

Study Title

**Bioinformatics Analysis of CP4 EPSPS Protein Sequence
Utilizing an Allergen Database**

Authors

Ronald E. Hileman, Ph.D. and James D. Astwood, Ph.D.

Study Completed On

September 29, 1999

Performing Laboratory

**Monsanto Company
Biotechnology Regulatory Sciences
700 Chesterfield Parkway North
St. Louis, MO 63198**

Laboratory Project ID

MSL Number: 16267

Study Number: 99-01-62-07

Statement of No Data Confidentiality Claim

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).

"We submit this material to the United States Environmental Protection Agency specifically under provisions contained in FIFRA as amended, and thereby consent to use and disclosure of this material by EPA according to FIFRA. In submitting this material to the EPA according to method and format requirements contained in PR Notice 86-5, we do not waive any protection of rights involving this material that would have been claimed by the company if this material had not been submitted to the EPA."

Company: Monsanto Company

Company Agent: _____

Title: _____

Signature: _____ Date: _____

MSL Number: 16267

Title: Bioinformatics Analysis of CP4 EPSPS Protein Sequence
Utilizing an Allergen Database

Facility: Monsanto Company
Biotechnology Regulatory Sciences
700 Chesterfield Parkway North
Chesterfield, Missouri 63198

Sponsor: Ravinder Sidhu

Authors: Ronald E. Hileman
James D. Astwood

Contributors: Richard E. Goodman
Conrad H. Halling

Study Start Date: September 13, 1999

Study Completion Date: September 29, 1999

Records Retention: All study specific raw data and final report will be retained at
Monsanto-St. Louis.

Signatures of Final Report Approval:

Author: R. E. Hileman

Date: 29 SEPT 1999

Author: J. D. Astwood

Date: Sept 29, 1999

Sponsor: Ravinder S. Sidhu

Date: 29 Sep 99

Table of Contents

Section	Page
Title Page.....	1
Statement of No Data Confidentiality Claims.....	2
Signatures of Final Report Approval.....	3
Table of Contents	4
Abbreviations and Definitions	6
1.0 Summary	7
2.0 Introduction	7
3.0 Purpose	8
4.0 Methods.....	9
4.1 Allergen and gliadin database preparation (UPDATE2).....	9
4.2 Allergen database search for sequences similar to CP4 EPSPS protein using FASTA	9
4.3 Allergen database epitope sequence comparisons to CP4 EPSPS protein sequences using IDENTITYSEARCH	10
5.0 Results and Discussion.....	10
5.1 Allergen and gliadin database (UPDATE2).....	10
5.2 CP4 EPSPS protein sequence is not similar to allergens and gliadins.....	10
5.3 CP4 EPSPS protein sequence does not contain immunologically significant allergen epitopes.....	11
6.0 Conclusions	11
7.0 References	12

Appendices

Appendix 1.	CP4 EPSPS protein sequence.....	14
Appendix 2.	Allergen and gliadin protein sequences.....	15
Appendix 3.	FASTA sequence similarity alignments of CP4 EPSPS protein with allergens and gliadins.....	27
Appendix 4.	Linear sequence scanning for immunologically significant similarities between CP4 EPSPS protein and allergens/gliadins	32

Abbreviations and Definitions

aa	amino acid
<i>cp4 epsps</i>	gene encoding the <i>Agrobacterium</i> sp. strain CP4 5-enolpyruvylshikimate-3-phosphate synthase protein
CP4 EPSPS	Protein produced from the <i>cp4 epsps</i> gene
EMBL	European Molecular Biology Laboratory
FASTA	Algorithm used to find local high scoring alignments between a pair of protein or nucleotide sequences
GCG	Genetics Computer Group
GenBank	A public genetic database maintained by the National Center for Biotechnology Information
GenPept	GenBank peptide database
IDENTITYSEARCH	Algorithm used to find immunologically relevant amino acid identities between a pair of protein sequences
NRL3D	National Research Laboratories, 3D Protein database at Brookhaven
PIR	Protein Information Resource
RR	Roundup Ready®
SwissProt	Translated sequences from EMBL database
UPDATE2	Allergen and gliadin database

1.0 Summary

In the safety assessment of proteins introduced into genetically modified crops, possible animal and human health effects are addressed in a multifaceted approach which involves addressing the history of safe use, directly assessing the oral acute toxicity of the introduced protein, assessing the sensitivity to rapid digestion and comparing the amino acid sequence of the introduced protein with proteins associated with toxicity, allergenicity or other health effects. A biologically relevant sequence similarity to a known allergen or gliadin (*i.e.* a sequence derived from a common ancestor gene and/or containing a potential allergenic epitope) may indicate that additional immunological assessments should be performed.

A database of protein sequences associated with allergy and coeliac disease was assembled from publicly available genetic databases (GenBank, EMBL, PIR and SwissProt) and from current literature. The keyword "allergen" was used to retrieve allergen sequences from the public domain databases. Additional unique allergens found in only current literature were appended creating a database containing 567 unique protein sequences. The amino acid sequence of the CP4 EPSPS protein was compared to these sequences using the sequence alignment tool FASTA. The test protein sequence, CP4 EPSPS, shared no structurally significant sequence similarity to sequences within the allergen database.

In addition, the amino acid sequence of the CP4 EPSPS protein was compared to the allergen database using an algorithm (IDENTITYSEARCH) that scans for a window of 8 linearly contiguous identical amino acids. The CP4 EPSPS protein sequence does not share 8 linearly contiguous amino acid identities to any sequence in the allergen database. These data establish that the CP4 EPSPS protein does not share any immunologically significant sequence similarity to proteins associated with allergy or coeliac disease.

2.0 Introduction

Monsanto Company has developed Roundup Ready[®] corn line NK603 which is tolerant to glyphosate (the active ingredient in Roundup[®] herbicide) at the whole plant level. Corn line NK603 contains a 5-enolpyruvylshikimate-3-phosphate synthase protein from *Agrobacterium* sp. strain CP4 (CP4 EPSPS). Corn plants that demonstrate commercial level tolerance to Roundup herbicide are called Roundup Ready[®] (RR). The CP4 EPSPS gene from *Agrobacterium* sp. strain CP4 has been completely sequenced and encodes a 47.6-kDa protein consisting of a single polypeptide of 455 amino acids (Padgett et al., 1996). The CP4 EPSPS protein is functionally similar to plant EPSPS enzymes but has a much reduced affinity for glyphosate (Padgett et al., 1993). In nontransgenic plants, glyphosate binds to the plant EPSPS enzyme and blocks the biosynthesis of aromatic amino acids thereby depriving plants of these essential nutrients (Haslam, 1993; Steinrucken and

Amrhein, 1980). In Roundup Ready plants, nutritional requirements for normal growth and development are met by the continued action of the glyphosate-tolerant CP4 EPSPS enzyme in the presence of glyphosate. A comprehensive safety assessment of the CP4 EPSPS protein has been described in the literature (Harrison et al., 1996).

Corn line NK603 was produced by transformation of corn tissue with a 6.7-kb linear DNA fragment PV-ZMGT32L derived from the plasmid vector PV-ZMGT32, using a particle acceleration method. Molecular analysis (Deng et al., 1999) has shown that corn line NK603 contains a single DNA insert consisting of two expression cassettes: the first CP4 EPSPS gene cassette, containing the CP4 EPSPS coding sequence under regulation of the rice actin promoter and intron (P-ract1/ract intron), a chloroplast transit peptide (CTP2) sequence, and a nopaline synthase (NOS) 3' polyadenylation sequence; and second CP4 EPSPS gene cassette, containing the CP4 EPSPS coding sequence under the regulation of the cauliflower mosaic virus (CaMV) enhanced 35S plant promoter (e35S), a maize heat-shock protein 70 (*Zmhsp70*) intron, CTP2 and the NOS 3' polyadenylation sequence.

This report describes the bioinformatics assessment of the CP4 EPSPS protein which demonstrated the absence of sequence homology as well as the absence of immunologically significant sequence similarity to proteins identified as allergens or gliadins. The complete CP4 EPSPS protein sequence is shown in Appendix 1. A database of 567 unique protein sequences identified as allergens or gliadins was assembled from publicly available genetic databases (GenBank, EMBL, PIR and SwissProt) and current literature. The extent of sequence similarity (*i.e.* structural similarity) to allergens and gliadins was assessed using the FASTA sequence alignment tool. A high degree of sequence similarity may indicate that the protein was derived from a common ancestor gene. Proteins structurally unrelated to allergens and gliadins may still contain immunologically significant epitopes. An immunologically significant sequence was defined as 8 linearly contiguous identical amino acids (Metcalf et al., 1996). CP4 EPSPS protein was assessed for immunologically relevant sequences using a linear identity computer search algorithm (IDENTITYSEARCH).

3.0 Purpose

The purpose of this study was to evaluate the amino acid sequence similarity of the CP4 EPSPS protein to protein sequences within an allergen database using bioinformatics techniques. Structural similarities were assessed using the sequence alignment tool, FASTA. Immunologically relevant similarities were assessed using an algorithm, IDENTITYSEARCH. A structurally relevant and/or immunologically relevant sequence similarity to a known allergen or gliadin protein (*i.e.* a sequence containing 8 or more linearly contiguous amino acid identities) may indicate that additional immunological assessments be initiated to address the relevance of the observed similarities.

4.0 Methods

- 4.1 *Allergen and gliadin database preparation (UPDATE2).* The allergen and gliadin database was assembled from public domain databases (Genbank and EMBL GenPept version 108, PIR and NRL3D version 56 and SwissProt version 36). A preliminary list (AllPeptides_121.strings) of 804 sequences was compiled with the keyword "allergen" using the STRINGSEARCH function of GCG, version 9.1. The preliminary list contained duplicate sequences as well as non-allergen "hits". Non-allergen "hits" (40 sequences) were identified by retrieving the entire flat file of each suspected sequence and the remainder of the preliminary list sorted according to allergen type. The modified list (AllPeptides_121_ver1.strings) was compared to a previous allergen database (Metcalf et al., 1996) and sequences from the previous allergen database not found in the modified list were merged, creating a list of 829 sequences (AllPeptides_121_ver2.strings). Duplicate sequences were identified and removed by performing a FASTA search on each individual sequence against the version 2 list (AllPeptides_121_ver2.strings). A sequence was considered unique if a single amino acid difference existed when compared to the remainder of the list. The file containing no duplicate sequences (AllPeptides_121_ver3.strings) contained a list of 487 unique allergen and gliadin protein sequences. This list was used to compile a preliminary allergen database, UPDATE2, using DATASET of the GCG software (version 10.0).

The allergen database was finalized by merging 80 additional allergen sequences identified by: i) comparison of the UPDATE2 sequences to the publicly available list located on the Internet site, <ftp://biobase.dk/resources/pub/who-iuis/allergen.list>; and ii) performing an extensive search of the current literature using the publicly available Internet PubMed and Entrez search engines. In cases where only the DNA sequence was retrieved, the amino acid code was obtained using TRANSLATE of the GCG software (version 10.0). These 80 additional sequences were appended using the APPEND command of DATASET in GCG (version 10.0), creating an allergen and gliadin database containing 567 unique protein sequences.

