Study Title

Bioinformatic Analysis of Proteins in Golden Rice 2 to Assess Potential Allergenic Cross-Reactivity

Authors

Richard E. Goodman, John Wise

Study Completed On

May 2, 2006

Performing Laboratory

Food Allergy Research and Resource Program
Food Science and Technology
University of Nebraska
143 Food Science & Technology
Lincoln, NE 68583-0955

Laboratory Project ID

Study Number: BIO-02-2006

Study Number:	BIO-02-2006
Title:	Bioinformatic Analysis of Proteins in Golden Rice 2 to Assess Potential Allergenic Cross-Reactivity
Facility:	Food Allergy Research and Resource Program Food Science and Technology University of Nebraska 143 Food Industry Complex Lincoln, NE 68583-0955 USA
Authors:	Richard E. Goodman, John Wise University of Nebraska
Study Start Date:	April 4, 2006
Study Completion Date:	June 5, 2006
Records Retention:	All study specific raw data and a copy of the final report will be retained at the Food Allergy Research and Resource Program, University of Nebraska.

Signatures of Final Report Approval:

Richard E. Goodman, Ph.D.	5 June 2006
Study Author	Date
Stephen L. Taylor, Ph.D.	5 June 2006
FARRP Co-Director	Date

Table of Contents

Section **Page**

Tables

Figures

Appendices

Abbreviations and Definitions

aa	Amino acid
AO6	http://www.AllergenOnline.com/ database version 6.0
Ber e 1	An allergenic 2S albumin from *Bertholletia excelsa* (positive control)
BLASTP	Algorithm used to find local high scoring alignments between a pair of protein sequences (using databases on Entrez)
CRTI	Carotene desaturase I from *Erwinia uredovora*
CTP	Chloroplast transit peptide from garden pea
CTP-CRTI	Fused CTP-CRTI
Entrez NCBI	A public genetic database maintained by the National Center for Biotechnology Information (NCBI) at the National Institutes of Health, Bethesda, MD. Protein entries in the Entrez search and retrieval system are maintained by the NCBI of the National Institutes of Health (U.S.A.)
FASTA3	Algorithm used to find local high scoring alignments between a pair or protein sequences (using the AllergenOnline database)
GI	A unique identification number assigned by NCBI to each sequence in the database
GR2	Golden Rice version 2 is a genetically modified rice that includes three genes introduced through biotechnology (see http://www.goldenrice.org).
PMI	Phosphomannose isomerase from *Escherichia coli*
PSY	Phytoene synthase from *Zea mays*

1.0 Summary

The three proteins expressed by the genes introduced into Golden Rice 2 (GR2) through genetic engineering were evaluated using bioinformatic approaches to identify any potential sequence matches to allergenic proteins that might indicate an elevated risk of allergic cross reactivity in consumers. Two sequence alignment and similarity scoring algorithms were used in these comparisons. A FASTA3 algorithm (Pearson, 2000) was used with the default scoring matrix (BLOSUM 50) to evaluate overall alignment of each query sequence compared to all sequences in AllergenOnline, looking for matches of low *E* score values (< 1e-7) and/or greater than 50% identity as an indication of potential cross-reactivity. FASTA3 was also used to search for any segment of 80 or more amino acids that aligned with a match of 35% identity or more compared to any sequence in AllergenOnline, as suggested as a lower limit for considering potential cross-reactivity (Codex, 2003). Finally, BLASTP was used to identify any significant similarity to any newly reported "allergen" sequences not found in AllergenOnline version 6.0, by searching the non-redundant (nr) sequences in the NCBI-Entrez Protein Database. FASTA and BLASTP algorithms perform relatively similar comparisons and although the scoring matrices and scoring penalties are slightly different. Both programs compare amino acid sequences (*i.e.,* primary protein structure), and the alignment data may be used to infer higher order structural similarities (*i.e.,* secondary and tertiary protein structures). Proteins that share a high degree of similarity throughout the entire length usually share secondary structure, common three-dimensional folds and functions. Because of the structural similarity, closely related homologoues may share immunological cross-reacivitity, including IgE binding. Proteins that contain two or more IgE epitopes, or proteins with single epitopes that are cross-linked may bind IgE on the surface of mast cells in sensitized (allergic) individuals. If a sufficient number of allergens are bound on a mast cell, it will degranulate, releasing immune mediators such as histamine and leukotrienes that induce the allergic reaction. Highly similar homologues are more likely to be bound by the same IgE, because of the increased likelihood that the IgE binding structures are conserved. It is therefore important to understand whether a protein that may be introduced into food products might be sufficiently similar to cause allergic cross-reactions.

The bioinformatics analysis is meant to identify proteins that are sufficiently similar to known allergens to suspect possible allergic cross-reactivity. In the case where a protein introduced into a GM plant matched an allergen to the extent that cross-reactivity was suspected (by accepted international criteria, e.g. at least 35% identity over an 80 amino acid match), an evaluation would be made as to whether a sufficient numer of individuals with allergies to the matched allergen could be identified to serve as donors for sera to test IgE binding, also whether the data demonstrating the allergenic risk of the matched allergen would indicate a significant potential risk posed by the potentially cross-reactive protein. If so, specific serum testing would be performed to test the possibility of cross-reactivity.

The work was performed in the Food Allergy Research and Resource Program (FARRP) laboratory of the University of Nebraska. Dr. Jorge Mayer of the Golden Rice Network

(http://www.goldenrice.org) was consulted to ensure the sequences were representative of those in GR2. Dr. Mayer assured us that we used the correct sequences.

Based on the criteria described, no biologically relevant sequence/structural similarities were identified by FASTA3 or BLASTP between any of the three query protein sequences in Golden Rice 2 and any known allergen(s).

2.0 Introduction

Grain from rice (*Oryza sativa*) is an important staple food in many countries, where there are rates of vitamin A insufficiency. Rice plants were transformed with genes encoding phytoene synthase (PSY), accession number U32636 from maize (*Zea mays*), carotene desaturase (CRTI), accession number D90087 from a bacterium (*Erwinia uredovora*), which was fused with the gene segment of a chloroplast transit peptide (CTP), amino acids 1-57 from the ribulose-1,5-bisphosphate decarboxylase gene (Misawa et al., 1993) of garden pea (*Pisum sativum*). Together these two genes increase levels of vitamin A precursor to levels that are sufficient for maintenance of health (Paine et al., 2005). The purpose of adding the CTP is to target newly synthesized CRTI to the chloroplast organelles in order to efficiently produce beta-carotene. Although the CTP is cleaved from CRTI as it is transferred across the chloroplast membrane, some small amount of the CTP may remain fused to the CRTI that has not yet transferred into chloroplasts, and we have taken this into consideration in the bioinformatics analysis by testing the complete coded peptide. In order to efficiently select appropriately transformed rice plants, a marker gene encoding the enzyme phosphomannose isomerase protein (PMI), accession number M15380 was transferred from *Escherichia coli*. An assessment of potential allergenicity for these three proteins is required before the Golden Rice 2 varieties could be approved for production of food and feed or consumption in non-producing areas (Codex Alimentarius Commission, 2003).

Most severe or life-threatening allergic reactions to foods are due to Type I (immediate hypersensitivity) reactions. Such reactions occur when a sufficient number of the individual allergenic proteins, each bound by two separate IgE antibodies, are attached to IgE receptors on the surface of mast cells, stimulating the cells to release histamine and other allergenic mediators. The two binding sites (epitopes) are generally different structures or sequences, unless two identical proteins are covalently linked (e.g. through a disulfide bond). Sensitization to produce the IgE antibodies requires a third allergen-specific recognition sequence within a protein, a CD4+ T-cell epitope, to provide appropriate stimulation to the IgE-producing B cells (Plaut and Zimmerman, 1993). Cross-reactions may be induced in sensitized individuals by a different protein if it contains IgE-binding epitopes that are identical or nearly identical to those of the sensitizing protein. However, cross-reactive ligands usually bind with lower affinity than the immunogen to the antibody and are therefore typically less effective at inducing an immune response (Berzofsky *et al.*, 1993).

The allergenicity assessment includes evaluation of the historical allergenicity of the source of the genes (Codex Alimentarius Commission, 2003). If the source is a major allergen individuals with allergies to the source would be asked to provide serum for IgE binding studies to determine if the protein expressed by the donated gene is an allergen. Extensive literature searches have not identified scientific data implicating either *E. uredovora* or *E. coli* as allergens. *Zea mays* (maize or corn) is known to cause allergies in a very limited number of individual consumers, due to the presence of a few specific proteins. Pastorello et al. (2000) identified a lipid transfer protein as the major food allergen of maize. A new study reports the 50 kDa gamma-zein protein may act as a cross-reactive allergen for a few individuals with IgE to the major almond allergen (Lee et al., 2005), which may be the same protein identified in an earlier study (Pasini et al., 2002). There are a few reports of allergy to garden peas (*Pisum sativum*), with clear IgE binding to a few major proteins (Sell et al., 2005), but no reports of IgE binding to the CTP, full-length or mature ribulose-1,5-bisphosphate carboxylase. There are very few reports of allergic reactions to maize and therefore the probability of that the PSY gene is an allergen is extremely small, and the likelihood of finding appropriate allergic individuals to test is also very small. Therefore, the focus of the allergenicity assessment is on the similarity of the three proteins to those of known allergens, based on a bioinformatics analysis.

A bioinformatics search can not be expected to identify allergens de novo. Instead it is meant to identify proteins that are sufficiently similar to a known allergen to suspect that specifically allergic individuals might suffer an allergic reaction because their IgE specific to the allergen would bind strongly to the potentially cross-reactive protein (Goodman et al., 2005). The results of the search are used to identify transgenic proteins that should be tested by specific serum screens, and to identify the specific allergic populations that might be at risk, and should be used as serum donors for specific testing (Codex Alimentarius Commission, 2003).

Screening the amino acid sequences of proteins introduced into plants for similarity to sequences of known allergens is one of many assessments performed to evaluate product safety (Metcalfe *et al.*, 1996). The extent of sequence similarities between a protein of interest and database sequences of allergens can be assessed using the FASTA sequence alignment tool. While the FASTA and BLASTP programs compare amino acid sequences (*i.e.*, primary protein structure) and provide optimum local sequence alignments. The results may be used to infer similarity of higher order structure (*i.e.*, secondary and tertiary protein structures). Proteins that share a high degree of similarity throughout the entire length are often homologous. Homologous proteins share secondary structure and common three-dimensional folds (Pearson, 1996). Homologous proteins are more likely to share allergenic cross-reactive conformational and linear epitopes than unrelated proteins; however, the degree of similarity between homologues varies widely. It is uncommon to find proteins that are cross-reactive and share less than 50% identity across the full length of the protein sequences, while it is common to find proteins with greater than 70% identity that cause cross-reactions (Aalberse, 2000).

The early bioinformatics comparisons relied on short amino-acid match comparisons designed to identify matches of eight contiguous amino acids (Metcalfe et al., 1996), based on the theoretical consideration that an IgE epitope may be around eight or so amino acids in length. While some IgE epitopes may be as short as five amino acids (Banerjee *et al.*, 1999; Beezhold *et al.*, 1999), the majority of characterized IgE-linear epitopes are eight amino acids or longer (Chatchatee *et al.*, 2001; Reese *et al.*, 1999; Shin *et al.*, 1998). Although many of these reports have demonstrated IgE binding, few have tested the affinity (avidity) of the binding, or the allergic significance of the *in vitro* binding, and it is clear from some reports that high affinity binding requires eight or more amino acids (Banerjee *et al.*, 1999; Rabjohn *et al.*, 1999). In order to induce an allergic reaction, at least two IgE binding epitopes are required on each allergenic protein (or two epitopes that may be identical must be in close enough proximity to cross-link IgE bound to receptors on mast cells or basophils). Because of concerns that some potential IgE epitopes are smaller than eight amino acids, the scientific advisory panel to FAO/WHO (2001) recommended reducing the size of a match to any contiguous six amino acids. However, as demonstrated in a number of bioinformatics papers, searches for six amino acids causes an unacceptably high rate of false positive identities (Hileman et al., 2002; Silvanovich et al., 2006). Further, the positive predictive value of finding identities as short as eight amino acids has not been demonstrated, and the likelihood of false positive matches is predicted to be relatively high, although lower than matches of six amino acids (Silvanovich et al., 2006; Bjorklund et al., 2005). The Codex Alimentarius Commission guidelines (2003) clearly state that the bioinformatics test and all other tests for the allergenicity assessment should be based on scientifically sound methods. They recommended a FASTA (or BLASTP) search with a threshold of concern of 35% identity as a primary bioinformatics method. Based on information in the scientific literature for sequence identities of clinically demonstrated cross-reactivities, it appears that there are very few cross-reactive pairs of proteins that would not be identified by a scanning window of 80 amino acids with a threshold of 35-45% identity (Goodman and Hefle, 2005; Bjorklund et al., 2005), to compare against a well founded allergen database. Based on the evidence, we are using FASTA alignments to evaluate the transgenic proteins.

In order to maximize the probability of finding any important match of the proteins in Golden Rice 2, compared to any allergen, we have used both an overall (full sequence length) FASTA, to find matches of greater than 50% identity as having an increased likelihood of cross-reactivity and we are using an 80 amino acid sliding window to scan the transgenic sequences against the allergen database (AllergenOnline, version 6.0, released January, 2006), using FASTA to search for matches of 35% identity or more. In addition, we have used BLASTP to search the current version of NCBI general protein sequences, with the delimiter of “allergen” to indentify matches to any sequences that were not in NCBI during the last update of AllergenOnline. Parameters of the search methods (default settings) and database characteristics are presented in the body of the report. In order to demonstrate that these search methods are reliable, we have also compared the sequence of a known cross-reactive major allergen (Ber e 1) of Brazil nut to database entries using the same criteria and bioinformatics tools.

3.0 Purpose

The purpose of this study was to identify potentially relevant matches in amino acid sequence between the three proteins in Golden Rice 2 and known and putative allergens, which might indicated an elevated risk of allergic reaction to those with specific existing allergies, if they consume grain or food products made from Golden Rice 2.

4.0 Methods

4.1 *Protein databases*

4.1.1 The AllergenOnline version 6.0 (updated January, 2006) database was used for the primary comparisons to allergens. The database is available at http://www.allergenonline.com/. The complete allergen list of 1537 known or putative allergens is shown in Appendix 1 (without sequences). This curated database is maintained by the Food Allergy Research and Resource Program of the University of Nebraska. It includes known and putative allergens that have identified from food, airway, contact and venom allergen sources. Many sequences that are designated as allergen or associated with allergy in the database entry have been filtered out based on specific criteria. All database entries are linked to sequences in the National Center for Biotechnology Information (NCBI) of the National Institutes of Health (NIH).

4.1.2 NCBI Entrez Protein Database. Protein entries in the Entrez search and retrieval system, maintained by the NCBI of the National Institutes of Health (U.S.A.), and were compiled from the following sources: SwissProt, PIR, PRF, PDB and translations from annotated coding regions in GenBank and RefSeq. The database is potentially updated or modified daily, and therefore the date of sequence searches by BLASTP is relevant to the dataset used in the BLASTP searches. For the present report, all BLASTP searches were run on 4 April, 2006. The database includes all non-redundant GenBank coding sequence translations, the PDB (protein database), SwissProt, PIR and PRF. The total database size on that date was 3,538,507 sequences. The searches with BLASTP were performed using sequences from all organisms, with a filter of "allergen" in the Entrez query, bringing the total number of sequences searched down to 2,529.

4.2 *Sequence database search strategies.*

4.2.1 FASTA3 overall search of AllergenOnline. The potential sequential and inferred structural similarity of each of the three proteins transferred into Golden Rice 2, and a positive control allergen, Ber e 1 (2S albumin from *Bertholletia excelsa*) was evaluated relative to all allergens in AllergenOnline version 6.0, by performing a FASTA3 search with these sequences using the default search and scoring criteria of Pearson (2000). The default scoring matrix is BLOSUM 50 (Henikoff and Henikoff, 1992 and 1996). The penalty for each gap inserted into query or searched sequences to obtain optimal alignments is calculated as (–q + –r*k), where q (10) is an initial penalty for each independent gap, r (2) is a penalty for each amino acid position within the gap and k is the number of amino acid positions within the gap (Reese and Pearson, 2002). The default word size (ktup) is two (Pearson, 2000). The FASTA3 version used in these searches was 3.4t25b1, dated November 12, 2004 (ftp://ftp.virginia.edu/pub/fasta/). Statistical values are calculated for each search and compared to expected values, as illustrated in the histogram of the computer output. Alignment of regions containing low sequence complexity may lead to irrelevant alignments and are expected to show skewed distributions and should be reanalyzed after removing the low complexity regions (Pearson, 2000). Very small expectation values (*E* values) indicate probable evolutionary homology, and structural similarity. While the *E* value default for FASTA3 is set to 10, a value that does not indicate significant similarity, distantly related sequences will generally have *E* values less than 0.01, and highly similar sequences that probably represent close homology are more likely to have *E* values less than 1e –7. If the statistical parameters calculated for an alignment between the query protein and any one allergen appear to indicate significant similarity, the percent identity over the length of the intact proteins may be evaluated for possible cross-reactivities in those sensitized to the matched allergen. As discussed by Aalberse (2000), a protein sharing greater than 70% identity over its length, relative to an allergen is likely to be cross-reactive, or share IgE binding. Those that have less than 50% identity are not very likely to be cross-reactive.

4.2.2 FASTA3 of AllergenOnline by 80 aa segments. The identification of relatively short regions of high identity shared by a query sequence and an allergen may indicate similarities that could also share IgE binding, or cross-reactivity. Based on the recommendation of Codex (2003), the FASTA3 algorithm was used to compare all possible contiguous amino acid segments of each of the four test proteins against all sequences listed in AllergenOnline. Every possible contiguous 80-amino acid sequence of each query protein was searched, beginning with amino acids 1-80, then 2-81, 3-82 and so on until the last 80 amino acid segment of each protein was compared with the database on AllergenOnline, using the same FASTA3 algorithm used for the overall-comparison. In this case, only the *E* values and percent identities [(# identical residues / 80 or more amino acids) * 100%)] were evaluated to consider potential cross-reactivity. Alignments of less than 80 amino acids in length were recalculated to normalize the identity to an 80 amino acid score, by increasing the denominator to 80, without altering the numerator. Therefore an alignment with 38 identical amino acids over a length of 40 (=95%), would be recalculated to 47.5%. The reason for the adjustment is that alignments less than 80 amino acids long may have very high identities, and would therefore be more likely to act as a cross-reactive allergen if the matched region represented an IgE epitope, than longer alignments of markedly lower identity scores. When the FASTA3 program inserts gaps in the query sequence to provide optimal alignment, the length of the aligment will exceed 80 amino acids. Rather than “correcting” the alignment identity scores, the same criterion of 35% identity is maintained as per the recommendation of Codex (2003). The rationale by Codex (2003) for recommending that alignments of >35% identity over segments as short as 80 amino acids is that proteins sometimes contain structural motifs that are comprised of sequences much shorter than the intact protein, and that these structural motifs may include a conformational IgE binding epitope. In such a situation, the overall sequence identity for the aligned proteins may be significantly less than 35%, even though a short region could contain an important cross-reactive epitope. This criterion is more conservative than empirical data would suggest is common for cross-reactive proteins (Aalbersee, 2000). It should also help to identify potentially cross-reactive proteins that are not true homologues of an allergen that have significant local identities that might provide an immunological target for IgE antibodies in those with allergies to the matched allergen. The current algorithm will identify any match containing 28 identical aligned, identical amino acids in an overlap of 80 amino acids.

The output of the 80 amino acid FASTA3 search includes a table of each allergen that was matched, and the total number of 80-amino acid matches of greater than 35% identity. An additional file of all alignments is maintained to allow location and further evaluation of any significant match.

4.2.3 BLASTP of NCBI Entrez *allergen*. BLASTP and FASTA3 are unique computer algorithms that provide similar local alignments and results if the appropriate scoring matrices and criteria are used. The BLASTP is available on the NCBI Entrez website (http://www.ncbi.nlm.nih.gov/BLAST/). The current version is BLASTP 2.2.13 (Nov-27-2005). A BLAST search was used comparing each complete query sequence against the entire Entrez Protein database, with a limit option selected to query entries for “allergen”, to align only with proteins identified as allergens. The purpose of this BLAST search is to ensure that a significant match with a newly discovered allergenic sequence that has not yet been entered into AllergenOnline, is not overlooked. The default conditions for Expectation value (10), word size (3), scoring matrix (BLOSUM62) and gap penalties (-11 existance, plus extension of -1) were used. The “low complexity” filter was turned off, meaning that irrelevant low complexity alignment matches may be identified as significant, so each identified match should be further evaluated to understand the significance. A comparison with the low complexity filter on may also be checked to determine if that is the cause of the match. Evaluation of the *E* value, the length of the alignment and the percent identity of any identified match is necessary to judge the significance of any alignment.

The *E* value (expectation score) calculated by BLASTP is similar to the value calculated by FASTA3. The *E* values and identities calculated for alignments from BLASTP should be judged similarly to those from FASTA3. Since the sequences in AllergenOnline should also be found in the Entrez Protein Database, it is likely that matches of high similarity by FASTA3 using AllergenOnline will also be identified during the BLASTP search. Sequences of allergens that were published since the compilation of AllergenOnline version 6.0 maybe identified as an alignment by BLASTP, that were not identified by FASTA3.

5.0 Results and Discussion. The complete results for all three of the proteins introduced into GR2: CTP-CRTI, PSY and PMI and the positive allergen control, Ber e 1, will be presented in order.

5.1 Bioinformatics results for CTP-CRTI. The full-length sequence of CTP-CRTI (549 amino acids) as expressed in Golden Rice 2 (Figure 1) includes a fused chloroplast transit peptide from Rubisco small subunit of pea (aa 1-57 of GI 20858, *Pisum sativum*), and the full-length coding region (492 aa) of the crtI gene identified in GI:22474502 from *Erwinia uredovora* (*Pantoea ananatis*) identified by Misawa et al. (1990). The fusion gene was created as a DNA construct described by Misawa et al. (1993). The transit peptide coding region was identified by Coruzzi et al. (1984), and is used to target the expressed CRTI protein to chloroplasts in GR2. While the transit peptide is cleaved from the translated protein during transmembrane movement, the intact protein may be present at some low concentration and the bioinformatics evaluation therefore includes the transit peptide. Detailed search results and the complete sequence are presented in Appendix 2.

CTP-CRTI

```
1   MASMISSSAV TTVSRASRGQ SAAVAPFGGL KSMTGFPVKK VNTDITSITS NGGRVKCMKP
61  TTVIGAGFGG LALAIRLQAA GIPVLLLEQR DKPGGRAYVY EDQGFTFDAG PTVITDPSAI
121 EELFALAGKQ LKEYVELLPV TPFYRLCWES GKVFNYDNDQ TRLEAQIQQF NPRDVEGYRQ
181 FLDYSRAVFK EGYLKLGTVP FLSFRDMLRA APQLAKLQAW RSVYSKVASY IEDEHLRQAF
241 SFHSLLVGGN PFATSSIYTL IHALEREWGV WFPRGGTGAL VQGMIKLFQD LGGEVVLNAR
301 VSHMETTGNK IEAVHLEDGR RFLTQAVASN ADVVHTYRDL LSQHPAAVKQ SNKLQTKRMS
361 NSLFVLYFGL NHHHDQLAHH TVCFGPRYRE LIDEIFNHDG LAEDFSLYLH APCVTDSSLA
421 PEGCGSYYVL APVPHLGTAN LDWTVEGPKL RDRIFAYLEQ HYMPGLRSQL VTHRMFTPFD
481 FRDQLNAYHG SAFSVEPVLT QSAWFRPHNR DKTITNLYLV GAGTHPGAGI PGVIGSAKAT
541 AGLMLEDLI
```

Figure 1. CTP-CRTI query sequence for GR2. The full-length protein (amino acid) sequence of the fused chloroplast transit peptide (CTP) from Pisum sativum (underlined) is followed immediately by the full-length sequence of the carotene desaturase enzyme I of Erwinia uredovora.

5.1.1 FASTA3 overall search of AllergenOnline version 6.0 with CTP-CRTI. Results of the FASTA3 search of the chloroplast transit peptide-carotene desaturase I protein against AllergenOnline version 6.0 did not identify any significant alignment with an allergen. The statistical-fit histogram indicated only minor deviations from the expected distribution of alignments, indicating the amino acid sequence is not unusually skewed. Scoring results for the four best scoring alignments demonstrate no significant matches with any allergen (see Table 1). The best scoring aligned sequence, a Bet v 1 homologue from rice, does not appear to be even a distantly related homologue of CRTI or the chloroplast transit peptide. The identity (24%) is markedly below the level that is likely to indicate cross-reactivity (< 50% identity, Aalberse, 2000) and it is also below the 35% identity level over 80 or more amino acids suggested by Codex (2003) as a match that may possibly be cross-reactive. A test of the CRTI sequence without the fused CTP resulted in the same alignments and scores (data not shown).

Table 1. **Overall FASTA3 search of AllergenOnline version 6.0 with CTP-CRTI.** Highest scoring alignments of known and putative allergens in AllergenOnline version 6.0, compared to the chloroplast transit peptide-carotene desaturase I protein, using FASTA3.

Sequence GI #	Organism	Description	Length	*E* score	% Identity	aa Alignment
15624049	*Oryza sativa* rice	Bet v 1 allergen-like protein	208	0.18	24.3	111
21335406	*Plodia interpunctella* moth	unnamed allergen	705	1.2	25.9	108
3021373	*Glycine max* soybean	profilin	131	2.1	24.7	89
729970	*Dermatophagoides farinae* house dust mite	HSP70 MAG29	145	2.8	44.4	36

5.1.2 FASTA3 of AllergenOnline by 80 aa segments of CTP-CRTI. The identity results from FASTA3 alignments of every possible 80 amino acid segment of chloroplast transit peptide-carotene desaturase I protein was compared individually to all sequences in AllergenOnline version 6.0 were all less than 35% over 80 amino acids. There were no matches of greater than 35% identity over 80 amino acids.

5.1.3 BLASTP of NCBI Entrez "allergen" with CRT-CRTI. The full-length chloroplast transit peptide-CRTI protein was compared to sequences in NCBI-Entrez, which were designated as "allergen" in the NCBI database on 4 April, 2006. There were only two alignments with *E* scores below "10", the default cutoff. Note that the best alignment was to a PR10-allergen-like protein from *Pseudomonas aeruginosa* (GI:88192990), with an *E*-score of approximately 1.3 and an identity match of 27% in an alignment of 73 amino acids. These results are markedly below the criteria that would be relevant for potential cross-reactivity.

Table 2. **BLASTP of NCBI Entrez "allergen" with CTP-CRTI.** The scoring alignments with *E* scores below 10, to putative allergens in the NCBI Entrez database on 4 April, 2006, were compared to thefull-length sequence of CTP-CRTI, using BLASTP.

Sequence GI #	Organism	Description	Length	*E* score	% Identity	aa Alignment
88192990	*Pseudomonas auriginosa* bacteria	PR-10-like protein Pa1206	157	1.3	27%	73
74016535	*Burkholderia ambifaria* bacteria	alkyl hydroperoxide reductase	218	3	31%	41

5.1.4 Bioinformatics Summary for CTP-CRTI. None of the results from the three bioinformatics searches with the chloroplast transit peptide-carotene desaturase I sequence meet the criteria that might suggest potential allergenic cross-reactivity.

5.2 Bioinformatics results for PSY. The full length sequence (410 amino acids) of the phytoene synthase protein (Figure 2) from *Zea mays* was used as the query sequence. Detailed search results and the complete sequence are presented in Appendix 3.

PSY

```
1   MAIILVRAAS PGLSAADSIS HQGTLQCSTL LKTKRPAARR WMPCSLLGLH PWEAGRPSPA
61  VYSSLPVNPA GEAVVSSEQK VYDVVLKQAA LLKRQLRTPV LDARPQDMDM PRNGLKEAYD
121 RCGEICEEYA KTFYLGTMLM TEERRRAIWA IYVWCRRTDE LVDGPNANYI TPTALDRWEK
181 RLEDLFTGRP YDMLDAALSD TISRFPIDIQ PFRDMIEGMR SDLRKTRYNN FDELYMYCYY
241 VAGTVGLMSV PVMGIATESK ATTESVYSAA LALGIANQLT NILRDVGEDA RRGRIYLPQD
301 ELAQAGLSDE DIFKGVVTNR WRNFMKRQIK RARMFFEEAE RGVTELSQAS RWPVWASLLL
361 YRQILDEIEA NDYNNFTKRA YVGKGKKLLA. LPVAYGKSLL LPCSLRNGQT
```

Figure 2. PSY query sequence for GR2. The full-length protein (amino acid) sequence of the phytoene synthase protein from *Zea mays*.

5.2.1 FASTA3 overall search of AllergenOnline with PSY. Results of the FASTA3 search of the PSY protein against AllergenOnline version 6.0 did not identify any significant alignment with any allergen. The statistical-fit histogram indicated only minor deviations from the expected distribution of alignments. Scoring results for the five best scoring alignments demonstrate no significant matches with any allergen (see Table 3). The aligned sequences may be due to chance alone as they are relatively short compared to the length of the proteins and are of low percent identities. The identities are markedly below that suggested by Codex (2003) for potential allergenic cross-reactivity (greater than 35% identity over 80 amino acids). They are markedly below 50% identity, which might represent a likely allergenically cross-reactive protein (Aalberse, 2000).

Table 3. **Overall FASTA3 search of AllergenOnline version 6.0 with PSY.** Highest scoring alignments of known and putative allergens in AllergenOnline version 5.0, compared to the PSY, using FASTA3.

Sequence GI #	Organism	Description	Length	*E* score	% Identity	aa Alignment
399672	*Equus caballus* Horse	horse serum albumin Equ c 3	607	0.98	29.5%	61
11277083	*Bos taurus* Cow	lipocalin-like protein Bos d 2.0102	156	1	25.2%	107
633938	*Canis familiaris* Dog	serum albumin	265	1.4	29.6%	54
2739154	*Schistosoma japonicum* Parasite	tegumental protein	191	1.7	29.%	62
11277082	*Bos taurus* cow	lipocalin Bos d 2.0103	156	2	40%	45

5.2.2 FASTA3 of AllergenOnline by 80 aa segments of PSY. The identity results from FASTA3 alignments of every possible 80 amino acid segment of PSY compared individually to all sequences in AllergenOnline version 6.0 were all less than 35% over 80 amino acids. There were no alignments of 35% identity or more over any 80 amino acid segment to any sequence in the AllergenOnline database.

5.2.3 BLASTP of NCBI Entrez "allergen" with PSY. The sequence of the PSY protein was compared to sequences in NCBI-Entrez, which were designated as "allergen" in the NCBI database on 4 April, 2006. The best alignment scores are shown in Table 4. The smallest *E* score was 1.7, with an alignment to the nitrile-specifier protein of Pieris rapae (a herbivorous moth). This alignment is not significant.

Table 4. **BLASTP of NCBI Entrez "allergen" with PSY.** The five highest scoring alignments of known or putative allergens in the NCBI Entrez database on 20 January, 2005, were compared to the mature form of PSY, using BLASTP.

Sequence GI #	Organism	Description	Length	*E* score	% Identity	aa Alignment
40288346	*Pieris rapae* white cabbage moth	nitrile specifier protein	632	1.7	29%	106
6900304	*Lepidoglyphus destructor* storage mite	tropomyosin	284	2.8	23%	95
2739154	*Shistosoma japonicum* parasite	tegumental protein	191	6.3	29%	62
2440053	*Dermatophagoides pteronyssinus* house dust mite	tropomyosin	284	6.3	27%	47
2353266	*Dermatophagoides pteronyssinus* house dust mite	tropomyosin	284	6.3	27%	47

5.2.5 Bioinformatics Summary for PSY.

None of the results from the three bioinformatics searches with phytoene synthase meet the criteria that might suggest potential allergenic cross-reactivity.

5.3 Bioinformatics results for PMI.

The full length sequence of phosphomannose isomerase protein from the *E. coli* gene is 391 amino acids long (Figure 4).

PMI

```
1   MQKLINSVQN YAWGSKTALT ELYGMENPSS QPMAELWMGA HPKSSSRVQN AAGDIVSLRD
61  VIESDKSTLL GEAVAKRFGE LPFLFKVLCA AQPLSIQVHP NKHNSEIGFA KENAAGIPMD
121 AAERNYKDPN HKPELVFALT PFLAMNAFRE FSEIVSLLQP VAGAHPAIAH FLQQPDAERL
181 SELFASLLNM QGEEKSRALA ILKSALDSQQ GEPWQTIRLI SEFYPEDSGL FSPLLLNVVK
241 LNPGEAMFLF AETPHAYLQG VALEVMANSD NVLRAGLTPK YIDIPELVAN VKFEAKPANQ
301 LLTQPVKQGA ELDFPIPVDD FAFSLHDLSD KETTISQQSA AILFCVEGDA TLWKGSQQLQ
361 LKPGESAFIA ANESPVTVKG HGRLARVYNK L
```

Figure 3. PMI query sequence for GR2. The full-length amino acid sequence of the phosphomannose isomerase protein from *E. coli*

5.3.1 FASTA3 overall search of AllergenOnline version 6.0 with PMI.

Results of the FASTA3 search of the phosphomannose isomerase protein, against AllergenOnline version 6.0 did not identify any significant alignment with an allergen. The statistical-fit histogram indicated only minor deviations from the expected distribution of alignments. Scoring results for the five highest scoring alignments demonstrate no significant matches with any allergen (see Table 5). The identities appear to be markedly below that suggested by Codex (2003) as potentially cross-reactivity (greater than 35% identity over 80 amino acids). They are markedly below 50% identity, which might represent a likely allergenically cross-reactive protein (Aalberse, 2000).

Table 5. Overall FASTA3 search of AllergenOnline version 6.0 with PMI. Highest scoring overall alignments of known and putative allergens in AllergenOnline version 6.0, compared to the full-length form of PMI using FASTA3.

Sequence GI #	Organism	Description	Length	*E* score	% Identity	aa Alignment
18770	*Glycine max* soybean	trypsin inhibitor subtype A	217	0.36	29.6%	152
18772	*Glycine max* soybean	trypsin inhibitor subtype AB	217	0.5	27.7%	148
1304264	*Triticum aestivum* wheat	alpha-gliadin	259	0.79	24.4%	82
256429	*Glycine max* soybean	Kunitz trypsin inhibitor	216	0.82	29.1%	151
1076486	*Glycine max* soybean	cim1 expansin	277	1.9	22.6%	93

5.3.2 FASTA3 of AllergenOnline by 80 aa segments of PMI. The identity scores from FASTA3 alignments of every possible 80 amino acid segment of phosphomannose isomerase, compared individually to all sequences in AllergenOnline version 6.0 were all less than 35% over 80 amino acids.

5.3.3 BLASTP of NCBI Entrez "allergen" with PMI. The full-length sequence of the phosphomannose isomerase protein was compared to sequences in NCBI-Entrez, which were designated as "allergen" in the NCBI database on 4 April, 2006. The smallest *E* score was 2.2, which was against the tegumental antigen of *Schistosoma japonica*, GI:2739154. There were only three alignments with significant *E* scores (<10). In addition, all of the percent identities were below 35% over 80 amino acids as suggested by Codex (2003) and were markedly below the minimum level of 50% identity expected for cross-reactive proteins (Aalberse, 2000).

Table 6. **BLASTP of NCBI Entrez "allergen" with PMI.** The three alignments of known or putative allergens in the NCBI Entrez database on 4 April, 2006, were compared to the full-length of PMI using BLASTP.

Sequence GI #	Organism	Description	Length	*E* score	% Identity	aa Alignment
2739154	*Schistosoma japonicum* parasite	tegumental antigent	191	2.3	36%	33
14279169	*Olea europaea* olive	beta-glucanase-like protein	460	6.6	21%	89
555616	*Glycine max* soybean	cytokinin induced message	277	8.5	22%	93

5.3.5. Bioinformatics Summary for PMI. None of the results from the three bioinformatics searches with the phosphomannose isomerase protein sequence meet the criteria that might suggest potential cross-reactivity.

5.4 Bioinformatics results for Ber e 1. The sequence of the Bertholletia excelsa 2S albumin allergen (Ber e 1) used as a positive allergen control in this report is 146 amino acids long (GI:112754) as shown in Figure 4.

Ber e 1

```
1   MAKISVAAAA LLVLMALGHA TAFRATVTTT VVEEENQEEC REQMQRQQML SHCRMYMRQQ
61  MEESPYQTMP RRGMEPHMSE CCEQLEGMDE SCRCEGLRMM MMRMQQEEMQ PRGEQMRRMM
121 RLAENIPSRC NLSPMRCPMG GSIAGF
```

Figure 4. Ber e 1 query sequence used as a positive allergen control. The sequence of this protein has no relationship to GR2, but is used in this report as a positive control to demonstrate bioinformatics results that would be expected with a protein that has been demonstrated to cause some clinical cross-reactivity in some patients, due to structural and sequence similarities.

5.4.1 FASTA3 overall search of AllergenOnline Ber e 1. Results of the FASTA3 search of the 2S albumin protein of Bertholletia excelsa; against AllergenOnline version 6.0 aligned with itself and other known allergenic proteins. The statistical-fit histogram indicated marked deviations from the expected distribution of alignments, with a very narrow and high bell-shaped curve compared to the random distribution expected for most sequences. The reason is that the database contains a high number of 2S albumin proteins. Scoring results for the five highest scoring alignments demonstrate significant matches with four allergens (see Table 7). The identities are above 35% identity over 80 amino acids, which was suggested by Codex (2003) as possibly representing a cross-reactive match. One alignment is markedly above 50% identity, which suggests a strong possibility of cross-reactivity (it is an isoform of the Brazil nut allergen, the others are around 40% identity over the full length, and there is some evidence that a few individuals may have cross-reactivity across these tree nuts, although it is hard to differentiate primary sensitization from cross-reactivity.

Table 7. **Overall FASTA3 search of AllergenOnline version 6.0 with Ber e 1.** Highest scoring overall alignments of known and putative allergens in AllergenOnline version 6.0, compared to the 2S albumin allergen Ber e 1 using FASTA3.

Sequence GI #	Organism	Description	Length	*E* score	% Identity	aa Alignment
112754	*Bertholletia excela* Brazil nut	2S albumin	146	1.3e-31	100%	146
17713	*Bertholletia excela* Brazil nut	2S albumin large SU	154	2.1e-20	69%	153
1794252	*Juglans regia* English walnut	2S albumin	139	3.8e-10	42%	138
31321942	*Juglans nigra* black walnut	2S albumin	161	5.5e-10	39%	145
28207731	*Carya illinoinensis* pecan	2S albumin	143	2.2e-9	40%	145

5.4.2 FASTA3 of AllergenOnline by 80 aa segments with Ber e 1. The identity scores from FASTA3 alignments of every possible 80 amino acid segment of Ber e 1, compared individually to all sequences in AllergenOnline version 6.0 showed 9 proteins (including two Brazil nut entries) had matches of greater than 35% over 80 amino acids. Although there are a few reports of cross-reactivity, or at least co-sensitization of some individuals with Brazil nut and walnut, sesame seed and peanut, there is not yet any direct published evidence that the 2S albumins are cross-reactive across these species. Based on the identity matches walnut, pecan and sesame seeds are the most likely candidates for cross-reactivity, which could be tested using sera from individuals with clinical reactivity to those species.

Table 8. FASTA3 search of all possible 80 amino acid segments of Ber e 1 compared to AllergenOnline version 6.0. Highest scoring overall alignments of known and putative allergens in AllergenOnline version 6.0, compared to the 2S albumin allergen Ber e 1 using FASTA3.

Sequence GI #	Organism	Description	Highest scoring 80mer (%)	# 80mers over 35%	Overall % Identity
112754	*Bertholletia excela* Brazil nut	2S albumin	100%	67	100%
17713	*Bertholletia excela* Brazil nut	2S albumin large SU	73%	67	69%
1794252	*Juglans regia* English walnut	2S albumin	49%	59	42%
28207731	*Carya illinoinensis* pecan	2S albumin	48%	59	40%
31321942	*Juglans nigra* black walnut	2S albumin	45%	59	39%

5.4.3 BLASTP of NCBI Entrez “allergen” with Ber e 1. The sequence of the 2S albumin from Brazil nut was compared to sequences in NCBI-Entrez, that were designated as “allergen” in the NCBI database. .

Table 9. **BLASTP of NCBI Entrez "allergen" with Ber e 1.** A number of significanct alignments were identified between the 2S albumin of Brazil nut and other allergens. The percent identity falls relatively rapidly and based on limited data, it would require serum testing to determine whether the proteins from these plant seeds are truly cross-reactive, though it should be suspected that the incidence of cross-reactivity would not be very high due to the limited number of reports of co-reactivity for those with allergies to some tree nuts.

Sequence GI #	Organism	Description	Length	*E* score	% Identity	aa Alignment
112754	Bertolletia excelsa Brazil nut	2S albumin	146	1e-83	100%	146
1794252	Juglans regia English walnut	2S albumin	139	2e-28	42%	136
31321942	Juglans nigra black walnut	2S albumin	161	4e-28	39%	143
28207731	Carya illinoinsis pecan	2S albumin	143	3e-26	41%	136
112762	Ricinus communis castor bean	2S albumin precursor	258	5e-16	38%	104

5.4.5. Bioinformatics Summary for Ber e 1. A few significant alignments were identified by all three methods of bioinformatics that match Ber e 1 with other 2S allergens. These results suggest that it would be important to test for IgE cross-reactivity if one were to consider moving this Brazil nut protein into any other source. These set of data also demonstrates the consistency between bioinformatics methods and helps to demonstrate the utility of evaluating the proteins of Golden Rice by the same methods.

6.0 Conclusions

Bioinformatics analyses were performed on the three proteins introduced into the rice genome to make Golden Rice 2, in order to produce biologically significant amounts of beta-carotene in this staple crop. Results from the three search methods or strategies were negative, that is the criteria for suspected corss-reactivity were not reached. This demonstrates that there is not expected to be any significant risk of cross-reactivity for those who are allergic to known allergens. In fact, based on these results, it would not be possible to identify allergic individuals who would be at a heightened risk. The AllergenOnline version 6.0 list of 1537 known or putative allergens is presented in Appendix 1. Detailed search results are presented in Appendices 2-4. Results for the positive allergen control are shown in Appendix 5. The results of these tests will be maintained as archived records at FARRP, the Department of Food Science & Technology, the University of Nebraska, Lincoln.

7.0 References

Aalberse RC. (2000). Structural biology of allergens. *J Allergy Clin Immunol* 106:228-238.

Altenbach SB, Pearson KW, Leung FW, Sun SSM. (1987). Cloning and sequence analysis of a cDNA encoding a Brazil nut protein exceptionally rich in methionine. *Plant Mol Biol* 8:239-250.

Banerjee B, Greenberger PA, Fink JN, Kurup VP. (1999). Conformational and linear B-cell epitopes of Asp f 2, a major allergen of Aspergillus fumigatus, bind differently to immunoglobulin E antibody in the sera of allergic bronchopulmonary aspergillosis patients. *Infect Immun* 67:2284-2291.

Beezhold DH, Hickey VL, Slater JE, Sussman GL. (1999). Human IgE-binding epitopes of the latex allergen Hev b 5. *J Allergy Clin Immunol* 103:1166-1172.

Beyer P, Al-Babili S, Ye Xudong, Lucca P, Schaub P, Welsch R, Potrykus I. (2002). Golden Rice: Introducing the β-carotene biosynthesis pathway into rice endosperm by genetic engineering to defeat vitamin A deficiency. J Nutr 132:506S-510S.

Bjorklund AK, Soeria-Atmadja D, Zorzet A, Hammerling U, Gustafsson MG. (2005). Supervised identification of allergen-representative peptides for in silico detection of potentially allergenic proteins. Bioinformatics 21(1):39-50.

Burks AW, Shin D, Cockrell G, Stanley JS, Helm RM, Bannon GA. (1997). Mapping and mutational analysis of the IgE-binding epitopes on Ara h 1, a legume vicilin protein and a major allergen in peanut hypersensitivity. *Eur J Biochem* 245:334-339.

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.

Coruzzi G, Broglie R, Edwards C, Chua N-H. (1984). Tissue-specific and light-regulated expression of a pea nuclear gene encoding the small subunit of ribulose-1,5-bisphosphate carboxylase. The EMBO Journal 3(8):1671-1679.

FAO/WHO. (2001). Evaluation of allergenicity of genetically modified foods. Report of a Joint FAO / WHO Expert Consultation on Allergenicity of Foods Derived from Biotechnology. Food and Agriculture Organization of the United Nations (FAO), Rome, Italy 2001.

Gendel SM. (1998). The use of amino acid sequence alignments to assess potential allergenicity of proteins used in genetically modified foods. *Adv Food Nutr Res* 42:45-62.

Goodman RE, Hefle SL. 2005. Gaining perspective on the allergenicity assessment of genetically modified crops. Expert Rev Clin Immunol 1(4):561-578.

Goodman RE, Hefle SL, Taylor SL, van Ree R. 2005. Assessing genetically modified crops to minimize the risk of increased food allergy: A review. Int Arch Allergy Immunol 137:153-166.

Goodman RE, Silvanovich A, Hileman RE, Bannon GA, Rice EA, Astwood JD. (2002). Bioinformatic methods for identifying known or potential allergens in the safety assessment of genetically modified crops. *Comments Toxicology* 8:251-269.

Henikoff, S., and Henikoff, J. G. (1992). Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci U S A* **89**:10915-10919.

Henikoff, J. G., and Henikoff, S. (1996). Blocks database and its applications. *Methods Enzymol* **266**:88-10.5

Hileman RE, Silvanovich A, Goodman RE, Rice EA, Holleschak G, Astwood JD, Helfe SL. (2002). Bioinformatic methods for allergenicity assessment using a comprehensive allergen database. *Int Arch Allergy Immunol* 128:280-291.

Lee SH, Benmoussa M, Sathe SK, Roux KH, Teuber SS, Hamaker BR. (2005). A 50 kDa maize gamma-zein has marked cross-reactivity with the almond major protein.

Miles JS, Guest JR. (1984). Nucleotide sequence and transcriptional start point of the phosphomannose isomerase gene (manA) of Escherichia coli. *Gene* 32(1-2):41-48.

Misawa Y, Nakagawa M, Kobayashi K, Yamano S, Izawa Y, Nakamura K, Harashima K. (1990). Elucidation of the *Erwinia uredovora* carotenoid biosynthetic pathway by functional analysis of gene products expressed in *Escherichia coli*. J Bacteriol 172(12):6704-6712.

Misawa N, Yamano S, Linden H, de Felipe MR, Lucas M, Ikenaga H, Sandmann G. (1993). Functional expression of the *Erwinia uredovora* carotenoid biosynthesis gene crtI in transgenic plants showing an increase of β-carotene biosynthesis activity and resistance to the bleaching herbicide norflurazon. The Plant J 4(5):833-840.

Paine JA, Shipton CA, Chaggar S, Howells RM, Kennedy MJ, Vernon G, Wright SY, Hinchcliffe E, Adams JL, Silverstone AL, Drake R. (2005). Imroving the nutritional value of Golden Rice through increased pro-vitamin A content. Nat Biotechnol 23(4):482-487.

Pasini G, Simonato B, Curioni A, Vincenzi S, Cristaudo A, Sntucci B, Peruffo AD, Giannattasio M. (2002). IgE-mediated allergy to corn: a 50 kDa protein, belonging ot the Reduced Soluble Proteins, is a major allergen. Allergy 57(2):98-106.

Pastorello EA, Farioli L, Pravettoni V, Ispano M, Scibola E, Trambaioli C, Giuffrida MG, Ansaloni R, Godovac-Zimmermann J, Conti A, Fortunato D, Ortolani C. (2000). The maize major allergen, which is responsible for food-induced allergic reactions, is a lipid transfer protein. J Allergy Clin Immunol 106(4):744-751.

Rabjohn P, Helm EM, Stanley JS, West CM, Sampson HA, Burks AW, Bannon GA. (1999). Molecular cloning and epitope analysis of the peanut allergen Ara h 3. *J Clin Invest* 103:535-542.

Reese JT, Pearson WR. (2002). Empirical determination of effective gap penalties for sequence comparison. Bioinformatics 18:1500-1507.

Sell M, Steinhart H, Paschke A. (2005). Influence of maturation on the alteration of allergenicity of green pea (*Pisum sativum* L.). J Agric Food Chem 53:1717-1722.

Silvanovich A, Nemeth MA, Song P, Herman R, Tagliani L, Bannon GA. (2006). The value of short amino acid sequence matches for prediction of protein allergenicity. Toxicol Sci 90(1):252-258.

