```

190      200      210      220      230      240
CV127_      WAWDSFPKRLKAQRKLLFLYVQLIPPNSKIKL
gi|217 EQSRHESIRAIISIIILQQQQQQQQQQQQQQQSIIQYQQQQPQQLGQCVSQPLQQQLQQQ
250      260      270      280      290      300

40      50      60      70      80
CV127_ KLLRSVLKNLIHPLQSQLNITTHRHMRSLILPNLSRVNINMNNLSPTS
..: : . :... : .: :... :...
gi|217 LGQQPQQQQLAHQI-AQLEVMTSIALRT--LPTMCNVNVPYETTTSVPLGVGIGVGVY
310      320      330      340      350

>>gi|51242679 gi|51242679|gb|AAT99258.1| pectin-methyltr (362 aa)
initn: 44 initl: 44 opt: 50 Z-score: 84.6 bits: 22.1 E(): 9.2
Smith-Waterman score: 50; 35.714% identity (64.286% similar) in 28 aa overlap
(22-49:24-47)

10      20      30      40      50
CV127_ WAWDSFPKRLKAQRKLLFLYVQLIPPNSKIKLKLRSVLKNLIHPLQSQLNITTHRM
::: : :... :
gi|512 MEEHVSMLLVGFVLINIAFTSIAQLIPPNPA---ELESWFQGAVKPVSEQGLEPSVVQ
10      20      30      40      50

60      70      80
CV127_ RSLILPNLSRVNINMNNLSPTS
gi|512 TESGGVETIEVRQDGSQKFKTISDAVKHVKVGNTKRVIIITIGPGEYREKVKIERLHPYIT
60      70      80      90      100      110

gi|512 LYGIDPKNRPTITFAGTAAEFQTVDSATVIVESDYSVGAHLIVTNSAPRPDGKRKGAQAG
120      130      140      150      160      170

gi|512 ALRISGDRAAFYNCKFTGFQDTCDDKGNHFFTDCTEGTVDFIFGEARSLYLNTLHV
180      190      200      210      220      230

gi|512 PGDPMAMITAHARKNADGVGGYSFVHCKVTGTGGTALLGRAWFDAARVVFSYCNLSDAAK
240      250      260      270      280      290

gi|512 PEGWSDNNKPEAQKTILFGEYKNTGPGAAPDKRAPYTKQLTEADAKTFTSLEYIEAAKWL
300      310      320      330      340      350

gi|512 LPPPKV
360

```

80 residues in 1 query sequences  
307888 residues in 1386 library sequences  
Scomplib [34.26]  
start: Wed Mar 11 06:07:04 2009 done: Wed Mar 11 06:07:05 2009  
Total Scan time: 0.100 Total Display time: 0.050

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

## CV127\_rc\_5894\_6169\_frame1, amino acids 1-80

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

Please cite:

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

CV127\_rc\_5894\_6169\_frame1\_1-80, 80 aa  
vs /n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

307888 residues in 1386 sequences  
Expectation\_n fit: rho(ln(x))= 3.2799+/-0.00339; mu= 11.1522+/- 0.177  
mean\_var=39.3396+/-12.489, 0's: 8 Z-trim: 8 B-trim: 50 in 1/42  
Lambda= 0.204484

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

	opt	bits	E(1386)	%_id	%_sim	alen
gi 548449 gi 548449 sp Q06478.1 PA1 ( 317)	60	24.3	1.7	0.429	0.619	21
gi 47605808 gi 47605808 sp Q7M1X5.1 ( 134)	54	22.1	3.3	0.280	0.640	25
gi 21926 gi 21926 emb CAA36063.1  u ( 295)	55	22.8	4.5	0.250	0.558	52
gi 15826578 gi 15826578 pdb 1JTI A ( 385)	55	22.9	5.4	0.250	0.614	44
gi 808969 gi 808969 emb CAA23682.1  ( 386)	55	22.9	5.4	0.250	0.614	44
gi 116174182 gi 116174182 emb CAL59 ( 323)	54	22.5	5.9	0.345	0.586	29
gi 62147665 gi 62147665 emb CAI7721 ( 340)	54	22.6	6.1	0.345	0.586	29
gi 4538529 gi 4538529 emb CAB39376. ( 81)	49	20.4	6.5	0.379	0.621	29
gi 21510 gi 21510 emb CAA31575.1  u ( 386)	54	22.6	6.6	0.261	0.565	46
gi 169500 gi 169500 gb AAA33819.1  ( 386)	54	22.6	6.6	0.261	0.565	46
gi 129293 gi 129293 sp P01012.2 OVA ( 386)	54	22.6	6.6	0.250	0.614	44
gi 7404406 gi 7404406 sp P81651.2 N ( 91)	48	20.2	8.6	0.375	0.750	24
gi 144228127 gi 144228127 gb ABO935 ( 133)	49	20.6	9	0.174	0.543	46
gi 17907758 gi 17907758 dbj BAB7944 ( 133)	49	20.6	9	0.174	0.543	46
gi 34811333 gi 34811333 pdb 1UHG D ( 385)	52	22.0	9.9	0.250	0.614	44
gi 158517845 gi 158517845 sp P15476 ( 386)	52	22.0	10	0.261	0.565	46

>>>CV127\_rc\_5894\_6169\_frame1\_1-80, 80 aa vs  
/n/na4/bioinfo/refsets/PUB\_AllergenOnlineJan2009\_V9.fasta library

>>gi|548449 gi|548449|sp|Q06478.1|PA11\_DOLMA RecName: Fu (317 aa)  
initn: 60 initl: 60 opt: 60 Z-score: 97.8 bits: 24.3 E(): 1.7  
Smith-Waterman score: 60; 42.857% identity (61.905% similar) in 21 aa overlap  
(58-78:31-51)

	10	20	30	40	50	60
CV127_	YFIAVSSFQTSVCAFLYTKRRRSSCTELKRMLDATFSQTRPKLILPIAISIFKPKLPIFL					
						:::
gi 548			RLIMFVGDPSSSNELDRFSVCPFSNDTVKMIFL			
			10	20	30	
	70	80				
CV127_	RDEDTSHLFFKKDKLNWHIV					
	::	::	::	::	::	::
gi 548	TRENRKHDFTYLTDMNRHNEFKKSIKRPVVFITHGFTSSATEKNFVAMSEALMHTGDFL					
	40	50	60	70	80	90
gi 548	IIMVDWRMAACTDEYPGLKYMFKYAAVGNTRLVGNFIAMIAKKLVEQYKVPMTNIRLVGH					
	100	110	120	130	140	150
gi 548	SLGAHISGFAGKRVQELKLGKFSEIIGLDPAGPSFKKNDCSERICETDAHYVQILHTSSN					
	160	170	180	190	200	210
gi 548	LGTERTLTGTVDFYINNGSNQPGCRYIIGETCSHTRAVKYFTECIRRECCLIGVPQSKNPQ					
	220	230	240	250	260	270



CV127\_ YFIAVSSFQTSVCAFLYTKRRRSSCTELKRMLDATFSQTRPKLILPIAISI  
 .. .: . . . . : :  
 gi|158 VKELYRGGLEPINFQTAADQARELINSWVESQTNGIIRNVLPSSVDSQTAMVLVNAI-V  
 130 140 150 160 170

60 70 80  
 CV127\_ FKPKLPIFLRDEDTSHLFFKKDKLNWHIV  
 : : . : : : . :  
 gi|158 FKGLWEKTFKDEDQAMPFRVTEQESKPVQMMYQIGLFRVASMASEKMKILELPFASGTM  
 180 190 200 210 220 230

gi|158 SMLVLLPDEVSGLEQLESIINFEKLTETWSSNVMEERKIKVYLPRMKMEEKYNLTSVLMA  
 240 250 260 270 280 290

gi|158 MGITDVFSSSANLSGISSAESLKISQAVHAAHAEINEAGTEVVGSAEAGVDAASVSEEF  
 300 310 320 330 340 350

gi|158 ADHPFLFCIKHIATNAVLFGRVCVSP  
 360 370 380

>>gi|808969 gi|808969|emb|CAA23682.1| unnamed protein pr (386 aa)  
 initn: 40 initl: 40 opt: 55 Z-score: 88.8 bits: 22.9 E(): 5.4  
 Smith-Waterman score: 55; 25.000% identity (61.364% similar) in 44 aa overlap  
 (28-71:158-200)

gi|808 MGSIGAASMEFCFDFVKELKVHHAENIFYCPIAIMSALAMVYLGAKDSTRTQINKVVR  
 10 20 30 40 50 60

gi|808 DKLPFGGDSIEAQCGTSVNVHSSLRDILNQITKPNVDVYSFSLASRLYAEERYPIPEY  
 70 80 90 100 110 120

10 20 30 40 50  
 CV127\_ YFIAVSSFQTSVCAFLYTKRRRSSCTELKRMLDATFSQTRPKLILPIAIS  
 .. .: . . . . : :  
 gi|808 CVKELYRGGLEPINFQTAADQARELINSWVESQTNGIIRNVLPSSVDSQTAMVLVNAI-  
 130 140 150 160 170

60 70 80  
 CV127\_ IFKPKLPIFLRDEDTSHLFFKKDKLNWHIV  
 : : . : : : . :  
 gi|808 VFKGLWEKTFKDEDQAMPFRVTEQESKPVQMMYQIGLFRVASMASEKMKILELPFASGT  
 180 190 200 210 220 230

gi|808 MSMLVLLPDEVSGLEQLESIINFEKLTETWSSNVMEERKIKVYLPRMKMEEKYNLTSVLM  
 240 250 260 270 280 290

gi|808 AMGITDVFSSSANLSGISSAESLKISQAVHAAHAEINEAGREVVGSAEAGVDAASVSEEF  
 300 310 320 330 340 350

gi|808 RADHPFLFCIKHIATNAVLFGRVCVSP  
 360 370 380

>>gi|116174182 gi|116174182|emb|CAL59819.1| hyaluronidas (323 aa)  
 initn: 50 initl: 50 opt: 54 Z-score: 88.1 bits: 22.5 E(): 5.9  
 Smith-Waterman score: 54; 34.483% identity (58.621% similar) in 29 aa overlap  
 (41-69:257-284)

gi|116 VITIWPKKGFSIYWNIPTHFCHNFGVYFKELKQFNIKYNSMNNFRGETISLFYDPGNFPS  
 10 20 30 40 50 60

gi|116 MVLLKNGTYEIRNEGVPQKGNLTIHLEQFTKELDEIYPKKIAGGIGVIHFNWRPIFRN  
 70 80 90 100 110 120

gi|116 VDNLKINKDISIDLVRKEHPKWDKSMIEKEASNRFETSAKIFMEKTLKLAKEIRKKTEWG  
130 140 150 160 170 180

CV127\_ YFIAVSSSFQTSVCAFLYTKRRRSS  
10 20

gi|116 YHGYPHCLSGSTDKPSFDCDALSMSENDKMSWLFNNQNVLLPSIYLNVLKPDEKIHVLQ  
190 200 210 220 230 240

CV127\_ CTELKRMLDATFSQTRPKLILPIAISIFKPKLPFLRDEDTSHLFFKKDKLNWHIV  
30 40 50 60 70 80  
:: .:: .. : ::: . :... :

gi|116 ERLKEAIRISKNLKHLPK-VLPYWWTYQDKESIFLTEADVKNFTKEILTNGADGIIWG  
250 260 270 280 290

gi|116 VSYELTDRKRCEKLKEYLMKILGP  
300 310 320

>>gi|62147665 gi|62147665|emb|CAI77218.1| hyaluronidase (340 aa)  
initn: 50 initl: 50 opt: 54 Z-score: 87.8 bits: 22.6 E(): 6.1  
Smith-Waterman score: 54; 34.483% identity (58.621% similar) in 29 aa overlap  
(41-69:257-284)

gi|621 DRTIWPKKGFSIYWNIPTHFCHNFGVYFKELKQFNIKYNMNNFRGETISLFYDPGNFPS  
10 20 30 40 50 60

gi|621 MVLLKNGTYEIRNEGVPQGNLTIHLEQFTKELDEIYPKKIAGGIGIVHFHNWRPIFRN  
70 80 90 100 110 120

gi|621 VDNLKINKDISIDLVRKEHPKWDKSMIEKEASNRFETSAKIFMEKTLKLAKEIRKKTEWG  
130 140 150 160 170 180

CV127\_ YFIAVSSSFQTSVCAFLYTKRRRSS  
10 20

gi|621 YHGYPHCLSGSTDKPSFDCDALSMSENDKMSWLFNNQNVLLPSIYLNVLKPDEKIHVLQ  
190 200 210 220 230 240

CV127\_ CTELKRMLDATFSQTRPKLILPIAISIFKPKLPFLRDEDTSHLFFKKDKLNWHIV  
30 40 50 60 70 80  
:: .:: .. : ::: . :... :

gi|621 ERLKEAIRISKNFKHLPK-VLPYWWTYQDKESIFLTEADVKNFTKEILTNGADGIIWG  
250 260 270 280 290

gi|621 VSYELTDRKRCEKLKEYLMKILGPVIAFKVTKAVKENTPLNF  
300 310 320 330 340

>>gi|4538529 gi|4538529|emb|CAB39376.1| Cop c1 allergen (81 aa)  
initn: 33 initl: 33 opt: 49 Z-score: 87.4 bits: 20.4 E(): 6.5  
Smith-Waterman score: 49; 37.931% identity (62.069% similar) in 29 aa overlap  
(41-68:22-49)

CV127\_ YFIAVSSSFQTSVCAFLYTKRRRSSCTELKRMLDATFSQTRPKLILPIAISIFKPKLPFLRDEDTSHLFFKKDKLNWHIV  
10 20 30 40 50  
: :... . :... :...  
gi|453 RFLPSSSHLNPQHLPWLHPAPVLLLPV-LPQLKPVAPHLL  
10 20 30 40

CV127\_ LRDEDTSHLFFKKDKLNWHIV  
60 70 80  
: :... :

gi|453 LLPLDTTTLHMPPLLQLQLPPLLSQGNPACSPKWLQLLVP  
50 60 70 80

>>gi|21510 gi|21510|emb|CAA31575.1| unnamed protein prod (386 aa)  
initn: 31 initl: 31 opt: 54 Z-score: 87.2 bits: 22.6 E(): 6.6  
Smith-Waterman score: 54; 26.087% identity (56.522% similar) in 46 aa overlap  
(3-46:150-195)

gi|215 MATTKSFLILFFMILATTSSTCAKLEEMVTVLSIDGGGIKGIIPAILLEFLEGQLQEVDN  
10 20 30 40 50 60

gi|215 NKDARLADYFDVIGGTSTGGLLTAMITTPNENNRPFAAAKDIVPFYFEHGPHIFNYSGSI  
70 80 90 100 110 120

CV127\_ 10 20 30  
YFIAVSSFQTSVCA-FLYTKRRRSSCTELK-RM  
.:.:.:. . . .: . . :.  
gi|215 LGPMYDGKYLQLVLEKLGQETRVHQALTEVAISSFDIKTNKPVIFTKSNLAKSPELDAKM  
130 140 150 160 170 180

CV127\_ 40 50 60 70 80  
LDATFSQTRPKLILPIAISIFKPKLPFLRDEDTSHLFFKKDKLNWHIV  
: .: . . :.  
gi|215 YDICYSTAAAPIYFPPHHFVTHTSNGARYEFNLVDGAVATVGDPALLSLSVATRLAQEDP  
190 200 210 220 230 240

gi|215 AFSSIKSLDYKQMLLLSLGTGTNSEFDKTYTAEAAKWGPLRWMLAIQQMTNAASSYMTD  
250 260 270 280 290 300

gi|215 YYISTVFQARHSQNNYLRVQENALNGTTTEMDDASEANMELLVQVGETLLKKPVSKDSPE  
310 320 330 340 350 360

gi|215 TYEEALKRFAKLLSDRKKLRANKASH  
370 380

>>gi|169500 gi|169500|gb|AAA33819.1| patatin (386 aa)  
initn: 31 initl: 31 opt: 54 Z-score: 87.2 bits: 22.6 E(): 6.6  
Smith-Waterman score: 54; 26.087% identity (56.522% similar) in 46 aa overlap  
(3-46:150-195)

gi|169 MATTKSFLILFFMILATTSSTCAKLEEMVTVLSIDGGGIKGIIPAILLEFLEGQLQEVDN  
10 20 30 40 50 60

gi|169 NKDARLADYFDVIGGTSTGGLLTAMITTPNENNRPFAAAKDIVPFYFEHGPHIFNYSGSI  
70 80 90 100 110 120

CV127\_ 10 20 30  
YFIAVSSFQTSVCA-FLYTKRRRSSCTELK-RM  
.:.:.:. . . .: . . :.  
gi|169 IGPMYDGKYLQLVLEKLGQETRVHQALTEVAISSFDIKTNKPVIFTKSNLAKSPELDAKM  
130 140 150 160 170 180

CV127\_ 40 50 60 70 80  
LDATFSQTRPKLILPIAISIFKPKLPFLRDEDTSHLFFKKDKLNWHIV  
: .: . . :.  
gi|169 YDICYSTAAAPIYFPPHYFITHTSNGDIYEFNLVDGGVATVGDPALLSLSVATRLAQEDP  
190 200 210 220 230 240

gi|169 AFSSIKSLDYKQMLLLSLGTGTNSEFDKTYTAEAAKWGPLRWMLAIQQMTNAASSYMTD  
250 260 270 280 290 300

gi|169 YYISTVFQARHSQNNYLRVQENALTGTTTEMDDASEANMELLVQVGETLLKKPVSKDSPE

```

310      320      330      340      350      360
gi|169 TYEEALKRFAKLLSDRKKLANKASY
      370      380

>>gi|129293 gi|129293|sp|P01012.2|OVAL_CHICK RecName: Fu (386 aa)
  initn: 39 initl: 39 opt: 54 Z-score: 87.2 bits: 22.6 E(): 6.6
Smith-Waterman score: 54; 25.000% identity (61.364% similar) in 44 aa overlap
(28-71:158-200)

gi|129 MGSIGAASMEFCFDVFKELKVHHANENIFYCPIAIMSALAMVYLGAKDSTRQINKVVRF
      10      20      30      40      50      60

gi|129 DKLPFGDSIEAQCGTSVNVHSSLRDILNQITKPNVYSFSLASRLYAEERYPILPPEYLQ
      70      80      90      100      110      120

      10      20      30      40      50
CV127_ YFIAVSSFQTSVCAFLYTKRRRSSCTELKRMLDATFSQTRPKLILPIAIS
      .. :. . . . :. :
gi|129 CVKELYRGGLEPINFQTAADQARELINSWVESQTNGIIRNVLPQSSVDSQTAMVLVNAI-
      130      140      150      160      170

      60      70      80
CV127_ IFKPKLPIFLRDEDTSHLFFKKDKLNWHIV
      . : . : : : . . :
gi|129 VFKGLWEKAFKDEDTQAMPFRVTEQESKPVQMMYQIGLFRVASMASEKMKILELPFASGT
      180      190      200      210      220      230

gi|129 MSMLVLLPDEVSGLQLESIINFEKLTWETSSNVMEERKIKVYLPRMKMEEKYNLTSVLM
      240      250      260      270      280      290

gi|129 AMGITDVFSSSANLSGISSAESLKISQAVHAAHAEINEAGREVVGSAEAGVDAASVSEEF
      300      310      320      330      340      350

gi|129 RADHPFLFCIKHIATNAVLFFGRCVSP
      360      370      380

>>gi|7404406 gi|7404406|sp|P81651.2|NLTP1_PRUAR RecName: (91 aa)
  initn: 42 initl: 42 opt: 48 Z-score: 85.2 bits: 20.2 E(): 8.6
Smith-Waterman score: 48; 37.500% identity (75.000% similar) in 24 aa overlap
(18-41:41-63)

      10      20      30
CV127_ YFIAVSSFQTSVCAFLYTKRRRSSCTELKRMLDATFS
      : : : : . : : : :
gi|740 ITCGQVSSSLAPCIGYVRGGGAVPPACCGIRNVNNLARTTPDRRTACNCLKQ-LSGSIS
      10      20      30      40      50

      40      50      60      70      80
CV127_ QTRPKLILPIAISIFKPKLPIFLRDEDTSHLFFKKDKLNWHIV
      . :
gi|740 GVNPNNAALPGKCGVNIPYKISASTNCATVK
      60      70      80      90

>>gi|144228127 gi|144228127|gb|ABO93594.1| BW10KD allerg (133 aa)
  initn: 45 initl: 45 opt: 49 Z-score: 84.8 bits: 20.6 E(): 9
Smith-Waterman score: 49; 17.391% identity (54.348% similar) in 46 aa overlap
(2-47:4-49)

      10      20      30      40      50
CV127_ YFIAVSSFQTSVCAFLYTKRRRSSCTELKRMLDATFSQTRPKLILPIAISIFKPKLPI
      :. . . . . : .. : : . : : . . . .

```

gi|144 MKLFLILAAASLLIVASHADSQMRSKCRKQMRMMEPQLEQCEGYMTMDMDADSMRGQEC  
10 20 30 40 50 60

60 70 80  
CV127\_ FLRDEDTSHLFFKKDKLNWHIV

gi|144 RSEESCMRGCCAMKEMDDECMCEWMKMMVQQQORGEMGEEDMRMVMRKMKQLPNKCGMGH  
70 80 90 100 110 120

gi|144 MRCHMGMGTRVYE  
130

>>gi|17907758 gi|17907758|dbj|BAB79444.1| BW8KD allergen (133 aa)  
initn: 45 initl: 45 opt: 49 Z-score: 84.8 bits: 20.6 E(): 9  
Smith-Waterman score: 49; 17.391% identity (54.348% similar) in 46 aa overlap  
(2-47:4-49)

10 20 30 40 50  
CV127\_ YFIAVSSFQTSVCAFLYTKRRRSSCTELKRMLDATFSQTRPKLILPIAISIFKPKLPI  
:. . . . . : .. :. : . :. . . . .

gi|179 MKLFLILAAASLLIVASHADSQMRSKCRKQMRMMEPQLEQCEGYMTMDMDDDSMRGREC  
10 20 30 40 50 60

60 70 80  
CV127\_ FLRDEDTSHLFFKKDKLNWHIV

gi|179 RSEESCMRGCCLAMKEMDDECMCEWMKMMVQQQORGEMGEEDMRMVMRKMKQLPNKCGMGH  
70 80 90 100 110 120

gi|179 MRCHMGIGTRDYE  
130

>>gi|34811333 gi|34811333|pdb|1UHG|D Chain D, Crystal St (385 aa)  
initn: 39 initl: 39 opt: 52 Z-score: 84.0 bits: 22.0 E(): 9.9  
Smith-Waterman score: 52; 25.000% identity (61.364% similar) in 44 aa overlap  
(28-71:157-199)

gi|348 GSIGAASMEFCFDVFKELKVHHANENIFYCPIAIMSALAMVYLGAkdSTRtQINKVVRFD  
10 20 30 40 50 60

gi|348 KLPFGFDXIEAQCGTSVNVHSSLRDILNQITKPNdVYSFSLASRLYAEERYPILPEYLQC  
70 80 90 100 110 120

10 20 30 40 50  
CV127\_ YFIAVSSFQTSVCAFLYTKRRRSSCTELKRMLDATFSQTRPKLILPIAISI  
.. :. : . . . . :. : .

gi|348 VKELYRGGLEPINFQTAADQARELINSWVESQTNGIIRNVLQPXSVDSQTAMVLVNAI-V  
130 140 150 160 170

60 70 80  
CV127\_ FKPKLPIFLRDEDTSHLFFKKDKLNWHIV  
:. . . . . :. :

gi|348 FKGLWEKAFKDEDtQAMPFRVTEQESKPVQMMYQIGLFRVASMASEKMKILELPFAXGTM  
180 190 200 210 220 230

gi|348 SMLVLLPDEVSGLEQLESIIINFEKLTEWTSsnVMEERIKIKVYLPRMKMEEKYNLTsvLMA  
240 250 260 270 280 290

gi|348 MGITDVfSSSAnLSGISSAEXLKISQAVHAAHAEINEAGREVVGXAEAGVDAASVSEEFr  
300 310 320 330 340 350

gi|348 ADHPFLFCIKHIATNAVLFfGRCVSP



```

360      370      380

>>gi|158517845 gi|158517845|sp|P15476.2|PATB1_SOLTU RecN (386 aa)
  initn: 31 initl: 31 opt: 52 Z-score: 84.0 bits: 22.0 E(): 10
Smith-Waterman score: 52; 26.087% identity (56.522% similar) in 46 aa overlap
(3-46:150-195)

gi|158 MATTKSFLILFFMILATTSSTCAKLEEMVTVLSIDGGGIKGIIPAILLEFLEGQLQEVDN
      10      20      30      40      50      60

gi|158 NKDARLADYFDVIGGTSTGGLLTAMITTPNENNRPFAAAKDIVPFYFEHGPHIFNYSGSI
      70      80      90     100     110     120

CV127_                                10      20      30
      YFIAVSSFQTSVCA-FLYTKRRRSSCTELK-RM
      .:.:.:. . . .:. . . :. . :
gi|158 FGPRYDGKYLQVLQEKLGETRVHQALTEVAISSFDIKTNKPVIFTKSNLAKSPELDAKM
      130     140     150     160     170     180

CV127_      40      50      60      70      80
      LDATFSQTRPKLILPIAISIFKPKLPIFLRDEDTSHLFFKKDKLNWHIV
      : .: . . .:
gi|158 YDICSIAAAPYFPPHFFVTHTSNGATYEFNLVDGGVATVGDPAALLSLSVATRLAQEDP
      190     200     210     220     230     240

gi|158 AFSSIKSLDYQMLLSLGTGTNSEFDKTYTAEAAKWGPLRWMLAIQQLTNAASSYMTD
      250     260     270     280     290     300

gi|158 YYISTVFQARHSQNNYLRVQENALTGTTTEMDDASEANMELLVQVGETLLKKPVSKDSPE
      310     320     330     340     350     360

gi|158 TYEEALKRFAKLLSNRKKLRANKASY
      370     380

```

```

80 residues in 1 query sequences
307888 residues in 1386 library sequences
Scomplib [34.26]
start: Wed Mar 11 06:47:41 2009 done: Wed Mar 11 06:47:41 2009
Total Scan time: 0.090 Total Display time: 0.030

```

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

## APPENDIX 2. TOXIN HOMOLOGY AND GENERAL PROTEIN HOMOLOGY SEARCH RESULTS

The BLASTP results for the search of the GenBank non-redundant database with the deduced amino acid sequences of the one ORF that produced any hits (CV127\_5319\_5915\_frame2) is shown below. All descriptions of proteins from the GenBank non-redundant database which show the most significant local homology to these ORFs were manually compared to known toxins which act on humans as listed in the United States Code of Federal Regulations ([40 CFR Part 725.421](#)) to ensure that neither shared significant homology to a known toxin. Of the 30 putative proteins encoded by ORFs created by the junctions of CV127-9 to the soybean genome, 29 did not show homology to any known protein sequence using the search parameters defined in this report; thus no results are reported for those 29 ORFs in this appendix.

### CV127\_5319\_5915\_frame2

BLASTP 2.2.18 [Mar-02-2008]

Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.* 25:3389-3402.

Reference for compositional score matrix adjustment: Altschul, Stephen F., John C. Wootton, E. Michael Gertz, Richa Agarwala, Aleksandr Morgulis, Alejandro A. Schaffer, and Yi-Kuo Yu (2005) "Protein database searches using compositionally adjusted substitution matrices", *FEBS J.* 272:5101-5109.

Query= CV127\_5319\_5915\_frame2  
(199 letters)

Database: All non-redundant GenBank CDS  
translations+PDB+SwissProt+PIR+PRF excluding environmental samples  
from WGS projects  
7,946,514 sequences; 2,739,991,458 total letters

Searching.....done

Sequences producing significant alignments:		Score (bits)	E Value
gb AAM92569.1	acetolactate synthase [Arabidopsis thaliana]	270	5e-71
ref NP_190425.1	CSR1 (CHLORSULFURON/IMIDAZOLINONE RESISTANT 1) ...	270	5e-71
gb AAK68759.1	acetolactate synthase [Arabidopsis thaliana]	270	6e-71
emb CAA35887.1	unnamed protein product [Arabidopsis thaliana]	270	6e-71
gb ABJ80681.1	acetolactate synthase catalytic subunit [Arabidop...	270	6e-71
gb AAT72502.1	AT3G48560 [Arabidopsis lyrata subsp. lyrata]	268	3e-70
pdb 1YBH A Chain A, Crystal Structure Of Arabidopsis Thaliana Ac...		263	7e-69
gb AAR07632.1	acetolactate synthase 1 [Camelina microcarpa]	261	2e-68
gb AAR07633.1	acetolactate synthase 1 [Camelina microcarpa]	261	2e-68
gb ACB12189.1	acetolactate synthase [Descurainia sophia]	261	3e-68
gb ACB12188.1	acetolactate synthase [Descurainia sophia]	261	3e-68
gb ACB12190.1	acetolactate synthase [Descurainia sophia]	260	3e-68
gb AAR06607.1	acetolactate synthase 2 [Camelina microcarpa]	260	5e-68
sp P27818.1	ILVB1_BRANA RecName: Full=Acetolactate synthase 1, c...	259	1e-67
sp P27819.1	ILVB3_BRANA RecName: Full=Acetolactate synthase 3, c...	258	1e-67
gb AAA62705.1	acetolactate synthase	258	3e-67
gb AAY46012.1	putative acetolactate synthase [Sinapis arvensis]	257	4e-67

gb	AA46005.1	putative acetolactate synthase [Sinapis arvensis]	257	4e-67
gb	AA46013.1	putative acetolactate synthase [Sinapis arvensis]	257	4e-67
gb	AA46004.1	putative acetolactate synthase [Sinapis arvensis]	257	4e-67
gb	AA45998.1	putative acetolactate synthase [Sinapis arvensis]...	257	4e-67
emb	CAC86692.1	putative acetolactate synthase [Raphanus raphani...	257	4e-67
gb	AA45995.1	putative acetolactate synthase [Sinapis arvensis]...	257	4e-67
gb	AA46011.1	putative acetolactate synthase [Sinapis arvensis]	257	4e-67
gb	AA46006.1	putative acetolactate synthase [Sinapis arvensis]...	257	4e-67
gb	AA46010.1	putative acetolactate synthase [Sinapis arvensis]	257	4e-67
gb	AA46008.1	putative acetolactate synthase [Sinapis arvensis]	257	5e-67
gb	AA46007.1	putative acetolactate synthase [Sinapis arvensis]	256	5e-67
emb	CAC86696.1	putative acetolactate synthase [Raphanus raphani...	256	6e-67
emb	CAC86697.1	putative acetolactate synthase [Raphanus raphani...	255	1e-66
emb	CAC86701.1	putative acetolactate synthase [Raphanus raphani...	254	3e-66
emb	CAC86698.1	putative acetolactate synthase [Raphanus raphani...	254	3e-66
emb	CAC86700.1	putative acetolactate synthase [Raphanus raphani...	254	3e-66
emb	CAC86695.1	putative acetolactate synthase [Raphanus raphani...	254	3e-66
emb	CAC86694.1	putative acetolactate synthase [Raphanus raphani...	254	3e-66
emb	CAC86703.1	putative acetolactate synthase [Raphanus raphani...	253	4e-66
emb	CAC86699.1	putative acetolactate synthase [Raphanus raphani...	253	7e-66
emb	CAC86702.1	putative acetolactate synthase [Raphanus raphani...	251	3e-65
emb	CAC86693.1	putative acetolactate synthase [Raphanus raphani...	247	5e-64
gb	AAK50821.1	AF363370_1 acetolactate synthase [Amaranthus powel...	227	4e-58
gb	AAK50820.1	AF363369_1 acetolactate synthase [Amaranthus retro...	227	4e-58
gb	ABS72164.1	acetolactate synthase [Amaranthus hypochondriacus]	227	5e-58
gb	ABM53021.1	acetolactate synthase [Amaranthus tuberculatus]	225	2e-57
gb	AAC69629.1	herbicide resistant acetolactate synthase precurs...	223	6e-57
gb	AAB67839.1	acetolactate synthase precursor [Amaranthus sp.]	222	2e-56
gb	EEF51778.1	acetolactate synthase, putative [Ricinus communis]	221	2e-56
gb	ACD62486.1	acetolactate synthase [Bassia scoparia]	221	2e-56
gb	ACD62506.1	acetolactate synthase [Bassia scoparia] >gi 18852...	221	2e-56
gb	ACD62501.1	acetolactate synthase [Bassia scoparia] >gi 18852...	221	2e-56
gb	ACD62517.1	acetolactate synthase [Bassia scoparia]	221	3e-56
gb	ACD62496.1	acetolactate synthase [Bassia scoparia]	221	3e-56
gb	ACD62492.1	acetolactate synthase [Bassia scoparia]	221	3e-56
gb	ACD62493.1	acetolactate synthase [Bassia scoparia]	221	3e-56
gb	ACD62515.1	acetolactate synthase [Bassia scoparia]	221	3e-56
gb	ACD62514.1	acetolactate synthase [Bassia scoparia]	221	3e-56
gb	ACD62503.1	acetolactate synthase [Bassia scoparia]	221	3e-56
gb	ACD62497.1	acetolactate synthase [Bassia scoparia]	221	3e-56
gb	ACD62495.1	acetolactate synthase [Bassia scoparia]	221	3e-56
gb	ACD62508.1	acetolactate synthase [Bassia scoparia]	221	3e-56
gb	ACD62504.1	acetolactate synthase [Bassia scoparia]	221	3e-56
gb	ACD62485.1	acetolactate synthase [Bassia scoparia]	221	3e-56
gb	ACD62490.1	acetolactate synthase [Bassia scoparia] >gi 18852...	221	3e-56
gb	ACD62484.1	acetolactate synthase [Bassia scoparia]	221	3e-56
gb	ACD62513.1	acetolactate synthase [Bassia scoparia]	221	3e-56
gb	ACD62494.1	acetolactate synthase [Bassia scoparia]	221	3e-56
gb	ACD62498.1	acetolactate synthase [Bassia scoparia] >gi 18852...	221	3e-56
gb	ACD62499.1	acetolactate synthase [Bassia scoparia]	221	3e-56
gb	ACD62500.1	acetolactate synthase [Bassia scoparia]	221	3e-56
gb	ABM53018.1	acetolactate synthase [Amaranthus tuberculatus]	221	4e-56
gb	ABM53019.1	acetolactate synthase [Amaranthus tuberculatus]	220	4e-56
gb	ACD62489.1	acetolactate synthase [Bassia scoparia] >gi 18852...	219	1e-55
ref	XP_002322262.1	predicted protein [Populus trichocarpa] >gi ...	219	1e-55
gb	ABM53020.1	acetolactate synthase [Amaranthus tuberculatus]	218	2e-55
gb	ACD62505.1	acetolactate synthase [Bassia scoparia]	218	2e-55
gb	ACD62509.1	acetolactate synthase [Bassia scoparia]	218	2e-55
gb	ACD62488.1	acetolactate synthase [Bassia scoparia]	218	2e-55
gb	ACD62487.1	acetolactate synthase [Bassia scoparia]	218	2e-55
gb	ACD62510.1	acetolactate synthase [Bassia scoparia]	218	2e-55
gb	ACD62511.1	acetolactate synthase [Bassia scoparia]	218	2e-55
emb	CAA87084.1	acetohydroxyacid synthase [Gossypium hirsutum]	218	2e-55
emb	CAA87083.1	acetohydroxyacid synthase [Gossypium hirsutum]	218	3e-55
sp	P09114.1	ILVB2_TOBAC RecName: Full=Acetolactate synthase 2, c...	217	3e-55
prf	1407140B	acetolactate synthase SuRB	217	4e-55
sp	P14874.1	ILVB2_BRANA RecName: Full=Acetolactate synthase 2, c...	217	5e-55
gb	AAT07322.1	acetohydroxyacid synthase 1 [Helianthus annuus]	216	7e-55
gb	ABR68867.1	acetohydroxyacid synthase [Solanum ptychanthum]	216	9e-55