- 4.2 *Allergen database search for sequences similar to CP4 EPSPS protein using FASTA.* The SHUFFLE function of GCG (version 10.0) was used to generate a randomized sequence of the CP4 EPSPS test protein. This randomized sequence was used to test for amino acid bias (negative control) in the similarity search of the parent CP4 EPSPS sequence against UPDATE2. The allergen and gliadin database was searched using FASTA (Pearson and Lipman, 1988). A significant sequence similarity was defined as a sequence identity of at least 8 linearly contiguous amino acids.

- 4.3** *Allergen database epitope sequence comparisons to CP4 EPSPS protein sequence using IDENTITYSEARCH.* An algorithm was developed to identify whether or not a linearly contiguous match of 8 amino acids existed between the test protein sequence and sequences within the allergen database (UPDATE2). The algorithm was run from a UNIX terminal window in the GCG software. This program compares the test protein sequence to each protein sequence in the allergen database using a sliding window of 8 amino acids. Specifically, the test sequence was compared to each allergen or gliadin protein sequence and if a match of 8 (or more) linearly contiguous amino acid identities were found, the match is printed in the file output. An epitope of 8 amino acids was chosen to represent the smallest likely immunologically significant IgE binding epitope (Metcalf et al., 1996).

5.0 Results and Discussion

- 5.1** *Allergen and gliadin database (UPDATE2).* A combination of keyword searches using the STRINGSEARCH function of GCG and Internet searches against publicly available genetic and literature databases was performed to create a database (UPDATE2) containing 567 unique allergen and gliadin protein sequences (see Section 4.1 for a full description). The entire list of proteins in UPDATE2 is shown in Appendix 2. This database represents a comprehensive list of allergens and gliadins.
- 5.2** *CP4 EPSPS protein sequence is not similar to allergens and gliadins.* A similarity search of the allergen and gliadin database (UPDATE2) revealed no sequences with significant similarity to the CP4 EPSPS protein. The FASTA sequence alignment tool was used to identify overall similarity (structural similarity) as well as identify local (immunologically relevant) similarity to proteins within the allergen and gliadin database. The criteria for a immunologically relevant similarity was defined as an exact match to 8 or more linearly contiguous amino acids (Metcalf et al., 1996).

The best sequence similarity generated to the CP4 EPSPS protein sequence and the database was an 82 amino acid overlap to the *Dermatophagoides farinae* (mite) protein, Der f 2. This protein shares 30.5% identical residues and 62.2% similar residues within the overlap. The level of similarity is not biologically relevant (Doolittle, 1990) and does not indicate structural homology. Further, a local similarity of 8 or more amino acids was not observed between the CP4 EPSPS protein sequence and sequences within the UPDATE2 database. These data demonstrate the lack of both structurally and immunologically relevant

sequence similarity between the CP4 EPSPS protein sequence and proteins within the allergen and gliadin database.

The complete FASTA output, including sequence alignments, for the CP4 EPSPS search is shown in Attachment 3. The negative control sequence (shuffled CP4 EPSPS protein sequence) also revealed no significant similarities. It was observed that the CP4 EPSPS protein was no more likely to share sequence similarity with allergens and gliadins than the randomized (amino acid shuffled) CP4 EPSPS protein sequence.

5.3 *CP4 EPSPS protein sequence does not contain immunologically significant allergen epitopes.* An algorithm was developed to compare the CP4 EPSPS protein to the allergen database. This program was used to scan for a match of 8 (or more) linearly contiguous identical amino acids between a test sequence and any protein in the allergen database. This algorithm may be more sensitive than the FASTA sequence alignment tool. The window size of 8 amino acids has been previously justified in the literature (Metcalf et al., 1996).

The complete IDENTITYSEARCH output file for the CP4 EPSPS protein is shown in Appendix 4. No sequences in the allergen database (UPDATE2) were identified to contain a match of 8 or more linear amino acids to the CP4 EPSPS protein.

6.0 Conclusions

The retrieval strategy employed in this study was effective and allowed the assembly of a comprehensive allergen and gliadin database. The FASTA sequence alignment tool was used to compare the CP4 EPSPS protein sequence to sequences contained in this database. Results of the FASTA sequence alignments demonstrated the lack of structurally relevant similarity between the CP4 EPSPS protein and any known allergen or gliadin. Additionally, a computer algorithm (IDENTITYSEARCH) was developed to determine if the CP4 EPSPS protein sequence shared a match of 8 or more linearly contiguous amino acid identities to any sequence within the allergen and gliadin database. Results from this linear scanning algorithm demonstrated the lack of potential immunologically relevant sequences in the CP4 EPSPS protein sequence. Combined, these data demonstrate that the CP4 EPSPS protein is not similar to proteins associated with allergens or coeliac disease.

7.0 References

- Breitenbach, M., Simon, B., Probst, G., Oberkofler, H., Ferreira, F., Briza, P., Achatz, G., Unger, A., Ebner, C., Kraft, D. and Hirschwehr, R. 1997. Enolases are highly conserved fungal allergens. *Int Arch Allergy Immunol* 113, 114-7.
- Daul, C. B., Morgan, J. E. and Lehrer, S. B. 1993. Hypersensitivity reactions to crustacea and mollusks. *Clin Rev Allergy* 11, 201-22.
- Deng, M. Y., Lirette, R. P., Cavato, T. A. and Sidhu, R. S. 1999. Molecular Characterization of Roundup Ready Corn Line NK603. Study number MSL-16214, an unpublished study conducted by Monsanto Company.
- Doolittle, R. F. 1990. Searching through sequence databases. *Methods in Enzymology* 183, 99-110.
- Gupta, N., Martin, B. M., Metcalfe, D. D. and Rao, P. V. 1996. Identification of a novel hydroxyproline-rich glycoprotein as the major allergen in *Parthenium* pollen. *J Allergy Clin Immunol* 98, 903-12.
- Harrison, L. A., Bailey, M. R., Naylor, M. W., Ream, J. E., Hammond, B. G., Nida, D. L., Burnette, B. L., Nickson, T. E., Mitsky, T. A., Taylor, M. L., Fuchs, R. L. and Padgett, S. R. 1996. The expressed protein in glyphosate-tolerant soybean, 5-enolpyruvylshikimate-3-phosphate synthase from *Agrobacterium* sp. strain CP4, is rapidly digested *in vitro* and is not toxic to acutely gavaged mice. *J Nutr* 126, 728-40.
- Haslam, E. 1993. Shikimic Acid: Metabolism and Metabolites John Wiley and Sons, Chichester, England.
- Ishikawa, M., Ishida, M., Shimakura, K., Nagashima, Y. and Shiomi, K. 1998. Purification and IgE-binding epitopes of a major allergen in the gastropod *Turbo cornutus*. *Biosci Biotechnol Biochem* 62, 1337-43.
- Jensen-Jarolim, E., Santner, B., Leitner, A., Grimm, R., Scheiner, O., Ebner, C. and Breiteneder, H. 1998. Bell peppers (*Capsicum annuum*) express allergens (profilin, pathogenesis-related protein P23 and Bet v 1) depending on the horticultural strain. *Int Arch Allergy Immunol* 116, 103-9.

- Klysner, S., Welinder, K. G., Lowenstein, H. and Matthiesen, F. 1992. Group V allergens in grass pollens: IV. Similarities in amino acid compositions and NH₂-terminal sequences of the group V allergens from *Lolium perenne*, *Poa pratensis* and *Dactylis glomerata*. Clin Exp Allergy 22, 491-7.
- 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, S165-86.
- Monsalve, R. I., Gonzalez de la Pena, M. A., Lopez-Otin, C., Fiandor, A., Fernandez, C., Villalba, M. and Rodriguez, R. 1997. Detection, isolation and complete amino acid sequence of an aeroallergenic protein from rapeseed flour. Clin Exp Allergy 27, 833-41.
- Padgett, S., Re, D., Eichholtz, D., Delannay, X., Fuchs, R., Kishore, G. and Fraley, R. 1996. New weed control opportunities: Development of soybeans with a Roundup Ready™ gene, Duke, S. O., ed. CRC, Boca Raton, FL.
- Padgett, S. R., Barry, G. F., Re, D. B., Weldon, M., Eichholtz, D. A., Kolacz, K. and Kishore, G. M. 1993. Purification, cloning and characterization of a highly glyphosate-tolerant EPSP synthase from *Agrobacterium* sp. strain CP4. Study number MSL-12738, an unpublished study conducted by Monsanto Company.
- Pearson, W. and Lipman, D. 1988. Improved tools for biological sequence comparison. Proc Natl Acad Sci USA 85, 2440-2448.
- Shen, H., Lin, W., Tam, M. F., Wang, S., Tzean, S., Huang, M. and Han, S. 1999. Characterization of allergens from *Penicillium oxalicum* and *P. notatum* by immunoblotting and N-terminal amino acid sequence analysis. Clin Exp Allergy 29, 642-51.
- Steinrucken, H. C. and Amrhein, N. 1980. The herbicide glyphosate is a potent inhibitor of 5-enolpyruvyl- shikimic acid-3-phosphate synthase. Biochem Biophys Res Commun 94, 1207-12.

Appendix 1. CP4 EPSPS protein sequence.

(filename = CP4_EPSPS.pep)

```
1  MLHGASSRPA TARKSSGLSG TVRIPGDKSI SHRSFMFGGL ASGETRITGL
51  LEGEDVINTG KAMQAMGARI RKEGDTWIID GVGNGGLLAP EAPLDFGNAA
101 TGCRLTMGLV GVYDFDSTFI GDASLTKRPM GRVLNPLREM GVQVKSEDGD
151 RLPVTLRGPK TPTPITYRVP MASAQVKSAV LLAGLNTPGI TTVIEPIMTR
201 DHTEKMLQGF GANLTVETDA DGVRTIRLEG RGKLTGQVID VPGDPSSTAF
251 PLVAALLVPG SDVTILNVLM NPTRTGLILT LQEMGADIEV INPRLAGGED
301 VADLRVRSST LKGVTVPEDR APSMIDEYPI LAVAAFAEG ATVMNGLEEL
351 RVKESDRLSA VANGLKLNGV DCDEGETSLV VRGRPDGKGL GNASGAAVAT
401 HLDHRIAMSF LVMGLVSENP VTVDDATMIA TSFPEFMDLM AGLGAKIELS
451 DTKAA
```

