

8) Hileman, R. E., E. A. Rice, R. E. Goodman and J. D. Astwood (2001).  
Bioinformatics evaluation of the Cry3Bb1 protein produced in corn event MON 863  
utilizing allergen, toxin and public domain protein databases. MSL-17140, an  
unpublished analysis prepared by Monsanto Company.



Study Title

**Bioinformatics Evaluation of the Cry3Bb1 Protein Produced in Corn Event  
MON 863 Utilizing Allergen, Toxin and Public Domain Protein Databases**

Authors

**Ronald E. Hileman, Ph.D., Elena A. Rice, Ph.D., Richard E. Goodman, Ph.D. and  
James D. Astwood, Ph.D.**

Study Completed On

**April 17, 2001**

Performing Laboratory

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

Laboratory Project ID

**MSL Number: 17140**

**Study Number: 01-01-39-23**

Monsanto Company  
Final Report  
Product Safety Center

Study No. 01-01-39-23  
MSL No. 17140  
Page 2 of 102

### Statement of No Data Confidentiality

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

"We submit this material to the United States Environmental Protection Agency specifically under the requirements set forth in FIFRA as amended, and consent to the use and disclosure of this material by EPA strictly in accordance with FIFRA. By submitting this material to EPA in accordance with the method and format requirements contained in PR Notice 86-5, we reserve and do not waive any rights involving this material that are or can be claimed by the company notwithstanding this submission to EPA."

Company: Monsanto Company  
Company Agent: \_\_\_\_\_  
Title: \_\_\_\_\_  
Signature: \_\_\_\_\_ Date: \_\_\_\_\_

©01

Proprietary Information of Monsanto Company



Monsanto Company  
Final Report  
Product Safety Center

Study No. 01-01-39-23  
MSL No. 17140  
Page 3 of 102

### Statement of Compliance

This bioinformatics analysis does not fall under the scope of the Good Laboratory Practice Standards as specified in 40 CFR Part 160 because it is not a study as defined in this part.

Submitter

Date:

Sponsor

Date:

Author

Date:

©01

Proprietary Information of Monsanto Company

Monsanto Company  
Final Report  
Product Safety Center

Study No. 01-01-B9-23  
MSL No. 17140  
Page 4 of 102

**MSL Number:** 17140

**Title:** Bioinformatics Evaluation of the Cry3Bb1 Protein Produced in Corn Event MON 863 Utilizing Allergen, Toxin and Public Domain Protein Databases

**Facility:** Monsanto Company  
Product Safety Center  
Biotechnology Regulatory Sciences  
700 Chesterfield Parkway North  
Chesterfield, Missouri 63198

**Sponsor:** Ravinder S. Sidhu

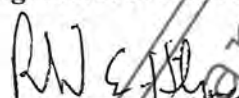
**Authors:** Ronald E. Hileman  
Elena A. Rice  
Richard E. Goodman  
James D. Astwood

**Study Start Date:** February 12, 2001

**Study Completion Date:** April 17, 2001

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

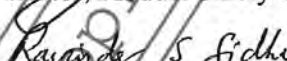
**Signatures of Final Report Approval:**

  
\_\_\_\_\_  
Author

Date: 17 APR 2001

  
\_\_\_\_\_  
Director, Product Safety Center

Date: April 17, 2001

  
\_\_\_\_\_  
Sponsor

Date: April 17, 2001

### Summary of Quality Control Review

This report was reviewed to ensure that it accurately reflects the raw data of the study. The raw data was audited for compliance to the Monsanto Company Guidelines for Keeping Research Records (GRR 10/1/99), and where applicable, to Monsanto SOP's.

*Paula A. Price*

Quality Assurance Specialist  
Monsanto Regulatory  
Monsanto Company

Date: *April 17, 2001*

## Table of Contents

Section	Page
Title Page.....	1
Statement of No Data Confidentiality Claims.....	2
Statement of Compliance .....	3
Signatures of Final Report Approval.....	4
Summary of Quality Control Review.....	5
Table of Contents .....	6
Abbreviations and Definitions .....	8
1.0 Summary .....	9
2.0 Introduction .....	10
3.0 Purpose.....	13
4.0 Methods.....	13
4.1 Sequence database preparation.....	13
4.2 Sequence database searches .....	14
5.0 Results and Discussion.....	16
5.1 Allergenic assessment of the Cry3Bb1 protein produced by corn event MON 863.....	16
5.2 Toxicity assessment of the Cry3Bb1 protein produced by corn event MON 863.....	17
5.3 Similarity of the Cry3Bb1 protein produced by corn event MON 863 to all publicly available proteins .....	17
6.0 Conclusions.....	18
7.0 References .....	18

## Tables

Table 1.	Amino acid diversity of Cry3Bb1 protein variants .....	11
Table 2.	Best similarities observed to corn event MON 863 Cry3Bb1.11098 protein sequence from FASTA searches against allergen/gliadin, toxin and the ALLPEPTIDES protein databases.....	16

## Figures

Figure 1.	The deduced sequence of the Cry3Bb1.11098 protein produced in corn event MON 863 .....	23
-----------	--	----

## Appendices

Appendix 1.	Allergen and gliadin sequence database (ALLERGEN3) .....	24
Appendix 2.	FASTA sequence similarity alignments of corn event MON 863 Cry3Bb1.11098 protein with allergens and gliadins.....	40
Appendix 3.	Search for immunologically relevant sequences .....	43
Appendix 4.	FASTA sequence similarity alignments of corn event MON 863 Cry3Bb1.11098 protein with toxins.....	44
Appendix 5.	FASTA sequence similarity similarities of corn event MON 863 Cry3Bb1.11098 protein with publicly available protein databases .....	100



### Abbreviations and Definitions

aa	amino acid
ALLERGEN3	Allergen and gliadin protein sequence database
ALLPEPTIDES	Protein sequence database comprised of GenPept, PIR and SwissProt
BLOCKS	A database of amino acid motifs found in protein families
BLOSUM	BLOcks SUBstitution Matrix, used to score similarities between pairs of distantly related of protein or nucleotide sequences
DATASET	Command used to create a GCG data library from a set of sequences in GCG format
EMBL	A public genetic database maintained by the European Molecular Biology Laboratory at the European Bioinformatics Institute, Hinxton, England
FASTA	Algorithm used to find local high scoring alignments between a pair of protein or nucleotide sequences
GCG	Genetics Computer Group
GenBank	A public genetic database maintained by the National Center for Biotechnology Information at the National Institutes of Health, Bethesda, MD
IDENTITYSEARCH	Algorithm used to find immunologically relevant amino acid identities between a pair of protein sequences, also referred to as ALLERGENSEARCH
NRL3D	National Research Laboratory's protein 3-dimensional protein database founded at Brookhaven National Laboratory and maintained by the RCSB
PDB	Protein Database
PIR	Protein Information Database
RCSB	Research Collaboratory for Structural Bioinformatics
STRINGSEARCH	Algorithm used to identify sequence entries by searching for character patterns, such as "toxin", in the annotation section of database flatfiles
SwissProt	Translated sequences from EMBL database
TOXIN4	Toxin protein sequence database
UNIX	A computer operating system originally termed "UNiplexed Information and Computing Service"

## 1.0 Summary

Corn rootworm protected corn event MON 863 was designed by the insertion of a gene encoding a modified *Bacillus thuringiensis* Cry3Bb1 protein into corn cells. Bioinformatic analyses were performed to assess this Cry3Bb1 protein for potential similarity to known allergens, toxins or other pharmacologically active proteins relevant to human and animal health. Protein sequence databases were assembled for this purpose and included allergen and gliadin (ALLERGEN3), toxin (TOXIN4) and the public domain (ALLPEPTIDES) sequence databases.

The FASTA sequence alignment tool was used to assess structural similarity. Although the FASTA program directly compares amino acid sequences (*i.e.* primary protein structure), the alignment data may be used to infer higher order structure (*i.e.* secondary and tertiary protein structure). 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. Therefore, proteins homologous to allergens are more likely to share cross-reactive allergenic epitopes than are unrelated proteins.

Structural similarities between the Cry3Bb1 protein sequence and the aligned database sequences were examined. The extent of each similarity was evaluated by visual inspection of the alignment, the calculated percent identity and *E* score value. The Cry3Bb1 protein sequence was also screened against the ALLERGEN3 sequence database for immunologically relevant similarity using a pair-wise comparison algorithm. In these analyses, sequences of eight (or more) linearly contiguous and identical amino acids were defined as immunologically relevant. The presence of such identities may point to the presence of potentially cross-reactive allergenic epitopes.

No biologically relevant structural similarities were observed between any known allergen or toxin and the Cry3Bb1 protein produced in corn event MON 863. Further, no immunologically relevant sequence similarities were observed between this Cry3Bb1 protein and proteins in the allergen and gliadin database. These data demonstrate the absence of both structurally and immunologically relevant similarities between allergens and the Cry3Bb1 protein produced in corn event MON 863.

Apart from expected similarities to other known crystal (Cry) proteins found in *Bacillus thuringiensis* and related species, no additional significant structural similarities were observed. The Cry family represents a diverse set of proteins derived from a common ancestral gene.

The results of these bioinformatics analyses indicate that the Cry3Bb1 protein produced by corn event MON 863 is not similar to known allergens, toxins or other pharmacologically active proteins relevant to animal or human health.

## 2.0 Introduction

Corn rootworm protected corn event MON 863 was designed by the insertion of a gene encoding a modified *Bacillus thuringiensis* Cry3Bb1 protein into corn cells. Plants producing the Cry3Bb1 protein are resistant to feeding damage by the coleopteran insect, corn rootworm. The wild type *cry3Bb1* gene (Accession No. M89794) was modified specifically to both increase bioactivity to corn rootworm as well as increase protein production in plants.

Initial modifications of the wild type *Bacillus thuringiensis* (*B.t.*) *cry3Bb1* gene (Donovan *et al.*, 1992) were performed using a *B.t.* cloning vector and expression system (English *et al.*, 2000). Bioactivity analyses were used to select variant *cry3Bb1* genes that produced Cry3Bb1 proteins in *B.t.* with increased activity to the larvae of a target insect, Southern corn rootworm. The modified *cry3Bb1* gene contained in *B.t.* strain EG11098 (producing the Cry3Bb1.11098 protein) was selected for corn transformation as part of this extensive process. Corn event MON 863 Cry3Bb1.11098 protein is the focus of this report. Other corn events producing the Cry3Bb1.11098 protein, Cry3Bb1.11231 as well as a separately modified *Bacillus thuringiensis* produced Cry3Bb1 protein are also described for clarity as shown in Table 1.

The Cry3Bb1.11098 protein produced in *B.t.* differs from the wild type Cry3Bb1 protein by 5 amino acids. When this *B.t.* gene was modified for expression in corn, the resulting gene is predicted to encode an additional two changes relative to the *B.t.* Cry3Bb1.11098 protein sequence. Thus the Cry3Bb1.11098 protein produced in corn event MON 863 differs from the *B.t.* produced Cry3Bb1.11098 protein by 2 amino acids and 7 amino acids relative to the wild type Cry3Bb1 protein.

Additional characterization of the DNA sequence of the *cry3Bb1* coding region in corn event MON 863 and Cry3Bb1 protein produced in corn event MON 863 confirmed these described amino acid differences relative to the deduced amino acid sequence of the wild type *cry3Bb1* gene (Hileman and Astwood, 2001).

**Table 1.** Amino acid diversity of Cry3Bb1 protein variants.

Cry3Bb1 Allele	Amino acid substitutions relative to wild type Cry3Bb1 protein <sup>a</sup>					
<i>B.t.</i> strain EG11098 Cry3Bb1.11098 protein	D165G	H231R	S311L	N313T	E317K	
<i>B.t.</i> strain EG11231 Cry3Bb1.11231 protein		H231R	S311L	N313T	E317K	
Corn event MON 863 Cry3Bb1.11098 protein <sup>b</sup>	A2	D166G	H232R	S312L	N314T	E318K Q349R
Corn event MON 853 Cry3Bb1.11231 protein <sup>b</sup>	A2		H232R	S312L	N314T	E318K

<sup>a</sup> wild type Cry3Bb1 protein corresponds to Accession No. M89794.

<sup>b</sup> Amino acid positions are shifted upwards of one amino acid relative to the *B.t.* proteins due to the introduction of an alanine at position 2.

Exposure to allergens in foods may pose a risk to sensitized, allergic individuals. Type I (immediate hypersensitivity) reactions occur when a sufficient quantity of individual allergenic proteins, each bound by two different IgE antibodies attached to IgE receptors on the surface of mast cells, stimulate the cells to release histamine and allergenic mediators. The two binding sites (epitopes) are generally different structures or sequences, unless two identical proteins are covalently linked. Sensitization to produce the IgE antibodies requires a third allergen specific recognition sequence, 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 have lower affinity than the immunogen for the antibody and are therefore typically less effective at inducing an immune response (Berzofsky *et al.*, 1993). A specific class of proteins, gliadins, are suspected of causing celiac disease (gluten-sensitive enteropathy), another immunologically mediated disease induced by the consumption of specific proteins by susceptible individuals (Ferguson, 1996).

A bioinformatic assessment is performed for each protein introduced into crops through the use of biotechnology to ensure that the introduced protein is not an allergen, not likely to act as a cross-reactive allergen, and is not a toxin. Screening the amino acid sequences of proteins introduced into plants by genetic engineering for similarity to sequences of known allergens and gliadins is one of many assessments performed to evaluate product safety (Metcalf *et al.*, 1996). Similarly, the amino acid sequence of introduced protein is also screened against known toxins and all known proteins in publicly available sequence databases to identify potential safety concerns.



The extent of sequence similarities between the Cry3Bb1 protein produced in corn event MON 863 and database sequences of allergens, gliadins, toxins and other proteins were assessed using the FASTA sequence alignment tool. Although the FASTA program directly compares amino acid sequences (*i.e.* primary protein structure), the alignment data may be used to infer higher order structure (*i.e.* secondary and tertiary protein structure). 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. The minimum requirement for allergenic cross-reactivity of conformational epitopes may be at least 50% identity across the full length of the protein sequences (Aalberse, 2000).

A second bioinformatics tool (IDENTITYSEARCH) was used to specifically search for immunologically relevant sequences in the corn event MON 863 Cry3Bb1.11098 protein. It is possible that proteins structurally unrelated to allergens and gliadins may still contain smaller immunologically significant epitopes. A sequence length of eight contiguous amino acids was chosen as a target to identify potentially cross reactive allergenic epitopes (Metcalf *et al.*, 1996). The IgE binding epitopes of only a few allergens have been fully mapped using sera from a few allergic individuals in *in vitro* IgE binding studies. While some IgE epitopes may be as short as 5 amino acids (Banerjee *et al.*, 1999; Beezhold *et al.*, 1999), the majority of characterized IgE linear epitopes are 8 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 8 or more amino acids (Banerjee *et al.*, 1999; Rabjohn *et al.*, 1999). Additionally, the CD4+ MHC II presented T-cell epitopes that are required for induction and maintenance of the allergic reaction are generally comprised of 12 or more amino acids (Plaut and Zimmerman, 1993).

In the absence of a complete description of IgE epitopes for all known allergens, a theoretical database of all potential epitopes for these same allergens can effectively be created by scanning all overlapping peptides (in this case 8 or more amino acids in length) of all the allergens of the database and comparing in pairwise fashion to all same-size potential peptides of the test protein, in this case Cry3Bb1, using computer software or scanning manually. This approach can be viewed as highly protective as most theoretical peptides do not represent *bona fide* epitopes. The use of an 8 amino acid sequence length represents a compromise to identify most potentially cross-reactive single epitopes, while reducing the probability of identifying a large number of irrelevant similarities that are unlikely to present a risk of cross-reactivity.



This report describes the bioinformatics assessment of the Cry3Bb1 protein produced by corn event MON 863 and demonstrates the absence of sequence similarity to proteins known to pose human health risks. Potential toxicity of the Cry3Bb1 protein was assessed by comparing the amino acid sequence to proteins associated with toxicity or other health effects. A database of protein sequences associated with toxicity was assembled from publicly available sequence databases (GenBank, EMBL, PIR and SwissProt). In addition, the Cry3Bb1 protein sequence was compared to all protein sequences resident in publicly available genetic databases to screen for structural similarity to other known proteins, including pharmacologically active proteins.

### 3.0 Purpose

The purpose of this study was to evaluate the amino acid sequence similarity between the corn event MON 863 Cry3Bb1 protein and proteins relevant to allergy, celiac disease, toxicity, or other pharmacological effect that had been assembled into sub-databases of publicly available sequences. Structural similarities to allergens, toxins and other proteins were assessed using the sequence alignment tool, FASTA. Immunologically relevant similarities were also assessed using a pair-wise comparison algorithm.

### 4.0 Methods

All analyses were performed using the UNIX based Genetics Computer Group (GCG) software package (version 10.0, Madison, WI) on a standard personal computer using the Windows 95 operating system and supported with Reflection X Client Manager network software (version 7.20, WRQ, Inc. Seattle, WA). Database construction was performed in the SeqLab environment of GCG. Searches were performed from an X terminal window.

#### 4.1 Sequence database preparation

A preliminary list of sequences was compiled using STRINGSEARCH (keyword = allergen). Non-allergen entries were identified by retrieving each flatfile and excluding the irrelevant entries from the list. The resulting list was compared to previously compiled allergen and gliadin databases (Hileman and Astwood, 1999a; Metcalfe *et al.*, 1996). Sequences present in the previous databases and missing in the STRINGSEARCH list were added to the list. Duplicate sequences were identified by performing a FASTA search on each individual sequence against the list and removed. A sequence was considered unique if a single amino acid difference existed when compared to the remainder of the list. Thus, the database may contain multiple isoforms of the same allergen (*e.g.* there are 36 entries for Bet v 1). The list was finalized by adding additional allergen sequences identified by: i) comparison of the list to allergens in the

publicly available list located on the Internet site, <ftp://biobase.dk/resources/pub/who-iuis/allergen.list>; and ii) performing an extensive search of the current literature using the publicly available Internet PubMed and Entrez information retrieval systems (Schuler *et al.*, 1996). For publications retrieved in which only an N-terminal sequence was reported, the sequence was added if 8 or more residues were provided. Newly identified allergens were assembled from public domain databases including Genbank and EMBL release 119 (Benson *et al.*, 1997; Stoesser *et al.*, 1997), PIR release 56 (George *et al.*, 1997), the NRL3D release 56 of RCSB PDB (Berman *et al.*, 2000; Bernstein *et al.*, 1977) and SwissProt version 38 (Bairoch and Apweiler, 1997). DATASET was used to compile the allergen and gliadin database (ALLERGEN3, shown in Appendix 1) that contains 659 separate protein sequences.

The toxin (TOXIN4) sequence database was assembled from public domain databases Genbank and EMBL release 108 (Benson *et al.*, 1997; Stoesser *et al.*, 1997), PIR release 56 (George *et al.*, 1997), the NRL3D (release 56) of RCSB PDB (Berman *et al.*, 2000; Bernstein *et al.*, 1977) and SwissProt release 36 (Bairoch and Apweiler, 1997) as previously described (Hileman and Astwood, 1999b). The ALLPEPTIDES sequence database was used to represent all currently known publicly available protein sequences and consisted of SwissProt release 38 and GenBank release 116.

#### 4.2 Sequence database searches

DNA sequence obtained from corn event MON 863 (Cavato and Lirette, 2001; Hileman and Astwood, 2001) was translated using the standard genetic code. The structural similarity of this deduced Cry3Bb1 protein sequence to sequences in each database (ALLPEPTIDES, TOXIN4 and ALLERGEN3) was assessed using the FASTA algorithm (Lipman and Pearson, 1985; Pearson and Lipman, 1988). Although it may be redundant to search both the TOXIN4 and ALLPEPTIDES databases for potential similarity to protein toxins, the ALLPEPTIDES database search was used to assess for potential similarity to other pharmacologically active proteins, such as prions, which may not have been annotated with the keyword "toxin".

The FASTA sequence alignment tool was used to assess structural similarity. Specific parameters used for these analyses included an expectation threshold of 10, a wordsize (*k-tuple*) of 2, a gap creation penalty of 12 and gap extension penalty of 2. The BLOSUM50 scoring matrix was used. The BLOSUM matrix series (Henikoff and Henikoff, 1992) was derived from a set of aligned, ungapped regions from protein families called the BLOCKS database. Sequences from each block were clustered. Two sequences were put into the same cluster whenever the percent of identical residues exceeded the specified level (Henikoff and Henikoff, 1996). Thus, the BLOSUM50 matrix will identify blocks of conserved residues that are at least 50% identical. BLOSUM50 works well for

sequence alignments with gaps and it has become a standard matrix for recognizing evolutionary distant relationships (Pearson, 2000).

The extent of similarity was evaluated by visual inspection of the alignment, the calculated percent identity and *E* score value. The *E* score (expectation score) reflects the degree of similarity and the value depends on the overall length of joined (gapped) local sequence alignments, the quality (percent identity, similarity) of the overlap and the size of the database. A larger *E* score value indicates a lower degree of similarity between the query sequence and the sequence from the database. For polypeptides, a sequence similarity may indicate sequence homology (*i.e.*, representing a sequence derived from a common ancestor gene with potentially homologous function). Sequences that shared extensive amino acid sequence identity and/or similarity throughout the entire alignment or identified domains were considered to be biologically relevant homologues. A similar amino acid was defined as a non-identical but physicochemically related amino acid. Similar amino acids are structurally related and share polar, hydrophobic, or charged states. Such substitutions are referred to as "conservative" since they are unlikely to change the structures, and by inference the functions, of homologous proteins.

In addition to the FASTA comparisons of each predicted peptide to allergens (to assess overall structural similarity), an 8-mer search was performed an algorithm called IDENTITYSEARCH (also referred to as ALLERGENSEARCH) was used to identify matches of 8 linearly contiguous amino acid identities between the query sequence and sequences within the allergen and gliadin database (ALLERGEN3). The algorithm was run from a UNIX terminal window in GCG. This program compares the query sequence to each protein sequence in the allergen database using a sliding window of 8 amino acids, with a 7 amino acid overlap relative to the preceding window. An epitope of 8 amino acids was chosen to represent the smallest typical allergenic epitope (Metcalf *et al.*, 1996).

Proteins structurally unrelated to allergens may still contain an immunologically relevant epitope (a short local overlap). IDENTITYSEARCH may be more accurate than FASTA for identification of immunologically relevant epitopes because it only requires that the specified window size (8 amino acid residues) match. Matches (if found) are printed in the file output. While the FASTA algorithm is a powerful tool for identification of sequence similarity, it is more appropriately used to assess structural similarity. FASTA comparisons are initiated by aligning a matched wordsize (*k*-tuple parameter) and the alignment extended based on the chosen scoring matrix. The final computed alignment is by default the highest scoring alignment. Thus, a short local overlap may be ignored for a higher scoring alignment containing a longer overlap but relatively few linearly contiguous amino acid identities.

## 5.0 Results and Discussion

Bioinformatic analyses were performed on the Cry3Bb1 protein produced by corn event MON 863 to assess the potential similarity to allergens, toxins or other pharmacologically active proteins. DNA sequence obtained from the *cry3Bb1* coding region of corn event MON 863 (Cavato and Lirette, 2001; Hileman and Astwood, 2001) was translated using the standard genetic code. The resulting protein sequence (Figure 1) was compared to three separate protein sequence databases. A summary of the best similarity from each analysis is shown in Table 2. Supporting dataset output files for each analysis are shown in the Appendices 2-5.

**Table 2.** Best similarities observed to corn event MON 863 Cry3Bb1.11098 protein sequence from FASTA searches against allergen/gliadin, toxin and the ALLPEPTIDES protein databases.

Database	Appendix	# Hits	Accession #	Description	E score	% Identity	aa Overlap
ALLERGEN3 <sup>a</sup>	2	6	S83343	Cyn d 1	3.8	23.6	55
TOXIN4 <sup>b</sup>	4	169	Q06117	Cry3Bb1	$2.7 \times 10^{-228}$	99.1	651
ALLPEPTIDES <sup>c</sup>	5	267	Q06117	Cry3Bb1	$5.7 \times 10^{-259}$	99.1	651

<sup>a</sup> ALLERGEN3 database corresponds to the allergen and gliadin protein database (Appendix 1).

<sup>b</sup> TOXIN4 database corresponds to the toxin protein database.

<sup>c</sup> The ALLPEPTIDES protein database was comprised of SwissProt release 38 and GenBank release 116.

### 5.1 Allergenic assessment of the Cry3Bb1 protein produced by corn event MON 863

Structural similarities shared between the corn event MON 863 Cry3Bb1.11098 protein and proteins in the allergen and gliadin database were evaluated using the FASTA sequence alignment tool. Identified proteins were ranked according to their degree of similarity (Appendix 2). The best similarity observed (Table 2) was to the Bermuda grass pollen allergen, Cyn d 1 (Accession No. S83343, *E* score value of 3.8). In this alignment, the overlap of 55 aa contained 2 gaps and was relatively short compared to the length of the allergen (>246 aa<sup>1</sup>), suggesting that the corn event MON 863 Cry3Bb1.11098 protein does not share homologous structure with Cyn d 1 (Doolittle, 1990). Frequently, alignments comprised of  $\geq 50\%$  identities in short overlaps (20-40 aa) occur by chance and do not indicate homology (Pearson, 1996). Further, the minimum similarity required for allergenic cross-reactivity is  $\geq 50\%$  identity across the entire length of the protein (Aalberse, 2000). Inspection of the remaining alignments (Appendix 2) also yielded poor *E* score values and did not suggest homologous structure or function (Pearson, 1996).

<sup>1</sup> The sequence of Cyn d 1 (Accession No. S83343) corresponds to a fragment of the complete protein.



No immunologically relevant sequences (8 contiguous amino acid identities) were detected when the Cry3Bb1 protein sequence was compared to the ALLERGEN3 sequence database (Appendix 3). Combined, these data demonstrate that the Cry3Bb1 protein produced in corn event MON 863 does not share structurally relevant or immunologically relevant sequence similarities with allergens or gliadins.

## 5.2 Toxicity assessment of the Cry3Bb1 protein produced by corn event MON 863

Structural similarity between the Cry3Bb1.11098 protein produced in corn event MON 863 and toxins were evaluated using the FASTA sequence alignment tool. Similarities identified were ranked according to their degree of similarity (Appendix 4). The best similarity observed (Table 2) was to the *Bacillus thuringiensis* Cry3Bb1 protein (Accession No. Q06117, *E* score value of  $2.7 \times 10^{-228}$ ). The Cry3Bb1 protein produced in corn event MON 863 was designed by modification of the wild type Cry3Bb1 gene. Thus a high degree of sequence similarity, indeed homology, was expected. Inspection of the other sequence alignments between the TOXIN4 database and the Cry3Bb1 protein revealed that 167 of the 169 entries contained in Appendix 4 were structurally related sequences from the Cry family of proteins. The homology shared between members of the Cry protein family has been previously described (Bravo, 1997; Crickmore *et al.*, 1998; Hileman and Astwood, 1999c). The remaining 2 entries corresponded to irrelevant entries. The first was to a hypothetical diphthamide synthesis protein encoded by a *Caenorhabditis elegans* cosmid clone (Accession No. Z37139, *E* score value of 8.9). This cosmid clone contained several hypothetical proteins, one of which contained the keyword toxin. The corresponding observed similarity, however, was not a toxin. The second entry was to a *Vibrio cholera* toxin transport protein (Accession No. P45773, *E* score value of 9.2). Again, this entry did not correspond to a toxin, but rather contained the keyword toxin in the annotation section. Regardless, both of these entries displayed high *E* score values that did not suggest a structurally relevant sequence similarity (Pearson, 1996).

Apart from expected similarities to other known crystal (Cry) protein found in *Bacillus thuringiensis* and related species, no significant structural similarities were observed. The Cry family represents a diverse set of proteins derived from a common ancestor gene. These results establish that the Cry3Bb1 protein produced in corn event MON 863 is not similar to any toxin relevant to human or animal health.

## 5.3 Similarity of the Cry3Bb1 protein produced by corn event MON 863 to all publicly available proteins

Structural similarity between the Cry3Bb1.11098 protein produced in corn event MON 863 and all publicly available proteins were evaluated using the FASTA sequence



alignment tool. Similarities identified were ranked according to their degree of similarity (Appendix 5). The best similarity observed (Table 2) was to the *Bacillus thuringiensis* Cry3Bb1 protein (Accession No. Q06117, *E* score value of  $5.7 \times 10^{-259}$ ). This result was anticipated, as described in Section 5.2. Appendix 5 contains 267 entries and 266 of these were structurally related sequences from the Cry family of proteins. The poorest scoring entry (*E* score value of 8.3) was not a Cry family protein and corresponded to an uncharacterized protein from *Drosophila melanogaster* (Accession No. Q9VB51). No structurally relevant sequence similarities were observed other than to the Cry family of proteins.

## 6.0 Conclusions

Bioinformatics analyses were performed on the Cry3Bb1.11098 protein produced in corn event MON 863. These analyses were used to assess the similarity between this protein and proteins described as an allergen, gliadin, toxin or any other relevant pharmacologically active protein. No structurally relevant sequence similarities with known allergens or gliadins were observed. No immunologically relevant sequence similarities of 8 linearly contiguous amino acid identities were observed suggesting that it is highly unlikely that this protein may contain immunologically cross-reactive allergenic epitopes. No structurally relevant sequence similarity was observed to proteins suspected to be toxic or pharmacologically active or relevant to human and animal health.

## 7.0 References

- Aalberse, R. C. (2000). Structural biology of allergens. *J Allergy Clin Immunol* **106**:228-238.
- Bairoch, A., and Apweiler, R. (1997). The SWISS-PROT protein sequence data bank and its supplement TrEMBL. *Nucl Acids Res* **25**:31-36.
- Banerjee, B., Greenberger, P. A., Fink, J. N., and Kurup, V. P. (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, D. H., Hickey, V. L., Slater, J. E., and Sussman, G. L. (1999). Human IgE-binding epitopes of the latex allergen Hev b 5. *J Allergy Clin Immunol* **103**:1166-1172.
- Benson, D. A., Boguski, M. S., Lipman, D. J., and Ostell, J. (1997). GenBank. *Nucl Acids Res* **25**:1-6.

- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000). The Protein Data Bank. *Nucl Acids Res* 28:235-242.
- Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977). The Protein Data Bank. A computer-based archival file for macromolecular structures. *Eur J Biochem* 80:319-324.
- Berzofsky, J. A., Berkower, I. J., and Epstein, S. L. (1993). Antigen-antibody interactions and monoclonal antibodies. In *Fundamental Immunology*, W. E. Paul, Ed. (New York, NY: Raven Press, Ltd), Chapter 12, pp. 421-465.
- Bravo, A. (1997). Phylogenetic relationships of *Bacillus thuringiensis* delta-endotoxin family proteins and their functional domains. *J Bacteriol* 179:2793-2801.
- Breitenbach, M., Simon, B., Probst, G., Oberkofler, H., Ferreira, F., Briza, P., Achatz, G., Unger, A., Ebner, C., Kraft, D., and Hirschwehr, R. (1997). Enolases are highly conserved fungal allergens. *Int Arch Allergy Immunol* 113:114-147.
- Cavato, T. A., and Lirette, R. P. (2001). PCR analysis and DNA sequence of the insert in corn rootworm event MON 863. (Monsanto Company, St. Louis, MO) Study 01-01-39-02, Technical Report MSL-17108.
- Chatchatee, P., Jarvinen, K. M., Bardina, L., Beyer, K., and Sampson, H. A. (2001). Identification of IgE- and IgG-binding epitopes on alpha(s1)-casein: differences in patients with persistent and transient cow's milk allergy. *J Allergy Clin Immunol* 107:379-383.
- Crickmore, N., Zeigler, D. R., Feitelson, J., Schnepf, E., Van Rie, J., Lereclus, D., Baum, J., and Dean, D. H. (1998). Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal crystal proteins. *Microbiol Mol Biol Rev* 62:807-813.
- Daul, C. B., Morgan, J. E., and Lehrer, S. B. (1993). Hypersensitivity reactions to crustacea and mollusks. *Clin Rev Allergy* 11:201-222.
- Donovan, W. P., Rupar, M. J., Slaney, A. C., Malvar, T., Gawron-Burke, M. C., and Johnson, T. B. (1992). Characterization of two genes encoding *Bacillus thuringiensis* insecticidal crystal proteins toxic to *Coleoptera* species. *Appl Environ Microbiol* 58:3921-3927.

Doolittle, R. F. (1990). Searching through sequence databases. *Methods Enzymol* **183**:99-110.

English, L. H., Brussock, S. M., Malvar, T. M., Bryson, J. W., Kulesza, C. A., Walters, F. S., Slatin, S. L., Von Tersch, M. A., and Romano, C. (2000). Nucleic acid segments encoding modified *Bacillus thuringiensis* coleopteran-toxic crystal proteins (United States Patent # 6,060,594: Ecogen, Inc., Langhorn, PA; Monsanto Company, St. Louis, MO).

Ferguson, A. (1996). Gluten-sensitive enteropathy (celiac disease). In Food allergy: Adverse reactions to foods and food additives, D. D. Metcalfe, H. A. Sampson, and R. A. Simon, Eds. (Cambridge, MA: Blackwell Science, Inc), Chapter 19, pp. 287-301.

George, D. G., Dodson, R. J., Garavelli, J. S., Haft, D. H., Hunt, L. T., Marzec, C. R., Orcutt, B. C., Sidman, K. E., Srinivasarao, G. Y., Yeh, L. S. L., Arminski, L. M., Ledley, R. S., Tsugita, A., and Barker, W. C. (1997). The Protein Information Resource (PIR) and the PIR-International Protein Sequence Database. *Nucl Acids Res* **25**:24-28.

Gupta, N., Martin, B. M., Metcalfe, D. D., and Rao, P. V. (1996). Identification of a novel hydroxyproline-rich glycoprotein as the major allergen in *Parthenium* pollen. *J Allergy Clin Immunol* **98**:903-912.

Henikoff, J. G., and Henikoff, S. (1996). Blocks database and its applications. *Methods Enzymol* **266**:88-105.

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

Hileman, R. E., and Astwood, J. D. (1999a). Bioinformatics analysis of CP4 EPSPS protein sequence utilizing an allergen database. (Monsanto Company, St. Louis, MO) Study 99-01-62-07, Technical Report MSL-16267.

Hileman, R. E., and Astwood, J. D. (1999b). Bioinformatics analysis of CP4 EPSPS protein sequence utilizing toxin and public domain genetic databases. (Monsanto Company, St. Louis, MO) Study 99-01-62-08, Technical Report MSL-16268.

Hileman, R. E., and Astwood, J. D. (1999c). Bioinformatics sequence analysis of *B.t.* protein 11098 and *B.t.* protein 11231 utilizing toxin and public domain genetic databases. (Monsanto Company, St. Louis, MO) Technical Report MSL-15870.

- Hileman, R. E., and Astwood, J. D. (2001). Additional characterization of the Cry3Bb1 protein produced in corn event MON 863. (Monsanto Company, St. Louis, MO) Technical Report MSL-17137.
- Ishikawa, M., Ishida, M., Shimakura, K., Nagashima, Y., and Shiomi, K. (1998). Purification and IgE-binding epitopes of a major allergen in the gastropod *Turbo cornutus*. *Biosci Biotechnol Biochem* **62**:1337-1343.
- Jensen-Jarolim, E., Santner, B., Leitner, A., Grimm, R., Scheiner, O., Ebner, C., and Breiteneder, H. (1998). Bell peppers (*Capsicum annuum*) express allergens (profilin, pathogenesis-related protein P23 and Bet v 1) depending on the horticultural strain. *Int Arch Allergy Immunol* **116**:103-109.
- Klysner, S., Welinder, K. G., Lowenstein, H., and Mathiesen, F. (1992). Group V allergens in grass pollens: IV. Similarities in amino acid compositions and NH<sub>2</sub>-terminal sequences of the group V allergens from *Lolium perenne*, *Poa pratensis* and *Dactylis glomerata*. *Clin Exp Allergy* **22**:491-497.
- Leitner, A., Jensen-Jarolim, E., Grimm, R., Wüthrich, B., Ebner, H., Scheiner, O., Kraft, D., and Ebner, C. (1998). Allergens in pepper and paprika: Immunologic investigation of the celery-birch-mugwort-spice syndrome. *Allergy* **53**:36-41.
- Lipman, D. J., and Pearson, W. R. (1985). Rapid and sensitive protein similarity searches. *Science* **227**:1435-1441.
- Metcalf, D. D., Astwood, J. D., Townsend, R., Sampson, H. A., Taylor, S. L., and Fuchs, R. L. (1996). Assessment of the allergenic potential of foods derived from genetically engineered crop plants. *Crit Rev Food Sci Nutr* **36**:S165-186.
- Monsalve, R. I., Gonzalez de la Pena, M. A., Lopez-Otin, C., Fiandor, A., Fernandez, C., Villalba, M., and Rodriguez, R. (1997). Detection, isolation and complete amino acid sequence of an aeroallergenic protein from rapeseed flour. *Clin Exp Allergy* **27**:833-841.
- Pearson, W. R. (1996). Effective protein sequence comparison. *Methods Enzymol* **266**:227-258.
- Pearson, W. R. (2000). Flexible sequence similarity searching with the FASTA3 program package. *Methods Mol Biol* **132**:185-219.



- Pearson, W. R., and Lipman, D. J. (1988). Improved tools for biological sequence comparison. *Proc Natl Acad Sci USA* **85**:2440-2448.
- Plaut, M., and Zimmerman, E. M. (1993). Allergy and mechanisms of hypersensitivity. In *Fundamental Immunology*, W. E. Paul, Ed. (New York, NY: Raven Press, Ltd), Chapter 40, pp. 1399-1425.
- Rabjohn, P., Helm, E. M., Stanley, J. S., West, C. M., Sampson, H. A., Burks, A. W., and Bannon, G. A. (1999). Molecular cloning and epitope analysis of the peanut allergen Ara h 3. *J Clin Invest* **103**:535-542.
- Reese, G., Ayuso, R., and Lehrer, S. B. (1999). Tropomyosin: an invertebrate pan-allergen. *Int Arch Allergy Immunol* **119**:247-258.
- Schuler, G. D., Epstein, J. A., Ohkawa, H., and Kans, J. A. (1996). Entrez: Molecular biology database and retrieval system. *Methods Enzymol* **266**:141-162.
- Shen, H., Lin, W., Tam, M. F., Wang, S., Tzean, S., Huang, M., and Han, S. (1999). Characterization of allergens from *Penicillium oxalicum* and *P. notatum* by immunoblotting and N-terminal amino acid sequence analysis. *Clin Exp Allergy* **29**:642-651.
- Shin, D. S., Compadre, C. M., Maleki, S. J., Kopper, R. A., Sampson, H., Huang, S. K., Burks, A. W., and Bannon, G. A. (1998). Biochemical and structural analysis of the IgE binding sites on Ara h1, an abundant and highly allergenic peanut protein. *J Biol Chem* **273**:13753-13759.
- Stoesser, G., Sterk, P., Tuli, M. A., Stoehr, P. J., and Cameron, G. N. (1997). The EMBL nucleotide sequence database. *Nucl Acids Res* **25**:7-13.



1 MANPNNRSEH DTIKVTPNSE LQTNHNQYPL ADNPNSTLBE LNYKEFLRMT  
51 EDSSTEVLND STVKDAVGTG ISVVGQILGV VGVPFAGALT SFYQSFLNTI  
101 WPSDADPWKA FMAQVEVLID KKIEEYAKSK ALAELOGLON NFEDYVNALN  
151 SWKKTPLSLR SKRSQGRIRE LFSQAESHFR NSMPSEAVSK FEVLFLPTYA  
201 QAANTHLLLL KDAQVFGEW GYSSDVAEF YRQLKLTOO YTDHCVNWYN  
251 VGLNGLRGST YDAWVKFNRF RREMTLTVLD LIVLFPFYDI RLYSKGVKTE  
301 LTRDIFTDPI FLLTTLQKYG PTFLSIENSI RKPFLFDYLO GIEFHTRLRP  
351 GYFGKDSFNY WSGNYVETRP SIGSSKTITS PFYGDKSTEP VQKLSFDGQK  
401 VYRTIANTDV AAWPNGKVYL GVTKVDFSQY DDQKNETSTQ TYDSKRNNGH  
451 VSAQDSIDQL PPETTDEPLE KAYSHQLNYA ECFLMQDRRG TIPFFTWTNR  
501 SVDFNTIDA EKITQLPVVK AYALSSGASI IEGPGFTGGN LLFLKESSNS  
551 IAKFKVTLNS AALLQRYRVR IRYASTTNLR LFVQNSNNDP LVIYINKTMN  
601 KDDDLTYQTF DLATNSNMG FSGDKNELII GAESFVSNEK IYIDKIEFIP  
651 VQL

**Figure 1.** The deduced sequence of the Cry3Bb1.11098 protein produced in corn event MON 863.

# Appendix 1. Allergen and gliadin sequence database (ALLERGEN3).

## Food Allergens - Plants

Species	Common name	Allergen	Comment	Acc. No.	aa length
Actinidia chinensis	Kiwi	Act c 1	Actinidin, cysteine protease	P00785	380
Apium graveolens	Celery	Api g 1	Similar to 17 kDa tree allergens	S63984	154
Apium graveolens	Celery	Api g 1.0201		CAA99992	108
Apium graveolens	Celery	Api g 2		Z75662	159
Apium graveolens	Celery	Api g 3	Chlorophyll a/b binding protein	CAA99993	177
Apium graveolens	Celery	Api g 4	Profilin	AF129423	134
Apium graveolens	Celery	Api g 5		P81943	18
Arachis hypogaeae	Peanut	Ara h 1	Vicilin, clone P17	P43237	614
Arachis hypogaeae	Peanut	Ara h 1	Vicilin, clone P41b	P43238	626
Arachis hypogaeae	Peanut	Ara h 2	Conglutin	L77197	157
Arachis hypogaeae	Peanut	Ara h 3	Glycinin	AF093541	507
Arachis hypogaeae	Peanut	Ara h 4	Glycinin	AAD47382	530
Arachis hypogaeae	Peanut	Ara h 5	Profilin	AAD55587	131
Arachis hypogaeae	Peanut	Ara h 6	Similar to conglutin	AAD56337	129
Arachis hypogaeae	Peanut	Ara h 7	Conglutin homolog	AAD56719	160
Arachis hypogaeae	Peanut		Lectin, phytohemagglutinin	S14765	273
Bertholletia excelsa	Brazil nut	Ber e 1	2S albumin, Be 2S1	X54490	146
Brassica juncea	Leaf mustard	Bra j 1-e	2S albumin	P80207	129
Brassica napis	Rapeseed	BnIII	2S albumin	(Monsalve <i>et al.</i> , 1997)	79
Brassica napis	Rapeseed	BnIII	2S albumin	(Monsalve <i>et al.</i> , 1997)	31
Capsicum annuum	Bell pepper		Profilin, pathogenesis-related protein P23	(Jensen-Jarolim <i>et al.</i> , 1998)	10
Carica papaya	Papaya		Papain	M15203	345
Corylus avellana	European hazelnut	Cor a 1.0401	Bet v 1 homolog	AF136945	161
Daucus carota	Carrot	Dau c 1	Pathogenesis-related protein GEA20	O04298	106
Daucus carota	Carrot	Dau c 1/1		Z84376	154
Daucus carota	Carrot	Dau c 1.2		Z81361	154
Daucus carota	Carrot	Dau c 1.3		Z81362	154
Glycine max	Soybean	Cim1 protein		S48032	277
Glycine max	Soybean	Gly m 1.0101	Soybean hull allergen, HPS	AAB34756	42
Glycine max	Soybean	Gly m 1.0102	Soybean hull allergen, HPS	AAB34756	39
Glycine max	Soybean	Gly m 2	Hull allergen	A57106	20
Glycine max	Soybean	Gly m 3	Profilin	O65809	131
Glycine max	Soybean		Lipoxygenase 1	S25064	839
Glycine max	Soybean		Lipoxygenase 2	A28161	865
Glycine max	Soybean		Alpha of beta-conglycinin	X17698	605
Glycine max	Soybean		A1aBx subunit of glycinin	X02985	496
Glycine max	Soybean		A5A4B3 subunit of glycinin	X02626	562
Glycine max	Soybean		G1 subunit of glycinin	X15121	496
Glycine max	Soybean		G2 subunit of glycinin	X15122	485
Glycine max	Soybean		G3 subunit of glycinin	X15123	484

Glycine max	Soybean		CG4 beta-conglycinin	S44893	439
Glycine max	Soybean		A3B4 subunit of glycinin	M10962	516
Glycine max	Soybean		Lectin Le1	K00821	285
Glycine max	Soybean		Kunitz trypsin inhibitor, Kti-S	X80039	208
Glycine max	Soybean		Trypsin inhibitor, Ti-a	X64447	217
Glycine max	Soybean		Trypsin inhibitor, Ti-b	X64448	217
Helianthus annuus	Sunflower	Hel a 1	Profilin	O81982	133
Helianthus annuus	Sunflower	Hel a 2	Profilin	Y15210	133
Hordeum vulgare	Barley	Hor v 15	Trypsin/alpha amylase inhibitor	P16968	146
Hordeum vulgare	Barley	Hor v 15	Trypsin/alpha amylase inhibitor	P32936	149
Hordeum vulgare	Barley		Dimeric protein BDP	S65605	20
Juglans regia	English walnut	Jug r 1	2S albumin	U66866	139
Juglans regia	English walnut	Jug r 2	Vicilin-like protein	AAF18269	593
Lycopersicon esculentum	tomato	LAT52	Ole e 1 homolog	P13447	161
Malus domestica	Apple	Mal d 1		O22517	159
Malus domestica	Apple	Mal d 1		JC4276	159
Malus domestica	Apple	Mal d 1	Variant	AF074721	159
Malus domestica	Apple	Mal d 1	Clone MDGC10	Z72425	159
Malus domestica	Apple	Mal d 1	Clone MDGC32	Z72426	160
Malus domestica	Apple	Mal d 1	Clone MDGC36	Z72427	160
Malus domestica	Apple	Mal d 1	Clone MDGC38	Z72428	160
Malus domestica	Apple	Mal d 1		X83672	159
Malus domestica	Apple	Mal d 2	Thaumatococcus homolog	AAC36740	245
Malus domestica	Apple	Mal d 3	Lipid transfer protein	AAF26450	115
Oryza sativa	Rice	Ory s 1	Similar to Lol p 1	Q40638	263
Oryza sativa	Rice	RAG1	Trypsin/alpha amylase inhibitor	Q01884	95
Oryza sativa	Rice	RAG2	Trypsin/alpha amylase inhibitor	Q01885	166
Oryza sativa	Rice	RA5	Trypsin/alpha amylase inhibitor	S31078	157
Oryza sativa	Rice	RA5b	Trypsin/alpha amylase inhibitor	Q40655	160
Oryza sativa	Rice	RA14		Q01882	165
Oryza sativa	Rice	RA14B	Trypsin/alpha amylase inhibitor	S59922	166
Oryza sativa	Rice	RA14d		D43657	111
Oryza sativa	Rice	RA14e		D43658	109
Oryza sativa	Rice	RA14f		D43659	113
Oryza sativa	Rice	RA16	Trypsin/alpha amylase inhibitor	S59924	157
Oryza sativa	Rice	RA17	Trypsin/alpha amylase inhibitor	S21157	162
Oryza sativa	Rice		Trypsin/alpha amylase inhibitor	Q40653	160
Oryza sativa	Rice			AF042200	157
Oryza sativa	Rice		Beta-expansin 2	AAF72991	192
Oryza sativa	Rice		Beta-expansin 1	AAF72990	192
Oryza sativa	Rice		Beta-expansin 3	AAF72983	194
Persea americana	Avocado	Per a 1	Endochitinase	P93680	326
Phaseolus vulgaris	Kidney bean	PR-1	Pathogenesis-related protein 1	S11929	156
Phaseolus vulgaris	Kidney bean	PR-2	Pathogenesis-related protein 2	S11930	155
Piper nigrum	Pepper	14 kDa		(Lefter et al., 1998)	27

Piper nigrum	Pepper	28 kDa		(Leitner <i>et al.</i> , 1998)	23
Pisum sativum	Garden pea		Pollen allergen-like protein	Q41043	258
Prunus armeniaca	Apricot	Pru ar 1	Bet v 1 homolog	U93165	160
Prunus armeniaca	Apricot	Pru ar 3	Nonspecific lipid transfer protein	P81651	91
Prunus avium	Cherry	Pru av 1	Bet v 1 homolog	U66076	160
Prunus avium	Cherry	Pru av 2	Thaumatococcus homolog	U32440	245
Prunus avium	Cherry	Pru av 4	Profilin	AF129425	131
Prunus domestica	Plum	Pru d 1	Lipid transfer protein 1	P82534	72
Prunus persica	Peach	Pru p 3	Lipid transfer protein 1	P81402	91
Ricinus communis	Castor bean	Ric c 1	2S albumin	P01089	258
Secale cereale	Rye	30 kDa allergen	Fragment	S38292	16
Secale cereale	Rye	Sec c 1	13.5 kDa asthma allergen, fragment	S65604	26
Sinapis alba	Yellow mustard	Sin a 1	Seed storage protein	P15322	127
Sinapis alba	White mustard	Sin a 1		PC1246	145
Sinapis alba	Yellow mustard	Sin a 1		S54101	145
Sinapis alba	White mustard	Sin a 1.0104		Q41277	145
Sinapis alba	White mustard	Sin a 1.0105		Q41278	145
Sinapis alba	White mustard	Sin a 1.0106		Q41279	145
Sinapis alba	White mustard	Sin a 1.0107		Q41280	145
Sinapis alba	White mustard	Sin a 1.0408		Q41281	145
Sinapis alba	White mustard	Sin a 1	Alpha-amylase inhibitor	PC1247	145
Solanum tuberosum	Potato	Sol t 1	Potatin, storage protein	P15476	377
Taraxacum officinale	Dandelion	RAP	Root allergen protein	AF036931	157
Triticum aestivum	Wheat	Tri a 3	Pollen allergen-like protein	Z50867	118
Triticum aestivum	Wheat		Pollen allergen-like protein	U91981	271
Zea mays	Maize	Zea m 1	Similar to Lol p 1	Q07154	191
Zea mays	Maize	Clone c13	Ole e 1 homolog	P33050	170
Zea mays	Maize		Pectate lyase, Clone Zm58.1	S43334	104
Zea mays	Maize		Pectate lyase	S43335	438
Zea mays	Maize	Zea m 14	Nonspecific lipid transfer protein	P19656	120

#### Giladins (wheat)

Triticum aestivum	Wheat		Alpha/beta gliadin MM1	X17361	307
Triticum aestivum	Wheat		Gliadin	X00627	286
Triticum aestivum	Wheat		Alpha/beta gliadin (pw1215)	X02538	296
Triticum aestivum	Wheat		Alpha/beta gliadin (pw8233)	X02569	286
Triticum aestivum	Wheat		Alpha/beta gliadin	X02540	313
Triticum aestivum	Wheat		Alpha gliadin	U08287	287
Triticum aestivum	Wheat		Gamma gliadin	M36999	302
Triticum aestivum	Wheat		Gamma gliadin B-precursor	M13713	291
Triticum aestivum	Wheat	pGliA-42	Alpha gliadin	K02068	318
Triticum aestivum	Wheat	Clone pA212	Alpha/beta gliadin class A-II	M10092	291
Triticum aestivum	Wheat	Clone pA42	Alpha/beta class A-V	M11073	319
Triticum aestivum	Wheat	pW8142	Alpha/beta gliadin	K03075	313
Triticum aestivum	Wheat	pW8233	Alpha/beta gliadin, class I	K03076	286