gb AAG40280.1 AF308649_1 acetolactate synthase [Solanum ptychant...	216	9e-55
gb AAG40279.1 AF308648_1 acetolactate synthase [Solanum ptychant...	216	9e-55
gb ABR68866.1 acetohydroxyacid synthase [Solanum ptychanthum]	216	1e-54
gb ABR68865.1 acetohydroxyacid synthase [Solanum ptychanthum]	216	1e-54
gb AAB60297.1 acetolactate synthase precursor	215	2e-54
gb AAA74913.1 acetolactate synthase precursor	214	2e-54
gb AAT07324.1 acetohydroxyacid synthase 1 [Helianthus annuus]	214	2e-54
gb AAT07323.1 acetohydroxyacid synthase 1 [Helianthus annuus]	214	3e-54
gb AAT07326.1 acetohydroxyacid synthase 1 [Helianthus annuus]	214	3e-54
gb AAT07325.1 acetohydroxyacid synthase 1 [Helianthus annuus]	214	3e-54
ref XP_002163737.1  PREDICTED: similar to 2-hydroxyacyl-CoA lyas...	213	7e-54
sp P09342.1 ILVB1_TOBAC RecName: Full=Acetolactate synthase 1, c...	213	9e-54
prf  1407140A acetolactate synthase SuRA	213	1e-53
gb AAT07327.1 acetohydroxyacid synthase 2 [Helianthus annuus]	211	3e-53
gb AAT07328.1 acetohydroxyacid synthase 2 [Helianthus annuus]	211	3e-53
gb ACF17639.1 putative acetolactate synthase [Capsicum annuum]	209	8e-53
dbj BAE97675.1 acetolactate synthase [Schoenoplectus juncooides]...	209	9e-53
gb ACF47582.1 acetolactate synthase 1 [Sonchus asper]	209	1e-52
gb ACF47583.1 acetolactate synthase 1 [Sonchus asper]	209	1e-52
gb AAO53551.1 acetohydroxyacid synthase [Triticum aestivum]	207	3e-52
gb AAO53549.1 acetohydroxyacid synthase [Triticum aestivum]	207	3e-52
dbj BAF57909.1 acetolactate synthase [Sagittaria trifolia]	207	5e-52
dbj BAE97677.1 acetolactate synthase [Schoenoplectus juncooides]...	206	9e-52
gb ABF66048.1 acetolactate synthase [Oryza sativa (indica culti...	206	1e-51
gb ABF66052.1 acetolactate synthase [Oryza sativa (indica culti...	206	1e-51
gb AA14282.1 acetolactate synthase [Oryza sativa (japonica cul...	206	1e-51
gb AA14281.1 acetolactate synthase [Oryza sativa] >gi 99079641...	206	1e-51
gb ABY57316.1 acetolactate synthase [Medicago littoralis]	205	2e-51
gb ABY57315.1 acetolactate synthase [Medicago littoralis]	204	2e-51
gb ABY57317.1 acetolactate synthase [Medicago truncatula]	204	2e-51
gb ABY57318.1 acetolactate synthase [Medicago truncatula]	204	3e-51
gb AA14283.1 acetolactate synthase [Oryza sativa]	204	4e-51
dbj BAF37288.1 acetolactate synthase [Monochoria vaginalis]	204	4e-51
dbj BAE53610.1 acetolactate synthase [Monochoria vaginalis]	204	4e-51
dbj BAE53611.1 acetolactate synthase [Monochoria vaginalis]	204	4e-51
dbj BAE53612.1 acetolactate synthase [Monochoria vaginalis]	204	4e-51
dbj BAE53605.1 acetolactate synthase [Monochoria vaginalis] >gi...	204	4e-51
dbj BAE53608.1 acetolactate synthase [Monochoria vaginalis]	204	4e-51
dbj BAE53591.1 acetolactate synthase [Monochoria vaginalis] >gi...	203	6e-51
dbj BAE53588.1 acetolactate synthase [Monochoria vaginalis]	203	6e-51
dbj BAE53593.1 acetolactate synthase [Monochoria vaginalis]	203	6e-51
dbj BAE53587.1 acetolactate synthase [Monochoria vaginalis] >gi...	203	6e-51
gb AAG30931.1 AF310684_1 acetolactate synthase precursor [Lolium...	203	7e-51
gb AAO53548.1 acetohydroxyacid synthase [Triticum aestivum]	202	1e-50
gb AAO53550.1 acetohydroxyacid synthase [Triticum aestivum]	202	1e-50
gb AAC14572.1 acetohydroxyacid synthase [Hordeum vulgare]	202	1e-50
gb ABM92357.2 acetolactate synthase [Cyperus difformis]	202	1e-50
gb AAM03119.1 AF488771_1 acetolactate synthase [Bromus tectorum]	202	2e-50
gb AAL93207.1 AF487459_1 acetolactate synthase [Bromus tectorum]	202	2e-50
emb CAD24801.2 acetolactate synthase [Alopecurus myosuroides]	201	2e-50
gb EEE57060.1 hypothetical protein OsJ_06870 [Oryza sativa Japo...	201	3e-50
emb CAE18088.1 acetolactate synthase [Papaver rhoeas]	201	3e-50
dbj BAB20813.1 acetolactate synthase [Oryza sativa (japonica cu...	201	4e-50
ref NP_001046931.1  Os02g0510200 [Oryza sativa (japonica cultiva...	201	4e-50
gb ACD74789.1 acetolactate synthase [Oryza sativa Indica Group]...	200	4e-50
gb ABF66051.1 acetolactate synthase [Oryza sativa (indica culti...	200	5e-50
dbj BAE53604.1 acetolactate synthase [Monochoria vaginalis]	199	1e-49
gb AAT07329.1 acetohydroxyacid synthase 3 [Helianthus annuus]	199	1e-49
emb CAP09635.1 acetolactate-synthase-N-DnaE intein-N fusion pro...	199	2e-49
gb EAY93888.1 hypothetical protein OsI_15664 [Oryza sativa Indi...	198	2e-49
gb ABQ85869.1 acetolactate synthase [Lolium rigidum]	198	2e-49
gb ABQ85870.1 acetolactate synthase [Lolium rigidum]	198	2e-49
dbj BAE53596.1 acetolactate synthase [Monochoria vaginalis]	198	2e-49
dbj BAE53598.1 acetolactate synthase [Monochoria vaginalis] >gi...	198	2e-49
dbj BAE53602.1 acetolactate synthase [Monochoria vaginalis]	198	2e-49
dbj BAE53597.1 acetolactate synthase [Monochoria vaginalis] >gi...	198	2e-49
gb EAZ30535.1 hypothetical protein OsJ_14582 [Oryza sativa Japo...	197	3e-49
emb CAH66432.1 OSIGBa0096P03.6 [Oryza sativa (indica cultivar-g...	197	3e-49
sp Q7XKQ8.2 ILVB2_ORYSJ RecName: Full=Probable acetolactate synt...	197	3e-49

sp Q41768.1 ILVB1_MAIZE RecName: Full=Acetolactate synthase 1, c...	197	5e-49
dbj BAF37287.1  acetolactate synthase [Monochoria vaginalis]	197	5e-49
gb ABF66049.1  acetolactate synthase [Oryza sativa (indica culti...	196	1e-48
sp Q41769.1 ILVB2_MAIZE RecName: Full=Acetolactate synthase 2, c...	194	3e-48
gb ACF87819.1  unknown [Zea mays]	194	3e-48
gb EEE60896.1  hypothetical protein OsJ_14580 [Oryza sativa Japo...	194	4e-48
gb ACG44246.1  acetolactate synthase III [Zea mays]	193	6e-48
emb CA049020.1  unnamed protein product [Vitis vinifera]	193	7e-48
gb ABR16079.1  unknown [Picea sitchensis]	179	1e-43
emb CAN73760.1  hypothetical protein [Vitis vinifera]	176	1e-42
gb ABR26179.1  acetolactate synthase III [Oryza sativa (indica c...	172	1e-41
ref XP_001759950.1  predicted protein [Physcomitrella patens sub...	171	2e-41
ref XP_001758473.1  predicted protein [Physcomitrella patens sub...	169	9e-41
dbj BAE53595.1  acetolactate synthase [Monochoria vaginalis]	163	6e-39
emb CA017032.1  unnamed protein product [Vitis vinifera]	162	1e-38
emb CA045071.1  unnamed protein product [Vitis vinifera]	160	4e-38
gb AAC04854.1  acetolactate synthase [Volvox carteril]	158	3e-37
gb AAB88296.1  acetolactate synthase [Volvox carteril]	158	3e-37
ref XP_001695168.1  acetolactate synthase, large subunit [Chlamy...	152	1e-35
gb AAB88292.1  acetolactate synthase [Chlamydomonas reinhardtii]	152	1e-35
gb AAC03784.1  acetolactate synthase [Chlamydomonas reinhardtii]	152	1e-35
emb CAL58226.1  acetolactate synthase 1 (ISS) [Ostreococcus tauri]	146	9e-34
ref XP_001421626.1  predicted protein [Ostreococcus lucimarinus ...	144	3e-33
emb CA042659.1  unnamed protein product [Vitis vinifera]	135	2e-30
ref NP_682086.1  acetolactate synthase 3 catalytic subunit [Ther...	135	2e-30
ref YP_001941018.1  Acetolactate synthase large subunit or other...	134	4e-30
ref ZP_01857674.1  acetolactate synthase III (Precursor) [Planct...	129	2e-28
ref YP_001930430.1  acetolactate synthase, large subunit, biosyn...	128	3e-28
ref YP_502706.1  acetolactate synthase, large subunit, biosynthe...	127	5e-28
ref NP_926225.1  acetolactate synthase 3 catalytic subunit [Gloe...	126	1e-27
ref ZP_01089063.1  acetolactate synthase III (Precursor) [Blasto...	125	2e-27
ref ZP_02963213.1  acetolactate synthase 1 catalytic subunit [Bi...	125	2e-27
ref YP_001111655.1  acetolactate synthase, large subunit, biosyn...	125	2e-27
ref YP_002249089.1  acetolactate synthase, large subunit, biosyn...	125	2e-27
ref ZP_01731805.1  acetolactate synthase III large subunit [Cyan...	125	3e-27
gb EDX84416.1  acetolactate synthase, large subunit, biosyntheti...	124	4e-27
ref ZP_02737021.1  acetolactate synthase III [Gemmata obscuriglo...	124	5e-27
ref NP_897837.1  acetolactate synthase 3 catalytic subunit [Syne...	124	6e-27
ref YP_001805892.1  acetolactate synthase [Cyanothecae sp. ATCC 5...	123	8e-27
ref YP_001046985.1  acetolactate synthase, large subunit, biosyn...	123	9e-27
ref ZP_01079763.1  acetolactate synthase [Synechococcus sp. RS99...	122	1e-26
ref ZP_03272318.1  acetolactate synthase, large subunit, biosynt...	122	2e-26
ref ZP_01621589.1  acetolactate synthase III large subunit [Lyng...	122	2e-26
ref YP_001232446.1  acetolactate synthase, large subunit, biosyn...	122	2e-26
ref YP_172076.1  acetolactate synthase 3 catalytic subunit [Syne...	122	2e-26
ref YP_399158.1  acetolactate synthase 3 catalytic subunit [Syne...	122	2e-26
ref ZP_00514506.1  Acetolactate synthase, large subunit, biosynt...	121	3e-26
ref YP_002380209.1  acetolactate synthase, large subunit, biosyn...	121	3e-26
ref NP_441297.1  acetolactate synthase 3 catalytic subunit [Syne...	121	3e-26
ref YP_001936332.1  acetolactate synthase large subunit [Heteros...	121	4e-26
ref YP_359378.1  acetolactate synthase, large subunit, biosynthe...	120	4e-26
ref YP_001293618.1  acetoxyhydroxyacid synthetase large subunit [R...	120	4e-26
ref YP_731202.1  acetolactate synthase 3 catalytic subunit [Syne...	120	5e-26
ref NP_870771.1  acetolactate synthase III [precursor] [Rhodopir...	120	5e-26
ref YP_002371135.1  acetolactate synthase, large subunit, biosyn...	120	8e-26
ref ZP_03142166.1  acetolactate synthase, large subunit, biosynt...	120	8e-26
ref YP_001227109.1  acetolactate synthase 3 catalytic subunit [S...	120	8e-26
ref YP_323265.1  acetolactate synthase 3 catalytic subunit [Anab...	120	9e-26
ref NP_488653.1  acetolactate synthase 3 catalytic subunit [Nost...	120	9e-26
ref YP_001517397.1  acetolactate synthase 3 catalytic subunit [A...	119	1e-25
ref YP_002508832.1  acetolactate synthase, large subunit, biosyn...	119	1e-25
ref YP_002250909.1  acetolactate synthase, large subunit, biosyn...	119	1e-25
ref YP_002481296.1  acetolactate synthase, large subunit, biosyn...	119	1e-25
ref YP_377644.1  acetolactate synthase 3 catalytic subunit [Syne...	119	1e-25
ref YP_002536969.1  acetolactate synthase, large subunit, biosyn...	119	1e-25
ref NP_053940.1  acetoxyhydroxyacid synthase large subunit [Porphy...	119	1e-25
ref YP_002327489.1  acetoxyhydroxy acid synthase large subunit [Va...	119	2e-25
ref YP_001211077.1  thiamine pyrophosphate-requiring enzymes [Pe...	119	2e-25
ref NP_849033.1  acetoxyhydroxyacid synthase large subunit [Cyanid...	119	2e-25

```

ref|ZP_01472584.1| acetolactate synthase III large subunit [Syne... 119 2e-25
ref|ZP_02850743.1| acetolactate synthase, large subunit, biosynt... 118 2e-25
ref|YP_002353080.1| acetolactate synthase, large subunit, biosyn... 118 2e-25
ref|ZP_03129981.1| acetolactate synthase, large subunit, biosynt... 118 2e-25
ref|YP_842552.1| acetolactate synthase 3 catalytic subunit [Meth... 118 2e-25
ref|ZP_01085428.1| acetolactate synthase III large subunit [Syne... 118 3e-25
ref|YP_002467404.1| acetolactate synthase, large subunit, biosyn... 118 3e-25
ref|YP_001030046.1| acetolactate synthase, large subunit [Methan... 118 3e-25
ref|YP_063568.1| acetohydroxyacid synthase large subunit [Gracil... 118 3e-25
ref|NP_050806.1| acetohydroxyacid synthase large subunit [Guilla... 118 3e-25
ref|YP_475476.1| acetolactate synthase 3 catalytic subunit [Syne... 118 3e-25
ref|YP_476898.1| acetolactate synthase 3 catalytic subunit [Syne... 118 3e-25
ref|ZP_01469460.1| acetolactate synthase III large subunit [Syne... 117 4e-25
gb|AAA26594.1| acetohydroxy acid synthase (AHAS) 117 4e-25
ref|YP_684797.1| acetolactate synthase, large subunit [unculture... 117 5e-25
ref|YP_001736021.1| acetolactate synthase, large subunit, biosyn... 117 5e-25
ref|YP_754808.1| acetolactate synthase large subunit [Syntrophom... 117 5e-25
ref|YP_002620543.1| acetolactate synthase, large subunit, biosyn... 117 6e-25
ref|YP_001568611.1| acetolactate synthase, large subunit, biosyn... 117 6e-25
ref|YP_303782.1| acetolactate synthase 3 catalytic subunit [Meth... 117 6e-25
ref|YP_722490.1| acetolactate synthase 3 catalytic subunit [Tric... 117 6e-25
ref|ZP_02179310.1| acetolactate synthase large subunit [Hydrogen... 116 1e-24
ref|NP_618663.1| acetolactate synthase 3 catalytic subunit [Meth... 116 1e-24
ref|NP_789495.1| acetolactate synthase 1 catalytic subunit [Trop... 116 1e-24
ref|YP_537012.1| acetohydroxyacid synthase large subunit [Porphy... 116 1e-24
ref|ZP_03155476.1| acetolactate synthase, large subunit, biosynt... 115 2e-24

```

```

>gb|AAM92569.1| acetolactate synthase [Arabidopsis thaliana]
Length = 670

```

```

Score = 270 bits (690), Expect = 5e-71, Method: Compositional matrix adjust.
Identities = 127/128 (99%), Positives = 128/128 (100%)

```

```

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHS D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D 81
          GSYPCDDELSLHMLGMHGT VYANYAVEHS D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D
Sbjct: 336 GSYPCDDELSLHMLGMHGT VYANYAVEHS D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D 395

```

```

Query: 82 IDSAEIGKNKTPHVSVC G D V K L A L Q G M N K V L E N R A E E L K L D F G V W R N E L N V Q K Q K F P L S F 141
          IDSAEIGKNKTPHVSVC G D V K L A L Q G M N K V L E N R A E E L K L D F G V W R N E L N V Q K Q K F P L S F
Sbjct: 396 IDSAEIGKNKTPHVSVC G D V K L A L Q G M N K V L E N R A E E L K L D F G V W R N E L N V Q K Q K F P L S F 455

```

```

Query: 142 KTFGEAVP 149
          KTFGEA+P
Sbjct: 456 KTFGEAIP 463

```

```

>ref|NP_190425.1| CSR1 (CHLORSULFURON/IMIDAZOLINONE RESISTANT 1) [Arabidopsis
thaliana]
sp|P17597.1|ILVB_ARATH RecName: Full=Acetolactate synthase, chloroplastic; AltName:
Full=Acetohydroxy-acid synthase; AltName: Full=ALS;
Flags: Precursor
emb|CAB62345.1| acetolactate synthase [Arabidopsis thaliana]
gb|AAW70386.1| At3g48560 [Arabidopsis thaliana]
prf||1501386B acetolactate synthase
Length = 670

```

```

Score = 270 bits (690), Expect = 5e-71, Method: Compositional matrix adjust.
Identities = 127/128 (99%), Positives = 128/128 (100%)

```

```

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHS D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D 81
          GSYPCDDELSLHMLGMHGT VYANYAVEHS D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D
Sbjct: 336 GSYPCDDELSLHMLGMHGT VYANYAVEHS D L L L A F G V R F D D R V T G K L E A F A S R A K I V H I D 395

```

```

Query: 82 IDSAEIGKNKTPHVSVC G D V K L A L Q G M N K V L E N R A E E L K L D F G V W R N E L N V Q K Q K F P L S F 141
          IDSAEIGKNKTPHVSVC G D V K L A L Q G M N K V L E N R A E E L K L D F G V W R N E L N V Q K Q K F P L S F
Sbjct: 396 IDSAEIGKNKTPHVSVC G D V K L A L Q G M N K V L E N R A E E L K L D F G V W R N E L N V Q K Q K F P L S F 455

```

```

Query: 142 KTFGEAVP 149
          KTFGEA+P

```



Sbjct: 456 KTFGEAIP 463

>gb|AAK68759.1| acetolactate synthase [Arabidopsis thaliana]  
Length = 670

Score = 270 bits (690), Expect = 6e-71, Method: Compositional matrix adjust.  
Identities = 127/128 (99%), Positives = 128/128 (100%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 336 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 395

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF  
Sbjct: 396 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 455

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 456 KTFGEAIP 463

>emb|CAA35887.1| unnamed protein product [Arabidopsis thaliana]  
Length = 670

Score = 270 bits (689), Expect = 6e-71, Method: Compositional matrix adjust.  
Identities = 127/128 (99%), Positives = 128/128 (100%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 336 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 395

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF  
Sbjct: 396 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 455

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 456 KTFGEAIP 463

>gb|ABJ80681.1| acetolactate synthase catalytic subunit [Arabidopsis thaliana]  
Length = 670

Score = 270 bits (689), Expect = 6e-71, Method: Compositional matrix adjust.  
Identities = 127/128 (99%), Positives = 128/128 (100%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 336 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 395

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF  
Sbjct: 396 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 455

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 456 KTFGEAIP 463

>gb|AAT72502.1| AT3G48560 [Arabidopsis lyrata subsp. lyrata]  
Length = 344

Score = 268 bits (684), Expect = 3e-70, Method: Compositional matrix adjust.  
Identities = 127/128 (99%), Positives = 128/128 (100%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 185 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 244

Query: 82 IDSAEIGKNKTPHVSVCADVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPPLSF 141  
 IDSAEIGKNKTPHVSVCADVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPPLSF  
 Sbjct: 245 IDSAEIGKNKTPHVSVCADVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPPLSF 304

Query: 142 KTFGEAVP 149  
 KTFGEA+P  
 Sbjct: 305 KTFGEAIP 312

>pdb|1YBH|A Chain A, Crystal Structure Of Arabidopsis Thaliana Acetohydroxyacid  
 Synthase In Complex With A Sulfonyleurea Herbicide  
 Chlorimuron Ethyl  
 pdb|1YHY|A Chain A, Crystal Structure Of Arabidopsis Thaliana Acetohydroxyacid  
 Synthase In Complex With A Sulfonyleurea Herbicide,  
 Metsulfuron Methyl  
 pdb|1YHZ|A Chain A, Crystal Structure Of Arabidopsis Thaliana Acetohydroxyacid  
 Synthase In Complex With A Sulfonyleurea Herbicide,  
 Chlorsulfuron  
 pdb|1YI0|A Chain A, Crystal Structure Of Arabidopsis Thaliana Acetohydroxyacid  
 Synthase In Complex With A Sulfonyleurea Herbicide,  
 Sulfometuron Methyl  
 pdb|1YI1|A Chain A, Crystal Structure Of Arabidopsis Thaliana Acetohydroxyacid  
 Synthase In Complex With A Sulfonyleurea Herbicide,  
 Tribenuron Methyl  
 pdb|1Z8N|A Chain A, Crystal Structure Of Arabidopsis Thaliana Acetohydroxyacid  
 Synthase In Complex With An Imidazolinone Herbicide,  
 Imazaquin  
 Length = 590

Score = 263 bits (671), Expect = 7e-69, Method: Compositional matrix adjust.  
 Identities = 126/128 (98%), Positives = 127/128 (99%)

Query: 22 GSYPCDDELSLHMLGMHGTYYANYAVEHSDLLAFGVRFDRTVGKLEAFASRAKIVHID 81  
 GSYPCDDELSLHMLGMHGTYYANYAVEHSDLLAFGVRFDRTVGKLEAFASRAKIVHID  
 Sbjct: 251 GSYPCDDELSLHMLGMHGTYYANYAVEHSDLLAFGVRFDRTVGKLEAFASRAKIVHID 310

Query: 82 IDSAEIGKNKTPHVSVCADVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPPLSF 141  
 IDSAEIGKNKTPHVSVCADVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPPLSF  
 Sbjct: 311 IDSAEIGKNKTPHVSVCADVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPPLSF 370

Query: 142 KTFGEAVP 149  
 KTFGEA+P  
 Sbjct: 371 KTFGEAIP 378

>gb|AAR07632.1| acetolactate synthase 1 [Camelina microcarpa]  
 Length = 668

Score = 261 bits (668), Expect = 2e-68, Method: Compositional matrix adjust.  
 Identities = 123/128 (96%), Positives = 127/128 (99%)

Query: 22 GSYPCDDELSLHMLGMHGTYYANYAVEHSDLLAFGVRFDRTVGKLEAFASRAKIVHID 81  
 GSYPCDDELSLHMLGMHGTYYANYAVEHSDLLAFGVRFDRTVGKLEAFASRAKIVHID  
 Sbjct: 334 GSYPCDDELSLHMLGMHGTYYANYAVEHSDLLAFGVRFDRTVGKLEAFASRAKIVHID 393

Query: 82 IDSAEIGKNKTPHVSVCADVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPPLSF 141  
 IDSAEIGKNKTPHVSVCADVKLALQGMNKVLENRAEELKLDGFGVWRNELNVQKQKFPPLSF  
 Sbjct: 394 IDSAEIGKNKTPHVSVCADVKLALQGMNKVLENRAEELKLDGFGVWRSELNEQKQKFPPLSF 453

Query: 142 KTFGEAVP 149  
 KTFGEA+P  
 Sbjct: 454 KTFGEAIP 461

>gb|AAR07633.1| acetolactate synthase 1 [Camelina microcarpa]  
 Length = 668

Score = 261 bits (667), Expect = 2e-68, Method: Compositional matrix adjust.



Identities = 123/128 (96%), Positives = 127/128 (99%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G+YPCDDELSLHMLGMHGT VYANY+VEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 334 GAYPCDDELSLHMLGMHGT VYANYSV EHSDDLAFGVRFDDRVTGKLEAFASRAKIVHID 393

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWR+ELN QKQKFP LSF  
Sbjct: 394 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRSELNEQKQKFP LSF 453

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 454 KTFGEAIP 461

>gb|ACB12189.1| acetolactate synthase [Descurainia sophia]  
Length = 566

Score = 261 bits (666), Expect = 3e-68, Method: Compositional matrix adjust.  
Identities = 124/128 (96%), Positives = 126/128 (98%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
GSYPCDDELSLHMLGMHGT VYANY+VEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 243 GSYPCDDELSLHMLGMHGT VYANYSV EHSDDLAFGVRFDDRVTGKLEAFASRAKIVHID 302

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLE RAEELKLD FGVWRNELN QKQKFP LSF  
Sbjct: 303 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLEKRAEELKLD FGVWRNELNEQKQKFP LSF 362

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 363 KTFGEAIP 370

>gb|ACB12188.1| acetolactate synthase [Descurainia sophia]  
Length = 566

Score = 261 bits (666), Expect = 3e-68, Method: Compositional matrix adjust.  
Identities = 124/128 (96%), Positives = 126/128 (98%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
GSYPCDDELSLHMLGMHGT VYANY+VEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 243 GSYPCDDELSLHMLGMHGT VYANYSV EHSDDLAFGVRFDDRVTGKLEAFASRAKIVHID 302

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLE RAEELKLD FGVWRNELN QKQKFP LSF  
Sbjct: 303 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLEKRAEELKLD FGVWRNELNEQKQKFP LSF 362

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 363 KTFGEAIP 370

>gb|ACB12190.1| acetolactate synthase [Descurainia sophia]  
Length = 566

Score = 260 bits (665), Expect = 3e-68, Method: Compositional matrix adjust.  
Identities = 124/128 (96%), Positives = 126/128 (98%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
GSYPCDDELSLHMLGMHGT VYANY+VEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 243 GSYPCDDELSLHMLGMHGT VYANYSV EHSDDLAFGVRFDDRVTGKLEAFASRAKIVHID 302

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLE RAEELKLD FGVWRNELN QKQKFP LSF  
Sbjct: 303 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLEKRAEELKLD FGVWRNELNEQKQKFP LSF 362

Query: 142 KTFGEAVP 149  
KTFGEA+P

Sbjct: 363 KTFGEAIP 370

>gb|AAR06607.1| acetolactate synthase 2 [Camelina microcarpa]  
Length = 665

Score = 260 bits (664), Expect = 5e-68, Method: Compositional matrix adjust.  
Identities = 122/128 (95%), Positives = 126/128 (98%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 81  
G+YPCDDELSLHMLGMHGT VYANY+VEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID  
Sbjct: 331 GAYPCDDELSLHMLGMHGT VYANYSV EHSDDL LA FGVRFD DRTGKLEAFASRAKIVHID 390

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENR EELKLD FGVWR+ELN QKQKFPLSF  
Sbjct: 391 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRGEELKLD FGVWRSELNEQKQKFPLSF 450

Query: 142 KTFGEAVP 149  
KTFGEA+P

Sbjct: 451 KTFGEAIP 458

>sp|P27818.1|ILVB1\_BRANA RecName: Full=Acetolactate synthase 1, chloroplastic; AltName:  
Full=Acetolactate synthase I; AltName:  
Full=Acetohydroxy-acid synthase I; AltName: Full=ALS I;  
Flags: Precursor  
emb|CAA77613.1| actohydroxyacid synthase I [Brassica napus]  
Length = 655

Score = 259 bits (661), Expect = 1e-67, Method: Compositional matrix adjust.  
Identities = 122/128 (95%), Positives = 126/128 (98%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID  
Sbjct: 321 GSYPCNDELSLQMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 380

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWR+EL+ QKQKFPLSF  
Sbjct: 381 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRSELSEQKQKFPLSF 440

Query: 142 KTFGEAVP 149  
KTFGEA+P

Sbjct: 441 KTFGEAIP 448

>sp|P27819.1|ILVB3\_BRANA RecName: Full=Acetolactate synthase 3, chloroplastic; AltName:  
Full=Acetolactate synthase III; AltName:  
Full=Acetohydroxy-acid synthase III; AltName: Full=ALS  
III; Flags: Precursor  
emb|CAA77615.1| acetohydroxyacid synthase III [Brassica napus]  
Length = 652

Score = 258 bits (660), Expect = 1e-67, Method: Compositional matrix adjust.  
Identities = 122/128 (95%), Positives = 126/128 (98%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID  
Sbjct: 318 GSYPCNDELSLQMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 377

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWR+EL+ QKQKFPLSF  
Sbjct: 378 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRSELSEQKQKFPLSF 437

Query: 142 KTFGEAVP 149  
KTFGEA+P

Sbjct: 438 KTFGEAIP 445

>gb|AAA62705.1| acetolactate synthase

Length = 599

Score = 258 bits (658), Expect = 3e-67, Method: Compositional matrix adjust.  
Identities = 122/128 (95%), Positives = 126/128 (98%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 265 GSYPCNDELSLQMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 324

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWR+EL+ QKQKFPLSF  
Sbjct: 325 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRSELSEQKQKFPLSF 384

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 385 KTFGEAIP 392

>gb|AA46012.1| putative acetolactate synthase [Sinapis arvensis]  
Length = 577

Score = 257 bits (657), Expect = 4e-67, Method: Compositional matrix adjust.  
Identities = 122/128 (95%), Positives = 126/128 (98%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 246 GSYPCNDELSLQMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 305

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWR+EL+ QKQKFPLSF  
Sbjct: 306 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRSELSEQKQKFPLSF 365

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 366 KTFGEAIP 373

>gb|AA46005.1| putative acetolactate synthase [Sinapis arvensis]  
Length = 573

Score = 257 bits (657), Expect = 4e-67, Method: Compositional matrix adjust.  
Identities = 122/128 (95%), Positives = 126/128 (98%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 242 GSYPCNDELSLQMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 301

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWR+EL+ QKQKFPLSF  
Sbjct: 302 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRSELSEQKQKFPLSF 361

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 362 KTFGEAIP 369

>gb|AA46013.1| putative acetolactate synthase [Sinapis arvensis]  
Length = 575

Score = 257 bits (657), Expect = 4e-67, Method: Compositional matrix adjust.  
Identities = 122/128 (95%), Positives = 126/128 (98%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 246 GSYPCNDELSLQMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 305

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWR+EL+ QKQKFPLSF  
Sbjct: 306 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRSELSEQKQKFPLSF 365

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 366 KTFGEAIP 373

>gb|AA46004.1| putative acetolactate synthase [Sinapis arvensis]  
Length = 576

Score = 257 bits (656), Expect = 4e-67, Method: Compositional matrix adjust.  
Identities = 122/128 (95%), Positives = 126/128 (98%)

Query: 22 GSYPCDELSTHMLGMHGTYYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGTYYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 246 GSYPCDELSTHMLGMHGTYYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 305

Query: 82 IDSAEIGKNKTPHVSVCADVKLALQGMNKVLENRAEELKLDGFWRNELNVQKQKPLSF 141  
IDSAEIGKNKTPHVSVCADVKLALQGMNKVLENRAEELKLDGFWR+EL+ QKQKPLSF  
Sbjct: 306 IDSAEIGKNKTPHVSVCADVKLALQGMNKVLENRAEELKLDGFWRSELSEKQKQKPLSF 365

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 366 KTFGEAIP 373

>gb|AA45998.1| putative acetolactate synthase [Sinapis arvensis]  
gb|AA45999.1| putative acetolactate synthase [Sinapis arvensis]  
Length = 574

Score = 257 bits (656), Expect = 4e-67, Method: Compositional matrix adjust.  
Identities = 122/128 (95%), Positives = 126/128 (98%)

Query: 22 GSYPCDELSTHMLGMHGTYYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGTYYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 246 GSYPCDELSTHMLGMHGTYYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 305

Query: 82 IDSAEIGKNKTPHVSVCADVKLALQGMNKVLENRAEELKLDGFWRNELNVQKQKPLSF 141  
IDSAEIGKNKTPHVSVCADVKLALQGMNKVLENRAEELKLDGFWR+EL+ QKQKPLSF  
Sbjct: 306 IDSAEIGKNKTPHVSVCADVKLALQGMNKVLENRAEELKLDGFWRSELSEKQKQKPLSF 365

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 366 KTFGEAIP 373

>emb|CAC86692.1| putative acetolactate synthase [Raphanus raphanistrum]  
Length = 585

Score = 257 bits (656), Expect = 4e-67, Method: Compositional matrix adjust.  
Identities = 121/128 (94%), Positives = 126/128 (98%)

Query: 22 GSYPCDELSTHMLGMHGTYYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGTYYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 251 GSYPCDELSTHMLGMHGTYYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 310

Query: 82 IDSAEIGKNKTPHVSVCADVKLALQGMNKVLENRAEELKLDGFWRNELNVQKQKPLSF 141  
IDSAEIGKNKTPHVSVCADVKLALQGMNK+LENRAEELKLDGFWR+EL+ QKQKPLSF  
Sbjct: 311 IDSAEIGKNKTPHVSVCADVKLALQGMNKILENRAEELKLDGFWRSELSEKQKQKPLSF 370

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 371 KTFGEAIP 378

>gb|AA45995.1| putative acetolactate synthase [Sinapis arvensis]  
gb|AA45996.1| putative acetolactate synthase [Sinapis arvensis]  
gb|AA45997.1| putative acetolactate synthase [Sinapis arvensis]  
gb|AA46000.1| putative acetolactate synthase [Sinapis arvensis]  
gb|AA46001.1| putative acetolactate synthase [Sinapis arvensis]

gb|AA46002.1| putative acetolactate synthase [Sinapis arvensis]  
gb|AA46003.1| putative acetolactate synthase [Sinapis arvensis]  
Length = 574

Score = 257 bits (656), Expect = 4e-67, Method: Compositional matrix adjust.  
Identities = 122/128 (95%), Positives = 126/128 (98%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHID  
Sbjct: 246 GSYPCNDELSLQMLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHID 305

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWR+EL+ QKQKFP LSF  
Sbjct: 306 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRSELSEQKQKFP LSF 365

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 366 KTFGEAIP 373

>gb|AA46011.1| putative acetolactate synthase [Sinapis arvensis]  
Length = 577

Score = 257 bits (656), Expect = 4e-67, Method: Compositional matrix adjust.  
Identities = 122/128 (95%), Positives = 126/128 (98%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHID  
Sbjct: 246 GSYPCNDELSLQMLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHID 305

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWR+EL+ QKQKFP LSF  
Sbjct: 306 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRSELSEQKQKFP LSF 365

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 366 KTFGEAIP 373

>gb|AA46006.1| putative acetolactate synthase [Sinapis arvensis]  
gb|AA46009.1| putative acetolactate synthase [Sinapis arvensis]  
Length = 567

Score = 257 bits (656), Expect = 4e-67, Method: Compositional matrix adjust.  
Identities = 122/128 (95%), Positives = 126/128 (98%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHID  
Sbjct: 236 GSYPCNDELSLQMLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHID 295

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWR+EL+ QKQKFP LSF  
Sbjct: 296 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRSELSEQKQKFP LSF 355

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 356 KTFGEAIP 363

>gb|AA46010.1| putative acetolactate synthase [Sinapis arvensis]  
Length = 576

Score = 257 bits (656), Expect = 4e-67, Method: Compositional matrix adjust.  
Identities = 122/128 (95%), Positives = 126/128 (98%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHID  
Sbjct: 245 GSYPCNDELSLQMLGMHGT VYANYAVEHSDLLLA FGVRFD DRVTGKLEAFASRAKIVHID 304

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWR+EL+ QKQKFPLSF  
Sbjct: 305 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRSELSEQKQKFPLSF 364

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 365 KTFGEAIP 372

>gb|AA46008.1| putative acetolactate synthase [Sinapis arvensis]  
Length = 575

Score = 257 bits (656), Expect = 5e-67, Method: Compositional matrix adjust.  
Identities = 122/128 (95%), Positives = 126/128 (98%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHID  
Sbjct: 244 GSYPCNDELSLQMLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHID 303

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWR+EL+ QKQKFPLSF  
Sbjct: 304 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRSELSEQKQKFPLSF 363

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 364 KTFGEAIP 371

>gb|AA46007.1| putative acetolactate synthase [Sinapis arvensis]  
Length = 574

Score = 256 bits (655), Expect = 5e-67, Method: Compositional matrix adjust.  
Identities = 122/128 (95%), Positives = 126/128 (98%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHID  
Sbjct: 244 GSYPCNDELSLQMLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHID 303