Appendix 2. Allergen and gliadin protein sequences

Food Allergens & Gliadins - Plants

Species	Common name	Allergen	Synonym/Function/Comment	Accession No.
Actinidia chinensis	Kiwi	Act c 1	Cysteine protease	P00785
Apium graveolens	Celery	Api g 1	Similar to 17 kD tree allergens	S63984
Apium graveolens	Celery	Api g 2		Z75662
Arachis hypogaeae	Peanut	Ara h 1	Clone P17	P43237
Arachis hypogaeae	Peanut	Ara h 1	Clone P41b	P43238
Arachis hypogaeae	Peanut	Ara h 2	Conglutin	L77197
Arachis hypogaeae	Peanut	Ara h 3	Glycinin	AF093541
Arachis hypogaeae	Peanut		Lectin, phytohemagglutinin	S14765
Bertholletia excelsa	Brazil nut	Ber e 1	2S albumin	X54490
Brassica juncea	Leaf mustard	Bra j 1-e	2S albumin	P80207
Brassica napis	Rapeseed	BnIII	2S albumin	(Monsalve et al., 1997)
Brassica napis	Rapeseed	BnIII	2S albumin	(Monsalve et al., 1997)
Capsicum annuum	Bell pepper		Profilin, pathogenesis-related protein P23	(Jensen-Jarolim et al., 1998)
Carica papaya	Papaya		Papain	M15203
Daucus carota	Carrot	Dau c 1/1		Z84376
Daucus carota	Carrot	Dau c 1.2		Z81361
Daucus carota	Carrot	Dau c 1.3		Z81362
Glycine max	Soybean	Cim1 protein		S48032
Glycine max	Soybean	Gly m 1.0101	Soybean hull allergen, HPS	AAB34755
Glycine max	Soybean	Gly m 1.0102	Soybean hull allergen, HPS	AAB34756
Glycine max	Soybean	Gly m 2	Hull allergen	A57106
Glycine max	Soybean		Lipoxygenase 1	S25064
Glycine max	Soybean		Lipoxygenase 2	A28161
Glycine max	Soybean		Alpha of beta-conglycinin	X17698
Glycine max	Soybean		A1aBx subunit of glycinin	X02985
Glycine max	Soybean		A5A4B3 subunit of glycinin	X02626
Glycine max	Soybean		G1 subunit of glycinin	X15121
Glycine max	Soybean		G2 subunit of glycinin	X15122
Glycine max	Soybean		G3 subunit of glycinin	X15123
Glycine max	Soybean		CG4 beta-conglycinin	S44893
Glycine max	Soybean		A3B4 subunit of glycinin	M10962
Glycine max	Soybean		Lectin Le1	K00821
Glycine max	Soybean		Kunitz trypsin inhibitor, Kti-S	X80039
Glycine max	Soybean		Trypsin inhibitor, Ti-a	X64447
Glycine max	Soybean		Trypsin inhibitor, Ti-b	X64448
Helianthus annuus	Sunflower	Hel a 1	Profilin	O81982
Helianthus annuus	Sunflower	Hel a 2	Profilin	Y15210
Hordeum vulgare	Barley	Hor v 1	trypsin/alpha amylase inhibitor	P16968
Hordeum vulgare	Barley	Hor v 1	trypsin/alpha amylase inhibitor	P32936
Hordeum vulgare	Barley		dimeric protein BDP	S65605
Juglans regia	English walnut	Jug r 1	2S albumin	U66866
Lycopersicon esculentum	tomato	LAT52	Ole e1 homolog	P13447
Malus domestica	Apple	Mal d 1		O22517
Malus domestica	Apple	Mal d 1		JC4276
Malus domestica	Apple	Mal d 1	Variant	AF074721
Malus domestica	Apple	Mal d 1	Clone MDGC10	Z72425
Malus domestica	Apple	Mal d 1	Clone MDGC32	Z72426
Malus domestica	Apple	Mal d 1	Clone MDGC36	Z72427

Malus domestica	Apple	Mal d 1	Clone MDG38	Z72428
Malus domestica	Apple	Mal d 1		X83672
Oryza sativa	Rice	Ory s 1	Similar to Lol P1	Q40638
Oryza sativa	Rice	RAG1	trypsin/alpha amylase inhibitor	Q01884
Oryza sativa	Rice	RAG2	trypsin/alpha amylase inhibitor	Q01885
Oryza sativa	Rice	RA5	trypsin/alpha amylase inhibitor	S31078
Oryza sativa	Rice	RA5b	trypsin/alpha amylase inhibitor	Q40655
Oryza sativa	Rice	RA14		Q01882
Oryza sativa	Rice	RA14B	trypsin/alpha amylase inhibitor	S59922
Oryza sativa	Rice	RA14d		D43657
Oryza sativa	Rice	RA14e		D43658
Oryza sativa	Rice	RA14f		D43659
Oryza sativa	Rice	RA16	trypsin/alpha amylase inhibitor	S59924
Oryza sativa	Rice	RA17	trypsin/alpha amylase inhibitor	S21157
Oryza sativa	Rice		trypsin/alpha amylase inhibitor	Q40653
Oryza sativa	Rice			AF042200
Persea americana	Avocado	Prs a 1	Endochitinase	P93680
Phaseolus vulgaris	Kidney bean	PR-1	Pathogenesis-related protein 1	S11929
Phaseolus vulgaris	Kidney bean	PR-2	Pathogenesis-related protein 2	S11930
Pisum sativum	Garden pea		pollen allergen-like protein	Q41043
Prunus armeniaca	Apricot	Pru ar 1	Bet v 1 homolog	U93165
Prunus avium	Cherry	Pru av 1	Bet v 1 homolog	U66076
Prunus persica	Peach	Pru p 1	Lipid transfer protein 1	P81402
Secale cereale	Rye	30K allergen	Fragment	S38292
Secale cereale	Rye	Sec c 1	Asthma allergen	S65604
Sinapis alba	Yellow mustard	Sin a 1	Seed storage protein	P15322
Sinapis alba	White mustard	Sin a 1		PC1246
Sinapis alba	White mustard	Sin a 1		S54101
Sinapis alba	White mustard	Sin a 1.0104		Q41277
Sinapis alba	White mustard	Sin a 1.0105		Q41278
Sinapis alba	White mustard	Sin a 1.0106		Q41279
Sinapis alba	White mustard	Sin a 1.0107		Q41280
Sinapis alba	White mustard	Sin a 1.0108		Q41281
Sinapis alba	White mustard	Sin a 1	alpha-amylase inhibitor	PC1247
Solanum tuberosum	Potato	Sol t 1	Patatin, storage protein	P15476
Taraxacum officinale	Dandelion	RAP	Root allergen protein	AF036931
Triticum aestivum	Wheat	Tri a 3	Pollen allergen-like protein	Z50867
Triticum aestivum	Wheat		pollen allergen-like protein	U91981
Zea mays	Maize	Zea m 1	Similar to Lol P1	Q07154
Zea mays	Maize	Clone c13	Ole e1 homolog	P33050
Zea mays	Maize		Pectate lyase, Clone Zm58.1	S43334
Zea mays	Maize		Pectate lyase	S43335
Gliadins from Wheat				
Triticum aestivum	Wheat		Alpha/beta gliadin MM1	X17361
Triticum aestivum	Wheat		gliadin	X00627
Triticum aestivum	Wheat		Alpha/beta gliadin (pw1215)	X02538
Triticum aestivum	Wheat		Alpha/beta gliadin (pw8233)	X02539
Triticum aestivum	Wheat		Alpha/beta gliadin	X02540
Triticum aestivum	Wheat		Alpha gliadin	U08287
Triticum aestivum	Wheat		Gamma gliadin	M36999
Triticum aestivum	Wheat		Gamma gliadin B-precursor	M13713
Triticum aestivum	Wheat	pGliA-42	Alpha gliadin	K02068
Triticum aestivum	Wheat	Clone pA212	Alpha/beta gliadin class A-II	M10092

Triticum aestivum	Wheat	Clone pA42	Alpha/beta class A-V	M11073
Triticum aestivum	Wheat	pW8142	Alpha/beta gliadin	K03075
Triticum aestivum	Wheat	pW8233	Alpha/beta gliadin, class I	K03076
Triticum aestivum	Wheat	Clone pA26	Alpha/beta gliadin, class A-I	M11074
Triticum aestivum	Wheat	Clone pA735	Alpha/beta gliadin, class A-IV	M11075
Triticum aestivum	Wheat	pA1235	Alpha/beta gliadin, class A-III	M11076
Triticum aestivum	Wheat	pB11-33	Gamma gliadin, class B-I	M11077
Triticum aestivum	Wheat	pB312	Gamma gliadin, class B-I	M11336
Triticum aestivum	Wheat		Gamma gliadin, class B-III	M11335
Triticum aestivum	Wheat		Gamma gliadin, variant	M16064
Triticum urartu	Wheat		Alpha/beta gliadin	M16496

Wheat lectins

Triticum aestivum	Wheat		Germ agglutinin isolectin A	M25536
Triticum aestivum	Wheat		Germ agglutinin isolectin D	M25537
Triticum durum	Wheat		Germ agglutinin (partial)	J02961

Wheat profilins

Triticum aestivum	Wheat	Clone TaPRO1	Profilin 1	X89825
Triticum aestivum	Wheat	Clone TaPRO2	Profilin 2	X89826
Triticum aestivum	Wheat	Clone TaPRO3	Profilin 3	X89827

Food Allergens - Animals

Bos taurus	Bovine	Bos d 4	Alpha-lactalbumin	J05147
Bos taurus	Bovine	Bos d 5	Beta-lactoglobulin	X14712
Bos taurus	Bovine	Bos d 6	BSA	M73993
Bos taurus	Bovine	Bos d 8	Alpha-s1 casein	M33123
Bos taurus	Bovine	Bos d 8	Alpha-s1 casein	M38641
Bos taurus	Bovine	Bos d 8	Alpha-s2 casein	M16644
Bos taurus	Bovine	Bos d 8	Beta casein	M15132
Bos taurus	Bovine	Bos d 8	Kappa casein, B2 variant	M36641
Gadus callarias	Baltic cod	Gad c 1	Parvalbumin beta, allergen M	P02622
Gallus gallus	Chicken	Gal d 1	ovomucoid	P01005
Gallus gallus	Chicken	Gal d 2	ovalbumin	P01012
Gallus gallus	Chicken	Gal d 3	ovotransferrin	P02789
Gallus gallus	Chicken	Gal d 4	Lysozyme C	P00698
Gallus gallus	Chicken		Ovalbumin Y gene	J00922
Gallus gallus	Chicken		vitellogenin II, Phosvitin	A92941
Gallus gallus	Chicken		Apovitellenin I	A93464
Metapenaeus ensis	Shrimp	Met e 1	Tropomyosin, partial sequence	U08008
Panulirus stimpsoni	Lobster	Pan s 1	Tropomyosin	AF030063
Penaeus aztecus	Shrimp	Pen a 1	Tropomyosin	(Daul et al., 1993)
Salmo salar	Atlantic salmon	Sal s 1	Parvalbumin beta 1	Q91482
Salmo salar	Atlantic salmon	Sal s 1	Parvalbumin beta 2	Q91483
Turbo cornutus	Gastropod	Tur c 1	Tropomyosin, peptide K20	(Ishikawa et al., 1998)
Turbo cornutus	Gastropod	Tur c 1	Tropomyosin, peptide K24	(Ishikawa et al., 1998)
Turbo cornutus	Gastropod	Tur c 1	Tropomyosin, peptide K19	(Ishikawa et al., 1998)
Turbo cornutus	Gastropod	Tur c 1	Tropomyosin, peptide K21	(Ishikawa et al., 1998)
Turbo cornutus	Gastropod	Tur c 1	Tropomyosin, peptides MT17, MT15, K23	(Ishikawa et al., 1998)