Appendix 1

Study No. 01-01-39-23

MSL-17140

Page 26 of 102

Triticum aestivum	Wheat	Clone pA26	Alpha/beta gliadin, class A-I	M11074	262
Triticum aestivum	Wheat	Clone pA735	Alpha/beta gliadin, class A-IV	M11075	297
Triticum aestivum	Wheat	pA1235	Alpha/beta gliadin, class A-III	M11076	282
Triticum aestivum	Wheat	pB11-33	Gamma gliadin, class B-I	M11077	304
Triticum aestivum	Wheat	pB312	Gamma gliadin, class B-I	M11336	323
Triticum aestivum	Wheat		Gamma gliadin, class B-III	M11335	244
Triticum aestivum	Wheat		Gamma gliadin, variant	M16064	327
Triticum urartu	Wheat		Alpha/beta gliadin	M16496	296

#### Lectins (wheat)

Triticum aestivum	Wheat		Germ agglutinin isolectin A	M25536	212
Triticum aestivum	Wheat		Germ agglutinin isolectin D	M25537	213
Triticum durum	Wheat		Germ agglutinin (partial)	J02961	186

#### Profilins (wheat)

Triticum aestivum	Wheat	Clone TaPRO1	Profilin 1	X89825	138
Triticum aestivum	Wheat	Clone TaPRO2	Profilin 2	X89826	141
Triticum aestivum	Wheat	Clone TaPRO3	Profilin 3	X89827	140

#### Food Allergens - Animals

Bos taurus	Bovine	Bos d 4	Alpha-lactalbumin	J05147	142
Bos taurus	Bovine	Bos d 5	Beta-lactoglobulin	X14712	178
Bos taurus	Bovine	Bos d 6	BSA	M73993	607
Bos taurus	Bovine	Bos d 8	Alpha-s1 casein	M33123	214
Bos taurus	Bovine	Bos d 8	Alpha-s1 casein	M38641	214
Bos taurus	Bovine	Bos d 8	Alpha-s2 casein-like	M16644	222
Bos taurus	Bovine	Bos d 8	Beta casein	AAA30430	224
Bos taurus	Bovine	Bos d 8	Kappa casein, B2 variant	M36641	190
Charybdis furiatus	Crab	Cha f 1	Tropomyosin	AAF35431	264
Chlamys nobilis	Clam	Chl n ?	Tropomyosin	AAG08989	173
Crassostrea virginica	Eastern oyster		Tropomyosin	AAG61869	160
Gadus callarias	Baltic cod	Gad c 1	Parvalbumin beta, allergen M	P02622	113
Gallus gallus	Chicken	Gal d 1	Ovomucoid	P01005	210
Gallus gallus	Chicken	Gal d 2	Ovalbumin	P01012	385
Gallus gallus	Chicken	Gal d 3	Ovotransferrin	P02789	705
Gallus gallus	Chicken	Gal d 4	Lysozyme C	P00898	147
Gallus gallus	Chicken	Gal d 5	Albumin	P19121	615
Gallus gallus	Chicken		Ovalbumin Y gene	J00922	388
Gallus gallus	Chicken		Vitellogenin II, Phosvitin	A92941	1850
Gallus gallus	Chicken		Apovitellenin I	A93464	106
Haliotis diversicolor	Mollusk	Hal d 1	Tropomyosin	AAG08987	168
Metapenaeus ensis	Shrimp	Met e 1	Tropomyosin	U08008	274
Panulirus stimpsoni	Lobster	Pan s 1	Tropomyosin	AF030063	274
Penaeus aztecus	Shrimp	Pen a 1	Tropomyosin	(Davi et al., 1993)	21



Perna viridis	Mussel	Per v 1	Tropomyosin	AAG08988	174
Salmo salar	Atlantic salmon	Sal s 1	Parvalbumin beta 1	Q91482	109
Salmo salar	Atlantic salmon	Sal s 1	Parvalbumin beta 2	Q91483	108
Turbo cornutus	Gastropod	Tur c 1	Tropomyosin, peptide K20	(Ishikawa <i>et al.</i> , 1998)	21
Turbo cornutus	Gastropod	Tur c 1	Tropomyosin, peptide K24	(Ishikawa <i>et al.</i> , 1998)	27
Turbo cornutus	Gastropod	Tur c 1	Tropomyosin, peptide K19	(Ishikawa <i>et al.</i> , 1998)	21
Turbo cornutus	Gastropod	Tur c 1	Tropomyosin, peptide K21	(Ishikawa <i>et al.</i> , 1998)	29
Turbo cornutus	Gastropod	Tur c 1	Tropomyosin, peptides MT17, MT15, K23	(Ishikawa <i>et al.</i> , 1998)	40

#### Aero-Allergens - Pollen

Agrostis alba	Bent grass	Agr a 1	Form 1, fragment	F58493	35
Agrostis alba	Bent grass	Agr a 1	Form 1, fragment	E37396	26
Agrostis alba	Bent grass	Agr a 1	Form 2, fragment	G58493	35
Alnus glutinosa	Alder	Aln g 1		S50892	160
Alnus glutinosa	Alder	Aln g 2		CAA76831	85
Ambrosia artemisiifolia	Short ragweed	Amb a 1.1	Antigen E	P27759	396
Ambrosia artemisiifolia	Short ragweed	Amb a 1.2	Antigen E	P27760	398
Ambrosia artemisiifolia	Short ragweed	Amb a 1.2	Antigen E	B53240	398
Ambrosia artemisiifolia	Short ragweed	Amb a 1.3	Antigen E	P27761	397
Ambrosia artemisiifolia	Short ragweed	Amb a 1.3	Antigen E	C53240	397
Ambrosia artemisiifolia	Short ragweed	Amb a 1.4	Antigen E	P28744	392
Ambrosia artemisiifolia	Short ragweed	Amb a 2	Antigen K	P27762	397
Ambrosia artemisiifolia	Short ragweed	Amb a 2	Antigen K	E53240	397
Ambrosia artemisiifolia	Short ragweed	Amb a 3	Allergen RA3	P00304	101
Ambrosia artemisiifolia	Short ragweed	Amb a 5	Allergen RA5, fragment	P02878	45
Ambrosia artemisiifolia	Short ragweed	Amb a 6	Allergen RA6	O04004	118
Ambrosia trifida	Giant ragweed	Amb t 5	Allergen RA5G	P10414	73
Ambrosia psilostachya	Western ragweed	Amb p 5	Clone B1	L24467	77
Ambrosia psilostachya	Western ragweed	Amb p 5	Clone A2	L24465	77
Ambrosia psilostachya	Western ragweed	Amb p 5	Clone B2	L24468	77
Ambrosia psilostachya	Western ragweed	Amb p 5	Clone A3	L24466	77
Ambrosia psilostachya	Western ragweed	Amb p 5	Clone B3	L24469	77
Anthoxanthum odoratum	Sweet vernal grass	Ant o 1	Fragment	G37396	26
Anthoxanthum odoratum	Sweet vernal grass	Ant o 1	Fragment, form 1	B58493	32
Anthoxanthum odoratum	Sweet vernal grass	Ant o 1	Fragment, form 2	C58493	32
Arabidopsis thaliana	Mouse-ear cress	Ara t 8	Profilin 1	3NUL	130
Arabidopsis thaliana	Mouse-ear cress	ESSA I AP2	Alternaria allergen similarity	Z99708	634
Arabidopsis thaliana	Mouse-ear cress	ESSA I	Fragment 7, Lol p1 similarity	Z97342	101
Arabidopsis thaliana	Mouse-ear cress	BAC F5114	Similar to Holcus major allergen	AC001229	276
Artemisia vulgaris	Mugwort	Art v 2	Fragments	A38624	71
Betula pendula	European White Birch	Bet v 1	Isoform at2	AJ002106	160
Betula pendula	European White Birch	Bet v 1	Isoform at12	AJ002107	160
Betula pendula	European White Birch	Bet v 1	Isoform at15	AJ002108	160
Betula pendula	European White Birch	Bet v 1	Isoform at76	AJ002110	160

Betula pendula	European White Birch	Bet v 1		O23746	159
Betula pendula	European White Birch	Bet v 1		O23747	159
Betula pendula	European White Birch	Bet v 1		O23748	159
Betula pendula	European White Birch	Bet v 1		O23749	159
Betula pendula	European White Birch	Bet v 1		O23750	159
Betula pendula	European White Birch	Bet v 1m/n		A57427	160
Betula pendula	European White Birch	Bet v 1.0103	Clone 2230	Z80099	160
Betula pendula	European White Birch	Bet v 1.0201		A55699	160
Betula pendula	European White Birch	Bet v 1.0301		B55699	160
Betula pendula	European White Birch	Bet v 1.0401		C55699	160
Betula pendula	European White Birch	Bet v 1.0501		D55699	160
Betula pendula	European White Birch	Bet v 1.0601		E55699	160
Betula pendula	European White Birch	Bet v 1.0701		F55699	160
Betula pendula	European White Birch	Bet v 1.0801		G55699	160
Betula pendula	European White Birch	Bet v 1.0901		H55699	160
Betula pendula	European White Birch	Bet v 1.1001		I55699	160
Betula pendula	European White Birch	Bet v 1.1502	Clone BVGC31	Z72432	160
Betula pendula	European White Birch	Bet v 1.1601	Clone BVGC181	Z72437	160
Betula pendula	European White Birch	Bet v 1.1701	Clone BVGC21	Z72430	160
Betula pendula	European White Birch	Bet v 1.1801	Clone BVGC25	Z72431	160
Betula pendula	European White Birch	Bet v 1.1901	Clone BVGC34	Z72433	160
Betula pendula	European White Birch	Bet v 1.2001	Clone BVGC45	Z72434	160
Betula pendula	European White Birch	Bet v 1.2101	Clone BVGC63	Z72435	160
Betula pendula	European White Birch	Bet v 1.2201	Clone BVGC681	Z72438	160
Betula pendula	European White Birch	Bet v 1.2301	Clone BVGC70	Z72436	160
Betula pendula	European White Birch	Bet v 1.2401	Clone 167	Z80100	160
Betula pendula	European White Birch	Bet v 1.2501	Clone 184	Z80101	160
Betula pendula	European White Birch	Bet v 1.2601	Clone 2225	Z80102	160
Betula pendula	European White Birch	Bet v 1.2701	Clone 2226	Z80103	160
Betula pendula	European White Birch	Bet v 1.2801	Clone 2227	Z80104	160
Betula pendula	European White Birch	Bet v 1.2901	Clone 2229	Z80105	160
Betula pendula	European White Birch	Bet v 1.3001	Clone 2301	Z80106	160
Betula pendula	White birch	Bet v 2	Profilin	P25816	133
Betula pendula	European White Birch	Bet v 2	Fragment	B45786	51
Betula pendula	European White Birch	Bet v 3	Calcium binding protein	S46233	205
Betula pendula	European White Birch	Bet v 4	Calcium binding protein	O04131	85
Betula pendula	European White Birch	Bet v 4		X87153	85
Betula pendula	European White Birch	Bet v 6	Isoflavone reductase homolog	A1047896	300
Betula pendula	European White Birch	Bet v 6.0101	Isoflavone reductase homolog	AAC05116	194
Brassica napus	Rape	Bra n 1	Pollen allergen group I	S65149	79
Brassica napus	Rape	Bra n 2	Pollen allergen group II	S65150	82
Brassica napus	Rape	Clone 42	Pollen allergen group II	S65144	83
Brassica napus	Rape	Clone 44	Pollen allergen group II	S65145	83
Brassica rapa	Turnip	Bra r 1	Calcium binding protein	S65151	79
Brassica rapa	Turnip	Bra r 2	Pollen allergen group II	S65152	83

Brassica rapa	Turnip	Clone 4	Pollen allergen group II	S65143	80
Brassica rapa	Turnip		Chitin-binding allergen, fragments	P81729	91
Carpinus betulus	Hornbeam	Car b 1	Isoform 1, Bet v 1 homolog	P38949	159
Carpinus betulus	Hornbeam	Car b 1	Isoform 2, Bet v 1 homolog	P38950	159
Carpinus betulus	Hornbeam	Car b 1.0103	Clone 380	Z80159	160
Carpinus betulus	Hornbeam	Car b 1.0104	Clone 383	Z80160	160
Carpinus betulus	Hornbeam	Car b 1.0105	Clone 372	Z80161	160
Carpinus betulus	Hornbeam	Car b 1.0106	Clone 541	Z80162	160
Carpinus betulus	Hornbeam	Car b 1.0107	Clone 385	Z80166	160
Carpinus betulus	Hornbeam	Car b 1.0108	Clone 295	Z80168	160
Carpinus betulus	Hornbeam	Car b 1.0301	Clone 563	Z80169	161
Carpinus betulus	Hornbeam	Car b 1.0302	Clone 2237	Z80170	161
Castanea sativa	European chestnut	Cas s 1	Bet v 1 homolog, fragment	PC2001	24
Castanea sativa	European chestnut	Cas s 5	Chitinase 1b	CAA64868	316
Catharanthus roseus	Madagascar periwinkle		Bet v 1 homolog	T10059	96
Chamaecyparis obtusa	Japanese cypress	Cha o 1		D45404	375
Chamaecyparis obtusa	Japanese cypress	Cha o 2	Polygalacturonidase	JC7100	357
Corylus avellana	European hazelnut	Cor a 1/1.1		S30055	160
Corylus avellana	European hazelnut	Cor a 1.0101		X70999	160
Corylus avellana	European hazelnut	Cor a 1.0102		X71000	160
Corylus avellana	European hazelnut	Cor a 1.0103		X70997	160
Corylus avellana	European hazelnut	Cor a 1.0104		X70998	160
Corylus avellana	European hazelnut	Cor a 1.0201	Clone CAGC10	Z72439	160
Corylus avellana	European hazelnut	Cor a 1.0301	Clone CAGC11	Z72440	160
Cryptomeria japonica	Japanese cedar	Cry j 1	Clone pCCI-1b	JC2124	374
Cryptomeria japonica	Japanese cedar	Cry j 1	Clone pCCI-2.2	JC2123	374
Cryptomeria japonica	Japanese cedar	Cry j 2	Pectinase	P43212	514
Cryptomeria japonica	Japanese cedar	Cry j 2		JC2498	514
Cucumis sativus	Cucumber	Cs-EXP1	Expansin S1	U30382	250
Cucumis sativus	Cucumber	Cs-EXP2	Expansin S2	U30460	258
Cupressus arizonica	Conifer	Cup a 1	Clone LUZ99.ARI4.1	GAB62551	264
Cupressus sempervirens	Conifer	Cup s 1	Clone 23	AAF72629	280
Cupressus sempervirens	Conifer	Cup s 1	Clone 18	AAF72628	280
Cupressus sempervirens	Conifer	Cup s 1	Clone 16	AAF72627	278
Cupressus sempervirens	Conifer	Cup s 1	Clone 15	AAF72626	278
Cupressus sempervirens	Conifer	Cup s 1	Clone 13	AAF72625	279
Cynodon dactylon	Bermuda grass	B1	B1 clone	A28046	71
Cynodon dactylon	Bermuda grass	B2	B2 clone	A28056	247
Cynodon dactylon	Bermuda grass	B4	B4 clone	A28050	73
Cynodon dactylon	Bermuda grass	Cyn d 1	Fragment	A61226	27
Cynodon dactylon	Bermuda grass	Cyn d 1	Clone 14c1 and CD1	S83343	246
Cynodon dactylon	Bermuda grass	Cyn d 7	Calcium binding protein	P94092	82
Cynodon dactylon	Bermuda grass	Cyn d 12	Profilin 1	Y08390	131
Dactylis glomerata	Orchard grass	Dac g 1	AgDg1, fragment	D58493	34
Dactylis glomerata	Orchard grass	Dac g 2	Similar to Lol p 1, fragment	Q41183	36

Dactylis glomerata	Orchard grass	Dac g 3	Fragment	P93124	96
Dactylis glomerata	Orchard grass	Dac g 3	Fragment	A60359	28
Dactylis glomerata	Orchard grass	Dac g 5a	Fragment	(Klysner <i>et al.</i> , 1992)	10
Dactylis glomerata	Orchard grass	Dac g 5b	Fragment	(Klysner <i>et al.</i> , 1992)	10
Festuca elatior	Reed fescue	Fes e 1	Fragment	A58493	35
Festuca elatior	Reed fescue	Fes e 1	Type A, fragment	C37396	17
Festuca elatior	Reed fescue	Fes e 1	Type B, fragment	D37396	20
Holcus lanatus	Velvet grass	Hol l 1.0101		Z27084	265
Holcus lanatus	Velvet grass	Hol l 1.0102		Z68893	248
Holcus lanatus	Velvet grass	Hol l 5	Group V allergen	Z97874	264
Holcus lanatus	Velvet grass	Hol l 5.02	Group V allergen	Z97875	240
Holcus lanatus	Velvet grass	30 kDa	Fragment	S38291	20
Hordeum vulgare	Barley	Hor v 9	Pollen allergen	JC5475	313
Juniperus ashei	Mountain cedar	Jun a 1-1		AF106662	367
Juniperus ashei	Mountain cedar	Jun a 1-2		AF106663	367
Juniperus ashei	Mountain cedar	Jun a 2		CAC05582	363
Juniperus ashei	Mountain cedar	Jun a 3	Pathogenesis-related protein	AF121776	225
Juniperus oxycedrus	Juniper	Jun o 2	EF-hand calcium binding	AF031471	165
Juniperus virginiana	Red cedar	Jun v 1-1		AAF80166	280
Juniperus virginiana	Red cedar	Jun v 1-2		AAF80164	279
Juniperus virginiana	Red cedar	Jun v 3-1		AAF80167	76
Juniperus virginiana	Red cedar	Jun v 3-2		AAF80165	76
Ligustrum vulgare	privet	L1		X77787	145
Ligustrum vulgare	privet	L10		X77788	145
Lolium perenne	Perennial ryegrass		Pollen allergen plb	A38582	308
Lolium perenne	Perennial ryegrass		50k allergen	S38288	20
Lolium perenne	Perennial ryegrass	Lol p 1	Group I, allergen R7	P14946	263
Lolium perenne	Perennial ryegrass	Lol p 1	Group I	S13614	263
Lolium perenne	Perennial ryegrass	Lol p 1.0102	Group I, clone 1A	A37881	252
Lolium perenne	Perennial ryegrass	Lol p 1.0101	Group I, clone 5A	B37881	263
Lolium perenne	Perennial ryegrass	Lol p 2	Group II, fragment	A48595	88
Lolium perenne	Perennial ryegrass	Lol p 2	Group II	Q40239	88
Lolium perenne	Perennial ryegrass	Lol p 2-A	Group II	P14947	97
Lolium perenne	Perennial ryegrass	Lol p 3	Group III	R14948	97
Lolium perenne	Perennial ryegrass	Lol p 4	Fragment	A60737	14
Lolium perenne	Perennial ryegrass	Lol p 5A	Similar to Poa P9/Phl P6	Q40240	308
Lolium perenne	Perennial ryegrass	Lol p 5B	Similar to Poa P9/Phl P6	Q40237	339
Lolium perenne	Perennial ryegrass	Lol p XI		A54002	134
Mercurialis annua		Mer a 1	Profilin	Y13271	133
Olea europea	Olive tree	Ole e 1.0102		Y12428	146
Olea europea	Olive tree	Ole e 1.0103		Y12427	146
Olea europea	Olive tree	Ole e 1.05		Y12426	146
Olea europea	Olive tree	Ole e 1		S75766	145
Olea europea	Olive tree	Ole e 1	Fragment	S36872	145
Olea europea	Olive tree	Ole e 2	Profilin 1	O24169	134

Olea europea	Olive tree	Ole e 2	Profilin 2	O24170	134
Olea europea	Olive tree	Ole e 2	Profilin 3	O24171	134
Olea europea	Olive tree	Ole3	Calcium-binding allergen	AF015810	84
Olea europea	Olive tree	Ole16	Fragment	I53806	137
Olea europea	Olive tree	Ole17	Fragment	E53806	136
Olea europea	Olive tree	Ole19	Fragment	F53806	136
Olea europea	Olive tree	Ole1c	Fragment	C53806	145
Olea europea	Olive tree	Ole20	Fragment	A38968	137
Olea europea	Olive tree	Ole26	Fragment	G53806	136
Olea europea	Olive tree	Ole33/Ole37	Fragment	D53806	136
Olea europea	Olive tree	Ole3c	Fragment	A53806	145
Olea europea	Olive tree	Ole5c	Fragment	B53806	145
Olea europea	Olive tree	Ole6	Fragment	H53806	136
Olea europea	Olive tree	Ole e 4	Fragment	P80741	24
Olea europea	Olive tree	Ole e 5	Superoxide dismutase, fragment	P80740	29
Olea europea	Olive tree	Ole e 6		O24172	50
Olea europea	Olive tree	Ole e 7	Fragment	P81430	19
Parietaria judaica		Par j 1.0101	Nonspecific lipid transfer protein	P43217	133
Parietaria judaica		Par j 1.0201	Nonspecific lipid transfer protein 1	Q40905	138
Parietaria judaica		Par j 2.0101	Nonspecific lipid transfer protein 2	P55958	133
Parthenium hysterophorus	Compositae weed	Par h 1	31 kDa hydroxyproline-rich glycoprotein	(Gupta et al., 1996)	91
Parietaria officinalis		Par o 1	Fragment	A53252	12
Phalaris Aquatica	Canary grass	Pha a 1	Similar to Lol p 1	Q41260	269
Phalaris Aquatica	Canary grass	Pha a 5.1	Similar to Poa p 9/Phl p 6	P56164	320
Phalaris Aquatica	Canary grass	Pha a 5.2	Similar to Poa p 9/Phl p 7	P56165	305
Phalaris Aquatica	Canary grass	Pha a 5.3	Similar to Poa p 9/Phl p 8	P56166	294
Phalaris Aquatica	Canary grass	Pha a 5.4	Similar to Poa p 9/Phl p 9	P56167	175
Phleum pratense	Common timothy	Phl p 1.0101		Z27090	263
Phleum pratense	Common timothy	Phl p 1.0102		X78813	263
Phleum pratense	Common timothy	Phl p 2		X75925	122
Phleum pratense	Common timothy	Phl p 5.0105	Clone 10022	AF061066	276
Phleum pratense	Common timothy	Phl p 5.0106	Clone 10027	AF061067	276
Phleum pratense	Common timothy	Phl p 5.0107	Clone 10029	AF061068	276
Phleum pratense	Common timothy	Phl p 5.0108	Clone 10030	AF061069	276
Phleum pratense	Common timothy	Phl p 5.0103	Group 5	AF069470	312
Phleum pratense	Common timothy	Phl p 5.0203	Group 5	AF069471	295
Phleum pratense	Common timothy	Phl p 5.0204	Group 5	AF069472	265
Phleum pratense	Common timothy	Phl p 5.0206	Group 5	AF069473	290
Phleum pratense	Common timothy	Phl p 5.0207	Group 5	AF069474	287
Phleum pratense	Common timothy	Phl p 5.0202	Isovariant 711I13	Z82985	281
Phleum pratense	Common timothy	Phl p 5.0104	Isovariant 618I18	Z82986	276
Phleum pratense	Common timothy	Phl p 5A	Similar to Poa P9/Phl p 6	Q40962	286
Phleum pratense	Common timothy	Phl p 5A		S32101	257
Phleum pratense	Common timothy	Phl p 5B	Similar to Poa p 9/Phl p 6	Q40963	284
Phleum pratense	Common timothy	Phl p 5B	Similar to Poa p 9/Phl p 7	S38584	280



Phleum pratense	Common timothy	Phl p 6	Isolate c223	Y16955	138
Phleum pratense	Common timothy	Phl p 6	Isolate c142	Y16956	138
Phleum pratense	Common timothy	Phl p 11	Profilin	P35079	131
Phleum pratense	Common timothy	Phl p 13	Polygalacturonase	AJ238848	394
Pinus radiata	Pine	PRE29		AF049068	293
Poa pratensis	Kentucky bluegrass	Poa p 9	Clone 31	C39098	373
Poa pratensis	Kentucky bluegrass	Poa p 9	Clone 41	A39098	333
Poa pratensis	Kentucky bluegrass	Poa p 9	KBG60	P22286	307
Poa pratensis	Kentucky bluegrass	Poa p 9	Clone 60	B39098	307
Poa pratensis	Kentucky bluegrass	Poa p 9	Clone 7.2	A60373	131
Pyrus communis	Pear	Pyr c 1	Bet v 1 homolog	AF057030	159
Pyrus communis	Pear	Pyr c 4	Profilin	AF129424	131
Pyrus communis	Pear	Pyr c 5	Isoflavone reductase	AF071477	308
Quercus alba	Oak	Qbe a 1	Bet v 1 homolog, fragment	D53288	24
Syringa vulgaris	Lilac	Syr v 1	Isoform 1	S43242	145
Syringa vulgaris	Lilac	Syr v 1	Isoform 2	S43243	145
Syringa vulgaris	Lilac	Syr v 1	Isoform 3	S43244	145

#### Aero-Allergens - Fungi

Alternaria alternata	Mold	Alt a 1		U62097	135
Alternaria alternata	Mold	Alt a 1		U82633	157
Alternaria alternata	Mold	Alt a 3	Hsp70	U87807	124
Alternaria alternata	Mold	Alt a 3	Hsp70	U87808	117
Alternaria alternata	Mold	Alt a 4	Protein sulfisomerase	X84217	433
Alternaria alternata	Mold	Alt a 6	Ribosomal protein	X78222	113
Alternaria alternata	Mold	Alt a 7	YCR4 protein	X78225	204
Alternaria alternata	Mold	Alt a 10	Aldehyde dehydrogenase	P42041	495
Alternaria alternata	Mold	Alt a 11	Enolase	(Breltenbach et al., 1997)	439
Alternaria alternata	Mold	Alt a 12	60S ribosomal protein P1	P49148	110
Aspergillus flavus	Mold	Asp fl 13	Alkaline serine protease, elastase	P35211	403
Aspergillus fumigatus	Mold	Asp f 1	Ribonuclease, mitogillin	P04389	176
Aspergillus fumigatus	Mold	Asp f 1	Ribonuclease, mitogillin	A46497	176
Aspergillus fumigatus	Mold	Asp f 1	Partial exon 2	AJ005037	125
Aspergillus fumigatus	Mold	Asp f 2		P79017	310
Aspergillus fumigatus	Mold	Asp f 3	Peroxisomal protein	U28722	250
Aspergillus fumigatus	Mold	Asp f 4	ABPA patient allergen	AJ001732	286
Aspergillus fumigatus	Mold	Asp f 5	Metalloprotease (MEP)	P46075	649
Aspergillus fumigatus	Mold	Asp f 6	Mn superoxide dismutase	U53561	221
Aspergillus fumigatus	Mold	Asp f 7		AJ223315	112
Aspergillus fumigatus	Mold	Asp f 8	Ribosomal protein P2	AJ224333	111
Aspergillus fumigatus	Mold	Asp f 9		AJ223327	302
Aspergillus fumigatus	Mold	Asp f 10	Aspartic protease	X85092	395
Aspergillus fumigatus	Mold	Asp f 12	Heat shock protein hsp1	P40292	441
Aspergillus fumigatus	Mold	Asp f 13	Alkaline serine protease	AJ002026	152

Aspergillus fumigatus	Mold	Asp f 17		AJ224865	197
Aspergillus niger	Baker's yeast	Asp n ?	Xylanase	Z84377	804
Aspergillus niger	Baker's yeast	Asp n 14	Beta-xylosidase	AF108944	804
Aspergillus oryzae		Asp o 13	Alkaline serine protease	P12547	403
Aspergillus oryzae		Asp o 21	TAKA-amylase A	JK0201	478
Candida albicans	Yeast	Can a 1	Alcohol dehydrogenase	P43067	350
Candida albicans	Yeast	Can a ?	Enolase 1	P30575	440
Candida boidinii	Yeast	Can b 2	Peroxisomal membrane protein A	P14292	166
Candida boidinii	Yeast	Can b 2	Peroxisomal membrane protein B	P14293	166
Cladisporium herbarum		Cla h 3	Aldehyde dehydrogenase	S43114	496
Cladisporium herbarum		Cla h 4	60S ribosomal protein P2	P42038	111
Cladisporium herbarum		Cla h 4	60S ribosomal protein P2	P42039	111
Cladisporium herbarum		Cla h ?	Heat shock protein hsp70	P40918	643
Cladisporium herbarum		Cla h 5	YCP4 protein	P42059	204
Cladisporium herbarum		Cla h 6	Enolase	P42040	440
Cladisporium herbarum		Cla h 12	60S ribosomal protein P1	P50344	110
Coprinus comatus	Shaggy cap	Cop c 1	Leucine zipper protein	AJ132235	81
Coprinus comatus	Shaggy cap	Cop c 2	Thioredoxin	AJ242791	106
Coprinus comatus	Shaggy cap	Cop c 3		AJ242792	328
Coprinus comatus	Shaggy cap	Cop c 5		AJ242793	141
Coprinus comatus	Shaggy cap	Cop c 7		AJ242794	152
Fusarium solani pisi		Fus s 2		P81010	8
Malassezia furfur	Mold	Mal f 1		X96486	350
Malassezia furfur	Mold	Mal f 2	MF1	JE0226	124
Malassezia furfur	Mold	Mal f 2	MF1	AB011804	177
Malassezia furfur	Mold	Mal f 3	MF2	JE0227	114
Malassezia furfur	Mold	Mal f 3	MF2	AB011805	166
Malassezia furfur	Mold	Mal f 4	Malate dehydrogenase homolog	AF084828	342
Malassezia furfur	Mold	Mal f 5		AJ011955	172
Malassezia furfur	Mold	Mal f 6	Cyclophilin homolog	AJ011956	162
Malassezia furfur	Mold	Mal f 7		AJ011957	187
Malassezia furfur	Mold	Mal f 8		AJ011958	132
Malassezia furfur	Mold	Mal f 9		AJ011959	87
Penicillium chrysogenum		Pen n 18	Vacuolar serine protease	AAE71379	371
Penicillium chrysogenum		68 kDa		S77837	117
Penicillium citrinum		Pen c 1	Alkaline serine protease	AF084846	397
Penicillium citrinum		Pen c 2	Alkaline serine protease	AF098517	457
Penicillium oxalicum		Pen o 18	34 kDa, similar to vacuolar serine protease	(Shen <i>et al.</i> , 1999)	20
Penicillium notatum		28 kDa	Peptide PN3, PN4	(Shen <i>et al.</i> , 1999)	20
Penicillium notatum		Pen n 18	32 kDa, similar to vacuolar serine protease	(Shen <i>et al.</i> , 1999)	15
Penicillium notatum		Pen n 20	68 kDa, N-acetyl glucosaminidase	Q02352	117
Trichophyton rubrum		Tri r 2	Serine protease	AF082515	412
Trichophyton rubrum		Tri r 4	Serine protease	AF082514	728
Trichophyton tonsurans		Tri t 4	83 kD hypersensitivity protein	P80514	26

# **Aero-Allergens - Animals**

Canis familiaris	Dog	Can f ?	Albumin, salivary gland allergen	S72946	265
Canis familiaris	Dog	Can f 1	Lipocalin	O18873	174
Canis familiaris	Dog	Can f 2	Lipocalin	O18874	180
Bos taurus	Bovine dander	BDA11	EF-hand calcium binding protein	Q28050	101
Bos taurus	Bovine dander	BDA20	Clones pPOT, pPOT10.2	L42867	172
Equus caballus	Horse dander	Equ c 2.0101	Lipocalin	P81216	29
Equus caballus	Horse dander	Equ c 2.0102	Lipocalin	P81217	19
Equus caballus	Horse	Equ c 1	Lipocalin	Q95182	187
Felis silvestris catus	Cat	Fel d 1-A		P30438	92
Felis silvestris catus	Cat	Fel d 1-A	Chain 1 precursor A	JC1136	92
Felis silvestris catus	Cat	Fel d 1-A	Chain 1 precursor B	JC1126	88
Felis silvestris catus	Cat	Fel d 1	Chain 1 short form	B56413	88
Felis silvestris catus	Cat	Fel d 1	Chain 1 long form	A56413	92
Felis silvestris catus	Cat	Fel d 1	Chain 2	M77341	109
Felis silvestris catus	Cat	Fel d 1	Chain 2 short form	JC1127	107
Felis silvestris catus	Cat	Fel d 1-A		P30439	88
Mus musculus	Mouse		Lacrimal gland protein	AF008595	93
Mus musculus	Mouse	MUP	Major urinary protein	M27068	184
Mus musculus	Mouse	MUP 1	Major urinary protein I	M16355	180
Mus musculus	Mouse	MUP 2	Major urinary protein II	M16356	180
Mus musculus	Mouse	MUP 4	Major urinary protein IV	M16358	178
Mus musculus	Mouse	MUP 5	Major urinary protein V	M16360	180
Rattus norvegicus	Rat	Rat n 1	Transferrin	P02767	147
Rattus norvegicus	Rat		Alpha 2u globulin	J00737	177

# **Aero-Allergens - Insects**

Blattella germanica	German cockroach	Bla g 1.0101		AF072219	412
Blattella germanica	German cockroach	Bla g 1.02		AF072220	492
Blattella germanica	German cockroach	Bla g 2	Aspartyl protease	A57164	352
Blattella germanica	German cockroach	Bla g 4	Calycin	U40767	182
Blattella germanica	German cockroach	Bla g 5	Glutathione S-aryltransferase	U92412	200
Chironomus thummi thummi	Midge	Chi t 1.01	Globin component III	P02229	151
Chironomus thummi thummi	Midge	Chi t 1.02	Globin component IV	P02230	151
Chironomus thummi thummi	Midge	Chi t 2.0101	Globin component I	P02221	158
Chironomus thummi thummi	Midge	Chi t 3	Globin component II beta	P02222	160
Chironomus thummi thummi	Midge	Chi t 4	Globin component III-A	P02231	151
Chironomus thummi thummi	Midge	Chi t 5	Globin component VI	P02224	162
Chironomus thummi thummi	Midge	Chi t 6.01	Globin component VII-A	P02226	145
Chironomus thummi thummi	Midge	Chi t 6.02	Globin component IX	P02223	161
Chironomus thummi thummi	Midge	Chi t 7	Globin component VII-B	P02225	161
Chironomus thummi thummi	Midge	Chi t 8	Globin component VIII	P02227	151
Chironomus thummi thummi	Midge	Chi t 9	Globin component X	P02228	151
Periplaneta americana	American cockroach	Per a 1.0101		AF072222	231

Periplaneta americana	American cockroach	Per a 1.0102	Cr-P11	U78970	228
Periplaneta americana	American cockroach	Per a 1.0103	Cr-P11	U69957	395
Periplaneta americana	American cockroach	Per a 1.0104	Cr-P11	U69261	274
Periplaneta americana	American cockroach	Per a 1.02	Cr-P11	U69260	446
Periplaneta americana	American cockroach	Per a 3.01	Cr-P1, clone C12	L40818	685
Periplaneta americana	American cockroach	Per a 3.0201	Cr-P1, clone C20	L40820	631
Periplaneta americana	American cockroach	Per a 3.0202	Cr-P1	L40819	470
Periplaneta americana	American cockroach	Per a 3.0203	Cr-P1, clone C28	L40821	393
Periplaneta americana	American cockroach	Per a 7	Tropomyosin	Y14854	284

#### Aero-Allergens - Mites

Acarus siro	Mite	Aca s 13	Partial sequence	AJ006774	64
Blomia tropicalis	Mite	Blot 5		U59102	134
Blomia tropicalis	Mite	Blot 12	Bt11a	U27479	144
Blomia tropicalis	Mite		Partial sequence	U27702	73
Blomia tropicalis	Mite	Blot 13	Fatty-acid binding protein, Bt6	U58106	130
Dermatoph. farinae	House dust mite	Der f 1	Thiol protease	P16311	321
Dermatoph. farinae	House dust mite	Der f 1		A61500	319
Dermatoph. farinae	House dust mite	Der f 1	Cysteine protease, Group I	X65196	212
Dermatoph. farinae	House dust mite	Der f 2	Complete sequence	D10448	146
Dermatoph. farinae	House dust mite	Der f 2	Partial sequence	D10449	138
Dermatoph. farinae	House dust mite	Der f 2	Group II allergen	S70378	142
Dermatoph. farinae	House dust mite	Der f 3	Trypsin family protease	P49275	259
Dermatoph. farinae	House dust mite	Der f 3	Trypsin family protease	Q94508	232
Dermatoph. farinae	House dust mite	Der f 6	Trypsin family protease, fragment	P49276	20
Dermatoph. farinae	House dust mite	Der f 7		Q26456	213
Dermatoph. farinae	House dust mite	Der f 10	Tropomyosin, MAG 44	D17682	299
Dermatoph. farinae	House dust mite	Der f 14	Apolipoprotein-like protein, MAG 3	D17686	349
Dermatoph. farinae	House dust mite	MAG		P39673	341
Dermatoph. farinae	House dust mite	MAG29	Heat shock protein hsp70	P39674	145
Dermatoph. microceras	House dust mite	Der m 1	Fragment	B27634	30
Dermatoph. pteronyssinus	House dust mite	Der p 1	Cysteine protease, group I	X65197	211
Dermatoph. pteronyssinus	House dust mite	Der p 1	Preproenzyme, complete sequence	U11695	320
Dermatoph. pteronyssinus	House dust mite	Der p 1	Cysteine protease	A31657	92
Dermatoph. pteronyssinus	House dust mite	Der p 1	Cysteine protease, fragment	B31657	23
Dermatoph. pteronyssinus	House dust mite	Der p 1	Cysteine protease, fragment	S03380	94
Dermatoph. pteronyssinus	House dust mite	Der p 1	Cysteine protease	JQ0337	245
Dermatoph. pteronyssinus	House dust mite	Der p 2		P49278	146
Dermatoph. pteronyssinus	House dust mite	Der p 3	Prepro-zymogen, complete	U11719	261
Dermatoph. pteronyssinus	House dust mite	Der p 3	Group III allergen	A39997	18
Dermatoph. pteronyssinus	House dust mite	Der p 4	Alpha amylase	AF144060	496
Dermatoph. pteronyssinus	House dust mite	Der p 5		P14004	132
Dermatoph. pteronyssinus	House dust mite	Der p 6	Trypsin family protease, chymotrypsin	P49277	20
Dermatoph. pteronyssinus	House dust mite	Der p 7	Complete sequence	U37044	215

Dermatoph. pteronyssinus	House dust mite	Der p 8	Glutathione S-transferase	S75286	219
Dermatoph. pteronyssinus	House dust mite	Der p 9	Collagenolytic serine protease	AAB50781	18
Dermatoph. pteronyssinus	House dust mite	Der p 10	Tropomyosin	O18416	284
Dermatoph. pteronyssinus	House dust mite	Der p 10	Tropomyosin	O16188	284
Dermatoph. pteronyssinus	House dust mite	Der p 10	Tropomyosin	AF016278	284
Euroglyphus maynei	House dust mite	Eur m 1	Group I, cysteine protease	X60073	211
Euroglyphus maynei	House dust mite	Eur m 3.0101	Trypsin family protease	AF047615	261
Euroglyphus maynei	House dust mite	Eur m 4	Alpha amylase	AF144061	521
Euroglyphus maynei	House dust mite	Eur m 14	Apolipoprotein-like protein	AF149827	1668
Glycyphagus domesticus	mite	Gly d 2		CAB59976	128
Glycyphagus domesticus	mite	Gly d 2.02	Isoform	CAB76459	125
Lepidoglyphus destructor	Storage mite	Lep d 1.0102		X89014	141
Lepidoglyphus destructor	Storage mite	Lep d 1.02		S66499	141
Tyrophagus putrescentiae	Dust mite		Group 2	Y12690	141

#### Venom & Salivary Allergens

Aedes aegypti	Yellowfever mosquito	Aed a 1	Salivary gland allergen	L12389	562
Aedes aegypti	Yellowfever mosquito	Aed a 2	Salivary gland allergen	M33157	321
Aedes aegypti	Yellowfever mosquito	Aed a 3	30 kDa salivary gland allergen	O01949	253
Anopheles gambiae	African malaria mosquito		gVAG protein, clone cF6	Y17702	60
Apis mellifera	Honeybee	Api m 1	Phospholipase A2	P00630	162
Apis mellifera	Honeybee	Api m 2	Hyaluronoglucosaminidase	Q08169	382
Apis mellifera	Honeybee	Api m 3	Melittin	P01501	70
Ctenocephalides felis	Cat flea	Cte f 1	Salivary antigen 1	AF102502	176
Ctenocephalides felis	Cat flea	Cte f 2		AAF65314	264
Dolichovespula arenaria	Yellow hornet	Dol a 5	Venom allergen 5	Q05108	203
Dolichovespula maculata	Whiteface hornet	Dol m 1.01	Phospholipase A1	Q06478	317
Dolichovespula maculata	Whiteface hornet	Dol m 1.02	Phospholipase A1 2	P53357	303
Dolichovespula maculata	Whiteface hornet	Dol m 2	Hyaluronoglucosaminidase	P49371	331
Dolichovespula maculata	Whiteface hornet	Dol m 5.01	Venom allergen 5.01	P10736	227
Dolichovespula maculata	Whiteface hornet	Dol m 5.02	Venom allergen 5.02	P10737	215
Myrmecia pilosula	Bulldog ant	Myr p 1		S28180	112
Myrmecia pilosula	Bulldog ant	Myr p 1		X70256	112
Myrmecia pilosula	Bulldog ant	Myr p 2		Q26464	75
Polistes annularis	Paper wasp	Pol a 5	Venom allergen 5	M96857	209
Polistes dominulus	Paper wasp	Pol d 5	Venom allergen 5	P81656	206
Polistes exclamans	Paper wasp	Pol e 5	Venom allergen 5	P35759	205
Polistes fuscatus	Paper wasp	Pol f 5	Venom allergen 5	F44583	205
Solenopsis geminata		Sol g 4.01		AAF65312	88
Solenopsis geminata		Sol g 4.02		AAF65313	90
Solenopsis invicta	Red fire ant	Sol i 2	Phospholipase, Venom allergen 2	A37330	138
Solenopsis invicta	Red fire ant	Sol i 3	Venom allergen 3	P35778	212
Solenopsis invicta	Red fire ant	Sol i 3	Venom allergen 3	AF012919	231
Solenopsis invicta	Red fire ant	Sol i 4	Venom allergen 4	P35777	117



Solenopsis richteri	Black fire ant	Sol r 2	Phospholipase, fragment	E60727	20
Solenopsis richteri	Black fire ant	Sol r 3	Venom allergen 3, fragment	P35779	20
Vespa crabro	European hornet	Vesp c 5.01	Antigen 5	G44583	202
Vespa crabro	European hornet	Vesp c 5.02	Antigen 5	H44583	202
Vespa madagascariensis	Hornet	Vesp m 5	Venom allergen 5	P81657	202
Vespula flavopilosa	Wasp	Ves f 5	Venom allergen 5	A44583	204
Vespula germanica	Wasp	Ves g 5	Venom allergen 5	B44583	204
Vespula maculifrons	Wasp	Ves m 1	Phospholipase A1	A44564	300
Vespula maculifrons	Wasp	Ves m 5	Venom allergen 5	P35760	204
Vespula pensylvanica	Wasp	Ves p 5	Venom allergen 5	C44583	204
Vespula squamosa	Wasp	Ves s 5	Venom allergen 5	D44583	205
Vespula vidua	Wasp	Ves vi 5		E44583	206
Vespula vulgaris	Wasp	Ves v 1	Phospholipase A1	L43561	336
Vespula vulgaris	Wasp	Ves v 2	Hyaluronoglucosaminidase	P49370	331
Vespula vulgaris	Wasp	Ves v 5	Venom allergen 5	M98858	227

#### Others - Contact Allergens

Hevea brasiliensis	Para rubber tree		Latex allergen	AJ223038	388
Hevea brasiliensis	Para rubber tree	Hev b 1	Rubber elongation factor	P15252	137
Hevea brasiliensis	Para rubber tree	Hev b 2	Beta 1,3 glucanase	S65077	374
Hevea brasiliensis	Para rubber tree	Hev b 3	Fragment	AAB33777	15
Hevea brasiliensis	Para rubber tree	Hev b 3	Fragment	AAB33778	18
Hevea brasiliensis	Para rubber tree	Hev b 3	Fragment	AAB33779	15
Hevea brasiliensis	Para rubber tree	Hev b 5	Latex allergen	U42640	151
Hevea brasiliensis	Para rubber tree	Hev b 6	Hevein	P02877	204
Hevea brasiliensis	Para rubber tree	Hev b 7	Latex patatin homolog	AJ223039	388
Hevea brasiliensis	Para rubber tree	Hev b 7	Latex patatin homolog	U80598	388
Hevea brasiliensis	Para rubber tree	Hev b 8	Profilin	Y15042	131
Hevea brasiliensis	Para rubber tree	Hev b 9	Enolase (isoform 1)	CAC00532	445
Hevea brasiliensis	Para rubber tree	Hev b 9	Enolase (isoform 2)	CAC00533	445
Hevea brasiliensis	Para rubber tree	Hev b 10	Mn superoxide dismutase	AJ249148	205

#### Others - Nematodes & Worms

Acanthamoeba castellanii		profilin ia	Actin-binding protein	1BRQ	125
Ancylostoma caninum	dog hookworm		Aspartic proteinase	JC5077	442
Angiostoma duodenale	Hookworm		Ancylostoma secreted protein 1	AF077402	425
Anisakis simplex		Ans1	Excretory gland allergen	A59069	17
Ascaridia galli		AG1	Fatty acid binding protein	AF065383	135
Ascaris lumbricoides	Common roundworm	ABA-1	ABA-1d1	U86100	134
Ascaris lumbricoides	Common roundworm	ABA-1	ABA-1dr1	U86097	267
Ascaris lumbricoides	Common roundworm	ABA-1	ABA-1d2	U86101	134
Ascaris lumbricoides	Common roundworm	ABA-1	ABA-1dr2	U86098	267
Ascaris lumbricoides	Common roundworm	ABA-1	ABA-1d3	U86102	184
Ascaris lumbricoides	Common roundworm	ABA-1	ABA-1dr3	U86099	267

Ascaris lumbricoides	Common roundworm	ABA-1	ABA-1r1	U86091	134
Ascaris lumbricoides	Common roundworm	ABA-1	ABA-1r2	U86092	134
Ascaris lumbricoides	Common roundworm	ABA-1	ABA-1r4	U86094	133
Ascaris lumbricoides	Common roundworm	ABA-1	ABA-1r6	U86096	133
Ascaris lumbricoides	Common roundworm	ABA-1	Fragment	B37188	41
Ascaris lumbricoides	Common roundworm	ABA-1	Fragment	A48576	395
Ascaris suum	Pig roundworm	ABA-1	Fatty acid binding protein	AF051702	1095
Ascaris suum	Pig roundworm	ABA-1		Q06811	395
Ascaris suum	Pig roundworm	ABA-1	Fragment	A37188	41
Ascaris suum	Pig roundworm	ABA-1	Fatty acid binding protein	L03211	395
Brugia malayi	Nematode		Vespid allergen antigen homolog	AF042088	220
Dictyocaulus vivipara	Bovine lungworm	DVA-1	ABA-1	Q24702	1557
Toxocara canis		TBA-1	ABA-1 homolog	B49139	36
Dirofilaria immitis			Venom allergen 5-like protein	AF001100	221
Loa loa	Filarial worm	LL20	15kDa ladder antigen	U03103	588
Meloidogyne incognita	Southern root-knot nematode	MSP-1	Secreted protein	AF013289	231
Necator americanus	Hookworm	ASP1	Ancylostoma secreted protein 1	AF079521	424
Onchocerca volvulus			Vespid allergen antigen homolog	AF042087	220
Strongyloides stercoralis			Allergen polypeptide homolog	AF035658	95

## Appendix 2. FASTA sequence similarity alignments of corn event MON 863 Cry3Bb1.11098 protein with allergens and gliadins.

!!SEQUENCE\_LIST 1.0

(Peptide) FASTA of: MON863.pep From: 1 to: 653 February 14, 2001 15:14

TO: allergen3.\* \* Sequences: 659 Symbols: 141,391 Word Size: 2

Databases searched: Release 1.0, Released on 13Oct2000, Formatted on 13Oct2000

Scoring matrix: GenRunData:blosom50.cmp  
Variable pamfactor used  
Gap creation penalty: 12  
Gap extension penalty: 2

Histogram Key:

Each histogram symbol represents 2 search set sequences  
z-scores computed from opt scores

z-score obs exp  
(=) (\*)

```
< 20 1 0:=
22 0 0:
24 0 0:
26 0 0:
28 0 0:
30 5 1.*==
32 3 4:==*
34 14 10:====*==
36 18 20:====*==
38 55 32:====*==
40 21 45:====*==
42 35 55:====*==
44 51 61:====*==
46 49 62:====*==
48 80 59:====*==
50 46 54:====*==
52 49 47:====*==
54 35 41:====*==
56 39 34:====*==
58 28 28:====*==
60 24 23:====*==
62 25 18:====*==
64 22 14:====*==
66 20 11:====*==
68 7 9:====*
70 9 7:====*
```

```
72 9 5:====
74 3 4:==*
76 4 3:==*
78 2 3:==*
80 3 2:==*
82 0 2:==*
84 1 1:==*
86 1 1:==*
88 0 1:==*
90 0 1:==*
92 0 0:
94 0 0:
96 0 0:
98 0 0:
100 0 0:
102 0 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:
```

Joining threshold: 38, opt. threshold: 26, opt. width: 16, reg.-scaled

The best scores are:

	init1	initn	opt	z-sc	E(658)...
AL:S83343					
! S83343 Cynodon dactylon Cyn d 1=maj...	62	62	70	85.7	3.8
AL:P02229					
! P02229 chironomus thummi thummi (md...	48	48	66	85.0	4.1
AL:U42640					
! U42640 Hevea brasiliensis Hevea bra...	32	32	62	80.4	7.4
AL:P14004					
! P14004 dermatophagoides pteronyssin...	31	31	61	80.4	7.5
AL:C53806					
! major allergen OLE1c - common olive...	30	30	61	79.6	8.2
AL:AJ223327					
! AJ223327 Aspergillus fumigatus Aspe...	30	30	66	79.4	8.4
\\End of List					

MON863.pep  
AL:S83343

LOCUS

DEFINITION

S83343\_1

Cyn d 1=major allergen (clone 14c1 and C81) (Cynodon dactylon=Bermuda grass, pollen, mRNA Partial, 759 nt) major allergen; This sequence comes from Fig. 3. Protein sequence is in conflict with the conceptual translation; mismatch(45[G->R])...