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWR+EL+ QKQKFPLSF  
Sbjct: 304 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRSELSEQKQKFPLSF 363

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 364 KTFGEAIP 371

>emb|CAC86696.1| putative acetolactate synthase [Raphanus raphanistrum]  
Length = 585

Score = 256 bits (655), Expect = 6e-67, Method: Compositional matrix adjust.  
Identities = 121/128 (94%), Positives = 125/128 (97%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHID  
Sbjct: 251 GSYPCNDELSLQMLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHID 310

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNK+LENRAEELKLD FGVWR+EL+ QKQKFPLSF  
Sbjct: 311 IDSAEIGKNKTPHVSVC GDVKLALQGMNKILENRAEELKLD FGVWRSELSEQKQKFPLSF 370

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 371 KTFGEAIP 378

>emb|CAC86697.1| putative acetolactate synthase [Raphanus raphanistrum]  
Length = 409

Score = 255 bits (652), Expect = 1e-66, Method: Compositional matrix adjust.  
Identities = 121/128 (94%), Positives = 125/128 (97%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 75 GSYPCNDELSLQMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 134

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FG VWRNENLVQKQKFPLSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNK+LENRAEELKLD FG VWR+EL QKQKFPLSF  
Sbjct: 135 IDSAEIGKNKTPHVSVC GDVKLALQGMNKILENRAEELKLD FG VWRSELKEQKQKFPLSF 194

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 195 KTFGEAIP 202

>emb|CAC86701.1| putative acetolactate synthase [Raphanus raphanistrum]  
Length = 585

Score = 254 bits (649), Expect = 3e-66, Method: Compositional matrix adjust.  
Identities = 119/128 (92%), Positives = 126/128 (98%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGT VYANY+VEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 251 GSYPCNDELSLQMLGMHGT VYANY+VEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 310

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FG VWRNENLVQKQKFPLSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMN++LENRAEELKLD FG VWR+EL+ QKQKFPLSF  
Sbjct: 311 IDSAEIGKNKTPHVSVC GDVKLALQGMNEILENRAEELKLD FG VWRSELSEQKQKFPLSF 370

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 371 KTFGEAIP 378

>emb|CAC86698.1| putative acetolactate synthase [Raphanus raphanistrum]  
Length = 585

Score = 254 bits (649), Expect = 3e-66, Method: Compositional matrix adjust.  
Identities = 119/128 (92%), Positives = 126/128 (98%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGT VYANY+VEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 251 GSYPCNDELSLQMLGMHGT VYANY+VEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 310

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FG VWRNENLVQKQKFPLSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMN++LENRAEELKLD FG VWR+EL+ QKQKFPLSF  
Sbjct: 311 IDSAEIGKNKTPHVSVC GDVKLALQGMNEILENRAEELKLD FG VWRSELSEQKQKFPLSF 370

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 371 KTFGEAIP 378

>emb|CAC86700.1| putative acetolactate synthase [Raphanus raphanistrum]  
Length = 585

Score = 254 bits (649), Expect = 3e-66, Method: Compositional matrix adjust.  
Identities = 119/128 (92%), Positives = 126/128 (98%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGT VYANY+VEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 251 GSYPCNDELSLQMLGMHGT VYANY+VEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 310

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FG VWRNENLVQKQKFPLSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMN++LENRAEELKLD FG VWR+EL+ QKQKFPLSF  
Sbjct: 311 IDSAEIGKNKTPHVSVC GDVKLALQGMNEILENRAEELKLD FG VWRSELSEQKQKFPLSF 370

Query: 142 KTFGEAVP 149

KTFGEA+P  
Sbjct: 371 KTFGEAIP 378

>emb|CAC86695.1| putative acetolactate synthase [Raphanus raphanistrum]  
Length = 585

Score = 254 bits (649), Expect = 3e-66, Method: Compositional matrix adjust.  
Identities = 119/128 (92%), Positives = 125/128 (97%)

Query: 22 GSYPCDDELSLHMLGMHGTYYANYAVEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGTYYANY+VEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 251 GSYPCNDELSLQMLGMHGTYYANYVSVEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHID 310

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNENLVQKQKFPLSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNK+LENR EELKLD FGVWR+EL+ QKQKFPLSF  
Sbjct: 311 IDSAEIGKNKTPHVSVC GDVKLALQGMNKILENRGEELKLD FGVWRSELSEQKQKFPLSF 370

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 371 KTFGEAIP 378

>emb|CAC86694.1| putative acetolactate synthase [Raphanus raphanistrum]  
Length = 585

Score = 254 bits (649), Expect = 3e-66, Method: Compositional matrix adjust.  
Identities = 119/128 (92%), Positives = 125/128 (97%)

Query: 22 GSYPCDDELSLHMLGMHGTYYANYAVEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGTYYANY+VEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 251 GSYPCNDELSLQMLGMHGTYYANYVSVEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHID 310

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNENLVQKQKFPLSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNK+LENR EELKLD FGVWR+EL+ QKQKFPLSF  
Sbjct: 311 IDSAEIGKNKTPHVSVC GDVKLALQGMNKILENRGEELKLD FGVWRSELSEQKQKFPLSF 370

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 371 KTFGEAIP 378

>emb|CAC86703.1| putative acetolactate synthase [Raphanus raphanistrum]  
Length = 585

Score = 253 bits (647), Expect = 4e-66, Method: Compositional matrix adjust.  
Identities = 119/128 (92%), Positives = 124/128 (96%)

Query: 22 GSYPCDDELSLHMLGMHGTYYANYAVEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
GSYPC+DE SL MLGMHGTYYANY+VEHSD LLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 251 GSYPCNDEXSLQMLGMHGTYYANYVSVEHSDXLLA FGVRFDDRVTKLEAFASRAKIVHID 310

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNENLVQKQKFPLSF 141  
IDSAEIGKNKTPHVSVC GDVKLALQGMNK+LENRAEELKLD FGVWR+ELN QKQKFPLSF  
Sbjct: 311 IDSAEIGKNKTPHVSVC GDVKLALQGMNKILENRAEELKLD FGVWRSELNEQKQKFPLSF 370

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 371 KTFGEAIP 378

>emb|CAC86699.1| putative acetolactate synthase [Raphanus raphanistrum]  
Length = 585

Score = 253 bits (646), Expect = 7e-66, Method: Compositional matrix adjust.  
Identities = 119/128 (92%), Positives = 125/128 (97%)

Query: 22 GSYPCDDELSLHMLGMHGTYYANYAVEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
GSYPC+DELSL MLGMHGTYYANY+VEHSDLLLAFGVRFDDRVTKLEA ASRAKIVHID



Sbjct: 251 GSYPCNDELSLQMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 310

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141

IDSAEIGKNKTPHVSVC GDVKLALQGMNK+LENRAEELKLD FGVWR+EL+ QKQKFP LSF

Sbjct: 311 IDSAEIGKNKTPHVSVC GDVKLALQGMNKILENRAEELKLD FGVWRSELSEQKQKFP LSF 370

Query: 142 KTFGEAVP 149

KTFGEA+P

Sbjct: 371 KTFGEAIP 378

>emb|CAC86702.1| putative acetolactate synthase [Raphanus raphanistrum]  
Length = 585

Score = 251 bits (640), Expect = 3e-65, Method: Compositional matrix adjust.  
Identities = 118/128 (92%), Positives = 124/128 (96%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 81

GSYPC+DE SL MLGMHGT VYANY+VEHSDLLLA FGVRFD DRTGKLEA ASRAKIVHID

Sbjct: 251 GSYPCNDEXSLQMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 310

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141

IDSAEIGKNKTPHVSVC GDVKLALQGMNK+LENRAEELKLD FGVWR+EL+ QKQKFP LSF

Sbjct: 311 IDSAEIGKNKTPHVSVC GDVKLALQGMNKILENRAEELKLD FGVWRSELSEQKQKFP LSF 370

Query: 142 KTFGEAVP 149

KTFGEA+P

Sbjct: 371 KTFGEAIP 378

>emb|CAC86693.1| putative acetolactate synthase [Raphanus raphanistrum]  
Length = 585

Score = 247 bits (630), Expect = 5e-64, Method: Compositional matrix adjust.  
Identities = 116/128 (90%), Positives = 123/128 (96%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 81

GS PC+DE SL MLGMHGT VYANY+VEHSDLLLA FGVRFD DRTGKLEAFASRAKIVH D

Sbjct: 251 GSXPCNDEXSLQMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHXD 310

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141

IDSAEIGKNKTPHVSVC GDVKLALQGMN+LENRAEELKLD FGVWR+EL+ QKQKFP LSF

Sbjct: 311 IDSAEIGKNKTPHVSVC GDVKLALQGMNEILENRAEELKLD FGVWRSELSEQKQKFP LSF 370

Query: 142 KTFGEAVP 149

KTFGEA+P

Sbjct: 371 KTFGEAIP 378

>gb|AAK50821.1|AF363370\_1 acetolactate synthase [Amaranthus powellii]  
Length = 669

Score = 227 bits (578), Expect = 4e-58, Method: Compositional matrix adjust.  
Identities = 104/128 (81%), Positives = 119/128 (92%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 81

G++PC D+LSLHMLGMHGT VYANYAV+ +DLLLA FGVRFD DRTGKLEAFASRAKIVHID

Sbjct: 335 GAFPCDDLSLHMLGMHGT VYANYAVDKADLLLA FGVRFD DRTGKLEAFASRAKIVHID 394

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141

IDSAEIGKNK PHVS+CGDVK+ALQG+NK+LE+R ++KLDF WR ELN QK+KFP LSF

Sbjct: 395 IDSAEIGKNKQPHVSICGDVKVALQGLNKILESRKGKVKLDFSNWREELNEQKKKFP LSF 454

Query: 142 KTFGEAVP 149

KTFG+A+P

Sbjct: 455 KTFGDAIP 462

>gb|AAK50820.1|AF363369\_1 acetolactate synthase [Amaranthus retroflexus]

Length = 669

Score = 227 bits (578), Expect = 4e-58, Method: Compositional matrix adjust.  
Identities = 104/128 (81%), Positives = 119/128 (92%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
G++PC D+LSLHMLGMHGT VYANYAV+ +DLLLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 335 GAFFPCTDDL SLHMLGMHGT VYANYAVDKADLLLA FGVRFDDRVTGKLEAFASRAKIVHID 394

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+CGDVK+ALQG+NK+LE+R ++KLDF WR ELN QK+KFPLSF  
Sbjct: 395 IDSAEIGKNKQPHVSIC G DVKVALQGLNKILESRKGKVKLDFSNWREELNEQKKKFPLSF 454

Query: 142 KTFGEAVP 149  
KTFG+A+P  
Sbjct: 455 KTFGDAIP 462

>gb|ABS72164.1| acetolactate synthase [Amaranthus hypochondriacus]  
Length = 669

Score = 227 bits (578), Expect = 5e-58, Method: Compositional matrix adjust.  
Identities = 104/128 (81%), Positives = 119/128 (92%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
G++PC D+LSLHMLGMHGT VYANYAV+ +DLLLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 335 GAFFPCTDDL SLHMLGMHGT VYANYAVDKADLLLA FGVRFDDRVTGKLEAFASRAKIVHID 394

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+CGDVK+ALQG+NK+LE+R ++KLDF WR ELN QK+KFPLSF  
Sbjct: 395 IDSAEIGKNKQPHVSIC G DVKVALQGLNKILESRKGKVKLDFSNWREELNEQKKKFPLSF 454

Query: 142 KTFGEAVP 149  
KTFG+A+P  
Sbjct: 455 KTFGDAIP 462

>gb|ABM53021.1| acetolactate synthase [Amaranthus tuberculatus]  
Length = 671

Score = 225 bits (573), Expect = 2e-57, Method: Compositional matrix adjust.  
Identities = 103/128 (80%), Positives = 118/128 (92%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
G++PC D+LSL MLGMHGT VYANYAV+ +DLLLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 337 GAFFPCTDDL SLQMLGMHGT VYANYAVDKADLLLA FGVRFDDRVTGKLEAFASRAKIVHID 396

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+CGD+K+ALQG+NK+LE+R +LKLDF WR ELN QK+KFPLSF  
Sbjct: 397 IDSAEIGKNKQPHVSIC G DIKVALQGLNKILESRKGKVKLDFSNWREELNEQKKKFPLSF 456

Query: 142 KTFGEAVP 149  
KTFG+A+P  
Sbjct: 457 KTFGDAIP 464

>gb|AAC69629.1| herbicide resistant acetolactate synthase precursor [Bassia  
scoparia]  
Length = 666

Score = 223 bits (569), Expect = 6e-57, Method: Compositional matrix adjust.  
Identities = 103/128 (80%), Positives = 117/128 (91%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 332 GAYPCNDL SLHMLGMHGT VYANYAVDKADLLLA FGVRFDDRVTGKLEAFASRAKIVHID 391

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +LKL++ WR EL QK+KFPLSF

Sbjct: 392 IDSAEIGKNKQPHVSICADVKYALKGMNKILESRKGKGLNLYSSWREELGEQKKKPLSF 451

Query: 142 KTFGEAVP 149  
KTFGEA+P

Sbjct: 452 KTFGEAIP 459

>gb|AAB67839.1| acetolactate synthase precursor [Amaranthus sp.]  
Length = 665

Score = 222 bits (565), Expect = 2e-56, Method: Compositional matrix adjust.  
Identities = 101/128 (78%), Positives = 118/128 (92%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G++PC D+LSL MLGMHGT VYANYAV+ +DLLAFGVRFDDRVTGKLEAFASRAKIVHID

Sbjct: 331 GAFFCTDDLSLQMLGMHGT VYANYAVDKADLLAFGVRFDDRVTGKLEAFASRAKIVHID 390

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKPLSF 141  
IDSAEIGKNK PHVS+CG++K+ALQG+NK+LE+R +LKLD F WR ELN QK+KPLSF

Sbjct: 391 IDSAEIGKNKQPHVSICGEIKVALQGLNKILESRKGKGLKLD FSNWREELNEQKKKPLSF 450

Query: 142 KTFGEAVP 149  
K+FG+A+P

Sbjct: 451 KSFDAIP 458

>gb|EEF51778.1| acetolactate synthase, putative [Ricinus communis]  
Length = 660

Score = 221 bits (564), Expect = 2e-56, Method: Compositional matrix adjust.  
Identities = 104/128 (81%), Positives = 115/128 (89%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G++P DELSL MLGMHGT VYANY+V+ SDLLAFGVRFDDRVTGKLEAFASRAKIVHID

Sbjct: 326 GAFPVGDELSLQMLGMHGT VYANYSVDKSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 385

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKPLSF 141  
IDSAEIGKNK PHVSVC G DVKLALQGMNK+LE++ + KLD F WR ELN QK K+PLSF

Sbjct: 386 IDSAEIGKNKQPHVSVC G DVKLALQGMNKILESKGAKSKLD FKA WREELNEQKVYPLSF 445

Query: 142 KTFGEAVP 149  
KTFG+A+P

Sbjct: 446 KTFDAIP 453

>gb|ACD62486.1| acetolactate synthase [Bassia scoparia]  
Length = 666

Score = 221 bits (564), Expect = 2e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLAFGVRFDDRVTGKLEAFASRAKIVHID

Sbjct: 332 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLAFGVRFDDRVTGKLEAFASRAKIVHID 391

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KPLSF

Sbjct: 392 IDSAEIGKNKQPHVSICADVKYALKGMNKILESRKGKGLNLYSSWREELGEQKKKPLSF 451

Query: 142 KTFGEAVP 149  
KTFGEA+P

Sbjct: 452 KTFGEAIP 459

>gb|ACD62506.1| acetolactate synthase [Bassia scoparia]  
gb|ACD62507.1| acetolactate synthase [Bassia scoparia]  
gb|ACD62518.1| acetolactate synthase [Bassia scoparia]  
Length = 666

Score = 221 bits (563), Expect = 2e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 332 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLAFGVRFDDRVTGKLEAFASRAKIVHID 391

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 392 IDSAEIGKNKQPHVSICADV KYALKGMNKILESRKGKLNLYSSWREELGEQKKKFPLSF 451

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 452 KTFGEAIP 459

>gb|ACD62501.1| acetolactate synthase [Bassia scoparia]  
gb|ACD62516.1| acetolactate synthase [Bassia scoparia]  
Length = 666

Score = 221 bits (563), Expect = 2e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 332 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLAFGVRFDDRVTGKLEAFASRAKIVHID 391

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 392 IDSAEIGKNKQPHVSICADV KYALKGMNKILESRKGKLNLYSSWREELGEQKKKFPLSF 451

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 452 KTFGEAIP 459

>gb|ACD62517.1| acetolactate synthase [Bassia scoparia]  
Length = 666

Score = 221 bits (563), Expect = 3e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 332 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLAFGVRFDDRVTGKLEAFASRAKIVHID 391

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 392 IDSAEIGKNKQPHVSICADV KYALKGMNKILESRKGKLNLYSSWREELGEQKKKFPLSF 451

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 452 KTFGEAIP 459

>gb|ACD62496.1| acetolactate synthase [Bassia scoparia]  
Length = 661

Score = 221 bits (563), Expect = 3e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 327 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLAFGVRFDDRVTGKLEAFASRAKIVHID 386

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 387 IDSAEIGKNKQPHVSICADV KYALKGMNKILESRKGKLNLYSSWREELGEQKKKFPLSF 446

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 447 KTFGEAIP 454

>gb|ACD62492.1| acetolactate synthase [Bassia scoparia]  
Length = 661

Score = 221 bits (563), Expect = 3e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 327 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLLA FGVRFDDRVTGKLEAFASRAKIVHID 386

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 387 IDSAEIGKNKQPHVSI CADVKYALKGMNKILESRKGKLN LNYSSWREELGEQKKKFPLSF 446

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 447 KTFGEAIP 454

>gb|ACD62493.1| acetolactate synthase [Bassia scoparia]  
Length = 666

Score = 221 bits (563), Expect = 3e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 332 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLLA FGVRFDDRVTGKLEAFASRAKIVHID 391

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 392 IDSAEIGKNKQPHVSI CADVKYALKGMNKILESRKGKLN LNYSSWREELGEQKKKFPLSF 451

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 452 KTFGEAIP 459

>gb|ACD62515.1| acetolactate synthase [Bassia scoparia]  
Length = 666

Score = 221 bits (563), Expect = 3e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 332 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLLA FGVRFDDRVTGKLEAFASRAKIVHID 391

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 392 IDSAEIGKNKQPHVSI CADVKYALKGMNKILESRKGKLN LNYSSWREELGEQKKKFPLSF 451

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 452 KTFGEAIP 459

>gb|ACD62514.1| acetolactate synthase [Bassia scoparia]  
Length = 666

Score = 221 bits (563), Expect = 3e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81

G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 332 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLLAFGVRFDDRVTGKLEAFASRAKIVHID 391

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 392 IDSAEIGKNKQPHVSICADV KYALKGMNKILES SRKGKLNLYSSWREELGEQKKKFPLSF 451

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 452 KTFGEAIP 459

>gb|ACD62503.1| acetolactate synthase [Bassia scoparia]  
Length = 650

Score = 221 bits (563), Expect = 3e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 316 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLLAFGVRFDDRVTGKLEAFASRAKIVHID 375

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 376 IDSAEIGKNKQPHVSICADV KYALKGMNKILES SRKGKLNLYSSWREELGEQKKKFPLSF 435

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 436 KTFGEAIP 443

>gb|ACD62497.1| acetolactate synthase [Bassia scoparia]  
Length = 666

Score = 221 bits (563), Expect = 3e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 332 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLLAFGVRFDDRVTGKLEAFASRAKIVHID 391

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 392 IDSAEIGKNKQPHVSICADV KYALKGMNKILES SRKGKLNLYSSWREELGEQKKKFPLSF 451

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 452 KTFGEAIP 459

>gb|ACD62495.1| acetolactate synthase [Bassia scoparia]  
Length = 666

Score = 221 bits (563), Expect = 3e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 332 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLLAFGVRFDDRVTGKLEAFASRAKIVHID 391

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 392 IDSAEIGKNKQPHVSICADV KYALKGMNKILES SRKGKLNLYSSWREELGEQKKKFPLSF 451

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 452 KTFGEAIP 459

>gb|ACD62508.1| acetolactate synthase [Bassia scoparia]  
Length = 666

Score = 221 bits (563), Expect = 3e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFD RVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLAFGVRFD RVTGKLEAFASRAKIVHID  
Sbjct: 332 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLAFGVRFD RVTGKLEAFASRAKIVHID 391

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMN K VLENRAEELKLD FGVWRNELNVQK QKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 392 IDSAEIGKNKQPHVSICADVKYALKGMNK ILESRKGLNLN YSSWREELGEQKKKFPLSF 451

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 452 KTFGEAIP 459

>gb|ACD62504.1| acetolactate synthase [Bassia scoparia]  
Length = 647

Score = 221 bits (563), Expect = 3e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFD RVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLAFGVRFD RVTGKLEAFASRAKIVHID  
Sbjct: 313 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLAFGVRFD RVTGKLEAFASRAKIVHID 372

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMN K VLENRAEELKLD FGVWRNELNVQK QKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 373 IDSAEIGKNKQPHVSICADVKYALKGMNK ILESRKGLNLN YSSWREELGEQKKKFPLSF 432

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 433 KTFGEAIP 440

>gb|ACD62485.1| acetolactate synthase [Bassia scoparia]  
Length = 666

Score = 221 bits (563), Expect = 3e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFD RVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLAFGVRFD RVTGKLEAFASRAKIVHID  
Sbjct: 332 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLAFGVRFD RVTGKLEAFASRAKIVHID 391

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMN K VLENRAEELKLD FGVWRNELNVQK QKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 392 IDSAEIGKNKQPHVSICADVKYALKGMNK ILESRKGLNLN YSSWREELGEQKKKFPLSF 451

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 452 KTFGEAIP 459

>gb|ACD62490.1| acetolactate synthase [Bassia scoparia]  
gb|ACD62491.1| acetolactate synthase [Bassia scoparia]  
Length = 666

Score = 221 bits (563), Expect = 3e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFD RVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLAFGVRFD RVTGKLEAFASRAKIVHID  
Sbjct: 332 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLAFGVRFD RVTGKLEAFASRAKIVHID 391

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMN K VLENRAEELKLD FGVWRNELNVQK QKFPLSF 141

IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 392 IDSAEIGKNKQPHVSICADV KYALKGMNKILES SRKGKLN LNYSSWREELGEQKKKFPLSF 451

Query: 142 KTFGEAVP 149  
                  KTFGEA+P

Sbjct: 452 KTFGEAIP 459

>gb|ACD62484.1| acetolactate synthase [Bassia scoparia]  
                  Length = 666

Score = 221 bits (563), Expect = 3e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
                  G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 332 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLLA FGVRFDDRVTGKLEAFASRAKIVHID 391

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
                  IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 392 IDSAEIGKNKQPHVSICADV KYALKGMNKILES SRKGKLN LNYSSWREELGEQKKKFPLSF 451

Query: 142 KTFGEAVP 149  
                  KTFGEA+P

Sbjct: 452 KTFGEAIP 459

>gb|ACD62513.1| acetolactate synthase [Bassia scoparia]  
                  Length = 666

Score = 221 bits (563), Expect = 3e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
                  G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 332 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLLA FGVRFDDRVTGKLEAFASRAKIVHID 391

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
                  IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 392 IDSAEIGKNKQPHVSICADV KYALKGMNKILES SRKGKLN LNYSSWREELGEQKKKFPLSF 451

Query: 142 KTFGEAVP 149  
                  KTFGEA+P

Sbjct: 452 KTFGEAIP 459

>gb|ACD62494.1| acetolactate synthase [Bassia scoparia]  
                  Length = 646

Score = 221 bits (563), Expect = 3e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
                  G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 312 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLLA FGVRFDDRVTGKLEAFASRAKIVHID 371

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
                  IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 372 IDSAEIGKNKQPHVSICADV KYALKGMNKILES SRKGKLN LNYSSWREELGEQKKKFPLSF 431

Query: 142 KTFGEAVP 149  
                  KTFGEA+P

Sbjct: 432 KTFGEAIP 439

>gb|ACD62498.1| acetolactate synthase [Bassia scoparia]  
gb|ACD62512.1| acetolactate synthase [Bassia scoparia]  
                  Length = 647



Score = 221 bits (563), Expect = 3e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 313 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLLLAFGVRFDDRVTKLEAFASRAKIVHID 372

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 373 IDSAEIGKNKQPHVSICADV KYALKGMNKILES SRKGKLNLYSSWREELGEQKKKFPLSF 432

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 433 KTFGEAIP 440

>gb|ACD62499.1| acetolactate synthase [Bassia scoparia]  
Length = 646

Score = 221 bits (562), Expect = 3e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 312 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLLLAFGVRFDDRVTKLEAFASRAKIVHID 371

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 372 IDSAEIGKNKQPHVSICADV KYALKGMNKILES SRKGKLNLYSSWREELGEQKKKFPLSF 431

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 432 KTFGEAIP 439

>gb|ACD62500.1| acetolactate synthase [Bassia scoparia]  
Length = 666

Score = 221 bits (562), Expect = 3e-56, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 332 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLLLAFGVRFDDRVTKLEAFASRAKIVHID 391

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 392 IDSAEIGKNKQPHVSICADV KYALKGMNKILES SRKGKLNLYSSWREELGEQKKKFPLSF 451

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 452 KTFGEAIP 459

>gb|ABM53018.1| acetolactate synthase [Amaranthus tuberculatus]  
Length = 668

Score = 221 bits (562), Expect = 4e-56, Method: Compositional matrix adjust.  
Identities = 101/128 (78%), Positives = 117/128 (91%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G++PC D+LSL MLGMHGT VYANYAV+ +DLLLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 334 GAFPC TDDLSLQMLGMHGT VYANYAVDKADLLLLAFGVRFDDRVTKLEAFASRAKIVHID 393

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+CGDVK+AL+G+N +LE+R ++KLDF WR ELN QK+KFPLSF  
Sbjct: 394 IDSAEIGKNKQPHVSICGDVKVALRGLNNILES SRKGKVLDFSNWREELNEQKKKFPLSF 453

Query: 142 KTFGEAVP 149

KTFG+A+P  
Sbjct: 454 KTFGDAIP 461

>gb|ABM53019.1| acetolactate synthase [Amaranthus tuberculatus]  
Length = 669

Score = 220 bits (561), Expect = 4e-56, Method: Compositional matrix adjust.  
Identities = 101/128 (78%), Positives = 117/128 (91%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 81  
G++PC D+LSL MLGMHGT VYANYAV+ +DLLLLAFGVRFD DRTGKLEAFASRAKIVHID  
Sbjct: 335 GAFPCDDLSLQMLGMHGT VYANYAVDKADLLLA FGVRFD DRTGKLEAFASRAKIVHID 394

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNK PHVS+CGDVK+AL+G+N +LE+R ++KLDF WR ELN QK+KFP LSF  
Sbjct: 395 IDSAEIGKNKQPHVSICGDVKVALRGLNNILESRKGKVKLDFSNWREELNEQKKKFP LSF 454

Query: 142 KTFGEAVP 149  
KTFG+A+P  
Sbjct: 455 KTFGDAIP 462

>gb|ACD62489.1| acetolactate synthase [Bassia scoparia]  
gb|ACD62502.1| acetolactate synthase [Bassia scoparia]  
Length = 666

Score = 219 bits (558), Expect = 1e-55, Method: Compositional matrix adjust.  
Identities = 101/128 (78%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLLLAFGVRFD+RVTGKLEAFASRAKIVHID  
Sbjct: 332 GAYPCNDLDSLHMLGMHGT VYANYAVDKADLLLA FGVRFD DRTGKLEAFASRAKIVHID 391

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFP LSF  
Sbjct: 392 IDSAEIGKNKQPHVSICADV KALKGMNKILESRKGKLNLYSSWREELGEQKKKFP LSF 451

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 452 KTFGEAIP 459

>ref|XP\_002322262.1| predicted protein [Populus trichocarpa]  
gb|EEF06389.1| predicted protein [Populus trichocarpa]  
Length = 650

Score = 219 bits (557), Expect = 1e-55, Method: Compositional matrix adjust.  
Identities = 104/128 (81%), Positives = 113/128 (88%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 81  
G++P DELSL MLGMHGT VYANYAV+ SDLLLLAFGVRFD DRTGKLEAFASRAKIVHID  
Sbjct: 316 GAFPVGDELSLQMLGMHGT VYANYAVDKSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 375

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNK PHVSVC GDVK+ALQGMNK+LE+R + K+DF WR ELN QK PLSF  
Sbjct: 376 IDSAEIGKNKQPHVSVC GDVKVALQGMNKILESRGAKGKMDFAWREELNEQKMNNP LSF 435

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 436 KTFGEAIP 443

>gb|ABM53020.1| acetolactate synthase [Amaranthus tuberculatus]  
Length = 669

Score = 218 bits (556), Expect = 2e-55, Method: Compositional matrix adjust.  
Identities = 100/128 (78%), Positives = 116/128 (90%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFD RVTGKLEAFASRAKIVHID 81  
G++PC D+LSL MLGMHGT VYANYAV+ +DLLAFGVRFD RVTGKLEAFASRAKIVHID  
Sbjct: 335 GAFPCDDLSLQMLGMHGT VYANYAVDKADLLAFGVRFD RVTGKLEAFASRAKIVHID 394

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNK PHVS+CGD K+AL+G+N +LE+R ++KLDF WR ELN QK+KFP LSF  
Sbjct: 395 IDSAEIGKNKQPHVSICGD XKVALRGLNNILESRKGKVKLDFSNWREELNEQKKKFP LSF 454

Query: 142 KTFGEAVP 149  
KTFG+A+P  
Sbjct: 455 KTFGDAIP 462

>gb|ACD62505.1| acetolactate synthase [Bassia scoparia]  
Length = 661

Score = 218 bits (556), Expect = 2e-55, Method: Compositional matrix adjust.  
Identities = 101/128 (78%), Positives = 115/128 (89%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFD RVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLAFGVRFD RVTGKLEAFASRAKIVHID  
Sbjct: 327 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLAFGVRFDXRV TKGLEAFASRAKIVHID 386

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFP LSF  
Sbjct: 387 IDSAEIGKNKQPHVSICADVKYALKGMNKILESRKGKLN LNYSSWREELGEQKKKFP LSF 446

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 447 KTFGEAIP 454

>gb|ACD62509.1| acetolactate synthase [Bassia scoparia]  
Length = 666

Score = 218 bits (555), Expect = 2e-55, Method: Compositional matrix adjust.  
Identities = 101/128 (78%), Positives = 115/128 (89%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFD RVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLAFGVRFD RVTGKLEAFASRAKIVHID  
Sbjct: 332 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLAFGVRFDXRV TKGLEAFASRAKIVHID 391

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFP LSF  
Sbjct: 392 IDSAEIGKNKQPHVSICADVKYALKGMNKILESRKGKLN LNYSSWREELGEQKKKFP LSF 451

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 452 KTFGEAIP 459

>gb|ACD62488.1| acetolactate synthase [Bassia scoparia]  
Length = 656

Score = 218 bits (555), Expect = 2e-55, Method: Compositional matrix adjust.  
Identities = 101/128 (78%), Positives = 115/128 (89%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFD RVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLAFGVRFD RVTGKLEAFASRAKIVHID  
Sbjct: 322 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLAFGVRFDXRV TKGLEAFASRAKIVHID 381

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFP LSF  
Sbjct: 382 IDSAEIGKNKQPHVSICADVKYALKGMNKILESRKGKLN LNYSSWREELGEQKKKFP LSF 441

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 442 KTFGEAIP 449

>gb|ACD62487.1| acetolactate synthase [Bassia scoparia]  
Length = 657

Score = 218 bits (555), Expect = 2e-55, Method: Compositional matrix adjust.  
Identities = 101/128 (78%), Positives = 115/128 (89%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD RVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLLLAFGVRFD RVTGKLEAFASRAKIVHID  
Sbjct: 323 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLLA FGVRFDXRV TGKLEAFASRAKIVHID 382

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 383 IDSAEIGKNKQPHVSICADV KYALKGMNKILESRKGKLN LNYSSWREELGEQKKKFPLSF 442

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 443 KTFGEAIP 450

>gb|ACD62510.1| acetolactate synthase [Bassia scoparia]  
Length = 666

Score = 218 bits (555), Expect = 2e-55, Method: Compositional matrix adjust.  
Identities = 101/128 (78%), Positives = 115/128 (89%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD RVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLLLAFGVRFD RVTGKLEAFASRAKIVHID  
Sbjct: 332 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLLA FGVRFDXRV TGKLEAFASRAKIVHID 391

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 392 IDSAEIGKNKQPHVSICADV KYALKGMNKILESRKGKLN LNYSSWREELGEQKKKFPLSF 451

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 452 KTFGEAIP 459

>gb|ACD62511.1| acetolactate synthase [Bassia scoparia]  
Length = 647

Score = 218 bits (555), Expect = 2e-55, Method: Compositional matrix adjust.  
Identities = 101/128 (78%), Positives = 115/128 (89%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD RVTGKLEAFASRAKIVHID 81  
G+YPC+D+LSLHMLGMHGT VYANYAV+ +DLLLLAFGVRFD RVTGKLEAFASRAKIVHID  
Sbjct: 313 GAYPCNDDLSLHMLGMHGT VYANYAVDKADLLLA FGVRFDXRV TGKLEAFASRAKIVHID 372

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+C DVK AL+GMNK+LE+R +L L++ WR EL QK+KFPLSF  
Sbjct: 373 IDSAEIGKNKQPHVSICADV KYALKGMNKILESRKGKLN LNYSSWREELGEQKKKFPLSF 432

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 433 KTFGEAIP 440

>emb|CAA87084.1| acetohydroxyacid synthase [Gossypium hirsutum]  
Length = 659

Score = 218 bits (555), Expect = 2e-55, Method: Compositional matrix adjust.  
Identities = 102/128 (79%), Positives = 113/128 (88%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD RVTGKLEAFASRAKIVHID 81  
G++P D+LSL MLGMHGT VYANYAV+ SDLLLA FGVRFD RVTGKLEAFASRAKIVHID  
Sbjct: 325 GAFPI SDDL SLQMLGMHGT VYANYAVDKS DLLLA FGVRFD RVTGKLEAFASRAKIVHID 384

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141

IDSAEIGKNK PHVSVC DVKLALQG+NK+LE + +L LD+ WR ELN QK KFPLS+  
 Sbjct: 385 IDSAEIGKNKQPHVSVCSDVKLALQGINKILETKVAKLNLDYSEWRQELNEQKLFPLSY 444

Query: 142 KTFGEAVP 149  
 KTFGEA+P

Sbjct: 445 KTFGEAIP 452

>emb|CAA87083.1| acetohydroxyacid synthase [Gossypium hirsutum]  
 Length = 659

Score = 218 bits (554), Expect = 3e-55, Method: Compositional matrix adjust.  
 Identities = 102/128 (79%), Positives = 112/128 (87%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFD RVTGKLEAFASRAKIVHID 81  
 G++P DELSL MLGMHGT VYANYAV+ SDLLAFGVRFD RVTGKLEAFASRAKIVHID  
 Sbjct: 325 GAFPISEDELSLQMLGMHGT VYANYAVDKSDLLAFGVRFD RVTGKLEAFASRAKIVHID 384

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
 IDSAEIGKNK PH+SVC DVKLALQG+NK+LE +L LD+ WR ELN QK KFPLS+  
 Sbjct: 385 IDSAEIGKNKQPHMSVCSDVKLALQGINKILETTGAKLNLDYSEWRQELNEQKLFPLSY 444

Query: 142 KTFGEAVP 149  
 KTFGEA+P

Sbjct: 445 KTFGEAIP 452

>sp|P09114.1|ILVB2\_TOBAC RecName: Full=Acetolactate synthase 2, chloroplastic; AltName:  
 Full=Acetolactate synthase II; AltName:  
 Full=Acetohydroxy-acid synthase II; AltName: Full=ALS  
 II; Flags: Precursor  
 emb|CAA30485.1| unnamed protein product [Nicotiana tabacum]  
 Length = 664

