MEAT ALLERGY AND THE ALLERGENIC COMPONENTS: UNDERLINING REASONS FOR THE ABSENCE OF CLINICAL PRESENTATION TO MEAT ANTIGENS DESPITE THE PRESENCE OF HIGH LEVELS OF SPECIFIC IGE

WONG KANG NING
(B. Sc. (Honours), NUS)

A THESIS SUBMITTED FOR
THE DEGREE OF MASTER OF SCIENCE
DEPARTMENT OF BIOLOGICAL SCIENCES
NATIONAL UNIVERSITY OF SINGAPORE
2006

Acknowledgments

My sincere gratitude to my supervisor, Dr Chew Fook Tim, for his advice and guidance. His constant ideas, understanding and support throughout the entire programme, and his invaluable contributions in the writing of this thesis were greatly appreciated.

My thanks to Dr Ong Tan Ching, Dr Shang Huishan and Dr Bi Xuezhi for their constructive ideas and kindly care on the immunoassays, molecular and proteomics aspects of this work.

A special mention of thanks to Ms Lim Yun Peng and Dr Li Kuo Bin for their technical assistance in handling the massive number of sequences during the allergenicity prediction.

Lastly, I would like to thank all the friends and colleagues in the Functional Genomics Laboratory Lab 1 and 3 for their care and help.

CONTENTS

Acknowledgments **i**

Table of Contents **ii**

List of Appendices **xii**

Summary **xiv**

List of Tables **xvi**

List of Figures **xix**

Abbreviations **27**

Abstract **29**

CHAPTER 1: INTRODUCTION

1.1 Allergy 1

1.1.1 Basic concepts of allergy 1

1.1.2 Hypersensitivity 1

1.1.3 Mechanism of Allergy – Type I (immediate) hypersensitivity 2

1.2 Food allergy 4

1.2.1 Food allergens 5

1.3 Meat allergy 7

1.3.1 Meat-based allergens 8

1.4 Trends in meat based allergies 11

1.5 Objectives 12

CHAPTER 2: DOT IMMUNOARRAY SYSTEM FOR DETECTION OF ALLERGEN-SPECIFIC IGES

2.1 INTRODUCTION 14

2.1.1 Techniques in allergy diagnosis 14

2.1.2 Advantages of *in vitro* techniques 15

2.1.3 Limitations of *in vitro* techniques 17

2.2 MATERIALS AND METHODS 18

2.2.1 Patients and sera 18

2.2.2 Skin Prick tests (SPTs) 18

2.2.3 Dotting apparatus 19

2.2.4 Support materials and washing buffers 20

2.2.5 Allergen extracts 20

2.2.6 Allergen immunoarray for the detection of specific IgE 21

2.2.7 Image analysis of immunoarray blots 22

2.2.8 Allergen immunoarray validation 23

2.2.9 Statistical analysis 24

2.3 RESULTS AND DISCUSSIONS 25

2.3.1 Skin prick test 25

2.3.2 Allergen immunoarray 25

2.3.2.1 Prevalence of meat-based allergy 25

2.3.2.2 IgE responses to pork among individuals (Malay Muslims) who do not consume pork 27

2.3.3 Allergen immunoarray validation 29

2.3.3.1 Performance of allergen immunoarray in terms of duplicates 29

2.3.3.2 Immunoarray *vs* ELISA 32

2.3.3.3 Self inhibition 33

2.3.4 Cross-reactivity 34

2.3.4.1 Prediction of pattern and potential for cross-reactivity 34

2.3.4.2 Validation of cross-reactivity via cross inhibition ELISA 36

2.4 CONCLUSION 38

CHAPTER 3: ALLERGEN PREDICTION USING A BIOINFORMATIC APPROACH

3.1 INTRODUCTION 40

3.1.1 Establishment of food safety guidelines 40

3.1.2 Allergen databases 41

3.1.3 Allergenicity prediction 43

3.1.4 Limitation of bioinformatics allergen prediction 45

3.1.5 Expressed Sequence Tagging in genome studies 45

3.1.6 Unigenes 47

3.2 MATERIALS AND METHODS 48

3.2.1 Data mining and content 48

3.2.2 Analysis of Sequence Similarity (Method 1) 48

3.2.3 Allergenicity prediction using wavelet transform (Method 2) 50

3.2.4 Cataloging of BLAST output into functional categories 50

3.3 RESULTS AND DISCUSSIONS 51

3.3.1 Allergen prediction based on sequence homology 51

3.3.1.1 Matched allergen profiles of the seven animal species 53

3.3.1.2 Performance of allergen prediction by sequence homology 64

3.3.2 Allergenicity prediction using wavelet transform 64

3.3.2.1 Performance of allergenicity prediction using wavelet transform 68

3.3.3 Comparison between sequence homology based and motif-based allergen prediction system 68

3.4 CONCLUSION 76

CHAPTER 4: IDENTIFICATION AND CHARACTERIZATION OF MEAT-BASED ALLERGENS USING A PROTEOMIC APPROACH

4.1 INTRODUCTION 77

4.1.1 Protein extraction from food sources 77

4.1.2 Methods used for protein separation and allergen isolation 78

4.1.3 Protein identification using Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) 79

4.2 MATERIALS AND METHODS 82

4.2.1 Patients and sera 82

4.2.2 Protein extraction 82

4.2.3 Gel Electrophoresis 82

4.2.4 Protein visualization and image analysis 84

4.2.5 Western Blotting Analysis 85

4.2.6 In-gel digestion of protein bands and spots 85

4.2.7 MALDI TOF/TOF MS/MS analysis 86

4.3 RESULTS AND DISCUSSIONS 87

4.3.1 1-Dimensional SDS-PAGE and immunoblots 87

4.3.2 2-Dimensional SDS-PAGE and immunoblots 91

4.3.3 Protein identification by MALDI-TOF_TOF 98

mass spectrometry

4.3.4 Comparison between bioinformatics and proteomics approach for allergen prediction and/or identification 105

4.4 CONCLUSION 107

CHAPTER 5: MOLECULAR CLONING AND IMMUNOGLOBULIN E (IGE) REACTIVITY OF PUTATIVE MEAT-BASED ALLERGENS

5.1 INTRODUCTION 108

5.1.1 Usage of recombinant allergens for research and diagnosis 108

5.1.2 Criteria for the production and characterization of recombinant allergens for clinical applications 109

5.2 MATERIALS AND METHODS 112

5.2.1 Bacterial strains 112

5.2.2 mRNA extraction 112

5.2.3 Molecular cloning of recombinant allergens 112

5.2.3.1 Bioinformatics analysis 112

5.2.3.2 RT-PCR to isolate full length clones of putative meat allergens 114

5.2.3.3 Cloning of PCR product into pGEMT®-Easy Vector 115

5.2.3.4 Transformation of *E. coli* strain XL1-Blue 115

5.2.3.5 Colony Screening 116

5.2.3.6 Culture of *E. coli* and Plasmid Extraction 116

5.2.3.7 Ligation Independent Cloning (LIC) of putative allergens into pET32a (+) expression vector 117

5.2.4 DNA sequencing 119

5.2.4.1 Automated sequencing 119

5.2.4.2 Purification of Automated sequencing products 119

5.2.4.3 Automated DNA sequencing analysis 120

5.2.5 Expression and purification of recombinant allergens 120

5.2.5.1 Sample induction and expression 120

5.2.5.2 Affinity purification of recombinant protein with pET-32a (+) His-Tag system 121

5.2.6 Recombinant proteins immunoarray 121

5.2.6.1 Patients and sera 121

5.2.6.2 Immunoarray 122

5.3 RESULTS AND DISCUSSIONS 122

5.3.1 Characterization of recombinant proteins 122

5.3.1.1 General Bioinformatics analysis of putative allergen sequences 122

5.3.1.2 Tropomyosins 123

5.3.1.3 Troponin 129
5.3.1.4 Myosin-light chain 133
5.3.1.5 Aldehyde dehydrogenase 135
5.3.1.6 Enolase 139
5.3.1.7 Heat shock proteins 142
5.3.2 Expression and purification of recombinant allergens 152
5.3.3 Recombinant proteins immunoarray 154
5.3.3.1 Prevalence of IgE-binding of crude and recombinant proteins 155
5.4 **CONCLUSION** **157**

CHAPTER 6: INVESTIGATIONS ON THE CROSS-REACTIVE CARBOHYDRATE DETERMINANTS (CCD) OF MEAT-BASED ALLERGENS

6.1 INTRODUCTION **158**
6.2 MATERIALS AND METHODS **160**
6.2.1 Patients and sera 160
6.2.2 Protein extraction 160
6.2.3 Enzymatic deglycosylation procedures 161
6.2.4 Immunoassays 161
6.2.4.1 Western blot analysis 161

6.2.4.2 Enzyme-linked immunosorbent assay (ELISA) 162

6.3 RESULTS AND DISCUSSIONS 162

6.3.1 Deglycosylation experiments 162

6.3.2 Immunoassays 164

6.3.2.1 Western blot analysis 164

6.3.2.2 ELISA 165

6.4 CONCLUSION 169

CHAPTER 7: BLOCKING IMMUNOGLOBULIN G (IGG) ANTIBODIES IN MEAT ALLERGY

7.1 INTRODUCTION 170

7.1.1 Specific immunotherapy (SIT) 170

7.1.2 Concept of blocking IgG antibodies 171

7.2 MATERIALS AND METHODS 172

7.2.1 Patients and sera 172

7.2.2 Allergen immunoarray for the detection of specific IgG 172

7.2.3 Plasma preparation 173

7.2.4 Isolation of peripheral blood mononuclear cells (PBMCs) 173

7.2.5 Immunoaffinity depletion of IgG from plasma 174

7.2.6 Preparation of meat antigens 175

7.2.7 Histamine-release assay 175

7.3 RESULTS AND DISCUSSIONS 177

7.3.1 Allergen immunoarray for the detection of specific IgG 177

7.3.2 Immunoaffinity depletion of IgG from plasma 179

7.2.3 Histamine-release assay 181

7.4 CONCLUSION 184

BIBLIOGRAPHY AND REFERENCES 186

Appendix I. Allergens dotted onto the array 206

Appendix II. Detail lists of sequence homology matches for beef with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology. 208

Appendix III. Detail lists of sequence homology matches for pork with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology. 212

Appendix IV. Detail lists of sequence homology matches for chicken with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology. 216

Appendix V. Detail lists of sequence homology matches for trout with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology. 219

Appendix VI. Detail lists of sequence homology matches for sheep with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology. 223

Appendix VII. Detail lists of sequence homology matches for goat with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology. 225

Appendix VIII. Detail lists of sequence homology matches for dog with query, subject, subject description, functional category, bit score, E-value and region of amino acid. 226

Appendix IX. Detail lists of sequence homology matches for cat with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology. 228

Appendix X. List of putative allergens predicted in pork using wavelet transform**.** 229

Appendix XI. List of putative allergens predicted in chicken using wavelet transform. 230

Appendix XII. List of putative allergens predicted in sheep using wavelet transform. 231

Appendix XIII. List of putative allergens predicted in dog using wavelet transform. 231

Appendix XIV. List of putative allergens predicted in cat using wavelet transform. 231

Appendix XV. Molecular cloning of Tropo 3. 232

Appendix XVI. Molecular cloning of TRNT. 233

Appendix XVII. Molecular cloning of Myo_L. 234

Appendix XVIII. Molecular cloning of ADH. 235

Appendix XIX. Molecular cloning of ENO 1. 236

Appendix XX. Molecular cloning of pHSP70. 237

Appendix XXI. Molecular cloning of bHSP70. 238

Appendix XXII. Molecular cloning of bHSP90. 239

Summary

This study aimed to identify and characterize meat-based allergens and also to elucidate the underlining reasons for the observed paradox of high abundance of IgE-binding to meats antigens in sera of allergic patients but no clinical presentation to these antigens.

Our study based on an dot-blot immunoarray showed that the frequency of IgE binding to 3 commonly consumed meat is especially high in 1096 allergic patients's sera [pork 46% (504/1096), beef 39% (428/1096), mutton 37% (403/1096)]. Cross-inhibition ELISA showed that these meats are cross-reactive. In order to Identify and characterize the meat-based allergens, a dual bioinformatics and proteomics approach was employed.

For the bioinformatics approach, allergenicity prediction was achieved by subjecting Unigenes sequences from cow, pig, chicken, trout, goat, sheep, cat and dog to both BLASTx algorithm and motif-based prediction. Many significant hits were found and many of these putative allergens (namely heat shock proteins, tropomyosins, aldehyde dehydrogenases, enolases and albumins) were similar across the species. The similarities seem to imply that there is a potential for cross-reactivity among these animal species. Additionally, nine of these putative allergens from cow and pig were cloned and expressed as recombinant proteins. However, they showed weak IgE-binding using patients' sera on the immunoarray. This could be attributed to the lack of post-translational modifications or incorrect folding of the protein.

The proteomics approach involved separation of protein extracts from cow, pig and goat by both 1D and 2D electrophoresis followed by immunoblotting using sera from meat-allergic patients. IgE-binding protein spots were excised and analyzed by MALDI-

TOF-TOF mass spectrometry. A total of 58 spots were identified and many of which were similar to those predicted as putative allergens in the bioinformatics approach.

Despite presence of high levels of meat specific IgEs, only 2 out of 18 patients tested via SPT were beef-positive. This indicates that the high levels of IgE may not have clinical relevance as they are unable to elicit *in vivo* histamine release. We hypothesized that the lack of clinical relevance was due to unspecific IgEs binding to CCDs in meat sources and/or *in vivo* IgG blocking of histamine release resulting in negative SPTs. In the CCD study, the crude meat extracts from beef and pork were deglycosylated and IgE-binding reactivity was validated by ELISA and immunoblots. Indeed, there was significant reduction in IgE-binding in deglycosylated samples suggesting that majority of the IgEs were binding to carbohydrate moieties. In the IgG blocking study, 25 patients with high IgE-binding to meats were shown to have significantly higher levels of meat specific IgG on the immunoarray. PBMCs, from two patients with both high IgE and IgG to meats, co-incubated with plasma (IgG depleted) and meat extracts were able to elicit histamine release which was not seen in the non-depleted IgG plasma suggesting the presence of blocking IgG inhibit histamine release.

In conclusion, the high IgE-binding to meat extracts is mainly due to presence of mammalian cross-reactive carbohydrate determinants (CCDs). Negative SPT is due to presence of “blocking” IgG antibodies which inhibits histamine release.

List of Tables

Chapter 1:

Table 1	**Cross-reactivity between food proteins and clinical cross-reactivity among members of plant and animal species** (adapted from Krishna *et al*., 2001).	6
Table 2	**Known allergen from *Bos taurus* listed on WHO/IUIS nomenclature system**	9

Chapter 2:

Table 1	**Intra-membrane and inter-membrane concordances**	30
Table 2	**Validation results between the immunoarray method versus the ELISA system.**	32

Chapter 3:

Table 1	**Major allergen-related data sources**	42
Table 2	**No. of putative allergens predicted for each species of animal based on sequence homology.**	52
Table 3	**Unigenes of pig, cow, chicken, goat, sheep, dog, and cat found to be significantly homologous to allergens from various organisms.** Ticks indicate the presence of allergen-homologous unigenes (not named) within the animal species.	54
Table 4	**Summarized list allergen homologues from the seven species of animals**	58
Table 5	**Allergen motifs.** The protein families were identified by using hmm search to search the Swiss-Prot with a profile HMM generated from the corresponding allergen motif.	65
Table 6	**No. of putative allergens predicted for each species of animal using wavelet transform allergen prediction system.**	66
Table 7	**An example of the list of putative allergens predicted in beef using wavelet transform**	67

Table 8	**Comparison of predicted putative allergens in beef by both bioinformatics systems**	71
Table 9	**Comparison of predicted putative allergens in pork by both bioinformatics systems**	72
Table 10	**Comparison of predicted putative allergens in chicken by both bioinformatics systems**	73
Table 11	**Comparison of predicted putative allergens in sheep by both bioinformatics systems**	74
Table 12	**Comparison of predicted putative allergens in dog by both bioinformatics systems**	75
Table 13	**Comparison of predicted putative allergens in cat by both bioinformatics systems**	75

Chapter 4:

Table 1	**Identification of proteins fro 2-DE of *S. scrofa* (Pig) after in-gel trypsin digestion by MALDI-TOF-TOF and NCBI database searching.** Missing spots were due to poor spectra, no significant matches, or keratin contaminations.	102
Table 2	**Identification of proteins fro 2-DE of *B. taurus* (cow) after in-gel trypsin digestion by MALDI-TOF-TOF and NCBI database searching.** Missing spots were due to poor spectra, no significant matches, or keratin contaminations.	103
Table 3	**Identification of proteins fro 2-DE of *O. aries* (goat) after in-gel trypsin digestion by MALDI-TOF-TOF and NCBI database searching.** Missing spots were due to poor spectra, no significant matches, or keratin contaminations.	104

Chapter 5:

Table 1	**List of putative allergens to be cloned**	113
Table 2	**List of specific forward and reverse conserved primers used for PCR amplification of desire gene**	114
Table 3	**List of universal primers used for colony screening**	116
Table 4	**List of Ek-LIC forward and reverse primers**	118

Table 5 **Detailed bioinformatics analyses of the putative allergens** 123

Table 6 **Estimated molecular weight of the expressed allergen with the fusion protein.** 153

List of Figures

Chapter 2:

Figure 1	**Images of the dotting apparatus (A) and the membrane dotted with allergens (B)**	19
Figure 2	**Process of image analysis of the immunoarray blots**	23
Figure 3	**Prevalence of allergy to meat and other animal products.** The cut-offs for low, med and high reactions are at 2SD, 4SD and 8SD respectively.	26
Figure 4	**Percentage of individuals possessing pork-specific IgE in various races**.	28
Figure 5	**Venn diagrams showing number and percentages of individuals possessing pork-specific and/or beef-specific IgE.** (A) Malay Muslims and (B) other races.	29
Figure 6	**Examples of intra-membrane and inter-membrane concordance bi-plots**	31
Figure 7	**Correlation of the ELISA versus immunoarray system.** Correlation coefficient, r was analyzed using Spearman's Correlation Test. p values: p = 0.05*.	33
Figure 8	**Graph showing self inhibition for pork (A), beef (B) and lamb (C).** The sera were selected based on positivity on both immunoarray and ELISA.	34
Figure 9	**Dendrogram showing relationships between the allergens used** (Done courtesy of Ms Mavis Low)**.**	35
Figure 10	**Correlation bi-plots between pork, beef and lamb.** Correlation coefficient, r was analyzed using Spearman's Correlation Test. p values: p = 0.01**.	36
Figure 11	**Graph showing percentage inhibition against amount of protein (micrograms) inhibitors.** Sera from three patients were inhibited with beef, mutton, pork, chicken, and rabbit protein. The ELISA plate was coated with pork protein.	37
Figure 12	**Graph showing percentage inhibition against amount of protein (micrograms) inhibitors.** The serum from P1 was inhibited with beef, mutton, pork, chicken, and rabbit protein.	38

The ELISA plate was coated with beef protein (A) and lamb (B).

Chapter 3:

Figure 1 **Flowchart of the entire prediction system based on sequence homology** 49

Figure 2 **Relationship between the total numbers of Unigene, nucleotide or protein sequences used for each species (bars) and the percentage of these sequences that match allergens (lines).** 52

Figure 3A **Pie chart of the allergen homologues from pig classified based on their biological function.** 59

Figure 3B **Pie chart of the allergen homologues from cow classified based on their biological function.** 60

Figure 3C **Pie chart of the allergen homologues from chicken classified based on their biological function.** 61

Figure 3D **Pie chart of the allergen homologues from goat classified based on their biological function.** 62

Figure 3E **Pie chart of the allergen homologues from sheep classified based on their biological function** 62

Figure 3F **Pie chart of the allergen homologues from cat classified based on their biological function.** 63

Figure 3G **Pie chart of the allergen homologues from dog classified based on their biological function.** 63

Figure 4 **Venn diagrams of putative allergens being predicted in both allergencity prediction systems for each animal species** 70

Chapter 4:

Figure 1 **Principle of matrix-assisted laser desorption/ionization mass spectrometry**. The analyte mixed with a saturated matrix solution forms crystals. The irradiation of this mixture by the laser induces the ionization of the matrix, desorption, transfer of protons from photo-excited matrix to analyte to form a protonated molecule (adapted from Marvin *et al*., 2003). 80

Figure 2	**Process of spots matching using the Bio-rad PDQuest software**	84
Figure 3	**1-D SDS-PAGE and immunoblotting analysis of proteins from *S. scrofa* (pig) extract**. Total protein: Coomassie stain for total protein analysis; Lane 1 (M): marker (kDa); Lane 2 (cM): commercial pork skin prick extract; Lane 3 (I): pig intestine PBS extract; Lane 4 (K): pig kidney PBS extract; Lane 5 (M): pork PBS extract. Immunoblots (1 – 8): 8 patients; Immunoblot 9: control subject; Immunoblot 10: blank control (secondary antibody only).	88
Figure 4	**1-D SDS-PAGE and immunoblotting analysis of proteins from *B. taurus* (cow) extract**. Lane P: Coomassie stain for total protein analysis; Lane 1 – 10: immunoblots with 10 patients' sera; Lane 11 and 12: immunoblots with 2 control subjects' sera; Lane 13 and 14: Blank controls (secondary antibody only).	89
Figure 5	**1-D SDS-PAGE and immunoblotting analysis of proteins from *O. aries* (goat) extract**. Lane P: Coomassie stain for total protein analysis; Lane 1 – 10: immunoblots with 10 patients' sera; Lane 11 and 12: immunoblots with 2 control subjects' sera; Lane 13 and 14: Blank controls (secondary antibody only).	90
Figure 6	**2-DE separation of *S. scrofa* (pig) proteins**. *S. scrofa* meat (pork) was extracted with TCA/acetone and dissolved in urea sample buffer before 2-D PAGE. First dimension: pH 3 – 10 NL; second dimension: 12% SDS-PAGE gel. Protein spots were visualized by Coomaisse blue staining. Isoelectric points and molecular weight (kDa) are indicated at the top and on the left side, respectively. An arrow with numeral indicates an IgE-binding spot identified by MALDI-TOF-TOF mass spectrometry.	92
Figure 7	**2-DE separation of *B. taurus* (cow) proteins.** *B. taurus* meat (beef) was extracted with TCA/acetone and dissolved in urea sample buffer before 2-D PAGE. First dimension: pH 3 – 10 NL; second dimension: 12% SDS-PAGE gel. Protein spots were visualized by Coomaisse blue staining. Isoelectric points and molecular weight (kDa) are indicated at the top and on the left side, respectively. An arrow with numeral indicates an IgE-binding spot identified by MALDI-TOF-TOF mass spectrometry.	93

Figure 8	**2-DE separation of *O. aries* (goat) proteins**. *O. aries* meat (mutton) was extracted with TCA/acetone and dissolved in urea sample buffer before 2-D PAGE. First dimension: pH 3 – 10 NL; second dimension: 12% SDS-PAGE gel. Protein spots were visualized by Coomaisse blue staining. Isoelectric points and molecular weight (kDa) are indicated at the top and on the left side, respectively. An arrow with numeral indicates an IgE-binding spot identified by MALDI-TOF-TOF mass spectrometry	94
Figure 9	**2-DE immnoblots of *S. scrofa* (pig) proteins**. A blotting membrane was probed with serum IgE from patients (A – F) and from control subject as negative control (G) Blank control (H) is probed with secondary antibody only.	95
Figure 10	**2-DE immnoblots of *B. taurus* (cow) proteins**. A blotting membrane was probed with serum IgE from patients (A – D) and from control subject as negative control (E) Blank control (F) is probed with secondary antibody only.	96
Figure 11	**2-DE immnoblots of *O. aries* (goat) proteins**. A blotting membrane was probed with serum IgE from patients (A – C) and from control subject as negative control (D) Blank control (E) is probed with secondary antibody only.	97
Figure 12	**Three-dimensional homology modeling of allergen Gal d 3 (ovotransferrin precursor-conalbumin) and other transferrins from (A) pig and (B) cow**. They show very high sequence and structural homology thus are candidate putative allergens.	98
Figure 13	**Venn diagram showing the comparison between bioinformatics and proteomics approach for allergen prediction and/or identification**	106

Chapter 5:

Figure 1	**Nucleotide and deduced amino acid sequence of Tropo 1.** The predicted initiation Met start and stop codon (TAA) is in red. Highlighted in yellow are the forward and reverse primer sequences. Highlighted in green are likely regions of tropomyosin IgE-binding epitopes based on previously known epitopes. Underlined is the tropomyosin signature at amino acid position 232 – 240.	125

Figure 2 **Nucleotide and deduced amino acid sequence of Tropo 3.** The predicted initiation Met start and stop codon (TAG) is in red. Highlighted in yellow are the forward and reverse primer sequences. Highlighted in green are likely regions of tropomyosin IgE-binding epitopes based on previously known epitopes. Underlined is the tropomyosin signature at amino acid position 196 – 204. 126

Figure 3 **Multiple sequence alignments between Tropo 1 and Tropo 3.** Amino acid with 100% identity colored in black and more than 50% homology colored in blue. Dots have been introduced to maximize the alignments. 127

Figure 4 **Molecular cloning of Tropo 1. (A) RNA extraction**: Agarose (1%) gel showing the total RNA extraction of meat from *Sus scrofa* using Trizol reagent in Lane 1. Distinct double bands were observed indicating integrity of the 28s and 18s RNA, however, there was an accumulation of 5S RNA. Nevertheless, the RNA was used for cDNA synthesis. **(B) PCR amplification of target tropomyosin gene** with gene specific primers using Expand long-template *Taq* DNA polymerase. Amplicons in Lane 1 and 2 corresponds to correct expected size of ~900 bp. PCR amplicons from both lanes were extracted and purified using QIA quick Gel extraction Kit (Qiagen) and ligated to pGEMT-Easy vector followed by transformation into XL-1 blue non-expression host. **(C) Colony screening of PCR inserts in pGEMT vector using SP6 and T7 primers:** A total of 10 colonies were screened for insert. Only five lanes were showed here (Lane: 1 to 5). Only 5 out of 10 clones showed the presence of insert with expected size of ~900 bp. **(D): PCR amplification of target gene from pGEM-T plasmids with correct insert using designed LIC primer adaptors.** Purified pGEM-T plasmid from clone 1 (Lane 1 of Fig C) was used. The PCR amplified band was gel extracted, and purified. The final digested DNA fragment was ligated into pET-32a (+) expression vector (Novagen, USA) using T4 DNA ligase (Invitrogen, USA) and transformed into XL1-Blue non-expression host cell. **(E): Colony screening of pET32a ligated insert** in transformed XL 1-blue non-expression host strain using LIC primers. Lane 1 to 5 corresponds to 5 clones chosen with the correct size of insert. **(F): Sub cloning of ligated Pet Vector plasmid into BL-21 (DE3) (Novagen, USA) expression host**. Again, colony screening was performed (lane 1 to 5). The clones were subsequently sequenced from both ends to check for correct reading frame. Clone 2 and Clone 4 were selected for protein expression. Glycerol stocks were made from those clones that were used for expression. 128

Figure 5	**Nucleotide and deduced amino acid sequence of TRNT.** The predicted initiation Met start and stop codon (TAG) is in red. Highlighted in yellow are the forward and reverse primer sequences. Highlighted in red is the predicted N-glycosylation site	130
Figure 6	**Multiple sequence alignments between TRNT, Bla g 6, Bet v 3, Bos d 3 and Ole e 3.** Amino acid with 100% identity colored in black, 75% homology colored in pink and 50% homology colored in blue**...** Dots have been introduced to maximize the alignments	132
Figure 7	**Nucleotide and deduced amino acid sequence of Myo_L.** The predicted initiation Met start and stop codon (TAG) is in red. Highlighted in yellow are the forward and reverse primer sequences. Highlighted in red are the predicted N-glycosylation sites.	134
Figure 8	**Nucleotide and deduced amino acid sequence of ADH.** The predicted initiation Met start and stop codon (TAG) is in red. Highlighted in yellow are the forward and reverse primer sequences. Highlighted in red is the predicted N-glycosylation site. Underlined are the conserved glutamic acid site and cysteine site which are located at amino acid positions 268 - 275 and 296 - 307 respectively.	135
Figure 9	**Multiple sequence alignments between ADH, Alt a 10 and Cla h 3.** Amino acid with 100% identity colored in black and more than 50% homology colored in blue. Dots have been introduced to maximize the alignments.	138
Figure 10	**Nucleotide and deduced amino acid sequence of ENO 1.** The predicted initiation Met start and stop codon (TAG) is in red. Highlighted in yellow are the forward and reverse primer sequences. Highlighted in red are the predicted N-glycosylation sites. Underlined is the enolase signature at amino acid positions 340 – 353.	140
Figure 11	**Nucleotide and deduced amino acid sequence of pHSP70.** The predicted initiation Met start and stop codon (TAG) is in red. Highlighted in yellow are the forward and reverse primer sequences. Highlighted in red are the predicted N-glycosylation sites. Underlined are the three heat shock hsp70 proteins family signatures at amino acid positions 9 – 16, 197 – 210, and 334 – 348.	143

Figure 12	**Nucleotide and deduced amino acid sequence of bHSP70.** The predicted initiation Met start and stop codon (TAG) is in red. Highlighted in yellow are the forward and reverse primer sequences. Highlighted in red are the predicted N-glycosylation sites. Underlined are the three heat shock hsp70 proteins family signatures at amino acid positions 9 – 16, 197 – 210, and 334 – 348.	146
Figure 13	**Nucleotide and deduced amino acid sequence of HSP90.** The predicted initiation Met start and stop codon (TAG) is in red. Highlighted in yellow are the forward and reverse primer sequences. Highlighted in red are the predicted N-glycosylation sites. Underlined is the heat shock hsp90 proteins family signature at amino acid positions 38 – 47.	149
Figure 14	**SDS-PAGE analyses of expression (A) and purification (B) of the recombinant allergens.**	153
Figure 15	**Recombinant proteins immunoarray plate format and immunodot blots.** (A) The plate format of the recombinant immunoarray where highlighted in blue, yellow and green are the controls, crude proteins and recombinant proteins respectively. (B) An example of the immunodot blots labeled +ve sera, -ve sera and blank (secondary antibody only). Dots in purple indicate positive IgE-binding.	155
Figure 16	**Percentage of IgE-binding frequency of crude and recombinant proteins**	156
Chapter 6:		
Figure 1	**Protein separation of non-deglycosylated and deglycosylated pork and beef crude extracts.** M: marker; ND: non-deglycosylated; N: deglycosylated. Indicated in arrows are bands or regions with noticeable changes before and after deglycosylation.	163
Figure 2	**Immunodetection after SDS-PAGE of pork (A) and beef (B) extracts (non-deglycosylated and deglycosylated) with patients' sera.** M: marker; ND: non-deglycosylated; N: deglycosylated. In both blots, membranes 1 – 7 were IgE immunoblots using 7 patients' sera. Membranes 8 were negative serum and membranes 9 were blanks with secondary antibody only.	165

Figure 3 **ELISA validation of IgE-binding for non-deglycosylated and deglycosylated pork extract with patients' sera.** 166

Figure 4 **Correlation bi-plots between meat antigens (pig and cow) and CCDs markers (bromelain and horseradish peroxidase).** Spearman's Correlation Test showed no correlation. 168

Chapter 7:

Figure 1 **Typical diagram showing Ficoll- Hypaque separation of whole blood sample.** Samples are separated based on density after centrifugation 174

Figure 2 **Principle of the histamine enzyme immunoassay (EIA) kit.** 176

Figure 3 **Comparison between levels of meat specific IgEs and IgGs.** A total of 25 sera were screened for meat specific IgG antibodies. (Conversion for IgE is 1 IU = 2.4ng/protein; for IgG is 1 IU = 0.8147 mg). The difference in magnitude between IgG and IgE is more than 10000 times. 178

Figure 4 **SDS-PAGE of plasma and IgG depleted plasma from two patients.** Lane 1: marker; Lane 2: Patient 1 IgG depleted plasma; Lane 3: Patient 1 plasma; Lane 4: Patient 2 IgG depleted plasma; Lane 5: Patient 2 plasma. Boxed region indicates the removal of IgG between 140 – 170 kDa when compared to non-IgG depleted plasma. 180

Figure 5 **ELISA confirmation of IgG removal from plasma samples** 180

Figure 6 **Histamine release from PBMCs induced by either pork or beef extracts from two patients (P1 and P2).** PBMCs were co-incubated with plasma (blue bars) or IgG depleted plasma (red bars). (A) Histamine release induced by beef extract with P1 plasma. (B) Histamine release induced by beef extract with P2 plasma. (C) Histamine release induced by pork extract with P1 plasma 182

List of abbreviations:

Chemicals and reagents:

AP Alkaline Phosphatase
BCIP 5-Bromo-4-Chloro-3-Indolyl Phosphate
BSA Bovine Serum Albumin
EB Elution Buffer
EDTA Ethylene Di-amine Tetra-Acetate
IPTG IsoPropyl-beta-D-ThioGalactopyranoside
LB broth Luria-Bertani broth
NBT Nitro-Blue Tetrazolium
NC Nitro-Cellulose
PBS Phosphate Buffered Saline
PBST Phosphate Buffered Saline, with Tween 20
SDS Sodium Dodecyl Sulphate

Units and measurements:

b	base
n	nano
c	centi
O.D.	Optical Density
Da	Dalton
p	pico
g	gram
pH	activity of H+ in solution (acidity)
Hz	Hertz
pmol	pico-mole
I.U.	International Units
rpm	revolutions per minute
k	kilo
V	Volt
l	litre
v/v	volume by volume
m	milli / metre
M	Mole
°C	degree Celsius
µ	micro

Assays and analytical tools:

BLAST	Basic Local Alignment Search Tool
BLOSUM	Blocks Substitution Matrix
DBPCFC	Double-Blind Placebo-Controlled Food Challenge

ELISA Enzyme-Linked Immuno-Sorbent Assay
FASTA FAST-All (sequence alignment program)
MALDI-TOF Matrix-assisted laser desorption-ionization-time of flight
MS Mass Spectrometry
PAGE Poly-Acrylamide Gel Electrophoresis
PCR Polymerase Chain Reaction
RAST Radio Allergo-Sorbent Test
SPT Skin Prick Test

Organizations:

FAO Food and Agricultural Organization (United Nations)
I.U.I.S. International Union of Immunological Societies
WHO World Health Organization

Others:

APC Antigen-Presenting Cell
CCDs Cross-reactive Carbohydrate Determinants
cDNA complementary DNA
cds coding sequence
Ek/LIC Enterokinase / Ligation-Independent Cloning
EST Expressed Sequence Tag
IFN Interferon
IgE Immunoglobulin E
IgG Immunoglobulin G
SD Standard Deviation

Abstract

Meat allergy and the allergenic components: Underlining reasons for the absence of clinical presentation to meat antigens despite the presence of high levels of specific IgE

Little is known about meat allergy and the allergenic components involved. Using a dot-blot immunoarray, we showed high IgE-binding frequency to beef (8%), pork (12%) and mutton (5%) in 1096 allergic patients' sera tested. High degree of cross-reactivity between the meat antigens was observed with inhibition ELISA. Identification and characterization of meat-based allergens were achieved using a dual bioinformatics (allergenicity prediction based on allergen-motif or sequence homology) and proteomics (2D electrophoresis and mass spectrometry) approach. Bioinformatics approach predicted 252 distinct putative allergens from six animal species whereas proteomics approach identified 56 IgE-reactive proteins from beef, pork and mutton. Despite presence of high levels of meat specific IgEs, only 2 out of 18 patients tested via SPT were beef-positive. The high IgE-binding to meat extracts is mainly due to presence of mammalian cross-reactive carbohydrate determinants (CCDs). Negative SPT is due to presence of "blocking" IgG antibodies which inhibits histamine release.

CHAPTER 1: INTRODUCTION

1.1 ALLERGY

1.1.1 Basic concepts of allergy

The word “allergy” actually derived from Greek, meaning “altered reactivity” (Arshad, 2002). The term allergy was first coined by Clemens von Pirquet in 1906 to distinguish between beneficial and harmful immune reactions (Roecken *et al*., 2004). However, allergy is a word that is as often misused as it is used correctly. Many people will assume all intolerance reactions, such as allergic, pseudoallergic, idiosyncratic, or toxic, are allergic reactions. Today one defines allergy as an inappropriate and harmful immune response against exogenous substances (allergens), which are normally harmless (Arshad, 2002). The chief actor in an allergic reaction to an allergen is the acquired, specific immune response. The initial exposure with a potential allergen may lead to sensitization of the exposed entity without producing any clinical symptoms. Antigen-specific lymphocytes and antibodies are produced. When the individual is exposed to the antigen again, an allergic reaction with clinical signs and symptoms can appear.

1.1.2 Hypersensitivity

Hypersensitivity is the abnormal or exaggerated response of the immune system, resulting in cellular and tissue damage (Arshad, 2002). Four or five types of hypersensitivity are often described. The four types of hypersensitivity reaction (Type I,

II, III, and IV) were expounded by Gell and Coombs (Gell and Coombs, 1963). Type V hypersensitivity (often used in Britain); termed "stimulatory" was later added to distinguish from Type II. The different types of hypersensitivity are not mutually exclusive as more than one type of immune response is often involved in hypersensitivity.

1.1.3 Mechanism of Allergy – Type I (immediate) hypersensitivity

The term "allergy" is basically used to refer to a type I immediate hypersensitivity reaction (Roitt *et al*., 1998). IgE antibodies mediate this reaction.

Antigens (or allergens) enter the body through the respiratory and gastrointestinal mucosa and the skin. Subsequently, the antigen-presenting cells (APCs) engulf the antigens and, after processing, present these to the naïve T cells (Th0). In non-atopic individuals, this process will mount a low-grade immunological response and produce allergen specific IgG_1 and IgG_4 antibodies (Kemeny *et al.,* 1989). Their T cells respond to the antigen with a modest degree of proliferation and production of IFN-γ that is typical of Th1 cells (Romagnani, 1991; Ebner *et al.,* 1995; Till *et al.*, 1997). In atopic individuals, this process stimulates the production of Th2 cells, which then secrete cytokines, IL-4 and IL-13. These cytokines cause proliferation and switching of B cells to IgE-producing B and plasma cells, specific to the antigen (Pene *et al.*, 1988; Finkelman *et al.*, 1988;
Punnonen *et al.*, 1993; Emson *et al.*, 1998). Some of these cells have a long life and are called memory cells. The IgE circulates in the blood in small quantities but mostly present in the tissues bound to high-affinity receptors (FcεR1) on the surface of mast cells

and low-affinity receptors (FcεR2) on eosinophils, macrophages and platelets (Roitt *et al*., 1998).

Upon re-exposure to allergen, cross-linking of allergen specific IgE occurs. This early response causes mast cell degranulation and the secretion of mediators such as histamine, tryptase, heparin, prostaglandins, leukotrienes, and bradykinin (Kinet, 1999). These mediators cause vascular dilation, increased permeability and attract cells into the tissues, thus leading to inflammation. The symptoms of immediate hypersensitivity reactions include erythema and urticaria on the skin, coughing, wheezing, sneezing, rhinorrhea, blocked nose, watery eyes, and more serious conditions such as asthma and anaphylaxis.

The late response takes place a few hours after the allergen exposure. Eosinophils are the most important cells at this stage but lymphocytes, mononuclear cells and neutrophils are also involved. Mast cells, lymphocytes and eosinophils secrete IL-3, IL-5, granulocyte-macrophage colony-stimulating factor (GM-CSF) and chemokines to promote eosinophils proliferation, activiation and survivial (Arshad, 2002). Upon activation, eosinophils release pre-formed and newly synthesized mediators such as eosinophilic cationic protein (ECP), major basic protein (MBP), leukotrienes and prostaglandins to enhance inflammation and prolong epithelial damage (Dombrowicz and Capron, 2001; The Allergy Report, 2000).

1.2 Food allergy

Recently, the World Allergy Organization (WAO) created a Nomenclature Review Committee to review the European Academy of Allergology and Clinical Immunology (EAACI) Nomenclature Position Statement (NPS) and to present a globally acceptable nomenclature for the field of allergy (Johansson *et al*., 2004). The appropriate term for food allergy is when immunologic mechanisms have been demonstrated. If IgE is involved in the reaction, the term IgE-mediated food allergy is appropriate. All other reactions should be referred to as non-allergic food hypersensitivity (Bruijnzeel-Koomen *et al*., 1995; Ortolani *et al*., 1999)

Food allergy is a major public health issue and it is among the most frequent health complaints of our time. Up to 8% of children and 2% of adults in westernized countries suffer from allergic reactions against various foods (Beyer and Teuber, 2004). Children with atopic disorders tend to have a higher prevalence of food allergy; about 35% of children with moderate to severe atopic dermatitis have IgE-mediated food allergy (Eigenmann *et al*., 1998) and about 6% of children with asthma have food-induced wheezing (Novembre *et al*., 1988). There is limited data on food allergy in Singapore, but in a questionnaire survey done by schoolchildren in 1997 estimated the prevalence to be of 4 – 5% (Hill *et al*., 1999)

Food allergy is a malfunction of the immune system in response to dietary antigens (Beyer and Teuber, 2004). It develops in genetically predisposed individuals when oral

tolerance fails to develop normally or break down (Sampson, 2003). The underlying immunologic mechanisms involved in oral tolerance induction have not been fully elucidated, but recent studies suggest that various antigen-presenting cells, especially intestinal epithelial cells and various dendritic cells, and regulatory T cells play a central role (Mowat, 2003).

1.2.1 Food allergens

Although hundreds of different foods are a part of the human diet, only a small number account for the vast majority of food allergic reactions. In young children, milk, eggs, peanuts, soybeans and wheat account for approximately 90% of hypersensitivity reactions whereas, in adults, peanuts, fish, shellfish and tree nuts account for approximately 85% of reactions (Krishna *et al*., 2001). Recently, due to the increased accessibility of fresh fruits and vegetables from various part of the world, there are more reported cases of allergic reaction to fruits (e.g. kiwi, apple, peach and pear) and vegetables (e.g. celery, carrot and lettuce) (Crespo and Rodriguez, 2003). The regional dietary habits and methods of food preparation also play a role in the prevalence of specific food allergies in various countries (Sampson, 2004). Processing of food may weaken or enhance allergenicity. For example, in the case of peanut allergy, dry roasting (180°C) of peanut have been shown to increase the allergenicity of peanut proteins (Maleki *et al*., 2000; Beyer *et al*., 2001; Maleki *et al*., 2003; Chung *et al*., 2003). Cross-reactivity is also an important issue in food allergy. However, when we talk about cross-reactivity, we have to distinguish sensitization from symptom elicitation. This is because

not all cross-reactive IgE antibodies give rise to clinical food allergy (Aalberse *et al*., 2001, 1993; van Ree *et al*., 2004). Table 1 shows the pattern of cross-reactivity between food proteins and clinical cross-reactivity among members of plant and animal species.

Table 1
Cross-reactivity between food proteins and clinical cross-reactivity among members of plant and animal species (adapted from Krishna *et al*., 2001).

Plant material	*Cross-reacting foods*	
Silver birch and hazel pollens	Parsnips Oranges Raw apples Onions Raw carrots	Tomatoes Raw potatoes Hazelnuts Raw celery
Grass pollen	Peaches Plums Apricots Cherries	
Mugwort pollen	Celery Apples Peanuts Kiwi fruit	
Ragweed pollen	Bananas Melon (watermelon, cantaloupe, honeydew)	
Latex	Bananas Chestnuts Avocados	Papayas Kiwi fruit Pineapples
Peanuts	Legumes (except lentils)	
Soya beans	Legumes	
Wheat	Other cereal grains	
Peanuts	Tree nuts	
Tree nuts	Other nuts	
Animal material	*Cross-reacting foods*	
Eggs	Chicken meat	
Cows' milk	Beef/veal	
Cows' milk	Goats' milk	
Beef/veal	Lamb	
Fish	Other fish species	

Sensitization to food allergens can occur in the gastrointestinal tract (considered traditional or class 1 food allergy) or as a consequence of an allergic sensitization to inhalant allergens (class 2 food allergy) (Breiteneder and Ebner, 2000). Food allergens are primarily water-soluble glycoproteins that have molecular weights ranging between 10 to 70 kDa and generally stable to heat, acid and proteases (Sampson, 1999). As stated in a review by Breiteneder and Radauer, as more allergenic proteins are being identified, isolated, and characterized, it has become apparent that similar types of animal and plant proteins make up the vast majority of food allergens (Breiteneder and Radauer, 2004).

1.3 Meat allergy

Meat is a main source of proteins in western diets and it is an important food for children as its high content of polyamines is involved in the development of children's gastrointestinal mucosa (Johnson, 1987). However, meat allergy has long been considered a rare pathology, occurring mainly in children (Restani *et al*., 1997). As the number of studies regarding the nature, epidemiology, and symptoms of meat allergy increases, it clearly indicates that current situations on meat allergy are under-reported and it may not be so rare. The prevalence of beef allergy ranges from 3% to 6.5% among children with atopic dermatitis and can be up to 20% in cow milk allergic children (Besler *et al*., 2001). Several studies reported an incidence of 1-2% of food induced anaphylactic reactions cause by ingestion of beef (Kanny *et al*., 1998; Biedermann *et al*., 1999). Reports of allergy to pork are relatively rare. Challenge proven allergy to pork

meat ranges from 0.6% to 2.6% in food allergic individuals (Besler *et al*., 2001). Also, several anaphylactic and fatal reactions have been described (Pavel and Comanescu, 1969; Wüthrich, 1996; Llatser *et al*., 1998; Drouet *et al*., 2001). Prevalence for lamb allergy has not been reported but anaphylactic reaction has been reported (Welt *et al*., 2005). Cross-reactivity among various meats has also been reported (Fiocchi *et al*., 1995; Restani *et al*., 1997; Restani *et al*., 2002; Mamikoglu, 2005). Studies on pork meat allergy have revealed a high frequency of concomitant allergy to cat epithelium (pork-cat syndrome) (Drouet and Sabbah, 1996). Further investigations showed that serum albumin is the common allergen and that the frequency of sensitization among cat-allergic patients was 14% to 23% for cat serum albumin and 3% to 10% for pig serum albumin (Hilger et al., 1997). Immunoglobulins are also involved in cross-reactivity between meats from different animal species and in milk/meat co-sensitization (Ayuso *et al*., 2000; Belser *et al*., 2001b; Restani, 2002). The high degree of structural similarity between albumins and immunoglobulins suggests that patients sensitized by one species are likely to react to several different animal meats and epithelia (Mamikoglu, 2005).

1.3.1 Meat-based allergens

Currently, only limited information exists on meat allergens and their IgE-binding patterns in meat-allergic patients. The allergens are usually divided into specific groups based on their biochemical composition, sequence homology and molecular weight. The classification is based on the system of nomenclature as recommended by the World Health Organization/International Union of Immunological Societies (WHO/IUIS). The first three letters denotes the genus, followed by the first letter of the species name and an Arabic

numeral. The Arabic numeral indicates the chronological order in which the allergen was isolated. For example, Gal d 1 is first isolated from *Gallus domesticus* (chicken). Allergens from different species of the same or different genus which share the common biochemical properties are usually considered to belong to the same group (WHO/IUS, 1994). However, the nomenclature for animal allergens is not well classified and established. Currently, very few animal meat-based allergens has been designated under the WHO/IUIS system of nomenclature. Only seven allergens from *Bos taurus/domesticus* (beef) have been listed (Table 2) and none from *Sus scrofa* (pork) and *Ovis aries* (mutton).

Table 2
Known allergen from *Bos taurus* listed on WHO/IUIS nomenclature system

Species name	Allergen name	Biochemical ID or obsolete name	MW kDa
Bos taurus	Bos d 2	Ag3, lipocalin	20
	Bos d 3	Ca-binding S100 hom.	11
	Bos d 4	alpha-lactalbumin	14.2
	Bos d 5	beta-lactoglobulin	18.3
	Bos d 6	serum albumin	67
	Bos d 7	immunoglobulin	160
	Bos d 8	caseins	20-30

Most of the cases of meat allergy are due to bovine serum albumin (BSA) (Bos d 6) sensitization, as demonstrated by double blind placebo-controlled food challenges (DBPCFC) (Fuentes *et al*., 2004). This protein has been described as one of the most important allergens in beef and it is also one of the most widely studied and applied protein in biochemistry (Werfel *et al*., 1997; Tanabe *et al*., 2002). Its complete amino acid sequence and three-dimensional conformation has been determined (Hirayama *et al*.,

1990; Holowachuk *et al*., 1991). The tertiary structure is made up of three domains, I, II, and III (1 – 190, 191 – 382, and 383 – 581) and it consists of nine separate disulfide-bonded loops connected by peptide links of 11 – 26 residues (Brown, 1975; Brown, 1977; Peter *et al*., 1977). This tertiary structure and repeating pattern of disulfides is conserved among serum albumins from other species including human (Gelamo and Tabak, 2000). Tanabe *et al*. has observed nine IgE-binding epitopes and three T-cell epitopes that were found to induce T cell proliferation (Tanabe *et al*., 2002).

Besides BSA, another major cross-reactive beef allergen is bovine gamma globulin (BGG) particularly the immunoglobulin G (IgG) (Bos d 7). This 160 kDa protein was detected in raw beef as an allergen in 83% of beef-allergic patients tested (Ayuso *et al*., 2000). BGG is heat stable at 60°C but show reduced antigenicity at 100°C (Han, *et al*., 2002). This is because heat treatment at 100°C results in heat-coagulation (precipitation) of the beef extract. Nevertheless, the precipitate is still able to induce IgE-binding with patients' sera indicating the persistent antigenicity of the allergen (Han *et al*., 2002).

There are also other meat-based proteins that have been reported in literatures but not listed under WHO/IUS. Among the muscle proteins, tropomyosin was found to be a weak meat allergen (Ayuso *et al*., 1999). Other beef proteins that appear to be allergenic are actin and the heat resistant myoglobin (Restani *et al*., 1997; Fuentes *et al*., 2004). For pork meat, allergens at molecular weights 67, 65, 51, 45, 43, 41, 40, 31, and 28 – 30 kDa have been reports in several literatures (Sabbah *et al*., 1994a; Sabbah *et al*., 1994b; Asero *et al*., 1997; Llatser *et al*., 1998; Benito *et al*., 2002; Atanaskovic-Marković *et al*., 2002).

Similarly, for chicken meat, allergens at molecular weights 150, 66, 50, 45, 33, 31, 28, 27, 24, 23, 21, 20, 17, and 13 kDa have been reported (Cahen *et al*., 1998; Ayuso *et al*., 1999).

1.4 Trends in meat-based allergies

In the 1980s, few reports in the literature had described hypersensitivity reactions to meats, the allergens involved, or the actual frequency of what was perceived to be a rare condition. Some nutritionists have even dubbed beef "allergy-safe", whereas others have included it in hypoallergenic diets (Crawford, 1980). Today, we are beginning to realize that meat allergy may not as infrequent as previously thought. Many studies have been done on bovine proteins in relationship with cow's milk allergy. For patients who suffer from multiple food allergies – including many infants and children with cow's milk allergy – the removal of bovine proteins from the diet may pose a cascade of nutritional problems and it is not supported in current opinion (Fiocchi and Restani, 2002).

It has been suggested that food allergy in childhood is an altogether different entity from food allergy in adults or adult-onset food allergy (Fiocchi *et al*., 2005). Evidence also suggests that beef allergy may involve two modalities whereby clinical reactivity to bovine proteins can be acquired depending on the timing of, and/or age at, sensitization (Ayuso *et al*., 2000; Fuentes *et al*., 2004). The suggestion is that in childhood bovine serum albumin (BSA) is the major allergen, whereas in adulthood immunoglobulins and myoglobin may be more or as important. The natural history of bovine protein allergy

proves that children outgrow their clinical symptoms, whereas adults do not and, this would also reflect different sensitization patterns (Fiocchi *et al*., 2005).

Diagnosis of meat allergy remains as an important issue. The golden rule for diagnosis of food allergy is still the food challenge-based assessments particularly the double-blind, placebo-controlled food challenge (DBPCFC). The diagnostic accuracy of skin prick tests (SPTs) in IgE-mediated adverse reactions to bovine proteins has been reviewed by Fiocchi *et al*. In the literature, the positive predictive accuracies of skin prick tests vary between 69% and 100% and the negative predictive accuracies between 20% to 86% for cow milks (Fiocchi *et al*., 2002b). Caution should be exercised before ascribing *in vitro* assays such as immunoblots and Pharmacia UniCAP for diagnosis of food allergy. Studies have shown that no signification correlation was observed between RAST-positivities to meat and serum albumin and the diameters of wheal in skin prick tests, thus underlining that not all sensitizing allergens are clinically relevant (Fiocchi *et al*., 1995). The relative roles of cross-reactivity and co-sensitization also limit the scope of inference that can be drawn for diagnosis.

1.5 Objectives

The purpose of this study is to elucidate the underlining reasons for the observed paradox of high abundance of IgE-binding to meats antigens in sera of allergic patients but no clinical presentation of any allergic response present to these antigens. These data are intended for improved meat allergy diagnosis and also better understanding of oral tolerance

since IgE specific against meat antigens seen to be present without any clinical relevance. Meat allergy may even serve as a model to understand immunotherapy in its natural form.

The specific aims are:

1. Determine the prevalence of detectable specific IgE to meat antigens in the local population using an allergen immunoarray.
2. Evaluate the cross-reactivity among various meat sources namely pork, beef, lamb, chicken and rabbit.
3. Identify and characterize meat-based allergens from pork, beef and lamb using both proteomics and bioinformatics approaches.
4. Clone and express putative allergens identified in previous study and assess IgE-binding reactivity of all the recombinant allergens via immuno-dot blot.
5. Investigate the presence of serum IgE binding cross-reactive carbohydrate determinants (CCDs) in meat antigens and it effects on IgE-binding upon deglycosylation.
6. Evaluate the presence of blocking IgG antibodies in plasma inhibiting histamine release and clinical presentation.

CHAPTER 2: DOT IMMUNOARRAY SYSTEM FOR DETECTION OF ALLERGEN-SPECIFIC IGES

2.1 INTRODUCTION

2.1.1 Techniques in allergy diagnosis

As in other medical conditions, a complete history and physical examination is mandatory for the diagnosis and management of patients with allergic disease. Demonstrating the presence of antigen-specific IgE antibodies and histamine is crucial in establishing the diagnosis of allergic disease. There are various *in vivo* and *in vitro* diagnostic tests for confirming an allergic disease and identifying the allergen.

There are many *in vivo* diagnostic tests available, for instance, skin prick test (SPT), prick-puncture test, modified prick test, prick-prick test, patch test, and intradermal test. Skin prick test is the most preferred method as it is quick, reliable, convenient and safe. Intradermal test although being more sensitive than skin prick test, is prone to false-positive reactions and anaphylaxis may occur. The mechanism of these *in vivo* tests requires the allergen, when introduced into the skin, to react with the IgE antibodies bound to the mast cells and release histamine. The histamine induces wheal and erythema formation on the skin. The size of the reaction depends on the degree of sensitization, amount of standardized allergen injected, number and releasability of mast cells and the reactivity of dermal tissue to histamine. Skin prick test is regarded as positive when the wheal size (mean wheal diameter) is at least 3 mm, with a positive control of >3 mm, in

the absence of a reaction to negative control. Reactions are then scored from 0 to 4+ based on wheal and erythema size and the presence of pseudopodia.

The most commonly used *in vitro* laboratory tests for diagnosis of allergic diseases are measurements of total or specific IgE antibodies. Immunoassay, using solid phase such as tube, sponges, beads or microtitre plate, are commonly used. These include radioallergosorbent (RAST), fluorescent allergosorbent (FAST), ADVIA Centaur®, AlaSTAT®, CARLA®, ENEA®, Hycor HY-TEC® and Pharmacia UniCAP® (Ricci *et al*., 2003; Hamilton *et al*., 1999; Plebani *et al*., 1998; Nolthe and DuBuske, 1997).

In *in vitro* immunoassays, immunosorbent matrices are utilized to bind to the allergens. Patient's serum, containing unknown quantities of IgE is added, and the IgE binds to the allergens. A radiolabelled, fluorescence-labelled or enzyme conjugated anti-human IgE antibody is then added to detect the presence of IgEs. The intensity of radioactivity, fluorescence or colour is directly related to the concentration of IgE present in the patient's serum. A standard curve is built in a parallel test with known quantities of the IgE antibodies to compare and quantify IgE present in the patient's serum.

2.1.2 Advantages of *in vitro* techniques

In vitro techniques are usually employed when *in vivo* testing is for some reason not practical or possible. For instance, a patient had taken a medication that would preclude skin testing (e.g. antihistamines or related drugs). Or the patient suffers from extensive

skin disease and/or has a history of uncontrolled asthma, anaphylaxis, etc. *in vitro* testing, in these scenarios, will allow the patients to avoid coming in contact with the allergens thus reducing the risk of eliciting allergic reactions and the discomfort experienced during testing. Immunoassays are also useful when allergenic substances (e.g. latex, industrial chemicals) are not available as a licensed extract for *in vivo* testing (Dolen, 2003).

The *in vivo* tests such as food challenge and skin testing are usually applied and the results are usually recorded by trained and licensed clinicians and/or technician. Unlike *in vitro* testing, which do not require skilled hands, skin testing is a difficult technique and it requires training and experience to perform with both accuracy and consistency (Nelson, 1994), and for interpretation of results (McCann and Ownby, 2002). Also, current specific IgE immunoassays are both qualitative and quantitative in nature. This provides additional information on the levels of specific IgE antibodies and its relevance in actual clinical situations.

Sample usage for *in vitro* methods is also kept at a minimal. A single drop of serum is sufficient for multiple assays. With current advances in protein-chip array and allergen array technology, the amount of protein required per assay fall in range of pg/nl (Lebrun *et al.*, 2005; Bacarese-Hamilton *et al.*, 2005). This improvement in technology will greatly reduce the cost of current allergy testing. Currently, the average cost of a single *in vitro* test in Singapore ranges between S$18.00 to $25.00 (charges by the National University Hospital and Allergy Laboratories Pte. Ltd.) depending on the system used and type of allergens tested.

2.1.3 Limitations of *in vitro* techniques

The limitations of *in vitro* techniques arise when there are discrepancies in results, the skin test is positive and the immunoassay is negative or vice versa. This discrepancy is usually thought more due to limitations of the immunoassay than to problems with skin testing (van der Zee *et al*., 1988). Immunoassays that lack sensitivity will produce many false negative results. Assay sensitivity can be improved by lowering the assay cutoff, but doing so can result in a loss of specificity (Dolen, 2003). Source materials that contain cross-reactive carbohydrate determinants (CCDs) with epitopes recognized by human IgE antibodies may produce clinically irrelevant false positive results (Mari *et al*., 1999; Foetisch *et al*., 2003; Ebo *et al*., 2004).

Generally, specific IgE immunoassays are known for a lack of standardization (Williams et al., 2000; Szeinbach et al., 2001), even though there are published guidelines for assay design, performance, standardization, and quality assurance (Matsson, 1997). True standardization is probably not possible because of varying sources for raw allergenic materials (Dolen, 1995), differing methods for binding allergen to a detection matrix, and different detection systems. However, this can be overcome by implementing strict quality control measures such as incorporation of various positive and negative controls with repeats. Also, the anti-human IgE antibody used to recognize human IgE captured by allergen should have essentially no cross-reactivity with other immunoglobulin classes.

2.2 MATERIALS AND METHODS

2.2.1 Patients and sera

In total, 1096 consecutive sera from patients suspected of having allergies through clinical symptoms over a period of one year from 2001, were screened. All sera were screened in duplicate.

2.2.2 Skin Prick tests (SPTs)

SPT tests: All tests are performed in the respiratory laboratory by an experienced technician using commercial allergen extracts and the GreerPick skin prick test device (Greer Laboratories, Lenoir North Carolina) and evaluated by an experienced paediatric allergist. A wheal diameter of 3 mm or more in excess of the negative control was considered a positive test result. The allergen extracts included in our panel are commercially produced by Greer (Greer Laboratories, Lenoir North Carolina), except for the *B. tropicalis* extract produced by the allergy and molecular immunology laboratory, National University of Singapore, Singapore. Our SPT panel includes: house dust mite mix (*D. farinae* 5000 AU/ml + *D. pteronyssinus* 5000 AU/ml, standardised.), *Blomia tropicalis* 0.2 mg protein/ml, 50% v/v glycerol, cockroach mix (*Periplaneta americana, Blattella germanica*), mixed feathers (chicken, duck, goose), canary feathers (*Serinus canaria*), kapok seeds, *Alternaria alternata, Curvularia spicifera, Cladosporium herbarum, Aspergillus fumigatus, Candida albicans*, cat hair (standardized cat hair, *Felis catus domesticus* 10,000 BAU/ml) and dog epithelia (*Canis familiaris*). Grass mix (9

grass mix, standardised), acacia, melaleuca, beefwood (Australian pine), oil palm, mango blossom, sage mix and weed mix. Food protein extracts include cow's milk, soybean, egg whole, egg white, peanut, sesame seed, rice, beef, pork, chicken, wheat, fish mix and shellfish mix.

2.2.3 Dotting apparatus

The dotting apparatus used in the allergen array consist of a 384-pin MULTI-BLOT™ replicator (VP386) and colony copier (VP380) from V&P Scientific (California) as well as single well plates (NUNC, USA). The replicator consists of 384 solid 1.19 mm diameter hydrophobic pins, designed to deliver 0.1 µl of liquid onto a membrane. The colony copier registers the replicator to the membrane on a single well plate for high density arrays by using four holes located on the rear of the copier frame. This results in an array of 1536 spots on a 7.5 × 11.5 cm membrane. Figure 1 shows examples of the dotting apparatus and the dotted membrane.

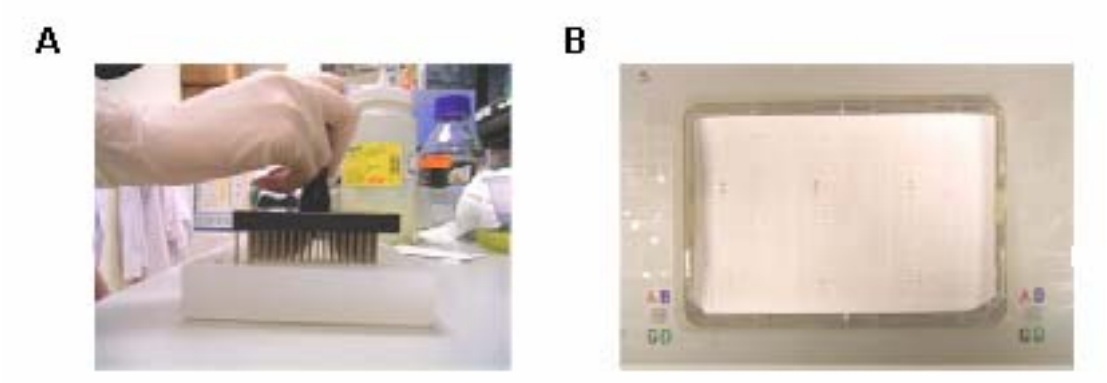


Figure 1.
Images of the dotting apparatus (A) and the membrane dotted with allergens (B)

2.2.4 Support materials and washing buffers

The Trans-Blot nitrocellulose membrane (Bio-Rad, USA) was used as the solid phase support. Dotted membranes were blocked with phosphate buffered saline (PBS) (37mM NaCl, 8.1mM Na_2HPO4, 2.7mM KCl, 1.4mM KH_2PO4, pH 7.4) containing 0.1% of Tween 20 (0.1 % PBS-T) by placing them on an orbital shaker. Similarly, the blocked membranes were washed with 0.05% PBS-T. Immunoarray tested membranes were scanned using conventional scanners (Hewlett-Packard). Dot intensities on the membranes were then measured using the Olympus MicroImage™ image analysis software (Media Cybernetics, 1999) by manually setting the threshold levels of the colour intensities.

2.2.5 Allergen extracts

Skin prick extracts from ALK-Abelló S.A. (Spain) and GREER Laboratories Incorporated (USA) were purchased. However, certain skin prick extracts purchased did not meet the minimum requirement of 0.2 mg ml^{-1} in total protein concentration. Thus, raw materials were purchased from local markets, Greer Laboratories Incorporated and Allergon AB (Sweden). Table 4.1 shows the allergens studied. All raw materials including the local allergens were then homogenized using a mortal and pestle after quick freezing with liquid nitrogen and suspended in PBS extraction buffer for 16 hrs at 4°C. The extracts were then centrifuged at 15,000 g for 15 minutes at 4°C. Supernatants were collected and the pellets discarded. Total protein concentration was then determined

using the BioRad protein assay kit (Bio-Rad Laboratories, USA) based on the Bradford method (Bradford, 1976).

2.2.6 Allergen immunoarray for the detection of specific IgE

Allergen extracts prepared at 0.2 mg ml^{-1} were filled into the 384-well plates. Membranes 7.5 × 11.5 cm in size were then placed onto single-well plates. Each membrane after dotting will consist of three replicates of a full set of allergens. The actual membrane size for a set of the allergen array was approximately 2.5 × 3.8 cm. Protein extracts were spotted at 1μg for each allergen on Trans-Blot nitrocellulose membrane (Bio-Rad, USA). After drying, membranes were blocked with 0.1% PBS-T at room temperature for 1 hour. After washing with 0.05% PBS-T, the membranes were incubated with the respective patients' sera. A total of 150μl of serum was diluted at 1:1 (v/v) PBS and incubated overnight at 4°C. Membranes were then washed followed by incubation with goat anti-human IgE ε-chain specific alkaline phosphatase conjugated antibody (Sigma, USA) at 1:1000 (v/v) dilution with PBS for 2.5 hours at room temperature. Washing was then performed. The positive binding of specific IgE to allergen was visualised by developing with BCIP/NBT (5-bromo-4chloro-3indolyl-phosphate / nitro-blue tetrazolium) colour substrate kit (Promega, USA) in alkaline phosphatase buffer (100mM Tris-HCl [pH 9.0], 150mM NaCl, 1mM $MgCl._2$). The membranes were then blot dried and scanned.

2.2.7 Image analysis of immunoarray blots

The reaction intensities were then measured using the Olympus MicroImage image analysis software (Media Cybernetics, USA). The images were processed through multiple morphological filters before the dot intensity readings were taken (Figure 4.1). First, the images were eroded using a 5 × 5 circle filter for a single pass to average out the colour of the dots. The dots on each membrane were then fitted onto a grid with 48 by 32 rings each with a diameter of 6 pixels by superimposition. Then a mask of the grid was created. The mask image was then filtered with a 5 × 5-circle single pass dilation followed by a 5 × 5-circle single pass closing to obtain solid circles. The logic operation 'AND' was performed (1st operand = eroded array image and 2nd operand = processed grid mask image). The final image, a superimposition of the processed grid image onto the array image, resulted in only the areas dotted with allergens is in its original colour. The rest of the image was black in colour. By using the command "automatic bright objects" which works automatically by setting the OD range of the image using the colour histogram, the dots were segmented from the background. Figure 2 shows the process of image analysis of the immunoarray blots.

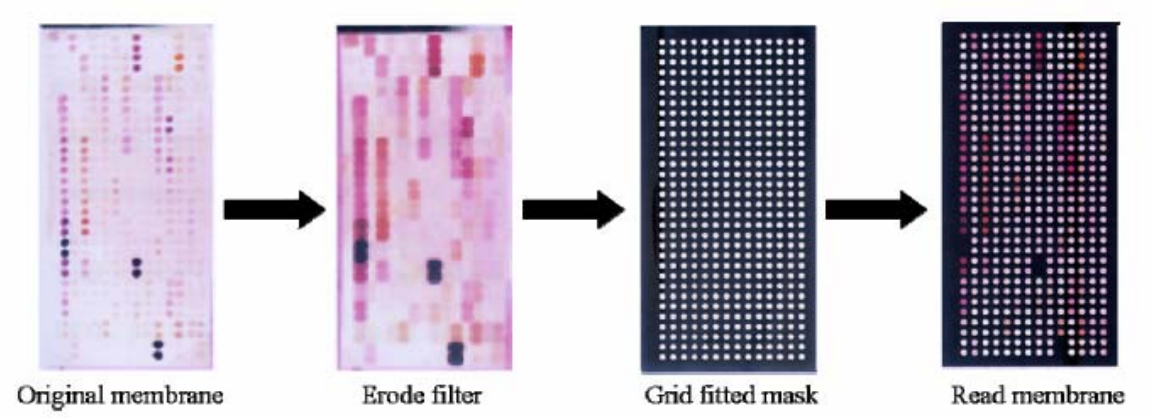


Figure 2.
Process of image analysis of the immunoarray blots

2.2.8 Allergen immunoarray validation

Enzyme-linked immunosorbent assay (ELISA) was used to validate the allergen array results. Sera samples were randomly chosen for validation. A total of 20ug of protein in carbonate/bicarbonate buffer (8mM Na_2CO_3, 17mM $NaHCO_3$, pH9.6) were coated overnight onto each well (NUNC, USA) at 4°C. Then the wells were washed 3 times with 0.05% PBS-T. Blocking was carried out using 0.1% PBS-T for 1 h at room temperature followed by washing. 25ul of patients' sera were then diluted 1:1 (v/v) with PBS and incubated in wells for overnight at 4°C. The wells were then washed and incubated with goat anti-human IgE conjugated with alkaline phosphatase (Sigma, USA) diluted 1:1000 (v/v) with PBS for 2 hours at room temperature. The wells were then washed 6 times before colour substrate *p*-Nitrophenyl phosphate disodium salt (pNPP) (Sigma, USA) was added and colour intensity read at 405ηm. For competitive inhibition ELISA, the patients' sera were first incubated with the inhibitors (pork, beef, lamb,

chicken and rabbit extracts) at concentration 0.2, 0.1, 0.05, 0.025, and 0.0125 mg/ml overnight at 4ºC before being added into the respective wells.

2.2.9 Statistical analyses

Analyses were done using SPSS 11.5 for Windows (SPSS, USA). Two standard deviations (SD) above negative reactions were used as the cut off points for positive results. Subsequently, 4 SD and 8 SD were used as cut off points for medium and high reactions respectively. All concordances were based on positive and negative reactions. Spearman's Correlation Test was used for all correlation analyses. Cluster Analysis was done using the SAS version 9.1 statistical programme (SAS, USA).

2.3 Results and Discussion

2.3.1 Skin prick test

A total of 50 patients from KK Women's and Children's Hospital were recruited for skin prick tests. Eighteen from the 50 have shown to possess relatively high level of specific IgEs to meat antigens, namely to pork and beef antigens. However, only two out of the 18 patients showed positive skin prick test to beef and none to pork. This result indicates that majority of the patient with high specific IgEs to meat antigen did not exhibit any clinical response and the presence of these meat specific IgE have no clinical relevance.

2.3.2 Allergen immunoarray

2.3.2.1 Prevalence of meat-based allergy

After the analysis of all 1096 sera (with a duplicate for each) the prevalence of specific IgE to meat-based antigens is summarized in the Figure 3 below:

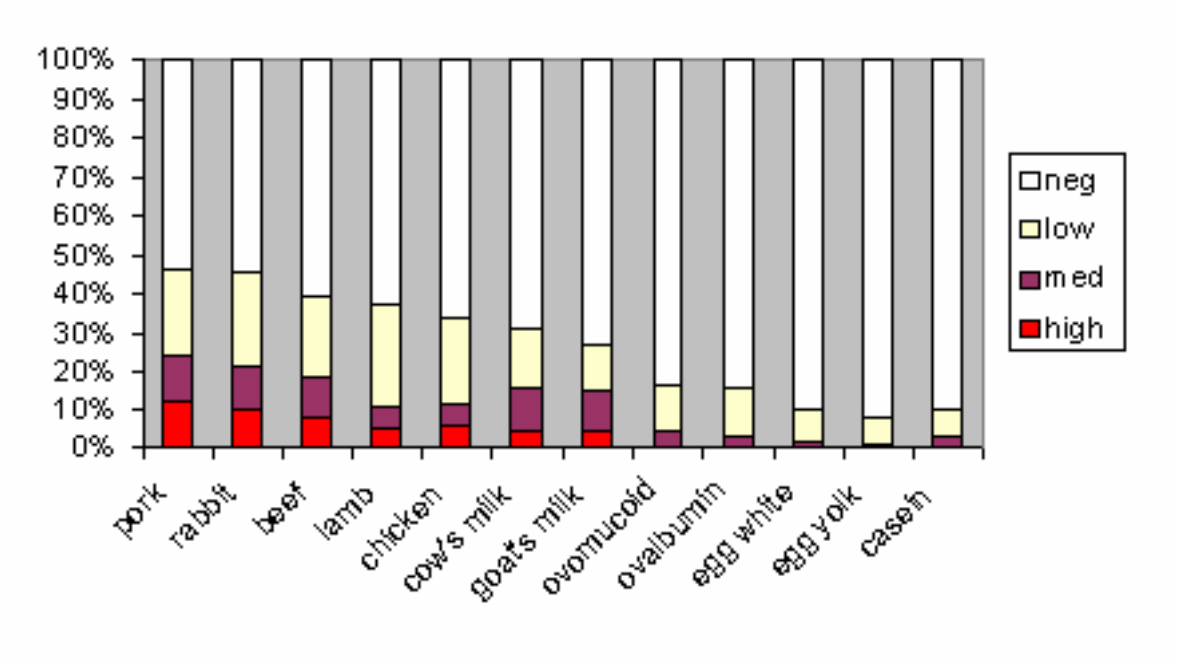


Figure 3.
Prevalence of allergy to meat and other animal products. The cut-offs for low, med and high reactions are at 2SD, 4SD and 8SD respectively.

The sensitivity profile of the 1096 allergic patients can give an idea of the trends of sensitization towards different meat-based antigens in the local context. In this way, further studies on meat allergens can be concentrated on the species that seemed to be the most antigenic. Because strong reactors are more likely to elicit symptoms clinically hence only these were considered. From Figure 3, we can gather that the most common meat-based allergen sources were the red meats mainly pork, beef and mutton, which have shown to have high IgE-binding in 12.13%, 7.94% and 4.74% of patients respectively. This was followed by other animal products such as milk (3.92%) and egg (0.73%). The prevalence of specific IgE to beef and pork we noted were close to double the prevalence of beef and pork allergy reported by Besler *et al*., 2001. This could be due to the differences in population studied and the experimental method administered. The methods employed by Besler *et al*. involved RAST as well as skin prick test which took

into consideration clinical relevance of the specific IgE, thus the prevalence in their framework was expected to be lower. In the case of milk and egg allergy, the prevalence has been reported to be between 1 – 3 %, which was similar to what we have observed in our immunoarray screens (Hill and Hosking, 1995; Kagan, 2003). These parallel observations increased the confidence and reliability of the data as they generally followed expected trends. Many other allergen sources such as plants, pollens, fungi, and insects were tested in the allergen immunoarray panel (listed in appendix I) but were not discussed in this study.

2.3.2.2 IgE responses to pork among individuals (Malay Muslims) who do not consume pork

We also evaluated if pork-specific IgE could be detected among individuals who have never consumed nor handled pork due to their religious beliefs (Malay Muslims), and evaluated if the presence of pork-specific IgE in these individuals were related to cross-reactivity of pork with other vertebrate meats. Forty-six sera from food allergic Malay Muslims were tested by dot blot for specific IgE antibodies to vertebrate meats (beef, lamb, pork, rabbit meat and chicken), and compared to the responses of 548 non-Muslim food allergic subjects (Chinese, Japanese, Caucasian and others). A total of 36.9% (17/46) of the Malay Muslims were found to have specific IgE to pork compared to 46.3% (259/559) among the non-Muslim (Figure 4).

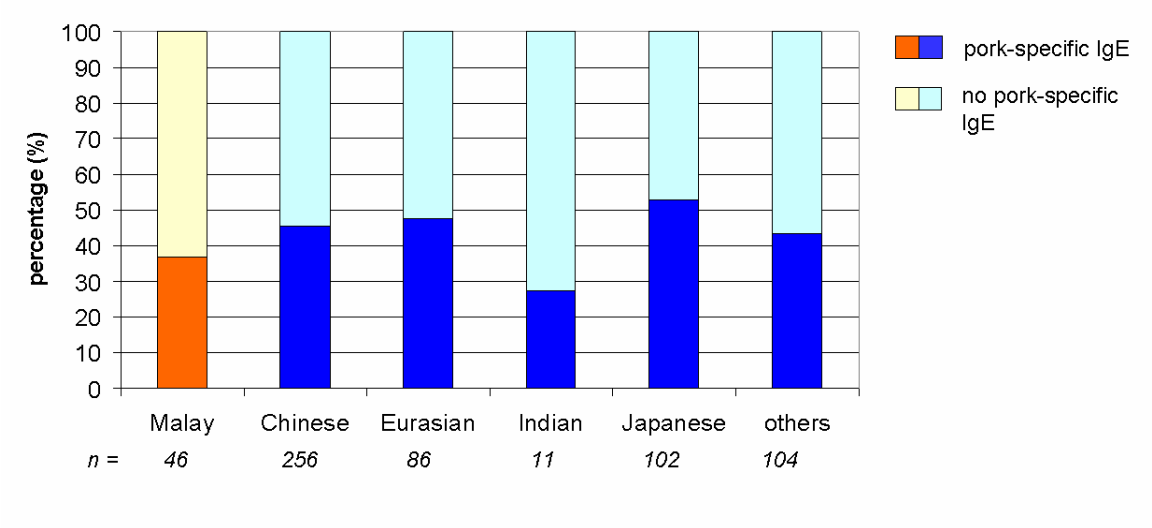


Figure 4.
Percentage of individuals possessing pork-specific IgE in various races.

When stratified by responses to beef, 88.2% (15/17) of pork-positive Malay Muslims were also positive to beef (69.9%, 181/259 of the pork-positive non-Muslims were also positive to beef (Figure 5). This strongly indicates possible existence of cross-reactive allergens within the meat of these vertebrates.

Noticeably the Indian population had a lower IgE-binding frequency to meat antigens. This could be attributed to their diet as a significant proportion of Indians were vegetarians thus they may have possessed lower levels of specific IgE to meat antigens. Another perspective to why the Indian population had a lower IgE-binding frequency to meat antigens may be related to the "hygiene hypothesis". Studies have shown that helminth infection can have a protective effect against atopic reactivity in a various settings in developing countries such as India and Venezuela (Lynch et al., 1987; Braun-Fahrlander, 2002). The helminth infection protect from allergy and asthma either through the saturation of Fcε receptors on mast cells and basophils (by polyclonal IgE) or by

induction of IgG blocking antibodies (Lau and Matricardi, 2006). These confounding factors may have affected the production of IgE against allergens thus preventing allergy.