Seq #	Species	Common_name	Allergen	Type	Comment	Gi_no	Acc_no	aa_length
1	Acanthamoeba castellanii	Parasitic protozoa	n/a	Protozoan	NULL	2781014	1PRQ	125
2	Acanthamoeba castellanii	Parasitic protozoa	n/a	Protozoan	NULL	9257090	1F2K_B	125
3	Acanthocheilonema viteae	Nematode	n/a	Worm (parasite)	Ladder protein	4102959	AAD01630	131
4	Acarus siro	Mite	Aca s 13	Aero Mite	Lipid binding protein	4049356	CAA07241	64
5	Actinidia chinensis	Kiwi	Act c 1	Food Plant	Actinidain protease	113285	P00785	380
6	Actinidia deliciosa	Kiwi	n/a	Food Plant	pKIWI501	450239	AAA53071	184
7	Actinidia deliciosa	Kiwi	n/a	Food Plant	Actinidin	166317	AAA32629	380
8	Actinidia deliciosa	Kiwi	n/a	Food Plant	Actinidain protease	15984	CAA34486	380
9	Actinidia deliciosa	Kiwi	Act c 2	Food Plant	(Act d 2?)	71057064	CAI38795	225
10	Actinidia deliciosa	Kiwi	n/a	Food Plant	NULL	40807635	AAR92223	116
11	Aedes aegypti	Yellow fever mosquito	Aed a 2	Venom or Salivary	D7 protein	118216	P18153	321
12	Aedes aegypti	Yellow fever mosquito	n/a	Venom or Salivary	Salivary gland protein - 30 kDa (Aed a 3)	2114497	AAB58417	253
13	Aedes aegypti	Yellow fever mosquito	Aed a 1	Venom or Salivary	Apyrase	556272	AAC37218	562
14	Aedes aegypti	Yellow fever mosquito	n/a	Venom or Salivary	NULL	18568322	AAL76031	215
15	Aedes aegypti	Yellow fever mosquito	n/a	Venom or Salivary	Putative 18.2 kDa secreted protein	18568332	AAL76036	158
16	Aedes albopictus	Asian tiger mosquito	n/a	Venom or Salivary	NULL	56417500	AAV90691	210
17	Aedes albopictus	Asian tiger mosquito	n/a	Venom or Salivary	NULL	56417502	AAV90692	210
18	Aedes albopictus	Asian tiger mosquito	n/a	Venom or Salivary	NULL	56417504	AAV90693	271
19	Aedes albopictus	Asian tiger mosquito	n/a	Venom or Salivary	NULL	56417506	AAV90694	266
20	Aedes albopictus	Asian tiger mosquito	n/a	Venom or Salivary	NULL	56417508	AAV90695	204
21	Agrostis alba	Bent grass	n/a	Aero Plant	Grass Group I (fragment) (Agr a 1)	7489357	F58493	35
22	Agrostis alba	Bent grass	n/a	Aero Plant	Grass Group I (fragment) (Agr a 1)	7489358	G58493	35
23	Agrostis alba	Bent grass	n/a	Aero Plant	Grass Group I (fragment) (Agr a 1)	320606	E37396	26
24	Alnus glutinosa	Alder	n/a	Aero Plant	Calcium-binding protein (Aln g 4)	3319651	CAA76831	85
25	Alnus glutinosa	Alder	Aln g 1	Aero Plant	PRP Bet v 1 family (Aln g 1)	261407	AAB24432	160
26	Alternaria alternata	Fungus	Alt a 1	Aero Fungi	Major allergen 1	1842045	AAB47552	157

27	Alternaria alternata	Fungus	n/a	Aero Fungi	NULL	45680856	AAS75297	157
28	Alternaria alternata	Fungus	Alt a 12	Aero Fungi	60S acidic ribosomal protein P1	1350779	P49148	110
29	Alternaria alternata	Fungus	Alt a 6	Aero Fungi	Enolase	14423684	Q9HDT3	438
30	Alternaria alternata	Fungus	Alt a 1	Aero Fungi	Major allergen 1	1421808	AAB03877	135
31	Alternaria alternata	Fungus	Alt a 7	Aero Fungi	Flavodoxin 1	467619	CAA55069	204
32	Alternaria alternata	Fungus	Alt a 5	Aero Fungi	60s Ribosomal P2 phosphoprotein	1850540	AAB48041	113
33	Alternaria alternata	Fungus	n/a	Aero Fungi	Nuclear transport factor 2	21748153	CAD38167	124
34	Alternaria alternata	Fungus	Alt a 10	Aero Fungi	Aldehyde dehydrogenase	1169290	P42041	495
35	Alternaria alternata	Fungus	Alt a 4	Aero Fungi	Disulfidisomerase (Thioredoxin)	1006624	CAA58999	433
36	Alternaria alternata	Fungus	Alt a 5	Aero Fungi	60s Ribosomal P2 phosphoprotein	467617	CAA55066	113
37	Alternaria alternata	Fungus	n/a	Aero Fungi		4097481	AAD00097	190
38	Alternaria alternata	Fungus	Alt a 1	Aero Fungi	Major allergen 1	21913174	AAM77471	115
39	Alternaria alternata	Fungus	Alt a 3	Aero Fungi	Heat shock 70 kDa protein	14423730	P78983	152
40	Alternaria arborescens	Fungus	n/a	Aero Fungi	NULL	49476521	AAT66594	137
41	Alternaria argyranthemi	Fungus	n/a	Aero Fungi	NULL	49476475	AAT66571	138
42	Alternaria blumeae	Fungus	n/a	Aero Fungi	NULL	49476497	AAT66582	137
43	Alternaria brassicae	Fungus	n/a	Aero Fungi	NULL	49476533	AAT66600	137
44	Alternaria brassicicola	Fungus	n/a	Aero Fungi	Alt b1	20279107	AAM18717	158
45	Alternaria capsici	Fungus	n/a	Aero Fungi	NULL	49476511	AAT66589	137
46	Alternaria carotiincultae	Fungus	n/a	Aero Fungi	NULL	49476489	AAT66578	137
47	Alternaria cetera	Fungus	n/a	Aero Fungi	NULL	49476471	AAT66569	138
48	Alternaria cheiranthi	Fungus	n/a	Aero Fungi	NULL	49476495	AAT66581	137
49	Alternaria cinerariae	Fungus	n/a	Aero Fungi	NULL	49476531	AAT66599	137
50	Alternaria conjuncta	Fungus	n/a	Aero Fungi	NULL	49476477	AAT66572	138
51	Alternaria crassa	Fungus	n/a	Aero Fungi	NULL	49476501	AAT66584	137
52	Alternaria cucumerina	Fungus	n/a	Aero Fungi	NULL	49476515	AAT66591	137
53	Alternaria dauci	Fungus	n/a	Aero Fungi	NULL	49476499	AAT66583	137
54	Alternaria dumosa	Fungus	n/a	Aero Fungi	NULL	49476525	AAT66596	137
55	Alternaria eryngii	Fungus	n/a	Aero Fungi	NULL	49476541	AAT66604	137
56	Alternaria euphorbiicola	Fungus	n/a	Aero Fungi	NULL	49476543	AAT66605	137
57	Alternaria japonica	Fungus	n/a	Aero Fungi	NULL	49476539	AAT66603	137
58	Alternaria limoniasperae	Fungus	n/a	Aero Fungi	NULL	49476527	AAT66597	137
59	Alternaria longipes	Fungus	n/a	Aero Fungi	NULL	49476523	AAT66595	137
60	Alternaria macrospora	Fungus	n/a	Aero Fungi	NULL	49476503	AAT66585	137
61	Alternaria metachromatica	Fungus	n/a	Aero Fungi	NULL	49476485	AAT66576	137
62	Alternaria mimicula	Fungus	n/a	Aero Fungi	NULL	49476535	AAT66601	137
63	Alternaria mouchaccae	Fungus	n/a	Aero Fungi	NULL	49476473	AAT66570	138
64	Alternaria oregonensis	Fungus	n/a	Aero Fungi	NULL	49476481	AAT66574	128
65	Alternaria petroselini	Fungus	n/a	Aero Fungi	NULL	49476491	AAT66579	137

66	Alternaria photistica	Fungus	n/a	Aero Fungi	NULL	49476479	AAT66573	138
67	Alternaria porri	Fungus	n/a	Aero Fungi	NULL	49476507	AAT66587	137
68	Alternaria pseudorostrata	Fungus	n/a	Aero Fungi	NULL	49476505	AAT66586	137
69	Alternaria radicina	Fungus	n/a	Aero Fungi	NULL	49476487	AAT66577	137
70	Alternaria smyrnii	Fungus	n/a	Aero Fungi	NULL	49476493	AAT66580	137
71	Alternaria solani	Fungus	n/a	Aero Fungi	NULL	49476513	AAT66590	137
72	Alternaria sonchi	Fungus	n/a	Aero Fungi	NULL	49476529	AAT66598	138
73	Alternaria tagetica	Fungus	n/a	Aero Fungi	NULL	49476509	AAT66588	137
74	Alternaria tenuissima	Fungus	n/a	Aero Fungi	NULL	49476519	AAT66593	137
75	Ambrosia artemisiifolia	Short ragweed	Amb a 1.2	Aero Plant	Antigen E	113476	P27760	398
76	Ambrosia artemisiifolia	Short ragweed	Amb a 6	Aero Plant	Lipid-transfer protein	14285595	O04004	118
77	Ambrosia artemisiifolia	Short ragweed	Amb a 1.3	Aero Plant	Antigen E	539049	C53240	397
78	Ambrosia artemisiifolia	Short ragweed	Amb a 1.2	Aero Plant	Antigen E	539048	B53240	398
79	Ambrosia artemisiifolia	Short ragweed	Amb a 1.3	Aero Plant	Antigen E	166443	AAA32669	397
80	Ambrosia artemisiifolia	Short ragweed	Amb a 1.1	Aero Plant	Antigen E	113475	P27759	396
81	Ambrosia artemisiifolia	Short ragweed	Amb a 2	Aero Plant	Antigen K	539051	E53240	397
82	Ambrosia artemisiifolia	Short ragweed	Amb a 2	Aero Plant	Antigen K	113479	P27762	397
83	Ambrosia artemisiifolia	Short ragweed	Amb a 1.4	Aero Plant	Antigen E	113478	P28744	392
84	Ambrosia artemisiifolia	Short ragweed	n/a	Aero Plant	NULL	34851178	AAP15201	131
85	Ambrosia artemisiifolia	Short ragweed	n/a	Aero Plant	NULL	34851182	AAP15203	133
86	Ambrosia artemisiifolia	Short ragweed	n/a	Aero Plant	NULL	34851180	AAP15202	131
87	Ambrosia artemisiifolia	Short ragweed	Amb a 1.3	Aero Plant	Antigen E	113477	P27761	397
88	Ambrosia artemisiifolia (elatior)	Short ragweed	Amb a 3	Aero Plant	Ra3	416636	P00304	101
89	Ambrosia artemisiifolia (elatior)	Short ragweed	Amb a 5	Aero Plant	Ra5	114090	P02878	45
90	Ambrosia psilostachya	Western ragweed	n/a	Aero Plant	Allergen p 5	515956	AAA20066	77
91	Ambrosia psilostachya	Western ragweed	n/a	Aero Plant	Allergen p 5	515957	AAA20068	77
92	Ambrosia psilostachya	Western ragweed	n/a	Aero Plant	Allergen p 5	515953	AAA20065	77
93	Ambrosia psilostachya	Western ragweed	n/a	Aero Plant	Allergen p 5	515954	AAA20067	77
94	Ambrosia psilostachya	Western ragweed	n/a	Aero Plant	Allergen p 5	515955	AAA20064	77
95	Ambrosia trifida	Giant ragweed	Amb t 5	Aero Plant	Allergen Ra5G	114091	P10414	73
96	Anacardium occidentale	Cashew	Ana o 3	Food Plant	2S albumin	24473800	AAL91665	138
97	Anacardium occidentale	Cashew	Ana o 1	Food Plant	Vicilin-like protein 7S cupin	21914823	AAM73730	538
98	Anacardium occidentale	Cashew	Ana o 1	Food Plant	Vicilin-like protein 7S cupin	21666498	AAM73729	536
99	Anacardium occidentale	Cashew	Ana o 2	Food Plant	Cupin 11S	25991543	AAN76862	457
100	Ananas comosus	Pineapple	Ana c 1	Food Plant	Profilin	14161637	AAK54835	131
101	Ancylostoma caninum	Dog hookworm	n/a	Worm (parasite)	Aspartic proteinase (fragment)	2144165	JC5077	442
102	Ancylostoma caninum	Dog hookworm	n/a	Worm (parasite)	Secreted protein 1	4884851	AAD31839	424
103	Ancylostoma caninum	Dog hookworm	n/a	Worm (parasite)	Secreted protein 2	3608493	AAC35986	218
104	Ancylostoma duodenale	Hookworm	n/a	Worm (parasite)	NULL	3719257	AAD13339	425
105	Anisakis simplex	Parasitic fish worm	Ani s 1	Worm (parasite)	Excretory gland protein (fragment)	7494508	A59069	17
106	Anisakis simplex	Parasitic fish worm	n/a	Worm (parasite)	NULL	31339067	BAC77154	194

107	Anisakis simplex	Parasitic fish worm	Ani s 4	Worm (parasite)	NULL	47605398	P83885	14
108	Anisakis simplex	Parasitic fish worm	n/a	Worm (parasite)	Troponin-like protein	6065738	CAB58171	161
109	Anisakis simplex	Parasitic fish worm	Ani s 3	Worm (parasite)	Tropomyosin	14423976	Q9NAS5	284
110	Anisakis simplex	Parasitic fish worm	n/a	Worm (parasite)	NULL	6065744	CAB58173	321
111	Anisakis simplex	Parasitic fish worm	Ani s 2	Worm (parasite)	Paramyosin	8117843	AAF72796	869
112	Anisakis simplex	Parasitic fish worm	Ani s 2	Worm (parasite)	Paramyosin	8453086	AAF75225	473
113	Anopheles gambiae	African malaria mosquito	n/a	Venom or Salivary	30 kDa protein	18389879	AAL68776	182
114	Anthoxanthum odoratum	Sweet vernal grass	n/a	Aero Plant	Group I allergen (form 2) (fragment)	7489360	C58493	32
115	Anthoxanthum odoratum	Sweet vernal grass	n/a	Aero Plant	Group 1 allergen (fragment)	320607	G37396	26
116	Anthoxanthum odoratum	Sweet vernal grass	n/a	Aero Plant	Group I allergen (form 1) (fragment)	7489359	B58493	32
117	Apis cerana	Indian honeybee	n/a	Venom or Salivary	Phospholipase A2	7435005	A59055	134
118	Apis cerana cerana	Indian honeybee	n/a	Venom or Salivary	NULL	24638082	Q9BMK4	134
119	Apis dorsata	Giant honeybee	n/a	Venom or Salivary	Melittin	69551	MEHBCD	26
120	Apis dorsata	Giant honeybee	n/a	Venom or Salivary	Phospholipase A2	7435004	B59055	134
121	Apis mellifera	Honeybee	n/a	Venom or Salivary	NULL	34921475	P83563	71
122	Apis mellifera	Honeybee	n/a	Venom or Salivary	NULL	61656214	NP_001013377	388
123	Apis mellifera	Honeybee	n/a	Venom or Salivary	NULL	58585250	NP_001011626	126
124	Apis mellifera	Honeybee	n/a	Venom or Salivary	Melittin	69552	MEHB2	27
125	Apis mellifera	Honeybee	Api m 1	Venom or Salivary	Phospholipase A2	24418862	P00630	167
126	Apis mellifera	Honeybee	n/a	Venom or Salivary	NULL	66534655	XP_624662	157
127	Apis mellifera	Honeybee	n/a	Venom or Salivary	Prepromelittin	5622	CAA26038	70
128	Apis mellifera	Honeybee	Api m 2	Venom or Salivary	Hyaluronidase	585279	Q08169	382
129	Apis mellifera	Honeybee	Api m 7	Venom or Salivary	Venom protease	22724911	AAN02286	405
130	Apium graveolens	Celery	n/a	Food Plant	NULL	32363125	P81943_3	24
131	Apium graveolens	Celery	Api g 1	Food Plant	PRP Bet v 1 family	2147642	S63984	154
132	Apium graveolens	Celery	Api g 4	Food Plant	Profilin	4761578	AAD29409	134
133	Apium graveolens	Celery	n/a	Food Plant	NULL	32363126	P81943_4	10
134	Apium graveolens	Celery	n/a	Food Plant	NADP-dependent malate dehydrogenase	6706333	CAB66003	570

135	Apium graveolens	Celery	n/a	Food Plant	Chlorophyll a-b binding protein	14423661	P92919	264
136	Apium graveolens	Celery	Api g 1.0201	Food Plant	PRP Bet v 1 family	1769847	CAA99992	159
137	Apium graveolens	Celery	n/a	Food Plant	Phosphoglyceromutase	6706331	CAB66002	559
138	Apium graveolens	Celery	n/a	Food Plant	NULL	32363124	P81943_2	30
139	Apium graveolens	Celery	n/a	Food Plant	NULL	33300921	P81943_1	22
140	Arachis hypogaea	Peanut	n/a	Food Plant	NULL	57118278	AAW34231	124
141	Arachis hypogaea	Peanut	Ara h 8	Food Plant	PR-10 Bet v 1 family	37499626	AAQ91847	157
142	Arachis hypogaea	Peanut	Ara h 2.02	Food Plant	Conglutin 2S albumin	26245447	AAN77576	172
143	Arachis hypogaea	Peanut	n/a	Food Plant	Glycinin (Ara h 3/Ara h 4)	21314465	AAM46958	538
144	Arachis hypogaea	Peanut	n/a	Food Plant	Mannose/glucose-binding lectin	951118	AAA74576	254
145	Arachis hypogaea	Peanut	n/a	Food Plant	Conglutin (patent sequence, partial synthetic?)	14347293	CAC41202	207
146	Arachis hypogaea	Peanut	Ara h 3	Food Plant	Glycinin (Cupin)	3703107	AAC63045	507
147	Arachis hypogaea	Peanut	n/a	Food Plant	NULL	46560476	AAT00596	428
148	Arachis hypogaea	Peanut	Ara h 4	Food Plant	Glycinin	5712199	AAD47382	530
149	Arachis hypogaea	Peanut	Ara h 1	Food Plant	Vicilin-like protein 7S (cupin)	1168391	P43238	626
150	Arachis hypogaea	Peanut	Ara h 5	Food Plant	Profilin	5902968	AAD55587	131
151	Arachis hypogaea	Peanut	n/a	Food Plant	NULL	46560472	AAT00594	303
152	Arachis hypogaea	Peanut	Ara h 6	Food Plant	Similar to Conglutin	5923742	AAD56337	129
153	Arachis hypogaea	Peanut	Ara h 1	Food Plant	Vicilin-like protein 7S (cupin)	1168390	P43237	614
154	Arachis hypogaea	Peanut	n/a	Food Plant	NULL	47933675	AAT39430	510
155	Arachis hypogaea	Peanut	Ara h 2	Food Plant	Conglutin-like	31322017	AAM78596	169
156	Arachis hypogaea	Peanut	n/a	Food Plant	Oleosin variant A	13161005	AAK13449	176
157	Arachis hypogaea	Peanut	n/a	Food Plant	Oleosin	47156059	AAT11925	176
158	Arachis hypogaea	Peanut	n/a	Food Plant	Oleosin variant B	13161008	AAK13450	176
159	Arachis hypogaea	Peanut	n/a	Food Plant	NULL	22135348	AAM93157	219
160	Arachis hypogaea	Peanut	n/a	Food Plant	Conglutin	17225991	AAL37561	144
161	Arachis hypogaea	Peanut	n/a	Food Plant	Agglutinin (lectin)	253289	AAB22817	273
162	Arachis hypogaea	Peanut	n/a	Food Plant	NULL	46560474	AAT00595	299
163	Arachis hypogaea	Peanut	Ara h 7	Food Plant	Similar to Conglutin	5931948	AAD56719	160
164	Argas reflexus	European pigeon tick	Arg r 1	Venom or Salivary	NULL	58371884	CAG26895	159
165	Artemisia vulgaris	Mugwort	Art v 1	Aero Plant	Major pollen allergen precursor	27818335	AAO24900	132
166	Arthroderma benhamiae	Fungus	n/a	Contact	Dipeptidyl aminopeptidase (see Trichophyton)	23894232	CAD23611	726
167	Arthroderma benhamiae	Fungus	n/a	Contact	Subtilase family (see Trichophyton)	23894248	CAD23615	405
168	Arthroderma benhamiae	Fungus	n/a	Contact	Subtilase family (see Trichophyton)	23894244	CAD23614	404
169	Arthroderma benhamiae	Fungus	n/a	Contact	Subtilase family (see Trichophyton)	23894240	CAD23613	292

170	Ascaridia galli	Nematode	n/a	Worm (parasite)	Fatty-acid/retinoid binding protein	3152922	AAC17174	135
171	Ascaris lumbricoides	Parasitic roundworm	n/a	Worm (parasite)	NULL	2735100	AAD13646	134
172	Ascaris lumbricoides	Parasitic roundworm	n/a	Worm (parasite)	ABA-1 (polymorphic retinol binding ptn)	2735114	AAB93837	134
173	Ascaris lumbricoides	Parasitic roundworm	n/a	Worm (parasite)	ABA-1 (polymorphic retinol binding ptn)	2735118	AAB93839	134
174	Ascaris lumbricoides	Parasitic roundworm	n/a	Worm (parasite)	ABA-1 (polymorphic retinol binding ptn)	2735098	AAD13645	134
175	Ascaris lumbricoides	Parasitic roundworm	n/a	Worm (parasite)	ABA-1 (polymorphic retinol binding ptn)	2735110	AAD13651	267
176	Ascaris lumbricoides	Parasitic roundworm	n/a	Worm (parasite)	ABA-1 (polymorphic retinol binding ptn)	2735108	AAD13650	267
177	Ascaris lumbricoides	Parasitic roundworm	n/a	Worm (parasite)	ABA-1 (polymorphic retinol binding ptn)	2735106	AAD13649	133
178	Ascaris lumbricoides	Parasitic roundworm	n/a	Worm (parasite)	ABA-1 (polymorphic retinol binding ptn)	2735096	AAD13644	134
179	Ascaris lumbricoides	Parasitic roundworm	n/a	Worm (parasite)	ABA-1 (polymorphic retinol binding ptn)	2735102	AAD13647	133
180	Ascaris lumbricoides	Parasitic roundworm	n/a	Worm (parasite)	ABA-1 (polymorphic retinol binding ptn)	2735112	AAD13652	267
181	Ascaris suum	Parasitic roundworm	Asc s 1	Worm (parasite)	Fatty-acid/retinoid binding protein (fragment)	299550	AAB26195	68
182	Ascaris suum	Parasitic roundworm	Asc s 1	Worm (parasite)	Fatty-acid/retinoid binding protein	159653	AAB41117	395
183	Ascaris suum	Parasitic roundworm	Asc s 1	Worm (parasite)	Fatty-acid/retinoid binding protein	2970629	AAC06015	1095
184	Aspergillus flavus	Fungus	n/a	Aero Fungi	Oryzin (serine protease)	464318	P35211	403
185	Aspergillus flavus	Fungus	n/a	Aero Fungi	Oryzin (serine protease)	7435651	JC7081	403
186	Aspergillus fumigatus	Fungus	n/a	Aero Fungi	NULL	42820661	CAF31974	1407
187	Aspergillus fumigatus	Fungus	Asp f 8	Aero Fungi	Ribosomal protein 2	6686524	CAB64688	111
188	Aspergillus fumigatus	Fungus	n/a	Aero Fungi		3219530	CAA07186	185
189	Aspergillus fumigatus	Fungus	n/a	Aero Fungi	Cellular serine proteinase (Subtilase)	2143220	CAA73782	495
190	Aspergillus fumigatus	Fungus	Asp f 1	Aero Fungi	Ribonuclease	322995	A46497	176
191	Aspergillus fumigatus	Fungus	n/a	Aero Fungi	NULL	54039254	P67875	176
192	Aspergillus fumigatus	Fungus	Asp f 16	Aero Fungi	(Similar to Asp f 9)	3643813	AAC61261	427
193	Aspergillus fumigatus	Fungus	Asp f 1	Aero Fungi	18 kDa protein (Ribonuclease)	9280360	AAF86369	150
194	Aspergillus fumigatus	Fungus	Asp f 9	Aero Fungi	Glycosyl hydrolase	2879890	CAA11266	302
195	Aspergillus fumigatus	Fungus	n/a	Aero Fungi	Pathogenesis-related protein	19309414	CAD27313	164
196	Aspergillus fumigatus	Fungus	Asp f 4	Aero Fungi		3005839	CAA04959	286
197	Aspergillus fumigatus	Fungus	Asp f 3	Aero Fungi	Peroxisomal-like protein	2769700	AAB95638	168
198	Aspergillus fumigatus	Fungus	n/a	Aero Fungi	Ribosomal protein L3	21215170	AAM43909	392
199	Aspergillus fumigatus	Fungus	Asp f 5	Aero Fungi	Metalloprotease (MEP)	3776613	CAA83015	634

200	Aspergillus fumigatus	Fungus	Asp f 7	Aero Fungi		2879888	CAA11255	112
201	Aspergillus fumigatus	Fungus	Asp f 12	Aero Fungi	Heat shock protein - 65 kDa	2851483	P40292	441
202	Aspergillus fumigatus	Fungus	Asp f 15	Aero Fungi	Protease	3005841	CAA05149	152
203	Aspergillus fumigatus	Fungus	Asp f 6	Aero Fungi	Manganese superoxide dismutase	1648970	AAB60779	221
204	Aspergillus fumigatus	Fungus	Asp f 17	Aero Fungi		2980819	CAA12162	197
205	Aspergillus fumigatus	Fungus	Asp f 2	Aero Fungi	Metalloprotease	664852	AAB07620	250
206	Aspergillus fumigatus	Fungus	Asp f 1	Aero Fungi	Ribonuclease	3021324	CAA06305	125
207	Aspergillus fumigatus	Fungus	Asp f 10	Aero Fungi	Aspergillopepsin I	963013	CAA59419	395
208	Aspergillus fumigatus	Fungus	Asp f 11	Aero Fungi	PPIase	5019414	CAB44442	178
209	Aspergillus fumigatus	Fungus	Asp f 2	Aero Fungi	Metalloprotease	2499791	P79017	310
210	Aspergillus fumigatus	Fungus	Asp f 22	Aero Fungi	Enolase	13925873	AAK49451	438
211	Aspergillus nidulans FGSC A4	Fungus	n/a	Aero Fungi	NULL	49128088	XP_412829	168
212	Aspergillus niger	Fungus	n/a	Aero Fungi	Serine protease	289172	AAA32702	533
213	Aspergillus niger	Fungus	Asp n 14	Aero Fungi	Xylosidase	2181180	CAB06417	804
214	Aspergillus niger	Fungus	Asp n 14	Aero Fungi	Beta-xylosidase	4235093	AAD13106	804
215	Aspergillus oryzae	Fungus	Asp o 13	Aero Fungi	Oryzin (Alkaline proteinase)	129235	P12547	403
216	Aspergillus oryzae	Fungus	Asp o 21	Aero Fungi	Alpha-amylase A (Taka-amylase A)	113779	P10529	499
217	Batillus cornutus	Japanese turban shell	n/a	Food Animal	Tropomyosin	7441399	JE0229	146
218	Bertholletia excelsa	Brazil nut	Ber e 1	Food Plant	2S sulfur-rich seed storage protein	112754	P04403	146
219	Bertholletia excelsa	Brazil nut	n/a	Food Plant	2S albumin large subunit	17713	CAA38363	154
220	Bertholletia excelsa	Brazil nut	Ber e 2	Food Plant	11S globulin	30313867	AAO38859	465
221	Beta vulgaris subsp. vulgaris	Beet	n/a	Food Plant	RS2 protein	11691639	CAC18641	158
222	Betula pendula	European white birch	n/a	Aero Plant	Fragment	320546	B45786	51
223	Betula pendula	European white birch	Bet v 1	Aero Plant	Isoform at42	4006955	CAA07324	160
224	Betula pendula	European white birch	Bet v 1	Aero Plant	Betv1	38492423	1LLT_A	159
225	Betula pendula	European white birch	Bet v 1	Aero Plant	Isoform at8	4006928	CAA07318	160
226	Betula pendula	European white birch	n/a	Aero Plant	NULL	4006947	CAA07320	120
227	Betula pendula	European white birch	n/a	Aero Plant	NULL	4006963	CAA07328	120
228	Betula pendula	European white birch	n/a	Aero Plant	Profilin	1942360	1CQA	133
229	Betula pendula	European white birch	Bet v 1	Aero Plant	Isoform at10	4006945	CAA07319	160
230	Betula pendula	European white birch	Bet v 1	Aero Plant	Isoform at37	4006953	CAA07323	160
231	Betula pendula	European white birch	n/a	Aero Plant	NULL	239734	AAB20452	43
232	Betula pendula	European white birch	Bet v 1	Aero Plant	Fragment	320545	A45786	51
233	Betula pendula	European white birch	Bet v 1.0301	Aero Plant	Bet v 1c	1361959	B55699	160
234	Betula pendula	European white birch	Bet v 1.1601	Aero Plant	Betv1	1321714	CAA96546	160
235	Betula pendula	European white birch	Bet v 1.1701	Aero Plant	Betv1	1321716	CAA96539	160
236	Betula pendula	European white birch	Bet v 1.1801	Aero Plant	Betv1	1321718	CAA96540	160
237	Betula pendula	European white birch	Bet v 1.1502	Aero Plant	Betv1	1321720	CAA96541	160
238	Betula pendula	European white birch	Bet v 1.1901	Aero Plant	Betv1	1321722	CAA96542	160

239	Betula pendula	European white birch	Bet v 1.2001	Aero Plant	Betv1	1321724	CAA96543	160
240	Betula pendula	European white birch	Bet v 1.2101	Aero Plant	Betv1	1321726	CAA96544	160
241	Betula pendula	European white birch	Bet v 1.2201	Aero Plant	Betv1	1321728	CAA96547	160
242	Betula pendula	European white birch	Bet v 1	Aero Plant	Betv1	11514622	1QMR_A	159
243	Betula pendula	European white birch	Bet v 6.0102	Aero Plant	Isoflavone reductase-like protein	10764491	AAG22740	308
244	Betula pendula	European white birch	Bet v 1.3001	Aero Plant	Betv1	1542873	CAB02161	160
245	Betula pendula	European white birch	Bet v 1	Aero Plant	Betv1	6980540	1B6F_A	159
246	Betula pendula	European white birch	Bet v 1.2401	Aero Plant	Betv1	1542861	CAB02155	160
247	Betula pendula	European white birch	Bet v 1	Aero Plant	Bet v 1d/h	1361960	C55699	160
248	Betula pendula	European white birch	Bet v 1.0501	Aero Plant	Bet v 1e	1361961	D55699	160
249	Betula pendula	European white birch	Bet v 1f/I	Aero Plant	Bet v 1f/i	1361962	E55699	160
250	Betula pendula	European white birch	Bet v 1.0701	Aero Plant	Bet v 1g	1361963	F55699	160
251	Betula pendula	European white birch	Bet v 1.0801	Aero Plant	Bet v 1j	1361964	G55699	160
252	Betula pendula	European white birch	Bet v 1.0901	Aero Plant	Bet v 1k	1361965	H55699	160
253	Betula pendula	European white birch	Bet v 1.1001	Aero Plant	Bet v 1l	1361966	I55699	160
254	Betula pendula	European white birch	Bet v 1m/n	Aero Plant	Bet v 1m/n	1361967	A57427	160
255	Betula pendula	European white birch	Bet v 1.0201	Aero Plant	Bet v 1b	1361968	A55699	160
256	Betula pendula	European white birch	Bet v 4	Aero Plant		7446007	JC5711	85
257	Betula pendula	European white birch	Bet v 1.2301	Aero Plant	Betv1	2414158	CAA96545	160
258	Betula pendula	European white birch	Bet v 1	Aero Plant	Betv1	4376221	CAA04828	159
259	Betula pendula	European white birch	Bet v 1x	Aero Plant	Betv1	30908931	AAP37482	21
260	Betula pendula	European white birch	Bet v 1	Aero Plant	1-Sc-1	534910	CAA54694	160
261	Betula pendula	European white birch	Bet v 1	Aero Plant	Betv1, isoform at45	4006957	CAA07325	160
262	Betula pendula	European white birch	Bet v 1	Aero Plant	Betv1, isoform at50	4006959	CAA07326	160
263	Betula pendula	European white birch	Bet v 1	Aero Plant	Betv1, isoform at59	4006961	CAA07327	160
264	Betula pendula	European white birch	Bet v 1	Aero Plant	Betv1, isoform at5	4006965	CAA07329	160
265	Betula pendula	European white birch	Bet v 1	Aero Plant	Betv1, isoform at7	4006967	CAA07330	160
266	Betula pendula	European white birch	Bet v 7	Aero Plant	Peptidylprolyl isomerase (cyclophilin)	21886603	CAC84116	173
267	Betula pendula	European white birch	Bet v 1	Aero Plant	Betv1	4376216	CAA04823	159
268	Betula pendula	European white birch	Bet v 1	Aero Plant	1-Sc-3	534898	CAA54696	160
269	Betula pendula	European white birch	Bet v 4	Aero Plant	Polcalcin	14423850	Q39419	85
270	Betula pendula	European white birch	Bet v 1	Aero Plant	Betv1	4376220	CAA04827	159
271	Betula pendula	European white birch	Bet v 1.0102	Aero Plant	Betv1	1542857	CAB02153	160
272	Betula pendula	European white birch	Bet v 1	Aero Plant	Betv1	4376222	CAA04829	159
273	Betula pendula	European white birch	Bet v 3	Aero Plant		1076247	S46233	205
274	Betula pendula	European white birch	Bet v 1.2901	Aero Plant	Betv1	1542871	CAB02160	160
275	Betula pendula	European white birch	Bet v 1.2801	Aero Plant	Betv1	1542869	CAB02159	160
276	Betula pendula	European white birch	Bet v 1.2701	Aero Plant	Betv1	1542867	CAB02158	160
277	Betula pendula	European white birch	Bet v 1.2601	Aero Plant	Betv1	1542865	CAB02157	160
278	Betula pendula	European white birch	Bet v 1 b1	Aero Plant	isoallergen bet v 1 b1	4590392	AAD26560	160
279	Betula pendula	European white birch	Bet v 1 b2	Aero Plant	isoallergen Bet v 1 b2	4590394	AAD26561	160
280	Betula pendula	European white birch	bet v 1 b3	Aero Plant	isoallergen bet v 1 b3	4590396	AAD26562	160

281	Betula pendula	European white birch	Bet v 1.2501	Aero Plant	Betv1	1542863	CAB02156	160
282	Betula pendula	European white birch	Bet v 1	Aero Plant	1-Sc-2	534900	CAA54695	159
283	Betula pendula	European white birch	Bet v 1	Aero Plant	Betv1	4376219	CAA04826	159
284	Betula pendula	European white birch	Bet v 1	Aero Plant	Betv1	2564224	CAA05188	160
285	Betula pendula	European white birch	n/a	Aero Plant	NULL	46014884	1H4B_A	84
286	Betula pendula	European white birch	Bet v 1	Aero Plant	Betv1	2564228	CAA05190	160
287	Betula pendula	European white birch	Bet v 1	Aero Plant	Betv1	2564222	CAA05187	160
288	Betula pendula	European white birch	Bet v 1	Aero Plant	Betv1	2564220	CAA05186	160
289	Betula pendula	European white birch	Bet v 2	Aero Plant	Profilin	130975	P25816	133
290	Betula platyphylla	Japanese white birch	n/a	Aero Plant	Bet vl jap2	12583683	BAB21490	160
291	Betula platyphylla	Japanese white birch	n/a	Aero Plant	Bet vl jap3	12583685	BAB21491	160
292	Betula platyphylla	Japanese white birch	n/a	Aero Plant	Bet vl jap1	12583681	BAB21489	160
293	Betula sp.	Birch	n/a	Aero Plant	NULL	298737	AAB25851	51
294	Betula sp.	Birch	n/a	Aero Plant	NULL	298736	AAB25850	51
295	Blattella germanica	German cockroach	Bla g 2	Aero Insect	Aspartic protease-like	1362590	A57164	352
296	Blattella germanica	German cockroach	n/a	Aero Insect	36 kDa protein (fragment)	544619	AAB29345	25
297	Blattella germanica	German cockroach	Bla g 1.0101	Aero Insect		4572592	AAD13530	412
298	Blattella germanica	German cockroach	n/a	Aero Insect	36 kDa protein (fragment)	544618	AAB29344	20
299	Blattella germanica	German cockroach	n/a	Aero Insect	NULL	6225491	O18598	204
300	Blattella germanica	German cockroach	Bla g 4	Aero Insect	Calycin (lipid binding protein)	1166573	AAA87851	182
301	Blattella germanica	German cockroach	n/a	Aero Insect	NULL	8101069	AAF72534	284
302	Blattella germanica	German cockroach	n/a	Aero Insect	NULL	62738637	1YG9_A	330
303	Blattella germanica	German cockroach	Bla g 1.02	Aero Insect		4240395	AAD13531	492
304	Blomia tropicalis	Mite	n/a	Aero Mite	Trypsin	25989482	AAM10779	266
305	Blomia tropicalis	Mite	Blo t 12	Aero Mite		902012	AAA78904	144
306	Blomia tropicalis	Mite	n/a	Aero Mite		25989484	AAM10780	338
307	Blomia tropicalis	Mite	n/a	Aero Mite	(Similar to Blo t 5)	915347	AAB49396	73
308	Blomia tropicalis	Mite	Blo t 11	Aero Mite	Paramyosin	21954740	AAM83103	875
309	Blomia tropicalis	Mite	Blo t 5	Aero Mite		4204917	AAD10850	134
310	Blomia tropicalis	Mite	Blo t 1	Aero Mite	Cysteine protease	14276828	AAK58415	221
311	Blomia tropicalis	Mite	Blo t 13	Aero Mite	Fatty-acid binding protein	1377859	AAC80579	130
312	Bombus terrestris	Bumblebee	n/a	Venom or Salivary	Phospholipase A2	14423832	P82971	136
313	Bos taurus	Bovine	n/a	Food Animal	Alpha-s1 casein	162927	AAA30478	76
314	Bos taurus	Bovine	Bos d 2.0103	Aero Animal	Lipocalin-like	11277082	B59225	156
315	Bos taurus	Bovine	n/a	Food Animal	Beta-casein A3	459292	AAB29137	224
316	Bos taurus	Bovine	Bos d 5	Food Animal	Beta-lactoglobulin	520	CAA32835	178
317	Bos taurus	Bovine	n/a	Food Animal	Beta-casein	162931	AAA30480	224
318	Bos taurus	Bovine	n/a	Food Animal	Alpha-s2-like casein	162929	AAA30479	222
319	Bos taurus	Bovine	Bos d 2.0102	Aero Animal	Lipocalin-like	11277083	A59225	156
320	Bos taurus	Bovine	n/a	Food Animal	Kappa-casein	162811	AAA30433	190
321	Bos taurus	Bovine	n/a	Food Animal	Epididymal secretory protein (19.5kD)	27806881	NP_776343	149
322	Bos taurus	Bovine	n/a	Food Animal	MD-2 protein	27806539	NP_776536	160

323	Bos taurus	Bovine	n/a	Food Animal	Beta-casein	162805	AAA30431	224
324	Bos taurus	Bovine	n/a	Food Animal	Alpha-s1 casein	162792	AAA30428	214
325	Bos taurus	Bovine	n/a	Food Animal	Alpha-s1 casein	162794	AAA30429	214
326	Bos taurus	Bovine	Bos d 4	Food Animal	Alpha-lactalbumin	162644	AAA30367	142
327	Bos taurus	Bovine	n/a	Food Animal	Beta-lactoglobulin	162750	AAA30413	14
328	Bos taurus	Bovine	n/a	Food Animal	Beta-lactoglobulin	162748	AAA30411	151
329	Bos taurus	Bovine	n/a	Food Animal	Alpha-s1 casein	162650	AAA62707	93
330	Bos taurus	Bovine	Bos d 6	Food Animal	Serum albumin (Bos t 6)	162648	AAA51411	607
331	Bos taurus	Bovine	n/a	Food Animal	Beta-lactoglobulin variant B	669061	CAA88303	178
332	Bos taurus	Bovine	Bos d 2	Aero Animal	Lipocalin-like, dander (BDA20)	886215	AAB08720	172
333	Bos taurus	Bovine	n/a	Food Animal	Kappa-casein	1228078	CAA33034	190
334	Bos taurus	Bovine	n/a	Food Animal	Beta-casein	162797	AAA30430	224
335	Bos taurus	Bovine	n/a	Aero Animal	NULL	61878171	XP_596033	93
336	Bos taurus	Bovine	n/a	Food Animal	Serum albumin	3336842	CAA76847	607
337	Bos taurus	Bovine	n/a	Aero Animal	NULL	61839003	XP_599492	116
338	Bos taurus	Bovine	n/a	Aero Animal	NULL	61858238	XP_606720	155
339	Bos taurus	Bovine	Bos d 4	Food Animal	Alpha-lactalbumin (Bos t 4)	295774	CAA29664	142
340	Bos taurus	Bovine	n/a	Unassigned	NULL	126796	P02687	169
341	Bos taurus	Bovine	Bos d 3	Aero Animal	S100 calcium-binding protein A7	2493414	Q28050	101
342	Bos taurus	Bovine	n/a	Aero Animal	NULL	61869494	XP_597522	133
343	Bos taurus	Bovine	n/a	Aero Animal	NULL	61819907	XP_594082	171
344	Brassica juncea	Mustard	Bra j 1	Food Plant		407610	AAB27813	92
345	Brassica juncea	Mustard	Bra j 1	Food Plant		407609	AAB27812	37
346	Brassica napus	Rape	n/a	Food Plant	2S albumin (pronapin)	26985163	AAN86249	109
347	Brassica napus	Rape	n/a	Aero Plant	Polcalcin Group II (Bet v 4 homologue)	2129801	S65144	83
348	Brassica napus	Rape	n/a	Aero Plant	Polcalcin Group II (Bet v 4 homologue)	2129802	S65145	83
349	Brassica napus	Rape	Bra n I	Aero Plant	Calcim binding protein Group I	2129799	S65149	79
350	Brassica napus	Rape	n/a	Aero Plant	NULL	59800145	P69198	83
351	Brassica rapa	Turnip	n/a	Aero Plant	Probable pollen coat protein	7488531	T14380	83
352	Brassica rapa	Turnip	Bra r 2	Aero Plant	Polcalcin Group II (Bet v 4 homologue)	2129805	S65143	80
353	Brassica rapa	Turnip	n/a	Aero Plant	Lipid transfer protein pollen	3062791	BAA25680	86
354	Brassica rapa subsp. rapa	Turnip	n/a	Aero Plant	NULL	11135129	O64432	123
355	Brassica rapa subsp. rapa	Turnip	Bra r 2	Food Plant	Chitin binding protein	32363457	P81729_1	33
356	Brassica rapa subsp. rapa	Turnip	Bra r 2	Food Plant	Chitin binding protein	32363458	P81729_2	34
357	Brassica rapa subsp.	Turnip	Bra r 2	Food Plant	Chitin binding protein	32363459	P81729_3	13

	rapa							
358	Brassica rapa subsp. rapa	Turnip	n/a	Aero Plant	NULL	59800146	P69199	83
359	Brassica rapa subsp. rapa	Turnip	Bra r 2	Food Plant	Chitin binding protein	32363461	P81729_5	8
360	Brassica rapa subsp. rapa	Turnip	n/a	Aero Plant	NULL	59800144	P69197	79
361	Brugia malayi	Parasitic nematode	n/a	Worm (parasite)	Filarial antigen	13310414	AAK18279	227
362	Brugia malayi	Parasitic nematode	n/a	Worm (parasite)	Larval allergen	24711753	AAN62757	248
363	Brugia malayi	Parasitic nematode	n/a	Worm (parasite)	Venom like protein	13095442	AAK12274	220
364	Candida albicans	Yeast	Cand a 1	Contact	Alcohol dehydrogenase	1168348	P43067	350
365	Candida albicans	Yeast	n/a	Contact	Fructose-bisphosphate aldolase (fragment)	18203509	Q9URB4	40
366	Candida albicans	Yeast	n/a	Contact	Enolase 1	232054	P30575	440
367	Candida albicans	Yeast	Cand a 3	Contact	NULL	37548637	AAN11300	236
368	Candida albicans	Yeast	n/a	Contact	Putative aminoacid permease	24461764	AAN62330	648
369	Candida albicans SC5314	Yeast	n/a	Contact	NULL	46431509	EAK91061	341
370	Candida albicans SC5314	Yeast	n/a	Contact	NULL	46431500	EAK91053	341
371	Candida boidinii	Yeast	Cand b 2	Contact	Peroxisomal membrane protein B	130361	P14293	167
372	Candida boidinii	Yeast	Cand b 2	Contact	Peroxisomal membrane protein A	130360	P14292	167
373	Canis familiaris	Dog	Can f 2	Aero Animal	Lipocalin	29292272	CAD82911	177
374	Canis familiaris	Dog	Can f 2	Aero Animal	Lipocalin	29292274	CAD82912	179
375	Canis familiaris	Dog	Can f 4	Aero Animal	NULL	60729635	A59491	13
376	Canis familiaris	Dog	Can f 3	Aero Animal	Albumin	3319897	CAA76841	585
377	Canis familiaris	Dog	Can f 1	Aero Animal	Lipocalin	3121745	O18873	174
378	Canis familiaris	Dog	Can f 2	Aero Animal	Lipocalin	3121746	O18874	180
379	Canis familiaris	Dog	Can f 3	Aero Animal	Albumin	6687188	CAB64867	608
380	Canis familiaris	Dog	Can f 3	Aero Animal	Albumin	633938	AAB30434	265
381	Capsicum annuum	Bell pepper	Cap a 2	Food Plant	Profilin	16555785	CAD10376	131
382	Capsicum annuum	Bell pepper	Cap a 1	Food Plant	Osmotin-like protein	16609959	CAC34055	246
383	Capsicum annuum	Bell pepper	n/a	Food Plant	Profilin (Cold induced protein)	40287498	AAR83863	18
384	Carica papaya	Papaya	n/a	Food Plant	Papain	167391	AAB02650	345
385	Carpinus betulus	Hornbeam	Car b 1.0301	Aero Plant	PRP (Bet v 1 family)	1545895	CAB02216	161
386	Carpinus betulus	Hornbeam	Car b 1.0105	Aero Plant	PRP (Bet v 1 family)	1545887	CAB02212	160
387	Carpinus betulus	Hornbeam	Car b 1	Aero Plant	PRP (Bet v 1 family)	402745	CAA47357	159
388	Carpinus betulus	Hornbeam	Car b 1	Aero Plant	PRP (Bet v 1 family)	730048	P38949	160
389	Carpinus betulus	Hornbeam	n/a	Aero Plant	NULL	239735	AAB20453	40
390	Carpinus betulus	Hornbeam	Car b 1.0103	Aero Plant	PRP (Bet v 1 family)	1545875	CAB02206	160
391	Carpinus betulus	Hornbeam	Car b 1	Aero Plant	PRP (Bet v 1 family)	730049	P38950	160
392	Carpinus betulus	Hornbeam	Car b 1.0104	Aero Plant	PRP (Bet v 1 family)	1545879	CAB02208	160

393	Carpinus betulus	Hornbeam	Car b 1.0302	Aero Plant	PRP (Bet v 1 family)	1545897	CAB02217	161
394	Carpinus betulus	Hornbeam	Car b 1.0104	Aero Plant	PRP (Bet v 1 family)	1545877	CAB02207	160
395	Carpinus betulus	Hornbeam	Car b 1	Aero Plant	PRP (Bet v 1 family)	1545891	CAB02214	160
396	Carpinus betulus	Hornbeam	Car b 1.0108	Aero Plant	PRP (Bet v 1 family)	1545893	CAB02215	160
397	Carya illinoinensis	Pecan	n/a	Food Plant	Seed storage protein	28207731	AAO32314	143
398	Castanea sativa	European chestnut	Cas s 1	Aero Plant	Pathogenesis-related protein family 10	16555781	CAD10374	160
399	Castanea sativa	European chestnut	Cas s 5	Food Plant	Chitinase Ib	1359600	CAA64868	316
400	Castanea sativa	European chestnut	n/a	Aero Plant	Fragment	543622	PC2001	24
401	Catharanthus roseus	Madagascar periwinkle	n/a	Aero Plant	Cytokinin-induced protein (fragment)	7442208	T10059	140
402	Cavia porcellus	Domestic guinea pig	Cav p 2	Unassigned	NULL	32363133	P83508	15
403	Cavia porcellus	Domestic guinea pig	n/a	Unassigned	NULL	544241	P35709	234
404	Cavia porcellus	Domestic guinea pig	n/a	Unassigned	NULL	119238	P22032	233
405	Cavia porcellus	Domestic guinea pig	Cav p 1	Unassigned	NULL	32469617	P83507	15
406	Chamaecyparis obtusa	Japanese cypress	n/a	Aero Plant		1514943	BAA08246	375
407	Chamaecyparis obtusa	Japanese cypress	n/a	Aero Plant	Polygalacturonase	7520823	JC7100	514
408	Charybdis feriatus	Crab	n/a	Food Animal	Tropomyosin	7024506	AAF35431	264
409	Chenopodium album	Pigweed	Che a 2	Aero Plant	profilin	29465666	AAL92870	131
410	Chenopodium album	Pigweed	Che a 3	Aero Plant	Polcalcin	29465668	AAL92871	86
411	Chenopodium album	Pigweed	Che a 1	Aero Plant		22074346	AAL07319	168
412	Chironomus kiiensis	Midge	n/a	Aero Insect	Tropomyosin	7321108	CAA09938	285
413	Chironomus thummi thummi	Midge	Chi t 7	Aero Insect	Globin CTT-VIIB-4	121246	P12292	161
414	Chironomus thummi thummi	Midge	Chi t 7	Aero Insect	Globin CTT-VIIB-3	121244	P12548	161
415	Chironomus thummi thummi	Midge	Chi t 8	Aero Insect	Globin CTT-VIII	121237	P02227	151
416	Chironomus thummi thummi	Midge	Chi t 6.01	Aero Insect	Globin CTT-VIIA	121236	P02226	145
417	Chironomus thummi thummi	Midge	Chi t 1.02	Aero Insect	Globin CTT-IV	121227	P02230	151
418	Chironomus thummi thummi	Midge	Chi t 9	Aero Insect	Globin CTT-X	121259	P02228	151
419	Chironomus thummi thummi	Midge	Chi t 7	Aero Insect	Globin CTT-VIIB-6	121248	P12549	161
420	Chironomus thummi thummi	Midge	Chi t 1.01	Aero Insect	Globin CTT-III (Erythrocruorin III)	121219	P02229	151
421	Chironomus thummi thummi	Midge	Chi t 4	Aero Insect	Globin CTT-IIIA	121256	P02231	151
422	Chironomus thummi thummi	Midge	Chi t 7	Aero Insect	Globin CTT-VIIB-5/CTT-VIIB-9	121247	P11583	161
423	Chironomus thummi thummi	Midge	Chi t 7	Aero Insect	Globin CTT-VIIB-7	121249	P12550	162
424	Chironomus thummi	Midge	Chi t 3	Aero Insect	Globin CTT-II beta	1707908	P02222	160

	thummi							
425	Chironomus thummi thummi	Midge	Chi t 5	Aero Insect	Globin CTT-VI	2506461	P02224	162
426	Chironomus thummi thummi	Midge	Chi t 6.02	Aero Insect	Globin CTT-IX	1707911	P02223	161
427	Chironomus thummi thummi	Midge	Chi t 2	Aero Insect	Globin CTT-I/CTT-IA (Erythrocruorin)	2506460	P02221	158
428	Citrus limon	Lemon	n/a	Unassigned	NULL	52783176	P84160	20
429	Citrus sinensis	Navel orange	n/a	Unassigned	NULL	54036219	P84177	10
430	Citrus sinensis	Navel orange	n/a	Unassigned	NULL	52783177	P84161	20
431	Citrus sinensis	Navel orange	n/a	Unassigned	NULL	52782810	P84159	25
432	Citrus sinensis	Navel orange	n/a	Unassigned	NULL	56000996	CAI23765	131
433	Citrus sinensis	Navel orange	Cit s 3	Unassigned	NULL	50199132	CAH03799	91
434	Coprinus comatus	Shaggy mane	Cop c 7	Food Fungi	(from cap and gills)	5689675	CAB52133	152
435	Coprinus comatus	Shaggy mane	Cop c 5	Food Fungi	(from cap and gills)	5689673	CAB52132	141
436	Coprinus comatus	Shaggy mane	Cop c 2	Food Fungi	Thioredoxin (from cap and gills)	5689669	CAB52130	106
437	Coprinus comatus	Shaggy mane	Cop c 1	Food Fungi	(from cap and gills)	4538529	CAB39376	81
438	Coprinus comatus	Shaggy mane	Cop c 3	Food Fungi	(from cap and gills)	5689671	CAB52131	328
439	Corylus avellana	European hazelnut	Cor a 1.0402	Food Plant	Cor a 1 (Bet v 1 homologue)	11762102	AAG40329	161
440	Corylus avellana	European hazelnut	Cor a 8	Food Plant	Lipid transfer protein	13507262	AAK28533	115
441	Corylus avellana	European hazelnut	Cor a 1.0404	Food Plant	Cor a 1 (Bet v 1 homologue)	11762106	AAG40331	161
442	Corylus avellana	European hazelnut	Cor a 1.0201	Aero Plant	Cor a 1 (Bet v 1 homologue)	1321731	CAA96548	160
443	Corylus avellana	European hazelnut	Cor a 9	Food Plant	11S globulin-like protein	18479082	AAL73404	515
444	Corylus avellana	European hazelnut	Cor a 2	Aero Plant	Profilin	12659208	AAK01236	131
445	Corylus avellana	European hazelnut	Cor a 10	Aero Plant	Putative luminal binding protein	10944737	CAC14168	668
446	Corylus avellana	European hazelnut	Cor a 1.0401	Food Plant	Cor a 1 (Bet v 1 homologue)	5726304	AAD48405	161
447	Corylus avellana	European hazelnut	n/a	Aero Plant	Putative LEA III protein isoform 1	14148979	CAC39160	109
448	Corylus avellana	European hazelnut	Cor a 1.0403	Food Plant	Cor a 1 (Bet v 1 homologue)	11762104	AAG40330	161
449	Corylus avellana	European hazelnut	n/a	Food Plant	Oleosin	29170509	AAO65960	140
450	Corylus avellana	European hazelnut	Cor a 1.0301	Aero Plant	Cor a 1 (Bet v 1 homologue)	1321733	CAA96549	160
451	Corylus avellana	European hazelnut	n/a	Aero Plant	Putative LEA III protein isoform 2	14148981	CAC39110	87
452	Corylus avellana	European hazelnut	Cor a 2	Aero Plant	Profilin	12659206	AAK01235	131
453	Corylus avellana	European hazelnut	Cor a I	Aero Plant	Cor a 1 (Bet v 1 homologue)	421788	S30055	160
454	Corylus avellana	European hazelnut	Cor a 1.0102	Aero Plant	Cor a 1 (Bet v 1 homologue)	22690	CAA50328	160
455	Corylus avellana	European hazelnut	Cor a 1.0101	Aero Plant	Cor a 1 (Bet v 1 homologue)	22688	CAA50327	160
456	Corylus avellana	European hazelnut	Cor a 1.0103	Aero Plant	Cor a 1 (Bet v 1 homologue)	22684	CAA50325	160
457	Corylus avellana	European hazelnut	Cor a 1.0104	Aero Plant	Cor a 1 (Bet v 1 homologue)	22686	CAA50326	160
458	Corylus avellana	European hazelnut	Cor a 11	Food Plant	48 kDa glycoprotein (vicilin like)	19338630	AAL86739	448
459	Crassostrea gigas	American oyster	n/a	Food Animal	Tropomyosin	15419048	AAK96889	233