Score = 217 bits (553), Expect = 3e-55, Method: Compositional matrix adjust.  
 Identities = 100/128 (78%), Positives = 114/128 (89%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFD RVTGKLEAFASRAKIVHID 81  
 G++P DELSL MLGMHGT VYANYAV+ SDLLAFGVRFD RVTGKLEAFASRAKIVHID  
 Sbjct: 330 GAFPTGDELSLSMLGMHGT VYANYAVDSSDLLAFGVRFD RVTGKLEAFASRAKIVHID 389

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
 IDSAEIGKNK PHVS+C D+KLALQG+N +LE++ +LKLD F WR EL VQK K+PL+F  
 Sbjct: 390 IDSAEIGKNKQPHVSICADIKLALQGLNSILESKEGKLKLD FSAWRQELTVQKVYPLNF 449

Query: 142 KTFGEAVP 149  
 KTFG+A+P

Sbjct: 450 KTFGDAIP 457

>prf||1407140B acetolactate synthase SuRB  
 Length = 664

Score = 217 bits (553), Expect = 4e-55, Method: Compositional matrix adjust.  
 Identities = 100/128 (78%), Positives = 114/128 (89%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFD RVTGKLEAFASRAKIVHID 81  
 G++P DELSL MLGMHGT VYANYAV+ SDLLAFGVRFD RVTGKLEAFASRAKIVHID  
 Sbjct: 330 GAFPTGDELSLSMLGMHGT VYANYAVDSSDLLAFGVRFD RVTGKLEAFASRAKIVHID 389

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
 IDSAEIGKNK PHVS+C D+KLALQG+N +LE++ +LKLD F WR EL VQK K+PL+F  
 Sbjct: 390 IDSAEIGKNKQPHVSICADIKLALQGLNSILESKEGKLKLD FSAWRQELTVQKVYPLNF 449

Query: 142 KTFGEAVP 149  
 KTFG+A+P

Sbjct: 450 KTFGDAIP 457

```
>sp|P14874.1|ILVB2_BRANA RecName: Full=Acetolactate synthase 2, chloroplastic; AltName:
      Full=Acetolactate synthase II; AltName:
      Full=Acetohydroxy-acid synthase II; AltName: Full=ALS
      II; Flags: Precursor
emb|CAA77614.1| acetohydroxyacid synthase II [Brassica napus]
emb|CAA34680.1| unnamed protein product [Brassica napus]
      Length = 637
```

Score = 217 bits (552), Expect = 5e-55, Method: Compositional matrix adjust.  
Identities = 108/133 (81%), Positives = 115/133 (86%), Gaps = 3/133 (2%)

```
Query: 18 FQHEGSYPCCDEE-LSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAK 76
      F GSYPCDEE SL MLGMHGT VYANYAVE+SDLLAFGVRFDDRVTKLEAFASRAK
Sbjct: 308 FMGLGSYPCCDEEFSLQMLGMHGT VYANYAVEYSDLLAFGVRFDDRVTKLEAFASRAK 367
```

```
Query: 77 IVHIDIDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQK 136
      IVHIDIDS EIGKNKTPHVSVC DV+LALQGMN+VLENR + LDFG WR ELN Q+ K
Sbjct: 368 IVHIDIDSTEIGKNKTPHVSVC DVQLALQGMNEVLENRRD--VLDFGEWRCELNEQRLK 425
```

```
Query: 137 FPLSFKTFGEAVP 149
      FPL +KTFGE +P
Sbjct: 426 FPLRYKTFGEEIP 438
```

```
>gb|AAT07322.1| acetohydroxyacid synthase 1 [Helianthus annuus]
      Length = 655
```

Score = 216 bits (551), Expect = 7e-55, Method: Compositional matrix adjust.  
Identities = 100/128 (78%), Positives = 111/128 (86%)

```
Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81
      G+Y P +LSLHMLGMHGT VYANYAV+ SDLLAFGVRFDDRVTKLEAFASRAKIVHID
Sbjct: 321 GAYPASSDLSLHMLGMHGT VYANYAVDKSDLLAFGVRFDDRVTKLEAFASRAKIVHID 380
```

```
Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141
      ID AEIGKNK PHVS+CGD+K+ALQG+NK+LE + LDF WR EL+ QK KFP LSF
Sbjct: 381 IDPAEIGKNKQPHVSICGDIKVALQGLNKILEEKNSVTNLDFSTWRKELDEQKMKFP LSF 440
```

```
Query: 142 KTFGEAVP 149
      KTFGEA+P
Sbjct: 441 KTFGEAIP 448
```

```
>gb|ABR68867.1| acetohydroxyacid synthase [Solanum ptychanthum]
      Length = 554
```

Score = 216 bits (550), Expect = 9e-55, Method: Compositional matrix adjust.  
Identities = 100/128 (78%), Positives = 112/128 (87%)

```
Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81
      G++P DELSL MLGMHGT VYANYAV+ SDLLAFGVRFDDRVTKLEAFASRAKIVHID
Sbjct: 230 GAFPTGDELSLQMLGMHGT VYANYAVDSSDLLAFGVRFDDRVTKLEAFASRAKIVHID 289
```

```
Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141
      IDSAEIGKNK PH S+C D+KLALQG+N +LE + +LKLD F WR ELN QK K+PL+F
Sbjct: 290 IDSAEIGKNKLPHASICADIKLALQGLNSILEGKEGKLKLD FSAWRQELNEQKVYPLNF 349
```

```
Query: 142 KTFGEAVP 149
      KTFGEA+P
Sbjct: 350 KTFGEAIP 357
```

```
>gb|AAG40280.1|AF308649_1 acetolactate synthase [Solanum ptychanthum]
gb|AAG40281.1|AF308650_1 acetolactate synthase [Solanum ptychanthum]
      Length = 567
```

Score = 216 bits (550), Expect = 9e-55, Method: Compositional matrix adjust.  
Identities = 100/128 (78%), Positives = 112/128 (87%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHID 81  
G++P DELSL MLGMHGT VYANYAV+ SDLLAFGVRFDDRV TGKLEAFASRAKIVHID  
Sbjct: 241 GAFPTGDELSLQMLGMHGT VYANYAVDSSDLLAFGVRFDDRV TGKLEAFASRAKIVHID 300

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNK PH S+C D+KLALQG+N +LE + +LKLD F WR ELN QK K+PL+F  
Sbjct: 301 IDSAEIGKNKLPHASICADIKLALQGLNSILEGKEGKLKLD FSAWRQELNEQKVYPLNF 360

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 361 KTFGEAIP 368

>gb|AAG40279.1|AF308648\_1 acetolactate synthase [Solanum ptychanthum]  
Length = 567

Score = 216 bits (550), Expect = 9e-55, Method: Compositional matrix adjust.  
Identities = 100/128 (78%), Positives = 112/128 (87%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHID 81  
G++P DELSL MLGMHGT VYANYAV+ SDLLAFGVRFDDRV TGKLEAFASRAKIVHID  
Sbjct: 241 GAFPTGDELSLQMLGMHGT VYANYAVDSSDLLAFGVRFDDRV TGKLEAFASRAKIVHID 300

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNK PH S+C D+KLALQG+N +LE + +LKLD F WR ELN QK K+PL+F  
Sbjct: 301 IDSAEIGKNKLPHASICADIKLALQGLNSILEGKEGKLKLD FSAWRQELNEQKVYPLNF 360

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 361 KTFGEAIP 368

>gb|ABR68866.1| acetohydroxyacid synthase [Solanum ptychanthum]  
Length = 565

Score = 216 bits (549), Expect = 1e-54, Method: Compositional matrix adjust.  
Identities = 100/128 (78%), Positives = 112/128 (87%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHID 81  
G++P DELSL MLGMHGT VYANYAV+ SDLLAFGVRFDDRV TGKLEAFASRAKIVHID  
Sbjct: 241 GAFPTGDELSLQMLGMHGT VYANYAVDSSDLLAFGVRFDDRV TGKLEAFASRAKIVHID 300

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNK PH S+C D+KLALQG+N +LE + +LKLD F WR ELN QK K+PL+F  
Sbjct: 301 IDSAEIGKNKLPHASICADIKLALQGLNSILEGKEGKLKLD FSAWRQELNEQKVYPLNF 360

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 361 KTFGEAIP 368

>gb|ABR68865.1| acetohydroxyacid synthase [Solanum ptychanthum]  
Length = 565

Score = 216 bits (549), Expect = 1e-54, Method: Compositional matrix adjust.  
Identities = 100/128 (78%), Positives = 112/128 (87%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHID 81  
G++P DELSL MLGMHGT VYANYAV+ SDLLAFGVRFDDRV TGKLEAFASRAKIVHID  
Sbjct: 241 GAFPTGDELSLQMLGMHGT VYANYAVDSSDLLAFGVRFDDRV TGKLEAFASRAKIVHID 300

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNK PH S+C D+KLALQG+N +LE + +LKLD F WR ELN QK K+PL+F  
Sbjct: 301 IDSAEIGKNKLPHASICADIKLALQGLNSILEGKEGKLKLD FSAWRQELNEQKVYPLNF 360

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 361 KTFGEAIP 368

>gb|AAB60297.1| acetolactate synthase precursor  
Length = 648

Score = 215 bits (547), Expect = 2e-54, Method: Compositional matrix adjust.  
Identities = 100/128 (78%), Positives = 112/128 (87%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 81  
G+YP +LSLHMLGMHGT VYANYAV+ SDLLLA FGVRFD DRTGKLEAFASRAKIVHID  
Sbjct: 314 GAYPASSDLSLHMLGMHGT VYANYAVDKSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 373

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNK PHVS+CGD+K+ALQG+NK+LE + LDF WR EL+ QK K+PLSF  
Sbjct: 374 IDSAEIGKNKQPHVSICGDIKVALQGLNKILEVKNSVTNLDFS NWRKELDEQKVYPLSF 433

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 434 KTFGEAIP 441

>gb|AAA74913.1| acetolactate synthase precursor  
Length = 648

Score = 214 bits (546), Expect = 2e-54, Method: Compositional matrix adjust.  
Identities = 100/128 (78%), Positives = 112/128 (87%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 81  
G+YP +LSLHMLGMHGT VYANYAV+ SDLLLA FGVRFD DRTGKLEAFASRAKIVHID  
Sbjct: 314 GAYPASSDLSLHMLGMHGT VYANYAVDKSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 373

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNK PHVS+CGD+K+ALQG+NK+LE + LDF WR EL+ QK K+PLSF  
Sbjct: 374 IDSAEIGKNKQPHVSICGDIKVALQGLNKILEVKNSVTNLDFS NWRKELDEQKVYPLSF 433

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 434 KTFGEAIP 441

>gb|AAT07324.1| acetohydroxyacid synthase 1 [Helianthus annuus]  
Length = 658

Score = 214 bits (546), Expect = 2e-54, Method: Compositional matrix adjust.  
Identities = 100/128 (78%), Positives = 111/128 (86%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 81  
G+YP +LSLHMLGMHGT VYANYAV+ SDLLLA FGVRFD DRTGKLEAFASRAKIVHID  
Sbjct: 324 GAYPASSDLSLHMLGMHGT VYANYAVDKSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 383

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID AEIGKNK PHVS+CGD+K+ALQG+NK+LE + LDF WR EL+ QK KFP LSF  
Sbjct: 384 IDPAEIGKNKQPHVSICGDIKVALQGLNKILEEKNSVTNLDFS NWRKELDEQKVYPLSF 443

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 444 KTFGEAIP 451

>gb|AAT07323.1| acetohydroxyacid synthase 1 [Helianthus annuus]  
Length = 654

Score = 214 bits (546), Expect = 3e-54, Method: Compositional matrix adjust.  
Identities = 100/128 (78%), Positives = 111/128 (86%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 81  
G+YP +LSLHMLGMHGT VYANYAV+ SDLLLA FGVRFD DRTGKLEAFASRAKIVHID  
Sbjct: 320 GAYPASSDLSLHMLGMHGT VYANYAVDKSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 379

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141



ID AEIGKNK PHVS+CGD+K+ALQG+NK+LE + LDF WR EL+ QK KFPLSF  
Sbjct: 380 IDPAEIGKNKQPHVSICGDIKVALQGLNKILEEKNSVTNLDFSNNWRKELDEQKVKFPLSF 439

Query: 142 KTFGEAVP 149  
KTFGEA+P

Sbjct: 440 KTFGEAIP 447

>gb|AAT07326.1| acetohydroxyacid synthase 1 [Helianthus annuus]  
Length = 652

Score = 214 bits (546), Expect = 3e-54, Method: Compositional matrix adjust.  
Identities = 100/128 (78%), Positives = 111/128 (86%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G+YP +LSLHMLGMHGT VYANYAV+ SDLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 318 GAYPASSDLSLHMLGMHGT VYANYAVDKSDLLAFGVRFDDRVTKLEAFASRAKIVHID 377

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKNVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
ID AEIGKNK PHVS+CGD+K+ALQG+NK+LE + LDF WR EL+ QK KFPLSF  
Sbjct: 378 IDPAEIGKNKQPHVSICGDIKVALQGLNKILEEKNSVTNLDFSNNWRKELDEQKVKFPLSF 437

Query: 142 KTFGEAVP 149  
KTFGEA+P

Sbjct: 438 KTFGEAIP 445

>gb|AAT07325.1| acetohydroxyacid synthase 1 [Helianthus annuus]  
Length = 652

Score = 214 bits (546), Expect = 3e-54, Method: Compositional matrix adjust.  
Identities = 100/128 (78%), Positives = 111/128 (86%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G+YP +LSLHMLGMHGT VYANYAV+ SDLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 318 GAYPASSDLSLHMLGMHGT VYANYAVDKSDLLAFGVRFDDRVTKLEAFASRAKIVHID 377

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKNVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
ID AEIGKNK PHVS+CGD+K+ALQG+NK+LE + LDF WR EL+ QK KFPLSF  
Sbjct: 378 IDPAEIGKNKQPHVSICGDIKVALQGLNKILEEKNSVTNLDFSNNWRKELDEQKVKFPLSF 437

Query: 142 KTFGEAVP 149  
KTFGEA+P

Sbjct: 438 KTFGEAIP 445

>ref|XP\_002163737.1| PREDICTED: similar to 2-hydroxyacyl-CoA lyase 1 [Hydra  
magnipapillata]  
gb|ABN08612.1| Thiamine pyrophosphate enzyme, central region [Medicago truncatula]  
Length = 655

Score = 213 bits (542), Expect = 7e-54, Method: Compositional matrix adjust.  
Identities = 101/128 (78%), Positives = 113/128 (88%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
GSYP DE SL MLGMHGT VYANYAV+ SDLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 321 GSYPSLDENSLQMLGMHGT VYANYAVDKSDLLAFGVRFDDRVTKLEAFASRAKIVHID 380

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKNVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
ID+AEIGKNK PHVSVC GD+KLAL+G+N+LE+ E KLD FGVWR ELN QK +FP+S+  
Sbjct: 381 IDAAEIGKNKQPHVSVC GD+KLAL+G+N+LE+ E KLD FGVWR ELN QK +FP+S+ 440

Query: 142 KTFGEAVP 149  
KTF EA+P

Sbjct: 441 KTFDEAIP 448

>sp|P09342.1|ILVB1\_TOBAC RecName: Full=Acetolactate synthase 1, chloroplastic; AltName:  
Full=Acetolactate synthase I; AltName:

```
Full=Acetohydroxy-acid synthase I; AltName: Full=ALS I;
Flags: Precursor
emb|CAA30484.1| unnamed protein product [Nicotiana tabacum]
prf||1501386A acetolactate synthase
Length = 667

Score = 213 bits (541), Expect = 9e-54, Method: Compositional matrix adjust.
Identities = 99/128 (77%), Positives = 112/128 (87%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81
      G++P  DELSL MLGMHGT VYANYAV+ SDLLAFGVRFDDRVTGKLEAFASRAKIVHID
Sbjct: 333 GAFPTGDELSLHMLGMHGT VYANYAVDSSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 392

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMN K VLENRAEELKLD FGVWRNELNVQKQK FPLSF 141
      IDSAEIGKNK PHVS+C D+KLALQG+N +LE++ +LKLD F WR EL QK K PL+F
Sbjct: 393 IDSAEIGKNKQPHVSICADIKLALQGLNS ILESKEGKLKLD FSAWRQELTEQKV K HPLNF 452

Query: 142 KTFGEAVP 149
      KTFG+A+P
Sbjct: 453 KTFGDAIP 460

>prf||1407140A acetolactate synthase SuRA
Length = 667

Score = 213 bits (541), Expect = 1e-53, Method: Compositional matrix adjust.
Identities = 99/128 (77%), Positives = 112/128 (87%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81
      G++P  DELSL MLGMHGT VYANYAV+ SDLLAFGVRFDDRVTGKLEAFASRAKIVHID
Sbjct: 333 GAFPTGDELSLHMLGMHGT VYANYAVDSSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 392

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMN K VLENRAEELKLD FGVWRNELNVQKQK FPLSF 141
      IDSAEIGKNK PHVS+C D+KLALQG+N +LE++ +LKLD F WR EL QK K PL+F
Sbjct: 393 IDSAEIGKNKQPHVSICADIKLALQGLNS ILESKEGKLKLD FSAWRQELTEQKV K HPLNF 452

Query: 142 KTFGEAVP 149
      KTFG+A+P
Sbjct: 453 KTFGDAIP 460

>gb|AAT07327.1| acetohydroxyacid synthase 2 [Helianthus annuus]
Length = 646

Score = 211 bits (536), Expect = 3e-53, Method: Compositional matrix adjust.
Identities = 101/128 (78%), Positives = 112/128 (87%), Gaps = 3/128 (2%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81
      G+YP  +LSLHMLGMHGT VYANYAV+ SDLLAFGVRFDDRVTGKLEAFASRAKIVHID
Sbjct: 315 GAYPASSDLSLHMLGMHGT VYANYAVDKSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 374

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMN K VLENRAEELKLD FGVWRNELNVQKQK FPLSF 141
      ID AEIGKNK PHVS+CGD+K+ALQG+NK+LE E+ LDF WR EL+ QK KFPLSF
Sbjct: 375 IDPAEIGKNKQPHVSICGDIKVALQGLNKILE---EKNSLDFSNWRKELDEQKV K FPLSF 431

Query: 142 KTFGEAVP 149
      KTFGEA+P
Sbjct: 432 KTFGEAIP 439

>gb|AAT07328.1| acetohydroxyacid synthase 2 [Helianthus annuus]
Length = 648

Score = 211 bits (536), Expect = 3e-53, Method: Compositional matrix adjust.
Identities = 101/128 (78%), Positives = 112/128 (87%), Gaps = 3/128 (2%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81
      G+YP  +LSLHMLGMHGT VYANYAV+ SDLLAFGVRFDDRVTGKLEAFASRAKIVHID
Sbjct: 317 GAYPASSDLSLHMLGMHGT VYANYAVDKSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 376
```

Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKNVLENRAEELKLDGFGVWRNELNVQKQKFPLSF 141  
ID AEIGKNK PHVS+CGD+K+ALQG+NK+LE E+ LDF WR EL+ QK KFPLSF  
Sbjct: 377 IDPAEIGKNKQPHVSICGDIKVALQGLNKILE--EKNSLDFSNNWRKELDEQKVKFPLSF 433

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 434 KTFGEAIP 441

>gb|ACF17639.1| putative acetolactate synthase [Capsicum annuum]  
Length = 653

Score = 209 bits (533), Expect = 8e-53, Method: Compositional matrix adjust.  
Identities = 98/129 (75%), Positives = 113/129 (87%), Gaps = 1/129 (0%)

Query: 22 GSYPCDDELSLHMLGMHGTVYANYAVEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G++P D+LSL MLGMHGTVYANYAV+ SDLLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 318 GAFTPGDDLSLQMLGMHGTVYANYAVDSSDLLLAFGVRFDDRVTKLEAFASRAKIVHID 377

Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKNVLENRAEEL-KLDGFGVWRNELNVQKQKFPLS 140  
IDSAEIGKNK PHVS+C D+KLALQG+N +LE + +L KLDG WR ELN QK K+PL+  
Sbjct: 378 IDSAEIGKNKQPHVSICADIKLALQGLNSILEGKEAKLKKLDGSAWRQELNEQVKYPLN 437

Query: 141 FKTFGEAVP 149  
+KTFG+A+P  
Sbjct: 438 YKTFGDAIP 446

>dbj|BAE97675.1| acetolactate synthase [Schoenoplectus juncoide]  
dbj|BAE97676.1| acetolactate synthase [Schoenoplectus juncoide]  
Length = 645

Score = 209 bits (532), Expect = 9e-53, Method: Composition-based stats.  
Identities = 94/128 (73%), Positives = 111/128 (86%)

Query: 22 GSYPCDDELSLHMLGMHGTVYANYAVEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G+YP +DELSL MLGMHGTVYANYAV+ +DLLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 311 GNYPNDELRLMLGMHGTVYANYAVDKADLLLAFGVRFDDRVTKLEAFASRAKIVHID 370

Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKNVLENRAEELKLDGFGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PH+S+C DVK +LQGMN++LE+ KLDG WR EL+ Q++ +PLS+  
Sbjct: 371 IDSAEIGKNKQPHLSICADVKSPLQGMNEILESTGLHRKLDGSSWRAELDEQRKAYPLSY 430

Query: 142 KTFGEAVP 149  
KTFGE +P  
Sbjct: 431 KTFGEIP 438

>gb|ACF47582.1| acetolactate synthase 1 [Sonchus asper]  
Length = 658

Score = 209 bits (531), Expect = 1e-52, Method: Compositional matrix adjust.  
Identities = 97/128 (75%), Positives = 111/128 (86%)

Query: 22 GSYPCDDELSLHMLGMHGTVYANYAVEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G+YP +LSL MLGMHGTVYANYAV+ SDLLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 324 GAYPASGDLRLQMLGMHGTVYANYAVDKSDLLLAFGVRFDDRVTKLEAFASRAKIVHID 383

Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKNVLENRAEELKLDGFGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PHVS+CGD+K+AL+G+N++LE R+E LDF WR EL+ QK PLSF  
Sbjct: 384 IDSAEIGKNKQPHVSICGDIKIALKGLNEILEKRSEMRNLDFSQRQELDEQKLTHPLSF 443

Query: 142 KTFGEAVP 149  
KTFG+A+P  
Sbjct: 444 KTFGDAIP 451

>gb|ACF47583.1| acetolactate synthase 1 [Sonchus asper]

Length = 658

Score = 209 bits (531), Expect = 1e-52, Method: Compositional matrix adjust.  
Identities = 97/128 (75%), Positives = 111/128 (86%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G+YP +LSL MLGMHGT VYANYAV+ SDLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 324 GAYPASGDL SLQMLGMHGT VYANYAVDKSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 383

Query: 82 IDSAEIGKNKTPHVSVC GVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNK PHVS+CGD+K+AL+G+N++LE R+E LDF WR EL+ QK PLSF  
Sbjct: 384 IDSAEIGKNKQPHVSICGDIKIALKGLNEILEKRSEMRNLDFSSWRQELDEQKLTHPLSF 443

Query: 142 KTFGEAVP 149  
KTFG+A+P  
Sbjct: 444 KTFGDAIP 451

>gb|AA053551.1| acetohydroxyacid synthase [Triticum aestivum]  
Length = 598

Score = 207 bits (528), Expect = 3e-52, Method: Composition-based stats.  
Identities = 94/128 (73%), Positives = 108/128 (84%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +DLLAFGVRFDDRVTGK+EAFASR+KIVHID  
Sbjct: 264 GNFPDSDPLSLRMLGMHGT VYANYAVDKADLLAFGVRFDDRVTGKIEAFASRSKIVHID 323

Query: 82 IDSAEIGKNKTPHVSVC GVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID AEIGKNK PHVS+C DVKLALQG+N +L + LDFG W EL+ QK++FPL F  
Sbjct: 324 IDPAEIGKNKQPHVSICADV K LALQGLNALLNGSKAQQGLDFGPWHKELDQQKREFPLGF 383

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 384 KTFGEAIP 391

>gb|AA053549.1| acetohydroxyacid synthase [Triticum aestivum]  
Length = 598

Score = 207 bits (528), Expect = 3e-52, Method: Composition-based stats.  
Identities = 94/128 (73%), Positives = 108/128 (84%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +DLLAFGVRFDDRVTGK+EAFASR+KIVHID  
Sbjct: 264 GNFPDSDPLSLRMLGMHGT VYANYAVDKADLLAFGVRFDDRVTGKIEAFASRSKIVHID 323

Query: 82 IDSAEIGKNKTPHVSVC GVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID AEIGKNK PHVS+C DVKLALQG+N +L + LDFG W EL+ QK++FPL F  
Sbjct: 324 IDPAEIGKNKQPHVSICADV K LALQGLNALLNGSKAQQGLDFGPWHKELDQQKREFPLGF 383

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 384 KTFGEAIP 391

>dbj|BAF57909.1| acetolactate synthase [Sagittaria trifolia]  
Length = 684

Score = 207 bits (526), Expect = 5e-52, Method: Compositional matrix adjust.  
Identities = 97/128 (75%), Positives = 111/128 (86%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
GS+P +LSL MLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 350 GSFTSSDLSLKM LGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 409

Query: 82 IDSAEIGKNKTPHVSVC GVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID AEIGKNK PHVS+CGD+KLAL+G+N++LE +LDF WR EL+ QK+KFPLS+  
Sbjct: 410 IDPAEIGKNKQPHVSICGDLKLAL EGINELLEETKIHEQLDFSSWRGELDEQKRKFPLSY 469

Query: 142 KTFGEAVP 149  
K FG+A+P  
Sbjct: 470 KKFGDAIP 477

>dbj|BAE97677.1| acetolactate synthase [Schoenoplectus juncooides]  
dbj|BAE97678.1| acetolactate synthase [Schoenoplectus juncooides]  
Length = 645

Score = 206 bits (524), Expect = 9e-52, Method: Compositional matrix adjust.  
Identities = 95/128 (74%), Positives = 111/128 (86%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G+YP +DELSL MLGMHGT VYANYAV+ +DLLLLAFGVRFDDRVTKLEAFASR+KIVHID  
Sbjct: 308 GNYPSNDELSRLMLGMHGT VYANYAVDKADLLAFGVRFDDRVTKLEAFASRSKIVHID 367

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
IDSAEIGKNK PH+S+C DVK +LQGMNK+LE+ KLD F WR EL+ QK+ +PL++  
Sbjct: 368 IDSAEIGKNKQPHLSICADV KPSLQGMNKILESTGLHRKLD FSSWRAELDEQKKAYPLNY 427

Query: 142 KTFGEAVP 149  
KTFGE +P  
Sbjct: 428 KTFGEEIP 435

>gb|ABF66048.1| acetolactate synthase [Oryza sativa (indica cultivar-group)]  
Length = 644

Score = 206 bits (523), Expect = 1e-51, Method: Composition-based stats.  
Identities = 92/128 (71%), Positives = 107/128 (83%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +DLLLLAFGVRFDDRVTK+EAFASRAKIVHID  
Sbjct: 310 GNFP SDDPLSLRMLGMHGT VYANYAVDKADLLAFGVRFDDRVTKIEAFASRAKIVHID 369

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID AEIGKNK PHVS+C DVKLALQG+N +L+ + DF W NEL+ QK++FPL +  
Sbjct: 370 IDPAEIGKNKQPHVSICADV KALQGLNALLDQSTTKTSSD FSAWHNELDQKREFP LGY 429

Query: 142 KTFGEAVP 149  
KTFGE +P  
Sbjct: 430 KTFGEEIP 437

>gb|ABF66052.1| acetolactate synthase [Oryza sativa (indica cultivar-group)]  
Length = 644

Score = 206 bits (523), Expect = 1e-51, Method: Composition-based stats.  
Identities = 92/128 (71%), Positives = 107/128 (83%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +DLLLLAFGVRFDDRVTK+EAFASRAKIVHID  
Sbjct: 310 GNFP SDDPLSLRMLGMHGT VYANYAVDKADLLAFGVRFDDRVTKIEAFASRAKIVHID 369

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID AEIGKNK PHVS+C DVKLALQG+N +L+ + DF W NEL+ QK++FPL +  
Sbjct: 370 IDPAEIGKNKQPHVSICADV KALQGLNALLDQSTTKTSSD FSAWHNELDQKREFP LGY 429

Query: 142 KTFGEAVP 149  
KTFGE +P  
Sbjct: 430 KTFGEEIP 437

>gb|AAX14282.1| acetolactate synthase [Oryza sativa (japonica cultivar-group)]  
Length = 644

Score = 206 bits (523), Expect = 1e-51, Method: Composition-based stats.  
Identities = 92/128 (71%), Positives = 107/128 (83%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +DLLLLAFGVRFDDRVTK+EAFASRAKIVHID  
Sbjct: 310 GNFPSDDPLSLRMLGMHGT VYANYAVDKADLLAFGVRFDDRVTKIEAFASRAKIVHID 369

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
ID AEIGKNK PHVS+C DVKLALQG+N +L+ + DF W NEL+ QK++FPL +  
Sbjct: 370 IDPAEIGKNKQPHVSICADV KLALQGLNALLDQSTTKTSSDFS AWHNELDQKREFPLGY 429

Query: 142 KTFGEAVP 149  
KTFGE +P  
Sbjct: 430 KTFGEEIP 437

>gb|AAX14281.1| acetolactate synthase [Oryza sativa]  
gb|ABF66050.1| acetolactate synthase [Oryza sativa (indica cultivar-group)]  
gb|ACD74787.1| acetolactate synthase [Oryza sativa Indica Group]  
gb|ACD74788.1| acetolactate synthase [Oryza sativa Indica Group]  
Length = 644

Score = 206 bits (523), Expect = 1e-51, Method: Composition-based stats.  
Identities = 92/128 (71%), Positives = 107/128 (83%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +DLLLLAFGVRFDDRVTK+EAFASRAKIVHID  
Sbjct: 310 GNFPSDDPLSLRMLGMHGT VYANYAVDKADLLAFGVRFDDRVTKIEAFASRAKIVHID 369

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
ID AEIGKNK PHVS+C DVKLALQG+N +L+ + DF W NEL+ QK++FPL +  
Sbjct: 370 IDPAEIGKNKQPHVSICADV KLALQGLNALLDQSTTKTSSDFS AWHNELDQKREFPLGY 429

Query: 142 KTFGEAVP 149  
KTFGE +P  
Sbjct: 430 KTFGEEIP 437

>gb|ABY57316.1| acetolactate synthase [Medicago littoralis]  
Length = 651

Score = 205 bits (521), Expect = 2e-51, Method: Compositional matrix adjust.  
Identities = 97/127 (76%), Positives = 111/127 (87%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
GSYP E SL MLGMHGT VYANYAV++SDLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 317 GSYPIGGEHLSMLGMHGT VYANYAVDNSDLLAFGVRFDDRVTKLEAFASRAKIVHID 376

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PH+S+C D+K+AL+G+N+VLE++ + KLDF WR ELNVQK KFPL F  
Sbjct: 377 IDSAEIGKNKIPHLSICADMKVALEGLNRVLESKGIKGLDFEAWRQELNVQKLFPLGF 436

Query: 142 KTFGEAV 148  
KTF +A+  
Sbjct: 437 KTFEDAI 443

>gb|ABY57315.1| acetolactate synthase [Medicago littoralis]  
Length = 651

Score = 204 bits (520), Expect = 2e-51, Method: Compositional matrix adjust.  
Identities = 97/127 (76%), Positives = 111/127 (87%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
GSYP E SL MLGMHGT VYANYAV++SDLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 317 GSYPIGGEHLSMLGMHGT VYANYAVDNSDLLAFGVRFDDRVTKLEAFASRAKIVHID 376

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PH+S+C D+K+AL+G+N+VLE++ + KLDF WR ELNVQK KFPL F  
Sbjct: 377 IDSAEIGKNKIPHLSICADMKVALEGLNRVLESKGIKGLDFEAWRQELNVQKLFPLGF 436

Query: 142 KTFGEAV 148  
KTF +A+  
Sbjct: 437 KTFEDAI 443

>gb|ABY57317.1| acetolactate synthase [Medicago truncatula]  
Length = 651

Score = 204 bits (520), Expect = 2e-51, Method: Compositional matrix adjust.  
Identities = 97/127 (76%), Positives = 111/127 (87%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
GSYP E SL MLGMHGT VYANYAV++SDLLLA FGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 317 GSYPIGGEHSLSLMLGMHGT VYANYAVDNSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 376

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FG VWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PH+S+C D+K+AL+G+N+VLE++ + KLDF WR ELNVQK KFPL F  
Sbjct: 377 IDSAEIGKNKIPHLSICADMKVLEGLNRVLESKGIKGLDFEAWRQELNVQKLFPLGF 436

Query: 142 KTFGEAV 148  
KTF +A+  
Sbjct: 437 KTFEDAI 443

>gb|ABY57318.1| acetolactate synthase [Medicago truncatula]  
Length = 651

Score = 204 bits (520), Expect = 3e-51, Method: Compositional matrix adjust.  
Identities = 97/127 (76%), Positives = 111/127 (87%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
GSYP E SL MLGMHGT VYANYAV++SDLLLA FGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 317 GSYPIGGEHSLSLMLGMHGT VYANYAVDNSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 376

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FG VWRNELNVQKQKFPLSF 141  
IDSAEIGKNK PH+S+C D+K+AL+G+N+VLE++ + KLDF WR ELNVQK KFPL F  
Sbjct: 377 IDSAEIGKNKIPHLSICADMKVLEGLNRVLESKGIKGLDFEAWRQELNVQKLFPLGF 436

Query: 142 KTFGEAV 148  
KTF +A+  
Sbjct: 437 KTFEDAI 443

>gb|AA14283.1| acetolactate synthase [Oryza sativa]  
Length = 644

Score = 204 bits (519), Expect = 4e-51, Method: Composition-based stats.  
Identities = 91/128 (71%), Positives = 107/128 (83%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +DLLLA FGVRFDDRVTGK+EAFAASRAKIVHID  
Sbjct: 310 GNFPDSDLPLSLRMLGMHGT VYANYAVDKADLLLA FGVRFDDRVTGKIEAFASRAKIVHID 369

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FG VWRNELNVQKQKFPLSF 141  
ID AEIGKNK PHVS+C DV+LALQG+N +L+ + DF W NEL+ QK++FPL +  
Sbjct: 370 IDPAEIGKNKQPHVSICADVELALQGLNALLDQSTTKTSSDFS AWHNELDQKREFPLGY 429

Query: 142 KTFGEAVP 149  
KTFGE +P  
Sbjct: 430 KTFGEEIP 437

>dbj|BAF37288.1| acetolactate synthase [Monochoria vaginalis]  
Length = 614

Score = 204 bits (518), Expect = 4e-51, Method: Composition-based stats.  
Identities = 92/128 (71%), Positives = 106/128 (82%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81

G YP D LSL MLGMHGT VYANYA++ +DLLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 285 G VYPLDGPLSLKMLGMHGT VYANYAIDKADLLLAFGVRFDDRVTGKLEAFASRAKIVHID 344

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
ID AEIGKNK PHVS+CGD+KLALQ MN+++E KLD F WR EL+ QK+ +PL +  
Sbjct: 345 IDPAEIGKNKQPHVSICGDIKLALQEMNEMIEENGIH SKLDFS AWREELDQQKKNYPLKY 404

Query: 142 KTFGEAVP 149  
KTFG+ +P  
Sbjct: 405 KTFGDLIP 412

>dbj|BAE53610.1| acetolactate synthase [Monochoria vaginalis]  
Length = 614

Score = 204 bits (518), Expect = 4e-51, Method: Composition-based stats.  
Identities = 92/128 (71%), Positives = 106/128 (82%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G YP D LSL MLGMHGT VYANYA++ +DLLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 285 G VYPLDGPLSLKMLGMHGT VYANYAIDKADLLLAFGVRFDDRVTGKLEAFASRAKIVHID 344

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
ID AEIGKNK PHVS+CGD+KLALQ MN+++E KLD F WR EL+ QK+ +PL +  
Sbjct: 345 IDPAEIGKNKQPHVSICGDIKLALQEMNEMIEENGIH SKLDFS AWREELDQQKKNYPLKY 404

Query: 142 KTFGEAVP 149  
KTFG+ +P  
Sbjct: 405 KTFGDLIP 412

>dbj|BAE53611.1| acetolactate synthase [Monochoria vaginalis]  
Length = 614

Score = 204 bits (518), Expect = 4e-51, Method: Composition-based stats.  
Identities = 92/128 (71%), Positives = 106/128 (82%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G YP D LSL MLGMHGT VYANYA++ +DLLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 285 G VYPLDGPLSLKMLGMHGT VYANYAIDKADLLLAFGVRFDDRVTGKLEAFASRAKIVHID 344