SCORES Initl: 62 Initn: 62 Opt: 70 z-score: 85.7 E(): 3.8  
 Smith-Waterman score: 70; 23.6% identity in 55 aa overlap

	540	550	560	570	580	590
MON863.pep	PGFTGGNLLFLKESNSIAKFKVTLNSAALLQRYRVIRYASTTNLRLFVQNSNDPLVI					
S83343	AYHFDLGRAGAMAKKGGEDLRKAGELTQRRVKCKYPSGKITPHIEKGSNDHYLA					
	110	120	130	140	150	160
	600	610	620	630	640	650
MON863.pep	YINKTMNKDDDLTYQTFLATNSNMGSDDNELIIGASSFVSNEKIYIDKIEFIPVQL					
S83343	LLVKYAAGDGNIV--AVDIKPRBSDEFIPKSSWGAIWRIDPKKPLKGFPSIRLTSEGGA					
	170	180	190	200	210	220

MON863.pep  
 AL:P02229

ID GLB3\_CHITH STANDARD; PRT; 151 AA.  
 AC P02229;  
 DT 21-JUL-1986 (REL. 01, CREATED)  
 DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE GLOBIN CTT-III PRECURSOR (ERYTHROCRUORIN III) . . .

SCORES Initl: 48 Initn: 48 Opt: 66 z-score: 85.0 E(): 4.1  
 Smith-Waterman score: 66; 35.1% identity in 37 aa overlap

	460	470	480	490	500	510
MON863.pep	DSIDQLPPETTDPLEKAYSHQLNYAECFLMQDRRGTFPFTWTHRSVDFFNTIDAEKIT					
P02229	VKGDVPGILYAVFKADPSIMAKFTQFAGKDLESIKGYAPFEIHANRIVGFFSKI----IG					
	40	50	60	70	80	
	520	530	540	550	560	570
MON863.pep	QLPVVKAYALSSGASIEGPGFTGGNLLFLKESNSIAKFKVTLNSAALLQRYRVIRYA					
P02229	ELPNIEADVNTFVASHKPRGVTHDQNNFRAGFVSVMKAHTDFAGAEAAWATLDTFFGM					
	90	100	110	120	130	140

MON863.pep  
 AL:U42640

LOCUS HBU42640\_1  
 DEFINITION Hevea brasiliensis latex allergen Hev b 5 mRNA, complete cds.  
 DATE 11-OCT-1996  
 ACCESSION U42640  
 NID g1480456  
 ORGANISM Hevea brasiliensis . . .

SCORES Initl: 32 Initn: 32 Opt: 62 z-score: 80.4 E(): 7.4

Appendix 2

Study No. 01-01-39-23

MSL-17140

Page 41 of 102

Smith-Waterman score: 62; 30.9% identity in 55 aa overlap

		10	20	30
MON863.pep	MANPNRSEHDTIKVTPNSELQTNHNQVPLADNPNST			
U42640	SAATALPKNETPEVTKAEETKTEPAAPPASEQTADATPEKEEPTAAPAEPEAPAPET-			
	10	20	30	40
	40	50	60	70
MON863.pep	LEELNYKEFLRMTEDSSTEVLNDSTVKDAVGTGISVVGQILGVVGVFPFAGALTSFYQSFL			
U42640	-EKAEVEKIEKTEEPAPPA-DQTTPEEKPAEPFVAEEEPKHETKETETEAPAPAEGB			
	70	80	90	100
				110
				120

MON863.pep  
 AL:P14004

ID DER5\_DERPT STANDARD; PRT; 132 AA.  
 AC P14004;  
 DT 01-JAN-1990 (REL. 13, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE MITE ALLERGEN DER P 5 (DER P V) (IGE-BINDING ALLERGEN) . . .

SCORES Initl: 31 Initn: 31 Opt: 61 z-score: 80.4 E(): 7.5  
 Smith-Waterman score: 61; 25.9% identity in 58 aa overlap

	60	70	80	90	100	110
MON863.pep	DNSTVKDAVGTGISVVGQILGVVGVFPFAGALTSFYQSFLNTIWPSSADPWKAFMAQVEVL					
P14004	MKFIIAFFVATLAVMTVSGEDKKHXYQNEFDL					
	10	20	30			
	120	130	140	150	160	170
MON863.pep	IDKKIEEYAKSKALA--ELQGLCNFEDYVNALNSWKKTPLSLRSKRSQGRIRELFSQAE					
P14004	LMERIHDOYKNGELALFVLEQDINHFEKPTREMKDKIVAEMDTIAMDGVGVLDRLM					
	40	50	60	70	80	90

MON863.pep  
 AL:C53806

F1:C53806 - major allergen OLE1c - common olive (fragment)  
 C:Species: Olea europea (common olive)  
 C:Date: 01-Dec-1995 #sequence\_revision 01-Dec-1995 #text\_change 25-Oct-1996  
 C:Accession: C53806  
 R;Villalba, M.; Batanero, E.; Monsalve, R.I.; Gonzalez de la Pena, M.A.;  
 Lahoz,  
 C.; Rodriguez, R. . . .

SCORES Initl: 30 Initn: 30 Opt: 61 z-score: 79.6 E(): 8.2  
 Smith-Waterman score: 61; 26.0% identity in 73 aa overlap

450 460 470 480 490 500  
 MON863.pep KRNNGHVSAQDSIDQLPPETTDEPLEKAYSHQLNYAECFLMQRDR--GTIPFFTWTHRS  
 C53806 CKDGENGKITFTTEVGVTAEGLYSMLIERDHKNFCEITLLSSSRKDCDEIPTGHWKPS

! Distributed over 1 thread.  
 ! Start time: Wed Feb 14 15:14:53 2001  
 ! Completion time: Wed Feb 14 15:14:59 2001

510 520 530 540 550  
 MON863.pep VDF-PNTIDA--EKITQLPVVKAYALSSGASITIEGPGFTGGNLLFLKESNSIAKFKVTL  
 C53806 VKFILNTVNGTTRTINPLGFFKKEALPKCPQVFNKLMYPPNM

! CPU time used:  
 ! Database scan: 0:00:00.4  
 ! Post-scan processing: 0:00:00.3  
 ! Total CPU time: 0:00:00.8  
 ! Output File: MON863.allergen3

560 570 580 590 600 610  
 MON863.pep NSAALLQRYRVRIYASTTNLALFVQNSNNDFFVYINKTMNKDDDTYQTFDLATTNSN

MON863.pep  
 AL:AJ223327

LOCUS AFAJ3327\_1  
 DEFINITION Aspergillus fumigatus mRNA for rAsp f 9 allergen.  
 DATE 27-FEB-1998  
 ACCESSION AJ223327  
 NID g2879889  
 ORGANISM Aspergillus fumigatus . . .

SCORES Init1: 30 Initn: 30 Opt: 66 z-score: 79.4 E(): 8.4  
 Smith-Waterman score: 66; 32.3% identity in 65 aa overlap

320 330 340 350 360 370  
 MON863.pep LOKYGPTFLSIENSIRKPHILFDYLQGIETRLRPGYFGK-DSFNWYSGNVETRPISGS  
 AJ223327 AAFGTGVVSSIVLESDDLDEVDWEVLGGDTTQVQTNFYFGKGDITTYDRGTYPVATPQET  
 110 120 130 140 150 160

380 390 400 410 420 430  
 MON863.pep SKTITSPFYGDKSTEPVQKLSFDGQKVYRTIANTDVAAWPNGKVYLGVTKVDFSQYDDQK  
 AJ223327 FHTYTIWTKDAVT-----WSIDGA-VVRTLTYNDAKGGTRFPQTPMRLRLGSGWAGGDPS  
 170 180 190 200 210 220

440 450 460 470 480 490  
 MON863.pep NETSTQTYDSKRNGHVSAQDSIDQLPPETTDEPLEKAYSHQLNYAECFLMQRDRGTIPF  
 AJ223327 NPKGTIEWAGGLTDYSAGPYTMVKSURIENANPAESYTSDNSGSWQSIKFDGSDVISS  
 230 240 250 260 270 280

©01

Proprietary

Information of Monsanto Company



**Appendix 3.** Search for immunologically relevant sequences.

```
allergensearch -INfile1=MON863.pep -INFILE2=allergen3:** -  
OUTfile=MON863_allergen3.allergensearch -WINDOWsize=8
```

```
Query sequence(s): MON863.pep  
Search sequence(s): allergen3:**  
Window size: 8
```

no sequences found

MON863\_allergen3.allergensearch Wed Feb 14 15:16:15 2001

..

# Appendix 4. FASTA sequence similarity alignments of corn event MON 863 Cry3Bb1.11098 protein with toxins.

SEQUENCE LIST 1 0

(Peptide) FASTA of: MON863-map from: 1 to: 651 February 12, 2001 11:30  
TO: TOXIN4 \* \* Sequences: 4,897 Symbols: 1,402,897 Word Size: 3  
Database searched: Release 1.0, Released on 28Dec99, Formatted on 28Dec99  
Scoring matrix: GenRunData: blosum50.cmp  
Variable pamfactor used  
Gap creation penalty: 12  
Gap extension penalty: 2

Histogram Key:  
Each histogram symbol represents 9 search set sequences  
Each inset symbol represents 4 search set sequences  
z-scores computed from opt scores

z-score	obs	exp
(n)	(*)	
< 20	20	0.***
22	0	0
24	0	0
26	0	0
28	0	1 *
30	8	6 *
32	2	24 = *
34	27	65 ***
36	84	134 *****
38	130	221*****
40	264	308*****
42	372	377*****
44	474	416*****
46	536	423*****
48	462	405*****
50	440	370*****
52	430	325*****
54	261	278*****
56	251	232*****
58	181	151*****
60	196	154*****
62	82	124*****
64	105	98*****
66	62	78*****
68	26	61***
70	18	48**
72	19	37**
74	13	29**
76	6	23**
78	8	18**
80	5	14**
82	4	10**
84	2	8*
86	2	6*
88	2	5*
90	0	4*
92	14	3**
94	2	2*
96	0	2**
98	0	1*
100	0	1*
102	0	1**
104	0	1*
106	0	0
108	0	0
110	0	0
112	0	0
114	2	0**
116	0	0
118	0	0
>120	165	0*****

Joining threshold: .38, opt. threshold: .26, opt. width: .16, req. scaled

The best scores are:	init1	initn	opt	z-sc	E(4508)
TO:Q06117					
! Q06117 bacillus thuringiensis (subs...	4232	4232	4232	4196.8	2.7e-228
TO:BTU031633_1					
! U01633 Bacillus thuringiensis Bacil...	4178	4178	4178	4143.3	2.5e-225
TO:A07234_1					
! A07234 Bacillus thuringiensis B.thu...	4028	4028	4028	3994.7	4.8e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366_1					
! M1123 bacillus thuringiensis Bacil...	4028	4028	4028	3994.6	4.9e-217
TO:CR70_BACTO					
! P17965 bacillus thuringiensis (subs...	4028	4028	4028	3994.6	4.9e-217
TO:BTU04366					

## Appendix 4

Study No. 01-01-39-23

MSL-17140

## Appendix 4

Study No. 01-01-39-23

MSL-17140

Page 46 of 102



SCORES Initl 4028 Inltn 4028 Opt 4028 z-score: 3994.6 E(): 4 9e-217  
Smith-Waterman score: 4028, 93.4% identity in 650 aa overlap

MON863.pap MANPNRSEHDITKVTNPSRLQTNHNOYPLADNPSTLEELNKEFLAMTDEDS  
S10228 MIRMGRKNNPNRSEHDITKVTNPSRLQTNHNOYPLADNPSTLEELNKEFLAMTADN

MON863.pap STEVLNSTVKDAVGTGIVVGGIQLGVVGFAGALTSFYQSFLNTWPSDADPWKAFMA  
S10228 STEVLNSTVKDAVGTGIVVGGIQLGVVGFAGALTSFYQSFLNTWPSDADPWKAFMA

MON863.pap QVVLVDKKIEEYAKSALAEQLQNNFEDVYNALDSWKKAPVNLRSRQDRIRLFS  
S10228 QVVLVDKKIEEYAKSALAEQLQNNFEDVYNALDSWKKAPVNLRSRQDRIRLFS

MON863.pap QABSHFNSMPSFAVSKFEVLPTTAAQANTHLLKDAQVGEWYSSDVAEYFR  
S10228 QABSHFNSMPSFAVSKFEVLPTTAAQANTHLLKDAQVGEWYSSDVAEYFR

MON863.pap QKLTQOYTHCVNHYVGLNLRGSTDYDAMKFNRRREMTLVLDLVLFPFYDRLV  
S10228 QKLTQOYTHCVNHYVGLNLRGSTDYDAMKFNRRREMTLVLDLVLFPFYDRLV

MON863.pap SKGVKTHLTDITPDIPLTLQKQYPTFLSISIRKPHLPDVLQIEFHTLRAPYF  
S10228 SKGVKTHLTDITPDIPLTLQKQYPTFLSISIRKPHLPDVLQIEFHTLRAPYF

MON863.pap GKDSFNNMEGVYETRPISGSKTISFFYQDKSTEPVQKLSFDQKVRTIANTOVAAM  
S10228 GKDSFNNMEGVYETRPISGSKTISFFYQDKSTEPVQKLSFDQKVRTIANTOVAAM

MON863.pap PKNVYLVGVKVFSDQDQKNETSTQYDSKRNKGVSAQDQLPPTTDEPLEKAY  
S10228 PKNVYLVGVKVFSDQDQKNETSTQYDSKRNKGVSAQDQLPPTTDEPLEKAY

MON863.pap SHQLWYACFLMDRGTIPFFTWHSVDFTNTIDAKITQLPVVYKALSSGASITRG  
S10228 SHQLWYACFLMDRGTIPFFTWHSVDFTNTIDAKITQLPVVYKALSSGASITRG

MON863.pap PPTGONLFLKSSNSIAKFKVTLNSAALLQRYVRIRYASTNLRFLVQNSNDPLVI  
S10228 PPTGONLFLKSSNSIAKFKVTLNSAALLQRYVRIRYASTNLRFLVQNSNDPLVI

MON863.pap YINKTMKDDLTQYCTPLATTNMGFSDGNELIAGSPVNSKIYIDKIEFIPVQ  
S10228 YINKTMKDDLTQYCTPLATTNMGFSDGNELIAGSPVNSKIYIDKIEFIPVQ

MON863.pap TO CR70\_BACTT  
S10228 TO CR70\_BACTT

FI.S10228 - parasporal crystal protein, coleopteran-active - *Bacillus thuringiensis* (fragment)  
N: Alternate names delta-endotoxin  
C: Species: *Bacillus thuringiensis*  
C: Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 20-Feb-1998  
C: Accession S10228

SCORES Initl 4028 Inltn 4028 Opt 4028 z-score: 3994.6 E(): 4.9e-217  
Smith-Waterman score 4028; 93.4% identity in 650 aa overlap

MON863.pap MANPNRSEHDITKVTNPSRLQTNHNOYPLADNPSTLEELNKEFLAMTDEDS  
S10228 MIRMGRKNNPNRSEHDITKVTNPSRLQTNHNOYPLADNPSTLEELNKEFLAMTADN

MON863.pap STEVLNSTVKDAVGTGIVVGGIQLGVVGFAGALTSFYQSFLNTWPSDADPWKAFMA  
S10228 STEVLNSTVKDAVGTGIVVGGIQLGVVGFAGALTSFYQSFLNTWPSDADPWKAFMA

MON863.pap QVVLVDKKIEEYAKSALAEQLQNNFEDVYNALDSWKKAPVNLRSRQDRIRLFS  
S10228 QVVLVDKKIEEYAKSALAEQLQNNFEDVYNALDSWKKAPVNLRSRQDRIRLFS

MON863.pap QABSHFNSMPSFAVSKFEVLPTTAAQANTHLLKDAQVGEWYSSDVAEYFR  
S10228 QABSHFNSMPSFAVSKFEVLPTTAAQANTHLLKDAQVGEWYSSDVAEYFR

MON863.pap QKLTQOYTHCVNHYVGLNLRGSTDYDAMKFNRRREMTLVLDLVLFPFYDRLV  
S10228 QKLTQOYTHCVNHYVGLNLRGSTDYDAMKFNRRREMTLVLDLVLFPFYDRLV

MON863.pap SKGVKTHLTDITPDIPLTLQKQYPTFLSISIRKPHLPDVLQIEFHTLRAPYF  
S10228 SKGVKTHLTDITPDIPLTLQKQYPTFLSISIRKPHLPDVLQIEFHTLRAPYF

MON863.pap GKDSFNNMEGVYETRPISGSKTISFFYQDKSTEPVQKLSFDQKVRTIANTOVAAM  
S10228 GKDSFNNMEGVYETRPISGSKTISFFYQDKSTEPVQKLSFDQKVRTIANTOVAAM

MON863.pap PKNVYLVGVKVFSDQDQKNETSTQYDSKRNKGVSAQDQLPPTTDEPLEKAY  
S10228 PKNVYLVGVKVFSDQDQKNETSTQYDSKRNKGVSAQDQLPPTTDEPLEKAY

MON863.pap SHQLWYACFLMDRGTIPFFTWHSVDFTNTIDAKITQLPVVYKALSSGASITRG  
S10228 SHQLWYACFLMDRGTIPFFTWHSVDFTNTIDAKITQLPVVYKALSSGASITRG

MON863.pap PPTGONLFLKSSNSIAKFKVTLNSAALLQRYVRIRYASTNLRFLVQNSNDPLVI  
S10228 PPTGONLFLKSSNSIAKFKVTLNSAALLQRYVRIRYASTNLRFLVQNSNDPLVI

MON863.pap YINKTMKDDLTQYCTPLATTNMGFSDGNELIAGSPVNSKIYIDKIEFIPVQ  
S10228 YINKTMKDDLTQYCTPLATTNMGFSDGNELIAGSPVNSKIYIDKIEFIPVQ

MON863.pap TO CR70\_BACTT  
S10228 TO CR70\_BACTT

ID CR70\_BACTT STANDARD, PRT. 644 AA.  
AC P07130; P21255;  
DT 01-APR-1988 (REL. 07, CREATED)  
DT 01-APR-1988 (REL. 07, LAST SEQUENCE UPDATE)  
DT 01-OCT-1995 (REL. 34, LAST ANNOTATION UPDATE)  
DE 70 KD CRYSTAL PROTEIN (DELTA ENDOTOXIN) (CRYSTALLINE ENTOMOCIDAL . . .

SCORES Initl 1640 Inltn 2941 Opt 2972 z-score: 2948.6 E(): 8.9e-159  
Smith-Waterman score: 2972, 68.4% identity in 649 aa overlap

MON863.pap MANPNRSEHDITKVTNPSRLQTNHNOYPLADNPSTLEELNKEFLAMTDEDS  
CR70\_BACTT MIRMGRKNNPNRSEHDITKVTNPSRLQTNHNOYPLADNPSTLEELNKEFLAMTADN

MON863.pap STVKDAVGTGIVVGGIQLGVVGFAGALTSFYQSFLNTWPSDADPWKAFMAQVEVLID  
CR70\_BACTT STVKDAVGTGIVVGGIQLGVVGFAGALTSFYQSFLNTWPSDADPWKAFMAQVEVLID

MON863.pap KKIEEYAKSALAEQLQNNFEDVYNALDSWKKAPVNLRSRQDRIRLFSQASHPFR  
CR70\_BACTT KKIEEYAKSALAEQLQNNFEDVYNALDSWKKAPVNLRSRQDRIRLFSQASHPFR

MON863.pap NSMPSFAVSKFEVLPTTAAQANTHLLKDAQVGEWYSSDVAEYFRQLKLPQ  
CR70\_BACTT NSMPSFAVSKFEVLPTTAAQANTHLLKDAQVGEWYSSDVAEYFRQLKLPQ

MON863.pap YTDHCVNHYVGLNLRGSTDYDAMKFNRRREMTLVLDLVLFPFYDRLVSKGVKTE  
CR70\_BACTT YTDHCVNHYVGLNLRGSTDYDAMKFNRRREMTLVLDLVLFPFYDRLVSKGVKTE

MON863.pap LTRDIPDIPFLTLQKQYPTFLSISIRKPHLPDVLQIEFHTLRAPYFQKDSHNY  
CR70\_BACTT LTRDIPDIPFLTLQKQYPTFLSISIRKPHLPDVLQIEFHTLRAPYFQKDSHNY

MON863.pap WSGNYVETRPISGSKTISFFYQDKSTEPVQKLSFDQKVRTIANTOVAAMPKLYL  
CR70\_BACTT WSGNYVETRPISGSKTISFFYQDKSTEPVQKLSFDQKVRTIANTOVAAMPKLYL

MON863.pap GVTKVFSDQDQKNETSTQYDSKRNKGVSAQDQLPPTTDEPLEKAYSHQALYA  
CR70\_BACTT GVTKVFSDQDQKNETSTQYDSKRNKGVSAQDQLPPTTDEPLEKAYSHQALYA

MON863.pap BCFMQRDRGTIPFFTWHSVDFTNTIDAKITQLPVVYKALSSGASITRG  
CR70\_BACTT BCFMQRDRGTIPFFTWHSVDFTNTIDAKITQLPVVYKALSSGASITRG

MON863.pap LLFLKSSNSIAKFKVTLNSAALLQRYVRIRYASTNLRFLVQNSNDPLVI  
CR70\_BACTT LLFLKSSNSIAKFKVTLNSAALLQRYVRIRYASTNLRFLVQNSNDPLVI

MON863.pap XQDLTQYCTPLATTNMGFSDGNELIAGSPVNSKIYIDKIEFIPVQ  
CR70\_BACTT XQDLTQYCTPLATTNMGFSDGNELIAGSPVNSKIYIDKIEFIPVQ

MON863.pap TO BACRYC\_1  
CR70\_BACTT TO BACRYC\_1

LOCUS BACRYC\_1  
DEFINITION *Bacillus thuringiensis* delta-endotoxin, 5' end,  
DATE 15-JUN-1989  
ACCESSION M78472  
MID g14793 . . .

SCORES Initl 1640 Inltn 2941 Opt 2972 z-score: 2948.6 E(): 9e-159  
Smith-Waterman score: 2972, 68.4% identity in 649 aa overlap

MON863.pap MANPNRSEHDITKVTNPSRLQTNHNOYPLADNPSTLEELNKEFLAMTDEDS  
BACRYC\_1 MIRMGRKNNPNRSEHDITKVTNPSRLQTNHNOYPLADNPSTLEELNKEFLAMTADN

MON863.pap STEVLNSTVKDAVGTGIVVGGIQLGVVGFAGALTSFYQSFLNTWPSDADPWKAFMA  
BACRYC\_1 STEVLNSTVKDAVGTGIVVGGIQLGVVGFAGALTSFYQSFLNTWPSDADPWKAFMA





Page 49 of 102

[illegible]

240 250 260 270 280 290  
MON863.psp Q03749  
AC Q03749.01 PRELIMINARY. PRT. 1138 AA.  
DT 01-NOV-1996 (TREMBL) 01. CREATED  
DT 01-NOV-1996 (TREMBL) 01. LAST SEQUENCE UPDATE  
DT 01-FEB-1997 (TREMBL) 02. LAST ANNOTATION UPDATE  
DE 129 KD CRYSTAL PROTEIN (DELTA ENDOTOXIN) (CRYSTALLINE ENTOMOCIDAL)  
SCORES Initl: 546 Initl: 1246 Opt: 1419 z-score: 1405.6 E(1) 7 9e-73  
Smith-Waterman score: 1419; 37.6% identity in 651 aa overlap  
10 20 30 40 50 60  
MON863.psp MANPNRSHSDITKVPNSELQTHNNQVPLADNPSTLERLKYKFLMHTDSSTVELDN  
Q03749 MNLNLDVSDSRTLNLSLN-TPYQKALSPKLNQYQDFLSTTERBOPALAS  
10 20 30 40 50  
MON863.psp STVKADVGTGISVVGQIGVGVVFPAGALSTPQSFNTMPSDADPKAFMAQVEVLID  
Q03749 GNT--AINTVVSVTGATGAGLVGASITIPFLKIAQGLWENKIDSEFTVEALID  
60 70 80 90 100 110  
MON863.psp KKIEEYAKKALRLQGLQNNFEDYVVALNSN---KKTPLSLSKRSQGRIRKELFSQAES  
Q03749 QKIEEYVVKKALRLQGLQNNFEDYVVALNSN---KKTPLSLSKRSQGRIRKELFSQAES  
120 130 140 150 160 170  
MON863.psp KRPNSPSPAVSKFVLPLPTQAQANLHLLKDAQVQGEWYSEDAEYFRQLKL  
Q03749 LRFNSPSPAVSKFVLPLPTQAQANLHLLKDAQVQGEWYSEDAEYFRQLKL  
170 180 190 200 210 220

240 250 260 270 280 290  
MON863.psp Q03749  
Q03749 TQQTTHCVNMYNVLNGLRGSTYDANKFNFRFRENMTLVLDLVLFPFYDIRLYSKGV  
ISEYSDHCTKWNYSGLSRNLGSTYEQNINYNFRFRENMTLVLDLVLFPFYDIRLYSKGV  
230 240 250 260 270 280  
MON863.psp Q03749  
Q03749 KTELTRDIPFDIPFLTLQKYGPTFLSIENS-IRKPHLPDYLGQIEFHTRLAPGV---F  
STQLRKRVITDPSVLSIENPDIGSPQMENTAITHPLVLDVLELYTYSKYKAFSHEI  
290 300 310 320 330 340  
MON863.psp Q03749  
Q03749 OKDSFNWNGNYVETSPSISGSKTITSPFYQDKSTEPVK--LSPDQKVVRTIANTOVA  
QDPLF-YNSAHKVSFKK-RQSNLYTTGIQ-KTSGISSGAYSPHNDIYRTLAAPSVV  
350 360 370 380 390 400 410  
MON863.psp Q03749  
Q03749 AMPMKVYLQVTKVDFSGYDDQKNETSTQYDSKRNHVSQAQSDIDLPRTTDEFLK  
VYPTQNY-GVQVSEYGVVGHVHYGDKNYDL-----TYSDIDQLPPD--GEPIHE  
410 420 430 440 450 460 470  
MON863.psp Q03749  
Q03749 AYEMOLNVAECFLNQ--DR-ROTIPTFTWTHASVDFPNTIDAEKITQLPVVKAVALSQA  
KTHRLCHATAIPSTPDYDNATIPFQWTHASVDFPNTIDAEKITQLPVVKAVALSQA  
480 490 500 510 520 530  
MON863.psp Q03749  
Q03749 EILBPOPTQGNLLFLKESNSIAKFKVTLNSAALLQRYVRYASTNKLKLFVQVNSN  
TVVKGPGTQGDVL-KRSTGYTGDIKATVNSP-LSQYRVRYATVNSQGFVYINDK  
540 550 560 570 580 590  
MON863.psp Q03749  
Q03749 NMFVLVIYN--KTNKDDDTLQTFDLATNNSNGSGDKNELIQAESFVNSKEIYDK  
ITLQKQNTVSTIGKXDLTYSQFYVSTTIQFDEPHKITHLSDLSNSFVYDS  
600 610 620 630 640  
MON863.psp Q03749  
Q03749 IREFPVQL  
IREFPVQVYVAKRKKLKAQAVNTLFTSGNALQKQVTDYKQGVSLVDCISDGLYN  
640 650 660 670 680 690  
MON863.psp Q03749  
Q03749 TO: A07236\_1  
LOCUS A07236\_1  
DEFINITION B.thuringiensis (strain P081245) gene  
DATE 04-OCT-1993  
ACCESSION A07236  
VERSION 490177  
ORGANISM Bacillus thuringiensis . . .  
SCORES Initl: 546 Initl: 1246 Opt: 1419 z-score: 1405.6 E(1) 7 9e-73  
Smith-Waterman score: 1419; 37.6% identity in 651 aa overlap  
10 20 30 40 50 60  
MON863.psp MANPNRSHSDITKVPNSELQTHNNQVPLADNPSTLERLKYKFLMHTDSSTVELDN  
A07236\_1 MNLNLDVSDSRTLNLSLN-TPYQKALSPKLNQYQDFLSTTERBOPALAS  
10 20 30 40 50  
MON863.psp STVKADVGTGISVVGQIGVGVVFPAGALSTPQSFNTMPSDADPKAFMAQVEVLID  
A07236\_1 GNT--AINTVVSVTGATGAGLVGASITIPFLKIAQGLWENKIDSEFTVEALID  
60 70 80 90 100 110  
MON863.psp KKIEEYAKKALRLQGLQNNFEDYVVALNSN---KKTPLSLSKRSQGRIRKELFSQAES  
A07236\_1 QKIEEYVVKKALRLQGLQNNFEDYVVALNSN---KKTPLSLSKRSQGRIRKELFSQAES  
120 130 140 150 160 170  
MON863.psp KRPNSPSPAVSKFVLPLPTQAQANLHLLKDAQVQGEWYSEDAEYFRQLKL  
A07236\_1 LRFNSPSPAVSKFVLPLPTQAQANLHLLKDAQVQGEWYSEDAEYFRQLKL  
170 180 190 200 210 220

240 250 260 270 280 290  
MON863.psp A07236\_1  
A07236\_1 TQQTTHCVNMYNVLNGLRGSTYDANKFNFRFRENMTLVLDLVLFPFYDIRLYSKGV  
ISEYSDHCTKWNYSGLSRNLGSTYEQNINYNFRFRENMTLVLDLVLFPFYDIRLYSKGV  
230 240 250 260 270 280  
MON863.psp A07236\_1  
A07236\_1 KTELTRDIPFDIPFLTLQKYGPTFLSIENS-IRKPHLPDYLGQIEFHTRLAPGV---F  
STQLRKRVITDPSVLSIENPDIGSPQMENTAITHPLVLDVLELYTYSKYKAFSHEI  
290 300 310 320 330 340  
MON863.psp A07236\_1  
A07236\_1 OKDSFNWNGNYVETSPSISGSKTITSPFYQDKSTEPVK--LSPDQKVVRTIANTOVA  
QDPLF-YNSAHKVSFKK-RQSNLYTTGIQ-KTSGISSGAYSPHNDIYRTLAAPSVV  
350 360 370 380 390 400 410  
MON863.psp A07236\_1  
A07236\_1 AMPMKVYLQVTKVDFSGYDDQKNETSTQYDSKRNHVSQAQSDIDLPRTTDEFLK  
VYPTQNY-GVQVSEYGVVGHVHYGDKNYDL-----TYSDIDQLPPD--GEPIHE  
410 420 430 440 450 460 470  
MON863.psp A07236\_1  
A07236\_1 AYEMOLNVAECFLNQ--DR-ROTIPTFTWTHASVDFPNTIDAEKITQLPVVKAVALSQA  
KTHRLCHATAIPSTPDYDNATIPFQWTHASVDFPNTIDAEKITQLPVVKAVALSQA  
480 490 500 510 520 530  
MON863.psp A07236\_1  
A07236\_1 EILBPOPTQGNLLFLKESNSIAKFKVTLNSAALLQRYVRYASTNKLKLFVQVNSN  
TVVKGPGTQGDVL-KRSTGYTGDIKATVNSP-LSQYRVRYATVNSQGFVYINDK  
540 550 560 570 580 590  
MON863.psp A07236\_1  
A07236\_1 NMFVLVIYN--KTNKDDDTLQTFDLATNNSNGSGDKNELIQAESFVNSKEIYDK  
ITLQKQNTVSTIGKXDLTYSQFYVSTTIQFDEPHKITHLSDLSNSFVYDS  
600 610 620 630 640  
MON863.psp A07236\_1  
A07236\_1 IREFPVQL  
IREFPVQVYVAKRKKLKAQAVNTLFTSGNALQKQVTDYKQGVSLVDCISDGLYN  
640 650 660 670 680 690  
MON863.psp A07236\_1  
A07236\_1 TO: BTH04368\_1  
LOCUS BTH04368\_1  
DEFINITION Bacillus thuringiensis kumamotoensis HD867 CryIII delta-  
endotoxin  
DATE 27-AUG-1994  
ACCESSION BTH04368  
VERSION 490177  
ORGANISM Bacillus thuringiensis . . .  
SCORES Initl: 546 Initl: 1259 Opt: 1416 z-score: 1402.6 E(1) 1.2e-72  
Smith-Waterman score: 1436; 38.1% identity in 649 aa overlap  
10 20 30 40 50 60  
MON863.psp MANPNRSHSDITKVPNSELQTHNNQVPLADNPSTLERLKYKFLMHTDSSTVELDN  
BTH04368\_1 MNLNLDVSDSRTLNLSLN-TPYQKALSPKLNQYQDFLSTTERBOPALAS  
10 20 30 40 50  
MON863.psp STVKADVGTGISVVGQIGVGVVFPAGALSTPQSFNTMPSDADPKAFMAQVEVLID  
BTH04368\_1 GNT--AINTVVSVTGATGAGLVGASITIPFLKIAQGLWENKIDSEFTVEALID  
60 70 80 90 100 110 120  
MON863.psp KKIEEYAKKALRLQGLQNNFEDYVVALNSN---KKTPLSLSKRSQGRIRKELFSQAES  
BTH04368\_1 QKIEEYVVKKALRLQGLQNNFEDYVVALNSN---KKTPLSLSKRSQGRIRKELFSQAES  
120 130 140 150 160 170 180  
MON863.psp KRPNSPSPAVSKFVLPLPTQAQANLHLLKDAQVQGEWYSEDAEYFRQLKL  
BTH04368\_1 LRFNSPSPAVSKFVLPLPTQAQANLHLLKDAQVQGEWYSEDAEYFRQLKL  
170 180 190 200 210 220



Page 51 of 102

[illegible]

## Appendix 4

Study No. 01-01-39-23

MSL-17140

Page 52 of 102





120 130 140 150 160 170 179  
MON863.pcp DKKIEYAKKALAELOQNNFEDYVNAISWKKTPLSLRKRSQGRIRLFSQAESHF  
S25383 NOKISTYARKKALADGLOGLALAVHDSLSWVG---NNNTTRARSVVKSOYIALELMF  
120 130 140 150 160  
180 190 200 210 220 230 239  
MON863.pcp RNSMPSFAVEKFEVLPTTYAQAANTHLLKDAQVGSSEWVSSEDAFEPYRROLKLTQ  
S25383 VOKLPSFAVSEVEVPLPTTYAQAANTHLLKDAQVGSSEWVSSEDAFEPYRROLKLTQ  
170 180 190 200 210 220 230  
240 250 260 270 280 290 299  
MON863.pcp QYTDRCVNMVYVGLNGLAGSTYDAWVKPFRPRENTLVLDLIVFPFYDIRLYSKGVKT  
S25383 DYSVHCVMYVGLNGLAGSTYDAWVKPFRPRENTLVLDLIVFPFYDIRLYSKGVKT  
230 240 250 260 270 280 290  
300 310 320 330 340 349  
MON863.pcp ELTRDIFTD-----PIFLTT-LQKYQPTPLSIENSI-RKPHLFYDLOIEFPHRLR  
S25383 ELTRDIFTD-----PIFLTT-LQKYQPTPLSIENSI-RKPHLFYDLOIEFPHRLR  
290 300 310 320 330 340 349  
350 360 370 380 390 400 409  
MON863.pcp POYFGDSNPNVWNTVTPSPISGSKTITSFPYQKSTEPVOKLSFDQKVTYRTIAMD  
S25383 R-WNTQVNMWNGHKLKFA-TIGOTLNISTQSTNTSINPT-LPTSRDVTYRTESLAG  
350 360 370 380 390 400 409  
410 420 430 440 450 460 460  
MON863.pcp VAAWPMKVYLVGVDFPS-QYDQKNETSTQYDSKNNHVSQAQSDIGLPPETDEP  
S25383 LNLFLTOPVN-GVPRVDFHMFVTHPIADNFYPOYAGIG-TOLQDSNELPPEATQGP  
410 420 430 440 450 460 460  
470 480 490 500 510 520 520  
MON863.pcp LKAYSHQNLVACFLMDRAGTIPFPTWTHRSVDFPTIDAKITQLVVKAVALSSGA  
S25383 NYESYSHLCHIGLSAHSVKAIV--YEWTHRADRTWTEPNTSITQIPVKAVALSSGA  
470 480 490 500 510 520 520  
530 540 550 560 570 580 580  
MON863.pcp SIIBGPOPTGNNLFLKSSNSIAKVTLSAALLQYRVIRYASTNLRLFVQNSNN  
S25383 AVVKGPOPTGNNLFLKSSNSIAKVTLSAALLQYRVIRYASTNLRLFVQNSNN  
530 540 550 560 570 580 580  
590 600 610 620 630 640 640  
MON863.pcp DFLVIYINKTKNDLDTYQTFDLATTSNMGVSGDKNELIQAESVVMKXIYDKIKYF  
S25383 AINQNFSAVNRGDELDTYQTFDLATTSNMGVSGDKNELIQAESVVMKXIYDKIKYF  
590 600 610 620 630 640 640  
650  
MON863.pcp IPVQL  
S25383 VVVEVYVYAEVDFKACQKVTALFTSTNPRGLKTDVQYKHIDQVSNLVESLSEDFYLDK  
650 660 670 680 690 700  
MON863.pcp TO BTCTRVV\_1  
LOCUS BTCTRVV\_1  
DEFINITION B.thuringiensis cryV gene for delta-endotoxin  
DATE 19-MAY-1992  
ACCESSION X62821  
KEYWORDS g10289  
ORGANISM Bacillus thuringiensis . . .  
SCORES Initl: 517 Inntn: 1265 Opt: 1320 z-score 1311.0 E(): 1.5e-67  
Smith-Waterman score 1434; 37.7% identity in 636 aa overlap  
10 20 30 40 50 59  
MON863.pcp MANPNRNSHDTIKVTPNSGLQTNQYPLADNPSTLRLNKLKELKTEGDSSTV-VLD  
S25383 MANPNRNSHDTIKVTPNSGLQTNQYPLADNPSTLRLNKLKELKTEGDSSTV-VLD  
BTCTRVV\_1 MKLNQDKQSPSSNAKVXKISTDLSKNTDLELQNIHNSDKHSEKVEVPPVS  
10 20 30 40 50 59  
60 70 80 90 100 110 119  
MON863.pcp NSTVKDAVQGISVVGQILGVVVPFAGALTSFYQSLMTIMPSEDADPKAFACQVLEI  
S25383 NSTVKDAVQGISVVGQILGVVVPFAGALTSFYQSLMTIMPSEDADPKAFACQVLEI  
BTCTRVV\_1 NSTVKDAVQGISVVGQILGVVVPFAGALTSFYQSLMTIMPSEDADPKAFACQVLEI  
60 70 80 90 100 110 119

120 130 140 150 160 170 179  
MON863.pcp DKKIEYAKKALAELOQNNFEDYVNAISWKKTPLSLRKRSQGRIRLFSQAESHF  
BTCTRVV\_1 NOKISTYARKKALADGLOGLALAVHDSLSWVG---NNNTTRARSVVKSOYIALELMF  
120 130 140 150 160  
180 190 200 210 220 230 239  
MON863.pcp RNSMPSFAVEKFEVLPTTYAQAANTHLLKDAQVGSSEWVSSEDAFEPYRROLKLTQ  
BTCTRVV\_1 VOKLPSFAVSEVEVPLPTTYAQAANTHLLKDAQVGSSEWVSSEDAFEPYRROLKLTQ  
170 180 190 200 210 220 230  
240 250 260 270 280 290 299  
MON863.pcp QYTDRCVNMVYVGLNGLAGSTYDAWVKPFRPRENTLVLDLIVFPFYDIRLYSKGVKT  
BTCTRVV\_1 DYSVHCVMYVGLNGLAGSTYDAWVKPFRPRENTLVLDLIVFPFYDIRLYSKGVKT  
240 250 260 270 280 290 299  
300 310 320 330 340 349  
MON863.pcp ELTRDIFTD-----PIFLTT-LQKYQPTPLSIENSI-RKPHLFYDLOIEFPHRLR  
BTCTRVV\_1 ELTRDIFTD-----PIFLTT-LQKYQPTPLSIENSI-RKPHLFYDLOIEFPHRLR  
290 300 310 320 330 340 349  
350 360 370 380 390 400 409  
MON863.pcp POYFGDSNPNVWNTVTPSPISGSKTITSFPYQKSTEPVOKLSFDQKVTYRTIAMD  
BTCTRVV\_1 R-WNTQVNMWNGHKLKFA-TIGOTLNISTQSTNTSINPT-LPTSRDVTYRTESLAG  
350 360 370 380 390 400 409  
410 420 430 440 450 460 460  
MON863.pcp VAAWPMKVYLVGVDFPS-QYDQKNETSTQYDSKNNHVSQAQSDIGLPPETDEP  
BTCTRVV\_1 LNLFLTOPVN-GVPRVDFHMFVTHPIADNFYPOYAGIG-TOLQDSNELPPEATQGP  
410 420 430 440 450 460 460  
470 480 490 500 510 520 520  
MON863.pcp LKAYSHQNLVACFLMDRAGTIPFPTWTHRSVDFPTIDAKITQLVVKAVALSSGA  
BTCTRVV\_1 NYESYSHLCHIGLSAHSVKAIV--YEWTHRADRTWTEPNTSITQIPVKAVALSSGA  
470 480 490 500 510 520 520  
530 540 550 560 570 580 580  
MON863.pcp SIIBGPOPTGNNLFLKSSNSIAKVTLSAALLQYRVIRYASTNLRLFVQNSNN  
BTCTRVV\_1 AVVKGPOPTGNNLFLKSSNSIAKVTLSAALLQYRVIRYASTNLRLFVQNSNN  
530 540 550 560 570 580 580  
590 600 610 620 630 640 640  
MON863.pcp DFLVIYINKTKNDLDTYQTFDLATTSNMGVSGDKNELIQAESVVMKXIYDKIKYF  
BTCTRVV\_1 AINQNFSAVNRGDELDTYQTFDLATTSNMGVSGDKNELIQAESVVMKXIYDKIKYF  
590 600 610 620 630 640 640  
650  
MON863.pcp IPVQL  
BTCTRVV\_1 VVVEVYVYAEVDFKACQKVTALFTSTNPRGLKTDVQYKHIDQVSNLVESLSEDFYLDK  
650 660 670 680 690 700  
MON863.pcp TO Q45750  
ID Q45750 PRELIMINARY. PRT: 719 AA.  
AC Q45750; DT 01-NOV-1996 (TRMBREL 01, CREATED)  
DT 01-NOV-1996 (TRMBREL 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1996 (TRMBREL 01, LAST ANNOTATION UPDATE)  
DR DELTA-ENDOTOXIN. . . .  
SCORES Initl: 517 Inntn: 1265 Opt: 1320 z-score 1311.0 E(): 1.5e-67  
Smith-Waterman score 1434; 37.7% identity in 636 aa overlap  
10 20 30 40 50 59  
MON863.pcp MANPNRNSHDTIKVTPNSGLQTNQYPLADNPSTLRLNKLKELKTEGDSSTV-VLD  
Q45750 MANPNRNSHDTIKVTPNSGLQTNQYPLADNPSTLRLNKLKELKTEGDSSTV-VLD  
10 20 30 40 50 59  
60 70 80 90 100 110 119  
MON863.pcp NSTVKDAVQGISVVGQILGVVVPFAGALTSFYQSLMTIMPSEDADPKAFACQVLEI  
Q45750 NSTVKDAVQGISVVGQILGVVVPFAGALTSFYQSLMTIMPSEDADPKAFACQVLEI  
60 70 80 90 100 110 119

120 130 140 150 160 170 179  
MON863.pcp DKKIEYAKKALAELOQNNFEDYVNAISWKKTPLSLRKRSQGRIRLFSQAESHF  
Q45750 NOKISTYARKKALADGLOGLALAVHDSLSWVG---NNNTTRARSVVKSOYIALELMF  
120 130 140 150 160  
180 190 200 210 220 230 239  
MON863.pcp RNSMPSFAVEKFEVLPTTYAQAANTHLLKDAQVGSSEWVSSEDAFEPYRROLKLTQ  
Q45750 VOKLPSFAVSEVEVPLPTTYAQAANTHLLKDAQVGSSEWVSSEDAFEPYRROLKLTQ  
170 180 190 200 210 220 230  
240 250 260 270 280 290 299  
MON863.pcp QYTDRCVNMVYVGLNGLAGSTYDAWVKPFRPRENTLVLDLIVFPFYDIRLYSKGVKT  
Q45750 DYSVHCVMYVGLNGLAGSTYDAWVKPFRPRENTLVLDLIVFPFYDIRLYSKGVKT  
240 250 260 270 280 290 299  
300 310 320 330 340 349  
MON863.pcp ELTRDIFTD-----PIFLTT-LQKYQPTPLSIENSI-RKPHLFYDLOIEFPHRLR  
Q45750 ELTRDIFTD-----PIFLTT-LQKYQPTPLSIENSI-RKPHLFYDLOIEFPHRLR  
290 300 310 320 330 340 349  
350 360 370 380 390 400 409  
MON863.pcp POYFGDSNPNVWNTVTPSPISGSKTITSFPYQKSTEPVOKLSFDQKVTYRTIAMD  
Q45750 R-WNTQVNMWNGHKLKFA-TIGOTLNISTQSTNTSINPT-LPTSRDVTYRTESLAG  
350 360 370 380 390 400 409  
410 420 430 440 450 460 460  
MON863.pcp VAAWPMKVYLVGVDFPS-QYDQKNETSTQYDSKNNHVSQAQSDIGLPPETDEP  
Q45750 LNLFLTOPVN-GVPRVDFHMFVTHPIADNFYPOYAGIG-TOLQDSNELPPEATQGP  
410 420 430 440 450 460 460  
470 480 490 500 510 520 520  
MON863.pcp LKAYSHQNLVACFLMDRAGTIPFPTWTHRSVDFPTIDAKITQLVVKAVALSSGA  
Q45750 NYESYSHLCHIGLSAHSVKAIV--YEWTHRADRTWTEPNTSITQIPVKAVALSSGA  
470 480 490 500 510 520 520  
530 540 550 560 570 580 580  
MON863.pcp SIIBGPOPTGNNLFLKSSNSIAKVTLSAALLQYRVIRYASTNLRLFVQNSNN  
Q45750 AVVKGPOPTGNNLFLKSSNSIAKVTLSAALLQYRVIRYASTNLRLFVQNSNN  
530 540 550 560 570 580 580  
590 600 610 620 630 640 640  
MON863.pcp DFLVIYINKTKNDLDTYQTFDLATTSNMGVSGDKNELIQAESVVMKXIYDKIKYF  
Q45750 AINQNFSAVNRGDELDTYQTFDLATTSNMGVSGDKNELIQAESVVMKXIYDKIKYF  
590 600 610 620 630 640 640  
650  
MON863.pcp IPVQL  
Q45750 VVVEVYVYAEVDFKACQKVTALFTSTNPRGLKTDVQYKHIDQVSNLVESLSEDFYLDK  
650 660 670 680 690 700  
MON863.pcp TO BACRYIE\_1  
LOCUS BACRYIE\_1  
DEFINITION Bacillus thuringiensis crystal protein (cryI Bts) gene, complete  
cda  
DATE 25-APR-1994  
ACCESSION L32026. 790 g41899 . . .  
SCORES Initl: 498 Inntn: 1303 Opt: 1322 z-score 1308.8 E(): 1.9e-67  
Smith-Waterman score 1447; 38.2% identity in 641 aa overlap  
10 20 30 40 50 59  
MON863.pcp MANPNRNSHDTIKVTPNSGLQTNQYPLADNPSTLRLNKLKELKTEGDSSTV-VLD  
BACRYIE\_1 MANPNRNSHDTIKVTPNSGLQTNQYPLADNPSTLRLNKLKELKTEGDSSTV-VLD  
10 20 30 40 50 59  
60 70 80 90 100 110 119  
MON863.pcp NSTVKDAVQGISVVGQILGVVVPFAGALTSFYQSLMTIMPSEDADPKAFACQVLEI  
BACRYIE\_1 NSTVKDAVQGISVVGQILGVVVPFAGALTSFYQSLMTIMPSEDADPKAFACQVLEI  
60 70 80 90 100 110 119

## Appendix 4

Study No. 01-01-39-23

MSL-17140

Page 55 of 102



MON863.pap 120 130 140 150 160 170  
EVLDIKKIERAKKALAEQGLQNNFEDYVNALNSMKTPLSLRKSKQRIRLEFSGA  
Q45706 120 130 140 150 160 170  
EELIDQKILDSVRSRAIDANSRAIVRYONALEDRKNP--HSTRSAALVKERFGNA

MON863.pap 180 190 200 210 220 230  
ESHFRNSMPSPFAVSKFEVLPTTAAQAANTHLLLLKDAQVQGEWYSSEDAVFYRRL  
Q45706 180 190 200 210 220 230  
EALRATNMGSPSOTNVTETPLPYTAAASLLHVRDVQYVQKMGYPONDIDLFYKEGV

MON863.pap 240 250 260 270 280 290  
KLTOGYTDCVNNYVGLNGLGSTYDAWVKFNRFRREKTLVLDLIVLFPFYDRLYSK  
Q45706 240 250 260 270 280 290  
SYTARYSDHCQVYVGLNGLGSTYDAWVKFNRFRREKTLVLDLIVLFPFYDRLYSK

MON863.pap 300 310 320 330 340  
GVKTELTRDITDPTPLTLQK-----YQ-----PTFLSIENIKKPHLFDYLOGIE  
Q45706 300 310 320 330 340  
ETWALTRDITDPTPLTLQK-----YQ-----PTFLSIENIKKPHLFDYLOGIE

MON863.pap 350 360 370 380 390 400  
FHTRLRPGYPOK-DSFNWAGNVVTRSPISGSS-KTITSPFYDGKSTFVKLEFDOOK-  
Q45706 350 360 370 380 390 400  
MYTSFRQN--GTISYVNWGOORLALSYVIGSSFNKYSVLGAGAEIIPV-----GQND

MON863.pap 410 420 430 440 450 460  
VYRTIAMTDAAMPKGVYLVGVKVSQYDQKNETSTQYDSEKRNKHVSQAQSDIDOL  
Q45706 410 420 430 440 450 460  
TYRVVV-TYIGRYTNS--LLGVNPTF-YFSNNTQKTSYK--KQFAGIKTIDISCREL

MON863.pap 470 480 490 500 510  
PPSTDEPLEKAYSGQVLAECFLNQDRGTI-----PFTVTHRSVDFPMIDAEKIDOL  
Q45706 470 480 490 500 510  
TYRYV--GTISYVNWGOORLALSYVIGSSFNKYSVLGAGAEIIPV-----GQND

MON863.pap 520 530 540 550 560 570  
PVVYALASOASISLQPO-FTQGNLLFLKESNSTAKFVYVNSAALQRYVIRYAS  
Q45706 520 530 540 550 560 570  
PINKASHTSGAGVNTQEGLYNGQPVKLDGSGGVNLRVATDAKAGARQYRIRYAS

MON863.pap 580 590 600 610 620 629  
TTNRLRFV-----QNSNDLVLYINKNKDDDTYQTPDLATNS--NMGFSDGNMELI  
Q45706 580 590 600 610 620 629  
DRAGKFTISGRAPENPATYSASIAATYNTMNASLTYSTAYARSOPINLQIGSGRTFD

MON863.pap 630 640 650  
IGABEFVNSKIIYIDKTEFPVQL  
Q45706 630 640 650  
ISTKBAQANLVDRIFPVNTLFBABEDLDVAKKAVNGLFTNKKALQTSVQYOVN

MON863.pap  
TO:BTU04366\_1

LOCUS  
DEFINITION Bacillus thuringiensis japonensis subsp. CryII delta-endotoxin  
gene, complete cds.  
DATE 27-AUG-1994  
ACCESSION U04366.1 g532523

SCORES Initl: 569 Initn: 1057 Opt: 1289 z-score: 1276.6 E(): 1.2e-65  
Smith-Waterman score: 1378; 36.5% identity in 680 aa overlap

MON863.pap 10 20 30 40 50 60  
MANPNNSBSEDTIKVTPNELQTNHNPVPLADNPSTLSELYKEFLKTEDESTEVLON  
BTU04366\_1 10 20 30 40 50 60  
MSPNNQNEVEIDALSPTSVDNSIRYPLANDQTNLQNNYKYDLQNTSTNALSRLN

MON863.pap 70 80 90 100 110  
S---TVKDAVGTVGIVVGGIIGVVPFAGALTSFYQSLNTWESDAD-WKAFMAOV  
BTU04366\_1 70 80 90 100 110  
POTFISAGQAVGTGIDVETIISGLTGVKVSISLQSLGLLWSPNNHNVQIPNHRV

MON863.pap 120 130 140 150 160 170  
EVLDIKKIERAKKALAEQGLQNNFEDYVNALNSMKTPLSLRKSKQRIRLEFSGA  
BTU04366\_1 120 130 140 150 160 170  
EELIDQKILDSVRSRAIDANSRAIVRYONALEDRKNP--HSTRSAALVKERFGNA

MON863.pap 180 190 200 210 220 230  
ESHFRNSMPSPFAVSKFEVLPTTAAQAANTHLLLLKDAQVQGEWYSSEDAVFYRRL  
BTU04366\_1 180 190 200 210 220 230  
EALRATNMGSPSOTNVTETPLPYTAAASLLHVRDVQYVQKMGYPONDIDLFYKEGV

MON863.pap 240 250 260 270 280 290  
KLTOGYTDCVNNYVGLNGLGSTYDAWVKFNRFRREKTLVLDLIVLFPFYDRLYSK  
BTU04366\_1 240 250 260 270 280 290  
SYTARYSDHCQVYVGLNGLGSTYDAWVKFNRFRREKTLVLDLIVLFPFYDRLYSK

MON863.pap 300 310 320 330 340  
GVKTELTRDITDPTPLTLQK-----YQ-----PTFLSIENIKKPHLFDYLOGIE  
BTU04366\_1 300 310 320 330 340  
ETWALTRDITDPTPLTLQK-----YQ-----PTFLSIENIKKPHLFDYLOGIE

MON863.pap 350 360 370 380 390 400  
FHTRLRPGYPOK-DSFNWAGNVVTRSPISGSS-KTITSPFYDGKSTFVKLEFDOOK-  
BTU04366\_1 350 360 370 380 390 400  
MYTSFRQN--GTISYVNWGOORLALSYVIGSSFNKYSVLGAGAEIIPV-----GQND

MON863.pap 410 420 430 440 450 460  
VYRTIAMTDAAMPKGVYLVGVKVSQYDQKNETSTQYDSEKRNKHVSQAQSDIDOL  
BTU04366\_1 410 420 430 440 450 460  
TYRVVV-TYIGRYTNS--LLGVNPTF-YFSNNTQKTSYK--KQFAGIKTIDISCREL

MON863.pap 470 480 490 500 510  
PPSTDEPLEKAYSGQVLAECFLNQDRGTI-----PFTVTHRSVDFPMIDAEKIDOL  
BTU04366\_1 470 480 490 500 510  
TYRYV--GTISYVNWGOORLALSYVIGSSFNKYSVLGAGAEIIPV-----GQND

MON863.pap 520 530 540 550 560 570  
PVVYALASOASISLQPO-FTQGNLLFLKESNSTAKFVYVNSAALQRYVIRYAS  
BTU04366\_1 520 530 540 550 560 570  
PINKASHTSGAGVNTQEGLYNGQPVKLDGSGGVNLRVATDAKAGARQYRIRYAS

MON863.pap 580 590 600 610 620 629  
TTNRLRFV-----QNSNDLVLYINKNKDDDTYQTPDLATNS--NMGFSDGNMELI  
BTU04366\_1 580 590 600 610 620 629  
DRAGKFTISGRAPENPATYSASIAATYNTMNASLTYSTAYARSOPINLQIGSGRTFD

MON863.pap 630 640 650  
IGABEFVNSKIIYIDKTEFPVQL  
BTU04366\_1 630 640 650  
ISTKBAQANLVDRIFPVNTLFBABEDLDVAKKAVNGLFTNKKALQTSVQYOVN

MON863.pap  
TO:140589

PI;140589 - cryII delta-endotoxin - Bacillus thuringiensis  
C.Species: Bacillus thuringiensis  
C.Date: 12-Aug-1996 Sequence revision 12-Aug-1996 Statx\_change 07-Feb-1997  
C.Accession: 140589  
R:Sato, R.; Takeuchi, K.; Ogihara, K.; Minami, N.; Kajii, Y.; Suzuki, N.,  
Hori,  
H.; Asano, S.; Ohba, M.; Iwahana, H.