460	Crassostrea virginica	Eastern oyster	n/a	Food Animal	Tropomyosin	3668408	AAC61869	160
461	Crocus sativus	Saffron crocus	n/a	Aero Plant	NULL	58700651	AAW81034	131
462	Crotalus durissus terrificus	Rattlesnake	n/a	Venom or Salivary	Phospholipase A2 inhibitor	501050	AAA19162	200
463	Cryptococcus neoformans var. neoformans JEC21	Fungus (yeast)	n/a	Aero Fungi	NULL	58265022	XP_569667	104
464	Cryptococcus neoformans var. neoformans JEC21	Fungus (yeast)	n/a	Aero Fungi	NULL	57227639	AAW44097	234
465	Cryptomeria japonica	Japanese cedar	n/a	Aero Plant	NULL	22830595	BAC15615	233
466	Cryptomeria japonica	Japanese cedar	n/a	Aero Plant	NULL	22830593	BAC15614	232
467	Cryptomeria japonica	Japanese cedar	Cry j 2	Aero Plant	Polygalacturonase	24898904	BAC23082	514
468	Cryptomeria japonica	Japanese cedar	n/a	Aero Plant	Isoflavone reductase-like protein CJP-6	19847822	AAK27264	306
469	Cryptomeria japonica	Japanese cedar	n/a	Aero Plant	NULL	22830597	BAC15616	230
470	Cryptomeria japonica	Japanese cedar	Cry j 1	Aero Plant	Pectate lyase	541802	JC2124	374
471	Cryptomeria japonica	Japanese cedar	Cry j 1	Aero Plant	Pectate lyase	19570315	BAB86286	374
472	Cryptomeria japonica	Japanese cedar	n/a	Aero Plant	NULL	56550550	BAD77932	281
473	Cryptomeria japonica	Japanese cedar	Cry j 2	Aero Plant	Polygalacturonase	1076242	JC2498	514
474	Cryptomeria japonica	Japanese cedar	Cry j 2	Aero Plant	Polygalacturonase	24898906	BAC23083	514
475	Cryptomeria japonica	Japanese cedar	Cry j 2	Aero Plant	Polygalacturonase	1171004	P43212	514
476	Cryptomeria japonica	Japanese cedar	n/a	Aero Plant	NULL	62149370	BAD93485	290
477	Cryptomeria japonica	Japanese cedar	n/a	Aero Plant	Pectate lyase	280372	A44773	20
478	Cryptomeria japonica	Japanese cedar	n/a	Aero Plant	NULL	62149368	BAD93484	278
479	Cryptomeria japonica	Japanese cedar	Cry j 2	Aero Plant	Polygalacturonase	24898908	BAC23084	514
480	Cryptomeria japonica	Japanese cedar	Cry j 1	Aero Plant	Pectate lyase	541803	JC2123	374
481	Cryptomeria japonica	Japanese cedar	n/a	Aero Plant	NULL	62149372	BAD93486	348
482	Ctenocephalides felis	Cat flea	n/a	Venom or Salivary	FS-H precursor	1575479	AAB09540	98
483	Ctenocephalides felis	Cat flea	Cte f 2	Venom or Salivary	Salivary antigen 2	7638032	AAF65314	264
484	Ctenocephalides felis	Cat flea	Cte f 1	Venom or Salivary	Salivary antigen 1	4336703	AAD17905	176
485	Cucumis melo	Muskmelon	Cuc m 3	Food Plant	NULL	46396596	P83834_1	21
486	Cucumis melo	Muskmelon	Cuc m 2	Food Plant	Profilin	31559374	CAD92666	131
487	Cucumis melo	Muskmelon	Cuc m 1	Food Plant	Pre-pro-cucumisin	807698	BAA06905	731
488	Cucumis melo	Muskmelon	Cuc m 3	Food Plant	NULL	46396597	P83834_2	10
489	Cucumis melo	Muskmelon	Cuc m 2	Food Plant	NULL	58263793	AAW69549	131
490	Cucumis melo	Muskmelon	Cuc m 3	Food Plant	NULL	46396598	P83834_3	10
491	Cucumis melo var. reticulatus	Netted muskmelon	Cuc m 2	Food Plant	Profilin	30142164	AAP13533	131
492	Cucumis melo var. reticulatus	Netted muskmelon	Cuc m 2	Food Plant	NULL	57021110	AAP13533	131

493	Cupressus arizonica	Arizona Cypress	n/a	Aero Plant	NULL	9929163	CAC05258	199
494	Cupressus arizonica	Arizona Cypress	Cup a 1	Aero Plant		19069497	CAC37790	367
495	Cupressus arizonica	Arizona Cypress	Cup a 1	Aero Plant		6562326	CAB62551	346
496	Cupressus sempervirens	Mediterranean Cypress	n/a	Aero Plant	NULL	38456230	AAR21075	225
497	Cupressus sempervirens	Mediterranean Cypress	Cup s 1.0105	Aero Plant		8101719	AAF72629	367
498	Cupressus sempervirens	Mediterranean Cypress	n/a	Aero Plant	NULL	38456228	AAR21074	225
499	Cupressus sempervirens	Mediterranean Cypress	Cup s 1.0104	Aero Plant		8101717	AAF72628	367
500	Cupressus sempervirens	Mediterranean Cypress	Cup s 1.0103	Aero Plant		8101715	AAF72627	367
501	Cupressus sempervirens	Mediterranean Cypress	Cup s 1.0102	Aero Plant		8101713	AAF72626	367
502	Cupressus sempervirens	Mediterranean Cypress	Cup s 1.0101	Aero Plant		8101711	AAF72625	367
503	Cynodon dactylon	Bermuda grass	Cyn d 1	Aero Plant	Group 1 grass	543623	A61226	27
504	Cynodon dactylon	Bermuda grass	Cyn d 23	Aero Plant		32344779	AAP80170	118
505	Cynodon dactylon	Bermuda grass	Cyn d 1	Aero Plant	Group 1 grass	14423757	O04701	246
506	Cynodon dactylon	Bermuda grass	Cyn d 15	Aero Plant		32344781	AAP80171	112
507	Cynodon dactylon	Bermuda grass	n/a	Aero Plant	B1 protein	1247373	CAA01909	71
508	Cynodon dactylon	Bermuda grass	n/a	Aero Plant	B4 protein	1247375	CAA01910	73
509	Cynodon dactylon	Bermuda grass	n/a	Aero Plant	B2 protein	1247377	CAA01911	247
510	Cynodon dactylon	Bermuda grass	Cyn d 1.0202	Aero Plant	Group 1 grass	16076693	AAL14077	262
511	Cynodon dactylon	Bermuda grass	n/a	Aero Plant	NULL	51950706	AAU15051	153
512	Cynodon dactylon	Bermuda grass	Cyn d 12	Aero Plant	Profilin 1	2154730	CAA69670	131
513	Cynodon dactylon	Bermuda grass	n/a	Aero Plant	2-phosphoglycerate dehydrogenase 2	37222051	AAQ17040	446
514	Cynodon dactylon	Bermuda grass	n/a	Aero Plant	Group 2 grass	4006978	CAA10346	122
515	Cynodon dactylon	Bermuda grass	Cyn d 1.0203	Aero Plant	Group 1 grass	16076697	AAL14079	262
516	Cynodon dactylon	Bermuda grass	Cyn d 1.0201	Aero Plant	Group 1 grass	15384338	AAK96255	244
517	Cynodon dactylon	Bermuda grass	Cyn d 1	Aero Plant	Group 1 grass	451274	AAB28566	25
518	Cynodon dactylon	Bermuda grass	Cyn d 1	Aero Plant	Group 1 grass	691726	AAB32317	34
519	Cynodon dactylon	Bermuda grass	Cyn d 7	Aero Plant	Calcium-binding protein	1871507	CAA62634	82
520	Cynodon dactylon	Bermuda grass	Cyn d 1	Aero Plant	Group 1 grass	451275	AAB28567	38
521	Cynodon dactylon	Bermuda grass	n/a	Aero Plant	NULL	41393750	AAS02108	522
522	Cynodon dactylon	Bermuda grass	Cyn d 1.0204	Aero Plant	Group 1 grass	10314021	AAF80379	244
523	Cynodon dactylon	Bermuda grass	Cyn d 1	Aero Plant	Group 1 grass	16076695	AAL14078	262
524	Cyprinus carpio	Carp	n/a	Food Animal	Parvalbumin	17977825	CAC83658	109
525	Cyprinus carpio	Carp	n/a	Food Animal	Parvalbumin	17977827	CAC83659	109
526	Dactylis glomerata	Orchard grass	Dac g 3	Aero Plant	DG3	280409	A60359	28
527	Dactylis glomerata	Orchard grass	Dac g 2	Aero Plant	Group 2 grass	4007040	CAA10345	122
528	Dactylis glomerata	Orchard grass	n/a	Aero Plant	NULL	32363464	P82946_1	12

529	Dactylis glomerata	Orchard grass	Dac g 1	Aero Plant	Group 1 grass	18093991	CAD20406	264
530	Dactylis glomerata	Orchard grass	Dac g 5	Aero Plant	Group 5 grass	18093971	CAD20405	265
531	Dactylis glomerata	Orchard grass	n/a	Aero Plant	NULL	33149333	AAP96759	240
532	Dactylis glomerata	Orchard grass	n/a	Aero Plant	NULL	32363466	P82946_3	17
533	Dactylis glomerata	Orchard grass	Dac g 3	Aero Plant	Group 3 grass	1825459	AAB42200	96
534	Dactylis glomerata	Orchard grass	n/a	Aero Plant	NULL	32363467	P82946_4	15
535	Dactylis glomerata	Orchard grass	Dac g 2	Aero Plant	Group 2 grass	1093120	2103117A	196
536	Dactylis glomerata	Orchard grass	Dac g 5	Aero Plant	Group 5 grass	14423124	AAK62278	290
537	Dactylis glomerata	Orchard grass	n/a	Aero Plant	NULL	32363465	P82946_2	11
538	Daucus carota	Carrot	Dau c 1.0101	Food Plant	PRP (Bet v 1 family)	1335877	AAB01092	168
539	Daucus carota	Carrot	Dau c 1.0105	Food Plant	PRP (Bet v 1 family)	2154736	CAB06416	154
540	Daucus carota	Carrot	Dau c 1.0103	Food Plant	PRP (Bet v 1 family)	2154732	CAB03715	154
541	Daucus carota	Carrot	Dau c 1.0201	Food Plant	PRP (Bet v 1 family)	18652047	AAL76932	154
542	Daucus carota	Carrot	Dau c 4	Food Plant	Profilin	18652049	AAL76933	134
543	Daucus carota	Carrot	n/a	Food Plant	PRP (Bet v 1 family)	19912791	BAB88129	154
544	Daucus carota	Carrot	Dau c 1.0102	Food Plant	PRP (Bet v 1 family)	1663522	BAA13604	154
545	Daucus carota	Carrot	Dau c 1.0104	Food Plant	PRP (Bet v 1 family)	2154734	CAB03716	154
546	Davidiella tassiana	Fungus	Cla h 6	Aero Fungi	Enolase	467660	CAA55070	440
547	Davidiella tassiana	Fungus	Cla h 10	Aero Fungi	Aldehyde dehydrogenase	467625	CAA55072	496
548	Davidiella tassiana	Fungus	Cla h 12	Aero Fungi	Ribosomal protein P1	1143425	CAA59463	110
549	Davidiella tassiana	Fungus	Cla h 5	Aero Fungi	60S acidic ribosomal protein P2	1173074	P42038	111
550	Davidiella tassiana	Fungus	Cla h 7	Aero Fungi		1168970	P42059	204
551	Davidiella tassiana	Fungus	n/a	Aero Fungi	Hydrophobin	22796153	CAD42710	105
552	Davidiella tassiana	Fungus	n/a	Aero Fungi	Putative nuclear transport factor 2	21748151	CAD38166	125
553	Davidiella tassiana	Fungus	n/a	Aero Fungi	Heat shock 70 kDa protein	729764	P40918	643
554	Davidiella tassiana	Fungus	Cla h 5	Aero Fungi	60S acidic ribosomal protein P2	21542440	P42039	111
555	Davidiella tassiana	Fungus	Cla h 6	Aero Fungi	Enolase	6015094	P42040	440
556	Dermatophagoides farinae	House dust mite	n/a	Aero Mite	Peptidase S1-like protein	14424450	P49276	279
557	Dermatophagoides farinae	House dust mite	Der f 2	Aero Mite		17978844	AAL47677	129
558	Dermatophagoides farinae	House dust mite	Der f 2	Aero Mite		546852	AAB30829	142
559	Dermatophagoides farinae	House dust mite	Der f 3	Aero Mite	Peptidase S1-like protein	2507248	P49275	259
560	Dermatophagoides farinae	House dust mite	Der f 2	Aero Mite		217308	BAA01241	138
561	Dermatophagoides farinae	House dust mite	Der f 2	Aero Mite		217306	BAA01240	146
562	Dermatophagoides farinae	House dust mite	n/a	Aero Mite	NULL	259012	AAB23973	20

563	Dermatophagoides farinae	House dust mite	n/a	Aero Mite	NULL	404371	AAB27594	20
564	Dermatophagoides farinae	House dust mite	Der f 10	Aero Mite	Mag44	1359436	BAA04557	299
565	Dermatophagoides farinae	House dust mite	Der f 7	Aero Mite	Mite Group 7 protein	2498299	Q26456	213
566	Dermatophagoides farinae	House dust mite	Der f 1	Aero Mite	Cysteine protease	27530349	BAC53948	321
567	Dermatophagoides farinae	House dust mite	n/a	Aero Mite	60 kDa protein	27550039	AAM19082	462
568	Dermatophagoides farinae	House dust mite	n/a	Aero Mite	NULL	55859466	CAI05848	146
569	Dermatophagoides farinae	House dust mite	Der f 1	Aero Mite		627141	A61500	319
570	Dermatophagoides farinae	House dust mite	Der f 2	Aero Mite	Fragment	627142	A61501	129
571	Dermatophagoides farinae	House dust mite	Der f 2	Aero Mite		13560629	AAK30133	170
572	Dermatophagoides farinae	House dust mite	n/a	Aero Mite	NULL	55859470	CAI05850	146
573	Dermatophagoides farinae	House dust mite	n/a	Aero Mite	Mag29	729970	P39674	145
574	Dermatophagoides farinae	House dust mite	n/a	Aero Mite	Mag	729979	P39673	341
575	Dermatophagoides farinae	House dust mite	Der f 1	Aero Mite	Cysteine protease	730035	P16311	321
576	Dermatophagoides farinae	House dust mite	Der f 3	Aero Mite	Trypsin-like protein	1314736	AAA99805	232
577	Dermatophagoides farinae	House dust mite	Der f 15	Aero Mite	98 kDa HDM	5815436	AAD52672	555
578	Dermatophagoides farinae	House dust mite	n/a	Aero Mite	NULL	55859468	CAI05849	146
579	Dermatophagoides farinae	House dust mite	Der f 16	Aero Mite	Gelsolin-like protein	21591547	AAM64112	480
580	Dermatophagoides farinae	House dust mite	Der f 14	Aero Mite	Mag3	1545803	BAA04558	349
581	Dermatophagoides farinae	House dust mite	Der f 11	Aero Mite	Paramyosin-like protein	13785807	AAK39511	692
582	Dermatophagoides farinae	House dust mite	n/a	Aero Mite	NULL	56378069	BAD74060	146
583	Dermatophagoides microceras	House dust mite	Der m 1	Aero Mite	Fragment	84699	B27634	30
584	Dermatophagoides pteronyssinus	House dust mite	n/a	Aero Mite	NULL	15072346	AAG02250	180
585	Dermatophagoides	House dust mite	Der p 4	Aero Mite	Alpha-amylase	5059162	AAD38942	496

	pteronyssinus							
586	Dermatophagoides pteronyssinus	House dust mite	Der p 10	Aero Mite	Tropomyosin	2440053	CAA75141	284
587	Dermatophagoides pteronyssinus	House dust mite	n/a	Aero Mite	NULL	3891991	1A9V	129
588	Dermatophagoides pteronyssinus	House dust mite	Der p 11	Aero Mite	NULL	37778944	AAO73464	875
589	Dermatophagoides pteronyssinus	House dust mite	Der p 6	Aero Mite	Peptidase S1-like protein	1352239	P49277	20
590	Dermatophagoides pteronyssinus	House dust mite	n/a	Aero Mite	Fragment (Der p 5)	102834	B31657	23
591	Dermatophagoides pteronyssinus	House dust mite	Der p 7	Aero Mite		10189811	CAC09234	215
592	Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite	Cysteine protease-like protein	84700	JQ0337	245
593	Dermatophagoides pteronyssinus	House dust mite	Der p 5	Aero Mite	Fragment	84702	S06734	148
594	Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite	Cysteine protease-like protein	21725572	CAD38367	222
595	Dermatophagoides pteronyssinus	House dust mite	Der p 5	Aero Mite	Fragment (Der p 5)	9072	CAA35692	148
596	Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite	Mite Group 2 protein	1352237	P49278	146
597	Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite		21725582	CAD38372	129
598	Dermatophagoides pteronyssinus	House dust mite	Der p 5	Aero Mite		913285	AAB32842	132
599	Dermatophagoides pteronyssinus	House dust mite	n/a	Aero Mite	NULL	102832	A39997	18
600	Dermatophagoides pteronyssinus	House dust mite	Der p 5	Aero Mite		28798085	CAD69036	132
601	Dermatophagoides pteronyssinus	House dust mite	n/a	Aero Mite	Collagenolytic serine protease	22595342	AAN02510	244
602	Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite		21725568	CAD38365	222
603	Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite		21725560	CAD38361	222
604	Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite		21725562	CAD38362	222
605	Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite		21725564	CAD38363	222
606	Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite		21725566	CAD38364	222
607	Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite	Cysteine protease-like protein	21725570	CAD38366	222

608	Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite		21725574	CAD38368	222
609	Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite		21725576	CAD38369	222
610	Dermatophagoides pteronyssinus	House dust mite	n/a	Aero Mite	NULL	61608445	AAX47076	216
611	Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite		21725580	CAD38371	222
612	Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite		29786835	CAD88532	302
613	Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite		21725584	CAD38373	129
614	Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite		21725586	CAD38374	129
615	Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite		21725588	CAD38375	129
616	Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite		21725590	CAD38376	129
617	Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite		21725592	CAD38377	129
618	Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite		21725594	CAD38378	129
619	Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite		21725596	CAD38379	129
620	Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite		21725600	CAD38381	129
621	Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite		21725602	CAD38382	129
622	Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite		21725604	CAD38383	129
623	Dermatophagoides pteronyssinus	House dust mite	n/a	Aero Mite	Glutathione S-transferase	807138	AAB32224	219
624	Dermatophagoides pteronyssinus	House dust mite	Der p 10	Aero Mite	Tropomyosin	2353266	AAB69424	284
625	Dermatophagoides pteronyssinus	House dust mite	n/a	Aero Mite	NULL	10189816	CAC09235	213
626	Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite		21725578	CAD38370	222
627	Dermatophagoides pteronyssinus	House dust mite	Der p 1	Aero Mite	Cysteine protease-like protein	511953	AAB60215	320
628	Dermatophagoides pteronyssinus	House dust mite	n/a	Aero Mite	NULL	1351935	P49274	19
629	Dermatophagoides pteronyssinus	House dust mite	Der p 3	Aero Mite	Trypsin-like protein	511476	AAA19973	261
630	Dermatophagoides	House dust mite	n/a	Aero Mite	Serine protease LM-1	37654735	AAN02511	261

	pteronyssinus							
631	Dermatophagoides pteronyssinus	House dust mite	Der p 2	Aero Mite		21465915	1KTJ_A	129
632	Dermatophagoides pteronyssinus	House dust mite	n/a	Aero Mite	Serine protease LM-1	31745576	AAP57077	273
633	Dermatophagoides pteronyssinus	House dust mite	Der p 7	Aero Mite		1045602	AAA80264	215
634	Dermatophagoides pteronyssinus	House dust mite	Der p 14	Aero Mite	Group 14 lipid-binding protein	20385544	AAM21322	1662
635	Dicranopteris curranii	Fern	n/a	Aero Plant	2S storage protein-like albumin	6573245	AAF17602	125
636	Dicranopteris linearis	Climbing fern	n/a	Aero Plant	2S storage protein-like albumin	6573243	AAF17601	125
637	Dicranopteris linearis	Climbing fern	n/a	Aero Plant	2S storage protein-like albumin	6573247	AAF17603	125
638	Dictyocaulus viviparus	Parasitic nematode	n/a	Worm (parasite)	Lipid-binding protein	2498317	Q24702	1557
639	Dirofilaria immitis	Parasitic nematode	n/a	Worm (parasite)	Venom antigen 5-like protein	2245508	AAB62535	221
640	Dolichovespula arenaria	Yellow jacket	Dol a 5	Venom or Salivary	Antigen 5	465052	Q05108	203
641	Dolichovespula maculata	Whiteface hornet	Dol m 2	Venom or Salivary	Hyaluronidase	1346322	P49371	331
642	Dolichovespula maculata	Whiteface hornet	Dol m 5	Venom or Salivary	Antigen 5 form 3	549186	P10737	215
643	Dolichovespula maculata	Whiteface hornet	Dol m 5	Venom or Salivary	Phospholipase A1 1	548449	Q06478	317
644	Dolichovespula maculata	Whiteface hornet	n/a	Venom or Salivary	Antigen 5	85299	B31085	213
645	Dolichovespula maculata	Whiteface hornet	Dol m 1	Venom or Salivary	Phospholipase A1 2	1709542	P53357	303
646	Dolichovespula maculata	Whiteface hornet	Dol m 5	Venom or Salivary	Antigen 5 form 2	137395	P10736	227
647	Embellisia allii	Fungus	n/a	Aero Fungi	NULL	49476559	AAT66613	138
648	Embellisia indefessa	Fungus	n/a	Aero Fungi	NULL	49476561	AAT66614	137
649	Embellisia novae-zelandiae	Fungus	n/a	Aero Fungi	NULL	49476563	AAT66615	136
650	Embellisia telluster	Fungus	n/a	Aero Fungi	NULL	49476565	AAT66616	138
651	Epicoccum nigrum	Fungus	n/a	Unassigned	NULL	24636821	P83340_1	12
652	Equus caballus	Horse	Equ c 3	Aero Animal	Preproalbumin	399672	CAA52194	607
653	Equus caballus	Horse	Equ c 1	Aero Animal	Lipocalin-like protein	3121758	Q95182	187
654	Equus caballus	Horse	n/a	Aero Animal	Latherin	20143977	AAM09530	228
655	Equus caballus	Horse	Equ c 2.0102	Aero Animal	Lipocalin-like protein (fragment)	3121756	P81217	19
656	Equus caballus	Horse	Equ c 2.0101	Aero Animal	Lipocalin-like protein (fragment)	3121755	P81216	29
657	Euroglyphus maynei	House dust mite	Eur m 2	Aero Mite	Mite group 2 protein	14423649	Q9TZZ2	145

658	Euroglyphus maynei	House dust mite	Eur m 14	Aero Mite	M-177	6492307	AAF14270	1668
659	Euroglyphus maynei	House dust mite	n/a	Aero Mite	Putative cysteine protease	3941388	AAC82351	321
660	Euroglyphus maynei	House dust mite	Eur m 2.0102	Aero Mite	Mite group 2 protein	3941386	AAC82350	135
661	Euroglyphus maynei	House dust mite	n/a	Aero Mite	Alpha-amylase	5059164	AAD38943	521
662	Euroglyphus maynei	House dust mite	n/a	Aero Mite	Peptidase S1-like protein	14423685	O97370	261
663	Euroglyphus maynei	House dust mite	n/a	Aero Mite	Cysteine protease	3941390	AAC82352	327
664	Fagopyrum esculentum	Buckwheat	n/a	Food Plant	Legumin-like protein	2317670	BAA21758	565
665	Fagopyrum esculentum	Buckwheat	n/a	Food Plant	NULL	61970231	AAX57578	127
666	Fagopyrum esculentum	Buckwheat	n/a	Food Plant	Legumin-like protein	2317674	BAA21760	504
667	Fagopyrum esculentum	Buckwheat	n/a	Food Plant	BW 8 kDaprotein	17907758	BAB79444	133
668	Fagopyrum esculentum	Buckwheat	n/a	Food Plant	Legumin-like protein	4895075	AAD32713	538
669	Fagopyrum gracilipes	Buckwheat	n/a	Food Plant	22 kDa storage protein	6979766	AAF34634	191
670	Fagopyrum tataricum	Buckwheat	n/a	Food Plant	Cupin	22353013	AAK97787	195
671	Farfantepenaeus aztecus	Brown shrimp	Pen a 1	Food Animal	Tropomyosin	632782	AAB31957	21
672	Felis catus	Cat	Fel d 3	Aero Animal	Cystatin	17939981	AAL49391	98
673	Felis catus	Cat	Fel d 1	Aero Animal	Chain 1 long form	1082944	A56413	92
674	Felis catus	Cat	Fel d 1	Aero Animal	Chain A	38492847	1PUO_A	170
675	Felis catus	Cat	Fel d 1	Aero Animal	Chain 1	163827	AAC41617	88
676	Felis catus	Cat	Fel d 1	Aero Animal	Chain 2 short form	423192	JC1127	107
677	Felis catus	Cat	Fel d 1	Aero Animal	Uteroglobin-like protein	1169666	P30439	88
678	Felis catus	Cat	n/a	Aero Animal	NULL	1169665	P30438	92
679	Felis catus	Cat	Fel d 2	Aero Animal	Albumin	886485	CAA59279	608
680	Felis catus	Cat	Fel d 4	Aero Animal	NULL	45775300	AAS77253	186
681	Felis catus	Cat	Fel d 1	Aero Animal	Salivary protein	163823	AAC41616	109
682	Felis catus	Cat	Fel d 1	Aero Animal	Beta chain	539716	B53283	20
683	Felis catus	Cat	Fel d 1	Aero Animal	Chain 1 leader B	1364212	CAA44343	88
684	Felis catus	Cat	Fel d 1	Aero Animal	Chain 1 precursor A	423190	JC1136	92
685	Felis catus	Cat	n/a	Aero Animal	Prostaglandin FP receptor	17224444	AAL36977	366
686	Felis catus	Cat	Fel d 1	Aero Animal	Chain 1 leader A	1364213	CAA44344	92
687	Felis catus	Cat	Fel d 1	Aero Animal	Chain 1 precursor B	423191	JC1126	88
688	Fragaria x ananassa	Strawberry	Fra a 1	Food Plant	NULL	60389905	Q5ULZ4_3	74
689	Fragaria x ananassa	Strawberry	n/a	Food Plant	NULL	54306612	AAV33460	80
690	Fragaria x ananassa	Strawberry	Fra a 1	Food Plant	NULL	60389904	Q5ULZ4_2	14
691	Fragaria x ananassa	Strawberry	Fra a 1	Food Plant	NULL	60389903	Q5ULZ4_1	12
692	Fraxinus excelsior	European ash	Fra e 1	Aero Plant	Pollen allergen	34978692	AAQ83588	146
693	Fraxinus excelsior	European ash	Fra e 1	Aero Plant	NULL	56122438	AAV74343	145
694	Fraxinus excelsior	European ash	Fra e 1	Aero Plant	NULL	33327133	AAQ08947	145
695	Fusarium culmorum	Fungus	Fus c 2	Aero Fungi	Thioredoxin-like protein	19879659	AAL79931	121
696	Fusarium culmorum	Fungus	n/a	Aero Fungi	Unnamed protein product	27965571	AAO25544	342
697	Fusarium culmorum	Fungus	n/a	Aero Fungi	HLH protein	25361513	AAN73248	450
698	Fusarium culmorum	Fungus	Fus c 1	Aero Fungi	60S acidic ribosomal protein P2	19879657	AAL79930	109
699	Gadus callarias	Baltic cod	Gad c 1	Food Animal	Parvalbumin beta	131112	P02622	113

700	Gadus morhua	Atlantic cod	n/a	Food Animal	Parvalbumin beta	14531016	AAK63087	109
701	Gadus morhua	Atlantic cod	n/a	Food Animal	Parvalbumin beta	14531014	AAK63086	109
702	Gallus gallus	Chicken	Gal d 2	Food Animal	Ovalbumin	212900	AAA68882	388
703	Gallus gallus	Chicken	Gal d 2	Food Animal	Ovalbumin	808969	CAA23682	386
704	Gallus gallus	Chicken	n/a	Food Animal	Vitellogenin	63885	CAA25027	81
705	Gallus gallus	Chicken	Gal d 2	Food Animal	Ovalbumin	129293	P01012	386
706	Gallus gallus	Chicken	Gal d 5	Food Animal	Preproalbumin	63748	CAA43098	615
707	Gallus gallus	Chicken	Gal d 2	Food Animal	Ovalbumin	63052	CAA23681	155
708	Gallus gallus	Chicken	n/a	Food Animal	Very low density lipoprotein II	211156	AAA48596	106
709	Gallus gallus	Chicken	Gal d 1	Food Animal	Ovomucoid	124757	P01005	210
710	Gallus gallus	Chicken	Gal d 3	Food Animal	Ovotransferrin	1351295	P02789	705
711	Gallus gallus	Chicken	n/a	Food Animal	NULL	34811333	1UHG_D	385
712	Gallus gallus	Chicken	Gal d 2	Food Animal	Ovalbumin	15826578	1JTI_A	385
713	Gallus gallus	Chicken	n/a	Food Animal	Vitellogenin	212881	AAA49139	1852
714	Gallus gallus	Chicken	n/a	Food Animal	NULL	50731634	XP_418301	155
715	Gallus gallus	Chicken	Gal d 4	Food Animal	Lysozyme fragment	212279	AAA48944	24
716	Gallus gallus	Chicken	Gal d 4	Food Animal	Lysozyme C	126608	P00698	147
717	Gallus gallus	Chicken	Gal d 3	Food Animal	Ovotransferrin	757851	CAA26040	705
718	Gallus gallus	Chicken	n/a	Food Animal	Vitellogenin	63887	CAA31942	1850
719	Gibberella zeae PH-1	Fungus	n/a	Unassigned	NULL	46122455	XP_385781	109
720	Gibberella zeae PH-1	Fungus	n/a	Unassigned	NULL	46137705	XP_390544	108
721	Gibberella zeae PH-1	Fungus	n/a	Unassigned	NULL	46108928	XP_381522	438
722	Glycine max	Soybean	Gly m 2	Food Plant	Hull protein fragment	1362049	A57106	20
723	Glycine max	Soybean	n/a	Food Plant	Kunitz trypsin inhibitor KTi2	256636	AAB23483	204
724	Glycine max	Soybean	n/a	Food Plant	Glycinin subunit G2	18637	CAA33216	485
725	Glycine max	Soybean	n/a	Food Plant	CG4 beta-conglycinin	256427	AAB23463	439
726	Glycine max	Soybean	n/a	Food Plant	Kunitz trypsin inhibitor KTi	256429	AAB23464	216
727	Glycine max	Soybean	n/a	Food Plant	Kunitz trypsin inhibitor KTi1	256635	AAB23482	203
728	Glycine max	Soybean	n/a	Food Plant	Glycinin subunit G1	18635	CAA33215	495
729	Glycine max	Soybean	Gly m 3	Food Plant	Profilin	3021373	CAA11755	131
730	Glycine max	Soybean	n/a	Food Plant	Gly m Bd 30 kDa (also called P34)	129353	P22895	379
731	Glycine max	Soybean	n/a	Food Plant	NULL	123506	P24337	80
732	Glycine max	Soybean	n/a	Food Plant	Cim1 protein	1076486	S48032	277
733	Glycine max	Soybean	Gly m 1.0101	Food Plant	Hydrophobic seed protein fragment	999355	AAB34755	42
734	Glycine max	Soybean	n/a	Food Plant	Trypsin inhibitor subtype A	18770	CAA45777	217
735	Glycine max	Soybean	Gly m 4	Food Plant	Pathogenesis-related protein	18744	CAA42646	158
736	Glycine max	Soybean	n/a	Food Plant	Glycinin	18641	CAA37044	562
737	Glycine max	Soybean	n/a	Food Plant	Gly m Bd 30 kDa (also called P34)	1199563	AAB09252	379
738	Glycine max	Soybean	n/a	Food Plant	Glycinin subunit G3	18639	CAA33217	481
739	Glycine max	Soybean	n/a	Food Plant	Bd 28 kDa	12697782	BAB21619	473
740	Glycine max	Soybean	n/a	Food Plant	Trypsin inhibitor subtype B	18772	CAA45778	217

741	Glycine max	Soybean	Gly m 3	Food Plant	Profilin 1	3914435	O65809	131
742	Glycine max	Soybean	n/a	Food Plant	Beta-conglycinin, alpha subunit	18536	CAA35691	605
743	Glycine max	Soybean	n/a	Food Plant	Gly m Bd 30 kDa (also called P34)	3097321	BAA25899	379
744	Glycine max	Soybean	n/a	Food Plant	Kunitz trypsin inhibitor	510515	CAA56343	208
745	Glycine max	Soybean	n/a	Food Plant	Glycinin A2B1a subunit	18609	CAA26575	485
746	Glycine max	Soybean	n/a	Food Plant	Beta-conglycinin, alpha subunit	169927	AAA33947	218
747	Glycine max	Soybean	n/a	Food Plant	Glycinin	18615	CAA26723	495
748	Glycine max	Soybean	n/a	Food Plant	Beta-conglycinin storage protein	169929	AAB01374	639
749	Glycine max	Soybean	n/a	Food Plant	Glycinin	169969	AAA33964	516
750	Glycine max	Soybean	n/a	Food Plant	Glycinin	169971	AAA33965	240
751	Glycine max	Soybean	n/a	Food Plant	Lectin prepeptide	170006	AAA33983	285
752	Glycine max	Soybean	n/a	Food Plant	Glycinin A5A4B3 subunits	732706	CAA26478	562
753	Glycine soja	Soybean	n/a	Food Plant	Glycinin A5A4B3 subunit	806556	CAA60533	563
754	Glycine soja	Soybean	n/a	Food Plant	Glycinin A3B4 subunit	736002	CAA55977	517
755	Glycyphagus domesticus	Storage mite	Gly d 2.0102	Aero Mite	Group II allergen	7160811	CAB76459	125
756	Glycyphagus domesticus	Storage mite	Gly d 2.0101	Aero Mite	Group II allergen	6179520	CAB59976	128
757	Haliotis diversicolor	Mollusk	n/a	Food Animal	Tropomyosin	9954249	AAG08987	284
758	Helianthus annuus	Sunflower	Hel a 2	Aero Plant	Profilin	3581965	CAA75506	133
759	Helianthus annuus	Sunflower	n/a	Food Plant	Seed 2S albumin	18821	CAA40015	141
760	Helix aspersa	Brown garden snail	n/a	Food Animal	Tropomyosin	4468224	CAB38044	284
761	Heterodera glycines	Nematode	n/a	Worm (parasite)	Putative polyprotein Hgg-14	18677166	AAL78215	288
762	Heterodera glycines	Nematode	n/a	Worm (parasite)	Vap-1 (plant parasite)	14326230	AAK60209	215
763	Heterodera glycines	Nematode	n/a	Worm (parasite)	NULL	13447461	AAK21961	212
764	Heterodera glycines	Nematode	n/a	Worm (parasite)	Venom-like protein	14211968	AAK55116	212
765	Hevea brasiliensis	Para rubber tree	Hev b 8	Contact	Profilin	3183706	CAA75312	131
766	Hevea brasiliensis	Para rubber tree	Hev b 10.0102	Contact	MnSOD	5777414	CAB53458	205
767	Hevea brasiliensis	Para rubber tree	Hev b 1	Contact	Rubber elongation factor protein	132270	P15252	138
768	Hevea brasiliensis	Para rubber tree	Hev b 8.0204	Contact	Profilin 6	14423856	Q9LEI8	131
769	Hevea brasiliensis	Para rubber tree	Hev b 8.0203	Contact	Profilin 5	14423858	Q9M7M8	131
770	Hevea brasiliensis	Para rubber tree	Hev b 8.0202	Contact	Profilin 4	14423859	Q9M7M9	131
771	Hevea brasiliensis	Para rubber tree	Hev b 7.02	Contact	Patatin-like protein	3087805	CAA11041	388
772	Hevea brasiliensis	Para rubber tree	Hev b 6	Contact	Prohevein	2832430	CAA05978	187
773	Hevea brasiliensis	Para rubber tree	Hev b 8.0201	Contact	Profilin 3	14423860	Q9M7N0	131
774	Hevea brasiliensis	Para rubber tree	Hev b 8.0102	Contact	Profilin 2	14423868	Q9STB6	131
775	Hevea brasiliensis	Para rubber tree	Hev b 7.01	Contact	Latex patatin homolog	1916805	AAC27724	388
776	Hevea brasiliensis	Para rubber tree	Hev b 3	Contact	Small rubber particle protein	14423933	O82803	204
777	Hevea brasiliensis	Para rubber tree	Hev b 6	Contact	Prohevein	123062	P02877	204
778	Hevea brasiliensis	Para rubber tree	Hev b 11	Contact	Putative class I chitinase	14575525	CAC42881	295

779	Hevea brasiliensis	Para rubber tree	n/a	Contact	Patatin-like protein	3288200	CAA11042	388
780	Hevea brasiliensis	Para rubber tree	Hev b 10.0101	Contact	MnSOD	348137	AAA16792	233
781	Hevea brasiliensis	Para rubber tree	Hev b 2	Contact	Beta-1,3-glucanase	32765543	AAP87281	374
782	Hevea brasiliensis	Para rubber tree	Hev b 12	Contact	Lipid transfer protein	20135538	AAL25839	116
783	Hevea brasiliensis	Para rubber tree	Hev b 5	Contact		1480457	AAC49447	151
784	Hevea brasiliensis	Para rubber tree	Hev b 9	Contact	Enolase, isoform 2	9581746	CAC00533	445
785	Hevea brasiliensis	Para rubber tree	Hev b 2	Contact	Beta-1,3-glucanase	1184668	AAA87456	374
786	Hevea brasiliensis	Para rubber tree	n/a	Contact	NULL	41581137	CAE85467	387
787	Hevea brasiliensis	Para rubber tree	Hev b 7	Contact	Patatin-like protein	6707018	AAF25553	388
788	Hevea brasiliensis	Para rubber tree	Hev b 10.0103	Contact	MnSOD	10862818	CAC13961	205
789	Hevea brasiliensis	Para rubber tree	Hev b 8	Contact	Profilin, chain A	11513601	1G5U_A	131
790	Hevea brasiliensis	Para rubber tree	Hev b 9	Contact	Enolase, isoform 1	9581744	CAC00532	445
791	Hevea brasiliensis	Para rubber tree	n/a	Contact	NULL	46410859	AAR98518	366
792	Hevea brasiliensis	Para rubber tree	Hev b 13	Contact	NULL	30909057	AAP37470	391
793	Hevea brasiliensis subsp. brasiliensis	Para rubber tree	Hev b 11	Contact	NULL	27526732	CAD24068	295
794	Holcus lanatus	Velvet grass	n/a	Aero Plant	Group II	4007084	CAA10347	122
795	Holcus lanatus	Velvet grass	Hol l 1.0102	Aero Plant	Group I	1167836	CAA93121	248
796	Holcus lanatus	Velvet grass	n/a	Aero Plant	30 kDa fragment	543482	S38291	20
797	Holcus lanatus	Velvet grass	n/a	Aero Plant	Group I	3860384	CAA10140	263
798	Holcus lanatus	Velvet grass	Hol l 1.0101	Aero Plant		414703	CAA81610	265
799	Holcus lanatus	Velvet grass	n/a	Aero Plant	Group V	2266625	CAB10765	264
800	Holcus lanatus	Velvet grass	n/a	Aero Plant	Group V	11991229	AAG42255	296
801	Holcus lanatus	Velvet grass	n/a	Aero Plant	Group V	2266623	CAB10766	240
802	Homarus americanus	American lobster	n/a	Food Animal	Tropomyosin	14285796	O44119	284
803	Homarus americanus	American lobster	n/a	Food Animal	Fast tropomyosin isoform	2660868	AAC48288	284
804	Hordeum vulgare	Barley	Hor v 15	Aero Plant	Alpha-amylase inhibitor BMAI-1	2506771	P16968	146
805	Hordeum vulgare	Barley	n/a	Food Plant	Amylase/protease inhibitor (LTP)	167077	AAA32970	117
806	Hordeum vulgare	Barley	n/a	Aero Plant	Pollen allergen	7442148	JC5475	313
807	Hordeum vulgare	Barley	n/a	Food Plant	Dimeric protein (fragment)	2146809	S65605	20
808	Hordeum vulgare	Barley	n/a	Food Plant	Alpha-amylase/trypsin inhibitor (CMb)	585290	P32936	149
809	Hordeum vulgare	Barley	n/a	Food Plant	LTP 1	19039	CAA42832	134
810	Hordeum vulgare	Barley	n/a	Food Plant	Trypsin inhibitor cme	1405736	CAA35188	144
811	Hordeum vulgare	Barley	n/a	Aero Plant	NULL	55859462	CAH92635	518
812	Hordeum vulgare subsp. vulgare	Barley	n/a	Food Plant	Alpha-amylase/trypsin inhibitor	18869	CAA31794	147
813	Hordeum vulgare subsp. vulgare	Barley	n/a	Food Plant	Alpha-amylase inhibitor component CMb	452323	CAA49556	149
814	Hordeum vulgare subsp.	Barley	n/a	Food Plant	Alpha-amylase inhibitor	439275	CAA49555	145

	vulgare				component Cma			
815	Hordeum vulgare subsp. vulgare	Barley	n/a	Food Plant	Trypsin inhibitor cme	19009	CAA46705	148
816	Hordeum vulgare subsp. vulgare	Barley	n/a	Food Plant	Alpha-amylase inhibitor component CMd	452325	CAA49557	171
817	Hordeum vulgare subsp. vulgare	Barley	n/a	Food Plant	Alpha-amylase inhibitor BDAI-1	3367714	CAA08836	152
818	Hordeum vulgare subsp. vulgare	Barley	n/a	Aero Plant	NULL	19003	CAA45085	146
819	Hordeum vulgare subsp. vulgare	Barley	n/a	Food Plant	Alpha-amylase inhibitor	18955	CAA41956	144
820	Humulus japonicus	Japanese hop	Hum j 1	Aero Plant	NULL	33113263	AAP94213	155
821	Humulus scandens	Japanese hop	n/a	Unassigned	NULL	34851176	AAP15200	131
822	Humulus scandens	Japanese hop	n/a	Unassigned	NULL	34851174	AAP15199	131
823	Ixodes ricinus	Castor bean tick	n/a	Venom or Salivary	Antibacterial protein	28564461	CAD68004	116
824	Juglans nigra	Black walnut	Jug n 2	Food Plant	NULL	31321944	AAM54366	481
825	Juglans nigra	Black walnut	Jug n 1	Food Plant	Albumin seed storage protein	31321942	AAM54365	161
826	Juglans regia	English walnut	Jug r 1	Food Plant	Albumin seed storage protein	1794252	AAB41308	139
827	Juglans regia	English walnut	Jug r 2	Food Plant	Vicilin-like protein	6580762	AAF18269	593
828	Juniperus ashei	Mountain cedar	Jun a 1	Aero Plant	Unnamed protein product	4138877	AAD03608	367
829	Juniperus ashei	Mountain cedar	Jun a 3	Aero Plant	Unnamed protein product	6940772	AAF31759	225
830	Juniperus ashei	Mountain cedar	Jun a 2	Aero Plant	Unnamed protein product	9955725	CAC05582	507
831	Juniperus ashei	Mountain cedar	n/a	Aero Plant	Elongation factor	21632054	AAK85129	172
832	Juniperus oxycedrus	Juniper	Jun o 4	Aero Plant	EF-hand calcium-binding protein	5391446	AAC15474	165
833	Juniperus oxycedrus	Juniper	n/a	Aero Plant	Putative pectate lyase	15139849	CAC48400	367
834	Juniperus rigida	Cedar	n/a	Aero Plant	NULL	38456222	AAR21071	225
835	Juniperus rigida	Cedar	n/a	Aero Plant	NULL	38456224	AAR21072	225
836	Juniperus virginiana	Red cedar	Jun v 1	Aero Plant	Unnamed protein product	8843917	AAF80164	367
837	Juniperus virginiana	Red cedar	Jun v 1	Aero Plant	Unnamed protein product	8843921	AAF80166	367
838	Juniperus virginiana	Red cedar	n/a	Aero Plant	NULL	51316532	Q9LD79	110
839	Kryptolebias marmoratus	Mangrove rivulus	n/a	Food Animal	NULL	50953781	AAT90497	109
840	Kryptolebias marmoratus	Mangrove rivulus	n/a	Food Animal	NULL	50953783	AAT90498	109
841	Lens culinaris	Lentil	Len c 1.0102	Food Plant	Seed storage protein	29539111	CAD87731	415
842	Lens culinaris	Lentil	Len c 1.0101	Food Plant	Seed storage protein	29539109	CAD87730	418
843	Lepidoglyphus destructor	Storage mite	Lep d 2	Aero Mite	Unnamed protein product	2147108	S66499	141
844	Lepidoglyphus destructor	Storage mite	Lep d 13	Aero Mite	Fatty acid-binding protein	14423714	Q9U5P1	131
845	Lepidoglyphus destructor	Storage mite	Lep d 5	Aero Mite	Mite group 5 protein	14423651	Q9U5P2	110
846	Lepidoglyphus destructor	Storage mite	Lep d 7	Aero Mite	Mite Group 7 protein	14423650	Q9U1G2	216
847	Lepidoglyphus destructor	Storage mite	Lep d 2	Aero Mite	NULL	1582223	2118249B	141
848	Lepidoglyphus destructor	Storage mite	Lep d 10	Aero Mite	Tropomyosin	14423956	Q9NFZ4	284
849	Lepidoglyphus destructor	Storage mite	n/a	Aero Mite	NULL	19702131	CAD20979	450
850	Lepidoglyphus destructor	Storage mite	Lep d 2	Aero Mite	Unnamed protein product	999462	CAA61419	141

851	Lepidoglyphus destructor	Storage mite	Lep d 2	Aero Mite	NULL	1582222	2118249A	141
852	Lepidoglyphus destructor	Storage mite	Lep d 2	Aero Mite	Unnamed protein product	21213900	CAD32314	141
853	Lepidoglyphus destructor	Storage mite	Lep d 2	Aero Mite	Unnamed protein product	21213898	CAD32313	141
854	Lepisma saccharina	Silverfish	n/a	Aero Mite	Tropomyosin	20387029	CAC84593	243
855	Lepisma saccharina	Silverfish	Lep s 1	Aero Mite	Tropomyosin	20387027	CAC84590	284
856	Lewia ethzedia	Fungus	n/a	Aero Fungi	NULL	49476483	AAT66575	138
857	Ligustrum vulgare	Privet	Lig v 1.0101	Aero Plant	Ole e 1 homolog	3256210	CAA54818	145
858	Ligustrum vulgare	Privet	Lig v 1.0102	Aero Plant	Ole e 1 homolog	3256212	CAA54819	145
859	Lilium longiflorum	Trumpet lily	n/a	Aero Plant	NULL	6425109	AAF08304	131
860	Lilium longiflorum	Trumpet lily	n/a	Aero Plant	NULL	6425107	AAF08303	131
861	Lilium longiflorum	Trumpet lily	n/a	Aero Plant	NULL	6425105	AAF08302	131
862	Litchi chinensis	Lychee nut	Lit c 1	Food Plant	Profilin	15809696	AAL07320	131
863	Litomosoides carinii	Nematode	n/a	Worm (parasite)	NULL	4102911	AAD01619	132
864	Loa loa	Nematode	n/a	Worm (parasite)	LL20 15 kDa ladder antigen	414553	AAB54037	588
865	Loa loa	Nematode	n/a	Worm (parasite)	Ladder protein	4102951	AAD01626	132
866	Lolium italicum	Italian rye grass	n/a	Aero Plant	Group II	4007636	CAA10350	122
867	Lolium perenne	Perennial ryegrass	Lol p 2	Aero Plant	Expansin-like protein	126386	P14947	97
868	Lolium perenne	Perennial ryegrass	Lol p 3	Aero Plant	Expansin-like protein	126387	P14948	97
869	Lolium perenne	Perennial ryegrass	n/a	Aero Plant	Fragment	82450	A23341	31
870	Lolium perenne	Perennial ryegrass	n/a	Aero Plant	NULL	55859464	CAH92637	423
871	Lolium perenne	Perennial ryegrass	Lol p 2	Aero Plant	Group II	939932	CAA51775	88
872	Lolium perenne	Perennial ryegrass	Lol p 5	Aero Plant	Ribonuclease	6634467	CAB64344	301
873	Lolium perenne	Perennial ryegrass	Lol p 5.0102	Aero Plant	Ribonuclease	100639	A38582	308
874	Lolium perenne	Perennial ryegrass	Lol p 1	Aero Plant	Group I	168314	AAA63278	252
875	Lolium perenne	Perennial ryegrass	Lol p 1	Aero Plant	Group I	100636	S13614	263
876	Lolium perenne	Perennial ryegrass	Lol p 1	Aero Plant	Expansin-like protein Group I	126385	P14946	263
877	Lolium perenne	Perennial ryegrass	Lol p 2	Aero Plant	Fragment	542132	A48595	88
878	Lolium perenne	Perennial ryegrass	Lol p 11	Aero Plant	Ole e 1 homolog	626029	A54002	134
879	Lolium perenne	Perennial ryegrass	Lol p 5	Aero Plant	Ribonuclease	4416516	AAD20386	301
880	Lolium perenne	Perennial ryegrass	Lol p 5	Aero Plant	Poa p 9/Phl p 6 homolog	2498582	Q40237	339
881	Lolium perenne	Perennial ryegrass	Lol p 1	Aero Plant	Clone 5A	100638	B37881	263
882	Lolium perenne	Perennial ryegrass	Lol p 5	Aero Plant	Poa p 9/Phl p 6 homolog	2498581	Q40240	308
883	Lolium perenne	Perennial ryegrass	n/a	Aero Plant	50 kDa fragment Group I	542131	S38288	20
884	Lupinus luteus	Yellow lupine	n/a	Food Plant	NULL	61680377	1XDF_B	157
885	Lupinus luteus	Yellow lupine	n/a	Food Plant	Pathogenesis-related protein, chain A	22218768	1ICX_A	155
886	Lupinus luteus	Yellow lupine	n/a	Food Plant	NULL	22218772	1IFV_B	155
887	Lycopersicon esculentum	Tomato	Lyc e 1	Food Plant	Profilin	16555787	CAD10377	131
888	Lycopersicon esculentum	Tomato	Lyc e 2.0102	Food Plant	Beta-fructofuranosidase	18542115	AAL75450	636
889	Lycopersicon esculentum	Tomato	n/a	Food Plant	Ole e 1 homolog	125887	P13447	161
890	Lycopersicon esculentum	Tomato	Lyc e 2.0101	Food Plant	Beta-fructofuranosidase	18542113	AAL75449	553
891	Lycopersicon esculentum	Tomato	Lyc e 1	Food Plant	Profilin	17224229	AAL29690	131
892	Macaca mulatta	Rhesus monkey	n/a	Unassigned	NULL	229507	750521A	14
893	Macrospora scirpicola	Fungus	n/a	Aero Fungi	NULL	49476555	AAT66611	138