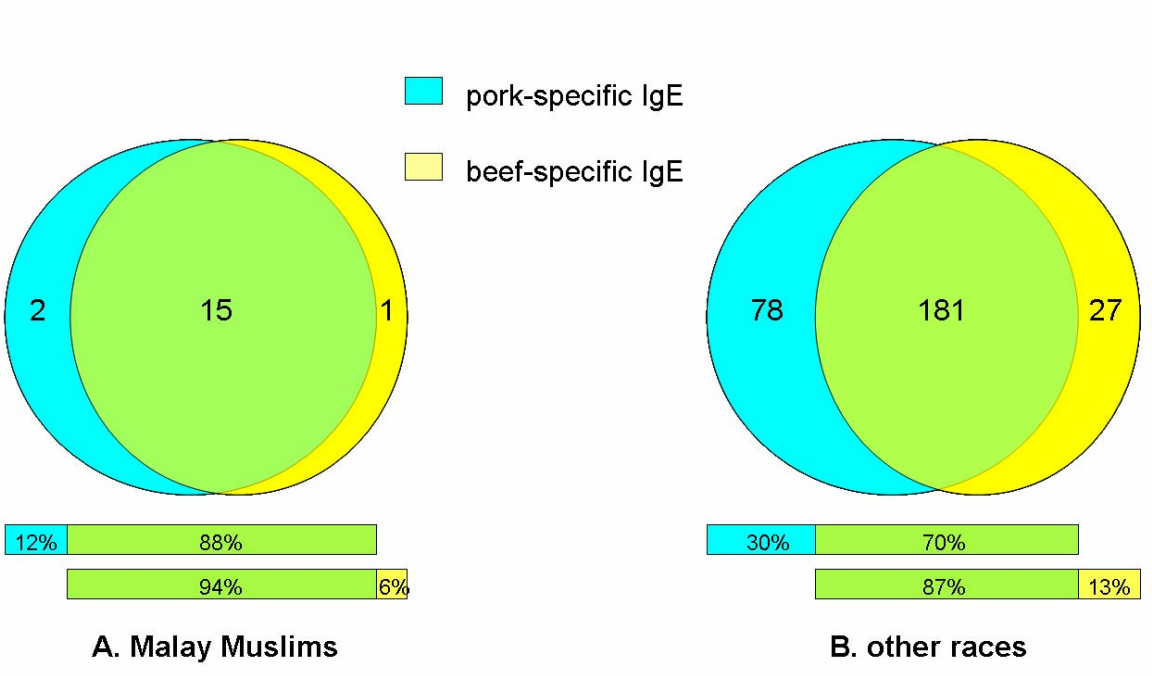


Figure 5.
Venn diagrams showing number and percentages of individuals possessing pork-specific and/or beef-specific IgE. (A) Malay Muslims and (B) other races.

2.3.3 Allergen immunoarray validation

2.3.3.1 Performance of allergen immunoarray in terms of duplicates

All 1096 sera screened on the allergen immunoarray were done in duplicates. Intra-membrane concordances were evaluated by comparing the duplicate dots of the same allergen on a single membrane while inter-membrane concordances were evaluated using results obtained on duplicate membranes. Table 1 shows the inter-membrane and intra membrane concordances for meat-based and animal product allergens on the immunoarray. Generally, the immunoarray showed good intra-membrane and inter-

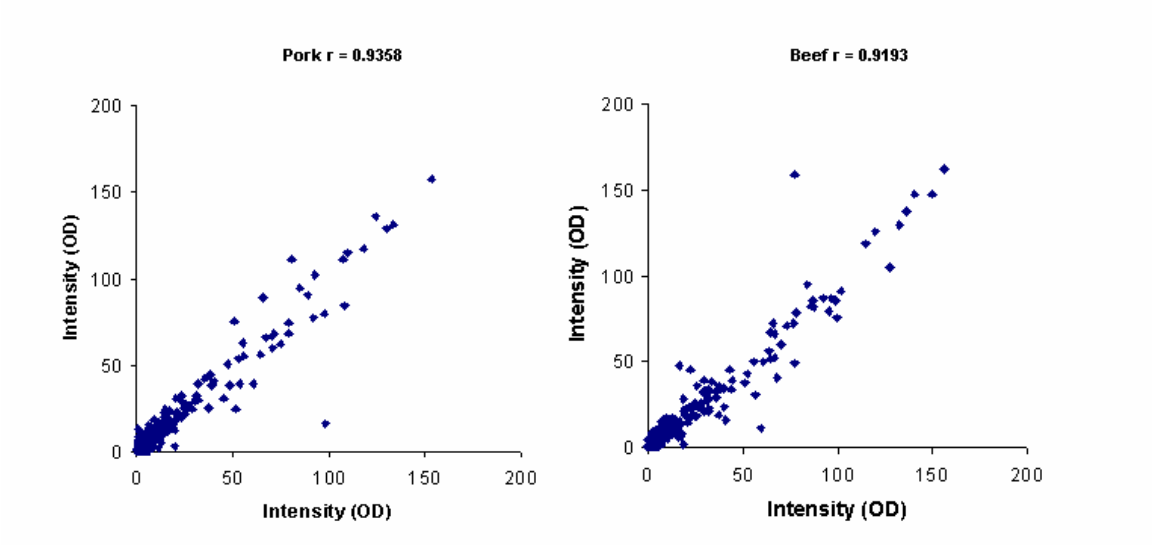
membrane concordances. The intra-membrane concordances ranged from 89.57 to 96.26% while the inter-membrane concordances, from 81.10 to 90.94%. The correlations of intensity values between duplicate dots (intra-membrane) of two selected species, pork and beef, are shown in Figure 6. The Spearman's ranked order correlation coefficient (r) of 0.9358 and 0.9193 respectively showed that the intra-membrane correlation of pork and beef are strong and there was low variance between duplicates. Similarly, for the dots between membranes (inter-membrane), pork and beef showed strong correlation coefficient at 0.6746 and 0.6887 respectively. Both good intra- and inter-membrane concordance and correlation demonstrated the consistency of the screen.

Table 1.
Intra-membrane and inter-membrane concordances

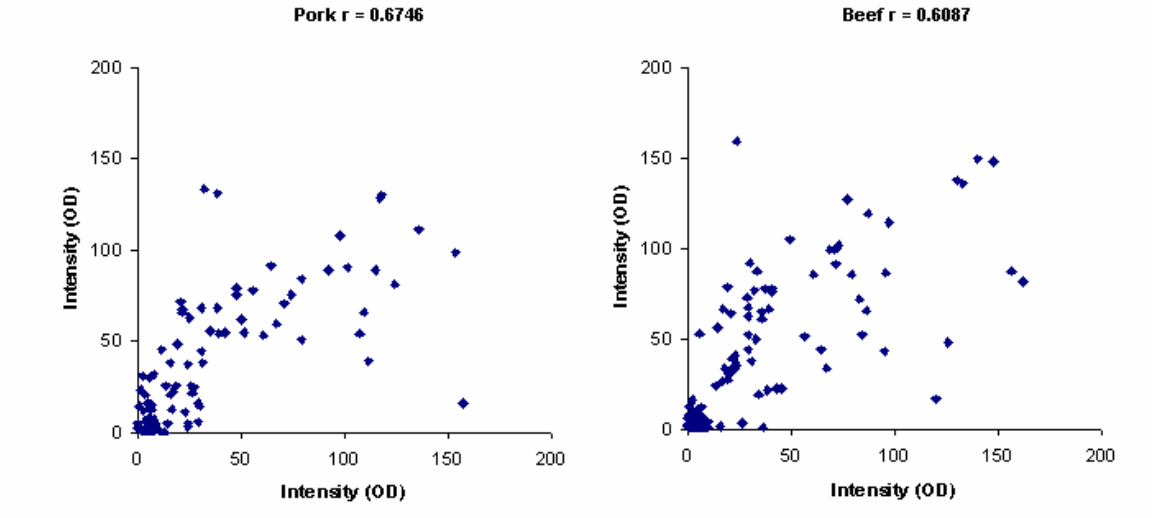
	Intra-membrane Concordance (%)			Inter-membrane Concordance (%)		
Allergen	**Positive (+/+)**	**Negative (-/-)**	**Total**	**Positive (+/+)**	**Negative (-/-)**	**Total**
Beef	90.57	88.48	89.57	75.18	88.03	81.10
Pork	95.18	91.88	93.90	83.65	97.89	88.98
Sheep	90.61	92.40	91.54	81.48	94.96	87.80
Chicken	90.41	94.46	92.72	77.78	90.41	85.04
Rabbit meat	95.22	92.56	94.09	76.82	93.20	83.46
Egg white	72.12	94.80	90.16	54.35	93.75	86.61
Egg yolk	84.15	97.18	95.08	67.50	83.18	80.71
Milk, cow	87.15	95.14	92.32	90.82	91.03	90.94
Milk, goat	90.67	95.24	93.50	80.37	92.52	87.40
Ovalbumin	94.12	96.92	96.26	74.65	97.27	90.94
Ovomucoid	77.22	97.67	94.49	79.17	82.52	81.89
Casein	92.08	93.12	92.91	70.00	91.75	86.61

Figure 6.
Examples of intra-membrane and inter-membrane concordance bi-plots

Intra-membrane bi-plots



Inter-membrane bi-plots



Spearman's Correlation Test analysis was used for the correlation studies. All plots shown are significantly correlated with a minimum of p = 0.05.

2.3.3.2 Immunoarray *vs* ELISA

For ELISA, a total of 3 allergens namely pork, beef and lamb were tested (Table 2). Varying degrees of concordances for pork (75.00%), beef (78.26%) and lamb (78.95%) were obtained. However, Spearman's Correlation Test for the three allergens tested ranged from no correlation for beef ($r = 0.0596$) and lamb ($r = 0.0647$) to moderate correlation for pork ($r = 0.432$, $p < 0.05$) (Figure 7). The reason for the weak correlation is still unknown. However, we hypothesized that it was due presence of blocking IgG in the sera of the patient. The presence of blocking IgG inhibited the binding of IgE to the antigen. For ELISA (sera/antibody in excess), the amount of sera used was much more compared to dot-blot (antigen in excess), thus the effects of blocking IgG antibody were amplified resulting in the discordance between the two techniques. The effect of blocking IgG will be further discussed in Chapter 7.

Table 2.
Validation results between the immunoarray method versus the ELISA system.

	Concordance of immunoarray versus ELISA (%)				
Allergens tested (n = 24 to 76)	+ / +	- / -	+ / -	- / +	Total
Pork	66.66	8.33	25.00	–	75.00
Beef	73.91	4.35	11.59	10.14	78.26
Lamb	76.32	2.63	21.05	–	78.95

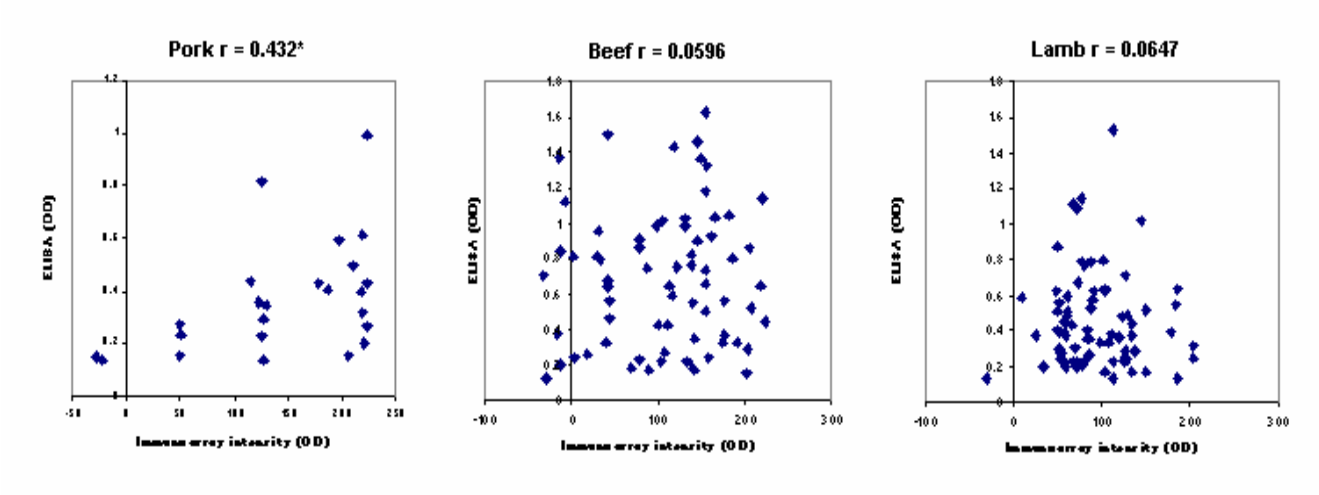


Figure 7.
Correlation of the ELISA versus immunoarray system. Correlation coefficient, r was analyzed using Spearman's Correlation Test. p values: p = 0.05*.

2.3.3.3 Self inhibition

Self inhibition was done to remove the ambiguity that the IgE-binding is due to non-specific binding. In self-inhibition ELISA, the serum is inhibited with the same proteins that the ELISA plate is coated with. Percentage inhibition is calculated by the formula:

$$\frac{(OD_{405} \text{ of positive control} - OD_{405})}{OD_{405} \text{ of positive control}} \times 100\%$$

Figure 8 illustrates a concentration-dependent response was observed when positive sera for pork, beef and mutton were self-inhibited respectively. This shows that sera that were screened positive by ELISA are indeed reacting to the proteins of interest, and it is not due to non-specific binding.

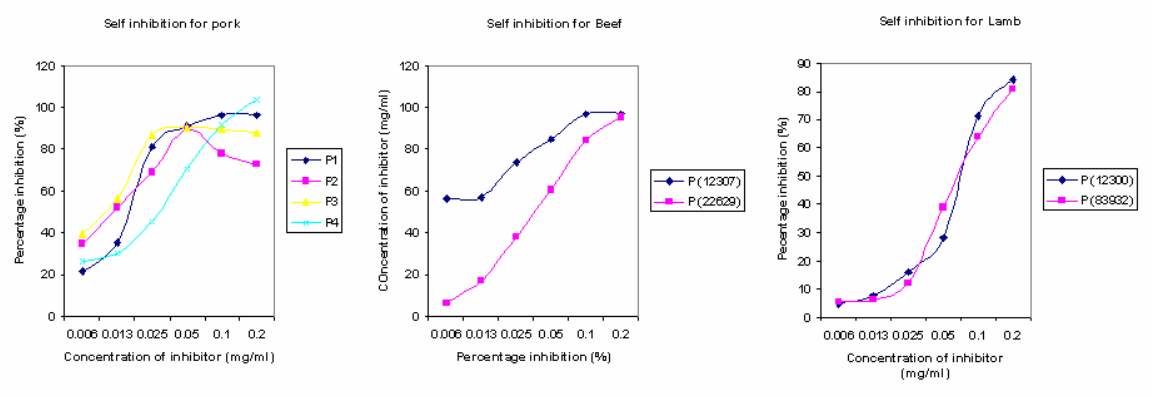


Figure 8.
Graph showing self inhibition for pork (A), beef (B) and lamb (C). The sera were selected based on positivity on both immunoarray and ELISA.

2.3.4 Cross-reactivity

2.3.4.1 Prediction of pattern and potential for cross-reactivity

After observing the results that have been obtained, certain patterns of relationships among the different food can be identified. This led to the suggestion of cross reactivity occurring in the study, where patients were frequently found to be sensitive to not just one, but several food belonging to the same group, especially if the proteins in them have epitope similarities.

To give us an idea of the food groups that would usually cross-react, a Cluster Analysis using the SAS statistical programme has been done (Figure 9). The dendrogram was obtained in a way similar to how a phylogenetic tree can be built in cladistics.

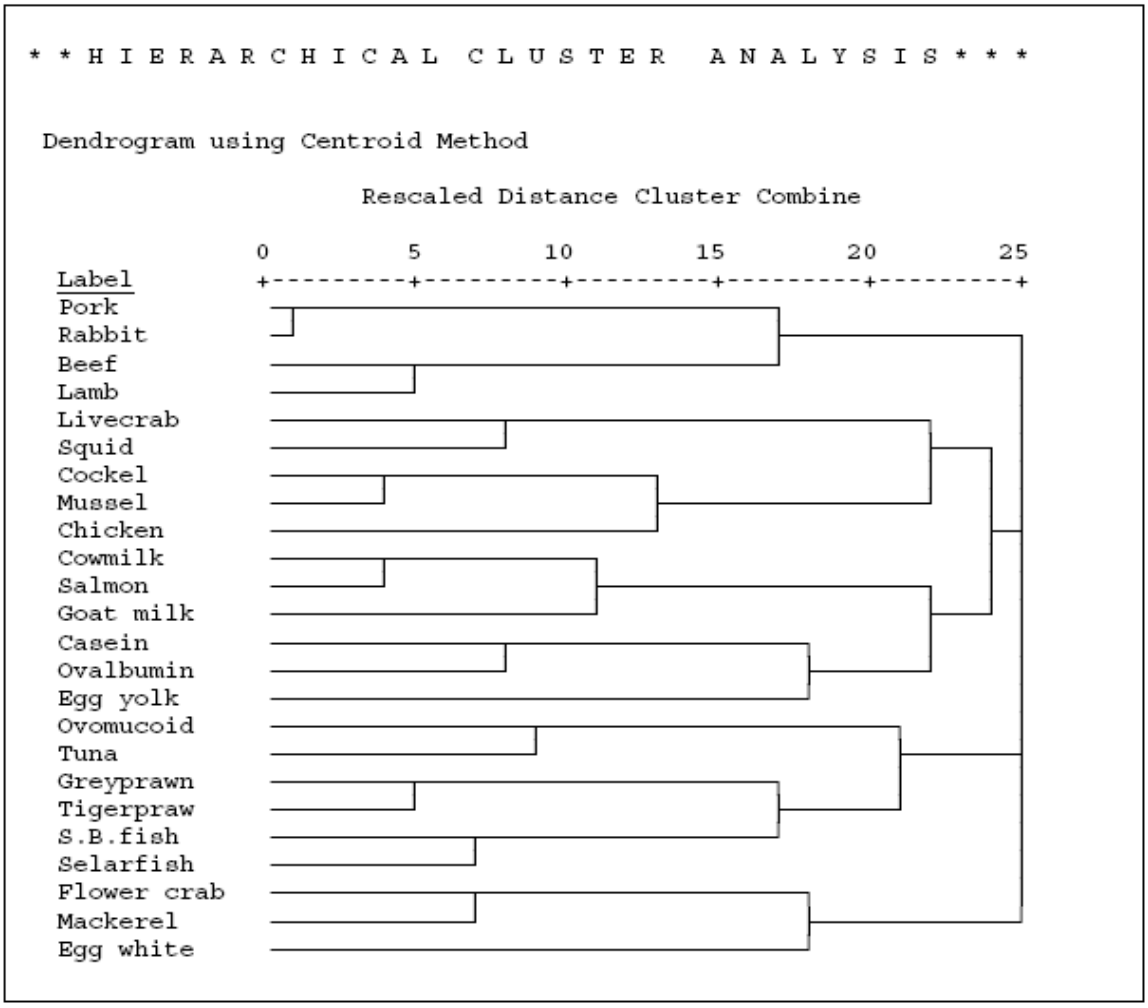


Figure 9.
Dendrogram showing relationships between the allergens used (Done courtesy of Ms Mavis Low)**.**

It was not unexpected to see that all the red meat were clustered together and most of the seafood were grouped together with members of the same phyla. To take it one step further, a few of these patterns which suggested cross-reactivities were singled out so that they could be more closely analyzed. In Figure 10, the cross-reactivity potential among the red meats can be observed. Patients who are allergic to pork will most likely react to lamb and beef as well.

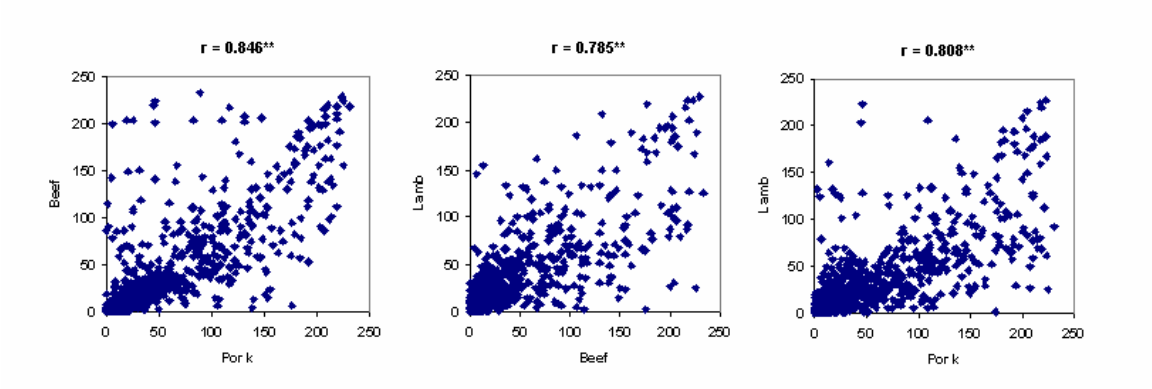


Figure 10.
Correlation bi-plots between pork, beef and lamb. Correlation coefficient, r was analyzed using Spearman's Correlation Test. p values: p = 0.01**.

2.3.4.2 Validation of cross-reactivity via cross inhibition ELISA

Competitive cross inhibition ELISA was done to determine whether IgE-binding was due to cross-reactivity or co-sensitization from the different allergen sources. Figure 11 shows the results of ELISA plates coated with pork and incubated with sera that were pre-incubated with pork, beef, lamb, rabbit and chicken. The results clearly indicated that the three sera tested elicited a similar concentration-dependent inhibition when pre-incubated with pork, beef and lamb. Close to 100% inhibition could be achieved with 0.2 mg/ml of inhibitor proteins (beef and lamb). Additionally, pork, beef and lamb do not cross-react with rabbit meat and chicken (except P3 which showed slight inhibition of close to 40% by rabbit meat). From the steepness of the curve, we can infer that beef is probably the primary sensitizer because it always attended close to 100% inhibition at a lower concentration of inhibitors. This can also explain why the Muslims have specific IgE to pork because they were primarily sensitized by beef.

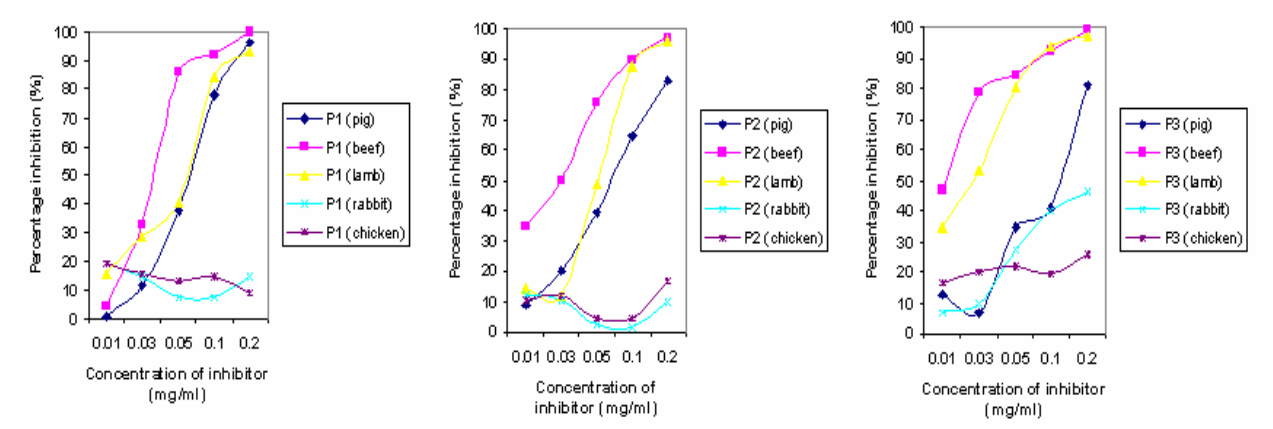


Figure 11.
Graph showing percentage inhibition against amount of protein (micrograms) inhibitors. Sera from three patients were inhibited with beef, mutton, pork, chicken, and rabbit protein. The ELISA plate was coated with pork protein.

To further validate the cross-inhibition results, the reserve was done whereby ELISA plates were coated with beef and lamb and tested with P1 sera pre-incubated with the inhibitors mentioned above. The outcome showed similar pattern of cross-reactivity among pork, beef and lamb (Figure 12). This reverse ELISA demonstrated that the three meats are indeed cross-reactive and also showed the pattern of cross-reactivity is reproducible. Another point to note was that the sera did not cross-react with chicken and rabbit also served as a control for steric hindrance. The fact that there was still a reaction when the sera are incubated with chicken protein showed that the reduced signals observed for sera incubated with beef or lamb were not due to steric hindrance.

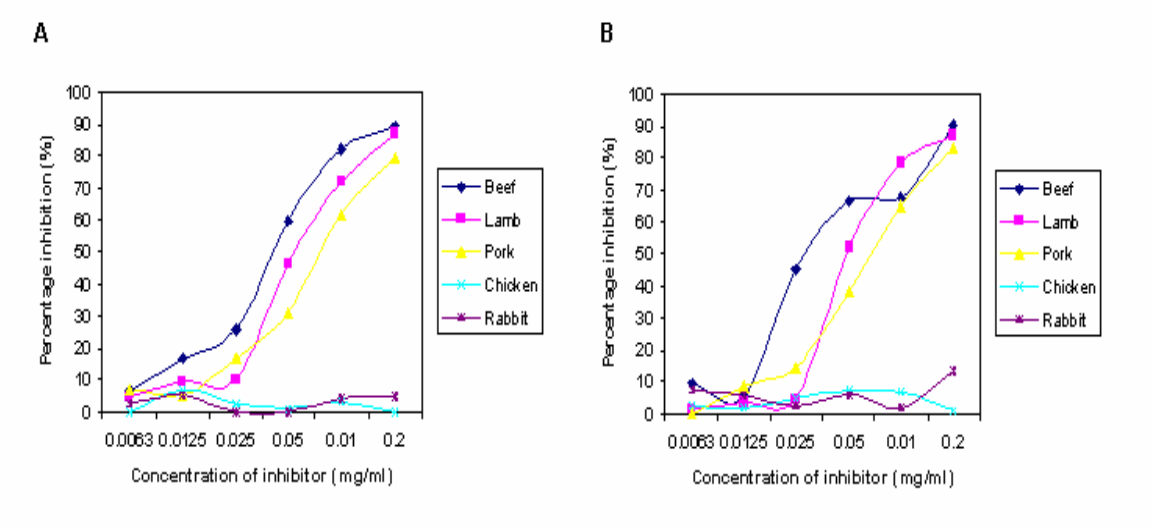


Figure 12.
Graph showing percentage inhibition against amount of protein (micrograms) inhibitors. The serum from P1 was inhibited with beef, mutton, pork, chicken, and rabbit protein. The ELISA plate was coated with beef protein (A) and lamb (B).

2.4 Conclusion

The dot-blot immunoarray system has been shown to be an effective screening system of more than 150 types of allergens in duplicate simultaneously. The amount of serum required is 150 µl per assay and amount is less than 1 µl of serum for each allergen tested. This mass screening method provided useful insights to possible cross-reactivity patterns that may exist. The reliability and dependability of the immunoarray were verified by showing: i) good intra- and inter membrane concordances and correlations; ii) the presence of moderate to high concentrations of specific IgE towards well-known allergenic species such as milk and egg; iii) good concordance between results of immunoarray and ELISA. In this study, specific IgEs to meat-based allergens were detected in the patients' sera samples. Highest number of IgE reactivity was found in the red meats namely pork,

beef and mutton. Self-inhibition ELISA demonstrated that the IgE binding to the meat antigen is a real phenomenon. However, the low occurrences of skin prick positive to meat antigens in patients with high level of specific to meats suggest that presences of these meat specific IgEs lack clinical relevance. Cross-reactivity among meat antigens has also been demonstrated with both Cluster analysis and cross competitive inhibition ELISA.

CHAPTER 3: ALLERGEN PREDICTION USING A BIOINFORMATIC APPROACH

3.1 INTRODUCTION

3.1.1 Establishment of food safety guidelines

The safety evaluation of food, especially genetically modified food has become an important issue. The potential allergenicity and the potency of proteins in our food require apprehensive assessment. There is, however, no universal and reliable test system for the evaluation of the allergenic potency of food products and a case-by-case approach is suggested (Wal, 1999). The first systematic attempt to develop a structured approach to assessment of the allergenic potential of novel food proteins was jointly developed by the International Food Biotechnology Council (IFBC) and the International Life Science Institute (ILSI) Allergy and Immunology Institute and published in 1996 (Metcalfe *et al*., 1996). The proposed IFBC/ILSI decision tree involves a range of test procedures, for example, structure analysis and amino acid sequence comparison, immunosorbent assays, skin prick tests, food challenges, study of physical and chemical properties like digestive and process (heat) stability, and consideration of the amount of the protein in the foods and consumption patterns. In 2001, a Food and Agriculture Organization (FAO) and World Health Organization (WHO) expert consultation modified the IFBC/ILSI decision tree strategy with additional guidelines (accessible at http://www.who.int/fsf/GMfood/) (FAO/WHO, 2001).

Notably, a comparison of the amino acid sequence of the novel protein with those of known allergens is suggested as an introductory step for transgenes without a history of allergy. This is because any homology between the transgenic protein and known allergens is regarded as a risk for potential cross-reactivity and/or immunogenicity (FAO/WHO, 2001). Hence, a standard method for sequence comparison has been defined. Briefly, a protein is considered allergenic if it shares more than 35% sequence similarity (window of 80 residues) or an identity of at least six contiguous amino acids with a known allergen (FAO/WHO, 2001). With amino acid sequence comparisons currently a criterion for allergenicity assessment, there is a need for specialized allergen databases and better bioinformatics methods that could help to classify and predict the allergenicity of a given protein.

3.1.2 Allergen databases

The number of characterized protein allergens is increasing steadily because of advances in genomic, proteomics, and molecular biology techniques. Hence, there is a need for allergy-related databases to facilitate the collection, access, and use of these data. The Allergen Nomenclature Sub-Committee of the International Union of Immunological Societies (IUIS) maintains a list of certified protein allergens (http://www.allergen.org/). Some 500 protein allergens and 270 isoforms had been classified, as of September 2005. These numbers may not be a true representation of the total as many unlisted allergens can be found in the literatures. Also, the database lack detailed structural and functional

information of the characterized allergens. Therefore, in recent years, more number of specialized allergy-related databases has appeared. Major allergen-related data sources are listed in Table. 1. The number of allergen entries varies greatly among the databases; they range from 140 in BIFS to 4500 in ALLALLERGY. Malandain commended that many of the sequences in the databases did not correspond to allergens and many were unproven IgE-reactive sequences either ill-defined or present as US patent data (1035, among them 82% had less than 40 amino acids) (Malandain, 2004). After treating redundancy, only 1052 different allergen sequences was found, a result corresponding with data stored in Allergome, the currently most carefully updated allergen database (Malandain, 2004).

Table 1. Major allergen-related data sources

Database	URL (http://)
IUIS	www.allergen.org
SWISS-PROT Allergen Index	www.expasy.ch/cgibin/lists?allergen.txt
BIFS	www.iit.edu/ sgendel/fa.htm
CSL	www.csl.gov.uk/allergen
FARRP	www.allergenonline.com
PROTALL	www.ifr.bbsrc.ac.uk/Protall
ALLALLERGY	www.allallergy.net
Asthma & Allergy	cooke.gsf.de/asthmagen/main.cfm
Allergome	www.allergome.org/
Informall	www.foodallergens.info/
International Immunogenetics Information System	www.imgt.cines.fr/

3.1.3 Allergenicity prediction

Despite the great number of presently identified allergenic proteins, it is still not known why only few and particular proteins provoke allergic reactions. Thus, a method for allergenicity prediction would be beneficial, especially to prevent the chance of generating new allergenic food products by biotechnology methods. Bioinformatics, more specifically, sequence analysis methods have an important role in the identification and prediction of allergenicity (Hileman *et al*., 2002, Gendel, 2002). Based on current recommendations of the FAO/WHO Expert Consultation, as outlined in the Codex alimentarius (FAO/WHO, 2003), two allergenicity predictive webtools namely Allermatch™ [http://allermatch.org] and AllerPredict [http://sdmc.i2r.astar.edu.sg/Templar/DB/Allergen/Predict/Predict.html] were developed. However, the FAO/WHO criteria have been shown to lack full predictive capability (Kleter and Peijnenburg, 2002; Zorzet *et al*., 2002; Soeria-Atmadja *et al*., 2004). The six amino-acid identity rule is not practical because it produces a large number of false positive hits. The criterion of having a minimum 35% sequence similarity is on the other hand too stringent to find most true allergens.

Analyses using protein motifs emerge as an alternative method for prediction of potential allergenicity (Stadler and Stadler, 2003). It involves the use of motif identification tool, such as MEME (Bailey and Elkan, 1994) to identify allergen motif. Subsequently, a query protein is predicted to be allergenic if its sequence matched an allergen motif with a high score. Similarly, Li *et al*. developed a motif approach allergenicity prediction

system based on a novel motif detecting tool using wavelet analysis (Li *et al*., 2004). Both motif-based prediction methods claimed to have achieved more than 90% precision and their motif-based system is a more superior alternative to the current FAO/WHO method. Using the entire Swiss-Prot as the query, both groups predicted around two thousand potential allergens.

Structural comparison of allergenic proteins can be used to predict allergenic cross-responses (Ipsen and Lowenstein, 1997), and eventually determine possible common characteristics of IgE recognition (Breiteneder and Ebner, 2000; Lascombe *et al*., 2000; Midoro-Horiuti *et al*., 1999; Soman *et al*., 2000; Wellhausen *et al*., 1996). The web-based Structural Database of Allergen Proteins (SDAP) [http://fermi.utmb.edu/SDAP/], which is currently the most ambitious of the molecular databases, was developed for this purpose (Ivanciuc *et al*., 2002; Ivanciuc *et al*., 2003a; Ivanciuc *et al*., 2003b). In addition to allergen sequences and structural links, this database has implemented some unique search capabilities to identify B- and T-cell epitopes and assess cross-reactivity with accompanying references.

Besides computational allergenicity prediction, the FAO/WHO decision tree strategy has also introduced animal testing the focused on the prediction of the sensitizing potential of novel protein (FAO/WHO, 2001). However, despite increasing research efforts in recent years, validated and widely accepted animal models are not available yet (Knippels and Penninks, 2005).

3.1.4 Limitation of bioinformatics allergen prediction

Bioinformatics is a fast and easy method when compared with the conventional allergen identification methods which require tedious laborious work. However, the accuracy of the allergenicity prediction by primary sequence comparison depends strictly on the quality of the data used for comparison. Hence, databases with accurate annotations and correct format of the sequences are essential for the outcome of the prediction. Also, prediction of novel allergens other then those with sequences deposited in the databases is not possible. Nevertheless, bioinformatics allergen prediction serves as an initial screen for putative candidates whereby actual characterization of the allergens still requires immunochemical assays.

3.1.5 Expressed Sequence Tagging in genome studies

Expressed Sequence Tagging was first proposed by Adams *et al.* as a means to rapidly increase the number of available sequences for the human genome project (Adams *et al*., 1991). The basic strategy involves selecting cDNA clones at random and performing a single, automated, sequencing read from one or both ends of their inserts. They introduced the term expressed sequence tag (EST) to refer to this new class of sequence, which is characterized by being short (typically about 400–600 bases) and relatively inaccurate (around 2% error). These short sequences act as gene tags for identifying each cDNA clone through sequence search homology (e.g. BLAST) in simple nucleotide and protein databases.

Partial sequencing of cDNAs to generate expressed sequence tags (ESTs) has been demonstrated to be a rapid and efficient way to establish a detailed profile of genes expressed in a tissue or cell type or even in an organism (Adams *et al*., 1991; Adams *et al*., 1993a; Adams *et al*., 1993b).]. In many cases, the information from a single sequencing reaction has been sufficient to assign a cDNA to a gene family based on sequence similarity (Waterston *et al*., 1992; McCombie *et al*., 1992; Okubo *et al*., 1992; Adams *et al*., 1993a; Khan *et al*., 1999), and provide a reliable and efficient place to start designing experiments to characterize gene function and tissue physiology.

In many organisms, a large part of the genome comprise of nongenic DNA such as repeats, pseudogenes and other non-coding sequences, which makes genome analysis difficult and labour-intensive. For instance, only 3% of the human genome codes for proteins, the function of the rest of the genome is unclear, but much of it may have no function. The easiest way to access protein coding regions is thus to sequence cDNAs. The EST philosophy follows the argument that by solely sequencing the most important protein coding region, it give the largest return of biological information per base pair sequenced (Adams *et al*., 1991; Marra *et al*., 1999) and it is the most cost effective. With EST libraries available in public databases, the rate of identification of new allergens could be accelerated and new findings such as homologous allergens crossing diverse species boundaries could rapidly be obtained.

3.1.6 Unigenes

The sheer number of EST sequences is extraordinary, indeed for most organisms much larger than the number of genes. A major challenge is to make putative gene assignments for these sequences because many of these genes are not well defined and annotated. Computationally, this can be thought of as a clustering problem in which the sequences are vertices that may be coalesced into clusters by establishing connections among them (Pontius *et al*., 2003).

UniGene is an experimental system for automatically partitioning GenBank sequences into a non-redundant set of gene-oriented clusters. Each UniGene cluster contains sequences that represent a unique gene, as well as related information such as the tissue types in which the gene has been expressed and map location. The UniGene Web site allows the user to view UniGene information on a per cluster, per sequence, or per library basis. UniGene is also the basis for three other NCBI resources: ProtEST [http://www.ncbi.nlm.nih.gov/UniGene/ProtEST/], a facility for browsing protein similarities; Digital Differential Display (DDD) [http://www.ncbi.nlm.nih.gov/UniGene/ddd.cgi?], for comparison of EST-based expression profiles; and HomoloGene [http://www.ncbi.nlm.nih.gov/HomoloGene/], which provides information about putative homology relationships. Therefore, Unigene may be a useful database for prediction and identification of putative allergen in various species.

3.2 MATERIALS AND METHODS

3.2.1 Data mining and content

Unigenes, nucleotide and protein sequences (in cases where Unigenes sequences are not available) from various species of animals were downloaded from National Center of Biotechnology Information (NCBI). Nucleotide sequences were used for goat and sheep while protein sequences were used for dog and cat instead. In total, 91,612 sequences have been downloaded (*Sus scrofa* (Pig) 19,590 unigenes, *Bos taurus* (Cow) 17,874 unigenes, *Gallus gallus* (Chicken) 8,834 unigenes, *Oncorhynchus mykiss* (Trout) 13,156 unigenes, *Ovis aries* (Sheep) 21,106 nucleotide sequences, *Canis familiaris* (Dog) 3,415 protein sequences, *Capra hircus* (Goat) 2,476 nucleotide sequences, and *Felis domesticus* (cat) 5,161 protein sequences).

3.2.2 Analysis of Sequence Similarity (Method 1)

The flowchart in Figure 1 describes the method that was used for allergenicity prediction. The search method corresponds with FAO/WHO guidelines described above with more stringent cutoffs. Basically, a sequence homology search using BLASTX (Claverie, 1992) with default parameters (matrix = Blosum 62, Gap open = 11, Gap extend = 1) against NCBI nr datadase was carried. This was done courtesy of Miss Lim Yun Ping from Singapore Biomedical Computing Resource (SBCR), Bioinformatics Institute (BII).

The BLAST output was manually inspected for matches holding the text “allergen” in the annotations. The results were then ranked according to their bit scores and E-values. If the alignment resulted in scores with bit scores above 100 and E-values less than a threshold of 1e -4, the query sequence is predicted to be allergenic.

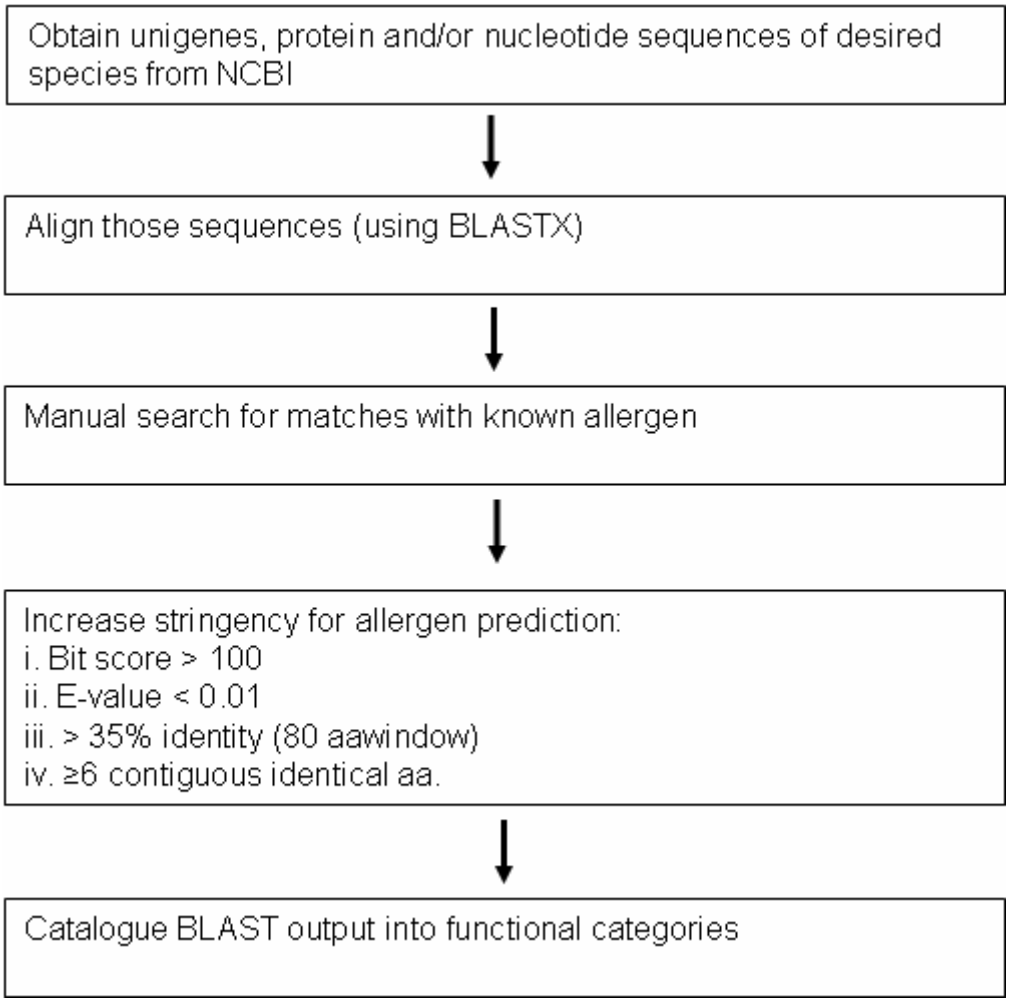


Figure 1.
Flowchart of the entire prediction system based on sequence homology

3.2.3 Allergenicity prediction using wavelet transform (Method 2)

This was done courtesy of Dr Li Kuo Bin according to Li *et al*., 2004. Briefly, a reference database of 664 allergen proteins was constructed by searching four allergen-specific databases (IUIS, Swiss-Prot's Allergen index, BIFS, and FARRP). Subsequently, clustering of these proteins into groups was done using ClustalW, PAM and T-Coffee software. Conserved motifs for each cluster were predicted using wavelet method and a profile for each predicted motif was built using HMMER package. Finally, putative allergens were predicted using HMMER and BLASTP based on a scoring system.

3.2.4 Cataloging of BLAST output into functional categories

For sequences with significant identities, a catalogue classifying them according to their putative biological functions was made. The categories in this catalogue were created based on the Mark Adam's (1993a) classification of human brain cDNA library with minor modifications. These categories are: cell signaling, general metabolism, homeostasis, structural protein, chaperone/stress-related protein, protein synthesis, transport protein and unknowns.

3.3 Results and Discussion

3.3.1 Allergen prediction based on sequence homology

The first main result of the work reported here is the analysis of the massive blast output of 91,612 sequences for homologues to know allergens. Careful evaluation was required to produce a thorough and accurate list of putative allergen for each species therefore substantial effort was put into this work. The arbitrary cut off values of bit scores above 100 and E-values less than a threshold of 1e -4 were required to reduce the number of false positive. The number of putative allergens predicted based on sequence homology for each species is listed in Table 2. The actual list of sequence homology matches for each species with query description, subject description, bit score, E-value and region of amino acid homology is attached in appendix II - VIX. From this result, we can deduce that the percentage of sequences that matched allergenic components fluctuate between 1 – 2 % for each species (Figure 2) suggesting that the number of allergenic components within each species saturate around 2% of its genome. Previous EST libraries studies done on house dust mites have shown that allergenic components made up 4 to 6 percent of their genome (Ong, 2003). The percentage estimated using EST library was expected to be higher because the library consisted of multiple copies of the same genes whereby in our study we employed unique copy genes (unigenes) which significantly reduced the redundancy rate.

Table 2.
No. of putative allergens predicted for each species of animal based on sequence homology.

Animal	No. of putative allergens
Pork	188
Beef	180
Trout	172
Chicken	123
Dog	73
Sheep	70
Cat	45
Goat	25

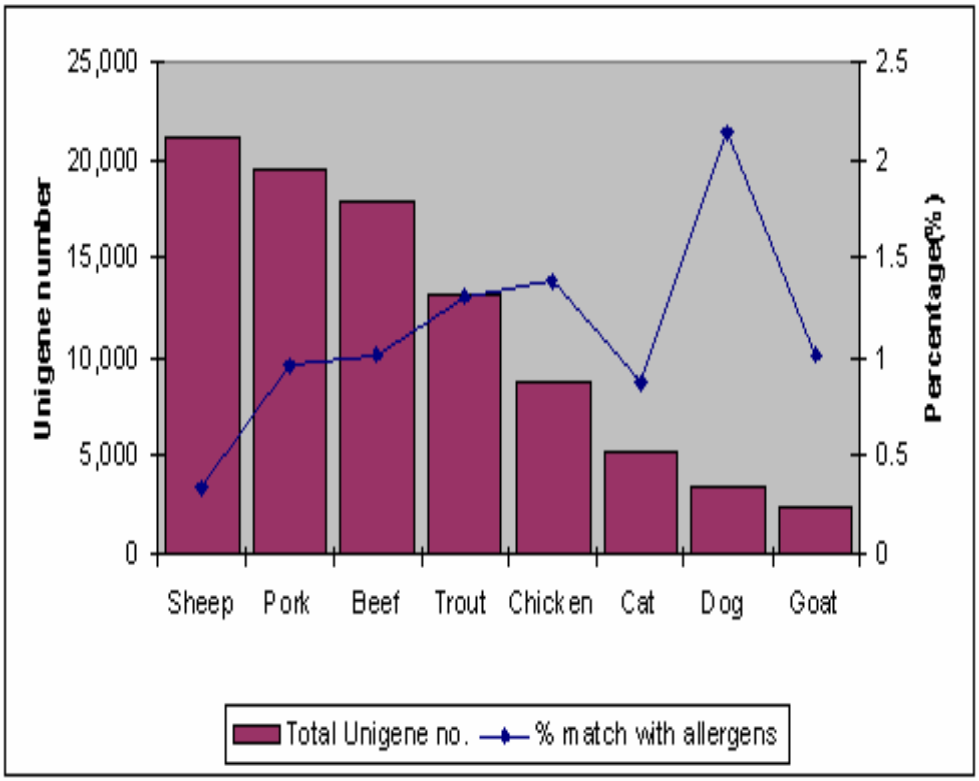


Figure 2.
Relationship between the total numbers of Unigene, nucleotide or protein sequences used for each species (bars) and the percentage of these sequences that match allergens (lines).

3.3.1.1 Matched allergen profiles of the seven animal species

Table 3 shows the type of plant, fungal, insect and animal allergen-homologues within the allergen homologous Unigene sequences. Majority of the unigenes were homologous to allergens from animal and insect. This was expected since the unigenes used were themselves of animal origin. In contrast, the unigenes homologous to allergens from plant origin was much lower. This illustrates the vast difference in allergen-type found between plants and animals hence suggesting that plant food allergens and animal food allergens are less likely cross-reactive. Fungal allergens represented a significant proportion among the allergen-homologous unigenes found. They were mainly stress-related and metabolic proteins such as heat shock proteins, dehydrogenases, superoxide dismutases, thioedoxin, etc. There proteins were evolutionary highly conserved thus they were found in almost all the seven species of animal and most likely cross-reactive. Among the homologous insect and animal allergens, the most commonly found were tropomyosins, serum albumin, beta-lactoglobulin, alpha-lactalbumin, egg white proteins, phospholipase and dust mite proteins (Group 3, 8 and 13). These proteins were again highly conserved evolutionary and most likely cross-reactive.

Table 3.
Unigenes of pig, cow, chicken, goat, sheep, dog, and cat found to be significantly homologous to allergens from various organisms. Ticks indicate the presence of allergen-homologous unigenes (not named) within the animal species.

ALLERGEN TYPE	ALLERGEN HOMOLOGUE	EXAMPLES	PIG	COW	CHICKEN	GOAT	SHEEP	DOG	CAT
Plant									
	Actinidain	Act c 1	✓	✓	✓			✓	
	Calcium-binding allergen	Ole e 8, Bet v 3	✓	✓	✓				
	Enolase	Hev b 9	✓	✓	✓		✓	✓	
Fungal									
	60S acidic ribosomal protein	Cla h 12, Alt a 6	✓	✓	✓	✓	✓		
	68 kDa allergen		✓	✓					
	Aldehyde dehydrogenase	Alt a 10, Cla h 3	✓	✓	✓		✓		
	Alkyl hydroperoxide reductase	Mal allergen	✓	✓	✓				
	Heat shock 70 kDa protein	Pen c 19, Cla h 4, Alt a 3	✓	✓	✓	✓	✓	✓	
	Heat shock protein 90	Asp f 12	✓	✓	✓	✓	✓		
	Malate dehydrogenase	Mal f 4	✓	✓	✓		✓		
	Peroxisomal membrane protein	Mal f 2, Asp f 3	✓	✓					
	Serine protease	Tri s 4	✓						
	Superoxide dismutase	Asp f 6	✓	✓	✓		✓		
	Thioredoxin	Cop c 2	✓	✓	✓	✓	✓		
Insect (Mite)									
	98kDa HDM allergen		✓	✓	✓	✓	✓		
	Adenosine diphosphatase	Aed a 1		✓					
	Allergen A precursor		✓						
	Allergen MAG29		✓	✓				✓	
	Apolipophorin	Der f 14			✓				
	Arginine kinase	Plo i 1	✓	✓	✓			✓	
	Calycin	Bla g 4					✓		
	Cysteine protease	Eur m 1				✓			
	Fatty acid-binding protein	Lep d 13, Blo t 13	✓	✓	✓	✓	✓		
	Gelsolin-like allergen	Der f 16	✓	✓	✓				
	Glutathione S-transferase	Der p 8, Bla g 5	✓	✓	✓	✓	✓	✓	
	Hyaluronidase	Pol a 2, Api m 2	✓	✓	✓		✓		
	Mite allergen	Lep d 1.02		✓	✓				
	Paramyosin	Blo t 11	✓	✓	✓				✓
	Phospholipase A1 1 precursor	Dol m I, Ves v 1	✓	✓	✓		✓	✓	✓
	Tropomyosin	Blo t 10, Lep d 10, Per a 7	✓	✓	✓	✓	✓	✓	
	Trypsin	Der f 3, Eur m 3	✓	✓	✓	✓		✓	✓
	Venom allergen	Dol m 5.01	✓	✓					

Table 3. (cont.)

ALLERGEN TYPE	ALLERGEN HOMOLOGUE	EXAMPLES	PIG	COW	CHICKEN	GOAT	SHEEP	DOG	CAT
Animal									
	Allergen dI chain C2A			✓		✓			
	Alpha, lactalbumin	Bos d 4	✓	✓		✓	✓	✓	
	Alpha-2u globulin	Rat n 1	✓	✓	✓			✓	
	Arginine kinase	Pen m 2	✓		✓			✓	
	Beta-lactoglobulin	Bos d 5	✓	✓		✓	✓	✓	✓
	Bovine dander	BDA20	✓	✓		✓			
	Cat-1	Fed d 1							✓
	Lipocalin	Equ c 1, Can f 1	✓	✓				✓	
	Lysozyme C precursor	Gal d 4		✓				✓	
	Ovalbumin	Gal d 2	✓	✓	✓		✓		
	Ovomucoid	Gal d 1	✓	✓	✓			✓	✓
	Parvalbumin	Gad m 1, Sco j 1, Sal s 1			✓				✓
	S100 calcium-binding protein A7	Bos d 3	✓	✓	✓		✓	✓	
	Serum albumin	Can f 3, Gal d 5, Bos d 6	✓	✓	✓		✓	✓	✓
	Serum transferrin	Gal d 3	✓	✓	✓	✓	✓		
	Tropomyosin	Hom a 1, Met e I, Cha f 1	✓	✓	✓	✓	✓	✓	
Others									
	21k allergen			✓			✓		
	C07A4.3		✓						
	C10G8.3				✓				
	T05A10.5		✓	✓					

The summarized list of allergen homologues from the seven species of animals is listed in Table 4. This table provides the proportion of unigenes homologous to all matched known allergens and it provides insight to their abundance within the genome of each species. The three most abundant allergens were calcium binding protein, ovalbumin and paramyosin. Calcium-binding proteins contain a variable number of motifs, termed EF-hands, which are made up of two perpendicularly placed alpha-helics and an inter-helical loop forming a single calcium-binding site (Valenta *et al*., 1998). Its main function is to transport calcium as well as to interact with a variety of ligands in a calcium-dependent manner. After parvalbumin, a three EF-hand fish allergen, calcium-binding allergens were discovered in pollens of trees, grasses and weeds and, recently, as autoallergens in man (Valenta *et al*., 1998). Although only a small percentage of atopic individuals display IgE reactivity to calcium-binding allergens, these allergens may be important because of their ability to cross-sensitize allergic individuals. Ovalbumin (a form of albumin) is member of a class of water-soluble, heat-coagulating proteins. Other examples of albumins include myogen of muscle, serum albumin of blood, lactalbumin of milk, legumelin of peas, and leucosin of wheat. The main functions of albumin are protein transport and maintenance of homeostasis within the organism. Ovalbumin is a highly abundant (more than 50%) major allergenic protein of hen's egg white. Due to ovalbumin's resemblance with serum albumin, it was expected to observe high number of unigenes homologues to ovalbumin among the seven animal species. Paramyosin, like the tropomyosin and myosin, is a protein associated with associated with muscles. Both tropomyosin and paramyosin composed of two identical coiled-coil alphahelices (Tsai *et al*., 1999). Unlike tropomyosin, paramyosin is only found in invertebrates and the

distribution and functions of paramyosin are quite diverse (Tsai *et al*., 1999). The fact that many unigenes from the seven animal species were found to be homologues to paramyosin suggested that there are many variants or isoforms of paramyosin or tropomyosins present in vertebrates. Illustrated in Figure 3A – 3G are the pie-charts of the allergen homologues from individual animal species classified based on their biological function. Most of the allergens were involved in metabolism and homeostasis. Additionally, most of these allergen homologues were associated with stress response. Researchers have shown that certain allergic disorders such as eczema and asthma are regulated, in part, by hormones and brain chemicals released into the bloodstream in response to stress (Wright *et al*., 2005). For instance chaperons like heat shock proteins and thioredoxin play an important role in protein synthesis and in the protection of cellular structures during stress-related processes (Gruehn *et al*., 2003). Reduced and carboxymethylated bovine serum albumin (rcm-BSA) have been shown to elicit a stress response (Mifflin and Cohen, 1994). Expression of enzymes such as superoxide dismutases and phospholipases has also been shown to up-regulate during fungal infection (Rementeria *et al*., 2005). Another study done on *Escherichia coli* has demonstrated that temperature stress could cause the accumulation of stress proteins such as disaggregation chaperones (DnaK and ClpB), components of the RNA ribosomal proteins and enolases (Lethanh *et al*., 2005).

Table 4.
Summarized list allergen homologues from the seven species of animals

Allergen matched	Beef	Pork	Chicken	Goat	Sheep	Dog	Cat
21k allergen	2	0	0	0	1	0	0
60S acidic ribosomal protein	4	7	1	1	6	0	0
68 kDa allergen	1	2	0	0	0	0	0
98kDa HDM allergen	4	4	2	1	3	0	0
Actinidin	9	7	5	0	0	2	0
Aldehyde dehydrogenase	8	7	5	0	1	0	0
alkyl hydroperoxide reductase	3	2	1	0	0	0	0
allergen [Malassezia sympodialis]	8	8	5	1	14	1	1
allergen Bos d 2.0103	1	0	0	0	0	0	0
Allergen Bla g 4 precursor	0	0	0	0	5	0	0
allergen dI chain C2A	1	0	0	1	0	0	0
allergen Lep d 1.02	1	0	1	0	0	0	0
allergen A precursor	0	1	0	0	0	0	0
Allergen MAG29	1	1	0	0	0	1	0
Allergen Mal f 2	0	1	0	0	0	0	0
alpha-2u globulin	1	3	1	0	0	1	0
Apyrase precursor	1	0	0	0	0	0	0
Arginine kinase	4	4	3	0	0	4	0
Beta-lactoglobulin precursor	1	2	0	6	6	6	14
C07A4.3 [Caenorhabditis elegans]	0	1	0	0	0	0	0
C10G8.3 [Caenorhabditis elegans]	0	0	1	0	0	0	0
calcium ion binding protein	23	26	7	0	2	1	0
Der f 3 mite allergen	9	12	6	1	0	12	2
Enolase	3	4	3	0	4	2	0
Fatty acid-binding protein	9	10	9	2	2	0	0
Fel d 1	0	0	0	0	0	0	17
gelsolin-like allergen	4	3	2	0	0	0	0
Glutathione S-transferase	3	7	6	1	8	1	0
group 14 allergen protein	0	0	1	0	0	0	0
Heat shock protein	9	11	6	2	5	3	0
Hyaluronidase	2	1	1	0	1	0	0
major allergenic protein Mal f 4	0	0	3	0	1	0	0
lactalbumin, alpha	1	1	0	2	3	2	0
Lysozyme C precursor	1	0	0	0	0	6	0
major allergen BDA20	3	1	0	1	0	0	0
Major allergen Can f 1 precursor	1	1	0	0	0	3	0
Minor allergen Can f 2 precursor	0	0	0	0	0	4	0
Major allergen Equ c 1 precursor	1	2	0	0	0	0	0
major allergenic protein Mal f 4	2	2	0	0	0	0	0
Mite group 1 allergen Eur m 1 precursor	0	0	0	1	0	0	0
Ovalbumin	21	18	7	0	2	0	2
Ovomucoid precursor	3	5	7	0	0	5	0
Ovotransferrin precursor	2	7	2	2	1	0	1
paramyosin allergen	13	5	19	0	0	0	2
Parvalbumin beta	0	0	3	0	0	0	2
Phospholipase A1 1 precursor	2	1	2	0	1	6	0
preproalbumin	0	0	0	0	0	1	0
Putative peroxiredoxin pmp20	1	2	0	0	0	0	4
Serum albumin precursor	5	4	2	0	1	11	0
Superoxide dismutase [Mn]	1	1	1	0	1	0	0
T05A10.5 [Caenorhabditis elegans]	1	1	0	0	0	0	0
Thioredoxin	4	4	4	1	1	0	0
tri s 4 allergen	0	1	0	0	0	0	0
Tropomyosin	5	6	5	2	1	1	0
venom allergen	1	2	2	0	0	0	0
Total	**180**	**188**	**123**	**25**	**70**	**73**	**45**

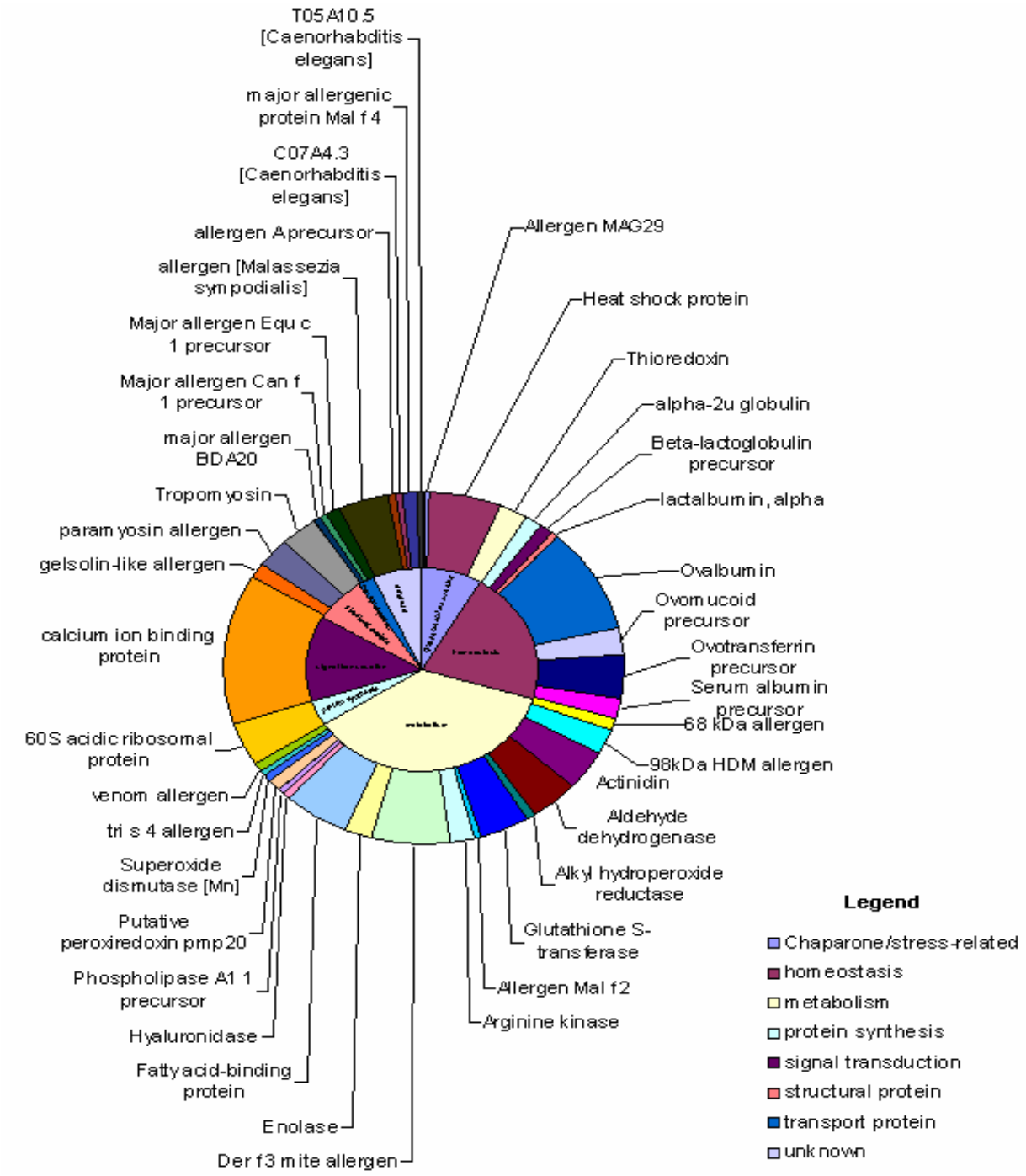


Figure 3A.
Pie chart of the allergen homologues from pig classified based on their biological function.

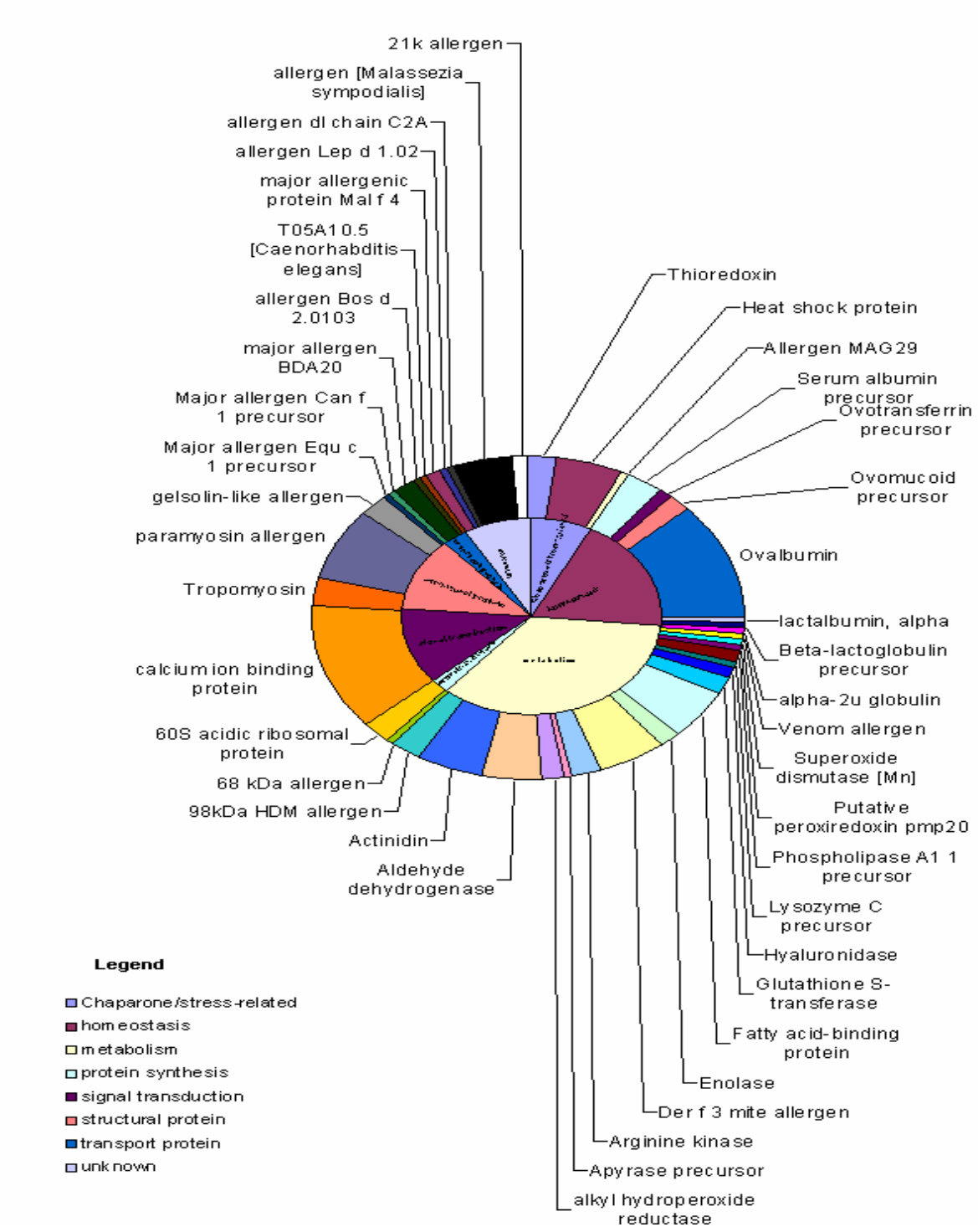


Figure 3B.
Pie chart of the allergen homologues from cow classified based on their biological function.

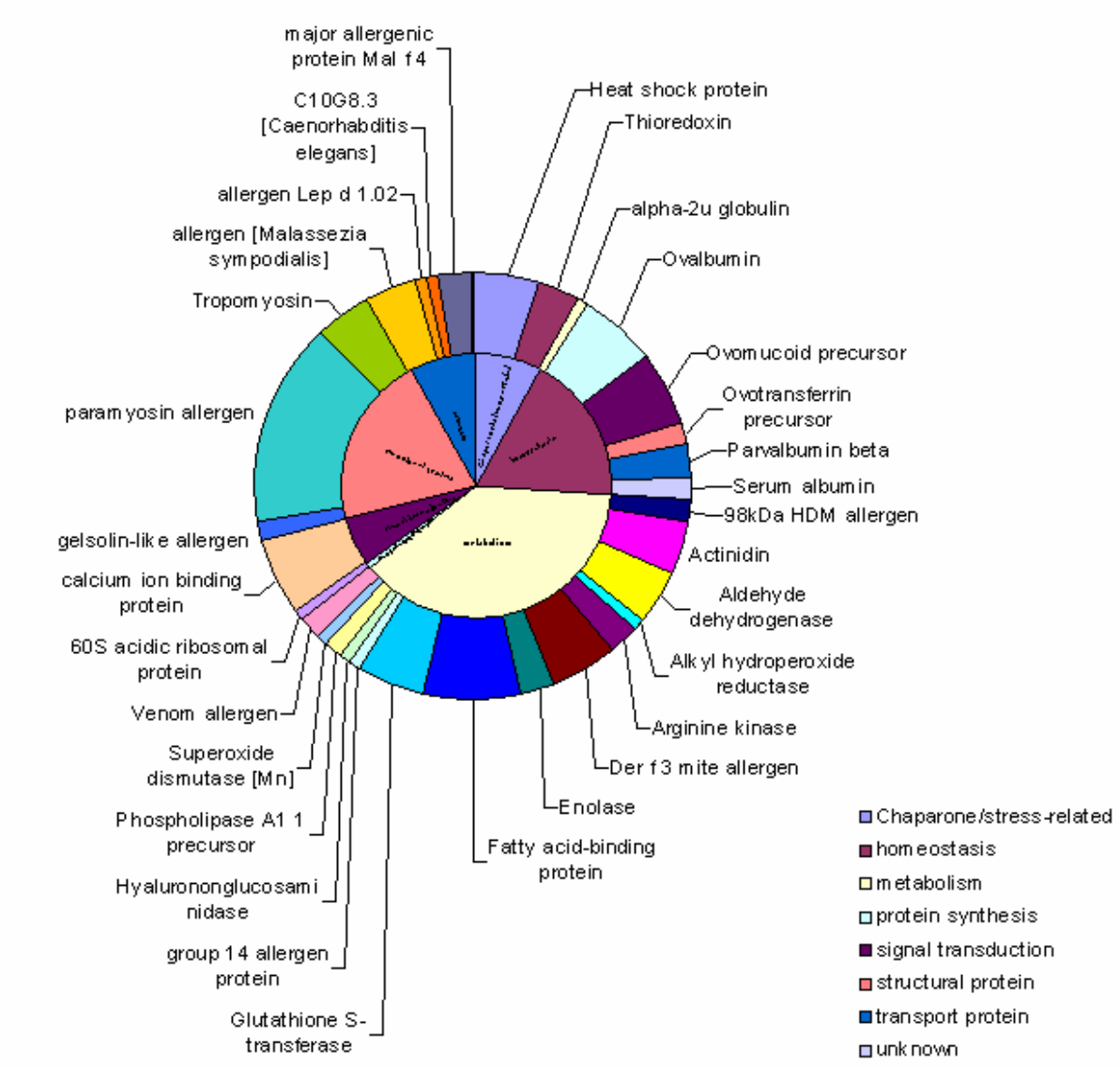


Figure 3C.
Pie chart of the allergen homologues from chicken classified based on their biological function.

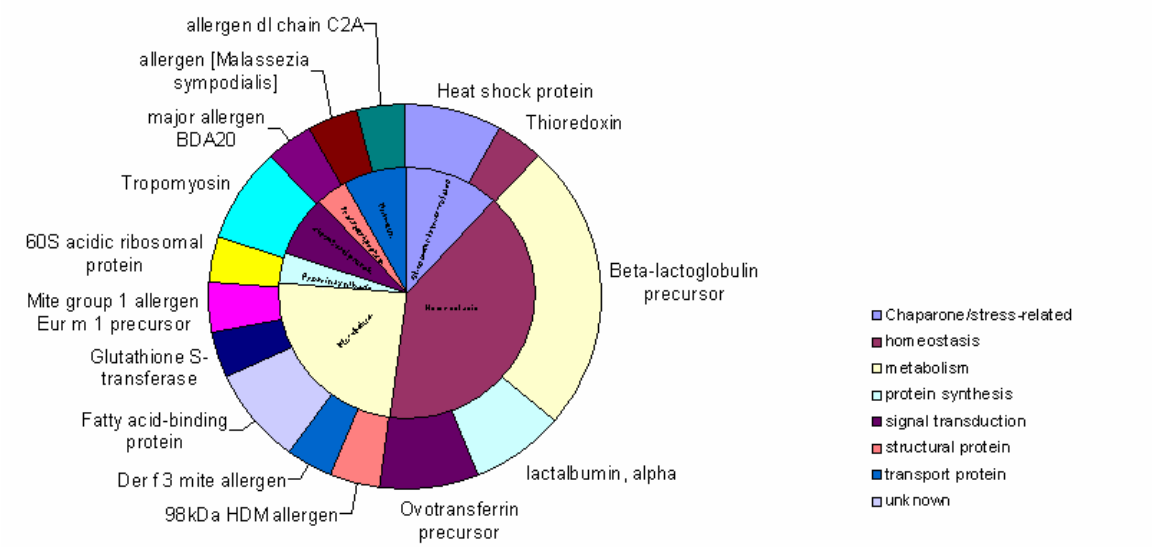


Figure 3D.
Pie chart of the allergen homologues from goat classified based on their biological function.

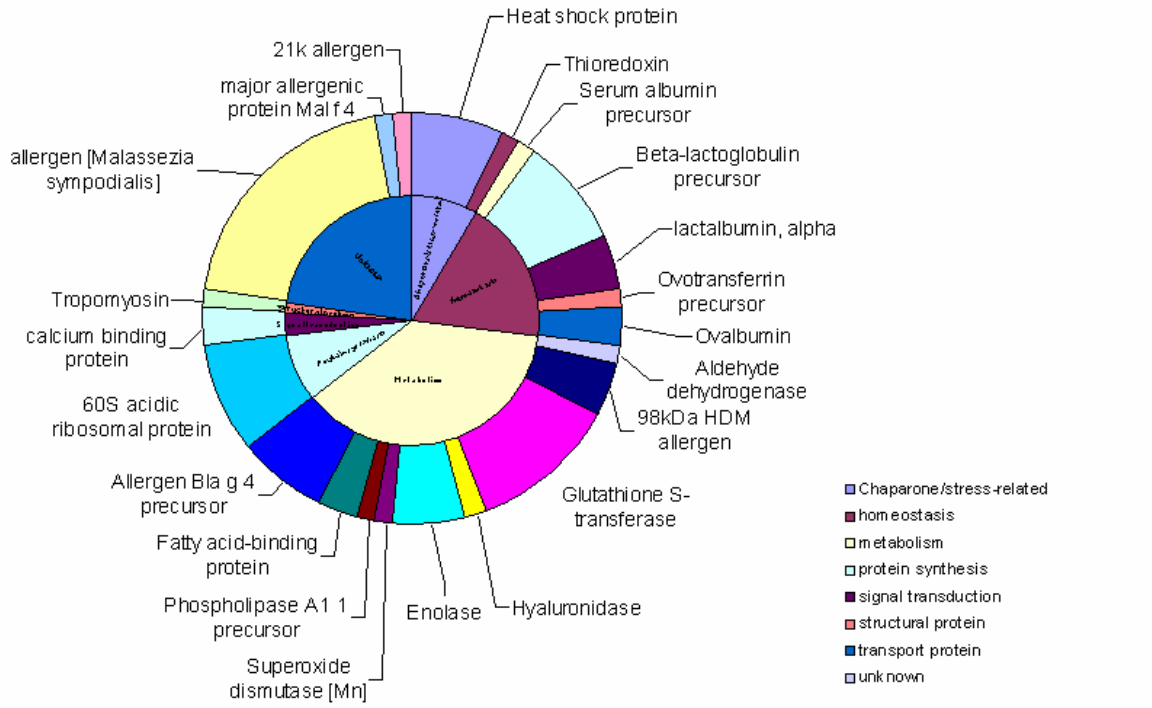


Figure 3E.
Pie chart of the allergen homologues from sheep classified based on their biological function

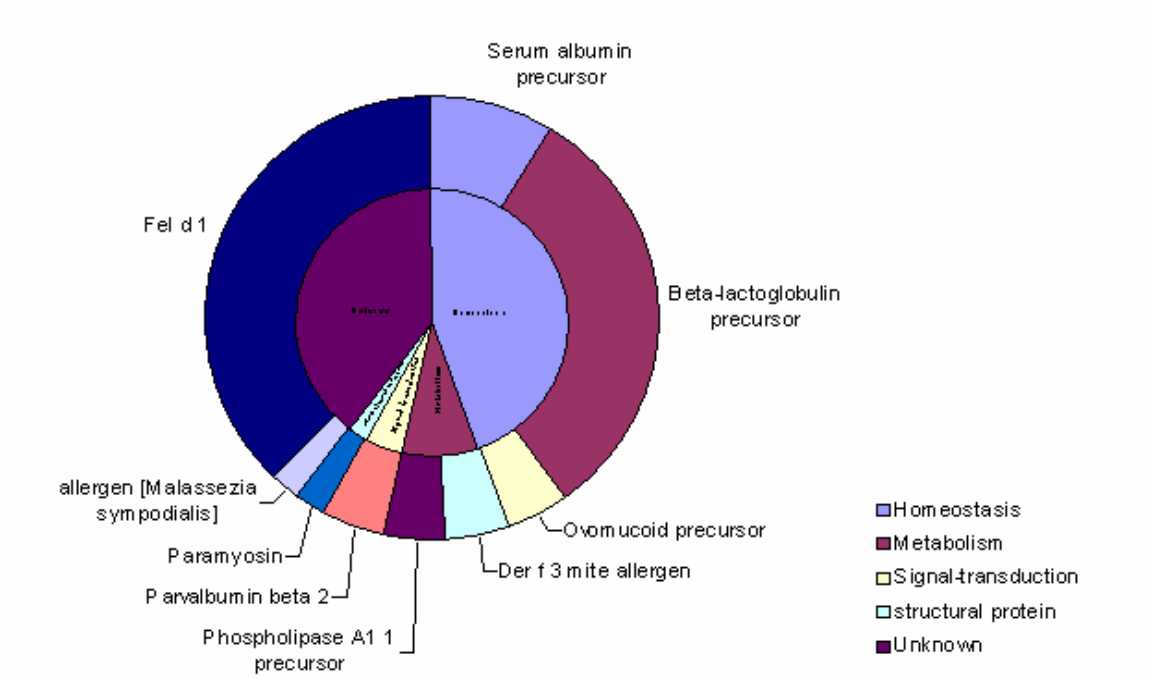


Figure 3F.
Pie chart of the allergen homologues from cat classified based on their biological function.