Aero-Allergens - Pollen

Agrostis alba	Bent grass	Agr a 1	Form 1, fragment	F58493
Agrostis alba	Bent grass	Agr a 1	Form 1, fragment	E37396
Agrostis alba	Bent grass	Agr a 1	Form 2, fragment	G58493
Alnus glutinosa	Alder	Aln g 1		S50892
Alnus glutinosa	Alder	Aln g 2		Y17713
Ambrosia artemisiifolia	Short ragweed	Amb a 1.1	Antigen E	P27759
Ambrosia artemisiifolia	Short ragweed	Amb a 1.2	Antigen E	P27760
Ambrosia artemisiifolia	Short ragweed	Amb a 1.2	Antigen E	B53240
Ambrosia artemisiifolia	Short ragweed	Amb a 1.3	Antigen E	P27761
Ambrosia artemisiifolia	Short ragweed	Amb a 1.3	Antigen E	C53240
Ambrosia artemisiifolia	Short ragweed	Amb a 1.4	Antigen E	P28744
Ambrosia artemisiifolia	Short ragweed	Amb a 2	Antigen K	P27762
Ambrosia artemisiifolia	Short ragweed	Amb a 2	Antigen K	E53240
Ambrosia artemisiifolia	Short ragweed	Amb a 3	Allergen RA3	P00304
Ambrosia artemisiifolia	Short ragweed	Amb a 5	Allergen RA5, fragment	P02878
Ambrosia artemisiifolia	Short ragweed	Amb a 6	Allergen RA6	O04004
Ambrosia trifida	Giant ragweed	Amb t 5	Allergen RA5G	P10414
Ambrosia psilostachya	Western ragweed	Amb p 5	Clone B1	L24467
Ambrosia psilostachya	Western ragweed	Amb p 5	Clone A2	L24465
Ambrosia psilostachya	Western ragweed	Amb p 5	Clone B2	L24468
Ambrosia psilostachya	Western ragweed	Amb p 5	Clone A3	L24466
Ambrosia psilostachya	Western ragweed	Amb p 5	Clone B3	L24469
Anthoxanthum odoratum	Sweet vernal grass	Ant o 1	Fragment	G37396
Arabidopsis thaliana	Mouse-ear cress	profilin i		3NUL
Arabidopsis thaliana	Mouse-ear cress	ESSA 1 AP2	Alternaria allergen similarity	Z99708
Arabidopsis thaliana	Mouse-ear cress	ESSA 1	Fragment 7, Lol p1 similarity	Z97342
Arabidopsis thaliana	Mouse-ear cress	BAC F5114	Similar to Holcus major allergen	AC001229
Artemisia vulgaris	Mugwort	Art v 2	Fragments	A38624
Betula verrucosa	White birch	Bet v 2	Profilin	P25816
Betula pendula	European White Birch	Bet v 1	isoform at2	AJ002106
Betula pendula	European White Birch	Bet v 1	isoform at12	AJ002107
Betula pendula	European White Birch	Bet v 1	isoform at15	AJ002108
Betula pendula	European White Birch	Bet v 1	isoform at76	AJ002110
Betula pendula	European White Birch	Bet v 1	Clone BVGC181	Z72437
Betula pendula	European White Birch	Bet v 1	Clone BVGC21	Z72430
Betula pendula	European White Birch	Bet v 1	Clone BVGC25	Z72431
Betula pendula	European White Birch	Bet v 1	Clone BVGC31	Z72432
Betula pendula	European White Birch	Bet v 1	Clone BVGC34	Z72433
Betula pendula	European White Birch	Bet v 1	Clone BVGC45	Z72434
Betula pendula	European White Birch	Bet v 1	Clone BVGC63	Z72435
Betula pendula	European White Birch	Bet v 1	Clone BVGC681	Z72438
Betula pendula	European White Birch	Bet v 1	Clone BVGC70	Z72436
Betula pendula	European White Birch	Bet v 1	Clone 2230	Z80099
Betula pendula	European White Birch	Bet v 1	Clone 167	Z80100
Betula pendula	European White Birch	Bet v 1	Clone 184	Z80101
Betula pendula	European White Birch	Bet v 1	Clone 2225	Z80102
Betula pendula	European White Birch	Bet v 1	Clone 2226	Z80103
Betula pendula	European White Birch	Bet v 1	Clone 2227	Z80104
Betula pendula	European White Birch	Bet v 1	Clone 2229	Z80105
Betula pendula	European White Birch	Bet v 1	Clone 2301	Z80106
Betula pendula	European White Birch	Bet v 1		O23746
Betula pendula	European White Birch	Bet v 1		O23747
Betula pendula	European White Birch	Bet v 1		O23748

Betula pendula	European White Birch	Bet v 1		O23749
Betula pendula	European White Birch	Bet v 1		O23750
Betula pendula	European White Birch	Bet v 1b		A55699
Betula pendula	European White Birch	Bet v 1c		B55699
Betula pendula	European White Birch	Bet v 1d/h		C55699
Betula pendula	European White Birch	Bet v 1e		D55699
Betula pendula	European White Birch	Bet v 1f/i		E55699
Betula pendula	European White Birch	Bet v 1g		F55699
Betula pendula	European White Birch	Bet v 1j		G55699
Betula pendula	European White Birch	Bet v 1k		H55699
Betula pendula	European White Birch	Bet v 1l		I55699
Betula pendula	European White Birch	Bet v 1m/n		A57427
Betula pendula	European White Birch	Bet v 2	Fragment	B45786
Betula pendula	European White Birch	Bet v 3	Calcium binding protein	S46233
Betula pendula	European White Birch	Bet v 4	Calcium binding protein	O04131
Betula pendula	European White Birch	Bet v 4		X87153
Betula pendula	European White Birch		Isoflavone reductase homolog	AF047896
Brassica napus	Rape	Bra n 1	Pollen allergen group I	S65149
Brassica napus	Rape	Bra n 2	Pollen allergen group II	S65150
Brassica napus	Rape	Clone 42	Pollen allergen group II	S65144
Brassica napus	Rape	Clone 44	Pollen allergen group II	S65145
Brassica rapa	Turnip	Bra r 1	Calcium binding protein	S65151
Brassica rapa	Turnip	Bra r 2	Pollen allergen group II	S65152
Brassica rapa	Turnip	Clone 4	Pollen allergen group II	S65143
Carpinus betulus	Hornbeam	Car b 1	Clone 380	Z80159
Carpinus betulus	Hornbeam	Car b 1	Clone 383	Z80160
Carpinus betulus	Hornbeam	Car b 1	Clone 372	Z80161
Carpinus betulus	Hornbeam	Car b 1	Clone 541	Z80162
Carpinus betulus	Hornbeam	Car b 1	Clone 385	Z80166
Carpinus betulus	Hornbeam	Car b 1	Clone 295	Z80168
Carpinus betulus	Hornbeam	Car b 1	Clone 563	Z80169
Carpinus betulus	Hornbeam	Car b 1	Clone 2237	Z80170
Carpinus betulus	Hornbeam	Car b 1	Isoform 1, Bet v 1 homolog	P38949
Carpinus betulus	Hornbeam	Car b 1	Isoform 2, Bet v 1 homolog	P38950
Castanea sativa	European chestnut	Cas s 1	Bet v 1 homolog, fragment	PC2001
Chamaecyparis obtusa	Japanese cypress	Cha o 1		D45404
Corylus avellana	European hazel	Cor a 1/5		X70999
Corylus avellana	European hazel	Cor a 1/6		X71000
Corylus avellana	European hazel	Cor a 1/11		S30055
Corylus avellana	European hazel	Cor a 1/11		X70997
Corylus avellana	European hazel	Cor a 1	Clone CAGC10	Z72439
Corylus avellana	European hazel	Cor a 1	Clone CAGC11	Z72440
Corylus avellana	European hazel	Cor a 1/16		X70998
Cryptomeria japonica	Japanese cedar	Cry j 1	Clone pCCI-15	JC2124
Cryptomeria japonica	Japanese cedar	Cry j 1	Clone pCCI-2-2	JC2123
Cryptomeria japonica	Japanese cedar	Cry j 2	Pectinase	P43212
Cryptomeria japonica	Japanese cedar	Cry j 2		JC2498
Cucumis sativus	Cucumber	Cs-EXP1	Expansin S1	U30382
Cucumis sativus	Cucumber	Cs-EXP2	Expansin S2	U30460
Cynodon dactylon	Bermuda grass	B1	B1 clone	A28046
Cynodon dactylon	Bermuda grass	B2	B2 clone	A28056
Cynodon dactylon	Bermuda grass	B4	B4 clone	A28050
Cynodon dactylon	Bermuda grass		Calcium binding protein	P94092
Cynodon dactylon	Bermuda grass	Cyn d 1	Fragment	A61226