894	Malassezia furfur	Yeast	Mala f 3	Contact	Unnamed protein product	7514251	JE0227	166
895	Malassezia furfur	Yeast	n/a	Contact	NULL	13959403	Q01940	350
896	Malassezia furfur	Yeast	Mala f 3	Contact	MF2	3445492	BAA32436	166
897	Malassezia furfur	Yeast	Mala f 2	Contact	MF1	3445490	BAA32435	177
898	Malassezia furfur	Yeast	Mala f 4	Contact	Unnamed protein product	4587985	AAD25927	342
899	Malassezia sympodialis	Yeast	Mala s 7	Contact	Unnamed protein product	4138175	CAA09885	187
900	Malassezia sympodialis	Yeast	Mala s 6	Contact	Unnamed protein product	4138173	CAA09884	162
901	Malassezia sympodialis	Yeast	Mala s 5	Contact	Unnamed protein product	4138171	CAA09883	172
902	Malassezia sympodialis	Yeast	Mala s 9	Contact	Unnamed protein product	19069920	CAA09887	342
903	Malassezia sympodialis	Yeast	n/a	Contact	NULL	58416210	CAI43283	618
904	Malassezia sympodialis	Yeast	Mala s 1	Contact	Unnamed protein product	1261972	CAA65341	350
905	Malassezia sympodialis	Yeast	Mala s 11	Contact	Manganese superoxide dismutase	28569698	CAD68071	237
906	Malassezia sympodialis	Yeast	Mala s 8	Contact	Unnamed protein product	7271239	CAA09886	179
907	Malus x domestica	Apple	Mal d 1	Food Plant	PR-10 (Bet v 1 homologue)	1313970	CAA96536	160
908	Malus x domestica	Apple	n/a	Food Plant	NULL	32709122	AAP86780	190
909	Malus x domestica	Apple	Mal d 1	Food Plant	PR-10 (Bet v 1 homologue)	2443824	AAD13683	159
910	Malus x domestica	Apple	Mal d 1	Food Plant	PR-10 (Bet v 1 homologue)	1313966	CAA96534	159
911	Malus x domestica	Apple	Mal d 4	Food Plant	Profilin 1 (GD4-1)	14423873	Q9XF40	131
912	Malus x domestica	Apple	n/a	Food Plant	NULL	38492338	AAR22488	115
913	Malus x domestica	Apple	n/a	Food Plant	Fragment	1478293	AAB35897	26
914	Malus x domestica	Apple	n/a	Food Plant	18 kDa Bet v 1 homolog fragment	1478292	AAB35896	25
915	Malus x domestica	Apple	Mal d 1	Food Plant	PR-10 (Bet v 1 homologue)	1313972	CAA96537	160
916	Malus x domestica	Apple	Mal d 2	Food Plant	Thaumatin-like protein	10334651	CAC10270	246
917	Malus x domestica	Apple	Mal d 1	Food Plant	PR-10 (Bet v 1 homologue)	2119763	JC4276	159
918	Malus x domestica	Apple	n/a	Food Plant	Fragment	1174276	AAB35328	37
919	Malus x domestica	Apple	Mal d 4	Food Plant	NULL	28881457	CAD46561	131
920	Malus x domestica	Apple	Mal d 1	Food Plant	PR-10 (Bet v 1 homologue)	1313968	CAA96535	160
921	Malus x domestica	Apple	Mal d 1	Food Plant	PR-10 (Bet v 1 homologue)	4590380	AAD26554	159
922	Malus x domestica	Apple	Mal d 3	Food Plant	Lipid transfer protein	6715522	AAF26450	115
923	Malus x domestica	Apple	Mal d 1	Food Plant	PR-10 (Bet v 1 homologue)	4590378	AAD26553	159
924	Malus x domestica	Apple	n/a	Food Plant	NULL	50659859	AAT80649	115
925	Malus x domestica	Apple	n/a	Food Plant	NULL	50659885	AAT80662	115
926	Malus x domestica	Apple	Mal d 1	Food Plant	PR-10 (Bet v 1 homologue)	4590364	AAD26546	159
927	Malus x domestica	Apple	Mal d 1	Food Plant	PR-10 (Bet v 1 homologue)	4590388	AAD26558	159
928	Malus x domestica	Apple	n/a	Food Plant	NULL	50659891	AAT80665	115
929	Malus x domestica	Apple	Mal d 1	Food Plant	PR-10 (Bet v 1 homologue)	4590382	AAD26555	159
930	Malus x domestica	Apple	Mal d 1	Food Plant	PR-10 (Bet v 1 homologue)	4590366	AAD26547	159
931	Malus x domestica	Apple	Mal d 1	Food Plant	PR-10 (Bet v 1 homologue)	4590376	AAD26552	159
932	Malus x domestica	Apple	Mal d 1	Food Plant	PR-10 (Bet v 1 homologue)	4590368	AAD26548	159
933	Malus x domestica	Apple	Mal d 1	Food Plant	PR-10 (Bet v 1 homologue)	21685277	CAD32318	158
934	Malus x domestica	Apple	Mal d 4	Food Plant	Profilin 2 (GD4-2)	14423874	Q9XF41	131
935	Malus x domestica	Apple	Mal d 4	Food Plant	Profilin 3 (GD4-5)	14423875	Q9XF42	131

936	Malus x domestica	Apple	n/a	Food Plant	NULL	50659889	AAT80664	115
937	Malus x domestica	Apple	Mal d 4	Food Plant	NULL	28881453	CAD46559	131
938	Malus x domestica	Apple	n/a	Food Plant	NULL	50659879	AAT80659	115
939	Malus x domestica	Apple	Mal d 4	Food Plant	NULL	28881455	CAD46560	131
940	Malus x domestica	Apple	Mal d 1	Food Plant	PR-10 (Bet v 1 homologue)	747852	CAA58646	159
941	Malus x domestica	Apple	Mal d 1	Food Plant	PR-10 (Bet v 1 homologue)	16555783	CAD10375	159
942	Malus x domestica	Apple	Mal d 1	Food Plant	PR-10 (Bet v 1 homologue)	3309647	AAC26136	159
943	Malus x domestica	Apple	Mal d 1	Food Plant	PR-10 (Bet v 1 homologue)	27922941	AAO25113	159
944	Manihot esculenta	Cassava	n/a	Food Plant	Pt2L4	21585695	AAM55492	177
945	Meloidogyne incognita	Nematode	n/a	Worm (parasite)	Secreted protein MSP-1	4102596	AAD01511	231
946	Mercurialis annua	Annual mercury grass	Mer a 1	Aero Plant	Profilin	2959898	CAA73720	133
947	Metapenaeus ensis	Shrimp	Met e 1	Food Animal	Tropomyosin	607633	AAA60330	274
948	Methanococcus maripaludis S2	Bacteria	n/a	Unassigned	bacterial protein	45047603	CAF30730	217
949	Mimachlamys nobilis	Noble scallop	n/a	Food Animal	Tropomyosin	9954253	AAG08989	284
950	Mus musculus	Mouse	n/a	Aero Animal	NULL	28476845	DAA00350	112
951	Mus musculus	Mouse	n/a	Aero Animal	NULL	28476849	DAA00352	108
952	Mus musculus	Mouse	n/a	Aero Animal	NULL	28476853	DAA00354	46
953	Mus musculus	Mouse	n/a	Aero Animal	NULL	10181188	NP_065588	93
954	Mus musculus	Mouse	n/a	Aero Animal	NULL	30315676	NP_803229	112
955	Mus musculus	Mouse	n/a	Aero Animal	NULL	30315678	NP_840093	112
956	Mus musculus	Mouse	n/a	Aero Animal	NULL	45331198	NP_919319	112
957	Mus musculus	Mouse	n/a	Aero Animal	NULL	45331208	NP_987098	112
958	Mus musculus	Mouse	Mus m 1	Aero Animal	Major urinary protein 6 (Alpha-2U-globulin)	20178291	P02762	180
959	Mus musculus	Mouse	n/a	Aero Animal	NULL	28476851	DAA00353	112
960	Mus musculus	Mouse	n/a	Aero Animal	NULL	56694673	AAW23035	250
961	Mus musculus	Mouse	n/a	Aero Animal	Androgen-binding protein delta	8926324	AAF81793	112
962	Mus sp.	Mouse	n/a	Aero Animal	DEC-205	1174278	AAB35330	25
963	Musa acuminata	Banana	Mus xp 1	Food Plant	Profilin	14161635	AAK54834	131
964	Myrmecia pilosula	Jumper ant	Myr p 1	Venom or Salivary		2133755	S65709	16
965	Myrmecia pilosula	Jumper ant	Myr p 1	Venom or Salivary	Venom allergen	312284	CAA49760	112
966	Myrmecia pilosula	Jumper ant	Myr p 2	Venom or Salivary		2498604	Q26464	75
967	Myrmecia pilosula	Jumper ant	Myr p 1	Venom or Salivary		345521	S28180	112
968	Myrmecia pilosula	Jumper ant	Myr p 2	Venom or Salivary		1587177	2206305A	75
969	Necator americanus	Hookworm	n/a	Worm (parasite)	Ancylostoma secreted protein 1	3396070	AAD13340	424
970	Necator americanus	Hookworm	n/a	Worm (parasite)	Calreticulin	3687326	CAA07254	403

971	Nectria haematococca mpVI	Fungus	n/a	Aero Fungi	NULL	3122132	P81010	8
972	Neurospora crassa	Fungus	n/a	Aero Fungi	Similar to Asp f 7	28950043	CAD70798	265
973	Neurospora crassa	Fungus	n/a	Aero Fungi	Asp f 9 homolog	28949979	CAD70740	301
974	Nicotiana tabacum	Tobacco	n/a	Aero Plant	NULL	59798467	Q8VWY6	84
975	Nicotiana tabacum	Tobacco	n/a	Aero Plant	NULL	57283137	CAE17316	559
976	Nicotiana tabacum	Tobacco	n/a	Aero Plant	NULL	57283139	CAE17317	520
977	Nicotiana tabacum	Tobacco	n/a	Aero Plant	Beta expansin-like protein	12330698	AAG52887	273
978	Nicotiana tabacum	Tobacco	n/a	Aero Plant	NULL	59798468	Q8VWY7	86
979	Nimbya caricis	Fungus	n/a	Aero Fungi	NULL	49476557	AAT66612	138
980	Olea europaea	Olive tree	Ole e 1.0105	Aero Plant	(Ole e 1)	2465127	CAA73038	146
981	Olea europaea	Olive tree	Ole e 1.0104	Aero Plant	OLE1c fragment (Ole e 1)	1362131	C53806	145
982	Olea europaea	Olive tree	Ole e 1	Aero Plant	OLE20 fragment (Ole e 1)	1362132	A38968	137
983	Olea europaea	Olive tree	n/a	Aero Plant	OLE19 fragment (Ole e 1)	1362130	F53806	136
984	Olea europaea	Olive tree	n/a	Aero Plant	OLE26 fragment (Ole e 1)	1362133	G53806	136
985	Olea europaea	Olive tree	n/a	Aero Plant	OLE17 fragment (Ole e 1)	1362129	E53806	136
986	Olea europaea	Olive tree	n/a	Aero Plant	OLE16 fragment (Ole e 1)	1362128	I53806	137
987	Olea europaea	Olive tree	n/a	Aero Plant	OLE33/OLE37 fragment (Ole e 1)	1362134	D53806	136
988	Olea europaea	Olive tree	Ole e 1	Aero Plant	(Ole e 1)	480443	S36872	145
989	Olea europaea	Olive tree	Ole e 1.0102	Aero Plant	OLE3c (Ole e 1)	1362135	A53806	145
990	Olea europaea	Olive tree	Ole e 1.0103	Aero Plant	OLE5c (Ole e 1)	1362136	B53806	145
991	Olea europaea	Olive tree	Ole e 1	Aero Plant	Fragment	100478	A36153	28
992	Olea europaea	Olive tree	n/a	Aero Plant	OLE6 fragment	1362137	H53806	136
993	Olea europaea	Olive tree	n/a	Aero Plant	NULL	33325115	AAQ08190	132
994	Olea europaea	Olive tree	n/a	Aero Plant	NULL	33329758	AAQ10281	131
995	Olea europaea	Olive tree	n/a	Aero Plant	NULL	33329756	AAQ10280	132
996	Olea europaea	Olive tree	Ole e 5	Aero Plant	Cu/Zn super-oxide dismutase	39840779	CAD21706	152
997	Olea europaea	Olive tree	n/a	Aero Plant	NULL	33329754	AAQ10279	132
998	Olea europaea	Olive tree	n/a	Aero Plant	NULL	33329752	AAQ10278	131
999	Olea europaea	Olive tree	n/a	Aero Plant	NULL	33329748	AAQ10276	129
1000	Olea europaea	Olive tree	Ole e 7	Aero Plant	Unnamed protein product	22002032	P81430	21
1001	Olea europaea	Olive tree	Ole e 9	Aero Plant	Beta-1,3-glucanase-like protein	14279169	AAK58515	460
1002	Olea europaea	Olive tree	n/a	Aero Plant	NULL	33329744	AAQ10274	131
1003	Olea europaea	Olive tree	n/a	Aero Plant	NULL	37548753	AAN18044	132
1004	Olea europaea	Olive tree	n/a	Aero Plant	NULL	33329732	AAQ10268	132
1005	Olea europaea	Olive tree	n/a	Aero Plant	NULL	33329750	AAQ10277	131
1006	Olea europaea	Olive tree	Ole e 1.0106	Aero Plant	(Ole e 1)	2465129	CAA73037	146
1007	Olea europaea	Olive tree	Ole e 8	Aero Plant	Calcium-binding protein	14423648	Q9M7R0	171
1008	Olea europaea	Olive tree	n/a	Aero Plant	NULL	33325111	AAQ08188	132
1009	Olea europaea	Olive tree	Ole e 3	Aero Plant	Calcium-binding allergen	3337403	AAD05375	84
1010	Olea europaea	Olive tree	Ole e 2	Aero Plant	Profilin 1	3914426	O24169	134
1011	Olea europaea	Olive tree	Ole e 2	Aero Plant	Profilin 2	3914427	O24170	134

1012	Olea europaea	Olive tree	Ole e 8	Aero Plant	Calcium-binding protein	6901654	AAF31152	171
1013	Olea europaea	Olive tree	Ole e 6	Aero Plant	Pollen allergen Ole e 6	14423643	O24172	50
1014	Olea europaea	Olive tree	Ole e 2	Aero Plant	Profilin 3	3914428	O24171	134
1015	Olea europaea	Olive tree	Ole e 1.0101	Aero Plant	Main olive allergen	13195753	AAB32652	130
1016	Olea europaea	Olive tree	n/a	Aero Plant	NULL	33329738	AAQ10271	132
1017	Olea europaea	Olive tree	n/a	Aero Plant	NULL	37724593	AAO22132	135
1018	Olea europaea	Olive tree	Ole e 5	Aero Plant	Cu/Zn super-oxide dismutase	2500822	P80740	29
1019	Olea europaea	Olive tree	Ole e 1.0107	Aero Plant	Profilin	2465131	CAA73036	146
1020	Olea europaea	Olive tree	Ole e 4	Aero Plant	NULL	32363448	P80741_1	10
1021	Olea europaea	Olive tree	Ole e 4	Aero Plant	NULL	32363449	P80741_2	14
1022	Olea europaea	Olive tree	Ole e 3	Aero Plant	NULL	37725377	AAO33897	52
1023	Olea europaea	Olive tree	n/a	Aero Plant	NULL	37724597	AAO22133	134
1024	Olea europaea	Olive tree	Ole e 10	Aero Plant	NULL	29465664	AAL92578	123
1025	Olea europaea var. sylvestris	Wild olive	n/a	Aero Plant	NULL	37725369	AAO33896	38
1026	Onchocerca cervicalis	Parasitic nematode	n/a	Worm (parasite)	NULL	4102953	AAD01627	133
1027	Onchocerca volvulus	Parasitic nematode	n/a	Worm (parasite)	Activation-associated secreted protein-1	5868902	AAB69625	220
1028	Onchocerca volvulus	Parasitic nematode	n/a	Worm (parasite)	Beta-galactoside-binding lectin	433317	AAA20541	280
1029	Onchocerca volvulus	Parasitic nematode	n/a	Worm (parasite)	Myosin-like protein	159881	AAA29414	343
1030	Onchocerca volvulus	Parasitic nematode	n/a	Worm (parasite)	Activation-associated secreted protein-2	11762066	AAG40311	224
1031	Onchocerca volvulus	Parasitic nematode	n/a	Worm (parasite)	Vespid homolog	2796175	AAB97282	220
1032	Oryza sativa	Rice	n/a	Aero Plant	Beta-expansin	8118430	AAF72987	275
1033	Oryza sativa	Rice	n/a	Aero Plant	Beta-expansin	8118425	AAF72985	286
1034	Oryza sativa	Rice	n/a	Food Plant	RA17 (Trypsin alpha-amylase inhibitor)	100691	S21157	162
1035	Oryza sativa	Rice	n/a	Food Plant	Trypsin alpha-amylase inhibitor (frag)	7438309	T02664	157
1036	Oryza sativa	Rice	n/a	Aero Plant	Beta-expansin	11346546	AAF72989	271
1037	Oryza sativa	Rice	n/a	Aero Plant	Beta-expansin	8118437	AAF72990	269
1038	Oryza sativa	Rice	n/a	Food Plant	Fragment	7438312	T03966	113
1039	Oryza sativa	Rice	n/a	Food Plant	Fragment (Trypsin alpha-amylase inhibitor)	7438311	T03965	109
1040	Oryza sativa	Rice	n/a	Food Plant	Trypsin alpha-amylase inhibitor (frag)	7438310	T03963	111
1041	Oryza sativa	Rice	n/a	Aero Plant	Beta-expansin	2224915	AAB61710	261
1042	Oryza sativa	Rice	n/a	Aero Plant	NULL	32363197	P83466	15
1043	Oryza sativa	Rice	n/a	Aero Plant	Beta-expansin	8118423	AAF72984	268
1044	Oryza sativa	Rice	n/a	Aero Plant	Beta-expansin	8118432	AAF72988	327
1045	Oryza sativa	Rice	Ory s 1	Aero Plant	Expansin-like protein	2498586	Q40638	263
1046	Oryza sativa	Rice	n/a	Food Plant	RA14 (Trypsin alpha-amylase inhibitor)	419798	S31080	165

1047	Oryza sativa	Rice	n/a	Food Plant	RA5 (Trypsin alpha-amylase inhibitor)	419799	S31078	157
1048	Oryza sativa	Rice	n/a	Food Plant	RA14C (Trypsin alpha-amylase inhibitor)	2118428	S59923	160
1049	Oryza sativa	Rice	n/a	Food Plant	RA5B (Trypsin alpha-amylase inhibitor)	2118429	S59925	160
1050	Oryza sativa	Rice	n/a	Food Plant	RA16 (Trypsin alpha-amylase inhibitor)	2118430	S59924	157
1051	Oryza sativa	Rice	n/a	Aero Plant	Beta-expansin	8118421	AAF72983	267
1052	Oryza sativa	Rice	n/a	Food Plant	RAG2 (Trypsin alpha-amylase inibitor)	419801	S31082	166
1053	Oryza sativa	Rice	n/a	Aero Plant	Beta-expansin	8118439	AAF72991	267
1054	Oryza sativa	Rice	n/a	Food Plant	Trypsin alpha-amylase inhibitor (frag)	2118427	S59922	166
1055	Oryza sativa	Rice	n/a	Aero Plant	Beta-expansin	8118428	AAF72986	275
1056	Oryza sativa (japonica cultivar-group)	Rice	n/a	Food Plant	Alpha-expansin-like protein	6069656	BAA85432	284
1057	Oryza sativa (japonica cultivar-group)	Rice	n/a	Unassigned	NULL	53791944	BAD54206	206
1058	Oryza sativa (japonica cultivar-group)	Rice	n/a	Food Plant	Trypsin alpha-amylase inhibitor (frag)	1398913	BAA07710	166
1059	Oryza sativa (japonica cultivar-group)	Rice	n/a	Aero Plant	Beta expansin-like protein	20502989	AAM22698	289
1060	Oryza sativa (japonica cultivar-group)	Rice	n/a	Food Plant	Trypsin alpha-amylase inhibitor (frag)	1398916	BAA07712	157
1061	Oryza sativa (japonica cultivar-group)	Rice	n/a	Food Plant	Trypsin alpha-amylase inhibitor (frag)	1398918	BAA07713	160
1062	Oryza sativa (japonica cultivar-group)	Rice	n/a	Unassigned	NULL	53791938	BAD54200	206
1063	Oryza sativa (japonica cultivar-group)	Rice	n/a	Aero Plant	NULL	52075924	BAD45870	117
1064	Oryza sativa (japonica cultivar-group)	Rice	n/a	Food Plant	Trypsin alpha-amylase inhibitor (frag)	1398915	BAA07711	160
1065	Oryza sativa (japonica cultivar-group)	Rice	n/a	Aero Plant	NULL	52075917	BAD45863	117
1066	Oryza sativa (japonica cultivar-group)	Rice	n/a	Aero Plant	NULL	52075915	BAD45861	117
1067	Oryza sativa (japonica cultivar-group)	Rice	n/a	Aero Plant	NULL	52075914	BAD45860	113
1068	Oryza sativa (japonica cultivar-group)	Rice	n/a	Aero Plant	NULL	52075913	BAD45859	117
1069	Oryza sativa (japonica cultivar-group)	Rice	n/a	Food Plant	NULL	45736177	BAD13223	269
1070	Oryza sativa (japonica cultivar-group)	Rice	n/a	Aero Plant	NULL	52075910	BAD45856	118

1071	Oryza sativa (japonica cultivar-group)	Rice	n/a	Aero Plant	NULL	52075909	BAD45855	118
1072	Oryza sativa (japonica cultivar-group)	Rice	n/a	Unassigned	NULL	52077087	BAD46118	158
1073	Oryza sativa (japonica cultivar-group)	Rice	n/a	Unassigned	NULL	52077086	BAD46117	146
1074	Oryza sativa (japonica cultivar-group)	Rice	n/a	Food Plant	Glyoxalase I	16580747	BAB71741	291
1075	Oryza sativa (japonica cultivar-group)	Rice	n/a	Unassigned	NULL	50252014	BAD27946	204
1076	Oryza sativa (japonica cultivar-group)	Rice	n/a	Aero Plant	NULL	50251728	BAD27648	155
1077	Oryza sativa (japonica cultivar-group)	Rice	n/a	Unassigned	NULL	50251668	BAD29692	180
1078	Oryza sativa (japonica cultivar-group)	Rice	n/a	Aero Plant	NULL	51536378	BAD37571	117
1079	Oryza sativa (japonica cultivar-group)	Rice	n/a	Food Plant	Trypsin alpha-amylase inhibitor (frag)	1304218	BAA07774	113
1080	Oryza sativa (japonica cultivar-group)	Rice	n/a	Food Plant	Trypsin alpha-amylase inhibitor (frag)	218193	BAA01998	165
1081	Oryza sativa (japonica cultivar-group)	Rice	n/a	Food Plant	Trypsin alpha-amylase inhibitor (frag)	2827316	AAB99797	157
1082	Oryza sativa (japonica cultivar-group)	Rice	n/a	Unassigned	NULL	20146453	BAB89233	364
1083	Oryza sativa (japonica cultivar-group)	Rice	n/a	Unassigned	NULL	15624049	BAB68102	208
1084	Oryza sativa (japonica cultivar-group)	Rice	n/a	Aero Plant	NULL	38175617	BAD01325	384
1085	Oryza sativa (japonica cultivar-group)	Rice	n/a	Aero Plant	NULL	45736119	BAD13150	82
1086	Oryza sativa (japonica cultivar-group)	Rice	n/a	Unassigned	NULL	57900264	BAD87082	139
1087	Oryza sativa (japonica cultivar-group)	Rice	n/a	Food Plant	NULL	21104587	BAB93180	251
1088	Oryza sativa (japonica cultivar-group)	Rice	n/a	Aero Plant	NULL	53792960	BAD54134	175
1089	Oryza sativa (japonica cultivar-group)	Rice	n/a	Food Plant	Trypsin alpha-amylase inhibitor (frag)	1304216	BAA07772	111
1090	Oryza sativa (japonica cultivar-group)	Rice	n/a	Unassigned	NULL	42408062	BAD09204	430
1091	Oryza sativa (japonica cultivar-group)	Rice	n/a	Food Plant	RA17 (Trypsin alpha-amylase inhibitor)	34900094	NP_911393	163
1092	Oryza sativa (japonica cultivar-group)	Rice	n/a	Food Plant	RAG2 (Trypsin alpha-amylase inibitor)	34900098	NP_911395	166
1093	Oryza sativa (japonica	Rice	n/a	Food Plant	RA16 (Trypsin alpha-amylase	34900104	NP_911398	157

	cultivar-group)				inhibitor)			
1094	Oryza sativa (japonica cultivar-group)	Rice	n/a	Food Plant	RA5B (Trypsin alpha-amylase inhibitor)	34900132	NP_911412	160
1095	Oryza sativa (japonica cultivar-group)	Rice	n/a	Aero Plant	Beta expansin-like protein	37536544	NP_922574	275
1096	Oryza sativa (japonica cultivar-group)	Rice	n/a	Aero Plant	NULL	19386815	BAB86193	185
1097	Oryza sativa (japonica cultivar-group)	Rice	n/a	Food Plant	Trypsin alpha-amylase inhibitor (frag)	218195	BAA01997	162
1098	Oryza sativa (japonica cultivar-group)	Rice	n/a	Unassigned	NULL	53792800	BAD53834	207
1099	Oryza sativa (japonica cultivar-group)	Rice	n/a	Food Plant	Trypsin alpha-amylase inhibitor (frag)	218197	BAA01996	157
1100	Oryza sativa (japonica cultivar-group)	Rice	n/a	Unassigned	NULL	49388537	BAD25659	207
1101	Oryza sativa (japonica cultivar-group)	Rice	n/a	Food Plant	Trypsin alpha-amylase inhibitor (frag)	1304217	BAA07773	109
1102	Oryza sativa (japonica cultivar-group)	Rice	n/a	Food Plant	NULL	23616947	BAC20650	160
1103	Oryza sativa (japonica cultivar-group)	Rice	n/a	Food Plant	Glyoxalase I	4126809	BAA36759	291
1104	Oryza sativa (japonica cultivar-group)	Rice	n/a	Aero Plant	NULL	54291022	BAD61700	313
1105	Oryza sativa (japonica cultivar-group)	Rice	n/a	Unassigned	NULL	42408066	BAD09208	431
1106	Panulirus stimpsoni	Lobster	n/a	Food Animal	Tropomyosin	14285797	O61379	274
1107	Parietaria judaica	Weed	Par j 1	Aero Plant	Lipid transfer protein 1	741844	2008179A	143
1108	Parietaria judaica	Weed	Par j 1.0101	Aero Plant	Lipid transfer protein 1	3915783	P43217	139
1109	Parietaria judaica	Weed	Par j 1.0201	Aero Plant	Lipid transfer protein 1	2497749	Q40905	138
1110	Parietaria judaica	Weed	Par j 3	Aero Plant	Profilin 1	14423876	Q9XG85	132
1111	Parietaria judaica	Weed	Par j 3	Aero Plant	Profilin 2	14423869	Q9T0M8	131
1112	Parietaria judaica	Weed	Par j 2.0102	Aero Plant	Lipid transfer protein 2	1532056	CAA65122	133
1113	Parietaria judaica	Weed	Par j 1.0102	Aero Plant	Lipid transfer protein 1	1532058	CAA65123	176
1114	Parietaria judaica	Weed	Par j 2.0101	Aero Plant	Lipid transfer protein 2	2497750	P55958	133
1115	Parietaria officinalis	Weed	n/a	Aero Plant	NULL	1836010	AAB46819	25
1116	Parietaria officinalis	Weed	Par o 1	Aero Plant	Lipid transfer protein	1311509	AAB36008	17
1117	Parietaria officinalis	Weed	Par o 1	Aero Plant	Lipid transfer protein	1311510	AAB36009	15
1118	Parietaria officinalis	Weed	Par o 1	Aero Plant	Lipid transfer protein	1836011	AAB46820	24
1119	Parietaria officinalis	Weed	Par o 1	Aero Plant	Lipid transfer protein	543659	A53252	12
1120	Parietaria officinalis	Weed	Par o 1	Aero Plant	Lipid transfer protein	1311512	AAB36011	15
1121	Parietaria officinalis	Weed	Par o 1	Aero Plant	Lipid transfer protein	1311513	AAB36012	30
1122	Parietaria officinalis	Weed	Par o 1	Aero Plant	Lipid transfer protein	1311511	AAB36010	15
1123	Penaeus monodon	Black tiger shrimp	Pen m 2	Food Animal	Arginine kinase	27463265	AAO15713	356
1124	Penicillium chrysogenum	Fungus	Pen ch 13	Aero Fungi	Alkaline serine protease	6684758	AAF23726	397
1125	Penicillium chrysogenum	Fungus	Pen ch 18	Aero Fungi	Vacuolar serine protease	7963902	AAF71379	494

1126	Penicillium chrysogenum	Fungus	Pen ch 18	Aero Fungi	Vacuolar serine protease	14215732	AAG44693	494
1127	Penicillium chrysogenum	Fungus	Pen ch 13	Aero Fungi	Alkaline serine protease	21069093	AAM33821	398
1128	Penicillium chrysogenum	Fungus	n/a	Aero Fungi	68 kDa protein	999009	AAB34785	117
1129	Penicillium citrinum	Fungus	Pen c 22	Aero Fungi	Enolase	13991101	AAK51201	438
1130	Penicillium citrinum	Fungus	n/a	Aero Fungi	Vacuolar serine protease	12005501	AAG44480	358
1131	Penicillium citrinum	Fungus	Pen c 24	Aero Fungi	elongation factor 1 beta	38326693	AAR17475	228
1132	Penicillium citrinum	Fungus	Pen c 3	Aero Fungi	Peroxisomal membrane protein	5326864	AAD42074	167
1133	Penicillium citrinum	Fungus	n/a	Aero Fungi	Pen c 1	4587983	AAD25926	397
1134	Penicillium citrinum	Fungus	Pen c 19	Aero Fungi	Heat shock 70 kDa protein	14423733	Q92260	503
1135	Penicillium citrinum	Fungus	n/a	Aero Fungi	Alkaline serine protease	4588118	AAD25995	457
1136	Penicillium oxalicum	Fungus	Pen o 18	Aero Fungi	Vacuolar serine protease	12005497	AAG44478	503
1137	Periplaneta americana	American cockroach	Per a 3.0201	Aero Insect	Similar to hemolymph proteins	1531589	AAB09632	631
1138	Periplaneta americana	American cockroach	n/a	Aero Insect	Ribosomal protein S12	21217441	AAM33784	139
1139	Periplaneta americana	American cockroach	n/a	Aero Insect	Rab11	21217443	AAM33785	204
1140	Periplaneta americana	American cockroach	Per a 7.0102	Aero Insect	Tropomyosin	4468639	CAB38086	284
1141	Periplaneta americana	American cockroach	Per a 1.0101	Aero Insect		4240399	AAD13533	231
1142	Periplaneta americana	American cockroach	Per a 3.01	Aero Insect	Hemolymph-like protein	1580792	AAB09629	685
1143	Periplaneta americana	American cockroach	Per a 3.0202	Aero Insect	Hemolymph-like protein	1580794	AAB62731	470
1144	Periplaneta americana	American cockroach	Per a 3.0203	Aero Insect	Hemolymph-like protein	1580797	AAB63595	393
1145	Periplaneta americana	American cockroach	Per a 1.0102	Aero Insect		2897849	AAC34312	228
1146	Periplaneta americana	American cockroach	n/a	Aero Insect	NULL	30144660	AAP13554	124
1147	Periplaneta americana	American cockroach	Per a 1	Aero Insect	Cr-PII	2580504	AAB82404	395
1148	Periplaneta americana	American cockroach	Per a 1	Aero Insect	Cr-PII allergen	2231297	AAC34736	446
1149	Periplaneta americana	American cockroach	Per a 1.0104	Aero Insect	Cr-PII	2253610	AAC34737	274
1150	Periplaneta americana	American cockroach	Per a 7.0101	Aero Insect	Tropomyosin	4378573	AAD19606	284
1151	Periplaneta fuliginosa	Smokybrown cockroach	n/a	Aero Insect	NULL	19310971	AAL86701	284
1152	Perna viridis	Asian green mussell	n/a	Food Animal	Tropomyosin	9954251	AAG08988	284
1153	Persea americana	Avocado	Pers a 1	Food Plant	Endochitinase	3201547	CAB01591	326
1154	Petroselinum crispum	Parsley	n/a	Food Plant	Pathogenesis-related protein 1	1843451	CAA67246	155
1155	Phalaris aquatica	Canary grass	n/a	Aero Plant	Fragment	409328	AAB27445	20
1156	Phalaris aquatica	Canary grass	n/a	Aero Plant	Similar to Poa p 9/Phl p 6 allergens	2498578	P56165	305
1157	Phalaris aquatica	Canary grass	n/a	Aero Plant	Similar to Poa p 9/Phl p 6 allergens	2498577	P56164	320
1158	Phalaris aquatica	Canary grass	n/a	Aero Plant	Similar to Poa p 9/Phl p 6 allergens	2498579	P56166	294
1159	Phalaris aquatica	Canary grass	n/a	Aero Plant	Similar to Poa p 9/Phl p 6 allergens	2498580	P56167	175
1160	Phalaris aquatica	Canary grass	Pha a 1	Aero Plant	Expansin-like protein	2498576	Q41260	269
1161	Phaseolus vulgaris	Kidney bean	n/a	Food Plant	Pathogenesis-related protein 2 (PvPR2)	21048	CAA43636	155

1162	Phaseolus vulgaris	Kidney bean	n/a	Food Plant	Pathogenesis-related protein 1 (PvPR1)	21044	CAA43637	156
1163	Phlebotomus papatasi	Sand fly	n/a	Venom or Salivary	30 kDa salivary protein	15963513	AAL11049	253
1164	Phlebotomus papatasi	Sand fly	n/a	Venom or Salivary	28 kDa salivary protein	15963511	AAL11048	254
1165	Phleum pratense	Common timothy	Phl p 11	Aero Plant	Group 11	23452313	AAN32987	143
1166	Phleum pratense	Common timothy	n/a	Aero Plant	Group V	21725632	CAD38397	287
1167	Phleum pratense	Common timothy	n/a	Aero Plant	Group V	21725630	CAD38396	287
1168	Phleum pratense	Common timothy	n/a	Aero Plant	Group V	21725628	CAD38395	287
1169	Phleum pratense	Common timothy	n/a	Aero Plant	Group V	21725622	CAD38392	287
1170	Phleum pratense	Common timothy	n/a	Aero Plant	Group V	21725624	CAD38393	287
1171	Phleum pratense	Common timothy	n/a	Aero Plant	NULL	54144334	CAD54671	508
1172	Phleum pratense	Common timothy	n/a	Aero Plant	NULL	54144332	CAD54670	508
1173	Phleum pratense	Common timothy	Phl p 5	Aero Plant	Group V fragment	626037	A61505	24
1174	Phleum pratense	Common timothy	n/a	Aero Plant	Group V	21725620	CAD38391	287
1175	Phleum pratense	Common timothy	n/a	Aero Plant	Group V	21725618	CAD38390	287
1176	Phleum pratense	Common timothy	n/a	Aero Plant	Group V	21725616	CAD38389	287
1177	Phleum pratense	Common timothy	n/a	Aero Plant	Group V	21725614	CAD38388	287
1178	Phleum pratense	Common timothy	n/a	Aero Plant	Group V	21725612	CAD38387	287
1179	Phleum pratense	Common timothy	n/a	Aero Plant	Group V	21725610	CAD38386	287
1180	Phleum pratense	Common timothy	n/a	Aero Plant	Group V	21725608	CAD38385	287
1181	Phleum pratense	Common timothy	n/a	Aero Plant	Group V	21725606	CAD38384	287
1182	Phleum pratense	Common timothy	n/a	Aero Plant	Group V	21725626	CAD38394	287
1183	Phleum pratense	Common timothy	Phl p 1	Aero Plant	Group 1	28373838	1N10_A	241
1184	Phleum pratense	Common timothy	n/a	Aero Plant	Group VI	3004469	CAA76558	106
1185	Phleum pratense	Common timothy	Phl p 6	Aero Plant	Group VI	3004467	CAA76557	138
1186	Phleum pratense	Common timothy	Phl p 5.0105	Aero Plant	Group V	3135497	AAC16525	276
1187	Phleum pratense	Common timothy	Phl p 5.0106	Aero Plant	Group V	3135499	AAC16526	276
1188	Phleum pratense	Common timothy	Phl p 5.0107	Aero Plant	Group V	3135501	AAC16527	276
1189	Phleum pratense	Common timothy	n/a	Aero Plant	Group I	481432	S38620	262
1190	Phleum pratense	Common timothy	Phl p 5	Aero Plant	Group V	481397	S38584	280
1191	Phleum pratense	Common timothy	Phl p 6	Aero Plant	Group VI	3004465	CAA76556	138
1192	Phleum pratense	Common timothy	Phl p 1.0102	Aero Plant	Group I	473360	CAA55390	263
1193	Phleum pratense	Common timothy	n/a	Aero Plant	Group V	3309047	AAC25998	287
1194	Phleum pratense	Common timothy	Phl p 5.0108	Aero Plant	Group V	3135503	AAC16528	276
1195	Phleum pratense	Common timothy	Phl p 1.0101	Aero Plant	Group I	3901094	CAA81613	263
1196	Phleum pratense	Common timothy	Phl p 12	Aero Plant	Profilin 1	464471	P35079	131
1197	Phleum pratense	Common timothy	Phl p 5	Aero Plant	Similar to Poa p 9/Phl p 6 allergens	2851457	Q40963	284
1198	Phleum pratense	Common timothy	Phl p 5.0103	Aero Plant	Group V	3309039	AAC25994	312
1199	Phleum pratense	Common timothy	Phl p 5	Aero Plant	Fragment Group V	422005	S32101	257
1200	Phleum pratense	Common timothy	n/a	Aero Plant	Group V	3309041	AAC25995	295
1201	Phleum pratense	Common timothy	Phl p 2	Aero Plant	Group II	415896	CAA53529	122

1202	Phleum pratense	Common timothy	n/a	Aero Plant	Group V	3309045	AAC25997	290
1203	Phleum pratense	Common timothy	Phl p 5.0104	Aero Plant	Ribonuclease	1684720	CAB05372	276
1204	Phleum pratense	Common timothy	Phl p 5.0202	Aero Plant	Ribonuclease	1684718	CAB05371	281
1205	Phleum pratense	Common timothy	Phl p 5	Aero Plant	Group 1	28948464	1L3P_A	102
1206	Phleum pratense	Common timothy	n/a	Aero Plant	Polcalcin	14423846	O82040	78
1207	Phleum pratense	Common timothy	n/a	Aero Plant	NULL	45823012	CAG24374	240
1208	Phleum pratense	Common timothy	n/a	Aero Plant	NULL	39841264	AAR31142	97
1209	Phleum pratense	Common timothy	n/a	Aero Plant	Group V	1092249	2023228A	285
1210	Phleum pratense	Common timothy	n/a	Aero Plant	NULL	28374072	1NLX_N	111
1211	Phleum pratense	Common timothy	n/a	Aero Plant	NULL	29500897	CAD87529	284
1212	Phleum pratense	Common timothy	Phl p 12	Aero Plant	Profilin 4	2415702	CAA70610	131
1213	Phleum pratense	Common timothy	Phl p 5.0101	Aero Plant	Group I fragment	398830	CAA52753	312
1214	Phleum pratense	Common timothy	n/a	Aero Plant	NULL	45108967	CAF32566	500
1215	Phleum pratense	Common timothy	Phl p 5	Aero Plant	Group V	2398757	CAA50281	286
1216	Phleum pratense	Common timothy	Phl p 13	Aero Plant	Polygalacturonase	4826572	CAB42886	394
1217	Phleum pratense	Common timothy	Phl p 5	Aero Plant	Group V	13430402	AAK25823	275
1218	Phleum pratense	Common timothy	Phl p 12	Aero Plant	Profilin 3	2415700	CAA70609	131
1219	Phleum pratense	Common timothy	n/a	Aero Plant	NULL	45108973	CAF32567	500
1220	Phoenix dactylifera	Date palm	Pho d 2	Aero Plant	Profilin	21322677	CAD10390	131
1221	Pieris rapae	Cabbage white	n/a	Unassigned	NULL	40288346	AAR84202	632
1222	Pinus radiata	Monterey pine	n/a	Aero Plant	NULL	7484482	T08115	293
1223	Pisum sativum	Pea	Pis s 1	Food Plant	NULL	42414627	CAF25232	415
1224	Pisum sativum	Pea	Pis s 1	Food Plant	NULL	42414629	CAF25233	415
1225	Pisum sativum	Pea	n/a	Aero Plant	NULL	1076531	S53082	258
1226	Plantago lanceolata	Narrow-leaved plantain	Pla l 1	Aero Plant	unnamed protein	14422359	CAC41633	131
1227	Plantago lanceolata	Narrow-leaved plantain	Pla l 1	Aero Plant	unnamed protein	14422361	CAC41634	131
1228	Plantago lanceolata	Narrow-leaved plantain	Pla l 1	Aero Plant	unnamed protein	14422363	CAC41635	131
1229	Plantago lanceolata	Narrow-leaved plantain	n/a	Aero Plant	unnamed protein	29163773	CAD80019	65
1230	Plantago lanceolata	Narrow-leaved plantain	n/a	Aero Plant	unnamed protein	29163771	CAD80018	65
1231	Plasmodium falciparum 3D7	Parasitic protozoa	n/a	Protozoan	NULL	23505219	CAD51999	171
1232	Platanus x acerifolia	London plane tree	Pla a 1	Aero Plant	Putative invertase inhibitor	26190140	CAD20556	179
1233	Platanus x acerifolia	London plane tree	Pla a 2	Aero Plant	NULL	49523394	CAE52833	377
1234	Pleospora herbarum	Fungus	n/a	Aero Fungi	NULL	49476469	AAT66568	137
1235	Pleospora tarda	Fungus	n/a	Aero Fungi	NULL	49476463	AAT66565	137
1236	Plodia interpunctella	Indian meal moth	n/a	Aero Insect	Arginine kinase	15886861	CAC85911	355
1237	Plodia interpunctella	Indian meal moth	n/a	Aero Insect	unnamed protein	21335404	CAD33268	285
1238	Plodia interpunctella	Indian meal moth	n/a	Aero Insect	unnamed protein	21335408	CAD33270	254
1239	Plodia interpunctella	Indian meal moth	n/a	Aero Insect	unnamed protein	21335406	CAD33269	705

1240	Poa pratensis	Kentucky bluegrass	Poa p 5	Aero Plant		11991227	AAG42254	303
1241	Poa pratensis	Kentucky bluegrass	n/a	Aero Plant	Clone 60	100733	B39098	307
1242	Poa pratensis	Kentucky bluegrass	n/a	Aero Plant	Clone 31	483123	C39098	373
1243	Poa pratensis	Kentucky bluegrass	n/a	Aero Plant	Clone 7.2	539056	A60373	131
1244	Poa pratensis	Kentucky bluegrass	Poa p 1	Aero Plant	Fragment	280414	A60372	20
1245	Poa pratensis	Kentucky bluegrass	n/a	Aero Plant	Clone 41	100732	A39098	333
1246	Poa pratensis	Kentucky bluegrass	n/a	Aero Plant	Ribonuclease	113562	P22286	307
1247	Poa pratensis	Kentucky bluegrass	Poa p 1	Aero Plant	Group 1	4090265	CAA10520	263
1248	Poa pratensis	Kentucky bluegrass	Poa p 1	Aero Plant	Fragment	320620	F37396	26
1249	Poa pratensis	Kentucky bluegrass	n/a	Aero Plant	Group 2	4007655	CAA10348	122
1250	Polistes annularis	Paper wasp	Pol a 5	Venom or Salivary	Venom allergen 5	160780	AAA29793	209
1251	Polistes annularis	Paper wasp	Pol a 2	Venom or Salivary	Hyaluronidase	14423735	Q9U6V9	367
1252	Polistes annularis	Paper wasp	Pol a 1	Venom or Salivary	Phospholipase A1	14423833	Q9U6W0	301
1253	Polistes dominulus	Paper wasp	n/a	Venom or Salivary	NULL	45510891	AAS67043	316
1254	Polistes dominulus	Paper wasp	n/a	Venom or Salivary	Venom serine protease	30909091	AAP37412	277
1255	Polistes dominulus	Paper wasp	n/a	Venom or Salivary	NULL	45510889	AAS67042	316
1256	Polistes dominulus	Paper wasp	n/a	Venom or Salivary	NULL	45510893	AAS67044	316
1257	Polistes dominulus	Paper wasp	n/a	Venom or Salivary	NULL	45510887	AAS67041	337
1258	Polistes dominulus	Paper wasp	Pol d 5	Venom or Salivary	NULL	51093377	AAT95010	227
1259	Polistes exclamans	Paper wasp	n/a	Venom or Salivary	NULL	51093375	AAT95009	226
1260	Polistes exclamans	Paper wasp	Pol e 5	Venom or Salivary	Venom allergen 5	549187	P35759	205
1261	Polistes fuscatus	Paper wasp	Pol f 5	Venom or Salivary	Venom allergen 5	627190	F44583	205
1262	Polistes gallicus	Paper wasp	Pol g 5	Venom or Salivary	NULL	25091511	P83377	206
1263	Polistes gallicus	Paper wasp	n/a	Venom or Salivary	NULL	41017429	P83542	42
1264	Polybia scutellaris	Wasp	n/a	Venom or Salivary	NULL	31747352	AAP57536	206
1265	Polybia scutellaris rioplatensis	Wasp	n/a	Venom or Salivary	NULL	47117356	Q7Z156	207
1266	Polypedilum nubifer	Chironomid	n/a	Aero Insect	Hemoglobin fragment	1703681	AAB38011	37
1267	Prunus armeniaca	Apricot	Pru ar 1	Food Plant	PRP (Bet v 1 family)	2677826	AAB97141	160
1268	Prunus armeniaca	Apricot	Pru ar 3	Food Plant	Lipid-transfer protein	7404406	P81651	91

1269	Prunus armeniaca	Apricot	n/a	Food Plant	NULL	4887129	AAD32205	168
1270	Prunus avium	Cherry	Pru av 3	Food Plant	Lipid transfer protein	6715520	AAF26449	117
1271	Prunus avium	Cherry	Pru av 1	Food Plant	NULL	44409496	AAS47037	160
1272	Prunus avium	Cherry	Pru av 1	Food Plant	PRP (Bet v 1 family)	1513216	AAC02632	160
1273	Prunus avium	Cherry	n/a	Food Plant	NULL	13787043	1E09_A	159
1274	Prunus avium	Cherry	Pru av 2	Food Plant	Thaumatin-like protein	1144346	AAB38064	245
1275	Prunus avium	Cherry	Pru av 1	Food Plant	NULL	44409474	AAS47036	160
1276	Prunus avium	Cherry	Pru av 4	Food Plant	Profilin	4761582	AAD29411	131
1277	Prunus avium	Cherry	Pru av 1	Food Plant	NULL	44409451	AAS47035	160
1278	Prunus avium	Cherry	Pru av 1	Food Plant	PRP (Bet v 1 family)	34809853	1H2O_A	159
1279	Prunus domestica	Plum	Pru d 3	Food Plant	Nonspecific lipid-transfer protein 1	9297015	P82534	91
1280	Prunus dulcis	Almond	Pru du 4	Food Plant	Profilin	24473794	AAL91662	131
1281	Prunus persica	Peach	Pru p 4.01	Food Plant	Profilin	27528310	CAD37201	131
1282	Prunus persica	Peach	Pru p 4.02	Food Plant	Profilin	27528312	CAD37202	131
1283	Prunus persica	Peach	Pru p 3	Food Plant	Nonspecific lipid-transfer protein 1	17974195	CAB96876	91
1284	Prunus persica	Peach	Pru p 3	Food Plant	Nonspecific lipid-transfer protein 1	3287877	P81402	91
1285	Psoroptes ovis	Sheep mite	n/a	Venom or Salivary	NULL	6625558	AAF19264	219
1286	Psoroptes ovis	Sheep mite	n/a	Venom or Salivary	Der p 1 homolog	27450737	AAO14671	263
1287	Psoroptes ovis	Sheep mite	n/a	Venom or Salivary	Allergen A	14388619	AAK61827	143
1288	Pyrus communis	Pear	Pyr c 5	Food Plant	Isoflavone reductase related protein	3243234	AAC24001	308
1289	Pyrus communis	Pear	n/a	Food Plant	Nonspecific lipid-transfer protein	14423813	Q9M5X6	115
1290	Pyrus communis	Pear	Pyr c 4	Food Plant	Profilin	4761580	AAD29410	131
1291	Pyrus communis	Pear	Pyr c 1	Food Plant	Similar to Bet v 2 protein	3044216	AAC13315	159
1292	Quercus alba	Oak	Que a 1	Aero Plant	Fragment	543675	D53288	24
1293	Rana esculenta	Frog	Ran e 1	Food Animal	Parvalbumin alpha	20796729	CAC83046	110
1294	Rana esculenta	Frog	Ran e 2	Food Animal	Parvalbumin beta	20797081	CAC95152	109
1295	Rana sp. CH-2001	Frog	n/a	Food Animal	Parvalbumin beta	20797085	CAC95153	109
1296	Rana sp. CH-2001	Frog	n/a	Food Animal	Parvalbumin alpha	20796733	CAC83047	110
1297	Rattus norvegicus	Rat	n/a	Aero Animal	Epididymal secretory protein 1	27465565	NP_775141	149
1298	Rattus norvegicus	Rat	Rat n 1	Aero Animal	Alpha-2u-globulin	111366	S05440	179
1299	Rattus norvegicus	Rat	n/a	Aero Animal	NULL	62649086	XP_578456	181
1300	Rattus norvegicus	Rat	n/a	Aero Animal	Similar to lacrimal gland protein	34855801	XP_218497	127
1301	Rattus norvegicus	Rat	Rat n 1	Aero Animal	Alpha-2u-globulin	127533	P02761	181
1302	Rhodotorula mucilaginosa	Fungus	n/a	Unassigned	NULL	37078092	Q870B9	439

1303	Ricinus communis	Castor bean	Ric c 1	Food Plant	2S albumin	112762	P01089	258
1304	Saccharomyces cerevisiae	Baker's yeast	n/a	Aero Fungi	NULL	15988101	1K0K_A	125
1305	Salmo salar	Salmon	Sal s 1	Food Animal	Parvalbumin beta 1	2493445	Q91482	109
1306	Salmo salar	Salmon	Sal s 1	Food Animal	Parvalbumin beta 2	18281421	Q91483	108
1307	Salsola kali	Thistle	n/a	Aero Plant	NULL	25090948	P83181_1	11
1308	Salsola kali	Thistle	Sal k 1	Aero Plant		22726221	AAN05083	320
1309	Salsola kali	Thistle	n/a	Aero Plant	NULL	25090950	P83181_3	9
1310	Salsola kali	Thistle	n/a	Aero Plant	NULL	25090951	P83181_4	14
1311	Salsola kali	Thistle	n/a	Aero Plant	NULL	25090949	P83181_2	8
1312	Sambucus nigra	European elder	n/a	Aero Plant		6561156	AAF16869	159
1313	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Unnamed protein	27462848	AAO15613	330
1314	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Cathepsin L-like protease	27462834	AAO15606	245
1315	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	NULL	27462842	AAO15610	131
1316	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Trypsin-like serine protease	38202329	AAR14089	248
1317	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Trypsin-like serine protease	38202317	AAR14083	251
1318	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Trypsin-like serine protease	38202313	AAR14081	260
1319	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	NULL	27462840	AAO15609	157
1320	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Trypsin-like serine protease	38202319	AAR14084	264
1321	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	NULL	27462838	AAO15608	275
1322	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Trypsin-like serine protease	38202321	AAR14085	256
1323	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Trypsin-like serine protease	38202323	AAR14086	251
1324	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Vitellogenin-like protein	27462844	AAO15611	174
1325	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Trypsin-like serine protease	38202327	AAR14088	259
1326	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Trypsin-like serine protease	38202315	AAR14082	271
1327	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Trypsin-like serine protease	38202345	AAR14097	261
1328	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Trypsin-like serine protease	38202333	AAR14091	257
1329	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Trypsin-like serine protease	38202335	AAR14092	250