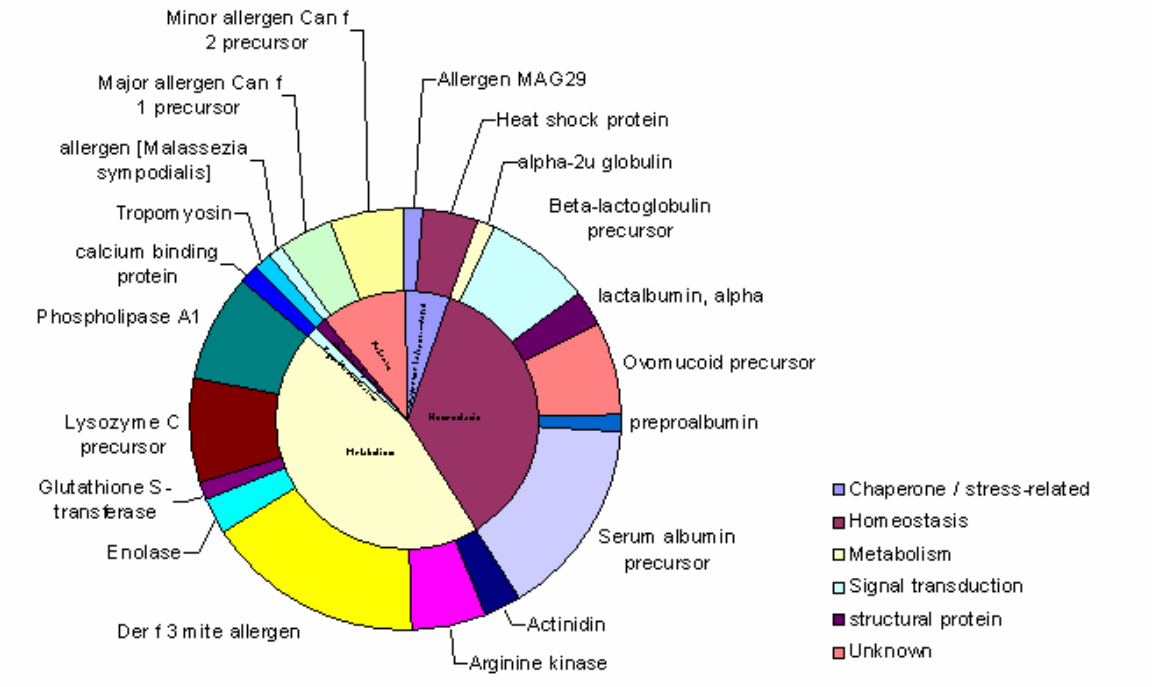


Figure 3G.
Pie chart of the allergen homologues from dog classified based on their biological function

3.3.1.2 Performance of allergen prediction by sequence homology

The recall rate of the known allergen in each individual species was used as an indication of the method's performance. For *B. taurus* (cow), there are currently 7 allergens (Bos d 2 to Bos d 8) listed on WHO/IUIS and the system was able to recall 5 of the allergens (71.4%). For *G. gallus* (chicken), there are 5 allergens (Gal d 1 to Gal d 5) listed and the system was able to recall 4 of the allergens (80%). For *C. familiaris* (dog), there are 4 allergens (Can f 1 to Can f 4) listed and 3 allergens were recalled (75%). For *F. domesticus* (cat), the number of successful recall was only 2 out of the listed 7 allergens (28%). This could be due to the quality and coverage of the sequences available on NCBI.

3.3.2 Allergenicity prediction using wavelet transform

This segment was done courtesy of Dr. Li Kuo Bin. All 664 reference allergens were used to predict putative allergen motifs. A total of 62 allergen motifs were predicted and listed in Table 5 (Adapted from Li *et al*., 2004).

Table 5.
Allergen motifs. The protein families were identified by using hmm search to search the Swiss-Prot with a profile HMM generated from the corresponding allergen motif.

Motif identifiers	Main protein families	Motif identifiers	Main protein families
motif-1	Profilin	motif-32	Lymphocyte antigen 75 precursor
motif-2	Glutenin low molecular weight	motif-33	Glutenin, high molecular weight
motif-3	Agglutinin	motif-34	Paramyosin
motif-4	Glutenin high molecular weight	motif-35	Glutenin, high molecular weight
motif-5	Glutenin high molecular weight	motif-36	Glutenin, high molecular weight
motif-6	Beta-expansin	motif-37	Enolase
motif-7	Major mite fecal allergen	motif-38	Glycinin
motif-8	Profilin	motif-39	Glycinin
motif-9	Beta-fructofuranosidas	motif-40	Gamma-gliadin B-I precursor
motif-10	Seed allergenic protein	motif-41	Calreticulin precursor
motif-11	Profilin	motif-42	Beta-conglycinin
motif-12	Aspartic protease inhibitor	motif-43	Glycinin precursor
motif-13	Tropomyosin	motif-44	Pollen allergen Amb a 5
motif-14	Tropomyosin; chlorophyll a-b binding proteins	motif-45	Parvalbumin
motif-15	Tropomyosin; heat shock protein	motif-46	Globin
motif-16	Gamma-gliadin	motif-47	Superoxide dismutase
motif-17	Profilin	motif-48	13S globulin seed storage protein 3 precursor
motif-18	Cysteine protease inhibitor	motif-49	Profilin
motif-19	Cysteine protease inhibitor	motif-50	Profilin
motif-20	Alpha/beta-gliadin	motif-51	Allergen Pha a 1 precursor
motif-21	Polygalacturonase	motif-52	Beta-expansin 1 precursor
motif-22	Gamma-gliadin precursor	motif-53	Allergen Hol l 1 precursor
motif-23	Polygalacturonase	motif-54	Patatin T5 precursor
motif-24	Polcalcin	motif-55	Keratin
motif-25	Patatin precursor	motif-56	Phospholipase A1
motif-26	Patatin precursor	motif-57	Lipid-transfer protein
motif-27	Aldehyde dehydrogenase	motif-58	Melittin precursor
motif-28	Hyaluronoglucosaminidase precursor	motif-59	Pathogenesis-related protein
motif-29	Alpha-amylase inhibitor	motif-60	Glutenin, low molecular weight
motif-30	13S globulin seed storage protein 3 precursor	motif-61	Allergen Amb a 1.1 precursor
motif-31	13S globulin seed storage protein 3 precursor	motif-62	Phospholipase A1

The downloaded unigenes, nucleotide and protein sequences from various species were put through the wavelet transform system for allergenic proteins prediction. The number of predicted allergens for each species varied from 34 proteins in beef to only 4 proteins in cat (Table 6). For example, tabulated in Table 7 is the list of putative allergens predicted in beef. The list was ranked based on the likelihood that the protein is allergenic as indicated by its E-values. The smaller is the E-value, the more likely the protein is allergenic. The potential allergens predicted from other species can be found in appendix X - XIV.

Table 6.
No. of putative allergens predicted for each species of animal using wavelet transform allergen prediction system.

Animal	No. of putative allergens
Beef	34
Chicken	34
Pork	22
Sheep	12
Dog	6
Cat	4

Table 7.
An example of the list of putative allergens predicted in beef using wavelet transform

Subject (hits)	Subject Description	Motif no.	E-values
P00711	Alpha-lactalbumin precursor (Lactose synthase B protein) (Allergen Bos d 4).	17	5.30E-36
Q28133	Allergen Bos d 2 precursor (Dander major allergen BDA20) (Dermal allergen BDA20).	31	5.30E-13
P02769	Serum albumin precursor (Allergen Bos d 6) (BSA).	34	6.00E-11
P42918	Calreticulin, brain isoform 2 precursor (CRP55) (Calregulin) (HACBP).	41	7.50E-08
P52193	Calreticulin, brain isoform 1 precursor (CRP55) (Calregulin) (HACBP).	41	1.00E-07
P19120	Heat shock cognate 71 kDa protein.	15	1.80E-07
Q28050	S100 calcium-binding protein A7 (Allergen Bos d 3) (Dander minor allergen BDA11) (Dermal allergen BDA11) (Calcium-binding protein in amniotic fluid 2) (CAAF2).	10	1.90E-07
P34933	HEAT SHOCK 70 KD PROTEIN 3.	15	1.30E-06
P02663	Alpha-S2 casein precursor [Contains: Casocidin-I].	16	1.80E-06
Q27975	HEAT SHOCK 70 KD PROTEIN 1 (HSP70-1).	15	6.70E-06
Q27965	HEAT SHOCK 70 KDA PROTEIN 2 (HSP70-2).	15	6.70E-06
P02754	Beta-lactoglobulin precursor (Beta-LG) (Allergen Bos d 5).	14	9.20E-06
P00792	Pepsin A precursor.	25	1.90E-05
P26882	40 kDa peptidyl-prolyl cis-trans isomerase (PPIase) (Rotamase) (Cyclophilin-40) (CYP-40) (Cyclophilin-related protein) (Estrogen receptor binding cyclophilin).	60	5.90E-05
Q09139	Fatty acid-binding protein, brain (B-FABP).	8	7.40E-05
Q9XSJ4	Alpha enolase (2-phospho-D-glycerate hydro-lyase) (Non-neural enolase) (NNE) (Enolase 1) (Phosphopyruvate hydratase).	42	9.80E-05
P04374	Peptidyl-prolyl cis-trans isomerase A (PPIase) (Rotamase) (Cyclophilin A) (Cyclosporin A-binding protein).	7	0.00011
P05786	Keratin, type II cytoskeletal 8 (Cytokeratin 8) (Cytokeratin A).	55	0.00021
P10790	Fatty acid-binding protein, heart (H-FABP) (Heart-type fatty acid-binding protein) (Mammary-derived growth inhibitor) (MDGI).	8	0.00023
P02662	Alpha-S1 casein precursor.	5	0.00044
P02668	Kappa casein precursor [Contains: Casoxin C; Casoxin 6; Casoxin A; Casoxin B; Casoplatelin].	12	0.0011
Q29443	Serotransferrin precursor (Transferrin) (Siderophilin) (Beta-1-metal binding globulin).	27	0.0011
P48820	Ran-binding protein 2 (RanBP2) (Nuclear pore complex protein Nup358) (Nucleoporin Nup358) (358 kDa nucleoporin) (P270).	7	0.0018
P02666	Beta casein precursor.	12	0.002
P48644	Aldehyde dehydrogenase 1A1 (Aldehyde dehydrogenase, cytosolic) (ALDH class 1) (ALHDII) (ALDH-E1).	28	0.0031
P25975	Cathepsin L precursor.	6	0.0037
P24627	Lactotransferrin precursor (Lactoferrin) [Contains: Lactoferricin B (Lfcin B)].	28	0.0041
P80209	Cathepsin D precursor.	25	0.0045
P80311	Peptidyl-prolyl cis-trans isomerase B precursor (PPIase) (Rotamase) (Cyclophilin B) (S-cyclophilin) (SCYLP).	60	0.0045
Q9BGI1	Peroxiredoxin 5, mitochondrial precursor (Prx-V).	14	0.0059
Q02373	NADH-ubiquinone oxidoreductase PDSW subunit (Complex I-PDSW) (CI-PDSW).	16	0.023
P09867	Heterogeneous nuclear ribonucleoprotein A1 (Helix-destabilizing protein) (Single-strand binding protein) (hnRNP core protein A1) (Unwinding protein 1) (UP1).	37	0.24
Q9TT96	Beta-1 adrenergic receptor (Beta-1 adrenoceptor) (Beta-1 adrenoreceptor).	51	0.26
P98107	E-selectin precursor (Endothelial leukocyte adhesion molecule 1) (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2).	3	0.51

3.3.2.1 Performance of allergenicity prediction using wavelet transform

For individual species, the recall of its own known allergens was relatively high again with the exception in cat. For *B. taurus* (cow), 6 out of the 7 (85%) listed allergens were recalled. Similarly, in the case of *G. gallus* (chicken), 4 out of the 5 (80%) listed allergens were recalled. For *C. familiaris* (dog), 3 out of the 4 (75%) listed allergens listed were recalled and in lastly for *F. domesticus* (cat) only 2 out of the listed 7 allergens (28%). Again this was due to the quality and coverage of the sequences available on NCBI.

3.3.3 Comparison between sequence homology based and motif-based allergen prediction system

In total, 197 distinct putative allergens were predicted by sequence homology from six species namely pig, cow, chicken, sheep, cat and dog. This number was obtained after filtering the repeated hits on the same allergen within each animal species. Similarly after filtering, the motif-based system predicted 90 distinct putative allergens from the same six animal species. The identities of the putative allergens predicted by both systems were quite dissimilar for each species (Table 8 - 13). The percentage of an allergen being predicted in both systems varies from 9.3 % in pig to 100 % in cat (Figure 4). On average, only about 15 % of the allergens were predicted in both systems with the exception in cat. The only possible reason for the difference in output of the prediction has to be attributed to the difference in mode and parameter of search because the input information was entirely the same for both systems. We know from literatures that allergen prediction based on sequence homology were prone to false positive results

(Kleter and Peijnenburg, 2002; Zorzet *et al*., 2002; Soeria-Atmadja *et al*., 2004). On the other hand, motif-based allergen prediction may be too stringent and prone to false negative results. Therefore, the two systems serve to complement each other to increase the confidence level of the prediction. With the results from the two systems, we are confident that those putative allergens predicted in both systems are likely real allergens. These proteins are namely heat shock proteins, serum albumins, enolases, and dehydrogenases across the six animal species. The rest of the putative allergens that were not predicted in both systems remind as allergen candidates until farther studies to confirm their allergenicity. Results from later studies using the proteomics approach will also serve as verification of the outcome in allergen prediction.

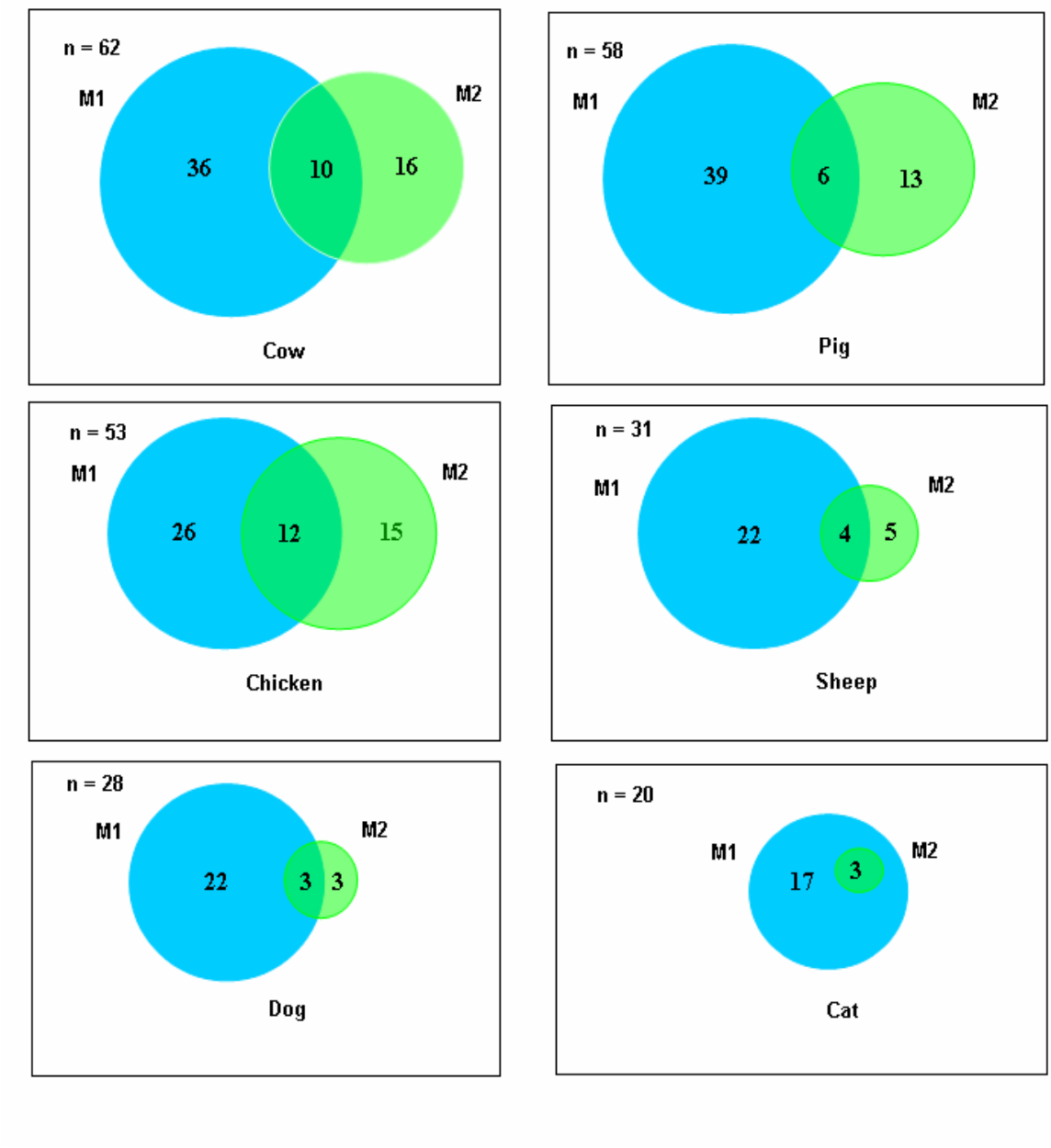


Figure 4.
Venn diagrams of putative allergens being predicted in both allergencity prediction systems for each animal species

Table 8. Comparison of predicted putative allergens in beef by both bioinformatics systems

Allergen	Method 1	Method 2
21k allergen [Anisakis simplex]	✓	
40 kDa peptidyl-prolyl cis-trans isomerase (Cyclophilin-related protein)		✓
60S acidic ribosomal proteins	✓	
68 kDa allergen [Penicillium chrysogenum]	✓	
98kDa HDM allergen [Dermatophagoides farinae]	✓	
Actinidain precursor (Actinidin) (Allergen Act c 1)	✓	
Aldehyde dehydrogenase (ALDDH)	✓	✓
Alkyl hydroperoxide reductases	✓	
Allergen [Malassezia sympodialis]	✓	
Allergen Bos d 2 precursor (Dander major allergen BDA20) (Dermal allergen BDA20).	✓	✓
Allergen dI chain C2A [Mus musculus]	✓	
Allergen Lep d 1.02	✓	
Allergen MAG29	✓	
Alpha-2u globulin PGCL1 [Rattus norvegicus]	✓	
Alpha-lactalbumin precursor (Lactose synthase B protein) (Allergen Bos d 4).	✓	✓
Alpha casein precursor.		✓
Apyrase precursor (Allergen Aed a 1)	✓	
Arginine kinase (AK) (Allergen Plo i 1)	✓	
Beta casein precursor.		✓
Beta-1 adrenergic receptor (Beta-1 adrenoceptor)		✓
Beta-lactoglobulin precursor (Beta-LG) (Allergen Bos d 5)	✓	✓
Calcium-binding allergens	✓	
Calreticulin		✓
Cathepsin precursors.		✓
Dust mite group 3 allergens	✓	
Enolase 1 (2-phosphoglycerate dehydratase 1) (2-phospho-D-glycerate hydro-lyase 1) (Allergen Hev b 9)	✓	✓
Enolase 2 (2-phosphoglycerate dehydratase 2) (2-phospho-D-glycerate hydro-lyase 2) (Allergen Hev b 9)	✓	
E-selectin precursor (Endothelial leukocyte adhesion molecule 1) (ELAM-1)		✓
Fatty acid-binding protein	✓	✓
Gelsolin-like allergen Der f 16 [Dermatophagoides farinae]	✓	
Glutathione S-transferase (GST class-mu) (Major allergen Der p 8) (P dp 15)	✓	
Heat shock 70 KD protein	✓	✓
Heat shock protein 90 (Allergen Asp f 12)	✓	
Heterogeneous nuclear ribonucleoprotein A1 (Helix-destabilizing protein)		✓
Hyaluronon glucosaminidase precursor (Hyaluronidase)	✓	
Kappa casein precursor [Contains: Casoxin C; Casoxin 6; Casoxin A; Casoxin B; Casoplatelin].		✓
Keratin, type II cytoskeletal 8 (Cytokeratin 8) (Cytokeratin A).		✓
Lactotransferrin precursor (Lactoferrin) [Contains: Lactoferricin B (Lfcin B)].		✓
Lysozyme C precursor (1,4-beta-N-acetylmuramidase C) (Allergen Gal d 4) (Gal d IV)	✓	
Major allergen BDA20 [Bos taurus]	✓	
Major allergen Can f 1 precursor (Allergen Dog 1)	✓	
Major allergen Equ c 1 precursor	✓	
major allergenic protein Mal f4 [Malassezia furfur]	✓	
NADH-ubiquinone oxidoreductase PDSW subunit (Complex I-PDSW) (CI-PDSW).		✓
Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	✓	
Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	✓	
Paramyosin allergen [Blomia tropicalis]	✓	
Paramyosin-like allergen [Dermatophagoides farinae]	✓	
Pepsin A precursor.		✓
Peptidyl-prolyl cis-trans isomerase (Cyclophilin-related protein)		✓
Peroxiredoxin 5, mitochondrial precursor (Prx-V).		✓
Phospholipase A1 1 precursor (Allergen Dol m 1.01) (Dol m I)	✓	
Putative peroxiredoxin pmp20 (Thioredoxin reductase) (Peroxisomal membrane protein pmp20) (Allergen Asp f 3)	✓	
Ran-binding protein 2 (RanBP2)		✓
S100 calcium-binding protein A7 (psoriasin 1) [Bos taurus]	✓	✓
Serum transferrin	✓	✓
Serum albumin precursor (Allergen Bos d 6) (BSA)	✓	✓
Superoxide dismutase [Mn], mitochondrial precursor (Allergen Asp f 6)	✓	
T05A10.5 [Caenorhabditis elegans]	✓	
Thioredoxin	✓	
Tropomyosins	✓	
Venom allergen 5.01 precursor	✓	

Table 9. Comparison of predicted putative allergens in pork by both bioinformatics systems

Allergen	Method 1	Method 2
60S acidic ribosomal proteins	✓	
68 kDa allergen [Penicillium chrysogenum]	✓	
98kDa HDM allergen [Dermatophagoides farinae]	✓	
Actinidain precursor (Actinidin) (Allergen Act c 1)	✓	
Aldehyde dehydrogenase (ALDDH)	✓	
Alkyl hydroperoxide reductase/ Thiol specific antioxidant/ Mal allergen [Nitrosomonas europaea ATCC 19718]	✓	
Allergen [Malassezia sympodialis]	✓	
Allergen A precursor [Psoroptes ovis]	✓	
Allergen Bla g 5 [Blattella germanica]	✓	
Allergen MAG29	✓	
Allergen Mal f 2 (MF1)	✓	
Allergen Pen m 2 [Penaeus monodon]	✓	
Alpha-2-HS-glycoprotein precursor (Fetuin-A).		✓
Alpha-2u globulin PGCL1 [Rattus norvegicus]	✓	✓
Alpha-lactalbumin precursor	✓	
Arginine kinase (AK) (Allergen Plo i 1)	✓	
Beta-lactoglobulin precursor (Beta-LG) (Allergen Bos d 5)	✓	
C07A4.3 [Caenorhabditis elegans]	✓	
Calcium-binding allergen Ole e 8 (PCA18/PCA23)	✓	
Cathepsin L precursor.		✓
Cytochrome P450 3A29 (CYPIIIA29).		✓
Dehydrogenase/reductase SDR family member 4		✓
Dust mite group 3 allergens	✓	
Enolase 2 (2-phosphoglycerate dehydratase 2) (2-phospho-D-glycerate hydro-lyase 2) (Allergen Hev b 9).	✓	
Fatty acid-binding proteins	✓	✓
Fumarate hydratase, mitochondrial (Fumarase).		✓
Gelsolin-like allergen Der f 16 [Dermatophagoides farinae]	✓	
Glutathione S-transferases	✓	
Heat shock 70 KD proteins	✓	✓
Heat shock protein 90 (Allergen Asp f 12)	✓	✓
Hyalurononglucosaminidase precursor (Hyaluronidase) (Allergen Pol a 2)	✓	
Inhibitor of carbonic anhydrase precursor.		✓
Lactotransferrin precursor (Lactoferrin).		✓
Major allergen BDA20 [Bos taurus]	✓	
Major allergen Can f 1 precursor (Allergen Dog 1)	✓	
Major allergen Equ c 1 precursor	✓	
Major allergenic protein Mal f4 [Malassezia furfur]	✓	
Malate dehydrogenase, mitochondrial precursor.		✓
Membrane associated progesterone receptor component 1.		✓
Myosin heavy chain, cardiac muscle beta isoform (MyHC-beta).		✓
NADP-dependent malic enzyme (NADP-ME) (Malic enzyme 1).		✓
Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	✓	
Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	✓	
Serum transferrin	✓	
Paramyosin allergen [Blomia tropicalis]	✓	
Paramyosin-like allergen [Dermatophagoides farinae]	✓	
Pepsin A precursor.		✓
Phospholipase A1 1 precursor (Allergen Dol m 1.01) (Dol m I)	✓	
Putative peroxiredoxin pmp20 (Thioredoxin reductase) (Allergen Asp f 3)	✓	
S100 calcium-binding protein A7 (psoriasin 1) [Bos taurus]	✓	
Serum albumin precursor.	✓	✓
Superoxide dismutase [Mn], mitochondrial precursor (Allergen Asp f 6)	✓	
T05A10.5 [Caenorhabditis elegans]	✓	
Thioredoxin (Allergen Cop c 2)	✓	
Tri s 4 allergen [Trichophyton schoenleinii]	✓	
Tropomyosins	✓	✓
venom allergen 5, LONg family member (lon-1) [Caenorhabditis elegans].	✓	
Von Ebner's gland protein precursor (VEG protein) (Lipocalin-1).		✓

Table 10. Comparison of predicted putative allergens in chicken by both bioinformatics systems

Allergen	Method 1	Method 2
40S ribosomal protein S27a.		✓
60S acidic ribosomal protein P1 (Allergen Alt a 12) (Alt a XII).	✓	
98kDa HDM allergen [Dermatophagoides farinae]	✓	
Actinidain precursor (Actinidin) (Allergen Act c 1)	✓	
Aldehyde dehydrogenase	✓	✓
Alkyl hydroperoxide reductase/ Thiol specific antioxidant/ Mal allergen [Nitrosomonas europaea ATCC 19718]	✓	
Allergen [Malassezia sympodialis]	✓	
Allergen Bla g 5 [Blattella germanica]	✓	
Allergen Lep d 1.02	✓	
Allergen Pen m 2 [Penaeus monodon]	✓	
Alpha-2u globulin PGCL1 [Rattus norvegicus]	✓	
Arginine kinase (AK) (Allergen Plo i 1)	✓	
Bromodomain adjacent to zinc finger domain 2B (Extracellular matrix protein F22).		✓
C10G8.3 [Caenorhabditis elegans]	✓	
Calcium-binding allergens	✓	
Calretinin (CR).		✓
Chromosome-associated kinesin KIF4A (Chromokinesin).		✓
Der f 3 mite allergen	✓	
DNA topoisomerase II, alpha isozyme.		✓
Enolase 1 (2-phosphoglycerate dehydratase 1) (2-phospho-D-glycerate hydro-lyase 1) (Allergen Hev b 9)	✓	✓
Enolase 2 (2-phosphoglycerate dehydratase 2) (2-phospho-D-glycerate hydro-lyase 2) (Allergen Hev b 9)	✓	✓
Fatty acid-binding proteins	✓	✓
Gamma enolase (2-phospho-D-glycerate hydro-lyase) (Neural enolase) (NSE).		✓
Gelsolin-like allergen Der f 16 [Dermatophagoides farinae]	✓	
GENE X PROTEIN (OVALBUMIN-RELATED).		✓
Gene Y protein (Ovalbumin-related).		✓
Glutathione S-transferases	✓	
Group 14 allergen protein [Dermatophagoides pteronyssinus]	✓	
Heat shock 70 kDa proteins	✓	✓
Heat shock protein 90	✓	✓
Hyaluronon glucosaminidase (Hyaluronidase) (Allergen Dol m 2) (Dol m II)	✓	
Keratin, type II cytoskeletal cochleal (Cytokeratin otokeratin).		✓
Major allergenic protein Mal f4 [Malassezia furfur]	✓	
Dust mite group 3 allergens	✓	
Myosin heavy chain		✓
Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	✓	✓
Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	✓	✓
Ovotransferrin precursor (Conalbumin) (Allergen Gal d 3) (Gal d III) (Serum transferrin)	✓	✓
Paramyosin allergen [Blomia tropicalis]	✓	
Paramyosin-like allergen [Dermatophagoides farinae]	✓	
Parvalbumins	✓	✓
Pepsin A precursor.		✓
Peptidyl-prolyl cis-trans isomerase B precursor (PPIase) Cyclophilin-related protein		✓
Phospholipase A1 precursor	✓	
Restin (Cytoplasmic linker protein-170) (CLIP-170).		✓
S100 calcium-binding protein A7 (psoriasin 1) [Bos taurus]	✓	
Serum albumin precursor	✓	✓
Superoxide dismutase [Mn], mitochondrial precursor (Allergen Asp f 6)	✓	
Thioredoxin (Allergen Cop c 2)	✓	
Triosephosphate isomerase (TIM) (Triose-phosphate isomerase).		✓
Tropomyosins	✓	✓
Venom allergens	✓	
Vitellogenin II precursor (Major vitellogenin)		✓

Table 11. Comparison of predicted putative allergens in sheep by both bioinformatics systems

Allergen	Method 1	Method 2
21k allergen like family member (5F762) [Caenorhabditis elegans]	✓	
60S acidic ribosomal protein P1 (Allergen Alt a 12) (Alt a XII)	✓	
60S acidic ribosomal protein P1 (Allergen Cla h 12) (Cla h XII)	✓	
60S acidic ribosomal protein P2 (Minor allergen Alt a 6) (Alt a VI)	✓	
60S acidic ribosomal protein P2 (Minor allergen Fus c 1)	✓	
98kDa HDM allergen [Dermatophagoides farinae]	✓	
Aldehyde dehydrogenase (ALDDH)	✓	✓
allergen [Malassezia sympodialis]	✓	
Allergen Bla g 4 precursor (Bla g IV)	✓	
Alpha casein precursors		✓
Beta casein precursor.		✓
Beta-lactoglobulin precursor	✓	✓
Enolase 2 (2-phosphoglycerate dehydratase 2) (2-phospho-D-glycerate hydro-lyase 2) (Allergen Hev b 9)	✓	
Fatty acid-binding proteins	✓	
Glutathione S-transferase	✓	
Heat shock 70 kDa protein (Allergen Cla h 4) (Cla h IV)	✓	
Heat shock protein HSP1 (65 kDa IgE-binding protein) (Allergen Asp f 12)	✓	
Hyaluronoglucosaminidase precursor (Hyaluronidase) (Allergen Api m 2)(Api m II)	✓	
Interleukin-3 precursor (IL-3) (Mast-cell growth factor) (MCGF).		✓
Keratin, type II		✓
Alpha, lactalbumin [Bos taurus]	✓	✓
major allergenic protein Mal f4 [Malassezia furfur]	✓	
Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	✓	
Ovotransferrin precursor (Conalbumin) (Allergen Gal d 3) (Gal d III) (Serum transferrin)	✓	
Phospholipase A1 1 precursor (Allergen Dol m 1.01) (Dol m I)	✓	
S100 calcium-binding protein A7 (psoriasin 1) (Dermal allergen BDA11)	✓	
Serum albumin precursor.	✓	✓
Superoxide dismutase [Mn], mitochondrial precursor (Allergen Asp f 6)	✓	
Thioredoxin (Allergen Cop c 2)	✓	
Trichohyalin.		✓
Tropomyosin	✓	

Table 12. Comparison of predicted putative allergens in dog by both bioinformatics systems

Allergen	Method 1	Method 2
Actinidain precursor (Actinidin) (Allergen Act c 1)	✓	
Allergen [Malassezia sympodialis]	✓	
Allergen MAG29	✓	
Allergen Pen m 2 [Penaeus monodon]	✓	
Alpha-2u globulin PGCL1 [Rattus norvegicus]	✓	
Arginine kinase (AK) (Allergen Plo i 1)	✓	
Beta-lactoglobulin precursor (Beta-LG) (Allergen Bos d 5)	✓	
Calnexin precursor (pp90)		✓
Der f 3 mite allergen	✓	
Enolase (2-phosphoglycerate dehydratase) (2-phospho-D-glycerate hydro-lyase) (Allergen Rho m 1)	✓	
Glutathione S-transferase (GST class-mu) (Major allergen Der p 8) (P dp 15)	✓	
Heat shock 70 kDa protein (Allergen Cla h 4) (Cla h IV)	✓	
Alpha, lactalbumin [Bos taurus]	✓	✓
Lysozyme C precursor (1,4-beta-N-acetylmuramidase C) (Allergen Gal d 4) (Gal d IV)	✓	
Major allergen Can f 1 precursor (Allergen Dog 1)	✓	✓
Minor allergen Can f 2 precursor (Allergen Dog 2)	✓	
Mite allergen Der f 3 precursor (Der f III)	✓	
Mite allergen Der p 3 precursor (Der p III).	✓	
Mite allergen Eur m 3 precursor	✓	
Occludin		✓
Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	✓	
Phospholipase A1 (Allergen Ves m 1) (Ves m I)	✓	
Phospholipase A1 precursor (Allergen Ves v 1) (Ves v I)	✓	
Preproalbumin (serum albumin) [Gallus gallus]	✓	
S100 calcium-binding protein A7 (psoriasin 1) [Bos taurus]	✓	
Serum albumin precursor (Allergen Can f 3)	✓	✓
Sex-determining region Y protein (Testis-determining factor)		✓
Tropomyosins	✓	

Table 13. Comparison of predicted putative allergens in cat by both bioinformatics systems

Allergen	Method 1	Method 2
Allergen [Malassezia sympodialis]	✓	
Beta-lactoglobulin precursor (Beta-LG) (Allergen Bos d 5)	✓	
Chain A, Crystal Structure Of Fel d 1	✓	
Der f 3 mite allergen	✓	
Major allergen chain 1 precursor A - cat	✓	
Major allergen chain 1 precursor B - cat	✓	
Major allergen chain 2 precursor, short form - cat	✓	
Major allergen Fel dI chain 1 long form precursor - cat	✓	
Major allergen I	✓	
Major allergen I polypeptide chain 1 major form precursor (Allergen Fel d 1-A)	✓	✓
Major allergen I polypeptide chain 2 precursor (Allergen Fel d 1-B) (Fel d I-B)	✓	✓
Major cat allergen Fel d I beta chain - cat (fragment)	✓	
Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	✓	
Paramyosin (Allergen Ani s 2)	✓	
Parvalbumin beta 2 (Major allergen Sal s 1)	✓	
Phospholipase A1 1 precursor (Allergen Dol m 1.01) (Dol m I)	✓	
Serum albumin precursor (Allergen Fel d 2)	✓	✓

3.4 Conclusion

Many allergen-homologous unigenes predicted based on sequence homology were found in each of eight animal species (pig, cow, chicken, goat, sheep, trout, dog and cat). Some of these are known allergens in the species concerned, but many more are putative new allergens. Many of these allergen-homologous unigenes are homologous to the same allergens across the eight animal species, implying that there is a possibility of cross-reactivity among the different species, by means of pan-allergens. The motif-based system also predicted many putative allergens from six animal species (pig, cow, chicken, sheep, dog and cat). However, the lists of putative allergens predicted by both systems were quite varied. Nevertheless, both the systems have achieved high recall rates for the known allergens with each animal species. The two allergenicity predictive systems serve to complement each other and those putative allergens predicted in both systems are likely ‘true’ allergens. The detailed list of predicted homologous allergen equipped with regions of sequence homology could be utilized for further research such as recombinant protein production and epitope mapping.

CHAPTER 4: IDENTIFICATION AND CHARACTERIZATION OF MEAT-BASED ALLERGENS USING A PROTEOMIC APPROACH

4.1 INTRODUCTION

Food allergies of Type-I hypersensitivity are IgE antibodies mediated and caused by certain proteins or glycoproteins, which are called food allergens. An analytical marker of allergens is the IgE-reactivity to these substances (Becker and Reese, 2001). The identification of food allergens is a priority in the management of food allergy, because of the need to obtain standardized extracts and pure allergens for diagnostic and therapeutic purposes (Pastorello and Trambaioli, 2001). It is thus important to develop methods for allergen extraction, separation and immunological detection methods to identify and characterize individual food allergens with minimum manipulation.

4.1.1 Protein extraction from food sources

Since a lot of different food sources are studied for their allergenic properties the first requirement for correct research procedure is to obtain good protein extracts, that is starting material with sufficiently high protein concentration and suitably low lipid and sugar content so as to allow for better protein separation (Pastorello and Trambaioli, 2001). The optimization of the extraction procedure depends on the nature of source

material (e.g. lipids and carbohydrates content) and purpose (e.g. protease activity). The use of different extraction buffers will influence the outcome of protein extraction. For instance, the pH of the extraction buffer greatly influence the extract composition and antigen yield (Niemeijer *et al*., 1996). Buffers with solid polyvinylpolypyrrolidone (PVPP) and 2 m*M* ethyl-enediaminetetraacetic acid disodium salt (EDTA) added, inhibited reactions between protein and phenolic compounds, especially for plant food (Loomis, 1974). Some studies have focused on evaluating the effects on protein separation by various protein extraction and sample preparation methods (Carpentier *et al*., 2005; Natarajan *et al*., 2005; Saravanan and Rose, 2004). The results suggested that trichloroacetic acid (TCA) extraction method is one of the better methods for protein extraction.

In the case of animal foods and/or meats, good extracts have often been achieved with simple grinding and incubation of the food in Phosphate Buffered Saline (PBS) solution to extract proteins contained in the raw materials (Asero *et al*., 1997; Llatser *et al*., 1998; Pastorello and Trambaioli, 2001).

4.1.2 Methods used for protein separation and allergen isolation

At present the most widely-used technique to identify allergenic molecules in raw extracts is one- and two-dimensional electrophoretic separation in sodium dodecylsulphate polyacrylamide gel (SDS-PAGE) followed by immunoblotting with allergenic patients' serum. This method has been used for almost all allergenic proteins

reported in publications (Pastorello and Trambaioli, 2001). In some cases the allergenic proteins are directly eluted from the gel. This one-step purification protocol was used for isolation of Mal d 1, a major apple allergen (Vieths et al., 1995). Besides SDS-PAGE coupled with immunoblotting, chromatographic techniques are also widely used in allergen isolation and identification. Chromatographic techniques include ion-exchange (anionic and cationic), gel-flitration, reversed-phase, hydrophobic interaction, and affinity chromatography. These techniques are usually used in tandem for better purity and increased specificity of the desired allergenic protein. For instance, the egg white allergens, parvalbumin from fish, lipid transfer proteins (LTPs) from various plants, were isolated by ion-exchange chromatography followed by gel filtration (Aas and jebsen, 1967; Elsayed and Aas, 1971; Anet *et al*., 1985; Hoffman, 1983; Ebbehøj *et al*., 1995; Pastorello *et al*., 1999; Pastorello *et al*., 2000; Pastorello *et al*., 1999). The famous profiling, a panallergen, was purified by affinity chromatography using a poly(L-proline) column (Valenta et al., 1991; Lindberg *et al*., 1988).

4.1.3 Protein identification using Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS)

Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS), was first introduced in 1988 by Tanaka *et al*. and independently by Hillenkamp and Karas. For the past decade, it has become a widely accepted and versatile method for post-separation protein identification. Its ability to desorb high-molecular-weight thermolabile molecules, its high accuracy and sensitivity, combined with its wide mass range (1–300 kDa), make

MALDI-TOF MS a promising tool for the identification of biomolecules in complex samples, including peptides, proteins, oligosaccharides and oligonucleotides (Marvin *et al*., 2003).

The general principle of MALDI-Tof-MS revolves around the rapid photo-volatilization of a sample embedded in a UV-absorbing matrix followed by time-of-flight mass spectrum analysis (Figure 1).

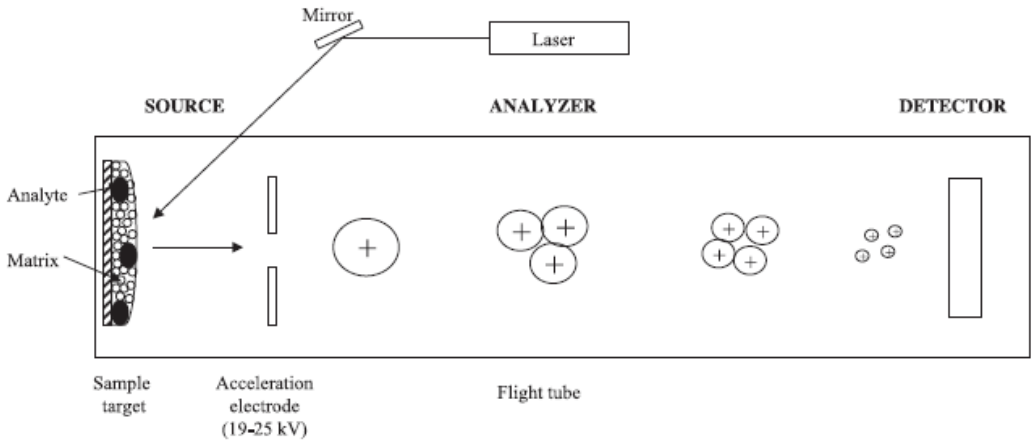


Figure 1.
Principle of matrix-assisted laser desorption/ionization mass spectrometry. The analyte mixed with a saturated matrix solution forms crystals. The irradiation of this mixture by the laser induces the ionization of the matrix, desorption, transfer of protons from photo-excited matrix to analyte to form a protonated molecule (adapted from Marvin *et al*., 2003).

MALDI-TOF-MS identification of proteins is carried out by the peptide mass mapping or peptide mass fingerprinting technique. This protein identification is based on the accurate mass measurement of a group of peptides derived from a protein by sequence-specific proteolysis (Marvin *et al*., 2003). After proteolysis with a specific protease (e.g. trypsin), proteins of different amino acid sequence produce a series of peptides masses, which are subsequently detected by sensitive ion detectors. The spectrum of identified peptide masses is unique for individual protein and is known as a mass fingerprint. Searching the selected masses from the fingerprint against protein sequence databases (e.g. NCBI and SwissProt- TrEMBL) using software (e.g. Mascot and MS-fit) enables the identification of most proteins. Many proteomic data have been obtained using a combination of two-dimensional polyacrylamide gel (2-D PAGE) and liquid chromatography (LC) with MALDI-TOF-MS. Recently, the technology of MS has improved to include tandem MS (MALDI-TOF-MS/MS), which provides actual amino acid sequences of the peptide fragments. This greatly increases the accuracy and specificity of protein identification.

4.2 MATERIALS AND METHODS

4.2.1 Patients and sera

The sera used in this study were selected from the previously pre-screened population mentioned in section 2.2.6. They were selected based on their IgE reactivity to the meat antigens and also the volume available for downstream studies. These sera were also confirmed by ELISA for their IgE reactivity towards the meat antigens.

4.2.2 Protein extraction

The crude extracts of pork, beef and lamb were extracted as described in section 2.2.3. Total protein concentration was then determined using the Bio-Rad protein assay kit (Bio-Rad, USA). TCA protein extraction method was also used in this study. Briefly, 10 ml of TCA-acetone (90% acetone, 10% TCA, 0.07% DTT) was added to the tissue powder, which was grinded with liquid nitrogen. Samples were vortexed, incubated at –20°C for 1 h, and centrifuged at 15,000 *g* for 15 min at 4°C. Protein pellets were washed 3 times with acetone/DTT and subsequently lypholized and dissolved in lysis buffer.

4.2.3 Gel Electrophoresis

Meat extracts were separated by electrophoresis according to the protocols for 40% acrylamide/bis solutions, 37.5:1 (Bio-Rad, USA) given by Laemmli (1970). A total of

35µg of sample was loaded with sample buffer (62.5mM Tris-HCl, pH 6.8, 2% SDS, 10% glycerol, 5% β-mercaptoethanol, and 0.01% bromophenol blue). For 2D electrophoresis, the first-dimension IEF was performed using either 7-cm or 17-cm IPG strips (Bio-Rad ReadyStrip IPG strips, pH 3-10) in the PROTEAN IEF System (Bio-Rad, USA). The 7-cm IPG strips were passively rehydrated for 16 hrs with 125µl of rehydration buffer (9.5M urea, 4% CHAPS, 100 mM dithiothreitol, 0.2% bio-lytes pH 3–10, 0.001% bromophenol blue) containing 75µg of protein whereas the 17-cm IPG strips were rehydrated with 350 µl of rehydration buffer containing 500 µg of protein. The IEF settings for 7-cm strips were 250 V rapid for 15 min, 8000 V gradient for 1 h, and 8000 V rapid to a total of 20 kVh. The IEF settings for 17-cm strips were 250 V rapid for 1 hr, 4000 V gradient for 3 hr, and 8000 V rapid to a total of 60 kVh. The focused strips were incubated with equilibration buffer 1 (50 mM Tris–HCl [pH 8.8], 6 M urea, 20% glycerol, 2% SDS, 0.002% bromophenol blue, 1% DTT) and equilibration buffer 2 (50 mM Tris–HCl [pH 8.8], 6 M urea, 20% glycerol, 2% SDS, 0.002% bromophenol blue, 2.5% iodoacetamide) for 15 min each and subsequently placed onto 12% polyacrylamide gel. The strips were overlaid with agarose sealing solution (0.25M trisbase, 1.92 M glycine, 1% SDS, 1% agarose, 0.002% bromophenol blue). The electrophoresis was performed either using the Mini PROTEAN 3 cell unit or the PROTEAN II xi system (Bio-Rad, USA). The Bio-Rad 10 to 250kDa standard was used as molecular weight makers.

4.2.4 Protein visualization and image analysis

Proteins were visualized either by Coomassie brilliant blue staining or by silver staining. For Coomassie blue staining, the gels were stained overnight in staining solution (20% methanol, 10% acetic acid, and 0.1% w/v Coomassie brilliant blue G-250 (Merck KGaA, Germany) and destained with destaining solution (10% methanol and 10% acetic acid). For silver staining, the gels were fixed for 1 hr in a fixing solution (50% methanol, 10% acedic acid), washed for 20 min in 30% ethanol and 20 min in Milli-Q water. After washing the gels were sensitized in 0.02% $Na_2S_2O_3$ for 1 min and incubated in chilled silver stain (0.2% $AgNO_3$, 0.02% formaldehyde) for 20 min. The gels were then washed with Milli-Q water and developed in developing solution (3% Na_2CO_3, 0.0005% $Na_2S_2O_3$, 0.05% formaldehyde) and stopped with 0.5% glycine solution. Stained gels were scanned and calibrated with Bio-RAD GS-700 Imaging Densitometer (Bio-Rad, USA). Image analysis was performed with PDQuest 2-D Analysis Software (Bio-Rad, USA). Spot detection was realized without spot editing. Spot matching for replicate gels consistency and spot identification on western blots was also done as shown in Figure 2.

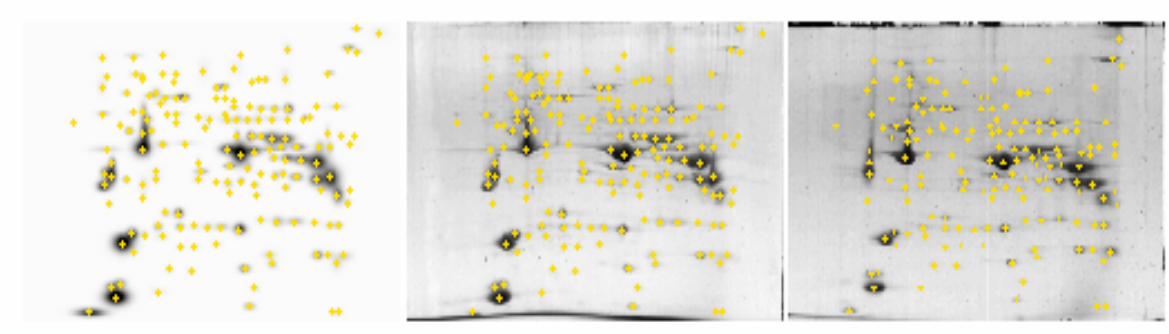


Figure 2.
Process of spots matching using the Bio-rad PDQuest software

4.2.5 Western Blotting Analysis

1D and 2D gels were transferred overnight using 20V at 4ºC to Hybond ECL nitrocellulose membranes (Amersham Life Science, USA) using the Mini Trans-blot Electrophoretic Transfer Cell (BioRad). The transferred membranes were blocked with 0.1% PBS-T at room temperature for 1 hour. After washing with 0.05% PBS-T, the membranes were incubated overnight at 4°C with the respective patients' sera. Membranes were then washed followed by incubation with goat anti-human IgE ε-chain specific alkaline phosphatase conjugated antibody (Sigma, USA) at 1:1000 (v/v) dilution with PBS for 2.5 hours at room temperature. Washing was then performed. The positive binding of specific IgE to allergen was visualised by developing with BCIP/NBT (5-bromo-4chloro-3indolyl-phosphate / nitro-blue tetrazolium) colour substrate kit (Promega, USA) in alkaline phosphatase buffer. The membranes were then blot dried and scanned.

4.2.6 In-gel digestion of protein bands and spots

Protein bands and spots which were found positive on the Western blots were subsequently excised from the corresponding gels. Destaining was subsequently carried out with 50% acetonitrile (ACN) containing 50mM ammonium bicarbonate. Reduction and alkylation were later performed on the gels with 10mM DTT followed by 55mM iodoacetamide. Dehydration was performed by adding 100% ACN to the excised gels.

The gels were subsequently digested overnight at 37°C with 15µl (or till the whole gel is totally covered) sequencing grade porcine trypsin (Promega, USA) (12.5 ng/µl). The resulting tryptic fragments were then extracted with 50% ACN and 5% Trifluoroacetic acid (TFA). After vacuum drying, the peptides were re-dissolved in 1 µl matrix solution (5mg/ml of α-cyano-4-hydroxycinnamic acid [CHCA], 0.1%TFA, 50% ACN), and spotted onto a stainless steel MALDI target plate.

4.2.7 MALDI TOF/TOF MS/MS analysis

The samples on the MALDI target plates were analyzed using the ABI 4700 Proteomics Analyzer (Applied Biosystems, USA). For MS analysis, on average 1000 shots were accumulated for each sample. MS/MS analyses were performed using nitrogen, at collision energy of 1 kV and collision gas pressure of 3.0 × 10-7 Torr. A stop condition was used to ensure at least 2,000 to 10,000 shots were combined to obtain a good spectrum. MASCOT search engine (version 1.9; Matrix Science) was used to search for all of the tandem mass spectra. The GPS ExplorerTM software (Applied Biosystems, USA) was used to transform the spectrum files for MASCOT search engine for peptide and protein identifications. Cysteine carbamidomethylation, N-terminal acetylation, pyroglutamation, and methionine oxidation were selected as variable modifications. An error of one missing cleavage was allowed. Precursor error tolerance was set to less than 150 ppm and MS/MS fragment error tolerance less than 0.2 Da.

4.3 Results and Discussion

4.3.1 1-Dimensional SDS-PAGE and immunoblots

Proteins were separated in 12% SDS-PAGE and either stained with Coomassie blue for total protein analysis or transferred onto membranes and immunolabeled with sera previously tested positive for meats on immunoarray to identify allergenic proteins of various meats (pork, beef and lamb). The IgE 1-DE immunoblots for various meat proteins using pre-screened patients' sera confirmed the immunoarray results. Figure 3 shows the SDS-PAGE and the results of the immunoblotting of pork for 8 patients and the control subject. IgE-binding proteins for pork meat were identified at >100, 67, 57, 50 – 40, 37, 30 and 28kDa. The identified immunoreactive bands were comparable with results shown by Sabbah *et al*., 1994a, Sabbah *et al*., 1994b, Asero *et al*., 1997, Llatser *et al*., 1998, Benito *et al*., 2002, and Atanaskovic-Marković *et al*., 2002. For pig intestine and kidney, many IgE-binding proteins ranging from 200 to 15 kDa were seen. Especially noticeable were immunoreactive bands at 90, 50, and 30 kDa in both pig intestine and kidney. The control subject and blank control (secondary antibody only) showed no IgE-binding. The general pattern of IgE-binding was similar among the 8 patients with variation only in IgE-binding intensity. Majority of the immunoreactive proteins were high molecular weight proteins more than 30 kDa.

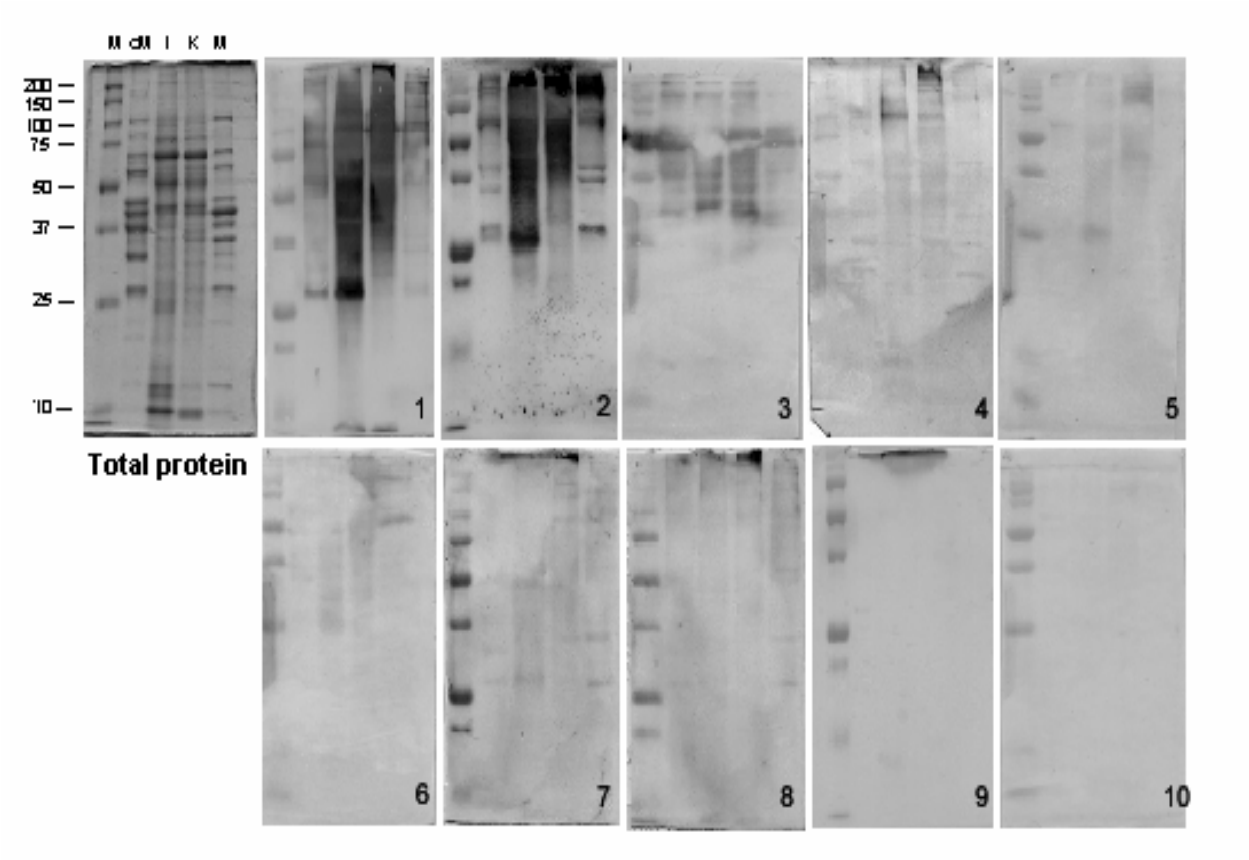


Figure 3.
1-D SDS-PAGE and immunoblotting analysis of proteins from *S. scrofa* (pig) extract. Total protein: Coomassie stain for total protein analysis; Lane 1 (M): marker (kDa); Lane 2 (cM): commercial pork skin prick extract; Lane 3 (I): pig intestine PBS extract; Lane 4 (K): pig kidney PBS extract; Lane 5 (M): pork PBS extract. Immunoblots (1 – 8): 8 patients; Immunoblot 9: control subject; Immunoblot 10: blank control (secondary antibody only).

Figure 4 shows the SDS-PAGE and the immunoblots of beef for 10 patients and the control subjects. IgE-binding proteins were identified at 180, 160, 67, 50, 31, 28, and 23 kDa. Among to two control patients tested, one (No. 12) showed IgE-binding suggesting that the particular control patient may not be a true negative patient. Nevertheless, the blank control (secondary antibody only) showed no IgE-binding. The identities of immunoreactive bands at 160 and 67 kDa are likely to be bovine gamma globulin (BGG) and bovine serum albumin respectively (Ayuso *et al*., 2000; Fuentes *et al*., 2004). The rest of the IgE-binding proteins are yet to be known. However, from the molecular weight, they are likely to be similar to those in pork.

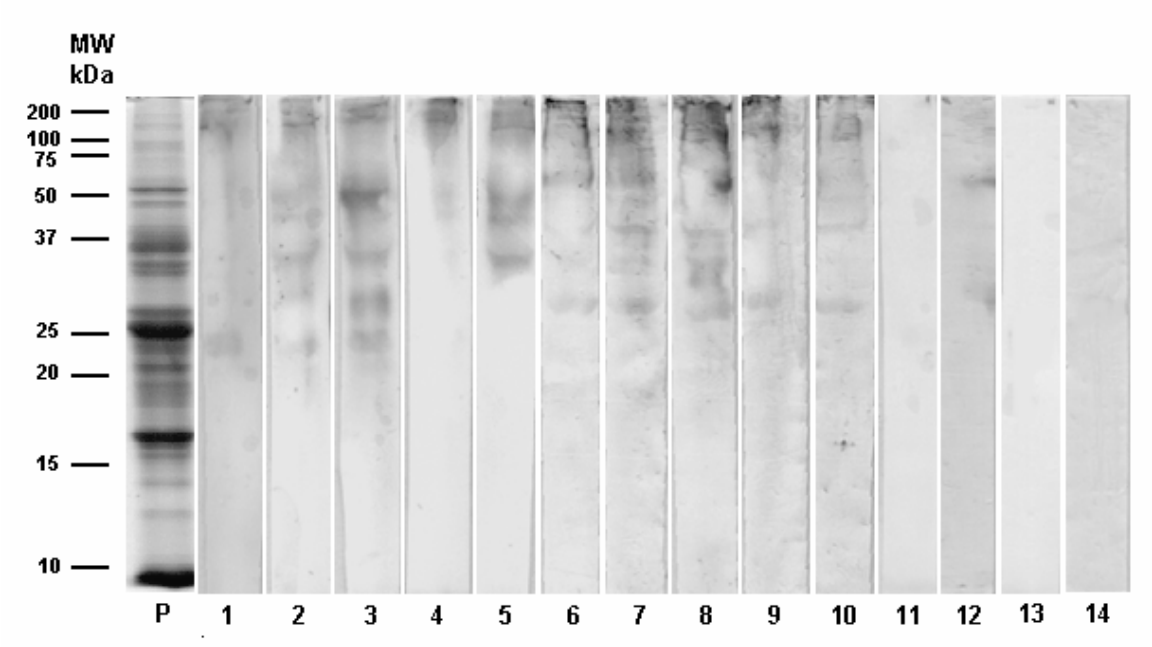


Figure 4. 1-D SDS-PAGE and immunoblotting analysis of proteins from *B. taurus* (cow) extract. Lane P: Coomassie stain for total protein analysis; Lane 1 – 10: immunoblots with 10 patients' sera; Lane 11 and 12: immunoblots with 2 control subjects' sera; Lane 13 and 14: Blank controls (secondary antibody only).

Figure 5 shows the SDS-PAGE and the immunoblots of lamb for 10 patients and the control subjects. IgE-binding proteins were identified at >100, 75, 50 - 45, 37, 31, 28, and 25 kDa. The control subjects and blank control (secondary antibody only) showed no IgE-binding. Again, the immunoreactive bands were of similar molecular weight to those in pork and beef suggesting similar allergenic proteins are present among various meat sources and they may be cross-reactive.

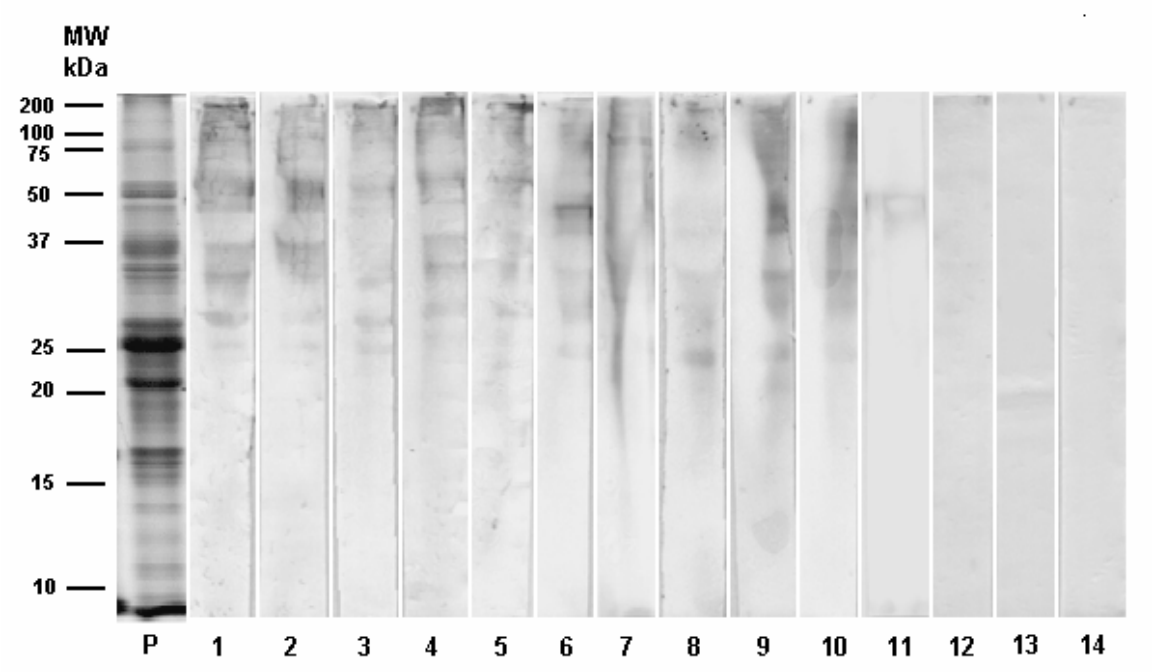


Figure 5. 1-D SDS-PAGE and immunoblotting analysis of proteins from *O. aries* (goat) extract. Lane P: Coomassie stain for total protein analysis; Lane 1 – 10: immunoblots with 10 patients' sera; Lane 11 and 12: immunoblots with 2 control subjects' sera; Lane 13 and 14: Blank controls (secondary antibody only)**.**

4.3.2 2-Dimensional SDS-PAGE and immunoblots

For better separation and resolution to allow accurate identification of allergens and their isoforms, the proteins from pork, beef and lamb were separated by means of 2-D PAGE and immunoblotted using patients and control subjects sera. Figure 6 – 8, show the typical Coomassie blue stained 2-DE maps of pork, beef and lamb proteins in the pH range 3 – 10 respectively. The three meat sources showed similar electrophoric profile with an average of approximately 250 protein spots distributed over the separation range of pI 3 – 10 and molecular size range of 6.5 – 200 kDa. Identification of the IgE-immunoreactive proteins for each patient was performed by matching the image of the Coomassie blue stained gel to the image of immunoblot membrane after immunodetection using PDQuest 2-D Analysis Software (Bio-Rad, USA). Both the reference gel and the gel for immunoblot were run under similar conditions side by side on the same day. Figure 9 – 11, show the 2-DE immunoblots patients and control subjects for pork, beef, and lamb respectively to give an overview of all identified allergenic proteins.

After image analysis, an average of 25 distinct spots or regions from each species were selected as being IgE immunoreactive proteins, and these were excised for MALDI-TOF-TOF analysis. The results of MALDI-TOF TOF identifications for pork, beef and lamb are shown in Table 1 - 3 respectively.

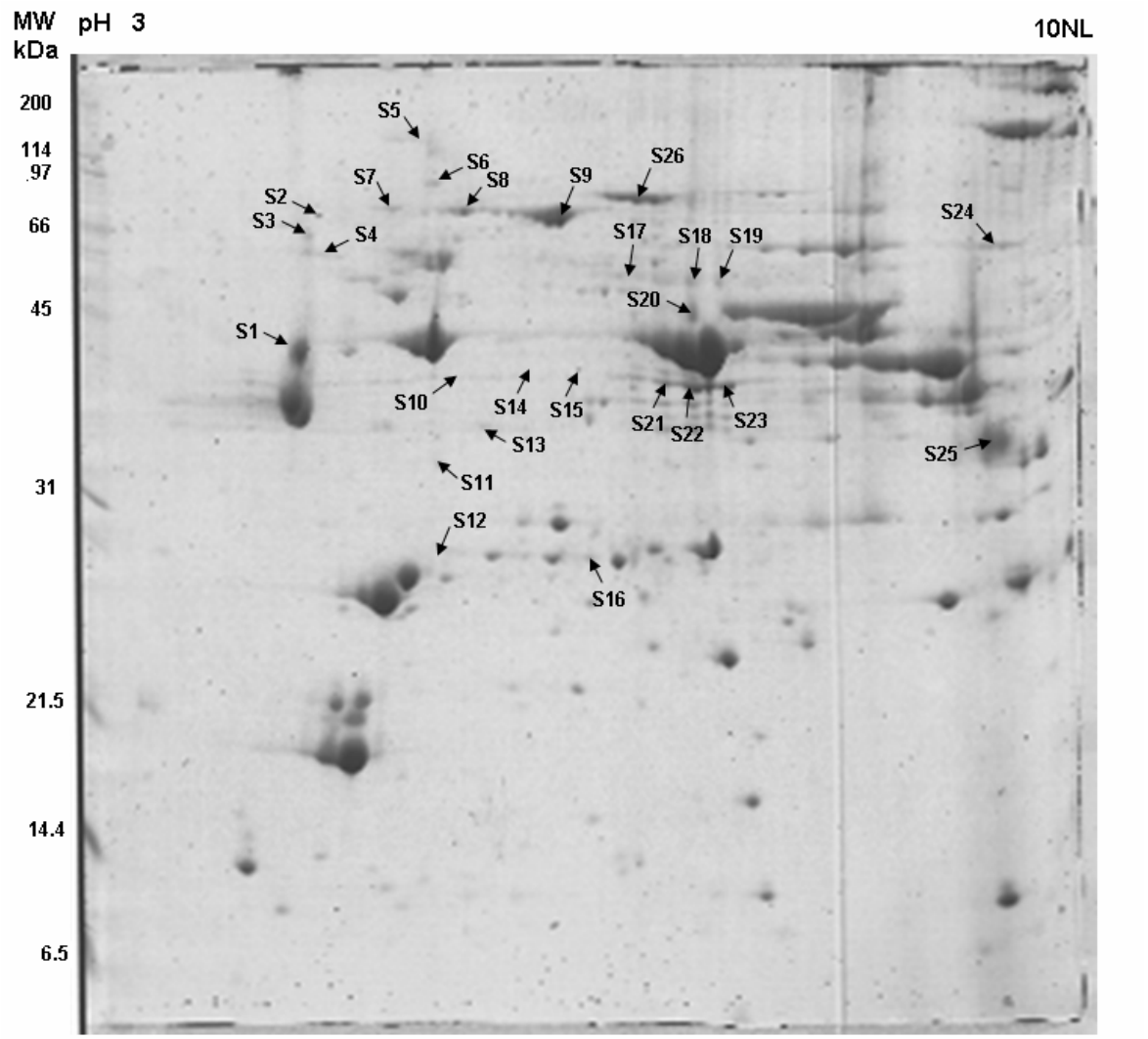


Figure 6. 2-DE separation of *S. scrofa* (pig) proteins. *S. scrofa* meat (pork) was extracted with TCA/acetone and dissolved in urea sample buffer before 2-D PAGE. First dimension: pH 3 – 10 NL; second dimension: 12% SDS-PAGE gel. Protein spots were visualized by Coomaisse blue staining. Isoelectric points and molecular weight (kDa) are indicated at the top and on the left side, respectively. An arrow with numeral indicates an IgE-binding spot identified by MALDI-TOF-TOF mass spectrometry.

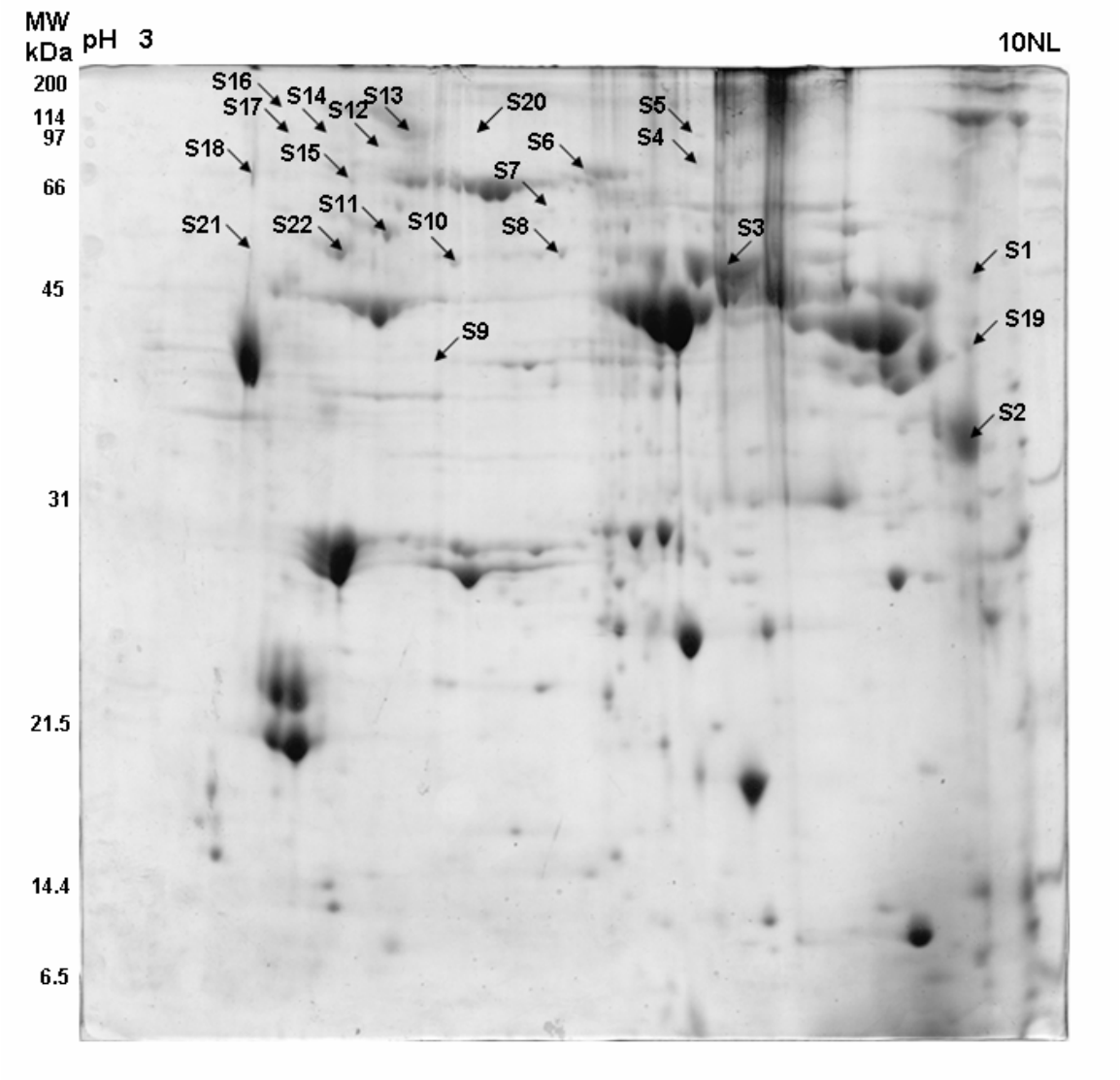


Figure 7. 2-DE separation of *B. taurus* (cow) proteins. *B. taurus* meat (beef) was extracted with TCA/acetone and dissolved in urea sample buffer before 2-D PAGE. First dimension: pH 3 – 10 NL; second dimension: 12% SDS-PAGE gel. Protein spots were visualized by Coomaisse blue staining. Isoelectric points and molecular weight (kDa) are indicated at the top and on the left side, respectively. An arrow with numeral indicates an IgE-binding spot identified by MALDI-TOF-TOF mass spectrometry.

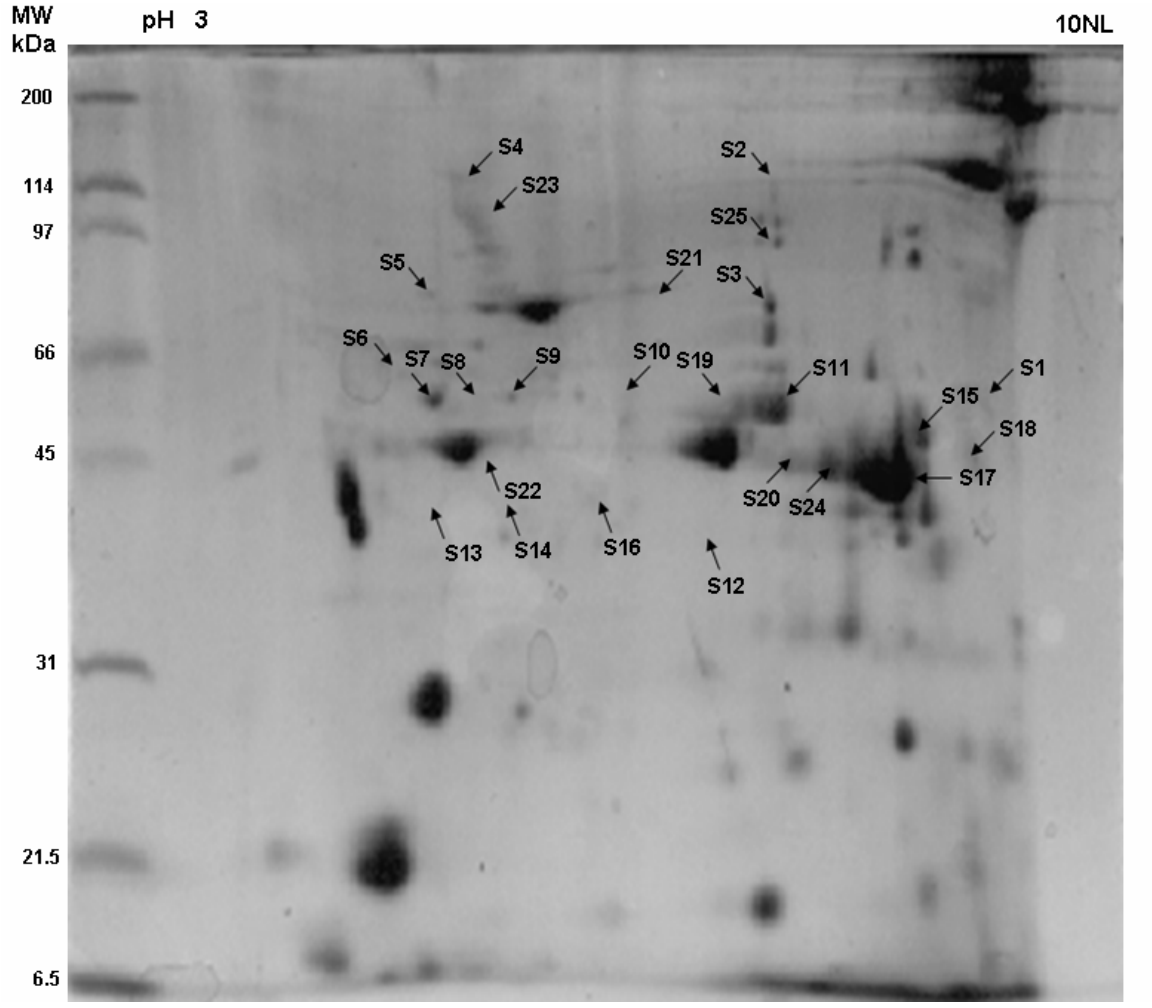


Figure 8. 2-DE separation of *O. aries* (goat) proteins. ***O. aries*** meat (mutton) was extracted with TCA/acetone and dissolved in urea sample buffer before 2-D PAGE. First dimension: pH 3 – 10 NL; second dimension: 12% SDS-PAGE gel. Protein spots were visualized by Coomaisse blue staining. Isoelectric points and molecular weight (kDa) are indicated at the top and on the left side, respectively. An arrow with numeral indicates an IgE-binding spot identified by MALDI-TOF-TOF mass spectrometry.

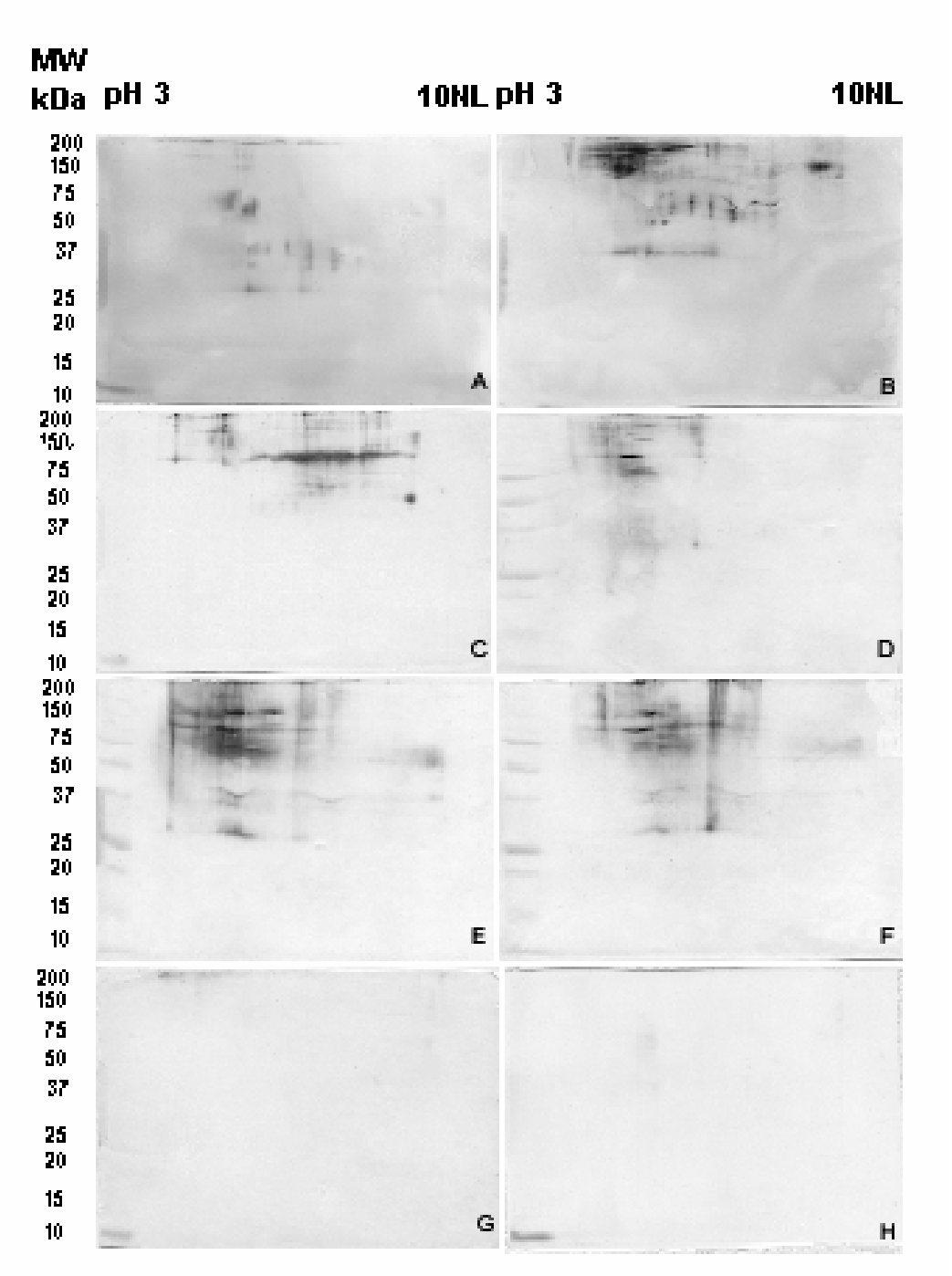


Figure 9. 2-DE immnoblots of *S. scrofa* (pig) proteins. A blotting membrane was probed with serum IgE from patients (A – F) and from control subject as negative control (G) Blank control (H) is probed with secondary antibody only.