SCORES Initl: 569 Initn: 1057 Opt: 1289 z-score: 1276.6 E(): 1.2e-65  
Smith-Waterman score: 1378; 36.5% identity in 680 aa overlap

MON863.pap 10 20 30 40 50 60  
MANPNNSBSEDTIKVTPNELQTNHNPVPLADNPSTLSELYKEFLKTEDESTEVLON  
I40589 10 20 30 40 50 60  
MSPNNQNEVEIDALSPTSVDNSIRYPLANDQTNLQNNYKYDLQNTSTNALSRLN

MON863.pap 70 80 90 100 110  
S---TVKDAVGTVGIVVGGIIGVVPFAGALTSFYQSLNTWESDAD-WKAFMAOV  
I40589 70 80 90 100 110  
POTFISAGQAVGTGIDVETIISGLTGVKVSISLQSLGLLWSPNNHNVQIPNHRV

MON863.pap 70 80 90 100 110  
S---TVKDAVGTVGIVVGGIIGVVPFAGALTSFYQSLNTWESDAD-WKAFMAOV  
I40589 70 80 90 100 110  
POTFISAGQAVGTGIDVETIISGLTGVKVSISLQSLGLLWSPNNHNVQIPNHRV

MON863.pap 120 130 140 150 160 170  
EVLDIKKIERAKKALAEQGLQNNFEDYVNALNSMKTPLSLRKSKQRIRLEFSGA  
I40589 120 130 140 150 160 170  
EELIDQKILDSVRSRAIDANSRAIVRYONALEDRKNP--HSTRSAALVKERFGNA

MON863.pap 180 190 200 210 220 230  
ESHFRNSMPSPFAVSKFEVLPTTAAQAANTHLLLLKDAQVQGEWYSSEDAVFYRRL  
I40589 180 190 200 210 220 230  
EALRATNMGSPSOTNVTETPLPYTAAASLLHVRDVQYVQKMGYPONDIDLFYKEGV

MON863.pap 240 250 260 270 280 290  
KLTOGYTDCVNNYVGLNGLGSTYDAWVKFNRFRREKTLVLDLIVLFPFYDRLYSK  
I40589 240 250 260 270 280 290  
SYTARYSDHCQVYVGLNGLGSTYDAWVKFNRFRREKTLVLDLIVLFPFYDRLYSK

MON863.pap 300 310 320 330 340  
GVKTELTRDITDPTPLTLQK-----YQ-----PTFLSIENIKKPHLFDYLOGIE  
I40589 300 310 320 330 340  
ETWALTRDITDPTPLTLQK-----YQ-----PTFLSIENIKKPHLFDYLOGIE

MON863.pap 350 360 370 380 390 400  
FHTRLRPGYPOK-DSFNWAGNVVTRSPISGSS-KTITSPFYDGKSTFVKLEFDOOK-  
I40589 350 360 370 380 390 400  
MYTSFRQN--GTISYVNWGOORLALSYVIGSSFNKYSVLGAGAEIIPV-----GQND

MON863.pap 410 420 430 440 450 460  
VYRTIAMTDAAMPKGVYLVGVKVSQYDQKNETSTQYDSEKRNKHVSQAQSDIDOL  
I40589 410 420 430 440 450 460  
TYRVVV-TYIGRYTNS--LLGVNPTF-YFSNNTQKTSYK--KQFAGIKTIDISCREL

MON863.pap 470 480 490 500 510  
PPSTDEPLEKAYSGQVLAECFLNQDRGTI-----PFTVTHRSVDFPMIDAEKIDOL  
I40589 470 480 490 500 510  
TYRYV--GTISYVNWGOORLALSYVIGSSFNKYSVLGAGAEIIPV-----GQND

MON863.pap 520 530 540 550 560 570  
PVVYALASOASISLQPO-FTQGNLLFLKESNSTAKFVYVNSAALQRYVIRYAS  
I40589 520 530 540 550 560 570  
PINKASHTSGAGVNTQEGLYNGQPVKLDGSGGVNLRVATDAKAGARQYRIRYAS

MON863.pap 580 590 600 610 620 629  
TTNRLRFV-----QNSNDLVLYINKNKDDDTYQTPDLATNS--NMGFSDGNMELI  
I40589 580 590 600 610 620 629  
DRAGKFTISGRAPENPATYSASIAATYNTMNASLTYSTAYARSOPINLQIGSGRTFD

MON863.pap  
TO:AF077326\_1

LOCUS  
DEFINITION Bacillus thuringiensis plasmid pNYC2383 stein BG158C2, CryIb-  
like delta-endotoxin gene, complete cds.  
DATE 31-JUL-1998  
ACCESSION AF077326  
MID g3360518

SCORES Initl: 499 Initn: 1088 Opt: 1188 z-score: 1171.2 E(): 9e-60  
Smith-Waterman score: 1295; 37.2% identity in 610 aa overlap

MON863.pap 10 20 30 40 50 60  
YPLADNPSTLSELYKEFLKTEDESTEVLONSKDAVNGVIVVGGIIGVVPFAG  
AF077326\_1 10 20 30 40 50 60  
IPAVVNSHAGNLSSTADTSDSLAKNDIDPFVNSAFVQVQVINTKRIKIGLVNHRV

MON863.pap 70 80 90 100 110  
YPLADNPSTLSELYKEFLKTEDESTEVLONSKDAVNGVIVVGGIIGVVPFAG  
AF077326\_1 70 80 90 100 110  
IPAVVNSHAGNLSSTADTSDSLAKNDIDPFVNSAFVQVQVINTKRIKIGLVNHRV





Page 58 of 102

MON863.pap 170 180 190 200 210 220  
RELFGAESHFRNSHPSFAVSKFVLFLPTAQAANTHLLKDAQVGEWYSSSDVA  
IC1Y Q-PFNKMSALTATPLAVQVPLSVVQAANHLHSLVLDVSPQGRWGFDAATIN  
230 240 250 260 270 280  
MON863.pap EFYRRLKLTQQTTHCVNMYVGLNGRGSTYDAWVKNFRPRRMTLTVLDLVLFPFY  
IC1Y Q-PFNKMSALTATPLAVQVPLSVVQAANHLHSLVLDVSPQGRWGFDAATIN  
290 300 310 320 330 340  
MON863.pap DIRLYSGVKVTELRDIFDIFLLTLQKVGPTT---LSIENSTRKPHLFDYLOQIEF  
IC1Y DSRHYPIRTVSQLTREIYTPV---LENFGDSFRMAQRIQNIQPHLMDLNLSITI  
350 360 370 380 390 400  
MON863.pap HTLRAPQYFGKDSFNYSNGNYVTRPSTGSEKTTTSPFYGDK-STEPVQKLFSDQKYYR  
IC1Y YDVHRO---FNYMGGHQTASPVGSGPEFAPFLFGAGNAAPVVLVSLTGLQIFR  
410 420 430 440 450 460  
MON863.pap TIANDT---VAAMPNGKVLGVTKVDSQYDDQKNETSTQYDEKRNHNSAQDSID  
IC1Y TLSSPLVRRILGSGPNNQBLFVLDGTEFSASLTNLPT-IV---RGRQTV---DGLD  
470 480 490 500 510 520  
MON863.pap QLPFETDEPLEKAYSHQIYACFLMQRGRTI---PFTVTHRSVDFPNTIDAEKI  
IC1Y VIFPQNGVPPRAGFSRLSHVT---MLSGAAGAVYTLRAFTSWHRSASFENNTIPSSQI  
530 540 550 560 570 580  
MON863.pap TQLPVVAYALSSGASIBOPQFTGNNLLFLKSSNSIAKPKVTLNSAALLQRYVRIRY  
IC1Y TQPLTKSNLGSQVGSVAGPFTGDLIR-KTSPQGLSTRVNI-TAPLSQRYVRIRY  
590 600 610 620 630 640  
MON863.pap ASTTNLALFVQNSNNDFLVIYINKTKNDDDLTYCTPDLATTNNMGFSQDKNELIGAR  
IC1Y APTTLQPTSTIDRPIQNGFSATMSSGNNLQSGSFRTVPTTFNFMSSSVFTLSAH  
650 660 670 680 690 700  
MON863.pap SFVSNKTYIDKIEFIPVGL  
IC1Y VFSNGNEVYIDRIFFVPAEVT  
710 720 730 740 750 760  
MON863.pap BTCRYAG\_1  
TO:JC2219  
P1.JC2219 - lepidopteran-specific protoxin - *Bacillus thuringiensis*  
C.Species: *Bacillus thuringiensis*  
C.Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 08-Sep-1997  
C.Accession: JC2219  
R.Udayasuriyan, V.; Nakamura, A.; Mori, H.; Masaki, H.; Uozumi, T.  
Bioneer Biotechnol. Biochem 58, 830-835, 1994  
SCORES In1: 350 In1n: 869 Opt: 1106 z-score: 1089.3 E(): 3.3e-55  
Smith-Waterman score: 1117, 34.1% identity in 590 aa overlap  
MON863.pap BSTEVLNSTRVKDAVGTGISVVGQILGVVGPVAGALTSFYQSFINTIN---PSDADFW  
JC2219 VEVLGGRSTRTVTPIDISLSLTQFLSEFVPGAG---FVLGLVDIINGIPQSQ---W  
110 120 130 140 150 160  
MON863.pap KAPMAQVVELDKKIEVYAKKALAEQLQGNFEDYVNAIWNKKTPLSRKRSQRI  
BTU43605\_1 DAPLVQQLINGRIFEFARNQALISLBLSNLQVYARSFREWADPTN-PALREEMRI  
170 180 190 200 210 220  
MON863.pap RELFGAESHFRNSHPSFAVSKFVLFLPTAQAANTHLLKDAQVGEWYSSSDVA  
BTU43605\_1 Q-PFNKMSALTATPLAVQVPLSVVQAANHLHSLVLDVSPQGRWGFDAATIN  
230 240 250 260 270 280  
MON863.pap EFYRRLKLTQQTTHCVNMYVGLNGRGSTYDAWVKNFRPRRMTLTVLDLVLFPFY  
BTCRYAG\_1 SRYNDLTRLIQNTDYAVRWYVGLRWVSPGSDRWVKNFRPRRMTLTVLDLVLFPFY

MON863.pap 230 240 250 260 270 280  
EFYRRLKLTQQTTHCVNMYVGLNGRGSTYDAWVKNFRPRRMTLTVLDLVLFPFY  
BTU43605\_1 SRYNDLTRLIQNTDYAVRWYVGLRWVSPGSDRWVKNFRPRRMTLTVLDLVLFPFY  
290 300 310 320 330 340  
MON863.pap DIRLYSGVKVTELRDIFDIFLLTLQKVGPTT---LSIENSTRKPHLFDYLOQIEF  
BTU43605\_1 DSRHYPIRTVSQLTREIYTPV---LENFGDSFRMAQRIQNIQPHLMDLNLSITI  
350 360 370 380 390 400  
MON863.pap HTLRAPQYFGKDSFNYSNGNYVTRPSTGSEKTTTSPFYGDK-STEPVQKLFSDQKYYR  
BTU43605\_1 YDVHRO---FNYMGGHQTASPVGSGPEFAPFLFGAGNAAPVVLVSLTGLQIFR  
410 420 430 440 450 460  
MON863.pap TIANDT---VAAMPNGKVLGVTKVDSQYDDQKNETSTQYDEKRNHNSAQDSID  
BTU43605\_1 TLSSPLVRRILGSGPNNQBLFVLDGTEFSASLTNLPT-IV---RGRQTV---DGLD  
470 480 490 500 510 520  
MON863.pap QLPFETDEPLEKAYSHQIYACFLMQRGRTI---PFTVTHRSVDFPNTIDAEKI  
BTU43605\_1 VIFPQNGVPPRAGFSRLSHVT---MLSGAAGAVYTLRAFTSWHRSASFENNTIPSSQI  
530 540 550 560 570 580  
MON863.pap TQLPVVAYALSSGASIBOPQFTGNNLLFLKSSNSIAKPKVTLNSAALLQRYVRIRY  
BTU43605\_1 TQPLTKSNLGSQVGSVAGPFTGDLIR-KTSPQGLSTRVNI-TAPLSQRYVRIRY  
590 600 610 620 630 640  
MON863.pap ASTTNLALFVQNSNNDFLVIYINKTKNDDDLTYCTPDLATTNNMGFSQDKNELIGAR  
BTU43605\_1 APTTLQPTSTIDRPIQNGFSATMSSGNNLQSGSFRTVPTTFNFMSSSVFTLSAH  
650 660 670 680 690 700  
MON863.pap SFVSNKTYIDKIEFIPVGL  
BTU43605\_1 VFSNGNEVYIDRIFFVPAEVT  
710 720 730 740 750 760  
MON863.pap BTCRYAG\_1  
TO:JC2219  
P1.JC2219 - lepidopteran-specific protoxin - *Bacillus thuringiensis*  
C.Species: *Bacillus thuringiensis*  
C.Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 08-Sep-1997  
C.Accession: JC2219  
R.Udayasuriyan, V.; Nakamura, A.; Mori, H.; Masaki, H.; Uozumi, T.  
Bioneer Biotechnol. Biochem 58, 830-835, 1994  
SCORES In1: 350 In1n: 869 Opt: 1106 z-score: 1089.3 E(): 3.3e-55  
Smith-Waterman score: 1117, 34.1% identity in 590 aa overlap  
MON863.pap BSTEVLNSTRVKDAVGTGISVVGQILGVVGPVAGALTSFYQSFINTIN---PSDADFW  
JC2219 VEVLGGRSTRTVTPIDISLSLTQFLSEFVPGAG---FVLGLVDIINGIPQSQ---W  
110 120 130 140 150 160  
MON863.pap KAPMAQVVELDKKIEVYAKKALAEQLQGNFEDYVNAIWNKKTPLSRKRSQRI  
JC2219 DAPLVQQLINGRIFEFARNQALISLBLSNLQVYARSFREWADPTN-PALREEMRI  
170 180 190 200 210 220  
MON863.pap RELFGAESHFRNSHPSFAVSKFVLFLPTAQAANTHLLKDAQVGEWYSSSDVA  
JC2219 Q-PFNKMSALTATPLAVQVPLSVVQAANHLHSLVLDVSPQGRWGFDAATIN  
230 240 250 260 270 280  
MON863.pap EFYRRLKLTQQTTHCVNMYVGLNGRGSTYDAWVKNFRPRRMTLTVLDLVLFPFY  
BTCRYAG\_1 SRYNDLTRLIQNTDYAVRWYVGLRWVSPGSDRWVKNFRPRRMTLTVLDLVLFPFY

MON863.pap 230 240 250 260 270 280  
EFYRRLKLTQQTTHCVNMYVGLNGRGSTYDAWVKNFRPRRMTLTVLDLVLFPFY  
JC2219 SRYNDLTRLIQNTDYAVRWYVGLRWVSPGSDRWVKNFRPRRMTLTVLDLVLFPFY  
290 300 310 320 330 340  
MON863.pap DIRLYSGVKVTELRDIFDIFLLTLQKVGPTT---LSIENSTRKPHLFDYLOQIEF  
JC2219 DSRHYPIRTVSQLTREIYTPV---LENFGDSFRMAQRIQNIQPHLMDLNLSITI  
350 360 370 380 390 400  
MON863.pap HTLRAPQYFGKDSFNYSNGNYVTRPSTGSEKTTTSPFYGDK-STEPVQKLFSDQKYYR  
JC2219 YDVHRO---FNYMGGHQTASPVGSGPEFAPFLFGAGNAAPVVLVSLTGLQIFR  
410 420 430 440 450 460  
MON863.pap TIANDT---VAAMPNGKVLGVTKVDSQYDDQKNETSTQYDEKRNHNSAQDSID  
JC2219 TLSSPLVRRILGSGPNNQBLFVLDGTEFSASLTNLPT-IV---RGRQTV---DGLD  
470 480 490 500 510 520  
MON863.pap QLPFETDEPLEKAYSHQIYACFLMQRGRTI---PFTVTHRSVDFPNTIDAEKI  
JC2219 VIFPQNGVPPRAGFSRLSHVT---MLSGAAGAVYTLRAFTSWHRSASFENNTIPSSQI  
530 540 550 560 570 580  
MON863.pap TQLPVVAYALSSGASIBOPQFTGNNLLFLKSSNSIAKPKVTLNSAALLQRYVRIRY  
JC2219 TQPLTKSNLGSQVGSVAGPFTGDLIR-KTSPQGLSTRVNI-TAPLSQRYVRIRY  
590 600 610 620 630 640  
MON863.pap ASTTNLALFVQNSNNDFLVIYINKTKNDDDLTYCTPDLATTNNMGFSQDKNELIGAR  
JC2219 APTTLQPTSTIDRPIQNGFSATMSSGNNLQSGSFRTVPTTFNFMSSSVFTLSAH  
650 660 670 680 690 700  
MON863.pap SFVSNKTYIDKIEFIPVGL  
JC2219 VFSNGNEVYIDRIFFVPAEVT  
710 720 730 740 750 760  
MON863.pap BTCRYAG\_1  
LOCUS BTCRYAG\_1  
DEFINITION *Bacillus thuringiensis* cryIA gene for parasporal crystal toxin;  
parasporal crystal toxin (AA 1-1176).  
DATE 12-Sep-1993  
ACCESSION X84355  
KEYWORDS 940566  
SCORES In1: 350 In1n: 863 Opt: 1100 z-score: 1089.3 E(): 3.3e-55  
Smith-Waterman score: 1111, 34.1% identity in 590 aa overlap  
MON863.pap BSTEVLNSTRVKDAVGTGISVVGQILGVVGPVAGALTSFYQSFINTIN---PSDADFW  
BTCRYAG\_1 VEVLGGRSTRTVTPIDISLSLTQFLSEFVPGAG---FVLGLVDIINGIPQSQ---W  
110 120 130 140 150 160  
MON863.pap KAPMAQVVELDKKIEVYAKKALAEQLQGNFEDYVNAIWNKKTPLSRKRSQRI  
BTCRYAG\_1 DAPLVQQLINGRIFEFARNQALISLBLSNLQVYARSFREWADPTN-PALREEMRI  
170 180 190 200 210 220  
MON863.pap RELFGAESHFRNSHPSFAVSKFVLFLPTAQAANTHLLKDAQVGEWYSSSDVA  
BTCRYAG\_1 Q-PFNKMSALTATPLAVQVPLSVVQAANHLHSLVLDVSPQGRWGFDAATIN  
230 240 250 260 270 280  
MON863.pap EFYRRLKLTQQTTHCVNMYVGLNGRGSTYDAWVKNFRPRRMTLTVLDLVLFPFY  
BTCRYAG\_1 SRYNDLTRLIQNTDYAVRWYVGLRWVSPGSDRWVKNFRPRRMTLTVLDLVLFPFY





MON863.pcp  
420 430 440 450 460 470  
CRYT\_BACTA  
420 430 440 450 460 470  
MON863.pcp  
520 530 540 550 560 570  
CRYT\_BACTA  
480 490 500 510 520 530  
MON863.pcp  
580 590 600 610 620 630  
CRYT\_BACTA  
540 550 560 570 580 590  
MON863.pcp  
640 650  
CRYT\_BACTA  
600 610 620 630 640 650

MON863.pcp  
TO:802134  
PI,802134 - parasporal crystal protein - *Bacillus thuringiensis* (strain aizawa)  
IC1  
N,Alternate names: delta-endotoxin IC1; entomocidal crystal protein  
C,Species: *Bacillus thuringiensis*  
A,Variant: strain aizawa IC1  
C,Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 08-Sep-1997

SCORES Initl: 409 Initn 993 Opt. 1088 z-score: 1077.5 E() 1.5e-54  
Smith-Waterman score: 1157; 35.0% identity in 585 aa overlap

MON863.pcp  
60 70 80 90 100  
802134  
30 40 50 60 70

MON863.pcp  
110 120 130 140 150 160  
802134  
80 90 100 110 120 130

MON863.pcp  
170 180 190 200 210 220  
802134  
140 150 160 170 180 190

MON863.pcp  
230 240 250 260 270 280  
802134  
200 210 220 230 240 250

MON863.pcp  
290 300 310 320 330 340  
802134  
260 270 280 290 300 310

MON863.pcp  
350 360 370 380 390 400  
802134  
310 320 330 340 350 360

MON863.pcp  
410 420 430 440 450 460  
802134  
370 380 390 400 410 420

MON863.pcp  
470 480 490 500 510  
802134  
420 430 440 450 460 470

MON863.pcp  
520 530 540 550 560 570  
802134  
480 490 500 510 520 530

MON863.pcp  
580 590 600 610 620 630  
802134  
540 550 560 570 580 590

MON863.pcp  
640 650  
802134  
600 610 620 630 640 650

MON863.pcp  
TO:802134  
LOCUS BCRV1C1\_1  
DEFINITION BCRV1C1\_1 gene for entomocidal delta-endotoxin,  
delta-endotoxin.  
DATE 12-SEP-1993  
ACCESSION X11833  
ID 940677  
SCORES Initl: 409 Initn: 993 Opt. 1088 z-score: 1077.5 E() 1.5e-54  
Smith-Waterman score: 1157; 35.0% identity in 585 aa overlap

MON863.pcp  
60 70 80 90 100  
BCRV1C1\_1  
30 40 50 60 70

MON863.pcp  
110 120 130 140 150 160  
BCRV1C1\_1  
80 90 100 110 120 130

MON863.pcp  
170 180 190 200 210 220  
BCRV1C1\_1  
140 150 160 170 180 190

MON863.pcp  
230 240 250 260 270 280  
BCRV1C1\_1  
200 210 220 230 240 250

MON863.pcp  
290 300 310 320 330 340  
BCRV1C1\_1  
260 270 280 290 300 310

MON863.pcp  
350 360 370 380 390 400  
BCRV1C1\_1  
310 320 330 340 350 360

MON863.pcp  
410 420 430 440 450 460  
BCRV1C1\_1  
370 380 390 400 410 420

MON863.pcp  
470 480 490 500 510  
BCRV1C1\_1  
420 430 440 450 460 470

MON863.pcp  
520 530 540 550 560 570  
BCRV1C1\_1  
480 490 500 510 520 530

MON863.pcp  
580 590 600 610 620 630  
BCRV1C1\_1  
540 550 560 570 580 590

MON863.pcp  
640 650  
BCRV1C1\_1  
600 610 620 630 640 650

MON863.pcp  
TO:CRY5\_BACTB  
ID CRY5\_BACTB STANDARD, PRT; 1155 AA  
AC P06577; DT 01-JAN-1988 (REL 06, CREATED)  
DT 01-JAN-1988 (REL 06, LAST SEQUENCE UPDATE)  
DT 01-MAY-1991 (REL 18, LAST ANNOTATION UPDATE)  
DE 130 KD CRYSTAL PROTEIN (DELTA ENDOTOXIN) (CRYSTALLINE ENTOMOICIDAL)

SCORES Initl: 409 Initn 989 Opt. 1084 z-score: 1073.6 E() 2.5e-54  
Smith-Waterman score: 1157; 34.9% identity in 585 aa overlap

MON863.pcp  
60 70 80 90 100  
CRY5\_BACTB  
30 40 50 60 70

MON863.pcp  
110 120 130 140 150 160  
CRY5\_BACTB  
80 90 100 110 120 130

MON863.pcp  
170 180 190 200 210 220  
CRY5\_BACTB  
140 150 160 170 180 190

MON863.pcp  
230 240 250 260 270 280  
CRY5\_BACTB  
200 210 220 230 240 250

MON863.pcp  
290 300 310 320 330 340  
CRY5\_BACTB  
260 270 280 290 300 310

MON863.pcp  
350 360 370 380 390 400  
CRY5\_BACTB  
310 320 330 340 350 360

MON863.pcp  
410 420 430 440 450 460  
CRY5\_BACTB  
370 380 390 400 410 420

MON863.pcp  
470 480 490 500 510  
CRY5\_BACTB  
420 430 440 450 460 470

MON863.pcp  
520 530 540 550 560 570  
CRY5\_BACTB  
480 490 500 510 520 530

LOCUS M0863.pap  
DEFINITION B. thuringiensis (berliner), crystalline entomocidal protoxin  
gene, complete cds;  
DATE 20-MAY-1987  
ACCESSION M13898..

SCORES Init1 409 Initn 989 Opt. 1084 z-score: 1073.6 E(-2.5e-54)  
Smith-Waterman score: 1153; 34.9% identity in 585 aa overlap

LOCUS M0863.pap  
DEFINITION B. thuringiensis (berliner), crystalline entomocidal protoxin  
gene, complete cds;  
DATE 20-MAY-1987  
ACCESSION M13898..

SCORES Init1 409 Initn 989 Opt. 1084 z-score: 1073.6 E(-2.5e-54)  
Smith-Waterman score: 1153; 34.9% identity in 585 aa overlap

LOCUS M0863.pap  
DEFINITION B. thuringiensis (berliner), crystalline entomocidal protoxin  
gene, complete cds;  
DATE 20-MAY-1987  
ACCESSION M13898..

SCORES Init1 409 Initn 989 Opt. 1084 z-score: 1073.6 E(-2.5e-54)  
Smith-Waterman score: 1153; 34.9% identity in 585 aa overlap

LOCUS M0863.pap  
DEFINITION B. thuringiensis (berliner), crystalline entomocidal protoxin  
gene, complete cds;  
DATE 20-MAY-1987  
ACCESSION M13898..

SCORES Init1 409 Initn 989 Opt. 1084 z-score: 1073.6 E(-2.5e-54)  
Smith-Waterman score: 1153; 34.9% identity in 585 aa overlap

LOCUS M0863.pap  
DEFINITION B. thuringiensis (berliner), crystalline entomocidal protoxin  
gene, complete cds;  
DATE 20-MAY-1987  
ACCESSION M13898..

SCORES Init1 409 Initn 989 Opt. 1084 z-score: 1073.6 E(-2.5e-54)  
Smith-Waterman score: 1153; 34.9% identity in 585 aa overlap

LOCUS M0863.pap  
DEFINITION B. thuringiensis (berliner), crystalline entomocidal protoxin  
gene, complete cds;  
DATE 20-MAY-1987  
ACCESSION M13898..

SCORES Init1 409 Initn 989 Opt. 1084 z-score: 1073.6 E(-2.5e-54)  
Smith-Waterman score: 1153; 34.9% identity in 585 aa overlap

LOCUS M0863.pap  
DEFINITION B. thuringiensis (berliner), crystalline entomocidal protoxin  
gene, complete cds;  
DATE 20-MAY-1987  
ACCESSION M13898..

SCORES Init1 409 Initn 989 Opt. 1084 z-score: 1073.6 E(-2.5e-54)  
Smith-Waterman score: 1153; 34.9% identity in 585 aa overlap

LOCUS M0863.pap  
DEFINITION B. thuringiensis (berliner), crystalline entomocidal protoxin  
gene, complete cds;  
DATE 20-MAY-1987  
ACCESSION M13898..

SCORES Init1 409 Initn 989 Opt. 1084 z-score: 1073.6 E(-2.5e-54)  
Smith-Waterman score: 1153; 34.9% identity in 585 aa overlap

LOCUS M0863.pap  
DEFINITION B. thuringiensis (berliner), crystalline entomocidal protoxin  
gene, complete cds;  
DATE 20-MAY-1987  
ACCESSION M13898..

SCORES Init1 409 Initn 989 Opt. 1084 z-score: 1073.6 E(-2.5e-54)  
Smith-Waterman score: 1153; 34.9% identity in 585 aa overlap



DEFINITION *B. thuringiensis* (var. *Kurstaki* HD1) gene for delta-endotoxin  
 DATE 25-JAN-1994  
 ACCESSION A09398  
 NID g420487  
 ORGANISM *Bacillus thuringiensis*  
 SCORES Initl 409 Initn 989 Opt 1084 z-score 1073.6 E() 2.5e-54  
 Smith-Waterman score 1153; 34.9% identity in 585 aa overlap  
 MON863.pap SSTEVLNSTVKDAVGTSVVGQILGVVPPAGALTSFYQSPLNTIN----PSADAPW  
 A09398\_1 VEVLGERISTGYTPIDISLSTQPLLSRFVPGAG----FVLGLVDIINGIPPSQ--W  
 MON863.pap 110 120 130 140 150 160  
 A09398\_1 110 120 130 140 150 160  
 MON863.pap KAFMAQVEVLIDKIKERYAKKALABLOGLQNFEDYVVALNSWKKTPSLRSKRSGRI  
 A09398\_1 DAFIVQEQINQRIEFARNQASRLBGLSNLYQIYAEFRWADPTN-PALAREMKI  
 MON863.pap 170 180 190 200 210 220  
 A09398\_1 170 180 190 200 210 220  
 MON863.pap RELFSAQSHFRNSMPSFAVSKFVFLPTAQAANHLILLKDAQVGEWYSSSDVA  
 A09398\_1 Q--FNDMSALTAIPAVQNYQVPLLSVYVQAANHLVLRLDVSVPQGRWFDATIN  
 MON863.pap 230 240 250 260 270 280  
 A09398\_1 230 240 250 260 270 280  
 MON863.pap ERYRRLKLTQQTTHCVMVNVGLNGRSTYDAWVKFRFRREMTLTVDLVLFPFY  
 A09398\_1 SRYNDLRIGMTTHAVWNTGLRWGPDSDRMKYNQFRELTLTVDLVLFPFY  
 MON863.pap 290 300 310 320 330 340  
 A09398\_1 290 300 310 320 330 340  
 MON863.pap DILNLYGKVKELTRDITPDIPLTLTKYQVPTLSIENSIRKPHLPDVLQOIFPTRL  
 A09398\_1 DSCYPTVNSQLTRITNVP-LENFDSGFSGAQIGSIRSPHLMIDINSITTYDA  
 MON863.pap 350 360 370 380 390 400  
 A09398\_1 350 360 370 380 390 400  
 MON863.pap RFGYFGKDSFNYMGNVETSPISGSKTITSFYQDK-STSPVQLSPD-GQKYRTIA  
 A09398\_1 HRGEY-----TWGSHIMASVQFSQFPTFLVGTNNAAPQRIVAQLOQVYRTLS  
 MON863.pap 410 420 430 440 450 460  
 A09398\_1 410 420 430 440 450 460  
 MON863.pap NTDVAAMPKQKYLGVTKVDFSGYDQK---NRTSTQYDSKRNHVSACDSTDLPPH  
 A09398\_1 ST-LYRRP---FNTGNNQGLSVLDOTEFAYGTSNLSAVYKSGTV---DSDLEIPQ  
 MON863.pap 470 480 490 500 510 520  
 A09398\_1 470 480 490 500 510 520  
 MON863.pap TTDPLEKAYSHQINACFLMDQRNGTI---PFFTWTRSVDFPNTIDAKKITQLPV  
 A09398\_1 TKTNLSGTSVVGKPGTGGDILR-RTSPQSTSLRVNI-TAPLSQRYRVRIRYASTIN  
 MON863.pap 530 540 550 560 570 580  
 A09398\_1 530 540 550 560 570 580  
 MON863.pap LALFVQNSNDFLVIYINKTKDKDDLTQYTFDLATTNSMGPFGDKNELIGASFSVN  
 A09398\_1 LQFTSTIDRPGNQFSAIKSSGSLQGGSPFTVQFTVPVFNQSSSVTLGAVFNMSG  
 MON863.pap 590 600 610 620 630 640  
 A09398\_1 590 600 610 620 630 640  
 MON863.pap KVIQDKIEFTVQL  
 A09398\_1 NEVYIDRIEFVPAVTFEAYDLERAQKAVNELFTSSNQGLKLTVDYHIDQVSNLVC  
 MON863.pap 650 660 670 680 690 700  
 A09398\_1 650 660 670 680 690 700

MON863.pap TO BACHDCRY\_1  
 LOCUS BACHDCRY\_1  
 DEFINITION *Bacillus thuringiensis* gene for delta-endotoxin, complete cds  
 DATE 16-APR-1998  
 ACCESSION D00117  
 NID g216279

ORGANISM *Bacillus thuringiensis*  
 SCORES Initl 409 Initn 989 Opt 1084 z-score 1073.6 E() 2.5e-54  
 Smith-Waterman score 1153; 34.9% identity in 585 aa overlap  
 MON863.pap SSTEVLNSTVKDAVGTSVVGQILGVVPPAGALTSFYQSPLNTIN----PSADAPW  
 BACHDCRY\_1 VEVLGERISTGYTPIDISLSTQPLLSRFVPGAG----FVLGLVDIINGIPPSQ--W  
 MON863.pap 110 120 130 140 150 160  
 BACHDCRY\_1 110 120 130 140 150 160  
 MON863.pap KAFMAQVEVLIDKIKERYAKKALABLOGLQNFEDYVVALNSWKKTPSLRSKRSGRI  
 BACHDCRY\_1 DAFIVQEQINQRIEFARNQASRLBGLSNLYQIYAEFRWADPTN-PALAREMKI  
 MON863.pap 170 180 190 200 210 220  
 BACHDCRY\_1 170 180 190 200 210 220  
 MON863.pap RELFSAQSHFRNSMPSFAVSKFVFLPTAQAANHLILLKDAQVGEWYSSSDVA  
 BACHDCRY\_1 Q--FNDMSALTAIPAVQNYQVPLLSVYVQAANHLVLRLDVSVPQGRWFDATIN  
 MON863.pap 230 240 250 260 270 280  
 BACHDCRY\_1 230 240 250 260 270 280  
 MON863.pap ERYRRLKLTQQTTHCVMVNVGLNGRSTYDAWVKFRFRREMTLTVDLVLFPFY  
 BACHDCRY\_1 SRYNDLRIGMTTHAVWNTGLRWGPDSDRMKYNQFRELTLTVDLVLFPFY  
 MON863.pap 290 300 310 320 330 340  
 BACHDCRY\_1 290 300 310 320 330 340  
 MON863.pap DILNLYGKVKELTRDITPDIPLTLTKYQVPTLSIENSIRKPHLPDVLQOIFPTRL  
 BACHDCRY\_1 DSCYPTVNSQLTRITNVP-LENFDSGFSGAQIGSIRSPHLMIDINSITTYDA  
 MON863.pap 350 360 370 380 390 400  
 BACHDCRY\_1 350 360 370 380 390 400  
 MON863.pap RFGYFGKDSFNYMGNVETSPISGSKTITSFYQDK-STSPVQLSPD-GQKYRTIA  
 BACHDCRY\_1 HRGEY-----TWGSHIMASVQFSQFPTFLVGTNNAAPQRIVAQLOQVYRTLS  
 MON863.pap 410 420 430 440 450 460  
 BACHDCRY\_1 410 420 430 440 450 460  
 MON863.pap NTDVAAMPKQKYLGVTKVDFSGYDQK---NRTSTQYDSKRNHVSACDSTDLPPH  
 BACHDCRY\_1 ST-LYRRP---FNTGNNQGLSVLDOTEFAYGTSNLSAVYKSGTV---DSDLEIPQ  
 MON863.pap 470 480 490 500 510 520  
 BACHDCRY\_1 470 480 490 500 510 520  
 MON863.pap TTDPLEKAYSHQINACFLMDQRNGTI---PFFTWTRSVDFPNTIDAKKITQLPV  
 BACHDCRY\_1 TKTNLSGTSVVGKPGTGGDILR-RTSPQSTSLRVNI-TAPLSQRYRVRIRYASTIN  
 MON863.pap 530 540 550 560 570 580  
 BACHDCRY\_1 530 540 550 560 570 580  
 MON863.pap LALFVQNSNDFLVIYINKTKDKDDLTQYTFDLATTNSMGPFGDKNELIGASFSVN  
 BACHDCRY\_1 LQFTSTIDRPGNQFSAIKSSGSLQGGSPFTVQFTVPVFNQSSSVTLGAVFNMSG  
 MON863.pap 590 600 610 620 630 640  
 BACHDCRY\_1 590 600 610 620 630 640  
 MON863.pap KVIQDKIEFTVQL  
 BACHDCRY\_1 NEVYIDRIEFVPAVTFEAYDLERAQKAVNELFTSSNQGLKLTVDYHIDQVSNLVC  
 MON863.pap 650 660 670 680 690 700  
 BACHDCRY\_1 650 660 670 680 690 700

Pl:JD0002 - parasporal crystal protein - *Bacillus thuringiensis*  
 N:Alternate names: delta-endotoxin-2; entomocidal protoxin, entomopathogenic  
 crystal protein  
 C:Species: *Bacillus thuringiensis*  
 C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 20-Mar-1998  
 C:Accession A90025; A91560, A90955; S14555; A24172, A29043, JD0002  
 SCORES Initl 409 Initn 989 Opt 1084 z-score 1073.6 E() 2.5e-54  
 Smith-Waterman score 1153; 34.9% identity in 585 aa overlap  
 Pl:JD0002 - parasporal crystal protein - *Bacillus thuringiensis*  
 N:Alternate names: delta-endotoxin-2; entomocidal protoxin, entomopathogenic  
 crystal protein  
 C:Species: *Bacillus thuringiensis*  
 C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 20-Mar-1998  
 C:Accession A90025; A91560, A90955; S14555; A24172, A29043, JD0002  
 SCORES Initl 409 Initn 989 Opt 1084 z-score 1073.6 E() 2.5e-54  
 Smith-Waterman score 1153; 34.9% identity in 585 aa overlap

MON863.pap 60 70 80 90 100  
 JD0002 SSTEVLNSTVKDAVGTSVVGQILGVVPPAGALTSFYQSPLNTIN----PSADAPW  
 VEVLGERISTGYTPIDISLSTQPLLSRFVPGAG----FVLGLVDIINGIPPSQ--W  
 MON863.pap 110 120 130 140 150 160  
 JD0002 110 120 130 140 150 160  
 MON863.pap KAFMAQVEVLIDKIKERYAKKALABLOGLQNFEDYVVALNSWKKTPSLRSKRSGRI  
 JD0002 DAFIVQEQINQRIEFARNQASRLBGLSNLYQIYAEFRWADPTN-PALAREMKI  
 MON863.pap 170 180 190 200 210 220  
 JD0002 170 180 190 200 210 220  
 MON863.pap RELFSAQSHFRNSMPSFAVSKFVFLPTAQAANHLILLKDAQVGEWYSSSDVA  
 JD0002 Q--FNDMSALTAIPAVQNYQVPLLSVYVQAANHLVLRLDVSVPQGRWFDATIN  
 MON863.pap 230 240 250 260 270 280  
 JD0002 230 240 250 260 270 280  
 MON863.pap ERYRRLKLTQQTTHCVMVNVGLNGRSTYDAWVKFRFRREMTLTVDLVLFPFY  
 JD0002 SRYNDLRIGMTTHAVWNTGLRWGPDSDRMKYNQFRELTLTVDLVLFPFY  
 MON863.pap 290 300 310 320 330 340  
 JD0002 290 300 310 320 330 340  
 MON863.pap DILNLYGKVKELTRDITPDIPLTLTKYQVPTLSIENSIRKPHLPDVLQOIFPTRL  
 JD0002 DSCYPTVNSQLTRITNVP-LENFDSGFSGAQIGSIRSPHLMIDINSITTYDA  
 MON863.pap 350 360 370 380 390 400  
 JD0002 350 360 370 380 390 400  
 MON863.pap RFGYFGKDSFNYMGNVETSPISGSKTITSFYQDK-STSPVQLSPD-GQKYRTIA  
 JD0002 HRGEY-----TWGSHIMASVQFSQFPTFLVGTNNAAPQRIVAQLOQVYRTLS  
 MON863.pap 410 420 430 440 450 460  
 JD0002 410 420 430 440 450 460  
 MON863.pap NTDVAAMPKQKYLGVTKVDFSGYDQK---NRTSTQYDSKRNHVSACDSTDLPPH  
 JD0002 ST-LYRRP---FNTGNNQGLSVLDOTEFAYGTSNLSAVYKSGTV---DSDLEIPQ  
 MON863.pap 470 480 490 500 510 520  
 JD0002 470 480 490 500 510 520  
 MON863.pap TTDPLEKAYSHQINACFLMDQRNGTI---PFFTWTRSVDFPNTIDAKKITQLPV  
 JD0002 TKTNLSGTSVVGKPGTGGDILR-RTSPQSTSLRVNI-TAPLSQRYRVRIRYASTIN  
 MON863.pap 530 540 550 560 570 580  
 JD0002 530 540 550 560 570 580  
 MON863.pap LALFVQNSNDFLVIYINKTKDKDDLTQYTFDLATTNSMGPFGDKNELIGASFSVN  
 JD0002 LQFTSTIDRPGNQFSAIKSSGSLQGGSPFTVQFTVPVFNQSSSVTLGAVFNMSG  
 MON863.pap 590 600 610 620 630 640  
 JD0002 590 600 610 620 630 640  
 MON863.pap KVIQDKIEFTVQL  
 JD0002 NEVYIDRIEFVPAVTFEAYDLERAQKAVNELFTSSNQGLKLTVDYHIDQVSNLVC  
 MON863.pap 650 660 670 680 690 700  
 JD0002 650 660 670 680 690 700

LOCUS BACHDCRY\_1  
 DEFINITION *Bacillus thuringiensis* berliner 1715 bac gene for delta  
 endotoxin; delta endotoxin NA 1-1151  
 DATE 12-SEP-1993  
 ACCESSION X04568  
 NID g42054  
 SCORES Initl 409 Initn 989 Opt 1084 z-score 1073.6 E() 2.5e-54  
 Smith-Waterman score 1153; 34.9% identity in 585 aa overlap  
 MON863.pap 60 70 80 90 100  
 FBDT2TOX\_1 SSTEVLNSTVKDAVGTSVVGQILGVVPPAGALTSFYQSPLNTIN----PSADAPW  
 VEVLGERISTGYTPIDISLSTQPLLSRFVPGAG----FVLGLVDIINGIPPSQ--W

MON863.pap  
BTBT2TOX\_1  
KATMAQVRLDKKIEETAKSALAELOGLQNTFEDYVNAIWSKKTPLSLRKSKSQRI  
DAFLVQIBOLINRIEFARNGATSRLEOLSNLYQIYASFRNEADPTN-PALAREMRI  
170 180 190 200 210 220  
MON863.pap  
BTBT2TOX\_1  
RELFSQARSHFNPSFPAVSKFEVLFLPTAQAANTHLLLLADAQVPGERWGYSSRDVA  
Q--PNDNSALTAIPFAVQNYQVLLSVYQAANLHLVLRVDSVFGORNGPDAATIN  
230 240 250 260 270 280 290  
MON863.pap  
BTBT2TOX\_1  
EFYRQKLKTOQYTDHCVMYVGLNLRGSTYDAWVKFNRFRMTCITVDLVLFPFY  
SRYNOLRLIGMTTHAVPNTLRLSRVSPDRSDRWITVDFRRLITVLDIVSLFPNY  
290 300 310 320 330 340 350  
MON863.pap  
BTBT2TOX\_1  
DIRLSYKGVKTELTDIPTDPTFLTLQKYGPTFLSIENSIRKPHLPDGLQIEFTL  
DRTYPIRTVSLQREIYTNV-LENFDGSPRAGAQIBOSTRSPHLMDSITITDA  
350 360 370 380 390 400 410  
MON863.pap  
BTBT2TOX\_1  
RPOYFQKDSFNYWGNVETSPISGSKTITSPFYQDK-STSPVQKLEP-OKQVYRTIA  
HGEY-----YMSGQIMASPVGSGPPTFLVGTGMNAAPQORIVAGLQGVYTLG  
410 420 430 440 450 460 470  
MON863.pap  
BTBT2TOX\_1  
NTDVAAMPNKKVYLVGVTKVDFSDQDQK---NETSTQYDSKRNNGHVSACDSIDQLPPE  
ST-LYRRP---FNIGLNQGLVLDGTFAVGTSSNLPSAVYKSGTV---DGLDEIFPQ  
470 480 490 500 510 520 530  
MON863.pap  
BTBT2TOX\_1  
TTDBLEKAYSHQNLVACFLNQDRGTI-----PFTVTHRSVDFPNTIDAKITOLPV  
NNVTPRQGFQSHRSHVSNFSPGSESVIIRAPMPSWIRSANFNINIPSGQITQPL  
520 530 540 550 560 570 580  
MON863.pap  
BTBT2TOX\_1  
VKAYALSQASIIEGPQGTGNNLFLKESNSIAKPKVTLNSAALLQRYVRIYASTTN  
TKSTWLSGTSVVGKGFQGDILA-RTPSQIETLAVNI-TAPLSQRYVRIYASTTN  
580 590 600 610 620 630 640  
MON863.pap  
BTBT2TOX\_1  
LRLVQNSMNDLVIYINKVNDKDDLTQTPPLATTSMKSPGSKNELIIGAESFVSN  
LQFTSTIDORPTNQNFSAHSGSGLQGGFRTVCTTPFNYSGNSVFTLSAIVFNSQ  
640 650 660 670 680 690 700  
MON863.pap  
BTBT2TOX\_1  
EKIYIDKIEFIPVOL  
NRVYIDRIEFVAVRVTGAEYDLERAKAVNHLFSSNGIQGLKTDVTHIDQVSNLVC  
600 610 620 630 640 650  
MON863.pap  
TO:CRYV\_BACTA  
ID CRYV\_BACTA STANDARD, PRT, 1155 AA.  
AC P09667;  
DT 01-MAR-1989 (REL 10, CREATED)  
DT 01-MAR-1989 (REL 10, LAST SEQUENCE UPDATE)  
DT 01-MAY-1991 (REL 18, LAST ANNOTATION UPDATE)  
DE 130 KD CRYSTAL PROTEIN (DELTA ENDOTOXIN) (CRYSTALLINE ENTOMOCIDAL . .  
SCORES Initl 409 Initn 984 Opt 1079 z-score: 1073.6 E( ) 2.5e-54  
Smith-Waterman score: 1153; 34.84 identity in 589 aa overlap  
MON863.pap  
CRYV\_BACTA  
SSTVGLNSTVKDAVGTGIVSVQVLGVVQVFPAGALSTFYQSLNTIW---PSADPW  
VEVLGGRISTGYTIDISLSTQFLLSEFVGAG---FVLGLVDITWIGLFPQSQ--W  
110 120 130 140 150 160  
MON863.pap  
CRYV\_BACTA  
KAPMAQVRLDKKIEETAKSALAELOGLQNTFEDYVNAIWSKKTPLSLRKSKSQRI  
DAFLVQIBOLINRIEFARNGATSRLEOLSNLYQIYASFRNEADPTN-PALAREMRI  
170 180 190 200 210 220  
MON863.pap  
CRYV\_BACTA  
RELFSQARSHFNPSFPAVSKFEVLFLPTAQAANTHLLLLADAQVPGERWGYSSRDVA  
Q--PNDNSALTAIPFAVQNYQVLLSVYQAANLHLVLRVDSVFGORNGPDAATIN  
230 240 250 260 270 280 290