1330	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Trypsin-like serine protease	38202337	AAR14093	259
1331	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Trypsin-like serine protease	38202339	AAR14094	259
1332	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Trypsin-like serine protease	38202341	AAR14095	263
1333	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Trypsin-like serine protease	38202343	AAR14096	233
1334	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Trypsin-like serine protease	38202331	AAR14090	262
1335	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Paramyosin	27462846	AAO15612	638
1336	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Trypsin-like serine protease	38202325	AAR14087	242
1337	Sarcoptes scabiei type hominis	Scabies mite	n/a	Venom or Salivary	Glutathione S-transferase	27462836	AAO15607	219
1338	Schedonorus arundinaceus	Tall fescue	n/a	Aero Plant	Type A fragment	320610	C37396	17
1339	Schedonorus arundinaceus	Tall fescue	n/a	Aero Plant	Type B fragment	320611	D37396	20
1340	Schedonorus arundinaceus	Tall fescue	n/a	Aero Plant	Group I fragment	7489364	A58493	35
1341	Schistosoma japonicum	Schistosoma	n/a	Protozoan	22.6 kDa tegumental antigen	2739154	AAC67308	191
1342	Schistosoma japonicum	Schistosoma	n/a	Protozoan	Putative profilin	29841461	AAP06493	129
1343	Schizophyllum commune	Fungus	n/a	Unassigned	NULL	386678	AAA16208	214
1344	Schizophyllum commune	Fungus	n/a	Unassigned	NULL	169865	AAA16207	204
1345	Scomber japonicus	Chub mackerel	n/a	Food Animal	Parvalbumin	29420793	BAC66618	109
1346	Secale cereale	Rye	n/a	Aero Plant	NULL	55859454	CAH92627	518
1347	Secale cereale	Rye	n/a	Aero Plant	NULL	55859456	CAH92630	520
1348	Secale cereale	Rye	n/a	Aero Plant	Fragment	2130094	S65604	26
1349	Secale cereale	Rye	n/a	Aero Plant	30 kDa fragment	542172	S38292	16
1350	Sesamum indicum	Sesame	Ses i 2	Food Plant	2S albumin	5381323	AAD42943	148
1351	Sesamum indicum	Sesame	Ses i 1	Food Plant	2S albumin	13183175	AAK15088	153
1352	Sesamum indicum	Sesame	Ses i 3	Food Plant	7S globulin	13183177	AAK15089	585
1353	Setaria cervi	Nematode	n/a	Worm (parasite)	Ladder protein	4102957	AAD01629	133
1354	Sinapis alba	White mustard	Sin a 1	Food Plant	2S Albumin	1009440	CAA62912	145
1355	Sinapis alba	White mustard	Sin a 1	Food Plant	2S Albumin	7545129	AAB25214	145
1356	Sinapis alba	White mustard	Sin a 1	Food Plant	2S Albumin	1009442	CAA62908	145
1357	Sinapis alba	White mustard	Sin a 1	Food Plant	2S Albumin	1009438	CAA62911	145
1358	Sinapis alba	White mustard	Sin a 1	Food Plant	2S Albumin	1009436	CAA62910	145
1359	Sinapis alba	White mustard	Sin a 1	Food Plant	2S Albumin	1009434	CAA62909	145
1360	Sinapis alba	White mustard	Sin a 1	Food Plant	2S Albumin	322674	PC1247	145
1361	Solanum tuberosum	Potato	Sola t 1	Food Plant	Patatin B1	129641	P15476	377
1362	Solanum tuberosum	Potato	n/a	Food Plant	Aspartic proteinase inhibitor (Sol t 4)	21413	CAA45723	217

1363	Solanum tuberosum	Potato	n/a	Food Plant	Patatin B2	21510	CAA31575	386
1364	Solanum tuberosum	Potato	n/a	Food Plant	Patatin (B2)	21512	CAA27571	386
1365	Solanum tuberosum	Potato	n/a	Food Plant	Patatin (B2)	21514	CAA27588	386
1366	Solanum tuberosum	Potato	Sola t 2	Food Plant	Aspartic protease inhibitor 11	124148	P16348	188
1367	Solanum tuberosum	Potato	Sola t 3	Food Plant	Cysteine protease inhibitor	20141344	P20347	222
1368	Solanum tuberosum	Potato	n/a	Food Plant	Patatin (B2)	169500	AAA33819	386
1369	Solanum tuberosum	Potato	Sola t 4	Food Plant	Serine protease inhibitor 7	20141714	P30941	221
1370	Solenopsis geminata	Fire ant	Sol g 4	Venom or Salivary	Venom allergen	7638028	AAF65312	137
1371	Solenopsis geminata	Fire ant	Sol g 4	Venom or Salivary	Venom allergen	7638030	AAF65313	137
1372	Solenopsis invicta	Red fire ant	Sol i 4	Venom or Salivary	Venom allergen IV	112559	C37330	117
1373	Solenopsis invicta	Red fire ant	n/a	Venom or Salivary	NULL	1336813	AAB36121	26
1374	Solenopsis invicta	Red fire ant	n/a	Venom or Salivary	NULL	51093373	AAT95008	346
1375	Solenopsis invicta	Red fire ant	Sol i 4	Venom or Salivary	Venom allergen IV	14424465	P35777	137
1376	Solenopsis invicta	Red fire ant	Sol i 2	Venom or Salivary	Venom allergen II	1079186	A37330	138
1377	Solenopsis invicta	Red fire ant	Sol i 3	Venom or Salivary	Venom allergen III	112558	B37330	212
1378	Solenopsis invicta	Red fire ant	Sol i 4	Venom or Salivary	Venom allergen	4038411	AAC97370	137
1379	Solenopsis invicta	Red fire ant	Sol i 3	Venom or Salivary	Venom allergen III	14424466	P35778	234
1380	Solenopsis invicta	Red fire ant	n/a	Venom or Salivary	NULL	1336809	AAB36117	58
1381	Solenopsis invicta	Red fire ant	n/a	Venom or Salivary	NULL	1336811	AAB36119	25
1382	Solenopsis invicta	Red fire ant	n/a	Venom or Salivary	NULL	1336812	AAB36120	26
1383	Solenopsis richteri	Black fire ant	n/a	Venom or Salivary	Phospholipase fragment	321043	E60727	20
1384	Solenopsis richteri	Black fire ant	n/a	Venom or Salivary	Venom allergen III fragment	321044	D60727	20
1385	Solenopsis richteri	Black fire ant	n/a	Venom or Salivary	Venom allergen	7512067	A58853	119
1386	Solenopsis richteri	Black fire ant	n/a	Venom or Salivary	Venom allergen III	6136163	P35779	211
1387	Stemphylium callistephi	Fungus	n/a	Aero Fungi	NULL	49476467	AAT66567	137
1388	Stemphylium vesicarium	Fungus	n/a	Aero Fungi	NULL	49476465	AAT66566	137
1389	Strongyloides stercoralis	Parasitic nematode	n/a	Worm (parasite)		2290394	AAB65142	144
1390	Strongyloides stercoralis	Parasitic nematode	n/a	Worm (parasite)		2290392	AAB65141	128

1391	Strongyloides stercoralis	Parasitic nematode	n/a	Worm (parasite)		2290390	AAB65140	90
1392	Strongyloides stercoralis	Parasitic nematode	n/a	Worm (parasite)		2290388	AAB65139	152
1393	Strongyloides stercoralis	Parasitic nematode	n/a	Worm (parasite)		2801529	AAB97359	156
1394	Strongyloides stercoralis	Parasitic nematode	n/a	Worm (parasite)		2801531	AAB97360	95
1395	Styela plicata	Sea squirt	n/a	Food Animal	NULL	58257626	BAD88796	126
1396	Stylonychia lemnae	Ciliate	n/a	Protozoan	Unnamed protein	10142	CAA34151	104
1397	Syringa vulgaris	Lilac	Syr v I.0101	Aero Plant	Isoform 1	631911	S43242	145
1398	Syringa vulgaris	Lilac	Syr v I.0103	Aero Plant	Isoform 3	631913	S43244	145
1399	Syringa vulgaris	Lilac	Syr v I.0102	Aero Plant	Isoform 2	631912	S43243	145
1400	Syringa vulgaris	Lilac	n/a	Aero Plant	Polycalcin	14423847	P58171	81
1401	Taraxacum officinale	Dandelion	n/a	Food Plant	Root protein (Bet v 1 homologue)	2707295	AAB92255	157
1402	Tenebrio molitor	Yellow mealworm	n/a	Unassigned	NULL	32967475	AAP92419	595
1403	Thaumetopoea pityocampa	Pine moth	Tha p 1	Aero Insect	Fragment	25528311	A59396	18
1404	Theragra chalcogramma	Alaska pollock	n/a	Food Animal	NULL	14531020	AAK63089	109
1405	Theragra chalcogramma	Alaska pollock	n/a	Food Animal	NULL	14531018	AAK63088	109
1406	Tityus serrulatus	Brazilian scorpion	n/a	Venom or Salivary	NULL	1173399	P45669	62
1407	Toxocara canis	Dog roundworm	n/a	Worm (parasite)	Ladder protein	4102955	AAD01628	133
1408	Toxocara canis	Dog roundworm	n/a	Worm (parasite)	TBA-1	1731859	BAA14015	140
1409	Toxocara canis	Dog roundworm	n/a	Worm (parasite)	TBA-1 fragment	477888	B49139	36
1410	Triatoma protracta	Western conenose	Tria p 1	Venom or Salivary	Procalin	15426413	AAF07903	169
1411	Trichophyton rubrum	Fungus	Tri r 2	Contact	Subtilase-like protein	5813790	AAD52013	412
1412	Trichophyton rubrum	Fungus	Tri r 4	Contact	Prolyl oligopeptidase-like protein	5813788	AAD52012	726
1413	Trichophyton schoenleinii	Fungus	n/a	Contact		23894227	CAD23374	726
1414	Trichophyton schoenleinii	Fungus	n/a	Contact		23894260	CAD23618	405
1415	Trichophyton tonsurans	Fungus	n/a	Unassigned	NULL	1708296	P80514	26
1416	Trichostrongylus colubriformis	Nematode	n/a	Worm (parasite)	Aspartyl protease inhibitor	28274792	AAO34715	228
1417	Triticum aestivum	Wheat	n/a	Gliadin	Gliadin	21757	CAA26383	296
1418	Triticum aestivum	Wheat	n/a	Gliadin	Pre-alpha-/beta gliadin A-V	170716	AAA34278	319
1419	Triticum aestivum	Wheat	n/a	Gliadin	Gamma gliadin B-III	170734	AAA34287	244
1420	Triticum aestivum	Wheat	n/a	Gliadin	Gamma gliadin	170738	AAA34289	327
1421	Triticum aestivum	Wheat	n/a	Food Plant	LMW glutenin	7442124	T05923	374
1422	Triticum aestivum	Wheat	n/a	Gliadin	Alpha gliadin	1304264	BAA12318	259
1423	Triticum aestivum	Wheat	n/a	Food Plant	LMW glutenin	7442133	T05910	286
1424	Triticum aestivum	Wheat	n/a	Food Plant	Hypothetical 11.8 k protein	100794	B27319	106
1425	Triticum aestivum	Wheat	n/a	Food Plant	HMW glutenin subunit A	170743	AAB02788	815
1426	Triticum aestivum	Wheat	n/a	Food Plant	Gliadin omega-5 (fragment exercise-induced anaphyl	11277193	A59156	32
1427	Triticum aestivum	Wheat	n/a	Aero Plant	NULL	3913017	P81496	27

1428	Triticum aestivum	Wheat	n/a	Aero Plant	Putative group V allergen	40644794	CAE53905	119
1429	Triticum aestivum	Wheat	n/a	Food Plant	Triosephosphate isomerase	11124572	CAC14917	253
1430	Triticum aestivum	Wheat	n/a	Aero Plant	Bakers asthma allergen	1311642	AAB35730	16
1431	Triticum aestivum	Wheat	n/a	Food Plant	HMW glutenin subunit 1A	21743	CAA43331	830
1432	Triticum aestivum	Wheat	n/a	Aero Plant	NULL	4007852	CAA10349	122
1433	Triticum aestivum	Wheat	n/a	Gliadin	NULL	62484809	CAI78902	285
1434	Triticum aestivum	Wheat	n/a	Food Plant	Serine protease inhibitor	1885350	CAA72273	399
1435	Triticum aestivum	Wheat	n/a	Food Plant	Pollen allergen homolog	4099919	AAD10496	271
1436	Triticum aestivum	Wheat	n/a	Food Plant	NULL	62550933	CAI79052	326
1437	Triticum aestivum	Wheat	n/a	Aero Plant	Ribonuclease	40644796	CAE53906	231
1438	Triticum aestivum	Wheat	n/a	Gliadin	Alpha/beta gliadin	21673	CAA35238	307
1439	Triticum aestivum	Wheat	n/a	Food Plant	Alpha amylase inhibitor	21701	CAA35598	145
1440	Triticum aestivum	Wheat	n/a	Gliadin	Gliadin	21765	CAA26385	313
1441	Triticum aestivum	Wheat	n/a	Food Plant	Alpha amylase inhibitor	21713	CAA35597	168
1442	Triticum aestivum	Wheat	n/a	Food Plant	HMW glutenin subunit 1By9	22090	CAA43361	705
1443	Triticum aestivum	Wheat	n/a	Food Plant	HMW glutenin subunit 10	21751	CAA31396	648
1444	Triticum aestivum	Wheat	n/a	Food Plant	Storage protein	21755	CAA25593	286
1445	Triticum aestivum	Wheat	n/a	Gliadin	Gliadin	21761	CAA26384	286
1446	Triticum aestivum	Wheat	n/a	Food Plant	LMW glutenin	1168171	AAB35353	30
1447	Triticum aestivum	Wheat	n/a	Food Plant	Seed storage protein	21769	CAA28208	154
1448	Triticum aestivum	Wheat	n/a	Food Plant	LMW glutenin	21773	CAA31685	307
1449	Triticum aestivum	Wheat	n/a	Food Plant	HMW glutenin	21779	CAA26847	660
1450	Triticum aestivum	Wheat	n/a	Food Plant	LMW glutenin	21783	CAA30570	356
1451	Triticum aestivum	Wheat	n/a	Food Plant	HMW glutenin	21793	CAA24934	39
1452	Triticum aestivum	Wheat	n/a	Food Plant	Trypsin alpha-amylase inhibitor (frag)	21711	CAA42453	143
1453	Triticum aestivum	Wheat	n/a	Food Plant	HMW gluten	897811	CAA24933	101
1454	Triticum aestivum	Wheat	n/a	Food Plant	HMW glutenin	736319	CAA27052	838
1455	Triticum aestivum	Wheat	n/a	Gliadin	Pre-alpha-/beta gliadin A-IV	170724	AAA34282	297
1456	Triticum aestivum	Wheat	n/a	Food Plant	Alpha-amylase inhibitor	629817	S38955	180
1457	Triticum aestivum	Wheat	n/a	Aero Plant	NULL	55859460	CAH92633	518
1458	Triticum aestivum	Wheat	n/a	Food Plant	Profilin	1008445	CAA61945	140
1459	Triticum aestivum	Wheat	n/a	Food Plant	Profilin	1008443	CAA61944	141
1460	Triticum aestivum	Wheat	n/a	Gliadin	Alpha/beta gliadin	170718	AAA34279	313
1461	Triticum aestivum	Wheat	n/a	Food Plant	Pollen allergen homolog	972513	CAA90746	118
1462	Triticum aestivum	Wheat	n/a	Food Plant	Lysosomal thiol	30793446	BAC76688	203
1463	Triticum aestivum	Wheat	n/a	Aero Plant	NULL	55859458	CAH92632	518
1464	Triticum aestivum	Wheat	n/a	Food Plant	LMW glutenin	886967	CAA59340	276
1465	Triticum aestivum	Wheat	n/a	Gliadin	Alpha/beta gliadin	170720	AAA34280	286
1466	Triticum aestivum	Wheat	n/a	Food Plant	LMW glutenin	886965	CAA59339	261
1467	Triticum aestivum	Wheat	n/a	Food Plant	LMW glutenin	886963	CAA59338	229
1468	Triticum aestivum	Wheat	n/a	Gliadin	Gamma gliadin	170736	AAA34288	251
1469	Triticum aestivum	Wheat	n/a	Gliadin	Pre-alpha-/beta gliadin A-I	170722	AAA34281	262
1470	Triticum aestivum	Wheat	n/a	Gliadin	Alpha gliadin	170728	AAA34284	186

1471	Triticum aestivum	Wheat	n/a	Gliadin	Pre-alpha/beta gliadin A-II	170712	AAA34276	291
1472	Triticum aestivum	Wheat	n/a	Gliadin	Alpha gliadin	170710	AAA34275	318
1473	Triticum aestivum	Wheat	n/a	Gliadin	Gamma gliadin B	170708	AAA34274	291
1474	Triticum aestivum	Wheat	n/a	Gliadin	Gamma gliadin	170702	AAA34272	302
1475	Triticum aestivum	Wheat	n/a	Gliadin	Gamma gliadin	170732	AAA34286	323
1476	Triticum aestivum	Wheat	n/a	Gliadin	Pre-gamma gliadin B-I	170730	AAA34285	304
1477	Triticum aestivum	Wheat	n/a	Gliadin	NULL	63252971	CAI78903	375
1478	Triticum aestivum	Wheat	n/a	Food Plant	Agglutinin isolectin D	170670	AAA34258	213
1479	Triticum aestivum	Wheat	n/a	Gliadin	Alpha gliadin	473876	AAA17741	287
1480	Triticum aestivum	Wheat	n/a	Food Plant	Agglutinin isolectin A	170666	AAA34256	212
1481	Triticum aestivum	Wheat	n/a	Gliadin	Pre-alpha-/beta gliadin A-III	170726	AAA34283	282
1482	Triticum aestivum	Wheat	n/a	Gliadin	Gamma gliadin	1063270	BAA11251	279
1483	Triticum aestivum	Wheat	n/a	Aero Plant	Putative pollen allergen	40644792	CAE53904	178
1484	Triticum aestivum	Wheat	n/a	Food Plant	Profilin	1052817	CAA61943	138
1485	Triticum turgidum	Wheat	n/a	Aero Plant	Bakers asthma allergen	100842	S19296	10
1486	Triticum turgidum	Wheat	n/a	Aero Plant	NULL	244610	AAB21323	18
1487	Triticum turgidum subsp. durum	Wheat	n/a	Food Plant	LMW glutenin	21930	CAA44473	285
1488	Triticum turgidum subsp. durum	Wheat	n/a	Food Plant	CM2 protein (Trypsin alpha-amylase)	21920	CAA39099	145
1489	Triticum turgidum subsp. durum	Wheat	n/a	Food Plant	Mature glutenin	21926	CAA36063	295
1490	Triticum turgidum subsp. durum	Wheat	n/a	Food Plant	Wheat germ agglutinin	170668	AAA34257	186
1491	Triticum turgidum subsp. durum	Wheat	n/a	Food Plant	Alpha-amylase inhibitor	100834	S16031	168
1492	Triticum turgidum subsp. durum	Wheat	n/a	Food Plant	Trypsin alpha-amylase inhibitor (frag)	21916	CAA34709	143
1493	Triticum urartu	Wheat	n/a	Food Plant	Unknown protein	806315	AAA66167	106
1494	Triticum urartu	Wheat	n/a	Gliadin	Gliadin	170740	AAA34290	296
1495	Tyrophagus putrescentiae	Dust mite	Tyr p 2	Aero Mite	Group 2 allergen	2182106	CAA73221	141
1496	Tyrophagus putrescentiae	Dust mite	Tyr p 13	Aero Mite	NULL	51860756	AAU11502	131
1497	Ulocladium alternariae	Fungus	n/a	Aero Fungi	NULL	49476547	AAT66607	138
1498	Ulocladium atrum	Fungus	n/a	Aero Fungi	NULL	49476551	AAT66609	137
1499	Ulocladium botrytis	Fungus	n/a	Aero Fungi	NULL	49476549	AAT66608	137
1500	Ulocladium chartarum	Fungus	n/a	Aero Fungi	NULL	49476553	AAT66610	137
1501	Ulocladium cucurbitae	Fungus	n/a	Aero Fungi	NULL	49476545	AAT66606	137
1502	Ustilago maydis 521	Fungus	n/a	Aero Fungi	NULL	49074410	XP_401341	162
1503	Verticillium dahliae	Fungus	n/a	Aero Fungi	NULL	42742375	AAS45248	403
1504	Verticillium dahliae	Fungus	n/a	Aero Fungi	NULL	42742377	AAS45249	297
1505	Vespa crabro	European hornet	Vesp c 5	Venom or Salivary	Venom allergen	627182	H44583	202

1506	Vespa crabro	European hornet	Vesp c 5	Venom or Salivary	Venom allergen	627181	G44583	202
1507	Vespa mandarinia	Wasp	Vesp m 5	Venom or Salivary	Venom allergen 5	6136165	P81657	202
1508	Vespula flavopilosa	Wasp	Ves f 5	Venom or Salivary	Venom allergen 5	627185	A44583	204
1509	Vespula germanica	Wasp	Ves g 5	Venom or Salivary	Venom allergen 5	627186	B44583	204
1510	Vespula maculifrons	Wasp	Ves m 1	Venom or Salivary	Phospholipase A1	482382	A44564	300
1511	Vespula maculifrons	Wasp	Ves m 5	Venom or Salivary	Venom allergen 5	549191	P35760	204
1512	Vespula pensylvanica	Wasp	Ves p 5	Venom or Salivary	Venom allergen 5	627187	C44583	204
1513	Vespula squamosa	Wasp	Ves s 5	Venom or Salivary	Venom allergen 5	627188	D44583	205
1514	Vespula vidua	Wasp	Ves vi 5	Venom or Salivary	Venom allergen 5	627189	E44583	206
1515	Vespula vulgaris	Wasp	Ves v 5	Venom or Salivary	Venom allergen 5	4826574	CAB42887	204
1516	Vespula vulgaris	Wasp	Ves v 2	Venom or Salivary	Hyaluronidase	1346323	P49370	331
1517	Vespula vulgaris	Wasp	Ves v 1	Venom or Salivary	Phospholipase A1	897647	AAB48072	336
1518	Vespula vulgaris	Wasp	Ves v 5	Venom or Salivary	Venom allergen 5	11514279	1QNX_A	209
1519	Vespula vulgaris	Wasp	Ves v 5	Venom or Salivary	Venom allergen 5	162551	AAA30333	227
1520	Vespula vulgaris	Wasp	n/a	Venom or Salivary	NULL	62147665	CAI77218	340
1521	Vitis sp.	Grape	n/a	Food Plant	Lipid transfer protein P1	1084312	S39034	38
1522	Vitis sp.	Grape	n/a	Food Plant	Lipid transfer protein P3	1084313	S39036	37
1523	Vitis sp.	Grape	Vit v 1	Food Plant	Nonspecific lipid-transfer protein P4	462719	P80274	37
1524	Wuchereria bancrofti	Nematode	n/a	Worm (parasite)	Vespid antigen homolog	4324680	AAD16985	220
1525	Wuchereria bancrofti	Nematode	n/a	Worm (parasite)	Cuticular endochitinase	7673688	AAF66988	504
1526	Wuchereria bancrofti	Nematode	n/a	Worm (parasite)	Tumor protein-like protein	14700054	AAK71499	181
1527	Wuchereria bancrofti	Nematode	n/a	Worm (parasite)	Vespid antigen homolog	4704758	AAD28256	220
1528	Zea mays	Corn	n/a	Aero Plant	Zm13 (Ole e 1 family)	1588669	2209273A	170
1529	Zea mays	Corn	n/a	Aero Plant	NULL	33188338	AAP97894	120
1530	Zea mays	Corn	n/a	Aero Plant	Pectate lyase homolog	405535	AAA16476	438
1531	Zea mays	Corn	n/a	Food Plant	Profilin	2642324	AAB86960	131
1532	Zea mays	Corn	Zea m 14	Food Plant	Nonspecific lipid-transfer protein	128388	P19656	120
1533	Zea mays	Corn	n/a	Aero Plant	Beta-expansin 1-like protein	28630923	AAO45608	269

1534	Zea mays	Corn	n/a	Aero Plant	Beta-expansin 1a	20138191	P58738	269
1535	Zea mays	Corn	n/a	Aero Plant	Beta-expansin 1-like protein	28630919	AAO45607	269
1536	Zea mays	Corn	n/a	Aero Plant	NULL	54111527	AAV28626	154
1537	Zea mays	Corn	n/a	Aero Plant	NULL	7449063	S43334	104

Appendix 2. CTP-CRTI: Sequence and Alignment Data

Amino acid sequence of the fused construct of CTP-CRTI. The construct and sequences are referenced in Paine et al., 2005. The CTP (underlined) is the chloroplast transit peptide from the Pisum sativum rubisco small subunit, as described by Misawa et al. 1993). The CRTI protein sequence is from translated nucleotide sequence accession D90087, GI:22474502, sequence of crtI. As noted by Paine, the selected version of the gene includes a one nucleotide change from the primary sequence entry, but the change did not alter the amino acid sequence of the protein.

```
1   MASMISSSAVTTVSRASRGQSAAVAPFGGLKSMTGFPVKKVNTDITSITSNGGRVKCMKP
61  TTVIGAGFGGLALAIRLQAAGIPVLLLEQRDKPGGRAYVYEDQGFTFDAGPTVITDPSAI
121 EELFALAGKQLKEYVELLPVTPFYRLCWESGKVFNYDNDQTRLEAQIQQFNPRDVEGYRQ
181 FLDYSRAVFKEGYLKLGTVPFLSFRDMLRAAPQLAKLQAWRSVYSKVASYIEDEHLRQAF
241 SFHSLLVGGNPFATSSIYTLIHALEREWGVWFPRGGTGALVQGMIKLFQDLGGEVVLNAR
301 VSHMETTGNKIEAVHLEDGRRFLTQAVASNADVVHTYRDLLSQHPAAVKQSNKLQTKRMS
361 NSLFVLYFGLNHHHDQLAHHTVCFGPRYRELIDEIFNHDGLAEDFSLYLHAPCVTDSSLA
421 PEGCGSYYVLAPVPHLGTANLDWTVEGPKLRDRIFAYLEQHYMPGLRSQLVTHRMFTPFD
481 FRDQLNAYHGSAFSVEPVLTQSAWFRPHNRDKTITNLYLVGAGTHPGAGIPGVIGSAKAT
541 AGLMLEDLI
```

Search Full-length using FASTA3 Query CTP-CRTI against AllergenOnline Database v6.0 (January, 2006)

```
Query Sequence:
>fused pea rubisco SSU transit peptide plus the CRTI from Erwinia
MASMISSSAVTTVSRASRGQSAAVAPFGGLKSMTGFPVKKVNTDITSITSNGGRVKCMKP
TTVIGAGFGGLALAIRLQAAGIPVLLLEQRDKPGGRAYVYEDQGFTFDAGPTVITDPSAI
EELFALAGKQLKEYVELLPVTPFYRLCWESGKVFNYDNDQTRLEAQIQQFNPRDVEGYRQ
FLDYSRAVFKEGYLKLGTVPFLSFRDMLRAAPQLAKLQAWRSVYSKVASYIEDEHLRQAF
SFHSLLVGGNPFATSSIYTLIHALEREWGVWFPRGGTGALVQGMIKLFQDLGGEVVLNAR
VSHMETTGNKIEAVHLEDGRRFLTQAVASNADVVHTYRDLLSQHPAAVKQSNKLQTKRMS
NSLFVLYFGLNHHHDQLAHHTVCFGPRYRELIDEIFNHDGLAEDFSLYLHAPCVTDSSLA
PEGCGSYYVLAPVPHLGTANLDWTVEGPKLRDRIFAYLEQHYMPGLRSQLVTHRMFTPFD
FRDQLNAYHGSAFSVEPVLTQSAWFRPHNRDKTITNLYLVGAGTHPGAGIPGVIGSAKAT
AGLMLEDLI

Searching: AllergenOnline Database v6.0 (January, 2006)

# fasta34t25b1/fasta34 -a -Q -3 -E10 -d5 -B /tmp/allergen16596.tmp allergen6.fasta
FASTA searches a protein or DNA sequence data bank
 version 3.4t25 Nov 12, 2004
Please cite:
 W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

Query library /tmp/allergen16596.tmp vs allergen6.fasta library
```

```
searching allergen6.fasta library

  1>>>fused pea rubisco SSU transit peptide plus the CRTI from Erwinia - 549 aa
 vs  allergen6.fasta library

       opt      E()
< 20     4     0:==
  22     0     0:           one = represents 3 library sequences
  24     0     0:
  26     0     0:
  28     2     0:=
  30     7     2:*==
  32    16     8:==*===
  34    33    22:=======*===
  36    58    46:===============*====
  38    61    75:=====================   *
  40   118   105:==================================*=====
  42   145   129:==========================================*======
  44   125   142:==========================================     *
  46   101   144:==================================             *
  48   125   138:==========================================   *
  50   138   126:=========================================*====
  52   129   111:====================================*======
  54    77    95:==========================     *
  56    60    79:====================      *
  58    54    65:==================   *
  60    45    53:===============  *
  62    34    42:============ *
  64    43    34:===========*===
  66    46    27:========*=======
  68    26    21:======*==
  70    25    16:=====*===
  72     9    13:=== *
  74    12    10:===*
  76    16     8:==*===
  78    11     6:=*==
  80     8     5:=*=
  82     2     4:=*
  84     3     3:*
  86     0     2:*
  88     0     2:*          inset = represents 1 library sequences
  90     0     1:*
  92     0     1:*         :*
  94     1     1:*         :*
  96     1     1:*         :*
  98     0     0:          *
 100     1     0:=         *=
 102     0     0:          *
 104     0     0:          *
 106     0     0:          *
 108     0     0:          *
 110     0     0:          *
 112     0     0:          *
 114     0     0:          *
 116     1     0:=         *=
 118     0     0:          *
>120     0     0:          *
 336538 residues in  1537 sequences
  Expectation_n fit: rho(ln(x))= 5.2945+/-0.00383; mu= 11.0509+/- 0.204
 mean_var=50.1475+/-12.288, 0's: 4 Z-trim: 5  B-trim: 228 in 2/45
 Lambda= 0.181113
 Kolmogorov-Smirnov  statistic: 0.0355 (N=27) at  62

FASTA (3.47 Mar 2004) function [optimized, BL50 matrix (15:-5)] ktup: 2
 join: 37, opt: 25, open/ext: -10/-2, width:  16
 Scan time:  0.120
The best scores are:                                      opt z-sc E(1537)
gi|15624049|dbj|BAB68102.1| Bet v I allergen-l ( 208) [f]   86  115.9    0.18
gi|21335406|emb|CAD33269.1| unnamed protein pr ( 705) [f]   82  101.2     1.2
gi|3021373|emb|CAA11755.1| profilin [Glycine m ( 131) [f]   70   96.8     2.1
```

```
gi|729970|sp|P39674|MAG29_DERFA Allergen MAG29 ( 145) [f]   69   94.6     2.8

>>gi|15624049|dbj|BAB68102.1| Bet v I allergen-like [Ory  (208 aa)
 initn:  65 init1:  65 opt:  86  Z-score: 115.9  bits: 29.8 E(): 0.18
Smith-Waterman score: 86;  24.324% identity (53.153% similar) in 111
aa overlap (242-347:38-148)

               10        20        30        40        50        60
fused  MASMISSSAVTTVSRASRGQSAAVAPFGGLKSMTGFPVKKVNTDITSITSNGGRVKCMKP

               70        80        90       100       110       120
fused  TTVIGAGFGGLALAIRLQAAGIPVLLLEQRDKPGGRAYVYEDQGFTFDAGPTVITDPSAI

              130       140       150       160       170       180
fused  EELFALAGKQLKEYVELLPVTPFYRLCWESGKVFNYDNDQTRLEAQIQQFNPRDVEGYRQ

              190       200       210       220       230       240
fused  FLDYSRAVFKEGYLKLGTVPFLSFRDMLRAAPQLAKLQAWRSVYSKVASYIEDEHLRQAF

gi|156                         MPYAAVRPSPPPQLSRPIGSGAGGGKACPAVPCEVA
                                       10        20        30

              250       260       270         280       290
fused  SFHSLLVGGNPFATSSIYTLIHALEREWGVW--FPRGGTGALVQGMIKLFQDLGGEV--V
        .:   ::..   .. . ..    .  :.:   : :  .        .: .  ::::  :
gi|156 RYHEHAVGAGQCFSTVVQAIAAPADAVWSVVRRFDRPQAYKKFIKSCRLVDGDGGEVGSV
         40        50        60        70        80        90

        300        310       320       330       340       350
fused  LNARV-SHMETTGNKIEAVHLEDGRRFLTQAVASNADVVHTYRDLLSQHPAAVKQSNKLQ
        ..:: : . .:... .   :.: :: :.  ....   . .::.. . : ::
gi|156 REVRVVSGLPATSSRERLEVLDDDRRVLSFRIVGGEHRLANYRSVTTVHEAAAPAMAVVV
        100       110       120       130       140       150

         360       370       380       390       400       410
fused  TKRMSNSLFVLYFGLNHHHDQLAHHTVCFGPRYRELIDEIFNHDGLAEDFSLYLHAPCVT

gi|156 ESYVVDVPPGNTWEETRVFVDTIVRCNLQSLARTVERLAPEAPRANGSIDHA
        160       170       180       190       200

         420       430       440       450       460       470
fused  DSSLAPEGCGSYYVLAPVPHLGTANLDWTVEGPKLRDRIFAYLEQHYMPGLRSQLVTHRM

         480       490       500       510       520       530
fused  FTPFDFRDQLNAYHGSAFSVEPVLTQSAWFRPHNRDKTITNLYLVGAGTHPGAGIPGVIG

         540
fused  SAKATAGLMLEDLI

>>gi|21335406|emb|CAD33269.1| unnamed protein product [P  (705 aa)
 initn:  39 init1:  39 opt:  82  Z-score: 101.2  bits: 28.9 E():  1.2
Smith-Waterman score: 82;  25.926% identity (60.185% similar) in 108
aa overlap (153-254:21-125)

               10        20        30        40        50        60
fused  MASMISSSAVTTVSRASRGQSAAVAPFGGLKSMTGFPVKKVNTDITSITSNGGRVKCMKP

               70        80        90       100       110       120
fused  TTVIGAGFGGLALAIRLQAAGIPVLLLEQRDKPGGRAYVYEDQGFTFDAGPTVITDPSAI

              130       140       150       160        170
fused  EELFALAGKQLKEYVELLPVTPFYRLCWESGKVFNYDNDQTR-LEAQIQQFNPR---DVE
                                       :: ::. .:: ::... :.. .    .:
gi|213             MKTVLILAGLVALAAGNTFPVFRYDHVETRKLEGDLLQYQSKFLSLLE
                           10        20        30        40

        180       190         200       210       220       230
fused  GYRQFLDYSRAVFK--EGYLKLGTVPFLSFRDMLRAAPQLAKLQAWRSVYSKVASYIEDE
       . :: .::    .:  .::  ....   : .: .::   : ..    ..:.   : . :.
gi|213 NVRQ-IDYEAEYYKVGKGYDIVASIENYSDQDAVRAFAGLREIGFMPKAYT--FSIFYDR
```

```
       50         60        70        80        90         100

          240       250       260       270       280       290
fused  HLRQAFSFHSLLVGGNPFATSSIYTLIHALEREWGVWFPRGGTGALVQGMIKLFQDLGGE
       . ..:  ...:. ... . :
gi|213 QREEAKIIYDLFYSAKDLDTFYKTVAYGRIYFNEYQFMYAFYAAIIQRSDTTGIVLPAPY
         110       120       130       140       150       160

          300       310       320       330       340       350
fused  VVLNARVSHMETTGNKIEAVHLEDGRRFLTQAVASNADVVHTYRDLLSQHPAAVKQSNKL

gi|213 ELYPEYFLNMYTIQRMYRTQMQSGIFNEEVASNYGIWKMDNNYYYYYNYSNPLTYRNQEY
         170       180       190       200       210       220

          360       370       380       390       400       410
fused  QTKRMSNSLFVLYFGLNHHHDQLAHHTVCFGPRYRELIDEIFNHDGLAEDFSLYLHAPCV

gi|213 RLSYLTEDIGWNSYYYYFHNLMPFWGKGEDFIGIFKERRGEFYYYFYQQLLSRYYLERLS
         230       240       250       260       270       280

          420       430       440       450       460       470
fused  TDSSLAPEGCGSYYVLAPVPHLGTANLDWTVEGPKLRDRIFAYLEQHYMPGLRSQLVTHR

gi|213 NGLGEIPDFSWYQPLRSGYYPAIYTSSAYPFAQRPNYYYMGTEENVDYIQFLDAQEKSFV
         290       300       310       320       330       340

          480       490       500       510       520       530
fused  MFTPFDFRDQLNAYHGSAFSVEPVLTQSAWFRPHNRDKTITNLYLVGAGTHPGAGIPGVI

gi|213 QFLQIGQFKAFKQDVDFRNSKSINFVGNFWQGNPDLYDKYGREVNYDDSYEIIARRVLGA
         350       360       370       380       390       400

          540
fused  GSAKATAGLMLEDLI

gi|213 APPTSDNYEFVPSALDFYQTSLRDPAFYMLYNKIMSYIVQYKEWLEPYDQEVLHYSGVKI
         410       420       430       440       450       460

gi|213 NDVSVGNLTTFFEYYDFNATNAVFLSDQEIQQQYSSFIVRQPRLNHEPFSVTIDVKSDVE
         470       480       490       500       510       520

gi|213 AEAYFKIFVGPKYDGEGRPLSLEDNWMNFVELDWFTHKLTSGQNKVERKSEEFFFFKEDS
         530       540       550       560       570       580

gi|213 VSMSKIYELLKQGQVPESMSEDYDSMPSRLMLPRGTPGGFPVQFFVFVYPYQALSKDLEA
         590       600       610       620       630       640

gi|213 MKNIILDNKPLGYPFDRPVEYPYLFLQPNMYFEDVNIYHRGPQYPWWSNGQFRLNEVPRQ
         650       660       670       680       690       700

>>gi|3021373|emb|CAA11755.1| profilin [Glycine max]       (131 aa)
 initn:  35 init1:  35 opt:  70  Z-score: 96.8  bits: 25.6 E():  2.1
Smith-Waterman score: 70;  24.719% identity (51.685% similar) in 89
aa overlap (456-542:5-90)

               10        20        30        40        50        60
fused  MASMISSSAVTTVSRASRGQSAAVAPFGGLKSMTGFPVKKVNTDITSITSNGGRVKCMKP

               70        80        90       100       110       120
fused  TTVIGAGFGGLALAIRLQAAGIPVLLLEQRDKPGGRAYVYEDQGFTFDAGPTVITDPSAI

              130       140       150       160       170       180
fused  EELFALAGKQLKEYVELLPVTPFYRLCWESGKVFNYDNDQTRLEAQIQQFNPRDVEGYRQ

              190       200       210       220       230       240
fused  FLDYSRAVFKEGYLKLGTVPFLSFRDMLRAAPQLAKLQAWRSVYSKVASYIEDEHLRQAF

              250       260       270       280       290       300
fused  SFHSLLVGGNPFATSSIYTLIHALEREWGVWFPRGGTGALVQGMIKLFQDLGGEVVLNAR
```

```
              310       320       330       340       350       360
fused  VSHMETTGNKIEAVHLEDGRRFLTQAVASNADVVHTYRDLLSQHPAAVKQSNKLQTKRMS

              370       380       390       400       410       420
fused  NSLFVLYFGLNHHHDQLAHHTVCFGPRYRELIDEIFNHDGLAEDFSLYLHAPCVTDSSLA

              430       440       450       460       470       480
fused  PEGCGSYYVLAPVPHLGTANLDWTVEGPKLRDRIFAYLEQHYMPGLRSQLVTHRMFTPFD
                                          ::...: . :.... .::  .   :
gi|302                                MSWQAYVDDHLLCGIEGNHLTHAAIIGQD
                                              10        20

              490       500       510         520       530
fused  FRDQLNAYHGSAFSVEPVLTQSAWFRPHNRDKTI--TNLYLVGAGTHPGAGIPGVIGSAK
           :..     :. : .   .: .   :.  ..  :.::: :.      : ::..  .:
gi|302 GSVWLQSTDFPQFKPEEI---TAIMNDFNEPGSLAPTGLYLGGTKYMVIQGEPGAVIRGK
      30        40           50        60        70        80

      540
fused  ATAGLMLEDLI
          :
gi|302 KGPGGVTVKKTGAALIIGIYDEPMTPGQCNMVVERLGDYLIDQGY
         90       100       110       120       130

>>gi|729970|sp|P39674|MAG29_DERFA Allergen MAG29          (145 aa)
 initn:  69 init1:  46 opt:  69  Z-score: 94.6  bits: 25.4 E():  2.8
Smith-Waterman score: 69;  44.444% identity (66.667% similar) in 36
aa overlap (514-542:96-131)

               10        20        30        40        50        60
fused  MASMISSSAVTTVSRASRGQSAAVAPFGGLKSMTGFPVKKVNTDITSITSNGGRVKCMKP

               70        80        90       100       110       120
fused  TTVIGAGFGGLALAIRLQAAGIPVLLLEQRDKPGGRAYVYEDQGFTFDAGPTVITDPSAI

              130       140       150       160       170       180
fused  EELFALAGKQLKEYVELLPVTPFYRLCWESGKVFNYDNDQTRLEAQIQQFNPRDVEGYRQ

              190       200       210       220       230       240
fused  FLDYSRAVFKEGYLKLGTVPFLSFRDMLRAAPQLAKLQAWRSVYSKVASYIEDEHLRQAF

              250       260       270       280       290       300
fused  SFHSLLVGGNPFATSSIYTLIHALEREWGVWFPRGGTGALVQGMIKLFQDLGGEVVLNAR

              310       320       330       340       350       360
fused  VSHMETTGNKIEAVHLEDGRRFLTQAVASNADVVHTYRDLLSQHPAAVKQSNKLQTKRMS

              370       380       390       400       410       420
fused  NSLFVLYFGLNHHHDQLAHHTVCFGPRYRELIDEIFNHDGLAEDFSLYLHAPCVTDSSLA

gi|729                                                           KD


              430       440       450       460       470       480
fused  PEGCGSYYVLAPVPHLGTANLDWTVEGPKLRDRIFAYLEQHYMPGLRSQLVTHRMFTPFD

gi|729 DIERMVKEAESYKEEDDKQRDRIAAKNSLEGYAFQMKATLDEEAIKSKVSEEDRKKILDK
             10        20        30        40        50        60

              490       500       510           520          530
fused  FRDQLNAYHGSAFSVEPVLTQSAWFRPHNRDKTITNLYL----VGAGTHPGA---GIPGV
                                        ::.::     .:::  ::.   :.::.
gi|729 VDEVLKWLDANALAEKDEFEHQRKELESVCNPIITKLYQQAGGAGAGGMPGGFPGGFPGT
             70        80        90       100       110       120

           540
fused  IGSAKATAGLMLEDLI
        ::. ..::
gi|729 DGSGGGAAGGDGGKSGPTIEEVD
            130       140
```

```
549 residues in 1 query   sequences
336538 residues in 1537 library sequences
 Scomplib [34t25]
 start: Thu Apr  6 16:47:48 2006 done: Thu Apr  6 16:47:48 2006
 Total Scan time:  0.120 Total Display time:  0.010

Function used was FASTA [version 3.4t25 Nov 12, 2004]
```

Search 80-mer Sliding Window using FASTA3 with Query CTP-CRTI against AllergenOnline Database v6.0 (January, 2006)

Database	AllergenOnline v6.0 (January, 2006)
Input Query	>fused pea rubisco SSU transit peptide plus the CRTI from Erwinia MASMISSSAVTTVSRASRGQSAAVAPFGGLKSMTGFPVKKVNTDITSITSNGGRVKCMKP TTVIGAGFGGLALAIRLQAAGIPVLLLEQRDKPGGRAYVYEDQGFTFDAGPTVITDPSAI EELFALAGKQLKEYVELLPVTPFYRLCWESGKVFNYDNDQTRLEAQIQQFNPRDVEGYRQ FLDYSRAVFKEGYLKLGTVPFLSFRDMLRAAPQLAKLQAWRSVYSKVASYIEDEHLRQAF SFHSLLVGGNPFATSSIYTLIHALEREWGVWFPRGGTGALVQGMIKLFQDLGGEVVLNAR VSHMETTGNKIEAVHLEDGRRFLTQAVASNADVVHTYRDLLSQHPAAVKQSNKLQTKRMS NSLFVLYFGLNHHHDQLAHHTVCFGPRYRELIDEIFNHDGLAEDFSLYLHAPCVTDSSLA PEGCGSYYVLAPVPHLGTANLDWTVEGPKLRDRIFAYLEQHYMPGLRSQLVTHRMFTPFD FRDQLNAYHGSAFSVEPVLTQSAWFRPHNRDKTITNLYLVGAGTHPGAGIPGVIGSAKAT AGLMLEDLI
Length	549
Number of 80 mers	470

Hit #	Defline	Best %ID	# Hits > 35%	Full Alignment			Links	
				E-val	%ID	length	NCBI	Details

NOTE: There were no matched identities of 35% or more over any 80-amino acid segment of CTP-CRTI.

BLASTP Search with Query CTP-CRTI against the NCBI Database Available on Thursday, April 06, 2006, using a key word delimiter “allergen” (limited search to 2,529 sequences).

```
Thursday, April 06, 2006

BLASTP 2.2.13 [Nov-27-2005]
Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
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: Schäffer, Alejandro A., L. Aravind, Thomas L. Madden, Sergei
Shavirin, John L. Spouge, Yuri I. Wolf, Eugene V. Koonin, and
Stephen F. Altschul (2001), "Improving the accuracy of PSI-BLAST
protein database searches with composition-based statistics
and other refinements", Nucleic Acids Res. 29:2994-3005.

RID: 1144361735-31184-110700561729.BLASTQ4


Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF excluding environmental samples
           3,538,507 sequences; 1,215,510,009 total letters
Query=  fused pea rubisco SSU transit peptide plus the CRTI from Erwinia
Length=549


                                                                    Score     E
Sequences producing significant alignments:                         (Bits)  Value

gi|88192990|pdb|2FFS|B  Chain B, Structure Of Pr10-Allergen-Li...   26.9    1.3
gi|74016535|ref|ZP_00687161.1|  Alkyl hydroperoxide reductase/...   25.8    3.0

ALIGNMENTS
>gi|88192990|pdb|2FFS|B Chain B, Structure Of Pr10-Allergen-Like Protein Pa1206 From
Pseudomonas Aeruginosa Pao1
 gi|88192989|pdb|2FFS|A Chain A, Structure Of Pr10-Allergen-Like Protein Pa1206 From
Pseudomonas Aeruginosa Pao1
Length=157

 Score = 26.9 bits (58),  Expect = 1.3, Method: Composition-based stats.
 Identities = 20/73 (27%), Positives = 30/73 (41%), Gaps = 2/73 (2%)

Query  394  EIFNHDGLAEDFSLYLHAPCVTDSSLAPEGCGSYYVLAPVPHLGTANLDWTVEGPKLRDR  453
            EI   DG      LYL    V D  +      ++Y + P   +   +LD T+E P+
Sbjct  47   EILVDDGDRLHRRLYLPGLVVEDEVVLKAPDSAHYSIKPSAEVAGGSLDXTIEEPEPGSL  106

Query  454  I--FAYLEQHYMP  464
               FAY  ++  P
Sbjct  107  FVRFAYCTRYLQP  119


>gi|74016535|ref|ZP_00687161.1| Alkyl hydroperoxide reductase/ Thiol specific
antioxidant/ Mal
allergen [Burkholderia ambifaria AMMD]
 gi|72610906|gb|EAO46852.1| Alkyl hydroperoxide reductase/ Thiol specific antioxidant/
Mal
allergen [Burkholderia ambifaria AMMD]
Length=218

 Score = 25.8 bits (55),  Expect = 3.0, Method: Composition-based stats.
 Identities = 13/41 (31%), Positives = 17/41 (41%), Gaps = 4/41 (9%)

Query  453  RIFAYLEQHYMPGLRSQLVTHRMFTPFDFRDQLNAYHGSAF  493
            RI   L QH+  G R   + +    PF    +  AYH   F
Sbjct  16   RILGTLHQHHPKGTRKMPIINTQIKPF----KATAYHNGEF  52
```

```
Database: All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF
excluding environmental samples
    Posted date:  Apr 4, 2006  3:06 AM
  Number of letters in database: 583,171
  Number of sequences in database:  2,529
Lambda     K      H
   0.321    0.137    0.413
Gapped
Lambda     K      H
   0.267   0.0410    0.140
Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 2529
Number of Hits to DB: 36784
Number of extensions: 1585
Number of successful extensions: 5
Number of sequences better than 10: 2
Number of HSP's better than 10 without gapping: 0
Number of HSP's gapped: 5
Number of HSP's successfully gapped: 2
Length of query: 549
Length of database: 583171
Length adjustment: 82
Effective length of query: 467
Effective length of database: 583171
Effective search space: 272340857
Effective search space used: 175495331
T: 11
A: 40
X1: 16 (7.4 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (20.4 bits)
S2: 51 (24.3 bits)
```

Appendix 3. PSY1: Sequence and Alignment Data

Amino acid sequence of the PSY1 protein. The construct and sequences are referenced in Paine et al., 2005.