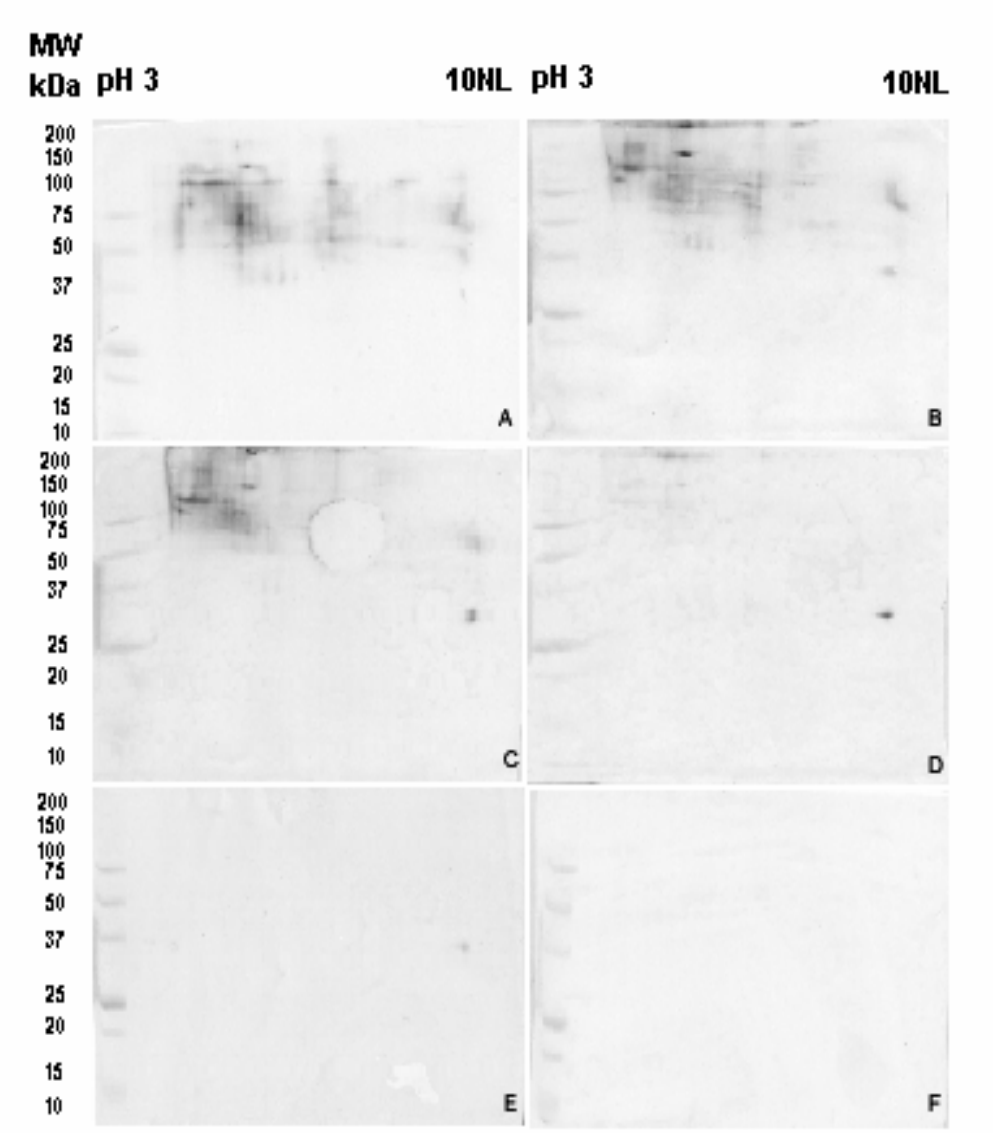


Figure 10. 2-DE immnoblots of *B. taurus* (cow) proteins. A blotting membrane was probed with serum IgE from patients (A – D) and from control subject as negative control (E) Blank control (F) is probed with secondary antibody only.

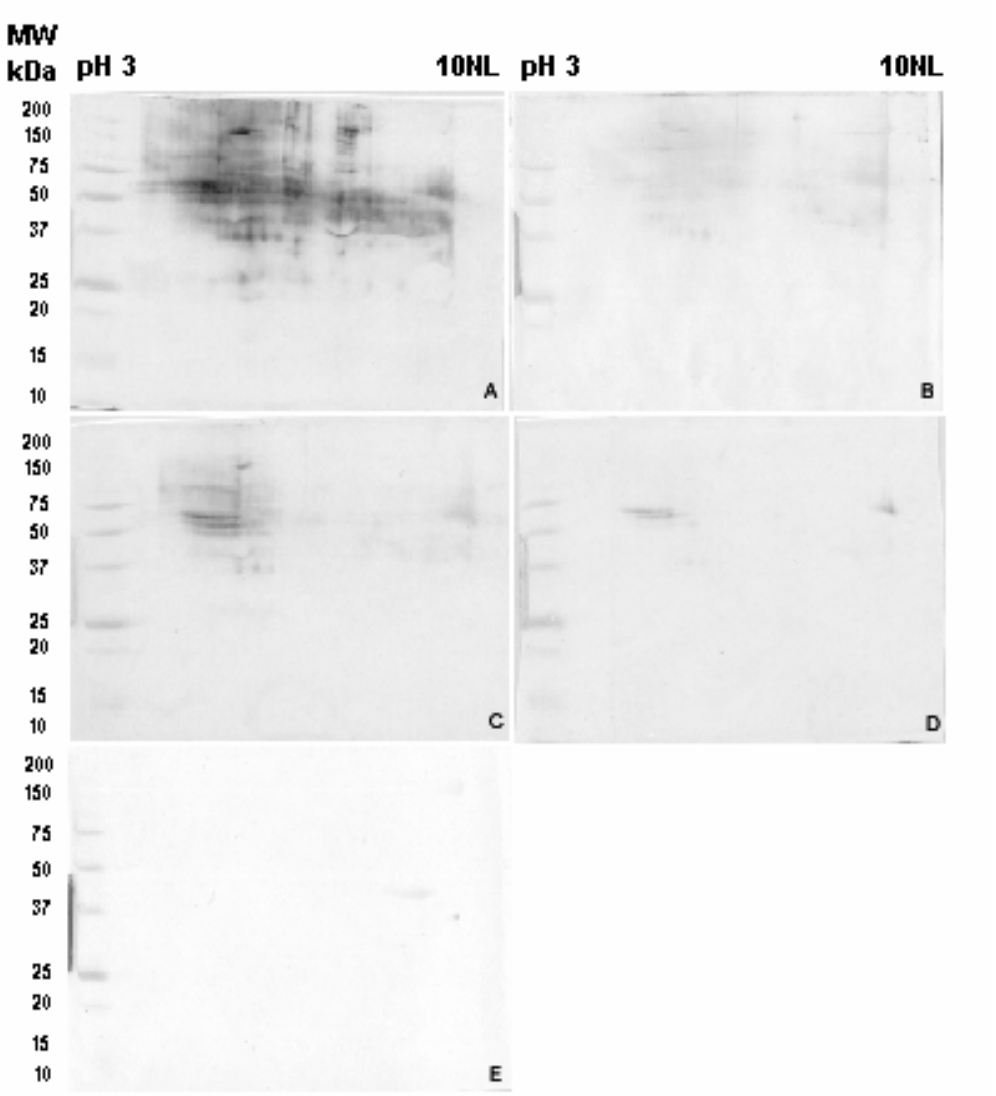


Figure 11. 2-DE immnoblots of *O. aries* (goat) proteins. A blotting membrane was probed with serum IgE from patients (A – C) and from control subject as negative control (D) Blank control (E) is probed with secondary antibody only.

4.3.3 Protein identification by MALDI-TOF_TOF mass spectrometry

The results from MALDI-TOF-TOF analysis for pork showed 25 out of 26 spots have significant homology (Mowse scores >75) with known peptide sequences from the database. Interestingly, many of the identified proteins are homologous to known allergens such as tropomyosins, heat shock proteins, serum albumin, IgG heavy chain, enolases, troponins, and transferrins. Most of the above mentioned identified proteins are highly conserved based on sequence homology and even structural conformation (except troponins), hence they are very likely to be cross-reactive. For instance, shown in Figure 12, is the structural similarity among transferrins from pig, cow and the know allergen Gal d 3 which is an ovotransferrin. The prevalence of hypersensitivity reactions to homologous allergens from other origins that share IgE epitopes on molecules is likely to become a significant clinical problem as a consequence of the general increase in allergic sensitization.

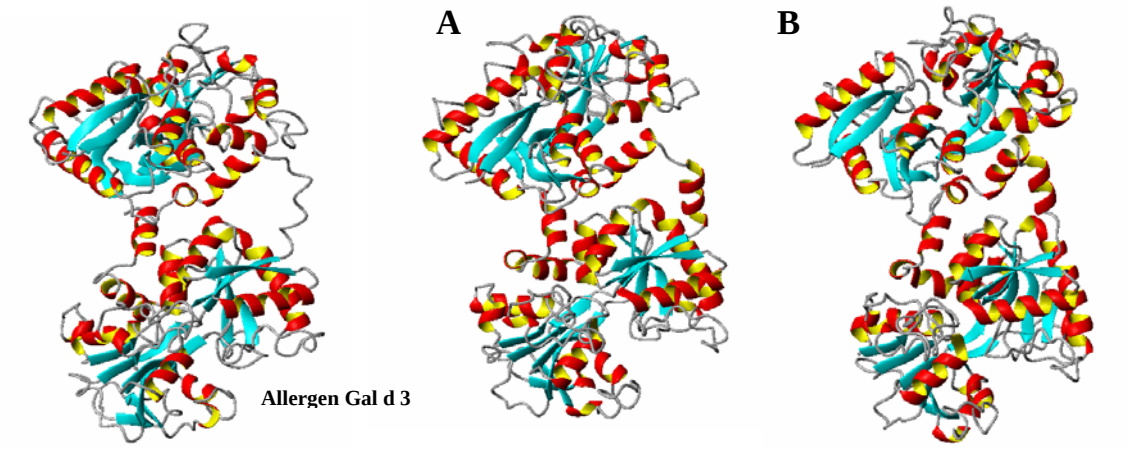


Figure 12. Three-dimensional homology modeling of allergen Gal d 3 (ovotransferrin precursor-conalbumin) and other transferrins from (A) pig and (B) cow. They show very high sequence and structural homology thus are candidate putative allergens.

Tropomyosins are highly conserved with approximately 75% sequence identity with other arthropod tropomyosins and they have been shown to be highly cross-reactive among shellfish and other invertebrate tropomyosins (Reese *et al*., 1997). These proteins are highly allergenic and hence have been termed pan allergens, and are responsible for a significant proportion of house dust mite and shellfish allergies. The purified natural protein has a molecular weight of 37 kDa (Aki et al., 1995; Reese *et al*., 1997), and the result 2-D immunoblotting demonstrated a similar molecular weight (Spot S1 at approximately 37 kDa) for porcine tropomyosin homologue (Figure 5). Spot S3, which was also identified as tropomyosin, could be a dimerized entity of tropomyosin at approximately 70 kDa. Structural study on tropomyosins have revealed a parallel two-stranded alpha-helical coiled-coil structure with a remarkable core hence suggest the likelihood of forming dimers (Brown *et al*., 2001)

Heat shock proteins (HSPs) are ubiquitously expressed highly conserved molecular chaperones that are classified based on their respective molecular masses. The HSP family includes: small HSPs (HSP25, HSP27, and HSP28 family), HSP40, HSP60, HSP70, HSP90 and HSP110 family. Unfortunately, HSPs are also common fungal allergens known to widely cross-reactive across the different families (Shen *et al*., 1997). Additionally, HSP70 are common allergens in barley, corn and hazel pollen (Chiung *et al*., 2000; Gruehn *et al*., 2003). Our 2-DE immunoblotting results have identified two spots S7 and S8 as heat shock protein 70 at the correct molecular weight. The fact that there were two or even more heat shock proteins present suggests that they are isoforms or variants of post-translational protein modification. To date, there are no studies done

on isoforms identification and post-translational protein modification with regards to meat (pork, beef, and mutton) proteome. However, considerable interest in milk proteomics in recent years because the occurrence of co- and post-translational protein modifications means that many gene products are present in milk as multiple protein forms (Natale *et al*., 2004). The spatial and temporal distribution of isoforms can play a crucial role in determining functionality whether in the context of development, maintenance of homeostasis or progression of allergic diseases (Natale *et al*., 2004). The difference in the two HSPs identified were likely due to glycosylation because spot S8 was glucose-related.

Enolase is another pan-allergen usually isolated among the molds (Breitenbach *et al.,* 1997; Simon-Nobbe *et al.,* 2000) and recently also identified as a cross-reacting allergen in latex (*Hevea. brasiliensis*), Hev b 9 (Wagner *et al.,* 2000). A typical molecular weight of enolase ranges from 47 to 55 kDa and the result from our 2-DE immunoblots have identified an enolase protein (Spot S20) at approximately 47 kDa. Troponin is a known allergen (Bla g 6) specifically found in cockroach *Blatella germanica* (Arruda *et al*., 1995). Troponin is allergenic probably due to its association with tropomyosin which is a common pan-allergen. This is because troponin and tropomyosin forms a complex during the process of muscle contraction and relaxation hence they may have region of similar conformational or linear epitopes. The results of our 2-DE immunoblots have identified three troponin spots (Spot S21, S22, and S23) indicating the presence of isoforms or variants of post-translational protein modification. However, the molecular weights of the identified troponins (approximately 36 kDa) were of a larger size compared to a typical

troponin of approximately 30 kDa. The slight difference in molecular weight observed may be due the slight variations in ionic charges of the protein which will affect the speed of migration in the SDS-PAGE gel resulting in a shift in molecular weight. Another reason for the shift in molecular weight may be due to post-translational protein modification such as glycosylation and phosphorylation. For instance, an additional N-glycan glycosylation caused a 3 kDa shift (Piva *et al*., 2002).

The degree of overlap among the proteins identified by mass spectrometry was very high among the three animal species. A total of 13 out of 58 proteins (22%) were found in all three animal species and 35 proteins (60%) were found in at least two species. This high degree of overlap indicates that the compositions of allergenic components are vastly similar among the three species and they are most likely cross-reactive. One interesting observation was many of the proteins identified in mutton are related to glycolysis or Kerbs' cycle pathway such as enolase, pyruvate dehydrogenase, fructose 1,6-bisphosphate aldolase, malate dehydrogenase and aconitase. Studies have shown that sensitization to microbial antigens was accompanied by the process of glycolysis and dehydrogenase activity (Pleshkova, 1978). Another study has shown that ovalbumin sensitized and challenged mice have a significant increase in nitrated protein and many of these proteins have been identified as the above mentioned enzymes involved in glycolysis and kerbs' cycle (Ghosh *et al*., 2006). Therefore, these enzymes may have a role in the process of sensitization and/or inflammatory response. Additionally, enolase has been demonstrated to be an allergen (Hev b 9) in latex suggesting that the other enzymes were likely allergenic.

Table 1.

Identification of proteins fro 2-DE of *S. scrofa* (Pig) after in-gel trypsin digestion by MALDI-TOF-TOF and NCBI database searching. Missing spots were due to poor spectra, no significant matches, or keratin contaminations.

Protein identified by MALDI TOF/TOF MS/MS							
Spot ID	**Calculated pI/Mr**	**Protein identity**	**Peptides matched**	**Sequence coverage (%)**	**MOWSE score**	**E-value**	**NCBI Accession number**
S1	4.66/32931	Tropomyosin 2, beta	5	26	271	2.00E-19	gi|11875203
S2	4.56/50913	PREDICTED: similar to RIKEN cDNA 1110030K22	13	19	212	1.6e-015	gi|62645352
S3	4.71/32732	Tropomyosin 1, alpha	18	40	161	2e-010	gi|57281697
S4	5.50/39199	Alpha-2-HS glycoprotein precursor (Fetuin-A)	10	30	85	0.0074	gi|231467
S5	5.3/42272	Actin skeletal muscle	14	23	171	2e-011	gi|217376
S6	5.13/89917	Valosin-containing protein	34	44	231	2.00E-17	gi|47523626
S7	5.24/70989	Heat shock protein 70kDa	27	42	321	2.00E-26	gi|123647
S8	5.38/81568	Heat shock protein 70kDa (Glucose-related)	21	30	181	2.00E-12	gi|55632315
S9	5.92/71362	Serum albumin precusor	26	32	299	3.20E-24	gi|833798
S11	5.36/31573	Actin capping protein Z	11	29	105	7.90E-05	gi|45382141
S12	5.02/22521	Myosin light chain	11	38	162	1.6e-010	gi|33563264
S13	5.36/31611	F-actin capping protein (CapZ beta)	13	34	99	0.00032	gi|1345668
S14	5.61/45547	Aminoacylase I	13	28	110	2.50E-05	gi|1845
S15	5.03/20336	Glycerol-3-phosphate dehydrogenase	8	39	110	2.50E-05	gi|2149959
S16	6.49/22985	Heat shock 27kDa protein 1	19	46	368	4e-031	gi|55926209
S17	6.82/52954	IgG heavy chain	10	25	94	0.00097	gi|47523192
S18	6.82/52954	IgG heavy chain	13	30	308	4.00E-25	gi|47523192
S19	6.82/52954	IgG heavy chain	7	12	133	1.30E-07	gi|47523192
S20	6.73/55735	Enolase 3	28	50	437	5e-038	gi|57086343
S21	7.74/29811	Troponin T	11	32	122	1.6e-006	gi|46389787
S22	7.74/29811	Troponin T	10	30	62	1.5	gi|46389787
S23	7.74/29811	Troponin T	12	32	78	0.044	gi|46389787
S24	9.12/55760	Titin immunoglobulin domain protein, myotilin	20	28	229	3.20E-17	gi|13529230
S25	9.3/35582	LIM protein	16	29	152	1.6e-009	gi|47523806
S26	6.73/78954	transferrin	35	47	603	1.3e-054	gi|833800

PMF software: Mascot

Table 2.
Identification of proteins fro 2-DE of *B. taurus* (cow) after in-gel trypsin digestion by MALDI-TOF-TOF and NCBI database searching. Missing spots were due to poor spectra, no significant matches, or keratin contaminations.

Protein identified by MALDI TOF/TOF MS/MS							
Spot ID	**Calculated pI/Mr**	**Protein identity**	**Peptides matched**	**Sequence coverage (%)**	**MOWSE score**	**E-value**	**NCBI Accession number**
S1	9.36/51598	Mitochondrial fatty acid beta-oxidation multienzyme complex	26	56	330	2.5E-27	gi|2832715
S2	8.63/35509	Four and a half LIM domains 1 protein, isoform C	16	30	261	2E-20	gi|61826953
S3	5.23/42376	Alpha actin	16	37	473	1.30E-41	gi|15825436
S4	8.08/86045	Aconitase	29	24	392	1.6E-033	gi|1351857
S6	6.75/79870	Transferrin	32	41	440	2.5E-038	gi|2501351
S8	6.44/47589	Enolase 1	13	23	189	2.00E-07	gi|27806645
S10	5.46/49852	Bovine Mitochondrial Cytochrome Bc1 Complex	26	52	252	1.60E-19	gi|3891848
S11	5.21/52587	Desmin	42	70	932	1.60E-87	gi|2959452
S12	5.14/89977	Valosin-containing protein	23	31	106	2.00E-09	gi|17865351
S13	5.37/112936	Beta-myosin heavy chain	50	43	447	5.00E-39	gi|49641
S14	5.38/81568	Heat shock 70kDa protein (glucose-related)	18	21	81	0.021	gi|55632315
S15	5.82/71244	Albumin	15	18	79	0.033	gi|162648
S18	4.68/32896	Tropomyosin	11	26	81	0.022	gi|19387215
S19	9.45/36543	Succinate-CoA ligase	12	20	127	5.00E-07	gi|61553261
S20	5.36/611249	Microtubule-actin crosslinking factor	46	8	79	0.029	gi|4887229
S22	4.97/51397	Bovine F1-Atpase, Chain E	28	56	598	4.00E-54	gi|3660252

PMF software: Mascot

Table 3.

Identification of proteins fro 2-DE of *O. aries* (goat) after in-gel trypsin digestion by MALDI-TOF-TOF and NCBI database searching. Missing spots were due to poor spectra, no significant matches, or keratin contaminations.

Protein identified by MALDI TOF/TOF MS/MS							
Spot ID	Calculated pI/Mr	Protein identity	Peptides matched	Sequence coverage (%)	MOWSE score	E-value	NCBI Accession number
S1	9.36/51598	Subunit beta of the mitochondrial fatty acid beta-oxidation multienzyme complex	19	39	172	1.6e-011	gi|2832715
S3	7.6/58393	Pyruvate kinase	26	38	305	7.90E-25	gi|2231167
S4	5.62/223764	Myosin heavy chain	30	17	109	3.2e-005	gi|21743235
S5	5.07/72473	Heat shock 70kDa protein (glucose-regulated protein)	25	35	185	7.90E-13	gi|38303969
S7	4.97/51397	Bovine F1-Atpase	31	65	505	7.90E-45	gi|3660252
S8	5.21/52587	Desmin	22	49	210	2.50E-15	gi|2959452
S9	5.46/49866	Chain A, Cytochrome Bc1 Complex From Bovine	23	50	202	1.6e-014	gi|51247182
S11	6.73/47337	Enolase 3, beta muscle	22	44	217	5.00E-16	gi|6679651
S14	5.38/39299	Pyruvate dehydrogenase	11	28	120	2.5e-006	gi|56090293
S15	8.64/44973	Phosphoglycerate kinase	25	56	333	1.30E-27	gi|52783777
S16	6.15/36700	Cytosolic malate dehydrogenase	16	37	212	1.6e-015	gi|61856478
S17	8.20/39586	Fructose 1,6-Bisphosphate Aldolase	22	59	269	3.20E-21	gi|6730621
S18	9.01/45003	Aminotransferase	16	48	111	2.00E-05	gi|222979
S19	6.73/47337	Enolase 3, beta muscle	22	44	255	7.90E-20	gi|6679651
S20	8.2/39586	Fructose-bisphosphate aldolase A (Muscle-type aldolase)	21	43	245	7.90E-19	gi|6730621
S24	8.3/39647	Fructose 1,6-Bisphosphate Aldolase	19	45	214	1.00E-15	gi|2781030
S25	7.25/86045	Aconitase	28	30	240	2.5e-018	gi|1351857

PMF software: Mascot

4.3.4 Comparison between bioinformatics and proteomics approach for allergen prediction and/or identification

In total, the proteomics approach has identified 58 allergenic proteins from the three animal species (pig, cow and goat). Of the 58 proteins, 17 (30%) proteins were predicted and identified by both bioinformatics and proteomics approaches (Figure 13). Additionally, out of the 17 proteins, 15 (88%) proteins were predicted in both the motif-based and sequence homology based prediction systems. This result demonstrates that when a protein is predicted as a putative allergen in both prediction systems, it is highly credible that the particular protein is allergenic. The remaining two proteins [tropomyosin in cow (M1) and alpha-2-HS glycoprotein precursor (Fetuin-A) in pig (M2)] were predicted one each by the two prediction systems. This indicates that one prediction system alone is still deficient and the combination of the two systems significantly increases the confidence level of the prediction. Nevertheless, predictive systems are still entry level whereby it still requires the proteomics tools to determine IgE-binding to confirm its role in allergenicity.

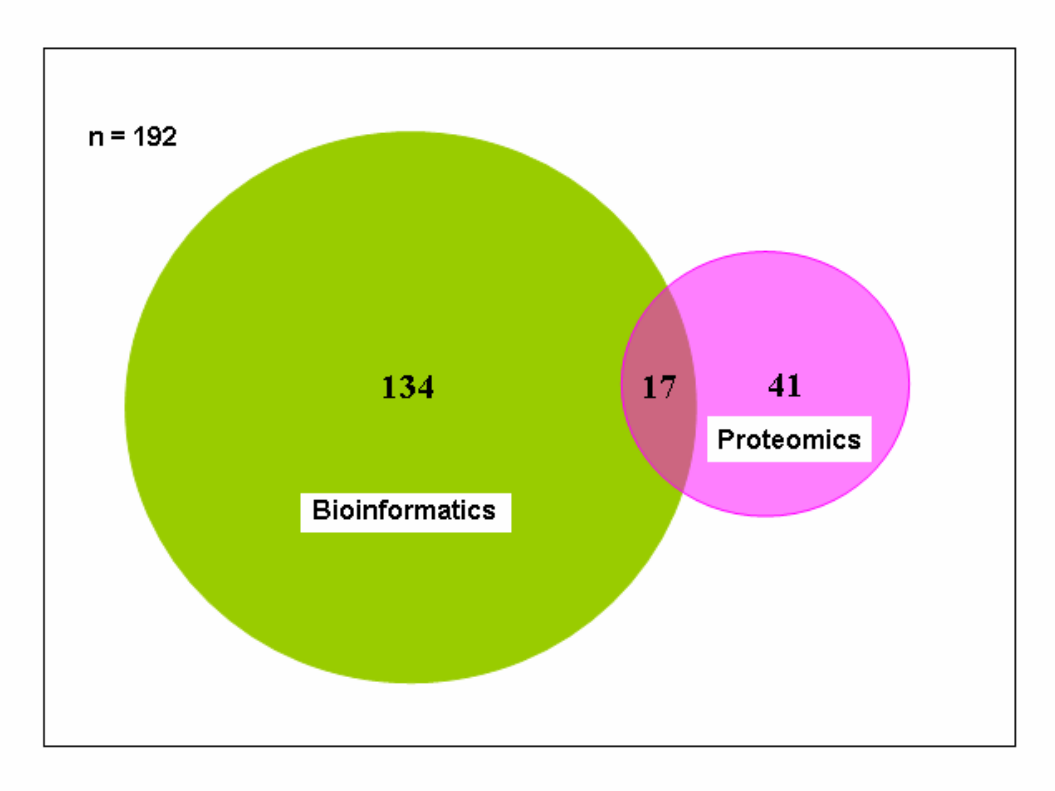


Figure 13.
Venn diagram showing the comparison between bioinformatics and proteomics approach for allergen prediction and/or identification

4.4 Conclusion

In this study, we are able to identify 58 allergens from cow, pig and goat. Many of them are homologous to well-known allergens such heat shock proteins, tropomyosins, dehydrogenases, serum albumins and transferrins. The composition of allergenic components is very similar among the three animal species indicating that these similar allergenic components contribute to the cross-reactivity among the three species. The allergens identified by the proteomics approach also confirm that those putative allergens predicted by both allergen prediction systems are indeed allergenic hence increasing the level of confidence of allergen prediction by the bioinformatics approach.

CHAPTER 5: MOLECULAR CLONING AND IMMUNOGLOBULIN E (IGE) REACTIVITY OF PUTATIVE MEAT-BASED ALLERGENS

5.1 INTRODUCTION

5.1.1 Usage of recombinant allergens for research and diagnosis

Currently, allergists still rely on natural allergenic products that were obtained from crude extracts for the diagnosis and treatment of allergic diseases. The natural extracts however are heterogeneous in nature containing many non-allergenic proteins, carbohydrates, lipids and other macromolecules. The allergens produced from natural sources vary in composition and content. Natural products are also at risk of being contaminated with allergens from other sources and can contain proteolytic enzymes (van der Veen *et al.,* 1996). The enzymes may be allergenic or non-allergenic, but in either case can cause degradation and loss of potency when administered together with other allergens during diagnosis or immunotherapy (Nelson *et al.,* 1996).

Recombinant DNA technology facilitates the characterization and analysis of the allergenic proteins; it also provides the basis for producing allergens and their derivatives for both diagnostic application and immunotherapy (Valenta, 2002; Schmid-Grendelmeier and Crameri, 2001). Further advantages of recombinant allergens and their derivatives include the production of preparations of consistent pharmaceutical quality;

the avoidance of problems of natural extract standardization; the inclusion of optimal concentrations of the important allergens; the exclusion of non-allergenic proteins; the avoidance of the possible risk of contamination by allergens from other sources; and exclusion of the risk of introducing infectious agents (Cromwell *et al*., 2004).

Beside diagnostic applications, purified allergens (natural or recombinant) are also essential research tools to investigate the cellular mechanisms of immediate hypersensitivity and the molecular basis of inflammatory reactions (Chapman *et al*., 2000). For instance, the IgE binding capacity, T cell responses and cytokine profiles of allergens such as the birch pollen Bet v 1 (Ferreira *et al.,* 1993; Ferreira *et al.,* 1996; Ferreira *et al.,* 1997) and some peanut allergens (Rabjohn *et al.,* 1999; Burks *et al.,* 1995) have been studied using recombinant proteins. The three-dimensional structure of allergens can be determined using recombinant protein as an alternative source to its natural counterpart. For example, the structure of the fully immunoreactive birch pollen allergen Bet v 1 has been solved by both nuclear magnetic resonance (NMR) and X-ray crystallography (Gajhede *et al*., 1996).

5.1.2 Criteria for the production and characterization of recombinant allergens for clinical applications

The choice of expression system is essential for the production of recombinant allergen. *Escherichia coli* is the most commonly used host for proteins expression when post-transcriptional modifications (e.g. glycosylation) is not necessary. Glycosylated proteins

can be produced in several eukaryotic expression systems (Schmidt and Hoffman, 2002) such as the yeast *Pichia pastoris*, *Baculovirus* in host insect cells, mammalian cells, and various plants including the tobacco plant *Nicotiana benthamiana* (Breiteneder *et al*., 2001) and barley (Horvath *et al*., 2000), but this is not to say that the glycosylation will be comparable with that of the natural glycoprotein.

Another important issue is solubility and correct folding of the recombinant allergen. For *E.coli* expression system, high-level expression of correctly folded and soluble proteins may be achieved for some allergens, but others may be expressed as insoluble inclusion bodies. Additionally, problems of aggregation may arise during the refolding process due to exposure of hydrophobic residues in unfolded or intermediate states. Therefore, an *in vitro* solubilisation/refolding strategy has to be established to recover the recombinant allergen in a soluble and correctly folded form. Correct folding of the recombinant protein can be inferred from similar circular dichroism (CD) spectrum profile, reactivity and/or activity as the native protein.

The final and most important criterion is safety issues on usage of recombinant allergens. There are potential risks associated with contaminants derived from the host cells. Such contaminants could conceivably have immunopathological effects and nucleic acid contaminants pose theoretical risks, including the possibility that they could be integrated into the host genome. Recombinant products derived from plant, insect and animal host cells, or from systems including animal products, pose an additional risk of viral infection (Cornwell *et al*., 2004). Nevertheless, since 1991, recombinant allergens such as Bet v 1 and Bet v 2 were successfully used for *in vitro* diagnosis of pollen allergy (Valenta *et al*.,

1991). Later, the first skin test with recombinant allergen (fungus *Aspergillus fumigatus* Asp f I/a) was described (Moser *et al.,* 1992). Several important allergens form various sources are now commercially available for diagnosis purposes, these include mite group 1, 2, 5 and 7 allergens, cat Fel d 1 and albumin, cockroach Bla g 1, 2, 4 and 5 allergens, birch pollen Bet v 1 and Bet v 2, peanut Ara h 1, Ara h 2 and Ara h 3 and etc (Chapman *et al.,* 2000).

5.2 MATERIALS AND METHODS

5.2.1 Bacterial strains

XL1-Blue *[N1]Δ(mcrA) 183Δ(mcrCB–hsdSMR-mrr)173 endA1 supE44 thi-1recA1 gyr 1A96 relA1 lac [F' proAB lacI_qZ ΔM15 Tn10 (Tet_r)]*

BL21 (DE3) *F- ompT hsdS_B($r_{-B}$$m_{-B}$) gal dcm (DE3) pLysS*

5.2.2 mRNA extraction

Fresh pork and beef were purchased from a local market. One gram of fresh meat tissue was ground in a baked overnight mortar filled with liquid nitrogen. After grinding, RNA isolation was performed using TRIZOL Reagent (Gibco-BRL, USA) according to the manufacturer's protocol. The purity of the RNA was judged by electrophoresis separation and spectrophotometer.

5.2.3 Molecular cloning of recombinant allergens

5.2.3.1 Bioinformatics analysis

The full-length nucleotide and protein sequences of the putative allergens listed in Table 1 were obtained from NCBI database.

Table 1.
List of putative allergens to be cloned

Sus scrofa	*Bos taurus*
Tropomyosin 1 alpha chain (Tropo 1)	Enolase 1 (ENO 1)
Tropomyosin 3 (Tropo 3)	Aldehyde dehydrogenase (ADH)
Myosin light chain (Myo_L)	Heat shock protein 70 kDa (bHSP70)
Troponin T fast skeletal muscle type (TRNT)	
Heat shock protein 70 (pHSP70)	
90-kDa heat shock protein alpha (HSP90)	

All nucleotide and amino acid sequences were analyzed and aligned using DNAMAN version 4.15 by Lynnon BioSoft. Prediction of the signal peptide cleavage site was analyzed using the software SignalP v1.1 (Nielsen *et al.*, 1997). N- glycosylation site prediction was done using NetNGlyc 1.0 Server (Blom *et al.,* 2004) using artificial neural networks to examine the sequence context of Asn-Xaa-Ser/ Thr sequences. Specific primers (Table 2) were designed at both ends of the sequence excluding the signal peptide. Oligonucleotide properties were calculated using the webtool Oligonucleotide Properties Calculator (http://www.basic.northwestern.edu/biotools/oligocalc.html).

Table 2.
List of specific forward and reverse conserved primers used for PCR amplification of desire gene

Types/Names of primer	Oligonucleotide primer sequence
Tropo 1 Forward	5'-ATGGACGCCATCAAGAAGAAG-3'
Tropo 1 Reverse	5'-TTATATGGAAGTCATATCGTTG-3'
Tropo 3 Forward	5'-ATGGCTGGAATCACCACCATCG-3'
Tropo 3 Reverse	5'-CTACATCTCATTCAGGTCGAGC-3'
Myo_L Forward	5'-ATGTCCTTCAGTGCTGACCAGATTG-3'
Myo_L Reverse	5'-TTAGATAGACATGATGTGCTTGAC-3'
TRNT Forward	5'-ATGTCGGACGAGGAAGTAGAACACG-3'
TRNT Reverse	5'-TTACTTCCACCTTCCGCCAACCTTG-3'
pHSP70 Forward	5'-ATGGCGAAGAGCGTGGCCATCG-3'
pHSP70 Reverse	5'-CTAATCCACCTCCTCGATGGTG-3'
ENO 1 Forward	5'-ATGTCCATCCTGAAGGTCCACG-3'
ENO 1 Reverse	5'-TTACTTGGCCAACGGGTTTCTGAAG-3'
ADH Forward	5'-ATGTCGTCCTCAGCCATGCC-3'
ADH Reverse	5'-TTATGAGTTCTTCTGAGAAATTTTGA-3'
bHSP70 Forward	5'-ATGGCGAAAAACATGGCTATCG-3'
bHSP70 Reverse	5'-CTAATCCACCTCCTCAATGGTGGG-3'
HSP90 Forward	5'-ATGCCCGAGGAGACCCAGGC-3'
HSP90 Reverse	5'-TCAGTCCACTTCCTCCATGCGGG-3'

5.2.3.2 RT-PCR to isolate full length clones of putative meat allergens

1 µg of total RNA was used for first strand cDNA synthesis using iScript cDNA synthesis kit (Bio-Rad, USA). The second strand of cDNA was synthesized by the use of specific forward and reverse conserved primers (Table 2) designed based on known sequence published from NCBI. PCR amplification was carried out using Expand long-template *Taq* DNA polymerase (Roche, USA). The PCR mixture contained 5 µl of 10× PCR buffer, 1 µl of 10 mM dNTP, 0.5 µl of 50 µM forward and reverse primer, 0.5 µl of *Taq* DNA polymerase, 41 µl of ddH2O, and 1.5 µl of cDNA template. The PCR cycling

profile was as follow: initial denaturation at 94 °C for 15 sec, followed by 32 cycles at 94 °C for 15 sec, 50 °C for 25 sec, 68 °C for 2 min and final extension at 68 °C for 7 min. PCR was carried out in PTC-100^{TM} Programmable Thermal Controller (MJ Research, Inc., USA). The PCR product was analyzed on 1.0% agarose gel. The DNA products with the appropriate size were excised and purified from the gel using QIAquick Gel Extraction Kit (QIAGEN Inc., USA) as according to manufacturer's instruction.

5.2.3.3 Cloning of PCR product into pGEMT®-Easy Vector

The extracted PCR products were ligated using pGEM®-T Easy Vector system (Promega, USA) following manufacturer's instruction. Briefly, One µl (~50 ng insert DNA) was ligated with 0.5 µl pGEM_-T Easy vector in 8.5 µl reaction mixture containing 5 µl 2x Rapid ligation buffer, 1 µl of T4 DNA ligase and 2.5 µl sterile deionized water. The mixture was incubated at 4 °C overnight and was subsequently used to transform *E. coli* XL1-Blue competent cells.

5.2.3.4 Transformation of *E. coli* strain XL1-Blue

Competent cells were prepared according to Sambrook and Russel (2001). For transformation, 100 µl of competent cells (XL1-Blue) were added to the ligation mixture and placed on ice for 40 min. The mixture was heat-shocked for 2 min at 42 ^{o}C and cooled on ice for 2 min. After adding 1 ml Luria-Bertani (LB) broth, the mixture was incubated for 45 min in a 37 ^{o}C shaker at 200 rpm. The bacterial culture was plated on

LB agar containing 100 μg/ml ampicillin, 80 μg/ml of 5-bromo-4-chloro-3-indolyl-beta-D galactopyranoside (Xgal) and 0.5mM of isopropyl 1-thio-β-Dgalactoside (IPTG). The plates were incubated overnight at 37 °C.

5.2.3.5 Colony Screening

Colonies were picked and employed as template in PCR screening to detect correct insert in the transformed cells. Clone specific vector primer i.e. SP6 and T7 were used. PCR was carried using *Taq* DNA polymerase (Fermentas, USA) according to manufacturer's instruction. Gel electrophoresis was performed to confirm the insert according with the expected size. Listed in Table 3 are the universal primers used for colony screening.

Table 3.
List of universal primers used for colony screening

Types/Names of primer	Oligonucleotide primer sequence
SP6	5'-ACAGCTATGACCATGATTACG-'3
T7	5'-AACGACGGCCAGTGAATTGTA-'3
pET32 Forward	5'-GTGCCACGCGGTTCTGGTATG-'3
pET32 Reverse	5'-TAGCAGCCGGATCTCAGTGGT-'3

5.2.3.6 Culture of *E. coli* and Plasmid Extraction

Only colonies with the correct insert were selected for inoculation overnight in 5 ml of LB broth at 37 °C with continuous shaking at 200 rpm. *E. coli* cultures were then centrifuged at 3000 x g for 15 min to obtain the cell pellet. Plasmid was extracted from

the cells using the QIAprep Miniprep kit (QIAGEN Inc., USA) according to manufacturer's instruction.

5.2.3.7 Ligation Independent Cloning (LIC) of putative allergens into pET32a (+) expression vector

LIC kit (Novagen, USA) was applied according to the manufacturer's protocol for directional cloning without the need of restriction enzyme digestion or ligation reaction. PCR products of putative or identified allergens with complementary overhangs to EK/LIC vector were created by building appropriate 5' extension into the primers (Table 4). The 5' of the primer was incorporated into the following sequence:

Sense primer: 5'- GAC GAC GAC AGG ATX* -insert specific sequence-3'

Antisense primer: 5'-GAG GAG AAG CCC GGT- insert specific sequence-3'

(* The first nucleotide of the insert specific sequence must complete the codon ATX)

Table 4.
List of Ek-LIC forward and reverse primers

Types/Names of primer	Oligonucleotide primer sequence
Ek-LIC Tropo 1 Forward	5'-GACGACGACAAGATGGACGCCATCAAGAAG-3'
Ek-LIC Tropo 1 Reverse	5'-GAGGAGAAGCCCGGTTTATATGGAAGTCAT-3'
Ek-LIC Tropo 3 Forward	5'-GACGACGACAAGATGGCTGGAATCACCACC-3'
Ek-LIC Tropo 3 Reverse	5'-GAGGAGAAGCCCGGTCTACATCTCATTCAG-3'
Ek-LIC Myo_L Forward	5'-GACGACGACAAGATGTCCTTCAGTGCTGAC-3'
Ek-LIC Myo_L Reverse	5'-GAGGAGAAGCCCGGTTTAGATAGACATGATG-3'
Ek-LIC TRNT Forward	5'-GACGACGACAAGATGTCGGACGAGGAAGTAG-3'
Ek-LIC TRNT Reverse	5'-GAGGAGAAGCCCGGTTTACTTCCACCTTCC-3'
Ek-LIC pHSP70 Forward	5'-GACGACGACAAGATGGCGAAGAGCGTGGCC-3'
Ek-LIC pHSP70 Reverse	5'-GAGGAGAAGCCCGGTCTAATCCACCTCCTC-3'
Ek-LIC ENO 1 Forward	5'-GACGACGACAAGATGTCCATCCTGAAGGTC-3'
Ek-LIC ENO 1 Reverse	5'-GAGGAGAAGCCCGGTTTACTTGGCCAACGGGTTTC-3'
Ek-LIC ADH Forward	5'-GACGACGACAAGATGTCGTCCTCAGCCATG-3'
Ek-LIC ADH Reverse	5'-GAGGAGAAGCCCGGTTTATGAGTTCTTCTGAGAAATTTTG-3'
Ek-LIC bHSP70 Forward	5'-GACGACGACAAGATGGCGAAAAACATGGCTATC-3'
Ek-LIC bHSP70 Reverse	5'-GAGGAGAAGCCCGGTCTAATCCACCTCCTCAATGG-3'
Ek-LIC HSP90 Forward	5'-GACGACGACAAGATGCCCGAGGAGACCCAG-3'
Ek-LIC HSP90 Reverse	5'-GAGGAGAAGCCCGGTTCAGTCCACTTCCTCCATGC-3'

PCR amplification was carried out as described in section 5.2.3 using specific LIC forward and reverse primers for individual allergen. The PCR product was analyzed on 1.0% agarose gel and subjected to gel extraction. The extracted PCR products were treated with T4 DNA polymerase in the presence of the appropriate dATP to generate the specific vector-compatible overhangs as indicated by the manufacturer's protocol. The inserts were then annealed into the Ek/LIC pET32a (+) expression vector according to the manufacturer's instructions with minor modifications. Briefly, 1 µl of Ek/LIC vector was added to 2 µl of T4 DNA polymerase treated Ek/LIC insert and incubated at 22 °C for 5 min. Subsequently, 1 µl of 25 mM EDTA was added to the mixture and incubate at 22 °C for another 5 min.

The constructed pET32a (+) expression vectors were transformed (mentioned in section 5.2.5) into *E. coli* strain BL21 (DE3) and the plates were incubated overnight at 37 °C. After a round of colony screening (mentioned in section 5.2.6) with pET forward and reverse primers, clones with the correct size were sent for DNA sequencing.

5.2.4 DNA sequencing

5.2.4.1 Automated sequencing

DNA sequencing reactions were performed as suggested in the ABI Prism^{TM} Dye Terminator cycle sequencing ready reaction kits (Perkin Elmer, USA) which involved a 20 µl mixture reaction containing 2 µl Big Dye ^{TM}, 3 µl of 2.5 X Sequencing Buffer (Perkin Elmer, USA), 300 ng PCR products and 3.2 pmole primer. Thermal cycling profile was set for denaturation at 96 °C for 30 sec, annealing at 50 °C for 15 sec, extension at 60 °C for 4 min and repeated for 25 cycles.

5.2.4.2 Purification of Automated sequencing products

Following cycle sequencing, 2 µl of 3M sodium acetate (pH4.6), 2 µl of 125 mM EDTA and 50 µl of absolute ethanol were added to the reaction mixture and incubated on ice for 10 minutes, followed by centrifugation for 20 min at 13,000 x g. The pellet was then washed twice with 500 µl 70% ethanol and centrifuged for 5 min each time. The pellet was thoroughly dried before DNA sequencing analysis.

5.2.4.3 Automated DNA sequencing analysis

The purified extension products were subjected to sequence analysis on an ABI Prism 377 DNA sequencer (Applied Biosystems, USA). The DNA sequencing service was provided by the Department of Biological Sciences, National University of Singapore (NUS).

5.2.5 Expression and purification of recombinant allergens

5.2.5.1 Sample induction and expression

Clones that are of the correct size and have been sequence confirmed were cultured overnight in 5 ml of LB broth at 37 °C with continuous shaking at 200 rpm. The cultures were then transferred into fresh 200ml LB broth with 100 μg/ml ampicillin and grown to OD_{600} at approximately 0.6. Thereafter, a final concentration of 0.5 mM of isopropyl-beta-D-thiogalactopyranoside (IPTG) was added to the bacterial cultures and incubated for another 4 hrs at 37 °C with continuous shaking at 200 rpm. The cells were harvested by centrifugation at 5000 x g for 10 min at 4 °C and re-suspended in 10 ml of binding buffer (w/o urea) containing 5 mM imidazole, 0.5M NaCl, 20 mM Tris-HCl, pH 7.9. The re-suspened cells were sonicated on ice for 5 x 1 min or until the lysates become slightly clear. Both the inclusion bodies and the cell lysates were separated from the supernatant by centrifuging at 15,000 x g for 30 min in 4 °C. The recombinant protein's degree of solubility was evaluated by SDS-PAGE for both supernatant and pellet. For protein in inclusion bodies, the pellet was dissolved in 1× binding buffer with 6 M urea.

5.2.5.2 Affinity purification of recombinant protein with pET-32a (+) His-Tag system

His-tagged recombinant proteins were then purified with affinity chromatography by using the His•Bind resin and buffer kit (Novagen, USA) according to the manufacturer's instructions. Briefly, the supernatants from either soluble or insoluble proteins were loaded onto approximately 10 ml of charged Ni-NTA resin (Novagen) packed in columns (Bio-Rad, USA). The bound proteins were washed 10 x with wash buffer (60 mM imidazole, 0.5 M NaCl and 20 mM Tris-HCl, pH 7.9) and eluted with 15 ml of elution buffer (0.5 M imidazole, 0.5 M NaCl and 20 mM Tris-HCl, pH 7.9). In the case of insoluble proteins, 6 M urea was added in the washing and elution buffers. The purified proteins were then quantified by the Bio-Rad protein assay (Bio-Rad, USA).

5.2.6 Recombinant proteins immunoarray

5.2.6.1 Patients and sera

A total of 80 sera from the previously pre-screened population mentioned in section 2.2.6 were selected for recombinant allergen screening. Positive (70) and negative (10) sera were selected based on positivity or negativity to meat crude proteins *via* allergen immunoarray screening performed previously.

5.2.6.2 Immunoarray

The procedure for the recombinant proteins immunoarray was essentially the same as described in the previous section 2.2.4.

5.3 Results and Discussion

5.3.1 Characterization of recombinant proteins

5.3.1.1 General Bioinformatics analysis of putative allergen sequences

A total of nine full-length nucleotide and protein sequences were obtained from NCBI database. Detailed bioinformatics analyses of the putative allergens are summarized in (Table 5). Prediction of the signal peptide cleavage site is based on the (-3, -1) rule as suggested by von Heijne (1986) and analysed using the software SignalP v1.1 (Nielsen *et al.*, 1997). All the nine putative allergens do not have signal peptides and most of them possess N-glycosylation sites except the tropomyosins.

Table 5.
Detailed bioinformatics analyses of the putative allergens

Putative allergen	CDS (bp)	CDS (amino acids)	Est. MW (Daltons)	Calculated pI	Signal Peptide	Glycosylation sites	cDNA sequence diagram
Tropo 1	855	284	32693	4.67	No	No	Figure 1
Tropo 3	747	248	29000	4.75	No	No	Figure 2
Myo_L	453	150	16711	4.63	No	Yes	Figure 7
TRNT	753	250	29795	7.74	No	Yes	Figure 5
pHSP70	1926	641	70019	5.6	No	Yes	Figure 11
ENO 1	1305	434	47223	6.44	No	Yes	Figure 10
ADH	1506	501	54771	6.24	No	Yes	Figure 8
bHSP70	1926	641	70152	5.68	No	Yes	Figure 12
HSP90	2202	733	84677	4.93	No	Yes	Figure 13

5.3.1.2 Tropomyosins

Tropomyosin belongs to a family of highly conserved proteins with multiple isoforms found in both muscle and non-muscle cells of all species of vertebrates and invertebrates. Its native structure consists of two parallel alpha-helical tropomyosin molecules that are wound around each other to form a coiled-coil dimer of approximately 66 kDa (Reese *et al.*,1999). The basic function of tropomyosin is regulation of muscle contraction and relaxation.

Tropomyosins are potent allergens. Many species especially arthropods are sources of the allergenic tropomyosins that can sensitize and induce IgE-mediated allergic reactions in humans. Due to its high sequence homology, tropomyosins have been shown to be the common allergens in the extracts of crustaceans, mollusks, insects (dust mites, cockroaches, moths, etc) and meat exhibiting cross-reactivity among these organisms (Martinez, *et.al*., 1997; Asturias, *et al.*, 1999; Ayuso *et al*., 1999). A total of eight

tropomyosin epitopes were known based on studies on Pen I 1 from Shrimp (Shanti *et al.,* 1993), Pen a 1 (*Penaeus azetus)* by Reese *et al.,*1997, and Oyster Cra g 1 (Ishikawa *et al.,* 1998) (Table). Thus, tropomyosins from *Sus scrofa* were cloned to determine their IgE-binding reactivity. Full length sequence genes of Tropo 1 (Figure1) and Tropo 3 (Figure 2) comprised of 855 bp and 747 bp encoding polypeptides of 284 and 248 amino acids respectively. Tropo 1 and Tropo 3 have relatively high sequence homology with overall identity at 60.55% (Figure 3). Also, the likely sites of IgE-binding epitopes were predicted based on the eight known tropomyosin epitopes.

>X66274 (Sus scrofa Tropomyosin 1 alpha chain (Alpha-tropomyosin)-cardiac)

```
1          ATGGACGCCATCAAGAAGAAGATGCAAATGCTCAAGCTCGACAAGGAGAACGCCTTGGAT
1           M  D  A  I  K  K  K  M  Q  M  L  K  L  D  K  E  N  A  L  D
61         CGAGCGGACGAGGCGGAGGCCGATAAGAAGGCGGCAGAGGACAGGAGCAAGCAGCTGGAA
21          R  A  D  E  A  E  A  D  K  K  A  A  E  D  R  S  K  Q  L  E
121        GATGAGCTGGTGTCGCTGCAAAAGAAGCTCAAGGCCACCGAAGATGAACTGGACAAATAT
41          D  E  L  V  S  L  Q  K  K  L  K  A  T  E  D  E  L  D  K  Y
181        TCCGAGGCTCTCAAAGATGCCCAGGAGAAGCTGGAGCTGGCGGAGAAAAAGGCCACCGAT
61          S  E  A  L  K  D  A  Q  E  K  L  E  L  A  E  K  K  A  T  D
241        GCTGAAGCCGATGTAGCTTCTCTGAACAGACGCATCCAACTGTTTGAGGAAGAGCTGGAC
81          A  E  A  D  V  A  S  L  N  R  R  I  Q  L  F  E  E  E  L  D
301        CGTGCCCAGGAGCGACTGGCAACAGCTTTACAGAAACTTGAGGAGGCTGAGAAGGCAGCA
101         R  A  Q  E  R  L  A  T  A  L  Q  K  L  E  E  A  E  K  A  A
361        GATGAGAGTGAGAGAGGCATGAAAGTCATTGAAAGCCGAGCCCAAAAGGATGAGGAGAAA
121         D  E  S  E  R  G  M  K  V  I  E  S  R  A  Q  K  D  E  E  K
```

421 ATGGAAATTCAGGAGATCCAGCTGAAAGAAGCCAAGCACATTGCTGAGGATGCCGACCGC

141 M E I Q E I Q L K E A K H I A E D A D R

481 AAGTATGAAGAGGTGGCCCGTAAGCTGGTCATCATTGAGAGTGACCTGGAACGTGCCGAG

161 K Y E E V A R K L V I I E S D L E R A E

541 GAGCGGGCTGAACTCTCAGAAGGCAAATGTGCCGAGCTTGAAGAAGAGTTGAAAACTGTG

181 E R A E L S E G K C A E L E E E L K T V

601 ACGAACAACTTGAAGTCACTGGAGGCTCAGGCTGAGAAGTACTCACAGAAGGAAGACAAA

201 T N N L K S L E A Q A E K Y S Q K E D K

661 TATGAAGAAGAGATCAAGGTCCTTTCTGACAAGCTGAAGGAGGCTGAAACTCGGGCCGAG

221 Y E E E I K V L S D K L K E A E T R A E

721 TTTGCAGAGAGGTCAGTAACTAAATTGGAGAAAAGCATTGATGACTTAGAAGACGAGCTG

241 F A E R S V T K L E K S I D D L E D E L

781 TACGCTCAGAAACTGAAGTACAAAGCCATCAGCGAGGAGCTGGACCACGCTCTCAACGAT

261 Y A Q K L K Y K A I S E E L D H A L N D

841 ATGACTTCCATATAA

281 M T S I *

Figure 1.
Nucleotide and deduced amino acid sequence of Tropo 1. The predicted initiation Met start and stop codon (TAA) is in red. Highlighted in yellow are the forward and reverse primer sequences. Highlighted in green are likely regions of tropomyosin IgE-binding epitopes based on previously known epitopes. Underlined is the tropomyosin signature at amino acid position 232 – 240.

>NM_001001632 (Sus scrofa tropomyosin 3 (TPM3)

1 ATGGCTGGAATCACCACCATCGAGGCGGTGAAGCGCAAGATCCAGGTTTTGCAGCAGCAG

1 M A G I T T I E A V K R K I Q V L Q Q Q

61 GCCGATGATGCAGAGGAGAGGGCCGAGCGCCTCCAGCGGGAAGTGGAGGGGGAAAGGCGG

21 A D D A E E R A E R L Q R E V E G E R R

121 GCCCGGGAACAGGCTGAGGCTGAGGTGGCCTCCTTGAACCGTAGGATCCAGCTGGTTGAA

41 A R E Q A E A E V A S L N R R I Q L V E

181 GAGGAGCTGGACCGCGCTCAAGAGCGCCTGGCCACTGCCCTGCAAAAGCTGGAGGAAGCT

61 E E L D R A Q E R L A T A L Q K L E E A

241 GAGAAAGCTGCTGATGAGAGCGAGAGAGGTATGAAGGTCATTGAAAACCGAGCCTTAAAA

81 E K A A D E S E R G M K V I E N R A L K

301 GATGAGGAAAAAATGGAACTCCAGGAAATCCAACTCAAAGAAGCTAAGCACATTGCAGAA

101 D E E K M E L Q E I Q L K E A K H I A E

361 GAGGCAGATAGGAAGTATGAAGAGGTGGCTCGTAAGTTGGTGATTATTGAGGGAGACTTG

121 E A D R K Y E E V A R K L V I I E G D L

421 GAACGCACAGAGGAGCGAGCTGAGCTGGCAGAGTCCCGTTGCCGAGAGATGGATGAGCAG

141 E R T E E R A E L A E S R C R E M D E Q

481 ATCAGACTGATGGACCAGAATCTGAAGTGTCTGAGTGCTGCTGAAGAAAAGTACTCTCAA

161 I R L M D Q N L K C L S A A E E K Y S Q

541 AAAGAAGACAAATATGAGGAAGAGATAAAGATTCTCACTGACAAACTCAAGGAGGCAGAG

181 K E D K Y E E E I K I L T D K L K E A E

601 ACCCGGGCCGAGTTTGCCGAGAGATCGGTAGCCAAGCTGGAAAAGACAATTGATGACTTG

201 T R A E F A E R S V A K L E K T I D D L

661 GAAGATAAACTGAAATGCACCAAAGAGGAGCACCTCTGTACACAAAGGATGCTGGACCAG

221 E D K L K C T K E E H L C T Q R M L D Q

721 ACTCTGCTCGACCTGAATGAGATGTAG

241 T L L D L N E M *

Figure 2.
Nucleotide and deduced amino acid sequence of Tropo 3. The predicted initiation Met start and stop codon (TAG) is in red. Highlighted in yellow are the forward and reverse primer sequences. Highlighted in green are likely regions of tropomyosin IgE-binding epitopes based on previously known epitopes. Underlined is the tropomyosin signature at amino acid position 196 – 204.

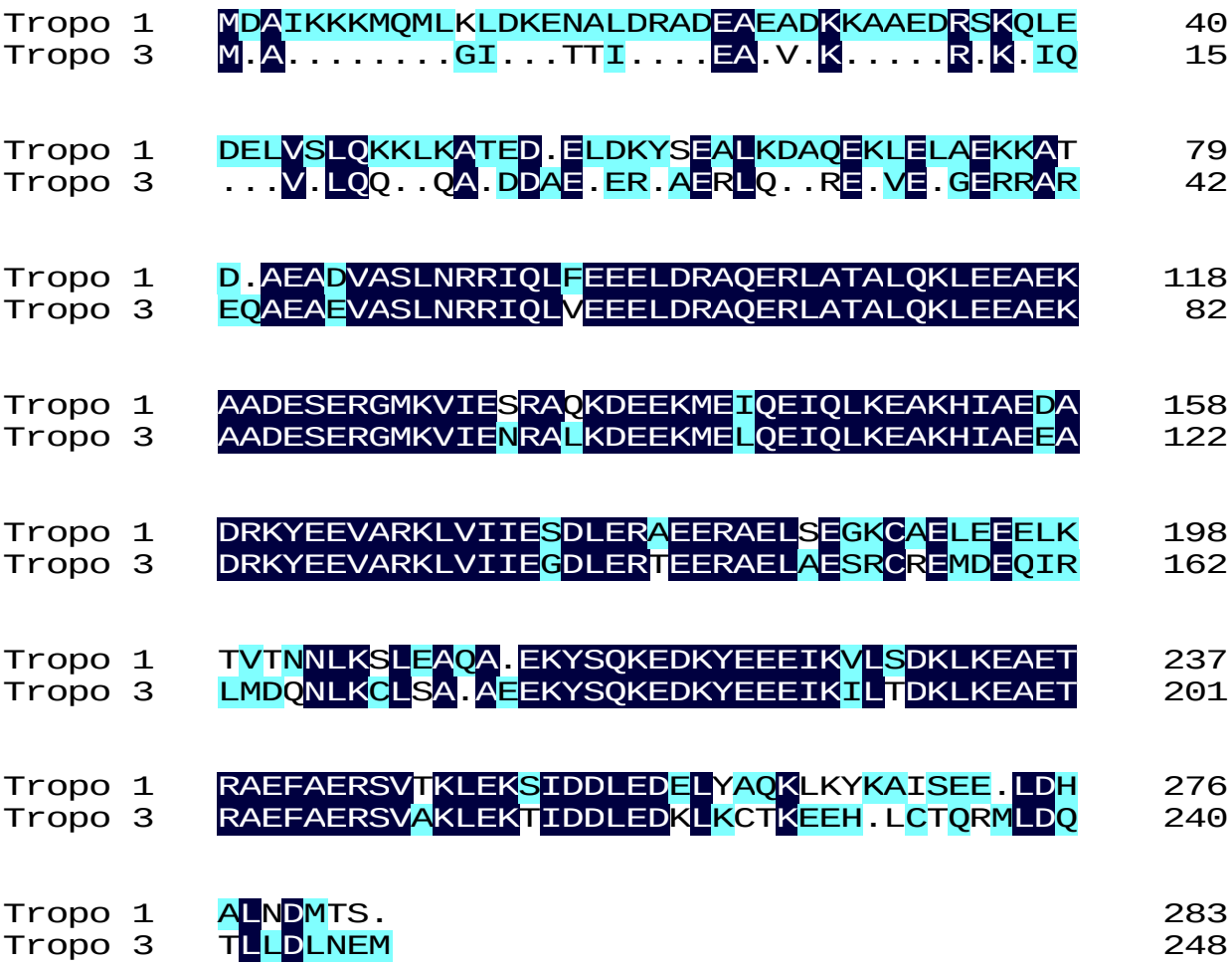


Figure 3.
Multiple sequence alignments between Tropo 1 and Tropo 3. Amino acid with 100% identity colored in black and more than 50% homology colored in blue. Dots have been introduced to maximize the alignments.

Total *Sus scrofa* RNA was used to amplify the cDNA coding for the tropomyosins using gene specific primers designed based on the known sequences. Various steps for cloning and agarose gel electrophoresis results were shown in Figure 4. Cloning and agarose gel electrophoresis results for other putative allergens are shown in appendix XV – XXII.

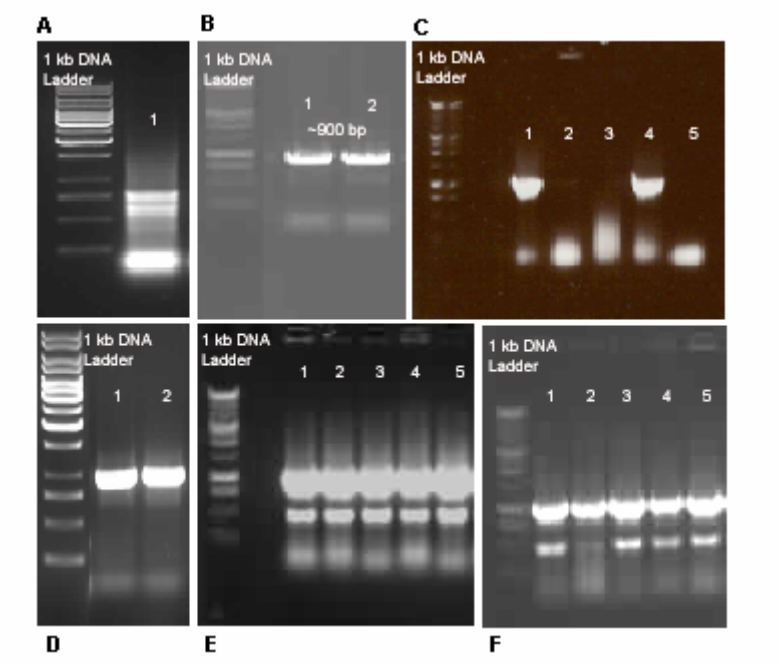


Figure 4.
Molecular cloning of Tropo 1. (A) RNA extraction: Agarose (1%) gel showing the total RNA extraction of meat from *Sus scrofa* using Trizol reagent in Lane 1. Distinct double bands were observed indicating integrity of the 28s and 18s RNA, however, there was an accumulation of 5S RNA. Nevertheless, the RNA was used for cDNA synthesis. **(B) PCR amplification of target tropomyosin gene** with gene specific primers using Expand long-template *Taq* DNA polymerase. Amplicons in Lane 1 and 2 corresponds to correct expected size of ~900 bp. PCR amplicons from both lanes were extracted and purified using QIA quick Gel extraction Kit (Qiagen) and ligated to pGEMT-Easy vector followed by transformation into XL-1 blue non-expression host. **(C) Colony screening of PCR inserts in pGEMT vector using SP6 and T7 primers:** A total of 10 colonies were screened for insert. Only five lanes were showed here (Lane: 1 to 5). Only 5 out of 10 clones showed the presence of insert with expected size of ~900 bp. **(D): PCR amplification of target gene from pGEM-T plasmids with correct insert using designed LIC primer adaptors.** Purified pGEM-T plasmid from clone 1 (Lane 1 of Fig C) was used. The PCR amplified band was gel extracted, and purified. The final digested DNA fragment was ligated into pET-32a (+) expression vector (Novagen, USA) using T4 DNA ligase (Invitrogen, USA) and transformed into XL1-Blue non-expression host cell. **(E): Colony screening of pET32a ligated insert** in transformed XL 1-blue non-expression host strain using LIC primers. Lane 1 to 5 corresponds to 5 clones chosen with the correct size of insert. **(F): Sub cloning of ligated Pet Vector plasmid into BL-21 (DE3) (Novagen, USA) expression host**. Again, colony screening was performed (lane 1 to 5). The clones were subsequently sequenced from both ends to check for correct reading frame. Clone 2 and Clone 4 were selected for protein expression. Glycerol stocks were made from those clones that were used for expression.

5.3.1.3 Troponin

Troponin is a protein complex that confers calcium sensitivity to muscle cells which is pivotal for muscle contraction (Wikipedia, 2006). It is found in both skeletal muscle and cardiac muscle, but the specific versions of troponin differ between types of muscle. Troponin is a component of thin filaments (along with actin and tropomyosin), and is the protein to which calcium binds that is essential for muscle contraction and relaxation regulation. Troponin has three subunits, TnC, TnI, and TnT. When calcium is bound to specific sites on TnC, the structure of the thin filament changes in such a manner that myosin (a molecular motor organized in muscle thick filaments) attaches to thin filaments and produces force and/or movement (Wikipedia, 2006). In the absence of calcium, tropomyosin interferes with this action of myosin, and therefore muscles remain relaxed.

Troponin has been identified as an allergen in cockroach *Blatella germanica* (Arruda *et al*., 1995). The identified cockroach allergen Bla g 6 is a troponin C whereas the putative allergen from *Sus scrofa* is a troponin T (TRNT). Both troponins are calcium-binding proteins which may be related to other calcium-binding allergen such as Bet v 3, Bos d 3 and Ole e 3. However, Bla g 6, Bet v 3, Bos d 3, and Ole e 3 showed weak overall sequence identity of 17.00%, 16.02%, 13.55% and 11.07% to TRNT respectively (Figure 6).

The full length sequence genes TRNT comprised of 753 bp encoding a 250 amino acid protein (Figure 5). One N-glycosylation site has been predicted in TRNT.

>NP_001001863 (Troponin T fast skeletal muscle type)

```
1          ATGTCGGACGAGGAAGTAGAACACGTCGAGGAAGAGTACGAGGAGGAAGAAGAGGCCCAG
1           M  S  D  E  E  V  E  H  V  E  E  E  Y  E  E  E  E  E  A  Q
61         GAGGAAGAGGAAGTTCAAGAAGAGGAGAAGCCGAGACCCAAACTCACTGCTCCTAAGATC
21          E  E  E  E  V  Q  E  E  E  K  P  R  P  K  L  T  A  P  K  I
121        CCGGAAGGGGAGAAAGTCGACTTTGATGACATCCAGAAGAAGCGCCAGAATAAGGACCTT
41          P  E  G  E  K  V  D  F  D  D  I  Q  K  K  R  Q  N  K  D  L
181        ATGGAGCTGCAGGCCCTCATCGACAGCCACTTCGAGGCTCGGAAGAAGGAGGAAGAGGAG
61          M  E  L  Q  A  L  I  D  S  H  F  E  A  R  K  K  E  E  E  E
241        CTGGTCGCTCTCAAGGAGAGAATCGAGAAGCGCCGTGCCGAGAGAGCCGAGCAGCAGAGG
81          L  V  A  L  K  E  R  I  E  K  R  R  A  E  R  A  E  Q  Q  R
301        ATCCGGGCTGAGAAGGAGCGGGAGCGCCAGAACAGGCTGGCGGAGGAGAAGGCCCGGCGG
101         I  R  A  E  K  E  R  E  R  Q  N  R  L  A  E  E  K  A  R  R
361        GAGGAGGAGGAAGCCAAGAGAAGGGCAGAGGACGACCTGAAGAAAAAGAAGGCGCTGTCC
121         E  E  E  E  A  K  R  R  A  E  D  D  L  K  K  K  K  A  L  S
421        TCCATGGGCGCCAACTACAGCAGCTACCTGGCCAAGGCCGAYCAGAAGCGAGGCAAGAAG
141         S  M  G  A  N  Y  S  S  Y  L  A  K  A  D  Q  K  R  G  K  K
481        CAGACGGCCCGGGAGATGAAGAAGAAGGTGCTGGCCGAGCGGAGGAAGCCCCTCAACATC
161         Q  T  A  R  E  M  K  K  K  V  L  A  E  R  R  K  P  L  N  I
541        GACCACCTCAGTGAGGACAAGCTGAGGGACAAGGCCAAGGAGCTCTGGGACGCCCTGTAC
181         D  H  L  S  E  D  K  L  R  D  K  A  K  E  L  W  D  A  L  Y
601        CARCTGGAGATTGACAAGTTCGAGTACGGGGAGAAGCTGAAGCGCCAGAAATACGACATC
201         Q  L  E  I  D  K  F  E  Y  G  E  K  L  K  R  Q  K  Y  D  I
661        ATCAACCTCAGAAGCCGCATCGACCAGGCCCAGAAGCACAGCAAGAAGGCCGGGACGACG
221         I  N  L  R  S  R  I  D  Q  A  Q  K  H  S  K  K  A  G  T  T
```

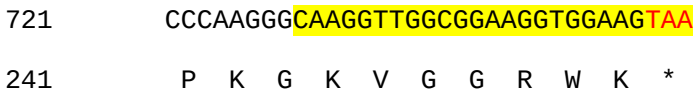


Figure 5.
Nucleotide and deduced amino acid sequence of TRNT. The predicted initiation Met start and stop codon (TAG) is in red. Highlighted in yellow are the forward and reverse primer sequences. Highlighted in red is the predicted N-glycosylation site.

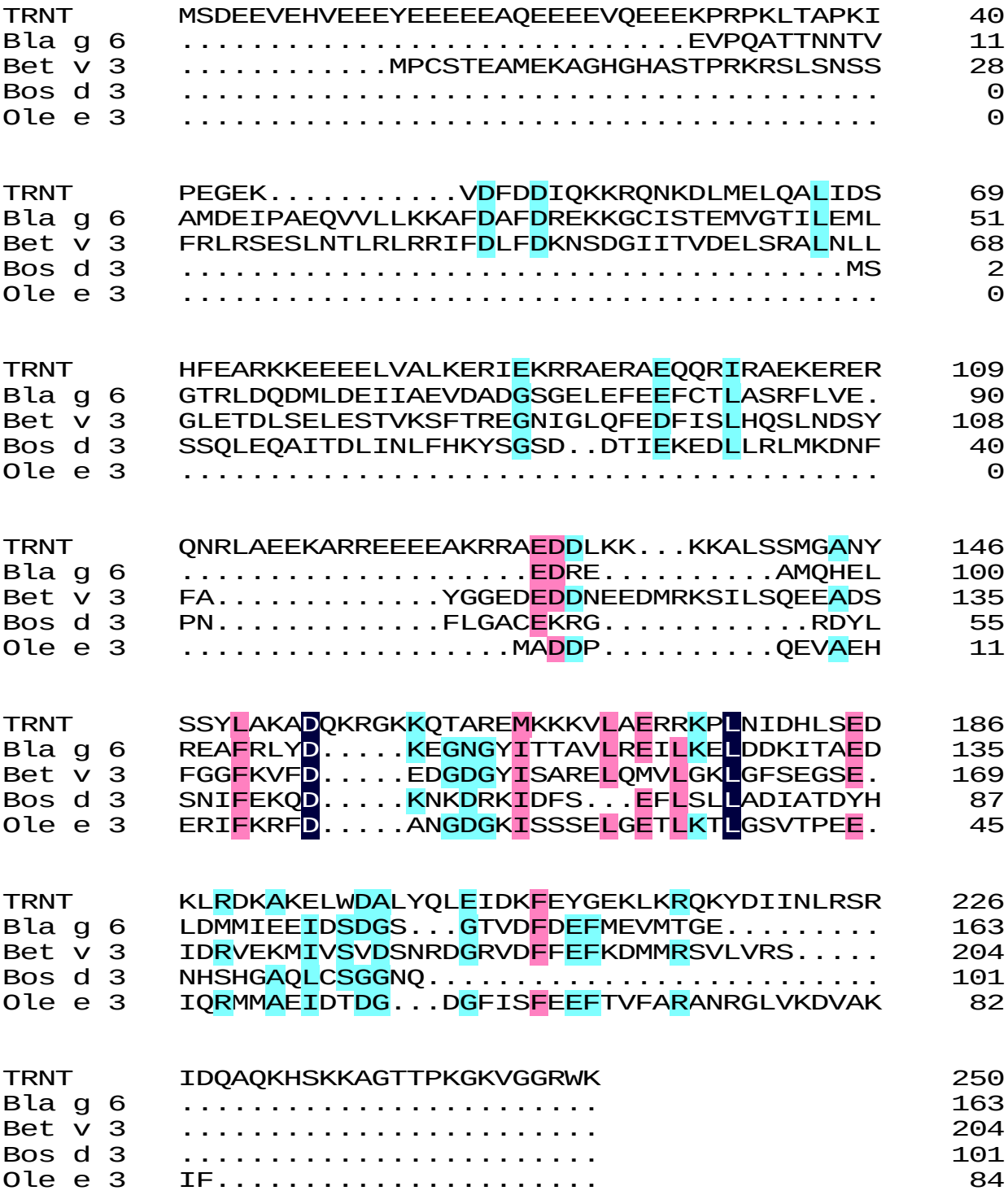


Figure 6.
Multiple sequence alignments between TRNT, Bla g 6, Bet v 3, Bos d 3 and Ole e 3.
Amino acid with 100% identity colored in black, 75% homology colored in pink and 50% homology colored in blue.. Dots have been introduced to maximize the alignments

5.3.1.4 Myosin-light chain

Myosin protein is a heteropolymer composed of two heavy chains (approximately 200,000 MW) and two pair of light chains. The light chain pair have been termed essential light chain 1 (25,000 MW) and regulatory light chain 2 (19,000 MW). The light chains that associate with each heavy chain in the neck region are calmodulin or calmodulin-like proteins. Light chains may provide structural support in neck domains and in some cases have a regulatory role. Another study has suggested that the essential light chain could form a link between thin and thick filaments to modulate the actomyosin interaction. The myosin light chain may play an important role in modulating the kinetics of catalytic activity in myosin which is mainly determined by the heavy chain structure (Sugiura *et al*., 2002).

Although currently no allergen has been classified as a myosin-light chain, our studies have shown that this protein is potentially allergenic. The possible account that myosin light chain could be allergenic may be related to its association with actin, which has been shown to bind IgE with meat allergic patients' sera (Bourne *et al*., 2005). Cross-linking of myosin from fast muscle with actin leads to, amongst others, the formation of a Myosin light chain -1F-actin complex (Andreev and Borejdo, 1995). The formation of Myosin light chain – 1F-actin complex indicates that myosin light chain may share similar conformational or linear epitopes with actin hence eliciting an immune response. Thus the myosin light chain cDNA clone was constructed by PCR from *S. scrofa* cDNA.

The cloned myosin light chain (Myo_L) comprised of 453 bp encoding a 150 amino acid protein (Figure 7). There are three N-glycosylation sites predicted in Myo_L.

>NM_214374 (*Sus scrofa* myosin light chain (MYL1)

```
1          ATGTCCTTCAGTGCTGACCAGATTGCTGAATTCAAGGAGGCATTTCTCCTCTTTGACAGA
1           M  S  F  S  A  D  Q  I  A  E  F  K  E  A  F  L  L  F  D  R
61         ACAGGCGAATGCAAGATCACCCTAAGCCAGGTTGGTGATGTCCTTCGGGCTCTGGGCACA
21          T  G  E  C  K  I  T  L  S  Q  V  G  D  V  L  R  A  L  G  T
121        AATCCCACCAATGCAGAGGTCAAGAAGGTTCTGGGAAACCCCAGCAATGAAGAGATGAAT
41          N  P  T  N  A  E  V  K  K  V  L  G  N  P  S  N  E  E  M  N
181        GCCAAGAAAATTGAGTTTGAACAATTCCTGCCTATGCTGCAAGCTATTTCCAACAACAAG
61          A  K  K  I  E  F  E  Q  F  L  P  M  L  Q  A  I  S  N  N  K
241        GACCAGGGAAGCTATGAAGACTTTGTTGAGGGTCTGCGTGTCTTTGACAAGGAAGGCAAT
81          D  Q  G  S  Y  E  D  F  V  E  G  L  R  V  F  D  K  E  G  N
301        GGTACAGTCATGGTGACTGAACTTCGTCATGTTCTAGCTACACTAGGTGAAAAGATGAAA
101         G  T  V  M  V  T  E  L  R  H  V  L  A  T  L  G  E  K  M  K
361        GAGGAAGAAGTGGAAGCCCTGATGGCAGGTCAAGAAGACTCCAATGGCTGCATCAACTAT
121         E  E  E  V  E  A  L  M  A  G  Q  E  D  S  N  G  C  I  N  Y
421        GAAGCCTTTGTCAAGCACATCATGTCTATCTAA
141         E  A  F  V  K  H  I  M  S  I  *
```

Figure 7.
Nucleotide and deduced amino acid sequence of Myo_L. The predicted initiation Met start and stop codon (TAG) is in red. Highlighted in yellow are the forward and reverse primer sequences. Highlighted in red are the predicted N-glycosylation sites.

5.3.1.5 Aldehyde dehydrogenase

Aldehyde dehydrogenases have been found in nearly every form of living thing. Their primary role in humans and other mammals is protecting the body from toxic compounds called aldehydes. Aldehyde dehydrogenase, often in tandem with alchohol dehydrogenase, acts in detoxifying a wide variety of organic compounds, toxins and pollutants hence it is also classify as a stress response protein. The bovine aldehyde dehydrogenase displayed high sequence homology to known aldehyde dehydrogenase fungal allergens from *Alternaria alternate* (Alt a 10) and *Cladosporium herbarum* (Cla h 3) at 49.3% and 50.5% respectively (Figure 9).

Thus, the aldehyde dehydrogenase cDNA clone (ADH) was constructed by PCR from *B. taurus* cDNA. The ADH sequence comprised of 1506 bp encoding a 501 amino acid protein (Figure 8). The PROSITE program (Hoffman *et al.,* 1999) detected a conserved glutamic acid and cysteine residues in the sequence which was implicated in the catalytic activity of mammalian aldehyde dehydrogenase. One N-glycosylation site was also predicted in ADH.