MON863.pap  
CRYV\_BACTA  
RELFSQARSHFNPSFPAVSKFEVLFLPTAQAANTHLLLLADAQVPGERWGYSSRDVA  
Q--PNDNSALTAIPFAVQNYQVLLSVYQAANLHLVLRVDSVFGORNGPDAATIN  
170 180 190 200 210 220  
MON863.pap  
CRYV\_BACTA  
EFYRQKLKTOQYTDHCVMYVGLNLRGSTYDAWVKFNRFRMTCITVDLVLFPFY  
SRYNOLRLIGMTTHAVPNTLRLSRVSPDRSDRWITVDFRRLITVLDIVSLFPNY  
230 240 250 260 270 280 290  
MON863.pap  
CRYV\_BACTA  
DIRLSYKGVKTELTDIPTDPTFLTLQKYGPTFLSIENSIRKPHLPDGLQIEFTL  
DRTYPIRTVSLQREIYTNV-LENFDGSPRAGAQIBOSTRSPHLMDSITITDA  
290 300 310 320 330 340 350  
MON863.pap  
CRYV\_BACTA  
RPOYFQKDSFNYWGNVETSPISGSKTITSPFYQDK-STSPVQKLEP-OKQVYRTIA  
HGEY-----YMSGQIMASPVGSGPPTFLVGTGMNAAPQORIVAGLQGVYTLG  
350 360 370 380 390 400 410  
MON863.pap  
CRYV\_BACTA  
NTDVAAMPNKKVYLVGVTKVDFSDQDQK---NETSTQYDSKRNNGHVSACDSIDQLPPE  
ST-LYRRP---FNIGLNQGLVLDGTFAVGTSSNLPSAVYKSGTV---DGLDEIFPQ  
410 420 430 440 450 460 470  
MON863.pap  
CRYV\_BACTA  
TTDBLEKAYSHQNLVACFLNQDRGTI-----PFTVTHRSVDFPNTIDAKITOLPV  
NNVTPRQGFQSHRSHVSNFSPGSESVIIRAPMPSWIRSANFNINIPSGQITQPL  
470 480 490 500 510 520 530  
MON863.pap  
CRYV\_BACTA  
VKAYALSQASIIEGPQGTGNNLFLKESNSIAKPKVTLNSAALLQRYVRIYASTTN  
TKSTWLSGTSVVGKGFQGDILA-RTPSQIETLAVNI-TAPLSQRYVRIYASTTN  
520 530 540 550 560 570 580  
MON863.pap  
CRYV\_BACTA  
LRLVQNSMNDLVIYINKVNDKDDLTQTPPLATTSMKSPGSKNELIIGAESFVSN  
LQFTSTIDORPTNQNFSAHSGSGLQGGFRTVCTTPFNYSGNSVFTLSAIVFNSQ  
580 590 600 610 620 630 640  
MON863.pap  
CRYV\_BACTA  
EKIYIDKIEFIPVOL  
NRVYIDRIEFVAVRVTGAEYDLERAKAVNHLFSSNGIQGLKTDVTHIDQVSNLVC  
640 650 660 670 680 690 700  
MON863.pap  
TO:CRYV\_BACTA  
ID CRYV\_BACTA STANDARD, PRT, 1155 AA.  
AC P21257; DT 01-MAY-1991 (REL 18, CREATED)  
DT 01-MAY-1991 (REL 18, LAST SEQUENCE UPDATE)  
DT 01-MAY-1991 (REL 18, LAST ANNOTATION UPDATE)  
DE 130 KD CRYSTAL PROTEIN (DELTA ENDOTOXIN) (CRYSTALLINE ENTOMOCIDAL . .  
SCORES Initl 409 Initn 984 Opt 1079 z-score: 1068.6 E( ) 4.6e-54  
Smith-Waterman score: 1149; 34.94 identity in 585 aa overlap  
MON863.pap  
CRYV\_BACTA  
SSTVGLNSTVKDAVGTGIVSVQVLGVVQVFPAGALSTFYQSLNTIW---PSADPW  
VEVLGGRISTGYTIDISLSTQFLLSEFVGAG---FVLGLVDITWIGLFPQSQ--W  
60 70 80 90 100  
MON863.pap  
CRYV\_BACTA  
KAPMAQVRLDKKIEETAKSALAELOGLQNTFEDYVNAIWSKKTPLSLRKSKSQRI  
DAFLVQIBOLINRIEFARNGATSRLEOLSNLYQIYASFRNEADPTN-PALAREMRI  
110 120 130 140 150 160  
MON863.pap  
CRYV\_BACTA  
RELFSQARSHFNPSFPAVSKFEVLFLPTAQAANTHLLLLADAQVPGERWGYSSRDVA  
Q--PNDNSALTAIPFAVQNYQVLLSVYQAANLHLVLRVDSVFGORNGPDAATIN  
170 180 190 200 210 220

MON863.pap  
CRYV\_BACTA  
EFYRQKLKTOQYTDHCVMYVGLNLRGSTYDAWVKFNRFRMTCITVDLVLFPFY  
SRYNOLRLIGMTTHAVPNTLRLSRVSPDRSDRWITVDFRRLITVLDIVSLFPNY  
230 240 250 260 270 280 290  
MON863.pap  
CRYV\_BACTA  
DIRLSYKGVKTELTDIPTDPTFLTLQKYGPTFLSIENSIRKPHLPDGLQIEFTL  
DRTYPIRTVSLQREIYTNV-LENFDGSPRAGAQIBOSTRSPHLMDSITITDA  
290 300 310 320 330 340 350  
MON863.pap  
CRYV\_BACTA  
RPOYFQKDSFNYWGNVETSPISGSKTITSPFYQDK-STSPVQKLEP-OKQVYRTIA  
HGEY-----YMSGQIMASPVGSGPPTFLVGTGMNAAPQORIVAGLQGVYTLG  
350 360 370 380 390 400 410  
MON863.pap  
CRYV\_BACTA  
NTDVAAMPNKKVYLVGVTKVDFSDQDQK---NETSTQYDSKRNNGHVSACDSIDQLPPE  
ST-LYRRP---FNIGLNQGLVLDGTFAVGTSSNLPSAVYKSGTV---DGLDEIFPQ  
410 420 430 440 450 460 470  
MON863.pap  
CRYV\_BACTA  
TTDBLEKAYSHQNLVACFLNQDRGTI-----PFTVTHRSVDFPNTIDAKITOLPV  
NNVTPRQGFQSHRSHVSNFSPGSESVIIRAPMPSWIRSANFNINIPSGQITQPL  
470 480 490 500 510 520 530  
MON863.pap  
CRYV\_BACTA  
VKAYALSQASIIEGPQGTGNNLFLKESNSIAKPKVTLNSAALLQRYVRIYASTTN  
TKSTWLSGTSVVGKGFQGDILA-RTPSQIETLAVNI-TAPLSQRYVRIYASTTN  
520 530 540 550 560 570 580  
MON863.pap  
CRYV\_BACTA  
LRLVQNSMNDLVIYINKVNDKDDLTQTPPLATTSMKSPGSKNELIIGAESFVSN  
LQFTSTIDORPTNQNFSAHSGSGLQGGFRTVCTTPFNYSGNSVFTLSAIVFNSQ  
580 590 600 610 620 630 640  
MON863.pap  
CRYV\_BACTA  
EKIYIDKIEFIPVOL  
NRVYIDRIEFVAVRVTGAEYDLERAKAVNHLFSSNGIQGLKTDVTHIDQVSNLVC  
640 650 660 670 680 690 700  
MON863.pap  
TO:CRYV\_BACTA  
P1:139838 - delta-endotoxin (5.3 class) - Bacillus thuringiensis  
C.Species: Bacillus thuringiensis  
C.Date: 19-Jul-1996 #sequence\_revision: 19-Jul-1996 #text\_change: 16-Feb-1997  
C.Accession: J39838  
R:K. J. A. Broussard, R. J. Prefontaine, G. J. Hanna, Z. J. Condie, J. A. Lau,  
P. C. K. . . .  
SCORES Initl 409 Initn 984 Opt 1079 z-score: 1068.6 E( ) 4.6e-54  
Smith-Waterman score: 1149; 34.94 identity in 585 aa overlap  
MON863.pap  
I39838  
SSTVGLNSTVKDAVGTGIVSVQVLGVVQVFPAGALSTFYQSLNTIW---PSADPW  
VEVLGGRISTGYTIDISLSTQFLLSEFVGAG---FVLGLVDITWIGLFPQSQ--W  
60 70 80 90 100  
MON863.pap  
I39838  
KAPMAQVRLDKKIEETAKSALAELOGLQNTFEDYVNAIWSKKTPLSLRKSKSQRI  
DAFLVQIBOLINRIEFARNGATSRLEOLSNLYQIYASFRNEADPTN-PALAREMRI  
110 120 130 140 150 160  
MON863.pap  
I39838  
RELFSQARSHFNPSFPAVSKFEVLFLPTAQAANTHLLLLADAQVPGERWGYSSRDVA  
Q--PNDNSALTAIPFAVQNYQVLLSVYQAANLHLVLRVDSVFGORNGPDAATIN  
170 180 190 200 210 220

MON863.pap  
I39838  
230 240 250 260 270 280  
EFTYRQLKLTQYTHVAVWVWVGLNGLRGSTYDAMVKNFRFRREMTLTVDLVLVLPFY  
SRVNDLTALIGNTYTHAVRWVWVGLNGLRGSTYDAMVKNFRFRREMTLTVDLVLVLPFY  
200 210 220 230 240 250  
MON863.pap  
I39838  
290 300 310 320 330 340  
DRLYSKGVKTLTRDITPDIPLLTQKQYPTFLSIENSRKPHLPDYLOGIEFTRL  
DRLYSKGVKTLTRDITPDIPLLTQKQYPTFLSIENSRKPHLPDYLOGIEFTRL  
260 270 280 290 300 310  
MON863.pap  
I39838  
350 360 370 380 390 400  
RPGYFGKDSFNYNGVYVTRPSIGSSNTISPPYCDK-STEPVQKLSFD-GQKYVRTIA  
HRGEY-----YWSGHIMAFVWVWVGLNGLRGSTYDAMVKNFRFRREMTLTVDLVLVLPFY  
310 320 330 340 350 360  
MON863.pap  
I39838  
410 420 430 440 450 460  
NTDVAAMPNGKVLVGVKVSQYDDK---NETSTQYVDSKRMHVSQAQSDIDOLPPE  
ST-LYRRP---FNIGINNQLSVLDOTEFAYGTSNLSNPVYKSGTV---DSLDIEPPO  
370 380 390 400 410 420  
MON863.pap  
I39838  
470 480 490 500 510  
TDEPLEKAYSHQLYABCFMQDRRTI---PFTVTHRSVDFNTIDAEKITQLPV  
NNVPPRQGFSGRLSHVWVWVGLNGLRGSTYDAMVKNFRFRREMTLTVDLVLVLPFY  
420 430 440 450 460 470  
MON863.pap  
I39838  
520 530 540 550 560 570  
VKAVALSRGATIRGPGFQGNLFLKESNSSTAKFVTLNSAALLQYVVRIRYASTIN  
TKSTNLSGTSVVGPGTGGDLR-RTSPOQISTLRVNI-TAPLSQRYVVRIRYASTIN  
480 490 500 510 520 530  
MON863.pap  
I39838  
580 590 600 610 620 630  
LRLVQNSNNDPLVIYINKPKDDLTQYPTFLATNSNMGFSGDNELIIGASPVSN  
LAPVCHMORPINCQFPAHNSGSLGSGFPHLQYPTFLATNSNMGFSGDNELIIGASPVSN  
540 550 560 570 580 590  
MON863.pap  
I39838  
640 650  
RKVIIDKIEPIVQL  
NEVIDIRFVPARVTPFAEDYERAKAVNELFTSSNQIQLTVDYHIDQVSNLVECLSD  
600 610 620 630 640 650  
MON863.pap  
TO BACETNRD12\_1  
LOCUS BACETNRD12\_1  
DEFINITION B. thuringiensis 5.3 class delta-endotoxin gene, complete cds,  
5.3 class delta-endotoxin  
DATE 14-NOV-1991  
ACCESSION M72423  
NID g142885  
SCORES Intit. 409 Inntn. 984 Opt. 1079 z-score: 1056.6 E( ) 4 6e-54  
Smith-Waterman score: 1148; 34.9% identity in 585 aa overlap  
MON863.pap  
I39838  
660 670 680 690 700  
BSTEVLNDSVVKDAVGTQISVVGQILGVVGFAGALSTFYQSFINTW-----PSADPW  
BACETNRD12\_1 VEVLOGERITQYTPIDISLSTQFLSRFPVAG---FVLGLDILWGTPQDQW  
30 40 50 60 70  
MON863.pap  
I39838  
710 720 730 740 750 760  
KAPMAQVVLIDKIEYPSKALAEQLQGNFEDYVALNSWKTPLSLKRSQORREL  
DAFVLOGERITQYTPIDISLSTQFLSRFPVAG---FVLGLDILWGTPQDQW  
70 80 90 100 110 120  
MON863.pap  
I39838  
170 180 190 200 210 220  
RELPSQASHFRNPKSPFAVSKFVLFTYAQAANTHLLKDAQVSGWSSYSDVA  
O-FNDMSALITATPLAVQVGVLLSVYQANLHLSVLKDVSTFGQVQFDAATTN  
140 150 160 170 180 190  
MON863.pap  
I39838  
230 240 250 260 270 280  
EFTYRQLKLTQYTHVAVWVWVGLNGLRGSTYDAMVKNFRFRREMTLTVDLVLVLPFY  
SRVNDLTALIGNTYTHAVRWVWVGLNGLRGSTYDAMVKNFRFRREMTLTVDLVLVLPFY  
200 210 220 230 240 250

MON863.pap  
I39838  
290 300 310 320 330 340  
DRLYSKGVKTLTRDITPDIPLLTQKQYPTFLSIENSRKPHLPDYLOGIEFTRL  
BACETNRD12\_1 DRLYSKGVKTLTRDITPDIPLLTQKQYPTFLSIENSRKPHLPDYLOGIEFTRL  
260 270 280 290 300 310  
MON863.pap  
I39838  
350 360 370 380 390 400  
RPGYFGKDSFNYNGVYVTRPSIGSSNTISPPYCDK-STEPVQKLSFD-GQKYVRTIA  
HRGEY-----YWSGHIMAFVWVWVGLNGLRGSTYDAMVKNFRFRREMTLTVDLVLVLPFY  
310 320 330 340 350 360  
MON863.pap  
I39838  
410 420 430 440 450 460  
NTDVAAMPNGKVLVGVKVSQYDDK---NETSTQYVDSKRMHVSQAQSDIDOLPPE  
ST-LYRRP---FNIGINNQLSVLDOTEFAYGTSNLSNPVYKSGTV---DSLDIEPPO  
370 380 390 400 410 420  
MON863.pap  
I39838  
470 480 490 500 510  
TDEPLEKAYSHQLYABCFMQDRRTI---PFTVTHRSVDFNTIDAEKITQLPV  
NNVPPRQGFSGRLSHVWVWVGLNGLRGSTYDAMVKNFRFRREMTLTVDLVLVLPFY  
420 430 440 450 460 470  
MON863.pap  
I39838  
520 530 540 550 560 570  
VKAVALSRGATIRGPGFQGNLFLKESNSSTAKFVTLNSAALLQYVVRIRYASTIN  
TKSTNLSGTSVVGPGTGGDLR-RTSPOQISTLRVNI-TAPLSQRYVVRIRYASTIN  
480 490 500 510 520 530  
MON863.pap  
I39838  
580 590 600 610 620 630  
LRLVQNSNNDPLVIYINKPKDDLTQYPTFLATNSNMGFSGDNELIIGASPVSN  
LAPVCHMORPINCQFPAHNSGSLGSGFPHLQYPTFLATNSNMGFSGDNELIIGASPVSN  
540 550 560 570 580 590  
MON863.pap  
I39838  
640 650  
RKVIIDKIEPIVQL  
NEVIDIRFVPARVTPFAEDYERAKAVNELFTSSNQIQLTVDYHIDQVSNLVECLSD  
600 610 620 630 640 650  
MON863.pap  
TO: Q03748  
ID Q03748 PRELIMINARY, PKT 1181 AA  
AC Q03748  
DT 01-NOV-1996 (TRIMBLREL, 01, CREATED)  
DT 01-NOV-1996 (TRIMBLREL, 01, LAST SEQUENCE UPDATE)  
DT 01-FEB-1997 (TRIMBLREL, 02, LAST ANNOTATION UPDATE)  
DE 134 KD CRYSTAL PROTEIN (DELTA ENDOTOXIN) (CRYSTALLINE ENTOMOXIN) . . .  
SCORES Intit. 350 Inntn. 959 Opt. 1067 z-score: 1056.6 E( ) 2 2e-53  
Smith-Waterman score: 1139; 34.5% identity in 582 aa overlap  
MON863.pap  
I39838  
660 670 680 690 700  
BSTEVLNDSVVKDAVGTQISVVGQILGVVGFAGALSTFYQSFINTW-----PSADPW  
BACETNRD12\_1 VEVLOGERITQYTPIDISLSTQFLSRFPVAG---FVLGLDILWGTPQDQW  
30 40 50 60 70  
MON863.pap  
I39838  
710 720 730 740 750 760  
KAPMAQVVLIDKIEYPSKALAEQLQGNFEDYVALNSWKTPLSLKRSQORREL  
DAFVLOGERITQYTPIDISLSTQFLSRFPVAG---FVLGLDILWGTPQDQW  
70 80 90 100 110 120  
MON863.pap  
I39838  
170 180 190 200 210 220  
RELPSQASHFRNPKSPFAVSKFVLFTYAQAANTHLLKDAQVSGWSSYSDVA  
O-FNDMSALITATPLAVQVGVLLSVYQANLHLSVLKDVSTFGQVQFDAATTN  
140 150 160 170 180 190  
MON863.pap  
I39838  
230 240 250 260 270 280  
EFTYRQLKLTQYTHVAVWVWVGLNGLRGSTYDAMVKNFRFRREMTLTVDLVLVLPFY  
SRVNDLTALIGNTYTHAVRWVWVGLNGLRGSTYDAMVKNFRFRREMTLTVDLVLVLPFY  
200 210 220 230 240 250

MON863.pap  
I39838  
360 370 380 390 400 409  
YFGKDSFNYNGVYVTRPSIGSSNTISPPYCDK-STEPVQKLSFD-GQKYVRTIANTD  
YF-----YWSGHIMAFVWVWVGLNGLRGSTYDAMVKNFRFRREMTLTVDLVLVLPFY  
320 330 340 350 360  
MON863.pap  
I39838  
410 420 430 440 450 460  
VAAMFNGKVLVGVKVSQYDDK---NETSTQYVDSKRMHVSQAQSDIDOLPPE  
---FYRNPFIIGINQLSVLDOTEFAYGTSNLSNPVYKSGTV---DSLDIEPPO  
370 380 390 400 410 420  
MON863.pap  
I39838  
470 480 490 500 510 520  
EPLKAYSHQLYABCFMQDRRTI---PFTVTHRSVDFNTIDAEKITQLPVKA  
VPRQGFSGRLSHVWVWVGLNGLRGSTYDAMVKNFRFRREMTLTVDLVLVLPFY  
420 430 440 450 460 470  
MON863.pap  
I39838  
530 540 550 560 570 580  
YALSGAIIISGPGTGGDLR-RTSPOQISTLRVNI-TAPLSQRYVVRIRYASTIN  
TNLSGTSVVGPGTGGDLR-RTSPOQISTLRVNI-TAPLSQRYVVRIRYASTIN  
480 490 500 510 520 530  
MON863.pap  
I39838  
590 600 610 620 630 640  
FVQNSNNDPLVIYINKPKDDLTQYPTFLATNSNMGFSGDNELIIGASPVSNKI  
ITSDIRPINCQFPAHNSGSLGSGFPHLQYPTFLATNSNMGFSGDNELIIGASPVSN  
540 550 560 570 580 590  
MON863.pap  
I39838  
650  
YIDKIEPIVQL  
YIDIRFVPARVTPFAEDYERAKAVNELFTSSNQIQLTVDYHIDQVSNLVECLSD  
600 610 620 630 640 650  
MON863.pap  
TO: A41052  
P1.A41052 - parasporal crystal protein - Bacillus thuringiensis (strain  
Aventi)  
C/Speckles. Bacillus thuringiensis  
C/Date. 03-Apr-1992 #sequence\_revision 03-Apr-1992 #text\_change 08-Sep-1997  
C/Accession. A41052  
R/Lee, C.S.; Aronson, A.I.  
J. Bacteriol. 173, 6635-6638, 1991 . . .  
SCORES Intit. 350 Inntn. 959 Opt. 1067 z-score: 1056.6 E( ) 2 2e-53  
Smith-Waterman score: 1139; 34.5% identity in 582 aa overlap  
MON863.pap  
I39838  
660 670 680 690 700  
BSTEVLNDSVVKDAVGTQISVVGQILGVVGFAGALSTFYQSFINTW-----PSADPW  
BACETNRD12\_1 VEVLOGERITQYTPIDISLSTQFLSRFPVAG---FVLGLDILWGTPQDQW  
30 40 50 60 70  
MON863.pap  
I39838  
710 720 730 740 750 760  
KAPMAQVVLIDKIEYPSKALAEQLQGNFEDYVALNSWKTPLSLKRSQORREL  
DAFVLOGERITQYTPIDISLSTQFLSRFPVAG---FVLGLDILWGTPQDQW  
70 80 90 100 110 120  
MON863.pap  
I39838  
170 180 190 200 210 220  
RELPSQASHFRNPKSPFAVSKFVLFTYAQAANTHLLKDAQVSGWSSYSDVA  
O-FNDMSALITATPLAVQVGVLLSVYQANLHLSVLKDVSTFGQVQFDAATTN  
140 150 160 170 180 190  
MON863.pap  
I39838  
230 240 250 260 270 280  
EFTYRQLKLTQYTHVAVWVWVGLNGLRGSTYDAMVKNFRFRREMTLTVDLVLVLPFY  
SRVNDLTALIGNTYTHAVRWVWVGLNGLRGSTYDAMVKNFRFRREMTLTVDLVLVLPFY  
200 210 220 230 240 250



MON863.pap 360 370 380 390 400 409  
YVQKSNVYVNSVYVETASIGSMTTTPPYGDK-STFVQKLEFD-GQKVRTIANTD  
A41052 YY-----YVSGHQTASPVGFSAPEFTTLYGTGMAAPQORIVAGLQGVYRTLST-

MON863.pap 410 420 430 440 450 460  
VAAMNKGVLGVTKVPSODDK---NETSTQTYDENKNNHVSQAQSDQLPPETTD  
A41052 ---PVRNPTTGINNRLSLVDGEPAYGSENLPSAVTKGT---DELSDIPQDNH

MON863.pap 470 480 490 500 510 520  
HPLREKAYSHQLNVAEFLMDORGTI---PPTVTHRSVDFPMTIDAEKITQLP  
A41052 VPPRQPSHRLSHVMSFASQNSSVETRAKPMFWHRAKNNILPSSQITQLTKS

MON863.pap 530 540 550 560 570 580  
YALSSQASSTZBOPPTGGILFLKESNSIAKFKVTINSAALLQRYVRIRYASTT  
A41052 TNLGSGTVVKQPOPTGGDLR-RTSPQGITLKVNI-TAPLQRYVRIRYASTTQF

MON863.pap 590 600 610 620 630 640  
FVQNNNDPLVIYINKTKKDDLTQYQDPLATNSNGFSDKNKDCGAKHNVNKKI  
A41052 HPSIDGRPIQGNFATMSGQNGQSGSFRTVPTTPPNFSGNSVPTLSNVSGDNY

MON863.pap 650  
YIDKINPIPVOL  
A41052 YIDRIEFPARTFPAEYDLERAQAVNALTFEPNQIGLKTVDYHIDQVNLVECLSD

MON863.pap 660 610 620 630 640 650

LOCUS BACENIAB\_1  
DEFINITION Bacillus thuringiensis alesti delta endotoxin gene, complete  
cds, DATE 16-MAY-1991  
ACCESSION M56252  
MID g142874  
ORGANISM Bacillus thuringiensis . . .

SCORES Initl: 350 Initn: 959 Opt: 1067 z-score: 1056.6 E(): 2.2e-53  
Smith-Waterman score 1139, 34.5% identity in 582 aa overlap

MON863.pap 60 70 80 90 100 110  
SSTEVLNMTVKDAMVTGISVVGQILGVGVFFAGALTSFYQSFINTIMPDAAD---PWAKFAQVE  
BACENIAB\_1 VEVLGGERITGYTIDISLALTOPLLSFVPGAG---FVLGLDLNGLVQPSQWDAF

MON863.pap 120 130 140 150 160 170  
MAZVVLIDKKIEYAKSALAKLQGLQNNFEDYVNLNSWKKTPSLSRKRSQRTIRRL  
BACENIAB\_1 LVQIQGLIEYAKSALAKLQGLQNNFEDYVNLNSWKKTPSLSRKRSQRTIRRL

MON863.pap 180 190 200 210 220 230  
FSQARSHFNKSPFVSKFVFLPTQAQANTHLLLLKDAQVFGESWYSSDVAEYF  
BACENIAB\_1 FNDMSALTAIPVQYQVPLLEVTVQAVNLHSLVLDVSVVQGNHGLDVATINSRY

MON863.pap 240 250 260 270 280 290  
RAQLKLTQYTCVNMVNLGSLSTYDAMVFNKFARENTLVLDLVLPFYDIAL  
BACENIAB\_1 NDLRLQYTCVNMVNLGSLSTYDAMVFNKFARENTLVLDLVLPFYDIAL

MON863.pap 300 310 320 330 340 350  
LYSKGVKTELZRDIFDPIFLTLQKYGPTFLSIENSIRKPHLFDYLGQIEFTRLRPQ  
BACENIAB\_1 TPIRTVSLTREITVFP-LNFTDQSGSAQIRISQISPHMADLNSITTYTDAHQ

MON863.pap 360 370 380 390 400 409  
YVQKSNVYVNSVYVETASIGSMTTTPPYGDK-STFVQKLEFD-GQKVRTIANTD  
BACENIAB\_1 YY-----YVSGHQTASPVGFSAPEFTTLYGTGMAAPQORIVAGLQGVYRTLST-

MON863.pap 410 420 430 440 450 460  
VAAMNKGVLGVTKVPSODDK---NETSTQTYDENKNNHVSQAQSDQLPPETTD  
BACENIAB\_1 ---PVRNPTTGINNRLSLVDGEPAYGSENLPSAVTKGT---DELSDIPQDNH

MON863.pap 470 480 490 500 510 520  
HPLREKAYSHQLNVAEFLMDORGTI---PPTVTHRSVDFPMTIDAEKITQLP  
BACENIAB\_1 VPPRQPSHRLSHVMSFASQNSSVETRAKPMFWHRAKNNILPSSQITQLTKS

MON863.pap 530 540 550 560 570 580  
YALSSQASSTZBOPPTGGILFLKESNSIAKFKVTINSAALLQRYVRIRYASTT  
BACENIAB\_1 TNLGSGTVVKQPOPTGGDLR-RTSPQGITLKVNI-TAPLQRYVRIRYASTTQF

MON863.pap 590 600 610 620 630 640  
FVQNNNDPLVIYINKTKKDDLTQYQDPLATNSNGFSDKNKDCGAKHNVNKKI  
BACENIAB\_1 HPSIDGRPIQGNFATMSGQNGQSGSFRTVPTTPPNFSGNSVPTLSNVSGDNY

MON863.pap 650  
YIDKINPIPVOL  
BACENIAB\_1 YIDRIEFPARTFPAEYDLERAQAVNALTFEPNQIGLKTVDYHIDQVNLVECLSD

MON863.pap 660 610 620 630 640 650

ID 032307 PRELIMINARY, PRT: 618 AA  
AC 032407 DT 01-JAN-1999 (TRANSLATED, 05, CREATED)  
DT 01-JAN-1998 (TRANSLATED, 05, LAST SEQUENCE UPDATE)  
DT 01-JAN-1988 (TRANSLATED, 05, LAST ANNOTATION UPDATE)  
DE MORQUITOCIDAL TOXIN .

SCORES Initl: 226 Initn: 752 Opt: 1051 z-score: 1045.3 E(): 9.2e-53  
Smith-Waterman score: 1076; 33.4% identity in 671 aa overlap

MON863.pap 10 20 30 40 50 59  
MANPNNSHDIITKVTPN-SELQTHNNOYPLADNPNSTLEELNPKYKSLHMTDGFSEVLD  
032307 MHTYQNNHNTILMAENSDNNTYPTPLANPOGLKMTYVLAQVYKYNHNER

MON863.pap 60 70 80 90 100 110  
NSTVKDAVGTGIVSVVGQILGVGVFFAGALTSFYQSFINTIMPDAAD---PWAKFAQVE  
032307 KASVRAGLQGLQIVSTVGFPGGSIILDTGLFYQIRLHNPEDDQVYQVQIMHVS

MON863.pap 120 130 140 150 160 170  
VLIDKKIEYAKSALAKLQGLQNNFEDYVNLNSWKKTPSLSRKRSQRTIRRL  
032307 DLIDKKIEYAKSALAKLQGLQNNFEDYVNLNSWKKTPSLSRKRSQRTIRRL

MON863.pap 180 190 200 210 220 230  
SHFNKSPF---AVSKFVFLPTQAQANTHLLLLKDAQVFGESWYSSDVAEYF  
032307 SDFMKARTVNSGSPYSLLELLPVYAGIANLHLLLRDAQVYQVQIMHVS

MON863.pap 240 250 260 270 280 290  
RAQLKLTQYTCVNMVNLGSLSTYDAMVFNKFARENTLVLDLVLPFYDIAL  
032307 IQLEKTKYETCYVNLNGLNDFR-TAGQVNFNRYKRENTLVLDLVLPFYDIAL

MON863.pap 300 310 320 330 340 350  
YSKGVKTELZRDIFDPIFLTLQKYGPTFLSIENSIRKPHLFDYLGQIEFTRLRPQ  
032307 YPTVKTELZRDIFDPIFLTLQKYGPTFLSIENSIRKPHLFDYLGQIEFTRLRPQ

MON863.pap 360 370 380 390 400 409  
YVQKSNVYVNSVYVETASIGSMTTTPPYGDK-STFVQKLEFD-GQKVRTIANTD  
032307 NSISVYVFLSGQNKSVYTHNSNTHR---GFRQDQDYQVQSTINPNSNY---VY

MON863.pap 410 420 430 440 450 460  
RTIANTOVAAMNKGVLGVTKVPSODDKNETSTQTYDENKNNHVSQAQSDQLP-  
032307 NLNTENYVETYPWGDG-VNITKXNFVTDN--NSSKELIYGAHRTNPPVTRD-PDLTN

MON863.pap 470 480 490 500 510  
PETTDSPLEKAYSHQLN---ABCFIMQORGTIPPTVTHRSVDFPMTIDAEKITQLP  
032307 KEGTSLAKYNDYVNLSTYMLNORTF-GQVRRHVS--FAPTHSSVDNNTIAAKITQIP

MON863.pap 520 530 540 550 560 570  
VVKAYALSSQASSTZBOPPTGGILFLKESNSIAKFKVTINSAALLQRYVRIRYASTT  
032307 VVKASSIMSISIEKOPPTGGDLVQMRADBSLTKRFAEL----LKKYVRIRYKCHY

MON863.pap 580 590 600 610 620 630  
NRLFLVQN-SMNDPLVIYINKTKKDDLTQYQDPLATNSNGFSDKNKDCGAKHNVNKKI  
032307 SSKILIRKMKBOYIQQIHNIISPTYGAFSYLESFTTITTEITFDLMEITYPY--GRQF

MON863.pap 640 650  
VEN-EKIVIDKIRPIVOL  
032307 VEDIPSLILKIEFLPN

MON863.pap 660

LOCUS BTJ0665\_1  
DEFINITION R.thuringiensis cry19a gene & ORF2.  
DATE 14-NOV-1997  
ACCESSION Y07603  
MID g264004  
ORGANISM Bacillus thuringiensis . .

SCORES Initl: 226 Initn: 752 Opt: 1051 z-score: 1045.3 E(): 9.2e-53  
Smith-Waterman score 1076; 33.4% identity in 671 aa overlap

MON863.pap 10 20 30 40 50 59  
MANPNNSHDIITKVTPN-SELQTHNNOYPLADNPNSTLEELNPKYKSLHMTDGFSEVLD  
BTJ0665\_1 MHTYQNNHNTILMAENSDNNTYPTPLANPOGLKMTYVLAQVYKYNHNER

MON863.pap 60 70 80 90 100 110  
NSTVKDAVGTGIVSVVGQILGVGVFFAGALTSFYQSFINTIMPDAAD---PWAKFAQVE  
BTJ0665\_1 KASVRAGLQGLQIVSTVGFPGGSIILDTGLFYQIRLHNPEDDQVYQVQIMHVS

MON863.pap 120 130 140 150 160 170  
VLIDKKIEYAKSALAKLQGLQNNFEDYVNLNSWKKTPSLSRKRSQRTIRRL  
BTJ0665\_1 DLIDKKIEYAKSALAKLQGLQNNFEDYVNLNSWKKTPSLSRKRSQRTIRRL

MON863.pap 180 190 200 210 220 230  
SHFNKSPF---AVSKFVFLPTQAQANTHLLLLKDAQVFGESWYSSDVAEYF  
BTJ0665\_1 SDFMKARTVNSGSPYSLLELLPVYAGIANLHLLLRDAQVYQVQIMHVS

MON863.pap 240 250 260 270 280 290  
RAQLKLTQYTCVNMVNLGSLSTYDAMVFNKFARENTLVLDLVLPFYDIAL  
BTJ0665\_1 IQLEKTKYETCYVNLNGLNDFR-TAGQVNFNRYKRENTLVLDLVLPFYDIAL

MON863.pap 300 310 320 330 340 350  
YSKGVKTELZRDIFDPIFLTLQKYGPTFLSIENSIRKPHLFDYLGQIEFTRLRPQ  
BTJ0665\_1 YPTVKTELZRDIFDPIFLTLQKYGPTFLSIENSIRKPHLFDYLGQIEFTRLRPQ

MON863.pap 360 370 380 390 400 409  
YVQKSNVYVNSVYVETASIGSMTTTPPYGDK-STFVQKLEFD-GQKVRTIANTD  
BTJ0665\_1 NSISVYVFLSGQNKSVYTHNSNTHR---GFRQDQDYQVQSTINPNSNY---VY



Page 67 of 102

580 590 600 610 620 630  
MON863.psp TTNLRLLFVQNSNDPFLVIYINKTKDKDDLTQYOTFLATTNSNGMFSOGKNEILIGAESF  
AC P05068, DT 13-AUG-1987 (REL 06, CREATED)  
DT 13-AUG-1987 (REL 05, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL 35, LAST ANNOTATION UPDATE)  
DE 133 KD CRYSTAL PROTEIN (DELTA ENDOTOXIN) (CRYSTALLINE ENTEROCIDIN)  
SCORES Initl 402 Initn 877 Opt 1038 z-score: 1027.8 E(): 8.6e-52  
Smith-Waterman score: 1107, 35.0% identity in 586 aa overlap  
60 70 80 90 100  
MON863.psp SSTEVLNNTVKDAVGTGISVVGQIGLVGVVFPAGALTSTFYQSFNTIM----PSDADPW  
CRYM\_BACTK VEVLQGRIRITGYTPIDISLSLTQFLLESFVPGAG---FVLGLVDITWIGIFGPGQ---W  
110 120 130 140 150 160  
MON863.psp KAPMAQVEVLIDKKIRYAKSKALAELOQLNNFEDYVNALNSWKKTPLSLRSKRSQGR  
CRYM\_BACTK DAFVLQIGLQRIEIRFARNGAISRLEGLSNLQIYAFSFRWEDPTN-PALREEMRI  
170 180 190 200 210 220  
MON863.psp RELFSCAESFRNSMSPFAVSKFEVLPLPTVQAANTHLLKDAQVFGESWYSSEDA  
CRYM\_BACTK Q--FNDKMSALTAITPLFVAVNYQVPLLSVTVQANHLISVLRLDVSFQGRWGAATIN  
230 240 250 260 270 280  
MON863.psp EFRYRQLKLTQOITDCHVWYVNLGLRSTYDAWVKNFRREMTLVLDLIVLPFY  
CRYM\_BACTK SYNDLTKLIGNTYAVWYVNLGLRSTYDAWVKNFRREMTLVLDLIVLPFY  
290 300 310 320 330 340  
MON863.psp DIRLSKQVKTETLTHDPTDPIFLTLQYQVPTFLIENSIRKPHLPDYLQIEFTRL  
CRYM\_BACTK DSRVYPIRTVSQLTREIYTNV-LNFDSFPGSAQGISRSIRSPHMLINSITITIDA  
350 360 370 380 390 400  
MON863.psp RPYVFOKDSFNYWNSYVSTRPSIGSKTITSPFYQDK-STEPVQKLSFD-CQKYNTIA  
CRYM\_BACTK HROYI-----YNSHQIMASPVVPSQPEPTFLYQTMGNAPQQRIVAGLQGVYRLS  
410 420 430 440 450 460  
MON863.psp NTVAAWPKGKVLGVTVKVSQYDDQK---NRTSTQYDSKNNHVSQAQSDIQLPPE  
CRYM\_BACTK ST-LYRPP---FNIGINNOQLSVLDTEFAVOTSSMLPFAVYKSGTV---DRLDIPPO  
470 480 490 500 510  
MON863.psp TIDPLEKAYSHQYFARCPIMQDRGII-----PFTVTKRISVDFPNTDAKIKTQLPV  
CRYM\_BACTK NNWVPRQSGHLSHVENRSGFSENGEIVIRAPFVSNIRASRFPNIIASDSITQIPA  
520 530 540 550 560 570  
MON863.psp VKAYALSQAISIEGPOPTGCHLLFKESNSIAK---FKVTLNSAALLQRYRVRIRYAS  
CRYM\_BACTK VGNVLFNG-SVLSQPGFQGGDLVNLNNSGNTQNGRQYEVPIRFTSTSTRYVVRIRYAS  
580 590 600 610 620 630  
MON863.psp TTNLRLLFVQNSNDPFLVIYINKTKDKDDLTQYOTFLATTNSNGMFSOGKNEILIGAESF  
CRYM\_BACTK VTPHNLVWNGNSISFNTVPATATSLNL--QSSDFGYFESANFTSGLN-IVGVNRP

640 650  
MON863.psp VSNKIVIDKIEFIPVQL  
CRYM\_BACTK SGTAGVILDRFIEPTVATLEAEYNLERAKAVNALPTSTKQLGLKTNVTDVHIDQVSNL  
MON863.psp TO.BTU87793\_1  
LOCUS BTU87793.1  
DEFINITION Bacillus thuringiensis kurstaki insecticidal delta-endotoxin  
CRYIA(c) (cryIA(c)) gene, complete cds.  
DATE 12-FEB-1997, ACCESSION U87793  
NID g1839245  
SCORES Initl 402 Initn 877 Opt 1038 z-score: 1027.8 E(): 8.6e-52  
Smith-Waterman score: 1107, 35.0% identity in 586 aa overlap  
60 70 80 90 100  
MON863.psp SSTEVLNNTVKDAVGTGISVVGQIGLVGVVFPAGALTSTFYQSFNTIM----PSDADPW  
BTU87793\_1 VEVLQGRIRITGYTPIDISLSLTQFLLESFVPGAG---FVLGLVDITWIGIFGPGQ---W  
110 120 130 140 150 160  
MON863.psp KAPMAQVEVLIDKKIRYAKSKALAELOQLNNFEDYVNALNSWKKTPLSLRSKRSQGR  
BTU87793\_1 DAFVLQIGLQRIEIRFARNGAISRLEGLSNLQIYAFSFRWEDPTN-PALREEMRI  
170 180 190 200 210 220  
MON863.psp RELFSCAESFRNSMSPFAVSKFEVLPLPTVQAANTHLLKDAQVFGESWYSSEDA  
BTU87793\_1 Q--FNDKMSALTAITPLFVAVNYQVPLLSVTVQANHLISVLRLDVSFQGRWGAATIN  
230 240 250 260 270 280  
MON863.psp EFRYRQLKLTQOITDCHVWYVNLGLRSTYDAWVKNFRREMTLVLDLIVLPFY  
BTU87793\_1 EFRYRQLKLTQOITDCHVWYVNLGLRSTYDAWVKNFRREMTLVLDLIVLPFY  
290 300 310 320 330 340  
MON863.psp DIRLSKQVKTETLTHDPTDPIFLTLQYQVPTFLIENSIRKPHLPDYLQIEFTRL  
BTU87793\_1 DSRVYPIRTVSQLTREIYTNV-LNFDSFPGSAQGISRSIRSPHMLINSITITIDA  
350 360 370 380 390 400  
MON863.psp RPYVFOKDSFNYWNSYVSTRPSIGSKTITSPFYQDK-STEPVQKLSFD-CQKYNTIA  
BTU87793\_1 HROYI-----YNSHQIMASPVVPSQPEPTFLYQTMGNAPQQRIVAGLQGVYRLS  
410 420 430 440 450 460  
MON863.psp NTVAAWPKGKVLGVTVKVSQYDDQK---NRTSTQYDSKNNHVSQAQSDIQLPPE  
BTU87793\_1 ST-LYRPP---FNIGINNOQLSVLDTEFAVOTSSMLPFAVYKSGTV---DRLDIPPO  
470 480 490 500 510  
MON863.psp TIDPLEKAYSHQYFARCPIMQDRGII-----PFTVTKRISVDFPNTDAKIKTQLPV  
BTU87793\_1 TIDPLEKAYSHQYFARCPIMQDRGII-----PFTVTKRISVDFPNTDAKIKTQLPV  
520 530 540 550 560 570  
MON863.psp VKAYALSQAISIEGPOPTGCHLLFKESNSIAK---FKVTLNSAALLQRYRVRIRYAS  
BTU87793\_1 VKAYALSQAISIEGPOPTGCHLLFKESNSIAK---FKVTLNSAALLQRYRVRIRYAS  
580 590 600 610 620 630  
MON863.psp TTNLRLLFVQNSNDPFLVIYINKTKDKDDLTQYOTFLATTNSNGMFSOGKNEILIGAESF  
BTU87793\_1 TTNLRLLFVQNSNDPFLVIYINKTKDKDDLTQYOTFLATTNSNGMFSOGKNEILIGAESF  
640 650  
MON863.psp VSNKIVIDKIEFIPVQL  
BTU87793\_1 SGTAGVILDRFIEPTVATLEAEYNLERAKAVNALPTSTKQLGLKTNVTDVHIDQVSNL

MON863.psp TO.BTU89872\_1  
LOCUS BTU89872.1  
DEFINITION Bacillus thuringiensis CryIAC delta-endotoxin gene, complete  
cds.  
DATE 18-MAR-1997  
ACCESSION U89872  
NID g188858  
ORGANISM Bacillus thuringiensis  
SCORES Initl 402 Initn 877 Opt 1038 z-score: 1027.8 E(): 8.6e-52  
Smith-Waterman score: 1107, 35.0% identity in 586 aa overlap  
60 70 80 90 100  
MON863.psp SSTEVLNNTVKDAVGTGISVVGQIGLVGVVFPAGALTSTFYQSFNTIM----PSDADPW  
BTU89872\_1 VEVLQGRIRITGYTPIDISLSLTQFLLESFVPGAG---FVLGLVDITWIGIFGPGQ---W  
110 120 130 140 150 160  
MON863.psp KAPMAQVEVLIDKKIRYAKSKALAELOQLNNFEDYVNALNSWKKTPLSLRSKRSQGR  
BTU89872\_1 DAFVLQIGLQRIEIRFARNGAISRLEGLSNLQIYAFSFRWEDPTN-PALREEMRI  
170 180 190 200 210 220  
MON863.psp RELFSCAESFRNSMSPFAVSKFEVLPLPTVQAANTHLLKDAQVFGESWYSSEDA  
BTU89872\_1 Q--FNDKMSALTAITPLFVAVNYQVPLLSVTVQANHLISVLRLDVSFQGRWGAATIN  
230 240 250 260 270 280  
MON863.psp EFRYRQLKLTQOITDCHVWYVNLGLRSTYDAWVKNFRREMTLVLDLIVLPFY  
BTU89872\_1 EFRYRQLKLTQOITDCHVWYVNLGLRSTYDAWVKNFRREMTLVLDLIVLPFY  
290 300 310 320 330 340  
MON863.psp DIRLSKQVKTETLTHDPTDPIFLTLQYQVPTFLIENSIRKPHLPDYLQIEFTRL  
BTU89872\_1 DSRVYPIRTVSQLTREIYTNV-LNFDSFPGSAQGISRSIRSPHMLINSITITIDA  
350 360 370 380 390 400  
MON863.psp RPYVFOKDSFNYWNSYVSTRPSIGSKTITSPFYQDK-STEPVQKLSFD-CQKYNTIA  
BTU89872\_1 HROYI-----YNSHQIMASPVVPSQPEPTFLYQTMGNAPQQRIVAGLQGVYRLS  
410 420 430 440 450 460  
MON863.psp NTVAAWPKGKVLGVTVKVSQYDDQK---NRTSTQYDSKNNHVSQAQSDIQLPPE  
BTU89872\_1 ST-LYRPP---FNIGINNOQLSVLDTEFAVOTSSMLPFAVYKSGTV---DRLDIPPO  
470 480 490 500 510  
MON863.psp TIDPLEKAYSHQYFARCPIMQDRGII-----PFTVTKRISVDFPNTDAKIKTQLPV  
BTU89872\_1 TIDPLEKAYSHQYFARCPIMQDRGII-----PFTVTKRISVDFPNTDAKIKTQLPV  
520 530 540 550 560 570  
MON863.psp VKAYALSQAISIEGPOPTGCHLLFKESNSIAK---FKVTLNSAALLQRYRVRIRYAS  
BTU89872\_1 VKAYALSQAISIEGPOPTGCHLLFKESNSIAK---FKVTLNSAALLQRYRVRIRYAS  
580 590 600 610 620 630  
MON863.psp TTNLRLLFVQNSNDPFLVIYINKTKDKDDLTQYOTFLATTNSNGMFSOGKNEILIGAESF  
BTU89872\_1 TTNLRLLFVQNSNDPFLVIYINKTKDKDDLTQYOTFLATTNSNGMFSOGKNEILIGAESF  
640 650  
MON863.psp VSNKIVIDKIEFIPVQL  
BTU89872\_1 SGTAGVILDRFIEPTVATLEAEYNLERAKAVNALPTSTKQLGLKTNVTDVHIDQVSNL

MON863 pep  
TO: ASCYRIC\_1  
LOCUS ASCYRIC\_1  
DEFINITION Artificial sequences mRNA for delta-endotoxin.  
DATE 20-JAN-1997  
ACCESSION X94103  
NID g120466  
ORGANISM synthetic construct

SCORES Initl: 445 Initn: 931 Opt: 1033 z-score: 1026.3 E(): 1.1e-51  
Smith-Waterman score: 1179, 37.2% identity in 646 aa overlap

MON863 pep 10 20 30 40 50 60  
SEHDITKVPNSLQTNHNGYPLADNPSTLSENYKHPKMTDSETE--VLN--ST  
ASCYRIC\_1 MEENNQNCIYVNCISNPBVLDDGERIST

MON863 pep 70 80 90 100 110 120  
VKDAVGTGIVVQGLGVGVFPAGALTSFYQSFLTWIPSDADPKAFMAQVVLIDKK  
ASCYRIC\_1 QNSSDIDSLGV-QFLVSNVPGGGLVGLDFVWGLVQVQ--WDAFLVGLSOLINER

MON863 pep 130 140 150 160 170 180  
ISEYAKKALAELOQONFEDYVNALNSMKKTPLSLSKRSQGRIRLSFQASHFRNS  
ASCYRIC\_1 IAEFARNAAIANLEGLONFNIVYFAFKWEEDP--NNPATRTVIRFDRILOLLESD

MON863 pep 190 200 210 220 230 240  
MSEFAVKEFVLPLTYAQAANTHLLKDAQVGEWGYSSDVAEFPYRQKLTOQYT  
ASCYRIC\_1 IPSPRISEFVPLGVYAQAANLHAIILRSDVPCRWGLTINNNVYALIRHIDEYA

MON863 pep 250 260 270 280 290 300  
DHCANVYVGLNGLSGSTYDANVKNFRRENFLTLVDLVLFPFYDIRLSKGVKTEL  
ASCYRIC\_1 DHCANTYVGLNGLSGSTYDANVKNFRRENFLTLVDLVLFPFYDIRLSKGVKTEL

MON863 pep 310 320 330 340 350  
RDITPDIFLLIT-LQYK--PTFLSIENS-IRKPHLPDLOQIEPHTLRNAYQKDSF  
ASCYRIC\_1 REVYTDPLINFPOLQVQVLPFTNVMESSATRNPLFDILNLTFT--DWFVGRN

MON863 pep 360 370 380 390 400 410  
MYNSGVYETRPISGSKTITSPFYQ-DKSTFVOKLSDQKVVYRTIANTDV--AAW  
ASCYRIC\_1 PYNGHVRVIS--SLIGGNTSPYVGRANQCPSPSTPNP-VPTLGNPTLALLQCPW

MON863 pep 420 430 440 450 460 470  
PNKRVYL-GVTKVDFQYVDQKNETSTQYVDSKRNHGVSAQSDIDQLPETTDEPLEKA  
ASCYRIC\_1 PAPPFNLKGVGVFE--TTNSFTY--RGRGTV--DSLTLFPEDNSVPPREG

MON863 pep 480 490 500 510 520  
YSHQLNAYCFIMQDRROTIFP--PTWTHRSVDFPNTIDAEKITQLPVVYKAYALS  
ASCYRIC\_1 YSHQLNAYCFIMQDRROTIFP--PTWTHRSVDFPNTIDAEKITQLPVVYKAYALS

MON863 pep 530 540 550 560 570 580  
QASITDOPGPGGGLFLKSSNIAKFKVTLKSAALLQYVVRVRYASTVTLRLPVQNS  
ASCYRIC\_1 QTSVITGPGGGLFLKSSNIAKFKVTLKSAALLQYVVRVRYASTVTLRLPVQNS

MON863 pep 590 600 610 620 630  
NNDPLV--IYIN--KTMKDDDLTYQTFPLATTNSNGFS--GDKNELIGAE  
ASCYRIC\_1 AASTGVGVGVVMPLOKMEIGENLSTFTFNSFRANPDIGISGQLPFGAG

MON863 pep 640 650  
SFVSNKVIYDKIEFTPVQL  
ASCYRIC\_1 SITSGR-LYIDKIEITLADATFEASDLERAQAKNALPFTSGITGDTVDVHIDQV

MON863 pep  
TO: BTA729\_1  
LOCUS BTA729\_1  
DEFINITION B.thuringiensis CryIC gene, ssp aizawai 7.29  
DATE 15-MAY-1996  
ACCESSION X96682  
NID g1234883  
ORGANISM Bacillus thuringiensis

SCORES Initl: 445 Initn: 931 Opt: 1033 z-score: 1026.3 E(): 1.1e-51  
Smith-Waterman score: 1179, 37.2% identity in 646 aa overlap

MON863 pep 10 20 30 40 50 60  
SEHDITKVPNSLQTNHNGYPLADNPSTLSENYKHPKMTDSETE--VLN--ST  
BTA729\_1 MEENNQNCIYVNCISNPBVLDDGERIST

MON863 pep 70 80 90 100 110 120  
VKDAVGTGIVVQGLGVGVFPAGALTSFYQSFLTWIPSDADPKAFMAQVVLIDKK  
BTA729\_1 QNSSDIDSLGV-QFLVSNVPGGGLVGLDFVWGLVQVQ--WDAFLVGLSOLINER

MON863 pep 130 140 150 160 170 180  
ISEYAKKALAELOQONFEDYVNALNSMKKTPLSLSKRSQGRIRLSFQASHFRNS  
BTA729\_1 IAEFARNAAIANLEGLONFNIVYFAFKWEEDP--NNPATRTVIRFDRILOLLESD

MON863 pep 190 200 210 220 230 240  
MSEFAVKEFVLPLTYAQAANTHLLKDAQVGEWGYSSDVAEFPYRQKLTOQYT  
BTA729\_1 IPSPRISEFVPLGVYAQAANLHAIILRSDVPCRWGLTINNNVYALIRHIDEYA

MON863 pep 250 260 270 280 290 300  
DHCANVYVGLNGLSGSTYDANVKNFRRENFLTLVDLVLFPFYDIRLSKGVKTEL  
BTA729\_1 DHCANTYVGLNGLSGSTYDANVKNFRRENFLTLVDLVLFPFYDIRLSKGVKTEL

MON863 pep 310 320 330 340 350  
RDITPDIFLLIT-LQYK--PTFLSIENS-IRKPHLPDLOQIEPHTLRNAYQKDSF  
BTA729\_1 REVYTDPLINFPOLQVQVLPFTNVMESSATRNPLFDILNLTFT--DWFVGRN

MON863 pep 360 370 380 390 400 410  
MYNSGVYETRPISGSKTITSPFYQ-DKSTFVOKLSDQKVVYRTIANTDV--AAW  
BTA729\_1 PYNGHVRVIS--SLIGGNTSPYVGRANQCPSPSTPNP-VPTLGNPTLALLQCPW

MON863 pep 420 430 440 450 460 470  
PNKRVYL-GVTKVDFQYVDQKNETSTQYVDSKRNHGVSAQSDIDQLPETTDEPLEKA  
BTA729\_1 PAPPFNLKGVGVFE--TTNSFTY--RGRGTV--DSLTLFPEDNSVPPREG

MON863 pep 480 490 500 510 520  
YSHQLNAYCFIMQDRROTIFP--PTWTHRSVDFPNTIDAEKITQLPVVYKAYALS  
BTA729\_1 YSHQLNAYCFIMQDRROTIFP--PTWTHRSVDFPNTIDAEKITQLPVVYKAYALS

MON863 pep 530 540 550 560 570 580  
QASITDOPGPGGGLFLKSSNIAKFKVTLKSAALLQYVVRVRYASTVTLRLPVQNS  
BTA729\_1 QTSVITGPGGGLFLKSSNIAKFKVTLKSAALLQYVVRVRYASTVTLRLPVQNS

MON863 pep 590 600 610 620 630  
NNDPLV--IYIN--KTMKDDDLTYQTFPLATTNSNGFS--GDKNELIGAE  
BTA729\_1 AASTGVGVGVVMPLOKMEIGENLSTFTFNSFRANPDIGISGQLPFGAG

MON863 pep 640 650  
SFVSNKVIYDKIEFTPVQL  
BTA729\_1 SITSGR-LYIDKIEITLADATFEASDLERAQAKNALPFTSGITGDTVDVHIDQV

MON863 pep  
TO: Q45725  
ID Q45725 PRELIMINARY, PR: 756 AA  
AC Q45725;  
DT 01-NOV-1996 (TREMBLEL 01, CREATED)  
DT 01-NOV-1996 (TREMBLEL 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1996 (TREMBLEL 01, LAST ANNOTATION UPDATE)  
DE DELTA-ENDOTOXIN (FRAGMENT).