Cynodon dactylon	Bermuda grass	Cyn d 1	Clone 14c1 and CD1	S83343
Cynodon dactylon	Bermuda grass	Cyn d 12	profilin 1	Y08390
Dactylis glomerata	Orchard grass	Dac g 1	AgDg1, fragment	D58493
Dactylis glomerata	Orchard grass	Dac g 2	Similar to Lol p 1, fragment	Q41183
Dactylis glomerata	Orchard grass	Dac g 3	Fragment	P93124
Dactylis glomerata	Orchard grass	Dac g 3	Fragment	A60359
Dactylis glomerata	Orchard grass	Dac g 5a	Fragment	(Klysner et al., 1992)
Dactylis glomerata	Orchard grass	Dac g 5b	Fragment	(Klysner et al., 1992)
Festuca elatior	Reed fescue	Fes e 1	fragment	A58493
Festuca elatior	Reed fescue	Fes e 1	Type A, fragment	C37396
Festuca elatior	Reed fescue	Fes e 1	Type B	D37396
Holcus lanatus	Velvet grass	Hol l 1		Z27084
Holcus lanatus	Velvet grass	Hol l 1		Z68893
Holcus lanatus	Velvet grass	Hol l 5	Group V allergen	Z97874
Holcus lanatus	Velvet grass	Hol l 5.02	Group V allergen	Z97875
Holcus lanatus	Velvet grass	30 kD	Fragment	S38291
Hordeum vulgare	Barley	Hor v 9		JC5475
Juniperus ashei	Mountain cedar	Jun a 1-1		AF106662
Juniperus ashei	Mountain cedar	Jun a 1-2		AF106663
Juniperus oxycedrus	Juniper	Jun o 2	EF-hand calcium binding	AF031471
Ligustrum vulgare	privet	L1		X77787
Ligustrum vulgare	privet	L10		X77788
Lolium perenne	Perennial ryegrass		Pollen allergen plb	A38582
Lolium perenne	Perennial ryegrass		50k allergen	S38288
Lolium perenne	Perennial ryegrass	Lol p 1	Group I, allergen R7	P14946
Lolium perenne	Perennial ryegrass	Lol p 1	Group I	S13614
Lolium perenne	Perennial ryegrass	Lol p 1	Group I, clone 1A	A37881
Lolium perenne	Perennial ryegrass	Lol p 1	Group I, clone 5A	B37881
Lolium perenne	Perennial ryegrass	Lol p 2	Group II, fragment	A48595
Lolium perenne	Perennial ryegrass	Lol p 2	Group II	Q40239
Lolium perenne	Perennial ryegrass	Lol p 2-A	Group II	P14947
Lolium perenne	Perennial ryegrass	Lol p 3	Group III	P14948
Lolium perenne	Perennial ryegrass	Lol p 4		A60737
Lolium perenne	Perennial ryegrass	Lol p 5A	Similar to Poa P9/Phl P6	Q40240
Lolium perenne	Perennial ryegrass	Lol p 5B	Similar to Poa P9/Phl P6	Q40237
Lolium perenne	Perennial ryegrass	Lol p X1		A54002
Mercurialis annua			Profilin	Y13271
Olea europea	Olive tree	Ole e 1.0102		Y12428
Olea europea	Olive tree	Ole e 1.0103		Y12427
Olea europea	Olive tree	Ole e 1.05		Y12426
Olea europea	Olive tree	Ole e 1		S75766
Olea europea	Olive tree	Ole e 1	Fragment	S36872
Olea europea	Olive tree	Ole e 2	profilin 1	Q24169
Olea europea	Olive tree	Ole e 2	profilin 2	Q24170
Olea europea	Olive tree	Ole e 2	profilin 3	Q24171
Olea europea	Olive tree	Ole3	Calcium-binding allergen	AF015810
Olea europea	Olive tree	Ole16	Fragment	I53806
Olea europea	Olive tree	Ole17	Fragment	E53806
Olea europea	Olive tree	Ole19	Fragment	F53806
Olea europea	Olive tree	Ole1c	Fragment	C53806
Olea europea	Olive tree	Ole20	Fragment	A38968
Olea europea	Olive tree	Ole26	Fragment	G53806
Olea europea	Olive tree	Ole33/Ole37	Fragment	D53806
Olea europea	Olive tree	Ole3c	Fragment	A53806

Olea europea	Olive tree	Ole5c	Fragment	B53806
Olea europea	Olive tree	Ole6	Fragment	H53806
Olea europea	Olive tree	Ole e 4		P80741
Olea europea	Olive tree	Ole e 5	Superoxide dismutase, fragment	P80740
Olea europea	Olive tree	Ole e 6		O24172
Parietaria judaica		Par j 1.0101	Lipid transfer protein	P43217
Parietaria judaica		Par j 1.0201	Lipid transfer protein 1	Q40905
Parietaria judaica		Par j 2.0101	Lipid transfer protein 2	P55958
Parthenium hysterophorus	Compositae weed	Par h 1	31 kDa hydroxyproline-rich glycoprotein	(Gupta et al., 1996)
Parietaria officinalis		Par o 1	Fragment	A53252
Phalaris Aquatica	Canary grass	Pha a 1	Similar to Lol P1	Q41260
Phalaris Aquatica	Canary grass	Pha a 5.1	Similar to Poa p 9/Phl p 6	P56164
Phalaris Aquatica	Canary grass	Pha a 5.2	Similar to Poa p 9/Phl p 7	P56165
Phalaris Aquatica	Canary grass	Pha a 5.3	Similar to Poa p 9/Phl p 8	P56166
Phalaris Aquatica	Canary grass	Pha a 5.4	Similar to Poa p 9/Phl p 9	P56167
Phleum pratense	Common timothy	Phl p 1		X78813
Phleum pratense	Common timothy	Phl p 1		Z27090
Phleum pratense	Common timothy	Phl p 2		X75925
Phleum pratense	Common timothy	Phl p 5.0105	Clone 10022	AF061066
Phleum pratense	Common timothy	Phl p 5.0106	Clone 10027	AF061067
Phleum pratense	Common timothy	Phl p 5.0107	Clone 10029	AF061068
Phleum pratense	Common timothy	Phl p 5.0108	Clone 10030	AF061069
Phleum pratense	Common timothy	Phl p 5.0103	Group 5	AF069470
Phleum pratense	Common timothy	Phl p 5.0203	Group 5	AF069471
Phleum pratense	Common timothy	Phl p 5.0204	Group 5	AF069472
Phleum pratense	Common timothy	Phl p 5.0206	Group 5	AF069473
Phleum pratense	Common timothy	Phl p 5.0207	Group 5	AF069474
Phleum pratense	Common timothy	Phl p 5	Isovariant 711I13	Z82985
Phleum pratense	Common timothy	Phl p 5	Isovariant 618I18	Z82986
Phleum pratense	Common timothy	Phl p 5A	Similar to Poa P9/Phl p 6	Q40962
Phleum pratense	Common timothy	Phl p 5A		S32101
Phleum pratense	Common timothy	Phl p 5B	Similar to Poa p 9/Phl p 6	Q40963
Phleum pratense	Common timothy	Phl p 5B	Similar to Poa p 9/Phl p 7	S38584
Phleum pratense	Common timothy	Phl p 6	Isolate c223	Y16955
Phleum pratense	Common timothy	Phl p 6	Isolate c142	Y16956
Phleum pratense	Common timothy	Phl p 11	Profilin	P35079
Pinus radiata	Pine	PRE29		AF049068
Poa pratensis	Kentucky bluegrass	Poa p 9	Clone 31	C39098
Poa pratensis	Kentucky bluegrass	Poa p 9	Clone 41	A39098
Poa pratensis	Kentucky bluegrass	Poa p 9	KBG60	P22286
Poa pratensis	Kentucky bluegrass	Poa p 9	Clone 60	B39098
Poa pratensis	Kentucky bluegrass	Poa p 9	Clone 7.2	A60373
Pyrus communis	Pear	Pyr c 1		AF057030
Pyrus communis	Pear	Pyr c 2	Isoflavone reductase	AF071477
Quercus alba	Oak	Que a 1	Bet v 1 homolog, fragment	D53288
Syringa vulgaris	Lilac	Syr v 1	Isoform 1	S43242
Syringa vulgaris	Lilac	Syr v 1	Isoform 2	S43243
Syringa vulgaris	Lilac	Syr v 1	Isoform 3	S43244

Aero-Allergens - Fungi

Alternaria alternata	Mold	Alt a 1		U62097
Alternaria alternata	Mold	Alt a 1		U82633
Alternaria alternata	Mold	Alt a 3	hsp70	U87807
Alternaria alternata	Mold	Alt a 3	hsp70	U87808
Alternaria alternata	Mold	Alt a 6	Ribosomal P2 phosphoprotein	U87806
Alternaria alternata	Mold		Enolase	(Breitenbach et al., 1997)
Alternaria alternata	Mold	Alt a 6	Ribosomal protein	X78222
Alternaria alternata	Mold	Alt a 7		X78225
Alternaria alternata	Mold	Alt a 10	Aldehyde dehydrogenase	P42041
Alternaria alternata	Mold	Alt a 12	60S ribosomal protein P1	P49148
Aspergillus fumigatus	Mold	Asp f 1	Ribonuclease, mitogillin	P04389
Aspergillus fumigatus	Mold	Asp f 1	Ribonuclease, mitogillin	A46497
Aspergillus fumigatus	Mold	Asp f 1	Partial exon 2	AJ005037
Aspergillus fumigatus	Mold	Asp f 2		P79017
Aspergillus fumigatus	Mold	rAsp f 4	ABPA patient allergen	AJ001732
Aspergillus fumigatus	Mold	Asp f 5	Metalloprotease (MEP)	P46075
Aspergillus fumigatus	Mold	Asp f 6	Mn superoxide dismutase	U53561
Aspergillus fumigatus	Mold	rAsp f 7		AJ223315
Aspergillus fumigatus	Mold	rAsp f 9		AJ223327
Aspergillus fumigatus	Mold	Asp f 10	Aspartic protease	X85092
Aspergillus fumigatus	Mold	Asp f 12	Heat shock protein hsp1	P40292
Aspergillus fumigatus	Mold	rAsp f 13		AJ002026
Aspergillus niger	Baker's yeast	Asp n ?	Xylanase	Z84377
Aspergillus oryzae		Asp o 2	TAKA-amylase A	JK0201
Aspergillus oryzae		Asp o 13	Alkaline serine protease	P12547
Candida albicans	Yeast	Can a 1	Alcohol dehydrogenase	P43067
Candida albicans	Yeast	Can a ?	Enolase 1	P30575
Candida boidinii	Yeast	Can b 2	peroxisomal membrane protein A	P14292
Candida boidinii	Yeast	Can b 2	peroxisomal membrane protein B	P14293
Cladisporium herbarum		Cla h 3	Aldehyde dehydrogenase	S43114
Cladisporium herbarum		Cla h 4	60S ribosomal protein P2	P42038
Cladisporium herbarum		Cla h 4	60S ribosomal protein P2	P42039
Cladisporium herbarum		Cla h ?	Heat shock 70 protein	P40918
Cladisporium herbarum		Cla h 5		P42059
Cladisporium herbarum		Cla h 6	Enolase	P42040
Cladisporium herbarum		Cla h 12	60S ribosomal protein P1	P50344
Fusarium solani pisi		Fus s ?		P81010
Malassezia furfur	Mold	Mal f 1		X96486
Malassezia furfur	Mold	Mal f 4	Malate dehydrogenase homolog	AF084828
Malassezia furfur	Mold	Mal f 5		AJ011955
Malassezia furfur	Mold	Mal f 6	cyclophilin homolog	AJ011956
Malassezia furfur	Mold			AJ011957
Malassezia furfur	Mold			AJ011958
Malassezia furfur	Mold			AJ011959
Penicillium citrinum		Pen c 1	Alkaline serine protease	AF084546
Penicillium citrinum		Pen c 2	Alkaline serine protease	AF098517
Penicillium oxalicum		34 kDa	Similar to vacuolar serine protease	(Shen et al., 1999)
Penicillium notatum		28 kDa	Peptide PN3, PN4	(Shen et al., 1999)
Penicillium notatum		32 kDa	Peptide PN2	(Shen et al., 1999)
Penicillium notatum		68 kDa		Q02352
Trichonophyton tonsurans			83 kD hypersensitivity protein	P80514