> NM_174239 (*Bos taurus* aldehyde dehydrogenase 1 family, member A1 (ALDH1A1)

```
1          ATGTCGTCCTCAGCCATGCCAGACGTACCTGCCCCACTCACCAATTTGCAGTTTAAATAT
1           M  S  S  S  A  M  P  D  V  P  A  P  L  T  N  L  Q  F  K  Y
61         ACTAAGATCTTCATAAACAATGAATGGCATAGTTCAGTGAGTGGTAAGAAATTTCCAGTC
```

21 T K I F I N N E W H S S V S G K K F P V

121 TTTAATCCTGCAACTGAGGAGAAACTCTGTGAGGTGGAAGAAGGAGATAAGGAGGATGTT

41 F N P A T E E K L C E V E E G D K E D V

181 GACAAAGCAGTGAAGGCTGTAAGACAAGCTTTTCAGATTGGCTCTCCATGGCGTACTATG

61 D K A V K A V R Q A F Q I G S P W R T M

241 GATGCTTCAGAGAGAGGACGGCTGTTAAACAAGTTGGCTGACTTAATTGAAAGAGATCAT

81 D A S E R G R L L N K L A D L I E R D H

301 CTGCTCCTGGCGACAATGGAGGCAATGAATGGTGGAAAACTATTTTCCAATGCATATCTG

101 L L L A T M E A M N G G K L F S N A Y L

361 ATGGATTTAGGAGGCTGCATAAAAACACTACGCTACTGTGCAGGCTGGGCTGACAAGATC

121 M D L G G C I K T L R Y C A G W A D K I

421 CAGGGCCGCACAATACCCATGGATGGAAACTTTTTTACATATACAAGAAGTGAGCCTGTT

141 Q G R T I P M D G N F F T Y T R S E P V

481 GGTGTGTGTGGCCAAATCATTCCTTGGAATTTCCCATTGCTCATGTTCCTCTGGAAGATA

161 G V C G Q I I P W N F P L L M F L W K I

541 GGGCCTGCCCTTAGCTGCGGAAACACAGTGGTTGTCAAACCAGCAGAGCAAACCCCTCTG

181 G P A L S C G N T V V V K P A E Q T P L

601 ACTGCTCTTCACATGGGATCTTTAATAAAAGAGGCAGGGTTTCCTCCTGGAGTAGTGAAT

201 T A L H M G S L I K E A G F P P G V V N

661 ATTGTCCCTGGTTATGGGCCTACTGCAGGGGCAGCCATTTCTTCTCACATGGATGTAGAC

221 I V P G Y G P T A G A A I S S H M D V D

721 AAAGTGGCCTTCACAGGATCGACAGAGGTTGGCAAACTGATCAAAGAAGCTGCTGGGAAA

241 K V A F T G S T E V G K L I K E A A G K

781 AGCAATCTGAAAAGGGTGTCCCTGGAACTCGGGGGAAAGAGTCCTTGCATTGTGTTTGCT

261 S N L K R V S L E L G G K S P C I V F A

841 GATGCCGACTTGGACAATGCTGTTGAATTTGCACACCAAGGAGTATTCTATCACCAGGGC

281 D A D L D N A V E F A H Q G V F Y H Q G

901 CAGTGTTGTATAGCTGCATCCCGTCTCTTTGTAGAAGAATCAATTTACGATGAGTTTGTT

301 Q C C I A A S R L F V E E S I Y D E F V

961 CGAAGGAGTGTTGAGCGGGCGAAAAAGTATGTTCTTGGAAATCCTCTGACCCCAGGAGTC

321 R R S V E R A K K Y V L G N P L T P G V

1021 AGTCAAGGCCCTCAGATTGATAAAGAACAATATGAAAAAATACTTGACCTCATTGAAAGT

341 S Q G P Q I D K E Q Y E K I L D L I E S

1081 GGGAAGAAGGAGGGGGCCAAGCTGGAATGTGGCGGAGGCCCTTGGGGGAATAAAGGCTAC

361 G K K E G A K L E C G G G P W G N K G Y

1141 TTTATCCAACCCACAGTTTTCTCTGATGTTACTGATGATATGCGCATTGCCAAAGAGGAG

381 F I Q P T V F S D V T D D M R I A K E E

1201 ATATTTGGACCTGTGCAGCAAATCATGAAGTTTAAGTCTTTAGATGATGTAATCAAGAGA

401 I F G P V Q Q I M K F K S L D D V I K R

1261 GCAAACAATACTTTCTATGGGTTATCTGCAGGAATATTTACCAATGATATTGATAAAGCC

421 A N N T F Y G L S A G I F T N D I D K A

1321 ATCACAGTCTCCTCTGCTTTGCAGTCTGGAACCGTGTGGGTGAACTGCTATAGTGTGGTA

441 I T V S S A L Q S G T V W V N C Y S V V

1381 TCTGCCCAGTGCCCCTTTGGTGGATTCAAGATGTCTGGAAATGGTCGAGAACTGGGAGAA

461 S A Q C P F G G F K M S G N G R E L G E

1441 TATGGTTTCCATGAATACACAGAAGTCAAGACGGTCACAATCAAAATTTCTCAGAAGAAC

481 Y G F H E Y T E V K T V T I K I S Q K N

1501 TCATAA

501 S *

Figure 8.
Nucleotide and deduced amino acid sequence of ADH. The predicted initiation Met start and stop codon (TAG) is in red. Highlighted in yellow are the forward and reverse primer sequences. Highlighted in red is the predicted N-glycosylation site. Underlined are the conserved glutamic acid site and cysteine site which are located at amino acid positions 268 - 275 and 296 - 307 respectively.

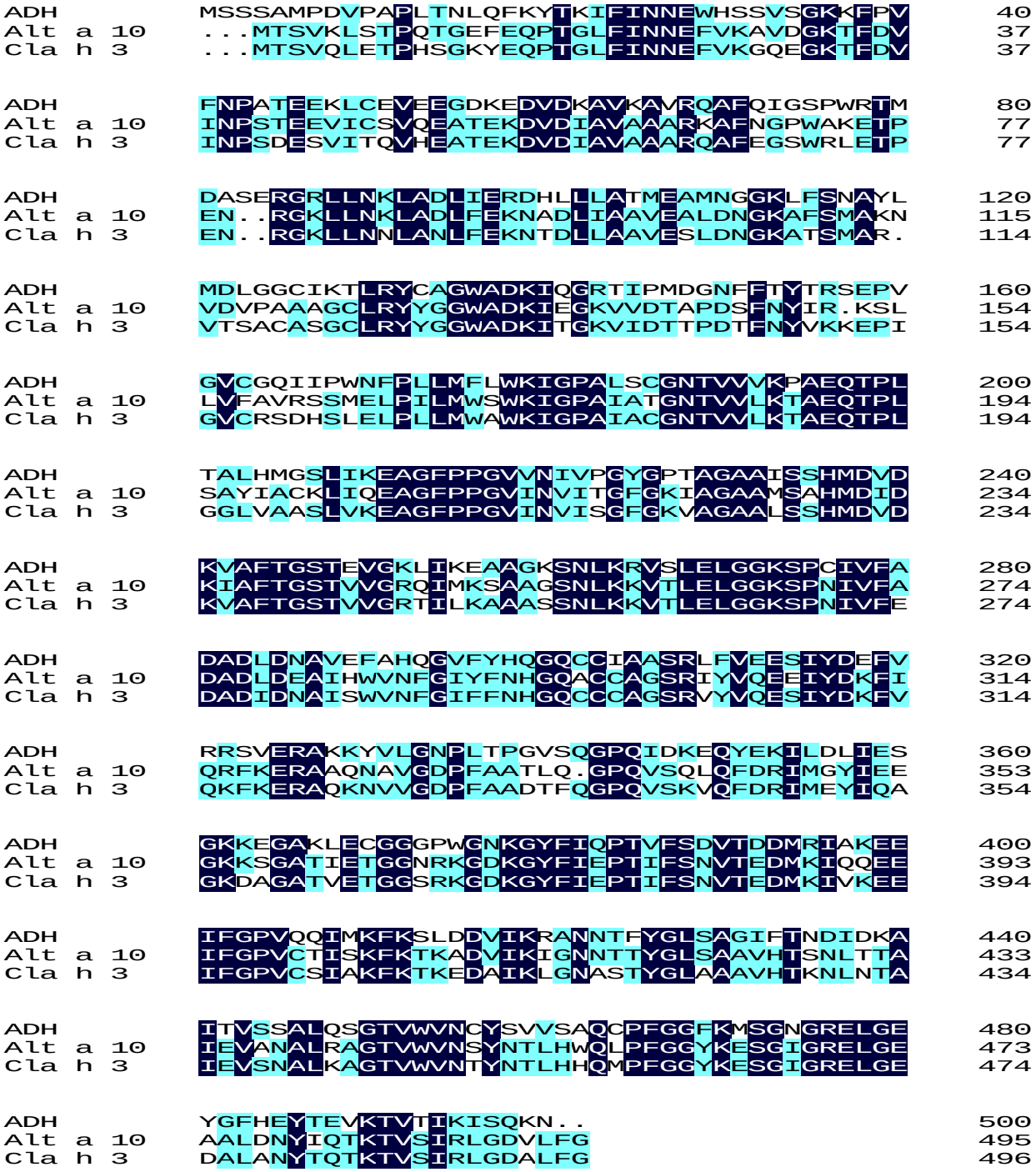


Figure 9.

Multiple sequence alignments between ADH, Alt a 10 and Cla h 3. Amino acid with 100% identity colored in black and more than 50% homology colored in blue. Dots have been introduced to maximize the alignments.

5.3.1.6 Enolase

Enolase is an enzyme that is required for glycolysis. Enolase catalyzes the conversion of 2-phosphoglycerate (2PG) to phosphoenolpyruvate (PEP), the penultimate step in the conversion of glucose to pyruvate (Wikipedia, 2006). The enzyme is a member of an enzyme superfamily (the enolase superfamily), whereby a group of evolutionarily related enzymes share similar reaction mechanisms but function on different substrates (with altered substrate affinity) and in different biochemical context. There are basically five distinct enolase isozymes, namely alpha enolase is ubiquitous in the cytoplasm of cells, beta is associated with muscle tissue and gamma is neuronal specific (Wikipedia, 2006).

Enolase is a phylogenetically conserved enzyme in terms of amino acid sequence. It is well known as a pan-allergen among the molds (Breitenbach *et al.,* 1997; Simon-Nobbe *et al.,* 2000) and recently also identified as a cross-reacting allergen in latex (*Hevea. brasiliensis*), Hev b 9 (Wagner *et al.,* 2000). Thus, enolase from *Bos taurus* is likely to be allergenic due to its high sequence homology.

The *Bos taurus* enolase (ENO 1) comprised of 1305 bp encoding a 434 amino acid protein (Figure 10). It possessed three predicted N-glycosyaltion sites and the unique enolase signature.

>NM_174049 (*Bos taurus* enolase 1 (ENO1)

```
1          ATGTCCATCCTGAAGGTCCACGCCAGAGAGATCTTTGACTCTCGTGGGAATCCCACCGTT
1           M  S  I  L  K  V  H  A  R  E  I  F  D  S  R  G  N  P  T  V
61         GAGGTTGATCTCTTCACCGCGAAAGGTCTCTTCAGAGCTGCTGTGCCCAGTGGCGCTTCA
21          E  V  D  L  F  T  A  K  G  L  F  R  A  A  V  P  S  G  A  S
121        ACTGGAATCTATGAGGCCCTGGAGCTCCGGGACAATGATAAGACGCGCTACATGGGGAAG
41          T  G  I  Y  E  A  L  E  L  R  D  N  D  K  T  R  Y  M  G  K
181        GGTGTCTCAAAGGCTGTTGAGCACATCAATAAAACTATTGCGCCTGCCCTGGTTAGCAAG
61          G  V  S  K  A  V  E  H  I  N  K  T  I  A  P  A  L  V  S  K
241        AAGTCGAACGTCGTGGAGCAGGAGAAGATCGACAAGCTGATGATAGAGATGGATGGCACA
81          K  S  N  V  V  E  Q  E  K  I  D  K  L  M  I  E  M  D  G  T
301        GAAAAAAAATCTAAGTTTGGTGCGAACGCCATCCTGGGCGTGTCCCTGGCTGTCTGCAAA
101         E  K  K  S  K  F  G  A  N  A  I  L  G  V  S  L  A  V  C  K
361        GCTGGTGCTGTGGAGAAGGGGGTGCCCCTCTACCGCCACATCGCCGACTTGGCTGGCAAT
121         A  G  A  V  E  K  G  V  P  L  Y  R  H  I  A  D  L  A  G  N
421        GCTGAGGTCATCCTGCCAGTTCCGGCTTTCAATGTCATCAACGGTGGCTCTCATGCTGGC
141         A  E  V  I  L  P  V  P  A  F  N  V  I  N  G  G  S  H  A  G
481        AACAAGCTGGCCATGCAGGAGTTTATGATCCTTCCTGTTGGGGCCGAAAACTTCCGGGAG
161         N  K  L  A  M  Q  E  F  M  I  L  P  V  G  A  E  N  F  R  E
541        GCCATGCGCATCGGAGCAGAGGTTTACCACAACCTGAAGAATGTCATCAAGGAGAAATAT
181         A  M  R  I  G  A  E  V  Y  H  N  L  K  N  V  I  K  E  K  Y
601        GGGAAGGATGCCACCAACGTGGGAGATGAGGGCGGCTTTGCCCCCAACATCCTGGAGAAC
201         G  K  D  A  T  N  V  G  D  E  G  G  F  A  P  N  I  L  E  N
661        AAAGAAGCCCTGGAGCTGCTGAAGAATGCCATCGGCAAGGCTGGCTACAGCGACAAGGTC
221         K  E  A  L  E  L  L  K  N  A  I  G  K  A  G  Y  S  D  K  V
721        GTCATCGGCATGGACGTAGCTGCCTCTGAGTTCTACAGGTCGGGCAAGTATGACCTGGAC
241         V  I  G  M  D  V  A  A  S  E  F  Y  R  S  G  K  Y  D  L  D
781        TTCAAGTCGCCCGATGACCCCAGCAGGTACATCACACCCGACGAGCTGGCCAACCTGTAC
```

```
261        F  K  S  P  D  D  P  S  R  Y  I  T  P  D  E  L  A  N  L  Y
841        AAGTCCTTCATCAGGGACTACCCAGTGGTGTCTATCGAAGATCCCTTCGACCAAGATGAC
281        K  S  F  I  R  D  Y  P  V  V  S  I  E  D  P  F  D  Q  D  D
901        TGGGAAGCTTGGCAGAAGTTCACTGCCAGCGCAGGGATCCAGGTGGTGGGGGATGATCTC
301        W  E  A  W  Q  K  F  T  A  S  A  G  I  Q  V  V  G  D  D  L
961        ACGGTGACAACCCCAAAGACGATCGCCAAGGGCGTGAACGAAAAATCCTGCAACTGCCTC
321        T  V  T  T  P  K  T  I  A  K  G  V  N  E  K  S  C  N  C  L
1021       CTGCTGAAAGTCAACCAGAATGGCTCTGTGACCGAGTCTCTTCAGGGGTGCAAGCTGGCC
341        L  L  K  V  N  Q  N  G  S  V  T  E  S  L  Q  G  C  K  L  A
1081       CATGCCAACGGGTGGGGCGTCATGGTTTCTCATCGTTCGGGGGAGACTGATGATACCTTC
361        H  A  N  G  W  G  V  M  V  S  H  R  S  G  E  T  D  D  T  F
1141       ATCGCTGAACTGGTGGTGGGGCTGTGCACTGGGCAGATCAAGAATGGTCCCCCTTGCCGT
381        I  A  E  L  V  V  G  L  C  T  G  Q  I  K  N  G  P  P  C  R
1201       ACTGAGCGCTTGGCCAAGTACAACCAGATCCTCAGAATTGAAGAGGAATTGGGCAGCAAG
401        T  E  R  L  A  K  Y  N  Q  I  L  R  I  E  E  E  L  G  S  K
1261       GCTAAGTTTGCCGGCAGGAACTTCAGAAACCCGTTGGCCAAGTAA
421        A  K  F  A  G  R  N  F  R  N  P  L  A  K  *
```

Figure 10.
Nucleotide and deduced amino acid sequence of ENO 1. The predicted initiation Met start and stop codon (TAG) is in red. Highlighted in yellow are the forward and reverse primer sequences. Highlighted in red are the predicted N-glycosylation sites. Underlined is the enolase signature at amino acid positions 340 – 353.

5.3.1.7 Heat shock proteins

Heat shock proteins (HSPs) are a group of proteins whereby expressions are up-regulated when the cells are exposed to elevated temperatures. This dramatic up-regulation of the heat shock proteins plays a key role in heat shock/stress response. Other factors that trigger production of high levels of heat shock proteins include exposure to different kinds of environmental stress conditions, such as infection, inflammation, exposure of the cell to toxins (ethanol, arsenic, trace metals and ultraviolet light, among many others), starvation, hypoxia (oxygen deprivation), nitrogen deficiency (in plants), or water deprivation (Wikipedia, 2006) . Heat shock proteins are also molecular chaperones for protein molecules which play an important role in protein-protein interactions such as folding and assisting in the establishment of proper protein conformation (shape) and prevention of unwanted protein aggregation (Wikipedia, 2006). This essential role is also known as protein maintenance.

Besides its essential role in stress response and protein maintanence, HSPs are also common fungal allergens known to widely cross reactive across the different families (Shen *et al*., 1997). Additionally, HSP70 are also common allergens in barley, corn and hazel pollen (Chiung *et al*., 2000; Gruehn *et al*., 2003). Thus, three heat shock proteins namely pHSP70, bHSP70 and HSP90 from both *Sus scrofa* and *Bos taurus* were cloned to determine their IgE-binding reactivity

The pHSP70 (Figure 11), bHSP70 (Figure 12)and HSP90 (Figure 13)cDNAs comprised of 1926 bp, 1926 bp and 2202 bp, which encoded a polypeptide of 641, 641, and 733 amino acids respectively. All the three heat shock proteins possessed N-glycosylation sites and the heat shock protein signature sequences.

>AY466608 (*Sus scrofa* heat shock protein 70.2 (hsp70.2)

```
1          ATGGCGAAGAGCGTGGCCATCGGCATCGACCTGGGCACCACGTACTCGTGCGTGGGGGTG
1           M  A  K  S  V  A  I  G  I  D  L  G  T  T  Y  S  C  V  G  V
61         TTCCAGCACGGCAAGGTGGAGATCATCGCCAACGACCAGGGCAACCGCACCACCCCCAGC
21          F  Q  H  G  K  V  E  I  I  A  N  D  Q  G  N  R  T  T  P  S
121        TACGTGGCCTTCACGGACACCGAGCGGCTGATCGGCGATGCGGCCAAGAACCAGGTGGCG
41          Y  V  A  F  T  D  T  E  R  L  I  G  D  A  A  K  N  Q  V  A
181        CTGAACCCGCAGAACACGGTGTTTGACGCGAAGCGGCTGATCGGGCGCAAGTTCGGCGAC
61          L  N  P  Q  N  T  V  F  D  A  K  R  L  I  G  R  K  F  G  D
241        CCGGTGGTGCAGGCGGACATGAAGCACTGGCCCTTCCGGGTGATCAACGACGGGGACAAG
81          P  V  V  Q  A  D  M  K  H  W  P  F  R  V  I  N  D  G  D  K
301        CCCAAGGTGCAGGTGAGCTACAAGGGCGAGACCAAGGCGTTCTACCCGGAGGAGATCTCG
101         P  K  V  Q  V  S  Y  K  G  E  T  K  A  F  Y  P  E  E  I  S
361        TCGATGGTGCTGACCAAGATGAAGGAGATCGCCGAGGCGTACCTGGGCCACCCGGTGAGC
121         S  M  V  L  T  K  M  K  E  I  A  E  A  Y  L  G  H  P  V  S
421        AACGCGGTGATCACGGTGCCGGCCTACTTCAACGACTCGCAGCGGCAGGCCACCAAGGAT
141         N  A  V  I  T  V  P  A  Y  F  N  D  S  Q  R  Q  A  T  K  D
481        GCGGGGGTGATCGCGGGGCTGAACGTGCTGCGGATCATCAACGAGCCCACGGCGGCGGCC
161         A  G  V  I  A  G  L  N  V  L  R  I  I  N  E  P  T  A  A  A
541        ATCGCCTACGGCCTGGACAGGACGGGCAAGGGGGAGCGCAACGTGCTGATCTTCGACCTG
181         I  A  Y  G  L  D  R  T  G  K  G  E  R  N  V  L  I  F  D  L
```

601 GGCGGGGGCACGTTCGACGTGTCCATCCTGACGATCGACGACGGCATCTTCGAGGTGAAG

201 G G G T F D V S I L T I D D G I F E V K

661 GCCACGGCGGGGGACACGCACCTGGGCGGCGAGGACTTCGACAACAGGCTGGTGAACCAC

221 A T A G D T H L G G E D F D N R L V N H

721 TTCGTGGAGGAGTTCAAGAGGAAGCACAAGAAGGACATCAGCCAGAACAAGCGGGCGGTG

241 F V E E F K R K H K K D I S Q N K R A V

781 AGGCGGCTGCGCACGGCCTGCGAGCGGGCCAAGAGGACCCTGTCGTCCAGCACACAGGCC

261 R R L R T A C E R A K R T L S S S T Q A

841 AGCCTGGAGATCGACTCCCTGTTCGAGGGCATCGACTTCTACACGTCCATCACCCGGGCG

281 S L E I D S L F E G I D F Y T S I T R A

901 CGCTTCGAGGAGCTGTGCTCGGACCTGTTCCGCAGCACCCTGGAGCCGGTGGAGAAGGCT

301 R F E E L C S D L F R S T L E P V E K A

961 CTGCGCGACGCGAAGCTGGACAAGGCCCAGATCCACGACCTGGTGCTGGTGGGGGGCTCG

321 L R D A K L D K A Q I H D L V L V G G S

1021 ACGCGCATCCCCAAGGTGCAGAAGCTGCTGCAGGACTTCTTCAACGGGCGCGACCTCAAC

341 T R I P K V Q K L L Q D F F N G R D L N

1081 AAGAGCATCAACCCGGACGAGGCGGTGGCGTATGGGGCGGCGGTGCAGGCGGCCATCCTG

361 K S I N P D E A V A Y G A A V Q A A I L

1141 ATGGGCGACAAGTCGGAGAACGTGCAGGACCTGCTGCTGCTGGACGTGGCCCCGCTGTCG

381 M G D K S E N V Q D L L L L D V A P L S

1201 CTGGGGCTGGAGACGGCCGGCGGCGTGATGACGGCGCTGATCAAGCGCAACTCCACCATC

401 L G L E T A G G V M T A L I K R N S T I

1261 CCCACCAAGCAGACGCAGATCTTCACCACGTACTCGGACAACCAGCCGGGCGTGCTGATC

421 P T K Q T Q I F T T Y S D N Q P G V L I

1321 CAGGTGTACGAGGGCGAGAGGGCCATGACGCGGGACAACAACCTGCTGGGGCGCTTCGAG

441 Q V Y E G E R A M T R D N N L L G R F E

1381 CTGAGCGGCATCCCGCCGGCCCCGCGGGGGGTGCCCCAGATCGAGGTGACCTTCGACATC

461 L S G I P P A P R G V P Q I E V T F D I

1441 GACGCCAACGGCATCCTGAACGTCACGGCGACGGACAAGAGCACGGGCAAGGCCAACAAG

481 D A N G I L N V T A T D K S T G K A N K

1501 ATCACCATCACCAACGACAAGGGCCGGCTGAGCAAGGAGGAGATCGAGCGCATGGTGCAG

501 I T I T N D K G R L S K E E I E R M V Q

1561 GAGGCGGAGAAGTACAAAGCGGAGGACGAGATCCAGCGCGAGAGGGTGTCGGCCAAGAAC

521 E A E K Y K A E D E I Q R E R V S A K N

1621 GCGCTGGAGTCGTACGCCTTCAACATGAAGAGCGCCGTGGAGGATGAGGGGCTCAAGGGC

541 A L E S Y A F N M K S A V E D E G L K G

1681 AAGATCAGCGAGGCGGACAAGAAGAAGGTGCTGGACAAGTGTCAGGAGGTGATTTCCTGG

561 K I S E A D K K K V L D K C Q E V I S W

1741 CTGGACGCCAACACGCTGGCCGAGAAGGACGAGTTTGAGCACAAGAGGAAGGAGCTGGAG

581 L D A N T L A E K D E F E H K R K E L E

1801 CAGGTGTGTAACCCCATCATCAGCGGACTGTACCAGGGGGCGGGTGGCCCCGGGGCTGGC

601 Q V C N P I I S G L Y Q G A G G P G A G

1861 GGCTTTGGGGCTCAGGCCCCCAAAGGGGGCTCTGGGTCTGGCCCCACCATCGAGGAGGTG

621 G F G A Q A P K G G S G S G P T I E E V

1921 GATTAG

641 D *

Figure 11.
Nucleotide and deduced amino acid sequence of pHSP70. The predicted initiation Met start and stop codon (TAG) is in red. Highlighted in yellow are the forward and reverse primer sequences. Highlighted in red are the predicted N-glycosylation sites. Underlined are the three heat shock hsp70 proteins family signatures at amino acid positions 9 – 16, 197 – 210, and 334 – 348.

>AY662497 (*Bos taurus* heat shock protein 70 kDa)

```
1          ATGGCGAAAAACATGGCTATCGGCATCGACCTGGGCACCACCTACTCCTGCGTAGGGGTG
1           M  A  K  N  M  A  I  G  I  D  L  G  T  T  Y  S  C  V  G  V
61         TTCCAGCACGGCAAGGTGGAGATCATCGCCAACGACCAGGGCAACCGCACCACCCCCAGC
21          F  Q  H  G  K  V  E  I  I  A  N  D  Q  G  N  R  T  T  P  S
121        TACGTGGCCTTCACCGATACCGAGCGGCTCATCGGGGATGCGGCCAAGAGCCAGGTGGCG
41          Y  V  A  F  T  D  T  E  R  L  I  G  D  A  A  K  S  Q  V  A
181        CTGAACCCGCAGAACACGGTGTTCGACGCGAAGCGCCTGATCGGCCGCAAGTTCGGAGAC
61          L  N  P  Q  N  T  V  F  D  A  K  R  L  I  G  R  K  F  G  D
241        CCGGTGGTGCAGTCGGACATGAAGCACTGGCCTTTCCGCGTCATCAACGACGGAGACAAG
81          P  V  V  Q  S  D  M  K  H  W  P  F  R  V  I  N  D  G  D  K
301        CCTAAGGTGCAGGTGAGCTACAAAGGGGAGACCAAGGCGTTCTACCCGGAGGAGATCTCG
101         P  K  V  Q  V  S  Y  K  G  E  T  K  A  F  Y  P  E  E  I  S
361        TCGATGGTGCTGACCAAGATGAAGGAGATCGCCGAGGCGTACCTGGGCCACCCGGTGACC
121         S  M  V  L  T  K  M  K  E  I  A  E  A  Y  L  G  H  P  V  T
421        AACGCGGTGATCACCGTGCCGGCCTACTTCAACGACTCGCAGCGGCAGGCCACCAAGGAC
141         N  A  V  I  T  V  P  A  Y  F  N  D  S  Q  R  Q  A  T  K  D
481        GCGGGGGTGATCGCGGGGCTGAACGTGCTGAGGATCATCAACGAGCCCACGGCCGCCGCC
161         A  G  V  I  A  G  L  N  V  L  R  I  I  N  E  P  T  A  A  A
541        ATCGCCTACGGCCTGGACAGGACGGGCAAGGGGGAGCGCAACGTGCTCATCTTTGATCTG
181         I  A  Y  G  L  D  R  T  G  K  G  E  R  N  V  L  I  F  D  L
601        GGAGGGGGCACGTTCGACGTGTCCATCCTGACGATCGACGACGGCATCTTCGAGGTGAAG
201         G  G  G  T  F  D  V  S  I  L  T  I  D  D  G  I  F  E  V  K
661        GCCACGGCCGGGGACACGCACCTGGGCGGGGAGGACTTCGACAACAGGCTGGTGAACCAC
221         A  T  A  G  D  T  H  L  G  G  E  D  F  D  N  R  L  V  N  H
721        TTCGTGGAGGAGTTCAAGAGGAAGCACAAGAAGGACATCAGCCAGAACAAGCGGGCCGTG
241         F  V  E  E  F  K  R  K  H  K  K  D  I  S  Q  N  K  R  A  V
```

781 AGGCGGCTGCGCACCGCATGCGAGCGGGCCAAGAGAACCTTGTCGTCCAGCACCCAGGCC

261 R R L R T A C E R A K R T L S S S T Q A

841 AGCCTGGAGATCGACTCCCTGTTCGAGGGCATCGACTTCTACACGTCCATCACCAGGGCG

281 S L E I D S L F E G I D F Y T S I T R A

901 CGGTTCGAGGAGCTGTGCTCCGACCTGTTCCGGAGCACCCTGGAGCCCGTGGAGAAGGCG

301 R F E E L C S D L F R S T L E P V E K A

961 CTACGCGACGCCAAGCTGGACAAGGCGCAGATCCACGACCTGGTCCTGGTGGGGGGCTCC

321 L R D A K L D K A Q I H D L V L V G G S

1021 ACCCGCATCCCCAAGGTGCAGAAGCTGCTGCAGGACTTCTTCAACGGGCGCGACCTCAAC

341 T R I P K V Q K L L Q D F F N G R D L N

1081 AAGAGCATCAACCCCGACGAGGCGGTGGCGTACGGGGCGGCGGTGCAGGCGGCCATCCTG

361 K S I N P D E A V A Y G A A V Q A A I L

1141 ATGGGGGACAAGTCGGAGAACGTGCAGGACCTGCTGTTGCTGGACGTGGCTCCCCTGTCG

381 M G D K S E N V Q D L L L L D V A P L S

1201 CTGGGACTGGAGACGGCCGGAGGCGTGATGACCGCCCTGATCAAGCGCAACTCCACCATC

401 L G L E T A G G V M T A L I K R N S T I

1261 CCCACGAAGCAGACGCAGATCTTCACCACCTACTCGGACAACCAGCCGGGCGTGCTGATC

421 P T K Q T Q I F T T Y S D N Q P G V L I

1321 CAGGTGTACGAGGGCGAGAGGGCCATGACGCGGGACAACAACCTGCTGGGGCGCTTCGAG

441 Q V Y E G E R A M T R D N N L L G R F E

1381 TTGAGCGGCATCCCGCCGGCCCCGCGGGGGGTGCCCCAGATCGAGGTGACCTTCGACATC

461 L S G I P P A P R G V P Q I E V T F D I

1441 GACGCCAATGGCATCCTGAACGTCACGGCCACGGACAAGAGCACGGGCAAGGCCAACAAG

481 D A N G I L N V T A T D K S T G K A N K

1501 ATCACCATCACCAACGACAAGGGCCGGCTGAGCAAGGAGGAGATCGAGCGCATGGTGCAG

501 I T I T N D K G R L S K E E I E R M V Q

1561 GAGGCGGAAAAGTACAAGGCGGAGGACGAGGTCCAGCGCGAGAGGGTGTCTGCCAAGAAC

521 E A E K Y K A E D E V Q R E R V S A K N

1621 GCGCTGGAGTCATACGCCTTCAACATGAAGAGCGCCGTGGAGGATGAGGGGCTGAAGGGC

```
541        A  L  E  S  Y  A  F  N  M  K  S  A  V  E  D  E  G  L  K  G

1681       AAGATCAGCGAGGCGGACAAGAAGAAGGTGCTGGACAAGTGCCAGGAGGTGATTTCCTGG

561        K  I  S  E  A  D  K  K  K  V  L  D  K  C  Q  E  V  I  S  W

1741       CTGGACGCCAACACCTTGGCGGAGAAGGACGAGTTTGAGCACAAGAGGAAGGAGCTGGAG

581        L  D  A  N  T  L  A  E  K  D  E  F  E  H  K  R  K  E  L  E

1801       CAGGTGTGTAACCCCATCATCAGCAGACTGTACCAGGGGGCGGGCGGCCCCGGGGCTGGC

601        Q  V  C  N  P  I  I  S  R  L  Y  Q  G  A  G  G  P  G  A  G

1861       GGCTTTGGGGCTCAGGGCCCTAAAGGGGGCTCTGGGTCTGGCCCCACCATTGAGGAGGTG

621        G  F  G  A  Q  G  P  K  G  G  S  G  S  G  P  T  I  E  E  V

1921       GATTAG

641        D  *
```

Figure 12.
Nucleotide and deduced amino acid sequence of bHSP70. The predicted initiation Met start and stop codon (TAG) is in red. Highlighted in yellow are the forward and reverse primer sequences. Highlighted in red are the predicted N-glycosylation sites. Underlined are the three heat shock hsp70 proteins family signatures at amino acid positions 9 – 16, 197 – 210, and 334 – 348.

>NM_213973 (90-kDa heat shock protein (HSP90)

```
1       ATGCCCGAGGAAACCCAGACCCAAGACCAGCCGATGGAGGAGGAGGAGGTGGAGACGTTC
1        M  P  E  E  T  Q  T  Q  D  Q  P  M  E  E  E  E  V  E  T  F
61      GCCTTCCAGGCGGAAATCGCCCAGTTGATGTCGTTGATCATCAACACTTTCTACTCGAAC
21       A  F  Q  A  E  I  A  Q  L  M  S  L  I  I  N  T  F  Y  S  N
121     AAGGAGATCTTTCTGAGGGAGCTCATTTCCAACTCGTCCGATGCTTTGGACAAGATCAGA
41       K  E  I  F  L  R  E  L  I  S  N  S  S  D  A  L  D  K  I  R
181     TACGAAAGCCTGACGGATCCCAGTAAACTAGACTCCGGGAAAGAGCTGCACATTAATCTC
61       Y  E  S  L  T  D  P  S  K  L  D  S  G  K  E  L  H  I  N  L
241     ATTCCGAACAAGCAAGACCGGACCCTCACGATAGTGGACACCGGCATCGGCATGACCAAG
81       I  P  N  K  Q  D  R  T  L  T  I  V  D  T  G  I  G  M  T  K
301     GCCGACTTGATCAATAACCTTGGTACGATCGCCAAGTCTGGGACCAAGGCGTTCATGGAG
101      A  D  L  I  N  N  L  G  T  I  A  K  S  G  T  K  A  F  M  E
361     GCTTTGCAGGCCGGTGCCGATATCTCGATGATTGGCCAGTTCGGTGTCGGCTTCTACTCT
121      A  L  Q  A  G  A  D  I  S  M  I  G  Q  F  G  V  G  F  Y  S
421     GCGTACCTGGTCGCTGAGAAAGTGACCGTTATCACCAAACACAACGATGACGAGCAGTAT
141      A  Y  L  V  A  E  K  V  T  V  I  T  K  H  N  D  D  E  Q  Y
481     GCCTGGGAGTCTTCTGCAGGAGGATCTTTCACCGTTAGGACAGACACAGGAGAACCTATG
161      A  W  E  S  S  A  G  G  S  F  T  V  R  T  D  T  G  E  P  M
541     GGTCGTGGAACAAAGGTTATTCTACATCTGAAAGAAGACCAAACTGAGTACTTGGAAGAA
181      G  R  G  T  K  V  I  L  H  L  K  E  D  Q  T  E  Y  L  E  E
601     AGGAGAATAAAGGAGATTGTGAAGAAACACTCTCAGTTTATTGGCTACCCCATTACTCTC
201      R  R  I  K  E  I  V  K  K  H  S  Q  F  I  G  Y  P  I  T  L
661     TTCGTGGAGAAGGAACGTGATAAAGAAGTCAGTGACGACGAGGCGGAAGAAAAGGAAGAC
221      F  V  E  K  E  R  D  K  E  V  S  D  D  E  A  E  E  K  E  D
721     AAAGAGGAAGAAAAGGAGAAAGAAGAGAAGGAATCTGAGGATAAACCGGAGATAGAAGAT
241      K  E  E  E  K  E  K  E  E  K  E  S  E  D  K  P  E  I  E  D
```

```
781       GTTGGTTCTGATGAAGAGGAAGAAGAAAAGAAGGATGGTGACAAGAAGAAGAAGAAGAAG
261        V  G  S  D  E  E  E  E  E  K  K  D  G  D  K  K  K  K  K  K
841       ATCAAGGAGAAGTATATTGATCAAGAGGAACTCAACAAGACAAAGCCTATCTGGACCAGA
281        I  K  E  K  Y  I  D  Q  E  E  L  N  K  T  K  P  I  W  T  R
901       AACCCCGATGACATCACTAATGAAGAGTACGGGGAGTTCTATAAGAGCTTGACCAATGAC
301        N  P  D  D  I  T  N  E  E  Y  G  E  F  Y  K  S  L  T  N  D
961       TGGGAGGATCACTTGGCTGTGAAGCACTTTTCAGTCGAAGGGCAGTTGGAGTTCAGAGCC
321        W  E  D  H  L  A  V  K  H  F  S  V  E  G  Q  L  E  F  R  A
1021      CTTCTTTTCGTCCCAAGACGCGCTCCTTTCGACTTATTTGAAAACAGAAAGAAGAAGAAC
341        L  L  F  V  P  R  R  A  P  F  D  L  F  E  N  R  K  K  K  N
1081      AACATCAAGCTGTATGTTCGCAGAGTGTTCATCATGGACAACTGCGAGGAGCTCATCCCT
361        N  I  K  L  Y  V  R  R  V  F  I  M  D  N  C  E  E  L  I  P
1141      GAGTATCTGAATTTCATTAGAGGCGTGGTGGACTCTGAGGATCTTCCTCTGAACATTTCT
381        E  Y  L  N  F  I  R  G  V  V  D  S  E  D  L  P  L  N  I  S
1201      CGTGAGATGTTGCAACAAAGCAAAATTTTGAAAGTCATCAGGAAGAATCTGGTCAAGAAA
401        R  E  M  L  Q  Q  S  K  I  L  K  V  I  R  K  N  L  V  K  K
1261      TGCTTGGAGCTCTTTACTGAATTGGCCGAAGATAAAGAGAATTACAAGAAGTTCTATGAG
421        C  L  E  L  F  T  E  L  A  E  D  K  E  N  Y  K  K  F  Y  E
1321      CAGTTCTCTAAAAATATTAAGCTTGGAATACATGAAGATTCTCAAAATCGGAAGAAGCTT
441        Q  F  S  K  N  I  K  L  G  I  H  E  D  S  Q  N  R  K  K  L
1381      TCCGAGCTGTTGAGGTACTACACTTCTGCTTCTGGCGACGAGATGGTTTCCCTCAAGGAC
461        S  E  L  L  R  Y  Y  T  S  A  S  G  D  E  M  V  S  L  K  D
1441      TATTGCACCAGAATGAAGGAGAACCAGAAACACATCTATTACATCACAGGTGAGACCAAG
481        Y  C  T  R  M  K  E  N  Q  K  H  I  Y  Y  I  T  G  E  T  K
1501      GACCAGGTGGCCAACTCGGCCTTCGTGGAACGTCTGCGGAAGCACGGCCTGGAGGTGATC
501        D  Q  V  A  N  S  A  F  V  E  R  L  R  K  H  G  L  E  V  I
1561      TACATGATCGAGCCCATCGATGAGTACTGTGTGCAGCAGCTGAAGGAGTTTGAGGGGAAG
521        Y  M  I  E  P  I  D  E  Y  C  V  Q  Q  L  K  E  F  E  G  K
1621      ACCTTAGTGTCAGTCACCAAAGAGGGCCTGGAGCTCCCGGAAGATGAAGAGGAGAAGAAG
```

```
541        T  L  V  S  V  T  K  E  G  L  E  L  P  E  D  E  E  E  K  K
1681       AAACAGGAGGAGAAGAAGACAAAGTTTGAAAACCTCTGCAAGATCATGAAGGACATCTTG
561        K  Q  E  E  K  K  T  K  F  E  N  L  C  K  I  M  K  D  I  L
1741       GAGAAGAAAGTCGAAAAGGTGGTTGTGTCGAACCGGTTGGTGACCTCACCGTGCTGCATT
581        E  K  K  V  E  K  V  V  V  S  N  R  L  V  T  S  P  C  C  I
1801       GTCACAAGCACATACGGCTGGACAGCAAACATGGAGAGAATCATGAAGGCTCAAGCCCTG
601        V  T  S  T  Y  G  W  T  A  N  M  E  R  I  M  K  A  Q  A  L
1861       AGAGACAACTCGACAATGGGTTACATGGCAGCAAAGAAGCACCTGGAGATAAATCCAGAC
621        R  D  N  S  T  M  G  Y  M  A  A  K  K  H  L  E  I  N  P  D
1921       CATTCCATCATCGAGACCCTGAGGCAAAAGGCAGAGGCGGACAAGAACGACAAGTCCGTG
641        H  S  I  I  E  T  L  R  Q  K  A  E  A  D  K  N  D  K  S  V
1981       AAGGATCTGGTCATCCTGCTGTACGAAACCGCTCTGCTGTCTTCTGGCTTCAGCCTGGAA
661        K  D  L  V  I  L  L  Y  E  T  A  L  L  S  S  G  F  S  L  E
2041       GATCCCCAGACGCACGCCAACAGGATCTACAGGATGATCAAACTTGGTCTTGGTATTGAT
681        D  P  Q  T  H  A  N  R  I  Y  R  M  I  K  L  G  L  G  I  D
2101       GAGGACGACCCCACCGCCGACGACAGCAGCGCTGCTGTGACGGAGGAGATGCCACCCCTG
701        E  D  D  P  T  A  D  D  S  S  A  A  V  T  E  E  M  P  P  L
2161       GAAGGGGACGACGACACGTCCCGCATGGAGGAAGTCGATTAG
721        E  G  D  D  D  T  S  R  M  E  E  V  D  *
```

Figure 13.
Nucleotide and deduced amino acid sequence of HSP90. The predicted initiation Met start and stop codon (TAG) is in red. Highlighted in yellow are the forward and reverse primer sequences. Highlighted in red are the predicted N-glycosylation sites. Underlined is the heat shock hsp90 proteins family signature at amino acid positions 38 – 47.

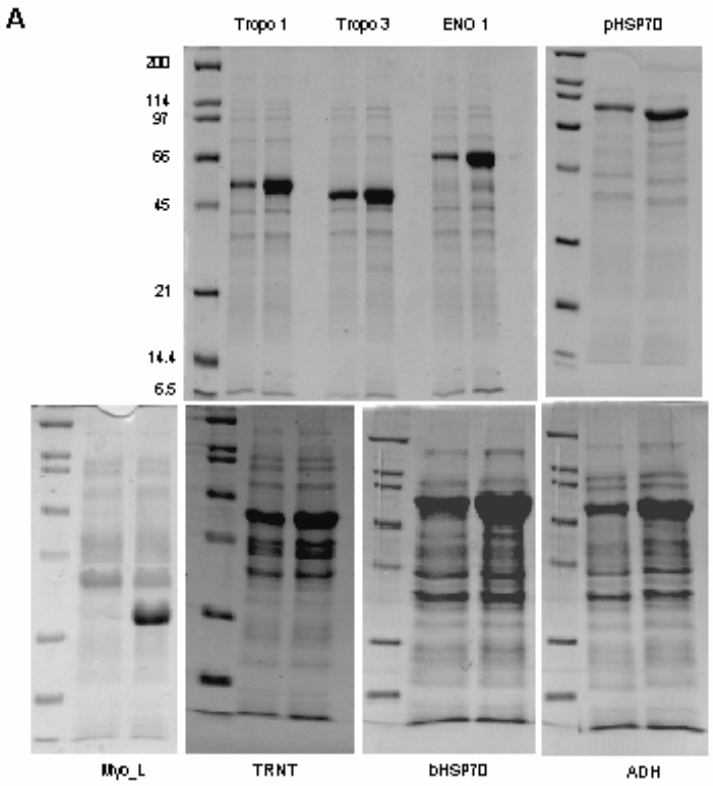
5.3.2 Expression and purification of recombinant allergens

The different clones encoding for the potential allergens were sub-cloned into the expression plasmid pET 32a and the proteins were expressed as His-tagged proteins. The pET System was chosen as it is a powerful system developed for cloning and expression of recombinant proteins in *E. coli*. The yield of the desired protein can comprise more than 50% of the total cell protein after a few hours of induction.

SDS-PAGE analysis showed the successful expression and purification for all of the recombinant proteins except HSP90 (Figure 14). Despite optimization by changing induction temperature and concentration of IPTG, HSP90 was still unable to be expressed. Thus, HSP90 was excluded in subsequent experiments. The molecular weights of the expressed recombinant fusion proteins were estimated by adding the molecular weight of each protein of interest with the molecular weight of fusion protein (17.3 kD). The molecular weight of fusion protein was obtained by estimated the molecular weight of deduce amino acids of fusion protein coding sequences in vector DNA. The molecular weight of the allergen of interest and its predicted mass after fusion with the fusion protein are shown in Table 6.

Table 6.
Estimated molecular weight of the expressed allergen with the fusion protein.

Putative allergen	Estimated MW cloned (kD)	Predicted MW together with fusion protein (kD)
Tropo 1	32.7	50
Tropo 3	29	46.3
Myo_L	16.7	34
TRNT	29.8	47.1
pHSP70	70	87.3
ENO 1	47.2	64.5
ADH	54.7	72
bHSP70	70.1	87.4



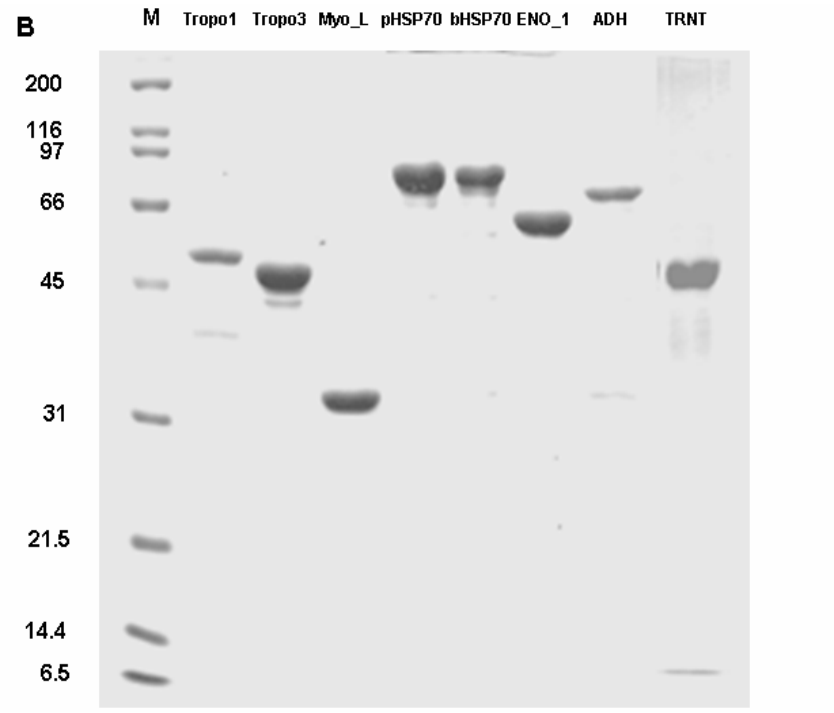


Figure 14.
SDS-PAGE analyses of expression (A) and purification (B) of the recombinant allergens.

5.3.3 Recombinant proteins immunoarray

Sera from 70 atopic patients that positively react to meat crude extracts in the previous allergen immunoarray analysis and 10 negative sera were tested in the recombinant proteins immunoarray to determine IgE reactivity to putative meat-based recombinant proteins. Various positive controls such as IgE standards and negative controls such as PBS buffer, elution buffer and *E. coli* proteins were added in the panel of screens. An example of the immunodot blot and the plate format are shown in Figure 15.

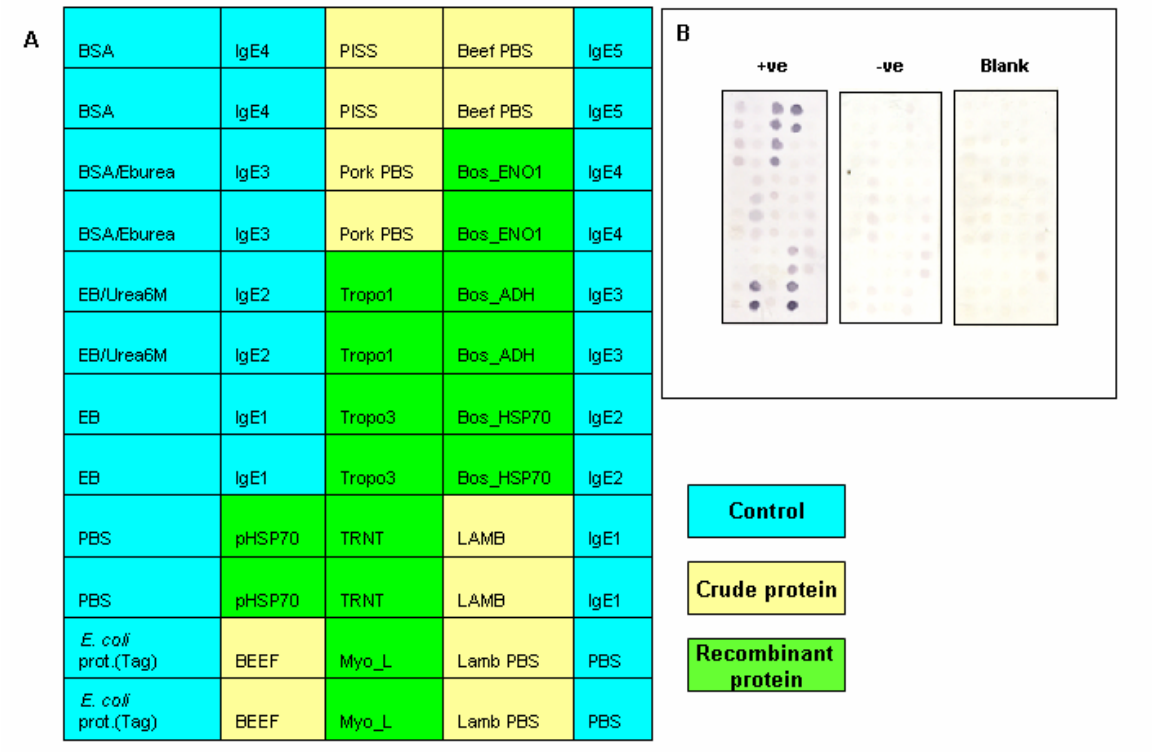


Figure 15.
Recombinant proteins immunoarray plate format and immunodot blots. (A) The plate format of the recombinant immunoarray where highlighted in blue, yellow and green are the controls, crude proteins and recombinant proteins respectively. (B) An example of the immunodot blots labeled +ve sera, -ve sera and blank (secondary antibody only). Dots in purple indicate positive IgE-binding.

5.3.3.1 Prevalence of IgE-binding of crude and recombinant proteins

Figure 16 shows the IgE-binding capacity (%) to both crude and recombinant proteins. As expected, close to 90% of the pre-screened sera reacted to any one type of meat (beef, lamb or pork) and these results strongly ascertain the reproducibility of the allergen immunoarray results. Majority of the meat reactive sera showed IgE-binding to BSA (49.3%), which is a major allergen in beef allergy (Fuentes *et al.*, 2004). However, the IgE-binding frequency was relatively low for the recombinant proteins ranging from

8.75% to 18.75%. This could be due to improper folding of the recombinant protein and also the lack of post-translation protein modification in an *E. coli* (prokaryotic) system. The lack of post-translation protein modification as glycosylation, phosphorylation or even the proper formation of disulphide bonds may lead to mis-folded protein and aggravate the formation of proteins in inclusion bodies which render them insoluble. Additionally, IgE-binding to certain proteins, especially glycoproteins, may occur at or be influenced by the glycosyl side-chains, which are known as cross-reactive carbohydrate determinants (CCD) (Mari *et al*., 1999; Mari, 2002). It is also well-known that allergens which are glycoproteins in particular can increase the antibodies titer (Thomas and Smith, 2002). Hence, this could be the reason why there was a marked reduction in IgE-binding for the recombinant proteins produced in *E. coli*.

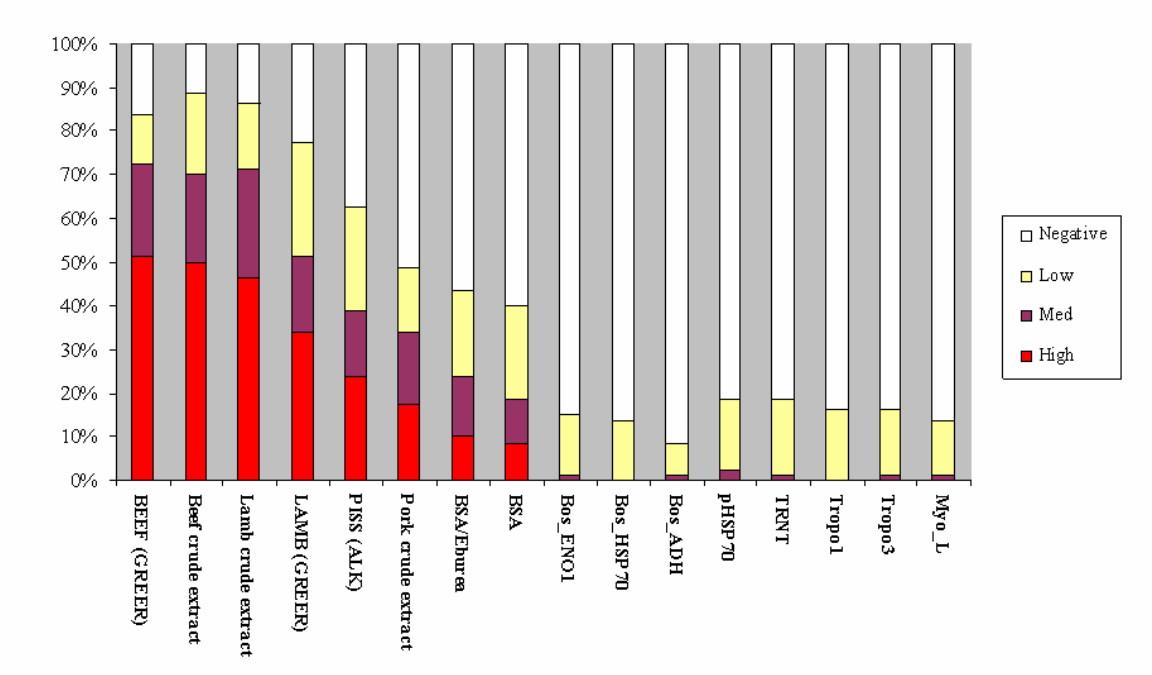


Figure 16.
Percentage of IgE-binding frequency of crude and recombinant proteins

On the other hand, CCDs are also known to reduce the specificity and diagnostic value of *in vivo* assays, and may be counter-productive in immunotherapy (Thomas and Smith, 2002; Vieths *et al.*, 2001). Therefore, we need to better understand the significances of CCDs in order to increase the sensitivity and specificity of allergy diagnosis especially *in vitro* diagnosis techniques. The effect of CCDs will be discussed farther in Chapter 6.

5.4 Conclusion

A total of nine allergens from *S. scrofa* (Tropomyosin 1 alpha chain, Tropomyosin 3, Myosin light chain, Troponin T fast skeletal muscle type, and Heat shock protein 70 kDa) and *B. taurus* (Enolase 1, Aldehyde dehydrogenase, and Heat shock protein 70 kDa) were successfully cloned, produced and purified as recombinant proteins. Their identities were confirmed by molecular weight and by DNA sequencing. Results from the recombinant proteins immunoarray showed relatively low IgE-binding in the 80 patients' sera tested. The low IgE-binding frequency could be due to the lack of post translational protein modification which resulted in the lost of CCDs and/or the improper folding of the recombinant protein.

CHAPTER 6: INVESTIGATIONS ON THE CROSS-REACTIVE CARBOHYDRATE DETERMINANTS (CCD) OF MEAT-BASED ALLERGENS

6.1 INTRODUCTION

The role of carbohydrates in allergy has been much discussed in recent years and is still controversial. Glycans often play important physiological roles e.g. in resistance to proteolytic degradation, transduction of information between cells and intercellular adhesion through ligand–receptor interactions (Gesundheit *et al*., 1987; Ashwell and Harford, 1982; Potskaldy *et al*., 1986). Indications for glycosylation can be deduced from the protein sequence. The sequence motif Asn-X (except of Pro)-Ser /Thr indicates probable *N*- glycosylation sites. (The prediction of *O*-linked glycans is much more difficult and less reliable. They are built up of only a few monosaccharide units and seem to be less important for the elucidation of allergy) (Petersen and Mundt, 2001).

In clinical relevance, due to the widespread occurrence of structurally similar CCDs in plants and invertebrates, IgE to carbohydrates often leads to false-positive diagnostic results when only the IgE binding to extracts or allergens (CAP, RAST, FAST, and immunoblot) but do not show clinical symptomology (Mari et al., 1999; van der Veen et al., 1997; van Ree, 1999). Many glycoproteins, such as bromelain, are monovalent glycoconjugates thus unable to cross-link IgE antibodies bound to the receptors of mast

cells and basophils to elicit clinical symptoms (Paschinger *et al*., 2005). However, there are also many glycoproteins with more than one N-linked glycan [eg, horseradish peroxidase (HRP)] or 3 potential glycoprotein allergens (eg, β-fructofuranosidase, polygalacturonase 2A and pectinesterase) that we have recently identified in tomato fruit extract (Westphal *et al*., 2003; Foetisch *et al*., 2001). Until now, there are no conclusive evidences to indicate that whether or not natural multivalent glycoproteins are able to induce histamine release and/or contribute to clinical symptoms.

The IgE-binding carbohydrates are often N-glycans carrying α(1,3)-fucose (linked to the proximal N-acetyl glucosamine) and/or β(1,2)-xylose (linked to the core mannose) commonly found in pollen and plant food glycoproteins (Ebo *et al*., 2004; Faye and Chrispeels, 1988). These are features absent in mammals and are, therefore, immunogenic (Wilson, 2002). The prevalence of anti-CCD IgE has been estimated to be 10–15% in patients with grass pollen allergy (van Ree and Aalberse, 1993) and increases up to 60%in patients with concomitant sensitization to pollen from trees, grasses and weeds (Mari *et al*., 1999). Invertebrate glycoproteins, present in hymenoptera venom, also bear IgE-binding α(1,3)-fucose-containing CCD (Tretter *et al*., 1993). About one in four honeybee–bumblebee venom and one in 10 yellow jacket venom allergic patients have been shown to be anti-CCD IgE positive (Kochuyt, 2000).

Besides CCDs, profilin is another protein that is known to elicit false-positive IgE-binding without clinical relevance. Profilins, a family of actin-binding proteins, are ubiquitous in eukaryotic cells and probably function as mediators of membrane–

cytoskeleton signaling (Ebo *et al*., 2004). Profilins are present in pollen (e.g. birch, mugwort, timothy grass), fruit (e.g. apple, cherry), vegetables (e.g. celery, peanut, carrot) and natural rubber latex (Vieths *et al*., 2002; Valenta *et al*., 1992; Vallier *et al*., 1995). The clinical role of profilins is still not clear. For instance, the birch profilin has been demonstrated to be a highly efficient sensitizer, but sensitization to birch profilin is rarely associated with symptoms (Pauli *et al*., 1996; Wensing *et al*., 2002)

6.2 MATERIALS AND METHODS

6.2.1 Patients and sera

A total of seven sera from the previously pre-screened population mentioned in section 2.2.6 were selected for investigation for presence of cross-reactive carbohydrate determinants (CCDs). The sera were selected on the basis that the patients have relatively high levels of specific IgE to meat antigens.

6.2.2 Protein extraction

The crude extracts of beef and pork were extracted as described in section 2.2.3. Total protein concentration was then determined using the BioRad protein assay kit (Bio-Rad, USA).

6.2.3 Enzymatic deglycosylation procedures

`Crude meat extracts from beef and pork were enzymatically deglycosylation with E-DEGLY kit (Sigma, USA) according to manufacturer's protocol with slight modifications. Briefly, 10 µl of 5X Reaction Buffer and 2.5 µl of Denaturation Solution were added to 100 µg of protein sample and incubated at 4 °C for 24 hrs. After which, 2.5 µl of TRITON X-100 solution was added along with 1 µl each of the PNGase F, O-Glycosidase, α-2(3,6,8,9) Neuraminidase, β(1-4)Galactosidase and β−N-Acetylglucosaminidase. The sample was incubated at 37 °C for 3 hrs and analyzed by SDS-PAGE for deglycosylation.

6.2.4 Immunoassays

6.2.4.1 Western blot analysis

The procedure for western blot was essentially the same as described in the previous section 4.2.5. Basically, the non-deglycosylated extracts and deglycosylated extracts were separated by 1D SDS-PAGE side by side for ease of profile comparison. The proteins were then transferred onto nitrocellulose membranes and assayed with patients' sera for IgE-binding.

6.2.4.2 Enzyme-linked immunosorbent assay (ELISA)

The procedure for enzyme-linked immunosorbent assay (ELISA) was similar as described in section 2.2.7.

6.3 Results and Discussion

6.3.1 Deglycosylation experiments

The deglycosylation procedure in crude extracts can only be interpreted, when monoclonal antibodies are available for identification of the glycoprotein and the mixture contains only a few glycoproteins (Petersen and Mundt, 2001). However, in this study, monoclonal antibodies are not available thus we have to assume that the changes in protein profile on the SDS-PAGE were due to the deglycosylation process. Figure 1 illustrates the deglycosylated SDS-PAGE gel for pork and beef with noticeable changes in protein profile after deglycosylation.

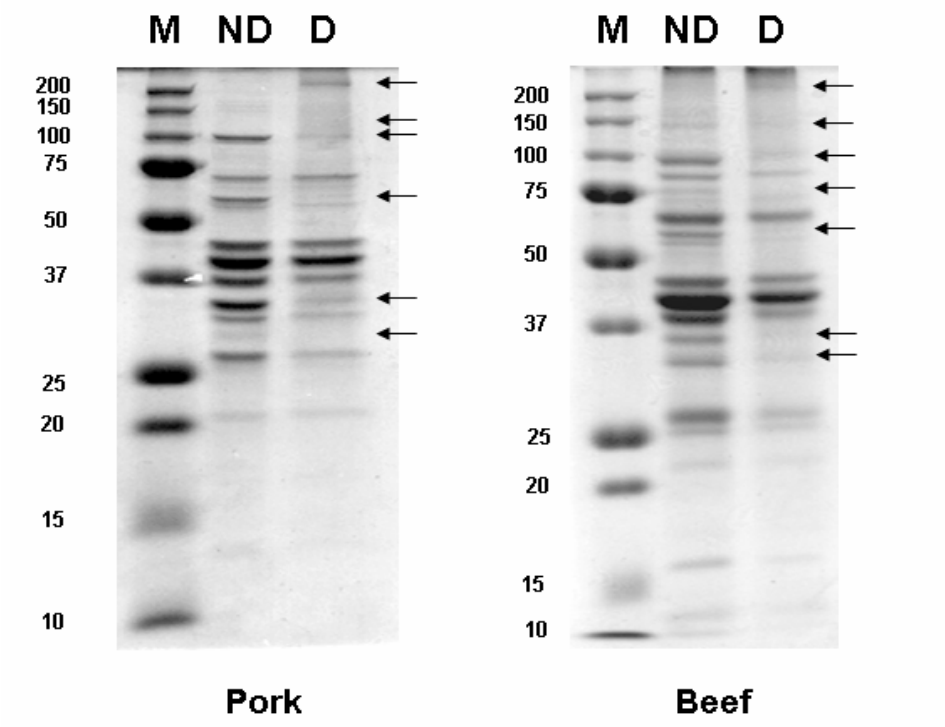


Figure 1.
Protein separation of non-deglycosylated and deglycosylated pork and beef crude extracts. M: marker; ND: non-deglycosylated; N: deglycosylated. Indicated in arrows are bands or regions with noticeable changes before and after deglycosylation.

6.3.2 Immunoassays

6.3.2.1 Western blot analysis

Results of immunoblotting after SDS-PAGE with patients' sera disclosed significant reduction in IgE-binding before and after deglycosylation. Reduction in IgE-binding is especially noticeable for high molecular weight reactive bands (<100 kDa) in both species' deglycosylated extracts with the exception in Patient 1 (Figure 2). One reactive band at approximately 52 kDa was retained in both non-deglycosylated and deglycosylated pork extract suggesting that the particular band has no carbohydrate moieties (Figure 2A). Based on previous experiment done in chapter four, the identity of this reactive band is IgG heavy chain. Computational N-glycosylation prediction using NetNGlyc 1.0 revealed that the potential of glycosylation for this protein is low with only one N-glycosylation site at amino acids 322 – 324. For patient 3 and 5, two additionally bands at 27 and 25 kDa were observed in the deglycosylated pork extracts (Figure 2A). These bands are likely reactive bands that were partially deglycosylated. In the case of beef western analysis, the patients showed relatively similar pattern in IgE-binding and no particular reactive band was retained in both non-deglycosylated and deglycosylated extracts (Figure 2B). Patient 6 and 7 even have total abolishment of IgE-binding in the beef deglycosylated extract and this clearly indicates that majority of the IgEs were binding to the carbohydrate determinants.

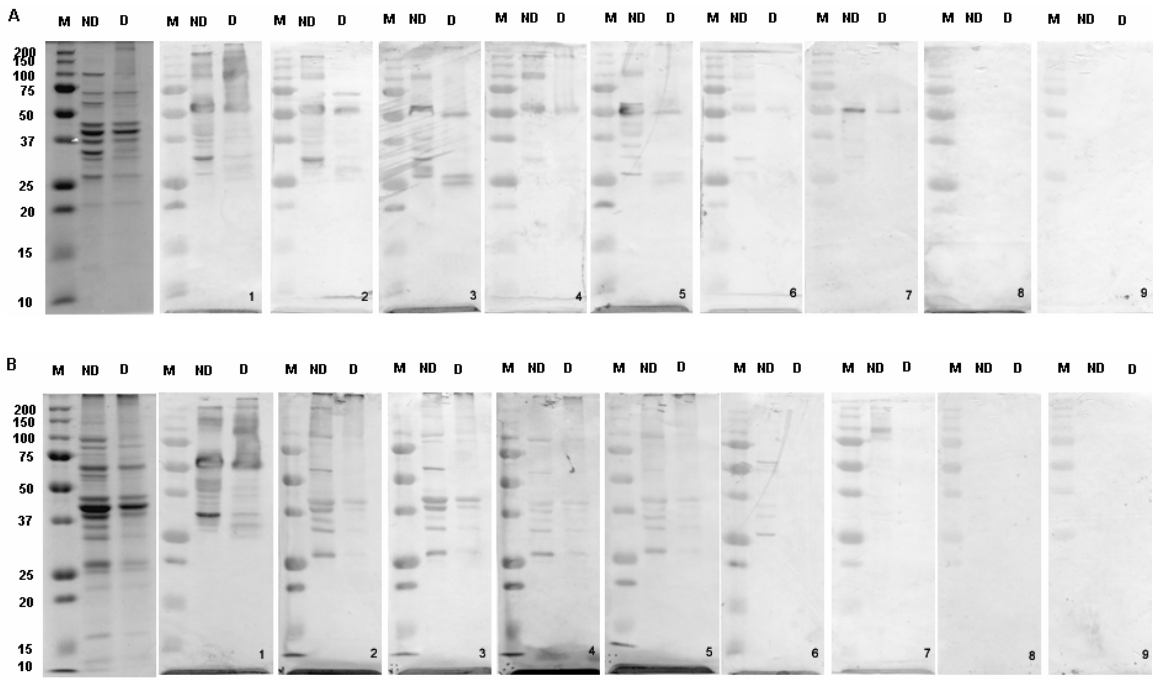


Figure 2.
Immunodetection after SDS-PAGE of pork (A) and beef (B) extracts (non-deglycosylated and deglycosylated) with patients' sera. M: marker; ND: non-deglycosylated; N: deglycosylated. In both blots, membranes 1 – 7 were IgE immunoblots using 7 patients' sera. Membranes 8 were negative serum and membranes 9 were blanks with secondary antibody only.

6.3.2.2 ELISA

ELISA was done to validate the results of the immunoblots. Figure 3 shows significant reduction in IgE-binding in deglycosylated extracts for individual patient. Sera from Patient 5 – 7 have reduced IgE-binding to basal level similar to that of the negative serum and blank. This result again confirm that majority of the IgEs were binding only to the carbohydrate determinants and not to the protein itself.

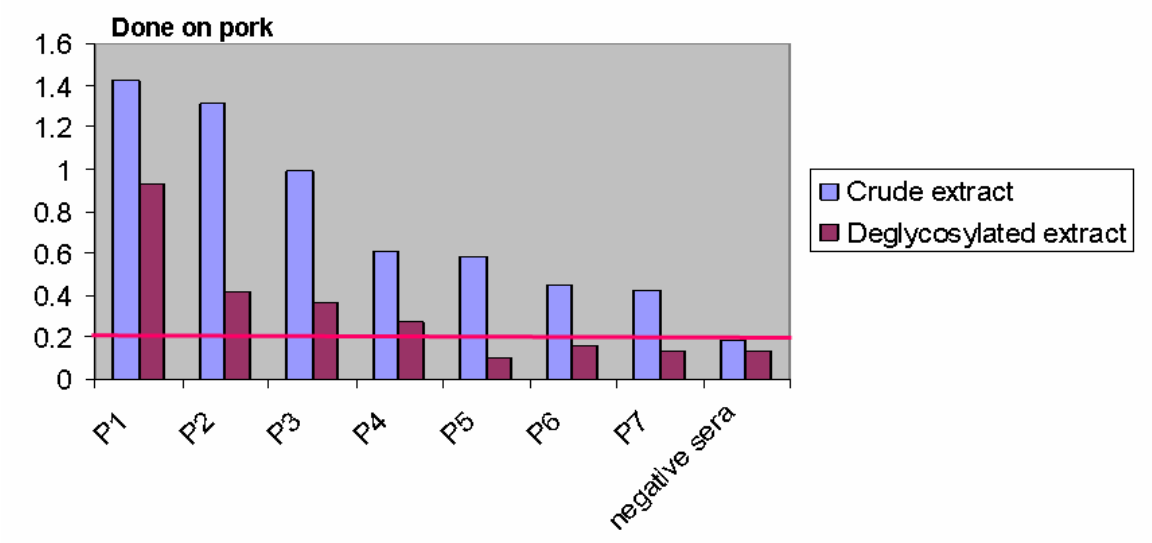


Figure 3.
ELISA validation of IgE-binding for non-deglycosylated and deglycosylated pork extract with patients' sera.

Physicians often rely on *in vitro* quantification of specific IgE to confirm the diagnosis of IgE-mediated allergic diseases. The patients are usually tested with allergenic extracts or molecules and those that show high IgE reactivity are assumed as the sensitizers. However, a positive specific IgE reaction may not always result in clinical presentation of allergic symptoms. The immunochemical cross-reactivity among different allergens sharing homologous epitopes has been identified as one of the main reasons for clinical irrelevant IgE and the most likely explanation for the lack of clinical presentation by some of these cross-reactive IgE antibodies may be related to epitope valency and low affinity (Ebo *et al*., 2004). Ubiquitous structures such as CCDs present on glycoproteins of mainly plants and hymenoptera venom have been known to induce the syntheses of IgE antibodies that lack biological activity (Ebo *et al*., 2004). The most relevant carbohydrate determinants for IgE reactivity is the α(1,3) fucosylation to the innermost N-acetylglucosamine and probably also the β(1,2) xylosylation to the trimannosyl core

residue (Faye and Chrispeels, 1988). These carbohydrate determinants are highly immunogenic in man (Ebo *et al*., 2004). The primary aim of this study is to investigate whether the high prevalence of IgE-binding to meat antigens is due to CCDs present on the antigens. Since this is the first study on mammalian-based CCDs and no unique marker for mammalian CCD reactivity is available, we have to rely on deglycosylation and reduction in IgE-binding to identified CCD sensitization in patients' sera with high IgE-binding to meat antigens.

Our data clearly show that deglycosylation reduces IgE-bindng in all seven patients tested. However, we have yet to identify the carbohydrate moieties where the anti-carbohydrate antibodies bind. Nevertheless, from the immunoarray screen results, we were able to deduce that the carbohydrate moieties were different from that of bromelain and horseradish peroxidase (standard control markers for CCDs). This was because the frequency of IgE binding for the meat namely pork and beef showed weak correlation with the two markers of CCDs (Figure 4). This indicates that the CCDs in mammalian meat or even vertebrates are unique and distinct from those found in plants. Further inhibition studies could be done by using neoglycoconjugates: bovine serum albumin carrying various forms of glycopeptides, N-glycans, or O-glycans to identify the actual mammalian carbohydrate moieties (Paschinger *et al*., 2005)

.

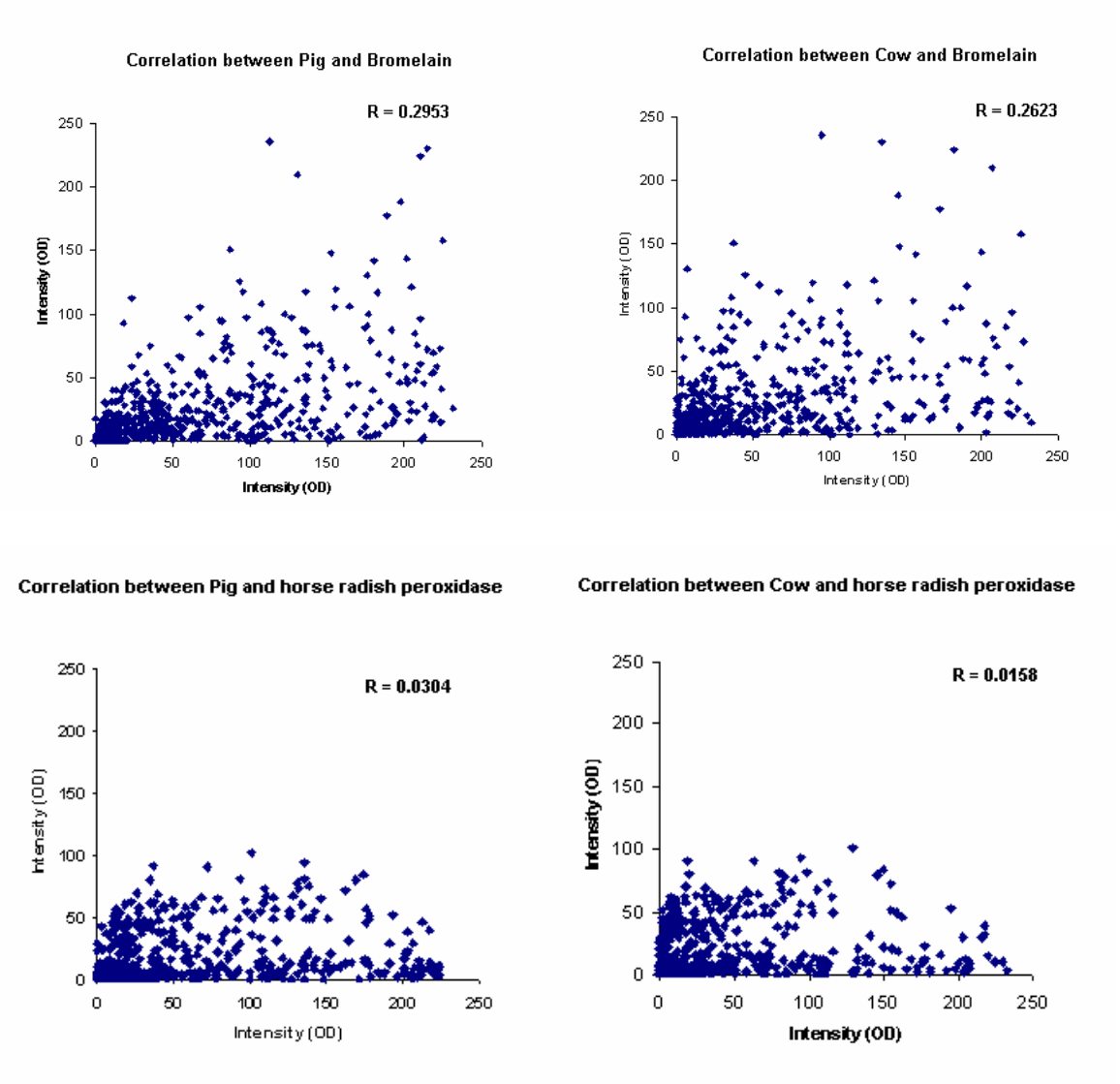


Figure 4
Correlation bi-plots between meat antigens (pig and cow) and CCDs markers (bromelain and horseradish peroxidase). Spearman's Correlation Test showed no correlation.

Besides inhibition studies, the detection and characterization of carbohydrate structures of glycoproteins can be achieved using a technique called hydrazinolysis, which is a starting point for detailed analysis of the carbohydrate structure. Hydrazinolysis basically involves acetone treatment followed by chromatographic procedures to purify the carbohydrate chains. Mass spectrometry is used to evaluate the monosaccharide composition and the exact carbohydrate structure can then be identified by NMR analysis. Oxford Glycosystem can also be used to determine the carbohydrate moieties by glycan sequencing followed by comparison with a glycan database to identify the particular structure.

6.4 Conclusion

This study confirms the presence of CCDs in crude meat antigens. Deglycosylation assay showed mark reduction in IgE-binding in both western blots and ELISA. Weak IgE-binding correlation between meat antigens and CCDs markers (bromelain and horseradish peroxidase) suggests that the carbohydrate moieties present in meat antigens are distinct from those present in plant. Further studies are required to identify and characterize in order to classify and understand this new form of mammalian CCD, which has never been reported in any literature.

CHAPTER 7: BLOCKING IMMUNOGLOBULIN G (IGG) ANTIBODIES IN MEAT ALLERGY

7.1 INTRODUCTION

7.1.1 Specific immunotherapy (SIT)

Specific immunotherapy (SIT) is an established treatment for selected allergic diseases including pollen, bee venom and dust mite allergy (Hardy *et al*., 2004). While the clinical efficacy of SIT is well-documented, the molecular mechanisms require further clarification. Current evidence suggests that specific immunotherapy effect several aspects of the immune system, including modulation of allergen-specific B cell and T cell responses. Studies on the effect of immunotherapy have demonstrated reduced basophil reactivity to allergens (Kimura *et al*., 1985), reductions in mucosal recruitment of inflammatory cells (Wilson *et al*., 2001), deviation of Th2 cytokine responses to allergens in favour of Th1 responses (Secrist *et al*., 1993; Jutel *et al*., 1995; McHugh *et al*., 1995; Akoum *et al*., 1996) and the induction of IL-10 producing regulatory T cells (Akdis *et al*., 1998; Bellinghausen *et al*., 1998; Francis *et al*., 2003). In addition, changes in levels of serum antibody in response to immunotherapy have been described, mostly as increases in allergen-specific IgG antibodies, particularly the IgG_4 isotype (Muller *et al*., 1989; McHugh *et al*., 1990; Ewan *et al*., 1993; Michils *et al*., 1997; van Neerven *et al*., 1999; Wachholz *et al*., 2003)

7.1.2 Concept of blocking IgG antibodies

Classical experiments performed by Cooke *et al*. in 1935 to investigate the protective mechanisms underlying allergen-specific immunotherapy demonstrated that allergen-specific IgG antibodies, termed blocking antibodies, can antagonize the cascade of allergic inflammation resulting from allergen recognition by IgE antibodies. The importance of such antibodies is still being debated. In general, SIT increases the level of allergen-specific IgG_1 and IgG_4 and these antibody concentrations are related to improvements in clinical outcome (Djurup and Osterballe, 1984; Moss *et al*., 1987; Gehlhar *et al*., 1999; Jutel *et al*., 2003). Frequently, however, successful SIT is not associated with a decreased concentration of allergen-specific IgE antibodies (Moss *et al*., 1987; Jutel *et al*., 1995; Gehlhar *et al*., 1999; Jutel *et al*., 2003), although this is not always the case (Gleich *et al*., 1982; Benjaponpitak *et al*., 1999; Flicker and Valenta, 2003).