This is the protein PSY1 Allele B73 (the main entry in NCBI has a difference (as it is the Q60 allele) is from Maize (U32636.1 GI:1098664), 410 amino acids.

```
1   MAIILVRAASPGLSAADSISHQGTLQCSTLLKTKRPAARRWMPCSLLGLHPWEAGRPSPA
61  VYSSLPVNPAGEAVVSSEQKVYDVVLKQAALLKRQLRTPVLDARPQDMDMPRNGLKEAYD
121 RCGEICEEYAKTFYLGTMLMTEERRRAIWAIYVWCRRTDELVDGPNANYITPTALDRWEK
181 RLEDLFTGRPYDMLDAALSDTISRFPIDIQPFRDMIEGMRSDLRKTRYNNFDELYMYCYY
241 VAGTVGLMSVPVMGIATESKATTESVYSAALALGIANQLTNILRDVGEDARRGRIYLPQD
301 ELAQAGLSDEDIFKGVVTNRWRNFMKRQIKRARMFFEEAERGVTELSQASRWPVWASLLL
361 YRQILDEIEANDYNNFTKRAYVGKGKKLLALPVAYGKSLLLPCSLRNGQT
```

Search Full-length using FASTA3 Query PSY1 against AllergenOnline Database v6.0 (January, 2006)

```
Query Sequence:
>PSY1 Allele B73
MAIILVRAASPGLSAADSISHQGTLQCSTLLKTKRPAARRWMPCSLLGLHPWEAGRPSPA
VYSSLPVNPAGEAVVSSEQKVYDVVLKQAALLKRQLRTPVLDARPQDMDMPRNGLKEAYD
RCGEICEEYAKTFYLGTMLMTEERRRAIWAIYVWCRRTDELVDGPNANYITPTALDRWEK
RLEDLFTGRPYDMLDAALSDTISRFPIDIQPFRDMIEGMRSDLRKTRYNNFDELYMYCYY
VAGTVGLMSVPVMGIATESKATTESVYSAALALGIANQLTNILRDVGEDARRGRIYLPQD
ELAQAGLSDEDIFKGVVTNRWRNFMKRQIKRARMFFEEAERGVTELSQASRWPVWASLLL
YRQILDEIEANDYNNFTKRAYVGKGKKLLALPVAYGKSLLLPCSLRNGQT

Searching: AllergenOnline Database v6.0 (January, 2006)

# fasta34t25b1/fasta34 -a -Q -3 -E10 -d5 -B /tmp/allergen16598.tmp allergen6.fasta
FASTA searches a protein or DNA sequence data bank
 version 3.4t25 Nov 12, 2004
Please cite:
 W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

Query library /tmp/allergen16598.tmp vs allergen6.fasta library
searching allergen6.fasta library

  1>>>PSY1 Allele B73 - 410 aa
 vs  allergen6.fasta library

       opt      E()
< 20     4     0:==
  22     0     0:           one = represents 3 library sequences
  24     0     0:
  26     0     0:
  28     0     0:
  30     2     2:*
  32     4     8:==*
  34    11    22:====   *
  36    41    46:============== *
  38    65    75:======================  *
```

```
  40   103   105:==================================*
  42   123   129:========================================= *
  44   160   142:===============================================*======
  46   150   144:===============================================*==
  48   150   138:=============================================*====
  50   115   126:=======================================  *
  52    92   111:===============================     *
  54    99    95:===============================*=
  56    69    79:=======================   *
  58    68    65:=====================*=
  60    64    53:=================*====
  62    39    42:=============*
  64    29    34:========== *
  66    31    27:========*==
  68    20    21:======*
  70    37    16:=====*=======
  72    19    13:====*==
  74     6    10:== *
  76     4     8:==*
  78     8     6:=*=
  80     1     5:=*
  82     5     4:=*
  84     4     3:*=
  86     0     2:*
  88     1     2:*          inset = represents 1 library sequences
  90     1     1:*
  92     2     1:*         :*=
  94     2     1:*         :*=
  96     4     1:*=        :*===
  98     1     0:=         *=
 100     1     0:=         *=
 102     2     0:=         *==
 104     0     0:          *
 106     0     0:          *
 108     0     0:          *
 110     0     0:          *
 112     0     0:          *
 114     0     0:          *
 116     0     0:          *
 118     0     0:          *
>120     0     0:          *
 336538 residues in  1537 sequences
  Expectation_n fit: rho(ln(x))= 4.9770+/-0.00311; mu= 9.9493+/- 0.160
 mean_var=54.9245+/-14.631, 0's: 4 Z-trim: 4  B-trim: 17 in 1/46
 Lambda= 0.173058
 Kolmogorov-Smirnov  statistic: 0.0252 (N=28) at  56

FASTA (3.47 Mar 2004) function [optimized, BL50 matrix (15:-5)] ktup: 2
 join: 37, opt: 25, open/ext: -10/-2, width:  16
 Scan time:  0.110
The best scores are:                                      opt z-sc E(1537)
gi|399672|emb|CAA52194.1| preproalbumin [Equus ( 607) [f]   81  102.8    0.98
gi|11277083|pir||A59225 allergen Bos d 2.0102  ( 156) [f]   74  102.5       1
gi|633938|gb|AAB30434.1| albumin [Canis famili ( 265) [f]   75  100.3     1.4
gi|2739154|gb|AAC67308.1| 22.6 kDa tegumental  ( 191) [f]   72   98.5     1.7
gi|11277082|pir||B59225 allergen Bos d 2.0103  ( 156) [f]   70   97.1       2
gi|886215|gb|AAB08720.1| major allergen BDA20  ( 172) [f]   70   96.5     2.2
gi|162648|gb|AAA51411.1| albumin [Bos taurus]  ( 607) [f]   76   96.1     2.3
gi|3336842|emb|CAA76847.1| bovine serum albumi ( 607) [f]   76   96.1     2.3
gi|886485|emb|CAA59279.1| albumin precursor [F ( 608) [f]   76   96.1     2.3
gi|6687188|emb|CAB64867.1| albumin [Canis fami ( 608) [f]   75   94.7     2.8
gi|3776613|emb|CAA83015.1| metalloprotease (ME ( 634) [f]   74   93.1     3.4
gi|3319897|emb|CAA76841.1| albumin [Canis fami ( 585) [f]   73   92.3     3.8
gi|16076695|gb|AAL14078.1| acidic Cyn d 1 isoa ( 262) [f]   67   89.6     5.4
gi|14423933|sp|O82803|SRPP_HEVBR Small rubber  ( 204) [f]   65   88.6     6.1
gi|2498577|sp|P56164|MP51_PHAAQ Major pollen a ( 320) [f]   65   85.5       9

>>gi|399672|emb|CAA52194.1| preproalbumin [Equus caballu  (607 aa)
 initn:  46 init1:  46 opt:  81  Z-score: 102.8  bits: 28.5 E(): 0.98
Smith-Waterman score: 81;  29.508% identity (60.656% similar) in 61
```

```
aa overlap (100-156:313-371)

gi|399 MKWVTFVSLLFLFSSAYSRGVLRRDTHKSEIAHRFNDLGEKHFKGLVLVAFSQYLQQCPF
               10        20        30        40        50        60

gi|399 EDHVKLVNEVTEFAKKCAADESAENCDKSLHTLFGDKLCTVATLRATYGELADCCEKQEP
               70        80        90       100       110       120

gi|399 ERNECFLTHKDDHPNLPKLKPEPDAQCAAFQEDPDKFLGKYLYEVARRHPYFYGPELLFH
              130       140       150       160       170       180

                                                10        20
PSY1                                    MAIILVRAASPGLSAADSISHQGTLQC

gi|399 AEEYKADFTECCPADDKLACLIPKLDALKERILLSSAKERLKCSSFQNFGERAVKAWSVA
              190       200       210       220       230       240

        30        40        50        60        70        80
PSY1   STLLKTKRPAARRWMPCSLLGLHPWEAGRPSPAVYSSLPVNPAGEAVVSSEQKVYDVVLK

gi|399 RLSQKFPKADFAEVSKIVTDLTKVHKECCHGDLLECADDRADLAKYICEHQDSISGKLKA
              250       260       270       280       290       300

        90       100         110       120         130       140
PSY1   QAALLKRQLRTPVLDARPQDM--DMPRNGLKEAYDRCGEICEEY--AKTFYLGTMLMTEE
                   . ... .:.  :.:  .   : :.  ::..:   ::  .:::.:.
gi|399 CCDKPLLQKSHCIAEVKEDDLPSDLPALAADFAEDK--EICKHYKDAKDVFLGTFLYEYS
              310       320       330         340       350

           150       160       170       180       190       200
PSY1   RRRAIWAIYVWCRRTDELVDGPNANYITPTALDRWEKRLEDLFTGRPYDMLDAALSDTIS
       ::.  ... .  :
gi|399 RRHPDYSVSLLLRIAKTYEATLEKCCAEADPPACYRTVFDQFTPLVEEPKSLVKKNCDLF
      360       370       380       390       400       410

           210       220       230       240       250       260
PSY1   RFPIDIQPFRDMIEGMRSDLRKTRYNNFDELYMYCYYVAGTVGLMSVPVMGIATESKATT

gi|399 EEVGEYDFQNALIVRYTKKAPQVSTPTLVEIGRTLGKVGSRCCKLPESERLPCSENHLAL
      420       430       440       450       460       470

           270       280       290       300       310       320
PSY1   ESVYSAALALGIANQLTNILRDVGEDARRGRIYLPQDELAQAGLSDEDIFKGVVTNRWRN

gi|399 ALNRLCVLHEKTPVSEKITKCCTDSLAERRPCFSALELDEGYVPKEFKAETFTFHADICT
      480       490       500       510       520       530

           330       340       350       360       370       380
PSY1   FMKRQIKRARMFFEEAERGVTELSQASRWPVWASLLLYRQILDEIEANDYNNFTKRAYVG

gi|399 LPEDEKQIKKQSALAELVKHKPKATKEQLKTVLGNFSAFVAKCCGREDKEACFAEEGPKL
      540       550       560       570       580       590

           390       400       410
PSY1   KGKKLLALPVAYGKSLLLPCSLRNGQT

gi|399 VASSQLALA
      600

>>gi|11277083|pir||A59225 allergen Bos d 2.0102 [importe  (156 aa)
 initn:  41 init1:  41 opt:  74  Z-score: 102.5  bits: 26.5 E():   1
Smith-Waterman score: 74;  25.234% identity (51.402% similar) in 107
aa overlap (115-215:35-134)

               10        20        30        40        50        60
PSY1   MAIILVRAASPGLSAADSISHQGTLQCSTLLKTKRPAARRWMPCSLLGLHPWEAGRPSPA

               70        80        90       100       110       120
PSY1   VYSSLPVNPAGEAVVSSEQKVYDVVLKQAALLKRQLRTPVLDARPQDMDMPRNGLKEAYD
                                                             :.. :
```

```
gi|112                     QETPAEIDPSKIPGEWRIIYAAADNKDKIVEGGPLRNYYR
                                   10        20        30        40

                130           140       150       160       170
PSY1   R--CGEICEEYAKTFYL---GT-MLMTEERRRAIWAIYVWCRRTDELVDGPNANYITPTA
       :  : . ::  . ::::   :: .:.::  .:    .::      :. :: :.  .  ..
gi|112 RIECINDCESLSITFYLKDQGTCLLLTEVAKRQEGYVYVL-----EF-DGTNTLEVIHVS
               50        60        70        80              90

          180       190       200       210       220       230
PSY1   LDRWEKRLEDLFTGRPYDMLDAALSDTISRFPIDIQPFRDMIEGMRSDLRKTRYNNFDEL
        .     .:. . :.    .  .:.   :  : ... ....
gi|112 ENMLVTYVEN-YDGERITKMTEGLAKGTSFTPEELEKYQQLNSERGVPNENIENLIKTDN
          100        110       120       130       140       150

          240       250       260       270       280       290
PSY1   YMYCYYVAGTVGLMSVPVMGIATESKATTESVYSAALALGIANQLTNILRDVGEDARRGR

gi|112 CPP


          300       310       320       330       340       350
PSY1   IYLPQDELAQAGLSDEDIFKGVVTNRWRNFMKRQIKRARMFFEEAERGVTELSQASRWPV

          360       370       380       390       400       410
PSY1   WASLLLYRQILDEIEANDYNNFTKRAYVGKGKKLLALPVAYGKSLLLPCSLRNGQT

>>gi|633938|gb|AAB30434.1| albumin [Canis familiaris]     (265 aa)
 initn:  46 init1:  46 opt:  75  Z-score: 100.3  bits: 26.9 E(): 1.4
Smith-Waterman score: 75;  29.630% identity (59.259% similar) in 54
aa overlap (109-160:111-162)

                 10        20        30        40        50
PSY1     MAIILVRAASPGLSAADSISHQGTLQCSTLLKTKRPAARRWMPCSLLGLHPWEAGRPS

gi|633 LSSAKERFKCASLQKFGDRAFKAWSVARLSQRFPKADFAEISKVVTDLTKVHKECCHGDL
               10        20        30        40        50        60

       60        70        80        90       100       110
PSY1   PAVYSSLPVNPAGEAVVSSEQKVYDVVLKQAALLKRQLRTPVLDARPQDMDMPRNGLKEA
                                                         :.:  .   .
gi|633 LECADDRADLAKYMCENQDSISTKLKECCDKPVLEKSQCLAEVERDELPGDLPSLAADFV
               70        80        90       100       110       120

      120         130       140       150       160       170
PSY1   YDRCGEICEEY--AKTFYLGTMLMTEERRRAIWAIYVWCRRTDELVDGPNANYITPTALD
        :.  :.:..:  ::  .:::.:.   ::.  ... .  : . :
gi|633 EDK--EVCKNYQEAKDVFLGTFLYEYSRRHPEYSVSLLLRLAKEYEATLEKCCATDDPPT
                130       140       150       160       170

        180       190       200       210       220       230
PSY1   RWEKRLEDLFTGRPYDMLDAALSDTISRFPIDIQPFRDMIEGMRSDLRKTRYNNFDELYM

gi|633 CYAKVLDEFKPLVDEPQNLVKTNCELFEKLGEYGFQNALLVRYTKKAPQVSTPTLVVEVS
      180       190       200       210       220       230

        240       250       260       270       280       290
PSY1   YCYYVAGTVGLMSVPVMGIATESKATTESVYSAALALGIANQLTNILRDVGEDARRGRIY

gi|633 RKLGKVGTKCCKKPESERMSCADDFLS
      240       250       260

        300       310       320       330       340       350
PSY1   LPQDELAQAGLSDEDIFKGVVTNRWRNFMKRQIKRARMFFEEAERGVTELSQASRWPVWA

        360       370       380       390       400       410
PSY1   SLLLYRQILDEIEANDYNNFTKRAYVGKGKKLLALPVAYGKSLLLPCSLRNGQT

>>gi|2739154|gb|AAC67308.1| 22.6 kDa tegumental antigen   (191 aa)
 initn:  36 init1:  36 opt:  72  Z-score: 98.5  bits: 26.1 E():  1.7
```

```
Smith-Waterman score: 72;  29.032% identity (51.613% similar) in 62
aa overlap (44-98:69-130)

                                        10        20        30
PSY1                                MAIILVRAASPGLSAADSISHQGTLQCSTLLKTKR

gi|273 MATTEYRLSLMEQFIRAFIEIDKDNNELIDKQELTKYCQQNQMDMKQIDPWIARFDTDKD
               10        20        30        40        50        60

          40        50        60               70        80
PSY1   PAARRWMPCSLLGLHPWEAGRPSPAV-------YSSLPVNPAGEAVVSSEQKVYDVVLKQ
               :  .::. ::. : .  .        :.::..    :.. :. : :..  :
gi|273 GKVSLEEFCRGFGLKVWEVRREKEELKKDKEGKVSTLPLDIQIIAATMSKAKQYNICCKF
               70        80        90       100       110       120

       90       100       110       120       130       140
PSY1   AALLKRQLRTPVLDARPQDMDMPRNGLKEAYDRCGEICEEYAKTFYLGTMLMTEERRRAI
         :: .  ::
gi|273 KELLDKTSRTGDEVRAVANDLKAFLDSEYGRVWQVIILTGSYWMNFSHEPFLSMQFKYSN
              130       140       150       160       170       180

      150       160       170       180       190       200
PSY1   WAIYVWCRRTDELVDGPNANYITPTALDRWEKRLEDLFTGRPYDMLDAALSDTISRFPID

gi|273 YVCLLWRTPSS
              190

      210       220       230       240       250       260
PSY1   IQPFRDMIEGMRSDLRKTRYNNFDELYMYCYYVAGTVGLMSVPVMGIATESKATTESVYS

      270       280       290       300       310       320
PSY1   AALALGIANQLTNILRDVGEDARRGRIYLPQDELAQAGLSDEDIFKGVVTNRWRNFMKRQ

      330       340       350       360       370       380
PSY1   IKRARMFFEEAERGVTELSQASRWPVWASLLLYRQILDEIEANDYNNFTKRAYVGKGKKL

      390       400       410
PSY1   LALPVAYGKSLLLPCSLRNGQT

>>gi|11277082|pir||B59225 allergen Bos d 2.0103 [importe  (156 aa)
 initn:  41 init1:  41 opt:  70  Z-score: 97.1  bits: 25.5 E():   2
Smith-Waterman score: 70;  40.000% identity (57.778% similar) in 45
aa overlap (115-153:35-79)

               10        20        30        40        50        60
PSY1   MAIILVRAASPGLSAADSISHQGTLQCSTLLKTKRPAARRWMPCSLLGLHPWEAGRPSPA

               70        80        90       100       110       120
PSY1   VYSSLPVNPAGEAVVSSEQKVYDVVLKQAALLKRQLRTPVLDARPQDMDMPRNGLKEAYD
                                                             :.. :
gi|112                     QETPAEIDPSKIPGEWRIIYAAADNKDKIVEGGPLRNYYR
                                   10        20        30        40

                130           140       150       160       170
PSY1   R--CGEICEEYAKTFYL---GT-MLMTEERRRAIWAIYVWCRRTDELVDGPNANYITPTA
       :  : . ::  . ::::   :: .:.::  .:    .::
gi|112 RIECINDCESLSITFYLKDQGTCLLLTEVAKRQEGYVYVLEFYGTNTLEVIHVSENMLVT
               50        60        70        80        90       100

          180       190       200       210       220       230
PSY1   LDRWEKRLEDLFTGRPYDMLDAALSDTISRFPIDIQPFRDMIEGMRSDLRKTRYNNFDEL

gi|112 YAENYDGERITKMTEGLAKGTSFTPEELEKYQQLNSERGVPNENIENLIKTDNCPP
              110       120       130       140       150

          240       250       260       270       280       290
PSY1   YMYCYYVAGTVGLMSVPVMGIATESKATTESVYSAALALGIANQLTNILRDVGEDARRGR

          300       310       320       330       340       350
PSY1   IYLPQDELAQAGLSDEDIFKGVVTNRWRNFMKRQIKRARMFFEEAERGVTELSQASRWPV
```

```
          360       370       380       390       400       410
PSY1   WASLLLYRQILDEIEANDYNNFTKRAYVGKGKKLLALPVAYGKSLLLPCSLRNGQT



410 residues in 1 query   sequences
336538 residues in 1537 library sequences
 Scomplib [34t25]
 start: Thu Apr  6 16:49:40 2006 done: Thu Apr  6 16:49:41 2006
 Total Scan time:  0.110 Total Display time:  0.010

Function used was FASTA [version 3.4t25 Nov 12, 2004]
```

Search 80-mer Sliding Window using FASTA3 with Query PSY1 against AllergenOnline Database v6.0 (January, 2006)

Database	AllergenOnline v6.0 (January, 2006)
Input Query	>PSY1 Allele B73 MAIILVRAASPGLSAADSISHQGTLQCSTLLKTKRPAARRWMPCSLLGLHPWEAGRPSPA VYSSLPVNPAGEAVVSSEQKVYDVVLKQAALLKRQLRTPVLDARPQDMDMPRNGLKEAYD RCGEICEEYAKTFYLGTMLMTEERRRAIWAIYVWCRRTDELVDGPNANYITPTALDRWEK RLEDLFTGRPYDMLDAALSDTISRFPIDIQPFRDMIEGMRSDLRKTRYNNFDELYMYCYY VAGTVGLMSVPVMGIATESKATTESVYSAALALGIANQLTNILRDVGEDARRGRIYLPQD ELAQAGLSDEDIFKGVVTNRWRNFMKRQIKRARMFFEEAERGVTELSQASRWPVWASLLL YRQILDEIEANDYNNFTKRAYVGKGKKLLALPVAYGKSLLLPCSLRNGQT
Length	410
Number of 80 mers	331

Hit #	Defline	Best %ID	# Hits > 35%	Full Alignment			Links	
				E-val	%ID	length	NCBI	Details

NOTE: There were no matched identities of 35% or more over any 80-amino acid segment of PSY1.

BLASTP Search with Query PSY1 against the NCBI Database Available on Thursday, April 06, 2006, using a key word delimiter “allergen” (limited search to 2,529 sequences).

```
Thursday, April 06, 2006

BLASTP 2.2.13 [Nov-27-2005]
Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
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.

RID: 1144362667-21716-121744964274.BLASTQ4


Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF excluding environmental samples
           3,538,507 sequences; 1,215,510,009 total letters
Query=  PSY1 Allele B73
Length=410


                                                                   Score     E
Sequences producing significant alignments:                        (Bits)  Value

gi|40288346|gb|AAR84202.1|  nitrile-specifier protein [Pieris rap  26.2    1.7
gi|6900304|emb|CAB71342.1|  Lep d 10 protein [Lepidoglyphus de...  25.4    2.8
gi|2739154|gb|AAC67308.1|  22.6 kDa tegumental antigen [Schistoso  24.3    6.3
gi|2440053|emb|CAA75141.1|  tropomyosin [Dermatophagoides pter...  24.3    6.3
gi|2353266|gb|AAB69424.1|  tropomyosin [Dermatophagoides pteronys  24.3    6.3
gi|5832805|emb|CAB55143.1|  Hypothetical protein Y116A8C.8 [Caeno  23.9    8.3
gi|14423976|sp|Q9NAS5|TPM_ANISI  Tropomyosin (Allergen Ani s 3)    23.9    8.3
gi|4468639|emb|CAB38086.1|  tropomyosin [Periplaneta americana...  23.9    8.3
gi|4378573|gb|AAD19606.1|  tropomyosin [Periplaneta americana]     23.9    8.3

ALIGNMENTS
>gi|40288346|gb|AAR84202.1| nitrile-specifier protein [Pieris rapae]
Length=632

 Score = 26.2 bits (56),  Expect = 1.7
 Identities = 31/106 (29%), Positives = 43/106 (40%), Gaps = 25/106 (23%)

Query  144  RRRAIWAIYVWCRRTDELVD-------GPNANYITPTALDRWEKRLEDLFTGRPYDMLDA  196
            +R  I+  Y W  R D L +           N I PT      +R     TGR    +
Sbjct  294  KRHGIFVAY-WIDRFDFLTEYFQKNDQAIGCNEIVPTEESHRSRRHP--MTGRT---MHT  347

Query  197  ALSDTISRFPI-DIQP-----------FRDMIEGMRSDLRKTRYNN  230
             L+DT+S  PI  +Q            F+  IEG+R+D  K  YN+
Sbjct  348  FLADTVSLLPIRQLQSLFNEKMETNAIFKKAIEGVRNDEFKELYND  393


>gi|6900304|emb|CAB71342.1| Lep d 10 protein [Lepidoglyphus destructor]
 gi|14423956|sp|Q9NFZ4|TPM_LEPDS Tropomyosin (Allergen Lep d 10)
Length=284

 Score = 25.4 bits (54),  Expect = 2.8
 Identities = 22/95 (23%), Positives = 41/95 (43%), Gaps = 10/95 (10%)

Query  264  ESVYSAALALGIANQLTNILRDVGEDARR--GRIYLPQDELAQAGLSDEDI--------F  313
            E   S   A G    L   ++ + ED  R  GR+ +   +L +A  S ++
Sbjct  73   EKEKSLQTAEGDVAALNRRIQLIEEDLERSEGRLKIATSKLEEASQSADESERMRKMLEH  132

Query  314  KGVVTNRWRNFMKRQIKRARMFFEEAERGVTELSQ  348
            + +        ++ Q+K ARM  E+A+R   E+++
Sbjct  133  RSITDEERMEGLESQLKEARMMAEDADRKYDEVAR  167
```

```
>gi|2739154|gb|AAC67308.1| 22.6 kDa tegumental antigen [Schistosoma japonicum]
Length=191

 Score = 24.3 bits (51),  Expect = 6.3
 Identities = 18/62 (29%), Positives = 26/62 (41%), Gaps = 7/62 (11%)

Query  44   CSLLGLHPWEAGRPSPAV-------YSSLPVNPAGEAVVSSEQKVYDVVLKQAALLKRQL  96
            C   GL  WE  R    +        S+LP++    A   S+ K Y++  K   LL +
Sbjct  69   CRGFGLKVWEVRREKEELKKDKEGKVSTLPLDIQIIAATMSKAKQYNICCKFKELLDKTS  128

Query  97   RT  98
            RT
Sbjct  129  RT  130


>gi|2440053|emb|CAA75141.1| tropomyosin [Dermatophagoides pteronyssinus]
 gi|14423954|sp|O18416|TPM_DERPT Tropomyosin (Allergen Der p 10)
Length=284

 Score = 24.3 bits (51),  Expect = 6.3
 Identities = 13/47 (27%), Positives = 25/47 (53%), Gaps = 9/47 (19%)

Query  302  LAQAGLSDEDIFKGVVTNRWRNFMKRQIKRARMFFEEAERGVTELSQ  348
            L    ++DE+  +G         ++ Q+K ARM  E+A+R   E+++
Sbjct  130  LEHRSITDEERMEG---------LENQLKEARMMAEDADRKYDEVAR  167


>gi|2353266|gb|AAB69424.1| tropomyosin [Dermatophagoides pteronyssinus]
Length=284

 Score = 24.3 bits (51),  Expect = 6.3
 Identities = 13/47 (27%), Positives = 25/47 (53%), Gaps = 9/47 (19%)

Query  302  LAQAGLSDEDIFKGVVTNRWRNFMKRQIKRARMFFEEAERGVTELSQ  348
            L    ++DE+  +G         ++ Q+K ARM  E+A+R   E+++
Sbjct  130  LEHRSITDEERMEG---------LENQLKEARMMAEDADRKYDEVAR  167


>gi|5832805|emb|CAB55143.1| Hypothetical protein Y116A8C.8 [Caenorhabditis
elegans] Length=247

 Score = 23.9 bits (50),  Expect = 8.3
 Identities = 17/50 (34%), Positives = 25/50 (50%), Gaps = 8/50 (16%)

Query  339  AERGVTELSQASRWPVWASLLLYRQILDEIEANDYNNFTKRAYVGKGKKL  388
            A+RG  E+++  RWP  + L    Q L E   N+Y N      +G GK +
Sbjct  123  AKRG--EIAEFRRWPRKSCLGHDMQRLRETYPNEYGN------IGDGKNI  164


>gi|14423976|sp|Q9NAS5|TPM_ANISI Tropomyosin (Allergen Ani s 3)
Length=284

 Score = 23.9 bits (50),  Expect = 8.3
 Identities = 10/26 (38%), Positives = 18/26 (69%), Gaps = 0/26 (0%)

Query  323  NFMKRQIKRARMFFEEAERGVTELSQ  348
            N ++ Q+K A+M  EEA+R   E+++
Sbjct  142  NTVESQLKEAQMLAEEADRKYDEVAR  167


>gi|4468639|emb|CAB38086.1| tropomyosin [Periplaneta americana]
 gi|14423957|sp|Q9UB83|TPM_PERAM Tropomyosin (Major allergen Per a 7)
Length=284

 Score = 23.9 bits (50),  Expect = 8.3
 Identities = 10/35 (28%), Positives = 20/35 (57%), Gaps = 0/35 (0%)

Query  314  KGVVTNRWRNFMKRQIKRARMFFEEAERGVTELSQ  348
            KG+      + ++ Q+K AR   EEA++   E+++
Sbjct  133  KGLADEERMDALENQLKEARFMAEEADKKYDEVAR  167
```

```
>gi|4378573|gb|AAD19606.1| tropomyosin [Periplaneta americana]
Length=284

 Score = 23.9 bits (50),  Expect = 8.3
 Identities = 10/35 (28%), Positives = 20/35 (57%), Gaps = 0/35 (0%)

Query  314  KGVVTNRWRNFMKRQIKRARMFFEEAERGVTELSQ  348
            KG+      + ++ Q+K AR   EEA++   E+++
Sbjct  133  KGLADEERMDALENQLKEARFMAEEADKKYDEVAR  167



Database: All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF
excluding environmental samples
    Posted date:  Apr 4, 2006  3:06 AM
  Number of letters in database: 583,171
  Number of sequences in database:  2,529
Lambda     K      H
   0.321    0.135    0.407
Gapped
Lambda     K      H
   0.267   0.0410    0.140
Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 2529
Number of Hits to DB: 25971
Number of extensions: 1070
Number of successful extensions: 3
Number of sequences better than 10: 0
Number of HSP's better than 10 without gapping: 0
Number of HSP's gapped: 3
Number of HSP's successfully gapped: 0
Length of query: 410
Length of database: 583171
Length adjustment: 79
Effective length of query: 331
Effective length of database: 583171
Effective search space: 193029601
Effective search space used: 126898780
T: 11
A: 40
X1: 16 (7.4 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (20.4 bits)
S2: 50 (23.9 bits)
```

Appendix 4. PMI: Sequence and Alignment Data

Amino acid sequence of the PMI protein. The construct and sequences are referenced in Paine et al., 2005.

This protein is the PMI marker from E. coli (Miles and Guest 1984).
GET REFERENCE>>>>>

```
1   MQKLINSVQNYAWGSKTALTELYGMENPSSQPMAELWMGAHPKSSSRVQNAAGDIVSLRD
61  VIESDKSTLLGEAVAKRFGELPFLFKVLCAAQPLSIQVHPNKHNSEIGFAKENAAGIPMD
121 AAERNYKDPNHKPELVFALTPFLAMNAFREFSEIVSLLQPVAGAHPAIAHFLQQPDAERL
181 SELFASLLNMQGEEKSRALAILKSALDSQQGEPWQTIRLISEFYPEDSGLFSPLLLNVVK
241 LNPGEAMFLFAETPHAYLQGVALEVMANSDNVLRAGLTPKYIDIPELVANVKFEAKPANQ
301 LLTQPVKQGAELDFPIPVDDFAFSLHDLSDKETTISQQSAAILFCVEGDATLWKGSQQLQ
361 LKPGESAFIAANESPVTVKGHGRLARVYNKL
```

```
LOCUS       P00946                   391 aa            linear   BCT 21-JUL-1986
DEFINITION  Mannose-6-phosphate isomerase (Phosphomannose isomerase) (PMI)
            (Phosphohexomutase).
ACCESSION   P00946
VERSION     P00946  GI:126724
DBSOURCE    swissprot: locus MANA_ECOLI, accession P00946;
            class: standard.
            xrefs: M15380.1, AAA24109.1, U00096.2, AAC74685.1, AP009048.1,
            BAA15361.1, U35067.1, AAA79050.1, ISECMP
            xrefs (non-sequence databases): HSSP:P34948, IntAct:P00946,
            SWISS-2DPAGE:P00946, GenomeReviews:U00096_GR, EchoBASE:EB0561,
            EcoGene:EG10566, BioCyc:EcoCyc:MANNPISOM-MONOMER,
            InterPro:IPR011051, InterPro:IPR001250, PANTHER:PTHR10309,
            Pfam:PF01238, PRINTS:PR00714, TIGRFAMs:TIGR00218, PROSITE:PS00965,
            PROSITE:PS00966
KEYWORDS    Complete proteome; Isomerase; Metal-binding; Zinc.
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1  (residues 1 to 391)
  AUTHORS   Miles,J.S. and Guest,J.R.
  TITLE     Nucleotide sequence and transcriptional start point of the
            phosphomannose isomerase gene (manA) of Escherichia coli
  JOURNAL   Gene 32 (1-2), 41-48 (1984)
   PUBMED   6397402
  REMARK    NUCLEOTIDE SEQUENCE [GENOMIC DNA].
            STRAIN=K12
REFERENCE   2  (residues 1 to 391)
  AUTHORS   Blattner,F.R., Plunkett,G. III, Bloch,C.A., Perna,N.T., Burland,V.,
            Riley,M., Collado-Vides,J., Glasner,J.D., Rode,C.K., Mayhew,G.F.,
            Gregor,J., Davis,N.W., Kirkpatrick,H.A., Goeden,M.A., Rose,D.J.,
            Mau,B. and Shao,Y.
  TITLE     The complete genome sequence of Escherichia coli K-12
  JOURNAL   Science 277 (5331), 1453-1474 (1997)
   PUBMED   9278503
  REMARK    NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
            STRAIN=K12 / MG1655
REFERENCE   3  (residues 1 to 391)
  AUTHORS   Aiba,H., Baba,T., Fujita,K., Hayashi,K., Inada,T., Isono,K.,
            Itoh,T., Kasai,H., Kashimoto,K., Kimura,S., Kitakawa,M.,
            Kitagawa,M., Makino,K., Miki,T., Mizobuchi,K., Mori,H., Mori,T.,
            Motomura,K., Nakade,S., Nakamura,Y., Nashimoto,H., Nishio,Y.,
            Oshima,T., Saito,N., Sampei,G., Seki,Y., Sivasundaram,S.,
            Tagami,H., Takeda,J., Takemoto,K., Takeuchi,Y., Wada,C.,
            Yamamoto,Y. and Horiuchi,T.
  TITLE     A 570-kb DNA sequence of the Escherichia coli K-12 genome
            corresponding to the 28.0-40.1 min region on the linkage map
```

```
  JOURNAL   DNA Res. 3 (6), 363-377 (1996)
   PUBMED   9097039
  REMARK    NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
            STRAIN=K12 / W3110
REFERENCE   4  (residues 1 to 391)
  AUTHORS   Robison,K., O'Keeffe,T. and Church,G.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (??-AUG-1995)
  REMARK    NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 378-391.
            STRAIN=K12 / EMG2
COMMENT     On May 27, 2005 this sequence version replaced gi:68444.
            [FUNCTION] Involved in the conversion of glucose to GDP-L-fucose,
            which can be converted to L-fucose, a capsular polysaccharide.
            [CATALYTIC ACTIVITY] D-mannose 6-phosphate = D-fructose
            6-phosphate.
            [COFACTOR] Binds 1 zinc ion per subunit (By similarity).
            [INTERACTION] P18196:minC; NbExp=1; IntAct=EBI-554045, EBI-554060;
            P75864:ycbY; NbExp=1; IntAct=EBI-554045, EBI-547718; P76161:ydfT;
            NbExp=1; IntAct=EBI-554045, EBI-554053; P45748:yrdC; NbExp=1;
            IntAct=EBI-554045, EBI-554067.
            [SUBCELLULAR LOCATION] Cytoplasm (Probable).
            [SIMILARITY] Belongs to the mannose-6-phosphate isomerase type 1
            family.
            -------------------------------------------------------------------
            ----Copyrighted by the UniProt Consortium, see
            http://www.uniprot.org/terms Distributed under the Creative Commons
            Attribution-NoDerivs License
            -------------------------------------------------------------------
            ----.
FEATURES             Location/Qualifiers
     source          1..391
                     /organism="Escherichia coli"
                     /db_xref="taxon:562"
     gene            1..391
                     /gene="manA"
                     /locus_tag="b1613"
                     /note="synonym: pmi"
     Protein         1..391
                     /gene="manA"
                     /locus_tag="b1613"
                     /product="Mannose-6-phosphate isomerase"
                     /EC_number="5.3.1.8"
     Region          1..391
                     /gene="manA"
                     /locus_tag="b1613"
                     /region_name="Mature chain"
                     /experiment="experimental evidence, no additional details
                     recorded"
                     /note="Mannose-6-phosphate isomerase.
                     /FTId=PRO_0000194230."
     Region          1..357
                     /gene="manA"
                     /locus_tag="b1613"
                     /region_name="Phosphomannose isomerase type I"
                     /note="PMI_typeI"
                     /db_xref="CDD:23170"
     Site            97
                     /gene="manA"
                     /locus_tag="b1613"
                     /site_type="metal-binding"
                     /inference="non-experimental evidence, no additional
                     details recorded"
                     /note="Zinc (By similarity)."
     Site            99
                     /gene="manA"
                     /locus_tag="b1613"
                     /site_type="metal-binding"
                     /inference="non-experimental evidence, no additional
                     details recorded"
                     /note="Zinc (By similarity)."
```

```
     Site            134
                     /gene="manA"
                     /locus_tag="b1613"
                     /site_type="metal-binding"
                     /inference="non-experimental evidence, no additional
                     details recorded"
                     /note="Zinc (By similarity)."
     Site            255
                     /gene="manA"
                     /locus_tag="b1613"
                     /site_type="metal-binding"
                     /inference="non-experimental evidence, no additional
                     details recorded"
                     /note="Zinc (By similarity)."
     Site            274
                     /gene="manA"
                     /locus_tag="b1613"
                     /site_type="active"
                     /inference="non-experimental evidence, no additional
                     details recorded"
                     /note="By similarity."
ORIGIN
        1 mqklinsvqn yawgsktalt elygmenpss qpmaelwmga hpksssrvqn aagdivslrd
       61 viesdkstll geavakrfge lpflfkvlca aqplsiqvhp nkhnseigfa kenaagipmd
      121 aaernykdpn hkpelvfalt pflamnafre fseivsllqp vagahpaiah flqqpdaerl
      181 selfasllnm qgeeksrala ilksaldsqq gepwqtirli sefypedsgl fsplllnvvk
      241 lnpgeamflf aetphaylqg valevmansd nvlragltpk yidipelvan vkfeakpanq
      301 lltqpvkqga eldfpipvdd fafslhdlsd kettisqqsa ailfcvegda tlwkgsqqlq
      361 lkpgesafia anespvtvkg hgrlarvynk l
//
```

Search Full-length using FASTA3 Query PMI against AllergenOnline Database v6.0 (January, 2006)

```
Query Sequence:
>PMI marker from E. coli
MQKLINSVQNYAWGSKTALTELYGMENPSSQPMAELWMGAHPKSSSRVQNAAGDIVSLRD
VIESDKSTLLGEAVAKRFGELPFLFKVLCAAQPLSIQVHPNKHNSEIGFAKENAAGIPMD
AAERNYKDPNHKPELVFALTPFLAMNAFREFSEIVSLLQPVAGAHPAIAHFLQQPDAERL
SELFASLLNMQGEEKSRALAILKSALDSQQGEPWQTIRLISEFYPEDSGLFSPLLLNVVK
LNPGEAMFLFAETPHAYLQGVALEVMANSDNVLRAGLTPKYIDIPELVANVKFEAKPANQ
LLTQPVKQGAELDFPIPVDDFAFSLHDLSDKETTISQQSAAILFCVEGDATLWKGSQQLQ
LKPGESAFIAANESPVTVKGHGRLARVYNKL

Searching: AllergenOnline Database v6.0 (January, 2006)

# fasta34t25b1/fasta34 -a -Q -3 -E10 -d5 -B /tmp/allergen16600.tmp allergen6.fasta
FASTA searches a protein or DNA sequence data bank
 version 3.4t25 Nov 12, 2004
Please cite:
 W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

Query library /tmp/allergen16600.tmp vs allergen6.fasta library
searching allergen6.fasta library

  1>>>PMI marker from E. coli - 391 aa
 vs  allergen6.fasta library

       opt      E()
< 20     4     0:==
  22     0     0:          one = represents 3 library sequences
```

```
  24     0     0:
  26     0     0:
  28     1     0:=
  30     4     2:*=
  32    18     8:==*===
  34    25    22:=======*=
  36    45    46:===============*
  38    73    75:========================*
  40   114   105:==================================*===
  42   110   129:=====================================     *
  44   127   142:===========================================    *
  46   123   144:=========================================      *
  48   120   138:========================================     *
  50   103   126:===================================      *
  52   163   111:====================================*==================
  54   126    95:===============================*==========
  56    76    79:==========================*
  58    50    65:=================    *
  60    39    53:=============    *
  62    56    42:=============*=====
  64    25    34:=========  *
  66    28    27:========*=
  68    29    21:======*===
  70    20    16:=====*=
  72    15    13:====*
  74     8    10:===*
  76     5     8:==*
  78     4     6:=*
  80     4     5:=*
  82     1     4:=*
  84     6     3:*=
  86     2     2:*
  88     2     2:*          inset = represents 1 library sequences
  90     2     1:*
  92     2     1:*         :*=
  94     2     1:*         :*=
  96     0     1:*         :*
  98     1     0:=         *=
 100     0     0:          *
 102     0     0:          *
 104     2     0:=         *==
 106     0     0:          *
 108     1     0:=         *=
 110     1     0:=         *=
 112     0     0:          *
 114     0     0:          *
 116     0     0:          *
 118     0     0:          *
>120     0     0:          *
 336538 residues in  1537 sequences
  Expectation_n fit: rho(ln(x))= 4.0320+/-0.00356; mu= 15.9742+/- 0.183
 mean_var=60.9582+/-16.593, 0's: 4 Z-trim: 6  B-trim: 81 in 1/46
 Lambda= 0.164270
 Kolmogorov-Smirnov  statistic: 0.0489 (N=29) at  50

FASTA (3.47 Mar 2004) function [optimized, BL50 matrix (15:-5)] ktup: 2
 join: 37, opt: 25, open/ext: -10/-2, width:  16
 Scan time:  0.120
The best scores are:                                      opt z-sc E(1537)
gi|18770|emb|CAA45777.1| trypsin inhibitor sub ( 217) [f]   85  110.6    0.36
gi|18772|emb|CAA45778.1| trypsin inhibitor sub ( 217) [f]   83  108.1     0.5
gi|1304264|dbj|BAA12318.1| alpha-gliadin [Trit ( 259) [f]   81  104.6    0.79
gi|256429|gb|AAB23464.1| Kunitz trypsin inhibi ( 216) [f]   80  104.2    0.82
gi|1076486|pir||S48032 cim1 protein - soybean  ( 277) [f]   76   97.8     1.9
gi|170740|gb|AAA34290.1| gliadin               ( 296) [f]   74   94.9     2.7
gi|170724|gb|AAA34282.1| pre-alpha-/beta-gliad ( 297) [f]   74   94.9     2.7
gi|14279169|gb|AAK58515.1|AF249675_1 beta-1,3- ( 460) [f]   74   92.7     3.6
gi|2739154|gb|AAC67308.1| 22.6 kDa tegumental  ( 191) [f]   70   92.1     3.9
gi|886215|gb|AAB08720.1| major allergen BDA20  ( 172) [f]   69   91.3     4.3
gi|18542115|gb|AAL75450.1|AF465613_1 minor all ( 636) [f]   74   91.0     4.5
```

```
gi|256636|gb|AAB23483.1| Kunitz trypsin inhibi ( 204) [f]   68   89.2     5.7
gi|433317|gb|AAA20541.1| beta-galactoside-bind ( 280) [f]   69   88.8     5.9
gi|170728|gb|AAA34284.1| alpha-type gliadin    ( 186) [f]   66   87.1     7.4
gi|162794|gb|AAA30429.1| alpha-S1-casein       ( 214) [f]   66   86.4     8.1
gi|21761|emb|CAA26384.1| unnamed protein produ ( 286) [f]   66   84.9     9.9
gi|21755|emb|CAA25593.1| unnamed protein produ ( 286) [f]   66   84.9     9.9
gi|170720|gb|AAA34280.1| alpha/beta-gliadin pr ( 286) [f]   66   84.9     9.9

>>gi|18770|emb|CAA45777.1| trypsin inhibitor subtype A [  (217 aa)
 initn:  37 init1:  37 opt:  85  Z-score: 110.6  bits: 28.4 E(): 0.36
Smith-Waterman score: 85;  29.605% identity (49.342% similar) in 152
aa overlap (246-387:8-148)

               10        20        30        40        50        60
PMI    MQKLINSVQNYAWGSKTALTELYGMENPSSQPMAELWMGAHPKSSSRVQNAAGDIVSLRD

               70        80        90       100       110       120
PMI    VIESDKSTLLGEAVAKRFGELPFLFKVLCAAQPLSIQVHPNKHNSEIGFAKENAAGIPMD

              130       140       150       160       170       180
PMI    AAERNYKDPNHKPELVFALTPFLAMNAFREFSEIVSLLQPVAGAHPAIAHFLQQPDAERL

              190       200       210       220       230       240
PMI    SELFASLLNMQGEEKSRALAILKSALDSQQGEPWQTIRLISEFYPEDSGLFSPLLLNVVK

gi|187                                                           MK


              250        260       270       280       290
PMI    LNPGEAMFLFAETPHAYL-QGVALEVMANSDNVLRAGLTPKYIDIPELVANVKFEAKPA-
            :.:::     .:: ...:  :. :  : :. : :  :  . ...:   ..: :.
gi|187 STIFFALFLFCAFTTSYLPSAIADFVLDNEGNPLENGGT--YYILSDITAFGGIRAAPTG
             10        20        30        40          50        60

      300        310              320       330       340       350
PMI    NQLLTQPVKQG-AELDFPI------PVD-DFAFSLHDLSDKETTISQQSAAILFCVEGDA
       :.     : :.  :::  :      :    :    : :: :  ..    :.:..:: :
gi|187 NERCPLTVVQSRNELDKGIGTIISSPYRIRFIAEGHPLSLKFDSF----AVIMLCV-GIP
               70        80        90       100           110

              360       370       380       390
PMI    TLWKGSQQLQLKPGESAFIAANESPVTVKGHGRLARVYNKL
       : :.  ..:   :.    .  .:.  .. :  :: ::
gi|187 TEWSVVEDLPEGPA----VKIGENKDAMDGWFRLERVSDDEFNNYKLVFCPQQAEDDKCG
         120           130       140       150       160       170

gi|187 DIGISIDHDDGTRRLVVSKNKPLVVQFQKLDKESLAKKNHGLSRSE
             180       190       200       210

>>gi|18772|emb|CAA45778.1| trypsin inhibitor subtype B [  (217 aa)
 initn:  37 init1:  37 opt:  83  Z-score: 108.1  bits: 27.9 E():  0.5
Smith-Waterman score: 83;  27.703% identity (54.054% similar) in 148
aa overlap (246-387:8-148)

               10        20        30        40        50        60
PMI    MQKLINSVQNYAWGSKTALTELYGMENPSSQPMAELWMGAHPKSSSRVQNAAGDIVSLRD

               70        80        90       100       110       120
PMI    VIESDKSTLLGEAVAKRFGELPFLFKVLCAAQPLSIQVHPNKHNSEIGFAKENAAGIPMD

              130       140       150       160       170       180
PMI    AAERNYKDPNHKPELVFALTPFLAMNAFREFSEIVSLLQPVAGAHPAIAHFLQQPDAERL

              190       200       210       220       230       240
PMI    SELFASLLNMQGEEKSRALAILKSALDSQQGEPWQTIRLISEFYPEDSGLFSPLLLNVVK

gi|187                                                           MK


              250        260       270       280       290
```

```
PMI    LNPGEAMFLFAETPHAYL-QGVALEVMANSDNVLRAGLTPKYIDIPELVANVKFEAKPA-
            :.:::     .:: ...:  :. :  : : .: :  :  . ...:   ..: :.
gi|187 STIFFALFLFCAFTTSYLPSAIADFVLDNEGNPLDSGGT--YYILSDITAFGGIRAAPTG
             10        20        30        40          50        60

      300        310        320       330         340       350
PMI    NQLLTQPVKQG-AELDFPI-PVDDFAFSLHDLSD-KETTISQQS-AAILFCVEGDATLWK
       :.     : :.  :::  :  . .  : .. ... .   .. .: :.:..:: :  : :.
gi|187 NERCPLTVVQSRNELDKGIGTIISSPFRIRFIAEGNPLRLKFDSFAVIMLCV-GIPTEWS
               70        80        90       100       110

          360       370       380       390
PMI    GSQQLQLKPGESAFIAANESPVTVKGHGRLARVYNKL
         ..:   :.    .  .:.  .: :  :. ::
gi|187 VVEDLPEGPA----VKIGENKDAVDGWFRIERVSDDEFNNYKLVFCTQQAEDDKCGDIGI
     120           130       140       150       160       170

gi|187 SIDHDDGTRRLVVSKNKPLVVQFQKVDKESLAKKNHGLSRSE
         180       190       200       210

>>gi|1304264|dbj|BAA12318.1| alpha-gliadin [Triticum aes  (259 aa)
 initn:  76 init1:  53 opt:  81  Z-score: 104.6  bits: 27.6 E(): 0.79
Smith-Waterman score: 81;  24.390% identity (56.098% similar) in 82
aa overlap (279-360:85-161)

               10        20        30        40        50        60
PMI    MQKLINSVQNYAWGSKTALTELYGMENPSSQPMAELWMGAHPKSSSRVQNAAGDIVSLRD

               70        80        90       100       110       120
PMI    VIESDKSTLLGEAVAKRFGELPFLFKVLCAAQPLSIQVHPNKHNSEIGFAKENAAGIPMD

              130       140       150       160       170       180
PMI    AAERNYKDPNHKPELVFALTPFLAMNAFREFSEIVSLLQPVAGAHPAIAHFLQQPDAERL

              190       200       210       220       230       240
PMI    SELFASLLNMQGEEKSRALAILKSALDSQQGEPWQTIRLISEFYPEDSGLFSPLLLNVVK

gi|130               VRVPVPQLQPQNPSQQQPQEQVPLVQQQQFLGQQQPFPPQQPYPQP
                             10        20        30        40

              250       260       270       280       290       300
PMI    LNPGEAMFLFAETPHAYLQGVALEVMANSDNVLRAGLTPKYIDIPELVANVKFEAKPANQ
                                             :.: .  : ... . . .  .:
gi|130 QPFPSQQPYLQLQPFPQPQLPYSQPQPFRPQQPYPQPQPQYSQPQEPISQQQQQQQQQQQ
         50        60        70        80        90       100

              310       320       330       340       350       360
PMI    LLTQPVKQGAELDFPIPVDDFAFSLHDLSDKETTISQQSAAILFCVEGDATLWKGSQQLQ
       .: : ..:  .:   ::  : ... :...  .. . :::.  :.       ::.  .: :
gi|130 ILQQILQQ--QL---IPCMDVVLQQHNIAHGRSQVLQQSTYQLLQELCCQHLWQIPEQSQ
        110            120       130       140       150       160

              370       380       390
PMI    LKPGESAFIAANESPVTVKGHGRLARVYNKL

gi|130 CQAIQNVVHAIILHQQQKQQQQPSSQVSFQQPLQQYPLGQGSFRPSQQNPQDQGSVQPQQ
             170       180       190       200       210       220

gi|130 LPQFEEIRNLALQTLPAMCNVYIPPYCTIAPFGIFGTN
             230       240       250

>>gi|256429|gb|AAB23464.1| Kunitz trypsin inhibitor; KTi  (216 aa)
 initn:  32 init1:  32 opt:  80  Z-score: 104.2  bits: 27.2 E(): 0.82
Smith-Waterman score: 80;  29.139% identity (49.007% similar) in 151
aa overlap (247-387:8-147)

               10        20        30        40        50        60
PMI    MQKLINSVQNYAWGSKTALTELYGMENPSSQPMAELWMGAHPKSSSRVQNAAGDIVSLRD

               70        80        90       100       110       120
```

```
PMI    VIESDKSTLLGEAVAKRFGELPFLFKVLCAAQPLSIQVHPNKHNSEIGFAKENAAGIPMD

              130       140       150       160       170       180
PMI    AAERNYKDPNHKPELVFALTPFLAMNAFREFSEIVSLLQPVAGAHPAIAHFLQQPDAERL

              190       200       210       220       230       240
PMI    SELFASLLNMQGEEKSRALAILKSALDSQQGEPWQTIRLISEFYPEDSGLFSPLLLNVVK

gi|256                                                            M


              250        260       270       280       290
PMI    LNPGEAMFLFAETPHAYL-QGVALEVMANSDNVLRAGLTPKYIDIPELVANVKFEAKPA-
             .:::     .:: ...:  :. :  : :. : :  :  . ...:   ..: :.
gi|256 KSTIFFLFLFCAFTTSYLPSAIADFVLDNEGNPLENGGT--YYILSDITAFGGIRAAPTG
              10        20        30        40          50

      300        310              320       330       340       350
PMI    NQLLTQPVKQG-AELDFPI------PVD-DFAFSLHDLSDKETTISQQSAAILFCVEGDA
       :.     : :.  :::  :      :    :    : :: :  ..    :.:..:: :
gi|256 NERCPLTVVQSRNELDKGIGTIISSPYRIRFIAEGHPLSLKFDSF----AVIMLCV-GIP
      60        70        80        90       100           110

              360       370       380       390
PMI    TLWKGSQQLQLKPGESAFIAANESPVTVKGHGRLARVYNKL
       : :.  ..:   :.    .  .:.  .. :  :: ::
gi|256 TEWSVVEDLPEGPA----VKIGENKDAMDGWFRLERVSDDEFNNYKLVFCPQQAEDDKCG
          120           130       140       150       160       170

gi|256 DIGISIDHDDGTRRLVVSKNKPLVVQFQKLDKESLAKKNHGLSRSE
              180       190       200       210

>>gi|1076486|pir||S48032 cim1 protein - soybean           (277 aa)
 initn:  78 init1:  52 opt:  76  Z-score: 97.8  bits: 26.4 E(): 1.9
Smith-Waterman score: 76;  22.581% identity (54.839% similar) in 93
aa overlap (176-268:182-274)

                     10        20        30        40        50
PMI          MQKLINSVQNYAWGSKTALTELYGMENPSSQPMAELWMGAHPKSSSRVQNAAGD

gi|107 MALTLQRALSELLTLIAILSIFLVIPSFCFNPKKLYNASYYSPSSSDWSPAVATWYGPAN
               10        20        30        40        50        60

           60        70        80        90       100       110
PMI    IVSLRDVIESDKSTLLGEAVAKRFGELPFLFKVLCAAQPLSIQVHPNKHNSEIGFAKENA

gi|107 GDGSEGGACGYGNAVGQPPFSSLISAGSPLIYDSGKGCGSCYEVKCTGNSACSGNPVKVV
               70        80        90       100       110       120

          120       130       140       150       160       170
PMI    AGIPMDAAERNYKDPNHKPELVFALTPFLAMNAFREFSEIVSLLQPVAGAHPAIAHFLQQ

gi|107 ITDECAGCGSDAQYHFDLSGNAFGAMAISGQDENLRNAGKINIQHRRIECNYPGRSIAFH
              130       140       150       160       170       180

          180       190       200       210       220       230
PMI    PDAERLSELFASLLNMQGEEKSRALAILKSALDSQQGEPWQTIRLISEFYPEDSGLFSPL
        :.   .: ::.:....  . . : . :: :::: . .  :        . . : : .:.
gi|107 VDSGSNQEYFATLVEYEDGDGDLAKVELKEALDSGSWDSMQQSWGAVWKFDKGSPLRAPF
              190       200       210       220       230       240

          240       250       260       270       280       290
PMI    LLNVVKLNPGEAMFLFAETPHAYLQGVALEVMANSDNVLRAGLTPKYIDIPELVANVKFE
        .... :. :...      : ..  : . . ..:
gi|107 SIKLTTLESGQTIVANNVIPAGWTPGQTYRSIVNFAT
              250       260       270

          300       310       320       330       340       350
PMI    AKPANQLLTQPVKQGAELDFPIPVDDFAFSLHDLSDKETTISQQSAAILFCVEGDATLWK
```

```
          360       370       380       390
PMI    GSQQLQLKPGESAFIAANESPVTVKGHGRLARVYNKL



391 residues in 1 query   sequences
336538 residues in 1537 library sequences
 Scomplib [34t25]
 start: Thu Apr  6 16:50:38 2006 done: Thu Apr  6 16:50:39 2006
 Total Scan time:  0.120 Total Display time:  0.010

Function used was FASTA [version 3.4t25 Nov 12, 2004]
```

Search 80-mer Sliding Window using FASTA3 with Query PMI against AllergenOnline Database v6.0 (January, 2006)

Database	AllergenOnline v6.0 (January, 2006)
Input Query	>PMI marker from E. coli MQKLINSVQNYAWGSKTALTELYGMENPSSQPMAELWMGAHPKSSSRVQNAAGDIVSLRD VIESDKSTLLGEAVAKRFGELPFLFKVLCAAQPLSIQVHPNKHNSEIGFAKENAAGIPMD AAERNYKDPNHKPELVFALTPFLAMNAFREFSEIVSLLQPVAGAHPAIAHFLQQPDAERL SELFASLLNMQGEEKSRALAILKSALDSQQGEPWQTIRLISEFYPEDSGLFSPLLLNVVK LNPGEAMFLFAETPHAYLQGVALEVMANSDNVLRAGLTPKYIDIPELVANVKFEAKPANQ LLTQPVKQGAELDFPIPVDDFAFSLHDLSDKETTISQQSAAILFCVEGDATLWKGSQQLQ LKPGESAFIAANESPVTVKGHGRLARVYNKL
Length	391
Number of 80 mers	312

Hit #	**Defline**	**Best %ID**	**# Hits > 35%**	**Full Alignment**			**Links**	
				E-val	**%ID**	**length**	**NCBI**	**Details**

NOTE: There were no matched identities of 35% or more over any 80-amino acid segment of PMI.