Aero-Allergens - Animals

Canis familiaris	Dog	Can f ?	Albumin, salivary gland allergen	S72946
Canis familiaris	Dog	Can f 1	Lipocalin	O18873
Canis familiaris	Dog	Can f 2	Lipocalin	O18874
Bos taurus	Bovine dander	BDA11	EF-hand calcium binding protein	Q28050
Bos taurus	Bovine dander	BDA20	Clones pPOT, pPOT10.2	L42867
Equus caballus	Horse dander	Equ c 2.0101	Lipocalin	P81216
Equus caballus	Horse dander	Equ c 2.0102	Lipocalin	P81217
Equus caballus	Horse	Equ c 1	Lipocalin	Q95182
Felis silvestris catus	Cat	Fel d 1-A		P30438
Felis silvestris catus	Cat	Fel d 1-A	Chain 1 precursor A	JC1136
Felis silvestris catus	Cat	Fel d 1-A	Chain 1 precursor B	JC1126
Felis silvestris catus	Cat	Fel d 1	Chain 1 short form	B56413
Felis silvestris catus	Cat	Fel d 1	Chain 1 long form	A56413
Felis silvestris catus	Cat	Fel d 1	Chain 2	M77341
Felis silvestris catus	Cat	Fel d 1	Chain 2 short form	JC1127
Felis silvestris catus	Cat	Fel d 1-A		P30439
Mus musculus	Mouse		Lacrimal gland protein	AF008595
Mus musculus	Mouse	MUP	Major urinary protein	M27068
Mus musculus	Mouse	MUP 1	Major urinary protein I	M16355
Mus musculus	Mouse	MUP 2	Major urinary protein II	M16356
Mus musculus	Mouse	MUP 4	Major urinary protein IV	M16358
Mus musculus	Mouse	MUP 5	Major urinary protein V	M16360
Rattus norvegicus	Rat	Rat n 1	Transthyretin	P02767
Rattus norvegicus	Rat		Alpha 2u globulin	J00737

Aero-Allergens - Insects & Mites

Blattella germanica	German cockroach	Bla g 2	Aspartyl protease	A57164
Blattella germanica	German cockroach	Bla g 4	Catlycin	U40767
Blattella germanica	German cockroach	Bla g 5	Glutathione S-aryltransferase	U92412
Chironomus thummi thummi	Midge	Chi t 1.01	Globin component III	P02229
Chironomus thummi thummi	Midge	Chi t 1.02	Globin component IV	P02230
Chironomus thummi thummi	Midge	Chi t 2.0101	Globin component I	P02221
Chironomus thummi thummi	Midge	Chi t 3	Globin component II beta	P02222
Chironomus thummi thummi	Midge	Chi t 4	Globin component III-A	P02231
Chironomus thummi thummi	Midge	Chi t 5	Globin component VI	P02224
Chironomus thummi thummi	Midge	Chi t 6.01	Globin component VII-A	P02226
Chironomus thummi thummi	Midge	Chi t 6.02	Globin component IX	P02223
Chironomus thummi thummi	Midge	Chi t 7	Globin component VII-B	P02225
Chironomus thummi thummi	Midge	Chi t 8	Globin component VIII	P02227
Chironomus thummi thummi	Midge	Chi t 9	Globin component X	P02228
Periplaneta americana	American cockroach	Per a 1	Cr-P1I	U69260
Periplaneta americana	American cockroach	Per a 1	Cr-P1I	U69261
Periplaneta americana	American cockroach	Per a 1	Cr-P1I	U69957
Periplaneta americana	American cockroach	Per a 1	Cr-P1I	U78970
Periplaneta americana	American cockroach	Per a 3	Cr-P1, clone C28	L40821
Periplaneta americana	American cockroach	Per a 3	Cr-P1	L40819
Periplaneta americana	American cockroach	Per a 3	Cr-P1, clone C20	L40820
Periplaneta americana	American cockroach	Per a 3	Cr-P1, clone C12	L40818
Acarus siro	Mite	Aca s 13	Partial sequence	AJ006774
Blomia tropicalis	Mite	Blo t 5		U59102
Blomia tropicalis	Mite	Blo t 12	Bt11a	U27479

Blomia tropicalis	Mite		Partial sequence	U27702
Blomia tropicalis	Mite	Blo t 13	Fatty-acid binding protein, Bt6	U58106
Dermatoph. farinae	House dust mite	Der f 1	Thiol protease	P16311
Dermatoph. farinae	House dust mite	Der f 1		A61500
Dermatoph. farinae	House dust mite	Der f 1	Cysteine protease, Group I	X65196
Dermatoph. farinae	House dust mite	Der f 2	Complete sequence	D10448
Dermatoph. farinae	House dust mite	Der f 2	Partial sequence	D10449
Dermatoph. farinae	House dust mite	Der f 2	Group II allergen	S70378
Dermatoph. farinae	House dust mite	Der f 3	Trypsin family protease	P49275
Dermatoph. farinae	House dust mite	Der f 3	Trypsin family protease	Q94508
Dermatoph. farinae	House dust mite	Der f 6	Trypsin family protease	P49276
Dermatoph. farinae	House dust mite	Der f 7		Q26456
Dermatoph. farinae	House dust mite	MAG		P39673
Dermatoph. farinae	House dust mite	MAG29	Heat shock protein 70	P39674
Dermatoph. microceras	House dust mite	Der m 1	Fragment	B27634
Dermatoph. pteronyssinus	House dust mite		Tropomyosin	AF016278
Dermatoph. pteronyssinus	House dust mite		Group III allergen	A39997
Dermatoph. pteronyssinus	House dust mite	Der p 1	Cysteine protease, group I	X65197
Dermatoph. pteronyssinus	House dust mite	Der p 1	Preproenzyme, complete sequence	U11695
Dermatoph. pteronyssinus	House dust mite	Der p 1	Cysteine protease	A31657
Dermatoph. pteronyssinus	House dust mite	Der p 1	Cysteine protease, fragment	B31657
Dermatoph. pteronyssinus	House dust mite	Der p 1	Cysteine protease, fragment	S03380
Dermatoph. pteronyssinus	House dust mite	Der p 1	Cysteine protease	JQ0337
Dermatoph. pteronyssinus	House dust mite	Der p 2		P49278
Dermatoph. pteronyssinus	House dust mite	Der p 3	Prepro-zymogen, complete	U11719
Dermatoph. pteronyssinus	House dust mite	Der p 4	Alpha amylase, fragment	P49274
Dermatoph. pteronyssinus	House dust mite	Der p 5		P14004
Dermatoph. pteronyssinus	House dust mite	Der p 6	Trypsin family protease, chymotrypsin	P49277
Dermatoph. pteronyssinus	House dust mite	Der p 7	Complete sequence	U37044
Dermatoph. pteronyssinus	House dust mite	Der p 8	Glutathione S-transferase	S75286
Dermatoph. pteronyssinus	House dust mite	Der p 9	Collagenolytic serine protease	AAB50781
Dermatoph. pteronyssinus	House dust mite	Der p 10	Tropomyosin	O18416
Dermatoph. pteronyssinus	House dust mite	Der p 10	Tropomyosin	O16188
Euroglyphus maynei	House dust mite	Eur m 1	Group I, cysteine protease	X60073
Euroglyphus maynei	House dust mite	Eur m 3.0101	Trypsin family protease	AF047615
Lepidoglyphus destructor	Storage mite	Lep d 1.0102		X89014
Lepidoglyphus destructor	Storage mite	Lep d 1.02		S66499
Tyrophagus putrescentiae	Dust mite		Group 2	Y12690

Venoms & Salivary Allergens

Aedes aegypti	Yellowfever mosquito	Aed a 1	salivary gland allergen	L12389
Aedes aegypti	Yellowfever mosquito	Aed a 2	salivary gland allergen	M33157
Aedes aegypti	Yellowfever mosquito	Aed a 3	30 kD salivary gland allergen	O01949
Anopheles gambiae	African malaria mosquito		gVAG protein, clone cF6	Y17702
Apis mellifera	Honeybee	Api m 1	Phospholipase A2	P00630
Apis mellifera	Honeybee	Api m 2	Hyaluronoglucosaminidase	Q08169
Apis mellifera	Honeybee	Api m 3	Melittin	P01501
Ctenocephalides felis	Cat flea	Cte f 1	Salivary antigen 1	AF102502
Dolichovespula arenaria	Yellow hornet	Dol a 5	Venom allergen 5	Q05108
Dolichovespula maculata	Whiteface hornet	Dol m 1.01	Phospholipase A1	Q06478
Dolichovespula maculata	Whiteface hornet	Dol m 1.02	Phospholipase A1 2	P53357
Dolichovespula maculata	Whiteface hornet	Dol m 2	Hyaluronoglucosaminidase	P49371
Dolichovespula maculata	Whiteface hornet	Dol m 5.01	Venom allergen 5.01	P10736
Dolichovespula maculata	Whiteface hornet	Dol m 5.02	Venom allergen 5.02	P10737

Myrmecia pilosula	Bulldog ant	Myr p 1		S28180
Myrmecia pilosula	Bulldog ant	Myr p 1		X70256
Myrmecia pilosula	Bulldog ant	Myr p 2		Q26464
Polistes annularis	Paper wasp	Pol a 5	Venom allergen 5	M98857
Polistes exclamans	Paper wasp	Pol e 5	Venom allergen 5	P35759
Polistes fuscatus	Paper wasp	Pol f 5	Venom allergen 5	F44583
Solenopsis invicta	Red fire ant	Sol i 2	Phospholipase, Venom allergen 2	A37330
Solenopsis invicta	Red fire ant	Sol i 3	Venom allergen 3	P35778
Solenopsis invicta	Red fire ant	Sol i 3	Venom allergen 3	AF012919
Solenopsis invicta	Red fire ant	Sol i 4	Venom allergen 4	P35777
Solenopsis richteri	Black fire ant	Sol r 2	Phospholipase, fragment	E60727
Solenopsis richteri	Black fire ant	Sol r 3	Venom allergen 3	P35779
Vespa crabro	European hornet	Vesp c 5.01	Antigen 5	G44583
Vespa crabro	European hornet	Vesp c 5.02	Antigen 5	H44583
Vespula flavopilosa	Wasp	Ves f 5	Venom allergen 5	A44583
Vespula germanica	Wasp	Ves g 5	Venom allergen 5	B44583
Vespula maculifrons	Wasp	Ves m 1	Phospholipase A1	A44564
Vespula maculifrons	Wasp	Ves m 5	Venom allergen 5	P35760
Vespula pensylvanica	Wasp	Ves p 5	Venom allergen 5	C44583
Vespula squamosa	Wasp	Ves s 5	Venom allergen 5	D44583
Vespula vidua	Wasp	Ves vi 5		E44583
Vespula vulgaris	Wasp	Ves v 1	Phospholipase A1	L43561
Vespula vulgaris	Wasp	Ves v 2	Hyaluronoglucosaminidase	P49370
Vespula vulgaris	Wasp	Ves v 5	Venom allergen 5	M98858

Others - Contact Allergens

Hevea brasiliensis	Para rubber tree		Latex allergen	AJ223038
Hevea brasiliensis	Para rubber tree	Hev b 1	Rubber elongation factor	P15252
Hevea brasiliensis	Para rubber tree	Hev b 2	beta 1,3-endoglucanase	S65077
Hevea brasiliensis	Para rubber tree	Hev b 3	Fragment	AAB33777
Hevea brasiliensis	Para rubber tree	Hev b 3	Fragment	AAB33778
Hevea brasiliensis	Para rubber tree	Hev b 3	Fragment	AAB33779
Hevea brasiliensis	Para rubber tree	Hev b 5	Latex allergen	U42640
Hevea brasiliensis	Para rubber tree	Hev b 6	Hevein	P02877
Hevea brasiliensis	Para rubber tree	Hev b 7	Latex patatin homolog	AJ223039
Hevea brasiliensis	Para rubber tree	Hev b 7	Latex patatin homolog	U80598
Hevea brasiliensis	Para rubber tree	Hev b 8	Profilin	Y15042