SCORES Initl: 445 Initn: 931 Opt: 1033 z-score: 1026.3 E(): 1.1e-51  
Smith-Waterman score: 1179, 37.2% identity in 646 aa overlap

MON863 pep 10 20 30 40 50 60  
SEHDITKVPNSLQTNHNGYPLADNPSTLSENYKHPKMTDSETE--VLN--ST  
Q45725 MEENNQNCIYVNCISNPBVLDDGERIST

MON863 pep 70 80 90 100 110 120  
VKDAVGTGIVVQGLGVGVFPAGALTSFYQSFLTWIPSDADPKAFMAQVVLIDKK  
Q45725 QNSSDIDSLGV-QFLVSNVPGGGLVGLDFVWGLVQVQ--WDAFLVGLSOLINER

MON863 pep 130 140 150 160 170 180  
ISEYAKKALAELOQONFEDYVNALNSMKKTPLSLSKRSQGRIRLSFQASHFRNS  
Q45725 IAEFARNAAIANLEGLONFNIVYFAFKWEEDP--NNPATRTVIRFDRILOLLESD

MON863 pep 190 200 210 220 230 240  
MSEFAVKEFVLPLTYAQAANTHLLKDAQVGEWGYSSDVAEFPYRQKLTOQYT  
Q45725 IPSPRISEFVPLGVYAQAANLHAIILRSDVPCRWGLTINNNVYALIRHIDEYA

MON863 pep 250 260 270 280 290 300  
DHCANVYVGLNGLSGSTYDANVKNFRRENFLTLVDLVLFPFYDIRLSKGVKTEL  
Q45725 DHCANTYVGLNGLSGSTYDANVKNFRRENFLTLVDLVLFPFYDIRLSKGVKTEL

MON863 pep 310 320 330 340 350  
RDITPDIFLLIT-LQYK--PTFLSIENS-IRKPHLPDLOQIEPHTLRNAYQKDSF  
Q45725 REVYTDPLINFPOLQVQVLPFTNVMESSATRNPLFDILNLTFT--DWFVGRN

MON863 pep 360 370 380 390 400 410  
MYNSGVYETRPISGSKTITSPFYQ-DKSTFVOKLSDQKVVYRTIANTDV--AAW  
Q45725 PYNGHVRVIS--SLIGGNTSPYVGRANQCPSPSTPNP-VPTLGNPTLALLQCPW

MON863 pep 420 430 440 450 460 470  
PNKRVYL-GVTKVDFQYVDQKNETSTQYVDSKRNHGVSAQSDIDQLPETTDEPLEKA  
Q45725 PAPPFNLKGVGVFE--TTNSFTY--RGRGTV--DSLTLFPEDNSVPPREG

MON863 pep 480 490 500 510 520  
YSHQLNAYCFIMQDRROTIFP--PTWTHRSVDFPNTIDAEKITQLPVVYKAYALS  
Q45725 YSHQLNAYCFIMQDRROTIFP--PTWTHRSVDFPNTIDAEKITQLPVVYKAYALS

MON863 pep 530 540 550 560 570 580  
QASITDOPGPGGGLFLKSSNIAKFKVTLKSAALLQYVVRVRYASTVTLRLPVQNS  
Q45725 QTSVITGPGGGLFLKSSNIAKFKVTLKSAALLQYVVRVRYASTVTLRLPVQNS

MON863 pep 590 600 610 620 630  
NNDPLV--IYIN--KTMKDDDLTYQTFPLATTNSNGFS--GDKNELIGAE  
Q45725 AASTGVGVGVVMPLOKMEIGENLSTFTFNSFRANPDIGISGQLPFGAG

MON863 pep 640 650  
SFVSNKVIYDKIEFTPVQL  
Q45725 SITSGR-LYIDKIEITLADATFEASDLERAQAKNALPFTSGITGDTVDVHIDQV



MON863 pep  
TO A49785  
PI: A49785 - parasporal protein cryIIA(c) - *Bacillus thuringiensis*  
auep  
kenyae (strain 80588-2)  
C: Species: *Bacillus thuringiensis* auep, kenyae  
C Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 next\_change 26-Feb-1998  
C Accession: A49785  
R: Von Terwich, M.A.; Robbins, H.L.; Jany, C.S.; Johnson, T.B.

SCORES Initl: 402 Inaln: 912 Opt: 1034 z-score: 1023.9 E(): 1.4e-51  
Smith-Waterman score: 1107; 34.9% identity in 585 aa overlap

MON863.pap  
A49785  
60 70 80 90 100  
SSTEVLNSTVKDAVGTGIVVQVLPVPPAALTSFYQSLNTN---PSDADPW  
VEVLGGRIETGYTPIDLSLSTQFLSEFVPGAG---FVLGLDVIWIFGPGQ---W

MON863.pap  
A49785  
110 120 130 140 150 160  
KAFMAQVVLIDKKISEYAKSKALASLQQLNNFEDYVNALNSNKKTPLSLRKSRGQRI  
DAFLVQTEQLNQRISFARMQATSLRGLNLQIYARSFREWADPTN-PALREMRI

MON863.pap  
A49785  
170 180 190 200 210 220  
RELFSQAESHFNMSPSFAVKFVLPTAQAANTHLLKDAQVFGESWYSSDEVA  
Q-FNDNSALTTATPLAVQNYQVPLSVQAANHLSEVLADSVFGQWGFDAATIN

MON863.pap  
A49785  
230 240 250 260 270 280  
EFYRRLKLTQOYTHCVNMYNGLNGASTYDAWKFRRPREMTLVLDLVLPFFY  
SRYNDLRLIGYTHAVVNTGLSRWQSDRWVYQFRELTLVLDAVALFPY

MON863.pap  
A49785  
290 300 310 320 330 340  
DIRLYSKGVKLTEDIFPILLLTQKYPTPLSTENSRKPHFLVQLQIEFTRL  
DSRRYPTIVSQTAEITVTPV-LNFQSGFSGSAQIESRSPHMDLSTIYTDA

MON863.pap  
A49785  
350 360 370 380 390 400  
RCYFGKDSFYMSQYVTRPSIGSKTTSPPYDNR-STBPVKLSPD-QOKVYRTIA  
HNGYV-----YSGHQIMASPVQSGPEPTPLVYTMNAAPQGRIVAGLQGVYKLS

MON863.pap  
A49785  
410 420 430 440 450 460  
MTDVAAMPNGKVVGLVTVDFSGYDDQK---NETSTQYVDSKRNHVSQAQSDIDLPFR  
ST-----FYRRPFHIGINMOQLSVLQTEFAYGTSENLPASVYKSGTV---DSDIDPQ

MON863.pap  
A49785  
470 480 490 500 510 519  
TTDEPLEKAYSHQANABCFMQRDRTG---TPPTMTLSDVDFNTIDAKKITQGLPV  
NNTVPPPGQFSHLSHVSGRSGSRAVSIIRAFNFIWASAFNFIASDGTQIPAV

MON863.pap  
A49785  
520 530 540 550 560 570  
KAYALSSQASIIQPGTQGNLLFKESNSIAK---FKVTLNSAALLQRYKRYAST  
KGNFLNG-SVLSGPGFTGGDLVRLNSGNNIQQNGYIEVPHFSTSTRYKRVYAV

MON863.pap  
A49785  
580 590 600 610 620 630  
TNLRFLVQSNDFVIYIMTKMKDDLTQTFPLATNMMNMFSGDNELLIGAESV  
TPHILVNMNMSISFEMTVATATSLNL---QSSDFYFESANFTSSLN-IVGVNRS

MON863.pap  
A49785  
640 650  
SNEKIYIDIEFIPVL  
GTAGVINDKFPFVATLEAEYHLRAKAVNALFTSTNGLQKTNVTDYHIDQVSNLV

MON863 pep  
TO Q45735  
ID Q45735 PRELIMINARY; PRT: 1177 AA.  
AC Q45735  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)  
DE DELTA-ENDOTOXIN. . . .

SCORES Initl: 402 Inaln: 912 Opt: 1034 z-score: 1023.9 E(): 1.4e-51  
Smith-Waterman score: 1107; 34.9% identity in 585 aa overlap

MON863 pep  
Q45735  
60 70 80 90 100  
SSTEVLNSTVKDAVGTGIVVQVLPVPPAALTSFYQSLNTN---PSDADPW  
VEVLGGRIETGYTPIDLSLSTQFLSEFVPGAG---FVLGLDVIWIFGPGQ---W

MON863 pep  
Q45735  
110 120 130 140 150 160  
KAFMAQVVLIDKKISEYAKSKALASLQQLNNFEDYVNALNSNKKTPLSLRKSRGQRI  
DAFLVQTEQLNQRISFARMQATSLRGLNLQIYARSFREWADPTN-PALREMRI

MON863 pep  
Q45735  
170 180 190 200 210 220  
RELFSQAESHFNMSPSFAVKFVLPTAQAANTHLLKDAQVFGESWYSSDEVA  
Q-FNDNSALTTATPLAVQNYQVPLSVQAANHLSEVLADSVFGQWGFDAATIN

MON863 pep  
Q45735  
230 240 250 260 270 280  
EFYRRLKLTQOYTHCVNMYNGLNGASTYDAWKFRRPREMTLVLDLVLPFFY  
SRYNDLRLIGYTHAVVNTGLSRWQSDRWVYQFRELTLVLDAVALFPY

MON863 pep  
Q45735  
290 300 310 320 330 340  
DIRLYSKGVKLTEDIFPILLLTQKYPTPLSTENSRKPHFLVQLQIEFTRL  
DSRRYPTIVSQTAEITVTPV-LNFQSGFSGSAQIESRSPHMDLSTIYTDA

MON863 pep  
Q45735  
350 360 370 380 390 400  
RCYFGKDSFYMSQYVTRPSIGSKTTSPPYDNR-STBPVKLSPD-QOKVYRTIA  
HNGYV-----YSGHQIMASPVQSGPEPTPLVYTMNAAPQGRIVAGLQGVYKLS

MON863 pep  
Q45735  
410 420 430 440 450 460  
MTDVAAMPNGKVVGLVTVDFSGYDDQK---NETSTQYVDSKRNHVSQAQSDIDLPFR  
ST-----FYRRPFHIGINMOQLSVLQTEFAYGTSENLPASVYKSGTV---DSDIDPQ

MON863 pep  
Q45735  
470 480 490 500 510 519  
TTDEPLEKAYSHQANABCFMQRDRTG---TPPTMTLSDVDFNTIDAKKITQGLPV  
NNTVPPPGQFSHLSHVSGRSGSRAVSIIRAFNFIWASAFNFIASDGTQIPAV

MON863 pep  
Q45735  
520 530 540 550 560 570  
KAYALSSQASIIQPGTQGNLLFKESNSIAK---FKVTLNSAALLQRYKRYAST  
KGNFLNG-SVLSGPGFTGGDLVRLNSGNNIQQNGYIEVPHFSTSTRYKRVYAV

MON863 pep  
Q45735  
580 590 600 610 620 630  
TNLRFLVQSNDFVIYIMTKMKDDLTQTFPLATNMMNMFSGDNELLIGAESV  
TPHILVNMNMSISFEMTVATATSLNL---QSSDFYFESANFTSSLN-IVGVNRS

MON863 pep  
Q45735  
640 650  
SNEKIYIDIEFIPVL  
GTAGVINDKFPFVATLEAEYHLRAKAVNALFTSTNGLQKTNVTDYHIDQVSNLV

MON863 pep  
TO: BACCRYIA\_1  
LOCUS BACCRYIA\_1  
DEFINITION B.thuringiensis delta-endotoxin gene, complete cds,

delta-endotoxin  
DATE 20-FEB-1991  
ACCESSION M35524; NID q142739.

SCORES Initl: 402 Inaln: 912 Opt: 1034 z-score: 1023.9 E(): 1.4e-51  
Smith-Waterman score: 1107; 34.9% identity in 585 aa overlap

MON863 pep  
BACCRYIA\_1  
60 70 80 90 100  
SSTEVLNSTVKDAVGTGIVVQVLPVPPAALTSFYQSLNTN---PSDADPW  
VEVLGGRIETGYTPIDLSLSTQFLSEFVPGAG---FVLGLDVIWIFGPGQ---W

MON863 pep  
BACCRYIA\_1  
110 120 130 140 150 160  
KAFMAQVVLIDKKISEYAKSKALASLQQLNNFEDYVNALNSNKKTPLSLRKSRGQRI  
DAFLVQTEQLNQRISFARMQATSLRGLNLQIYARSFREWADPTN-PALREMRI

MON863 pep  
BACCRYIA\_1  
170 180 190 200 210 220  
RELFSQAESHFNMSPSFAVKFVLPTAQAANTHLLKDAQVFGESWYSSDEVA  
Q-FNDNSALTTATPLAVQNYQVPLSVQAANHLSEVLADSVFGQWGFDAATIN

MON863 pep  
BACCRYIA\_1  
230 240 250 260 270 280  
EFYRRLKLTQOYTHCVNMYNGLNGASTYDAWKFRRPREMTLVLDLVLPFFY  
SRYNDLRLIGYTHAVVNTGLSRWQSDRWVYQFRELTLVLDAVALFPY

MON863 pep  
BACCRYIA\_1  
290 300 310 320 330 340  
DIRLYSKGVKLTEDIFPILLLTQKYPTPLSTENSRKPHFLVQLQIEFTRL  
DSRRYPTIVSQTAEITVTPV-LNFQSGFSGSAQIESRSPHMDLSTIYTDA

MON863 pep  
BACCRYIA\_1  
350 360 370 380 390 400  
RCYFGKDSFYMSQYVTRPSIGSKTTSPPYDNR-STBPVKLSPD-QOKVYRTIA  
HNGYV-----YSGHQIMASPVQSGPEPTPLVYTMNAAPQGRIVAGLQGVYKLS

MON863 pep  
BACCRYIA\_1  
410 420 430 440 450 460  
MTDVAAMPNGKVVGLVTVDFSGYDDQK---NETSTQYVDSKRNHVSQAQSDIDLPFR  
ST-----FYRRPFHIGINMOQLSVLQTEFAYGTSENLPASVYKSGTV---DSDIDPQ

MON863 pep  
BACCRYIA\_1  
470 480 490 500 510 519  
TTDEPLEKAYSHQANABCFMQRDRTG---TPPTMTLSDVDFNTIDAKKITQGLPV  
NNTVPPPGQFSHLSHVSGRSGSRAVSIIRAFNFIWASAFNFIASDGTQIPAV

MON863 pep  
BACCRYIA\_1  
520 530 540 550 560 570  
KAYALSSQASIIQPGTQGNLLFKESNSIAK---FKVTLNSAALLQRYKRYAST  
KGNFLNG-SVLSGPGFTGGDLVRLNSGNNIQQNGYIEVPHFSTSTRYKRVYAV

MON863 pep  
BACCRYIA\_1  
580 590 600 610 620 630  
TNLRFLVQSNDFVIYIMTKMKDDLTQTFPLATNMMNMFSGDNELLIGAESV  
TPHILVNMNMSISFEMTVATATSLNL---QSSDFYFESANFTSSLN-IVGVNRS

MON863 pep  
BACCRYIA\_1  
640 650  
SNEKIYIDIEFIPVL  
GTAGVINDKFPFVATLEAEYHLRAKAVNALFTSTNGLQKTNVTDYHIDQVSNLV

MON863 pep  
TO: Q03742

ID Q03742 PRELIMINARY; PRT: 1188 AA  
AC Q03742; DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-FEB-1997 (TREMBLREL. 02, LAST ANNOTATION UPDATE)  
DE 139 KD CRYSTAL PROTEIN (DELTA ENDOTOXIN) (CHETADENE ENDOTOXIN)

SCORES Initl: 445 Inaln: 931 Opt: 1033 z-score: 1022.8 E(): 1.4e-51  
Smith-Waterman score: 1182; 37.2% identity in 646 aa overlap



SCORES Init1: 445 Initn: 931 Opt: 1033 z-score: 1022.8 E(): 1.6e-5  
Smith-Waterman score: 1182; 37.2% identity in 646 aa overlap

DATE 30-JAN-1998  
ACCESSION U43606 . . .

Smith-Waterman score: 1099; 34.7% identity in 585 aa overlap

MON863.pcp SSTEVLNMTVKDVGTVGVVGGILGVVGFAGALTSEYQSLNTW----PSDADPW  
S11445 VEVLGGRITGYTPIDISLSLQFLSEFVPGAG----FVLGLVDIINGIFPGSQ--W  
MON863.pcp KAFMAQVEVLIDKKIEYAKSALAELOQNFEDYVNALNSMKKTPSLRSKRSQORI  
S11445 DAFVLGQELINQRIEFAHQAISRLBOLNLSYVARSFHEADPTN-PALREEMKI  
MON863.pcp RELFSQASHFNSMPEFAVSKFEVLFLPTTGAQANTHLLLLDAQVFGESGYSESDVA  
S11445 Q--FNDMSALTAIPLLAVQNYQVPLSVVQAANLHLSVLSDVSVFGQRWFDAAIN  
MON863.pcp EPYRRQLKLTQQYTHCVNMYVGLNGRSTYDAMVFNFRREMTLVLDLVLPFY  
S11445 SRYNDLRLIGNYDYAVRWYVGLERWGPDSRWVYVNFOPRRLTLVLDLVALPNY  
MON863.pcp DIRLYSGVKTELTRDITDPIFLTLTQKVPFTLSIENIRKPHLPDYLOIEFHTRL  
S11445 DERRYPITRTVQSLTREITNPNV-LENFDOSFRGSAQOISIRSPHLMOLNISTITTYDA  
MON863.pcp RPOYFQKDSFNSMPEFAVSKFEVLFLPTTGAQANTHLLLLDAQVFGESGYSESDVA  
S11445 HNGYV-----YMSGHQIMASPVQFSOPEFTFPLVGTNMAAPQORIVAQLOOQVYRTLS  
MON863.pcp TTDPEPLKAYSHQINAYARCFIMQDRGT----IPFTTWHRSVDFNTIDAKKITOLPVV  
S11445 NNVPFPGQFPHLSHVSFMSRGSSESVIIRAPMSWTHRSAPFNIIASDSITQIPAV  
MON863.pcp KAYALSSGASIIETPGTGGMLFLTESSTIAK---FKVTLSAALLQRYVRYRYAST  
S11445 KGNFLFNG-SVLSGQFTGDLVRLNSGNNIQRGYLWVFIHYSTYRYVRYASV  
MON863.pcp TNLRLPVQNSNDFLVIYNKIMKDDLTQYTFOLATTNNGKPSGDKNELIQAESFV  
S11445 TPIHLNVMNSSLIFENTVPATATSLDL--QSSDFQYPSAMAFSSLON-LVQVWPS  
MON863.pcp SNEKIYIDKIEFIPVQUL  
S11445 OTAGVVIDRPFIPVATLBAEYNLE  
MON863.pcp TO:BACKURS\_1  
LOCUS BACKURS\_1  
DEFINITION Bacillus thuringiensis gene, complete CDS.  
DATE 15-JUL-1991  
ACCESSION M73249  
KEYWORDS 0143125  
ORGANISM Bacillus thuringiensis . .  
SCORES Inchi: 394 Inchi: 869 Opt: 1031 z-score: 1020 9 E(): 2 1e-51  
Smith-Waterman score: 1100, 34 84 identity in 585 aa overlap  
MON863.pcp SSTEVLNMTVKDVGTVGVVGGILGVVGFAGALTSEYQSLNTW----PSDADPW  
BACKURS\_1 VEVLGGRITGYTPIDISLSLQFLSEFVPGAG----FVLGLVDIINGIFPGSQ--W

MON863.pcp KAFMAQVEVLIDKKIEYAKSALAELOQNFEDYVNALNSMKKTPSLRSKRSQORI  
BACKURS\_1 DAFVLGQELINQRIEFAHQAISRLBOLNLSYVARSFHEADPTN-PALREEMKI  
MON863.pcp RELFSQASHFNSMPEFAVSKFEVLFLPTTGAQANTHLLLLDAQVFGESGYSESDVA  
BACKURS\_1 Q--FNDMSALTAIPLLAVQNYQVPLSVVQAANLHLSVLSDVSVFGQRWFDAAIN  
MON863.pcp EPYRRQLKLTQQYTHCVNMYVGLNGRSTYDAMVFNFRREMTLVLDLVLPFY  
BACKURS\_1 SRYNDLRLIGNYDYAVRWYVGLERWGPDSRWVYVNFOPRRLTLVLDLVALPNY  
MON863.pcp DIRLYSGVKTELTRDITDPIFLTLTQKVPFTLSIENIRKPHLPDYLOIEFHTRL  
BACKURS\_1 DERRYPITRTVQSLTREITNPNV-LENFDOSFRGSAQOISIRSPHLMOLNISTITTYDA  
MON863.pcp RPOYFQKDSFNSMPEFAVSKFEVLFLPTTGAQANTHLLLLDAQVFGESGYSESDVA  
BACKURS\_1 HNGYV-----YMSGHQIMASPVQFSOPEFTFPLVGTNMAAPQORIVAQLOOQVYRTLS  
MON863.pcp TTDPEPLKAYSHQINAYARCFIMQDRGT----IPFTTWHRSVDFNTIDAKKITOLPVV  
BACKURS\_1 NNVPFPGQFPHLSHVSFMSRGSSESVIIRAPMSWTHRSAPFNIIASDSITQIPAV  
MON863.pcp KAYALSSGASIIETPGTGGMLFLTESSTIAK---FKVTLSAALLQRYVRYRYAST  
BACKURS\_1 KGNFLFNG-SVLSGQFTGDLVRLNSGNNIQRGYLWVFIHYSTYRYVRYASV  
MON863.pcp TNLRLPVQNSNDFLVIYNKIMKDDLTQYTFOLATTNNGKPSGDKNELIQAESFV  
BACKURS\_1 TPIHLNVMNSSLIFENTVPATATSLDL--QSSDFQYPSAMAFSSLON-LVQVWPS  
MON863.pcp SNEKIYIDKIEFIPVQUL  
BACKURS\_1 OTAGVVIDRPFIPVATLBAEYNLE  
MON863.pcp TO:BACKURS\_1  
LOCUS BACKURS\_1  
DEFINITION Bacillus thuringiensis (cryIa(c3)) gene, complete CDS  
DATE 15-JUL-1991  
ACCESSION M73249  
KEYWORDS 0143125  
ORGANISM Bacillus thuringiensis . .  
SCORES Inchi: 394 Inchi: 869 Opt: 1026 z-score: 1016.0 E(): 4e-51  
Smith-Waterman score: 1099, 34 74 identity in 585 aa overlap  
MON863.pcp SSTEVLNMTVKDVGTVGVVGGILGVVGFAGALTSEYQSLNTW----PSDADPW  
Q03743 VEVLGGRITGYTPIDISLSLQFLSEFVPGAG----FVLGLVDIINGIFPGSQ--W  
MON863.pcp KAFMAQVEVLIDKKIEYAKSALAELOQNFEDYVNALNSMKKTPSLRSKRSQORI  
Q03743 DAFVLGQELINQRIEFAHQAISRLBOLNLSYVARSFHEADPTN-PALREEMKI  
MON863.pcp RELFSQASHFNSMPEFAVSKFEVLFLPTTGAQANTHLLLLDAQVFGESGYSESDVA  
Q03743 Q--FNDMSALTAIPLLAVQNYQVPLSVVQAANLHLSVLSDVSVFGQRWFDAAIN  
MON863.pcp EPYRRQLKLTQQYTHCVNMYVGLNGRSTYDAMVFNFRREMTLVLDLVLPFY  
Q03743 SRYNDLRLIGNYDYAVRWYVGLERWGPDSRWVYVNFOPRRLTLVLDLVALPNY  
MON863.pcp DIRLYSGVKTELTRDITDPIFLTLTQKVPFTLSIENIRKPHLPDYLOIEFHTRL  
Q03743 DERRYPITRTVQSLTREITNPNV-LENFDOSFRGSAQOISIRSPHLMOLNISTITTYDA  
MON863.pcp RPOYFQKDSFNSMPEFAVSKFEVLFLPTTGAQANTHLLLLDAQVFGESGYSESDVA  
Q03743 HNGYV-----YMSGHQIMASPVQFSOPEFTFPLVGTNMAAPQORIVAQLOOQVYRTLS  
MON863.pcp TTDPEPLKAYSHQINAYARCFIMQDRGT----IPFTTWHRSVDFNTIDAKKITOLPVV  
Q03743 NNVPFPGQFPHLSHVSFMSRGSSESVIIRAPMSWTHRSAPFNIIASDSITQIPAV  
MON863.pcp KAYALSSGASIIETPGTGGMLFLTESSTIAK---FKVTLSAALLQRYVRYRYAST  
Q03743 KGNFLFNG-SVLSGQFTGDLVRLNSGNNIQRGYLWVFIHYSTYRYVRYASV  
MON863.pcp TNLRLPVQNSNDFLVIYNKIMKDDLTQYTFOLATTNNGKPSGDKNELIQAESFV  
Q03743 TPIHLNVMNSSLIFENTVPATATSLDL--QSSDFQYPSAMAFSSLON-LVQVWPS  
MON863.pcp SNEKIYIDKIEFIPVQUL  
Q03743 OTAGVVIDRPFIPVATLBAEYNLE  
MON863.pcp TO:BACKURS\_1  
LOCUS BACKURS\_1  
DEFINITION Bacillus thuringiensis (cryIa(c3)) gene, complete CDS  
DATE 15-JUL-1991  
ACCESSION M73249  
KEYWORDS 0143125  
ORGANISM Bacillus thuringiensis . .  
SCORES Inchi: 394 Inchi: 869 Opt: 1026 z-score: 1016.0 E(): 4e-51  
Smith-Waterman score: 1099, 34 74 identity in 585 aa overlap

MON863.pcp RELFSQASHFNSMPEFAVSKFEVLFLPTTGAQANTHLLLLDAQVFGESGYSESDVA  
Q03743 Q--FNDMSALTAIPLLAVQNYQVPLSVVQAANLHLSVLSDVSVFGQRWFDAAIN  
MON863.pcp EPYRRQLKLTQQYTHCVNMYVGLNGRSTYDAMVFNFRREMTLVLDLVLPFY  
Q03743 SRYNDLRLIGNYDYAVRWYVGLERWGPDSRWVYVNFOPRRLTLVLDLVALPNY  
MON863.pcp DIRLYSGVKTELTRDITDPIFLTLTQKVPFTLSIENIRKPHLPDYLOIEFHTRL  
Q03743 DERRYPITRTVQSLTREITNPNV-LENFDOSFRGSAQOISIRSPHLMOLNISTITTYDA  
MON863.pcp RPOYFQKDSFNSMPEFAVSKFEVLFLPTTGAQANTHLLLLDAQVFGESGYSESDVA  
Q03743 HNGYV-----YMSGHQIMASPVQFSOPEFTFPLVGTNMAAPQORIVAQLOOQVYRTLS  
MON863.pcp TTDPEPLKAYSHQINAYARCFIMQDRGT----IPFTTWHRSVDFNTIDAKKITOLPVV  
Q03743 NNVPFPGQFPHLSHVSFMSRGSSESVIIRAPMSWTHRSAPFNIIASDSITQIPAV  
MON863.pcp KAYALSSGASIIETPGTGGMLFLTESSTIAK---FKVTLSAALLQRYVRYRYAST  
Q03743 KGNFLFNG-SVLSGQFTGDLVRLNSGNNIQRGYLWVFIHYSTYRYVRYASV  
MON863.pcp TNLRLPVQNSNDFLVIYNKIMKDDLTQYTFOLATTNNGKPSGDKNELIQAESFV  
Q03743 TPIHLNVMNSSLIFENTVPATATSLDL--QSSDFQYPSAMAFSSLON-LVQVWPS  
MON863.pcp SNEKIYIDKIEFIPVQUL  
Q03743 OTAGVVIDRPFIPVATLBAEYNLE  
MON863.pcp TO:BACKURS\_1  
LOCUS BACKURS\_1  
DEFINITION Bacillus thuringiensis (cryIa(c3)) gene, complete CDS  
DATE 15-JUL-1991  
ACCESSION M73249  
KEYWORDS 0143125  
ORGANISM Bacillus thuringiensis . .  
SCORES Inchi: 394 Inchi: 869 Opt: 1026 z-score: 1016.0 E(): 4e-51  
Smith-Waterman score: 1099, 34 74 identity in 585 aa overlap  
MON863.pcp SSTEVLNMTVKDVGTVGVVGGILGVVGFAGALTSEYQSLNTW----PSDADPW  
BACCRYIACC\_1 VEVLGGRITGYTPIDISLSLQFLSEFVPGAG----FVLGLVDIINGIFPGSQ--W  
MON863.pcp KAFMAQVEVLIDKKIEYAKSALAELOQNFEDYVNALNSMKKTPSLRSKRSQORI  
BACCRYIACC\_1 DAFVLGQELINQRIEFAHQAISRLBOLNLSYVARSFHEADPTN-PALREEMKI  
MON863.pcp RELFSQASHFNSMPEFAVSKFEVLFLPTTGAQANTHLLLLDAQVFGESGYSESDVA  
BACCRYIACC\_1 Q--FNDMSALTAIPLLAVQNYQVPLSVVQAANLHLSVLSDVSVFGQRWFDAAIN  
MON863.pcp EPYRRQLKLTQQYTHCVNMYVGLNGRSTYDAMVFNFRREMTLVLDLVLPFY  
BACCRYIACC\_1 SRYNDLRLIGNYDYAVRWYVGLERWGPDSRWVYVNFOPRRLTLVLDLVALPNY  
MON863.pcp DIRLYSGVKTELTRDITDPIFLTLTQKVPFTLSIENIRKPHLPDYLOIEFHTRL  
BACCRYIACC\_1 DERRYPITRTVQSLTREITNPNV-LENFDOSFRGSAQOISIRSPHLMOLNISTITTYDA  
MON863.pcp RPOYFQKDSFNSMPEFAVSKFEVLFLPTTGAQANTHLLLLDAQVFGESGYSESDVA  
BACCRYIACC\_1 HNGYV-----YMSGHQIMASPVQFSOPEFTFPLVGTNMAAPQORIVAQLOOQVYRTLS  
MON863.pcp TTDPEPLKAYSHQINAYARCFIMQDRGT----IPFTTWHRSVDFNTIDAKKITOLPVV  
BACCRYIACC\_1 NNVPFPGQFPHLSHVSFMSRGSSESVIIRAPMSWTHRSAPFNIIASDSITQIPAV  
MON863.pcp KAYALSSGASIIETPGTGGMLFLTESSTIAK---FKVTLSAALLQRYVRYRYAST  
BACCRYIACC\_1 KGNFLFNG-SVLSGQFTGDLVRLNSGNNIQRGYLWVFIHYSTYRYVRYASV  
MON863.pcp TNLRLPVQNSNDFLVIYNKIMKDDLTQYTFOLATTNNGKPSGDKNELIQAESFV  
BACCRYIACC\_1 TPIHLNVMNSSLIFENTVPATATSLDL--QSSDFQYPSAMAFSSLON-LVQVWPS  
MON863.pcp SNEKIYIDKIEFIPVQUL  
BACCRYIACC\_1 OTAGVVIDRPFIPVATLBAEYNLE  
MON863.pcp TO:BACKURS\_1  
LOCUS BACCRYIACC\_1  
DEFINITION Bacillus thuringiensis (cryIa(c3)) gene, complete CDS  
DATE 15-JUL-1991  
ACCESSION M73249  
KEYWORDS 0143125  
ORGANISM Bacillus thuringiensis . .  
SCORES Inchi: 394 Inchi: 869 Opt: 1026 z-score: 1016.0 E(): 4e-51  
Smith-Waterman score: 1099, 34 74 identity in 585 aa overlap

	250	260	270	280	290	300
MON863.pcp	DMCAVYAPRGLNGLGKSGTIDAMVAFHFRREKLTULVGLIVLPFFPDIIRLKYGVKATLL					
S00944	DMCAVYAPRGLNGLPKSTYQDMITLHRLRDLTLVLDAFAFFPNVNNRRPFIQPGVGLT					
	210	220	230	240	250	260
MON863.pcp	RDITFDEIFLLLT-LKRGV--PTFLSENS-IRKPHLELYLQIEFHTLRPGVYKGSK					
S00944	REVTYDILNPNFLOGLVAGLPTFSSSRINRPHFLDIANNITPTL-----DNFVSGR					
	270	280	290	300	310	
MON863.pcp	360	370	380	390	400	410
MON863.pcp	NYNGSYVETRAPSGSSKSTISFPGV-DKSTPEVKQLSDQGVYVTIAMDI-----AA					
S00944	PTGGGHRVIE-LIGGKNTYFPGVREAPLSTFSSVPTFLNSPTLRLOQPK					
	320	330	340	350	360	370
MON863.pcp	420	430	440	450	460	470
MON863.pcp	PNKRGVLTGTVKVDSPGSDQKNKSTCTGYDKNNKHGVSAQDSIQLPSTTDEPLEKA					
S00944	PAFPFNLRGVGVSEFS-----TFNSFTY--RGRGV--DSLTLPEPDNSVPRKE					
	380	390	400	410	420	
MON863.pcp	480	490	500	510	520	
MON863.pcp	YSHQLNVAEFLCQDROGRTFP-----FTWHSRVDFNTIDARKITQLPVWKAVALGS					
S00944	YSHRLCHAT-FVQ--RSGT-PFLTGVGVSMWDSRALTNTIDPRKINQIPLKGVKFRG					
	430	440	450	460	470	480
MON863.pcp	530	540	550	560	570	580
MON863.pcp	GAISIEGPGTGGNGLFLKSSNSIAKFKVITNSAALLORYVRIVASTNLTPLVQNS					
S00944	GTSLVGVDFPGDGLLR-----KSTGKNTGPGWNNIP-ITQVALGFIYASSDRARVVLQ					
	490	500	510	520	530	
MON863.pcp	600	610	620	630		
MON863.pcp	NDPLV--IYIR--KTVGKDDQLVYGLPLATTNSWGS-----GDKHMLIGAI					
S00944	AASITGVQGVVNMFLQTKHNGMLSTSTFRYTFDSNPFPRANPDITGLSGDPLFCA					
	540	550	560	570	580	590
MON863.pcp	640	650				
MON863.pcp	SFVNSRKTIYDKTEFIPFVL					
S00944	-----					

[illegible]

```

      10      20      30      40      50      60
MON863.pep SEQDITVFNSELTQTHNNOYPLGNPNFTSEELKYKFLTLDSDSTE--VLDT--FV
      70      80      90      100      110      120
CRY1_BACTE MRENNQNCITPYNLSNPBEVLIDQKIS
      130      140      150      160      170      180
MON863.pep VKNQVITLSEVQGLVFPVFAITLSTYQSLNITMPSDADKWKAFVACVFLVDKI
      190      200      210      220      230      240
CRY1_BACTE GNSSIDSLSELVQPLVSNFVERGGFGLVGLVGVGVGVGQ--NDALFVQTEQLTNE
      250      260      270      280      290      300
MON863.pep IERYAKSKALAEQLQLOLNPHQYVIALNSKKTKPLSLKRSKRSRITPQBSRHRN
      310      320      330      340      350      360
CRY1_BACTE IASFAKNAATLNNKGRNNNTVFAKPEKED--NPPETVITLDFPILQDLSRL
      370      380      390      400      410      420
MON863.pep KPGFASTGKVELFLPTTAAQANLHLLLLKDAQVTFEYMGSESDVAFYRGLKTLTQ
      430      440      450      460      470      480
CRY1_BACTE ITPSFIRSEVFLVSLVYANMLALRLDSVIGFGRGATVNNVYNNRLHFLDQV
      490      500      510      520      530      540
MON863.pep
      550      560      570      580      590      600
CRY1_BACTE

```



250 260 270 280 290 300  
MON863.pap DRGVNPNVGLNGRGFTLAWVRKRRFRREMLTVLVLPPFYDIALYSGVKTBLT  
CRYT\_BACTE DHCATNNGIANKPKSTVQMTITLRLRLTLVLDIAAFPMYDNRAPYIQVQQLT  
210 220 230 240 250 260  
310 320 330 340 350  
MON863.pap ADIPTDIFLLTT-LKVG--PTLSIENS-IRKPHFDYDQIEFHTLRPGYFGKDSF  
CRYT\_BACTE REVVTGMLNPNGLSVAQGTAVVNSRIRKPHLFDITLMLTIT--DNFSVORN  
270 280 290 300 310  
360 370 380 390 400 410  
MON863.pap NYWSONVETKPSIGSKTITSPYDQ-DKSPEDVOKLAFDQKVVRTIAMTV--AAW  
CRYT\_BACTE FYWQKHVLS--SLIOGNITATVGRINAEPPFSTHGP-VFTLSNPTLRLLOQPM  
320 330 340 350 360 370  
420 430 440 450 460 470  
MON863.pap PDKGVYL-GVTKVDSQVDDQKNETSTQYDNRKNGHVSQAQSDIDQLBETIDSLKA  
CRYT\_BACTE PAPPFNLRGVDSQVDS--TPHTSPY--NKGSTV--DEQLPFDNSVDFHNS  
380 390 400 410 420  
480 490 500 510 520  
MON863.pap YSHQVAKCFMLQDQKGTIFF--PTWTRSVDFPMIDARKITQPVYVVALSS  
CRYT\_BACTE YSHKLCHAT-FVQ--KNOT-PFLTGVVSWDRSATLNTIDPERINGIPLVKKRVVQ  
430 440 450 460 470 480  
530 540 550 560 570 580  
MON863.pap GASIEGPPPTGQNLFIKESNSIAKFTVTLNRAALLQRYVRIYASTNLRFLVONS  
CRYT\_BACTE GTVITDPTDGGDLR-LNFGQDFVLOVQINSP-ITDGLRFRYASRDARVILVIG  
490 500 510 520 530  
590 600 610 620 630  
MON863.pap NNDFIV--IYIN--KYMKKDDLTQYOTDLATNNSNGFS--GDKNELIIGAE  
CRYT\_BACTE AASTGVGGVSNVMPLOKMEIGENLSTRTYTFDSKPSFRANPDITDISEQPLPAG  
540 550 560 570 580 590  
640 650  
MON863.pap SFVSEKIKYIDKIEFIPVQL  
CRYT\_BACTE SISSGE-LYIDKIKILLADATFNKASUDERAKAVNALPSSNQILKLTQVTDVYHIDQVS  
600 610 620 630 640 650  
MON863.pap TO:CRVU\_BACTK  
ID CRVU\_BACTK STANDARD. PRT. 1156 AA.  
AC P09663. DT 01-MAR-1989 (REL. 10, CREATED)  
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)  
DT 01-MAY-1991 (REL. 16, LAST ANNOTATION UPDATE)  
DE 131 KD CRYSTAL PROTEIN (DELTA ENDOTOXIN) (CRYSTALLINE ENTEROCOCCAL . . .  
SCORES Initl: 406 Inltn: 867 Opt: 1012 z-score: 1002.2 E(1) 2.3e-50  
Smith-Waterman score 1033, 33.0% identity in 595 aa overlap  
60 70 80 90 100  
MON863.pap STSEVLNSTVKIAVOTGTSVQVQILGVGVFPFAGALTSFYQSLNTIW--PSIDAPW  
CRYT\_BACTE VEVLQGERIFQYTPIDISLSLTQFLLEFPVQAG--FVLGLVDIINGIIFQSQ--W  
30 40 50 60 70  
110 120 130 140 150 160  
MON863.pap KATPAQVLEVLIDKIKIEYAKSKALAELOQNNFEDYVNAISNKKTPSLRSKRSQRI  
CRYT\_BACTE DAPLVQGLDQNLQIEFARKQATSRGLGSLNLYQIYAESFRWEDADPTN-PALREKRI  
80 90 100 110 120 130  
170 180 190 200 210 220  
MON863.pap RELPSQASHRSNNSPFAVSKFEVLFLFTYQAANTLALLKDAQVFGGEWYSSRDA  
CRYT\_BACTE Q--FNDKNSALTAIFLFAVQYGVLLSVVQAAHLNLSVLVDVSGVQGWGFDAAITN  
140 150 160 170 180 190  
230 240 250 260 270 280  
MON863.pap EFRRLKLTQOYTHCVNWNVGLNGRGFTLAWVRKRRFRREMLTVLVLPPFY  
CRYT\_BACTE SRYNLTIRLIGNYTHDRVRYNTGLRWGPDSDNRVRYQGFRELTLVLDIVSLPNNY  
200 210 220 230 240 250

290 300 310 320 330 340  
MON863.pap DIRLYSGVKTBLTADIFTDIPFLITLTKYQPTFLSINSTRKPHLPDYLOQIEFHTRL  
CRYT\_BACTE DSRTPYPTVTSQLTREIYINPV-LENFDQSPGSAQIGSGISPSHMLDILNLSITVYDA  
260 270 280 290 300  
350 360 370 380 390 400  
MON863.pap RPOVFKDPSFNWNSQVYVETSPISGSKTITSPFYQDK-STEPVOKLSFD-QKVVYRTIA  
CRYT\_BACTE HRGEY-----YWSHQIMASPVQSGPSTPFLYGTMGNAAPQRLVAQLQGVYRTLS  
310 320 330 340 350 360  
410 420 430 440 450 460  
MON863.pap NTDVAAMPNGKVLGVTKVDFSQYDDQKNET-STQYDNRKNGHVSQAQSDIDQLBETIT  
CRYT\_BACTE ST-LYRPP--FNIGINNQLSGPDGGRICLNNLKFQGPVYTKAEFVDSDEIPQNN  
370 380 390 400 410  
470 480 490 500 510  
MON863.pap DRLPKAYSHQ-LNVAECFLMDQRRGTIPFF-----TWTRSVDFPTNIDARKITQPV  
CRYT\_BACTE NVPPROGFSHRCLATVSKYSGFSNSVSVTRAPNDSNWTCSAEFGDVPSSQITQPL  
420 430 440 450 460 470  
520 530 540 550 560 570  
MON863.pap VKATLSSQGLTIDKIPGPGQNLFLKESNSIAKFTVTLNRAALLQRYVRIYASTN  
CRYT\_BACTE TYLQGLNLSVTVVGLQFQGDIL--ELTQOISTLRVNI-TAPLEGRVRIYASTN  
480 490 500 510 520 530  
580 590 600 610 620 630  
MON863.pap LRLFVQNSNDFLVIYINKTNKDDLTQYOTDLATNNSNGFSGDKNELIIGAEFVSN  
CRYT\_BACTE LQFMTSIDGRPTNGQNFSAHSRGNLQSRSPATVQPTTPPFTSNGSSVPTLSAHVFNQ  
540 550 560 570 580 590  
640 650  
MON863.pap EKIYIDKIEFIPVQL  
CRYT\_BACTE NEVYIDKIEFVPAVTFEAYEDLRAQKAVNLFSSNQILKLTQVTDVYHIDQVSNLVC  
600 610 620 630 640 650  
MON863.pap TO:RACHQA\_1  
LOCUS RACHQA\_1  
DEFINITION B.thuringiensis insecticidal endotoxin gene, complete cds,  
insecticidal endotoxin  
DATE 01-SEP-1988  
ACCESSION M12661  
KEYWORDS g143226 . . .  
SCORES Initl: 406 Inltn: 867 Opt: 1012 z-score: 1002.2 E(1) 2.3e-50  
Smith-Waterman score 1033, 33.0% identity in 595 aa overlap  
60 70 80 90 100  
MON863.pap STSEVLNSTVKIAVOTGTSVQVQILGVGVFPFAGALTSFYQSLNTIW--PSIDAPW  
RACHQA\_1 VEVLQGERIFQYTPIDISLSLTQFLLEFPVQAG--FVLGLVDIINGIIFQSQ--W  
30 40 50 60 70  
110 120 130 140 150 160  
MON863.pap KATPAQVLEVLIDKIKIEYAKSKALAELOQNNFEDYVNAISNKKTPSLRSKRSQRI  
RACHQA\_1 DAPLVQGLDQNLQIEFARKQATSRGLGSLNLYQIYAESFRWEDADPTN-PALREKRI  
80 90 100 110 120 130  
170 180 190 200 210 220  
MON863.pap RELPSQASHRSNNSPFAVSKFEVLFLFTYQAANTLALLKDAQVFGGEWYSSRDA  
RACHQA\_1 Q--FNDKNSALTAIFLFAVQYGVLLSVVQAAHLNLSVLVDVSGVQGWGFDAAITN  
140 150 160 170 180 190  
230 240 250 260 270 280  
MON863.pap EFRRLKLTQOYTHCVNWNVGLNGRGFTLAWVRKRRFRREMLTVLVLPPFY  
RACHQA\_1 SRYNLTIRLIGNYTHDRVRYNTGLRWGPDSDNRVRYQGFRELTLVLDIVSLPNNY  
200 210 220 230 240 250

290 300 310 320 330 340  
MON863.pap DIRLYSGVKTBLTADIFTDIPFLITLTKYQPTFLSINSTRKPHLPDYLOQIEFHTRL  
A29838 DSRTPYPTVTSQLTREIYINPV-LENFDQSPGSAQIGSGISPSHMLDILNLSITVYDA  
260 270 280 290 300  
350 360 370 380 390 400  
MON863.pap RPOVFKDPSFNWNSQVYVETSPISGSKTITSPFYQDK-STEPVOKLSFD-QKVVYRTIA  
A29838 HRGEY-----YWSHQIMASPVQSGPSTPFLYGTMGNAAPQRLVAQLQGVYRTLS  
310 320 330 340 350 360  
410 420 430 440 450 460  
MON863.pap NTDVAAMPNGKVLGVTKVDFSQYDDQKNET-STQYDNRKNGHVSQAQSDIDQLBETIT  
A29838 ST-LYRPP--FNIGINNQLSGPDGGRICLNNLKFQGPVYTKAEFVDSDEIPQNN  
370 380 390 400 410  
470 480 490 500 510  
MON863.pap DRLPKAYSHQ-LNVAECFLMDQRRGTIPFF-----TWTRSVDFPTNIDARKITQPV  
A29838 NVPPROGFSHRCLATVSKYSGFSNSVSVTRAPNDSNWTCSAEFGDVPSSQITQPL  
420 430 440 450 460 470  
520 530 540 550 560 570  
MON863.pap VKATLSSQGLTIDKIPGPGQNLFLKESNSIAKFTVTLNRAALLQRYVRIYASTN  
A29838 TYLQGLNLSVTVVGLQFQGDIL--ELTQOISTLRVNI-TAPLEGRVRIYASTN  
480 490 500 510 520 530  
580 590 600 610 620 630  
MON863.pap LRLFVQNSNDFLVIYINKTNKDDLTQYOTDLATNNSNGFSGDKNELIIGAEFVSN  
A29838 LQFMTSIDGRPTNGQNFSAHSRGNLQSRSPATVQPTTPPFTSNGSSVPTLSAHVFNQ  
540 550 560 570 580 590  
640 650  
MON863.pap EKIYIDKIEFIPVQL  
A29838 NEVYIDKIEFVPAVTFEAYEDLRAQKAVNLFSSNQILKLTQVTDVYHIDQVSNLVC  
600 610 620 630 640 650  
MON863.pap TO:RACHQA\_1  
LOCUS RACHQA\_1  
DEFINITION B.thuringiensis insecticidal endotoxin gene, complete cds,  
insecticidal endotoxin  
DATE 01-SEP-1988  
ACCESSION M12661  
KEYWORDS g143226 . . .  
SCORES Initl: 406 Inltn: 867 Opt: 1012 z-score: 1002.2 E(1) 2.3e-50  
Smith-Waterman score 1033, 33.0% identity in 595 aa overlap  
60 70 80 90 100  
MON863.pap STSEVLNSTVKIAVOTGTSVQVQILGVGVFPFAGALTSFYQSLNTIW--PSIDAPW  
RACHQA\_1 VEVLQGERIFQYTPIDISLSLTQFLLEFPVQAG--FVLGLVDIINGIIFQSQ--W  
30 40 50 60 70  
110 120 130 140 150 160  
MON863.pap KATPAQVLEVLIDKIKIEYAKSKALAELOQNNFEDYVNAISNKKTPSLRSKRSQRI  
RACHQA\_1 DAPLVQGLDQNLQIEFARKQATSRGLGSLNLYQIYAESFRWEDADPTN-PALREKRI  
80 90 100 110 120 130  
170 180 190 200 210 220  
MON863.pap RELPSQASHRSNNSPFAVSKFEVLFLFTYQAANTLALLKDAQVFGGEWYSSRDA  
RACHQA\_1 Q--FNDKNSALTAIFLFAVQYGVLLSVVQAAHLNLSVLVDVSGVQGWGFDAAITN  
140 150 160 170 180 190  
230 240 250 260 270 280  
MON863.pap EFRRLKLTQOYTHCVNWNVGLNGRGFTLAWVRKRRFRREMLTVLVLPPFY  
RACHQA\_1 SRYNLTIRLIGNYTHDRVRYNTGLRWGPDSDNRVRYQGFRELTLVLDIVSLPNNY  
200 210 220 230 240 250  
290 300 310 320 330 340  
MON863.pap DIRLYSGVKTBLTADIFTDIPFLITLTKYQPTFLSINSTRKPHLPDYLOQIEFHTRL  
RACHQA\_1 DSRTPYPTVTSQLTREIYINPV-LENFDQSPGSAQIGSGISPSHMLDILNLSITVYDA  
260 270 280 290 300



MON863.pap 310 320 330 340 350 360  
RDTPTDPIFLATT-LQKVG--PTFLSIENS-IRKPHLPDYLOQIEFHTLRPOYFGKDSF  
BACHSQA\_1 310 320 330 340 350 360  
RDTPTDPIFLATT-LQKVG--PTFLSIENS-IRKPHLPDYLOQIEFHTLRPOYFGKDSF  
MON863.pap 360 370 380 390 400 410  
NWSGNTVTRPSIGSKTISFPFG-DKSTSPVQKLSFGQVYRT--IANTDVAAMPN  
BACHSQA\_1 360 370 380 390 400 410  
NWSGNTVTRPSIGSKTISFPFG-DKSTSPVQKLSFGQVYRT--IANTDVAAMPN  
MON863.pap 420 430 440 450 460 470  
GKVVYL---GVTKVDFSVQDDQKNETSTQYDGRNNGHVSQAQSDIDQLPPTTDEPLEKA  
BACHSQA\_1 420 430 440 450 460 470  
GKVVYL---GVTKVDFSVQDDQKNETSTQYDGRNNGHVSQAQSDIDQLPPTTDEPLEKA  
MON863.pap 480 490 500 510 520 530  
YSHGLVABCFMADQKRTIFF-----FTWHRSVDFVFTDARKITOLPVQKVALES  
BACHSQA\_1 480 490 500 510 520 530  
YSHGLVABCFMADQKRTIFF-----FTWHRSVDFVFTDARKITOLPVQKVALES  
MON863.pap 540 550 560 570 580 590  
GASITGPGFTGGNLLFLKESNSIAKPKVTNSAALLQRYVRIRYASTTNLRFVONS  
BACHSQA\_1 540 550 560 570 580 590  
GASITGPGFTGGNLLFLKESNSIAKPKVTNSAALLQRYVRIRYASTTNLRFVONS  
MON863.pap 600 610 620 630 640 650  
NNDPLV---IYIN---KTMKDDDLTYQTFDLATNSNGPS-----GDKNSLIGAE  
BACHSQA\_1 600 610 620 630 640 650  
NNDPLV---IYIN---KTMKDDDLTYQTFDLATNSNGPS-----GDKNSLIGAE

Fl.004181 - parasporal crystal protein - *Bacillus thuringiensis* (strain  
aizawa  
7.29) (fragment)  
N: Alternate names: delta-endotoxin  
C: Species: *Bacillus thuringiensis*  
C: Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 08-Sep-1997  
C: Accession: 004181 ..