BLASTP Search with Query PMI against the NCBI Database Available on Thursday, April 06, 2006, using a key word delimiter “allergen” (limited search to 2,529 sequences).

```
Thursday, April 06, 2006

BLASTP 2.2.13 [Nov-27-2005]
Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
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: Schäffer, Alejandro A., L. Aravind, Thomas L. Madden, Sergei
Shavirin, John L. Spouge, Yuri I. Wolf, Eugene V. Koonin, and
Stephen F. Altschul (2001), "Improving the accuracy of PSI-BLAST
protein database searches with composition-based statistics
and other refinements", Nucleic Acids Res. 29:2994-3005.

RID: 1144362899-20844-112600970753.BLASTQ4


Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF excluding environmental samples
           3,538,507 sequences; 1,215,510,009 total letters
Query=  PMI marker from E. coli
Length=391


                                                                   Score     E
Sequences producing significant alignments:                        (Bits)  Value

gi|2739154|gb|AAC67308.1|  22.6 kDa tegumental antigen [Schistoso  25.8    2.3
gi|14279169|gb|AAK58515.1|  beta-1,3-glucanase-like protein [Olea  24.3    6.6
gi|555616|gb|AAA50175.1|  cytokinin induced message                23.9    8.5

ALIGNMENTS
>gi|2739154|gb|AAC67308.1| 22.6 kDa tegumental antigen [Schistosoma japonicum]
Length=191

 Score = 25.8 bits (55),  Expect = 2.3, Method: Composition-based stats.
 Identities = 12/33 (36%), Positives = 17/33 (51%), Gaps = 0/33 (0%)

Query  192  GEEKSRALAILKSALDSQQGEPWQTIRLISEFY  224
            G+E       LK+ LDS+ G  WQ I L   ++
Sbjct  131  GDEVRAVANDLKAFLDSEYGRVWQVIILTGSYW  163


>gi|14279169|gb|AAK58515.1| beta-1,3-glucanase-like protein [Olea europaea]
Length=460

 Score = 24.3 bits (51),  Expect = 6.6, Method: Composition-based stats.
 Identities = 19/89 (21%), Positives = 41/89 (46%), Gaps = 11/89 (12%)

Query  172  LQQPDAERLSELFASLLNMQGEEKSRALAILKSALDSQQGEPWQTIRLISEFYPEDSGLF  231
            L   D + +S+L  ++ N+Q    + +L          +      + ++S+ YP  SG+F
Sbjct  125  LTSGDQKLISQLLPAMQNVQNALNAASLG------GKVKVSTVHAMAVLSQSYPPSSGVF  178

Query  232  SPLLLNVVKL-----NPGEAMFLFAETPH  255
            +P L + +K      +  +A F+ +  P+
Sbjct  179  NPGLGDTMKALLQFQSANDAPFMISPYPY  207


>gi|555616|gb|AAA50175.1| cytokinin induced message
Length=277
```

```
 Score = 23.9 bits (50),  Expect = 8.5, Method: Composition-based stats.
 Identities = 21/93 (22%), Positives = 36/93 (38%), Gaps = 0/93 (0%)

Query  176  DAERLSELFASLLNMQGEEKSRALAILKSALDSQQGEPWQTIRLISEFYPEDSGLFSPLL  235
            D+    E FA+L+  +  +   A   LK ALDS   +  Q        + + S L +P
Sbjct  182  DSGSNQEYFATLVEYEDGDGLAKVELKEALDSGSWDSMQQSWGAVWKFDKGSPLRAPFS  241

Query  236  LNVVKLNPGEAMFLFAETPHAYLQGVALEVMAN  268
            + +  L  G+ +      P  +  G     + N
Sbjct  242  IKLTTLESGQTIVANNVIPAGWTPGQTYRSIVN  274



Database: All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF
excluding environmental samples
    Posted date:  Apr 4, 2006  3:06 AM
  Number of letters in database: 583,171
  Number of sequences in database:  2,529
Lambda     K      H
   0.316    0.132    0.379
Gapped
Lambda     K      H
   0.267   0.0410    0.140
Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 2529
Number of Hits to DB: 23967
Number of extensions: 992
Number of successful extensions: 3
Number of sequences better than 10: 0
Number of HSP's better than 10 without gapping: 0
Number of HSP's gapped: 3
Number of HSP's successfully gapped: 0
Length of query: 391
Length of database: 583171
Length adjustment: 79
Effective length of query: 312
Effective length of database: 583171
Effective search space: 181949352
Effective search space used: 119614560
T: 11
A: 40
X1: 16 (7.3 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (20.4 bits)
S2: 50 (23.9 bits)
```

Appendix 5. Positive Control Ber e 1: Sequence and Alignment Data

Amino acid sequence of the Ber e 1 protein. The construct and sequences are referenced in Paine et al., 2005.

Positive Control sequence Ber e 1 (146 aa)see Altenbach et al., 1987.

```
1   MAKISVAAAALLVLMALGHATAFRATVTTTVVEEENQEECREQMQRQQMLSHCRMYMRQQ
61  MEESPYQTMPRRGMEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGEQMRRMM
121 RLAENIPSRCNLSPMRCPMGGSIAGF
```

```
LOCUS       P04403                   146 aa            linear   PLN 07-FEB-2006
DEFINITION  2S sulfur-rich seed storage protein precursor (Allergen Ber e 1)
            [Contains: 2S sulfur-rich seed storage protein small chain (2S
            albumin small subunit); 2S sulfur-rich seed storage protein large
            chain 1B (2S albumin large subunit)].
ACCESSION   P04403
VERSION     P04403  GI:112754
DBSOURCE    swissprot: locus 2SS_BEREX, accession P04403;
            class: standard.
            extra accessions:P04402,Q9LRC2,created: Mar 20, 1987.
            sequence updated: Oct 1, 1989.
            annotation updated: Feb 7, 2006.
            xrefs: M17146.1, AAA33010.1, X57027.1, CAA40343.1, X57028.1,
            CAA40344.1, X54490.1, CAA38362.1, X54491.1, CAA38363.1, AB044391.1,
            BAA96554.1, A13818.1, CAA01131.1, A25802, S14946, 1GYS_A, 1GYS_B
            xrefs (non-sequence databases): InterPro:IPR003612,
            InterPro:IPR006106, InterPro:IPR000617, Pfam:PF00234,
            PRINTS:PR00808, PRINTS:PR00496, SMART:SM00499
KEYWORDS    3D-structure; Allergen; Direct protein sequencing; Polymorphism;
            Pyrrolidone carboxylic acid; Seed storage protein; Signal; Storage
            protein.
SOURCE      Bertholletia excelsa (Brazil nut)
  ORGANISM  Bertholletia excelsa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; Ericales; Lecythidaceae; Bertholletia.
REFERENCE   1  (residues 1 to 146)
  AUTHORS   Altenbach,S.B., Pearson,K.W., Leung,F.W. and Sun,S.S.M.
  TITLE     Cloning and sequence analysis of a cDNA encoding a Brazil nut
            protein exceptionally rich in methionine
  JOURNAL   Plant Mol. Biol. 8, 239-250 (1987)
  REMARK    NUCLEOTIDE SEQUENCE [MRNA].
REFERENCE   2  (residues 1 to 146)
  AUTHORS   Bassuener,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (??-DEC-1990)
  REMARK    NUCLEOTIDE SEQUENCE [GENOMIC DNA].
REFERENCE   3  (residues 1 to 146)
  AUTHORS   Gander,E.S., Holmstroem,K.O., De Paiva,G.R., De Castro,L.A.,
            Carneiro,M. and Grossi de Sa,M.F.
  TITLE     Isolation, characterization and expression of a gene coding for a
            2S albumin from Bertholletia excelsa (Brazil nut)
  JOURNAL   Plant Mol. Biol. 16 (3), 437-448 (1991)
   PUBMED   1840683
  REMARK    NUCLEOTIDE SEQUENCE [GENOMIC DNA].
REFERENCE   4  (residues 1 to 146)
  AUTHORS   Yamauchi,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (??-JUN-2000)
```

```
  REMARK    NUCLEOTIDE SEQUENCE [GENOMIC DNA].
REFERENCE   5  (residues 1 to 146)
  AUTHORS   Ampe,C., Van Damme,J., de Castro,L.A., Sampaio,M.J., Van Montagu,M.
            and Vandekerckhove,J.
  TITLE     The amino-acid sequence of the 2S sulphur-rich proteins from seeds
            of Brazil nut (Bertholletia excelsa H.B.K.)
  JOURNAL   Eur. J. Biochem. 159 (3), 597-604 (1986)
   PUBMED   3758080
  REMARK    PROTEIN SEQUENCE OF 37-64 AND 70-142.
REFERENCE   6  (residues 1 to 146)
  AUTHORS   Alcocer,M.J., Murtagh,G.J., Bailey,K., Dumoulin,M., Meseguer,A.S.,
            Parker,M.J. and Archer,D.B.
  TITLE     The disulphide mapping, folding and characterisation of recombinant
            Ber e 1, an allergenic protein, and SFA8, two sulphur-rich 2S plant
            albumins
  JOURNAL   J. Mol. Biol. 324 (1), 165-175 (2002)
   PUBMED   12421566
  REMARK    3D-STRUCTURE MODELING, GLYCOSYLATION, AND DISULFIDE BONDS.
COMMENT     On or before Mar 15, 2005 this sequence version replaced gi:81558,
            gi:99608.
            [FUNCTION] This is a 2S seed storage protein.
            [SUBUNIT] The mature protein consists of a small and a large chain
            linked by disulfide bonds.
            [ALLERGEN] Causes an allergic reaction in human.
            [SIMILARITY] Belongs to the 2S seed storage albumins family.
FEATURES             Location/Qualifiers
     source          1..146
                     /organism="Bertholletia excelsa"
                     /db_xref="taxon:3645"
     gene            1..146
                     /gene="BE2S1"
     gene            1..146
                     /gene="BE2S2"
     Protein         1..146
                     /gene="BE2S1"
                     /product="2S sulfur-rich seed storage protein precursor"
     Region          1..22
                     /gene="BE2S1"
                     /region_name="Signal"
                     /experiment="experimental evidence, no additional details
                     recorded"
     Region          23..36
                     /gene="BE2S1"
                     /region_name="Propeptide"
                     /experiment="experimental evidence, no additional details
                     recorded"
                     /note="/FTId=PRO_0000032105."
     Region          37..64
                     /gene="BE2S1"
                     /region_name="Mature chain"
                     /experiment="experimental evidence, no additional details
                     recorded"
                     /note="2S sulfur-rich seed storage protein small chain.
                     /FTId=PRO_0000032106."
     Site            37
                     /gene="BE2S1"
                     /site_type="modified"
                     /experiment="experimental evidence, no additional details
                     recorded"
                     /note="Pyrrolidone carboxylic acid."
     Region          38..46
                     /gene="BE2S1"
                     /region_name="Helical region"
                     /experiment="experimental evidence, no additional details
                     recorded"
     Region          38..39
                     /gene="BE2S1"
                     /region_name="Conflict"
                     /experiment="experimental evidence, no additional details
                     recorded"
```

```
                     /note="EE -> QQ (in Ref. 5)."
     Bond            bond(40,92)
                     /gene="BE2S1"
                     /bond_type="disulfide"
                     /experiment="experimental evidence, no additional details
                     recorded"
                     /note="Interchain (between small and large chains)."
     Region          47
                     /gene="BE2S1"
                     /region_name="Hydrogen bonded turn"
                     /experiment="experimental evidence, no additional details
                     recorded"
     Region          51
                     /gene="BE2S1"
                     /region_name="Beta-strand region"
                     /experiment="experimental evidence, no additional details
                     recorded"
     Region          52..53
                     /gene="BE2S1"
                     /region_name="Hydrogen bonded turn"
                     /experiment="experimental evidence, no additional details
                     recorded"
     Bond            bond(53,81)
                     /gene="BE2S1"
                     /bond_type="disulfide"
                     /experiment="experimental evidence, no additional details
                     recorded"
                     /note="Interchain (between small and large chains)."
     Region          54..59
                     /gene="BE2S1"
                     /region_name="Helical region"
                     /experiment="experimental evidence, no additional details
                     recorded"
     Region          60..61
                     /gene="BE2S1"
                     /region_name="Hydrogen bonded turn"
                     /experiment="experimental evidence, no additional details
                     recorded"
     Region          65..69
                     /gene="BE2S1"
                     /region_name="Propeptide"
                     /experiment="experimental evidence, no additional details
                     recorded"
                     /note="/FTId=PRO_0000032107."
     Region          70..142
                     /gene="BE2S1"
                     /region_name="Mature chain"
                     /experiment="experimental evidence, no additional details
                     recorded"
                     /note="2S sulfur-rich seed storage protein large chain 1B.
                     /FTId=PRO_0000032108."
     Bond            bond(82,130)
                     /gene="BE2S1"
                     /bond_type="disulfide"
                     /experiment="experimental evidence, no additional details
                     recorded"
     Region          91
                     /gene="BE2S1"
                     /region_name="Variant"
                     /experiment="experimental evidence, no additional details
                     recorded"
                     /note="S -> E (in variant 1A)."
     Bond            bond(94,137)
                     /gene="BE2S1"
                     /bond_type="disulfide"
                     /experiment="experimental evidence, no additional details
                     recorded"
     Region          102..103
                     /gene="BE2S1"
                     /region_name="Conflict"
```

```
                     /experiment="experimental evidence, no additional details
                     recorded"
                     /note="MR -> RM (in Ref. 4)."
     Region          107
                     /gene="BE2S1"
                     /region_name="Conflict"
                     /experiment="experimental evidence, no additional details
                     recorded"
                     /note="E -> K (in Ref. 4)."
     Region          122
                     /gene="BE2S1"
                     /region_name="Conflict"
                     /experiment="experimental evidence, no additional details
                     recorded"
                     /note="L -> M (in Ref. 5)."
     Region          126
                     /gene="BE2S1"
                     /region_name="Conflict"
                     /experiment="experimental evidence, no additional details
                     recorded"
                     /note="I -> L (in Ref. 5)."
     Region          143..146
                     /gene="BE2S1"
                     /region_name="Propeptide"
                     /experiment="experimental evidence, no additional details
                     recorded"
                     /note="/FTId=PRO_0000032109."
ORIGIN
        1 makisvaaaa llvlmalgha tafratvttt vveeenqeec reqmqrqqml shcrmymrqq
       61 meespyqtmp rrgmephmse cceqlegmde scrceglrmm mmrmqqeemq prgeqmrrmm
      121 rlaenipsrc nlspmrcpmg gsiagf
//
```

Search Full-length using FASTA3 Query Ber e 1 against AllergenOnline Database v6.0 (January, 2006)

```
Query Sequence:
>Positive Control sequence Ber e 1 (146 aa)
MAKISVAAAALLVLMALGHATAFRATVTTTVVEEENQEECREQMQRQQMLSHCRMYMRQQ
MEESPYQTMPRRGMEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGEQMRRMM
RLAENIPSRCNLSPMRCPMGGSIAGF


Searching: AllergenOnline Database v6.0 (January, 2006)

# fasta34t25b1/fasta34 -a -Q -3 -E10 -d5 -B /tmp/allergen16603.tmp allergen6.fasta
FASTA searches a protein or DNA sequence data bank
 version 3.4t25 Nov 12, 2004
Please cite:
 W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

Query library /tmp/allergen16603.tmp vs allergen6.fasta library
searching allergen6.fasta library

  1>>>Positive Control sequence Ber e 1 (146 aa)  - 146 aa
 vs  allergen6.fasta library

       opt      E()
< 20     6     0:==
  22     0     0:          one = represents 5 library sequences
  24     0     0:
  26     0     0:
  28     0     0:
```

```
  30     0     2:*
  32     0     8: *
  34     0    22:    *
  36     0    45:        *
  38     0    74:              *
  40    12   104:===                 *
  42    72   127:===============          *
  44   174   140:===========================*=======
  46   238   143:============================*===================
  48   275   137:===========================*===========================
  50   227   125:========================*=====================
  52   142   110:=====================*=======
  54    96    94:==================*=
  56    66    78:============== *
  58    36    64:========    *
  60    36    52:========  *
  62    16    42:====    *
  64    10    33:==    *
  66     1    26:=    *
  68     5    21:=   *
  70     5    16:=  *
  72     4    13:= *
  74     8    10:=*
  76     1     8:=*
  78     4     6:=*
  80     5     5:*
  82     5     4:*
  84     1     3:*
  86     4     2:*
  88     1     2:*          inset = represents 1 library sequences
  90     4     1:*
  92     5     1:*         :*====
  94    11     1:*==       :*==========
  96    15     1:*==       :*==============
  98     5     0:=         *=====
 100     3     0:=         *===
 102     6     0:==        *======
 104     7     0:==        *=======
 106     3     0:=         *===
 108     2     0:=         *==
 110     0     0:          *
 112     7     0:==        *=======
 114     0     0:          *
 116     0     0:          *
 118     0     0:          *
>120    19     0:====      *===================
 336538 residues in  1537 sequences
  Expectation_n fit: rho(ln(x))= 5.2218+/-0.00456; mu= 6.1286+/- 0.236
 mean_var=259.0922+/-88.572, 0's: 6 Z-trim: 24  B-trim: 270 in 2/45
 Lambda= 0.079680
 Kolmogorov-Smirnov  statistic: 0.1953 (N=25) at  42

FASTA (3.47 Mar 2004) function [optimized, BL50 matrix (15:-5)] ktup: 2
 join: 36, opt: 24, open/ext: -10/-2, width:  16
 Scan time:  0.060
The best scores are:                                      opt z-sc E(1537)
gi|112754|sp|P04403|2SS_BEREX 2S sulfur-rich s ( 146) [f] 1010  657.5 1.3e-31
gi|17713|emb|CAA38363.1| 2S albumin [Bertholle ( 154) [f]  686  456.0 2.1e-20
gi|1794252|gb|AAB41308.1| albumin seed storage ( 139) [f]  389  271.9 3.8e-10
gi|31321942|gb|AAM54365.1| 2S albumin seed sto ( 161) [f]  385  268.9 5.5e-10
gi|28207731|gb|AAO32314.1| putative allergen I ( 143) [f]  367  258.1 2.2e-09
gi|5381323|gb|AAD42943.1|AF091841_1 2S albumin ( 148) [f]  306  220.1 2.9e-07
gi|18821|emb|CAA40015.1| pre-pro-seed albumin  ( 141) [f]  295  213.4 6.8e-07
gi|24473800|gb|AAL91665.1| 2s albumin [Anacard ( 138) [f]  285  207.3 1.5e-06
gi|112762|sp|P01089|2SS_RICCO 2S albumin precu ( 258) [f]  267  194.1 8.2e-06
gi|13183175|gb|AAK15088.1|AF240005_1 2S albumi ( 153) [f]  254  187.7 1.9e-05
gi|17907758|dbj|BAB79444.1| BW8KD allergen pro ( 133) [f]  234  175.7 8.6e-05
gi|26985163|gb|AAN86249.1|AF448054_1 recombina ( 109) [f]  207  159.6 0.00068
gi|17225991|gb|AAL37561.1|AF366561_1 conglutin ( 144) [f]  195  151.2   0.002
gi|14347293|emb|CAC41202.1| unnamed protein pr ( 207) [f]  175  137.6   0.011
```

```
gi|61970231|gb|AAX57578.1| BW 16kDa allergen [ ( 127) [f]  167  134.2   0.018
gi|57118278|gb|AAW34231.1| Ara h 6 allergen [A ( 124) [f]  159  129.3   0.033
gi|31322017|gb|AAM78596.1| allergen Ara h 2 is ( 169) [f]  155  125.8   0.051
gi|26245447|gb|AAN77576.1| allergen Ara h 2.02 ( 172) [f]  155  125.8   0.052
gi|5923742|gb|AAD56337.1|AF092846_1 allergen A ( 129) [f]  144  119.9    0.11
gi|21757|emb|CAA26383.1| unnamed protein produ ( 296) [f]  138  113.5    0.25
gi|21711|emb|CAA42453.1| CM 17 protein precurs ( 143) [f]  134  113.3    0.26
gi|170740|gb|AAA34290.1| gliadin               ( 296) [f]  137  112.8    0.27
gi|2118430|pir||S59924 allergen RA16 precursor ( 157) [f]  133  112.4    0.29
gi|1398916|dbj|BAA07712.1| allergenic protein  ( 157) [f]  133  112.4    0.29
gi|1398913|dbj|BAA07710.1| allergenic protein  ( 166) [f]  133  112.2    0.29
gi|2118427|pir||S59922 allergen RA14B precurso ( 166) [f]  133  112.2    0.29
gi|170726|gb|AAA34283.1| pre-alpha-/beta-gliad ( 282) [f]  131  109.3    0.43
gi|23616947|dbj|BAC20650.1| putative allergeni ( 160) [f]  125  107.4    0.55
gi|21916|emb|CAA34709.1| precursor (AA -24 to  ( 143) [f]  122  105.9    0.67
gi|170718|gb|AAA34279.1| alpha/beta-gliadin pr ( 313) [f]  126  105.8    0.67
gi|21765|emb|CAA26385.1| unnamed protein produ ( 313) [f]  126  105.8    0.67
gi|170724|gb|AAA34282.1| pre-alpha-/beta-gliad ( 297) [f]  124  104.8    0.77
gi|452323|emb|CAA49556.1| CMb component of tet ( 149) [f]  120  104.5    0.79
gi|585290|sp|P32936|IAAB_HORVU Alpha-amylase/t ( 149) [f]  120  104.5    0.79
gi|218197|dbj|BAA01996.1| allergenic protein [ ( 157) [f]  120  104.3    0.81
gi|34900104|ref|NP_911398.1| allergen RA16 [Or ( 157) [f]  120  104.3    0.81
gi|419799|pir||S31078 seed allergen RA5 - rice ( 157) [f]  120  104.3    0.81
gi|34900098|ref|NP_911395.1| seed allergen RAG ( 166) [f]  119  103.5     0.9
gi|419801|pir||S31082 seed allergen RAG2 - ric ( 166) [f]  119  103.5     0.9
gi|439275|emb|CAA49555.1| CMa, component of te ( 145) [f]  118  103.4    0.92
gi|21920|emb|CAA39099.1| CM2 protein [Triticum ( 145) [f]  118  103.4    0.92
gi|21783|emb|CAA30570.1| unnamed protein produ ( 356) [f]  122  102.9    0.97
gi|21673|emb|CAA35238.1| unnamed protein produ ( 307) [f]  120  102.2     1.1
gi|1063270|dbj|BAA11251.1| gamma-gliadin precu ( 279) [f]  119  101.9     1.1
gi|21751|emb|CAA31396.1| high molecular weight ( 648) [f]  121  100.4     1.4
gi|21779|emb|CAA26847.1| unnamed protein produ ( 660) [f]  121  100.3     1.4
gi|5931948|gb|AAD56719.1|AF091737_1 allergen [ ( 160) [f]  113   99.9     1.4
gi|170730|gb|AAA34285.1| gamma-gliadin B-I pre ( 304) [f]  116   99.7     1.5
gi|170732|gb|AAA34286.1| gamma-gliadin         ( 323) [f]  116   99.5     1.5
gi|22090|emb|CAA43361.1| HMW glutenin subunit  ( 705) [f]  120   99.5     1.5
gi|1405736|emb|CAA35188.1| trypsin inhibitor c ( 144) [f]  111   99.0     1.6
gi|19009|emb|CAA46705.1| CMe [Hordeum vulgare  ( 148) [f]  111   98.9     1.6
gi|170712|gb|AAA34276.1| pre-alpha-/beta-gliad ( 291) [f]  112   97.4       2
gi|7442124|pir||T05923 glutenin low molecular  ( 374) [f]  113   97.2       2
gi|21701|emb|CAA35598.1| unnamed protein produ ( 145) [f]  108   97.1       2
gi|2827316|gb|AAB99797.1| allergenic protein [ ( 157) [f]  108   96.9     2.1
gi|7438309|pir||T02664 allergen - rice         ( 157) [f]  108   96.9     2.1
gi|2118429|pir||S59925 allergen RA5B precursor ( 160) [f]  108   96.8     2.1
gi|34900132|ref|NP_911412.1| allergen RA5B pre ( 160) [f]  108   96.8     2.1
gi|1398918|dbj|BAA07713.1| allergenic protein  ( 160) [f]  108   96.8     2.1
gi|34900094|ref|NP_911393.1| seed allergen RA1 ( 163) [f]  108   96.8     2.1
gi|419798|pir||S31080 seed allergen RA14 - ric ( 165) [f]  108   96.7     2.2
gi|218193|dbj|BAA01998.1| allergenic protein [ ( 165) [f]  108   96.7     2.2
gi|100691|pir||S21157 seed allergen RA17 - ric ( 162) [f]  107   96.2     2.3
gi|218195|dbj|BAA01997.1| allergenic protein [ ( 162) [f]  107   96.2     2.3
gi|170702|gb|AAA34272.1| gamma gliadin precurs ( 302) [f]  110   96.0     2.4
gi|7438312|pir||T03966 allergenic protein - ri ( 113) [f]  104   95.5     2.5
gi|1304218|dbj|BAA07774.1| allergenic protein  ( 113) [f]  104   95.5     2.5
gi|1398915|dbj|BAA07711.1| allergenic protein  ( 160) [f]  105   95.0     2.7
gi|2118428|pir||S59923 allergen RA14C precurso ( 160) [f]  105   95.0     2.7
gi|62550933|emb|CAI79052.1| putative LMW-glute ( 326) [f]  108   94.5     2.9
gi|1304216|dbj|BAA07772.1| allergenic protein  ( 111) [f]  102   94.3     2.9
gi|7438310|pir||T03963 allergenic protein - ri ( 111) [f]  102   94.3     2.9
gi|100834|pir||S16031 alpha-amylase inhibitor, ( 168) [f]  104   94.2       3
gi|21713|emb|CAA35597.1| unnamed protein produ ( 168) [f]  104   94.2       3
gi|18869|emb|CAA31794.1| alpha-amylase/trypsin ( 147) [f]  103   94.0     3.1
gi|21773|emb|CAA31685.1| unnamed protein produ ( 307) [f]  106   93.5     3.3
gi|21930|emb|CAA44473.1| LMW glutenin [Triticu ( 285) [f]  105   93.1     3.4
gi|170736|gb|AAA34288.1| gamma-gliadin         ( 251) [f]  104   92.9     3.5
gi|18955|emb|CAA41956.1| alpha-amylase inhibit ( 144) [f]  101   92.8     3.6
gi|170708|gb|AAA34274.1| gamma-gliadin B precu ( 291) [f]  104   92.4     3.8
gi|3367714|emb|CAA08836.1| BDAI-1; Barley dime ( 152) [f]  100   92.0     3.9
gi|170734|gb|AAA34287.1| gamma gliadin B-III   ( 244) [f]  102   91.7     4.1
gi|170738|gb|AAA34289.1| gamma-gliadin         ( 327) [f]  102   90.8     4.6
```

```
gi|886967|emb|CAA59340.1| low molecular weight ( 276) [f]  101   90.7     4.7
gi|886965|emb|CAA59339.1| low molecular weight ( 261) [f]  100   90.3     4.9
gi|452325|emb|CAA49557.1| CMd component of tet ( 171) [f]   97   89.8     5.3
gi|7442133|pir||T05910 glutenin low molecular  ( 286) [f]   99   89.3     5.5
gi|6580762|gb|AAF18269.1|AF066055_1 vicilin-li ( 593) [f]   99   87.0     7.5
gi|6573247|gb|AAF17603.1|AF204966_1 2S storage ( 125) [f]   89   85.8     8.7
gi|7438311|pir||T03965 allergenic protein - ri ( 109) [f]   88   85.6     8.9
gi|1304217|dbj|BAA07773.1| allergenic protein  ( 109) [f]   88   85.6     8.9
gi|62484809|emb|CAI78902.1| putative gamma-gli ( 285) [f]   93   85.6     8.9

>>gi|112754|sp|P04403|2SS_BEREX 2S sulfur-rich seed stor  (146 aa)
 initn: 1010 init1: 1010 opt: 1010  Z-score: 657.5  bits: 127.6 E(): 1.3e-31
Smith-Waterman score: 1010;  100.000% identity (100.000% similar) in 146
aa overlap (1-146:1-146)

               10        20        30        40        50        60
Positi MAKISVAAAALLVLMALGHATAFRATVTTTVVEEENQEECREQMQRQQMLSHCRMYMRQQ
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
gi|112 MAKISVAAAALLVLMALGHATAFRATVTTTVVEEENQEECREQMQRQQMLSHCRMYMRQQ
               10        20        30        40        50        60

               70        80        90       100       110       120
Positi MEESPYQTMPRRGMEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGEQMRRMM
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
gi|112 MEESPYQTMPRRGMEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGEQMRRMM
               70        80        90       100       110       120

              130       140
Positi RLAENIPSRCNLSPMRCPMGGSIAGF
       ::::::::::::::::::::::::::
gi|112 RLAENIPSRCNLSPMRCPMGGSIAGF
              130       140

>>gi|17713|emb|CAA38363.1| 2S albumin [Bertholletia exce  (154 aa)
 initn: 729 init1: 220 opt: 686  Z-score: 456.0  bits: 90.4 E(): 2.1e-20
Smith-Waterman score: 686;  69.281% identity (86.275% similar) in 153
aa overlap (1-144:1-152)

               10        20        30               40        50
Positi MAKISVAAAALLVLMALGHATAFRATVTTTVVEE--EN-----QEECREQMQRQQMLSHC
       :::.::.:::::.:..::.:::::.:::::. ::  ::      ...:::::.:::.:.::
gi|177 MAKMSVVAAALLALLVLGQATAFRTTVTTTLEEEQEENPRGRSEQQCREQMERQQQLNHC
               10        20        30        40        50        60

            60         70         80        90       100       110
Positi RMYMRQQMEESPYQT-MP-RRGMEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQEEMQP
       :::.::::::::::.  : ::: :::..::::::: ::: :::::::::. : :.:::.
gi|177 RMYLRQQMEESPYQNPRPLRRGEEPHLDECCEQLERMDEMCRCEGLRMMLRR-QREEMEL
               70        80        90       100       110

             120       130       140
Positi RGEQMRRMMRLAENIPSRCNLSPMRCPMGGSIAGF
       .::::.:.:: :::. :::::::.::::::  :
gi|177 QGEQMQRIMRKAENLLSRCNLSPQRCPMGGYTAWL
     120       130       140       150

>>gi|1794252|gb|AAB41308.1| albumin seed storage protein  (139 aa)
 initn: 403 init1: 136 opt: 389  Z-score: 271.9  bits: 56.2 E(): 3.8e-10
Smith-Waterman score: 389;  42.029% identity (76.087% similar) in 138
aa overlap (8-139:1-134)

               10        20        30              40        50
Positi MAKISVAAAALLVLMALGHATAFRATVTTTVVEEENQ------EECREQMQRQQMLSHCR
              :: :..:. ...:.:::.:.::  ..:. .      : ::::.:::: :.::.
gi|179        AALLVALLFVANAAAFRTTITTMEIDEDIDNPRRRGEGCREQIQRQQNLNHCQ
                      10        20        30        40        50

           60        70        80        90       100       110
Positi MYMRQQMEESPYQTMPRRGMEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGE
       .:.::: . . :.   .:    :. .::.::  :::.:.:::::... :.::..   :::
```

```
gi|179 YYLRQQSRSGGYDEDNQR---QHFRQCCQQLSQMDEQCQCEGLRQVVRRQQQQQGL-RGE
            60        70           80        90       100

          120       130       140
Positi QMRRMMRLAENIPSRCNLSPMRCPMGGSIAGF
       .:..:.. :...:..:..: .:: .
gi|179 EMEEMVQSARDLPNECGISSQRCEIRRSWF
     110       120       130

>>gi|31321942|gb|AAM54365.1| 2S albumin seed storage pro  (161 aa)
 initn: 399 init1: 123 opt: 385  Z-score: 268.9  bits: 55.9 E(): 5.5e-10
Smith-Waterman score: 385;  39.310% identity (75.862% similar) in 145
aa overlap (1-139:16-156)

                              10        20        30
Positi                MAKISVAAAALLVLMALGHATAFRATVTTTVVEEENQ------EE
                      ::.... :: :..:. ...:.:::.:.::  ..:. .      :
gi|313 RHEARKCIFHTFSLTMARLATLAALLVALLFVANAAAFRTTITTMEIDEDIDNPRRRGEG
               10        20        30        40        50        60

      40        50        60        70        80        90
Positi CREQMQRQQMLSHCRMYMRQQMEESPYQTMPRRGMEPHMSECCEQLEGMDESCRCEGLRM
       :.::.:::: :.::..:.::: . . :.   .:    :. .::.::  ..:.:.:::::.
gi|313 CQEQIQRQQNLNHCQYYLRQQSRSGGYDEDNQR---QHFRQCCQQLSQIEEQCQCEGLRQ
               70        80        90          100       110

     100       110       120       130       140
Positi MMMRMQQEEMQPRGEQMRRMMRLAENIPSRCNLSPMRCPMGGSIAGF
        . :.::..   :::.:..:.. :...:..:..: .:: .
gi|313 AVRRQQQQGL-RGEEMEEMVQSARDLPKECGISSQRCEIRRSWF
       120        130       140       150       160

>>gi|28207731|gb|AAO32314.1| putative allergen I1 [Carya  (143 aa)
 initn: 390 init1: 141 opt: 367  Z-score: 258.1  bits: 53.7 E(): 2.2e-09
Smith-Waterman score: 367;  40.000% identity (72.414% similar) in 145
aa overlap (1-139:1-138)

               10        20        30              40        50
Positi MAKISVAAAALLVLMALGHATAFRATVTTTVVEEENQ------EECREQMQRQQMLSHCR
       ::..   :: :..:. ...:.:::.:.::  ..:. .      : ::::.::::.:..:.
gi|282 MARV---AALLVALLFVANAAAFRTTITTMEIDEDIDNPRRRGESCREQIQRQQYLNRCQ
                  10        20        30        40        50

           60        70        80        90       100       110
Positi MYMRQQMEESPYQTMPRRGMEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGE
        :.::: . . :.   .:    :. .::.::  :.:.:.:::::. . :.::.:   :::
gi|282 DYLRQQCRSGGYDEDNQR---QHFRQCCQQLSQMEEQCQCEGLRQAV-RQQQQEEGIRGE
        60        70           80        90       100        110

          120       130       140
Positi QMRRMMRLAENIPSRCNLSPMRCPMGGSIAGF
       .:..:.. : ..:..:..:   : .
gi|282 EMEEMVQCASDLPKECGISSRSCEIRRSWF
           120       130       140



146 residues in 1 query   sequences
336538 residues in 1537 library sequences
 Scomplib [34t25]
 start: Thu Apr  6 16:51:57 2006 done: Thu Apr  6 16:51:57 2006
 Total Scan time:  0.060 Total Display time:  0.010

Function used was FASTA [version 3.4t25 Nov 12, 2004]
```

Search 80-mer Sliding Window using FASTA3 with Query Ber e 1 against AllergenOnline Database v6.0 (January, 2006)

Database	AllergenOnline v6.0 (January, 2006)

Input Query	`>Positive Control sequence Ber e 1 (146 aa)`
	`MAKISVAAAALLVLMALGHATAFRATVTTTVVEEENQEECREQMQRQQMLSHCRMYMRQQ`
	`MEESPYQTMPRRGMEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGEQMRRMM`
	`RLAENIPSRCNLSPMRCPMGGSIAGF`

Length	146
Number of 80 mers	67

Hit #	Defline	Best %ID	# Hits > 35%	Full Alignment			Links	
				E-val	%ID	length	NCBI	Details
1	gi|112754|sp|P04403|2SS_BEREX 2S sulfur-rich seed	100.00%	67of67	1.3e-31	100.00%	146	gi|112754	GO!
2	gi|17713|emb|CAA38363.1| 2S albumin [Bertholletia	73.20%	67of67	2.1e-20	69.30%	153	gi|17713	GO!
3	gi|1794252|gb|AAB41308.1| albumin seed storage pr	48.78%	59of67	3.8e-10	42.00%	138	gi|1794252	GO!
4	gi|28207731|gb|AAO32314.1| putative allergen I1 [	47.50%	59of67	2.2e-09	40.00%	145	gi|28207731	GO!
5	gi|31321942|gb|AAM54365.1| 2S albumin seed storag	45.03%	59of67	5.5e-10	39.30%	145	gi|31321942	GO!
6	gi|5381323|gb|AAD42943.1|AF091841_1 2S albumin pr	43.75%	67of67	2.9e-07	39.20%	148	gi|5381323	GO!
7	gi|112762|sp|P01089|2SS_RICCO 2S albumin precurso	38.80%	39of67	8.2e-06	38.50%	104	gi|112762	GO!
8	gi|18821|emb|CAA40015.1| pre-pro-seed albumin [He	38.80%	40of67	6.8e-07	36.20%	149	gi|18821	GO!
9	gi|13183175|gb|AAK15088.1|AF240005_1 2S albumin [	38.50%	22of67	1.9e-05	35.40%	147	gi|13183175	GO!