Others - Nematodes & Worms

Acanthamoeba castellanii		profilin ia	actin-binding protein	1PRQ
Ancylostoma caninum	dog hookworm		aspartic proteinase	JC5077
Anglostoma duodenale	Hookworm		ancylostoma secreted protein 1	AF077402
Ascaridia galli		AG1	Fatty acid binding protein	AF065383
Ascaris lumbricoides	Common roundworm	ABA-1	aba-1d1	ALU86100
Ascaris lumbricoides	Common roundworm	ABA-1	aba-1dr1	ALU86097
Ascaris lumbricoides	Common roundworm	ABA-1	aba-1d2	ALU86101
Ascaris lumbricoides	Common roundworm	ABA-1	aba-1dr2	ALU86098
Ascaris lumbricoides	Common roundworm	ABA-1	aba-1d3	ALU86102
Ascaris lumbricoides	Common roundworm	ABA-1	aba-1dr3	ALU86099
Ascaris lumbricoides	Common roundworm	ABA-1	aba-1r1	ALU86091
Ascaris lumbricoides	Common roundworm	ABA-1	aba-1r2	ALU86092
Ascaris lumbricoides	Common roundworm	ABA-1	aba-1r4	ALU86094
Ascaris lumbricoides	Common roundworm	ABA-1	aba-1r6	ALU86096

<i>Ascaris lumbricoides</i>	Common roundworm	ABA-1	Fragment	B37188
<i>Ascaris lumbricoides</i>	Common roundworm	ABA-1	Fragment	A48576
<i>Ascaris suum</i>	Pig roundworm	ABA-1	Fatty acid binding protein	AF051702
<i>Ascaris suum</i>	Pig roundworm	ABA-1		Q06811
<i>Ascaris suum</i>	Pig roundworm	ABA-1	Fragment	A37188
<i>Ascaris suum</i>		ABA-1	Fatty acid binding protein	L03211
<i>Brugia malayi</i>			vespid allergen antigen homolog	AF042088
<i>Dictyocaulus viviparus</i>	Bovine lungworm	DVA-1	ABA-1	Q24702
<i>Toxocara canis</i>		TBA-1	ABA-1 homolog	B49139
<i>Dirofilaria immitis</i>			Venom allergen 5-like protein	AF001100
<i>Loa loa</i>	Filarial worm	LL20	15kDa ladder antigen	U03103
<i>Necator americanus</i>	Hookworm	ASP1	Ancylostoma secreted protein 1	AF079521
<i>Onchocerca volvulus</i>			Vespid allergen antigen homolog	AF042087
<i>Strongyloides stercoralis</i>			Allergen polyprotein homolog	AF035658

Appendix 3. FASTA sequence similarity alignments of CP4 EPSPS protein with allergens and gliadins

(Peptide) FASTA of: CP4_EPSPS.pep from: 1 to: 455 September 13, 1999 12:08

TO: UPDATE2:* *
Sequences: 567
Symbols: 118,659
Word Size: 2

Databases searched: Release 1.0, Released on 12May1999, Formatted on 12May1999

Scoring matrix: GenRunData:blosun50.cmp
Variable pamfactor used
Gap creation penalty: 12
Gap extension penalty: 2

Histogram Key:

Each histogram symbol represents 1 search set sequences
z-scores computed from opt scores

z score	obs (=)	exp (*)
< 20	2	0:==
22	0	0:
24	0	0:
26	0	0:
28	0	0:
30	1	1:*
32	8	3:==*=====
34	13	8:=====*
36	8	17:=====
38	30	28:=====*
40	29	39:=====
42	25	47:=====
44	56	52:=====*
46	50	53:=====*
48	45	51:=====*
50	51	46:=====*
52	50	41:=====*
54	41	35:=====*
56	34	29:=====*
58	25	24:=====*
60	14	19:=====
62	21	16:=====*
64	13	12:=====*
66	9	10:=====*
68	2	8:=====
70	8	6:=====*
72	6	5:=====*
74	12	4:=====*
76	1	3:=====

78	2	2:==*
80	4	2:==**
82	2	1:*=
84	0	1:*
86	0	1:*
88	1	1:*
90	1	0:==
92	0	0:
94	0	0:
96	0	0:
98	2	0:==
100	0	0:
102	0	0:
104	1	0:==
106	0	0:
108	0	0:
110	1	0:==
112	0	0:
114	0	0:
116	0	0:
118	0	0:
>120	0	0:

Joining threshold: 37, opt. threshold: 25, opt. width: 16, reg. scaled

The best scores are:

	initl	initn	opt	z-sc	E(565)...
UP:S70378_1					
! S70378 Dermatophagoides farinae Der...	36	36	83	109.6	0.15
UP:DEPDER2_1					
! D10448 Dermatophagoides farinae Der...	36	36	79	104.5	0.29
UP:S63984					
! major allergen Api g 1 - Apium grav...	38	38	75	99.2	0.57
UP:DEPDER3_1					
! D10449 Dermatophagoides farinae Der...	36	36	73	97.6	0.71
UP:RNMG_ASPRE					
! P04389 aspergillus restrictus, and ...	52	52	68	89.8	1.9
UP:A46497					
! major allergen I - Aspergillus fumi...	52	52	68	89.8	1.9
UP:AH11_ARAHY					
! P43237 arachis hypogaea (peanut). a...	44	44	69	82.2	5.1
UP:AH12_ARAHY					
! P43238 arachis hypogaea (peanut). a...	44	44	69	82.1	5.2
UP:B39098					
! allergen Poa p IX (clone 60) - Kent...	58	58	64	81.0	5.9
UP:MP93_POAPR					
! P22286 poa pratensis (kentucky blue...	58	58	64	81.0	5.9
UP:O81982					
! O81982 helianthus annuus (common su...	35	35	59	80.8	6.1
UP:GB_PLN2_HAPROFILN					
! TRANSLATE of: input_30.rs(fhaprofil...	35	35	59	80.7	6.1
UP:PROF_BETVE					
! P25816 betula verrucosa (white birch...	36	36	57	78.3	8.3
UP:MP01_ORYSA					
! Q40638 oryza sativa (rice). major p...	31	31	60	77.2	9.6
\\End of List					

CP4_EPSPS.pep
UP:S70378_1

LOCUS S70378_1
DEFINITION Der f II=group II major allergen [Dermatophagoides
farinae=mites, mRNA Partial, 517 nt]; group II major allergen; This
sequence comes from Fig. 4.
DATE 22 SEP-1994 . . .

SCORES Initl: 36 Initn: 36 Opt: 83 z score: 109.6 E(): 0.15
Smith-Waterman score: 83; 30.5% identity in 82 aa overlap

```

      190      200      210      220      230
CP4_EPSPS.pe LLAGLNTPGITTVIEPIMTRDHTKMLQGFGANLTVET--DAD-GVRTIRLEGRGKLTGQ
S70378_1      QVDVKDCANNEIKKVMVDGCHGSDPCIHRGKPFLEALFDANQNTKTAKIEIKASLDGL
              20       30       40       50       60       70
```

```

      240      250      260      270      280      290
CP4_EPSPS.pe VIDVPGDPSSTAFPLVAALLVPGSDVTI---LNV-LMNPTRTGLILTLQEMGADIEVINP
S70378_1      EIDVPG-IDTNACHFVKCPLVKGQQYDIKYTWNVPKIAPKSENVVVTVKLGIDNGVLACA
              80       90      100      110      120      130
```

```

      300      310      320      330      340      350
CP4_EPSPS.pe RLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAFAEGATVMNGLEELRVK
S70378_1      IATHGKIRD
              140
```

CP4_EPSPS.pep
UP:DEPDER2_1

LOCUS DEPDER2_1
DEFINITION Dermatophagoides farinae mRNA for mite allergen Der f II
precursor, complete cds, clone:pFL2.
DATE 02-FEB-1999
ACCESSION D10448
NID g217305 . . .

SCORES Initl: 36 Initn: 36 Opt: 79 z score: 104.5 E(): 0.29
Smith-Waterman score: 79; 29.3% identity in 82 aa overlap

```

      190      200      210      220      230
CP4_EPSPS.pe LLAGLNTPGITTVIEPIMTRDHTKMLQGFGANLTVET--DAD-GVRTIRLEGRGKLTGQ
DEPDER2_1      QVDVKDCANNEIKKVMVDGCHGSDPCIHRGKPFLEALFDANQNTKTAKIEIKASLDGL
              20       30       40       50       60       70
```

```

      240      250      260      270      280      290
CP4_EPSPS.pe VIDVPGDPSSTAFPLVAALLVPGSDVTI---LNV-LMNPTRTGLILTLQEMGADIEVINP
DEPDER2_1      EIDVPG-IDTNACHFMKCPLVKGQQYDIKYTWNVPKIAPKSENVVVTVKLGIDNGVLACA
              80       90      100      110      120      130
```

```

      300      310      320      330      340      350
CP4_EPSPS.pe RLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAFAEGATVMNGLEELRVK
```

DEPDER2_1 IATHGKIRD
140

CP4_EPSPS.pep
UP:S63984

P1:S63984 - major allergen Api g 1 - Apium graveolens
C:Species: Apium graveolens
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 08-Sep-1997
C:Accession: S63984; S63974; S52851
R:Breiteneder, H.; Hoffmann-Sommergruber, K.; O'Riordain, G.; Susani, M.;
Ahorn, H.; Ebner, C.; Kraft, D.; Scheiner, O. . . .

SCORES Initl: 38 Initn: 38 Opt: 75 z-score: 99.2 E(): 0.57
Smith-Waterman score: 75; 22.6% identity in 137 aa overlap

```

      160      170      180      190      200      210
CP4_EPSPS.pe PKTPTPTITYRVPMASAVQKSAVLLAGLNTPGITTVIEPIMTRDHTKMLQGF--GANLTV
S63984      MGVSQTHVLELTSSVSAEKIFQGGFVIDVDTVL
              10       20       30
```

```

      220      230      240      250      260      270
CP4_EPSPS.pe ETDADGV-RTIRLEGRG-KLTGGQVIDVP-GDPSSTAFPLVAALLVPGSDVTILNVLMNPT
S63984      PKAAPGAYKSVEIKGDDGGPGTLKIITLPDGGPITTM-----TLRIDGVNKEALTFDYSVI
              40       50       60       70       80
```

```

      280      290      300      310      320      330
CP4_EPSPS.pe RTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKG-VTVPEDRAPSMIDEYPILA
S63984      DGDILLGFIESIENHVVLVPTADGGSICKTTAIFHT--KGDVAVPEENIKYANEQNTALF
              90      100      110      120      130      140
```

```

      340      350      360      370      380      390
CP4_EPSPS.pe VAAAFAGATVMNGLEELRVKESDRLSAVANGLKLVGDCDEGETSLVVRGRPDGKGLGN
S63984      KALEAYLIAN
              150
```

CP4_EPSPS.pep
UP:DEPDER3_1

LOCUS DEPDER3_1
DEFINITION Dermatophagoides farinae mRNA for mite allergen Der f II
precursor, partial cds, clone:pFL11.
DATE 02-FEB-1999
ACCESSION D10449
NID g217307 . . .