One way in which IgG may block IgE-mediated mechanisms is by inhibiting basophile histamine release by either direct competition with IgE for allergen binding or possibly by the binding of IgE-allergen-IgG complexs to the lower affinity IgG receptor (FcεRIIb) with resulting co-aggregation with the high affinity IgE receptor (FcεRI) and the consequent inhibition of IgE receptor triggering (Coggeshall, 1998; Ravetch and Bolland, 2001; Zhu *et al*., 2002). Another mechanism of blocking antibodies is via inhibition of IgE facilitated allergen presentation by B cells to allergen –specific T cells. This results

in reduced T cell proliferation and cytokine production by specific T cells (van Neeren *et al*., 1999; Wachholz *et al*., 2003; van Neeren *et al*., 2004)

7.2 MATERIALS AND METHODS

7.2.1 Patients and sera

A total of 25 sera from the previously pre-screened population mentioned in section 2.2.6 were selected for specific IgG screening. Positive and negative sera were selected based on positivity or negativity to meat crude proteins *via* allergen immunoarray screening performed previously. Eventually, plasma and peripheral blood mononuclear cells (PBMCs) from two patients with both high levels of meat specific IgE and IgG antibodies were collected.

7.2.2 Allergen immunoarray for the detection of specific IgG

The procedure for allergen immunoarray was essentially the same as described in the previous section 2.2.4 except the secondary antibody used was goat anti-human IgG -chain specific alkaline phosphatase conjugated antibody (Sigma, USA) diluted at 1:5000 (v/v) with PBS. Image analysis of the immunoblots was also performed as mentioned in section 2.2.5.

7.2.3 Plasma preparation

Whole blood samples were collected in EDTA vacutainer tubes (Becton-Dickinson, USA). Invert the tubes to mix the samples well with the anticoagulant. Centrifuge the samples at 2000 x g for 15 minutes at room temperature. After centrifugation, the plasma was removed and stored in aliquots of 500 µl at -20 °C until use. The remaining cells were used for peripheral blood mononuclear cells isolation.

7.2.4 Isolation of peripheral blood mononuclear cells (PBMCs)

Peripheral blood mononuclear cells were isolated by Ficoll-Hypaque centrifugation (Amersham Pharmacia, USA). Basically, the 3 ml of remaining cells from previous section was diluted with 3 ml of PBS. The 6 ml of diluted cells was subsequently overlaid with 3 ml of Ficoll- Hypaque solution and centrifuged for 15 min at 750 x g with no brakes. The upper layer was removed and the buffy coat containing the PBMCs was transferred into a fresh 15 ml tube (Figure 1). The cells were washed twice by adding 10 ml of RPMI 1640 medium (Sigma, USA) and centrifuged for 10 min at 750 x g at room temperature. The cells pellet were re-suspended in growth medium (RPMI 1640 medium containing 10% fetal calf serum (FCS) and 0.01% (w/v) each of penicillin and streptomycin) and checked for cell viability and number. Cell viability was visualized with Trypan blue dye (Sigma, USA) (1:2 dilution with PBS) and cell number was determined with a hemacytometer (). The isolated PBMCs were diluted with culture

medium to 1 x 10^6 cells/ml either for storage at -80 °C with 10% dimethyl sulphoxide (DMSO) (Sigma, USA) or culture in a humidified atmosphere at 37 °C with 5% of CO_2.

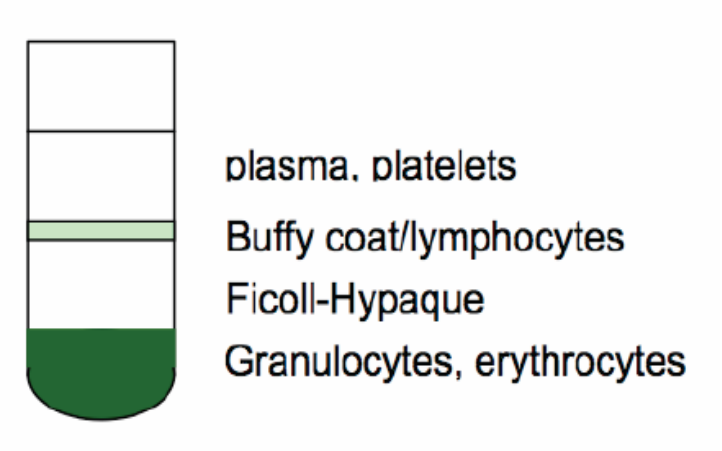


Figure 1.
Typical diagram showing Ficoll- Hypaque separation of whole blood sample.
Samples are separated based on density after centrifugation

7.2.5 Immunoaffinity depletion of IgG from plasma

Plasma instead of serum was used in this experiment because the process of blood coagulation triggers sticky platelets to release histamine (Masini *et al*., 1998). Thus, to prevent presence of endogenous histamine in serum, plasma samples were used. All plasma samples were centrifuged multiple times at 13,000 x g for 5 min to remove all clotted protein and lipid. IgG was depleted by loading 1 ml plasma aliquots onto 5 ml of Protein G PLUS-Agarose beads (Calbiochem, Germany) filled column (Bio-Rad, USA). The Protein G beads were used because it has high affinity for IgG antibodies and no affinity to IgE antibodies. The IgG depleted plasma was recovered from the flow-through and stored at -20 °C until use. Depletion of IgG was analyzed by SDS-PAGE. As a

control, PBS was loaded onto Protein G agarose column as above to exclude the contaminants from Protein G agarose.

7.2.6 Preparation of meat antigens

The crude extracts of pork, beef and chicken was prepared as described in section 2.2.3. To ensure the samples were sterile and free from contaminants, all extracts were syringe filtered at 0.22 (Milipore, USA).

7.2.7 Histamine-release assay

Crude extracts from pork, beef and chicken were diluted with PBS to concentrations 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1} or 1 mg/ml. Based on prior knowledge that the patients did not have IgE reactivity to chicken, thus, chicken was used as a negative non-allergen control. The PBMCs were first checked for cell viability with Trypan blue dye (Sigma) before aliquoting about 1 x 10^5 cell/tube in duplicates for histamine release assay. Tubes containing 1 x 10^5 cells each were centrifuged at 750 x g for 10 min to remove the culture medium. The each tube containing cells was then co-incubated with 100 µl of plasma (either with or without IgG) with different concentration of various crude extracts or PBS control. The samples were incubated at 37 °C for 1 hr to induce histamine release from basophilic granulocytes. The released histamine in the supernatant was subsequently determined using the histamine enzyme immunoassay kit (SPI-BIO, France) according to the manufacturer's instructions. The principle of the histamine enzyme immunoassay

(EIA) is based on the competition between unlabelled derivatized histamine and acetylcholinesterase (AChE) linked to histamine (tracer) for limited anti-histamine antibody sites (Figure 2). Other controls in the assay include the tests for endogenous histamine in the plasma and the extracts itself. Therefore, the final analyzed result would involve subtracting the reading from various controls namely the PBS extract control and the endogenous histamine controls as mentioned above.

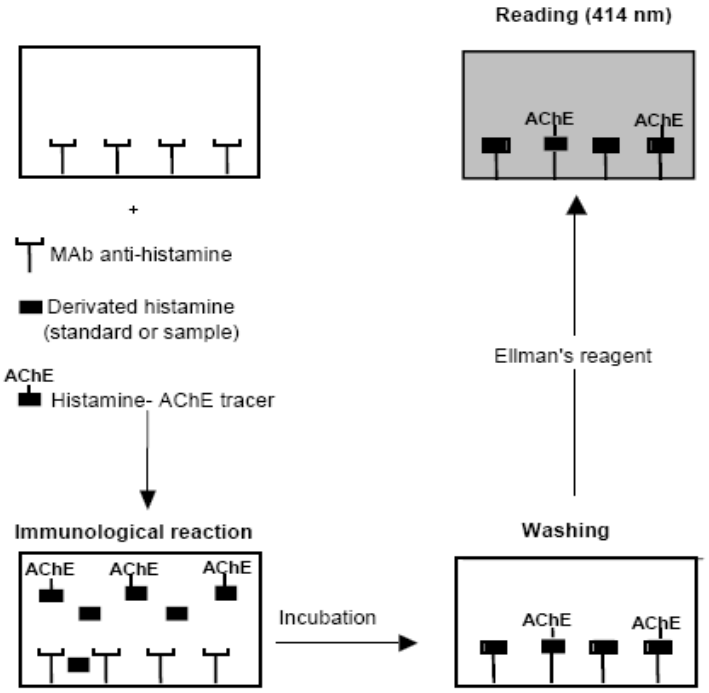
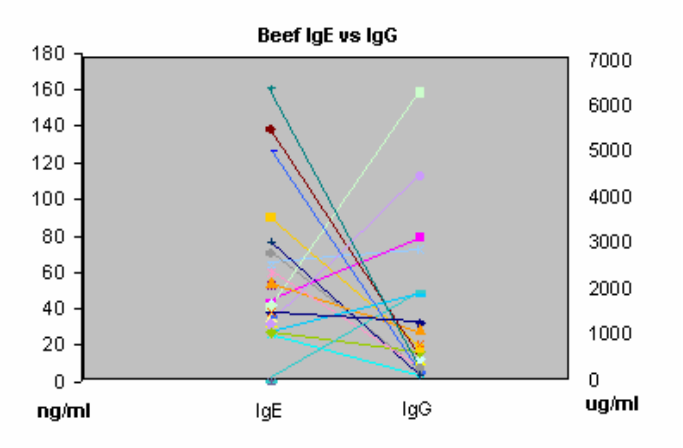


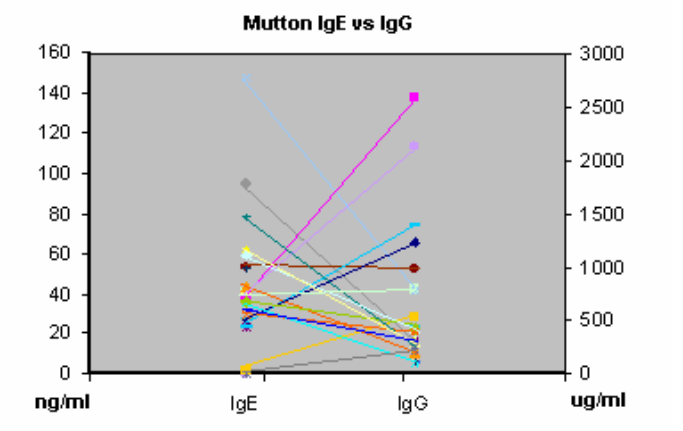
Figure 2.
Principle of the histamine enzyme immunoassay (EIA) kit.

7.3 Results and Discussion

7.3.1 Allergen immunoarray for the detection of specific IgG

A total of 25 sera with differential IgE-binding to pork, beef and lamb were analyzed for their level of specific IgG to the three meats. The reason for choosing sera with differential IgE-binding was to observe the distribution of specific meat IgG among patients with different atopy background and not only restricting to those with high level of specific IgE to meat. The levels of specific IgG for beef, pork and lamb were significantly higher ($p < 0.001$) than their respective specific IgE level (Figure 3). Interestingly, there seemed to be an inverse correlation for the levels of IgE and IgG antibodies. However, statistical analysis showed that the correlation is not significant ($p < 0.326$) thus we are unable to draw any conclusions. Nevertheless, the presence of specific meat IgG in such a high magnitude (>10000 fold) implies that the patients may have high abundance of blocking IgG antibodies which can antagonize the cascade of allergic inflammation resulting from allergen recognition by IgE antibodies.





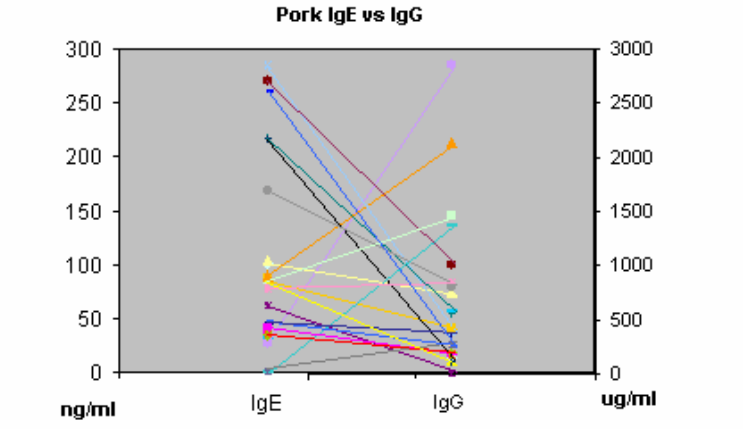


Figure 3. Comparison between levels of meat specific IgEs and IgGs. A total of 25 sera were screened for meat specific IgG antibodies. (Conversion for IgE is 1 IU = 2.4ng/protein; for IgG is 1 IU = 0.8147 mg). The difference in magnitude between IgG and IgE is more than 10000 times.

7.3.2 Immunoaffinity depletion of IgG from plasma

The IgG antibodies were removed from plasma using the Protein G PLUS-Agarose beads. Figure 4 shows the SDS-PAGE of the flow-through with IgG antibodies depleted from two patients' plasma. The molecular weight of IgGs (IgG_1, IgG_2, IgG_3 and IgG_4) ranges from 140 – 170 kDa (Gergely, 1967). The bands at molecular weights 140 – 170 kDa corresponding to molecular weights of IgG antibodies were removed in the IgG depleted plasma. Also, one additional band was noticed at approximately 30 kDa in the IgG depleted plasma. This band is probably due to contamination from the protein G agarose resin as its molecular weight corresponds with the molecular weight of protein G (Bjorck and Kronvall, 1984). Nevertheless, PBS was loaded onto the Protein G agarose column as control to exclude the contaminants from Protein G agarose in subsequent experiment. ELISA was also used to confirm the removal of IgG from the plasma samples (Figure 5).

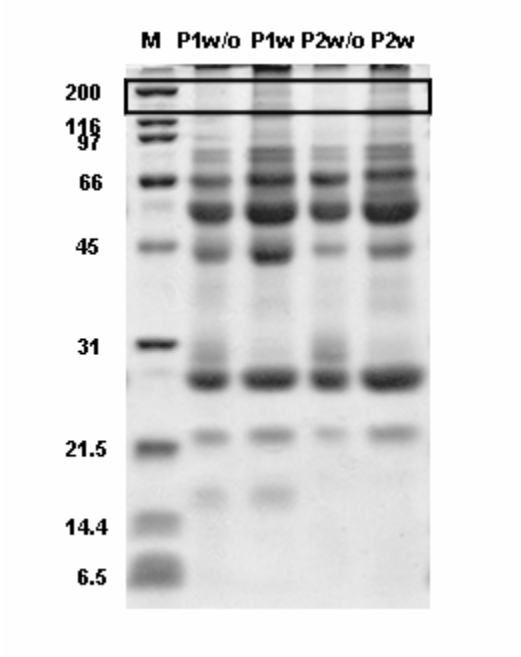


Figure 4.
SDS-PAGE of plasma and IgG depleted plasma from two patients. Lane 1: marker; Lane 2: Patient 1 IgG depleted plasma; Lane 3: Patient 1 plasma; Lane 4: Patient 2 IgG depleted plasma; Lane 5: Patient 2 plasma. Boxed region indicates the removal of IgG between 140 – 170 kDa when compared to non-IgG depleted plasma.

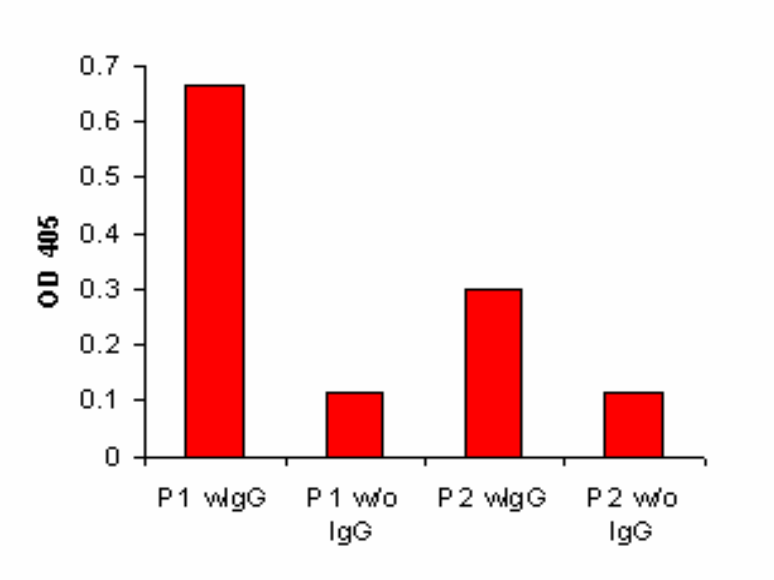


Figure 5.
ELISA confirmation of IgG removal from plasma samples

7.3.3 Histamine-release assay

When plasma (with or with IgG) co-incubated with allergen extracts were tested for the ability to trigger histamine release from human PBMCs, only samples with IgG depleted induced significant histamine release. This is especially prominent when IgG deplete samples were challenged with beef extract (Figure 6A and 6B)). A dose-response increase in histamine release was observed in both patients. On the other hand, in the case of plasma with IgG, only trace amount of histamine (< 2 ng/ml) was released when challenged with 1.0 mg/ml of beef extract. This strongly indicates the presence of IgG inhibits histamine release from human PBMCs. For pork, only trace amount of histamine (≤ 1 ng/ml) was released when challenged with 0.1 and 1.0 mg/ml of pork extract co-incubated with IgG depleted plasma from P1 (Figure 6C). No histamine was released from samples co-incubated with plasma containing IgG. The difference in levels of histamine release between pork and beef antigens may indicate that beef antigens are more potent allergens and the presence of IgG play a more significant role in inhibiting histamine release in beef allergy.

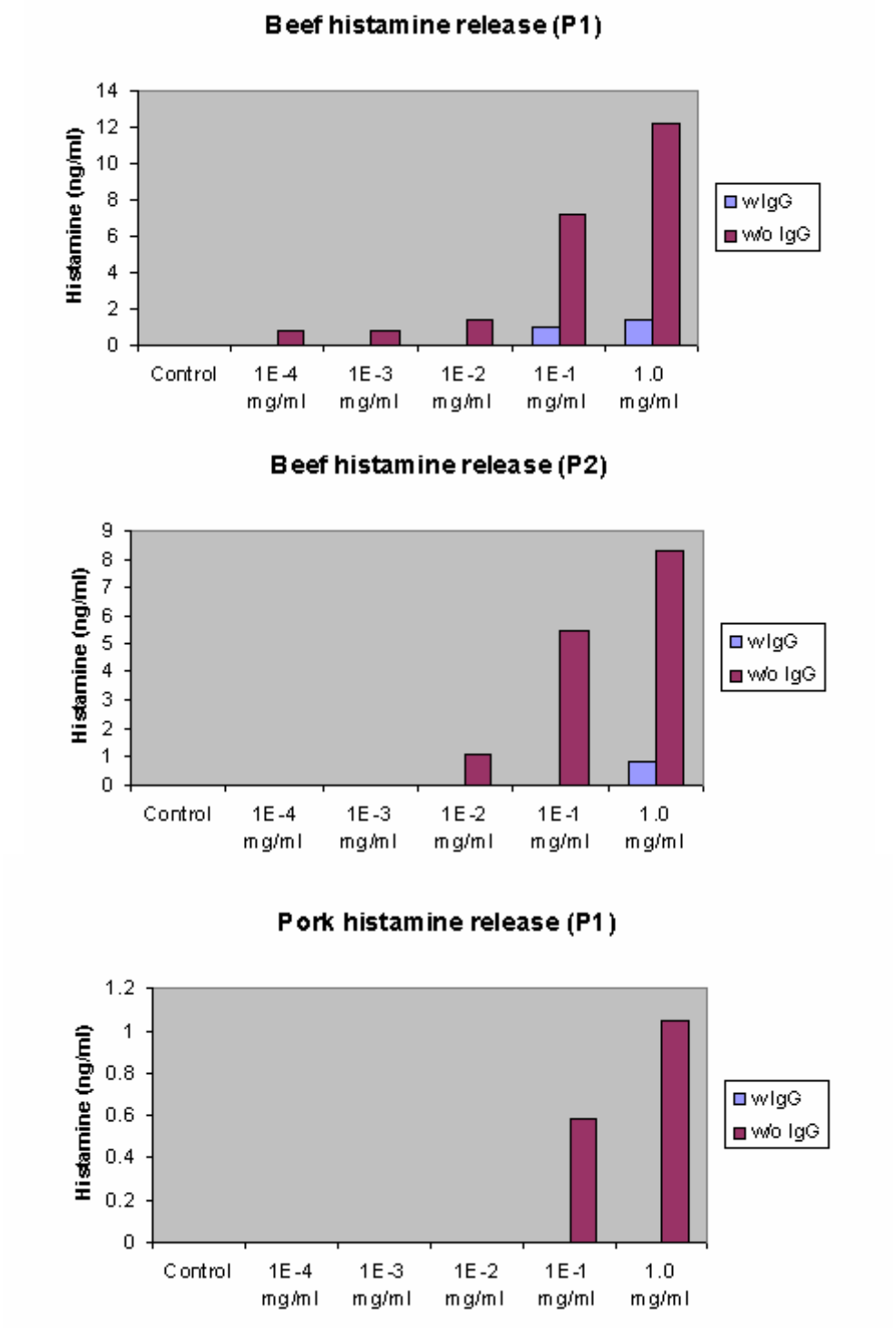


Figure 6.
Histamine release from PBMCs induced by either pork or beef extracts from two patients (P1 and P2). PBMCs were co-incubated with plasma (blue bars) or IgG depleted plasma (red bars). (A) Histamine release induced by beef extract with P1 plasma. (B) Histamine release induced by beef extract with P2 plasma. (C) Histamine release induced by pork extract with P1 plasma.

From this study, we can infer that the protective role of "blocking" IgG is the crucial factor which determines whether a patient develops allergic symptoms. Similarly in the case of SIT, increase in allergen-specific "blocking" IgG is one of the immunological mechanisms associated with successful SIT. The induction of IgG antibodies with blocking activity may have a protective role not only through the inhibition of allergen-induced, IgE-mediated release of inflammatory mediators from mast cells and basophils, but also through the inhibition of IgE-facilitated antigen presentation to T cells (Wachholz and Durham, 2004). Besides quantitative level of IgG, the qualitative changes such as affinity of IgG are also important in the allergen-specific IgG antibody response which retards allergic response (Svirshchevskaya *et al*., 2004). The production of IgG, particularly the IgG_4 subclass, is regulated by IL-10 which in turn reduce IL-4-induced IgE synthesis, while increasing allergen-specific γ4 transcription and IgG_4 production by B cells with an elevated IgG4:IgE ratio. Although SIT has no effect on antibody affinity of allergen-specific IgE, IgG_1 or IgG_4, study have shown that allergic patients with high-affinity IgG_1 and IgG_4 antibodies report less symptoms than patients with low affinity antibodies (Jakobsen *et al*., 2005). The mechanism of "blocking" IgG has yet to be fully resolved as studies are still ongoing to confirm whether the increase in IgG correlates well or not with reduce clinical symptoms. In the case whereby increases in IgG failed to predict the clinical response, it was hypothesized that the IgG may reflect high allergen exposure rather than play a causal role in successful immunotherapy (Muller *et al*., 1989; Ewan *et al*., 1993; van Ree *et al*., 1997)

Besides the blocking effect of IgG, the affinity and specificity of IgE also might play a role in histamine release. Recently, a strong correlation was found between the affinity of the IgE for its antigen and the sensitivity of the histamine release (Foetisch *et al*., 2003). Therefore the IgE antibody affinities to Der p 2 varied approximately 30-fold among 21 patients with mite allergy (Mita *et al*., 2000). Thus, in our case, the affinity of specific IgE for beef may be higher than that of pork resulting in a higher degree of histamine release since in PBMCs challenged with beef.

In the future, we hope to look at the effect on "blocking" IgG on T-cell proliferation and cytokine production. We would also like to explore the mechanism of "blocking" IgG antibodies by deciphering the methods of clearance and inhibition of the allergens. We hypothesize that the specific IgG antibodies are either involved in formation of IgG antibody-allergen complex for direct clearance from the immune system or the complex inhibits mast cell degranulation via inhibition of low affinity Fcε II receptor

7.4 Conclusion

In conclusion, we have shown that patients with high level of specific IgEs to meat antigen also have high if not higher level of meat specific IgGs indicating the presence of "blocking" IgG antibodies. IgG antibodies are successful depleted from plasma sample using protein G agarose resin. The PBMCs co-incubated with plasma sample with IgG depleted and the allergen has been shown to elicit histamine release suggesting that the presence of IgG antibodies inhibits histamine release. The *in vitro* inhibition of histamine

release by IgG antibodies also explains why there was a low prevalence of skin prick positive patients to meat antigens even though they themselves have high level of meat specific IgEs. In summary, the observed paradox of high abundance of IgE-binding to meats antigens but no clinical presentation to these antigens is due to the presence of unspecific meat CCDs and blocking IgG antibodies which suppress the allergic responses.

Bibliography and References

Aalberse RC, Akkerdaas J, van Ree R. (2001). Cross-reactivity of IgE antibodies to allergens. *Allergy* 56(6):478-90.

Adams MD, Soares MB, Kerlavage AR, Fields C, Venter JC (1993b). Rapid cDNA sequencing (expressed sequence tags) from a directionally cloned human infant brain cDNA library. *Nature Genetics* 4:373-380

Adams MD, Kelly JM, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, Merril CR, Wu A, Olde B, Moreno RF, Kerlavage AR, McCombie WR, Venter JC. (1991). Complementaey DNA sequencing: Expressed sequence tags and Human Genome Project. *Science* 252:1651-6.

Adams MD, Kerlavage AR, Fields C, Venter JC. (1993a). 3400 new expressed sequence tags identify diversity of transcripts in human brain. *Nature Genetics* 4:256-267.

Akdis CA, Blesken T, Akdis M, Wuthrich B, Blaser K. (1998). Role of interleukin 10 in specific immunotherapy. *J Clin Invest* 102:98–106.

Aki T, Kodama T, Fujikawa A, Miura K, Shigeta S, Wada T, Jyo T, Murooka Y, Oka S, Ono K. (1995). Immunochemical characterization of recombinant and native tropomyosins as a new allergen from the house dust mite, *Dermatophagoides farinae*. *J Allergy Clin Immunol* 96: 74–83.

Akoum H, Tsicopoulos A, Vorng H, Wallaert B, Dessaint JP, Joseph M, Hamid Q, Tonnel AB. (1996).Venom immunotherapy modulates interleukin-4 and interferon-gamma messenger RNA expression of peripheral T lymphocytes. *Immunology* 87:593–8.

Andreev OA and Borejdo J. (1995) Binding of heavy-chain and essential light-chain 1 of S1 to actin depends on the degree of saturation of F-actin filaments with S1. *Biochemistry* 34: 14829–33.

Arruda LK, Vailes LD, Benjamin DC, Chapman MD. (1995). Molecular cloning of German cockroach (Blattella germanica) allergens. *Int Arch Allergy Immunol.* 107(1-3):295-7.

Arshad SH. (2002) Allergy: An illustrated colour text. *Churchill livingstone, Edinburgh, London, New York, Philadelphia, Sydney, Toronto*.

Asero R, Mistrello G, Falagiani P. (1997) Oral allergy syndrome from pork. *Allergy*. 52(6):684-6.

Ashwell G and Harford J. (1982). Carbohydrate-specific receptors of the liver. *Annu Rev Biochem* 51:531-54.

Asturias JA, Gomez-Bayon N, Arilla MC, Martinez A, Palacios R, Sanchez-Gascon F, Martinez J (1999). Molecular characterization of American cockroach tropomyosin (*Periplaneta americana* allergen 7), a cross-reactive allergen. *J Immunol* 162 (7): 4342-4348.

Atanaskovic-Markovic M, Gavrovic-Jankulovic M, Jankov RM, Vuekovic O, Nestorovic B. (2002). Food allergy to pork meat. *Allergy* 57(10):960-1.

Ayuso R, Lehrer SB, Lopez M, Reese G, Ibanez MD, Esteban MM, Ownby DR, Schwartz H. (2000). Identification of bovine IgG as a major cross-reactive vertebrate meat allergen. *Allergy* 55(4):348-54.

Ayuso R, Lehrer SB, Tanaka L, Ibanez MD, Pascual C, Burks AW, Sussman GL, Goldberg B, Lopez M, Reese G. (1999). IgE antibody response to vertebrate meat proteins including tropomyosin. *Ann Allergy Asthma Immunol* 83(5):399-405.

Bacarese-Hamilton T, Gray J, Ardizzoni A, Crisanti A. (2005). Allergen microarrays. *Methods Mol Med* 114:195-207.

Bailey, T. L. and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 2, 28-36

Bellinghausen I, Metz G, Enk AH, Christmann S, Knop J, Saloga J. (1997). Insect venom immunotherapy induces interleukin-10 production and a Th2- to-Th1 shift, and changes surface marker expression in venom-allergic subjects. *Eur J Immunol* 27:1131–9.

Benito MJ, Rodriguez M, Nunez F, Asensio MA, Bermudez ME, Cordoba JJ. (2002). Purification and characterization of an extracellular protease from Penicillium chrysogenum Pg222 active against meat proteins. *Appl Environ Microbiol* 68(7):3532-6.

Benjaponpitak S, Oro A, Maguire P, Marinkovich V, DeKruyff RH, Umetsu DT. (1999). The kinetics of change in cytokine production by CD4 T cells during conventional allergen immunotherapy. *J Allergy Clin Immunol* 103:468–75.

Besler M, Fiocchi A, Restani P. (2001). Allergen data collection: Pork (*Sus scrofa*) Internet symposium on food allergens 3(4): 185-92. http://www.food-allergens.de

Besler M, Fiocchi A, Restani P. (2001b). Allergen data collection: Beef (*bos domesticus*) Internet symposium on food allergens 3(4): 171-84. http://www.food-allergens.de

Beyer K, Morrow E, Li XM, Bardina L, Bannon GA, Burks AW, Sampson HA. (2001). Effects of cooking methods on peanut allergenicity. *J Allergy Clin Immunol* 107:1077-81.

Beyer K, Teuber S. (2004). The mechanism of food allergy: what do we know today? *Curr Opin Allergy Clin Immunol*. 4(3):197-9.

Biedermann T, Schopf P, Rueff F, Przybilla B. (1999). Exertion-induced anaphylaxis after eating pork and beef. *Dtsch Med Wochenschr* 16;124(15):456-8.

Bjorck L and Kronvall G. (1984). Purification and some properties of streptococcal protein G, a novel IgG-binding reagent. *J Immunol* 133(2):969-74.

Blom N, Sicheritz-Ponten T, Gupta R, Gammeltoft S, Brunak S. (2004). Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* 4(6):1633-49.

Bourne HC, Restani P, Moutzouris M, Katelaris CH. (2005). An unusual pattern of meat allergy. *Allergy* 60(5):706-7.

Braun-Fahrlander C. (2002). Does the 'hygiene hypothesis' provide an explanation for the relatively low prevalence of asthma in Bangladesh. *Int J Epidemiol* .31(2):488 9.

Breitenbach M., Simon B., Probst G., Oberkofler H., Ferreira F., Briza P., Achatz G.,Unger A., Ebner C., Kraft D., Hirschwehr R. (1997) *Int Arch Allergy Immnol* 113: 114-117.

Breiteneder H, Ebner C. (2000). Molecular and biochemical classification of plant-derived food allergens. *J Allergy Clin Immunol* 106(1 Pt 1):27-36

Breiteneder H, Krebitz M, Wiedermann U, Wagner B, Essl D, Steinkellner H, Turpen TH, Ebner C, Buck D, Niggemann B, Scheiner O. (2001). Rapid production of recombinant allergens in Nicotiana benthamiana and their impact on diagnosis and therapy. *Int Arch Allergy Immunol* 124(1-3):48-50.

Breiteneder H and Radauer C. (2004). A classification of plant food allergens. *J Allergy Clin Immunol* 113(5):821-30

Brown JH, Kim KH, Jun G, Greenfield NJ, Dominguez R, Volkmann N, Hitchcock DeGregori SE, Cohen C. (2001). Deciphering the design of the tropomyosin molecule. *Proc Natl Acad Sci U S A* 17;98(15):8496-501.

Brown JR (1975) Structure of bovine serum albumin (abstract). *Fed Proc* 34:2105

Brown JR (1977) Serum albumin: amino acid sequence; in Rosenoer VM, Oratz M, Rothschild MA (eds): Albumin structure, function and uses. *Oxford, Pergamon* pp 27-51.

Bruijnzeel-Koomen C, Ortolani C, Aas K, Bindslev-Jensen C, Bjorksten B, Moneret Vautrin D, Wuthrich B. (1995). Adverse reactions to food. European Academy of Allergology and Clinical Immunology Subcommittee. *Allergy* 50(8):623-35.

Burks A., Cockrell G, Stanley J, Helm R, Bannon G. (1995). Recombinant peanut allergen Ara h 1 expression and IgE binding in patients with peanut hypersensitivity. *The Journal of Clinical Investigation* 96:1715-1721.

Cahen YD, Fritsch R, Wuthrich B. (1998). Food allergy with monovalent sensitivity to poultry meat. *Clin Exp Allergy* 28(8):1026-30.

Chapman M.D., Smith A.M., Vailes L.D., Arruda L.K., Dhanarajj V., Pomes A. (2000). Recombinant allergens for diagnosis and therapy of allergic disease. *J Allergy Clin Immunol* 106: 409-418.

Chiung YM, Lin BL, Yeh CH, Lin CY. (2000). Heat shock protein (hsp 70)-related epitopes are common allergenic determinants for barley and corn antigens. *Electrophoresis* 21(2):297-300.

Chung SY, Butts CL, Maleki SJ, Champagne E. (2003). Linking peanut allergenicity to the processes of maturation, curing, and roasting. *J Agric Food Chem* 51:4273-7.

Claverie JM. (1992). Identifying coding exons by similarity search: alu-derived and other potentially misleading protein sequences. *Genomics* 12(4):838-41.

Coggeshall KM. (1998). Inhibitory signaling by B cell Fc gamma RIIb. *Curr Opin Immunol* 10(3):306-12.

Cooke RA, Bernhard JH, Hebald S, Stull A. (1935). Serological evidence of immunity with co-existing sensitization in hay fever type of human allergy. *J Exp Med* 62: 733-750.

Crawford LV. Allergy diets. In: Bierman CW, Pearlman DS (eds). (1980). Allergic disease of infancy. Childhood and adolescence. *Philadelphia. PA: WB Saunders* pp 394-400.

Crespo JF and Rodriguez J. (2003). Food allergy in adulthood. *Allergy* 58(2):98-113.

Cromwell O, Suck R, Kahlert H, Nandy A, Weber B, Fiebig H. (2004). Transition of recombinant allergens from bench to clinical application. *Methods* 32(3):300 12.

Devey, M. E., Wilson DV,Wheeler AW. (1976). The IgG subclasses of antibodies to grass pollen allergens produced in hay fever patients during hyposensitization. *Clin. Allergy* 6:227.

Djurup R and Osterballe O. (1984). IgG subclass antibody response in grass pollen-allergic patients undergoing specific immunotherapy. Prognostic value of serum IgG subclass antibody levels early in immunotherapy. *Allergy* 39:433–41.

Dolen WK (2003). IgE antibody in serum ─ detection and diagnostic significance. *Allergy* 58: 717-723

Dolen WK. (1995). Allergen extract standardization: reality, myth, or dream*? Ann Allergy Asthma Immunol* 75:81–82

Dombrowicz D. and Capron M. (2001). Eosinophils, allergy and parasites. *Curr Opin Immunol* 13:716-720.

Drouet M, Sabbah A, Le Sellin J, Bonneau JC, Gay G, Dubois-Gosnet C. (2001). Fatal anaphylaxis after eating wild boar meat in a patient with pork-cat syndrome. *Allerg Immunol (Paris)* 33(4):163-5.

Drouet M and Sabbah A. (1996). The pork/cat syndrome or crossed reactivity between cat epithelia and pork meat. *Monogr Allergy* 32:164-73.

Ebner C, Schenk S, Najafian N, Siemann U, Steiner R, Fischer GW, Hoffmann K, Szepfalusi Z, Scheiner O, Kraft D. (1995). Nonallergic individuals recognized the same T cell epitopes of Bet v 1, the major birch pollen allergen, as atopic patients. *J Immunol* 154:1932-1940.

Ebo DG, Hagendorens MM, Bridts CH, De Clerck LS, Stevens WJ. (2004). Sensitization to cross-reactive carbohydrate determinants and the ubiquitous protein profilin: mimickers of allergy. *Clin Exp Allergy* 34(1):137-44

Eigenmann PA, Sicherer SH, Borkowski TA, Cohen BA, Sampson HA. (1998). Prevalence of IgE-mediated food allergy among children with atopic dermatitis. *Pediatrics* 101(3):E8.

Emson CL, Bell SE, Jones A, Wisden W, McKenzie AN. (1998). Interleukin (IL)-4-independent induction of immunoglobulin (Ig)E, and perturbation of T cell development in transgenic mice expressing IL-13. *J Exp Med* 188(2): 399-404.

Ewan PW, Deighton J, Wilson AB, Lachmann PJ. (1993). Venom-specific IgG antibodies in bee and wasp allergy: lack of correlation with protection from stings. *Clin Exp Allergy* 23:647–60.

FAO/WHO, (2001). Evaluation of allergenicity of genetically modified foods. Report of a Joint FAO/WHO Expert Consultation on Allergenicity of Foods derived from Biotechnology; 22–25 January 2001; Rome, Italy.

Faye L and Chrispeels M.J. (1988). Common antigenic determinants in the glycoproteins of plants, molluscs and insects. *Glycoconjugate J* 5: 245.

Ferreira F, Hirtenlehner K, Briza P, Breiteneder H, Scheiner O, Kraft D, Breitenbach M, Ebner C. (1997). Isoforms of atopic allergens with reduced allergenicity but conserved T cell antigenicity: Possible use for specific immunotherapy. *Int Arch Allergy Immunol* 113:125-127.

Ferreira F, Hirtenlehner K, Jilek A, Godnik-Cvar J, Breiteneder H, Grimm R, Hoffmann-Sommergruder K, Scheiner O, Kraft D, Breitenbach M, Rheinberger H.J, Ebner C. (1996). Dissection of immunoglobulin E and T lymphocyte reactivity of isoforms of the major birch pollen allergen Bet v 1: Potential use of hypoallergenic isoforms for immunotherapy. *J Exp Med* 183:599-609.

Ferreira F.D, Hoffmann-Sommergruber K, Breiteneder H, Pettenburger K, Ebner C, Sommergruber W, Steiner R., Bohle B, Sperr WR, Valent P, Kungl AJ, Breitenbach M, Kraft D, Scheiner O. (1993). Purification and characterization of recombinant Bet v 1, the major birch pollen allergen. *The Journal of Biological Chemistry* 268: 19574-19580.

Finkelman FD, Katona IM, Urban JF..Jr, Holmes J, Ohara J, Tung AS, Sample JV, Paul WE. (1988). IL-4 is required to generate and sustain *in vivo* IgE responses. *J Immunol* 141:2335-2341.

Fiocchi A, Bouygue GR, Restani P, Bonvini G, Startari R, Terracciano L. (2002) Accuracy of skin prick tests in IgE-mediated adverse reactions to bovine proteins. *Ann Allergy Asthma Immunol* 89(6 Suppl 1):26-32.

Fiocchi A, Restani P, Bouygue GR, Martelli A. (2005). Beef allergy in adults and children. Allergy 60(1):126.

Fiocchi A, Restani P, Riva E, Qualizza R, Bruni P, Restelli AR, Galli CL. (1995). Meat allergy: I--Specific IgE to BSA and OSA in atopic, beef sensitive children. J *Am Coll Nutr* 14(3):239-44.

Fiocchi A and Restani P. (2002). Adverse reactions to bovine proteins--then and now. *Ann Allergy Asthma Immunol*. 89(6 Suppl 1):1-2.

Flicker S and Valenta R. (2003). Renaissance of the blocking antibody concept in type I allergy. *Int Arch Allergy Immunol*; 132:13–24.

Foetisch K, Son DY, Altmann F, Aulepp H, Conti A, Haustein D. (2001). Tomato (*Lycopersicon esculentum*) allergens in pollen-allergic patients. *Eur Food Res Technol* 213:259-66.

Foetisch K, Westphal S, Lauer I, Retzek M, Altmann F, Kolarich D, Scheurer S, Vieths S. (2003). Biological activity of IgE specific for cross-reactive carbohydrate determinants. *J Allergy Clin Immunol* 111(4):889-96.

Francis JN, Till SJ, Durham SR. (2003). Induction of IL-101CD41CD251 T cells by grass pollen immunotherapy. *J Allergy Clin Immunol* 111:1255–61.

Fuentes MM, Palacios R, Garces MM, Caballero ML, Moneo I. (2004). Isolation and characterization of a heat-resistant beef allergen: myoglobin. *Allergy* 59(3):327-31.

Gajhede M, Osmark P, Poulsen FM, Ipsen H, Larsen J, van Neerven J, Schou C, Lowenstein H, Spangfort MD. (1996). X-ray and NMR structure of Bet v 1, the origin of birch pollen allergy. *Nature Structural Biology* 3:1040-5.

Gehlhar K, Schlaak M, Becker W, Bufe A. (1999). Monitoring allergen immunotherapy of pollen-allergic patients: the ratio of allergenspecific IgG4 to IgG1 correlates with clinical outcome. *Clin Exp Allergy* 29:497–506.

Gelamo EL and Tabak M. (2000). Spectroscopic studies on the interaction of bovine (BSA) and human (HSA) serum albumins with ionic surfactants. *Spectrochim Acta A Mol Biomol Spectrosc* 56A(11):2255-71.

Gell PGH and Coombs RRA. (1963) The classification of allergic reactions underlying disease. In Clinical Aspects of Immunology (Coombs RRA and Gell PGH, eds) *Blackwell Science*, UK.

Gendel, SM. (2002). Sequence analysis for assessing potential allergenicity. *Ann N.Y. Acad. Sci* 964:87-98.

Gergely J. (1967)."Structural studies of Igs". *Immunochemistry* 4,101

Gesundheit N, Fink DL, Silverman LA, Weintraub BD. (1987). Effect of thyrotropin-releasing hormone on the carbohydrate structure of secreted mouse thyrotropin. Analysis by lectin affinity chromatography. *J Biol Chem* 262(11):5197-203

Ghosh S, Janocha AJ, Aronica MA, Swaidani S, Comhair SA, Xu W, Zheng L, Kaveti S, Kinter M, Hazen SL, Erzurum SC. (2006). Nitrotyrosine proteome survey in asthma identifies oxidative mechanism of catalase inactivation. *J Immunol.*176(9):5587-97.

Gleich GJ, Zimmermann EM, Henderson LL, Yunginger JW. (1982). Effect of immunotherapy on immunoglobulin E and immunoglobulin G antibodies to ragweed antigens: a six-year prospective study. *J Allergy Clin Immunol* 70:261–71.

Green P, Shownkeen R, Metzstein N, Hawkins T, Wilson R, Berks M, Du Gruehn S, Suphioglu C, O'Hehir RE, Volkmann D. (2003). Molecular cloning and characterization of hazel pollen protein (70 kD) as a luminal binding protein (BiP): a novel cross-reactive plant allergen. *Int Arch Allergy Immunol* 131(2):91-100.

Hamilton RG, Biagini RE, Krieg EF. (1999). Diagnostic performance of Food and Drug Administration-cleared serologic assays for natural rubber latex-specific IgE antibody. The Multi-Center Latex Skin Testing Study Task Force. *J Allergy Clin Immunol* 103(5 Pt 1):925-30.

Han GD, Matsuno M, Ikeuchi Y, Suzuki A. (2002). Effects of heat and high-pressure treatments on antigenicity of beef extract. *Biosci Biotechnol Biochem* 66(1):202-5.

Hardy CL, Rolland JM, O'Hehir RE. (2004). Blocking antibodies in allergen immunotherapy: the Yin and Yang. *Clin Exp Allergy* 34(4):510-2.

Hileman RE, Silvanovich A, Goodman RE, Rice EA, Holleschak G, Astwood JD, Hee, SL. (2002). Bioinformatic methods for allergenicity assessment using a comprehensive allergen database. *Int. Arch. Allergy Immunol* 128(4):280-291.

Hilger C, Kohnen M, Grigioni F, Lehners C, Hentges F. (1997). Allergic cross-reactions between cat and pig serum albumin. Study at the protein and DNA levels. *Allergy* 52(2):179-87.

Hill DJ, Hosking CS, Heine RG. (1999). Clinical spectrum of food allergy in children in Australia and South-East Asia: identification and targets for treatment. *Ann Med* 31(4):272-81.

Hill DJ and Hosking CS. (1995). The cow milk allergy complex: overlapping disease profiles in infancy. *Eur J Clin Nutr* 49 Suppl 1:S1-12.

Hirayama K, Akashi S, Furuya M, Fukuhara K. (1990). Rapid confirmation and revision of the primary structure of bovine serum albumin by ESIMS and Frit-FAB LC/MS. *Biochem Biophys Res Commun* 14;173(2):639-46.

Holowachuk EW, Stoltenborg JK, Reed RG, Peter T Jr. (1991). Bovine serum albumin: cDNA sequence and expression. Submitted to EMBL data library.

Horvath H, Huang J, Wong O, Kohl E, Okita T, Kannangara CG, von Wettstein D. (2000). The production of recombinant proteins in transgenic barley grains. *Proc Natl Acad Sci U S A* 97(4):1914-9.

Ipsen H and Lowenstein H. (1997). Basic features of crossreactivity in tree and grass pollen allergy. *Clin Rev Allergy Immunol* 15(4):389-96

Ishikawa M, Ishida M, Shimakura K, Nagashima Y, Shiomi K.(1998). Purification and IgE-binding epitopes of a major allergen in the gastropod Turbo cornutus. .*Biosci Biotechnol Biochem* 62(7):1337-43

Ivanciuc O, Mathura V, Midoro-Horiuti T, Braun W, Goldblum RM, Schein CH. (2003b). Detecting potential IgE-reactive sites on food proteins using a sequence and structure database, SDAP-food. *J Agric Food Chem* 51(16):4830-7.

Ivanciuc O, Schein CH, Braun W. (2002). Data mining of sequences and 3D structures of allergenic proteins. *Bioinformatics* 18(10):1358-64.

Ivanciuc O, Schein CH, Braun W. (2003a) SDAP: database and computational tools for allergenic proteins. *Nucleic Acids Res* 31(1):359-62.

Potskaldy JM, Rouiller DG, Grunberger G, Baxter RC, McElduff A, Gordon P. (1986). Glycosylation defects alter insulin but not insulin-like growth factor I binding to Chinese hamster ovary cells. *J Biol Chem* 261(30):14076-81.

Jakobsen CG, Bodtger U, Poulsen LK, Roggen EL. (2005). Vaccination for birch pollen allergy: comparison of the affinities of specific immunoglobulins E, G1 and G4 measured by surface plasmon resonance. *Clin Exp Allergy*.35(2):193-8.

Johansson SG, Bieber T, Dahl R, Friedmann PS, Lanier BQ, Lockey RF, Motala C, Ortega Martell JA, Platts-Mills TA, Ring J, Thien F, Van Cauwenberge P, Williams HC. (2004).Revised nomenclature for allergy for global use: Report of the Nomenclature Review Committee of the World Allergy Organization, October 2003. *J Allergy Clin Immunol* 113(5):832-6.

Johnson LR: Regulation of the gastrointestinal growth. (1997). In Makhlouf GM, (ed); "Physiology of the gastrointestinal tract 2nd ed". *Raven Press*, New York pp 301-333.

Jutel M, Akdis M, Budak F, Aebischer-Casaulta C, Wrzyszcz M, Blaser K, Akdis CA. (2003). IL-10 and TGF-beta cooperate in the regulatory T cell response to mucosal allergens in normal immunity and specific immunotherapy. *Eur J Immunol* 33:1205–14.

Jutel M, Pichler WJ, Skrbic D, Urwyler A, Dahinden C, Muller UR. (1995). Bee venom immunotherapy results in decrease of IL-4 and IL-5 and increase of IFN-gamma secretion in specific allergenstimulated T cell cultures. *J Immunol* 154:4187–94.

Kagan RS. (2003). Food Allergy: An Overview. *Environmental Health Perspectives* Volume 111(2).

Kanny G, Puygrenier J, Beaudoin E, Moneret-Vautrin DA. (1994). Alimentary anaphylactic shock: implication of penicillin residues. *Allerg Immunol (Paris)* 26(5):181-3.

Kemeny DM, Urbanek R, Ewan P, McHugh S, Richards D, Patel S, Lessof MH. (1989). The subclass of IgG antibody in allergic disease: II. The IgG subclass of antibodies produced following natural exposure to dust mites and grass pollen in atopic and non-atopic individuals. *Clin Exp Allergy* 19:545-549

Khan J, Saal LH, Bittner ML, Chen Y, Trent JM, Meltzer PS. (1999). Expression profiling in cancer using cDNA microarrays. *Electrophoresis* 20:223-229.

Kimura I, Tanizaki Y, Goda Y, Komagoe H, Kitani H. (1985). Decrease in reactivity of basophils by immunotherapy with housedust extract. *Clin Allergy* 15:1–7.

Kinet J.P. (1999). The high affinity IgE receptor (Fc-epsilon RI): from physiology to pathology. *Annu Rev Immunol* 17:931-972.

Kleter GA and Peijnenburg AA. (2002). Screening of transgenic proteins expressed in transgenic food crops for the presence of short amino acid sequences identical to potential, IgE - binding linear epitopes of allergens. *BMC Struct Biol*. 2:8.

Knippels LM and Penninks AH. (2005). Recent advances using rodent models for predicting human allergenicity. *Toxicol Appl Pharmacol* 207(2 Suppl):157-60.

Kochuyt AM. (2000). Insect venom anaphylaxis: from diagnosis to therapy. Thesis. Medical Faculty, Univeristy Leuven, Leuven, Belgium.

Krishna TM, Mavroleon G, Holgate S. (2001). Essentials of allergy. *Martin Dunitz*, London. UK.

Lascombe MB, Gregoire C, Poncet P, Tavares GA, Rosinski-Chupin I, Rabillon J, Goubran-Botros H, Mazie JC, David B, Alzari PM. (2000). Crystal structure of the allergen Equ c 1. A dimeric lipocalin with restricted IgE-reactive epitopes. *J Biol Chem* 275(28):21572-7.

Lau S, Matricardi PM. Worms, asthma, and the hygiene hypothesis. (2006). *Lancet* 367(9522):1556-8.

Lebrun SJ, Petchpud WN, Hui A, McLaughlin CS. (2005).Development of a sensitive, colorometric microarray assay for allergen-responsive human IgE. *J Immunol Methods* 300(1-2):24-31.

Lethanh H, Neubauer P, Hoffmann F. (2005). The small heat-shock proteins IbpA and IbpB reduce the stress load of recombinant Escherichia coli and delay degradation of inclusion bodies. *Microb Cell Fact* 4(1):6.

Li KB, Issac P, Krishnan A. (2004). Predicting allergenic proteins using wavelet transform. *Bioinformatics* 20(16):2572-8.

Llatser R, Polo F, De La Hoz F, Guillaumet B. (1998). Alimentary allergy to pork. Crossreactivity among pork kidney and pork and lamb gut. *Clin Exp Allergy* 28(8):1021-5.

Lynch NR, Lopez RI, Di Prisco-Fuenmayor MC, Hagel I, Medouze L, Viana G, Ortega C, Prato G. (1987). Allergic reactivity and socio-economic level in a tropical environment. *Clin Allergy* 17: 199–207.

Malandain H. (2004).Basic immunology, allergen prediction, and bioinformatics. *Allergy*. 59(9):1011-2.

Maleki SJ, Chung SY, Champagne ET, Raufman JP. (2000). The effects of roasting on the allergenic properties of peanut proteins. *J Allergy Clin Immunol* 106:763-8.

Maleki SJ, Viquez OM, Jacks T, Champagne ET, Chung SY, Landry SJ. (2003). The major peanut allergen, Ara h 2, functions as a trypsin inhibitor, and roasting enhances this function. *J Allergy Clin Immunol* 112:190-5.

Mamikoglu B. (2005). Beef, pork, and milk allergy (cross reactivity with each other and pet allergies). *Otolaryngol Head Neck Surg* 133(4):534-7.

Mari A, Iacovacci P, Afferni C, Barletta B, Tinghino R, Di Felice G, Pini C. (1999). Specific IgE to cross-reactive carbohydrate determinants strongly affect the in vitro diagnosis of allergenic diseases. *J Allergy Clin Immunol* 103:1005-11.

Mari, A. (2002). IgE to cross-reactive carbohydrate determinants: analysis of the distribution and appraisal of the in vivo and in vitro reactivity. *Int. Arch. Allergy Immunol*. 129: 286 – 295.

Marra MA, Kucaba TA, Hillier LW, Waterston RH. (1999). High-throughput plasmid DNA purification for 3 cents per sample. *Nucleic Acids Research* 27: e37.

Martinez A, Martinez J, Palacios R, Panzani R. (1997). Importance of tropomyosin in allergy to household arthropods. Cross-reactivity with other invertebrate extracts. *Allergologia et Immunopathologia* 25:118-126.

Marvin LF, Roberts MA, Fay LB. (2003). Matrix-assisted laser desorption/ionization time-of flight mass spectrometry in clinical chemistry. *Clin Chim Acta* 337(1-2):11-21.

Masini E, Di Bello MG, Cappugi P, Berni L, Mirabella C, Mannaioni PF. (1997). Platelet aggregation and platelet histamine release by immunological stimulation in atopic patients: modulation by nitric oxide. *Inflamm Res.* 46 Suppl 1:S81-2

Matsson P. (1997). Evaluation methods and analytical performance characteristics of immunological assays for human immunoglobulin E (IgE) antibodies of defined allergen specificities. Wayne, PA: NCCLS.

McCann WA and Ownby DR. (2002). The reproducibility of the allergy skin test scoring and interpretation by board-certified/board-eligible allergists. *Ann Allergy Asthma Immunol* 89(4):368-71.

McCombie WR, Heiner C, Kelley JM, Fitzgerald MG, Gocayne JD. (1992). Rapid and reliable fluorescent cycle sequencing of double-stranded templates. *DNA Seq* 2(5):289-96.

McHugh SM, Deighton J, Stewart AG, Lachmann PJ, Ewan PW. (1995). Bee venom immunotherapy induces a shift in cytokine responses from a TH- 2 to a TH-1 dominant pattern: comparison of rush and conventional immunotherapy. *Clin Exp Allergy* 25:828–38.

McHugh SM, Lavelle B, Kemeny DM, Patel S, Ewan PW. (1990). A placebo-controlled trial of immunotherapy with two extracts of Dermatophagoides pteronyssinus in allergic rhinitis, comparing clinical outcome with changes in antigen-specific IgE, IgG, and IgG subclasses. *J Allergy Clin Immunol* 86:521–31.

Metcalfe DD, Astwood JD, Townsend R, Sampson HA, Taylor SL, Fuchs RL. (1996). Assessment of the allergenic potential of foods from genetically engineered crop plants. *Crit. Rev. Food Sci. Nutr* 36(S): 165– 186.

Michils A, Ledent C, Mairesse M, Gossart B, Duchateau J. (1997). Wasp venom immunotherapy changes IgG antibody specificity. *Clin Exp Allergy* 27:1036–42.

Midoro-Horiuti T, Goldblum RM, Kurosky A, Wood TG, Schein CH, Brooks EG. (1999). Molecular cloning of the mountain cedar (Juniperus ashei) pollen major allergen, Jun a 1. *J Allergy Clin Immunol* 104(3 Pt 1):613-7

Mifflin LC and Cohen RE. (1994). Characterization of denatured protein inducers of the heat shock (stress) response in Xenopus laevis oocytes. *J Biol Chem* 269(22):15710-7.

Mita H, Yasueda H, Akiyama K. (2000). Affinity of IgE antibody to antigen influences allergen-induced histamine release. *Clin Exp Allergy* 30:1582-9.

Moser M, Crameri R, Menz G, Scheneider T, Dudler T, Virchow C, Gmachl M, Blaser K, Suter M. (1992). Cloning and expression of recombinant *Aspergillus fumigatus* allergen I/a (rAsp f I/a) with IgE binding and type I skin test activity. *J Immunol* 149:454-460.

Moss RB, Hsu YP, Kwasnicki JM, Sullivan MM, Reid MJ. (1987). Isotypic and antigenic restriction of the blocking antibody response to ryegrass pollen: correlation of rye group I antigen-specific IgG1 with clinical response. *J Allergy Clin Immunol* 79:387–98.

Mowat AM. (2003). Anatomical basis of tolerance and immunity to intestinal antigens. *Nat Rev Immunol* 3(4):331-41.

Muller U, Helbling A, Bischof M. (1989). Predictive value of venomspecific IgE, IgG and IgG subclass antibodies in patients on immunotherapy with honey bee venom. *Allergy* 44:412–8.

Natale M, Bisson C, Monti G, Peltran A, Garoffo LP, Valentini S, Fabris C, Bertino E, Coscia A, Conti A. (2004). Cow's milk allergens identification by two-dimensional immunoblotting and mass spectrometry. *Mol Nutr Food Res* 48(5):363-9.

Nelson HS, Ikle D, Buchmeier A. (1996). Studies of allergen extract stability: The effects of dilution and mixing. *J Allergy Clin Immunol* 98:382-288.

Nelson HS. (1994). Variables in allergy skin testing. *Allergy Proceedings* 15: 265-268

Nolte H and DuBuske LM (1997). Performance characteristics of a new automated enzyme immunoassay for the measurement of allergen-specific IgE — Summary of the probability outcomes comparing results of allergen skin testing to results obtained with the HYTEC system and CAP system. *Ann Allergy Asthma Immunol* 79: 27-34

Novembre E, de Martino M, Vierucci A. (1988). Foods and respiratory allergy. *J Allergy Clin Immunol* 81(5 Pt 2):1059-65.

Okubo K, Hori N, Matoba R, Niiyama T, Fukushima A, Kojima Y, Matsubara K. (1992). Large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression. *Nature Genetics* 2:173-179.

Ong ST. (2003). Expressed sequence tags analysis of major allergens producing dust musts and molecular characterization of their allergens. PhD. Thesis submitted to National University of Singapore.

Ortolani C, Bruijnzeel-Koomen C, Bengtsson U, Bindslev-Jensen C, Bjorksten B, Host A, Ispano M, Jarish R, Madsen C, Nekam K, Paganelli R, Poulsen LK, Wuthrich B. (1999). Controversial aspects of adverse reactions to food. European Academy of Allergology and Clinical Immunology (EAACI) Reactions to Food Subcommittee. *Allergy* 54(1):27-45.

Paschinger K, Fabini G, Schuster D, Rendic D, Wilson IB. (2005). Definition of immunogenic carbohydrate epitopes. *Acta Biochim Pol* 52(3):629-32.

Pauli G, Oster JP, Deviller P, Heiss S, Bessot JC, Susani M, Ferreira F, Kraft D, Valenta R. (1996). Skin testing with recombinant allergens rBet v 1 and birch profilin, rBet v2: diagnostic value for birch pollen and associated allergies. *J Allergy Clin Immunol* 97:1100–9.

Pavel AC and Comanescu A. (1969). Anaphylactic shock caused by pork. *Rev Med Chir Soc Med Nat Iasi* 73(2):489-92.

Pène J, Rousset F, Briere F, Chretien I, Bonnefoy JY, Spit H, Yokota T, Arai N, Arai K. (1988). IgE production by normal human lymphocytes is induced by interleukin 4 and suppressed by interferons α, γ and prostaglandin E2. *Proc Natl Acad Sci U S A* 65: 6880.

Peters T Jr, Feldhoff RC, Reed RG. (1977). Immunochemical studies of fragments of bovine serum albumin. *J Biol Chem* 252(23):8464-8.

Petersen A and Mundt C. (2001). Investigations on the carbohydrate moieties of glycoprotein allergens. *J Chromatogr B Biomed Sci Appl* 756(1-2):141-50.

Piva M, Moreno JI, Sharpe-Timms KL. (2002). Glycosylation and over-expression of endometriosis-associated peritoneal haptoglobin. *Glycoconj J* 19(1):33-41.

Plebani M, Bernardi D, Basso D, Borghesan F, Faggian D. (1998). Measurement of specific immunoglobulin E: intermethod comparison and standardization. *Clinical Chemistry* 44:1974-1979.

Pleshkova SM. (1978). Oxidation-reduction processes during sensitization and the allergic reaction to microbial antigens. *Bulletin of Experimental Biology and Medicine* 86(4):1363 – 1366

Pontius JU, Wagner L, Schuler GD. (2003). UniGene: a unified view of the transcriptome. In: The NCBI Handbook. Bethesda (MD): National Center for Biotechnology Information.

Punnonen J, Aversa G, Cocks BG, McKenzie AN, Mehon S, Zurawski G, de waal Malefyt R, de Vries J.E. (1993). Interleukin 13 induces interleukin 4- independent

IgG4 and IgE synthesis and CD23 expression by human B cells. *Proc Natl Acad Sci U S A* 90; 3730-3734.

Rabjohn P, Helm E, Stanley J, West C, Sampson H, Burks A, Bannon G. (1999). Molecular cloning and epitope analysis of the peanut allergen Ara h 3. *The Journal of Clinical Investigation* 103:535-542.

Ravetch JV and Bolland S. (2001). IgG Fc receptors. *Annu Rev Immunol* 19:275-90.

Reese G, Ayuso R, Lehrer SB. (1999). Tropomyosin: an invertebrate pan-allergen. *Int Arch Allergy Immunol* 119(4):247-58.

Reese G, Jeoung BJ, Daul CB & Lehrer SB. (1997). Characterisation of recombinant shrimp allergen Pen a 1 (tropomyosin). *Int Arch Allergy Immunol* 113: 240–242.

Rementeria A, Lopez-Molina N, Ludwig A, Vivanco AB, Bikandi J, Ponton J, Garaizar J. (2005). Genes and molecules involved in Aspergillus fumigatus virulence. *Rev Iberoam Micol* 22(1):1-23.

Restani P, Beretta B, Fiocchi A, Ballabio C, Galli CL. (2002). Cross-reactivity between mammalian proteins. *Ann Allergy Asthma Immunol* 89(6 Suppl 1):11-5.

Restani P, Fiocchi A, Beretta B, Velona T, Giovannini M, Galli CL. (1997). Meat allergy: III--Proteins involved and cross-reactivity between different animal species. *J Am Coll Nutr* 16(4):383-9.

Ricci G, Capelli M, Miniero R, Menna G, Zannarini, Dillon P and Masi M. (2003). A comparisonof different allergometric tests, skin prick test, Pharmacia UniCAP® and ADVIA Centaur®, for diagnosis of allergic disease in children. *Allergy* 58: 38-45

Roecken M, Grevers G, and Burgdorf W. (2004). Color atlas of allergic diseases. *Georg Thieme Velag, Stuttgart, Germany*.

Roitt IM, Brostoff J, Male DK, (1998). Immunology. *London: Mosby*

Romagnani S. (1991). Human TH1 and TH2 subsets: doubt no more. *Immunology Today* 12:256-257.

Sabbah A, Lauret MG, Chene J, Boutet S, Drouet M. (1994a). The pork-cat syndrome or crossed allergy between pork meat and cat epithelia (2). *Allerg Immunol (Paris)* 26(5):173-4, 177-80.

Sabbah A, Rousseau C, Lauret MG, Drouet M. (1994b). The pork-cat syndrome: RAST inhibition test with Fel d 1. *Allerg Immunol (Paris)* 26(7):259-60.

Sambrook J. and Russell D.V. (2001). *Molecular Cloning: A Laboratory Manual, 3rd Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York*.

Sampson HA. (2003). 9. Food allergy. *J Allergy Clin Immunol* 111(2 Suppl):S540-7.

Sampson HA. (1999). Food allergy. Part 1: immunopathogenesis and clinical disorders. *J Allergy Clin Immunol* 103(5 Pt 1):717-28.

Sampson HA. (2004).Update on food allergy. *J Allergy Clin Immunol* 113(5):805-19.

Schmid-Grendelmeier P and Crameri R. (2001). Recombinant allergens for skin testing. 1 *Int Arch Allergy Immunol* 25:96-111.

Schmidt M, Hoffman DR. (2002). Expression systems for production of recombinant allergens. *Int Arch Allergy Immunol* 128(4):264-70.

Secrist H, Chelen CJ, Wen Y, Marshall JD, Umetsu DT. (1993). Allergen immunotherapy decreases interleukin 4 production in CD41 T cells from allergic individuals. *J Exp Med* 178: 2123–30.

Shanti KN, Martin BM, Nagpal S, Metcalfe DD, Rao PV. (1993). Identification of tropomyosin as the major shrimp allergen and characterization of its IgE binding epitopes. *J Immunol* 151:5354-63.

Shen HD, Au LC, Lin WL, Liaw SF, Tsai JJ, Han SH. (1997). Molecular cloning and expression of a *Penicillium citrinum* allergen with sequence homology and antigenic crossreactivity to a hsp 70 human heat shock protein. *Clin Exp Allergy* 27(6):682-90

Simon-Nobbe B, Pobst G, Kajava AV, Oberkofler H, Susani M, Crameri R, Ferreira F, Ebner C, Breitenbach M. (2000) IgE binding epitopes of enolases, a class of highly conserved fungal allergens*. J Allergy Clin Immunol* 106, 887-895.

Soeria-Atmadja D, Zorzet A, Gustafsson MG, Hammerling U. (2004). Statistical evaluation of local alignment features predicting allergenicity using supervised classification algorithms. *Int Arch Allergy Immunol* 133(2):101-12.

Soman KV, Midoro-Horiuti T, Ferreon JC, Goldblum RM, Brooks EG, Kurosky A, Braun W, Schein CH. (2000). Homology modeling and characterization of IgE binding epitopes of mountain cedar allergen Jun a 3. *Biophys J* 79(3):1601-9.

Stadler Mb and Stadler BM. (2003). Allergenicity prediction by protein sequence. *FASEB J* 17(9):1141-3.

Sugiura S, Minamisawa S, Kawada T. (2002). Cardiovascular Physiome: Linking Genes and Molecules to Function - Molecular Physiology of Myosin. The 66th Annual

Scientific Meeting of the Japanese Circulation Society Symposium Basic 1. April 24 - 26

Svirshchevskaya EV, Viskova N, Shevchenko M, Alekseeva L, Marchenko A, Benevolensky S, Kurup VP. (2004). High-affinity IgG to a major A. fumigatus allergen, Asp f 2, retards allergic response. *Med Sci Monit* 10(10):BR371-80.

Szeinbach SL, Barnes JH, Sullivan TJ, Williams PB. (2001). Precision and accuracy of commercial laboratories_ ability to classify positive and/or negative allergen-specific IgE results. *Ann Allergy Asthma Immunol* 86:373– 381.

Tanabe S, Kobayashi Y, Takahata Y, Morimatsu F, Shibata R, Nishimura T. (2002). Some human B and T cell epitopes of bovine serum albumin, the major beef allergen. *Biochem Biophys Res Commun* 293(5):1348-53.

Thomas WR and Smith W. (2002). Editorial: How good are carbohydrates as allergens? *Clin Exp Allergy* 32, 658 – 661.

Till S, Durham S, Dickason R, Huston D, Bungre J, Walker S, Robinson D, Kay AB, Corrigan C. (1997). IL-13 production by allergen stimulated T cells is increased in allergic disease and associated with IL-5 but not IFN-gamma expression. *Immunology* 91:53-57.

Tretter V, Altmann F, Kubelka V, Marz L, Becker WM. (1993). Fucose alpha 1,3-linked to the core region of glycoprotein N-glycans creates an important epitope for IgE from honeybee venom allergic individuals. *Int Arch Allergy Immunol* 102:259–66.

Tsai L, Sun Y, Chao P, Ng H, Hung M, Hsieh K, Liaw S, Chua K. (1999). Sequence analysis and expression of a cDNA clone encoding a 98-kDa allergen in *Dermatophagoides farinae*. *Clin Exp Allergy* 29(12):1606-13.

Valenta R, Duchene M, Ebner C, Valent P, Sillaber C, Deviller P, Ferreira F, Tejkl M, Edelmann H, Kraft D, Scheiner O. (1992). Profilins constitute a novel family of functional plant pan-allergens. *J Exp Med* 175:377–85.

Valenta R, Hayek B, Seiberler S, Bugajska-Schretter A, Niederberger V, Twardosz A, Natter S, Vangelista L, Pastore A, Spitzauer S, Kraft D. (1998). Calcium-binding allergens: from plants to man. *Int Arch Allergy Immunol* 117(3):160-6.

Valenta R. (2002). The future of antigen-specific immunotherapy of allergy. *Nature Reviews: Immunology* 2:446-453.

Valenta R., Duchêne M., Vrtala S., Birkner T., Ebner C., Hirschwehr R., Breitenbach

Vallier P, Balland S, Harf R, Valenta R, Deviller P. (1995). Identificationof profilin as an IgE-binding component in latex from *Hevea brasiliensis*: clinical implications. *Clin Exp Allergy* 25:332–9.

van der Veen M.J., Nukder M., Witteman A.M., van Ree R., Aalberse R.C., Jansen H.M. (1996). False positive skin prick test responses to commercially available dog dander extracts caused by contamination with house dust mite (*Dermatophagoides pteronyssinus*) allergens. *J Allergy Clin Immunol* 98:1028-1034.

van der Veen MJ, van Ree R, Aalberse RC, Akkerdaas J, Koppelman SJ, Jansen HM, van der Zee JS. (1997). Poor biologic activity of cross-reactive IgE directed to carbohydrate determinants of glycoproteins. *J Allergy Clin Immunol* 100:327-34.

van der Zee JS, de Groot H, van Swieten P, Jansen HM, Aalberse RC. (1988). Discrepancies between the skin test and IgE antibody assays: study of mediator release, complement activation in vitro, and occurrence of allergen-specific IgG. *J Allergy Clin Immunol* 82:270–281.

van Neerven RJ, Arvidsson M, Ipsen H, Sparholt SH, Rak S, Wurtzen PA. (2004). A double-blind, placebo-controlled birch allergy vaccination study: inhibition of CD23 mediated serum-immunoglobulin E-facilitated allergen presentation. *Clin Exp Allergy* 34(3):420-8.

van Neerven RJ, Wikborg T, Lund G, Jacobsen B, Brinch-Nielsen A, Arnved J, Ipsen H. (1999). Blocking antibodies induced by specific allergy vaccination prevent the activation of CD41 T cells by inhibiting serum-IgE facilitated allergen presentation. *J Immunol* 163:2944–52.

van Ree R and Aalberse RC. (1993). Pollen–vegetable food cross-reactivity: serological and clinical relevance of cross-reactive IgE. *J Clin Immunoassay* 16:124–30.

van Ree R. (2004). Clinical importance of cross-reactivity in food allergy. *Curr Opin Allergy Clin Immunol* 4(3):235-40.

van Ree R. (1999). Specific IgE without clinical allergy. *J Allergy Clin Immunol* 103:1000-1.

van Ree, R., W. A. van Leeuwen, P. H. Dieges, R. G. van Wijk, N. de Jong, P. Z. Brewczyski, A. M. Kroon, P. P. Schilte, K. Y. Tan, I. F. Simon-Licht, et al. (1997). Measurement of IgG and IgG antibodies against purified grass pollen allergens (Lol p 1, 2, 3 and 5) during immunotherapy. *Clin. Exp. Allergy* 27:68.

Vieths S, Scheurer S, Ballmer-Weber B. (2002). Current understanding of cross-reactivity of food allergens and pollen. *Ann N Y Acad Sci* 964:47–68.

Vieths, S, Scheurer S. Reindl J, Lüttkopf D, Wangorsch A, Kästner M, Haase T, Haustein. D. (2001). Optimized allergen extracts and recombinant allergens in diagnostic applications. *Allergy* 56(Supp. 67): 78 – 82.

Wachholz PA and Durham SR. (2004). Mechanisms of immunotherapy: IgG revisited. *Curr Opin Allergy Clin Immunol* 4(4):313-8.

Wachholz PA, Soni NK, Till SJ, Durham SR. (2003). Inhibition of allergen-IgE binding to B cells by IgG antibodies after grass pollen immunotherapy. *J Allergy Clin Immunol* 112: 915–22.

Wagner S, Breiteneder H, Simon-Nobbe B, Susani M, Krebitz M, Niggemann B, Brehler R, Scheiner O, Hoffmann-Sommergruber K. (2000). Hev b 9, an enolase and a new cross-reactive allergen from Hevea latex and mold. *Eur. J. Biochem* 267, 7006-7014.

Wal JM. (1999). Assessment of allergic potential of (novel) foods. *Nahrung* 43, 168.

Waterston R, Martin C, Craxton M, Huynh C, Coulson A, Hillier L, Durbin R, Wellhausen A, Schoning B, Petersen A, Vieths S. (1996). IgE binding to a new cross-reactive structure: a 35 kDa protein in birch pollen, exotic fruit and other plant foods. *Z Ernahrungswiss* 35(4):348-55.

Welt K, Hinrichs R, Ott S, Thalmann M, Dieckmannken J, Schneider LA, Staib G, Scharffetter-Kochanek K. (2005). Anaphylaxis after the ingestion of lamb meat. *Allergy* 60(4):545.

Wensing M, Akkerdaas JH, van Leeuwen WA, Stapel SO, Bruijnzeel-Koomen CA, Aalberse RC, Bast BJ, Knulst AC, van Ree R. (2002). IgE to Bet v 1 and profilin: cross-reactivity patterns and clinical relevance. *J Allergy Clin Immunol* 110:435–42.

Werfel SJ, Cooke SK, Sampson HA. (1997). Clinical reactivity to beef in children allergic to cow's milk. *J Allergy Clin Immunol* 99(3):293-300.

Westphal S, Kolarich D, Foetisch K, Lauer I, Altmann F, Conti A, Crespo JF, Rodriguez J, Enrique E, Vieths S, Scheurer S. (2003). Molecular characterization and allergenic activity of Lyc e 2 (®-fructofuranosidase), a glycosylated allergen of tomato. *Eur J Biochem* 270:1327-37.

WHO/IUS Allergen Nomenclature Subcommittee World Health Organization, Geneva, Switzerland. (1994). Allergen Nomenclature. *Bulletin of the World Health Organization* 72:797-806.

Wikipedia – The free encyclopedia. Wikimedia Foundation, Inc. http://en.wikipedia.org/wiki/Main_Page (Accessed 23 May 2006)

Williams PB, Barnes JH, Szeinbach SL, Sullivan TJ. (2000). Analytic precision and accuracy of commercial immunoassays for specific IgE: establishing a standard. *J Allergy Clin Immunol* 105:1221–1230.

Wilson DR, Irani AM, Walker SM, Jacobson MR, Mackay IS, Schwartz LB, Durham SR. (2001). Grass pollen immunotherapy inhibits seasonal increases in basophils and eosinophils in the nasal epithelium. *Clin Exp Allergy* 31:1705–13.

Wilson IBH. (2002) Glycosylation of proteins in plants and invertebrates. *Curr Opin Struct Biol* 12: 569–577.

Wright RJ, Cohen RT, Cohen S. (2005). The impact of stress on the development and expression of atopy. *Curr Opin Allergy Clin Immunol* 5(1):23-29.

Wüthrich B. (1996). Allergy to meat proteins in adults. *Allergologie* 19:130-4.

Zhu D, Kepley CL, Zhang M, Zhang K, Saxon A. (2002). A novel human immunoglobulin Fc gamma Fc epsilon bifunctional fusion protein inhibits Fc epsilon RI-mediated degranulation. *Nat Med* 8(5):518-21.

Zorzet A, Gustafsson M, Hammerling U. (2002). Prediction of food protein allergenicity: a bioinformatic learning systems approach. *In Silico Biol*. 2(4):525-34.

Appendix I. Allergens dotted onto the array.

Allergen type	Species
Pollen	*Acacia auriculiformis* (acacia)[c], *Agrostis alba* (bent grass)[a], *Agropyron repens* (quack grass)[a], *Alnus glutinosa* (black alder)[a], *Alopecurus prantesis* (foxtail, meadow)[a], *Ambrosia artemisiifolia* (annual ragweed)[d], *Amaranthus hybridus* (careless weed)[d], *Ambrosia trifida* (tall ragweed)[a], *Acer negundo* (box elder)[d], *Anthxanthum odoratum* (sweet vernal grass)[a], *Atriplex polycarpa* (allscale)[d], *Arecastrum romanzo ffianum* (queen palm)[d], *Artemisia vulgaris* (common mugwort)[a], *Avena sativa* (cultivated oats)[a], *Baccharis halimifolia* (eastern baccharis)[d], *Betula verrucosa* (white birch)[d], *Bromus mollis* (spear grass)[a], *Brassica* spp. (brassica pollen)[d], *Carpinus betulus* (hornbeam)[a], *Casuarina equisetifolia* (Australian pine)[d], *Calluna vulgaris* (heather)[f], C*henopodium album* (lamb's quarter)[a], *Chrysanthemum leucanthemum* (ox eye daisy)[a], *Corylus avellana* (hazel)[f], *Cryptomeria japonica (*Japanese cedar)[d], *Cupressus arizonica* (Arizona cypress)[f], C*upressus sempervirens* (Italian cypress)[f], *Cynodon dactylon* (Bermuda grass)[d], *Dahlia cultorum* (dahlia)[a], *Dactylis glomerata* (orchard grass)[f], *Elaeis guineensis* (oil palm)[c], *Eucalyptus globules* (bluegum)[d], *Fagus sylvatica* (European beech)[f], *Festuca prantesis* (meadow fescue)[f], *Fraxinus excelsior* (ash)[a], *Holcus lanatus* (velvet grass)[a], *Hordeum vulgare* (cultivated barley)[a], *Humulus lupulus* (hops)[a], *Juniperus asheisabinoides* (mountain cedar)[d], *Ligustrum vulgare* (common privet)[f], L*olium perenne* (perennial rye grass)[f], *Medicago sativa* (alfalfa)[a], *Olea europea* (olive)[f], *Parietaria judaica* (wall pellitory)[f], *Populus deltoides* (eastern cottonwood)[d], *Phragmites communis* (reed)[f], *Philadelphus coronarius* (syringa)[a], *Phlenum pratesense*(Timothy grass)[b], *Pinus radiata* (pine)[a], *Platanus acerfolia* (plane tree)[d], *Plantago lanceolata* (English plantain)[a], *Populus nigra* (black poplar)[a], *Poa pratensis* (Kentucky bluegrass)[f], *Podocarpus polystachyus* (sea teak)[d], *Pinus strobus* (eastern white pine)[d], *Populus trichocarpa* (black cottonwood)[d], *Quercus alba* (white oak)[d], *Quercus ilex* (live oak)[f], *Quercus robur* (red oak)[a], *Robinia pseudoacacia* (false acacia)[a], *Rumex acetosella* (sorrell)[a], *Salsola kali* (Saltwalt or Russian thistle)[a], *Sambucus nigra* (European elder)[a], *Salix viminalis* (willow)[a], *Secale cereale* (cultivated rye)[a], *Schinus molle* (pepper tree)[d], *Sorghum halepense* (Johnson grass)[d], *Solidago virgaurea* (golden rod)[a], *Syringa vulgaris* (lilac)[a], *Tamarix gallica* (salt cedar)[d], *Taraxacum officinale* (dandelion)[a], *Tilia cordata* (linden)[a], *Triticum aestivum/sativum* (cultivated wheat)[a], *Ulmus americana* (American elm)[d], *Ulmus minor* (English elm)[a], *Urtica dioica* (nettle)[f], *Zea mays* (corn)[d]
Fungi	*Alternaria alternata*[d], *Aspergillus flavus* [d], *Aspergillus fumigatus* [a], *Aspergillus niger* [c], *Aspergillus terreus* [d], *Botrytis cinerea* [f], *Candida albicans* [d], *Cladosporium cladosporioides* [c], *Cladosporium fulvum* [a], *Cladosporium herbarum* [c], *Corenyspora cassicola* [c], *Curvularia brachyspora*[c], *Curvularia fallax*[c], *Curvularia inequalis*[c], *Curvularia lunata*[c], *Curvularia pallescences*[c], *Curvularia spicifera*[a], *Drechslerea/Bipolaris sorokiana* [c], *Fusarium moniliforme* [a], *Fusarium solani* [d], *Malazessia furfur* [c], *Mucor mucedo* [a], *Penicillium brevicompactum* [f], *Penicillium chrysogenum* [d], *Penicillium expansum* [a], *Penicillium notatum* [c], *Penicillium roqueforti* [a], *Rhizopus nigricans* [a], *Saccharomyces cerevisae* [a], *Stemphylium botryosum* [f], *Trichoderma viride* [d], *Trichophyton mentagrophytes* [f], *Trichophyton rubrum* [d], *Ustilago tritici*[f]

Allergen sources: [a]ALK-Abelló S.A., [b]GREER Laboratories Incorporated, [c]local sources, [d]raw materials from GREER Laboratories Incorporated, [e]raw materials for Allergon AB and [g]Sigma.