BLASTP Search with Query Ber e 1 against the NCBI Database Available on Thursday, April 06, 2006, using a key word delimiter “allergen” (limited search to 2,529 sequences).

```
Thursday, April 06, 2006

BLASTP 2.2.13 [Nov-27-2005]
Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
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.

RID: 1144361683-25428-161321806055.BLASTQ4


Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF excluding environmental samples
           3,538,507 sequences; 1,215,510,009 total letters
Query=  Positive Control sequence Ber e 1 (146 aa)
Length=146


                                                                   Score     E
Sequences producing significant alignments:                        (Bits)  Value

gi|112754|sp|P04403|2SS_BEREX  2S sulfur-rich seed storage pro...   300    1e-83
gi|1794252|gb|AAB41308.1|  albumin seed storage protein precursor   116    2e-28
gi|31321942|gb|AAM54365.1|  2S albumin seed storage protein [Jugl   115    4e-28
gi|28207731|gb|AAO32314.1|  putative allergen I1 [Carya illinoine   109    3e-26
gi|112762|sp|P01089|2SS_RICCO  2S albumin precursor (Allergen ...  75.9    5e-16
gi|17907758|dbj|BAB79444.1|  BW8KD allergen protein [Fagopyrum es  67.8    1e-13
gi|15418705|gb|AAK96887.1|  allergen II [Arachis hypogaea]         54.3    1e-09
gi|26245447|gb|AAN77576.1|  allergen Ara h 2.02 [Arachis hypogaea  49.7    3e-08
gi|28912768|gb|AAO61750.1|  seed storage protein [Arachis hypogae  46.2    4e-07
gi|61970231|gb|AAX57578.1|  BW 16kDa allergen [Fagopyrum esculent  45.4    7e-07
gi|83416591|gb|ABC18306.1|  16 kDa allergen [Fagopyrum esculentum  45.1    9e-07
gi|31322017|gb|AAM78596.1|  allergen Ara h 2 isoform [Arachis hyp  44.3    1e-06
gi|5931948|gb|AAD56719.1|  allergen [Arachis hypogaea]             41.6    1e-05
gi|1009438|emb|CAA62911.1|  allergen sin a 1.0106 [Sinapis alba]   39.7    4e-05
gi|1009442|emb|CAA62908.1|  allergen sin a 1.0108 [Sinapis alba]   39.7    4e-05
gi|7545129|gb|AAB25214.2|  major allergen Sin a I [Sinapis alb...  39.7    4e-05
gi|121095|sp|P04726|GDA6_WHEAT  Alpha/beta-gliadin clone PW1215 p  39.3    5e-05
gi|1009440|emb|CAA62912.1|  allergen sin a 1.0107 [Sinapis alba]   38.1    1e-04
gi|1009436|emb|CAA62910.1|  allergen sin a 1.0105 [Sinapis alba]   38.1    1e-04
gi|1009434|emb|CAA62909.1|  allergen sin a 1.0104 [Sinapis alba]   37.7    1e-04
gi|548657|sp|Q01881|RA05_ORYSA  Seed allergenic protein RA5 precu  37.4    2e-04
gi|5923742|gb|AAD56337.1|  allergen Arah6 [Arachis hypogaea]       37.0    2e-04
gi|121094|sp|P04725|GDA5_WHEAT  Alpha/beta-gliadin A-V precursor   36.6    3e-04
gi|121096|sp|P04727|GDA7_WHEAT  Alpha/beta-gliadin clone PW8142 p  36.6    3e-04
gi|121093|sp|P04724|GDA4_WHEAT  Alpha/beta-gliadin A-IV precursor  36.2    4e-04
gi|121092|sp|P04723|GDA3_WHEAT  Alpha/beta-gliadin A-III precurso  35.8    5e-04
gi|34900132|ref|NP_911412.1|  allergen RA5B precursor [Oryza s...  35.0    9e-04
gi|121098|sp|P18573|GDA9_WHEAT  Alpha/beta-gliadin MM1 precursor   34.7    0.001
gi|34900098|ref|NP_911395.1|  seed allergen RAG2 [Oryza sativa...  34.7    0.001
gi|585290|sp|P32936|IAAB_HORVU  Alpha-amylase/trypsin inhibito...  34.3    0.002
gi|6580762|gb|AAF18269.1|  vicilin-like protein precursor [Juglan  33.9    0.002
gi|407610|gb|AAB27813.1|  Bra j IE large chain=allergen [Brass...  33.1    0.003
gi|121091|sp|P04722|GDA2_WHEAT  Alpha/beta-gliadin A-II precursor  33.1    0.003
gi|67464993|sp|P02863|GDA0_WHEAT  Alpha/beta-gliadin precursor (P  28.1    0.11
gi|121090|sp|P04721|GDA1_WHEAT  Alpha/beta-gliadin A-I precursor   27.7    0.14
gi|21954740|gb|AAM83103.1|  paramyosin allergen [Blomia tropic...  27.3    0.19
gi|121097|sp|P04728|GDA8_WHEAT  Alpha/beta-gliadin clone PTO-A10   26.9    0.24
gi|2739154|gb|AAC67308.1|  22.6 kDa tegumental antigen [Schistoso  23.9    2.1
gi|6680437|ref|NP_032399.1|  interleukin 9 [Mus musculus]          23.5    2.7
gi|42559558|sp|O97192|TPM_HELAS  Tropomyosin (Allergen Hel as 1)   23.1    3.5
gi|13785807|gb|AAK39511.1|  paramyosin-like allergen [Dermatop...  23.1    3.5
gi|37778944|gb|AAO73464.1|  HDM allergen [Dermatophagoides pteron  22.7    4.6
gi|31321944|gb|AAM54366.1|  vicilin seed storage protein [Juglans  22.7    4.6
gi|69937363|ref|ZP_00632029.1|  Allergen V5/Tpx-1 related [Par...  22.7    4.6
gi|9954251|gb|AAG08988.1|  tropomyosin [Perna viridis]             22.3    6.0
gi|8118425|gb|AAF72985.1|  beta-expansin [Oryza sativa]            21.9    7.8
gi|28476851|tpg|DAA00353.1|  TPA: TPA_exp: allergen dI chain C2D   21.9    7.8
gi|37536762|ref|NP_922683.1|  beta-expansin EXPB4 [Oryza sativ...  21.9    7.8
```

```
ALIGNMENTS
>gi|112754|sp|P04403|2SS_BEREX 2S sulfur-rich seed storage protein precursor (Allergen
Ber e
1) [Contains: 2S sulfur-rich seed storage protein small chain
(2S albumin small subunit); 2S sulfur-rich seed storage protein
large chain 1B (2S albumin large subunit)]
Length=146

 Score =  300 bits (767),  Expect = 1e-83
 Identities = 146/146 (100%), Positives = 146/146 (100%), Gaps = 0/146 (0%)

Query  1    MAKISVAAAALLVLMALGHATAFRATVTTTVVEEENQEECREQMQRQQMLSHCRMYMRQQ  60
            MAKISVAAAALLVLMALGHATAFRATVTTTVVEEENQEECREQMQRQQMLSHCRMYMRQQ
Sbjct  1    MAKISVAAAALLVLMALGHATAFRATVTTTVVEEENQEECREQMQRQQMLSHCRMYMRQQ  60

Query  61   MEESPYQTMPRRGMEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGEQMRRMM  120
            MEESPYQTMPRRGMEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGEQMRRMM
Sbjct  61   MEESPYQTMPRRGMEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGEQMRRMM  120

Query  121  RLAENIPSRCNLSPMRCPMGGSIAGF  146
            RLAENIPSRCNLSPMRCPMGGSIAGF
Sbjct  121  RLAENIPSRCNLSPMRCPMGGSIAGF  146


>gi|1794252|gb|AAB41308.1| albumin seed storage protein precursor [Juglans regia]
Length=139

 Score =  116 bits (291),  Expect = 2e-28
 Identities = 58/136 (42%), Positives = 85/136 (62%), Gaps = 10/136 (7%)

Query  8    AAALLVLMALGHATAFRATVTTTVVEEE------NQEECREQMQRQQMLSHCRMYMRQQM  61
            AA L+ L+ + +A AFR T+TT  ++E+        E CREQ+QRQQ L+HC+ Y+RQQ
Sbjct  1    AALLVALLFVANAAAFRTTITTMEIDEDIDNPRRRGEGCREQIQRQQNLNHCQYYLRQQS  60

Query  62   EESPYQTMPRRGMEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGEQMRRMMR  121
                Y    +R    H  +CC+QL  MDE C+CEGLR  ++R QQ++   RGE+M  M++
Sbjct  61   RSGGYDEDNQR---QHFRQCCQQLSQMDEQCQCEGLR-QVVRRQQQQQGLRGEEMEEMVQ  116

Query  122  LAENIPSRCNLSPMRC  137
             A ++P+ C +S  RC
Sbjct  117  SARDLPNECGISSQRC  132


>gi|31321942|gb|AAM54365.1| 2S albumin seed storage protein [Juglans nigra]
Length=161

 Score =  115 bits (289),  Expect = 4e-28
 Identities = 57/143 (39%), Positives = 88/143 (61%), Gaps = 10/143 (6%)

Query  1    MAKISVAAAALLVLMALGHATAFRATVTTTVVEEE------NQEECREQMQRQQMLSHCR  54
            MA+++  AA L+ L+ + +A AFR T+TT  ++E+        E C+EQ+QRQQ L+HC+
Sbjct  16   MARLATLAALLVALLFVANAAAFRTTITTMEIDEDIDNPRRRGEGCQEQIQRQQNLNHCQ  75

Query  55   MYMRQQMEESPYQTMPRRGMEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGE  114
             Y+RQQ     Y    +R    H  +CC+QL  ++E C+CEGLR   +R QQ++   RGE
Sbjct  76   YYLRQQSRSGGYDEDNQR---QHFRQCCQQLSQIEEQCQCEGLR-QAVRRQQQQQGLRGE  131

Query  115  QMRRMMRLAENIPSRCNLSPMRC  137
            +M  M++ A ++P  C +S  RC
Sbjct  132  EMEEMVQSARDLPKECGISSQRC  154


>gi|28207731|gb|AAO32314.1| putative allergen I1 [Carya illinoinensis]
Length=143

 Score =  109 bits (273),  Expect = 3e-26
 Identities = 56/136 (41%), Positives = 81/136 (59%), Gaps = 10/136 (7%)

Query  8    AAALLVLMALGHATAFRATVTTTVVEEE------NQEECREQMQRQQMLSHCRMYMRQQM  61
            AA L+ L+ + +A AFR T+TT  ++E+        E CREQ+QRQQ L+ C+ Y+RQQ
```

```
Sbjct  5    AALLVALLFVANAAAFRTTITTMEIDEDIDNPRRRGESCREQIQRQQYLNRCQDYLRQQC  64

Query  62   EESPYQTMPRRGMEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGEQMRRMMR  121
                Y    +R    H  +CC+QL  M+E C+CEGLR   +R QQ+E   RGE+M  M++
Sbjct  65   RSGGYDEDNQR---QHFRQCCQQLSQMEEQCQCEGLR-QAVRQQQQEEGIRGEEMEEMVQ  120

Query  122  LAENIPSRCNLSPMRC  137
             A ++P  C +S   C
Sbjct  121  CASDLPKECGISSRSC  136


>gi|112762|sp|P01089|2SS_RICCO 2S albumin precursor (Allergen Ric c 1/3) [Contains:
Allergen
Ric c 3 small chain (4.7 kDa napin-like protein small chain)
(RS1A) (CB-1A small chain); Allergen Ric c 3 large chain (RL1)
(CB-1A large chain); Allergen Ric c 1 small chain (2S albumin
small chain) (4 kDa napin-like protein small chain) (RS2B);
Allergen Ric c 1 large chain (2S albumin large chain)
(7.3 kDa napin-like protein large chain) (RL2)]
Length=258

 Score = 75.9 bits (185),  Expect = 5e-16
 Identities = 40/104 (38%), Positives = 57/104 (54%), Gaps = 7/104 (6%)

Query  36   NQEECREQMQRQQMLSHCRMYMRQQMEESPYQTMPRRG--MEPHMSECCEQLEGMDESCR  93
            +Q+ CR Q+Q QQ L  C+ Y++QQ+        PRR    E  +  CC+ L+ M   CR
Sbjct  158  SQQGCRGQIQEQQNLRQCQEYIKQQVSGQG----PRRSDNQERSLRGCCDHLKQMQSQCR  213

Query  94   CEGLRMMMMRMQQEEMQPRGEQMRRMMRLAENIPSRCNLSPMRC  137
            CEGLR   +  QQ + Q +G+ +    R A N+PS C +SP  C
Sbjct  214  CEGLR-QAIEQQQSQGQLQGQDVFEAFRTAANLPSMCGVSPTEC  256


 Score = 62.4 bits (150),  Expect = 5e-12
 Identities = 42/153 (27%), Positives = 80/153 (52%), Gaps = 19/153 (12%)

Query  1    MAKI--SVAAAALLVLMALGHATAFRATVTTTVVEEENQE-------ECREQMQRQQMLS  51
            MAK+  ++A  ++L+ +    + A+R T+TT  ++E   E       +CR+++QR+  LS
Sbjct  1    MAKLIPTIALVSVLLFIIANASFAYRTTITTIEIDESKGEREGSSSQQCRQEVQRKD-LS  59

Query  52   HCRMYMRQQMEE-SPYQTMPRRGMEPH-------MSECCEQLEGMDESCRCEGLRMMMMR  103
             C  Y+RQ     SP + + R   + +       + +CC Q++ + + C+CE ++  +
Sbjct  60   SCERYLRQSSSRRSPGEEVLRMPGDENQQQESQQLQQCCNQVKQVRDECQCEAIK-YIAE  118

Query  104  MQQEEMQPRGEQMRRMMRLAENIPSRCNLSPMR  136
             Q ++ Q  GE+  R+ + A  I S C +  MR
Sbjct  119  DQIQQGQLHGEESERVAQRAGEIVSSCGVRCMR  151


>gi|17907758|dbj|BAB79444.1| BW8KD allergen protein [Fagopyrum esculentum]
Length=133

 Score = 67.8 bits (164),  Expect = 1e-13
 Identities = 41/140 (29%), Positives = 69/140 (49%), Gaps = 17/140 (12%)

Query  3    KISVAAAALLVLMALGHATAFRATVTTTVVEEENQEECREQMQRQQ-MLSHCRMYMRQQM  61
            K+ +  AA  +L+   HA            + + + +CR+QM+  +  L  C  YM   M
Sbjct  2    KLFLILAAASLLIVASHA------------DSQMRSKCRKQMRMMEPQLEQCEGYMTMDM  49

Query  62   -EESPYQTMPRRGMEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGEQMRRMM  120
             ++   +    R  E  M  CC  ++ MD+ C CE ++MM+   QQ+  +   E MR +M
Sbjct  50   MDDDSMRGRECRSEESCMRGCCLAMKEMDDECMCEWMKMMV---QQQRGEMGEEDMRMVM  106

Query  121  RLAENIPSRCNLSPMRCPMG  140
            R  + +P++C +  MRC MG
Sbjct  107  RKMKQLPNKCGMGHMRCHMG  126


>gi|15418705|gb|AAK96887.1| allergen II [Arachis hypogaea]
Length=156
```

```
 Score = 54.3 bits (129),  Expect = 1e-09
 Identities = 44/160 (27%), Positives = 74/160 (46%), Gaps = 31/160 (19%)

Query  1    MAKISVAAAALLVLMALGHATAFRATVTTTVVEEENQEECREQMQRQQMLSHCRMYMRQQ  60
            MAK+++  A  L L+A  HA+A +        E +    C+ Q++R   L  C   ++ Q+
Sbjct  1    MAKLTILVALALFLLA-AHASARQQW------ELQGDRRCQSQLERAN-LRPCEQHLMQK  52

Query  61   ME--------------ESPYQTMP--RRGM--EPHMSECCEQLEGMD--ESCRCEGLRMM  100
            ++              + PY   P  RRG     H   CC +L   +  + C CE L+
Sbjct  53   IQRDEDSYERDPYSPSQDPYSPSPYDRRGAGSSQHQERCCNELNEFENNQRCMCEALQ-Q  111

Query  101  MMRMQQEEMQPRGEQMRRMMRLAENIPSRCNL-SPMRCPM  139
            +M  Q + +Q R +Q ++  R   N+P +C L +P RC +
Sbjct  112  IMENQSDRLQGR-QQEQQFKRELRNLPQQCGLRAPQRCDL  150


>gi|26245447|gb|AAN77576.1| allergen Ara h 2.02 [Arachis hypogaea]
Length=172

 Score = 49.7 bits (117),  Expect = 3e-08
 Identities = 44/172 (25%), Positives = 73/172 (42%), Gaps = 43/172 (25%)

Query  1    MAKISVAAAALLVLMALGHATAFRATVTTTVVEEENQEECREQMQRQQMLSHCRMYMRQQ  60
            MAK+++  A  L L+A  HA+A +        E +    C+ Q++R   L  C   ++ Q+
Sbjct  1    MAKLTILVALALFLLA-AHASARQQW------ELQGDRRCQSQLERAN-LRPCEQHLMQK  52

Query  61   MEE-----------------SPYQTMPRR-------------GMEPHMSECCEQLEGMD-  89
            ++                  SP Q   RR             G   H   CC +L   +
Sbjct  53   IQRDEDSYGRDPYSPSQDPYSPSQDPDRRDPYSPSPYDRRGAGSSQHQERCCNELNEFEN  112

Query  90   -ESCRCEGLRMMMMRMQQEEMQPRGEQMRRMMRLAENIPSRCNL-SPMRCPM  139
             + C CE L+  +M  Q + +Q R +Q ++  R   N+P +C L +P RC +
Sbjct  113  NQRCMCEALQ-QIMENQSDRLQGR-QQEQQFKRELRNLPQQCGLRAPQRCDL  162


>gi|28912768|gb|AAO61750.1| seed storage protein [Arachis hypogaea]
Length=166

 Score = 46.2 bits (108),  Expect = 4e-07
 Identities = 42/170 (24%), Positives = 71/170 (41%), Gaps = 43/170 (25%)

Query  3    KISVAAAALLVLMALGHATAFRATVTTTVVEEENQEECREQMQRQQMLSHCRMYMRQQME  62
            K+++  A  L L+A  HA+A +        E +    C+ Q++R   L  C   ++ Q+++
Sbjct  1    KLTILVALALFLLA-AHASARQQW------ELQGDRRCQSQLERAN-LRPCEQHLMQKIQ  52

Query  63   E-----------------SPYQTMPRR-------------GMEPHMSECCEQLEGMD--E  90
                              SP Q   RR             G   H   CC +L   +  +
Sbjct  53   RDEDSYGRDPYSPSQDPYSPSQDPDRRDPYSPSPYDRRGAGSSQHQERCCNELNEFENNQ  112

Query  91   SCRCEGLRMMMMRMQQEEMQPRGEQMRRMMRLAENIPSRCNL-SPMRCPM  139
             C CE L+  +M  Q + +Q R +Q ++  R   N+P +C L +P RC +
Sbjct  113  RCMCEALQ-QIMENQSDRLQGR-QQEQQFKRELRNLPQQCGLRAPQRCDL  160


>gi|61970231|gb|AAX57578.1| BW 16kDa allergen [Fagopyrum esculentum]
Length=127

 Score = 45.4 bits (106),  Expect = 7e-07
 Identities = 30/119 (25%), Positives = 57/119 (47%), Gaps = 13/119 (10%)

Query  33   EEENQEECREQMQRQQM-LSHCRMYMRQQMEESPY-QTMPR------RGMEPHMSECCEQ  84
            E +   +C  Q++  +  L  C  Y+   + +  Y + + R      +  E  M  CC
Sbjct  9    ETQMSSKCMRQVKMNEPHLKKCNRYIAMDILDDKYAEALSRVEGEGCKSEESCMRGCCVA  68

Query  85   LEGMDESCRCEGLRMMMMRMQQEEMQPRGEQM-RRMMRLAENIPSRCNLSPMRCPMGGS  142
            ++ MD+ C CE ++MM+    + +    GE++ +  +R  + +PS+C LS M C   G+
Sbjct  69   MKEMDDECVCEWMKMMV----ENQKGRIGERLIKEGVRDLKELPSKCGLSEMECGSRGN  123
```

```
>gi|83416591|gb|ABC18306.1| 16 kDa allergen [Fagopyrum esculentum]
Length=149

 Score = 45.1 bits (105),  Expect = 9e-07
 Identities = 35/149 (23%), Positives = 69/149 (46%), Gaps = 14/149 (9%)

Query  3    KISVAAAALLVLMALGHATAFRATVTTTVVEEENQEECREQMQRQQM-LSHCRMYMRQQM  61
            K+ +  A   +L+A   AT +       + E +   +C  Q++  +  L  C  Y+   +
Sbjct  2    KLFIILATATLLIAATQAT-YPRDEGFDLGETQMSSKCMRQVKMNEPHLKKCNRYIAMDI  60

Query  62   EESPY-QTMPR------RGMEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGE  114
             +  Y + + R      +  E  M  CC  ++ MD+ C CE ++MM+    + +    GE
Sbjct  61   LDDKYAEALSRVEGEGCKSEESCMRGCCVAMKEMDDECVCEWMKMMV----ENQKGRIGE  116

Query  115  QM-RRMMRLAENIPSRCNLSPMRCPMGGS  142
            ++ +  +R  + +PS+C LS + C   G+
Sbjct  117  RLIKEGVRDLKELPSKCGLSELECGSRGN  145


>gi|31322017|gb|AAM78596.1| allergen Ara h 2 isoform [Arachis hypogaea]
Length=169

 Score = 44.3 bits (103),  Expect = 1e-06
 Identities = 41/169 (24%), Positives = 70/169 (41%), Gaps = 43/169 (25%)

Query  4    ISVAAAALLVLMALGHATAFRATVTTTVVEEENQEECREQMQRQQMLSHCRMYMRQQMEE  63
            +++  A  L L+A  HA+A +        E +    C+ Q++R   L  C  ++ Q+++
Sbjct  1    LTILVAPALFLLA-AHASARQQW------ELQGDRRCQSQLERAN-LRPCEQHLMQKIQR  52

Query  64   -----------------SPYQTMPRR-------------GMEPHMSECCEQLEGMD--ES  91
                             SP Q   RR             G   H   CC +L   +  +
Sbjct  53   DEDSYGRDPYSPSQDPYSPSQDPDRRDPYSPSPYDRRGAGSSQHQERCCNELNEFENNQR  112

Query  92   CRCEGLRMMMMRMQQEEMQPRGEQMRRMMRLAENIPSRCNL-SPMRCPM  139
            C CE L+  +M  Q + +Q R +Q ++  R   N+P +C L +P RC +
Sbjct  113  CMCEALQ-QIMENQSDRLQGR-QQEQQFKRELRNLPQQCGLRAPQRCDL  159


>gi|5931948|gb|AAD56719.1| allergen [Arachis hypogaea]
Length=160

 Score = 41.6 bits (96),  Expect = 1e-05
 Identities = 32/133 (24%), Positives = 54/133 (40%), Gaps = 19/133 (14%)

Query  1    MAKISVAAAALLVLMALGHATAFR----ATVTTTVVEEENQEECREQMQRQQMLSHCRMY  56
            M K+S+  A L  L+ +  AT +     +  +         ++C+ Q+QR   L  C  +
Sbjct  2    MVKLSILVALLGALLVVASATRWDPDRGSRGSRWDAPSRGDDQCQRQLQRAN-LRPCEEH  60

Query  57   MR--------QQMEESPYQTMPRRGMEPHMSE------CCEQLEGMDESCRCEGLRMMMM  102
            MR        Q+ +E PY     RG +P  S+      CC +L     + RC    +  +
Sbjct  61   MRRRVEQEQEQEQDEYPYSRRGSRGRQPGESDENQEQRCCNELNRFQNNQRCMCQALQQI  120

Query  103  RMQQEEMQPRGEQ  115
               Q    P G++
Sbjct  121  LQNQSFWVPAGQE  133


>gi|1009438|emb|CAA62911.1| allergen sin a 1.0106 [Sinapis alba]
Length=145

 Score = 39.7 bits (91),  Expect = 4e-05
 Identities = 31/135 (22%), Positives = 59/135 (43%), Gaps = 28/135 (20%)

Query  39   ECREQMQRQQMLSHCRMYMRQQMEES---PYQTMPRR-----GME---------PHMSEC  81
            +CR++ Q+ Q L  C+ ++ +Q  +S   P  T+         ME         P + +C
Sbjct  9    KCRKEFQQAQHLRACQQWLHKQAMQSGSGPSWTLDDEFDFEDDMENPQGPQQRPPLLQQC  68

Query  82   CEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGEQ----------MRRMMRLAENIPSRCN  131
            C +L   +  C C  L+     ++Q+  Q  G+Q          + R+ + A ++P  CN
Sbjct  69   CNELHQEEPLCVCPTLKGASKAVKQQVRQQLGQQGQQGPQVQHVISRIYQTATHLPKVCN  128
```

```
Query  132  LSPMR-CPMGGSIAG  145
            +  +  CP   ++ G
Sbjct  129  IPQVSVCPFKKTMPG  143


>gi|1009442|emb|CAA62908.1| allergen sin a 1.0108 [Sinapis alba]
Length=145

 Score = 39.7 bits (91),  Expect = 4e-05
 Identities = 31/135 (22%), Positives = 59/135 (43%), Gaps = 28/135 (20%)

Query  39   ECREQMQRQQMLSHCRMYMRQQMEES---PYQTMPRR-----GME---------PHMSEC  81
            +CR++ Q+ Q L  C+ ++ +Q  +S   P  T+         ME         P + +C
Sbjct  9    KCRKEFQQAQHLRACQQWLHKQAMQSGSGPSWTLDDEFDFEDDMENPQGPQQKPPLLQQC  68

Query  82   CEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGEQ----------MRRMMRLAENIPSRCN  131
            C +L   +  C C  L+     ++Q+  Q  G+Q          + R+ + A ++P  CN
Sbjct  69   CNELHQEEPLCVCPTLKGASKAVKQQVRQQLGQQGQQGPQVQHVISRIYQTATHLPKVCN  128

Query  132  LSPMR-CPMGGSIAG  145
            +  +  CP   ++ G
Sbjct  129  IPQVSVCPFKKTMPG  143


>gi|7545129|gb|AAB25214.2| major allergen Sin a I [Sinapis alba]
 gi|51338758|sp|P15322|ALL1_SINAL Allergen Sin a 1 precursor (Sin a I) [Contains:
Allergen Sin
a 1 small chain; Allergen Sin a 1 large chain]
Length=145

 Score = 39.7 bits (91),  Expect = 4e-05
 Identities = 31/135 (22%), Positives = 59/135 (43%), Gaps = 28/135 (20%)

Query  39   ECREQMQRQQMLSHCRMYMRQQMEES---PYQTMPRR-----GME---------PHMSEC  81
            +CR++ Q+ Q L  C+ ++ +Q  +S   P  T+         ME         P + +C
Sbjct  9    KCRKEFQQAQHLRACQQWLHKQAMQSGSGPSWTLDDEFDFEDDMENPQGPQQRPPLLQQC  68

Query  82   CEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGEQ----------MRRMMRLAENIPSRCN  131
            C +L   +  C C  L+     ++Q+  Q  G+Q          + R+ + A ++P  CN
Sbjct  69   CNELHQEEPLCVCPTLKGASKAVKQQVRQQLGQQGQQGPHLQHVISRIYQTATHLPKVCN  128

Query  132  LSPMR-CPMGGSIAG  145
            +  +  CP   ++ G
Sbjct  129  IRQVSVCPFKKTMPG  143


>gi|121095|sp|P04726|GDA6_WHEAT Alpha/beta-gliadin clone PW1215 precursor (Prolamin)
Length=296

 Score = 39.3 bits (90),  Expect = 5e-05
 Identities = 23/96 (23%), Positives = 47/96 (48%), Gaps = 14/96 (14%)

Query  33   EEENQEECREQMQRQQMLSH--------CRMYMRQQ--MEESPYQTMPRRGMEPHMSECC  82
            +++ Q++ ++Q Q+QQ+L          CR  + QQ  +  +  Q + +   +P    CC
Sbjct  127  QQQQQQQQQQQQQQQQILQQILQQQLIPCRDVVLQQHNIAHARSQVLQQSTYQPLQQLCC  186

Query  83   EQLEGMDESCRCEGL----RMMMMRMQQEEMQPRGE  114
            +QL  + E  RC+ +      +++  QQ + QP  +
Sbjct  187  QQLWQIPEQSRCQAIHNVVHAIILHQQQRQQQPSSQ  222


>gi|1009440|emb|CAA62912.1| allergen sin a 1.0107 [Sinapis alba]
Length=145

 Score = 38.1 bits (87),  Expect = 1e-04
 Identities = 29/135 (21%), Positives = 62/135 (45%), Gaps = 28/135 (20%)

Query  39   ECREQMQRQQMLSHCRMYMRQQMEES---PYQTMPRR-----GME---------PHMSEC  81
            +CR++ Q+ Q L  C+ ++ +Q  +S   P  T+         ME         P + +C
```

```
Sbjct  9    KCRKEFQQAQHLRACQQWLHKQAMQSGSGPSWTLDDEFDFEDDMENPQGPQQRPPLLQQC  68

Query  82   CEQLEGMDESCRC---EGLRMMMMRMQQEEMQPRGEQ-------MRRMMRLAENIPSRCN  131
            C +L   +  C C   +G    + +  +++++ +G+Q       + R+ + A ++P  CN
Sbjct  69   CNELHQEEPLCVCPTLKGASKAVKQQVRQQLEQQGQQGPHLQHVISRIYQTATHLPKVCN  128

Query  132  LSPMR-CPMGGSIAG  145
            +  +  CP   ++ G
Sbjct  129  IPQVSVCPFKKTMPG  143


>gi|1009436|emb|CAA62910.1| allergen sin a 1.0105 [Sinapis alba]
Length=145

 Score = 38.1 bits (87),  Expect = 1e-04
 Identities = 29/135 (21%), Positives = 62/135 (45%), Gaps = 28/135 (20%)

Query  39   ECREQMQRQQMLSHCRMYMRQQMEES---PYQTMPRR-----GME---------PHMSEC  81
            +CR++ Q+ Q L  C+ ++ +Q  +S   P  T+         ME         P + +C
Sbjct  9    KCRKEFQQAQHLRACQQWLHKQAMQSGSGPSWTLEGEFDFEDDMENPQGPQQRPPLLQQC  68

Query  82   CEQLEGMDESCRC---EGLRMMMMRMQQEEMQPRGEQ-------MRRMMRLAENIPSRCN  131
            C +L   +  C C   +G    + +  +++++ +G+Q       + R+ + A ++P  CN
Sbjct  69   CNELHQEEPLCVCPTLKGASKAVKQQVRQQLEQQGQQGPHLQHVISRIYQTATHLPKVCN  128

Query  132  LSPMR-CPMGGSIAG  145
            +  +  CP   ++ G
Sbjct  129  IPQVSVCPFKKTMPG  143


>gi|1009434|emb|CAA62909.1| allergen sin a 1.0104 [Sinapis alba]
Length=145

 Score = 37.7 bits (86),  Expect = 1e-04
 Identities = 29/135 (21%), Positives = 62/135 (45%), Gaps = 28/135 (20%)

Query  39   ECREQMQRQQMLSHCRMYMRQQMEES---PYQTMPRR-----GME---------PHMSEC  81
            +CR++ Q+ Q L  C+ ++ +Q  +S   P  T+         ME         P + +C
Sbjct  9    KCRKEFQQAQHLRACQQWLHKQAMQSGSGPSWTLDDEFDFEDDMENPQGPQQRPPLLQQC  68

Query  82   CEQLEGMDESCRC---EGLRMMMMRMQQEEMQPRGEQ-------MRRMMRLAENIPSRCN  131
            C +L   +  C C   +G    + +  +++++ +G+Q       + R+ + A ++P  CN
Sbjct  69   CNELHQEEPLCVCPTLKGASKAVKQQVRQQLEQQGQQGPHLQHVISRIYQTATHLPRVCN  128

Query  132  LSPMR-CPMGGSIAG  145
            +  +  CP   ++ G
Sbjct  129  IRQVSVCPFQKTMPG  143


>gi|548657|sp|Q01881|RA05_ORYSA Seed allergenic protein RA5 precursor
Length=157

 Score = 37.4 bits (85),  Expect = 2e-04
 Identities = 35/133 (26%), Positives = 55/133 (41%), Gaps = 28/133 (21%)

Query  28   TTTVVEEENQEECREQMQRQQ--------MLSHCRMYMRQQMEESPYQTMPRRGMEPHMS  79
            T T+ E  +Q++      R Q         L  CR  +++Q   S          E
Sbjct  22   TATMAEYHHQDQVVYTRARCQPGMGYPMYSLPRCRALVKRQCRGSA------AAAEQVRR  75

Query  80   ECCEQLEGMDES-CRCEGLRMMMMRMQQEEMQP----------RGEQMRRMMRLAENIPS  128
            +CC QL  +D+S CRCE +  M+  + +E   P          RG +   + R A ++P+
Sbjct  76   DCCRQLAAVDDSWCRCEAISHMLGGIYRELGAPDVGHPMSEVFRGCRRGDLERAAASLPA  135

Query  129  RCNLSPMRCPMGG  141
             CN+     P GG
Sbjct  136  FCNVD---IPNGG  145


>gi|5923742|gb|AAD56337.1| allergen Arah6 [Arachis hypogaea]
Length=129
```

```
 Score = 37.0 bits (84),  Expect = 2e-04
 Identities = 24/86 (27%), Positives = 42/86 (48%), Gaps = 6/86 (6%)

Query  57   MRQQMEESPYQTMPRRGMEPHMSECCEQLEGMD--ESCRCEGLRMMMMRMQQEEMQPRGE  114
            M +Q +   Y     R  +     CC++L  M+  + C CE L+  +M  Q + +Q R +
Sbjct  40   MGEQEQYDSYNFGSTRSSD-QQQRCCDELNEMENTQRCMCEALQ-QIMENQCDGLQDR-Q  96

Query  115  QMRRMMRLAENIPSRCNL-SPMRCPM  139
             ++   R   N+P +CN  +P RC +
Sbjct  97   MVQHFKRELMNLPQQCNFGAPQRCDL  122


 Score = 22.3 bits (46),  Expect = 6.0
 Identities = 11/58 (18%), Positives = 23/58 (39%), Gaps = 3/58 (5%)

Query  35   ENQEECREQMQRQQMLSHCRMYMRQQMEESPYQTMPRRGMEPHMSECCEQLEGMDESC  92
            + Q+ C +++   +    C     QQ+ E+    +  R M  H      +L  + + C
Sbjct  58   DQQQRCCDELNEMENTQRCMCEALQQIMENQCDGLQDRQMVQHFK---RELMNLPQQC  112


>gi|121094|sp|P04725|GDA5_WHEAT Alpha/beta-gliadin A-V precursor (Prolamin)
Length=319

 Score = 36.6 bits (83),  Expect = 3e-04
 Identities = 22/107 (20%), Positives = 53/107 (49%), Gaps = 14/107 (13%)

Query  34   EENQEECREQMQRQQMLSH--------CRMYMRQQ--MEESPYQTMPRRGMEPHMSECCE  83
            ++ Q++ ++Q Q+QQ+L           CR  + QQ  +  +  Q + +   +     CC+
Sbjct  125  QQQQQQQQQQQQQQQILQQILQQQLIPCRDVVLQQHNIAHASSQVLQQSTYQLLQQLCCQ  184

Query  84   QLEGMDESCRCEGL----RMMMMRMQQEEMQPRGEQMRRMMRLAENI  126
            QL  + E  +C+ +      ++M  QQ++ Q + +Q+++  +  + +
Sbjct  185  QLLQIPEQSQCQAIHNVAHAIIMHQQQQQQQEQKQQLQQQQQQQQQL  231


 Score = 28.1 bits (61),  Expect = 0.11
 Identities = 18/100 (18%), Positives = 44/100 (44%), Gaps = 9/100 (9%)

Query  33   EEENQEECREQMQRQQMLSHCRMYMRQQMEESPYQTMPRRGMEPHMSECCEQLEGMDESC  92
            +++ Q+E ++Q+Q+QQ         +QQ+++   Q   +   +    +  +Q      S
Sbjct  210  QQQQQQEQKQQLQQQQQ-------QQQQLQQQQQQQQQQPSSQVSFQQPQQQYPSSQVSF  262

Query  93   RCEGLRMMMMRMQQEEMQPRGEQMRRMMRLAENIPSRCNL  132
            +   L        Q +  P+  ++R +    + +P+ CN+
Sbjct  263  QPSQLNPQAQGSVQPQQLPQFAEIRNL--ALQTLPAMCNV  300


 Score = 25.8 bits (55),  Expect = 0.54
 Identities = 21/73 (28%), Positives = 35/73 (47%), Gaps = 8/73 (10%)

Query  10  ALLVLMALGHATAFRATVTTTVVEEENQEECREQ---MQRQQMLSHCRMYMRQQMEESPY  66
           ALL ++A    TA R  V     +  +Q++ +EQ   +Q+QQ     + +  QQ    PY
Sbjct  8   ALLAIVATTATTAVRVPVPQLQPQNPSQQQPQEQVPLVQQQQFPGQQQQFPPQQ----PY  63

Query  67  -QTMPRRGMEPHM  78
            Q  P    +P++
Sbjct  64  PQPQPFPSQQPYL  76


>gi|121096|sp|P04727|GDA7_WHEAT Alpha/beta-gliadin clone PW8142 precursor (Prolamin)
Length=313

 Score = 36.6 bits (83),  Expect = 3e-04
 Identities = 22/103 (21%), Positives = 52/103 (50%), Gaps = 14/103 (13%)

Query  33   EEENQEECREQMQRQQMLSH--------CRMYMRQQ--MEESPYQTMPRRGMEPHMSECC  82
            +++ Q++ ++Q Q+QQ+L           CR  + QQ  +  +  Q + +   +     CC
Sbjct  123  QQQQQQQQQQQQQQQQILQQILQQQLIPCRDVVLQQHNIAHASSQVLQQSTYQLLQQLCC  182
```

```
Query  83   EQLEGMDESCRCEGL----RMMMMRMQQEEMQPRGEQMRRMMR  121
            +QL  + E  RC+ +      ++M  Q+++ Q + +Q +++ +
Sbjct  183  QQLLQIPEQSRCQAIHNVVHAIIMHQQEQQQQLQQQQQQQLQQ  225


 Score = 24.3 bits (51),  Expect = 1.6
 Identities = 12/54 (22%), Positives = 24/54 (44%), Gaps = 0/54 (0%)

Query  11   LLVLMALGHATAFRATVTTTVVEEENQEECREQMQRQQMLSHCRMYMRQQMEES  64
            LL +       A    V   ++ ++ Q++  +Q Q+QQ+    +   +QQ   S
Sbjct  185  LLQIPEQSRCQAIHNVVHAIIMHQQEQQQQLQQQQQQQLQQQQQQQQQQQQPSS  238


 Score = 24.3 bits (51),  Expect = 1.6
 Identities = 16/97 (16%), Positives = 43/97 (44%), Gaps = 9/97 (9%)

Query  36   NQEECREQMQRQQMLSHCRMYMRQQMEESPYQTMPRRGMEPHMSECCEQLEGMDESCRCE  95
            +Q+E ++Q+Q+QQ     +   +QQ ++ P   +  +  +          +   ++ + +
Sbjct  207  HQQEQQQQLQQQQQQQLQQQQQQQQQQQQPSSQVSFQQPQQQYPSSQGSFQPSQQNPQAQ  266

Query  96   GLRMMMMRMQQEEMQPRGEQMRRMMRLAENIPSRCNL  132
            G         Q +  P+  ++R +    + +P+ CN+
Sbjct  267  G-------SVQPQQLPQFAEIRNL--ALQTLPAMCNV  294


>gi|121093|sp|P04724|GDA4_WHEAT Alpha/beta-gliadin A-IV precursor (Prolamin)
Length=297

 Score = 36.2 bits (82),  Expect = 4e-04
 Identities = 22/101 (21%), Positives = 49/101 (48%), Gaps = 14/101 (13%)

Query  32   VEEENQEECREQMQRQQMLSH--------CRMYMRQQ--MEESPYQTMPRRGMEPHMSEC  81
            + ++ Q++ ++Q Q+QQ+L         CR  + QQ  +     Q + +   +     C
Sbjct  121  ISQQQQQQQQQQQQQQQILQQILQQQLIPCRDVVLQQHSIAHGSSQVLQQSTYQLVQQFC  180

Query  82   CEQLEGMDESCRCEGL----RMMMMRMQQEEMQPRGEQMRR  118
            C+QL  + E  RC+ +      +++  QQ++ Q + +Q ++
Sbjct  181  CQQLWQIPEQSRCQAIHNVVHAIILHQQQQQQQQQQQQQQQ  221


 Score = 23.5 bits (49),  Expect = 2.7
 Identities = 20/72 (27%), Positives = 33/72 (45%), Gaps = 5/72 (6%)

Query  11  LLVLMALGHATAFRATVTTTVVEEENQEECREQMQRQQMLSHCRMYMRQQMEESPYQTMP  70
           +L L A+   TA  A V   V + + Q   ++Q Q+Q  L   + +  QQ    P Q  P
Sbjct  6   ILALRAIVATTATIA-VRVPVPQLQPQNPSQQQPQKQVPLVQQQQFPGQQQPFPPQQPYP  64

Query  71  RR----GMEPHM  78
           ++      +P+M
Sbjct  65  QQQPFPSQQPYM  76


>gi|121092|sp|P04723|GDA3_WHEAT Alpha/beta-gliadin A-III precursor (Prolamin)
Length=282

 Score = 35.8 bits (81),  Expect = 5e-04
 Identities = 20/89 (22%), Positives = 46/89 (51%), Gaps = 8/89 (8%)

Query  33   EEENQEECREQMQRQQMLSHCRMYMRQQ--MEESPYQTMPRRGMEPHMSECCEQLEGMDE  90
            +++ Q++  +Q+ +QQ++  CR  + QQ  +  +  Q + +   +     CC+QL  + E
Sbjct  121  QQQQQQQTLQQILQQQLIP-CRDVVLQQHNIAHASSQVLQQSSYQQLQQLCCQQLFQIPE  179

Query  91   SCRCEGLR-----MMMMRMQQEEMQPRGE  114
              RC+ +      +++   QQ++ QP  +
Sbjct  180  QSRCQAIHNVVHAIILHHHQQQQQPSSQ  208


 Score = 23.9 bits (50),  Expect = 2.1
 Identities = 19/66 (28%), Positives = 31/66 (46%), Gaps = 9/66 (13%)
```

```
Query  10  ALLVLMALGHATAFRATVTTTVVEEENQEECREQ---MQRQQMLSHCRMYMRQQMEESPY  66
           ALL ++A    +A R  V     +  +Q++ +EQ   MQ+QQ       +  QQ +  P
Sbjct  8   ALLAIVATTATSAVRVPVPQLQPQNPSQQQPQEQVPLMQQQQ------QFPGQQEQFPPQ  61

Query  67  QTMPRR  72
           Q  P +
Sbjct  62  QPYPHQ  67


>gi|34900132|ref|NP_911412.1| allergen RA5B precursor [Oryza sativa (japonica cultivar-
group)]
 gi|23495787|dbj|BAC19997.1| allergen RA5B precursor [Oryza sativa (japonica cultivar-
group)]
Length=160

 Score = 35.0 bits (79),  Expect = 9e-04
 Identities = 23/73 (31%), Positives = 36/73 (49%), Gaps = 14/73 (19%)

Query  80   ECCEQLEGMDES-CRCEGLRMMMMRMQQEEMQP----------RGEQMRRMMRLAENIPS  128
            +CC QL  +D+S CRCE +  M+  + +E   P          RG +   + R A ++P+
Sbjct  79   DCCRQLAAVDDSWCRCEAISHMLGGIYRELGAPDVGHPMSEVFRGCRRGDLERAAASLPA  138

Query  129  RCNLSPMRCPMGG  141
             CN+     P GG
Sbjct  139  FCNVD---IPNGG  148


>gi|121098|sp|P18573|GDA9_WHEAT Alpha/beta-gliadin MM1 precursor (Prolamin)
Length=307

 Score = 34.7 bits (78),  Expect = 0.001
 Identities = 24/105 (22%), Positives = 50/105 (47%), Gaps = 30/105 (28%)

Query  33   EEENQEECREQMQRQQMLSH--------CRMYMRQQ----------MEESPYQTMPRRGM  74
            +++ Q++ ++Q Q+QQ+L          CR  + QQ          +++S YQ + +
Sbjct  135  QQQQQQQKQQQQQQQQILQQILQQQLIPCRDVVLQQHSIAYGSSQVLQQSTYQLVQQL--  192

Query  75   EPHMSECCEQLEGMDESCRCEGL----RMMMMRMQQEEMQPRGEQ  115
                  CC+QL  + E  RC+ +      +++  QQ++ Q + +Q
Sbjct  193  ------CCQQLWQIPEQSRCQAIHNVVHAIILHQQQQQQQQQQQQ  231


>gi|34900098|ref|NP_911395.1| seed allergen RAG2 [Oryza sativa (japonica cultivar-group)]
 gi|23616951|dbj|BAC20654.1| seed allergen RAG2 [Oryza sativa (japonica cultivar-group)]
 gi|548671|sp|Q01885|RAG2_ORYSA Seed allergenic protein RAG2 precursor
Length=166

 Score = 34.7 bits (78),  Expect = 0.001
 Identities = 37/141 (26%), Positives = 63/141 (44%), Gaps = 18/141 (12%)

Query  6    VAAAALLVLMALGHATAFRATVTTTVVEEENQEECREQM-QRQQMLSHCRMYMRQQMEES  64
            V +A LL+++++  ATA  A      V      E C+  M       L  CR  +++Q
Sbjct  7    VFSALLLIIVSVLAATATMADHHKDQVVYSLGERCQPGMGYPMYSLPRCRAVVKRQCVG-  65

Query  65   PYQTMPRRGMEPHM-SECCEQLEGMDES-CRCEGLRMMMMRMQQE-----------EMQP  111
                 P   ++  +  +CC QL  +D+S CRC  L  M+  + +E           E+ P
Sbjct  66   --HGAPGGAVDEQLRQDCCRQLAAVDDSWCRCSALNHMVGGIYRELGATDVGHPMAEVFP  123

Query  112  RGEQMRRMMRLAENIPSRCNL  132
             G +   + R A ++P+ CN+
Sbjct  124  -GCRRGDLERAAASLPAFCNV  143


>gi|585290|sp|P32936|IAAB_HORVU Alpha-amylase/trypsin inhibitor CMb precursor
(Chloroform/methanol-soluble
protein CMb)
Length=149

 Score = 34.3 bits (77),  Expect = 0.002
 Identities = 26/92 (28%), Positives = 37/92 (40%), Gaps = 14/92 (15%)
```

```
Query  50   LSHCRMYMRQQMEESPYQTMPRRGMEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQEEM  109
            L  CR Y+ QQ         P    +    +CC +L  + + CRC+ LR  M R  + +
Sbjct  41   LPSCRDYVEQQACRIETPGPPYLAKQ----QCCGELANIPQQCRCQALRFFMGRKSRPDQ  96

Query  110  Q--------PRGEQMRRMMRLAENIPSRCNLS  133
                     PR  QM  +  L    P  CNL+
Sbjct  97   SGLMELPGCPREVQMDFVRILV--TPGFCNLT  126


>gi|6580762|gb|AAF18269.1| vicilin-like protein precursor [Juglans regia]
Length=593

 Score = 33.9 bits (76),  Expect = 0.002
 Identities = 31/126 (24%), Positives = 50/126 (39%), Gaps = 25/126 (19%)

Query  37   QEECREQMQRQQMLSHCRMYMRQQMEE---SPYQTMPRRG--------------MEPHMS  79
            QE CR Q Q Q+    C++   +++EE   S  +   RRG               E
Sbjct  22   QEYCRRQGQGQRQQQQCQIRCEERLEEDQRSQEERERRRGRDVDDQNPRDPEQRYEQCQQ  81

Query  80   ECCEQLEGMDESC---RCEGLRMMMMRMQQ-----EEMQPRGEQMRRMMRLAENIPSRCN  131
            +C  Q  G +++    RCE  R    R +Q     ++ Q +  + +R  ++ E  P R
Sbjct  82   QCERQRRGQEQTLCRRRCEQRRQQEERERQRGRDRQDPQQQYHRCQRRCQIQEQSPERQR  141

Query  132  LSPMRC  137
                RC
Sbjct  142  QCQQRC  147


 Score = 26.6 bits (57),  Expect = 0.32
 Identities = 21/88 (23%), Positives = 39/88 (44%), Gaps = 8/88 (9%)

Query  34   EENQEECREQMQRQ-----QMLSHCRMYMRQQMEESPYQT-MPRRGMEPHMSECCEQLEG  87
            E+  E+C++Q +RQ     Q L   R   R+Q EE   Q    R+  +     C  + +
Sbjct  73   EQRYEQCQQQCERQRRGQEQTLCRRRCEQRRQQEERERQRGRDRQDPQQQYHRCQRRCQI  132

Query  88   MDESCRCEGLRMMMMRMQQEEMQPRGEQ  115
             ++S   E  R    R +++  + +G +
Sbjct  133  QEQS--PERQRQCQQRCERQYKEQQGRE  158


 Score = 22.7 bits (47),  Expect = 4.6
 Identities = 12/38 (31%), Positives = 18/38 (47%), Gaps = 8/38 (21%)

Query  35   ENQEECREQMQRQQMLSHCRMYMRQQMEESPYQTMPRR  72
            E Q +C+++ +RQ        Y  QQ  E   +  PRR
Sbjct  138  ERQRQCQQRCERQ--------YKEQQGRERGPEASPRR  167


>gi|407610|gb|AAB27813.1| Bra j IE large chain=allergen [Brassica juncea=oriental-mustard,
Peptide, 92 aa]
 gi|32363446|sp||P80207_2 [Segment 2 of 2] Allergen Bra j 1-E (Bra j I) [Contains:
Allergen
Bra j 1-E small chain; Allergen Bra j 1-E large chain]
 gi|479902|pir||S35592 major allergen large chain - leaf mustard
Length=92

 Score = 33.1 bits (74),  Expect = 0.003
 Identities = 19/81 (23%), Positives = 39/81 (48%), Gaps = 11/81 (13%)

Query  76   PHMSECCEQLEGMDESCRCEGLR------MMMMRMQQEEMQPRGEQMR----RMMRLAEN  125
            P + +CC +L   +  C C  L++         +R Q ++   +G+Q++    R+ + A +
Sbjct  10   PLLQQCCNELHQEEPLCVCPTLKGASKAVKQQIRQQGQQQGQQGQQLQHEISRIYQTATH  69

Query  126  IPSRCNLSPMR-CPMGGSIAG  145
            +P  CN+  +  CP   ++ G
Sbjct  70   LPRVCNIPRVSICPFQKTMPG  90


>gi|121091|sp|P04722|GDA2_WHEAT Alpha/beta-gliadin A-II precursor (Prolamin)
```

```
Length=291

 Score = 33.1 bits (74),  Expect = 0.003
 Identities = 25/114 (21%), Positives = 52/114 (45%), Gaps = 29/114 (25%)

Query  32   VEEENQEECREQMQRQQMLSH--------CRMYMRQQ----------MEESPYQTMPRRG  73
            + ++ Q++ ++Q Q+QQ+L          CR  + QQ          ++ES YQ + +
Sbjct  114  ISQQQQQQQQQQQQQQQQILQQILQQQLIPCRDVVLQQHNIAHGSSQVLQESTYQLVQQL-  172

Query  74   MEPHMSECCEQLEGMDESCRCEGLRMMMMRM---QQEEMQPRGEQMRRMMRLAE  124
                   CC+QL  + E  RC+ +  ++  +   QQ     + +Q ++   L++
Sbjct  173  -------CCQQLWQIPEQSRCQAIHNVVHAIILHQQHHHHQQQQQQQQQQPLSQ  219


 Score = 26.2 bits (56),  Expect = 0.41
 Identities = 21/73 (28%), Positives = 35/73 (47%), Gaps = 8/73 (10%)

Query  10  ALLVLMALGHATAFRATVTTTVVEEENQEECREQ---MQRQQMLSHCRMYMRQQMEESPY  66
           ALL ++A    TA R  V    ++  +Q++ +EQ   +Q QQ     + +  QQ    PY
Sbjct  8   ALLAIVATTATTAVRVPVPQLQLQNPSQQQPQEQVPLVQEQQFQGQQQPFPPQQ----PY  63

Query  67  -QTMPRRGMEPHM  78
            Q  P    +P++
Sbjct  64  PQPQPFPSQQPYL  76


>gi|67464993|sp|P02863|GDA0_WHEAT Alpha/beta-gliadin precursor (Prolamin)
Length=286

 Score = 28.1 bits (61),  Expect = 0.11
 Identities = 16/96 (16%), Positives = 46/96 (47%), Gaps = 22/96 (22%)

Query  33   EEENQEECREQMQRQQMLSHCRMYMRQQ---------MEESPYQTMPRRGMEPHMSECCE  83
            +++ Q++  +Q+ +QQ++    + ++Q          +++S YQ +           CC+
Sbjct  127  QQQQQQQILQQILQQQLIPCMDVVLQQHNIAHGRSQVLQQSTYQLLQEL--------CCQ  178

Query  84   QLEGMDESCRCEGLR-----MMMMRMQQEEMQPRGE  114
             L  + E  +C+ +      +++ + Q+++ QP  +
Sbjct  179  HLWQIPEQSQCQAIHNVVHAIILHQQQKQQQQPSSQ  214


>gi|121090|sp|P04721|GDA1_WHEAT Alpha/beta-gliadin A-I precursor (Prolamin)
Length=262

 Score = 27.7 bits (60),  Expect = 0.14
 Identities = 22/73 (30%), Positives = 36/73 (49%), Gaps = 8/73 (10%)

Query  10  ALLVLMALGHATAFRATVTTTVVEEENQEECREQ---MQRQQMLSHCRMYMRQQMEESPY  66
           ALL ++A    TA R  V     +  +Q++ +EQ   +Q+QQ L   + +  QQ    PY
Sbjct  8   ALLAIVATTATTAVRVPVPQLQPQNPSQQQPQEQVPLVQQQQFLGQQQPFPPQQ----PY  63

Query  67  -QTMPRRGMEPHM  78
            Q  P    +P++
Sbjct  64  PQPQPFPSQQPYL  76


 Score = 27.7 bits (60),  Expect = 0.14
 Identities = 16/96 (16%), Positives = 46/96 (47%), Gaps = 22/96 (22%)

Query  33   EEENQEECREQMQRQQMLSHCRMYMRQQ---------MEESPYQTMPRRGMEPHMSECCE  83
            +++ Q++  +Q+ +QQ++    + ++Q          +++S YQ +           CC+
Sbjct  128  QQQQQQQIIQQILQQQLIPCMDVVLQQHNIVHGKSQVLQQSTYQLLQEL--------CCQ  179

Query  84   QLEGMDESCRCEGLR-----MMMMRMQQEEMQPRGE  114
             L  + E  +C+ +      +++ + Q+++ QP  +
Sbjct  180  HLWQIPEQSQCQAIHNVVHAIILHQQQKQQQQPSSQ  215


>gi|21954740|gb|AAM83103.1| paramyosin allergen [Blomia tropicalis]
 gi|42559485|sp|Q8MUF6|MYSP_BLOTA Paramyosin (Allergen Blo t 11)
```

```
Length=875

 Score = 27.3 bits (59),  Expect = 0.19
 Identities = 24/95 (25%), Positives = 39/95 (41%), Gaps = 8/95 (8%)

Query  19   HATAFRATVTTTVVEEENQEECREQMQRQQMLSHCRMYMRQQMEESPYQTMP------RR  72
            HA      + +  V+ E + E R +++RQ   ++      +   E+  Q         RR
Sbjct  256  HAHTLEVELESLKVQLEEESEARLELERQLTKANGDAASWKSKYEAELQAHVDEVEELRR  315

Query  73   GMEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQE  107
             M   +SE  EQLE +    +C  L     R+Q E
Sbjct  316  KMAQKISEYGEQLEALLN--KCSALEKQKARLQSE  348


 Score = 25.4 bits (54),  Expect = 0.71
 Identities = 16/60 (26%), Positives = 25/60 (41%), Gaps = 0/60 (0%)

Query  67   QTMPRRGMEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGEQMRRMMRLAENI  126
            Q +   G++ H  E   QL+   +       R   +  + EEM+   EQ  R  R AE +
Sbjct  571  QALQITGLQAHYDEVHRQLQQAVDQLGVTQRRCQALTAELEEMRVNLEQALRAKRAAEQM  630


>gi|121097|sp|P04728|GDA8_WHEAT Alpha/beta-gliadin clone PTO-A10 (Prolamin)
Length=186

 Score = 26.9 bits (58),  Expect = 0.24
 Identities = 20/106 (18%), Positives = 45/106 (42%), Gaps = 32/106 (30%)

Query  33   EEENQEECREQMQRQQMLSH--------CRMYMRQQ----------MEESPYQTMPRRGM  74
            +++ Q++ ++Q Q QQ+L          C   + QQ          +++S YQ +
Sbjct  18   QQQQQQQQQQQQQEQQILQQILQQQLIPCMDVVLQQHNIAHGRSQVLQQSTYQLLQEL--  75

Query  75   EPHMSECCEQLEGMDESCRCEGLR------MMMMRMQQEEMQPRGE  114
                  CC+ L  + E  +C+ +       ++  + Q+++ QP  +
Sbjct  76   ------CCQHLWQIPEQSQCQAIHNVVHAIILHQQQQKQQQQPSSQ  115


>gi|2739154|gb|AAC67308.1| 22.6 kDa tegumental antigen [Schistosoma japonicum]
Length=191

 Score = 23.9 bits (50),  Expect = 2.1
 Identities = 14/76 (18%), Positives = 35/76 (46%), Gaps = 8/76 (10%)

Query  35   ENQEECREQMQRQQMLSHCRMYMRQQMEESPYQTMPRRGMEPHMSECCEQLEGMDESCRC  94
            E  ++  E + +Q++  +C+       +  P+        +  +S        ++E CR
Sbjct  20   EIDKDNNELIDKQELTKYCQQNQMDMKQIDPWIARFDTDKDGKVS--------LEEFCRG  71

Query  95   EGLRMMMMRMQQEEMQ  110
             GL++  +R ++EE++
Sbjct  72   FGLKVWEVRREKEELK  87


>gi|6680437|ref|NP_032399.1| interleukin 9 [Mus musculus]
Length=144

 Score = 23.5 bits (49),  Expect = 2.7
 Identities = 15/57 (26%), Positives = 26/57 (45%), Gaps = 2/57 (3%)

Query  92   CRCEGLRMMMMRMQQEEMQPRGEQMRRMMRLAENI--PSRCNLSPMRCPMGGSIAGF  146
            C  EGL  +    Q+  + P   +++R++ + +NI  PS     P    M G+   F
Sbjct  68   CYREGLLQLTNATQKSRLLPVFHRVKRIVEVLKNITCPSFSCEKPCNQTMAGNTLSF  124


>gi|42559558|sp|O97192|TPM_HELAS Tropomyosin (Allergen Hel as 1)
Length=284

 Score = 23.1 bits (48),  Expect = 3.5
 Identities = 9/27 (33%), Positives = 17/27 (62%), Gaps = 0/27 (0%)

Query  98   RMMMMRMQQEEMQPRGEQMRRMMRLAE  124
```

```
            +M+ M+M++E    R EQ+ + +R  E
Sbjct  7    KMLAMKMEKENALDRAEQVEQKLRDCE  33


>gi|13785807|gb|AAK39511.1| paramyosin-like allergen [Dermatophagoides farinae]
 gi|42559514|sp|Q967Z0|MYSP_DERFA Paramyosin (Allergen Der f 11) (Antigen Df642)
Length=692

 Score = 23.1 bits (48),  Expect = 3.5
 Identities = 14/53 (26%), Positives = 23/53 (43%), Gaps = 0/53 (0%)

Query  74   MEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGEQMRRMMRLAENI  126
            ++ H  E   QL+   +       R   ++ + EEM+   EQ  R  R AE +
Sbjct  492  LQAHYDEVHRQLQQAVDQLGVTQRRCQALQAELEEMRIALEQANRAKRQAEQL  544


>gi|37778944|gb|AAO73464.1| HDM allergen [Dermatophagoides pteronyssinus]
Length=875

 Score = 22.7 bits (47),  Expect = 4.6
 Identities = 14/53 (26%), Positives = 23/53 (43%), Gaps = 0/53 (0%)

Query  74   MEPHMSECCEQLEGMDESCRCEGLRMMMMRMQQEEMQPRGEQMRRMMRLAENI  126
            ++ H  E   QL+   +       R   ++ + EEM+   EQ  R  R AE +
Sbjct  578  LQAHYDEVHRQLQQAVDQLGVTQRRCQALQAELEEMRIALEQASRAKRQAEQL  630


>gi|31321944|gb|AAM54366.1| vicilin seed storage protein [Juglans nigra]
Length=481

 Score = 22.7 bits (47),  Expect = 4.6
 Identities = 12/38 (31%), Positives = 18/38 (47%), Gaps = 8/38 (21%)

Query  35  ENQEECREQMQRQQMLSHCRMYMRQQMEESPYQTMPRR  72
           E Q +C+++ +RQ        Y  QQ  E   +  PRR
Sbjct  26  ERQRQCQQRCERQ--------YKEQQGRERGPEASPRR  55


>gi|69937363|ref|ZP_00632029.1| Allergen V5/Tpx-1 related [Paracoccus denitrificans
PD1222]
 gi|69151326|gb|EAN64530.1| Allergen V5/Tpx-1 related [Paracoccus denitrificans PD1222]
Length=199

 Score = 22.7 bits (47),  Expect = 4.6
 Identities = 16/76 (21%), Positives = 29/76 (38%), Gaps = 0/76 (0%)

Query  3    KISVAAAALLVLMALGHATAFRATVTTTVVEEENQEECREQMQRQQMLSHCRMYMRQQME  62
            KI+    A +    LG   A RA V    + +    +   ++  + M +  R +
Sbjct  49   KITPREEAEIPGRVLGQINALRANVAAPAMMQSPMLDAAAKVHARDMAAQNRAWHFGSDG  108

Query  63   ESPYQTMPRRGMEPHM  78
             SP   + R+G   H+
Sbjct  109  SSPLDRVRRQGYSGHL  124


>gi|9954251|gb|AAG08988.1| tropomyosin [Perna viridis]
Length=284

 Score = 22.3 bits (46),  Expect = 6.0
 Identities = 8/32 (25%), Positives = 19/32 (59%), Gaps = 0/32 (0%)

Query  98   RMMMMRMQQEEMQPRGEQMRRMMRLAENIPSR  129
            +M+ M+M+++    R EQ+ + +R  E   ++
Sbjct  7    KMVAMKMEKKNALDRAEQLEQKLRETEEAKAK  38


>gi|8118425|gb|AAF72985.1| beta-expansin [Oryza sativa]
Length=286

 Score = 21.9 bits (45),  Expect = 7.8
```

```
 Identities = 12/22 (54%), Positives = 14/22 (63%), Gaps = 2/22 (9%)

Query  7   AAAALLVLMALGHATA--FRAT  26
           AAA  L L+A+GH  A  F AT
Sbjct  9   AAAVFLSLLAVGHCAAADFNAT  30


>gi|28476851|tpg|DAA00353.1| TPA_exp: allergen dI chain C2D [Mus musculus]
Length=112

 Score = 21.9 bits (45),  Expect = 7.8
 Identities = 13/34 (38%), Positives = 19/34 (55%), Gaps = 0/34 (0%)

Query  18  GHATAFRATVTTTVVEEENQEECREQMQRQQMLS  51
           G  + F ATV   V  E+ QE  RE+ Q+  +L+
Sbjct  44  GDLSQFYATVAERVAFEKIQECFREEGQKTIILN  77


>gi|37536762|ref|NP_922683.1| beta-expansin EXPB4 [Oryza sativa (japonica cultivar-
group)]
 gi|14165334|gb|AAK55466.1| beta-expansin (EXPB4) [Oryza sativa (japonica cultivar-
group)]
 gi|31433457|gb|AAP54970.1| beta-expansin [Oryza sativa (japonica cultivar-group)]
Length=286

 Score = 21.9 bits (45),  Expect = 7.8
 Identities = 12/22 (54%), Positives = 14/22 (63%), Gaps = 2/22 (9%)

Query  7   AAAALLVLMALGHATA--FRAT  26
           AAA  L L+A+GH  A  F AT
Sbjct  9   AAAVFLSLLAVGHCAAADFNAT  30



  Database: All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF excluding
environmental samples
    Posted date:  Apr 4, 2006  3:06 AM
  Number of letters in database: 583,171
  Number of sequences in database:  2,529
Lambda     K      H
   0.323    0.129    0.383
Gapped
Lambda     K      H
   0.267   0.0410    0.140
Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 2529
Number of Hits to DB: 6922
Number of extensions: 221
Number of successful extensions: 1
Number of sequences better than 10: 1
Number of HSP's better than 10 without gapping: 0
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Length of query: 146
Length of database: 583171
Length adjustment: 69
Effective length of query: 77
Effective length of database: 583171
Effective search space: 44904167
Effective search space used: 31467590
T: 11
A: 40
X1: 16 (7.5 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 40 (20.0 bits)
S2: 45 (21.9 bits)
```