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
ID AEIGKNK PHVS+CGD+KLALQ MN+++E KLD F WR EL+ QK+ +PL +  
Sbjct: 345 IDPAEIGKNKQPHVSICGDIKLALQEMNEMIEENGIH SKLDFS AWREELDQQKKNYPLKY 404

Query: 142 KTFGEAVP 149  
KTFG+ +P  
Sbjct: 405 KTFGDLIP 412

>dbj|BAE53612.1| acetolactate synthase [Monochoria vaginalis]  
Length = 614

Score = 204 bits (518), Expect = 4e-51, Method: Composition-based stats.  
Identities = 92/128 (71%), Positives = 106/128 (82%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G YP D LSL MLGMHGT VYANYA++ +DLLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 285 G VYPLDGPLSLKMLGMHGT VYANYAIDKADLLLAFGVRFDDRVTGKLEAFASRAKIVHID 344

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
ID AEIGKNK PHVS+CGD+KLALQ MN+++E KLD F WR EL+ QK+ +PL +  
Sbjct: 345 IDPAEIGKNKQPHVSICGDIKLALQEMNEMIEENGIH SKLDFS AWREELDQQKKNYPLKY 404

Query: 142 KTFGEAVP 149  
KTFG+ +P  
Sbjct: 405 KTFGDLIP 412



```
>dbj|BAE53605.1| acetolactate synthase [Monochoria vaginalis]
dbj|BAE53606.1| acetolactate synthase [Monochoria vaginalis]
dbj|BAE53607.1| acetolactate synthase [Monochoria vaginalis]
dbj|BAE53609.1| acetolactate synthase [Monochoria vaginalis]
Length = 614

Score = 204 bits (518), Expect = 4e-51, Method: Composition-based stats.
Identities = 92/128 (71%), Positives = 106/128 (82%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81
          G YP D LSL MLGMHGT VYANYA++ +DLLAFGVRFDDRVTGKLEAFASRAKIVHID
Sbjct: 285 GVYPLDGP LSLKMLGMHGT VYANYAIDKADLLAFGVRFDDRVTGKLEAFASRAKIVHID 344

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMN K VLENRAEELKLD FGVWRNELNVQKQK FPLSF 141
          ID AEIGKNK PHVS+CGD+KLALQ MN+++E K LDF WR EL+ QK+ +PL +
Sbjct: 345 IDPAEIGKNKQPHVSICGDIKLALQEMNEMIEENGIH SKLDFS A WREELDQQKKNYPLKY 404

Query: 142 KTFGEAVP 149
          KTFG+ +P
Sbjct: 405 KTFGDLIP 412

>dbj|BAE53608.1| acetolactate synthase [Monochoria vaginalis]
Length = 614

Score = 204 bits (518), Expect = 4e-51, Method: Composition-based stats.
Identities = 92/128 (71%), Positives = 106/128 (82%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81
          G YP D LSL MLGMHGT VYANYA++ +DLLAFGVRFDDRVTGKLEAFASRAKIVHID
Sbjct: 285 GVYPLDGP LSLKMLGMHGT VYANYAIDKADLLAFGVRFDDRVTGKLEAFASRAKIVHID 344

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMN K VLENRAEELKLD FGVWRNELNVQKQK FPLSF 141
          ID AEIGKNK PHVS+CGD+KLALQ MN+++E K LDF WR EL+ QK+ +PL +
Sbjct: 345 IDPAEIGKNKQPHVSICGDIKLALQEMNEMIEENGIH SKLDFS A WREELDQQKKNYPLKY 404

Query: 142 KTFGEAVP 149
          KTFG+ +P
Sbjct: 405 KTFGDLIP 412

>dbj|BAE53591.1| acetolactate synthase [Monochoria vaginalis]
dbj|BAE53592.1| acetolactate synthase [Monochoria vaginalis]
Length = 642

Score = 203 bits (517), Expect = 6e-51, Method: Composition-based stats.
Identities = 92/128 (71%), Positives = 106/128 (82%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81
          G YP D LSL MLGMHGT VYANYA++ +DLLAFGVRFDDRVTGKLEAFASRAKIVHID
Sbjct: 308 GVYPLDGP LSLKMLGMHGT VYANYAIDKADLLAFGVRFDDRVTGKLEAFASRAKIVHID 367

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMN K VLENRAEELKLD FGVWRNELNVQKQK FPLSF 141
          ID AEIGKNK PHVS+CGD+KLALQ MN+++E K LDF WR EL+ QK+ +PL +
Sbjct: 368 IDPAEIGKNKQPHVSICGDIKLALQEMNEMIEESGIHNKLD FSA WREELDQQKKNYPLEY 427

Query: 142 KTFGEAVP 149
          KTFG+ +P
Sbjct: 428 KTFGDLIP 435

>dbj|BAE53588.1| acetolactate synthase [Monochoria vaginalis]
Length = 642

Score = 203 bits (517), Expect = 6e-51, Method: Composition-based stats.
Identities = 92/128 (71%), Positives = 106/128 (82%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81
          G YP D LSL MLGMHGT VYANYA++ +DLLAFGVRFDDRVTGKLEAFASRAKIVHID
```

Sbjct: 308 GVPYLDGPLSLKMLGMHGT VYANYAIDKADLLLAFGVRFDDRV TGKLEAFASRAKIVHID 367

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141

ID AEIGKNK PHVS+CGD+KLALQ MN+++E KLD F WR EL+ QK+ +PL +

Sbjct: 368 IDPAEIGKNKQPHVSICGDIKLALQEMNEMIEESGIHNKLD FSAWREELDQQKKNYPLEY 427

Query: 142 KTFGEAVP 149

KTFG+ +P

Sbjct: 428 KTFGDLIP 435

>dbj|BAE53593.1| acetolactate synthase [Monochoria vaginalis]

Length = 642

Score = 203 bits (517), Expect = 6e-51, Method: Composition-based stats.  
Identities = 92/128 (71%), Positives = 106/128 (82%)

Query: 22 GSYPCDDEL SLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRV TGKLEAFASRAKIVHID 81

G YP D LSL MLGMHGT VYANYA++ +D LLA FGVRFDDRV TGKLEAFASRAKIVHID

Sbjct: 308 GVPYLDGPLSLKMLGMHGT VYANYAIDKADLLLA FGVRFDDRV TGKLEAFASRAKIVHID 367

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141

ID AEIGKNK PHVS+CGD+KLALQ MN+++E KLD F WR EL+ QK+ +PL +

Sbjct: 368 IDPAEIGKNKQPHVSICGDIKLALQEMNEMIEESGIHNKLD FSAWREELDQQKKNYPLEY 427

Query: 142 KTFGEAVP 149

KTFG+ +P

Sbjct: 428 KTFGDLIP 435

>dbj|BAE53587.1| acetolactate synthase [Monochoria vaginalis]

dbj|BAE53589.1| acetolactate synthase [Monochoria vaginalis]

dbj|BAE53590.1| acetolactate synthase [Monochoria vaginalis]

dbj|BAE53594.1| acetolactate synthase [Monochoria vaginalis]

Length = 642

Score = 203 bits (517), Expect = 6e-51, Method: Composition-based stats.  
Identities = 92/128 (71%), Positives = 106/128 (82%)

Query: 22 GSYPCDDEL SLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRV TGKLEAFASRAKIVHID 81

G YP D LSL MLGMHGT VYANYA++ +D LLA FGVRFDDRV TGKLEAFASRAKIVHID

Sbjct: 308 GVPYLDGPLSLKMLGMHGT VYANYAIDKADLLLA FGVRFDDRV TGKLEAFASRAKIVHID 367

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141

ID AEIGKNK PHVS+CGD+KLALQ MN+++E KLD F WR EL+ QK+ +PL +

Sbjct: 368 IDPAEIGKNKQPHVSICGDIKLALQEMNEMIEESGIHNKLD FSAWREELDQQKKNYPLEY 427

Query: 142 KTFGEAVP 149

KTFG+ +P

Sbjct: 428 KTFGDLIP 435

>gb|AAG30931.1|AF310684\_1 acetolactate synthase precursor [Lolium multiflorum]

Length = 640

Score = 203 bits (516), Expect = 7e-51, Method: Composition-based stats.  
Identities = 92/128 (71%), Positives = 106/128 (82%)

Query: 22 GSYPCDDEL SLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRV TGKLEAFASRAKIVHID 81

G++P DD LSL MLGMHGT VYANYAV+ +D LLA FGVRFDDRV TGK+EAFASR+KIVHID

Sbjct: 306 GNFPSSDDPLSLRMLGMHGT VYANYAVDKADLLLA FGVRFDDRV TGKIEAFASRSKIVHID 365

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141

ID AEIGKNK PHVS+C DVKLALQG+N VL + DF W +EL QK++FPL +

Sbjct: 366 IDPAEIGKNKQPHVSICADV KLALQGLNAVL TGSKCDKSFDFASWHDELEQQKREFPLGY 425

Query: 142 KTFGEAVP 149

KTFGEA+P

Sbjct: 426 KTFGEAIP 433

>gb|AA053548.1| acetohydroxyacid synthase [Triticum aestivum]  
Length = 598

Score = 202 bits (515), Expect = 1e-50, Method: Compositional matrix adjust.  
Identities = 94/128 (73%), Positives = 108/128 (84%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +DLLAFGVRFDDRVTGK+EAFASR+KIVHID  
Sbjct: 264 GNFPSDDPLSLRMLGMHGT VYANYAVDKADLLAFGVRFDDRVTGKIEAFASRSKIVHID 323

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID AEIGKNK PHVS+C DVKLALQG+N +L + LDFG W EL+ QK++FPL F  
Sbjct: 324 IDPAEIGKNKQPHVSICADV KLALQGLNDLLNGSKAQGLDFGPWHKELDQKREFPLGF 383

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 384 KTFGEAIP 391

>gb|AA053550.1| acetohydroxyacid synthase [Triticum aestivum]  
Length = 598

Score = 202 bits (515), Expect = 1e-50, Method: Compositional matrix adjust.  
Identities = 94/128 (73%), Positives = 108/128 (84%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +DLLAFGVRFDDRVTGK+EAFASR+KIVHID  
Sbjct: 264 GNFPSDDPLSLRMLGMHGT VYANYAVDKADLLAFGVRFDDRVTGKIEAFASRSKIVHID 323

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID AEIGKNK PHVS+C DVKLALQG+N +L + LDFG W EL+ QK++FPL F  
Sbjct: 324 IDPAEIGKNKQPHVSICADV KLALQGLNDLLNGSKAQGLDFGPWHKELDQKREFPLGF 383

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 384 KTFGEAIP 391

>gb|AAC14572.1| acetohydroxyacid synthase [Hordeum vulgare]  
Length = 541

Score = 202 bits (514), Expect = 1e-50, Method: Compositional matrix adjust.  
Identities = 93/128 (72%), Positives = 108/128 (84%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +DLLAFGVRFDDRVTGK+EAFASR+KIVHID  
Sbjct: 207 GNFPSDDPLSLRMLGMHGT VYANYAVDKADLLAFGVRFDDRVTGKIEAFASRSKIVHID 266

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID AEIGKNK PHVS+C DVKLALQG+N +L + LDFG W EL+ QK++FPL +  
Sbjct: 267 IDPAEIGKNKQPHVSICADV KLALQGLNLLSGSKAQGLDFGPWHKELDQKREFPLGY 326

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 327 KTFGEAIP 334

>gb|ABM92357.2| acetolactate synthase [Cyperus difformis]  
Length = 569

Score = 202 bits (513), Expect = 1e-50, Method: Compositional matrix adjust.  
Identities = 93/128 (72%), Positives = 109/128 (85%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G++PCD+ L L +LGMHGT VYANYAV+ +DLLAFGVRFDDRVTGKLEAFASR+KIVHID  
Sbjct: 236 GNFPCEDEPLCLRLMLGMHGT VYANYAVDKADLLAFGVRFDDRVTGKLEAFASRSKIVHID 295

Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKNVLENRAEELKLDGFGVWRNELNVQKQKFPPLSF 141  
ID AEIGKNK PHVS+C DVK ALQGMN++LE+ KLDF WR EL+ QK+ +PLS+  
Sbjct: 296 IDPAEIGKNKQPHVSICADV KPALQGMNQILESSGVHKKLDFSSWRAELDEQKKTYPPLSY 355

Query: 142 KTFGEAVP 149  
KTFGE +P  
Sbjct: 356 KTFGEEIP 363

>gb|AAM03119.1|AF488771\_1 acetolactate synthase [Bromus tectorum]  
Length = 583

Score = 202 bits (513), Expect = 2e-50, Method: Compositional matrix adjust.  
Identities = 93/128 (72%), Positives = 107/128 (83%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +D LLA FGVRFD DRTGK+EAFASR+KIVHID  
Sbjct: 249 GNFP SDDPLSLRMLGMHGT VYANYAVDKADLLLA FGVRFD DRTGKIEAFASRSKIVHID 308

Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKNVLENRAEELKLDGFGVWRNELNVQKQKFPPLSF 141  
ID AEIGKNK PHVS+C DVKLALQG+N +L + LDFG W+ EL QK+ FPL +  
Sbjct: 309 IDPAEIGKNKQPHVSICADV KPALQGLNDLLNGSKAQKSLDFGPWQEELEQQKRTFPLGY 368

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 369 KTFGEAIP 376

>gb|AAL93207.1|AF487459\_1 acetolactate synthase [Bromus tectorum]  
Length = 583

Score = 202 bits (513), Expect = 2e-50, Method: Compositional matrix adjust.  
Identities = 93/128 (72%), Positives = 107/128 (83%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +D LLA FGVRFD DRTGK+EAFASR+KIVHID  
Sbjct: 249 GNFP SDDPLSLRMLGMHGT VYANYAVDKADLLLA FGVRFD DRTGKIEAFASRSKIVHID 308

Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKNVLENRAEELKLDGFGVWRNELNVQKQKFPPLSF 141  
ID AEIGKNK PHVS+C DVKLALQG+N +L + LDFG W+ EL QK+ FPL +  
Sbjct: 309 IDPAEIGKNKQPHVSICADV KPALQGLNDLLNGSKAQKSLDFGPWQEELEQQKRTFPLGY 368

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 369 KTFGEAIP 376

>emb|CAD24801.2| acetolactate synthase [Alopecurus myosuroides]  
Length = 640

Score = 201 bits (512), Expect = 2e-50, Method: Composition-based stats.  
Identities = 92/129 (71%), Positives = 109/129 (84%), Gaps = 1/129 (0%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +D LLA FGVRFD DRTGK+EAFASR+KIVHID  
Sbjct: 305 GNFP SDDPLSLRMLGMHGT VYANYAVDRADLLLA FGVRFD DRTGKIEAFASRSKIVHID 364

Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKNVLENRAEELKL-DFGVWRNELNVQKQKFPPLS 140  
ID AEIGKNK PHVS+C DVK+ALQG+N +L N ++ K DF W EL+ QK++FPL  
Sbjct: 365 IDPAEIGKNKQPHVSICADV KIALQGLNTLLLNESKTRKCSDFSSWHEELDQQKREFPLG 424

Query: 141 FKTFGEAVP 149  
+KTFGEA+P  
Sbjct: 425 YKTFGEAIP 433

>gb|EEE57060.1| hypothetical protein OsJ\_06870 [Oryza sativa Japonica Group]  
Length = 547

Score = 201 bits (511), Expect = 3e-50, Method: Compositional matrix adjust.  
Identities = 92/128 (71%), Positives = 107/128 (83%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFD RVTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +DLLLLAFGVRFD RVTGK+EAFASRAKIVHID  
Sbjct: 213 GNFP SDDPLSLRMLGMHGT VYANYAVDKADLLAFGVRFD RVTGKIEAFASRAKIVHID 272

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID AEIGKNK PHVS+C DVKLALQ G+N +L+ + DF W NEL+ QK++FPL +  
Sbjct: 273 IDPAEIGKNKQPHVSICADVKLALQGLNALLQ QSTTKTSSDFS AWHNELDQKREFPLGY 332

Query: 142 KTFGEAVP 149  
KTFGE +P  
Sbjct: 333 KTFGEEIP 340

>emb|CAE18088.1| acetolactate synthase [Papaver rhoeas]  
Length = 662

Score = 201 bits (510), Expect = 3e-50, Method: Compositional matrix adjust.  
Identities = 92/128 (71%), Positives = 110/128 (85%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFD RVTGKLEAFASRAKIVHID 81  
G++ C D LSL MLGMHGT VYANYAV+ +DLLLLAFGVRFD RVTGK+EAFASR+KIVHID  
Sbjct: 328 GAHSCTDVL SLQMLGMHGT VYANYAVDKADLLAFGVRFD RVTGKIEAFASRSKIVHID 387

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID AEIGKNK PHVS+C D+ +AL+G+NK+LE + +LKLD F WR EL QK+ +PLS+  
Sbjct: 388 IDPAEIGKNKQPHVSICADMGVALKGLNKLLEEKMSKLKLD FVSWREELMEQKKNYPLSY 447

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 448 KTFGEAIP 455

>dbj|BAB20813.1| acetolactate synthase [Oryza sativa (japonica cultivar-group)]  
Length = 644

Score = 201 bits (510), Expect = 4e-50, Method: Compositional matrix adjust.  
Identities = 92/128 (71%), Positives = 107/128 (83%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFD RVTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +DLLLLAFGVRFD RVTGK+EAFASRAKIVHID  
Sbjct: 310 GNFP SDDPLSLRMLGMHGT VYANYAVDKADLLAFGVRFD RVTGKIEAFASRAKIVHID 369

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID AEIGKNK PHVS+C DVKLALQ G+N +L+ + DF W NEL+ QK++FPL +  
Sbjct: 370 IDPAEIGKNKQPHVSICADVKLALQGLNALLQ QSTTKTSSDFS AWHNELDQKREFPLGY 429

Query: 142 KTFGEAVP 149  
KTFGE +P  
Sbjct: 430 KTFGEEIP 437

>ref|NP\_001046931.1| Os02g0510200 [Oryza sativa (japonica cultivar-group)]  
sp|Q6K2E8.1|ILVB1\_ORYSJ RecName: Full=Acetolactate synthase 1, chloroplastic; AltName:  
Full=Acetohydroxy-acid synthase 1; Flags: Precursor  
dbj|BAB20812.1| acetolactate synthase [Oryza sativa (japonica cultivar-group)]  
dbj|BAD23668.1| acetolactate synthase [Oryza sativa Japonica Group]  
dbj|BAF08845.1| Os02g0510200 [Oryza sativa (japonica cultivar-group)]  
Length = 644

Score = 201 bits (510), Expect = 4e-50, Method: Compositional matrix adjust.  
Identities = 92/128 (71%), Positives = 107/128 (83%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFD RVTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +DLLLLAFGVRFD RVTGK+EAFASRAKIVHID  
Sbjct: 310 GNFP SDDPLSLRMLGMHGT VYANYAVDKADLLAFGVRFD RVTGKIEAFASRAKIVHID 369

Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKLLENRAEELKLDGFGVWRNELNVQKQKFPPLSF 141  
ID AEIGKNK PHVS+C DVKLALQG+N +L+ + DF W NEL+ QK++FPL +  
Sbjct: 370 IDPAEIGKNKQPHVSICADVKLALQGLNALLQSTTKTSSDFSAAHNELDQQKREFPLGY 429

Query: 142 KTFGEAVP 149  
KTFGE +P  
Sbjct: 430 KTFGEEIP 437

>gb|ACD74789.1| acetolactate synthase [Oryza sativa Indica Group]  
gb|ACD74790.1| acetolactate synthase [Oryza sativa Indica Group]  
Length = 644

Score = 200 bits (509), Expect = 4e-50, Method: Compositional matrix adjust.  
Identities = 92/128 (71%), Positives = 107/128 (83%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFD RVTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +D LLLAFGVRFD RVTGK+EAFASRAKIVHID  
Sbjct: 310 GNFPD DPLSLRMLGMHGT VYANYAVDKADLLAFGVRFD RVTGKIEAFASRAKIVHID 369

Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKLLENRAEELKLDGFGVWRNELNVQKQKFPPLSF 141  
ID AEIGKNK PHVS+C DVKLALQG+N +L+ + DF W NEL+ QK++FPL +  
Sbjct: 370 IDPAEIGKNKQPHVSICADVKLALQGLNALLQSTTKTSSDFSAAHNELDQQKREFPLGY 429

Query: 142 KTFGEAVP 149  
KTFGE +P  
Sbjct: 430 KTFGEEIP 437

>gb|ABF66051.1| acetolactate synthase [Oryza sativa (indica cultivar-group)]  
Length = 644

Score = 200 bits (509), Expect = 5e-50, Method: Compositional matrix adjust.  
Identities = 92/128 (71%), Positives = 107/128 (83%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFD RVTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +D LLLAFGVRFD RVTGK+EAFASRAKIVHID  
Sbjct: 310 GNFPD DPLSLRMLGMHGT VYANYAVDKADLLAFGVRFD RVTGKIEAFASRAKIVHID 369

Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKLLENRAEELKLDGFGVWRNELNVQKQKFPPLSF 141  
ID AEIGKNK PHVS+C DVKLALQG+N +L+ + DF W NEL+ QK++FPL +  
Sbjct: 370 IDPAEIGKNKQPHVSICADVKLALQGLNALLQSTTKTSSDFSAAHNELDQQKREFPLGY 429

Query: 142 KTFGEAVP 149  
KTFGE +P  
Sbjct: 430 KTFGEEIP 437

>dbj|BAE53604.1| acetolactate synthase [Monochoria vaginalis]  
Length = 614

Score = 199 bits (506), Expect = 1e-49, Method: Composition-based stats.  
Identities = 91/128 (71%), Positives = 104/128 (81%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFD RVTGKLEAFASRAKIVHID 81  
G YP D LSL M GMHGT VYANYA++ +D LLLAFGVRFD RVTGKLEAFASRAKIVHID  
Sbjct: 285 GVYPLDGPLSLKMWGMHGT VYANYAIDKADLLAFGVRFD RVTGKLEAFASRAKIVHID 344

Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKLLENRAEELKLDGFGVWRNELNVQKQKFPPLSF 141  
ID AEIGKNK PHVS+CGD+KLALQ MN+++E KLDF WR EL+ QK+ PL +  
Sbjct: 345 IDPAEIGKNKQPHVSICGDIKLALQEMNEMIEENGISKLDFSAAWREELDQQKKNPLKY 404

Query: 142 KTFGEAVP 149  
KTFG+ +P  
Sbjct: 405 KTFGDLIP 412

>gb|AAT07329.1| acetohydroxyacid synthase 3 [Helianthus annuus]  
Length = 646

Score = 199 bits (506), Expect = 1e-49, Method: Compositional matrix adjust.  
Identities = 93/128 (72%), Positives = 108/128 (84%), Gaps = 1/128 (0%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G+YP +LSLHMLGMHGT VYANYA++ SDLLAFGVRFDDRVTK+EAFASRAKIVHID  
Sbjct: 314 GTPGSHDLSLHMLGMHGT VYANYAIDKSDLLAFGVRFDDRVTKIEAFASRAKIVHID 373

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID AEIGKNK PH S+CGD+K ALQG+NK+LE R E+L DF W+ E+ QK PLS+  
Sbjct: 374 IDPAEIGKNKQPHFSICGDIKALQGLNKILE-RGEDLDFDFSPWKEEVMNQKASNPLSY 432

Query: 142 KTFGEAVP 149  
KTFG+A+P  
Sbjct: 433 KTFGDAIP 440

>emb|CAP09635.1| acetolactate-synthase-N-DnaE intein-N fusion protein  
[Transformation Vector pICH13688]  
Length = 558

Score = 199 bits (505), Expect = 2e-49, Method: Compositional matrix adjust.  
Identities = 103/131 (78%), Positives = 107/131 (81%), Gaps = 5/131 (3%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 336 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 395

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAE--ELKLD FGVWRNELNVQKQKFP L 139  
IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLEN + E L FG L V+ P+  
Sbjct: 396 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENDVKFAEYCLSFGT--EILTVEYGPLPI 453

Query: 140 SFKTFGEAVPC 150  
K E + C  
Sbjct: 454 G-KIVSEEINC 463

>gb|EAY93888.1| hypothetical protein OsI\_15664 [Oryza sativa Indica Group]  
Length = 662

Score = 198 bits (504), Expect = 2e-49, Method: Composition-based stats.  
Identities = 89/128 (69%), Positives = 107/128 (83%), Gaps = 1/128 (0%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV++D LLLA GVRFDDRVTK+EAFASRAKIVH+D  
Sbjct: 328 GNFP SDDPLSLRMLGMHGT VYANYAVDNAD LLLALGVRFDDRVTKVEAFASRAKIVHVD 387

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID +E+GKNK PHVS+C DVKLALQGMN LE + + LDF WR+EL +K +FPL +  
Sbjct: 388 IDPSELGKNKQPHVSICADV K LALQGMNATLEQQQRK-NLDFS AWRSELEKKKA EFPLGY 446

Query: 142 KTFGEAVP 149  
+TFGE +P  
Sbjct: 447 RTFGEEIP 454

>gb|ABQ85869.1| acetolactate synthase [Lolium rigidum]  
Length = 516

Score = 198 bits (504), Expect = 2e-49, Method: Compositional matrix adjust.  
Identities = 92/128 (71%), Positives = 105/128 (82%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +D LLLAFGVRFDDRVTK+EAFASR+KIVHID  
Sbjct: 201 GNFP SDDPLSLRMLGMHGT VYANYAVDKAD LLLAFGVRFDDRVTKIEAFASRSKIVHID 260

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID AEIGKNK PHVS+C DVKLALQG+N +L DFG W EL QK++FPL +  
Sbjct: 261 IDPAEIGKNKQPHVSICADV K LALQGLNALLTGTKAHKS FDFGS WHEELQKREFFPLGY 320

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 321 KTFGEAIP 328

>gb|ABQ85870.1| acetolactate synthase [Lolium rigidum]  
Length = 516

Score = 198 bits (504), Expect = 2e-49, Method: Compositional matrix adjust.  
Identities = 92/128 (71%), Positives = 105/128 (82%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +DLLAFGVRFDDRVTGK+EAFASR+KIVHID  
Sbjct: 201 GNFPDSDPLSLRMLGMHGT VYANYAVDKADLLAFGVRFDDRVTGKIEAFASRSKIVHID 260

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
ID AEIGKNK PHVS+C DVKLALQG+N +L DFG W EL QK++FPL +  
Sbjct: 261 IDPAEIGKNKQPHVSICADV KLALQGLNALLTGTKAHKSFD FGSWHEELEQQKREFPLGY 320

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 321 KTFGEAIP 328

>dbj|BAE53596.1| acetolactate synthase [Monochoria vaginalis]  
Length = 604

Score = 198 bits (503), Expect = 2e-49, Method: Composition-based stats.  
Identities = 89/128 (69%), Positives = 105/128 (82%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G YP D LSL MLGMHGT VYANY+++ +DLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 270 GVYPLDAHL SLKMLGMHGT VYANYSIDKADLLAFGVRFDDRVTGKLEAFASRAKIVHID 329

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
ID AEIGKNK PHVS+C D+KLALQ MNK++E KL+F WR EL+ K+ +PL++  
Sbjct: 330 IDPAEIGKNKQPHVSICADIKLALQEMNKIIEESGIYNKLNFSAWREELDQHKKNYPLNY 389

Query: 142 KTFGEAVP 149  
KTFG+ +P  
Sbjct: 390 KTFGDLIP 397

>dbj|BAE53598.1| acetolactate synthase [Monochoria vaginalis]  
dbj|BAE53601.1| acetolactate synthase [Monochoria vaginalis]  
Length = 604

Score = 198 bits (503), Expect = 2e-49, Method: Composition-based stats.  
Identities = 89/128 (69%), Positives = 105/128 (82%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G YP D LSL MLGMHGT VYANY+++ +DLLAFGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 270 GVYPLDAHL SLKMLGMHGT VYANYSIDKADLLAFGVRFDDRVTGKLEAFASRAKIVHID 329

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
ID AEIGKNK PHVS+C D+KLALQ MNK++E KL+F WR EL+ K+ +PL++  
Sbjct: 330 IDPAEIGKNKQPHVSICADIKLALQEMNKIIEESGIYNKLNFSAWREELDQHKKNYPLNY 389

Query: 142 KTFGEAVP 149  
KTFG+ +P  
Sbjct: 390 KTFGDLIP 397

>dbj|BAE53602.1| acetolactate synthase [Monochoria vaginalis]  
Length = 604

Score = 198 bits (503), Expect = 2e-49, Method: Composition-based stats.  
Identities = 89/128 (69%), Positives = 105/128 (82%)



Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
G YP D LSL MLGMHGT VYANY+++ +DLLLLA FGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 270 G VYPLDAHL SLKMLGMHGT VYANYSIDKADLLLA FGVRFDDRVTGKLEAFASRAKIVHID 329

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID AEIGKNK PHVS+C D+KLALQ MNK++E KL+F WR EL+ K+ +PL++  
Sbjct: 330 IDPAEIGKNKQPHVSICADIKLALQEMNKIIEESGIYNKLNFSAWREELDQHKKNYPLNY 389

Query: 142 KTFGEAVP 149  
KTFG+ +P  
Sbjct: 390 KTFGDLIP 397

>dbj|BAE53597.1| acetolactate synthase [Monochoria vaginalis]  
dbj|BAE53599.1| acetolactate synthase [Monochoria vaginalis]  
dbj|BAE53600.1| acetolactate synthase [Monochoria vaginalis]  
dbj|BAE53603.1| acetolactate synthase [Monochoria vaginalis]  
Length = 604

Score = 198 bits (503), Expect = 2e-49, Method: Composition-based stats.  
Identities = 89/128 (69%), Positives = 105/128 (82%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
G YP D LSL MLGMHGT VYANY+++ +DLLLLA FGVRFDDRVTGKLEAFASRAKIVHID  
Sbjct: 270 G VYPLDAHL SLKMLGMHGT VYANYSIDKADLLLA FGVRFDDRVTGKLEAFASRAKIVHID 329

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID AEIGKNK PHVS+C D+KLALQ MNK++E KL+F WR EL+ K+ +PL++  
Sbjct: 330 IDPAEIGKNKQPHVSICADIKLALQEMNKIIEESGIYNKLNFSAWREELDQHKKNYPLNY 389

Query: 142 KTFGEAVP 149  
KTFG+ +P  
Sbjct: 390 KTFGDLIP 397

>gb|EAZ30535.1| hypothetical protein OsJ\_14582 [Oryza sativa Japonica Group]  
Length = 610

Score = 197 bits (502), Expect = 3e-49, Method: Composition-based stats.  
Identities = 90/131 (68%), Positives = 108/131 (82%), Gaps = 3/131 (2%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+++DLLLLA GVRFDDRVTGK+EAFASRAKIVH+D  
Sbjct: 293 GNFPDSDPLSLRMLGMHGT VYANYAVDNADLLLA LGVRFDDRVTGKVEAFASRAKIVHVD 352

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENR---AEELKLD FGVWRNELNVQKQKFP 138  
ID +E+GKNK PHVS+C DVKLALQGMN +LE + A LDF WR+EL +K +FP  
Sbjct: 353 IDPSELGKNKQPHVSICADV KLALQGMNAMLEEQSAAAARKNLDFS AWRSELEKKKVEFP 412

Query: 139 LSFKTFGEAVP 149  
L ++TFGE +P  
Sbjct: 413 LGYRTFGEEIP 423

>emb|CAH66432.1| OSIGBa0096P03.6 [Oryza sativa (indica cultivar-group)]  
Length = 663

Score = 197 bits (502), Expect = 3e-49, Method: Composition-based stats.  
Identities = 90/131 (68%), Positives = 108/131 (82%), Gaps = 3/131 (2%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+++DLLLLA GVRFDDRVTGK+EAFASRAKIVH+D  
Sbjct: 324 GNFPDSDPLSLRMLGMHGT VYANYAVDNADLLLA LGVRFDDRVTGKVEAFASRAKIVHVD 383

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENR---AEELKLD FGVWRNELNVQKQKFP 138  
ID +E+GKNK PHVS+C DVKLALQGMN +LE + A LDF WR+EL +K +FP  
Sbjct: 384 IDPSELGKNKQPHVSICADV KLALQGMNAMLEEQSAAAARKNLDFS AWRSELEKKKVEFP 443

Query: 139 LSFKTFGEAVP 149  
L ++TFGE +P  
Sbjct: 444 LGYRTFGEEIP 454

>sp|Q7XKQ8.2|ILVB2\_ORYSJ RecName: Full=Probable acetolactate synthase 2, chloroplastic;  
AltName: Full=Acetohydroxy-acid synthase 2; Flags:  
Precursor  
emb|CAE05539.2| OSJNBa0053B21.13 [Oryza sativa (japonica cultivar-group)]  
gb|EAY93893.1| hypothetical protein OsI\_15666 [Oryza sativa Indica Group]  
Length = 663

Score = 197 bits (502), Expect = 3e-49, Method: Composition-based stats.  
Identities = 90/131 (68%), Positives = 108/131 (82%), Gaps = 3/131 (2%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD RVTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV++D LLLA GVRFD RVTGK+EAFASRAKIVH+D  
Sbjct: 324 GNFP SDDPLSLRMLGMHGT VYANYAVDNAD LLLALGVRFD RVTGKVEAFASRAKIVHVD 383

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENR---AEELKLD FGVWRNELNVQKQKFP 138  
ID +E+GKNK PHVS+C DVKLALQGMN +LE + A LDF WR+EL +K +FP  
Sbjct: 384 IDPSELGKNKQPHVSICADV KLALQGMNAMLEEQSAAAARKNLDFS AWRSELEKKKVEFP 443

Query: 139 LSFKTFGEAVP 149  
L ++TFGE +P  
Sbjct: 444 LGYRTFGEEIP 454

>sp|Q41768.1|ILVB1\_MAIZE RecName: Full=Acetolactate synthase 1, chloroplastic; AltName:  
Full=Acetohydroxy-acid synthase 1; Flags: Precursor  
emb|CAA45116.1| acetohydroxyacid synthase [Zea mays]  
Length = 638

Score = 197 bits (501), Expect = 5e-49, Method: Composition-based stats.  
Identities = 89/127 (70%), Positives = 104/127 (81%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD RVTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +D LLLA GVRFD RVTGK+EAFASRAKIVH+D  
Sbjct: 304 GNFP SDDPLSLRMLGMHGT VYANYAVDKAD LLLALGVRFD RVTGKIEAFASRAKIVHVD 363

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID AEIGKNK PHVS+C DVKLALQGMN +LE + DFG W +EL+ QK++FPL +  
Sbjct: 364 IDPAEIGKNKQPHVSICADV KLALQGMNALLEGSTSKKS FDFGSWNELDQKQKREFPLGY 423

Query: 142 KTFGEAV 148  
KT E +  
Sbjct: 424 KTSNEEI 430

>dbj|BAF37287.1| acetolactate synthase [Monochoria vaginalis]  
Length = 642

Score = 197 bits (500), Expect = 5e-49, Method: Compositional matrix adjust.  
Identities = 91/128 (71%), Positives = 106/128 (82%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD RVTGKLEAFASRAKIVHID 81  
G YP D LSL MLGMHGT VYANYA++ +D LLLA FGVRFD +RVTGKLEAFASRAKIVHID  
Sbjct: 308 GVP L DGPLSLKMLGMHGT VYANYAIDKAD LLLA FGVRFD RVTGKLEAFASRAKIVHID 367

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID AEIGKNK PHVS+CGD+KLALQ MN+++E KLDF WR EL+ QK+ +PL +  
Sbjct: 368 IDPAEIGKNKQPHVSICGDIKLALQEMNEMIEESGIHNKLD FSAWREELDQKKNYPLEY 427

Query: 142 KTFGEAVP 149  
KTFG+ +P  
Sbjct: 428 KTFGDLIP 435

>gb|ABF66049.1| acetolactate synthase [Oryza sativa (indica cultivar-group)]