Appendix I. Allergens dotted onto the array (cont.)

Allergen type	Species
Mites	*Acarus siro*[c], *Austroglycyphagus geniculatus* [c], *Blomia tropicalis* [c], *Dermatophagoides farinae* [c], *Dermatophagoides pteronyssinus* [c], *Glycophagus domesticus* [c], *Lepidoglyphus destructor* [c], *Suidasia medanensis* [c], *Tyrophagus putrescentiae* [c]
Epithelial tissue/dander	budgerigar (*Melopsittacus undulatus)*[a], cat (*Felis domesticus)*[e], cow (*Bos taurus*) [a], dog (*Canis familiaris*) [a], feather mix (chicken and duck) (*Pullus gallinaceus* and *Anas platyrhynchas*)[a], goose (*Anser anser*) [a], goat (*Capra hircus*)[b], guinea pig (*Cavia porcellus*) [a], hamster (*Cricetus cricetus*) [a], horse (*Equus caballus*) [a] and rabbit (*Oryctolagus cuniculus*) [a]
Food (animal origin)	banana prawn (*Penaeus merguiensis*)[c], beef (*Bos taurus*)[b], casein[b], chicken (*Pullus gallinaceus*) [a], cockles (*Anadara granosa*)[c], egg white [a], egg yolk [a], mackerel fish (*Scomberomorus sp.*)[b], milk, cow (*Bos taurus)* [a], milk, goat (*Capra hircus*) [a], mud crab (*Scylla olivacea*)[c], mussels (*Perna viridis*)[c], ovalbumin [a], ovomucoid [a], pork (*Sus scrofa)*[a], rabbit (*Oryctolagus cuniculus*)[a], salmon fish (*Oncorhynchus sp.*)[d], sea bream fish (*Nemipterus furcosus*)[c], selar fish (*Atule mate*)[c], sheep (*Ovis aries)*[b], squid (*Photololigo duvaucelii*)[c], swimming crab (*Portunus pelagicus*)[c], tiger prawn (*Penaeus monodon*)[c], tuna fish (*Thunnus sp.*)[d]
Food (plant origin)	apple (*Malus domestica)*[b], banana (*Musa* hybrids*)*[c], broccoli (*Brassica oleracea* var. *botrytis*)[b], cabbage (*Brassica oleracea* var. *capitata*)[b], cacao (*Theobroma cacao)*[a], carrot (*Daucus carota)*[c], chard (*Beta vulgaris* var. *cicla)*[a], corn flour (*Zea mays*)[a], garlic (*Allium sativum)*[a], gliadine (*Triticum aestivum*)[a], hazelnut (*Corylus avellana)*[a], kiwi (*Actinidia chinensis*)[c], orange (*Citrus sinensis)*[c], peach (*Prunus persica)*[a], peanut (*Arachys hypogaea)*[a], potato (*Solanum tuberosum*)[c], rice flour (*Oryza sativa*)[a], soya bean (*Glycine max*)[c], spinach (*Spinacia oleracea*[a], strawberry (*Fragaria vesca)*[c], sunflower seed (*Helianthus annuus)*[a], tofu (*Glycine max*)[c], walnut (*Junglans regia)*[a], wheat flour (*Triticum aestivum*)[a]
Insects	American cockroach (*Periplaneta americana*)[b], fire ant (*Solenopsis invicta*)[b], German cockroach (*Blatella germanica*)[b], mosquito (*Culicidae* sp.)[b], oriental cockroach (*Blatta orientalis*)[a]
Venoms	honeybee (*Apis mellifera*)[b], hornet (*Dolichovespula* spp.)[b], wasp (*Polistes* spp.)[b], yellowjacket (*Vespula* spp.)[b]
Others	horseradish peroxidase [g], latex[b], bromelain[g]
Controls	NIBSC IgE standard (positive control), bovine serum albumin (protein control)[g], extraction buffer (negative control)

Allergen sources: [a] ALK-Abelló S.A., [b]GREER Laboratories Incorporated, [c]local sources, [d]raw materials from GREER Laboratories Incorporated, [e]raw materials for Allergon AB and [g]Sigma

Appendix II. Detail lists of sequence homology matches for beef with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology.

Query	Subject (hits)	Subject Description	Functional category	Bit Score	E-value	Identities	Positives	6 amino acid homology
Bt#S15645050	sp|P02769|	Serum albumin precursor (Allergen Bos d 6) (BSA)	homeostasis	1216	0	592/607 (97%)	592/607 (97%)	highly homologous
Bt#S12073040	sp|P40918|	Heat shock 70 kDa protein (Allergen Cla h 4) (Cla h IV)	Chaparone/stress-related	860	0			highly homologous
Bt#S11933280	sp|P40918|	Heat shock 70 kDa protein (Allergen Cla h 4) (Cla h IV)	Chaparone/stress-related	845	0			highly homologous
Bt#S14887313	sp|P02789|	Ovotransferrin precursor (Conalbumin) (Allergen Gal d 3)	homeostasis	687	0	354/694 (51%)	470/694 (67%)	highly homologous
Bt#S15645044	sp|P02789|	Ovotransferrin precursor (Conalbumin) (Allergen Gal d 3)	homeostasis	684	0	362/696 (52%)	471/696 (67%)	highly homologous
Bt#S12072680	sp|Q9LEI9|	Enolase 2	metabolism	549	1.00E-155	289/443 (65%)	346/443 (78%)	highly homologous
Bt#S16822592	sp|P40292|	Heat shock protein 90 (Heat shock protein hsp1)	Chaparone/stress-related	529	1.00E-148	269/444 (60%)	335/444 (75%)	highly homologous
Bt#S12072842	sp|P40108|	Aldehyde dehydrogenase (ALDDH) (Allergen Cla h 3)	metabolism	492	1.00E-138	248/478 (51%)	329/478 (68%)	highly homologous
Bt#S12002881	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	436	1.00E-120			
Bt#S12072372	sp|P40292|	Heat shock protein 90 (Heat shock protein hsp1)	Chaparone/stress-related	365	2.00E-99	190/421 (45%)	280/421 (66%)	IKLYVRRVFITDD, SRETLQQ, LAKLLR, NMERIMKAQA, KKTFEI
Bt#S12377745	ref|NP_777186.1|	major allergen BDA20 [Bos taurus]	transport protein	353	2.00E-96	172/172 (100%)	172/172 (100%)	highly homologous
Bt#S12072572	sp|P02754|	Beta-lactoglobulin precursor (Beta-LG) (Allergen Bos d 5)	homeostasis	316	2.00E-85	158/161 (98%)	158/161 (98%)	highly homologous
Bt#S12373715	sp|Q9LEJ0|	Enolase 1	metabolism	300	2.00E-80	156/231 (67%)	181/231 (78%)	highly homologous
Bt#S12073013	ref|NP_776803.1|	lactalbumin, alpha [Bos taurus]	homeostasis	276	2.00E-73	131/142 (92%)	131/142 (92%)	highly homologous
Bt#S12072260	sp|Q95PM9|	Arginine kinase (AK) (Allergen Plo i 1)	metabolism	274	2.00E-72	152/330 (46%)	211/330 (63%)	DLFDPII, QQQLIDDHFLF, FLVWVNEEDHLR, CPSNLGT
Bt#S12072749	sp|P50635|	Apyrase precursor (ATP-diphosphatase)	metabolism	263	6.00E-69	172/540 (31%)	274/540 (50%)	LGNHEFD
Bt#S14905395	sp|Q92260|	Heat shock 70 kDa protein (Allergen Pen c 19)	Chaparone/stress-related	262	5.00E-69	125/183 (68%)	153/183 (83%)	highly homologous
Bt#S12072373	gb|AAD52672.1|	98kDa HDM allergen [Dermatophagoides farinae]	metabolism	258	2.00E-67	141/378 (37%)	209/378 (55%)	FDGLDLDWEYPG, WVGYDD
Bt#S12073064	sp|Q95PM9|	Arginine kinase (AK) (Allergen Plo i 1)	metabolism	255	2.00E-66	148/341 (43%)	202/341 (59%)	QQQLIDDHFLF, FLVWVNEEDH, CPTNLGT, VHIKLP
Bt#S15812499	gb|AAD52672.1|	98kDa HDM allergen [Dermatophagoides farinae]	metabolism	249	8.00E-65	137/366 (37%)	207/366 (56%)	FDGLDL
Bt#S11933200	gb|AAD52672.1|	98kDa HDM allergen [Dermatophagoides farinae]	metabolism	243	1.00E-62	165/489 (33%)	250/489 (51%)	FDGLDL, WVGYDD
Bt#S12072246	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	242	1.00E-62	131/389 (33%)	218/389 (56%)	SALAMV, VLVNAI, KPVQMM, ADHPFLF
Bt#S12072666	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	226	6.00E-58			
Bt#S17487029	sp|Q95182|	Major allergen Equ c 1 precursor	transport protein	220	2.00E-56	111/173 (64%)	130/173 (75%)	highly homologous
Bt#S14837289	sp|P40108|	Aldehyde dehydrogenase (ALDDH) (Allergen Cla h 3)	metabolism	213	2.00E-54	108/240 (45%)	158/240 (65%)	LELGGKSP
Bt#S14561797	sp|P46419|	Glutathione S-transferase (GST class-mu)	metabolism	213	6.00E-54	97/199 (48%)	138/199 (69%)	ILGYWDIRG, WLNEKF, LGLDFPNLPY
Bt#S12072465	ref|NP_777021.1|	S100 calcium-binding protein A7 (psoriasin 1) [Bos taurus]	signal transduction	208	1.00E-53	101/101 (100%)	101/101 (100%)	highly homologous
Bt#S11933345	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	198	8.00E-49	100/157 (63%)	113/157 (71%)	NFRALCTGEKGFG, FHRVIPDF, PGLLSMAN, SQFFIT , KHVVFG
Bt#S15341477	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	196	7.00E-55	94/123 (76%)	99/123 (80%)	highly homologous
Bt#S12072771	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	195	1.00E-48	101/165 (61%)	114/165 (69%)	DVVPKT, NFRALCTGEKG, GLLSMANAG, NTNGSQFFITTV, LDGKHVVFG
Bt#S12072795	gb|AAM64112.1|	gelsolin-like allergen Der f 16 [Dermatophagoides farinae]	structural protein	195	5.00E-48	119/363 (32%)	192/363 (52%)	ELVPVP, LLHVKG
Bt#S14866901	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	190	2.00E-47	95/163 (58%)	118/163 (72%)	DVVPKTA, FADENF, PGLLSMAN, WLDGKHVVFG
Bt#S11932912	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	190	4.00E-47	97/214 (45%)	138/214 (64%)	CGGCWAFSA, DYWIVKNSW
Bt#S11933051	sp|Q92450|	Superoxide dismutase [Mn]	metabolism	189	1.00E-46	96/203 (47%)	128/203 (63%)	KFNGGGHINHS, WEHAYYLQY, IWNVINW
Bt#S16972488	sp|P46419|	Glutathione S-transferase (GST class-mu)	metabolism	188	8.00E-47	84/199 (42%)	134/199 (67%)	LGYWDIRG, LDFPNLPY
Bt#S12072533	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	185	1.00E-45			
Bt#S12073122	sp|P49275|	Mite allergen Der f 3 precursor	metabolism	185	8.00E-46	99/247 (40%)	138/247 (55%)	HFCGGS, DSCQGDSGGPV, GIVSWGYGCA
Bt#S12373623	sp|Q25456|	Tropomyosin (Allergen Met e 1) (Met e I)	structural protein	178	1.00E-43	94/191 (49%)	128/191 (67%)	ADRKYDEVARKL, ELEEEL, NNLKSLE, RAEFAERSV
Bt#S15971742	sp|Q9U6V9|	Hyaluronon glucosaminidase precursor	metabolism	177	4.00E-43	110/324 (33%)	169/324 (52%)	FMEETLKL, LFPSVY, PQLGNL
Bt#S15645043	sp|P00698|	Lysozyme C precursor (1,4-beta-N-acetylmuramidase C)	metabolism	174	3.00E-42	79/148 (53%)	109/148 (73%)	STDYGI, WWCNDG
Bt#S15460514	gb|AAM64112.1|	gelsolin-like allergen Der f 16 [Dermatophagoides farinae]	structural protein	173	6.00E-42	118/361 (32%)	179/361 (49%)	PVQHRE
Bt#S12072768	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	172	5.00E-42	87/158 (55%)	106/158 (67%)	TGEKGFGY, SMANAG, TNGSQFFITTV, WLDGKHV
Bt#S12073038	sp|Q08169|	Hyaluronon glucosaminidase precursor	metabolism	172	1.00E-41	113/338 (33%)	170/338 (50%)	nil
Bt#S12072261	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	169	1.00E-40			
Bt#S14878741	sp|Q92260|	Heat shock 70 kDa protein (Allergen Pen c 19)	Chaparone/stress-related	165	4.00E-46			
Bt#S17486502	ref|NP_777186.1|	major allergen BDA20 [Bos taurus]	transport protein	163	2.00E-39	81/156 (51%)	108/156 (69%)	EGGPLR
Bt#S12072961	ref|NP_240013.1|	alkyl hydroperoxide reductase	metabolism	159	1.00E-37	83/172 (48%)	110/172 (63%)	DFTFVCPTE, GLALRG, GEVCPA
Bt#S14769516	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	157	7.00E-37			
Bt#S12072754	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	157	1.00E-36			

Appendix II. Detail lists of sequence homology matches for beef with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology (cont.).

Query	Subject (hits)	Subject Description	Functional category	Bit Score	E-value	Identities	Positives	6 amino acid homology
Bt#S12072270	ref|NP_842456.1|	Alkyl hydroperoxide reductase	metabolism	156	5.00E-37	81/175 (46%)	111/175 (63%)	DFTFVCPTE, GEVCPA
Bt#S12072400	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	156	2.00E-36	112/379 (29%)	188/379 (49%)	nil
Bt#S12373644	sp|Q95PM9|	Arginine kinase (AK) (Allergen Plo i 1)	metabolism	154	1.00E-36	77/167 (46%)	108/167 (64%)	CPTNLGT, VYDISN
Bt#S12069906	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	148	6.00E-35			
Bt#S12069907	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	146	3.00E-34			
Bt#S12373656	sp|O44119|	Tropomyosin (Allergen Hom a 1)	structural protein	143	3.00E-33	78/168 (46%)	106/168 (63%)	AADESER, EVARKL, NNLKSLE, RAEFAERSV
Bt#S11933075	gb|AAA99805.1|	Der f 3 mite allergen	metabolism	141	2.00E-32			
Bt#S12072960	ref|NP_842456.1|	Alkyl hydroperoxide reductase	metabolism	141	2.00E-32	75/180 (41%)	106/180 (58%)	DFTFVCPTE, GEVCPA
Bt#S11965762	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	138	6.00E-32	72/212 (33%)	124/212 (58%)	KPVQMM
Bt#S12072401	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	137	3.00E-31			
Bt#S12072361	sp|P39675|	Mite allergen Der p 3 precursor (Der p III)	metabolism	136	6.00E-31			
Bt#S11972834	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	135	6.00E-31			
Bt#S12378832	sp|P39675|	Mite allergen Der p 3 precursor (Der p III)	metabolism	135	2.00E-30			
Bt#S12377802	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	134	9.00E-31	64/174 (36%)	104/174 (59%)	KPVQMM
Bt#S12072762	sp|P39675|	Mite allergen Der p 3 precursor (Der p III)	metabolism	132	3.00E-29			
Bt#S12065996	sp|P49064|	Serum albumin precursor (Allergen Fel d 2)	homeostasis	132	7.00E-30	58/125 (46%)	82/125 (65%)	NRRPCFS, VDETYVP, TEEQLKTV
Bt#S12072765	gb|AAA99805.1|	Der f 3 mite allergen	metabolism	129	2.00E-28			
Bt#S14846682	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	126	4.00E-28	65/186 (34%)	106/186 (56%)	ADHPFLF
Bt#S12378774	sp|P40108|	Aldehyde dehydrogenase (ALDDH) (Allergen Cla h 3)	metabolism	124	4.00E-28	57/126 (45%)	87/126 (69%)	GQCCCAGS, ESIYDKF
Bt#S15184002	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	115	3.00E-24			
Bt#S12375436	sp|P35747|	Serum albumin precursor (Allergen Equ c 3)	homeostasis	115	9.00E-25	64/180 (35%)	90/180 (50%)	FTFHAD
Bt#S14870192	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	114	1.00E-24			
Bt#S12378278	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	112	9.00E-25	61/145 (42%)	89/145 (61%)	nil
Bt#S12372716	sp|P39674|	Allergen MAG29	Chaparone/stress-related	112	4.00E-24	49/74 (66%)	63/74 (85%)	LAEKDEFEH, VCNPII
Bt#S11933030	sp|Q06478|	Phospholipase A1 1 precursor (Allergen Dol m 1.01)	metabolism	112	4.00E-23	77/276 (27%)	127/276 (46%)	SLGAHA, GLDPAGP
Bt#S12072353	gb|AAK39511.1|	paramyosin-like allergen [Dermatophagoides farinae]	structural protein	110	2.00E-22	153/731 (20%)	320/731 (43%)	nil
Bt#S12072399	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	109	2.00E-22	96/388 (24%)	173/388 (44%)	nil
Bt#S11963276	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	104	1.00E-21			
Bt#S12044428	sp|P39675|	Mite allergen Der p 3 precursor (Der p III)	metabolism	103	2.00E-21			
Bt#S12072611	sp|Q9UW02|	Thioredoxin (Allergen Cop c 2)	Chaparone/stress-related	102	4.00E-21	47/105 (44%)	69/105 (65%)	ATWCGPC
Bt#S12073098	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13) (Bt6)	metabolism	101	1.00E-20	48/125 (38%)	74/125 (59%)	STFKNTEI
Bt#S12073097	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	99	5.00E-20	52/127 (40%)	80/127 (62%)	STFKNTEI
Bt#S14887281	sp|P46419|	Glutathione S-transferase (GST class-mu)	metabolism	96	9.00E-19	62/207 (29%)	99/207 (47%)	nil
Bt#S14885468	sp|P49275|	Mite allergen Der f 3 precursor	metabolism	94	2.00E-18			
Bt#S12072302	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	93	1.00E-17			
Bt#S12375811	sp|P40108|	Aldehyde dehydrogenase (ALDDH) (Allergen Cla h 3)	metabolism	92	9.00E-18			
Bt#S12072282	sp|O43099|	Putative peroxiredoxin pmp20 (Thioredoxin reductase)	structural protein	91	3.00E-17	56/176 (31%)	85/176 (48%)	PGAFTP
Bt#S12072858	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	90	5.00E-17			
Bt#S17479898	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	89	7.00E-17			
Bt#S12072243	sp|P42037|	60S acidic ribosomal protein P2 (Minor allergen Alt a 6)	protein synthesis	89	1.00E-17	48/115 (41%)	62/115 (53%)	SVGIEAD, DMGFGLFD
Bt#S17485445	sp|Q92260|	Heat shock 70 kDa protein (Allergen Pen c 19)	Chaparone/stress-related	88	1.00E-16	49/159 (30%)	80/159 (50%)	nil
Bt#S12373902	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	87	4.00E-16			
Bt#S14903074	gb|AAD25927.1|	major allergenic protein Mal f4 [Malassezia furfur]	unknown	86	3.00E-22	54/128 (42%)	79/128 (61%)	EVVKAK, GAGSATLSMAYA
Bt#S14876172	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	86	5.00E-16	43/136 (31%)	76/136 (55%)	nil
Bt#S14885930	sp|Q06478|	Phospholipase A1 1 precursor (Allergen Dol m 1.01)	metabolism	86	5.00E-16	51/149 (34%)	80/149 (53%)	SLGAHISGFAG
Bt#S17486445	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	86	5.00E-16	43/119 (36%)	70/119 (58%)	LGEEFEE
Bt#S12072759	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	84	5.00E-15			
Bt#S12378546	sp|P40108|	Aldehyde dehydrogenase (ALDDH) (Allergen Cla h 3)	metabolism	84	3.00E-15			
Bt#S17484852	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	84	2.00E-15	48/137 (35%)	76/137 (55%)	GEEFEE

Appendix II. Detail lists of sequence homology matches for beef with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology (cont.).

Query	Subject (hits)	Subject Description	Functional category	Bit Score	E-value	Identities	Positives	6 amino acid homology
Bt#S12059808	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	82	4.00E-15			
Bt#S12072859	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	82	9.00E-15			nil
Bt#S15645006	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	82	2.00E-14	48/137 (35%)	72/137 (52%)	
Bt#S12065896	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	80	6.00E-14			AIKKKMQ, LNRRIQL
Bt#S12041279	sp|O18416|	Tropomyosin (Allergen Der p 10)	structural protein	80	4.00E-14	58/209 (27%)	85/209 (40%)	TGQIKTGAPCRSERLAKYNQL, RIEEELG
Bt#S14838136	sp|Q9LEJ0|	Enolase 1	metabolism	80	2.00E-14	38/45 (84%)	40/45 (88%)	nil
Bt#S11932968	sp|P01005|	Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	homeostasis	79	1.00E-13	45/131 (34%)	77/131 (58%)	
Bt#S12072330	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	78	4.00E-13			GEEFEE
Bt#S14874248	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	78	1.00E-13	45/131 (34%)	77/131 (58%)	nil
Bt#S12072807	gb|AAM64112.1|	gelsolin-like allergen Der f 16 [Dermatophagoides farinae]	structural protein	77	5.00E-12	91/391 (23%)	157/391 (40%)	
Bt#S12072308	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	77	1.00E-12			STFKNTEI
Bt#S11937254	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13) (Bt6)	metabolism	77	2.00E-13	37/89 (41%)	56/89 (62%)	QDCFNE
Bt#S17483233	ref|NP_803230.1|	allergen dI chain C2A [Mus musculus]	unknown	76	3.00E-13	38/93 (40%)	51/93 (54%)	
Bt#S12072667	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	76	3.00E-12			VGHYTQ
Bt#S12378420	sp|P10736|	Venom allergen 5.01 precursor (Antigen 5 form 2)	metabolism	76	9.00E-14	47/132 (35%)	68/132 (51%)	nil
Bt#S14561803	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13) (Bt6)	metabolism	76	1.00E-13	44/130 (33%)	69/130 (53%)	
Bt#S14902851	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	75	9.00E-13			NGTGGKSIY
Bt#S14878999	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	74	7.00E-17	35/59 (59%)	39/59 (66%)	
Bt#S12072476	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	74	2.00E-12			DHPFLF
Bt#S14847131	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	72	8.00E-12	39/104 (37%)	52/104 (50%)	
Bt#S12373387	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	72	8.00E-12			
Bt#S12072818	sp|Q9UW02|	Thioredoxin (Allergen Cop c 2)	Chaparone/stress-related	70	8.00E-11			
Bt#S12072861	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	69	1.00E-10			nil
Bt#S17486673	ref|NP_777186.1|	major allergen BDA20 [Bos taurus]	transport protein	69	6.00E-11	42/155 (27%)	77/155 (49%)	KPKATE, EGPKLVVSTQT
Bt#S11906978	sp|P02769|	Serum albumin precursor (Allergen Bos d 6) (BSA)	homeostasis	69	2.00E-11	41/75 (54%)	44/75 (58%)	
Bt#S14886676	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	68	1.00E-10			EGGPLRNYYR, LEKYQQLNSERGVPNENIENLIKTDNCPP
Bt#S17485312	pir||B59225	allergen Bos d 2.0103 [imported] - bovine	transport protein	67	1.00E-10	51/126 (40%)	63/126 (50%)	LGEEFEE
Bt#S12073100	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13) (Bt6)	metabolism	67	2.00E-10	38/127 (29%)	62/127 (48%)	nil
Bt#S12376903	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	66	5.00E-10	32/90 (35%)	51/90 (56%)	ADHPFLF
Bt#S11931473	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	66	3.00E-10	39/149 (26%)	81/149 (54%)	VGHYTQ
Bt#S12378419	ref|NP_509802.1|	T05A10.5 [Caenorhabditis elegans]	metabolism	65	3.00E-10	34/75 (45%)	40/75 (53%)	LFAKAL, SDDDMGFGLFD
Bt#S14835144	sp|P50344|	60S acidic ribosomal protein P1 (Allergen Cla h 12)	protein synthesis	65	7.00E-10	42/113 (37%)	51/113 (45%)	nil
Bt#S12072562	prf||2118249B	allergen Lep d 1.02	unknown	64	4.00E-09	36/110 (32%)	56/110 (50%)	
Bt#S16058522	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	62	1.00E-07			
Bt#S14879707	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	62	8.00E-09			nil
Bt#S11906148	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	61	8.00E-09	41/166 (24%)	74/166 (44%)	nil
Bt#S14870498	ref|NP_777021.1|	S100 calcium-binding protein A7 (psoriasin 1) [Bos taurus]	signal transduction	60	2.00E-08	31/91 (34%)	53/91 (58%)	SVGIEAD, KLASVP
Bt#S11894291	sp|P42037|	60S acidic ribosomal protein P2 (Minor allergen Alt a 6)	protein synthesis	60	2.00E-09	29/56 (51%)	39/56 (69%)	nil
Bt#S12072915	sp|Q9NFZ4|	Tropomyosin (Allergen Lep d 10)	structural protein	59	6.00E-07	49/233 (21%)	107/233 (45%)	nil
Bt#S14882439	sp|P49064|	Serum albumin precursor (Allergen Fel d 2)	homeostasis	58	2.00E-07	42/172 (24%)	72/172 (41%)	
Bt#S14881268	sp|Q9UW02|	Thioredoxin (Allergen Cop c 2)	Chaparone/stress-related	58	1.00E-07			
Bt#S11978542	sp|P42041|	Aldehyde dehydrogenase (ALDDH) (Allergen Alt a 10)	metabolism	57	9.00E-08			
Bt#S12072722	gb|AAD25927.1|	major allergenic protein Mal f4 [Malassezia furfur]	unknown	56	2.00E-06			nil
Bt#S12072244	ref|NP_671747.1|	alpha-2u globulin PGCL1 [Rattus norvegicus]	homeostasis	56	6.00E-07	38/156 (24%)	74/156 (47%)	nil
Bt#S14879787	gb|AAM64112.1|	gelsolin-like allergen Der f 16 [Dermatophagoides farinae]	structural protein	55	1.00E-06	48/174 (27%)	81/174 (46%)	
Bt#S12062724	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	55	6.00E-07			WLDGKHVVFG
Bt#S12371599	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	54	3.00E-06	26/50 (52%)	30/50 (60%)	FHKYSG
Bt#S12007486	ref|NP_777021.1|	S100 calcium-binding protein A7 (psoriasin 1) [Bos taurus]	signal transduction	54	2.00E-06	32/100 (32%)	52/100 (52%)	
Bt#S12375345	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	54	2.00E-14			nil

Appendix II. Detail lists of sequence homology matches for beef with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology (cont.).

Query	Subject (hits)	Subject Description	Functional category	Bit Score	E-value	Identities	Positives	6 amino acid homology
Bt#S14844653	sp|P40108|	Aldehyde dehydrogenase (ALDDH) (Allergen Cla h 3)	metabolism	54	7.00E-07	25/80 (31%)	45/80 (56%)	nil
Bt#S13584272	sp|P01005|	Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	homeostasis	53	9.00E-07	23/41 (56%)	26/41 (63%)	
Bt#S14872311	sp|P40292|	Heat shock protein 90 (Heat shock protein hsp1)	Chaparone/stress-related	53	4.00E-06	43/164 (26%)	79/164 (48%)	
Bt#S17484640	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	52	1.00E-05			
Bt#S12073034	gb|AAK39511.1|	paramyosin-like allergen [Dermatophagoides farinae]	structural protein	51	1.00E-04			nil
Bt#S12072631	sp|O18873|	Major allergen Can f 1 precursor (Allergen Dog 1)	transport protein	51	4.00E-05			TSETPK
Bt#S15828154	sp|Q9UB83|	Tropomyosin (Major allergen Per a 7)	structural protein	51	5.00E-05	39/158 (24%)	71/158 (44%)	
Bt#S11989097	gb|AAD52672.1|	98kDa HDM allergen [Dermatophagoides farinae]	metabolism	50	3.00E-04	31/84 (36%)	34/84 (40%)	
Bt#S12072484	gb|AAK39511.1|	paramyosin-like allergen [Dermatophagoides farinae]	structural protein	50	1.00E-04			nil
Bt#S12072508	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	50	5.00E-05			GCGGNANRF
Bt#S12072417	ref|NP_777021.1|	S100 calcium-binding protein A7 (psoriasin 1) [Bos taurus]	signal transduction	50	2.00E-05	29/81 (35%)	46/81 (56%)	nil
Bt#S12070644	dbj|BAC77154.1|	21k allergen [Anisakis simplex]	unknown	49	4.00E-05	27/64 (42%)	35/64 (54%)	
Bt#S14879262	gb|AAB34785.1|	68 kDa allergen [Penicillium chrysogenum]	metabolism	49	5.00E-05	35/100 (35%)	46/100 (46%)	
Bt#S12378066	gb|AAK39511.1|	paramyosin-like allergen [Dermatophagoides farinae]	structural protein	49	6.00E-05			nil
Bt#S11882487	gb|AAK39511.1|	paramyosin-like allergen [Dermatophagoides farinae]	structural protein	49	3.00E-05			
Bt#S14877298	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	49	1.00E-04	53/225 (23%)	102/225 (45%)	PVCGTDGVTY
Bt#S12072905	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	49	8.00E-05			
Bt#S13584198	sp|P01005|	Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	homeostasis	49	2.00E-05	18/41 (43%)	27/41 (65%)	
Bt#S12072720	emb|CAA09883.1|	allergen [Malassezia sympodialis]	signal transduction	48	6.00E-04			FHKYSG
Bt#S14884075	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	48	1.00E-04			
Bt#S12072466	ref|NP_777021.1|	S100 calcium-binding protein A7 (psoriasin 1) [Bos taurus]	signal transduction	48	6.00E-05	26/81 (32%)	43/81 (53%)	
Bt#S12043764	sp|P43187|	Calcium-binding allergen Bet v 3 (Bet v III)	signal transduction	48	9.00E-05			nil
Bt#S14905591	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	47	2.00E-04			LFAKAL
Bt#S12072412	ref|NP_777021.1|	S100 calcium-binding protein A7 (psoriasin 1) [Bos taurus]	signal transduction	47	5.00E-05	27/91 (29%)	47/91 (51%)	
Bt#S14882897	sp|P49148|	60S acidic ribosomal protein P1	protein synthesis	47	2.00E-04	26/61 (42%)	36/61 (59%)	LFDPIIED
Bt#S12376787	sp|P78983|	Heat shock 70 kDa protein (Allergen Alt a 3)	Chaparone/stress-related	47	2.00E-08			TWCGPC
Bt#S14855781	sp|Q95PM9|	Arginine kinase (AK) (Allergen Plo i 1)	metabolism	47	3.00E-07	20/31 (64%)	24/31 (77%)	
Bt#S14846535	sp|Q9UW02|	Thioredoxin (Allergen Cop c 2)	Chaparone/stress-related	47	2.00E-04	21/50 (42%)	29/50 (58%)	
Bt#S17482076	dbj|BAC77154.1|	21k allergen [Anisakis simplex]	unknown	46	5.00E-04			
Bt#S12375538	gb|AAK39511.1|	paramyosin-like allergen [Dermatophagoides farinae]	structural protein	46	8.00E-04			FVDWIE
Bt#S14902829	sp|Q9M7R0|	Calcium-binding allergen Ole e 8 (PCA18/PCA23).	signal transduction	46	7.00E-04			nil
Bt#S14886698	sp|P49275|	Mite allergen Der f 3 precursor	metabolism	45	2.00E-04	18/25 (72%)	18/25 (72%)	
Bt#S14874727	sp|P40108|	Aldehyde dehydrogenase (ALDDH) (Allergen Cla h 3)	metabolism	40	7.00E-05	16/32 (50%)	24/32 (75%)	

Appendix III. Detail lists of sequence homology matches for pork with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology.

Query	Subject (hits)	Subject Description	Functional category	Bit Score	E-value	Identities	Positives	6 amino acid homology
Ssc#S6091528	sp|P02769|	Serum albumin precursor (Allergen Bos d 6) (BSA)	homeostasis	1032	0	483/604 (79%)	543/604 (89%)	highly homologous
Ssc#S6091475	sp|P40918|	Heat shock 70 kDa protein (Allergen Cla h 4) (Cla h IV)	Chaparone/stress-related	815	0	406/640 (63%)	498/640 (77%)	highly homologous
Ssc#S6091477	sp|P40918|	Heat shock 70 kDa protein (Allergen Cla h 4) (Cla h IV)	Chaparone/stress-related	745	0	377/618 (61%)	475/618 (76%)	highly homologous
Ssc#S6091427	sp|P02789|	Ovotransferrin precursor (Conalbumin) (Allergen Gal d 3)	homeostasis	707	0	362/700 (51%)	484/700 (69%)	highly homologous
Ssc#S6091238	sp|P02789|	Ovotransferrin precursor (Conalbumin) (Allergen Gal d 3)	homeostasis	701	0	354/700 (50%)	479/700 (68%)	highly homologous
Ssc#S6091718	sp|P02789|	Ovotransferrin precursor (Conalbumin) (Allergen Gal d 3)	homeostasis	657	0	349/715 (48%)	471/715 (65%)	highly homologous
Ssc#S6090900	sp|P40292|	Heat shock protein 90 (Heat shock protein hsp1)	Chaparone/stress-related	529	1.00E-149	269/444 (60%)	335/444 (75%)	highly homologous
Ssc#S6089974	sp|P19121|	Serum albumin precursor (Alpha-livetin) (Allergen Gal d 5)	homeostasis	493	1.00E-138	239/614 (38%)	367/614 (59%)	YEIARRHP, YEYSRRHP, NRRPCFS
Ssc#S6090444	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	445	1.00E-123	259/772 (33%)	422/772 (54%)	nil
Ssc#S6090711	sp|P40292|	Heat shock protein 90 (Heat shock protein hsp1)	Chaparone/stress-related	365	3.00E-99	190/421 (45%)	279/421 (66%)	IKLYVRRVFITDD, KGVVDS, LAKLLR, NMERIMKAQA, KKTFEI
Ssc#S6077154	sp|Q92260|	Heat shock 70 kDa protein (Allergen Pen c 19)	Chaparone/stress-related	346	6.00E-96			
Ssc#S14768430	sp|P40918|	Heat shock 70 kDa protein (Allergen Cla h 4) (Cla h IV)	Chaparone/stress-related	319	3.00E-86			
Ssc#S6091740	gb|AAD25927.1|	major allergenic protein Mal f4 [Malassezia furfur]	unknown	270	8.00E-71	149/294 (50%)	198/294 (67%)	highly homologous
Ssc#S16767837	sp|Q9LEI9|	Enolase 2	metabolism	256	2.00E-67	126/179 (70%)	149/179 (83%)	highly homologous
Ssc#S17511701	sp|Q9LEI9|	Enolase 2	metabolism	248	5.00E-65	123/178 (69%)	147/178 (82%)	highly homologous
Ssc#S6089598	sp|Q9NAS5|	Tropomyosin (Allergen Ani s 3).	structural protein	246	5.00E-64	134/280 (47%)	177/280 (63%)	highly homologous
Ssc#S6091381	gb|AAD52672.1|	98kDa HDM allergen [Dermatophagoides farinae]	metabolism	244	3.00E-63	137/378 (36%)	214/378 (56%)	WVGYDD
Ssc#S6090433	gb|AAD52672.1|	98kDa HDM allergen [Dermatophagoides farinae]	metabolism	240	6.00E-62	167/489 (34%)	246/489 (50%)	FDGLDL, WVGYDD
Ssc#S6091242	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	219	9.00E-56	126/345 (36%)	190/345 (55%)	CWAFSA
Ssc#S6091086	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	217	4.00E-55	131/340 (38%)	182/340 (53%)	YWIVKNSW
Ssc#S17525633	sp|P02754|	Beta-lactoglobulin precursor (Beta-LG) (Allergen Bos d 5)	homeostasis	215	5.00E-55	108/167 (64%)	133/167 (79%)	highly homologous
Ssc#S6091512	sp|P02754|	Beta-lactoglobulin precursor (Beta-LG) (Allergen Bos d 5)	homeostasis	214	1.00E-54	107/167 (64%)	132/167 (79%)	highly homologous
Ssc#S6090496	sp|Q95182|	Major allergen Equ c 1 precursor	transport protein	212	8.00E-54	106/172 (61%)	131/172 (76%)	IEENGSM, NIIDLTKI
Ssc#S6091491	gb|AAM64112.1|	gelsolin-like allergen Der f 16 [Dermatophagoides farinae]	structural protein	208	4.00E-52	121/366 (33%)	199/366 (54%)	nil
Ssc#S6091341	sp|O18873|	Major allergen Can f 1 precursor (Allergen Dog 1)	transport protein	207	1.00E-52	97/153 (63%)	130/153 (84%)	highly homologous
Ssc#S6090478	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	206	7.00E-52	115/307 (37%)	172/307 (56%)	nil
Ssc#S6091747	ref|NP_776803.1|	lactalbumin, alpha [Bos taurus]	homeostasis	204	1.00E-51	95/142 (66%)	111/142 (78%)	highly homologous
Ssc#S17525050	sp|Q9LEI9|	Enolase 2	metabolism	199	2.00E-50	98/135 (72%)	113/135 (83%)	highly homologous
Ssc#S17510611	sp|Q9LEI9|	Enolase 2	metabolism	196	4.00E-49	95/121 (78%)	108/121 (89%)	highly homologous
Ssc#S6072833	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	196	2.00E-49	99/221 (44%)	137/221 (61%)	CWAFSA, LISLSEQ, CGIATM
Ssc#S17513259	sp|P40108|	Aldehyde dehydrogenase (ALDDH) (Allergen Cla h 3)	metabolism	191	8.00E-48	99/217 (45%)	133/217 (61%)	LELGGKSP
Ssc#S17518674	sp|Q92450|	Superoxide dismutase [Mn]	metabolism	189	4.00E-47	95/203 (46%)	125/203 (61%)	KFNGGGHINHS, WEHAYYLQY, IWNVINW
Ssc#S6081972	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	187	2.00E-46	97/232 (41%)	146/232 (62%)	VLVNAI
Ssc#S6090009	sp|Q9U6V9|	Hyalurononglucosaminidase precursor	metabolism	184	7.00E-45	109/331 (32%)	167/331 (50%)	NGGIPQ , FMQETLKL
Ssc#S6091332	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	180	5.00E-44	115/380 (30%)	202/380 (53%)	nil
Ssc#S17515670	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	175	7.00E-43	88/158 (55%)	106/158 (67%)	TGEKGFGY, SMANAG, TNGSQFFITTV, WLDGKHVVFG
Ssc#S17510471	sp|P46419|	Glutathione S-transferase (GST class-mu)	metabolism	174	2.00E-42	80/191 (41%)	125/191 (65%)	LDFPNLPY, FPNLKA
Ssc#S6074566	sp|P40292|	Heat shock protein 90 (Heat shock protein hsp1)	Chaparone/stress-related	165	8.00E-40	85/167 (50%)	116/167 (69%)	ANMERIMKA
Ssc#S6090480	sp|P49275|	Mite allergen Der f 3 precursor (Der f III)	metabolism	164	1.00E-57	76/90 (84%)	81/90 (90%)	highly homologous
Ssc#S17524919	sp|P02769|	Serum albumin precursor (Allergen Bos d 6) (BSA)	homeostasis	164	8.00E-39	96/242 (39%)	139/242 (57%)	LTAAHC
Ssc#S17518399	sp|Q92260|	Heat shock 70 kDa protein (Allergen Pen c 19)	Chaparone/stress-related	159	4.00E-37	105/382 (27%)	184/382 (48%)	nil
Ssc#S6091317	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	159	3.00E-38			
Ssc#S17511369	sp|Q95PM9|	Arginine kinase (AK) (Allergen Plo i 1)	metabolism	154	1.00E-42	77/167 (46%)	108/167 (64%)	CPSNLGT, VYDISN
Ssc#S17526670	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	151	7.00E-36	75/125 (60%)	89/125 (71%)	TGEKGFGY, SMANAG, TNGSQFFITTV, WLDGKHVVFG
Ssc#S17519014	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	149	4.00E-35			
Ssc#S6090124	sp|O97370|	Mite allergen Eur m 3 precursor	metabolism	148	2.00E-34	92/241 (38%)	128/241 (53%)	LTAAHCV, DSCQGDSGGP, GIVSWG
Ssc#S6090257	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	147	5.00E-34	108/379 (28%)	187/379 (49%)	nil
Ssc#S6090888	sp|P49275|	Mite allergen Der f 3 precursor (Der f III)	metabolism	143	6.00E-33			
Ssc#S6090178	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	143	1.00E-32	90/244 (36%)	129/244 (52%)	CQGDSGGP, GIVSWG

Appendix III. Detail lists of sequence homology matches for pork with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology (cont.).

Query	Subject (hits)	Subject Description	Functional category	Bit Score	E-value	Identities	Positives	6 amino acid homology
Ssc#S6091834	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	141	3.00E-32	116/392 (29%)	197/392 (50%)	nil
Ssc#S6090504	sp|P40292|	Heat shock protein 90 (Heat shock protein hsp1)	Chaparone/stress-related	140	7.00E-33	66/94 (70%)	85/94 (90%)	highly homologous
Ssc#S17512167	sp|P39675|	Mite allergen Der f 3 precursor (Der f III)	metabolism	139	3.00E-32	78/162 (48%)	99/162 (61%)	GGKDSCQGDSGGPVV, GIVSWGYGCA
Ssc#S6090943	sp|O44119|	Tropomyosin (Allergen Hom a 1)	structural protein	138	1.00E-31	79/213 (37%)	111/213 (52%)	VAALNRRIQ, AEEADRKY, EVARKL
Ssc#S17508597	sp|P40108|	Aldehyde dehydrogenase (ALDDH) (Allergen Cla h 3)	metabolism	133	3.00E-30	63/144 (43%)	100/144 (69%)	PFGGYK
Ssc#S17511216	sp|P49822|	Serum albumin precursor (Allergen Can f 3)	homeostasis	132	4.00E-29			
Ssc#S17524777	sp|P02789|	Ovotransferrin precursor (Conalbumin) (Allergen Gal d 3)	homeostasis	132	6.00E-30	70/168 (41%)	94/168 (55%)	EEGPKL
Ssc#S6091419	gb|AAA99805.1|	Der f 3 mite allergen	metabolism	132	3.00E-30	66/151 (43%)	100/151 (66%)	DFELLC
Ssc#S6081806	gb|AAO15713.1|	allergen Pen m 2 [Penaeus monodon]	structural protein	131	7.00E-38	65/110 (59%)	80/110 (72%)	LLSMAN, TNGSQFFITT
Ssc#S14765627	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	131	1.00E-29	70/157 (44%)	94/157 (59%)	HNDNKTFLVW, NEEDHLR, VHIKLP
Ssc#S16769360	sp|P49275|	Mite allergen Der f 3 precursor (Der f III)	metabolism	125	4.00E-28	68/192 (35%)	105/192 (54%)	HFCGGS, GDSGGP
Ssc#S6090453	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	125	3.00E-27	97/383 (25%)	186/383 (48%)	nil
Ssc#S6072984	sp|P42041|	Aldehyde dehydrogenase (ALDDH) (Allergen Alt a 10)	metabolism	118	7.00E-26	55/122 (45%)	82/122 (67%)	EEIFGPV, PFGGYK
Ssc#S6072240	sp|O44119|	Tropomyosin (Allergen Hom a 1)	structural protein	116	1.00E-27	58/135 (42%)	85/135 (62%)	AEEADRKY, EVARKL, RAEFAERSV
Ssc#S6091320	sp|Q06478|	Phospholipase A1 1 precursor (Allergen Dol m 1.01)	metabolism	110	1.00E-22	77/276 (27%)	126/276 (45%)	GLDPAGP, GYSLGAHA
Ssc#S6055440	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	107	2.00E-22	55/90 (61%)	61/90 (67%)	LLSMAN, SQFFIT, KHVVFG
Ssc#S17507675	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	105	4.00E-22	47/72 (65%)	58/72 (80%)	PGLLSMAN, GPNTNGSQFF, WLDGKHVVFGEV
Ssc#S17513373	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	103	2.00E-21	58/127 (45%)	78/127 (61%)	STFKNTEI
Ssc#S17526046	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	103	2.00E-21	63/188 (33%)	100/188 (53%)	nil
Ssc#S16764840	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	103	2.00E-21	58/127 (45%)	78/127 (61%)	STFKNTEI
Ssc#S14890189	sp|Q9UW02|	Thioredoxin (Allergen Cop c 2)	Chaparone/stress-related	100	2.00E-20	47/125 (37%)	75/125 (60%)	STFKNTEI
Ssc#S17526117	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13) (Bt6)	metabolism	100	3.00E-20	47/125 (37%)	75/125 (60%)	STFKNTEI
Ssc#S6090271	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13) (Bt6)	metabolism	100	2.00E-20	63/200 (31%)	105/200 (52%)	AKLLRY
Ssc#S17515904	sp|P40292|	Heat shock protein 90 (Heat shock protein hsp1)	Chaparone/stress-related	100	3.00E-20	44/95 (46%)	64/95 (67%)	ATWCGPC
Ssc#S6090227	ref|NP_777186.1|	major allergen BDA20 [Bos taurus]	transport protein	99	3.00E-20	54/157 (34%)	84/157 (53%)	nil
Ssc#S6090217	sp|Q95PM9|	Arginine kinase (AK) (Allergen Plo i 1)	metabolism	96	7.00E-19			
Ssc#S6089956	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	96	6.00E-19	53/122 (43%)	77/122 (63%)	CPSNLGT
Ssc#S16514177	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	96	4.00E-19	50/153 (32%)	77/153 (50%)	nil
Ssc#S16515301	ref|NP_671747.1|	alpha-2u globulin PGCL1 [Rattus norvegicus]	homeostasis	96	5.00E-19	63/218 (28%)	110/218 (50%)	nil
Ssc#S17510722	sp|P40108|	Aldehyde dehydrogenase (ALDDH) (Allergen Cla h 3)	metabolism	95	1.00E-18	49/85 (57%)	60/85 (70%)	YGLAAAV, AGTVWVN, PFGGYK, GRELGE
Ssc#S6090499	sp|O43099|	Putative peroxiredoxin pmp20 (Thioredoxin reductase)	metabolism	94	2.00E-18	57/176 (32%)	87/176 (49%)	PGAFTP
Ssc#S17526279	sp|O43099|	Putative peroxiredoxin pmp20 (Thioredoxin reductase)	metabolism	93	4.00E-18	56/176 (31%)	86/176 (48%)	PGAFTP
Ssc#S6057638	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	91	1.00E-17	50/93 (53%)	62/93 (66%)	GPNTNGS
Ssc#S16517353	sp|O44119|	Tropomyosin (Allergen Hom a 1)	structural protein	89	6.00E-17	64/194 (32%)	85/194 (43%)	ICNYGPAGN
Ssc#S17526881	ref|NP_741138.1|	venom allergen 5, LONg family member (lon-1)	metabolism	89	5.00E-17	52/144 (36%)	80/144 (55%)	RKYEEVARKL
Ssc#S17527194	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	87	3.00E-16			
Ssc#S17526951	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	87	2.00E-16			
Ssc#S6054734	sp|P42037|	60S acidic ribosomal protein P2 (Minor allergen Alt a 6)	protein synthesis	86	1.00E-25	43/91 (47%)	60/91 (65%)	GEVCPA
Ssc#S6076895	ref|NP_842456.1|	Alkyl hydroperoxide reductase	metabolism	86	2.00E-16	47/115 (40%)	62/115 (53%)	GNTSPSA, SVGIEAD
Ssc#S16767408	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	85	5.00E-16	45/138 (32%)	70/138 (50%)	nil
Ssc#S17526356	ref|NP_671747.1|	alpha-2u globulin PGCL1 [Rattus norvegicus]	homeostasis	85	1.00E-15	50/136 (36%)	74/136 (54%)	nil
Ssc#S17526677	sp|P02789|	Ovotransferrin precursor (Conalbumin) (Allergen Gal d 3)	homeostasis	84	1.00E-15	43/113 (38%)	70/113 (61%)	nil
Ssc#S17527193	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	82	3.00E-15			
Ssc#S17506573	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	80	4.00E-14			
Ssc#S6085810	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	80	4.00E-14			
Ssc#S17509918	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	80	4.00E-14			
Ssc#S14900653	ref|NP_741138.1|	venom allergen 5, LONg family member (lon-1)	metabolism	79	4.00E-14	52/158 (32%)	69/158 (43%)	CNYGPAGN
Ssc#S6084992	sp|Q9M7R0|	Calcium-binding allergen Ole e 8 (PCA18/PCA23)	signal transduction	78	3.00E-14	42/106 (39%)	64/106 (60%)	nil
Ssc#S16767712	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	78	1.00E-13	39/123 (31%)	65/123 (52%)	nil

Appendix III. Detail lists of sequence homology matches for pork with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology (cont.).

Query	Subject (hits)	Subject Description	Functional category	Bit Score	E-value	Identities	Positives	6 amino acid homology
Ssc#S17501675	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	77	1.00E-13			
Ssc#S14765758	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13) (Bt6)	metabolism	77	3.00E-13	45/131 (34%)	76/131 (58%)	GKEFEED, RTLSTFRN
Ssc#S6081386	sp|O97370|	Mite allergen Eur m 3 precursor	metabolism	77	2.00E-13	41/84 (48%)	55/84 (65%)	GDSGGP
Ssc#S6055999	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	77	2.00E-13	44/130 (33%)	69/130 (53%)	GEKVKTVV, EGDNK
Ssc#S6089848	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	77	5.00E-20	36/60 (60%)	41/60 (68%)	EKGFGY, GKSIYG
Ssc#S17526279	sp|P56577|	Allergen Mal f 2 (MF1)	metabolism	76	5.00E-13	35/75 (46%)	46/75 (61%)	DFTFVCPTE
Ssc#S16516418	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	76	4.00E-13			
Ssc#S6091287	sp|O18598|	Glutathione S-transferase (GST class-sigma)	metabolism	76	5.00E-13	41/120 (34%)	62/120 (51%)	PGAFTP
Ssc#S16515479	ref|NP_842456.1|	Alkyl hydroperoxide reductase	metabolism	76	8.00E-13	60/195 (30%)	96/195 (49%)	nil
Ssc#S6053690	sp|P02789|	Ovotransferrin precursor (Conalbumin) (Allergen Gal d 3)	homeostasis	75	3.00E-13	39/94 (41%)	60/94 (63%)	nil
Ssc#S17527544	sp|P40108|	Aldehyde dehydrogenase (ALDDH) (Allergen Cla h 3)	metabolism	74	1.00E-12	36/63 (57%)	46/63 (73%)	AGTVWVN, PFGGYK, GRELGE
Ssc#S16516911	sp|O97370|	Mite allergen Eur m 3 precursor	metabolism	74	2.00E-12	36/67 (53%)	47/67 (70%)	CQGDSGGP
Ssc#S6091286	gb|AAB72147.1|	allergen Bla g 5 [Blattella germanica]	metabolism	73	6.00E-12	59/195 (30%)	95/195 (48%)	nil
Ssc#S14767224	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	72	9.00E-12	35/114 (30%)	64/114 (56%)	nil
Ssc#S16516041	sp|P46419|	Glutathione S-transferase (GST class-mu)	metabolism	72	6.00E-12	32/87 (36%)	54/87 (62%)	nil
Ssc#S6053232	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	72	5.00E-12	39/127 (30%)	65/127 (51%)	nil
Ssc#S16514502	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	72	6.00E-12	42/128 (32%)	71/128 (55%)	nil
Ssc#S6091224	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	72	7.00E-12			
Ssc#S16765201	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	72	3.00E-11	76/384 (19%)	163/384 (42%)	nil
Ssc#S6081229	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	71	1.00E-11			
Ssc#S6061469	sp|P42037|	60S acidic ribosomal protein P2 (Minor allergen Alt a 6)	protein synthesis	70	6.00E-12	36/62 (58%)	44/62 (70%)	KLASVP
Ssc#S17504676	ref|NP_509707.1|	C07A4.3 [Caenorhabditis elegans]	unknown	70	3.00E-11	44/150 (29%)	76/150 (50%)	HNEYRQK
Ssc#S17503360	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	69	6.00E-11			
Ssc#S6090938	emb|CAD23374.1|	tri s 4 allergen [Trichophyton schoenleinii]	metabolism	68	9.00E-10	58/243 (23%)	107/243 (44%)	GSTGFGQ
Ssc#S17526971	sp|P50344|	60S acidic ribosomal protein P1 (Allergen Cla h 12)	protein synthesis	67	2.00E-10	42/113 (37%)	52/113 (46%)	LFAKAL, SDDDMGFGLFD
Ssc#S17526972	sp|P50344|	60S acidic ribosomal protein P1 (Allergen Cla h 12)	protein synthesis	67	2.00E-10	42/113 (37%)	52/113 (46%)	LFAKAL, SDDDMGFGLFD
Ssc#S17527394	sp|P50344|	60S acidic ribosomal protein P1 (Allergen Cla h 12)	protein synthesis	67	2.00E-10	42/113 (37%)	52/113 (46%)	LFAKAL, SDDDMGFGLFD
Ssc#S17527395	sp|P50344|	60S acidic ribosomal protein P1 (Allergen Cla h 12)	protein synthesis	67	2.00E-10	42/113 (37%)	52/113 (46%)	LFAKAL, SDDDMGFGLFD
Ssc#S17513058	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	66	4.00E-10			
Ssc#S16768146	sp|Q9UW02|	Thioredoxin (Allergen Cop c 2)	Chaparone/stress-related	65	1.00E-09			
Ssc#S6091057	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13) (Bt6)	metabolism	65	5.00E-10	44/108 (40%)	54/108 (50%)	GDSGG, GLVSWG, GIYTRV
Ssc#S6072210	sp|P50344|	60S acidic ribosomal protein P1 (Allergen Cla h 12)	protein synthesis	65	2.00E-09			
Ssc#S6061819	sp|P40918|	Heat shock 70 kDa protein (Allergen Cla h 4) (Cla h IV)	Chaparone/stress-related	65	9.00E-10			
Ssc#S17506578	sp|P39675|	Mite allergen Der f 3 precursor (Der f III)	metabolism	65	1.00E-09	37/128 (28%)	71/128 (55%)	nil
Ssc#S16512438	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	65	8.00E-10	35/90 (38%)	50/90 (55%)	nil
Ssc#S17526921	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	65	1.00E-09	42/113 (37%)	52/113 (46%)	ESDDDMGFGLFD
Ssc#S14889738	ref|NP_777021.1|	S100 calcium-binding protein A7 (psoriasin 1) [Bos taurus]	signal transduction	64	2.00E-09	33/91 (36%)	54/91 (59%)	nil
Ssc#S17526498	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	64	2.00E-09			
Ssc#S17526761	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	63	3.00E-09			
Ssc#S17526114	sp|P01005|	Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	homeostasis	62	4.00E-09			
Ssc#S17526115	sp|P01005|	Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	homeostasis	62	5.00E-09	44/171 (25%)	70/171 (40%)	nil
Ssc#S17525077	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	62	5.00E-09	44/171 (25%)	70/171 (40%)	nil
Ssc#S16771069	gb|AAD52672.1|	98kDa HDM allergen [Dermatophagoides farinae]	metabolism	61	2.00E-08	39/120 (32%)	59/120 (49%)	WVGYDD
Ssc#S6081164	sp|P49275|	Mite allergen Der f 3 precursor (Der f III)	metabolism	60	4.00E-08	38/139 (27%)	66/139 (47%)	WVLTAAH, DNDIAL
Ssc#S14889469	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	60	2.00E-08	26/59 (44%)	38/59 (64%)	nil
Ssc#S6053822	sp|O97370|	Mite allergen Eur m 3 precursor	metabolism	59	3.00E-08	29/69 (42%)	40/69 (57%)	QGDSGGP
Ssc#S6090725	gb|AAK61827.1|	allergen A precursor [Psoroptes ovis]	unknown	59	2.00E-07	35/112 (31%)	53/112 (47%)	nil
Ssc#S17514592	sp|P42041|	Aldehyde dehydrogenase (ALDDH) (Allergen Alt a 10)	metabolism	58	1.00E-07			
Ssc#S14900438	ref|NP_777021.1|	S100 calcium-binding protein A7 (psoriasin 1) [Bos taurus]	signal transduction	58	4.00E-12	40/138 (28%)	70/138 (50%)	nil

Appendix III. Detail lists of sequence homology matches for pork with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology (cont.).

Query	Subject (hits)	Subject Description	Functional category	Bit Score	E-value	Identities	Positives	6 amino acid homology
Ssc#S6074621	gb|AAB72147.1|	allergen Bla g 5 [Blattella germanica]	metabolism	58	3.00E-08	30/91 (32%)	52/91 (57%)	nil
Ssc#S17514843	gb|AAA99805.1|	Der f 3 mite allergen	metabolism	58	1.00E-07			
Ssc#S17526817	sp|P02789|	Ovotransferrin precursor (Conalbumin) (Allergen Gal d 3)	homeostasis	57	3.00E-07	54/217 (24%)	102/217 (47%)	nil
Ssc#S17509917	sp|O18598|	Glutathione S-transferase (GST class-sigma)	metabolism	57	3.00E-12	26/60 (43%)	42/60 (70%)	HAVVARDD
Ssc#S6091335	ref|NP_671747.1|	alpha-2u globulin PGCL1 [Rattus norvegicus]	homeostasis	57	2.00E-07	41/161 (25%)	74/161 (45%)	nil
Ssc#S6054592	sp|P39674|	Allergen MAG29	Chaparone/stress-related	56	3.00E-07	52/189 (27%)	88/189 (46%)	nil
Ssc#S6014546	sp|O18598|	Glutathione S-transferase (GST class-sigma)	metabolism	56	4.00E-07			
Ssc#S16518057	gb|AAM64112.1|	gelsolin-like allergen Der f 16 [Dermatophagoides farinae]	structural protein	56	2.00E-06	63/277 (22%)	117/277 (42%)	nil
Ssc#S17501633	gb|AAK39511.1|	paramyosin-like allergen [Dermatophagoides farinae]	structural protein	56	6.00E-13	46/169 (27%)	81/169 (47%)	nil
Ssc#S6091678	gb|AAD25927.1|	major allergenic protein Mal f4 [Malassezia furfur]	unknown	56	2.00E-07	26/57 (45%)	38/57 (66%)	VCNPII
Ssc#S17527560	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	55	5.00E-07			
Ssc#S16517598	sp|Q9M7R0|	Calcium-binding allergen Ole e 8 (PCA18/PCA23)	signal transduction	54	2.00E-06			
Ssc#S6090814	sp|P01005|	Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	homeostasis	52	2.00E-06	20/41 (48%)	28/41 (68%)	PVCGTDGVTY
Ssc#S17526077	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	52	1.00E-05	59/220 (26%)	102/220 (46%)	nil
Ssc#S16770867	gb|AAK39511.1|	paramyosin-like allergen [Dermatophagoides farinae]	structural protein	52	7.00E-06			
Ssc#S6090946	ref|NP_509802.1|	T05A10.5 [Caenorhabditis elegans]	unknown	51	2.00E-05	28/102 (27%)	50/102 (49%)	nil
Ssc#S16512766	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	51	6.00E-06	29/67 (43%)	33/67 (49%)	nil
Ssc#S17526404	sp|Q9NAS5|	Tropomyosin, muscle (Allergen Ani s 3)	structural protein	50	2.00E-05	31/110 (28%)	46/110 (41%)	nil
Ssc#S17525633	sp|Q95182|	Major allergen Equ c 1 precursor	transport protein	50	5.00E-05	30/141 (21%)	64/141 (45%)	nil
Ssc#S17514554	sp|P01005|	Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	homeostasis	50	3.00E-05	29/81 (35%)	46/81 (56%)	nil
Ssc#S17527413	ref|NP_777021.1|	S100 calcium-binding protein A7 (psoriasin 1) [Bos taurus]	signal transduction	50	2.00E-05	31/157 (19%)	79/157 (50%)	nil
Ssc#S14889652	ref|NP_777021.1|	S100 calcium-binding protein A7 (psoriasin 1) [Bos taurus]	signal transduction	50	2.00E-05	29/81 (35%)	46/81 (56%)	nil
Ssc#S6076164	sp|Q9UW02|	Thioredoxin (Allergen Cop c 2)	Chaparone/stress-related	49	1.00E-04	22/66 (33%)	42/66 (63%)	nil
Ssc#S6059348	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	49	2.00E-04	33/103 (32%)	49/103 (47%)	nil
Ssc#S6084107	gb|AAB34785.1|	68 kDa allergen [Penicillium chrysogenum]	metabolism	49	7.00E-05	32/99 (32%)	46/99 (46%)	nil
Ssc#S6091507	gb|AAB34785.1|	68 kDa allergen [Penicillium chrysogenum]	metabolism	49	1.00E-05	29/74 (39%)	37/74 (50%)	ADHPFLF
Ssc#S6074301	sp|Q9UW02|	Thioredoxin (Allergen Cop c 2)	Chaparone/stress-related	48	1.00E-04	23/50 (46%)	32/50 (64%)	nil
Ssc#S17526070	sp|Q95PM9|	Arginine kinase (AK) (Allergen Plo i 1)	metabolism	48	2.00E-04	21/71 (29%)	39/71 (54%)	nil
Ssc#S17518044	ref|NP_777021.1|	S100 calcium-binding protein A7 (psoriasin 1) [Bos taurus]	signal transduction	48	3.00E-05	25/75 (33%)	43/75 (57%)	nil
Ssc#S6052752	gb|AAM64112.1|	gelsolin-like allergen Der f 16 [Dermatophagoides farinae]	structural protein	48	7.00E-05	25/81 (30%)	43/81 (53%)	FHKYSG
Ssc#S6090612	sp|Q9M7R0|	Calcium-binding allergen Ole e 8 (PCA18/PCA23)	signal transduction	47	1.00E-04	30/82 (36%)	46/82 (56%)	nil
Ssc#S6052231	sp|P40108|	Aldehyde dehydrogenase (ALDDH) (Allergen Cla h 3)	metabolism	47	4.00E-04	48/260 (18%)	110/260 (42%)	nil
Ssc#S17516187	sp|P01005|	Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	homeostasis	47	8.00E-05	19/33 (57%)	26/33 (78%)	nil
Ssc#S14893538	ref|NP_777021.1|	S100 calcium-binding protein A7 (psoriasin 1) [Bos taurus]	signal transduction	47	2.00E-04	18/50 (36%)	29/50 (58%)	nil
Ssc#S17526542	ref|NP_777021.1|	S100 calcium-binding protein A7 (psoriasin 1) [Bos taurus]	signal transduction	47	3.00E-04	28/90 (31%)	53/90 (58%)	FHKYSG
Ssc#S16517063	gb|AAK39511.1|	paramyosin-like allergen [Dermatophagoides farinae]	structural protein	47	3.00E-04	30/97 (30%)	48/97 (49%)	nil
Ssc#S6084171	sp|Q9NAS5|	Tropomyosin, muscle (Allergen Ani s 3)	structural protein	46	5.00E-04	59/266 (22%)	110/266 (41%)	nil
Ssc#S6029777	ref|NP_777021.1|	S100 calcium-binding protein A7 (psoriasin 1) [Bos taurus]	signal transduction	46	3.00E-04	28/90 (31%)	52/90 (57%)	FHKYSG
Ssc#S6074754	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	46	7.00E-04	24/61 (39%)	35/61 (57%)	nil
Ssc#S6091596	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13) (Bt6)	metabolism	45	5.00E-04	35/134 (26%)	58/134 (43%)	nil
Ssc#S16762892	gb|AAD52672.1|	98kDa HDM allergen [Dermatophagoides farinae]	metabolism	35	0.99	23/73 (31%)	31/73 (42%)	nil

Appendix IV. Detail lists of sequence homology matches for chicken with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology.

Query	Subject (hits)	Subject Description	Functional category	Bit Score	E-value	Identities	Positives	6 amino acid homology
Gga#S7088995	sp|P02789|	Ovotransferrin precursor (Conalbumin) (Allergen Gal d 3)	homeostasis	1419	0	686/705 (97%)	688/705 (97%)	highly homologous
Gga#S7088912	sp|P19121|	Serum albumin precursor (Alpha-livetin) (Allergen Gal d 5)	homeostasis	1262	0	615/615 (100%)	615/615 (100%)	highly homologous
Gga#S7088142	sp|P40918|	Heat shock 70 kDa protein (Allergen Cla h 4) (Cla h IV)	Chaparone/stress-related	852	0			
Gga#S7089514	sp|P40918|	Heat shock 70 kDa protein (Allergen Cla h 4) (Cla h IV)	Chaparone/stress-related	744	0			
Gga#S7088619	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	728	0	371/386 (96%)	371/386 (96%)	highly homologous
Gga#S7088538	sp|Q9LEJ0|	Enolase 1	metabolism	595	1.00E-169	315/442 (71%)	363/442 (82%)	highly homologous
Gga#S7087857	sp|Q9LEI9|	Enolase 2	metabolism	583	1.00E-165	306/444 (68%)	360/444 (81%)	highly homologous
Gga#S7088537	sp|Q9LEJ0|	Enolase 1	metabolism	547	1.00E-154	290/442 (65%)	342/442 (77%)	highly homologous
Gga#S7088879	sp|P40292|	Heat shock protein 90 (Heat shock protein hsp1)	Chaparone/stress-related	512	1.00E-143	260/445 (58%)	325/445 (73%)	highly homologous
Gga#S7088128	sp|P40108|	Aldehyde dehydrogenase (ALDDH) (Allergen Cla h 3)	metabolism	511	1.00E-143	257/476 (53%)	337/476 (70%)	highly homologous
Gga#S7087329	sp|P40108|	Aldehyde dehydrogenase (ALDDH) (Allergen Cla h 3)	metabolism	488	1.00E-136	248/478 (51%)	330/478 (69%)	highly homologous
Gga#S7086749	sp|P40108|	Aldehyde dehydrogenase (ALDDH) (Allergen Cla h 3)	metabolism	479	1.00E-133	241/477 (50%)	318/477 (66%)	highly homologous
Gga#S7086560	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	467	1.00E-129			
Gga#S7089425	sp|P01005|	Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	homeostasis	452	1.00E-126	210/210 (100%)	210/210 (100%)	highly homologous
Gga#S7086621	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	450	1.00E-124			
Gga#S7087643	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	442	1.00E-122			
Gga#S7088770	sp|P02789|	Ovotransferrin precursor (Conalbumin) (Allergen Gal d 3)	homeostasis	428	1.00E-118	266/706 (37%)	379/706 (53%)	highly homologous
Gga#S7087480	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	369	1.00E-100	235/757 (31%)	378/757 (49%)	KCSSLEK
Gga#S7088958	gb|AAO15713.1|	allergen Pen m 2 [Penaeus monodon]	metabolism	275	1.00E-72	150/340 (44%)	209/340 (61%)	LSSLEGE, QQKLIDDHFLF, HNDNKTFLVWVNEEDHLR
Gga#S7087899	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	271	3.00E-71	146/409 (35%)	233/409 (56%)	nil
Gga#S7088730	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	263	8.00E-69			
Gga#S10817582	gb|AAD52672.1|	98kDa HDM allergen [Dermatophagoides farinae]	metabolism	258	2.00E-67	145/400 (36%)	216/400 (54%)	
Gga#S10817582	gb|AAD52672.1|	98kDa HDM allergen [Dermatophagoides farinae]	metabolism	258	2.00E-67	145/400 (36%)	216/400 (54%)	DWEYPGS
Gga#S7042222	sp|Q92260|	Heat shock 70 kDa protein (Allergen Pen c 19)	Chaparone/stress-related	257	3.00E-67	125/180 (69%)	149/180 (82%)	TKDNNLLGKF, LTGIPPAPRGVPQIEVTFD, NKITITNDKGRLSK, EAEKYKAEDEA
Gga#S7089575	sp|Q95PM9|	Arginine kinase (AK) (Allergen Plo i 1)	metabolism	256	8.00E-67	146/337 (43%)	201/337 (59%)	QQQLIDDHFLF, KTFLVW, CPTNLGT, VYDISN
Gga#S7089006	gb|AAO15713.1|	allergen Pen m 2 [Penaeus monodon]	metabolism	255	1.00E-66	147/340 (43%)	198/340 (58%)	QQQLIDDHFLF, HNDNKTFLVW, NEEDHLR, VHIKLP
Gga#S7089219	sp|Q9NAS5|	Tropomyosin, muscle (Allergen Ani s 3)	structural protein	239	6.00E-62	128/282 (45%)	180/282 (63%)	MDAIKKKMQ, AEEADRKYDEVARKL, LKEAETRAEFAERSV
Gga#S7048071	gb|AAD25927.1|	major allergenic protein Mal f4 [Malassezia furfur]	unknown	222	7.00E-61	116/193 (60%)	140/193 (72%)	highly homologous
Gga#S7088010	gb|AAM64112.1|	gelsolin-like allergen Der f 16 [Dermatophagoides farinae]	structural protein	209	3.00E-52	127/363 (34%)	193/363 (53%)	LVPVPK
Gga#S7089351	gb|AAM64112.1|	gelsolin-like allergen Der f 16 [Dermatophagoides farinae]	structural protein	204	5.00E-51	123/361 (34%)	185/361 (51%)	LLHVKGKK
Gga#S7088077	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	199	6.00E-52			
Gga#S7088756	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	198	3.00E-49	115/385 (29%)	203/385 (52%)	nil
Gga#S6758316	emb|CAA09884.1|	allergen [Malassezia sympodialis]	structural protein	192	7.00E-48	96/163 (58%)	118/163 (72%)	DVVPKTA, FADENF, PGLLSMAN, WLDGKHVVFG
Gga#S7087325	sp|Q92450|	Superoxide dismutase [Mn]	metabolism	188	2.00E-46	95/198 (47%)	121/198 (61%)	KFNGGGHINH, WEHAYYLQY, IWNVINW
Gga#S7089362	sp|O44119|	Tropomyosin (Allergen Hom a 1)	structural protein	187	9.00E-46	107/282 (37%)	147/282 (52%)	AIKKKMQ, LNRRIQL, AEEADRKYDEVARKL, NNLKSLE, RAEFAERSV
Gga#S7007685	sp|Q9N2R3|	Tropomyosin (Allergen Cha f 1) (Cha f I)	structural protein	178	2.00E-43	105/238 (44%)	144/238 (60%)	AIKKKMQ, AEGEVAALNRRIQL, AEEADRKY
Gga#S7024633	sp|P01005|	Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	homeostasis	176	9.00E-43	77/193 (39%)	118/193 (61%)	HDGECK, PVCGTDG, TYDNEC, VCGTDGVTY
Gga#S7088303	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	176	9.00E-43	115/365 (31%)	184/365 (50%)	nil
Gga#S7088670	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	174	2.00E-42	119/391 (30%)	191/391 (48%)	nil
Gga#S7020852	sp|P46419|	Glutathione S-transferase (GST class-mu)	metabolism	172	2.00E-42	81/182 (44%)	122/182 (67%)	GPAPDFD, LGLDFPNLPY, LGYWDI
Gga#S7088640	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	163	4.00E-39			
Gga#S6955753	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	160	4.00E-38			
Gga#S7040668	sp|P42041|	Aldehyde dehydrogenase (ALDDH) (Allergen Alt a 10)	metabolism	149	8.00E-35	73/147 (49%)	96/147 (65%)	EEIFGPV, ALRAGTVWVN, PFGGYK
Gga#S7088968	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	147	9.00E-34	105/381 (27%)	180/381 (47%)	nil
Gga#S7088301	sp|P39675|	Mite allergen Der p 3 precursor (Der p III)	metabolism	147	1.00E-33			
Gga#S7087868	sp|P02769|	Serum albumin precursor (Allergen Bos d 6) (BSA)	homeostasis	143	9.00E-33	109/463 (23%)	196/463 (42%)	nil
Gga#S6752375	sp|Q91483|	Parvalbumin beta 2 (Major allergen Sal s 1)	homeostasis	140	2.00E-32	66/102 (64%)	87/102 (85%)	AADSFN, DQDKSGFIEE, AETKAFLA, DGDGKIGV
Gga#S7046462	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	136	1.00E-41	68/99 (68%)	74/99 (74%)	KPGLLSMANAGP, TNGSQFFITTV, LDGKHVVFG
Gga#S16073703	sp|P39675|	Mite allergen Der p 3 precursor (Der p III)	metabolism	135	1.00E-30			

Appendix IV. Detail lists of sequence homology matches for chicken with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology (cont.).

Query	Subject (hits)	Subject Description	Functional category	Bit Score	E-value	Identities	Positives	6 amino acid homology
Gga#S12377738	gb|AAA99805.1|	Der f 3 mite allergen	metabolism	131	3.00E-29			
Gga#S7051840	sp|P49275|	Mite allergen Der f 3 precursor (Der f III)	metabolism	122	3.00E-38	71/183 (38%)	93/183 (50%)	HFCGGS, KGYPGVYT
Gga#S7088024	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	120	3.00E-25			
Gga#S7089676	sp|P59747|	Parvalbumin beta (Allergen Sco j 1)	homeostasis	119	5.00E-26	56/108 (51%)	77/108 (71%)	SGFIEE, DGDGKIG
Gga#S16508721	sp|P01005|	Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	homeostasis	117	9.00E-26	63/141 (44%)	85/141 (60%)	nil
Gga#S7087864	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	116	1.00E-23	176/835 (21%)	348/835 (41%)	nil
Gga#S6937106	sp|P40918|	Heat shock 70 kDa protein (Allergen Cla h 4) (Cla h IV)	Chaparone/stress-related	110	3.00E-23	65/211 (30%)	113/211 (53%)	AVITVPAYFN
Gga#S7088951	sp|Q06478|	Phospholipase A1 1 precursor (Allergen Dol m 1.01)	metabolism	109	2.00E-22	78/274 (28%)	128/274 (46%)	GHSLGAH, GLDPAGP,
Gga#S7087186	gb|AAK39511.1|	paramyosin-like allergen [Dermatophagoides farinae]	structural protein	106	4.00E-21			
Gga#S7021690	sp|O18598|	Glutathione S-transferase (GST class-sigma)	metabolism	104	3.00E-21	47/97 (48%)	66/97 (68%)	PSYKLTY
Gga#S6967219	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	103	5.00E-21			
Gga#S7006405	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	103	6.00E-21			
Gga#S16575233	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	102	8.00E-21	53/126 (42%)	75/126 (59%)	STFKNTEI, FKLGEEF
Gga#S7050592	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	102	1.00E-20	53/126 (42%)	75/126 (59%)	STFKNTEI, FKLGEEF
Gga#S6767541	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II)	homeostasis	98	3.00E-19	51/135 (37%)	77/135 (57%)	ANSGISS, ADHPFLF
Gga#S14760878	sp|P39675|	Mite allergen Der p 3 precursor (Der p III)	metabolism	94	2.00E-18			
Gga#S7023254	sp|O97370|	Mite allergen Eur m 3 precursor	metabolism	91	2.00E-17	58/197 (29%)	98/197 (49%)	LTAAHCV, GDSGGP
Gga#S7046109	sp|Q9M7R0|	Calcium-binding allergen Ole e 8 (PCA18/PCA23)	signal transduction	91	4.00E-17			
Gga#S7087231	pir||B37330	venom allergen III - red imported fire ant	metabolism	89	4.00E-16	60/183 (32%)	87/183 (47%)	YLVCNY
Gga#S7087199	sp|P35781|	Venom allergen 5.01 (Antigen 5-1) (Ag5-1)	metabolism	89	1.00E-16	60/178 (33%)	88/178 (49%)	DEVKDY, HYTQMVWA, YLVCNY
Gga#S7035521	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13) (Bt6)	metabolism	86	2.00E-17	43/107 (40%)	63/107 (58%)	STFKNTEI, FKLGEEF
Gga#S7030500	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13) (Bt6)	metabolism	85	2.00E-15	44/126 (34%)	71/126 (56%)	GEEFEE
Gga#S7089649	sp|P01005|	Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	homeostasis	84	4.00E-14			
Gga#S7041077	sp|Q90YK9|	Parvalbumin beta (Allergen Gad m 1)	homeostasis	83	6.00E-15	46/107 (42%)	60/107 (56%)	SGFIEEEEELK
Gga#S7089204	gb|AAK39511.1|	paramyosin-like allergen [Dermatophagoides farinae]	structural protein	81	7.00E-14	107/459 (23%)	204/459 (44%)	nil
Gga#S6760447	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	81	3.00E-14			
Gga#S6755446	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	81	2.00E-14			
Gga#S7022623	ref|NP_842456.1|	Alkyl hydroperoxide reductase	metabolism	77	4.00E-29	38/76 (50%)	48/76 (63%)	DFTFVCPTE, GEVCPA
Gga#S16508680	sp|Q9UW02|	Thioredoxin (Allergen Cop c 2)	Chaparone/stress-related	76	1.00E-13	34/77 (44%)	48/77 (62%)	nil
Gga#S7014316	gb|AAB72147.1|	allergen Bla g 5 [Blattella germanica]	metabolism	74	4.00E-12	54/194 (27%)	102/194 (52%)	nil
Gga#S7086812	gb|AAK39511.1|	paramyosin-like allergen [Dermatophagoides farinae]	structural protein	73	2.00E-11	83/381 (21%)	167/381 (43%)	nil
Gga#S7089578	ref|NP_671747.1|	alpha-2u globulin PGCL1 [Rattus norvegicus]	homeostasis	73	1.00E-11	41/148 (27%)	67/148 (45%)	nil
Gga#S7088757	sp|P01005|	Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	homeostasis	73	9.00E-12	51/188 (27%)	74/188 (39%)	nil
Gga#S7089243	sp|O18598|	Glutathione S-transferase (GST class-sigma)	metabolism	72	3.00E-11	57/208 (27%)	107/208 (51%)	nil
Gga#S7050015	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	72	1.00E-11	38/132 (28%)	74/132 (56%)	GEEFEED
Gga#S7088830	gb|AAK39511.1|	paramyosin-like allergen [Dermatophagoides farinae]	structural protein	70	3.00E-10			
Gga#S6910547	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	69	1.00E-10	46/108 (42%)	57/108 (52%)	nil
Gga#S7087904	gb|AAB72147.1|	allergen Bla g 5 [Blattella germanica]	metabolism	69	1.00E-10	47/163 (28%)	84/163 (51%)	nil
Gga#S7087705	gb|AAK39511.1|	paramyosin-like allergen [Dermatophagoides farinae]	structural protein	69	5.00E-10	81/407 (19%)	164/407 (40%)	nil
Gga#S6829946	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	67	3.00E-10	33/90 (36%)	55/90 (61%)	nil
Gga#S7046376	sp|P40918|	Heat shock 70 kDa protein (Allergen Cla h 4) (Cla h IV)	Chaparone/stress-related	65	9.00E-18			
Gga#S7089336	gb|AAK39511.1|	paramyosin-like allergen [Dermatophagoides farinae]	structural protein	64	2.00E-08			
Gga#S6943084	sp|P49148|	60S acidic ribosomal protein P1 (Allergen Alt a 12)	protein synthesis	64	4.00E-09	35/62 (56%)	42/62 (67%)	LFAKAL
Gga#S7089207	gb|AAK39511.1|	paramyosin-like allergen [Dermatophagoides farinae]	structural protein	62	6.00E-08	81/386 (20%)	159/386 (41%)	nil
Gga#S7046470	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	62	7.00E-09			
Gga#S7045877	sp|P40108|	Aldehyde dehydrogenase (ALDDH) (Allergen Cla h 3)	metabolism	62	5.00E-11	50/178 (28%)	82/178 (46%)	KEEIFGP
Gga#S14761110	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	60	1.00E-08			
Gga#S6922012	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	59	1.00E-07	34/70 (48%)	38/70 (54%)	TGEKGF, TGGKSIYG, DFMLQGGDFT
Gga#S7087550	gb|AAK39511.1|	paramyosin-like allergen [Dermatophagoides farinae]	structural protein	59	1.00E-06	72/309 (23%)	133/309 (43%)	QENTEL

Appendix IV. Detail lists of sequence homology matches for chicken with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology (cont.).