SCORES Init1: 36 Initn: 36 Opt: 73 z-score: 97.6 E(): 0.71
 Smith-Waterman score: 73; 38.0% identity in 50 aa overlap

	190	200	210	220	230
CP4_EPSPS.pe	LLAGLNTPGITTVIEPI	IMTRDHTKMLQGFGANLTVET--DAD	GVRTIRLEGRGKLTGQ		
DEPDER3_1	QVDVKDCANNEIKKVMVDGCHGSDPCI	IHRGKPF	TLEALFDANQNTKTAKIEIKASLDGL		
	20	30	40	50	60
	240	250	260	270	280
CP4_EPSPS.pe	VIDVPG-DPSSTAF---PLVAA	LLVPGSDVTILNVL	MNPNTRTGLILTLQEMGADIEVINP		
DEPDER3_1	EIDVPGIDTNACHFMKCPLVK	GQQYDAKYTNVVKIAPKSENVVVTVKLVGDNGVLACAI			
	80	90	100	110	120
					130

CP4_EPSPS.pep
 UP:RNMG_ASPRE

ID RNMG_ASPRE STANDARD; PRT; 176 AA.
 AC P04389; P19792;
 DT 20-MAR-1987 (REL. 04, CREATED)
 DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE RIBONUCLEASE MITOGILLIN PRECURSOR (EC 3.1.27.-) (RESTRICTOCIN) (MAJOR

SCORES Init1: 52 Initn: 52 Opt: 68 z-score: 89.8 E(): 1.9
 Smith-Waterman score: 68; 41.7% identity in 24 aa overlap

	50	60	70	80	90	100
CP4_EPSPS.pe	ITGLLEGEDVINTGKAMQAMGARIRKEGDTWI	IDGV-GNGGLLAPEAPLDFGNAATGCRL				
RNMG_ASPRE	KRLLYSQAKAESNSHHAPLSDGKTGSSYPHWFT	TNGYDNGKLIKGRTPIKFGKADCDRPP				
	50	60	70	80	90	100
	110	120	130	140	150	160
CP4_EPSPS.pe	TMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPL	REMGVQVKSEDDGRLPVTLRGPKTPTPI				
RNMG_ASPRE	KHSQNGMGKDDHYLLEFPTFPDGHDKFDSKPKED	PGPARVIYTPNKFVFCGIVAHQRG				
	110	120	130	140	150	160

CP4_EPSPS.pep
 UP:A46497

P1:A46497 - major allergen I - Aspergillus fumigatus
 N:Alternate names: 18K IgE binding protein; allergen I/a; IgE-binding
 ribotoxin; ribonuclease mitogillin (EC 3.1.-.-) I; ribonucleotoxin-related
 protein
 C:Species: Aspergillus fumigatus
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 20-Mar-1998

SCORES Init1: 52 Initn: 52 Opt: 68 z-score: 89.8 E(): 1.9
 Smith-Waterman score: 68; 41.7% identity in 24 aa overlap

	50	60	70	80	90	100
CP4_EPSPS.pe	ITGLLEGEDVINTGKAMQAMGARIRKEGDTWI	IDGV-GNGGLLAPEAPLDFGNAATGCRL				
A46497	KRLLYNQAKAESNSHHAPLSDGKTGSSYPHWFT	TNGYDNGKLIKGRTPIKFGKADCDRPP				
	50	60	70	80	90	100
	110	120	130	140	150	160
CP4_EPSPS.pe	TMGLVGVDYDFDSTFIGDASLTKRPMGRVLNPL	REMGVQVKSEDDGRLPVTLRGPKTPTPI				
A46497	KHSQNGMGKDDHYLLEFPTFPDGHDKFDSKPKED	PGPARVIYTPNKFVFCGIVAHQRG				
	110	120	130	140	150	160

CP4_EPSPS.pep
 UP:AH11_ARAHY

ID AH11_ARAHY STANDARD; PRT; 614 AA.
 AC P43237;
 DT 01-NOV-1995 (REL. 32, CREATED)
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE ALLERGEN ARA H 1, CLONE P17 (ARA H I).

SCORES Init1: 44 Initn: 44 Opt: 69 z-score: 82.2 E(): 5.1
 Smith-Waterman score: 69; 35.4% identity in 48 aa overlap

	160	170	180	190	200	210
CP4_EPSPS.pe	TLRGPKTPTPITYRVPMASQVKS	AVLLAGLNTPGITTVIEPI	IMTRDHTKMLQGFAN-			
AH11_ARAHY	HALRIPSGFISYILNRHDNQLRVAKISMPVNT	PGQFEDFFPASSRDQS-SYLQGS	SRNT			
	260	270	280	290	300	310
	220	230	240	250	260	270
CP4_EPSPS.pe	LTVETDA DGVRTIRLEGRGKLTGQVIDVPGD	PSSTAFPLVAA	LLVPGSDVTILNVL			
AH11_ARAHY	LEAAFNAEFNEIRVLL	LEENAGGEQEERGQRRRSTRSSDNEGVIVKVSKEHVQELTKHAK				
	320	330	340	350	360	370

CP4_EPSPS.pep
 UP:AH12_ARAHY

ID AH12_ARAHY STANDARD; PRT; 626 AA.
 AC P43238;
 DT 01-NOV-1995 (REL. 32, CREATED)
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE ALLERGEN ARA H 1, CLONE P41B (ARA H I).

SCORES Init1: 44 Initn: 44 Opt: 69 z-score: 82.1 E(): 5.2
 Smith-Waterman score: 69; 35.4% identity in 48 aa overlap

```
ID      MP93_POAPR      STANDARD;      PRT;      307 AA.
AC      P22286;
DT      01-AUG-1991 (REL. 19, CREATED)
DT      01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT      01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
```

```

              440              450
CP4_EPSPS.pe FPEFMDLMAGLGAKIELSDTKAA
              :  ::  |  :
O81982       LGDYL-LEQGM
              130

```

CP4_EPSPS.pep
UP:GB_PLN2_HAPROFILN

TRANSLATE of: input_30.rsfl(haprofiln) check: 1235 from: 1 to: 402
generated symbols 1 to: 134.
Description: Y15210 Helianthus annuus mRNA for profilin. 9/1998
Accession/ID: Y15210

=====General comments=====

LOCUS HAPROFILN 402 bp mRNA PLN 10-SEP-1998

SCORES Initl: 35 Initn: 35 Opt: 59 z-score: 80.7 E(): 6.1
Smith-Waterman score: 68; 24.8% identity in 101 aa overlap

```

      320      330      340      350      360      370
CP4_EPSPS.pe EDRAPSMID EYPI LA VAAFAEGATVMNGLEELR--VKESDRLSAVA-NGLKLNGVD--C
              ||::  :||  |::  :||  :|:  :|:
GB_PLN2_HAPR GTGQHLTSAAILGLDGTVWAQSAKFPQFKPEEMKGIIEKFDEAGTLAPTGMFIAGAKYVM
              20      30      40      50      60      70
```

```

      380      390      400      410      420      430
CP4_EPSPS.pe DEGETSLVVRGRPDGKGLGNASGAATHLDHRIAMSFVLMGLVSENPTVDDATMIATS
              :||  :|:|:|  ||  |  :  :  :  :|:|:|  :|:|:|  :|:|:|
GB_PLN2_HAPR LQGEFGAVIRGK---KGAG---GICI-----KKTGQAMINGIYDE-PVAPQCNMVVER
              80      90      100      110      120
```

```

      440      450
CP4_EPSPS.pe FPEFMDLMAGLGAKIELSDTKAA
              :  :  :  |  |:
GB_PLN2_HAPR LGDYL-LEQGM
              130
```

CP4_EPSPS.pep
UP:PROF_BETVE

ID PROF_BETVE STANDARD; PRT; 133 AA.
AC P25816;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PROFILIN (POLLEN ALLERGEN BET V 2) (BET V II). . . .

SCORES Initl: 36 Initn: 36 Opt: 57 z-score: 78.3 E(): 8.3
Smith-Waterman score: 67; 27.6% identity in 76 aa overlap

```

      340      350      360      370      380      390
CP4_EPSPS.pe AAFAEGATVMNGLEELRVKESDRLSAVANGKLNGVD--CDEGETSLVVRGRPDGKGLG
              ||:|:|:|  :||:|:|  :||:|:|  :||:|:|  :||:|:|
PROF_BETVE AQSSFPQFKPQEITGIMKDFEEPGHLAPTGLHLGGIKYMVIIQGEAGAVIRGK---KGS
              40      50      60      70      80      90
```

```

      400      410      420      430      440      450
CP4_EPSPS.pe NASGAATHLDHRIAMSFVLMGLVSENPTVDDATMIATSFPEFMDLMAGLGAKIELSD
              |  :  :  :|:|  :|:|  :|:|  :|:|  :|:|
PROF_BETVE ---GITI-----KKTGQALVFG-IYEEPTPGQCNMVERLG DYLIQGL
              100      110      120      130
```

CP4_EPSPS.pe TKA

CP4_EPSPS.pep
UP:MPO1_ORYSA

ID MPO1_ORYSA STANDARD; PRT; 263 AA.
AC Q40638;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE MAJOR POLLEN ALLERGEN ORY S 1 PRECURSOR (ORY S I). . . .

SCORES Initl: 31 Initn: 31 Opt: 60 z-score: 77.2 E(): 9.6
Smith-Waterman score: 60; 36.6% identity in 41 aa overlap

```

      90      100      110      120      130      140
CP4_EPSPS.pe LLAP EAPLDFGNAATGCRLTMGLVGVDFTFIGDASLTKRPMG----RVLNPLREMGV
              :|  ||  |  ||  |  ||  |  ||  |  :|  :|
MPO1_ORYSA SSSLLACVVVAAMVSPSPAGHPKVPPGNITTSYGDKWLEARPPGMVRPVLAP-KDNGG
              10      20      30      40      50      60

      150      160      170      180      190      200
CP4_EPSPS.pe QVKSEGDGRLPVTLRGPKTPTPIYRVPMASQVKSALLAGLNTPGITTVEIPIMTRDH
              :|  |:  |
MPO1_ORYSA ACGYKDVDKAPFLGMNSCGNDPIFKDGKCGSCFEIKCSKPEACSDKPALIHVTD MNDEP
              70      80      90      100      110      120
```

! Distributed over 1 thread.
! Start time: Mon Sep 13 12:08:03 1999
! Completion time: Mon Sep 13 12:08:07 1999

! CPU time used:
! Database scan: 0:00:00.3
! Post-scan processing: 0:00:00.4
! Total CPU time: 0:00:00.8
! Output File: CP4_EPSPS.UPDATE2