Length = 644

Score = 196 bits (497), Expect = 1e-48, Method: Composition-based stats.  
Identities = 88/128 (68%), Positives = 104/128 (81%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +D LLA FGVRFDDRVTGK+EAFASRAK +D  
Sbjct: 310 GNFP SDDPLSLRMLGMHGT VYANYAVDKADLLLA FGVRFDDRVTGKIEAFASRAKSLVD 369

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
ID AEIGKNK PHVS+C DVKLALQG+N +L+ + DF W NEL+ QK++FPL +  
Sbjct: 370 IDPAEIGKNKQPHVSICADV KLALQGLNALLDQSTTKTSSDFS AWHNELDQQKREFPLGY 429

Query: 142 KTFGEAVP 149  
KTFGE +P  
Sbjct: 430 KTFGEEIP 437

>sp|Q41769.1|ILVB2\_MAIZE RecName: Full=Acetolactate synthase 2, chloroplastic; AltName:  
Full=Acetohydroxy-acid synthase 2; Flags: Precursor  
emb|CAA45117.1| acetohydroxyacid synthase [Zea mays]  
Length = 638

Score = 194 bits (493), Expect = 3e-48, Method: Compositional matrix adjust.  
Identities = 90/127 (70%), Positives = 104/127 (81%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +D LLA FGVRFDDRVTGK+EAFARA KIVHID  
Sbjct: 304 GNFP SDDPLSLRMLGMHGT VYANYAVDKADLLLA FGVRFDDRVTGKIEAFAGRAKIVHID 363

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
ID AEIGKNK PHVS+C DVKLALQGMN +LE + DFG W +EL+ QK++FPL +  
Sbjct: 364 IDPAEIGKNKQPHVSICADV KLALQGMNTLLEGSTSKKSFDG SWHDEL DQQKREFPLGY 423

Query: 142 KTFGEAV 148  
K F E +  
Sbjct: 424 KIFNEEI 430

>gb|ACF87819.1| unknown [Zea mays]  
Length = 638

Score = 194 bits (493), Expect = 3e-48, Method: Compositional matrix adjust.  
Identities = 90/127 (70%), Positives = 104/127 (81%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +D LLA FGVRFDDRVTGK+EAFARA KIVHID  
Sbjct: 304 GNFP SDDPLSLRMLGMHGT VYANYAVDKADLLLA FGVRFDDRVTGKIEAFAGRAKIVHID 363

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
ID AEIGKNK PHVS+C DVKLALQGMN +LE + DFG W +EL+ QK++FPL +  
Sbjct: 364 IDPAEIGKNKQPHVSICADV KLALQGMNTLLEGSTSKKSFDG SWHDEL DQQKREFPLGY 423

Query: 142 KTFGEAV 148  
K F E +  
Sbjct: 424 KIFNEEI 430

>gb|EEE60896.1| hypothetical protein OsJ\_14580 [Oryza sativa Japonica Group]  
Length = 547

Score = 194 bits (492), Expect = 4e-48, Method: Compositional matrix adjust.  
Identities = 89/128 (69%), Positives = 107/128 (83%), Gaps = 1/128 (0%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRVTGKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+++D LLA GVRFDDRVTGK+EAFASRAKIVH+D  
Sbjct: 213 GNFP SDDPLSLRMLGMHGT VYANYAVDNADLLLA LGVRFDDRVTGKVEAFASRAKIVHVD 272

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141

ID +E+GKNK PHVS+C DVKLALQGMN LE + + LDF WR+EL +K +FPL +  
Sbjct: 273 IDPSELGKNKQPHVSICADVKLALQGMNATLEQQQRK-NLDFSARWSELEKKKAEFPLGY 331

Query: 142 KTFGEAVP 149  
+TFGE +P

Sbjct: 332 RTFGEEIP 339

>gb|ACG44246.1| acetolactate synthase III [Zea mays]  
Length = 547

Score = 193 bits (491), Expect = 6e-48, Method: Compositional matrix adjust.  
Identities = 89/127 (70%), Positives = 104/127 (81%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G++P DD LSL MLGMHGT VYANYAV+ +DLLLL GVRFDDRVTK+EAFASRAKIVH+D  
Sbjct: 213 GNFPDPLSLRMLGMHGT VYANYAVDKADLLLLALGVRFDDRVTKIEAFASRAKIVHVD 272

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPPLSF 141  
ID AEIGKNK PHVS+C DVKLALQGMN +LE + DFG W +EL+ QK++FPL +  
Sbjct: 273 IDPAEIGKNKQPHVSICADVKLALQGMNALLEGSTSKSFDFGSWNDELDDQKREFPLGY 332

Query: 142 KTFGEAV 148  
KT E +  
Sbjct: 333 KTSNEEI 339

>emb|CAO49020.1| unnamed protein product [Vitis vinifera]  
Length = 657

Score = 193 bits (490), Expect = 7e-48, Method: Compositional matrix adjust.  
Identities = 93/128 (72%), Positives = 104/128 (81%), Gaps = 7/128 (5%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G++PC D+LSL MLGMHGT+YANYAV+ SDLLAFGVRFDDRVTK+EAFASRAKIVHID  
Sbjct: 330 GAFPCTDKLSLQMLGMHGTMYANYAVDRSDLLAFGVRFDDRVTKIEAFASRAKIVHID 389

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPPLSF 141  
ID AEIGKNK PHVS AL GMN +LE + K DF WR+ELN QK K+PLSF  
Sbjct: 390 IDPAEIGKNKQPHVS-----ALTGMNSLLEETGAKSKYDFSSWRDELNEQKAKYPLSF 442

Query: 142 KTFGEAVP 149  
KTFGEA+P  
Sbjct: 443 KTFGEAIP 450

>gb|ABR16079.1| unknown [Picea sitchensis]  
Length = 669

Score = 179 bits (454), Expect = 1e-43, Method: Compositional matrix adjust.  
Identities = 86/128 (67%), Positives = 104/128 (81%), Gaps = 2/128 (1%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
GS+P E SL MLGMHGT VYANYAV+ +DLLLLAFGVRFDDRVTKLE+FASRA IVHID  
Sbjct: 337 GSFPSSEKSLGMLGMHGT VYANYAVDKADLLLLAFGVRFDDRVTKLESFASRASIVHID 396

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPPLSF 141  
ID AEIGKNK PH+S+C DV+LALQG+N +++ R E + DF W +EL QK K+PL++  
Sbjct: 397 IDPAEIGKNKQPHISLCADVQLALQGLNYLMKKR--EDRPDFYAWIDELEEQKTKWPLTY 454

Query: 142 KTFGEAVP 149  
+ F EA+P  
Sbjct: 455 QNFEEAIP 462

>emb|CAN73760.1| hypothetical protein [Vitis vinifera]  
Length = 586

Score = 176 bits (446), Expect = 1e-42, Method: Compositional matrix adjust.

Identities = 85/141 (60%), Positives = 103/141 (73%), Gaps = 13/141 (9%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G + C D+LSLHMLGMHGT+ ANYAV+ SDLLLAFGVRFDDRVTKG+EAF A IVHID  
Sbjct: 244 GIFACTDDLSLHMLGMHGTIQANYAVDRSDLLLAFGVRFDDRVTKVEAFARNATIVHID 303

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKL-----DFGVWRN 128  
ID AEIGKNK PH+S+C DVKLAL+G+N +LE A + + W  
Sbjct: 304 IDPAEIGKNKPHLSICTDVKLALLEGINTILEKNAAKQPTAENKRKGKTKFNDNVSAWIE 363

Query: 129 ELNVQKQKFPLSFKTFGEAVP 149  
E++ QK+K+P S+KTFGEA+P  
Sbjct: 364 EIDEQKEKYPASYKTFGEAIP 384

>gb|ABR26179.1| acetolactate synthase III [Oryza sativa (indica cultivar-group)]  
Length = 151

Score = 172 bits (437), Expect = 1e-41, Method: Compositional matrix adjust.  
Identities = 79/110 (71%), Positives = 92/110 (83%)

Query: 40 TVYANYAVEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGKNKTPHVSVC 99  
TVYANYAV+ +DLLLAFGVRFDDRVTKG+EAFASRAKIVHIDID AEIGKNK PHVS+C  
Sbjct: 6 TVYANYAVDKADLLLAFGVRFDDRVTKIEAFASRAKIVHIDIDPAEIGKNKQPHVSICA 65

Query: 100 DVKLALQGMNKVLENRAEELKLDFGVWRNELNVQKQKFPLSFKTFGEAVP 149  
DVKLALQG+N +L+ + DF W NEL+ QK++FPL +KTFGE +P  
Sbjct: 66 DVKLALQGLNALLDQSTTKTSSDFSAAHNELDQQKREFPLGYKTFGEIP 115

>ref|XP\_001759950.1| predicted protein [Physcomitrella patens subsp. patens]  
gb|EDQ75075.1| predicted protein [Physcomitrella patens subsp. patens]  
Length = 583

Score = 171 bits (434), Expect = 2e-41, Method: Compositional matrix adjust.  
Identities = 81/130 (62%), Positives = 101/130 (77%), Gaps = 1/130 (0%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G++P D+ L MLGMHGT VYANYAV+++D+LLAFGVRFDDRVTKLE+FA SRA IVHID  
Sbjct: 249 GTFPASDKRYLSMLGMHGT VYANYAVDNADMLLAFGVRFDDRVTKLESFASRASIVHID 308

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLDFGVWRNELNVQKQKFPLSF 141  
ID AEIGKNK PH+S+C DV+LAL G+NK+++ E + DF WR EL+ K K+P+ F  
Sbjct: 309 IDPAEIGKNKQPHISICADVQLALAGLNKLIKEGPAE-RPDFSAWRAELDGVKWKWPMKF 367

Query: 142 KTFGEAVPCP 151  
F +V P  
Sbjct: 368 PKFDNSVIVP 377

>ref|XP\_001758473.1| predicted protein [Physcomitrella patens subsp. patens]  
gb|EDQ76875.1| predicted protein [Physcomitrella patens subsp. patens]  
Length = 618

Score = 169 bits (429), Expect = 9e-41, Method: Compositional matrix adjust.  
Identities = 79/130 (60%), Positives = 98/130 (75%), Gaps = 1/130 (0%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G++P DE L MLGMHGT VYANYA+++SD+LLAFGVRFDDRVTKLE+FA SRA IVHID  
Sbjct: 284 GTFPASDERYLSMLGMHGT VYANYAIDNSDMLLAFGVRFDDRVTKLESFASRASIVHID 343

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLDFGVWRNELNVQKQKFPLSF 141  
ID AEIGKNK PH+S+C DV+LAL G+NK+++ + F WR EL+ K+K+P+ F  
Sbjct: 344 IDPAEIGKNKQPHISICADVQLALAGLNKLIKE-GPATRPSFSAWRKELDGVKEKWPMPKF 402

Query: 142 KTFGEAVPCP 151  
V P  
Sbjct: 403 PKLDSQVIVP 412

>dbj|BAE53595.1| acetolactate synthase [Monochoria vaginalis]  
Length = 566

Score = 163 bits (413), Expect = 6e-39, Method: Compositional matrix adjust.  
Identities = 78/101 (77%), Positives = 86/101 (85%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G YP D LSL MLGMHGT VYANYA++ +DLLLAFGVRFDDRVTKLEAFASRAKIVHID  
Sbjct: 232 GVYPLDGPLSLKMLGMHGT VYANYAIDKADLLLAFGVRFDDRVTKLEAFASRAKIVHID 291

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD 122  
ID AEIGKNK PHVS+CGD+KLALQ MN+++E KLD  
Sbjct: 292 IDPAEIGKNKQPHVSICGDIKLALQEMNEMIEESGIHNKLD 332

>emb|CA017032.1| unnamed protein product [Vitis vinifera]  
Length = 866

Score = 162 bits (411), Expect = 1e-38, Method: Compositional matrix adjust.  
Identities = 81/128 (63%), Positives = 95/128 (74%), Gaps = 2/128 (1%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G +PC D LSLHMLGMHGT+ ANYAV+ SDLLLAFGVRFDDRVTK+EAF A IVHID  
Sbjct: 539 GIFPCTDGLSLHMLGMHGTIQANYAVDRSDLLLAFGVRFDDRVTKVEAFARNATIVHID 598

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPPLSF 141  
ID AEIGKNK PH+S+C DVKLAL+G+N +LE A K + + P S+  
Sbjct: 599 IDPAEIGKNKKPHLSICTDVKLALLEGINTILEKNAA--KQPTAENKRGKGTKFNDNPPSY 656

Query: 142 KTFGEAVP 149  
+TFGEA+P  
Sbjct: 657 RTFGEAIP 664

>emb|CA045071.1| unnamed protein product [Vitis vinifera]  
Length = 463

Score = 160 bits (406), Expect = 4e-38, Method: Compositional matrix adjust.  
Identities = 80/128 (62%), Positives = 95/128 (74%), Gaps = 2/128 (1%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G + C D+LSLHMLGMHGT+ ANYAV+ SDLLLAFGVRFDDRVTK+EAF A IVHID  
Sbjct: 136 GIFACTDDLSLHMLGMHGTIQANYAVDRSDLLLAFGVRFDDRVTKVEAFARNATIVHID 195

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPPLSF 141  
ID AEIGKNK PH+S+C DVKLAL+G+N +LE A K + + P S+  
Sbjct: 196 IDPAEIGKNKKPHLSICTDVKLALLEGINTILEKNAA--KQPTAENKRGKGTKFNDNPPSY 253

Query: 142 KTFGEAVP 149  
+TFGEA+P  
Sbjct: 254 RTFGEAIP 261

>gb|AAC04854.1| acetolactate synthase [Volvox carteri]  
Length = 681

Score = 158 bits (399), Expect = 3e-37, Method: Composition-based stats.  
Identities = 77/131 (58%), Positives = 96/131 (73%), Gaps = 2/131 (1%)

Query: 18 FQHEGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTKLEAFASRAKI 77  
F G P +D L MLGMHGT V ANYAV+ +DLL+A GVRFDDRVTK+L+AFASRA+I  
Sbjct: 332 FMGLGVVPAEDPNHLQMLGMHGTVAANYAVDQADLLVALGVRFDDRVTKGRDFAFASRAI 391

Query: 78 VHIDIDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKF 137  
VH+DID+AEI KNKT HV VCGDVK AL+ +N++LE AE L F WR EL ++ +F  
Sbjct: 392 VHVDIDAAEISKNTAHVPVCGDVKQALRHLNRMLE--AEPLSDRFVAWRAELAAKRAEF 449

Query: 138 PLSFKTFGEAV 148

PL + +A+  
Sbjct: 450 PLRYPQRDDAI 460

>gb|AAB88296.1| acetolactate synthase [Volvox carteri]  
Length = 681

Score = 158 bits (399), Expect = 3e-37, Method: Composition-based stats.  
Identities = 77/131 (58%), Positives = 96/131 (73%), Gaps = 2/131 (1%)

Query: 18 FQHEGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKI 77  
F G P +D L MLGMHGT V ANYAV+ +DLL+A GVRFDDRV TG+L+AFASRA+I  
Sbjct: 332 FMGLGVVPAEDPNHLQMLGMHGTVAANYAVDQADLLVALGVRFDDRV TGRLDAFASRARI 391

Query: 78 VHIDIDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKF 137  
VH+DID+AEI KNKT HV VCGDVK AL+ +N++LE AE L F WR EL ++ +F  
Sbjct: 392 VHVIDIDAAEISKNTAHVPVCGDVKQALRHLNRMLE--AEPLSDRFVAVRAELAAKRAEF 449

Query: 138 PLSFKTFGEAV 148  
PL + +A+  
Sbjct: 450 PLRYPQRDDAI 460

>ref|XP\_001695168.1| acetolactate synthase, large subunit [Chlamydomonas reinhardtii]  
gb|EDP01876.1| acetolactate synthase, large subunit [Chlamydomonas reinhardtii]  
Length = 640

Score = 152 bits (385), Expect = 1e-35, Method: Composition-based stats.  
Identities = 76/132 (57%), Positives = 95/132 (71%), Gaps = 3/132 (2%)

Query: 18 FQHEGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKI 77  
F G P D L MLGMHGT V+ANYAV+ +DLL+A GVRFDDRV TGKL+AFA+RA+I  
Sbjct: 290 FMGLGVVPSTDPNHLQMLGMHGT V FANYAVDQADLLVALGVRFDDRV TGKLDFAFAARARI 349

Query: 78 VHIDIDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD -FGVWRNELNVQKQK 136  
VHIDID+AEI KNKT HV VCGDVK AL +N++L AE L D + WR EL ++ +  
Sbjct: 350 VHVIDIDAAEISKNTAHVPVCGDVKQALSHLNRLLA--AEPLPADKWAGWRAELAAKRAE 407

Query: 137 FPLSFKTFGEAV 148  
FP+ + +A+  
Sbjct: 408 FPMRYPQRDDAI 419

>gb|AAB88292.1| acetolactate synthase [Chlamydomonas reinhardtii]  
Length = 683

Score = 152 bits (385), Expect = 1e-35, Method: Composition-based stats.  
Identities = 76/132 (57%), Positives = 95/132 (71%), Gaps = 3/132 (2%)

Query: 18 FQHEGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKI 77  
F G P D L MLGMHGT V+ANYAV+ +DLL+A GVRFDDRV TGKL+AFA+RA+I  
Sbjct: 333 FMGLGVVPSTDPNHLQMLGMHGT V FANYAVDQADLLVALGVRFDDRV TGKLDFAFAARARI 392

Query: 78 VHIDIDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD -FGVWRNELNVQKQK 136  
VHIDID+AEI KNKT HV VCGDVK AL +N++L AE L D + WR EL ++ +  
Sbjct: 393 VHVIDIDAAEISKNTAHVPVCGDVKQALSHLNRLLA--AEPLPADKWAGWRAELAAKRAE 450

Query: 137 FPLSFKTFGEAV 148  
FP+ + +A+  
Sbjct: 451 FPMRYPQRDDAI 462

>gb|AAC03784.1| acetolactate synthase [Chlamydomonas reinhardtii]  
Length = 683

Score = 152 bits (385), Expect = 1e-35, Method: Composition-based stats.  
Identities = 76/132 (57%), Positives = 95/132 (71%), Gaps = 3/132 (2%)

Query: 18 FQHEGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKI 77

F G P D L MLGMHGTV+ANYAV+ +DLL+A GVRFDDRVTGKL+AFA+RA+I  
 Sbjct: 333 FMGLGVVPSTDPNHLQMLGMHGTVFANYAVDQADLLVALGVRFDDRVTGKLDFAAARARI 392  
 Query: 78 VHIDIDSAEIGKNKTPHVSVCQDVKLALQGMNKVLENRAEELKLD-FGVWRNELNVQKQK 136  
 VHIDID+AEI KNKT HV VCGDVK AL +N++L AE L D + WR EL ++ +  
 Sbjct: 393 VHIDIDAAEISKNTAHVPVCGDVKQALSHLNRLLA--AEPLPADKWAGWRAELAARAE 450  
 Query: 137 FPLSFKTFGEAV 148  
 FP+ + +A+  
 Sbjct: 451 FPMRYPQRDDAI 462

>emb|CAL58226.1| acetolactate synthase 1 (ISS) [Ostreococcus tauri]  
 Length = 679  
 Score = 146 bits (369), Expect = 9e-34, Method: Compositional matrix adjust.  
 Identities = 70/117 (59%), Positives = 89/117 (76%), Gaps = 1/117 (0%)

Query: 22 GSYPCDDELSLHMLGMHGTVYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
 G++P D+LSL MLGMHG VYANYA++ +DLLAFGVRFDDRVTGKL FA RA IVHID  
 Sbjct: 331 GTFPQSDLLSLMLGMHGA VYANYAIDSADLLAFGVRFDDRVTGKLEAFKRAAIVHID 390  
 Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP 138  
 ID AE+GKNK P S+ ++K AL+ +NK++ A L DF WR E+ V++++FP  
 Sbjct: 391 IDPAELGKNKKPFCSIFSNIKPALRSLNKLIALDAANLP-DFSDWRAEIEVKRKEFP 446

>ref|XP\_001421626.1| predicted protein [Ostreococcus lucimarinus CCE9901]  
 gb|ABO99919.1| predicted protein [Ostreococcus lucimarinus CCE9901]  
 Length = 610  
 Score = 144 bits (364), Expect = 3e-33, Method: Compositional matrix adjust.  
 Identities = 70/117 (59%), Positives = 89/117 (76%), Gaps = 2/117 (1%)

Query: 22 GSYPCDDELSLHMLGMHGTVYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
 G++P D+LSL MLGMHG VYANYA++ +DLLAFGVRFDDRVTGKL FA RA IVHID  
 Sbjct: 263 GTFPQSDLLSLMLGMHGA VYANYAIDSADLLAFGVRFDDRVTGKLEAFKRAAIVHID 322  
 Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP 138  
 ID AE+GKNK P S+ ++K AL+ +NK++ E+ DF WR ++VQ++KFP  
 Sbjct: 323 IDPAELGKNKKPFCSIFSNIKPAKLTNLKLA--LEKDMPDFSEWRAAIDVQRKKFP 377

>emb|CA042659.1| unnamed protein product [Vitis vinifera]  
 Length = 616  
 Score = 135 bits (340), Expect = 2e-30, Method: Compositional matrix adjust.  
 Identities = 71/128 (55%), Positives = 84/128 (65%), Gaps = 27/128 (21%)

Query: 22 GSYPCDDELSLHMLGMHGTVYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
 G +PC D+LSLHMLGMHGT+ ANYAV+ SDLLAFGVRFDDRVTGK+E+FA A IVHID  
 Sbjct: 314 GIFPCTDLSLHMLGMHGTIQANYAVDRSDLLAFGVRFDDRVTGKVESFARNATIVHID 373  
 Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPPLSF 141  
 ID AEIGKNK PH+S+C D A EN K+ ++  
 Sbjct: 374 IDPAEIGKNKKPHLSICTDQPTA-----EN-----KRGKGTNY 406

Query: 142 KTFGEAVP 149  
 +TFGEA+P  
 Sbjct: 407 RTFGEAIP 414

>ref|NP\_682086.1| acetolactate synthase 3 catalytic subunit [Thermosynechococcus  
 elongatus BP-1]  
 dbj|BAC08848.1| acetohydroxy acid synthase [Thermosynechococcus elongatus BP-1]  
 Length = 602  
 Score = 135 bits (339), Expect = 2e-30, Method: Composition-based stats.  
 Identities = 65/132 (49%), Positives = 86/132 (65%), Gaps = 7/132 (5%)



Query: 21 EGSYPCDELHSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHI 80  
+G++P LS+ MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+ FASRAK++HI  
Sbjct: 246 KGAFPEHPLSVGMLGMHGTAYANFAVSECDLLIAGGARFDDRVTKGLDEFASRAKVIHI 305

Query: 81 DIDS AEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLS 140  
DID AE+GKN+ P V + G VK LQ + K +E+ + ++ W + + KQ +PL  
Sbjct: 306 DIDPAEVGKNRVPDVPVIGSVKPVLLQQLKHIEDSGQAVEPQTQAWLERITIWQDYPL- 364

Query: 141 FKTFGEAVPCPS 152  
VP PS  
Sbjct: 365 -----IVPQPS 370

>ref|YP\_001941018.1| Acetolactate synthase large subunit or other thiamine  
pyrophosphate-requiring enzyme [Methylokorus infernorum  
V4]  
gb|ACD84421.1| Acetolactate synthase large subunit or other thiamine  
pyrophosphate-requiring enzyme [Methylacidiphilum  
infernorum V4]  
Length = 589

Score = 134 bits (338), Expect = 4e-30, Method: Compositional matrix adjust.  
Identities = 66/130 (50%), Positives = 94/130 (72%), Gaps = 2/130 (1%)

Query: 22 GSYPCDELHSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHI 81  
GS+P + LSL LGMHG+VY+N+AV+ SDLLAFGVRFDDRVTK ++AFA +A+IVHI  
Sbjct: 261 GSFPENHPLSLKWLGMHGSVYSNFAVDQSDLLAFGVRFDDRVTKGNVQAFQAKARIVHI 320

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
ID++E+ KNK + + GD+K AL+ +N+++E R + + +W+ E+ K K+P +  
Sbjct: 321 IDNSELNKNKRVDPILGDIKEALRLNRLIEERKFQ-SPGYSLWQEEIQAWKIKYPFRY 379

Query: 142 KTFGEAVPCP 151  
K F E V P  
Sbjct: 380 KKF-EGVIMP 388

>ref|ZP\_01857674.1| acetolactate synthase III (Precursor) [Planctomyces maris DSM 8797]  
gb|EDL56438.1| acetolactate synthase III (Precursor) [Planctomyces maris DSM 8797]  
Length = 593

Score = 129 bits (323), Expect = 2e-28, Method: Compositional matrix adjust.  
Identities = 65/127 (51%), Positives = 85/127 (66%), Gaps = 3/127 (2%)

Query: 22 GSYPCDELHSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHI 81  
G +P +D LSL MLGMHGT VYANYAV +DLLAFGVRFDDRVTKLE FA KIVH+D  
Sbjct: 257 GVFPGEDPLSLDMLGMHGT VYANYAVNEADLLAFGVRFDDRVTKLEEFKAKHGKIVHVD 316

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
ID +E+ KNK H+ + D+K L +N+ + + ++L W + K+KFPL +  
Sbjct: 317 IDPSELQKNKEAHIPINADLKHVLTLENEATD--DDL P-QVDSWLAQCKEWEKFKPLKY 373

Query: 142 KTFGEAV 148  
G+ +  
Sbjct: 374 PELGDVM 380

>ref|YP\_001930430.1| acetolactate synthase, large subunit, biosynthetic type  
[Sulfurihydrogenibium sp. YO3AOP1]  
gb|ACD65876.1| acetolactate synthase, large subunit, biosynthetic type  
[Sulfurihydrogenibium sp. YO3AOP1]  
Length = 582

Score = 128 bits (321), Expect = 3e-28, Method: Compositional matrix adjust.  
Identities = 65/122 (53%), Positives = 81/122 (66%)

Query: 21 EGSYPCDELHSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHI 80  
+G++P D LSLHMLGMHGT YAN AV HSDLL+A G RFDDRVTKG+ FA AKI+HI

Sbjct: 246 KGAFPETDPLSLHMLGMHGTYYANMAVYHSDLLIAVGARFDDRVTGKINEFAPEAKIIHI 305

Query: 81 DIDS AEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLS 140  
DID A I K T V + GDVK L+ + K LE + E W ++N K+K PL+

Sbjct: 306 DIDPASISKITITVDVPIVG DVKNVLRKLIKELEEKPIEWIAAREQWLKQINNEWKEKHPLN 365

Query: 141 FK 142

++

Sbjct: 366 YR 367

>ref|YP\_502706.1| acetolactate synthase, large subunit, biosynthetic type  
[Methanospirillum hungatei JF-1]  
gb|ABD40987.1| acetolactate synthase, large subunit [Methanospirillum hungatei  
JF-1]  
Length = 563

Score = 127 bits (319), Expect = 5e-28, Method: Compositional matrix adjust.  
Identities = 67/137 (48%), Positives = 86/137 (62%), Gaps = 5/137 (3%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHI 81  
G P D L+L MLGMHGT YANYA+ SDLL + GVRFDDRVTGK+EAFA +AKI+HID  
Sbjct: 242 GCIPTDHPNLGMLGMHGT EYANYAITESDLLFSIGVRFDDRVTGKIEAFAPQAKIIHI 301

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
ID AEIGKNK P V + GD K L + K + + W ++ + K+ PL +  
Sbjct: 302 IDPAEIGKNKQDPVPIVGDAKSVLS DILKKIPESCPHPE-----WSEKVMWKKNHPLRY 356

Query: 142 KTFGEAVPCPSKKT VHD 158

++ G+ P KT+ D

Sbjct: 357 RSDGKLYPQFVIKTLAD 373

>ref|NP\_926225.1| acetolactate synthase 3 catalytic subunit [Gloeobacter violaceus  
PCC 7421]  
dbj|BAC91220.1| acetohydroxyacid synthetase large subunit [Gloeobacter violaceus  
PCC 7421]  
Length = 596

Score = 126 bits (316), Expect = 1e-27, Method: Compositional matrix adjust.  
Identities = 62/125 (49%), Positives = 83/125 (66%), Gaps = 4/125 (3%)

Query: 21 EGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHI 80  
+G++P D L++ MLGMHGT YAN+AV DLL+A G RFDDRVTGKL+ FA+RA+++HI  
Sbjct: 245 KGAFPEDHPLAVGMLGMHGTAYANFAVTECDLLIAVGARFDDRVTGKLDEFAARARVIHI 304

Query: 81 DIDS AEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLS 140  
DID AE+GKN+ P V + GDVKL LQ + K L E + W + K ++PL  
Sbjct: 305 DIDPAEIVGKNRGPEVPIVG DVKLVLDLIKKL----IEFQPRTRAWLERIERWKAIEYPLE 360

Query: 141 FKTFG 145

+ G

Sbjct: 361 VPSDG 365

>ref|ZP\_01089063.1| acetolactate synthase III (Precursor) [Blastopirellula marina DSM  
3645]  
gb|EAQ82178.1| acetolactate synthase III (Precursor) [Blastopirellula marina DSM  
3645]  
Length = 590

Score = 125 bits (314), Expect = 2e-27, Method: Compositional matrix adjust.  
Identities = 66/127 (51%), Positives = 82/127 (64%), Gaps = 5/127 (3%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHI 81  
GS+P DD LSL MLGMHGT VYANYAV++ DLL++ GVRFDDRVTGKL FA AKI+HID  
Sbjct: 256 GSFPGDDPLSLDMLGMHGT VYANYAVQNC DLLISLGVRFDDRVTGKLAFAKHAKIIHI 315

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141

ID +EI KNK H V D+K AL+ + ++E D W + K+ PL +  
 Sbjct: 316 IDPSEINKKLAHFPVVSCLKPALRMLCDIVEKPE-----DISEWVAQCAKWKDEPLKY 370

Query: 142 KTFGEAV 148  
 T E +

Sbjct: 371 DTSFEGI 377

>ref|ZP\_02963213.1| acetolactate synthase 1 catalytic subunit [Bifidobacterium animalis  
 subsp. lactis HN019]  
 ref|YP\_002469142.1| acetolactate synthase, large subunit, biosynthetic type  
 [Bifidobacterium animalis subsp. lactis AD011]  
 gb|EDT89371.1| acetolactate synthase 1 catalytic subunit [Bifidobacterium animalis  
 subsp. lactis HN019]  
 gb|ACL28566.1| acetolactate synthase, large subunit, biosynthetic type  
 [Bifidobacterium animalis subsp. lactis AD011]  
 Length = 642

Score = 125 bits (314), Expect = 2e-27, Method: Composition-based stats.  
 Identities = 63/122 (51%), Positives = 80/122 (65%), Gaps = 1/122 (0%)

Query: 22 GSYPCDDESLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
 G P DD L MLGMHGT+ A AV+ SDLL+A G RFDDRVTKLEAFA A+++HID  
 Sbjct: 271 GIVPDDDDPAVLGMLGMHGTIAATGAVQRSDDLVAIGARFDDRVTKLEAFAPGARVIHID 330

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQG-MNKVLENRAEELKLD FGVWRNELNVQKQKFPLS 140  
 ID AEIGKN+TP V + GDVK L M ++ A K + W +N +K+PL+  
 Sbjct: 331 IDPAEIGKNRTPDVPIVGDVKT VLSALMPRIERQHAIHGKPNLTTWGLINSWVEKYPLT 390

Query: 141 FK 142  
 ++  
 Sbjct: 391 YQ 392

>ref|YP\_001111655.1| acetolactate synthase, large subunit, biosynthetic type  
 [Desulfotomaculum reducens MI-1]  
 gb|ABO48830.1| acetolactate synthase, large subunit [Desulfotomaculum reducens  
 MI-1]  
 Length = 554

Score = 125 bits (314), Expect = 2e-27, Method: Compositional matrix adjust.  
 Identities = 63/127 (49%), Positives = 86/127 (67%), Gaps = 6/127 (4%)

Query: 22 GSYPCDDESLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
 G +P D L+L MLGMHGT YAN+A+ DLL+A GVRFDDRVTKG+E+FA AKI+HID  
 Sbjct: 242 GGFPGDHALALGMLGMHGT K YANFAICDCDLLIAVGVRFDDRVTKVESFAPNAKIIHID 301

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
 ID AE+GKN V + G+VKL LQ + + ++ R E W++++ K+++PL F  
 Sbjct: 302 IDPAELGKNVRVDVPIAGNVKLV LQQLLERVQTRIFE-----AWQHKVAQWKKEYPLDF 355

Query: 142 KTFGEAV 148  
 GE +  
 Sbjct: 356 DENG E GL 362

>ref|YP\_002249089.1| acetolactate synthase, large subunit, biosynthetic type  
 [Thermodesulfovibrio yellowstonii DSM 11347]  
 gb|ACI21111.1| acetolactate synthase, large subunit, biosynthetic type  
 [Thermodesulfovibrio yellowstonii DSM 11347]  
 Length = 576

Score = 125 bits (314), Expect = 2e-27, Method: Compositional matrix adjust.  
 Identities = 68/134 (50%), Positives = 90/134 (67%), Gaps = 8/134 (5%)

Query: 22 GSYPCDDESLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
 GS+P ELSL MLGMHGT YAN AV++SDL++A G+RFDDRVTKG +AFA AKI+HID  
 Sbjct: 243 GSFPGTHELSLGMLGMHGTYYANMAVQNSDLIVAIGMRFDDRVTKGTD A FAPNAKIIHID 302

Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKLLENRAEELKLDLDF----GVWRNELNVQKQKF 137  
ID I KN + + GDV LQ +NK+L+ EE+K + W ++N K++  
Sbjct: 303 IDPTSIRKNVRVDIPVIGDVSRVLQVLNKLK---EEVKPQWKEIRQAWLKQINQWKKER 359

Query: 138 PLSFKTFGEAVPCP 151  
PL+++ F EAV P  
Sbjct: 360 PLTYE-FDEAVIKP 372

>ref|ZP\_01731805.1| acetolactate synthase III large subunit [Cyanotheca sp. CCY0110]  
gb|EAS88774.1| acetolactate synthase III large subunit [Cyanotheca sp. CCY0110]  
Length = 619

Score = 125 bits (313), Expect = 3e-27, Method: Composition-based stats.  
Identities = 61/132 (46%), Positives = 85/132 (64%), Gaps = 9/132 (6%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHID 81  
GS+ LS+ MLGMHGT YAN+AV DLL+A G RFDDRTVGKLE FASRA+++HID  
Sbjct: 263 GSFDEHHPISVGMMLGMHGTAYANFAVSECDLLIAVGSRFDDRTVGKLEEFASRARVIHID 322

Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKLLENRAEELKLDLDFGV-----WRNELNVQKQK 136  
ID AE+GKN+ P V + GDV+ L+ L R+ E+ + W N +N +++  
Sbjct: 323 IDPAEVGNRLPEVPIVGDVRQVLEQ---LLQRSREIDVPLDAEKTTPWLNRIINRWREQ 378

Query: 137 FPLSFKTFGEAV 148  
+PL+ + ++  
Sbjct: 379 YPLTAPHYDHSI 390

>gb|EDX84416.1| acetolactate synthase, large subunit, biosynthetic type  
[Synechococcus sp. PCC 7335]  
Length = 589

Score = 124 bits (312), Expect = 4e-27, Method: Compositional matrix adjust.  
Identities = 60/118 (50%), Positives = 79/118 (66%), Gaps = 6/118 (5%)

Query: 30 LSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHIDIDSAEIGK 89  
L++ MLGMHGT YAN+AV DLL+A G RFDDRTVGKL+ FAS AK++HIDID AE+GK  
Sbjct: 270 LAVGMLGMHGTAYANFAVSECDLLVAIGARFDDRTVGKLDEFASSAKVIHIDIDPAEVGK 329

Query: 90 NKTTPHVSVCQDVKLALQGMNKLLENRAEELKLDLDF---VWRNELNVQKQKFPPLSFKTF 144  
N+ P V + GDVK L ++L+ + + LD WR+++ KQ FPL +  
Sbjct: 330 NRVPVPIVGDVKAVLA---RMLDRLSGQAALDVNRSQSWRDQIETWKQDFPLVAPVY 384

>ref|ZP\_02737021.1| acetolactate synthase III [Gemmata obscuriglobus UQM 2246]  
Length = 590