SCORES Initl: 445 Initn: 924 Opt: 1000 z-score: 993.0 E(): 7.6e-50  
Smith-Waterman score: 1146; 36.7% identity in 646 aa overlap

MON863.pap 10 20 30 40 50 60  
SHNDTIKVTNPSLQTMNQYPLADNPSTLRSLYKFLRMTEDSSTE--VLDN---ST  
BACHSQA\_1 10 20 30 40 50 60  
SHNDTIKVTNPSLQTMNQYPLADNPSTLRSLYKFLRMTEDSSTE--VLDN---ST  
MON863.pap 70 80 90 100 110 120  
VKDAVGOTISVVOGLQVVPFAGALTSFYQSFMTWPSDADPMKAFNAQVEVLIDKK  
BACHSQA\_1 70 80 90 100 110 120  
VKDAVGOTISVVOGLQVVPFAGALTSFYQSFMTWPSDADPMKAFNAQVEVLIDKK  
MON863.pap 130 140 150 160 170 180  
IEEYAKSALAEGLQGNFEDYVNAKNSKKTPLSLRSKRSQGRIRHLPQASRHPNS  
BACHSQA\_1 130 140 150 160 170 180  
IEEYAKSALAEGLQGNFEDYVNAKNSKKTPLSLRSKRSQGRIRHLPQASRHPNS  
MON863.pap 190 200 210 220 230 240  
NPSFAVSKFVLFLPTTAQAANTHLLKDAQVGSSEGVSSDVAEPYRQKLTQOYT  
BACHSQA\_1 190 200 210 220 230 240  
NPSFAVSKFVLFLPTTAQAANTHLLKDAQVGSSEGVSSDVAEPYRQKLTQOYT  
MON863.pap 250 260 270 280 290 300  
DRCVWNTVGLNGRSTYDAWVKNFRRENTLTVDLVLFPFYDIRLYSGVKTELT  
BACHSQA\_1 250 260 270 280 290 300  
DRCVWNTVGLNGRSTYDAWVKNFRRENTLTVDLVLFPFYDIRLYSGVKTELT

MON863.pap 310 320 330 340 350  
RDTPTDPIFLATT-LQKVG--PTFLSIENS-IRKPHLPDYLOQIEFHTLRPOYFGKDSF  
BACHSQA\_1 310 320 330 340 350  
RDTPTDPIFLATT-LQKVG--PTFLSIENS-IRKPHLPDYLOQIEFHTLRPOYFGKDSF  
MON863.pap 360 370 380 390 400 410  
NWSGNTVTRPSIGSKTISFPFG-DKSTSPVQKLSFGQVYRT--IANTDVAAMPN  
BACHSQA\_1 360 370 380 390 400 410  
NWSGNTVTRPSIGSKTISFPFG-DKSTSPVQKLSFGQVYRT--IANTDVAAMPN  
MON863.pap 420 430 440 450 460 470  
GKVVYL---GVTKVDFSVQDDQKNETSTQYDGRNNGHVSQAQSDIDQLPPTTDEPLEKA  
BACHSQA\_1 420 430 440 450 460 470  
GKVVYL---GVTKVDFSVQDDQKNETSTQYDGRNNGHVSQAQSDIDQLPPTTDEPLEKA  
MON863.pap 480 490 500 510 520 530  
YSHGLVABCFMADQKRTIFF-----FTWHRSVDFVFTDARKITOLPVQKVALES  
BACHSQA\_1 480 490 500 510 520 530  
YSHGLVABCFMADQKRTIFF-----FTWHRSVDFVFTDARKITOLPVQKVALES  
MON863.pap 540 550 560 570 580 590  
GASITGPGFTGGNLLFLKESNSIAKPKVTNSAALLQRYVRIRYASTTNLRFVONS  
BACHSQA\_1 540 550 560 570 580 590  
GASITGPGFTGGNLLFLKESNSIAKPKVTNSAALLQRYVRIRYASTTNLRFVONS  
MON863.pap 600 610 620 630 640 650  
NNDPLV---IYIN---KTMKDDDLTYQTFDLATNSNGPS-----GDKNSLIGAE  
BACHSQA\_1 600 610 620 630 640 650  
NNDPLV---IYIN---KTMKDDDLTYQTFDLATNSNGPS-----GDKNSLIGAE  
MON863.pap 660 670 680 690 700 710  
SPVNEKIYIDKISPIVQL  
BACHSQA\_1 660 670 680 690 700 710  
SPVNEKIYIDKISPIVQL  
MON863.pap 720 730 740 750 760 770  
SISGGE-LYIDKIKILADATSSASLERAQKVAALFTSSNOIGATDVTNHIQVS  
BACHSQA\_1 720 730 740 750 760 770  
SISGGE-LYIDKIKILADATSSASLERAQKVAALFTSSNOIGATDVTNHIQVS  
MON863.pap 780 790 800 810 820 830  
ID CRWY\_BACTA STANDARD; PRT: 823 AA  
AC P10237;  
DT 01-MAR-1989 (REL. 10, CREATED)  
RT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)  
TM 01-MAY-1991 (REL. 18, LAST ANNOTATION UPDATE)  
DE 130 KD CRYSTAL PROTEIN (DELTA ENDOTOXIN) (CRYSTALLIN ENTOXINOCIDAL ..

MON863.pap 310 320 330 340 350  
RDTPTDPIFLATT-LQKVG--PTFLSIENS-IRKPHLPDYLOQIEFHTLRPOYFGKDSF  
BACHSQA\_1 310 320 330 340 350  
RDTPTDPIFLATT-LQKVG--PTFLSIENS-IRKPHLPDYLOQIEFHTLRPOYFGKDSF  
MON863.pap 360 370 380 390 400 410  
NWSGNTVTRPSIGSKTISFPFG-DKSTSPVQKLSFGQVYRT--IANTDVAAMPN  
BACHSQA\_1 360 370 380 390 400 410  
NWSGNTVTRPSIGSKTISFPFG-DKSTSPVQKLSFGQVYRT--IANTDVAAMPN  
MON863.pap 420 430 440 450 460 470  
GKVVYL---GVTKVDFSVQDDQKNETSTQYDGRNNGHVSQAQSDIDQLPPTTDEPLEKA  
BACHSQA\_1 420 430 440 450 460 470  
GKVVYL---GVTKVDFSVQDDQKNETSTQYDGRNNGHVSQAQSDIDQLPPTTDEPLEKA  
MON863.pap 480 490 500 510 520 530  
YSHGLVABCFMADQKRTIFF-----FTWHRSVDFVFTDARKITOLPVQKVALES  
BACHSQA\_1 480 490 500 510 520 530  
YSHGLVABCFMADQKRTIFF-----FTWHRSVDFVFTDARKITOLPVQKVALES  
MON863.pap 540 550 560 570 580 590  
GASITGPGFTGGNLLFLKESNSIAKPKVTNSAALLQRYVRIRYASTTNLRFVONS  
BACHSQA\_1 540 550 560 570 580 590  
GASITGPGFTGGNLLFLKESNSIAKPKVTNSAALLQRYVRIRYASTTNLRFVONS  
MON863.pap 600 610 620 630 640 650  
NNDPLV---IYIN---KTMKDDDLTYQTFDLATNSNGPS-----GDKNSLIGAE  
BACHSQA\_1 600 610 620 630 640 650  
NNDPLV---IYIN---KTMKDDDLTYQTFDLATNSNGPS-----GDKNSLIGAE  
MON863.pap 660 670 680 690 700 710  
SPVNEKIYIDKISPIVQL  
BACHSQA\_1 660 670 680 690 700 710  
SPVNEKIYIDKISPIVQL  
MON863.pap 720 730 740 750 760 770  
SISGGE-LYIDKIKILADATSSASLERAQKVAALFTSSNOIGATDVTNHIQVS  
BACHSQA\_1 720 730 740 750 760 770  
SISGGE-LYIDKIKILADATSSASLERAQKVAALFTSSNOIGATDVTNHIQVS  
MON863.pap 780 790 800 810 820 830  
ID CRWY\_BACTA STANDARD; PRT: 823 AA  
AC P10237;  
DT 01-MAR-1989 (REL. 10, CREATED)  
RT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)  
TM 01-MAY-1991 (REL. 18, LAST ANNOTATION UPDATE)  
DE 130 KD CRYSTAL PROTEIN (DELTA ENDOTOXIN) (CRYSTALLIN ENTOXINOCIDAL ..



Page 77 of 102







MON863.pap 290 300 310 320 330 340 350  
IFTPDIFLLTQKYGPF-----TFLSIENSI-RKPHLPDYLQIEFHTLRP  
BT39536 290 300 310 320 330 340 350  
VYTDLPVSSBSLPELRLCRLWHTSAMFENLEAISSPHLFDYINLMAYTGSFS

MON863.pap 360 370 380 390 400 410  
GYFGRSPFNWQNVVTRPEIGSEKTISSPYGDKSTPEVKLEFDQGVYRTANTDV  
BT39536 360 370 380 390 400 410  
VHLLTDLISLWIGHSVSSLLASGPTVLRNNGS-TTSVNYFSNDRDYQVINTSHT

MON863.pap 420 430 440 450 460 470 480 490 500 510 519  
A-AWPNKGVYLVTKVPSOYDQKNETSTQYDSEKRNHVSQAQ--SIDQLPPEITDE  
BT39536 420 430 440 450 460 470 480 490 500 510 519  
GLGPNAPLF-GITRAQF--YDQ-----TYSVTQNALTCRQYNSIDELPDLNPR

MON863.pap 520 530 540 550 560 570 579  
KAYALSGASITRPGTQGNLLFLKSSNSIAKFKVTLNSAALLQRYVRIYASTNL  
BT39536 520 530 540 550 560 570 579  
KSPFVACTVVRGPTGGDLR-RTGVGTFTIRVR-TTABLQRYRIKRFASSTNL

MON863.pap 580 590 600 610 620 630 640 650 660 670 680  
RLFVONENDFLVIYIN-KTNNKDDLYTQYDPLATNSNMFGSGDKNELIGASFPVS  
BT39536 580 590 600 610 620 630 640 650 660 670 680  
-FICIGVGDGVYDFPQRTNMRGDELRYSEATREPTFNFPOPELISVFANAFSA

MON863.pap 640 650  
NEKVIYDIKIEFIPVL  
BT39536 640 650  
GQSVYFDRIEIPVNPABAKEDLEAKKAVASLFTATDGLQNVKYQVQDAANLVSC

MON863.pap TO.339536  
p1.339536 - parasporal crystal protein-related protein - *Bacillus thuringiensis*  
N:Alternate names delta-endotoxin-related protein  
C:Species *Bacillus thuringiensis*  
C.Date: 07-Oct-1994 #sequence\_revision 01-Dec-1995 #text\_change 16-Feb-1997  
C.Accession 339536  
R.Shevlev, A.B.; Dvarivsky, M.A.; Karasin, A.I.; Kogun, Y.N.; Chestukhina, .

SCORES Initl. 291 Initn. 1088 Opt: 751 z-score: 743.7 E(1): 5.0e-36  
Smith-Waterman score. 1231; 34.0% identity in 644 aa overlap

MON863.pap 20 30 40 50 60 70  
TIKVTNBSLQTNWQVPLADNPSLEELNYKEFLNMTK-DSETEVLNDSFVKAQVQ  
339536 20 30 40 50 60 70  
FLLLEHFLHVFELKTCIMAFPLTKLSYKYDLEMSBDYIDSYINPQVTOLOTO

MON863.pap 80 90 100 110 120 129  
ISVVOQLGVVGFAGALTSPYQSLNTMPS-DADPKAFMAQVRLIDKKIEYAKS  
339536 80 90 100 110 120 129  
SDIVAVVVAALGVGGLTDFLPLGLPGLMDNQAWKEFTISQERLISRLISDQVR

MON863.pap 130 140 150 160 170 180 189  
KALALQLOQLQNPEDYVNLNSKKTPLSLASRQGRIRLFSQAESHPRNKSPPAVS  
339536 130 140 150 160 170 180 189  
TALDLQLOQLQNPEDYVNLNSKKTPLSLASRQGRIRLFSQAESHPRNKSPPAVS

MON863.pap 190 200 210 220 230 240  
-KPEVLPFLPTAQAANTHLLKDAQVPSBNGYSSDVAEFPYRGLKL-TOQYTH  
339536 190 200 210 220 230 240  
PQQRQFQQLLVYQAANLHLLADAKYGAARGLRBOQNLVLSQVTRDYTH

MON863.pap 250 260 270 280 290 300  
CVWVYVQNLNGRSTIDAMVVRPREMTUVDLVLVFPFYDRLYSKGVKLETRD  
339536 250 260 270 280 290 300  
CVWVYVQNLNGRSTIDAMVVRPREMTUVDLVLVFPFYDRLYSKGVKLETRD

MON863.pap 310 320 330 340 350  
IFTPDIFLLTQKYGPF-----TFLSIENSI-RKPHLPDYLQIEFHTLRP  
339536 310 320 330 340 350  
VYTDLPVSSBSLPELRLCRLWHTSAMFENLEAISSPHLFDYINLMAYTGSFS

MON863.pap 360 370 380 390 400 410  
GYFGRSPFNWQNVVTRPEIGSEKTISSPYGDKSTPEVKLEFDQGVYRTANTDV  
339536 360 370 380 390 400 410  
VHLLTDLISLWIGHSVSSLLASGPTVLRNNGS-TTSVNYFSNDRDYQVINTSHT

MON863.pap 420 430 440 450 460 470 480 490 500 510 519  
A-AWPNKGVYLVTKVPSOYDQKNETSTQYDSEKRNHVSQAQ--SIDQLPPEITDE  
339536 420 430 440 450 460 470 480 490 500 510 519  
GLGPNAPLF-GITRAQF--YDQ-----TYSVTQNALTCRQYNSIDELPDLNPR

MON863.pap 520 530 540 550 560 570 579  
KAYALSGASITRPGTQGNLLFLKSSNSIAKFKVTLNSAALLQRYVRIYASTNL  
339536 520 530 540 550 560 570 579  
KSPFVACTVVRGPTGGDLR-RTGVGTFTIRVR-TTABLQRYRIKRFASSTNL

MON863.pap 580 590 600 610 620 630 640 650 660 670 680  
RLFVONENDFLVIYIN-KTNNKDDLYTQYDPLATNSNMFGSGDKNELIGASFPVS  
339536 580 590 600 610 620 630 640 650 660 670 680  
-FICIGVGDGVYDFPQRTNMRGDELRYSEATREPTFNFPOPELISVFANAFSA

MON863.pap 640 650  
NEKVIYDIKIEFIPVL  
339536 640 650  
GQSVYFDRIEIPVNPABAKEDLEAKKAVASLFTATDGLQNVKYQVQDAANLVSC

MON863.pap TO.339536  
p1.339536 - parasporal crystal protein-related protein - *Bacillus thuringiensis*  
N:Alternate names delta-endotoxin-related protein  
C:Species *Bacillus thuringiensis*  
C.Date: 07-Oct-1994 #sequence\_revision 01-Dec-1995 #text\_change 16-Feb-1997  
C.Accession 339536  
R.Shevlev, A.B.; Dvarivsky, M.A.; Karasin, A.I.; Kogun, Y.N.; Chestukhina, .

SCORES Initl. 291 Initn. 1088 Opt: 751 z-score: 743.7 E(1): 5.0e-36  
Smith-Waterman score. 1231; 34.0% identity in 644 aa overlap

MON863.pap 20 30 40 50 60 70  
TIKVTNBSLQTNWQVPLADNPSLEELNYKEFLNMTK-DSETEVLNDSFVKAQVQ  
339536 20 30 40 50 60 70  
FLLLEHFLHVFELKTCIMAFPLTKLSYKYDLEMSBDYIDSYINPQVTOLOTO

MON863.pap 80 90 100 110 120 129  
ISVVOQLGVVGFAGALTSPYQSLNTMPS-DADPKAFMAQVRLIDKKIEYAKS  
339536 80 90 100 110 120 129  
SDIVAVVVAALGVGGLTDFLPLGLPGLMDNQAWKEFTISQERLISRLISDQVR

MON863.pap 130 140 150 160 170 180 189  
KALALQLOQLQNPEDYVNLNSKKTPLSLASRQGRIRLFSQAESHPRNKSPPAVS  
339536 130 140 150 160 170 180 189  
TALDLQLOQLQNPEDYVNLNSKKTPLSLASRQGRIRLFSQAESHPRNKSPPAVS

MON863.pap 190 200 210 220 230 240  
-KPEVLPFLPTAQAANTHLLKDAQVPSBNGYSSDVAEFPYRGLKL-TOQYTH  
339536 190 200 210 220 230 240  
PQQRQFQQLLVYQAANLHLLADAKYGAARGLRBOQNLVLSQVTRDYTH

MON863.pap 300 310 320 330 340 350  
YSGVKTELTRDIFDPI--FLLTQKYGPTFLSIENSI-RKPHLPDYLQIEFHTLRP  
BT39536 300 310 320 330 340 350  
YPIQKSELTRDIFDPI--FLLTQKYGPTFLSIENSI-RKPHLPDYLQIEFHTLRP

MON863.pap 360 370 380 390 400 410  
GYFGRSPFNWQNVVTRPEIGSEKTISSPYGDKSTPEVKLEFDQGVYRTANTDV  
BT39536 360 370 380 390 400 410  
VHLLTDLISLWIGHSVSSLLASGPTVLRNNGS-TTSVNYFSNDRDYQVINTSHT

MON863.pap 420 430 440 450 460 470 480 490 500 510 519  
A-AWPNKGVYLVTKVPSOYDQKNETSTQYDSEKRNHVSQAQ--SIDQLPPEITDE  
BT39536 420 430 440 450 460 470 480 490 500 510 519  
GLGPNAPLF-GITRAQF--YDQ-----TYSVTQNALTCRQYNSIDELPDLNPR

MON863.pap 520 530 540 550 560 570 579  
KAYALSGASITRPGTQGNLLFLKSSNSIAKFKVTLNSAALLQRYVRIYASTNL  
BT39536 520 530 540 550 560 570 579  
KSPFVACTVVRGPTGGDLR-RTGVGTFTIRVR-TTABLQRYRIKRFASSTNL

MON863.pap 580 590 600 610 620 630 640 650 660 670 680  
RLFVONENDFLVIYIN-KTNNKDDLYTQYDPLATNSNMFGSGDKNELIGASFPVS  
BT39536 580 590 600 610 620 630 640 650 660 670 680  
-FICIGVGDGVYDFPQRTNMRGDELRYSEATREPTFNFPOPELISVFANAFSA

MON863.pap 640 650  
NEKVIYDIKIEFIPVL  
BT39536 640 650  
GQSVYFDRIEIPVNPABAKEDLEAKKAVASLFTATDGLQNVKYQVQDAANLVSC

MON863.pap TO.339536  
p1.339536 - parasporal crystal protein-related protein - *Bacillus thuringiensis*  
N:Alternate names delta-endotoxin-related protein  
C:Species *Bacillus thuringiensis*  
C.Date: 07-Oct-1994 #sequence\_revision 01-Dec-1995 #text\_change 16-Feb-1997  
C.Accession 339536  
R.Shevlev, A.B.; Dvarivsky, M.A.; Karasin, A.I.; Kogun, Y.N.; Chestukhina, .

SCORES Initl. 291 Initn. 1088 Opt: 751 z-score: 743.7 E(1): 5.0e-36  
Smith-Waterman score. 1231; 34.0% identity in 644 aa overlap

MON863.pap 20 30 40 50 60 70  
TIKVTNBSLQTNWQVPLADNPSLEELNYKEFLNMTK-DSETEVLNDSFVKAQVQ  
339536 20 30 40 50 60 70  
FLLLEHFLHVFELKTCIMAFPLTKLSYKYDLEMSBDYIDSYINPQVTOLOTO

MON863.pap 80 90 100 110 120 129  
ISVVOQLGVVGFAGALTSPYQSLNTMPS-DADPKAFMAQVRLIDKKIEYAKS  
339536 80 90 100 110 120 129  
SDIVAVVVAALGVGGLTDFLPLGLPGLMDNQAWKEFTISQERLISRLISDQVR

MON863.pap 130 140 150 160 170 180 189  
KALALQLOQLQNPEDYVNLNSKKTPLSLASRQGRIRLFSQAESHPRNKSPPAVS  
339536 130 140 150 160 170 180 189  
TALDLQLOQLQNPEDYVNLNSKKTPLSLASRQGRIRLFSQAESHPRNKSPPAVS

MON863.pap 190 200 210 220 230 240  
-KPEVLPFLPTAQAANTHLLKDAQVPSBNGYSSDVAEFPYRGLKL-TOQYTH  
339536 190 200 210 220 230 240  
PQQRQFQQLLVYQAANLHLLADAKYGAARGLRBOQNLVLSQVTRDYTH



Page 81 of 102

	340	350	360	370	380	390
MON863 .pep	YLQIGSEHTA--LRGYPQKDFKSPYMSGNYVSTRFSGISGKTSISFPYGDKSTPEQKLS					
	: :	: :	: :	: :	: :	: :
G99031	FLNNIISTGLSGLTSPFSTDAARWVGS--RDRISPAHSGYITELSGQHTATATL					
	: :	: :	: :	: :	: :	: :
	350	360	370	380	390	400
MON863 .pep	PDGQVPTTITADLAWNGYVQGLQTVKVRQDQDKHETSTGSDGR--RNG--HVS					
	: :	: :	: :	: :	: :	: :
G99031	--GRNIFPKVQSG--ACNLNDTTL--GVNRAVTF--YHDASEGQSGRVYEGYITATIDNPR					
	: :	: :	: :	: :	: :	: :
	410	420	430	440	450	
MON863 .pep	AQGEIDQLPSTDEPRKAYSHOLYARCF----LMQDRGRTIFPFTTHRSVDFFN					
	: :	: :	: :	: :	: :	: :
G99031	QGNITLKPENEDTPEDYTLSTLTITGLKQVARNRSGVSGKGLARNR					
	: :	: :	: :	: :	: :	: :
	460	470	480	490	500	
MON863 .pep	510	520	530	540	550	560
	TIQASKITLPPVKVAYLSSAGSIIEQPPQGNIILFLKSSNSIAKFVTLNAGAILQR					
	: :	: :	: :	: :	: :	: :
G99031	TINPRTGLTGLTKRGTQVSGVDFGGLG--GRTHDSGLGKLVQVRF--PLLHGLQ					
	: :	: :	: :	: :	: :	: :
	520	530	540	550	560	570
MON863 .pep	570	580	590	600	610	620
	YRVRIVASTTILRLFLQMSNGYFVIYIN--YKNNKDDQLVGLDATTSSSEFSD					
	: :	: :	: :	: :	: :	: :
G99031	YRVRIVASTTHRLSLV--NGCFGTIGNLTPLNRLEDGLGSPFAIRFETSIPTPLS					
	: :	: :	: :	: :	: :	: :
	580	590	600	610	620	

```

      630      640      650      660      670      680
MON863 pep
TG BACCRYIB_1

LOCUS      BACCRYIB_1
DEFINITION Bacillus thuringiensis crystal protein (cryI Et4) gene, complete
            cds.
DATE      26-APR-1994
ACCESSION L27019
NUC      gct4891 . . .

```

```

SCORES   Intrl  416   Intrl:  301   Ope  668   z-score: 661.3 R(1) 2 2e-31
Smith-Waterman score: 1116; 34.4 bits identified in 578 aa overlap

      60      70      80      90     100     110
MON863.pep  SSTEELNGTVKDAVDVGVQVGLGVGVPPAGALITSPYGFQSLINWPSDADFAWPH
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
BACCRVTB_1  EVLLDGRHILFDIDPLFVSLSLQLQLINSPVGGQFISGLVKNIGALRPFSE--KDLFL
      20      30      40      50      60      70

MON863.pep  180      130      140      150      160      170
AQMVLIDALNIEYAKSALALQGLQWVDFVVAINGNKNKTLPLRSKRSQOETRELL
BACCRVTB_1  AQRIELRLDNLVTRAKITSELGLRGNYVIAQKWEKNSPD--DNEAKSGVLDKRF
      80      90      100      110      120      130

MON863.pep  180      190      200      210      220      230
DQAEHPHNPSPSPFAVKFPEFLPLFQAQANLLELLLDQAQVGEWGSSESDVAEYFA
BACCRVTB_1  RILDLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLN
      140      150      160      170      180      190

MON863.pep  240      250      260      270      280      290
RQMLKLTQDCHNVCVGNLRLNGLVDFVDAWVYVNPFRKATVITVLQILVLPPTVIRL
BACCRVTB_1  RQIRLEIFENRQVQVITNLESLRQFSAIQRIYVQFRKLTITLQIALVFNQSDRL
      200      210      220      230      240      250

MON863.pep  300      310      320      330      340      350
YEGVKVTLTDIDFTDPT--FLLTQKQYQPTLSIENSRNPLQDQOQEPHNRFP
BACCRVTB_1  YPIQFISLTLRTEVYFVYVQVINGNIGLITLTQQQITRPLHIDFNMLKYVLS--
      260      270      280      290      300      310

MON863.pep  320      330      340      350      360      370
GYQKDFSNVNSGNVYTRFPLSGSSKSTISPYDQK--STAPVQKLSF--DQKQVRYTAN

```



















Page 87 of 102

160 270 360 450 540 630 720 810 900 990  
 MN863 pep LPPTTDEPLKAYSHQLMYAKVQDQDRQTIPFTTTHRSVDFNTIDAKYKTLPLVV  
 Q45882 YTHSLSMILANLNKKIQIIM-----DTGHSYSY-SWYKQIDGNYISKLNQIPLV  
 410 520 630 740 850 960  
 MN863 pep KAYALSSG-----ASIIERPGPTGQGNLLFLK--EESMII--AKP-----KVTL-----NSAALLQ  
 Q45882 KEVKLKRRHYSEIYSVTKGPTQGLILKRVHKPNQIPAQYMNKKITIPKTKFPAQSG  
 460 470 480 490 500 510  
 MN863 pep RYRVRYIYASTNLRFLVQNMNDFLVIYINKTKMKDDO-----LYQYOTPLATLNNMNF  
 Q45882 DFLVRLCYARNNDIGLRLKGLKNTVITLQQPTTTHNRLVLDYQKYPNHLPTSLT  
 520 530 540 550 560 570  
 MN863 pep SQEKEELIQAESPVSNKVIITDKISPTVQL  
 Q45882 SQSIDELYLFLFYSTYDGNFDFPKLSIPTPTANYC  
 580 590 600 610  
 MN863 pep  
 TO:JC6033  
 P1:JC6033 - monosaccharide protein Cbm71 - Clostridium bifementans  
 C:Species: Clostridium bifementans  
 C:Data: 1-Dec-1996 #sequence\_revision 31-Dec-1996 #text\_change 09-Sep-1997  
 C:Accession JCG033, P06009  
 J:Barloy, F.; Delecluse, A.; Nicolas, L.; Lecadet, M.M.  
 J: Bacteriol. 178, 3059-3105, 1996. . .  
 SCORRS Init1. 169 Initn. 316 Opt. 459 x-score 459.2 E(1) 4 la-20  
 Smith-Waterman score. 582; 26 1% identity in 579 an overlap  
 MN863 pep TEVLNLSFKDAVGTGIVLSVQGLGVVAFPAALTSFY--QSPLENTLV--PSDADWKA  
 JC6033 IFSTMLKFSQGAHVAKVVIDITIMLSNMNFMILQDPTDITSLIMDPOHNEIPEIS  
 10 20 30 40 50 60  
 MN863 pep FMAQVEVLIDDKISETAKSKALRLQGLQGNFDFVVALNKNKKTLPLSKRSRGORTRE  
 JC6033 KIDGQSTITKNLSAQTKSGLLSNSEFGLKFKYINAFRSW-----IDATNPNEDDQVY  
 70 80 90 100 110 120  
 MN863 pep LPSGAESHPSNPSFASVEFLPTTAAQANLHLLDAQVPGSHV-----GYSS  
 JC6033 RFDKGLKSENNINPEKHYEVVLTETITGIANLALLLADGMIYGDANWLRYGDSFD  
 130 140 150 160 170 180  
 MN863 pep EDVAFYRQKLTQVQVTCVNVNVLKGRISGTDYGVKFNRRFRREMTLVRLDVL  
 JC6033 QP--SYNVVLQKTRFINDCLMYVITGLNKLDPNNMWDITRYCSFMTYILKMSI  
 190 200 210 220 230  
 MN863 pep FFFYVIRKSGKQKTELTQIDPTDTELLTFLQYQPTFLSINIRKIRHLYDQLOIE  
 JC6033 CYDIAKVIDPTKQVYKQVSDPNNFIMILPEKVMYK-----SPSLSTFLFIS  
 240 250 260 270 280 290  
 MN863 pep FTKLLKPGYFKDSFNNNGVYVSRPSISSMSTTSPYQDN--STENKQIFSDQGVY  
 JC6033 FTTNKSQNFILNHRVHGVQVLTN-----SLST-----HYNGYSNYMSEVNSADIKAY  
 300 310 320 330 340  
 MN863 pep RTIANTVAAAMQGVYLVGTVKDFSDVQD--KNESQTVQSKN--NRHVSQAQSDGID  
 JC6033 SH--HYTNVQNMSTSVKSLKTLKNNHNVYSGESGNDITFTRQKTSIMN--ES  
 350 360 370 380 390 400









MON863 pep 400 410 420 430 440 450  
GOKVPSIANTDVAAMPKVVYDQTKVDFEQ--YDQKNSTSTQYDSKUNNGHVSQAQD  
B29838 400 410 420 430 440 450  
TQNPFFVYHR-----DIDKVELVHRHNSDIYKIFFNSSESVFRYSSNSTIENNYK  
MON863 pep 460 470 480 490 500 509  
SIDQAPPETDAPLAKAYSHHLYABC---FIMODHGTIPFFTWHRSDVDFFTID  
B29838 460 470 480 490 500 509  
RTDENVIPROKNNES--TQVTLSTYIKEDNYFSVVRHARVA--PEWHTSVDFOPTID  
MON863 pep 510 520 530 540 550 560 569  
ARKITQLPVKKAYALASGAGTKEPQGTGGHLLPKESNSIAKFKVTLSAALLQRYVRIRYASTTNLKL  
B29838 510 520 530 540 550 560 569  
LQNTQIALAKAKVESKTKVGPONTQDGLVILKQDMDFVR--LKNVB--ROYOV  
MON863 pep 570 580 590 600 610 620  
RIRYASTTNLRLVQNSNNDPLVIYINKMKDD---DIATYOTPLANT--INENNGFS  
B29838 570 580 590 600 610 620  
RIRYATNAP-KTTFVLGIDTISVELPSTSRQNPRAQDLTADPSKYTFPRYTHYK  
MON863 pep 630 640 650  
GDKNHLIIQAESFVSNKSIYIDKIEPIPVOL  
B29838 630 640 650  
GEDLMTLYQTPNHSVNIYIDKIEPIPIQSVLDYTEKQIEKTKIVNDLVNKKVLA  
MON863 pep TO:Q45789  
ID Q45789 PRELIMINARY, PRT, 297 AA  
AC Q45789  
DT 01-NOV-1996 (TREMBLREL 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1996 (TREMBLREL 01, LAST ANNOTATION UPDATE)  
DE DELTA-ENDOTOXIN (AA 429-726) (FRAGMENT) . . .  
SCORES Initl 170 Initn 326 Opt 394 z-score: 400.4 E(): 7.7e-17  
Smith-Waterman score: 394; 39 1% identity in 161 aa overlap  
MON863 pep 470 480 490 500 510 520  
ETTDEPLEKAYSHQLYACFLMDQRGTIPFFTWHRSDVDFMTIDAEKITQLPVVKAY  
BACCRVB\_1 470 480 490 500 510 520  
RLSHVSNFRSGFSSSVSEIIRAMPFSMHRASSTKILPSQITQIPUKST  
MON863 pep 530 540 550 560 570 580  
ALSSGASITTEPQGTGGHLLPKESNSIAKFKVTLSAALLQRYVRIRYASTTNLKL  
BACCRVB\_1 530 540 550 560 570 580  
NLGGSTSVVKGPGTGGDLR-RTSGQGLSTLAVNI-TAPLQRYVRIRYASTTNLQLH  
MON863 pep 590 600 610 620 630 640  
VQNSNNPLVIYINKMKDDDLTYOTPLATTNSNKGFGGCKNELTGAESFVSNKSIY  
BACCRVB\_1 590 600 610 620 630 640  
TSIDORIINQGNFSATMSGSSNLQSGSTFVGTTFPFGNSGSEVTLSAHVNSGNEVY  
MON863 pep 650  
IDKIEPIPVOL  
BACCRVB\_1 650  
IDRIEFPVAVTFEARYDLERAKQAVNELFTSSNGTGLKCDVTDYHIDQVNLVECLSD  
MON863 pep TO:Q45789  
LOCUS BACCRVB\_1  
DEFINITION Bacillus thuringiensis delta-endotoxin (CryVB) gene, 5' end.  
DATE 04-DEC-1999  
ACCESSION U07087  
NID g142771  
ORGANISM Bacillus thuringiensis  
SCORES Initl 202 Initn 280 Opt 303 z-score: 299.2 E(): 3.3e-11  
Smith-Waterman score: 397; 24.3% identity in 602 aa overlap  
MON863 pep 50 60 70 80 90 99  
LNYKFLNMTEDSTEVLSNVTXDAVGTGLVVGQILVVG-VPPAGALSYQGLFNT  
BACCRVB\_1 50 60 70 80 90 99  
AFENKKNKAGKOKDLLEVAVTYITTRIDPDAVKGLVLLTPEVTVASASTVAF  
MON863 pep 100 110 120 130 140 150  
IWP-----SDADPNKAFMAQVVLIDKIEYAKSKALAEQGLQNNFEDYVNLNSW  
BACCRVB\_1 100 110 120 130 140 150  
IWPKIPGDKPAKNIIFSLKPOIALIQQDITVQDAINQKFDLSOKTINLVYATDNN  
MON863 pep 160 170 180 190 200 210  
KKTPLRLSKRSQGLRELFOAESHPFMSNPSTAVKSEVLFPLTYAQAANTHLLKLD  
BACCRVB\_1 160 170 180 190 200 210  
DYV-----TAKTQ-----LETLASILTADLSIVPEQVETVYKAVANALLKLD  
MON863 pep 220 230 240 250 260 270  
AQVPEESGYSSEDAFVYRQGLKT-QVYTCRVNMYVGLNGKSTYDAWKVNRFR  
BACCRVB\_1 220 230 240 250 260 270  
ATVNAEKLPQSDKEV-DTHKCTVNTINHTFVAVKAPLGLKFKELDVNSYKAKNYI  
MON863 pep 280 290 300 310 320 330  
RNTLVLDLIVLFFPYTILYKSKGVETLTDITPDIPLTLTKQVPTFLSNGSTR  
BACCRVB\_1 280 290 300 310 320 330  
KNTBVLQDLVALMPTDPDHYKQVSEIETIT--SSPIV-----QVPKKNQNTSSIV  
MON863 pep 340 350 360 370 380  
KPHLPDLQ-----ISFTR-----LRQVYFG-KDFNYSNVSQVSTPAGSKITTSF  
BACCRVB\_1 340 350 360 370 380  
PSDLPHY-OGDLVKLEFSTKTDNDGAKIPGIRMTF--KSPHETHTVDPFVYTCQSG  
MON863 pep 390 400 410 420 430 440  
FYDKESTFVKLSFQGVYRTIANTDVAAMPKVVYDQTKVDFEQYDQKNSTSTQYDSKUNNGHVSQAQD  
BACCRVB\_1 390 400 410 420 430 440  
NIGRGSNPTP-IDLNPITSTCIRNSFYKATAQSEVL--VNV-----KDTQGYA  
SCORES Initl 170 Initn 326 Opt 394 z-score: 400.4 E(): 7.7e-17

Smith-Waterman score 394, 39 1% identity in 161 aa overlap  
MON863 pep 470 480 490 500 510 520  
ETTDEPLEKAYSHQLYACFLMDQRGTIPFFTWHRSDVDFMTIDAEKITQLPVVKAY  
BACCRVB\_1 470 480 490 500 510 520  
RLSHVSNFRSGFSSSVSEIIRAMPFSMHRASSTKILPSQITQIPUKST  
MON863 pep 530 540 550 560 570 580  
ALSSGASITTEPQGTGGHLLPKESNSIAKFKVTLSAALLQRYVRIRYASTTNLKL  
BACCRVB\_1 530 540 550 560 570 580  
NLGGSTSVVKGPGTGGDLR-RTSGQGLSTLAVNI-TAPLQRYVRIRYASTTNLQLH  
MON863 pep 590 600 610 620 630 640  
VQNSNNPLVIYINKMKDDDLTYOTPLATTNSNKGFGGCKNELTGAESFVSNKSIY  
BACCRVB\_1 590 600 610 620 630 640  
TSIDORIINQGNFSATMSGSSNLQSGSTFVGTTFPFGNSGSEVTLSAHVNSGNEVY  
MON863 pep 650  
IDKIEPIPVOL  
BACCRVB\_1 650  
IDRIEFPVAVTFEARYDLERAKQAVNELFTSSNGTGLKCDVTDYHIDQVNLVECLSD  
MON863 pep TO:Q45789  
LOCUS BACCRVB\_1  
DEFINITION Bacillus thuringiensis delta-endotoxin (CryVB) gene, 5' end.  
DATE 04-DEC-1999  
ACCESSION U07087  
NID g142771  
ORGANISM Bacillus thuringiensis  
SCORES Initl 202 Initn 280 Opt 303 z-score: 299.2 E(): 3.3e-11  
Smith-Waterman score: 397; 24.3% identity in 602 aa overlap  
MON863 pep 50 60 70 80 90 99  
LNYKFLNMTEDSTEVLSNVTXDAVGTGLVVGQILVVG-VPPAGALSYQGLFNT  
BACCRVB\_1 50 60 70 80 90 99  
AFENKKNKAGKOKDLLEVAVTYITTRIDPDAVKGLVLLTPEVTVASASTVAF  
MON863 pep 100 110 120 130 140 150  
IWP-----SDADPNKAFMAQVVLIDKIEYAKSKALAEQGLQNNFEDYVNLNSW  
BACCRVB\_1 100 110 120 130 140 150  
IWPKIPGDKPAKNIIFSLKPOIALIQQDITVQDAINQKFDLSOKTINLVYATDNN  
MON863 pep 160 170 180 190 200 210  
KKTPLRLSKRSQGLRELFOAESHPFMSNPSTAVKSEVLFPLTYAQAANTHLLKLD  
BACCRVB\_1 160 170 180 190 200 210  
DYV-----TAKTQ-----LETLASILTADLSIVPEQVETVYKAVANALLKLD  
MON863 pep 220 230 240 250 260 270  
AQVPEESGYSSEDAFVYRQGLKT-QVYTCRVNMYVGLNGKSTYDAWKVNRFR  
BACCRVB\_1 220 230 240 250 260 270  
ATVNAEKLPQSDKEV-DTHKCTVNTINHTFVAVKAPLGLKFKELDVNSYKAKNYI  
MON863 pep 280 290 300 310 320 330  
RNTLVLDLIVLFFPYTILYKSKGVETLTDITPDIPLTLTKQVPTFLSNGSTR  
BACCRVB\_1 280 290 300 310 320 330  
KNTBVLQDLVALMPTDPDHYKQVSEIETIT--SSPIV-----QVPKKNQNTSSIV  
MON863 pep 340 350 360 370 380  
KPHLPDLQ-----ISFTR-----LRQVYFG-KDFNYSNVSQVSTPAGSKITTSF  
BACCRVB\_1 340 350 360 370 380  
PSDLPHY-OGDLVKLEFSTKTDNDGAKIPGIRMTF--KSPHETHTVDPFVYTCQSG  
MON863 pep 390 400 410 420 430 440  
FYDKESTFVKLSFQGVYRTIANTDVAAMPKVVYDQTKVDFEQYDQKNSTSTQYDSKUNNGHVSQAQD  
BACCRVB\_1 390 400 410 420 430 440  
NIGRGSNPTP-IDLNPITSTCIRNSFYKATAQSEVL--VNV-----KDTQGYA

MON863 pep 450 460 470 480 490 500  
YDSKUNNGHVSQAQDSTIDQLPPTTDEPLEKAYSHQLYACFLMDQRGTIPFFTWHRSDVDFMTIDAEKITQLPVVKAYALSSGASILEE-PGF-TGGNLLFLKESNSIAKFKVTLSAALLQRYVRIRYASTTNLKL  
BACCRVB\_1 450 460 470 480 490 500  
FAQAPTGG-----AMHSPISDQAP-----EGHKLNIYTSPTDRLDFINVTIL--S  
MON863 pep 510 520 530 540 550 559  
VDFFMTIDAEKITQLPVVKAYALSSGASILEE-PGF-TGGNLLFLKESNSIAKFKVTLSAALLQRYVRIRYASTTNLKL  
BACCRVB\_1 510 520 530 540 550 559  
TPTINELSTKYGFPARKGYIKNGIKYKPEYINGAQPVLNQGOTLIFEH----  
MON863 pep 560 570 580 590 600 610  
SAALLQRYVRIRYASTTNLKLFGNSNNDPLVIYINKMKDD-DLTYOTPLATTNSN  
BACCRVB\_1 560 570 580 590 600 610  
-ASKTAQTITIRYASTQGTQGTFLDNGLOTLNPTPSINGYVTCNIGRVOLYTIQSY  
MON863 pep 620 630 640 650  
MGPSDCKNELTGAESFVSNKSIYIDKIEPIPVOL  
BACCRVB\_1 620 630 640 650  
TITEGNTLQIQHNDK-----NKGVLDRISFVPKDGLQSDSPQSPFVHSTIIFDKSST  
MON863 pep 650 660 670 680 690 700  
TWSNKHYSKIHLEGGSTTSQGSYFNRLNLNLFPTDTPNMRNHTHVNNQDNVYDQKDV  
BACCRVB\_1 650 660 670 680 690 700  
ID Q45754 PRELIMINARY, PRT: 1257 AA  
AC Q45754  
DT 01-NOV-1996 (TREMBLREL 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1996 (TREMBLREL 01, LAST ANNOTATION UPDATE)  
DE DELTA-ENDOTOXIN (FRAGMENT) . . .  
SCORES Initl 202 Initn 230 Opt 303 z-score: 299.2 E(): 3.3e-11  
Smith-Waterman score: 397; 24.3% identity in 602 aa overlap  
MON863 pep 50 60 70 80 90 99  
LNYKFLNMTEDSTEVLSNVTXDAVGTGLVVGQILVVG-VPPAGALSYQGLFNT  
Q45754 50 60 70 80 90 99  
AFENKKNKAGKOKDLLEVAVTYITTRIDPDAVKGLVLLTPEVTVASASTVAF  
MON863 pep 100 110 120 130 140 150  
IWP-----SDADPNKAFMAQVVLIDKIEYAKSKALAEQGLQNNFEDYVNLNSW  
Q45754 100 110 120 130 140 150  
IWPKIPGDKPAKNIIFSLKPOIALIQQDITVQDAINQKFDLSOKTINLVYATDNN  
MON863 pep 160 170 180 190 200 210  
KKTPLRLSKRSQGLRELFOAESHPFMSNPSTAVKSEVLFPLTYAQAANTHLLKLD  
Q45754 160 170 180 190 200 210  
DYV-----TAKTQ-----LETLASILTADLSIVPEQVETVYKAVANALLKLD  
MON863 pep 220 230 240 250 260 270  
AQVPEESGYSSEDAFVYRQGLKT-QVYTCRVNMYVGLNGKSTYDAWKVNRFR  
Q45754 220 230 240 250 260 270  
ATVNAEKLPQSDKEV-DTHKCTVNTINHTFVAVKAPLGLKFKELDVNSYKAKNYI  
MON863 pep 280 290 300 310 320 330  
RNTLVLDLIVLFFPYTILYKSKGVETLTDITPDIPLTLTKQVPTFLSNGSTR  
Q45754 280 290 300 310 320 330  
KNTBVLQDLVALMPTDPDHYKQVSEIETIT--SSPIV-----QVPKKNQNTSSIV  
MON863 pep 340 350 360 370 380  
KPHLPDLQ-----ISFTR-----LRQVYFG-KDFNYSNVSQVSTPAGSKITTSF  
Q45754 340 350 360 370 380  
PSDLPHY-OGDLVKLEFSTKTDNDGAKIPGIRMTF--KSPHETHTVDPFVYTCQSG  
MON863 pep 390 400 410 420 430 440  
FYDKESTFVKLSFQGVYRTIANTDVAAMPKVVYDQTKVDFEQYDQKNSTSTQYDSKUNNGHVSQAQD  
Q45754 390 400 410 420 430 440  
NIGRGSNPTP-IDLNPITSTCIRNSFYKATAQSEVL--VNV-----KDTQGYA



MON863.pap 450 460 470 480 490  
DSKNNKHVSAQSDIDLPPTTDEFLKAYSHQNYACFLMDQRROTIP--  
A29913 450 460 470 480 490  
DSKNNKHVSAQSDIDLPPTTDEFLKAYSHQNYACFLMDQRROTIP--  
MON863.pap 500 510 520 530  
VDFPNTIDAEKITOLPVVYKAYALSSGASII--BGPPTGNNLLFLKSSNSIAKFKVLNS  
A29913 500 510 520 530  
VDFPNTIDAEKITOLPVVYKAYALSSGASII--BGPPTGNNLLFLKSSNSIAKFKVLNS  
MON863.pap 540 550 560 570 580 590  
OGMLPLKSSNSIAKFKVLNSAALLORVNVNASTWRLPVOGNHOPNIYINK  
A29913 540 550 560 570 580 590  
OGMLPLKSSNSIAKFKVLNSAALLORVNVNASTWRLPVOGNHOPNIYINK  
MON863.pap 600 610 620 630 640 650  
THKKDDLTQYTFDLATNSINNOFSGDKNEITIGASVFNKSIYIDKIEPIVOL  
A29913 600 610 620 630 640 650  
THKKDDLTQYTFDLATNSINNOFSGDKNEITIGASVFNKSIYIDKIEPIVOL  
MON863.pap 580 590  
TTNDNDNDGARFRFY  
A29913 580 590  
TTNDNDNDGARFRFY  
MON863.pap  
TO S17402  
P1.S17402 - parasporal crystal protein cryIIC - Bacillus thuringiensis  
plasmid  
N.Alternate names: delta-endotoxin  
C.Species: Bacillus thuringiensis  
C.Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Sep-1997  
C.Accession: S17402  
R.Wu, D. Cao, X. L. Bai, Y. Y. Aronson, A. I. . .  
SCORES Initl 140 Inith 140 Opt 276 z-score: 277.8 E(): 5.2e-10  
Smith-Waterman score: 300; 21.2% identity in 603 aa overlap  
MON863.pap 50 60 70 80 90  
KEFLRMKEDSTEVLSNSTVKDAVGTISVVGQILVVG--VPFAGALTS--FYQSFLN  
S17402 50 60 70 80 90  
KEFLRMKEDSTEVLSNSTVKDAVGTISVVGQILVVG--VPFAGALTS--FYQSFLN  
MON863.pap 100 110 120 130 140 150  
TIMPSDA-DPWKAFMAOVVLLDKIKIEYAKSALAELOGQNN--FEDYVNAIWSKK  
S17402 100 110 120 130 140 150  
TIMPSDA-DPWKAFMAOVVLLDKIKIEYAKSALAELOGQNN--FEDYVNAIWSKK  
MON863.pap 160 170 180 190 200 210  
TPLSLRSKRSQRIELFSQASHPFNKSPFAVSEKRVFLPTVQAQANTHLLKDAQ  
S17402 160 170 180 190 200 210  
TPLSLRSKRSQRIELFSQASHPFNKSPFAVSEKRVFLPTVQAQANTHLLKDAQ  
MON863.pap 220 230 240 250 260 270  
VFGESVYSSDVAEFYRQGLKQYTDRCVNTDNLGNKSTYDAWVKFRPAREX  
S17402 220 230 240 250 260 270  
VFGESVYSSDVAEFYRQGLKQYTDRCVNTDNLGNKSTYDAWVKFRPAREX  
MON863.pap 280 290 300 310 320  
TDLVLDLV--LFPFYDI--RLYSKGVKTELTRDIPDIPLLTLQKYQTFLS  
S17402 280 290 300 310 320  
TDLVLDLV--LFPFYDI--RLYSKGVKTELTRDIPDIPLLTLQKYQTFLS  
MON863.pap 330 340 350 360 370 380  
TENSATKPHLDYLOIE-FHRLAFQYFGRDSFNNSGVYVTR--PSIGSKTITSP  
S17402 330 340 350 360 370 380  
TENSATKPHLDYLOIE-FHRLAFQYFGRDSFNNSGVYVTR--PSIGSKTITSP  
MON863.pap 390 400 410 420 430 440  
FYQKSTHVPKLSFGQKVVYRTIANTDVAANPKVGLVGVKDFSOYDQKNSTGT  
S17402 390 400 410 420 430 440  
FYQKSTHVPKLSFGQKVVYRTIANTDVAANPKVGLVGVKDFSOYDQKNSTGT

MON863.pap 450 460 470 480 490 500  
YDSKNNKHVSAQSDIDLPPTTDEFLKAYSHQNYACFLMDQRROTIPPTWTHRS  
S17402 450 460 470 480 490 500  
YDSKNNKHVSAQSDIDLPPTTDEFLKAYSHQNYACFLMDQRROTIPPTWTHRS  
MON863.pap 510 520 530 540 550 560  
VDFPNTIDAEKITOLPVVYKAYALSSGASII--BGPPTGNNLLFLKSSNSIAKFKVLNS  
S17402 510 520 530 540 550 560  
VDFPNTIDAEKITOLPVVYKAYALSSGASII--BGPPTGNNLLFLKSSNSIAKFKVLNS  
MON863.pap 570 580 590 600 610 620  
AALLQRYKVRIRYASTTNLALFVQNSNDPLVIYINKWRRDDLTQYTFDLATNSINNO  
S17402 570 580 590 600 610 620  
AALLQRYKVRIRYASTTNLALFVQNSNDPLVIYINKWRRDDLTQYTFDLATNSINNO  
MON863.pap 630 640 650  
FSGDKNEITIGASVFNKSIYIDKIEPIVOL  
S17402 630 640 650  
FSGDKNEITIGASVFNKSIYIDKIEPIVOL  
MON863.pap 590 600 610 620  
ASANTVPLDTQYTFDLATNSINNOFSGDKNEITIGASVFNKSIYIDKIEPIVOL  
S17402 590 600 610 620  
ASANTVPLDTQYTFDLATNSINNOFSGDKNEITIGASVFNKSIYIDKIEPIVOL  
MON863.pap  
TO:BTCRYIIC\_3  
LOCUS BTCRYIIC\_3  
DEFINITION B. thuringiensis plasmid DNA for cryIIC operon and gene for  
cryIIC delta-endotoxin.  
DATE 29-AUG-1991  
ACCESSION X57252  
ID Q45743  
SCORES Initl 140 Inith 140 Opt 276 z-score: 277.8 E(): 5.2e-10  
Smith-Waterman score: 300; 21.2% identity in 603 aa overlap  
MON863.pap 50 60 70 80 90  
KEFLRMKEDSTEVLSNSTVKDAVGTISVVGQILVVG--VPFAGALTS--FYQSFLN  
S17402 50 60 70 80 90  
KEFLRMKEDSTEVLSNSTVKDAVGTISVVGQILVVG--VPFAGALTS--FYQSFLN  
MON863.pap 100 110 120 130 140 150  
TIMPSDA-DPWKAFMAOVVLLDKIKIEYAKSALAELOGQNN--FEDYVNAIWSKK  
S17402 100 110 120 130 140 150  
TIMPSDA-DPWKAFMAOVVLLDKIKIEYAKSALAELOGQNN--FEDYVNAIWSKK  
MON863.pap 160 170 180 190 200 210  
TPLSLRSKRSQRIELFSQASHPFNKSPFAVSEKRVFLPTVQAQANTHLLKDAQ  
S17402 160 170 180 190 200 210  
TPLSLRSKRSQRIELFSQASHPFNKSPFAVSEKRVFLPTVQAQANTHLLKDAQ  
MON863.pap 220 230 240 250 260 270  
VFGESVYSSDVAEFYRQGLKQYTDRCVNTDNLGNKSTYDAWVKFRPAREX  
S17402 220 230 240 250 260 270  
VFGESVYSSDVAEFYRQGLKQYTDRCVNTDNLGNKSTYDAWVKFRPAREX  
MON863.pap 280 290 300 310 320  
TDLVLDLV--LFPFYDI--RLYSKGVKTELTRDIPDIPLLTLQKYQTFLS  
S17402 280 290 300 310 320  
TDLVLDLV--LFPFYDI--RLYSKGVKTELTRDIPDIPLLTLQKYQTFLS  
MON863.pap 330 340 350 360 370 380  
TENSATKPHLDYLOIE-FHRLAFQYFGRDSFNNSGVYVTR--PSIGSKTITSP  
S17402 330 340 350 360 370 380  
TENSATKPHLDYLOIE-FHRLAFQYFGRDSFNNSGVYVTR--PSIGSKTITSP  
MON863.pap 390 400 410 420 430 440  
FYQKSTHVPKLSFGQKVVYRTIANTDVAANPKVGLVGVKDFSOYDQKNSTGT  
S17402 390 400 410 420 430 440  
FYQKSTHVPKLSFGQKVVYRTIANTDVAANPKVGLVGVKDFSOYDQKNSTGT

MON863.pap 510 520 530 540 550 560  
VDFPNTIDAEKITOLPVVYKAYALSSGASII--BGPPTGNNLLFLKSSNSIAKFKVLNS  
BTCRYIIC\_3 510 520 530 540 550 560  
VDFPNTIDAEKITOLPVVYKAYALSSGASII--BGPPTGNNLLFLKSSNSIAKFKVLNS  
MON863.pap 570 580 590 600 610 620  
AALLQRYKVRIRYASTTNLALFVQNSNDPLVIYINKWRRDDLTQYTFDLATNSINNO  
BTCRYIIC\_3 570 580 590 600 610 620  
AALLQRYKVRIRYASTTNLALFVQNSNDPLVIYINKWRRDDLTQYTFDLATNSINNO  
MON863.pap 630 640 650  
FSGDKNEITIGASVFNKSIYIDKIEPIVOL  
BTCRYIIC\_3 630 640 650  
FSGDKNEITIGASVFNKSIYIDKIEPIVOL  
MON863.pap 590 600 610 620  
ASANTVPLDTQYTFDLATNSINNOFSGDKNEITIGASVFNKSIYIDKIEPIVOL  
BTCRYIIC\_3 590 600 610 620  
ASANTVPLDTQYTFDLATNSINNOFSGDKNEITIGASVFNKSIYIDKIEPIVOL  
MON863.pap  
TO:Q45743  
ID Q45743 PRELIMINARY; PRT. 622 AA.  
AC Q45743; DT 01-NOV-1996 (TRMBLREL 01, CREATED)  
DT 01-NOV-1996 (TRMBLREL 01, LAST SEQUENCE UPDATE)  
DE 01-NOV-1996 (TRMBLREL 01, LAST ANNOTATION UPDATE)  
DE CRYIIC DELTA-ENDOTOXIN  
SCORES Initl 140 Inith 140 Opt 276 z-score: 277.8 E(): 5.2e-10  
Smith-Waterman score: 300; 21.2% identity in 603 aa overlap  
MON863.pap 50 60 70 80 90  
KEFLRMKEDSTEVLSNSTVKDAVGTISVVGQILVVG--VPFAGALTS--FYQSFLN  
Q45743 50 60 70 80 90  
KEFLRMKEDSTEVLSNSTVKDAVGTISVVGQILVVG--VPFAGALTS--FYQSFLN  
MON863.pap 100 110 120 130 140 150  
TIMPSDA-DPWKAFMAOVVLLDKIKIEYAKSALAELOGQNN--FEDYVNAIWSKK  
Q45743 100 110 120 130 140 150  
TIMPSDA-DPWKAFMAOVVLLDKIKIEYAKSALAELOGQNN--FEDYVNAIWSKK  
MON863.pap 160 170 180 190 200 210  
TPLSLRSKRSQRIELFSQASHPFNKSPFAVSEKRVFLPTVQAQANTHLLKDAQ  
Q45743 160 170 180 190 200 210  
TPLSLRSKRSQRIELFSQASHPFNKSPFAVSEKRVFLPTVQAQANTHLLKDAQ  
MON863.pap 220 230 240 250 260 270  
VFGESVYSSDVAEFYRQGLKQYTDRCVNTDNLGNKSTYDAWVKFRPAREX  
Q45743 220 230 240 250 260 270  
VFGESVYSSDVAEFYRQGLKQYTDRCVNTDNLGNKSTYDAWVKFRPAREX  
MON863.pap 280 290 300 310 320  
TDLVLDLV--LFPFYDI--RLYSKGVKTELTRDIPDIPLLTLQKYQTFLS  
Q45743 280 290 300 310 320  
TDLVLDLV--LFPFYDI--RLYSKGVKTELTRDIPDIPLLTLQKYQTFLS  
MON863.pap 330 340 350 360 370 380  
TENSATKPHLDYLOIE-FHRLAFQYFGRDSFNNSGVYVTR--PSIGSKTITSP  
Q45743 330 340 350 360 370 380  
TENSATKPHLDYLOIE-FHRLAFQYFGRDSFNNSGVYVTR--PSIGSKTITSP  
MON863.pap 390 400 410 420 430 440  
FYQKSTHVPKLSFGQKVVYRTIANTDVAANPKVGLVGVKDFSOYDQKNSTGT  
Q45743 390 400 410 420 430 440  
FYQKSTHVPKLSFGQKVVYRTIANTDVAANPKVGLVGVKDFSOYDQKNSTGT  
MON863.pap 450 460 470 480 490 500  
YDSKNNKHVSAQSDIDLPPTTDEFLKAYSHQNYACFLMDQRROTIPPTWTHRS  
Q45743 450 460 470 480 490 500  
YDSKNNKHVSAQSDIDLPPTTDEFLKAYSHQNYACFLMDQRROTIPPTWTHRS  
MON863.pap 510 520 530 540 550 560  
VDFPNTIDAEKITOLPVVYKAYALSSGASII--BGPPTGNNLLFLKSSNSIAKFKVLNS  
Q45743 510 520 530 540 550 560  
VDFPNTIDAEKITOLPVVYKAYALSSGASII--BGPPTGNNLLFLKSSNSIAKFKVLNS

570 580 590 600 610 620  
MON863.pep AALLORFVIRYASTTNLNFVNNDLVIYINKNKDDDLTYQTFDLATTNENMG  
Q45712 NST--ALYLVGEGSSTVIT--MNVVKAAMVTTHNDGLNARFSDINGNV  
540 550 560 570 580  
MON863.pep PQQKQNLIGABCFVMSGLYIKLPIVUL  
Q45712 ASANTVPLDITQFMNPPQFNNINPVNLPFLY  
590 600 610 620  
MON863.pep TO-BACCRIIIIA\_1  
LOCUS BACCRIIIIA\_1  
DEFINITION Bacillus thuringiensis insecticidal toxin (cryIIIA) gene, 5'  
end  
DATE 30-JUN-1993  
ACCESSION L03393  
NID g304150  
ORGANISM Bacillus thuringiensis . .  
SCORES Initl: 255 Initn: 255 Opt: 255 z-score: 276.5 E(): 4.2e-10  
Smith-Waterman score: 255, 79.2% identity in 48 aa overlap  
MON863.pep MANPNRSHSDTIKVTNSELQNNQYPLADNPSTLEELNYKEFLNTEDESTEVLON  
BACCRIIIIA\_1 MANPNRSHSDTIKVTNSELQNNQYPLADNPSTLEELNYKEFLNTEDESTEVLON  
70 80 90 100 110 120  
MON863.pep STVKDAVGTSISVVGQILGVVG-VFAGALTSPFYQSFLNTIMPS--  
MON863.pep TO-BTTOXD\_1  
LOCUS BTTOXD\_1  
DEFINITION B. thuringiensis 130kd mosquito-larvicidal delta-endotoxin gene  
fragment:  
endotoxin (AA 1-142) (577 is 1st base in codon).  
DATE 14-OCT-1993  
ACCESSION X05692 .  
SCORES Initl: 103 Initn: 180 Opt: 237 z-score: 250.6 E(): 1.7e-08  
Smith-Waterman score: 237, 28.4% identity in 134 aa overlap  
MON863.pep MANPNRSHSDTIKVTNSELQNNQYPLADNPSTLEELNYKEFLNTEDESTEVLON  
BTTOXD\_1 MNSGYPLANLOGSNNHTYKDLAKCHNQQYGVNP  
10 20 30 40 50 60  
MON863.pep STVKDA-VGTISVVGQILGVVG-VFAGALTSPFYQSFLNTIMPSDAD--WIAFMACV  
BTTOXD\_1 AAINSSGVSTALKVAGALKVNP-AOTLVLSAVLPLALPTPTTPRNVNDFMT  
70 80 90 100 110  
MON863.pep EVLIDKIKIEYAKALAEQLQNNFEDYVNALNSVKTPLSLRSKRSGQRIELPSQA  
BTTOXD\_1 DMLDQVTAYVETDRAKSTVVKVLDQITKTFPFWKASPNQSY  
100 110 120 130 140  
MON863.pep ESHFNSNPSGAVSGSEVLEFLPTAQAANTHLLLLKDAQVPSGEMYSSEDAVFYRRQL  
MON863.pep TO-Q45712  
ID Q45712 PRELIMINARY, PRF, 1245 AA.  
AC Q45712, DT 01-NOV-1996 (TREMBLREL 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1996 (TREMBLREL 01, LAST ANNOTATION UPDATE)  
DE DELTA-ENDOTOXIN (FRAGMENT).  
SCORES Initl: 100 Initn: 136 Opt: 211 z-score: 208.1 E(): 4e-06  
Smith-Waterman score: 282, 21.9% identity in 648 aa overlap

50 60 70 80 90 100  
MON863.pep RMCEDESTEVLNDSVTKDAVGTSISVVGQILGVVG-VFAGALTSPFYQSFLNTIMPS--  
Q45712 QKLLFEDHLTIAMNLYTKGLDYFALTAKSISLIGFIPGAEAVPPINKFVDFVWPKLFG  
50 60 70 80 90 100  
MON863.pep -----DADPWKAPMAQVEVLIDKIKIEYAKALAEQLQNNFEDYVNALNSVKT  
Q45712 ANTRGKQQLFNAIMDAVNHVKNVPLSTNLTKTIEALQNLGLFQNAIQVAICQGS  
110 120 130 140 150 160  
MON863.pep TPLSLR-----SKRSQRIRELPSQAESHFRNSMPST-----AVSKFEV  
Q45712 TPRVNFQNCCTPCNMPQCKDGLKVASRFDIANSQFQHLPEFKNPMSDSTQEFKR  
170 180 190 200 210 220  
MON863.pep -----LFLPTTAQAANTHLLLLKDAQVPSGEMYSSEDAVFYRRQL-LTQYTDHCVNM  
Q45712 TSVLELPLMTTATLHLLLYGYIEFNTKWNPNHGYLNNKVELQQLINHSYETVRS  
230 240 250 260 270 280  
MON863.pep YNVLNGLRGSTYDAMVKNRPRENTLVLDLVLFPFYDRLYSKGVKTELRDITPD  
Q45712 FLQFLPLNNRSGSVNAYNRYVNRVNCIDIAATWPTTDTNNHYNHCKGLDLRLIED  
290 300 310 320 330 340  
MON863.pep ---DIFLLITLQKY-GPTFLSTENSRKPLFDYLOQIEPHRLARQYFGKDSFNYSGN  
Q45712 TAPFIEYTTGDSGPE-----HNSITPMNLDWPTSTYQHS-----FVSGDLYSRKE  
350 360 370 380 390 400  
MON863.pep VYSTRISGSSKTIYS--EYQKSTEFVOKLST--DQNVYRTIANTVAANPKGVYLG  
Q45712 LQQLDAEATYNSNNCHPVLISYTDLSRYVDGNDQPTTNNCHNSYADITPLV  
400 410 420 430 440 450  
MON863.pep VTKVDFSQYDDQKNESTQYDSKRNKHVSADSIDQI--PPETDEPKAYSKGLNYA  
Q45712 NARHLVNAKSLQVSELVVSQVNOGSGSCIDAMITRLAPQSKN--EERPKINVL  
460 470 480 490 500 510  
MON863.pep ECFLNODRGTIPFFTWHRSDVFF-----NTI-DARKITQLPV--VKAYALSQASIIED  
Q45712 YPIETETVNGTQGNLGVISAVPMELVPSNIGVDNADTKLPLTQKQPFPRKYSGEYNN  
520 530 540 550 560 570  
MON863.pep PGFT-----GNNLLFLKESSENSIAKFKVTLNSAALLQRYVRIRYASTNMLRFVONS  
Q45712 ROISLVREKINGNAKLS-NSQSVG-IQITQTK--OKYELKRYAS-----KGD  
580 590 600 610 620  
MON863.pep NDFPLVIYINKNKDDDLTYQTFDLATTNENMGPSQDKNELI-----IGAEF--  
Q45712 NN-----VYFNVDSNPFRNSISFG-STESVVGQVGSNGKYLKSTITVEIPAGSYVH  
630 640 650 660 670  
MON863.pep VSENE--KIYIDKIEPIVQVL  
Q45712 ITQVGSSEDLRLIEVPKIQGFQDNNHLCDCNPNVDTCDFCCVCTSLTDCDCNPR  
680 690 700 710 720 730  
MON863.pep TO-BTU19725\_1  
LOCUS BTU19725\_1  
DEFINITION Bacillus thuringiensis F88G3 delta endotoxin gene, partial cds.  
DATE 21-JUN-1995  
ACCESSION U19725  
NID g86263  
ORGANISM Bacillus thuringiensis . .  
SCORES Initl: 100 Initn: 136 Opt: 211 z-score: 208.1 E(): 4e-06  
Smith-Waterman score: 282, 21.9% identity in 648 aa overlap

50 60 70 80 90 100  
MON863.pep RMTEDESTEVLNDSVTKDAVGTSISVVGQILGVVG-VFAGALTSPFYQSFLNTIMPS--  
BTU19725\_1 QKLLFEDHLTIAMNLYTKGLDYFALTAKSISLIGFIPGAEAVPPINKFVDFVWPKLFG  
50 60 70 80 90 100  
MON863.pep -----DADPWKAPMAQVEVLIDKIKIEYAKALAEQLQNNFEDYVNALNSVKT  
BTU19725\_1 ANTRGKQQLFNAIMDAVNHVKNVPLSTNLTKTIEALQNLGLFQNAIQVAICQGS  
110 120 130 140 150 160  
MON863.pep TPLSLR-----SKRSQRIRELPSQAESHFRNSMPST-----AVSKFEV  
BTU19725\_1 TPRVNFQNCCTPCNMPQCKDGLKVASRFDIANSQFQHLPEFKNPMSDSTQEFKR  
170 180 190 200 210 220  
MON863.pep -----LFLPTTAQAANTHLLLLKDAQVPSGEMYSSEDAVFYRRQL-LTQYTDHCVNM  
BTU19725\_1 TSVLELPLMTTATLHLLLYGYIEFNTKWNPNHGYLNNKVELQQLINHSYETVRS  
230 240 250 260 270 280  
MON863.pep YNVLNGLRGSTYDAMVKNRPRENTLVLDLVLFPFYDRLYSKGVKTELRDITPD  
BTU19725\_1 FLQFLPLNNRSGSVNAYNRYVNRVNCIDIAATWPTTDTNNHYNHCKGLDLRLIED  
290 300 310 320 330 340  
MON863.pep ---DIFLLITLQKY-GPTFLSTENSRKPLFDYLOQIEPHRLARQYFGKDSFNYSGN  
BTU19725\_1 TAPFIEYTTGDSGPE-----HNSITPMNLDWPTSTYQHS-----FVSGDLYSRKE  
350 360 370 380 390 400  
MON863.pep VYSTRISGSSKTIYS--EYQKSTEFVOKLST--DQNVYRTIANTVAANPKGVYLG  
BTU19725\_1 LQQLDAEATYNSNNCHPVLISYTDLSRYVDGNDQPTTNNCHNSYADITPLV  
400 410 420 430 440 450  
MON863.pep VTKVDFSQYDDQKNESTQYDSKRNKHVSADSIDQI--PPETDEPKAYSKGLNYA  
BTU19725\_1 NARHLVNAKSLQVSELVVSQVNOGSGSCIDAMITRLAPQSKN--EERPKINVL  
460 470 480 490 500 510  
MON863.pep ECFLNODRGTIPFFTWHRSDVFF-----NTI-DARKITQLPV--VKAYALSQASIIED  
BTU19725\_1 YPIETETVNGTQGNLGVISAVPMELVPSNIGVDNADTKLPLTQKQPFPRKYSGEYNN  
520 530 540 550 560 570  
MON863.pep PGFT-----GNNLLFLKESSENSIAKFKVTLNSAALLQRYVRIRYASTNMLRFVONS  
BTU19725\_1 ROISLVREKINGNAKLS-NSQSVG-IQITQTK--OKYELKRYAS-----KGD  
580 590 600 610 620 630  
MON863.pep NDFPLVIYINKNKDDDLTYQTFDLATTNENMGPSQDKNELI-----IGAEF--  
BTU19725\_1 NN-----VYFNVDSNPFRNSISFG-STESVVGQVGSNGKYLKSTITVEIPAGSYVH  
640 650 660 670 680 690  
MON863.pep VSENE--KIYIDKIEPIVQVL  
BTU19725\_1 ITQVGSSEDLRLIEVPKIQGFQDNNHLCDCNPNVDTCDFCCVCTSLTDCDCNPR  
700 710 720 730  
MON863.pep TO-Q45753  
ID Q45753 PRELIMINARY, PRF, 1298 AA.  
AC Q45753, DT 01-NOV-1996 (TREMBLREL 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1996 (TREMBLREL 01, LAST ANNOTATION UPDATE)  
DE DELTA-ENDOTOXIN (FRAGMENT).  
SCORES Initl: 67 Initn: 152 Opt: 164 z-score: 161.3 E(): 0.006  
Smith-Waterman score: 268, 20.6% identity in 695 aa overlap



LOCUS BACDRYVAB\_1  
DEFINITION Bacillus thuringiensis delta-endotoxin (CryA(b)) gene, 5' end.  
DATE 30-MAY-1995  
ACCESSION L07026

MON863.pap MANPNRSEHDITKVNSSELQTHNQYPLADNPSTLEEL--NYKEFLRMTEDEST  
Q45753 MATLNELVPSVYVNLAYTPSPFLDAGTQATPADL--TAYEQLLNKLEKINAGTYSKA  
10 20 30 40 50  
60 70 80 90 100  
MON863.pap --EVLNSTVKDAVGTGISV-VGQILGVGVFPAGALTSFYOSFLNTIWPSDADP-----  
Q45753 IADVLKGIPIIDTINQYVNLGLSLITLAVPEIGITPTPIGLFAALNKHDAPPPNAK  
60 70 80 90 100 110  
120 130 140 150 160  
MON863.pap --WKAFAQVEVLIDKKIEYAKSALAELOGLNFFEDYVNAISWKKTPSLRSKRQ  
Q45753 DIFEMKPAIQEMIDRTLTADQOTFLNGEISGLNLAARYQSTMDIQSH--GGFNKVD  
120 130 140 150 160 170  
180 190 200 210 219  
MON863.pap ORIKELFSQA---ESFPMNSMPSFAVSKF---EVLFLTYAQAANTHLLKDAQVQEE  
Q45753 OLKKFPTDEVLSNFSYTDRLVPVITNTADRTLLGLPYTALASMLMLRDIITKQPT  
180 190 200 210 220 230  
240 250 260 270 280 290  
MON863.pap WQYS--SEDAVFYRQLKLT-QQYTDHCVMYVNLGL--RGSTYDAMVFNPRREM  
Q45753 WQKINFTPDADISFPTDIKNNKLYSKTYVDFQGLASYOTPSDLESAKKQYIEIM  
240 250 260 270 280 290  
300 310 320 330 340 350  
MON863.pap TLTVLDELIVLFPFDYLYEYKQ---VATLRTDITPTDPLLT---TLQKYQ-----P  
Q45753 TTHCLDFAFLPTFDYLYEYKQ---VATLRTDITPTDPLLT---TLQKYQ-----P  
300 310 320 330 340 350  
360 370 380 390 400 410  
MON863.pap TFLSIENSTRKPHLDYLOGIEFHTRAKPGYFKDPSND--WS---GNVETR--PSIGSS  
Q45753 NYZMCKAFPMKE--RIKQPKLYPSWRAQYQGLLOPYLWALISVQDSVETRLVQGLFV  
360 370 380 390 400 410  
420 430 440 450 460 470  
MON863.pap KITTSFY--GDKSTSPVKLSFDGQKVVRTIANTDVAANMKVLYGVTKVDFSOYDDOK  
Q45753 DPOAGNYVSDISNFIQINMDTKT---PPQAGSWNTNLRGVSGLSLFRQDTR  
420 430 440 450 460 470  
480 490 500 510 520 530  
MON863.pap NISTSTOT--YDSKRNHGVSAQSDIQGLPPTTDEPLKAYSHQINABCFIMQDRRTI  
Q45753 LSAGNGGFPADITYSLPATHYLSYLYGTPTQSDN---YSGHVG--ALVGVSTPQEAZL  
480 490 500 510 520 530  
540 550 560 570 580 590  
MON863.pap PPFVTHRGVDFPNTIDAKITQLPVVAYALSSQASITIEGPOFGNMLFLKESNSIA  
Q45753 PNLIGQDEQONVTHG---PPFKA---SYQGVTVV--HWLQANMLKSPQ--ETG  
530 540 550 560 570 580  
590 600 610 620 630  
MON863.pap KFKVTLNSAALLQRYVRIRYASTNLRFL--VQNSNDPLVIYNTKNDKDDLYOTF  
Q45753 IPIYNTSG---SYQICRYASNDNTNVPFVDTGAGNFIQOINASTVNDNTGVQGA  
580 590 600 610 620 630  
640 650 660 670 680 690  
MON863.pap DLATTNSNMFGSDKNELIQAESF---VSNB---KIYIDKIEFIPVQL  
Q45753 NGVTVKSIATNTSFTVKIPIKTIINVLNGSSDQVLDRIEFLBSNVTFIWNYS  
640 650 660 670 680 690  
700 710 720 730 740 750  
Q45753 TTGSANLIPALPMTSTSDKALTGSMTIGRTTPNSDALLRFKTYDTOTIPIPGSG  
700 710 720 730 740 750  
MON863.pap TO: BACDRYVAB\_1

LOCUS g142769  
ORGANISM Bacillus thuringiensis  
SCORES Initl: 67 Initn: 152 Opt: 164 z-score: 161.3 E(): 0.0016  
Smith-Waterman score: 268, 20.6% identity in 695 aa overlap

MON863.pap MANPNRSEHDITKVNSSELQTHNQYPLADNPSTLEEL--NYKEFLRMTEDEST  
BACDRYVAB\_1 MATLNELVPSVYVNLAYTPSPFLDAGTQATPADL--TAYEQLLNKLEKINAGTYSKA  
10 20 30 40 50  
60 70 80 90 100  
MON863.pap --EVLNSTVKDAVGTGISV-VGQILGVGVFPAGALTSFYOSFLNTIWPSDADP-----  
BACDRYVAB\_1 IADVLKGIPIIDTINQYVNLGLSLITLAVPEIGITPTPIGLFAALNKHDAPPPNAK  
60 70 80 90 100 110  
120 130 140 150 160  
MON863.pap --WKAFAQVEVLIDKKIEYAKSALAELOGLNFFEDYVNAISWKKTPSLRSKRQ  
BACDRYVAB\_1 DIFEMKPAIQEMIDRTLTADQOTFLNGEISGLNLAARYQSTMDIQSH--GGFNKVD  
120 130 140 150 160 170  
180 190 200 210 219  
MON863.pap ORIKELFSQA---ESFPMNSMPSFAVSKF---EVLFLTYAQAANTHLLKDAQVQEE  
BACDRYVAB\_1 OLKKFPTDEVLSNFSYTDRLVPVITNTADRTLLGLPYTALASMLMLRDIITKQPT  
180 190 200 210 220 230  
240 250 260 270 280 290  
MON863.pap WQYS--SEDAVFYRQLKLT-QQYTDHCVMYVNLGL--RGSTYDAMVFNPRREM  
BACDRYVAB\_1 WQKINFTPDADISFPTDIKNNKLYSKTYVDFQGLASYOTPSDLESAKKQYIEIM  
240 250 260 270 280 290  
300 310 320 330 340 350  
MON863.pap TLTVLDELIVLFPFDYLYEYKQ---VATLRTDITPTDPLLT---TLQKYQ-----P  
BACDRYVAB\_1 TTHCLDFAFLPTFDYLYEYKQ---VATLRTDITPTDPLLT---TLQKYQ-----P  
300 310 320 330 340 350  
360 370 380 390 400 410  
MON863.pap TFLSIENSTRKPHLDYLOGIEFHTRAKPGYFKDPSND--WS---GNVETR--PSIGSS  
BACDRYVAB\_1 NYZMCKAFPMKE--RIKQPKLYPSWRAQYQGLLOPYLWALISVQDSVETRLVQGLFV  
360 370 380 390 400 410  
420 430 440 450 460 470  
MON863.pap KITTSFY--GDKSTSPVKLSFDGQKVVRTIANTDVAANMKVLYGVTKVDFSOYDDOK  
BACDRYVAB\_1 DPOAGNYVSDISNFIQINMDTKT---PPQAGSWNTNLRGVSGLSLFRQDTR  
420 430 440 450 460 470  
480 490 500 510 520 530  
MON863.pap NISTSTOT--YDSKRNHGVSAQSDIQGLPPTTDEPLKAYSHQINABCFIMQDRRTI  
BACDRYVAB\_1 LSAGNGGFPADITYSLPATHYLSYLYGTPTQSDN---YSGHVG--ALVGVSTPQEAZL  
480 490 500 510 520 530  
540 550 560 570 580 590  
MON863.pap PPFVTHRGVDFPNTIDAKITQLPVVAYALSSQASITIEGPOFGNMLFLKESNSIA  
BACDRYVAB\_1 PNLIGQDEQONVTHG---PPFKA---SYQGVTVV--HWLQANMLKSPQ--ETG  
530 540 550 560 570 580  
590 600 610 620 630 640  
MON863.pap KFKVTLNSAALLQRYVRIRYASTNLRFL--VQNSNDPLVIYNTKNDKDDLYOTF  
BACDRYVAB\_1 IPIYNTSG---SYQICRYASNDNTNVPFVDTGAGNFIQOINASTVNDNTGVQGA  
580 590 600 610 620 630  
640 650 660 670 680 690  
MON863.pap DLATTNSNMFGSDKNELIQAESF---VSNB---KIYIDKIEFIPVQL  
BACDRYVAB\_1 NGVTVKSIATNTSFTVKIPIKTIINVLNGSSDQVLDRIEFLBSNVTFIWNYS  
640 650 660 670 680 690  
700 710 720 730 740 750  
BACDRYVAB\_1 TTGSANLIPALPMTSTSDKALTGSMTIGRTTPNSDALLRFKTYDTOTIPIPGSG  
700 710 720 730 740 750  
MON863.pap TO: BACDRYVAB\_1

LOCUS BACDRYVAA\_1  
DEFINITION Bacillus thuringiensis delta-endotoxin (CryA(a)) gene, 5' end.  
DATE 30-MAY-1995  
ACCESSION L07025  
NID g142668  
ORGANISM Bacillus thuringiensis

SCORES Initl: 59 Initn: 99 Opt: 164 z-score: 160.7 E(): 0.0017  
Smith-Waterman score: 281, 21.5% identity in 701 aa overlap

MON863.pap MANPNRSEHDITKVNSSELQTHNQYPLADNPSTLEEL--NYKEFLRMTEDEST  
BACDRYVAA\_1 MATLNELVPSVYVNLAYTPSPFLDAGTQATPADL--TAYEQLLNKLEKINAGTYSKA  
10 20 30 40 50  
60 70 80 90 100  
MON863.pap --EVLNSTVKDAVGTGISV-VGQILGVGVFPAGALTSFYOSFLNTIWPSDADP-----  
BACDRYVAA\_1 IADVLKGIPIIDTINQYVNLGLSLITLAVPEIGITPTPIGLFAALNKHDAPPPNAK  
60 70 80 90 100 110  
120 130 140 150 160  
MON863.pap --WKAFAQVEVLIDKKIEYAKSALAELOGLNFFEDYVNAISWKKTPSLRSKRQ  
BACDRYVAA\_1 DIFEMKPAIQEMIDRTLTADQOTFLNGEISGLNLAARYQSTMDIQSH--GGFNKVD  
120 130 140 150 160 170  
180 190 200 210 219  
MON863.pap ORIKELFSQA---ESFPMNSMPSFAVSKF---EVLFLTYAQAANTHLLKDAQVQEE  
BACDRYVAA\_1 OLKKFPTDEVLSNFSYTDRLVPVITNTADRTLLGLPYTALASMLMLRDIITKQPT  
180 190 200 210 220 230  
240 250 260 270 280 290  
MON863.pap WQYS--SEDAVFYRQLKLT-QQYTDHCVMYVNLGL--RGSTYDAMVFNPRREM  
BACDRYVAA\_1 WQKINFTPDADISFPTDIKNNKLYSKTYVDFQGLASYOTPSDLESAKKQYIEIM  
240 250 260 270 280 290  
300 310 320 330 340 350  
MON863.pap TLTVLDELIVLFPFDYLYEYKQ---VATLRTDITPTDPLLT---TLQKYQ-----P  
BACDRYVAA\_1 TTHCLDFAFLPTFDYLYEYKQ---VATLRTDITPTDPLLT---TLQKYQ-----P  
300 310 320 330 340 350  
360 370 380 390 400 410  
MON863.pap TFLSIENSTRKPHLDYLOGIEFHTRAKPGYFKDPSND--WS---GNVETR--PSIGSS  
BACDRYVAA\_1 NYZMCKAFPMKE--RIKQPKLYPSWRAQYQGLLOPYLWALISVQDSVETRLVQGLFV  
360 370 380 390 400 410  
420 430 440 450 460 470  
MON863.pap KITTSFY--GDKSTSPVKLSFDGQKVVRTIANTDVAANMKVLYGVTKVDFSOYDDOK  
BACDRYVAA\_1 DPOAGNYVSDISNFIQINMDTKT---PPQAGSWNTNLRGVSGLSLFRQDTR  
420 430 440 450 460 470  
480 490 500 510 520 530  
MON863.pap NISTSTOT--YDSKRNHGVSAQSDIQGLPPTTDEPLKAYSHQINABCFIMQ  
BACDRYVAA\_1 LSAGNGGFPADITYSLPATHYLSYLYGTPTQSDN---YSGHVG--ALVGVSTPQEAZL  
480 490 500 510 520 530  
540 550 560 570 580 590  
MON863.pap PPFVTHRGVDFPNTIDAKITQLPVVAYALSSQASITIEGPOFGNMLFLKESNSIA  
BACDRYVAA\_1 PNLIGQDEQONVTHG---PPFKA---SYQGVTVV--HWLQANMLKSPQ--ETG  
530 540 550 560 570 580  
590 600 610 620 630 640  
MON863.pap KFKVTLNSAALLQRYVRIRYASTNLRFL--VQNSNDPLVIYNTKNDKDDLYOTF  
BACDRYVAA\_1 IPIYNTSG---SYQICRYASNDNTNVPFVDTGAGNFIQOINASTVNDNTGVQGA  
580 590 600 610 620 630  
640 650 660 670 680 690  
MON863.pap DLATTNSNMFGSDKNELIQAESF---VSNB---KIYIDKIEFIPVQL  
BACDRYVAA\_1 NGVTVKSIATNTSFTVKIPIKTIINVLNGSSDQVLDRIEFLBSNVTFIWNYS  
640 650 660 670 680 690  
700 710 720 730 740 750  
BACDRYVAA\_1 TTGSANLIPALPMTSTSDKALTGSMTIGRTTPNSDALLRFKTYDTOTIPIPGSG  
700 710 720 730 740 750  
MON863.pap TO: BACDRYVAA\_1



Page 97 of 102





©01

msmatch(86[C-D1]). . .

SCORES Initl: 62 Inlt: 62 Opt: 96 z-score: 114.3 E(): 0.66  
Smith-Waterman score: 96, 39.0% identity in 48 aa overlap

MON863.pep 60 70 80 90 100  
SSTEVLDNKKVAVGQISVVGQILGVGVFFGALTSFYQSLNTIIV---PSDADPW  
S72313\_1 VEVGQERISQYSEDISISLQFLLSETVFGAG---FVLGLIDLINCFVGPQ---W  
110 120 130 140 150 160  
MON863.pep KAPMAQGVVLDKKEEYAKSKALNLOGHNNVEDYVHALNSWNTPLSLRSKRSGQRI  
S72313\_1 DAPLVQIQEQLISQRIEF  
80 90

MON863.pep  
TO:CBCL4B1\_6

LOCUS CBCL4B1\_6  
DEFINITION Caenorhabditis elegans comid C14B1, complete sequence,  
similar to diphtheria toxin resistance protein, cDNA  
EST yk376a7.3 comes from this gene; cDNA EST yk376a7.5  
comes from this gene  
DATE 31-JUL-1998 . .

SCORES Initl: 39 Inlt: 39 Opt: 87 z-score: 94.1 E(): 8.9  
Smith-Waterman score: 87, 23.5% identity in 226 aa overlap

MON863.pep 140 150 160 170 180 190  
LNNFEDYVNAIWSKKTPLSLRSKRSGQRIELFSQAESHFRNSMPSFA-VSKFEVL--  
CBCL4B1\_6 SHLIDCVKENFQKRLVWVSTVQFIPSLQTLRTTFKDDSSIRIDIPQCKPLSPGEVLGC  
140 150 160 170 180 190

MON863.pep 200 210 220 230 240 250  
FLPTTAAQANTHLLLLKQGVFGSENGYSSEDVAF---YRQLKLTQCTYDHCYNWYN  
CBCL4B1\_6 TSPRLDAEKYDAIVLVQGRFHLISIKHNPEIAFQVDPYER--KLTRFYDHLAKRN  
200 210 220 230 240 250

MON863.pep 260 270 280 290 300  
-VG-LNGLRGSTYDANVKFNRFRRENTLVLDLIVLFFPYDIRLYSGVK--TELTRDIF  
CBCL4B1\_6 RTGSTEIAKCTTFLQIQ-QLLQKQNKVVKEL-----EALAKRKKKFLNVLLEIF  
260 270 280 290 300

MON863.pep 310 320 330 340 350  
TDPFLLTFL---QKYGTFPLGIR-NSIRKXHLFDY---LQKEPHTLRAP-GYFG  
CBCL4B1\_6 PEKLAHFREVDCNVQVACPR-LSIDWGTQPKPLLYPFELAVLADNISVPSDHWMDYYS  
310 320 330 340 350 360

MON863.pep 360 370 380 390 400 410  
KDSFNYMSGNYVETRPISGSSKTIISPFYQDKSTEPVQKLSFDGQKVYRTIANTDVAAMP  
CBCL4B1\_6 NESLQPTWMDNEANMKREKRKPHIVVREAS  
370 380 390

MON863.pep  
TO:GNPG\_VIBCH

ID GNPG\_VIBCH STANDARD; PRT, 146 AA.  
AC P45773  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE GENERAL SECRETION PATHWAY PROTEIN G PRECURSOR (CHOLERA TOXIN SECRETION  
. . .

SCORES Initl: 60 Inlt: 60 Opt: 79 z-score: 93.8 E(): 9.2  
Smith-Waterman score: 88, 27.3% identity in 143 aa overlap

MON863.pep 250 260 270 280 290 300  
DHCVMWYVGLNGLRGSTYDANVKFNRFRRENTLVLDLIVLFFPYDIRLYSGVKVTELT  
GNPG\_VIBCH MKIMRKQGTGFTLLVAVVVVLQILASFTVVVNLONKEKADQ  
10 20 30 40

MON863.pep 310 320 330 340 350 360  
RDIPTDIPPLLTLOKYGPTFLSIENIRKPHLPDYLOQIEPHTRLRPGYPGKDSFNYWS  
GNPG\_VIBCH KAV-TOIVALENALMY-----KLDSVY-PTT-D--QGLEALVT-KPT--NPEPRNYRE  
50 60 70 80 90

MON863.pep 370 380 390 400 410 420  
GNVYETRP--IGSKTITSPFYQDKSTEPVQKLSFDGQKVYRTIANTDVAAMPNGKVYL  
GNPG\_VIBCH GQYIKRLPDLFWNDVQYLEP--GKAGTIDVTTLGADGGGQGG--TGADIGNWNIQDFQ  
100 110 120 130 140

MON863.pep 430 440 450 460 470 480  
GVTKVDFSGYDDQKNETSTQTYDSKRNGHVBAQDSIDQLPPHTTDMPLKAYSHQLNYA

\* Distributed over 1 thread.  
\* Start time: Mon Feb 12 11 29:58 2001  
\* Completion time: Mon Feb 12 11 31:01 2001  
\* CPU time used:  
\* Database scan: 0 00 03.3  
\* Post-scan processing: 0 00 28.6  
\* Total CPU time: 0 00 32.0  
\* Output file: MON863\_TOXIN



[illegible]

TREMBL\_MAIN:Q97905  
 ' 087905 Bacillus thuringiensis subsp... 247 585 525 593.0 1.8e-25  
 GP\_BCT1:BTU0819\_1  
 ' 08188 Bacillus thuringiensis subsp... 247 585 525 593.0 1.8e-25  
 TREMBL\_MAIN:Q9X682  
 ' Q9X682 Bacillus thuringiensis (subs... 215 423 514 577.2 1.4e-24  
 GP\_BCT1:AF132428\_1  
 ' AF132428 Bacillus thuringiensis sub... 215 423 514 577.2 1.4e-24  
 TREMBL\_MAIN:Q87906  
 ' 087906 Bacillus thuringiensis subsp... 220 487 472 533.1 4e-22  
 GP\_BCT1:BTU08189\_1  
 ' 08189 Bacillus thuringiensis subsp... 220 487 472 533.1 4e-22  
 GP\_BCT1:AB023293\_1  
 ' AB023293 Bacillus thuringiensis Bac... 225 390 463 524.5 2.8e-21  
 TREMBL\_MAIN:Q96597  
 ' Q96597 Bacillus thuringiensis 34kd... 125 303 463 521.5 2.8e-21  
 TREMBL\_MAIN:Q45882  
 ' Q45882 Clostridium bifementans cb... 169 316 459 519.1 2.4e-21  
 GP\_BCT1:CBM71\_1  
 ' X94146 Clostridium bifementans C.b... 169 316 459 519.1 2.4e-21  
 GP\_BCT1:CBM72\_1  
 ' X9478 Clostridium bifementans C.b... 132 381 455 514.5 4.4e-21  
 TREMBL\_MAIN:Q05102  
 ' Q05102 Clostridium bifementans cb... 132 381 455 514.5 4.4e-21  
 GP\_BCT1:BACCRY18\_1  
 ' X72253 Bacillus thuringiensis Bacil... 213 733 423 733.9 8e-19  
 TREMBL\_MAIN:Q03745  
 ' Q03745 Bacillus thuringiensis 139... 213 733 423 733.9 8e-19  
 SWISSPROT:CR71\_BACTK  
 ' P08662 Bacillus thuringiensis (subs... 240 602 416 469.8 1.4e-18  
 GP\_BCT1:BACMSQB\_1  
 ' M12662 Bacillus thuringiensis B.thu... 240 602 416 469.8 1.4e-18  
 TREMBL\_MAIN:Q45789  
 ' Q45789 Bacillus thuringiensis delt... 170 326 394 450.5 1.6e-17  
 GP\_BCT1:BTTOXDA7\_1  
 ' X16315 Bacillus thuringiensis Bacil... 170 326 394 450.5 1.6e-17  
 TREMBL\_MAIN:Q45754  
 ' Q45754 Bacillus thuringiensis delt... 202 230 303 337.7 3.1e-11  
 GP\_BCT1:BACCRYV8\_1  
 ' L07027 Bacillus thuringiensis Bacil... 202 230 303 337.7 3.1e-11  
 SWISSPROT:CR72\_BACTK  
 ' P21253 Bacillus thuringiensis (subs... 154 154 283 319.8 3e-10  
 GP\_BCT1:BACCRYB2\_1  
 ' N01738 Bacillus thuringiensis B.thu... 154 154 283 319.8 3e-10  
 GP\_BCT1:BACCRYB1\_3  
 ' M23723 Bacillus thuringiensis Bacil... 154 154 283 319.8 3e-10  
 GP\_BCT1:BACCRYB2\_1  
 ' X03724 Bacillus thuringiensis B.thu... 153 153 280 316.4 4.7e-10  
 SWISSPROT:CR72\_BACTK  
 ' P21254 Bacillus thuringiensis (subs... 153 153 280 316.4 4.7e-10  
 GP\_BCT1:BTXCRY28\_1  
 ' X55416 Bacillus thuringiensis B th... 153 153 280 316.4 4.7e-10  
 TREMBL\_MAIN:Q02764  
 ' Q02764 Bacillus thuringiensis inae... 154 154 278 314.2 6.3e-10  
 GP\_BCT1:BCRCRY1\_1  
 ' AF047038 Bacillus thuringiensis Bac... 154 154 278 314.2 6.3e-10  
 TREMBL\_MAIN:Q9RM89  
 ' Q9RM89 Bacillus thuringiensis cry2... 154 154 277 314.0 6.4e-10  
 GP\_BCT1:BTNH132465\_1  
 ' A0132463 Bacillus thuringiensis Bac... 154 154 277 314.0 6.4e-10  
 TREMBL\_MAIN:Q966M4  
 ' Q966M4 Bacillus thuringiensis cry2... 154 154 277 313.0 7.3e-10  
 GP\_BCT1:BTNH132464\_1  
 ' A0132464 Bacillus thuringiensis Bac... 154 154 277 313.0 7.3e-10  
 GP\_BCT1:BTXCRY1IC\_3  
 ' X57252 Bacillus thuringiensis B.thu... 140 140 276 312.0 8.3e-10  
 TREMBL\_MAIN:Q45743  
 ' Q45743 Bacillus thuringiensis cry1... 140 140 276 312.0 8.3e-10  
 GP\_BCT1:AF200816\_1  
 ' AF200816 Bacillus thuringiensis Bac... 155 155 276 311.9 8.4e-10  
 TREMBL\_MAIN:Q9RMG3  
 ' Q9RMG3 Bacillus thuringiensis crya... 155 155 276 311.9 8.4e-10  
 GP\_BCT1:BTNH132465\_1  
 ' A0132465 Bacillus thuringiensis Bac... 154 154 275 310.8 9.7e-10  
 TREMBL\_MAIN:Q966N5  
 ' Q966N5 Bacillus thuringiensis cry2... 154 154 275 310.8 9.7e-10  
 GP\_BCT1:BACCRYV11A\_1  
 ' L01393 Bacillus thuringiensis Bacil... 255 255 255 305.6 1.9e-09  
 TREMBL\_MAIN:Q81518  
 ' 81518 Bacillus thuringiensis 70 k... 255 255 255 305.6 1.9e-09  
 GP\_BCT1:BCRCRYVPL\_2  
 ' X99049 Pannabacillus popilliae B.po... 141 166 250 281.7 4e-08  
 TREMBL\_MAIN:Q45358  
 ' Q45358 paenibacillus popilliae par... 141 166 250 281.7 4e-08

GP\_BCT1:BTTOXD\_1  
 ' X05692 Bacillus thuringiensis B. th... 103 180 237 278.0 6.5e-08  
 TREMBL\_MAIN:Q45712  
 ' Q45712 Bacillus thuringiensis delt... 100 136 211 233.8 1.9e-05  
 GP\_BCT1:BTU19725\_1  
 ' U19725 Bacillus thuringiensis Bacil... 100 136 211 233.8 1.9e-05  
 TREMBL\_MAIN:Q9RNE6  
 ' Q9RNE6 Bacillus thuringiensis endo... 88 88 186 220.5 0.0001  
 GP\_BCT1:AF182196\_1  
 ' AF182196 Bacillus thuringiensis Bac... 88 88 186 220.5 0.0001  
 GP\_BCT1:BACCRYVAB\_1  
 ' L07028 Bacillus thuringiensis Bacil... 67 152 164 180.4 0.018  
 TREMBL\_MAIN:Q45753  
 ' Q45753 Bacillus thuringiensis delt... 67 152 164 180.4 0.018  
 GP\_BCT1:BACCRYVAA\_1  
 ' L07029 Bacillus thuringiensis Bacil... 59 99 164 179.9 0.019  
 TREMBL\_MAIN:Q45760  
 ' Q45760 Bacillus thuringiensis delt... 59 99 164 179.9 0.019  
 TREMBL\_MAIN:Q45710  
 ' Q45710 Bacillus thuringiensis delt... 65 102 155 170.8 0.061  
 GP\_BCT1:BTU13955\_1  
 ' U13955 Bacillus thuringiensis Bacil... 65 102 155 170.8 0.061  
 TREMBL\_MAIN:Q00705  
 ' Q00705 Bacillus thuringiensis d-en... 81 81 149 167.1 0.098  
 GP\_BCT1:AF017416\_1  
 ' AF017416 Bacillus thuringiensis Bac... 81 81 149 167.1 0.098  
 TREMBL\_MAIN:Q45730  
 ' Q45730 Bacillus thuringiensis moeq... 115 115 140 157.2 0.35  
 GP\_BCT1:BTXCRY1B\_1  
 ' X06940 Bacillus thuringiensis B.thu... 115 115 140 157.2 0.35  
 GP\_BCT1:BCRCRYM0\_4  
 ' M17373 Bacillus thuringiensis B.thu... 80 136 130 146.7 1.3  
 SWISSPROT:CR72\_BACTK  
 ' M21256 Bacillus thuringiensis (subs... 80 136 130 146.7 1.3  
 TREMBL\_MAIN:Q45755  
 ' Q45755 Bacillus thuringiensis delt... 58 110 122 136.1 5.2  
 GP\_BCT1:BACCRYV6\_2  
 ' L07023 Bacillus thuringiensis Bacil... 58 110 122 136.1 5.2  
 TREMBL\_MAIN:Q9V861  
 ' Q9V861 Chrocephala melanogaster (acu... 56 56 108 132.5 8.3  
 \\End of List

Registration of Monsanto Company