Score = 124 bits (311), Expect = 5e-27, Method: Compositional matrix adjust.  
Identities = 65/123 (52%), Positives = 82/123 (66%), Gaps = 3/123 (2%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRTVGKLEAFASRAKIVHID 81  
G++P D L L MLGMHGT VY+NYA+ +DLLAFGVRFDDRTVGKL FA KIVH+D  
Sbjct: 256 GNFPADHYLCLQMLGMHGT VYSNYAINDADLLAFGVRFDDRTVGKLAFAKHGKIVHVD 315

Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKLLENRAEELKLDLDFGVWRNELNVQKQKFP 138  
ID +EI KNK HV V D+K AL G+N +L E + +L + W +++ + P  
Sbjct: 316 IDKSEIHKNKFAHVPVHSDLKHHLNALLKEKNADLTAGGRYTDWWRQVDWARDAP 375

Query: 139 LSF 141  
L F  
Sbjct: 376 LKF 378

>ref|NP\_897837.1| acetolactate synthase 3 catalytic subunit [Synechococcus sp. WH  
8102]  
sp|Q7U5G1.1|ILVB\_SYNPX RecName: Full=Acetolactate synthase large subunit; Short=AHAS;  
AltName: Full=Acetohydroxy-acid synthase large subunit;  
Short=ALS

emb|CAE08261.1| acetolactate synthase [Synechococcus sp. WH 8102]  
Length = 617

Score = 124 bits (310), Expect = 6e-27, Method: Composition-based stats.  
Identities = 59/120 (49%), Positives = 81/120 (67%), Gaps = 2/120 (1%)

Query: 21 EGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTGKLEAFASRAKIVHI 80  
+G++ +D LS+ MLGMHGT YAN+AV DLL+A G RFDDRVTGKL+ FA RA++VH  
Sbjct: 264 KGAFDENDALSVGMLGMHGTAYANFAVTECDLLIAVGARFDDRVTGKLDTFAPRARVVHF 323

Query: 81 DIDS AEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLS 140  
+ID AEIGKN+ V+V GD+ L+L M ++ R E + W +N K ++PL+  
Sbjct: 324 EIDPAEIGKNRKADVAVLGDLGLSLARMVEISLQRTAEPRT--AAWLERINTWKDRYPLT 381

>ref|YP\_001805892.1| acetolactate synthase [Cyanothecae sp. ATCC 51142]  
gb|ACB53826.1| acetolactate synthase [Cyanothecae sp. ATCC 51142]  
Length = 619

Score = 123 bits (309), Expect = 8e-27, Method: Composition-based stats.  
Identities = 60/123 (48%), Positives = 81/123 (65%), Gaps = 9/123 (7%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTGKLEAFASRAKIVHI 81  
GS+ LS+ MLGMHGT YAN+AV DLL+A G RFDDRVTGKL+ FASRA+++HID  
Sbjct: 263 GSFDEHHP LSVGMLGMHGTAYANFAVSECDLLIAVGARFDDRVTGKLDEFASRARVIHID 322

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGV-----WRNELNVQKQK 136  
ID AE+GKN+ P V + GDV+ L+ L R+ E+ + W N +N +++  
Sbjct: 323 IDPAEVGKNRLPEVPIVGDVRQVLEQ----LLQRSREIDVPLDAEKTTPWLNRIINRWREQ 378

Query: 137 FPL 139  
+PL  
Sbjct: 379 YPL 381

>ref|YP\_001046985.1| acetolactate synthase, large subunit, biosynthetic type  
[Methanoculleus marisnigri JR1]  
gb|ABN57003.1| acetolactate synthase, large subunit [Methanoculleus marisnigri  
JR1]  
Length = 558

Score = 123 bits (308), Expect = 9e-27, Method: Compositional matrix adjust.  
Identities = 61/121 (50%), Positives = 75/121 (61%), Gaps = 5/121 (4%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTGKLEAFASRAKIVHI 81  
G+ P D L L MLGMHGT ANYAV DLL+A GVRFDDRVTGK+E FA A ++HID  
Sbjct: 242 GAVPGDHPLCLGMLGMHGTQSANYAVTECDLLVAVGVRFDDRVTGKIETFAPNAAVIHID 301

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLS 141  
ID AEIGKNK V + GDVK LQ ++ R D W + + K ++PL +  
Sbjct: 302 IDPAEIGKNKKVDVPIVGDVKAVLQAFLSRMQKRG-----DTAAWVSRIGAWKAQYPLGY 356

Query: 142 K 142  
+  
Sbjct: 357 R 357

>ref|ZP\_01079763.1| acetolactate synthase [Synechococcus sp. RS9917]  
gb|EAQ69739.1| acetolactate synthase [Synechococcus sp. RS9917]  
Length = 590

Score = 122 bits (307), Expect = 1e-26, Method: Compositional matrix adjust.  
Identities = 61/131 (46%), Positives = 84/131 (64%), Gaps = 2/131 (1%)

Query: 9 KFFIFFHFFQHEGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTGKL 68  
+F I +G++ +D LS+ MLGMHGT YAN+AV DLL+A G RFDDRVTGKL  
Sbjct: 225 RFQIPVTTTLMGKGAFDENDPLSVGMLGMHGTAYANFAVTECDLLIAVGARFDDRVTGKL 284

Query: 69 EAFASRAKIVHIDIDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRN 128

+ FA RA+++H +ID AEIGKN+ V+V GD+ L+L M ++ R EE + W  
Sbjct: 285 DTFAPRARVIHFIEDPAEIGKNRQADVAVLGDLGLSLARMVEMSLQRHEEPRT--AAWLQ 342

Query: 129 ELNVQKQKFPL 139  
+ KQ++PL

Sbjct: 343 RIEAWKQRYPL 353

>ref|ZP\_03272318.1| acetolactate synthase, large subunit, biosynthetic type  
[Arthrospira maxima CS-328]  
gb|EDZ96099.1| acetolactate synthase, large subunit, biosynthetic type  
[Arthrospira maxima CS-328]  
Length = 612

Score = 122 bits (306), Expect = 2e-26, Method: Composition-based stats.  
Identities = 62/134 (46%), Positives = 86/134 (64%), Gaps = 11/134 (8%)

Query: 21 EGSYPCDELHMLGMLGHTVYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHI 80  
+GS+ LS+ MLGMHGT YAN+AV DLL+A G RFDDRVTKL+ FASRAK++HI  
Sbjct: 255 KGSFDEKHPLSVGMMLGHTAYANFAVSECDDLLIAGARFDDRVTKLDEFASRAKVIHI 314

Query: 81 DIDS AEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLDG-----VVRNELNVQK 134  
DID AE+GKN+TP V + GDV+ L + + R E+ D G W +N +  
Sbjct: 315 DIDPAEVEGKNRTPEVPIVGDVQRVLDLLR----RCREIG-DVGNDNQTSWLERINRWR 369

Query: 135 QKFPLSFKTFGEAV 148  
+ +PL ++ +++  
Sbjct: 370 EDYPLVVPYSYDSL 383

>ref|ZP\_01621589.1| acetolactate synthase III large subunit [Lyngbya sp. PCC 8106]  
gb|EAW36367.1| acetolactate synthase III large subunit [Lyngbya sp. PCC 8106]  
Length = 612

Score = 122 bits (306), Expect = 2e-26, Method: Composition-based stats.  
Identities = 57/129 (44%), Positives = 83/129 (64%), Gaps = 1/129 (0%)

Query: 21 EGSYPCDELHMLGMLGHTVYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHI 80  
+G++ + L++ MLGMHGT YAN+AV DLL+A G RFDDRVTKLE FASRA+++HI  
Sbjct: 255 KGAFNENHPLAVGMLGMLGHTAYANFAVSECDDLLIAGARFDDRVTKLEEFASRAEVIHI 314

Query: 81 DIDS AEIGKNKTPHVSVC GDVKLAL-QGMNKVLENRAEELKLDGFWVRNELNVQKQKFPL 139  
DID AE+GKN+ P V + GDV+ L + + E + W N +N ++ +PL  
Sbjct: 315 DIDPAEVEGKNRAPQVPIVGDVQRVLDLLRRCRETEEPAPQTQTEAWLNINRWREDYPL 374

Query: 140 SFKTFGEAV 148  
+ +++  
Sbjct: 375 EVPEYPSDSL 383

>ref|YP\_001232446.1| acetolactate synthase, large subunit, biosynthetic type [Geobacter  
uraniireducens Rf4]  
gb|ABQ27873.1| acetolactate synthase, large subunit [Geobacter uraniireducens Rf4]  
Length = 566

Score = 122 bits (305), Expect = 2e-26, Method: Compositional matrix adjust.  
Identities = 64/124 (51%), Positives = 80/124 (64%), Gaps = 3/124 (2%)

Query: 22 GSYPCDELHMLGMLGHTVYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHI 81  
GS+P DD LSL +LGMHGT YAN AV + DLL+A G RFDDRVTK+ +FA AKI+HID  
Sbjct: 242 GSFPEDDPLSLGLGMLGHTYANMAVSNCDDLLVAVGARFDDRVTKIASFAPHAKIIHID 301

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEEL---KLDGFWVRNELNVQKQKFPL 138  
+D I KN + + GDVK L M KVLE + ++L K W E+ K K P  
Sbjct: 302 VDPTSIKKNVRVDLPVIGDVKDLTQMLKVLEEQQDKLDDFKKSIVPWTEIEGWKAKHP 361

Query: 139 LSKF 142  
+S+K  
Sbjct: 362 MSYK 365

>ref|YP\_172076.1| acetolactate synthase 3 catalytic subunit [Synechococcus elongatus  
PCC 6301]  
dbj|BAD79556.1| acetolactate synthase [Synechococcus elongatus PCC 6301]  
Length = 612

Score = 122 bits (305), Expect = 2e-26, Method: Composition-based stats.  
Identities = 65/169 (38%), Positives = 93/169 (55%), Gaps = 20/169 (11%)

Query: 21 EGSYPCDELSTLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHI 80  
+G++ L+L MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+ FAS+A+++H+  
Sbjct: 255 KGAFDERHPLALGMLGMHGTAYANFAVSECDLLIAGGARFDDRVTKGLDEFASKAQVIHV 314

Query: 81 DIDS AEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGV---WRNELNVQKQK 136  
DID AE+GKN+ P V + GDV+ Q +N++L E++ D W + + K++  
Sbjct: 315 DIDPAEVGKNRVP EPIVGDVR---QVLNELLARAE EQISADDATRTQPWLDRIAYWKRE 371

Query: 137 FPLS-----FKTFGEAVPCPSKKT VHDPSYEWVGPIRREGP 172  
+PL + GEA P T W + GP  
Sbjct: 372 YPLQIPYYADVLS PQVIHSLGEAAPDAYFTT DVGQHQMWAAQFIKNGP 420

>ref|YP\_399158.1| acetolactate synthase 3 catalytic subunit [Synechococcus elongatus  
PCC 7942]  
gb|ABB56171.1| acetolactate synthase, large subunit [Synechococcus elongatus PCC  
7942]  
prf||1611501A acetolactate synthase  
Length = 612

Score = 122 bits (305), Expect = 2e-26, Method: Composition-based stats.  
Identities = 65/169 (38%), Positives = 93/169 (55%), Gaps = 20/169 (11%)

Query: 21 EGSYPCDELSTLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHI 80  
+G++ L+L MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+ FAS+A+++H+  
Sbjct: 255 KGAFDERHPLALGMLGMHGTAYANFAVSECDLLIAGGARFDDRVTKGLDEFASKAQVIHV 314

Query: 81 DIDS AEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGV---WRNELNVQKQK 136  
DID AE+GKN+ P V + GDV+ Q +N++L E++ D W + + K++  
Sbjct: 315 DIDPAEVGKNRVP EPIVGDVR---QVLNELLARAE EQISADDATRTQPWLDRIAYWKRE 371

Query: 137 FPLS-----FKTFGEAVPCPSKKT VHDPSYEWVGPIRREGP 172  
+PL + GEA P T W + GP  
Sbjct: 372 YPLQIPYYADVLS PQVIHSLGEAAPDAYFTT DVGQHQMWAAQFIKNGP 420

>ref|ZP\_00514506.1| Acetolactate synthase, large subunit, biosynthetic type  
[Crocospaera watsonii WH 8501]  
gb|EAM52344.1| Acetolactate synthase, large subunit, biosynthetic type  
[Crocospaera watsonii WH 8501]  
Length = 619

Score = 121 bits (304), Expect = 3e-26, Method: Composition-based stats.  
Identities = 60/121 (49%), Positives = 82/121 (67%), Gaps = 5/121 (4%)

Query: 22 GSYPCDELSTLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
GS+ LS+ MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+ FASRA+++HID  
Sbjct: 263 GSFDEHHPLSVGMLGMHGTAYANFAVSECDLLIAGGARFDDRVTKGLDEFASRARVIHID 322

Query: 82 IDS AEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD---FGVWRNELNVQKQKFP 138  
ID AE+GKN+ P V + GDV L+ + + +R ++ LD W + +N +Q++P  
Sbjct: 323 IDPAEVGKNRRPDVPIVGDVAQVLEQL--LQRSREIDVPLDPQKTEAWLHRINRWRQYQ 380

Query: 139 L 139  
L  
Sbjct: 381 L 381

>ref|YP\_002380209.1| acetolactate synthase, large subunit, biosynthetic type [Cyanothec

sp. PCC 7424]  
gb|ACK73341.1| acetolactate synthase, large subunit, biosynthetic type [Cyanothecae  
sp. PCC 7424]  
Length = 619

Score = 121 bits (304), Expect = 3e-26, Method: Composition-based stats.  
Identities = 61/128 (47%), Positives = 82/128 (64%), Gaps = 1/128 (0%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G++ LS+ MLGMHGT YAN+AV DLL+A G RFDDRVTKLE FASRAK+VHID  
Sbjct: 263 GAFDEHHP LSVGMLGMHGTAYANFAVTECDLLIAVGARFDDRVTKLEEFASRAKVHID 322

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGM-NKVLENRAEELKLD FGVWRNENLVQKQKFP L 140  
ID AE+GK + P V + GDV+ L+ M +V E + W +N ++ +PL  
Sbjct: 323 IDPAEVGKVRAPDVP IVDVQVLEQMLQVRREMDYPTVPERTKDWLERINRWREDYPLM 382

Query: 141 FKTFGEAV 148  
T+ +++  
Sbjct: 383 VPTYPD SI 390

>ref|NP\_441297.1| acetolactate synthase 3 catalytic subunit [Synechocystis sp. PCC  
6803]  
dbj|BAA17977.1| acetohydroxy acid synthase [Synechocystis sp. PCC 6803]  
Length = 621

Score = 121 bits (303), Expect = 3e-26, Method: Composition-based stats.  
Identities = 63/142 (44%), Positives = 87/142 (61%), Gaps = 11/142 (7%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G++ + LS+ MLGMHGT YAN+AV DLL+A G RFDDRVTKL+ FASRAK++HID  
Sbjct: 265 GAFDENHP LSVGMLGMHGTAYANFAVSECDLLIAVGARFDDRVTKLDEFASRAKVIHID 324

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKNVLENRAEELK L-----DFGVWRNENLVQKQK 136  
ID AE+GKN+ P V + GDV+ L+ L RA EL W N ++ +  
Sbjct: 325 IDPAEVGKNRAPDVP IVDVHVLEQ----LLQRAELDYPTHTPHTTQAWLNRI DHWRD 380

Query: 137 FPLSFKTFGEAVPCPSKKT VHD 158  
+PL + + + ++ VH+  
Sbjct: 381 YPLQVPHYEDTI--APQEVVHE 400

>ref|YP\_001936332.1| acetolactate synthase large subunit [Heterosigma akashiwo]  
gb|ABV65938.1| acetolactate synthase large subunit [Heterosigma akashiwo]  
gb|ABV70079.1| acetolactate synthase large subunit [Heterosigma akashiwo]  
Length = 608

Score = 121 bits (303), Expect = 4e-26, Method: Compositional matrix adjust.  
Identities = 70/182 (38%), Positives = 103/182 (56%), Gaps = 15/182 (8%)

Query: 18 FQHEGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKI 77  
+ +G++ + L L MLGMHGT +AN+AV DLL+A G RFDDRVTKL+ FA+ AKI  
Sbjct: 251 LKGKGAFDENHYLHLGMLGMHGTAFANFAVNECDLLIAVGARFDDRVTKLDEF AAGAKI 310

Query: 78 VHIDIDSAEIGKNKTPHVSVC GDVKLALQGMNKNVLENRAEELKLD FGV-WRNENLVQKQK 136  
+HIDID AEI KN P V++ GDVK+ LQ + ++ N ++ + + WR+ +N ++  
Sbjct: 311 IHIDIDPAEIAKNCIPQVALIGDVKIILQELIEIYTNNIDQYEEKSTLPWRHRINAWQEA 370

Query: 137 FPLSFKTFGEAVPCPSKKT VHDPSYEWVGPIRREGPPSPMREDKLTQNTHLIRTYKFVFF 196  
+PL +P P + +Y I G +P RE T + + + F  
Sbjct: 371 YPL-----LIPAPEE-----AYSPQQVINEIGKVAP-RETIFTTVDVGHQMWAAQFL 416

Query: 197 KC 198  
KC  
Sbjct: 417 KC 418

>ref|YP\_359378.1| acetolactate synthase, large subunit, biosynthetic type  
[Carboxydotherrmus hydrogenoformans Z-2901]



gb|ABB14028.1| acetolactate synthase, large subunit, biosynthetic type  
[Carboxydotherrmus hydrogenoformans Z-2901]  
Length = 554

Score = 120 bits (302), Expect = 4e-26, Method: Compositional matrix adjust.  
Identities = 63/122 (51%), Positives = 79/122 (64%), Gaps = 10/122 (8%)

Query: 23 SYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRV TGKLEAFASRAKIVHIDI 82  
S+P D L L MLG+HGT YAN AV D+L+ G VRF DRV TG+L FA +AKI+HID+  
Sbjct: 242 SFPQDSPLFLGMLGLHGTRYANLAVTECDVLIGLVRFADRV TGELSGFAPKAKIIHIDV 301

Query: 83 DSAEIGKNKTPHVSVC GDVKLALQGMNKVLE--NRAEELKLD FGVWRNELNVQKQKFPLS 140  
D AEIGKN V + GDVK LQ M K +E NR E W +++N KQ+FPPL  
Sbjct: 302 DPAEIGKNVRADVPIVGDVKNVLQEMLKQIEPQNRQE-----WLSQINTWKQEFPLK 353

Query: 141 FK 142  
++  
Sbjct: 354 YE 355

>ref|YP\_001293618.1| acetohydroxyacid synthetase large subunit [Rhodomonas salina]  
gb|ABO70722.1| acetohydroxyacid synthetase large subunit [Rhodomonas salina]  
Length = 574

Score = 120 bits (302), Expect = 4e-26, Method: Compositional matrix adjust.  
Identities = 58/110 (52%), Positives = 75/110 (68%)

Query: 30 LSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRV TGKLEAFASRAKIVHIDIDSAEIGK 89  
L L MLGMHGT VYANYAV DLL+A G RFDDRV TGKL+ FA A++VHIDID AE+GK  
Sbjct: 263 LCLGMLGMHGT VYANYAVSECDLLIALGARFDDRV TGKLDEFACHAQV VHIDIDPAEIVGK 322

Query: 90 NKT PHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPL 139  
N+ P V + GDVK +++ + VL+ WR L+ ++++PL  
Sbjct: 323 NRMPQVGIVGDVKESVKAILDVLKKEDYLEPEQTQAWRQRLDRWRKEYPL 372

>ref|YP\_731202.1| acetolactate synthase 3 catalytic subunit [Synechococcus sp.  
CC9311]  
gb|ABI45115.1| acetolactate synthase, large subunit, biosynthetic type  
[Synechococcus sp. CC9311]  
Length = 581

Score = 120 bits (302), Expect = 5e-26, Method: Compositional matrix adjust.  
Identities = 60/134 (44%), Positives = 86/134 (64%), Gaps = 2/134 (1%)

Query: 8 KKFFIFFHFFQHEGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRV TGK 67  
++F I +G++ + L+L MLGMHGT YAN+AV DLL+A G RFDDRV TGK  
Sbjct: 239 ERFQIPVTTLMGKGAFDENHPLALGMLGMHGTAYANFAVTDCLLIAVGARFDDRV TGK 298

Query: 68 LEAFASRAKIVHIDIDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWR 127  
LD+ FA +AK++H +ID AE+GKN+ P V V GDV L+L + V ++R +L W  
Sbjct: 299 LDTFAPKAKVIHFEIDPAEIVGKNRCPDVVVLGDVGLSLAQL--VDQSRPHSAELTSSWL 356

Query: 128 NELNVQKQKFPLSF 141  
++N K+ +PL+  
Sbjct: 357 EQINSWKELYPLTI 370

>ref|NP\_870771.1| acetolactate synthase III [precursor] [Rhodopirellula baltica SH 1]  
emb|CAD77848.1| acetolactate synthase III [Precursor] [Rhodopirellula baltica SH 1]  
Length = 619

Score = 120 bits (302), Expect = 5e-26, Method: Compositional matrix adjust.  
Identities = 62/120 (51%), Positives = 77/120 (64%), Gaps = 6/120 (5%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRV TGKLEAFASRAKIVHID 81  
G+ DD SL LGMHG YANYAV DLL+A G VRFDDRV TGK+EAF A KI+H+D  
Sbjct: 283 GAVSPDDPRSLDWLGMHGAAYANYAVRDCDLLIALGVRFDDRV TGKVEAFADAKIIHVD 342

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMN KVLNRAEELKLD FGVWRNELNVQKQK FPLSF 141  
IDS+E+ KNK H+ V GDVK L +NK+++ E W+ K K+PL +  
Sbjct: 343 IDSSELNKNKQAHIPVRGDVKDLVQLN KIVQPPEIE-----AWQKTCTDLKAKYPLKY 396

>ref|YP\_002371135.1| acetolactate synthase, large subunit, biosynthetic type [Cyanothec  
sp. PCC 8801]  
gb|ACK64979.1| acetolactate synthase, large subunit, biosynthetic type [Cyanothec  
sp. PCC 8801]  
Length = 619

Score = 120 bits (300), Expect = 8e-26, Method: Composition-based stats.  
Identities = 60/132 (45%), Positives = 83/132 (62%), Gaps = 9/132 (6%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G++ LS+ MLGMHGT YAN+AV DLL+A G RFDDRVTGKL+ FAS AK++HID  
Sbjct: 263 GAFDEHHPLSVAMLGMHGTAYANFAVSECDLLIAVGARFDDRVTGKLDEFASHAKVIHID 322

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMN KVLNRAEELKLD FGV-----WRNELNVQKQK 136  
ID AE+GKN+TP V + GDV+ L+ L RA E+ W + +N +++  
Sbjct: 323 IDPAEVGKNRTPDVP IVDVRRVLEQ---LLQ RAREIDCPTHA EKTQPWLSRINRWREQ 378

Query: 137 FPLSFKTFGEAV 148  
+PL + ++  
Sbjct: 379 YPLVIPHYDHSI 390

>ref|ZP\_03142166.1| acetolactate synthase, large subunit, biosynthetic type [Cyanothec  
sp. PCC 8802]  
gb|EDY03144.1| acetolactate synthase, large subunit, biosynthetic type [Cyanothec  
sp. PCC 8802]  
Length = 619

Score = 120 bits (300), Expect = 8e-26, Method: Composition-based stats.  
Identities = 60/132 (45%), Positives = 83/132 (62%), Gaps = 9/132 (6%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G++ LS+ MLGMHGT YAN+AV DLL+A G RFDDRVTGKL+ FAS AK++HID  
Sbjct: 263 GAFDEHHPLSVAMLGMHGTAYANFAVSECDLLIAVGARFDDRVTGKLDEFASHAKVIHID 322

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMN KVLNRAEELKLD FGV-----WRNELNVQKQK 136  
ID AE+GKN+TP V + GDV+ L+ L RA E+ W + +N +++  
Sbjct: 323 IDPAEVGKNRTPDVP IVDVRRVLEQ---LLQ RAREIDCPTHA EKTQPWLSRINRWREQ 378

Query: 137 FPLSFKTFGEAV 148  
+PL + ++  
Sbjct: 379 YPLVIPHYDHSI 390

>ref|YP\_001227109.1| acetolactate synthase 3 catalytic subunit [Synechococcus sp.  
RCC307]  
emb|CAK27756.1| Acetolactate synthase large subunit [Synechococcus sp. RCC307]  
Length = 615

Score = 120 bits (300), Expect = 8e-26, Method: Composition-based stats.  
Identities = 57/120 (47%), Positives = 81/120 (67%), Gaps = 2/120 (1%)

Query: 21 EGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHI 80  
+G + +D L++ MLGMHGT YAN+AV DLL+A G RFDDRVTGKL+ FA RAK++H  
Sbjct: 262 KGVFDENDPLAVGMLGMHGTAYANFAVTD CDLLIAVGARFDDRVTGKLDTFAPRAKVIHF 321

Query: 81 DIDSAEIGKNKTPHVSVC GDVKLALQGMN KVLNRAEELKLD FGVWRNELNVQKQK FPLS 140  
+ID AE+GKN+ P V+V GDV+ +L+ + + + R E W ++ KQ +PL+  
Sbjct: 322 EIDPAEVGKNRKP DVAVLGDVRRSLEELLQSQQRHVEPHT--AAWLARIDSWKQNYPLT 379

>ref|YP\_323265.1| acetolactate synthase 3 catalytic subunit [Anabaena variabilis ATCC  
29413]  
gb|ABA22370.1| acetolactate synthase, large subunit [Anabaena variabilis ATCC

29413]  
Length = 632

Score = 120 bits (300), Expect = 9e-26, Method: Compositional matrix adjust.  
Identities = 59/121 (48%), Positives = 79/121 (65%), Gaps = 1/121 (0%)

Query: 30 LSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGK 89  
LSL MLGMHGT YAN+AV DLL+ G RFDDRVTKGL+ FASRAK++HIDID AE+GK  
Sbjct: 285 LSLGMLGMHGTAYANFAVTDCLLICVGARFDDRVTKGLDEFASRAKVIHIDIDPAEVGK 344

Query: 90 NKTPHVSVC GDV-KLALQGMNKVLENRAEELKLDGFWVRNELNVQKQKFPLSFKTFGEAV 148  
N+ P V + GDV K+ L + + + A+ W N +N ++ +PL +++  
Sbjct: 345 NRIPEVPIVG DVKVLDDLRRCKQTGAKNTPNQNEWLNINRWREDYPLVIPQHPDSI 404

Query: 149 P 149  
P  
Sbjct: 405 P 405

>ref|NP\_488653.1| acetolactate synthase 3 catalytic subunit [Nostoc sp. PCC 7120]  
dbj|BAB76312.1| acetohydroxy acid synthase [Nostoc sp. PCC 7120]  
Length = 632

Score = 120 bits (300), Expect = 9e-26, Method: Compositional matrix adjust.  
Identities = 59/121 (48%), Positives = 79/121 (65%), Gaps = 1/121 (0%)

Query: 30 LSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHIDIDSAEIGK 89  
LSL MLGMHGT YAN+AV DLL+ G RFDDRVTKGL+ FASRAK++HIDID AE+GK  
Sbjct: 285 LSLGMLGMHGTAYANFAVTDCLLICVGARFDDRVTKGLDEFASRAKVIHIDIDPAEVGK 344

Query: 90 NKTPHVSVC GDV-KLALQGMNKVLENRAEELKLDGFWVRNELNVQKQKFPLSFKTFGEAV 148  
N+ P V + GDV K+ L + + + A+ W N +N ++ +PL +++  
Sbjct: 345 NRIPEVPIVG DVKVLDDLRRCKQTSAKTTPNQNEWLNINRWREDYPLVVPQHPDSI 404

Query: 149 P 149  
P  
Sbjct: 405 P 405

>ref|YP\_001517397.1| acetolactate synthase 3 catalytic subunit [Acaryochloris marina  
MBIC11017]  
gb|ABW28081.1| acetolactate synthase, large subunit, biosynthetic type  
[Acaryochloris marina MBIC11017]  
Length = 602

Score = 119 bits (299), Expect = 1e-25, Method: Compositional matrix adjust.  
Identities = 60/130 (46%), Positives = 80/130 (61%)

Query: 10 FFIFHFFHQHEGSYP CDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLE 69  
F I +GS+ LS+ MLGMHGT YAN+AV DLL+A G RFDDRVTKGL+  
Sbjct: 235 FHIPVTTTLMGKGSFDEAHPLSVGMLGMHGTAYANFAVSECDLLIAVGARFDDRVTKGLD 294

Query: 70 AFASRAKIVHIDIDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLDGFWVRNE 129  
FASRAK++HIDID AE+GKN+ P V + G VK L + K+ + + WR+  
Sbjct: 295 EFASRAKVIHIDIDPAEVGKNRAPTVPIVGSVKPVLVELLKIAKQSSPPDTAQTQEWRSR 354

Query: 130 LNVQKQKFPL 139  
+ ++ +PL  
Sbjct: 355 VERWQRDYPL 364

>ref|YP\_002508832.1| acetolactate synthase, large subunit, biosynthetic type  
[Halothermothrix orenii H 168]  
gb|ACL69837.1| acetolactate synthase, large subunit, biosynthetic type  
[Halothermothrix orenii H 168]  
Length = 555

Score = 119 bits (299), Expect = 1e-25, Method: Compositional matrix adjust.  
Identities = 60/121 (49%), Positives = 81/121 (66%), Gaps = 6/121 (4%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHI 81  
G YP +D LSL M GMHGT YANYA+ +DLL+A GVRFDDRVTGK+E FA AKI+HID  
Sbjct: 242 GIYPENDFLSLGMFGMGTTRYANYAISEADLLVAVGVRFDDRVTGKIEEFAPDAKIIHI 301

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPPLSF 141  
ID AEI KN H+ + GD + L+ + ++E++ + G W ++ K+K PL +  
Sbjct: 302 IDPAEICKNVEAHIPVGDARNILKELPLIEHK-----ERGAWLQQIKEWKEKNPLKY 355

Query: 142 K 142  
+  
Sbjct: 356 E 356

>ref|YP\_002250909.1| acetolactate synthase, large subunit, biosynthetic type  
[Dictyoglomus thermophilum H-6-12]  
gb|ACI19136.1| acetolactate synthase, large subunit, biosynthetic type  
[Dictyoglomus thermophilum H-6-12]  
Length = 557

Score = 119 bits (298), Expect = 1e-25, Method: Compositional matrix adjust.  
Identities = 58/128 (45%), Positives = 80/128 (62%), Gaps = 6/128 (4%)

Query: 21 EGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHI 80  
+G+ P ELS +GMHG YANYA+ SDL++A GVRF DR TGK+E FA AKI+HI  
Sbjct: 241 KGAIPEHETELSCGFIGMHGAAYANYAINDSDLLIAIGVRFSDRSTGKVETFPAPNAKIIHI 300

Query: 81 DIDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPPLS 140  
DID AEIGKN P+V + D K AL+ + +V+E + +W ++ K K+PL  
Sbjct: 301 DIDPAEIGKNVQPYVPVIVADAKRALEKLEIIEVVEPKTNP-----IWWEKIKDWKNKYPLR 354

Query: 141 FKTFGEAV 148  
++ E +  
Sbjct: 355 YRMSNEVI 362

>ref|YP\_002481296.1| acetolactate synthase, large subunit, biosynthetic type [Cyanothecae  
sp. PCC 7425]  
gb|ACL42935.1| acetolactate synthase, large subunit, biosynthetic type [Cyanothecae  
sp. PCC 7425]  
Length = 602

Score = 119 bits (298), Expect = 1e-25, Method: Composition-based stats.  
Identities = 62/135 (45%), Positives = 84/135 (62%), Gaps = 9/135 (6%)

Query: 21 EGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHI 80  
+G++ + L++ MLGMHGT YAN+AV DLL+A G RFDDRVTGKL+ FASRAK++HI  
Sbjct: 246 KGAFDENHPLAVGMLGMHGTAYANFAVSECDDLLIAVGARFDDRVTGKLDEFASRAKVIHI 305

Query: 81 DIDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPPL 139  
DID AE+GKN+ P V + G V+ L + + L E KL W ++ KQ +PL  
Sbjct: 306 DIDPAEVGKNRVPEVPIVGSVQPVLAELLR-LAREGEPAKLGQTRAWLERIDRWKQDYPL 364

Query: 140 SFKTFGEAVPCPSKK 154  
VP P+ +  
Sbjct: 365 -----VVPYPADR 372

>ref|YP\_377644.1| acetolactate synthase 3 catalytic subunit [Synechococcus sp.  
CC9902]  
gb|ABB26600.1| acetolactate synthase, large subunit [Synechococcus sp. CC9902]  
Length = 618

Score = 119 bits (298), Expect = 1e-25, Method: Composition-based stats.  
Identities = 55/120 (45%), Positives = 81/120 (67%), Gaps = 2/120 (1%)

Query: 21 EGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHI 80  
+G++ ++ L++ MLGMHGT YAN+AV DLL+A G RFDDRVTGKL+ FA RAK++H  
Sbjct: 265 KGAFDENESLAVGMLGMHGTAYANFAVTECDLLIAVGARFDDRVTGKLDTFAPRAKVIHF 324

Query: 81 DIDS AEIGKNKTPHVS VCGDV K LALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP L S 140  
+ID AE+GKN++P V+V GD+ L++ + ++ R E W + KQ +PL+  
Sbjct: 325 EIDPAE VGNRSPEVAVLG DLSL SMARLVQLSLGR--EADPHTAAWLKRIQSWKQHYPLT 382

>ref|YP\_002536969.1| acetolactate synthase, large subunit, biosynthetic type [Geobacter  
sp. FRC-32]  
gb|ACM19868.1| acetolactate synthase, large subunit, biosynthetic type [Geobacter  
sp. FRC-32]  
Length = 566

Score = 119 bits (298), Expect = 1e-25, Method: Compositional matrix adjust.  
Identities = 59/130 (45%), Positives = 82/130 (63%), Gaps = 3/130 (2%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRV TGKLEAFASRAKIVHI D 81  
G++P D+ LSL +LGMHGT YAN AV H DLL+A G RFDDRV TGK+ +FA A+I+HID  
Sbjct: 242 GAFPEDNPLSLGLGMHGTYYANMAVSHCDLLIAVGARFDDRV TGKIASFAPHAQIIHID 301

Query: 82 IDSAEIGKNKTPHVS VCGDV K LALQGMNKVLENRAE E---LKLDFGVWRNELNVQKQKFP 138  
+D I KN + + GDVK L+ + KVL +AE+ + W E+ K K P  
Sbjct: 302 VDPTS IKNVRVDLP IVGDV K DVLKQLLKVLAGQAEKAATFQKQIVPWSQEIEAWKAKHP 61

Query: 139 LSFKTFGEAV 148  
+++K+ +  
Sbjct: 362 MTKSSASVI 371

>ref|NP\_053940.1| acetohydroxyacid synthase large subunit [Porphyra purpurea]  
sp|P69683.1|ILVB\_PORPU RecName: Full=Acetolactate synthase large subunit; Short=AHAS;  
AltName: Full=Acetohydroxy-acid synthase large subunit;  
Short=ALS  
sp|P69684.1|ILVB\_PORUM RecName: Full=Acetolactate synthase large subunit; Short=AHAS;  
AltName: Full=Acetohydroxy-acid synthase large subunit;  
Short=ALS  
gb|AAA03052.1| acetolactate synthase  
gb|AAC08216.1| acetohydroxyacid synthase large subunit [Porphyra purpurea]  
Length = 590

Score = 119 bits (298), Expect = 1e-25, Method: Compositional matrix adjust.  
Identities = 54/119 (45%), Positives = 78/119 (65%)

Query: 21 EGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRV TGKLEAFASRAKIVHI 80  
+G + D E L MLGMHGT YAN+AV DLL+A G RFDDRV TGKL+ FA A+++H+  
Sbjct: 254 KGIFNEDSEFC LGMHGTAYANFAVSECDLLIALGARFDDRV TGKLDEFACNAQVIHV 313

Query: 81 DIDS AEIGKNKTPHVS VCGDV K LALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP L 139  
DID AE+GKN+ P V++ GDV + + +L+N + W+ ++ +Q++PL  
Sbjct: 314 DIDPAE VGNRIPQVAIVGDVTEVVTSLLNLLKNNFKPYPEQIISWQERIHRWRQYPL 372

>ref|YP\_002327489.1| acetohydroxy acid synthase large subunit [Vaucheria litorea]  
gb|ACF70906.1| acetohydroxy acid synthase large subunit [Vaucheria litorea]  
Length = 590

Score = 119 bits (297), Expect = 2e-25, Method: Compositional matrix adjust.  
Identities = 69/164 (42%), Positives = 89/164 (54%), Gaps = 33/164 (20%)

Query: 21 EGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFDDRV TGKLEAFASRAKIVHI 80  
+GS+ + L+L MLGMHGT YAN+AV + D L L L A G RFDDRV TGKLE FAS A I+HI  
Sbjct: 254 KGSFDENHPLALGMLGMHGTAYANFAVSNCDLLLAIGARFDDRV TGKLENFASNAVIIHI 313

Query: 81 DIDS AEIGKNKTPHVS VCGDV K LALQGMNKVLENRAEELKLD----- 122  
DID AEI KNK +S+ GDVK K+L E K+D  
Sbjct: 314 DIDPAEISKNSQLSILGDVK-----KILPQILEHSKIDPILNLFDTDLMGVKNNEI 366

Query: 123 -FGVWRNELNVQKQKFP LSFKTFGEAVPCPSKKT VHDPSYEWVG 165  
W ++N+ K+ +PLS +P P+K ++G  
Sbjct: 367 VINAWLKKINLWKEFYPLS-----IPIPTKSLSPQEVINYIG 403

```
>ref|YP_001211077.1| thiamine pyrophosphate-requiring enzymes [Pelotomaculum
thermopropionicum SI]
dbj|BAF58708.1| thiamine pyrophosphate-requiring enzymes [Pelotomaculum
thermopropionicum SI]
Length = 555

Score = 119 bits (297), Expect = 2e-25, Method: Compositional matrix adjust.
Identities = 61/127 (48%), Positives = 79/127 (62%), Gaps = 6/127 (4%)

Query: 23  SYPCDELHSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTVGKLEAFASRAKIVHIDI 82
      +P +  LSL MLGMHGT YAN+AV  DLL+A G RFDDRVTGKLE FA  A+I+HIDI
Sbjct: 245  GFPGNHPLSLGMLGMHGT KYANFAVCECDLLIAVGARFDDRVTGKLETFAPPEARIIHIDI 304

Query: 83  DSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLDFGVWRNELNVQKQKFPLSFK 142
      D AEIGKN  + + GDVK  L  + ++L  E  WR ++  K+++PL++
Sbjct: 305  DPAEIGKNVRVDIPIVGDVKRVLSQLLEILRPGLRE-----AWREKIEAWKKEYPLTYC 358

Query: 143  TFGEAVP 149
      G  P
Sbjct: 359  EQGRLKP 365

>ref|NP_849033.1| acetohydroxyacid synthase large subunit [Cyanidioschyzon merolae
strain 10D]
dbj|BAC76195.1| acetolactate synthase large subunit [Cyanidioschyzon merolae strain
10D]
Length = 569

Score = 119 bits (297), Expect = 2e-25, Method: Compositional matrix adjust.
Identities = 61/118 (51%), Positives = 75/118 (63%), Gaps = 4/118 (3%)

Query: 26  CDDE--LSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTVGKLEAFASRAKIVHIDI 82
      C DE  L L MLGMHGT YAN+AV  DLL+A G RFDDRVTGKL+ FAS A+++HIDI
Sbjct: 244  CFDETHPLYLGLMGMHGTAYANFAVSECDLLIALGARFDDRVTGKLDEFASSAQVIHIDI 303

Query: 83  DSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLDFG-VWRNELNVQKQKFPL 139
      D+AEIGKN+ P +++C DVKL LQ  + D  WR +  K +PL
Sbjct: 304  DAAEIGKNRIPQLAICSDVKLVLQHFLAATSIAPNLVANDETYAWRKRIARWKHDYPL 361

>ref|ZP_01472584.1| acetolactate synthase III large subunit [Synechococcus sp. RS9916]
gb|EAU73275.1| acetolactate synthase III large subunit [Synechococcus sp. RS9916]
Length = 624

Score = 119 bits (297), Expect = 2e-25, Method: Composition-based stats.
Identities = 58/131 (44%), Positives = 87/131 (66%), Gaps = 5/131 (3%)

Query: 21  EGSYPCDELHSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTVGKLEAFASRAKIVHI 80
      +G++  +D LS+ MLGMHGT YAN+AV  DLL+A G RFDDRVTGKL+ FA RA+++H
Sbjct: 271  KGAFDENDSL SVGLMGMHGTAYANFAVTECDLLIAVGARFDDRVTGKLDTFAPRARVIHF 330

Query: 81  DIDSAEIGKNKTPHVSVC GDVKLALQGMNKVLE-NRAEELKLDFGVWRNELNVQKQKFPL 139
      +ID AE+GK +  V+V GD+  Q M +++E +R  +++  W  ++  K+++PL
Sbjct: 331  EIDPAEVGKTRAEVAVLGDLA---QSMAQLVERSRQDIRPTTAAWLAKITTWEQYPL 387

Query: 140  SFKT-FGEAVP 149
      + T  GE  P
Sbjct: 388  TVPTPEGEIYP 398

>ref|ZP_02850743.1| acetolactate synthase, large subunit, biosynthetic type
[Paenibacillus sp. JDR-2]
gb|EDS49383.1| acetolactate synthase, large subunit, biosynthetic type
[Paenibacillus sp. JDR-2]
Length = 581

Score = 118 bits (296), Expect = 2e-25, Method: Compositional matrix adjust.
```

Identities = 60/121 (49%), Positives = 78/121 (64%), Gaps = 5/121 (4%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 81  
G +P ++L L M GMHGT AN +++ +DLL+ G RFDDRVTGKL FA AKIVHID

Sbjct: 263 GGFPSGNDLWLGMPGMHGT YTANKSIQSADLLINIGARFDDRVTGKLAGFAPLAKIVHID 322

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID AEIGKN + + GDVK LQ +N+ L RA++ WR E+ K ++P S+

Sbjct: 323 IDPAEIGKNVPTDIPVIGDVKTVLQQVNQ-LAKRADKAD----AWRQEIQESKAQY PFSY 377

Query: 142 K 142  
K

Sbjct: 378 K 378

>ref|YP\_002353080.1| acetolactate synthase, large subunit, biosynthetic type  
[Dictyoglomus turgidum DSM 6724]

gb|ACK42466.1| acetolactate synthase, large subunit, biosynthetic type  
[Dictyoglomus turgidum DSM 6724]  
Length = 557

Score = 118 bits (296), Expect = 2e-25, Method: Compositional matrix adjust.  
Identities = 58/122 (47%), Positives = 78/122 (63%), Gaps = 6/122 (4%)

Query: 21 EGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHI 80  
+GS P ELS +GMHG YANYA+ SDL++A GVRF DR TGK+E FA AKI+HI

Sbjct: 241 KGSIPETHELSCGFIGMGAAYANYAINDSDLI IAGVRFS DRSTGKVETFAPNAKIIHI 300

Query: 81 DIDS AEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 140  
DID AEIGKN P+V + D K AL+ + +V+E + + W ++ K K+PL

Sbjct: 301 DIDPAEIGKNVNPYPVIVADAKRALERLVEVVEPKINAM-----WWEKIKEWKTKYPLR 354

Query: 141 FK 142  
++

Sbjct: 355 YR 356

>ref|ZP\_03129981.1| acetolactate synthase, large subunit, biosynthetic type  
[Chthoniobacter flavus Ellin428]

gb|EDY19461.1| acetolactate synthase, large subunit, biosynthetic type  
[Chthoniobacter flavus Ellin428]  
Length = 591

Score = 118 bits (296), Expect = 2e-25, Method: Compositional matrix adjust.  
Identities = 63/124 (50%), Positives = 82/124 (66%), Gaps = 4/124 (3%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 81  
G +P ELSL LGMHGT VYAN AV +DLLLA GVRFD DRTGK+ FA IVHID

Sbjct: 241 GCFPETHELKSLKWLGMHGT VYANNAVN EADLLLAIGVRFD DRTGKISEFAKHGTIVHID 300

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVL-ENRAEELKLD FGV---WRNELNVQKQK F 137  
ID++EI KNK + + DVK AL +N++L E+ A+ ++ F W ++N ++ F

Sbjct: 301 IDNSEINKNKVVKLPILSDVKYALGRNLQ LLEESGAKRVQSGFKAFPDWYVKINAWREAF 360

Query: 138 PLSF 141  
P +F

Sbjct: 361 PFTF 364

>ref|YP\_842552.1| acetolactate synthase 3 catalytic subunit [Methanosaeta thermophila  
PT]

gb|ABK13912.1| acetolactate synthase, large subunit [Methanosaeta thermophila PT]  
Length = 557

Score = 118 bits (296), Expect = 2e-25, Method: Compositional matrix adjust.  
Identities = 58/120 (48%), Positives = 81/120 (67%), Gaps = 6/120 (5%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTGKLEAFASRAKIVHID 81  
G +P D LSL MLGMHGT YAN AV+ +D+LLA G RFDDRVTGK+ +FA +A+I+HID

Sbjct: 242 GCFPTDHPHSLGMLGMHGTKYANLAVQEADVLLAVGARFDDRVTKIASFAPKARIHID 301

Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKLLENRAEELKLDGFGVWRNELNVQKQKFPPLSF 141  
+D+AEIGKN + V GD K+AL + K ++ R + W ++ K+++PL +

Sbjct: 302 VDAAEIGKNVRVDIPVVGDAKIALTELLKHVQQRP-----WTEWNEKILSWKREYPLRY 355

>ref|ZP\_01085428.1| acetolactate synthase III large subunit [Synechococcus sp. WH 5701]  
gb|EAQ74912.1| acetolactate synthase III large subunit [Synechococcus sp. WH 5701]  
Length = 590

Score = 118 bits (296), Expect = 3e-25, Method: Compositional matrix adjust.  
Identities = 56/136 (41%), Positives = 87/136 (63%), Gaps = 2/136 (1%)

Query: 8 KKKFIFFFHFFQHEGSYPCDDELHMLGMHGTVYANYAVEHSDLLAFGVRFDRTVGK 67  
++F + +G++ L++ MLGMHGT YAN+AV DLL+A G RFDDRTVG+  
Sbjct: 224 ERFLQPLVTTLMGKGAFDERHPLAVGMLGMHGTAYANFAVTECDLLIAAGARFDDRTVGR 283

Query: 68 LEAFASRAKIVHIDIDSAEIGKNKTPHVSVCQDVKLALQGMNKLLENRAEELKLDGFGVWR 127  
L++FA RA+++HIDID+AE+GKN+ P V + DV+LAL+ + + ++ + G W  
Sbjct: 284 LDSFAPRAQVIHIDIDAEEVGKNRLPEVPISDVRLALEALLEASVGESDNGRT--GPWL 341

Query: 128 NELNVQKQKFPPLSFKT 143  
++ K+ +PL T  
Sbjct: 342 ERISRWKRHYPLVIPT 357

>ref|YP\_002467404.1| acetolactate synthase, large subunit, biosynthetic type [Candidatus  
Methanosphaerula palustris E1-9c]  
gb|ACL17681.1| acetolactate synthase, large subunit, biosynthetic type [Candidatus  
Methanosphaerula palustris E1-9c]  
Length = 558

Score = 118 bits (295), Expect = 3e-25, Method: Compositional matrix adjust.  
Identities = 57/120 (47%), Positives = 76/120 (63%), Gaps = 5/120 (4%)

Query: 22 GSYPCDDELHMLGMHGTVYANYAVEHSDLLAFGVRFDRTVGKLEAFASRAKIVHID 81  
G+ P D L+L MLGMHGT YAN+AV +DLL+A GVRFDRTVGKL FA A+++HID  
Sbjct: 242 GAVPGDHPLNLGMLGMHGTAYANFAVTEADLLIAVGVRFDRTVGKLSFAPGARVIHID 301

Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKLLENRAEELKLDGFGVWRNELNVQKQKFPPLSF 141  
+D AEIGKN+ P V + GD K L + L+ + + W ++ K+ PL +  
Sbjct: 302 VDPAEIGKNRQPEVPVIGDAKCVLGELLSRLQKKQSS-----SAWLEQVAAWKRDHPLRY 356

>ref|YP\_001030046.1| acetolactate synthase, large subunit [Methanocorpusculum labreanum  
Z]  
gb|ABN06779.1| acetolactate synthase, large subunit [Methanocorpusculum labreanum  
Z]  
Length = 567

Score = 118 bits (295), Expect = 3e-25, Method: Compositional matrix adjust.  
Identities = 62/114 (54%), Positives = 73/114 (64%), Gaps = 11/114 (9%)

Query: 22 GSYPCDDELHMLGMHGTVYANYAVEHSDLLAFGVRFDRTVGKLEAFASRAKIVHID 81  
G+ P D L+L MLGMHGT YANYAV +DLL+A G RFDDRTVGKL FA+ AK++HID  
Sbjct: 244 GAIPADHPLNLGMLGMHGTAYANYAVSEADLLIAVGARFDDRTVGKLSHFATHAKVIHID 303

Query: 82 IDSAEIGKNKTPHVSVCQDVKLALQGMNKLLENRA-----EELKLDGFGVWRN 128  
ID AEIGKN P V + GD K L M + EN E++KL WR  
Sbjct: 304 IDPAEIGKNVNPDPVIGDAKSVLADMICLAENNGCISEPWLEQVKL----WRT 353

>ref|YP\_063568.1| acetohydroxyacid synthase large subunit [Gracilaria tenuistipitata  
var. liui]  
gb|AAT79643.1| acetolactate synthase large subunit [Gracilaria tenuistipitata var.  
liui]  
Length = 596



Score = 118 bits (295), Expect = 3e-25, Method: Compositional matrix adjust.  
Identities = 56/114 (49%), Positives = 77/114 (67%), Gaps = 2/114 (1%)

Query: 28 DELSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTVGKLEAFASRAKIVHIDIDSAEI 87  
D L L MLGMHGT YAN+AV DLL+A G RFDDRVTGKL+ FA AK++HIDID AEI  
Sbjct: 262 DNLCLGMLGMHGTAYANFAVSECDLLIALGARFDDRVTGKLDEFACNAKVIHIDIDPAEI 321

Query: 88 GKNKTPHVSVC G DVKLALQGMNKVLENRAEEL--KLD FGVWRNELNVQKQKFPL 139  
GKN+ P V++ GDVK + + L+N + + W++ +++ + ++PL  
Sbjct: 322 GKNRVPQVAILGDVKQVMVQILSYLDNSSNFVNSQQTYSWKDRISMWRDQYPL 375

>ref|NP\_050806.1| acetohydroxyacid synthase large subunit [Guillardia theta]  
sp|O78518.1|ILVB\_GUITH RecName: Full=Acetolactate synthase large subunit; Short=AHAS;  
AltName: Full=Acetohydroxy-acid synthase large subunit;  
Short=ALS  
gb|AAC35740.1| acetohydroxyacid synthetase large subunit [Guillardia theta]  
Length = 575

Score = 118 bits (295), Expect = 3e-25, Method: Compositional matrix adjust.  
Identities = 55/110 (50%), Positives = 77/110 (70%)

Query: 30 LSLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTVGKLEAFASRAKIVHIDIDSAEIGK 89  
LSL MLGMHGT VYANYAV DLL+A G RFDDRVTGKL+ FA A+++H+DID AEIGK  
Sbjct: 262 LSLGMLGMHGT VYANYAVSECDLLIALGARFDDRVTGKLDEFACHAQVIHVIDIDPAEIGK 321

Query: 90 NKT PHVSVC G DVKLALQGMNKVLENRAEELKLDFGVWRNELNVQKQKFPL 139  
N+TP + + G++K ++ + + L+N WR+ + ++++PL  
Sbjct: 322 NRTPQIGIVGEIKDFVRDLIECLKNDINFDSESQAWRSRIIRWRKEYPL 371

>ref|YP\_475476.1| acetolactate synthase 3 catalytic subunit [Synechococcus sp.  
JA-3-3Ab]  
gb|ABD00213.1| acetolactate synthase, large subunit, biosynthetic type  
[Synechococcus sp. JA-3-3Ab]  
Length = 598

Score = 118 bits (295), Expect = 3e-25, Method: Composition-based stats.  
Identities = 60/133 (45%), Positives = 82/133 (61%), Gaps = 2/133 (1%)

Query: 21 EGSYPCDEL SLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTVGKLEAFASRAKIVHI 80  
+G++ +D LS+ MLGMHGT YAN+AV DLL+A G RFDDRVTGKL+ FA AK++HI  
Sbjct: 243 KGAFDENDPLSVGMLGMHGTAYANFAVTECDLLIAVGARFDDRVTGKLDEFKAKHAKVIHI 302

Query: 81 DIDS AEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLDFGVWRNELNVQKQKFPLS 140  
DID AE+GKN+TP V + GDV+ L + L ++ W + K+ +PL  
Sbjct: 303 DIDPAEVGKNRTPDVPIVGDVRTVLVELLAQL-SKMPHTPGQTQAWLERIERWKRDPQLQ 361

Query: 141 FKTFGEAVPCPSK 153  
+ E V P +  
Sbjct: 362 VPRY-EGVIAPQQ 373

>ref|YP\_476898.1| acetolactate synthase 3 catalytic subunit [Synechococcus sp.  
JA-2-3B'a(2-13)]  
gb|ABD01635.1| acetolactate synthase, large subunit, biosynthetic type  
[Synechococcus sp. JA-2-3B'a(2-13)]  
Length = 598

Score = 118 bits (295), Expect = 3e-25, Method: Composition-based stats.  
Identities = 60/133 (45%), Positives = 82/133 (61%), Gaps = 2/133 (1%)

Query: 21 EGSYPCDEL SLHMLGMHGT VYANYAVEHSDLLLA FGVRFD DRTVGKLEAFASRAKIVHI 80  
+G++ +D LS+ MLGMHGT YAN+AV DLL+A G RFDDRVTGKL+ FA AK++HI  
Sbjct: 243 KGAFDENDPLSVGMLGMHGTAYANFAVTECDLLIAVGARFDDRVTGKLDEFKAKHAKVIHI 302

Query: 81 DIDS AEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLDFGVWRNELNVQKQKFPLS 140  
DID AE+GKN+TP V + GDV+ L + L ++ W + K+ +PL  
Sbjct: 303 DIDPAEVGKNRTPDVPIVGDVRTVLVELLAQL-SKMPHTPGQTQAWLERIERWKRDPQLQ 361

Query: 141 FKTFGEAVPCPSK 153  
+ E V P +  
Sbjct: 362 VPRY-EGVIAPQQ 373

>ref|ZP\_01469460.1| acetolactate synthase III large subunit [Synechococcus sp. BL107]  
gb|EAU70474.1| acetolactate synthase III large subunit [Synechococcus sp. BL107]  
Length = 587

Score = 117 bits (294), Expect = 4e-25, Method: Compositional matrix adjust.  
Identities = 55/121 (45%), Positives = 82/121 (67%), Gaps = 2/121 (1%)

Query: 21 EGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHI 80  
+G++ + L++ MLGMHGT YAN+AV DLL+A G RFDDRVTGKL+ FA RAK++H  
Sbjct: 265 KGAFDENAPLAVGMLGMHGTAYANFAVTECDLLIAGGARFDDRVTGKLDTFAPRAKVIHF 324

Query: 81 DIDS AEIGKNKTPHVS VCGDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLS 140  
+ID AE+GKN+T V+V GD+ L+++ + ++ R + W ++ KQ++PLS  
Sbjct: 325 EIDPAEVEGKNRTAEVAVLGDLSLSMERLVQLSMER--QANPSTAAWLEQIQDWKQRYPLS 382

Query: 141 F 141

Sbjct: 383 V 383

>gb|AAA26594.1| acetohydroxy acid synthase (AHAS)  
Length = 612

Score = 117 bits (294), Expect = 4e-25, Method: Composition-based stats.  
Identities = 60/134 (44%), Positives = 83/134 (61%), Gaps = 11/134 (8%)

Query: 21 EGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHI 80  
+GS+ LS+ MLGMHGT YAN+AV D L+A G RFDDRVTGKL+ F SRAK++HI  
Sbjct: 255 KGSFDEKHP LSVGMLGMHGTAYANFAVSECDFLIAGGARFDDRVTGKLDEF G SRAKVIHI 314

Query: 81 DIDS AEIGKNKTPHVS VCGDVKLALQGMNKVLENRAEELKLD FGV-----VWRNELNVQK 134  
DID AE+GKN+TP V + GDV+ L + + R E+ D G W +N  
Sbjct: 315 DIDPAEVEGKNRTPEVP IGVGDVRQVLIHLLR----RCREIG-DVGNDNQTSWLERINRWP 369

Query: 135 QKFPLSFKTFGEAV 148  
+ +PL ++ +++  
Sbjct: 370 EDYPLVVP SYSDSL 383

>ref|YP\_684797.1| acetolactate synthase, large subunit [uncultured methanogenic  
archaeon RC-I]  
emb|CAJ35471.1| acetolactate synthase, large subunit [uncultured methanogenic  
archaeon RC-I]  
Length = 566

Score = 117 bits (293), Expect = 5e-25, Method: Compositional matrix adjust.  
Identities = 62/122 (50%), Positives = 82/122 (67%), Gaps = 6/122 (4%)

Query: 21 EGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHI 80  
+G++P D LSL MLGMHGT YANY++ DLL+A G RF DRV T G K L E + F A AKIVHI  
Sbjct: 246 KGAFPEDHPLSLGMLGMHGT RYANYSIIDCDLLIAGARFSDRV T G K L E S F A P N A K I V H I 305

Query: 81 DIDS AEIGKNKTPHVS VCGDVKLALQGMNKVLENRAEELKLD FGV-WRNELNVQKQKFPL 139  
DID AE+GKN V + GD K+ L+ + L +LKL W +++ K+++PL  
Sbjct: 306 DIDPAEMGKNVVDVPIVGDAK VILKKIISTL-----KLKLPKSAEWLKKIDLWKREYPL 360

Query: 140 SF 141

F  
Sbjct: 361 DF 362

>ref|YP\_001736021.1| acetolactate synthase, large subunit, biosynthetic type  
[Synechococcus sp. PCC 7002]

gb|ACB00766.1| acetolactate synthase, large subunit, biosynthetic type  
[Synechococcus sp. PCC 7002]  
Length = 591

Score = 117 bits (293), Expect = 5e-25, Method: Compositional matrix adjust.  
Identities = 61/132 (46%), Positives = 84/132 (63%), Gaps = 9/132 (6%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G++ + LS+ MLGMHG+ YAN+AV DLL+A G RFDDRVTKGL+ FAS AK++HID  
Sbjct: 263 GAFDEHNPLSVGMLGMHGSAYANFAVSECDLLIAVGARFDDRVTKGLDEFASEAKVIHID 322

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLDG----VWRNELNVQKQK 136  
ID AE+GKN+ P V++ GDV+ L M + L + L G W N++ K+K  
Sbjct: 323 IDPAEVGKNRLPEVAIVGDVRHVLIELMLQQLGKG---LPLGNKPTTEWLNKIQGWKEK 378

Query: 137 FPLSFKTFGEAV 148  
PL F +++  
Sbjct: 379 HPLVAPRFADSI 390

>ref|YP\_754808.1| acetolactate synthase large subunit [Syntrophomonas wolfei subsp.  
wolfei str. Goettingen]  
gb|ABI69437.1| acetolactate synthase, large subunit [Syntrophomonas wolfei subsp.  
wolfei str. Goettingen]  
Length = 562

Score = 117 bits (293), Expect = 5e-25, Method: Compositional matrix adjust.  
Identities = 60/130 (46%), Positives = 80/130 (61%), Gaps = 3/130 (2%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHID 81  
G + + LSL MLGMHG+ YANYA+ DLL+A GVRFDDRVTK++ FA A+++HID  
Sbjct: 242 GGFGLNSYLSLGMHGTTRYANYAIGECDDLLIAVGVRFDDRVTKIDTFAPHARVIHID 301

Query: 82 IDSAEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKLDG VWRNELNVQKQK FPLSF 141  
ID+AEIGKN V + G VK L +N+ LE E + W ++ K+++PL +  
Sbjct: 302 IDAAEIGKNVEVEVPIVQVKEVLAAINQRLAIESE--ELNEWHETIHRWKEEYPLRY 358

Query: 142 KTFGEAVPCP 151  
E P  
Sbjct: 359 GDSSEGRIMP 368

>ref|YP\_002620543.1| acetolactate synthase, large subunit, biosynthetic type  
[Microcoleus chthonoplastes PCC 7420]  
gb|EDX75556.1| acetolactate synthase, large subunit, biosynthetic type  
[Microcoleus chthonoplastes PCC 7420]  
Length = 619

Score = 117 bits (293), Expect = 6e-25, Method: Composition-based stats.  
Identities = 59/132 (44%), Positives = 82/132 (62%), Gaps = 7/132 (5%)

Query: 21 EGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTKLEAFASRAKIVHI 80  
+G++ LS+ MLGMHG+ YAN+AV DLL+A G RFDDRVTKGL+ FA AK++HI  
Sbjct: 262 KGAFDESHPLSVGMLGMHGTAYANFAVSECDLLIAVGARFDDRVTKGLDQFACHAKVIHI 321

Query: 81 DIDS AEIGKNKTPHVSVC G DVKLALQGMNKVLENRAEELKL----DFGVWRNELNVQKQK 136  
DID AE+GKN+ P V + GDV+ Q + K+L E+ L VW + +Q+  
Sbjct: 322 DIDPAEVGKNRAPQVPIVGDVQ---QVLAKLLRRYEEQDILADPNQTQVWLKRIQRWRQE 378

Query: 137 FPLSFKTFGEAV 148  
+PL + + +  
Sbjct: 379 YPLEAPQYPTI 390

>ref|YP\_001568611.1| acetolactate synthase, large subunit, biosynthetic type [Petrogla  
mobilis SJ95]  
gb|ABX32288.1| acetolactate synthase, large subunit, biosynthetic type [Petrogla  
mobilis SJ95]  
Length = 569

Score = 117 bits (293), Expect = 6e-25, Method: Compositional matrix adjust.  
Identities = 57/121 (47%), Positives = 80/121 (66%), Gaps = 6/121 (4%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHID 81  
G +P D ELSL MLGMHGT+YANYA+ +DL++ GVRFDDRV TGKLE FA AKI+HID  
Sbjct: 242 GVFPEDHLESLKMLGMHGTMYANYAISEADLIIGVGVRFDDRV TGKLETFAPHAKIIHID 301

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
ID AEI KN + + G+ K L + +++ ++ W ++ K++FPL++  
Sbjct: 302 IDPAEINKNVKVDIPIVGNAKNVLDKLIPLIKT-----IERKEWLEQIKEWKR RFP LNY 355

Query: 142 K 142  
+  
Sbjct: 356 E 356

>ref|YP\_303782.1| acetolactate synthase 3 catalytic subunit [Methanosarcina barkeri str. Fusaro]  
gb|AAZ69202.1| acetolactate synthase, large subunit [Methanosarcina barkeri str. Fusaro]  
Length = 588

Score = 117 bits (293), Expect = 6e-25, Method: Compositional matrix adjust.  
Identities = 56/130 (43%), Positives = 81/130 (62%), Gaps = 6/130 (4%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHID 81  
G+ P + L + MLGMHGT ANYA++ SDL++A G RFDDRV TGKLE+FA A+++HID  
Sbjct: 247 GAIPTEHPLHVGMGLGMHGT KSANYAIQESDLIIAVGARFDDRV TGKLESFAPNARVIHID 306

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFP LSF 141  
+D AEI KN + + GD K L+ +NK ++ E W ++ K+++PL++  
Sbjct: 307 VDPAEISKNVNVQIPIVGDAKQVLKSLNKYIKCKSE-----AWLEKIKQWKKEYPLTY 360

Query: 142 KTFGEAVPCP 151  
K V P  
Sbjct: 361 QMSSSDVIMP 370

>ref|YP\_722490.1| acetolactate synthase 3 catalytic subunit [Trichodesmium erythraeum IMS101]  
gb|ABG52017.1| acetolactate synthase, large subunit [Trichodesmium erythraeum IMS101]  
Length = 587

Score = 117 bits (293), Expect = 6e-25, Method: Compositional matrix adjust.  
Identities = 63/146 (43%), Positives = 88/146 (60%), Gaps = 9/146 (6%)

Query: 8 KKFFIFFHFFHQHEGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGK 67  
++F I +G++ LS+ MLGMHGT YANYAV DLL+A G RFDDRV TGK  
Sbjct: 242 ERFNIPVTTTLMGKGAFNEHHPLSVGMLGMHGTAYANYAVSECDLLIAGV GARFDDRV TGK 301

Query: 68 LEAFASRAKIVHIDIDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGF--- 124  
L+ FAS AK++HIDID AE+GKN++P V + GDV+ L M LE R +++ +  
Sbjct: 302 LDQFASYAKVIHIDIDPAEIVGKNRSPEVPIVGDVVRQVLIDM---LE-RCQKIDTNVSSNQ 357

Query: 125 --VWRNELNVQKQKFP LSFKTFGEAV 148  
W + KQ +PL + +++  
Sbjct: 358 TQTWLERIKRWKQDYPLVVP HYS DSM 383

>ref|ZP\_02179310.1| acetolactate synthase large subunit [Hydrogenivirga sp. 128-5-R1-1]  
gb|EDP73924.1| acetolactate synthase large subunit [Hydrogenivirga sp. 128-5-R1-1]  
Length = 567

Score = 116 bits (291), Expect = 1e-24, Method: Compositional matrix adjust.  
Identities = 61/121 (50%), Positives = 77/121 (63%)

Query: 21 EGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRV TGKLEAFASRAKIVHI 80

+G++P D L+LHMLGMHGT YAN AV +SDLL+A G RFDDRVTGK++ FA AKI+HI  
Sbjct: 245 KGAFPEDHPLALHMLGMHGTYYANMAVYNSDLLIAVGSRFDDRVTGKIDFAPEAKIIHI 304

Query: 81 DIDS AEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLS 140  
DID A I K V + GDVK LQ + K LE + E W + K + PL+

Sbjct: 305 DIDPASISKITINVDPVIGDVKKVLQKLIKELERKPVEWDEARES WIKLIEKWKIEHPLT 364

Query: 141 F 141

+  
Sbjct: 365 Y 365

>ref|NP\_618663.1| acetolactate synthase 3 catalytic subunit [Methanosarcina  
acetivorans C2A]  
gb|AAM07143.1| acetolactate synthase, large subunit [Methanosarcina acetivorans  
C2A]  
Length = 564

Score = 116 bits (290), Expect = 1e-24, Method: Compositional matrix adjust.  
Identities = 57/122 (46%), Positives = 81/122 (66%), Gaps = 10/122 (8%)

Query: 23 SYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHIDI 82  
+ P + L + MLGMHG YANYAV+ SDL++A G RFDDRVTGKLE+FA A+++HID+  
Sbjct: 248 AIPTEHPLYVGMMLGMHCKYANYAVQESD LIIAVGARFDDRVTGKLESFAPNARVIHIDV 307

Query: 83 DSAEIGKNKTPHVSVC GDVKLALQGMNKVLE--NRAEELKLD FGVWRNELNVQKQKFPLS 140  
D AEI KN HV + GD K L+ + K ++ N AE W ++N K+++PL+  
Sbjct: 308 DP AEISKNVKVHVPVIGDAKQILKSLIKYIQCNSAE-----WIEKINQWKKEYPLA 359

Query: 141 FK 142

++  
Sbjct: 360 YR 361

>ref|NP\_789495.1| acetolactate synthase 1 catalytic subunit [Tropheryma whippiei  
TW08/27]  
emb|CAD67233.1| acetolactate synthase [Tropheryma whippiei TW08/27]  
Length = 581

Score = 116 bits (290), Expect = 1e-24, Method: Compositional matrix adjust.  
Identities = 60/120 (50%), Positives = 76/120 (63%), Gaps = 3/120 (2%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G P + L+L M GMHGT V A A++ SDLL+A G RFDDRVTG+L++FA A++VH+D  
Sbjct: 246 GVLPSNRLNLGMPGMHGTVAAILALQKSDLLIALGTRFDDRVTGRLDSFAQNARVVHVD 305

Query: 82 IDSAEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLSF 141  
ID AEIGKN+ P + V GDVK L +N ENR D W L +Q +PL F  
Sbjct: 306 IDPAEIGKNRRPDLFVIGDVKEVLGYINSHSENRTTG--DLTAWRSLEKMRQIYPLVF 362

>ref|YP\_537012.1| acetohydroxyacid synthase large subunit [Porphyra yezoensis]  
sp|Q1XDF6.1|ILVB\_PORYE RecName: Full=Acetolactate synthase large subunit; Short=AHAS;  
AltName: Full=Acetohydroxy-acid synthase large subunit;  
Short=ALS  
dbj|BAE92455.1| acetohydroxyacid synthase large subunit [Porphyra yezoensis]  
Length = 590

Score = 116 bits (290), Expect = 1e-24, Method: Compositional matrix adjust.  
Identities = 53/119 (44%), Positives = 76/119 (63%)

Query: 21 EGSYPCDDELSLHMLGMHGT VYANYAVEHSDLLAFGVRFDDRVTGKLEAFASRAKIVHI 80  
+G + D + L MLGMHGT YAN+AV DLL+A G RFDDRVTGKL+ FA A+++H+  
Sbjct: 254 KGIFNEDSDYCLGMLGMHGTAYANFAVSECDLLIALGARFDDRVTGKLDEFACNAQVIHV 313

Query: 81 DIDS AEIGKNKTPHVSVC GDVKLALQGMNKVLENRAEELKLD FGVWRNELNVQKQKFPL 139  
DID AE+GKN+ P V++ GDV + + +L+ W+ +N +Q++PL  
Sbjct: 314 DIDPAEIVGNRIPQVAIVGDVAEIVVSEILNLLKTSFPPYPEQIISWQERINRWRQYPL 372

>ref|ZP\_03155476.1| acetolactate synthase, large subunit, biosynthetic type [Cyanothec  
sp. PCC 7822]  
gb|EDX97152.1| acetolactate synthase, large subunit, biosynthetic type [Cyanothec  
sp. PCC 7822]  
Length = 619

Score = 115 bits (289), Expect = 2e-24, Method: Composition-based stats.  
Identities = 58/128 (45%), Positives = 80/128 (62%), Gaps = 1/128 (0%)

Query: 22 GSYPCDDELSLHMLGMHGT VYANYAVEHSDLLLAFGVRFDDRVTGKLEAFASRAKIVHID 81  
G++ LSL MLGMHGT YAN+AV DLL+A G RFDDRVTGKLE FA AK++HID  
Sbjct: 263 GAFDEHHPLSLGMLGMHGTAYANFAVTECDLLIAVGARFDDRVTGKLEEFARNAKVIHID 322

Query: 82 IDSAEIGKNKTPHVSVC GDVKLAL-QGMNKVLENRAEELKLD FGVWRNELNVQKQKFPLS 140  
ID AE+GK + P V + GDV+ L Q + +V E W + +N ++ +PL  
Sbjct: 323 IDPAEVGKVRGPDVPIVGDVRQVLEQILQVRVIDYPTNSTRTKAWLDRINRWREDYPLM 382

Query: 141 FKTFGEAV 148  
++ +++  
Sbjct: 383 VPSYPDSI 390

Database: All non-redundant GenBank CDS  
translations+PDB+SwissProt+PIR+PRF excluding environmental samples  
from WGS projects  
Posted date: Mar 2, 2009 5:57 PM  
Number of letters in database: 2,739,991,458  
Number of sequences in database: 7,946,514

Lambda	K	H
0.324	0.139	0.436

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62  
Gap Penalties: Existence: 11, Extension: 1  
Number of Sequences: 7946514  
Number of Hits to DB: 1,100,448,215  
Number of extensions: 44460244  
Number of successful extensions: 92720  
Number of sequences better than 1.0: 2285  
Number of HSP's gapped: 92672  
Number of HSP's successfully gapped: 2285  
Length of query: 199  
Length of database: 2,739,991,458  
Length adjustment: 128  
Effective length of query: 71  
Effective length of database: 1,722,837,666  
Effective search space: 122321474286  
Effective search space used: 122321474286  
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: 41 (22.0 bits)  
S2: 84 (37.0 bits)