Query	Subject (hits)	Subject Description	Functional category	Bit Score	E-value	Identities	Positives	6 amino acid homology
Gga#S16567772	sp|Q9NAS5|	Tropomyosin, muscle (Allergen Ani s 3)	structural protein	59	1.00E-07	33/159 (20%)	84/159 (52%)	nil
Gga#S7050535	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	59	6.00E-08	30/101 (29%)	57/101 (56%)	nil
Gga#S14753739	sp|P01005|	Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	homeostasis	58	1.00E-07	27/57 (47%)	33/57 (57%)	nil
Gga#S7086605	gb|AAM21322.1|	group 14 allergen protein [Dermatophagoides pteronyssinus]	metabolism	57	3.00E-06	84/361 (23%)	143/361 (39%)	nil
Gga#S7088409	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	57	2.00E-06			
Gga#S7089429	sp|P01005|	Ovomucoid precursor (Allergen Gal d 1) (Gal d I)	homeostasis	57	8.00E-08	24/24 (100%)	24/24 (100%)	NKCNFCNAVVESNGTLTLSHFGKC
Gga#S7087161	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13) (Bt6)	metabolism	57	1.00E-07	33/130 (25%)	67/130 (51%)	nil
Gga#S6773814	sp|O18598|	Glutathione S-transferase (GST class-sigma)	metabolism	56	7.00E-07	49/153 (32%)	73/153 (47%)	nil
Gga#S7088954	gb|AAD25927.1|	major allergenic protein Mal f4 [Malassezia furfur]	unknown	55	4.00E-06			
Gga#S7087511	gb|AAK39511.1|	paramyosin-like allergen [Dermatophagoides farinae]	structural protein	55	8.00E-06			
Gga#S7089062	ref|NP_777021.1|	S100 calcium-binding protein A7 (psoriasin 1) [Bos taurus]	signal transduction	55	8.00E-07	29/91 (31%)	52/91 (57%)	nil
Gga#S7026285	sp|Q9UW02|	Thioredoxin (Allergen Cop c 2)	Chaparone/stress-related	55	2.00E-06	34/106 (32%)	56/106 (52%)	nil
Gga#S7035595	prf||2118249B	allergen Lep d 1.02	unknown	54	5.00E-06	30/106 (28%)	53/106 (50%)	nil
Gga#S7089017	sp|Q9M7R0|	Calcium-binding allergen Ole e 8 (PCA18/PCA23)	signal transduction	54	6.00E-06	47/172 (27%)	78/172 (45%)	nil
Gga#S6920649	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	53	5.00E-06			
Gga#S7086632	sp|Q9UW02|	Thioredoxin (Allergen Cop c 2)	Chaparone/stress-related	53	2.00E-05			
Gga#S6928896	sp|P49369|	Phospholipase A1 precursor (Allergen Ves v 1) (Ves v I)	metabolism	52	1.00E-05	22/58 (37%)	32/58 (55%)	nil
Gga#S7087316	gb|AAD25927.1|	major allergenic protein Mal f4 [Malassezia furfur]	unknown	51	4.00E-05			
Gga#S14762459	sp|P49371|	Hyalurononglucosaminidase (Hyaluronidase) (Allergen Dol m 2)	metabolism	51	2.00E-05	45/170 (26%)	60/170 (35%)	nil
Gga#S7088064	sp|Q9NAS5|	Tropomyosin, muscle (Allergen Ani s 3)	structural protein	51	6.00E-05			
Gga#S7088883	ref|NP_173866.1|	calcium ion binding [Arabidopsis thaliana]	signal transduction	49	2.00E-04			
Gga#S7036572	ref|NP_504417.1|	C10G8.3 [Caenorhabditis elegans]	unknown	49	2.00E-04			
Gga#S6944910	sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	metabolism	49	2.00E-04			
Gga#S6930035	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13) (Bt6)	metabolism	48	2.00E-04	26/75 (34%)	40/75 (53%)	nil
Gga#S7031814	sp|Q9UW02|	Thioredoxin (Allergen Cop c 2)	Chaparone/stress-related	47	7.00E-04			

Appendix V. Detail lists of sequence homology matches for trout with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology.

Query	Subject (hits)	Subject Description	Functional category	Bit Score	E-value	Identities	Positives	6 amino acid homology
Omy#S15341035	sp|P02789|	Ovotransferrin precursor (Conalbumin)	homeostasis	645	0.00E+00	354/701 (50%)	461/701 (65%)	highly homologous
Omy#S15341314	sp|Q92260|	Heat shock 70 kDa protein (Allergen Pen c 19)	chaparone/stress-related	409	1.00E-113			
Omy#S15341215	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	378	1.00E-103	218/681 (32%)	367/681 (53%)	
Omy#S18160274	sp|Q9LEJ0|	Enolase 1	metabolism	313	2.00E-84	171/252 (67%)	194/252 (76%)	Highly homologous
Omy#S18097301	sp|Q92260|	Heat shock 70 kDa protein (Allergen Pen c 19)	Chaparone/stress-related	312	3.00E-84	159/250 (63%)	188/250 (75%)	highly homologous
Omy#S18098044	gb|AAD25927.1|	major allergenic protein Mal f4 [Malassezia furfur]	unknown	258	8.00E-68	148/253 (58%)	174/253 (68%)	Highly homologous
Omy#S15341258	sp|Q95PM9|	Arginine kinase (AK) (Allergen Plo i 1)	metabolism	251	3.00E-65	147/364 (40%)	206/364 (56%)	DPIIED, QQQLIDDHFLF, KTFLVWVNEEDHLR, CPTNLGT
Omy#S15330950	sp|Q9NAS5|	Tropomyosin, muscle (Allergen Ani s 3)	structural protein	232	4.00E-60	125/225 (55%)	157/225 (69%)	highly homologous
Omy#S18163682	sp|Q92260|	Heat shock 70 kDa protein (Allergen Pen c 19)	chaparone/stress-related	227	2.00E-58	109/159 (68%)	135/159 (84%)	Highly homologous
Omy#S18098124	sp|P40108|	Aldehyde dehydrogenase (ALDDH)	metabolism	221	1.00E-56	117/256 (45%)	158/256 (61%)	KKEPIGVC, GNTVVLK, LELGGKSP
Omy#S15340856	sp|P00785|	Actinidain precursor (Actinidin)	metabolism	220	6.00E-56			
Omy#S15241444	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	220	2.00E-56	106/157 (67%)	120/157 (76%)	DVVPKTA, FHRVIP, QGGDFT, NGTGGKSIYG, LSMANAGPNTNGSQFFI, WLDGKHVVFG
Omy#S18101805	sp|Q9LEI9|	Enolase 2	metabolism	220	2.00E-56	105/143 (73%)	123/143 (86%)	highly homologous
Omy#S15302100	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	214	2.00E-54	101/157 (64%)	116/157 (73%)	highly homologous
Omy#S18093011	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	209	4.00E-53	100/157 (63%)	113/157 (71%)	highly homologous
Omy#S18102422	sp|Q92450|	Superoxide dismutase [Mn]	metabolism	207	1.00E-52	100/194 (51%)	129/194 (66%)	FNGGGHINH, QGSGWGWL, WEHAYYLQY, IWNVINW
Omy#S18100439	sp|P40918|	Heat shock 70 kDa protein (Allergen Cla h 4)	chaparone/stress-related	194	2.00E-48	104/250 (41%)	150/250 (60%)	nil
Omy#S18100604	sp|Q9LEJ0|	Enolase 1	metabolism	192	7.00E-48	93/132 (70%)	115/132 (87%)	Highly homologous
Omy#S18151560	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	188	8.00E-47	96/163 (58%)	116/163 (71%)	highly homologous
Omy#S18160728	sp|Q92260|	Heat shock 70 kDa protein (Allergen Pen c 19)	chaparone/stress-related	186	4.00E-46	96/194 (49%)	133/194 (68%)	LTGIPPAPRGVPQIEVTF, NKITITND, LESYAYSLKN
Omy#S15341266	sp|P00698|	Lysozyme C precursor	metabolism	186	4.00E-46	78/127 (61%)	104/127 (81%)	NTQATNRNTDGSTDYGI, DGRTPG, AWVAWR
Omy#S15298097	gb|AAO15713.1|	allergen Pen m 2 [Penaeus monodon]	metabolism	184	1.00E-45	100/249 (40%)	143/249 (57%)	QQKLIDDHFLF, KTFLVWVNEEDHLR, CPTNLGT
Omy#S15279465	sp|P42041|	Aldehyde dehydrogenase (ALDDH)	metabolism	184	1.00E-45	95/210 (45%)	125/210 (59%)	LFINNE, ATGNTVV, AEQTPL, GFPPGV
Omy#S15314908	sp|Q91482|	Parvalbumin beta 1 (Major allergen Sal s 1)	homeostasis	183	3.00E-45	94/109 (86%)	96/109 (88%)	highly homologous
Omy#S15276554	sp|P00785|	Actinidain precursor (Actinidin)	metabolism	181	2.00E-44			
Omy#S15263229	sp|Q91483|	Parvalbumin beta 2 (Major allergen Sal s 1)...	homeostasis	180	2.00E-44	92/108 (85%)	93/108 (86%)	highly homologous
Omy#S15289021	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	179	4.00E-44	88/147 (59%)	105/147 (71%)	NFRALCTG, GTGGKSIYG, PGLLSMAN, SQFFIT
Omy#S15296653	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2)	homeostasis	179	4.00E-44	99/269 (36%)	155/269 (57%)	SALAMV, VLVNAI
Omy#S15317026	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	178	5.00E-44	89/165 (53%)	113/165 (68%)	LCTGEKG, LLSMAN, TNGSQFFITT
Omy#S15300801	sp|P49275|	Mite allergen Der f 3 precursor	metabolism	178	1.00E-43			
Omy#S15261700	sp|Q91483|	Parvalbumin beta 2 (Major allergen Sal s 1)...	homeostasis	178	7.00E-44	92/108 (85%)	93/108 (86%)	highly homologous
Omy#S15299978	sp|O61379|	Tropomyosin (Allergen Pan s 1)	structural protein	175	7.00E-43	98/217 (45%)	127/217 (58%)	NRRIQL
Omy#S18101278	sp|P00785|	Actinidain precursor (Actinidin)	metabolism	171	2.00E-41	87/208 (41%)	111/208 (53%)	YWIVKNSW
Omy#S15299885	gb|AAA99805.1|	Der f 3 mite allergen	metabolism	170	3.00E-41			
Omy#S18097897	ref|NP_842456.1|	Alkyl hydroperoxide reductase	metabolism	166	3.00E-40	84/173 (48%)	112/173 (64%)	DFTFVCPTE, GEVCPA
Omy#S18098073	sp|P39675|	Mite allergen Der p 3 precursor	metabolism	164	1.00E-39			
Omy#S15298319	sp|P39675|	Mite allergen Der p 3 precursor	metabolism	161	1.00E-38			
Omy#S15253908	pdb|1FSK|	B Chain B, Fab Fragment of IgG antibody		158	1.00E-37	86/217 (39%)	123/217 (56%)	LTQSPKS, TLTIS+VQAED
Omy#S15289106	sp|Q95PM9|	Arginine kinase (AK) (Allergen Plo i 1)	metabolism	158	1.00E-37	84/185 (45%)	110/185 (59%)	QQQLIDDHFL, KTFLVWVNEEDHLR, CPTNLGT
Omy#S15295934	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	157	2.00E-37	81/140 (57%)	93/140 (66%)	TGEKGFGY, GKSIYG, LSMANAG, TNGSQFFITT, WLDGKHVVFG
Omy#S15340985	sp|P02789|	Ovotransferrin precursor (Conalbumin)	homeostasis	156	8.00E-37	107/355 (30%)	156/355 (43%)	GDVAFVKHTT
Omy#S15291877	ref|NP_842456.1|	Alkyl hydroperoxide reductase	metabolism	153	4.00E-36	80/174 (45%)	109/174 (62%)	DFTFVCPTE, GEVCPA
Omy#S15328178	sp|Q870B9|	Enolase	metabolism	149	5.00E-55	72/125 (57%)	97/125 (77%)	highly homologous
Omy#S15340913	ref|NP_842456.1|	Alkyl hydroperoxide reductase	metabolism	146	6.00E-34	74/167 (44%)	104/167 (62%)	DFTFVCPTE, GEVCPA
Omy#S15297807	sp|P39674|	Allergen MAG29	chaparone/stress-related	146	4.00E-34	66/101 (65%)	85/101 (84%)	DIERMV, VCNPIITKLYQ
Omy#S18092321	sp|Q91483|	Parvalbumin beta 2 (Major allergen Sal s 1)...	homeostasis	146	2.00E-34	75/108 (69%)	83/108 (76%)	highly homologous
Omy#S15344074	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2)	homeostasis	145	3.00E-33			
Omy#S18163799	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	145	5.00E-34			
Omy#S15247823	sp|Q08169|	Hyaluronoglucosaminidase precursor	metabolism	144	2.00E-33	85/269 (31%)	146/269 (54%)	NGGVPQ

Appendix V. Detail lists of sequence homology matches for trout with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology (cont.).

Query	Subject (hits)	Subject Description	Functional category	Bit Score	E-value	Identities	Positives	6 amino acid homology
Omy#S15301263	sp|P49275|	Mite allergen Der f 3 precursor	metabolism	142	7.00E-33			
Omy#S15286468	sp|P49275|	Mite allergen Der f 3 precursor	metabolism	142	5.00E-33			
Omy#S18150128	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	140	2.00E-32			
Omy#S15239535	ref|NP_842456.1|	Alkyl hydroperoxide reductase	metabolism	138	1.00E-31	72/170 (42%)	102/170 (60%)	DFTFVCPTE, GEVCPA
Omy#S15297133	gb|AAO15713.1|	allergen Pen m 2 [Penaeus monodon]	metabolism	137	6.00E-39	81/192 (42%)	113/192 (58%)	nil
Omy#S18148000	sp|P40292|	Heat shock protein HSP1	Chaparone/stress-related	137	2.00E-31	70/155 (45%)	104/155 (67%)	NMERIMKAQA
Omy#S15261218	sp|P00785|	Actinidain precursor (Actinidin)	metabolism	136	3.00E-31			
Omy#S15300863	sp|P49275|	Mite allergen Der f 3 precursor	metabolism	136	5.00E-31			
Omy#S15324732	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	135	9.00E-31	65/102 (63%)	78/102 (76%)	highly homologous
Omy#S15300340	sp|P40292|	Heat shock protein HSP1	chaparone/stress-related	135	1.00E-30	72/155 (46%)	101/155 (65%)	ANMERIMKAQALRD,
Omy#S15329061	gb|AAA99805.1|	Der f 3 mite allergen	metabolism	127	1.00E-28			
Omy#S18163842	sp|Q92260|	Heat shock 70 kDa protein (Allergen Pen c 19)	chaparone/stress-related	125	7.00E-28			
Omy#S18165126	pdb|1FSK|	B Chain B, Fab Fragment of IgG antibody		124	2.00E-27	68/151 (45%)	93/151 (61%)	DFTLTIS, PPSSEQL, SSTLTL
Omy#S15296242	sp|P40292|	Heat shock protein HSP1	chaparone/stress-related	124	3.00E-27	64/136 (47%)	90/136 (66%)	ANMERIMKAQA
Omy#S15320563	sp|P49275|	Mite allergen Der f 3 precursor	metabolism	119	5.00E-26			
Omy#S18145010	sp|Q9NFZ4|	Tropomyosin (Allergen Lep d 10)	structural protein	119	4.00E-26	60/109 (55%)	81/109 (74%)	KLKEAE, RAEFAERSV
Omy#S18145976	sp|P42041|	Aldehyde dehydrogenase (ALDDH)	metabolism	117	2.00E-25	54/116 (46%)	80/116 (68%)	EEIFGPV, PFGGYK
Omy#S18104480	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	117	2.00E-29			
Omy#S18153844	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	116	3.00E-39			TNGSQFFITT, LCTGEKG, DFMLQGGDF
Omy#S15287234	sp|Q92260|	Heat shock 70 kDa protein (Allergen Pen c 19)	chaparone/stress-related	116	4.00E-25	65/135 (48%)	86/135 (63%)	NKITITNDKGRLSKE, EAEKYK
Omy#S18159676	sp|P00785|	Actinidain precursor (Actinidin)	metabolism	113	3.00E-24			
Omy#S15315310	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2)	homeostasis	113	3.00E-24			
Omy#S15316954	sp|Q9N2R3|	Tropomyosin (Allergen Cha f 1)	structural protein	111	1.00E-23	56/127 (44%)	83/127 (65%)	
Omy#S15319952	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2)	homeostasis	110	3.00E-23	68/220 (30%)	118/220 (53%)	MLVLLPDE
Omy#S15252829	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13)	metabolism	107	2.00E-22	52/123 (42%)	78/123 (63%)	STFKNTEI, FKLGEEF
Omy#S18160812	sp|Q06478|	Phospholipase A1 1 precursor	metabolism	106	4.00E-22	68/213 (31%)	108/213 (50%)	SLGAHISGF, LDPAGP
Omy#S15281917	sp|Q9M7R0|	Calcium-binding allergen Ole e 8	signal transduction	105	1.00E-21			
Omy#S15297301	ref|NP_173866.1|	polcalcin (calcium-binding pollen allergen)	signal transduction	105	1.00E-21			
Omy#S15310703	sp|Q90YK9|	Parvalbumin beta (Allergen Gad m 1)	homeostasis	104	1.00E-21	54/107 (50%)	71/107 (66%)	DQDKSG, LFLQNFS
Omy#S18158866	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	102	4.00E-21			
Omy#S18092374	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13)	metabolism	102	6.00E-21	48/115 (41%)	73/115 (63%)	STFKNTEI, FKLGEEF
Omy#S15340855	sp|Q06478|	Phospholipase A1 1 precursor	metabolism	102	3.00E-20	78/283 (27%)	122/283 (43%)	SLGAHV, GLDPAGP
Omy#S15274427	sp|P59747|	Parvalbumin beta (Allergen Sco j 1)	homeostasis	98	1.00E-19	52/106 (49%)	61/106 (57%)	GSFDHKKFF, SGFIEEEELK, DAETKAFL
Omy#S15301757	emb|CAA09883.1|	allergen [Malassezia sympodialis]	unknown	95	1.00E-18	48/107 (44%)	70/107 (65%)	PGAFTP, ELKSKGV, VNDAFVM
Omy#S18102363	sp|P42037|	60S acidic ribosomal protein P2	protein synthesis	95	1.00E-18	48/113 (42%)	63/113 (55%)	LGGNTSPS, SVGIEA, GKDINE, DMGFGLFD
Omy#S15320006	sp|Q92260|	Heat shock 70 kDa protein (Allergen Pen c 19)	chaparone/stress-related	94	2.00E-18	46/69 (66%)	54/69 (78%)	TKDNNLL, LTGIPPAP, QIEVTFD
Omy#S18096761	sp|Q9UW02|	Thioredoxin (Allergen Cop c 2)	chaparone/stress-related	94	2.00E-18	42/92 (45%)	63/92 (68%)	nil
Omy#S18161350	sp|Q25456|	Tropomyosin (Allergen Met e 1)	structural protein	94	3.00E-18	55/107 (51%)	65/107 (60%)	LERAEERAE, NNLKSLE, RAEFAERSV
Omy#S18152048	emb|CAA09883.1|	allergen [Malassezia sympodialis]	unknown	93	6.00E-18	49/121 (40%)	74/121 (61%)	PGAFTP, NDAFVM
Omy#S15258008	sp|Q9UW02|	Thioredoxin (Allergen Cop c 2)	chaparone/stress-related	91	2.00E-17	43/104 (41%)	65/104 (62%)	ATWCGPC
Omy#S15275396	sp|P40292|	Heat shock protein HSP1	chaparone/stress-related	90	3.00E-17	48/107 (44%)	72/107 (67%)	nil
Omy#S18151973	sp|P59747|	Parvalbumin beta (Allergen Sco j 1)	homeostasis	89	7.00E-18	44/73 (60%)	49/73 (67%)	nil
Omy#S15266961	sp|P42037|	60S acidic ribosomal protein P2	protein synthesis	87	1.00E-16	46/114 (40%)	61/114 (53%)	LESVGIEAD, DMGFGLFD
Omy#S15276631	ref|NP_173866.1|	polcalcin (calcium-binding pollen allergen)	signal transduction	85	1.00E-15			
Omy#S18102906	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13)	metabolism	84	1.00E-15	41/125 (32%)	67/125 (53%)	STFKNTE
Omy#S18095855	sp|P40108|	Aldehyde dehydrogenase (ALDDH)	metabolism	83	4.00E-15			
Omy#S15298004	emb|CAA09884.1|	allergen [Malassezia sympodialis]	unknown	82	1.00E-14	48/102 (47%)	60/102 (58%)	nil
Omy#S18094395	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13)	metabolism	81	2.00E-14	42/128 (32%)	74/128 (57%)	GEEFEE
Omy#S15297826	sp|P39675|	Mite allergen Der p 3 precursor	metabolism	80	5.00E-15			

Appendix V. Detail lists of sequence homology matches for trout with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology (cont.).

Query	Subject (hits)	Subject Description	Functional category	Bit Score	E-value	Identities	Positives	6 amino acid homology
Omy#S18099206	sp|P42037|	60S acidic ribosomal protein P2	protein synthesis	80	5.00E-14	37/61 (60%)	51/61 (83%)	LGGNTSPS, SVGIEA, GKDINE
Omy#S15311381	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	78	1.00E-13	42/107 (39%)	58/107 (54%)	EIKFKL
Omy#S18095529	sp|P00785|	Actinidain precursor (Actinidin)	metabolism	75	9.00E-13			
Omy#S15281065	sp|P39675|	Mite allergen Der p 3 precursor	metabolism	75	7.00E-13			
Omy#S15289030	sp|P40292|	Heat shock protein HSP1	Chaparone/stress-related	75	1.00E-13	41/103 (39%)	66/103 (64%)	nil
Omy#S18160615	gb|AAM83103.1|	paramyosin allergen [Blomia tropicalis]	structural protein	75	3.00E-13			
Omy#S18094587	gb|AAD52672.1|	98kDa HDM allergen [Dermatophagoides farinae]	metabolism	73	6.00E-22	38/99 (38%)	60/99 (60%)	NVMTYD
Omy#S15296122	sp|Q9NFZ4|	Tropomyosin (Allergen Lep d 10)	structural protein	73	3.00E-12	35/57 (61%)	45/57 (78%)	KLKEAE, RAEFAERSV
Omy#S18101599	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13)	metabolism	72	6.00E-12	41/134 (30%)	77/134 (57%)	GEEFEE
Omy#S15291492	sp|Q9NFZ4|	Tropomyosin (Allergen Lep d 10)	structural protein	72	7.00E-12	36/66 (54%)	46/66 (69%)	RAEFAERSV
Omy#S15330831	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2)	homeostasis	71	2.00E-11			
Omy#S15298184	sp|Q9M7R0|	Calcium-binding allergen Ole e 8	signal transduction	70	3.00E-11			
Omy#S15301789	sp|P39675|	Mite allergen Der p 3 precursor	metabolism	70	8.00E-12	33/64 (51%)	42/64 (65%)	VSWGYGCA
Omy#S15337042	sp|O18598|	Glutathione S-transferase	metabolism	70	5.00E-11			
Omy#S15291988	sp|Q9M7R0|	Calcium-binding allergen Ole e 8	signal transduction	69	6.00E-11			
Omy#S15336377	gb|AAB72147.1|	allergen Bla g 5 [Blattella germanica]	metabolism	69	5.00E-11	50/162 (30%)	76/162 (46%)	nil
Omy#S15296648	sp|P49275|	Mite allergen Der f 3 precursor	metabolism	69	1.00E-10			
Omy#S15292324	sp|P00785|	Actinidain precursor (Actinidin)	metabolism	68	2.00E-10			
Omy#S15329842	sp|P00785|	Actinidain precursor (Actinidin)	metabolism	68	1.00E-10	46/152 (30%)	70/152 (46%)	nil
Omy#S18150434	emb|CAD23374.1|	tri s 4 allergen [Trichophyton schoenleinii]	metabolism	68	2.00E-10	56/229 (24%)	96/229 (41%)	
Omy#S15291186	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2)	homeostasis	67	3.00E-10			
Omy#S18164789	ref|XP_135574.1|	similar to Major urinary protein 6 precursor	unknown	67	4.00E-10	47/181 (25%)	86/181 (47%)	nil
Omy#S15298379	prf||2118249B	allergen Lep d 1.02	unknown	66	9.00E-10	41/128 (32%)	64/128 (50%)	nil
Omy#S18160554	sp|Q95PM9|	Arginine kinase (AK) (Allergen Plo i 1)	metabolism	66	4.00E-10	36/83 (43%)	50/83 (60%)	DPIIED
Omy#S15309334	sp|P49148|	60S acidic ribosomal protein P1	protein synthesis	65	9.00E-10	42/110 (38%)	52/110 (47%)	DMGFGLFD
Omy#S15238745	sp|P00785|	Actinidain precursor (Actinidin)	metabolism	64	2.00E-09			
Omy#S18157610	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	64	2.00E-09	37/128 (28%)	68/128 (53%)	nil
Omy#S18092678	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	64	2.00E-09	33/77 (42%)	47/77 (61%)	STFKNTEI
Omy#S18147957	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	64	2.00E-09	34/101 (33%)	58/101 (57%)	LSTFRN, GEEFEE
Omy#S15316324	ref|XP_135574.1|	similar to Major urinary protein 6 precursor	unknown	64	4.00E-09	40/154 (25%)	75/154 (48%)	nil
Omy#S18093730	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2)	homeostasis	63	4.00E-09			
Omy#S15299468	sp|P39675|	Mite allergen Der p 3 precursor	metabolism	62	2.00E-10			
Omy#S18162445	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13)	metabolism	60	4.00E-08	33/130 (25%)	61/130 (46%)	nil
Omy#S15273089	gb|AAM64112.1|	gelsolin-like allergen Der f 16	structural protein	60	3.00E-08	40/133 (30%)	66/133 (49%)	nil
Omy#S15298661	sp|Q95PM9|	Arginine kinase (AK) (Allergen Plo i 1)	metabolism	60	1.00E-08	35/91 (38%)	46/91 (50%)	CPTNLGT
Omy#S18149611	sp|P42039|	60S acidic ribosomal protein P2	protein synthesis	60	5.00E-08	26/66 (39%)	43/66 (65%)	GKDINE
Omy#S18094102	sp|O97370|	Mite allergen Eur m 3 precursor	transport protein	59	7.00E-08			
Omy#S18095061	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13)	metabolism	59	9.00E-08	28/88 (31%)	51/88 (57%)	nil
Omy#S18165005	gb|AAM64112.1|	gelsolin-like allergen Der f 16	structural protein	59	1.00E-07	39/136 (28%)	67/136 (49%)	nil
Omy#S15259426	sp|P42037|	60S acidic ribosomal protein P2	protein synthesis	59	1.00E-10	27/46 (58%)	37/46 (80%)	LGGNTSPS, SVGIEA
Omy#S18147972	ref|XP_135574.1|	similar to Major urinary protein 6 precursor	unknown	59	1.00E-07	31/103 (30%)	53/103 (51%)	nil
Omy#S18145916	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	57	1.00E-07	33/126 (26%)	64/126 (50%)	nil
Omy#S15280016	sp|Q9UW02|	Thioredoxin (Allergen Cop c 2)	chaparone/stress-related	56	6.00E-07			
Omy#S15299958	sp|Q9M7R0|	Calcium-binding allergen Ole e 8	signal transduction	55	1.00E-06			
Omy#S15301794	ref|NP_173866.1|	polcalcin (calcium-binding pollen allergen)	signal transduction	55	1.00E-06			
Omy#S15272111	gb|AAB34785.1|	68 kDa allergen [Penicillium chrysogenum]	unknown	54	1.00E-06	31/88 (35%)	49/88 (55%)	nil
Omy#S15278000	sp|P00785|	Actinidain precursor (Actinidin)	metabolism	54	3.00E-06	29/74 (39%)	43/74 (58%)	KYGKSY
Omy#S15278259	sp|P00785|	Actinidain precursor (Actinidin)	metabolism	54	4.00E-06	29/74 (39%)	43/74 (58%)	KYGKSY
Omy#S15259662	sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13)	metabolism	54	2.00E-06	33/124 (26%)	63/124 (50%)	nil

Appendix V. Detail lists of sequence homology matches for trout with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology (cont.).

Query	Subject (hits)	Subject Description	Functional category	Bit Score	E-value	Identities	Positives	6 amino acid homology
Omy#S18159014	ref|XP_135574.1|	similar to Major urinary protein 6 precursor	unknown	54	4.00E-06	37/154 (24%)	72/154 (46%)	nil
Omy#S15250021	gb|AAD52672.1|	98kDa HDM allergen [Dermatophagoides farinae]	metabolism	53	5.00E-06	25/56 (44%)	36/56 (64%)	nil
Omy#S15325049	sp|Q9M7R0|	Calcium-binding allergen Ole e 8	signal transduction	52	1.00E-05			
Omy#S18159084	sp|O18598|	Glutathione S-transferase	metabolism	52	2.00E-10	32/98 (32%)	50/98 (51%)	nil
Omy#S18149744	sp|P01005|	Ovomucoid precursor (Allergen Gal d 1)	homeostasis	52	1.00E-05	23/63 (36%)	33/63 (52%)	nil
Omy#S15325715	gb|AAM64112.1|	gelsolin-like allergen Der f 16	structural protein	51	2.00E-05	28/77 (36%)	42/77 (54%)	nil
Omy#S18150484	sp|P46419|	Glutathione S-transferase	metabolism	51	2.00E-05	34/137 (24%)	66/137 (48%)	nil
Omy#S15340590	ref|NP_505373.1|	21k allergen like (5J634)	unknown	50	3.00E-05	44/155 (28%)	65/155 (41%)	nil
Omy#S15287595	ref|NP_173866.1|	polcalcin (calcium-binding pollen allergen)	signal transduction	50	2.00E-05			
Omy#S18157215	ref|NP_777021.1|	S100 calcium-binding protein A7 (psoriasin 1)	signal transduction	50	2.00E-05	27/83 (32%)	50/83 (60%)	nil
Omy#S15311702	ref|NP_777021.1|	S100 calcium-binding protein A7 (psoriasin 1)	signal transduction	50	2.00E-05	27/83 (32%)	49/83 (59%)	nil
Omy#S18163138	sp|Q9M7R0|	Calcium-binding allergen Ole e 8	signal transduction	49	9.00E-05	42/130 (32%)	63/130 (48%)	IDTDKDGF
Omy#S15336978	sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	metabolism	49	1.00E-04	26/57 (45%)	33/57 (57%)	EIKFKL
Omy#S15318302	sp|P01012|	Ovalbumin (Plakalbumin) (Allergen Gal d 2)	homeostasis	49	1.00E-05	24/75 (32%)	37/75 (49%)	
Omy#S15321446	gb|AAD25927.1|	major allergenic protein Mal f4 [Malassezia furfur]	unknown	48	2.00E-04			
Omy#S15253358	ref|NP_777021.1|	S100 calcium-binding protein A7 (psoriasin 1)	signal transduction	48	9.00E-05	26/83 (31%)	49/83 (59%)	nil
Omy#S15299946	sp|Q9UW02|	Thioredoxin (Allergen Cop c 2) .	chaparone/stress-related	48	2.00E-04			
Omy#S18093805	gb|AAD52672.1|	98kDa HDM allergen [Dermatophagoides farinae]	metabolism	47	9.00E-05	28/51 (54%)	33/51 (64%)	nil
Omy#S18096822	sp|O24171|	Profilin 3 (Pollen allergen Ole e 2)	structural protein	47	4.00E-04	47/143 (32%)	69/143 (48%)	MSWQSYV
Omy#S15272298	ref|NP_777021.1|	S100 calcium-binding protein A7 (psoriasin 1)	signal transduction	47	2.00E-04	26/83 (31%)	47/83 (56%)	FHKYSG
Omy#S18150137	sp|Q91482|	Parvalbumin beta 1 (Major allergen Sal s 1)	homeostasis	46	7.00E-04	27/73 (36%)	39/73 (53%)	nil
Omy#S18161972	gb|AAD32205.1|	putative allergen protein [Prunus armeniaca]	unknown	46	5.00E-04	42/148 (28%)	62/148 (41%)	nil
Omy#S18162187	gb|AAM64112.1|	gelsolin-like allergen Der f 16	structural protein	45	2.00E-04	23/65 (35%)	39/65 (60%)	nil
Omy#S15288908	ref|NP_173866.1|	polcalcin (calcium-binding pollen allergen)	signal transduction	45	9.00E-04			
Omy#S18146465	gb|AAL92870.1|	pollen allergen Che a 2 [Chenopodium album]	unknown	39	0.075	43/146 (29%)	64/146 (43%)	GLHLGG

Appendix VI. Detail lists of sequence homology matches for sheep with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology.

Query	Hit	Hit Description	Function	Bit-score	E-value	Identity	Positivity	6 amino acid homology
gi|1386|emb|X17055.1|OASERALB[1386]	gi|1351907|sp|P02769|ALBU_BOVIN[1351907]	Serum albumin precursor (Allergen Bos d 6)	Homeostasis	1181	0.0	561/607 (92%)	588/607 (96%)	Highly homologous
gi|527681|gb|U12761.1|OAU12761[527681]	gi|729320|sp|P40108|DHAL_CLAHE[729320]	Aldehyde dehydrogenase (ALDDH) (Allergen Cla h 3)	Metabolism	494	e-138	248/478 (51%)	330/478 (69%)	Highly homologous
gi|1315|emb|X04520.1|OALGLBR[1315]	gi|125910|sp|P02754|LACB_BOVIN[125910]	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	323	3e-87	159/167 (95%)	165/167 (98%)	Highly homologous
gi|1309|emb|X06367.1|OALACA[1309]	gi|27805979|ref|NP_776803.1|[27805979]	lactalbumin, alpha [Bos taurus]	Homeostasis	262	6e-69	123/142 (86%)	126/142 (88%)	Highly homologous
gi|602293|gb|U17988.1|OAU17988[602293]	gi|5815436|gb|AAD52672.1|AF178772_1[5815436]	98kDa HDM allergen [Dermatophagoides farinae]	Metabolism	235	2e-60	145/421 (34%)	226/421 (53%)	FDGLDL, WVGYDD
gi|37576791|gb|AY392761.1|[37576791]	gi|5815436|gb|AAD52672.1|AF178772_1[5815436]	98kDa HDM allergen [Dermatophagoides farinae]	Metabolism	231	2e-59	132/366 (36%)	199/366 (54%)	FDGLDL
gi|31084950|gb|CD286907.1|[31084950]	gi|4138173|emb|CAA09884.1|[4138173]	allergen [Malassezia sympodialis]	Unknown	231	7e-60	112/157 (71%)	124/157 (78%)	TGEKGFGY, QGGDFT , NGTGGKSIYG ,LSMANAGPNTNGSQFFI ,WLDGKHVVFG
gi|31085783|gb|CD287740.1|[31085783]	gi|4138173|emb|CAA09884.1|[4138173]	allergen [Malassezia sympodialis]	Unknown	223	2e-57	108/153 (70%)	120/153 (78%)	TGEKGFGY , QGGDFT , NGTGGKSIYG , LSMANAGPNTNGSQFFI , WLDGKHVVF
gi|885600|gb|U16719.1|OAU16719[885600]	gi|5815436|gb|AAD52672.1|AF178772_1[5815436]	98kDa HDM allergen [Dermatophagoides farinae]	Metabolism	222	2e-56	140/421 (33%)	220/421 (52%)	FDGLDL, WVGYDD,
gi|33180696|gb|CF117600.1|[33180696]	gi|4138173|emb|CAA09884.1|[4138173]	allergen [Malassezia sympodialis]	Unknown	219	2e-56	106/147 (72%)	117/147 (79%)	TGEKGFGY , QGGDFT , NGTGGKSIYG , LSMANAGPNTNGSQFFI , WLDGKHVVFG
gi|31085773|gb|CD287730.1|[31085773]	gi|4138173|emb|CAA09884.1|[4138173]	allergen [Malassezia sympodialis]	Unknown	219	2e-56	108/157 (68%)	120/157 (76%)	TGEKGFGY , QGGDFT , NGTGGKSIYG , LSMANAGPNTNGSQ , LDGKHVVFG
gi|31087233|gb|CD289190.1|[31087233]	gi|4138173|emb|CAA09884.1|[4138173]	allergen [Malassezia sympodialis]	Unknown	200	3e-50	97/136 (71%)	106/136 (77%)	QGGDFT , NGTGGKSIYG , GPNTNGSQFFI , WLDGKHVVFG
gi|29893401|gb|AY251270.1|[29893401]	gi|4138173|emb|CAA09884.1|[4138173]	allergen [Malassezia sympodialis]	Unknown	176	1e-43	84/110 (76%)	91/110 (82%)	TGEKGFGY, QGGDFT, NGTGGKSIYG , LSMANAGPNTNGSQFFI , WLDGKHVVFG
gi|31086436|gb|CD288393.1|[31086436]	gi|1170095|sp|P46419|GTM1_DERPT[1170095]	Glutathione S-transferase (GST class-mu)	Metabolism	174	1e-42	78/141 (55%)	106/141 (75%)	ILGYWDIRG , WLNEKF , LGLDFPNLPY
gi|18699095|gb|AF411974.1|[18699095]	gi|14423735|sp|Q9U6V9|HUGA_POLAN[14423735]	Hyaluronoglucosaminidase precursor	Metabolism	174	4e-42	105/335 (31%)	170/335 (50%)	LFPSVY
gi|33177820|gb|CF116165.1|[33177820]	gi|14423687|sp|Q9LEI9|ENO2_HEVBR[14423687]	Enolase 2	Metabolism	172	2e-42	81/113 (71%)	96/113 (84%)	DLYKSF, VSIEDPFDQDDWE , VTNPKR, LLLKVNQIGSVTES, SHRSGETEDTFIADL
gi|31086489|gb|CD288446.1|[31086489]	gi|1170095|sp|P46419|GTM1_DERPT[1170095]	Glutathione S-transferase (GST class-mu)	Metabolism	165	4e-40	76/154 (49%)	106/154 (68%)	WLNEKF , LGLDFPNLPY
gi|31086018|gb|CD287975.1|[31086018]	gi|1170095|sp|P46419|GTM1_DERPT[1170095]	Glutathione S-transferase (GST class-mu)	Metabolism	162	1e-39	73/133 (54%)	98/133 (73%)	WLNEKF , LGLDFPNLPY , ILGYWDIRG
gi|31086149|gb|CD288106.1|[31086149]	gi|729764|sp|P40918|HS70_CLAHE[729764]	Heat shock 70 kDa protein (Allergen Cla h 4)	Chaperone/stress-related	161	5e-39	74/97 (76%)	88/97 (90%)	GIDLGTTYSCVG, EIIANDQGNRTTPS, VAFTDTERLIGD, AKNQVA, NTVFDAKRLIGR
gi|33182088|gb|CF118296.1|[33182088]	gi|4138173|emb|CAA09884.1|[4138173]	allergen [Malassezia sympodialis]	Unknown	160	1e-38	76/110 (69%)	86/110 (78%)	NFRALCTGEKGFG, FHRVIPDF, PGLLSMAN, SQFFIT
gi|31085538|gb|CD287495.1|[31085538]	gi|1170095|sp|P46419|GTM1_DERPT[1170095]	Glutathione S-transferase (GST class-mu)	Metabolism	156	1e-37	71/133 (53%)	95/133 (71%)	WLNEKF, LGLDFPNLPY
gi|7331112|gb|AF233075.1|AF233075[7331112]	gi|14423687|sp|Q9LEI9|ENO2_HEVBR[14423687]	Enolase 2	Metabolism	153	1e-36	79/110 (71%)	87/110 (79%)	HAGNKLAMQEFMILPVGASSF, GAEVYHHLK, DATNVGDEGGFAP, KVVIGMDVAASEFY
gi|33181791|gb|CF118146.1|[33181791]	gi|2851483|sp|P40292|HS82_ASPFU[2851483]	Heat shock protein HSP1	Chaperone/stress-related	145	3e-34	74/149 (49%)	103/149 (69%)	ANMERIMKAQALRD
gi|31084868|gb|CD286825.1|[31084868]	gi|1170095|sp|P46419|GTM1_DERPT[1170095]	Glutathione S-transferase (GST class-mu)	Metabolism	144	6e-34	65/132 (49%)	91/132 (68%)	LGYWDIRG, LDFPNLPY
gi|7274395|gb|AF233351.1|AF233351[7274395]	gi|4587985|gb|AAD25927.1|AF084828_1[4587985]	major allergenic protein Mal f4 [Malassezia furfur]	Unknown	144	6e-34	80/137 (58%)	95/137 (68%)	ISNPVNST, AEVFKK, FGVTTL, NVPVIGGH, EVVKAK
gi|31084662|gb|CD286619.1|[31084662]	gi|1170095|sp|P46419|GTM1_DERPT[1170095]	Glutathione S-transferase (GST class-mu)	Metabolism	143	1e-33	63/129 (48%)	90/129 (69%)	LGYWDIRG, LDFPNLPY
gi|20799497|gb|AF483003.1|[20799497]	gi|14423687|sp|Q9LEI9|ENO2_HEVBR[14423687]	Enolase 2	Metabolism	142	3e-43	69/94 (73%)	80/94 (85%)	VSLAVCK, LAMQEFMILPVGASSF, GAEVYH
gi|33179463|gb|CF116982.1|[33179463]	gi|2851483|sp|P40292|HS82_ASPFU[2851483]	Heat shock protein HSP1	Chaperone/stress-related	131	4e-30	60/103 (58%)	84/103 (81%)	KLGIHED, IYYITGES , QLKEFDGK
gi|31084020|gb|CD285977.1|[31084020]	gi|1170095|sp|P46419|GTM1_DERPT[1170095]	Glutathione S-transferase (GST class-mu)	Metabolism	121	3e-31	55/102 (53%)	72/102 (70%)	WLNEKF, LGLDFPNLPY
gi|31086912|gb|CD288869.1|[31086912]	gi|12231036|sp|Q92450|SODM_ASPFU[12231036]	Superoxide dismutase [Mn]	Metabolism	115	3e-25	52/91 (57%)	69/91 (75%)	KFNGGGHINHS
gi|31085365|gb|CD287322.1|[31085365]	gi|4138173|emb|CAA09884.1|[4138173]	allergen [Malassezia sympodialis]	Unknown	114	6e-25	56/75 (74%)	60/75 (80%)	TGEKGFGY , QGGDFT, NGTGGKSIYG
gi|11610598|dbj|AB052168.1|[11610598]	gi|27805979|ref|NP_776803.1|[27805979]	lactalbumin, alpha [Bos taurus]	Homeostasis	114	6e-24	50/52 (96%)	50/52 (96%)	FHTSGYDTQAIVQNNDSTEYGLFQINNKIWCKDDQNPHS, NICNISCD
gi|1812|emb|X68308.1|OOLPLIP[1812]	gi|548449|sp|Q06478|PA11_DOLMA[548449]	Phospholipase A1 1 precursor (Allergen Dol m 1.01)	Metabolism	112	3e-23	77/276 (27%)	127/276 (46%)	GLDPAGP
gi|31086186|gb|CD288143.1|[31086186]	gi|4138173|emb|CAA09884.1|[4138173]	allergen [Malassezia sympodialis]	Unknown	105	2e-39	50/70 (71%)	57/70 (81%)	LSMANAGPNTNGSQFFI , WLDGKHVVFG
gi|397946|emb|Z25864.1|OATHIORD[397946]	gi|11135375|sp|Q9UW02|THIO_COPCM[11135375]	Thioredoxin (Allergen Cop c 2)	Chaperone/stress-related	102	4e-21	47/105 (44%)	69/105 (65%)	ATWCGPC
gi|33178709|gb|CF116609.1|[33178709]	gi|4138173|emb|CAA09884.1|[4138173]	allergen [Malassezia sympodialis]	Unknown	101	3e-21	51/97 (52%)	59/97 (60%)	TGEKGFGY
gi|31085791|gb|CD287748.1|[31085791]	gi|4138173|emb|CAA09884.1|[4138173]	allergen [Malassezia sympodialis]	Unknown	97.4	6e-20	61/114 (53%)	68/114 (59%)	LSMANAGPNTNGSQFFI
gi|1217|emb|X07005.1|OABLG2[1217]	gi|125910|sp|P02754|LACB_BOVIN[125910]	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	93.2	2e-18	45/48 (93%)	48/48 (100%)	SLAMAASDISLLDAQSAPLRVYVEELKPTPEG, LEILLQKW
gi|45126132|emb|CQ772360.1|	gi|125910|sp|P02754|LACB_BOVIN[125910]	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	93.2	8e-17	45/48 (93%)	48/48 (100%)	SLAMAASDISLLDAQSAPLRVYVEELKPTPEG, LEILLQKW
gi|31086320|gb|CD288277.1|[31086320]	gi|4138173|emb|CAA09884.1|[4138173]	allergen [Malassezia sympodialis]	Unknown	86.7	5e-40	40/51 (78%)	42/51 (82%)	TGEKGFGY , QGGDFT , NGTGGKSIYG
gi|7288117|dbj|AB040058.1|[7288117]	gi|27805979|ref|NP_776803.1|[27805979]	lactalbumin, alpha [Bos taurus]	Homeostasis	84	8e-16	41/44 (93%)	43/44 (97%)	MMSFVSLLLVGILFHATQAEQLTKCEVF, YGGVSLPE
gi|33178236|gb|CF116373.1|[33178236]	gi|41688715|sp|Q8TFM9|RLA2_FUSCU[41688715]	60S acidic ribosomal protein P2 (Minor allergen Fus c 1)	Protein sysnthesis	82.4	2e-15	39/62 (62%)	52/62 (83%)	SVGIEAD
gi|31086798|gb|CD288755.1|[31086798]	gi|2851483|sp|P40292|HS82_ASPFU[2851483]	Heat shock protein HSP1	Chaperone/stress-related	82	7e-15	44/113 (38%)	73/113 (64%)	RIMKAQA , KKTFEI
gi|31085577|gb|CD287534.1|[31085577]	gi|4138173|emb|CAA09884.1|[4138173]	allergen [Malassezia sympodialis]	Unknown	80.9	9e-21	37/45 (82%)	38/45 (84%)	TGEKGFGY , QGGDFT, NGTGGKSIY
gi|33180730|gb|CF117617.1|[33180730]	gi|42559584|sp|Q23939|TPM_DERFA[42559584]	Tropomyosin (Allergen Der f 10)	Structural protein	80.9	5e-15	41/105 (39%)	62/105 (59%)	MEAIKKKMQ
gi|2318025|gb|AF012019.1|AF012019[2318025]	gi|1351295|sp|P02789|TRFE_CHICK[1351295]	Ovotransferrin precursor (Conalbumin) (Allergen Gal d 3)	Homeostasis	80.1	1e-14	35/61 (57%)	44/61 (72%)	FDEYFS, GCAPGS
gi|33182374|gb|CF118439.1|[33182374]	gi|1173071|sp|P42037|RLA2_ALTAL[1173071]	60S acidic ribosomal protein P2 (Minor allergen Alt a 6)	Protein sysnthesis	79.3	2e-14	38/62 (61%)	50/62 (80%)	VGIEAD
gi|31086369|gb|CD288326.1|[31086369]	gi|129293|sp|P01012|OVAL_CHICK[129293]	Ovalbumin (Plakalbumin) (Allergen Gal d 2)	Homeostasis	77.4	1e-13	42/155 (27%)	83/155 (53%)	SALAMV

Appendix VI. Detail lists of sequence homology matches for sheep with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology (cont.).

Query	Hit	Hit Description	Function	Bit-score	E-value	Identity	Positivity	6 amino acid homology
gi|31085636|gb|CD287593.1|[31085636]	gi|2851483|sp|P40292|HS82_ASPFU[2851483]	Heat shock protein HSP1	Chaperone/stress-related	77	2e-22	43/134 (32%)	76/134 (56%)	AKLLRY
gi|31086791|gb|CD288748.1|[31086791]	gi|37078092|sp|Q870B9|ENO_RHORB[37078092]	Enolase	Metabolism	67.4	1e-10	30/46 (65%)	39/46 (84%)	DSRGNPTVEV, VPSGASTG
gi|26000691|gb|AY157617.1|[26000691]	gi|14423698|sp|Q17284|FABP_BLOTA[14423698]	Fatty acid-binding protein (Allergen Blo t 13)	Metabolism	67.4	3e-09	31/58 (53%)	40/58 (68%)	STFKNTEI
gi|31084308|gb|CD286265.1|[31084308]	gi|1710589|sp|P50344|RLA1_CLAHE[1710589]	60S acidic ribosomal protein P1 (Allergen Cla h 12)	Protein sysnthesis	67	2e-10	42/113 (37%)	52/113 (46%)	LFAKAL, SDDDMGFGLFD
gi|31084215|gb|CD286172.1|[31084215]	gi|1710589|sp|P50344|RLA1_CLAHE[1710589]	60S acidic ribosomal protein P1 (Allergen Cla h 12)	Protein sysnthesis	66.6	2e-10	42/113 (37%)	52/113 (46%)	LFAKAL, SDDDMGFGLFD
gi|31086627|gb|CD288584.1|[31086627]	gi|1710589|sp|P50344|RLA1_CLAHE[1710589]	60S acidic ribosomal protein P1 (Allergen Cla h 12)	Protein sysnthesis	66.6	1e-10	42/113 (37%)	52/113 (46%)	LFAKAL , SDDDMGFGLFD
gi|33178800|gb|CF116653.1|[33178800]	gi|1350779|sp|P49148|RLA1_ALTAL[1350779]	60S acidic ribosomal protein P1 (Allergen Alt a 12)	Protein sysnthesis	62	7e-09	34/62 (54%)	42/62 (67%)	LFAKAL
gi|33181640|gb|CF118070.1|[33181640]	gi|4138173|emb|CAA09884.1|[4138173]	allergen [Malassezia sympodialis]	Unknown	56.6	1e-07	26/48 (54%)	36/48 (75%)	WLDGKHVVFG
gi|33182193|gb|CF118349.1|[33182193]	gi|129293|sp|P01012|OVAL_CHICK[129293]	Ovalbumin (Plakalbumin) (Allergen Gal d 2)	Homeostasis	55.8	2e-07	40/131 (30%)	62/131 (47%)	AMVYLGA
gi|27127169|gb|AF473547.1|[27127169]	gi|14423714|sp|Q9U5P1|FABP_LEPDS[14423714]	Fatty acid-binding protein (Allergen Lep d 13)	Metabolism	54.3	8e-07	26/49 (53%)	32/49 (65%)	STFKNTEI
gi|33178677|gb|CF116593.1|[33178677]	gi|27807077|ref|NP_777021.1|[27807077]	S100 calcium-binding protein A7 (psoriasin 1)	Signal transduction	52.8	2e-06	31/94 (32%)	51/94 (54%)	FHKYSG
gi|1219|emb|X07006.1|OABLG3[1219]	gi|125910|sp|P02754|LACB_BOVIN[125910]	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	51.6	6e-06	24/24 (100%)	24/24 (100%)	ENGECAQKKIIAEKTKIPAVFKID
gi|7595909|gb|AF241828.1|AF241828[7595909]	gi|17565876|ref|NP_504413.1|[17565876]	21k allergen like family member (5F762)	Unknown	46.2	0.002	25/61 (40%)	34/61 (54%)	GNKNNF
gi|31086544|gb|CD288501.1|[31086544]	gi|1170095|sp|P46419|GTM1_DERPT[1170095]	Glutathione S-transferase (GST class-mu)	Metabolism	45.8	2e-04	20/35 (57%)	26/35 (74%)	ILGYWDIRG
gi|4836365|gb|AF117693.1|AF117693[4836365]	gi|125910|sp|P02754|LACB_BOVIN[125910]	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	45.4	3e-04	20/22 (90%)	22/22 (100%)	FDKALKALPMHI, FNPTQLE
gi|1214|emb|X07004.1|OABLG1[1214]	gi|125910|sp|P02754|LACB_BOVIN[125910]	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	42.4	0.003	19/22 (86%)	20/22 (90%)	IVTQTMKGLDIQKV
gi|31087172|gb|CD289129.1|[31087172]	gi|1705483|sp|P54962|BLG4_BLAGE[1705483]	Allergen Bla g 4 precursor (Bla g IV)	Metabolism	37	0.24	16/20 (80%)	17/20 (85%)	VLATDYENYA
gi|33179271|gb|CF116886.1|[33179271]	gi|1705483|sp|P54962|BLG4_BLAGE[1705483]	Allergen Bla g 4 precursor (Bla g IV)	Metabolism	37	0.21	16/20 (80%)	17/20 (85%)	VLATDYENYA
gi|31085828|gb|CD287785.1|[31085828]	gi|1705483|sp|P54962|BLG4_BLAGE[1705483]	Allergen Bla g 4 precursor (Bla g IV)	Metabolism	37	0.18	16/20 (80%)	17/20 (85%)	VLATDYENYA
gi|31087082|gb|CD289039.1|[31087082]	gi|1705483|sp|P54962|BLG4_BLAGE[1705483]	Allergen Bla g 4 precursor (Bla g IV)	Metabolism	35.4	0.39	15/20 (75%)	16/20 (80%)	VLATDYENYA
gi|33178119|gb|CF116315.1|[33178119]	gi|27807077|ref|NP_777021.1|[27807077]	S100 calcium-binding protein A7 (psoriasin 1)	Signal transduction	35.4	0.8	17/46 (36%)	27/46 (58%)	FHKYSG

Appendix VII. Detail lists of sequence homology matches for goat with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology.

Query	Hit	Hit Description	Function	Bit-score	e-value	Identity	Positivity	6 animo acid homology
gi|6063528|dbj|AB033604.1|	gi|729764|sp|P40918|	Heat shock 70 kDa protein (Allergen Cla h 4)	Chaperone/stress-related	833	0	415/610 (68%)	489/610 (80%)	Highly homologous
gi|1280044|gb|U53857.1|	gi|1351295|sp|P02789|	Ovotransferrin precursor (Conalbumin) (Allergen Gal d 3)	Homeostasis	672	0	356/696 (51%)	465/696 (66%)	Highly homologous
gi|556806|emb|X78902.1|	gi|1351295|sp|P02789|	Ovotransferrin precursor (Conalbumin) (Allergen Gal d 3)	Homeostasis	667	0	355/696 (51%)	463/696 (66%)	Highly homologous
gi|437751|emb|Z19569.1|	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	324	2.00E-87	160/167 (95%)	164/167 (98%)	Highly homologous
gi|967|emb|X58471.1|	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	324	1.00E-87	160/167 (95%)	164/167 (98%)	Highly homologous
gi|23957583|gb|AF548366.1|	gi|2851483|sp|P40292|	Heat shock protein HSP1	Chaperone/stress-related	305	1.00E-81	145/191 (75%)	168/191 (87%)	LNKTKPIWTRNP , NDWEDHLAVKHFSVEGQLEFRA , RAPFDLFE , KNNIKLYVRRVFI , GVVDSEDLPLN , FSKNIKLGIHED
gi|979|emb|X05149.1|	gi|27805979|ref|NP_776803.1|	lactalbumin, alpha [Bos taurus]	Homeostasis	263	2.00E-69	124/142 (87%)	127/142 (89%)	Highly homologous
gi|19526602|gb|AY081150.1|	gi|5815436|gb|AAD52672.1|	98kDa HDM allergen [Dermatophagoides farinae]	Metabolism	249	1.00E-64	138/366 (37%)	206/366 (56%)	FDGLDL
gi|31088464|gb|CD051781.1|	gi|28195402|ref|NP_777186.1|	major allergen BDA20 [Bos taurus]	Transport	238	1.00E-61	113/172 (65%)	134/172 (77%)	AQETPAEIDPSK , KIVEGGPLR , AKGTSFT , KYQQLNSERG , PNENIE
gi|1292783|emb|Z71867.1|	gi|4138173|emb|CAA09884.1|	allergen [Malassezia sympodialis]	Unknown	202	1.00E-51	97/127 (76%)	103/127 (81%)	TGEKGFGY , QGGDFT , NGTGGKSIYG , LSMANAGPNTNGSQFFI , WLDGKHVVFG
gi|8489486|gb|AF247645.1|	gi|1170095|sp|P46419|	Glutathione S-transferase (GST class-mu)	Metabolic	184	1.00E-45	85/185 (45%)	123/185 (66%)	LGLDFPNLPY
gi|164106|gb|M63868.1|	gi|27805979|ref|NP_776803.1|	lactalbumin, alpha [Bos taurus]	Homeostasis	114	1.00E-23	50/52 (96%)	50/52 (96%)	Highly homologous
gi|39939389|gb|AY466498.1|	gi|14423698|sp|Q17284|	Fatty acid-binding protein (Allergen Blo t 13)	Metabolism	103	1.00E-21	49/125 (39%)	75/125 (60%)	STFKNTEI
|1292789|emb|Z71861.1|	gi|11135375|sp|Q9UW02|	Thioredoxin (Allergen Cop c 2)	Chaperone/stress-related	102	2.00E-21	47/105 (44%)	69/105 (65%)	ATWCGPC
gi|601887|dbj|D43752.1|	gi|1314736|gb|AAA99805.1|	Der f 3 mite allergen	Metabolic	99.4	3.00E-20	62/156 (39%)	81/156 (51%)	GDSGGP
gi|494966|emb|Z33881.1|	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	95.1	2.00E-17	46/48 (95%)	48/48 (100%)	Highly homologous
gi|31088632|gb|CD051949.1|	gi|14423714|sp|Q9U5P1|	Fatty acid-binding protein (Allergen Lep d 13)	Metabolism	81.3	1.00E-14	46/121 (38%)	74/121 (61%)	STFKNTE
gi|31088724|gb|CD052041.1|	gi|14423976|sp|Q9NAS5|	Tropomyosin (Allergen Ani s 3)	Structural	61.6	9.00E-09	30/52 (57%)	40/52 (76%)	LKEAET , AEFAERSV
gi|31088725|gb|CD052042.1|	gi|14423976|sp|Q9NAS5|	Tropomyosin (Allergen Ani s 3)	Structural	61.6	9.00E-09	30/52 (57%)	40/52 (76%)	LKEAET , AEFAERSV
gi|1292801|emb|Z71876.1|	gi|1350779|sp|P49148|	60S acidic ribosomal protein P1 (Allergen Alt a 12)	Protein synthesis	60.5	8.00E-09	33/61 (54%)	41/61 (67%)	LFAKAL
gi|31088838|gb|CD052155.1|	gi|45331208|ref|NP_987098.1|	allergen dl chain C2A [Mus musculus]	Unknown	55.8	3.00E-07	27/65 (41%)	38/65 (58%)	QDCFNE
gi|31088468|gb|CD051785.1|	gi|14424447|sp|P25780|	Mite group 1 allergen Eur m 1 precursor (Eur m I)	Metabolic	50.1	6.00E-05	20/50 (40%)	31/50 (62%)	YWIVRNSW
gi|619077|emb|Z47079.1|	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	43.5	0.002	20/20 (100%)	20/20 (100%)	VYVEELKPTPEGDLEILLQK
gi|437753|emb|Z19570.1|	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	42	0.003	19/22 (86%)	20/22 (90%)	IVTQTMKGLDIQKV
gi|12641614|emb|AJ292058.1|	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	38.5	0.072	17/19 (89%)	18/19 (94%)	IVTQTMKGLDIQK

Appendix VIII. Detail lists of sequence homology matches for dog with query, subject, subject description, functional category, bit score, E-value and region of amino acid.

Query	Hit	Hit Description	Function	Bit-score	E-value	Identity	Positivity	6 amino acid homology
gi|13124699|sp|P49822|	gi|13124699|sp|P49822|	Serum albumin precursor (Allergen Can f 3)	Homeostasis	1248	0	608/608 (100%)	608/608 (100%)	Highly homologous
gi|22531688|dbj|BAC10663.1|	gi|13124699|sp|P49822|	Serum albumin precursor (Allergen Can f 3)	Homeostasis	1246	0	606/608 (99%)	608/608 (100%)	Highly homologous
gi|6687188|emb|CAB64867.1|	gi|13124699|sp|P49822|	Serum albumin precursor (Allergen Can f 3)	Homeostasis	1246	0	606/608 (99%)	608/608 (100%)	Highly homologous
gi|3319897|emb|CAA76841.1|	gi|13124699|sp|P49822|	Serum albumin precursor (Allergen Can f 3)	Homeostasis	1193	0	579/584 (99%)	580/584 (99%)	Highly homologous
gi|32813265|dbj|BAC79353.1|	gi|729764|sp|P40918|	Heat shock 70 kDa protein (Allergen Cla h 4)	Chaperone / stress-related	916	0	458/640 (71%)	543/640 (84%)	Highly homologous
gi|32813271|dbj|BAC79356.1|	gi|729764|sp|P40918|	Heat shock 70 kDa protein (Allergen Cla h 4)	Chaperone / stress-related	916	0	458/640 (71%)	543/640 (84%)	Highly homologous
gi|17298186|dbj|BAB78505.1|	gi|729764|sp|P40918|	Heat shock 70 kDa protein (Allergen Cla h 4)	Chaperone / stress-related	904	0	455/640 (71%)	540/640 (84%)	Highly homologous
gi|2147092|pir||I46986	gi|13124699|sp|P49822|	Serum albumin precursor (Allergen Can f 3)	Homeostasis	538	e-152	264/265 (99%)	264/265 (99%)	Highly homologous
gi|22218072|dbj|BAC07513.1|	gi|45383974|ref|NP_990592.1|	preproalbumin (serum albumin) [Gallus gallus]	Homeostasis	514	e-144	245/613 (39%)	374/613 (61%)	ARRHPFLYAP , EYSRRH , YANRRPCF
gi|3121746|sp|O18874|	gi|3121746|sp|O18874|	Minor allergen Can f 2 precursor (Allergen Dog 2)	Unknown	371	e-102	180/180 (100%)	180/180 (100%)	Highly homologous
gi|29292270|emb|CAD82910.1|	gi|3121746|sp|O18874|	Minor allergen Can f 2 precursor (Allergen Dog 2)	Unknown	371	e-102	180/180 (100%)	180/180 (100%)	Highly homologous
gi|2598976|gb|AAC48795.1|	gi|3121746|sp|O18874|	alpha-2u globulin PGCL1 [Rattus norvegicus]	Homeostasis	371	e-102	180/180 (100%)	180/180 (100%)	Highly homologous
gi|29292274|emb|CAD82912.1|	gi|3121746|sp|O18874|	Minor allergen Can f 2 precursor (Allergen Dog 2).	Unknown	364	e-100	177/179 (98%)	177/179 (98%)	Highly homologous
gi|29292272|emb|CAD82911.1|	gi|3121746|sp|O18874|	Minor allergen Can f 2 precursor (Allergen Dog 2)	Unknown	358	3.00E-98	172/173 (99%)	173/173 (100%)	Highly homologous
gi|3121745|sp|O18873|	gi|3121745|sp|O18873|	Major allergen Can f 1 precursor (Allergen Dog 1).	Unknown	348	3.00E-95	174/174 (100%)	174/174 (100%)	Highly homologous
gi|29292148|emb|CAD82909.1|	gi|3121746|sp|O18874|	Major allergen Can f 1 precursor (Allergen Dog 1)	Unknown	348	2.00E-95	174/174 (100%)	174/174 (100%)	Highly homologous
gi|2598974|gb|AAC48794.1|	gi|3121745|sp|O18873|	Major allergen Can f 1 precursor (Allergen Dog 1)	Unknown	348	3.00E-95	174/174 (100%)	174/174 (100%)	Highly homologous
gi|633938|gb|AAB30434.1|	gi|13124699|sp|P49822|	Serum albumin precursor (Allergen Can f 3)	Homeostasis	296	4.00E-79	131/263 (49%)	186/263 (70%)	VHKECC , EDKEVCK , YEYSRRHPE
gi|125303|sp|P05123|	gi|25453077|sp|Q95PM9|	Arginine kinase (AK) (Allergen Plo i 1).	Metabolic	283	8.00E-75	157/341 (46%)	219/341 (64%)	DLFDPII, QQQLIDDHFLF , NKTFLVW , NEEDHLR
gi|89027|pir||A24686	gi|25453077|sp|Q95PM9|	Arginine kinase (AK) (Allergen Plo i 1)	Metabolic	275	1.00E-72	154/330 (46%)	211/330 (63%)	DLFDPII , QQQLIDDHFLF , NKTFLVW , NEEDHLR
gi|125292|sp|P05124|	gi|27463265|gb|AAO15713.1|	allergen Pen m 2 [Penaeus monodon]	Homeostasis	275	1.00E-72	154/340 (45%)	208/340 (60%)	LFDPIIED , LIDDHFLF , HNDNKTFLVW , NEEDHLR , VHIKLP
gi|320114|pir||B24686	gi|27463265|gb|AAO15713.1|	allergen Pen m 2 [Penaeus monodon]	Homeostasis	253	5.00E-66	146/340 (42%)	197/340 (57%)	LFDPIIED , LIDDHFLF , HNDNKTFLVW , NEEDHLR , VHIKLP
gi|27497538|gb|AAO13009.1|	gi|113285|sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	Metabolic	238	2.00E-61	126/303 (41%)	181/303 (59%)	CWAFSA
gi|8699209|gb|AAF78600.1|	gi|4138173|emb|CAA09884.1|	allergen [Malassezia sympodialis]	Unknown	225	3.00E-58	109/154 (70%)	121/154 (78%)	TGEKGFGY , QGGDFT , NGTGGKSIYG , LSMANAGPNTNGSQFFI , WLDGKHVVFG
gi|10185020|emb|CAC08809.1|	gi|113285|sp|P00785|	Actinidain precursor (Actinidin) (Allergen Act c 1)	Metabolic	223	7.00E-57	130/342 (38%)	192/342 (56%)	CWAFSA , WIVKNSW
gi|423167|pir||S33877	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	220	9.00E-57	100/162 (61%)	138/162 (85%)	QKVAGTW , AMAASDISLLD , APLRVY
gi|462472|sp|P33685|	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	220	9.00E-57	100/162 (61%)	138/162 (85%)	QKVAGTW , AMAASDISLLD , APLRVY
gi|448346|prf||1916447D	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	220	9.00E-57	100/162 (61%)	138/162 (85%)	QKVAGTW , AMAASDISLLD , APLRVY
gi|1082934|pir||S33878	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	213	2.00E-54	95/160 (59%)	134/160 (83%)	QKVAGTW , AMAASDISLLD , APLRVY ,
gi|462474|sp|P33686|	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	213	2.00E-54	95/160 (59%)	134/160 (83%)	QKVAGTW , AMAASDISLLD , APLRVY
gi|448347|prf||1916447E	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	213	2.00E-54	95/160 (59%)	134/160 (83%)	QKVAGTW , AMAASDISLLD , APLRVY
gi|136411|sp|P06872|	gi|2507248|sp|P49275|	Mite allergen Der f 3 precursor (Der f III)	Metabolic	189	5.00E-47	100/248 (40%)	144/248 (58%)	HFCGGS , DSCQGDSGGPVV , GIVSWGYGCA , NFVDWI
gi|67552|pir||TRDG	gi|2507248|sp|P49275|	Mite allergen Der f 3 precursor (Der f III)	Metabolic	189	5.00E-47	100/248 (40%)	144/248 (58%)	HFCGGS , DSCQGDSGGPVV , GIVSWGYGCA , NFVDWI
gi|27923790|sp|Q9N2G9|	gi|27805979|ref|NP_776803.1|	lactalbumin, alpha [Bos taurus]	Homeostasis	187	3.00E-47	83/127 (65%)	98/127 (77%)	FHTSGYDTQ , YGLFQI, WCKDDQN , KKILDK , YWLAHK , LCSEKL
gi|7959046|dbj|BAA95930.1|	gi|27805979|ref|NP_776803.1|	lactalbumin, alpha [Bos taurus]	Homeostasis	187	3.00E-47	83/127 (65%)	98/127 (77%)	FHTSGYDTQ , YGLFQI , WCKDDQN , KKILDK , YWLAHK , LCSEKL
gi|136406|sp|P06871|	gi|729315|sp|P39675|	Mite allergen Der p 3 precursor (Der p III).	Metabolic	174	2.00E-42	96/239 (40%)	133/239 (55%)	GGKDSCQGDSGGPV
gi|67551|pir||TRDGC	gi|729315|sp|P39675|	Mite allergen Der p 3 precursor (Der p III).	Metabolic	174	2.00E-42	96/239 (40%)	133/239 (55%)	GGKDSCQGDSGGPV
gi|117612|sp|P04813|	gi|729315|sp|P39675|	Mite allergen Der p 3 precursor (Der p III).	Metabolic	158	1.00E-37	84/231 (36%)	132/231 (56%)	GDSGGP , VGIVSWG
gi|108088|pir||A21195	gi|729315|sp|P39675|	Mite allergen Der p 3 precursor (Der p III).	Metabolic	158	1.00E-37	84/231 (36%)	132/231 (56%)	GDSGGP , VGIVSWG
gi|1085423|pir||S48641	gi|126608|sp|P00698|	Lysozyme C precursor	Metabolic	157	4.00E-38	69/130 (53%)	90/130 (69%)	ESNFNTQA
gi|20150098|pdb|1I56|A	gi|126608|sp|P00698|	Lysozyme C precursor	Metabolic	157	4.00E-38	69/130 (53%)	90/130 (69%)	ESNFNTQA
gi|9257149|pdb|1QQY|A	gi|126608|sp|P00698|	Lysozyme C precursor	Metabolic	157	4.00E-38	69/130 (53%)	90/130 (69%)	ESNFNTQA
gi|8928188|sp|P81708|	gi|126608|sp|P00698|	Lysozyme C precursor	Metabolic	157	4.00E-38	69/130 (53%)	90/130 (69%)	ESNFNTQA
gi|13787135|pdb|1EL1|B	gi|126608|sp|P00698|	Lysozyme C precursor	Metabolic	157	4.00E-38	69/130 (53%)	90/130 (69%)	ESNFNTQA
gi|684994|gb|AAB31794.1|	gi|126608|sp|P00698|	Lysozyme C precursor IV)	Metabolic	157	4.00E-38	69/130 (53%)	90/130 (69%)	ESNFNTQA
gi|4454073|emb|CAA05126.1|	gi|1314736|gb|AAA99805.1|	Der f 3 mite allergen	Metabolic	145	1.00E-33	87/240 (36%)	128/240 (53%)	LTAAHC , GDSGGP
gi|482952|pir||B32410	gi|14423685|sp|O97370|	Mite allergen Eur m 3 precursor.	Metabolic	139	9.00E-32	95/252 (37%)	135/252 (52%)	LTAAHCV , VSWGYGC
gi|136423|sp|P19236|	gi|14423685|sp|O97370|	Mite allergen Eur m 3 precursor	Metabolic	139	9.00E-32	95/252 (37%)	135/252 (52%)	LTAAHCV , VSWGYGC
gi|3318843|pdb|1RP1|	gi|1709545|sp|P51528|	Phospholipase A1 (Allergen Ves m 1)	Metabolic	112	1.00E-23	79/259 (30%)	116/259 (44%)	LIGHSLGAHV

Appendix VIII. Detail lists of sequence homology matches for dog with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology (cont.).

Query	Hit	Hit Description	Function	Bit-score	E-value	Identity	Positivity	6 amino acid homology
gi|126316|sp|P06857|	gi|1709545|sp|P51528|	Phospholipase A1 (Allergen Ves m 1)	Metabolic	112	2.00E-23	80/261 (30%)	117/261 (44%)	LIGHSLGAHV
gi|164048|gb|AAA30885.1|	gi|1709545|sp|P51528|	Phospholipase A1 (Allergen Ves m 1)	Metabolic	110	6.00E-23	79/261 (30%)	117/261 (44%)	LIGHSLGAHV
gi|67162|pir||LIDG	gi|1709545|sp|P51528|	Phospholipase A1 (Allergen Ves m 1)	Metabolic	110	6.00E-23	79/261 (30%)	117/261 (44%)	LIGHSLGAHV
gi|124847|sp|P01002|	gi|124757|sp|P01005|I	Ovomucoid precursor (Allergen Gal d 1)	Homeostasis	102	2.00E-21	50/119 (42%)	68/119 (57%)	RPLCGSD
gi|476547|pir||TIDGS	gi|124757|sp|P01005|I	Ovomucoid precursor (Allergen Gal d 1)	Homeostasis	102	2.00E-21	50/119 (42%)	68/119 (57%)	RPLCGSD
gi|163944|gb|AAA30840.1|	gi|1709545|sp|P51528|	Phospholipase A1 (Allergen Ves m 1)	Metabolic	99.4	1.00E-19	74/238 (31%)	104/238 (43%)	LIGHSLGAHV
gi|3914452|sp|Q28278|	gi|1314736|gb|AAA99805.1|	Der f 3 mite allergen	Metabolic	95.1	6.00E-19	58/155 (37%)	81/155 (51%)	GDSGGP
gi|1304048|dbj|BAA07808.1|	gi|1314736|gb|AAA99805.1|	Der f 3 mite allergen	Metabolic	95.1	6.00E-19	58/155 (37%)	81/155 (51%)	GDSGGP
gi|38374003|gb|AAR19224.1|	gi|1314736|gb|AAA99805.1|	Der f 3 mite allergen	Metabolic	87.4	1.00E-16	52/171 (30%)	87/171 (50%)	DSCQGDSGGP
gi|543068|pir||A48292	gi|37078092|sp|Q870B9|	Enolase	Metabolic	85.5	7.00E-15	48/84 (57%)	60/84 (71%)	VPSGASTG , TIAPALI
gi|402558|emb|CAA48914.1|	gi|37078092|sp|Q870B9|	Enolase	Metabolic	85.5	7.00E-15	48/84 (57%)	60/84 (71%)	VPSGASTG , TIAPALI
gi|3288714|dbj|BAA31256.1|	gi|729970|sp|P39674|	Allergen MAG29	Chaperone / stress-related	71.2	5.00E-12	41/67 (61%)	42/67 (62%)	VCNPIIT , SGPTIEEVD
gi|229545|prf||753699A	gi|124757|sp|P01005|	Ovomucoid precursor (Allergen Gal d 1)	Homeostasis	68.2	4.00E-11	38/135 (28%)	63/135 (46%)	PICGTD
gi|423181|pir||S29749	gi|13124699|sp|P49822|	Serum albumin precursor (Allergen Can f 3)	Homeostasis	53.5	1.00E-06	24/24 (100%)	24/24 (100%)	Highly homologous
gp-4	gi|124757|sp|P01005|	Ovomucoid precursor (Allergen Gal d 1)	Homeostasis	52	4.00E-06	19/41 (46%)	27/41 (65%)	NECLLC
gi|68720|pir||TIDGA	gi|124757|sp|P01005|	Ovomucoid precursor (Allergen Gal d 1)	Homeostasis	52	4.00E-06	19/41 (46%)	27/41 (65%)	NECLLC
gi|1683343|gb|AH004615.1|	gi|13124699|sp|P49822|	Serum albumin precursor (Allergen Can f 3).	Homeostasis	50.4	1.00E-04	24/28 (85%)	24/28 (85%)	NPGFPPLVAPEPDAL
gi|693831|gb|AAB32129.1|	gi|13124699|sp|P49822|	Serum albumin precursor (Allergen Can f 3).	Homeostasis	50.4	1.00E-05	24/28 (85%)	24/28 (85%)	NPGFPPLVAPEPDAL
gi|5738962|dbj|BAA83419.1|	gi|27807077|ref|NP_777021.1|	S100 calcium-binding protein A7 (psoriasin 1)	Signal transduction	46.2	2.00E-04	23/81 (28%)	43/81 (53%)	FHKYSG
gi|3822535|gb|AAC69882.1|	gi|1352699|sp|P49369|	Phospholipase A1 precursor (Allergen Ves v 1)	Metabolic	42.7	0.002	21/42 (50%)	28/42 (66%)	RNTRLVGQ
gi|2952306|gb|AAC05499.1|	gi|42559584|sp|Q23939|	Tropomyosin (Allergen Der f 10)	Structural	40.8	0.008	20/28 (71%)	21/28 (74%)	ELEEEL
gi|693830|gb|AAB32128.1|	gi|13124699|sp|P49822|	Serum albumin precursor (Allergen Can f 3).	Homeostasis	39.3	0.022	17/17 (100%)	17/17 (100%)	EAYKSEIAHRYNDLGEE
gi|693833|gb|AAB32131.1|	|13124699|sp|P49822|	Serum albumin precursor (Allergen Can f 3).	Homeostasis	38.1	0.05	17/17 (100%)	17/17 (100%)	VLDEFKPLVDEPQNLVK
gi|238998|gb|AAB20343.1|	gi|1170095|sp|P46419|	Glutathione S-transferase (GST class-mu)	Metabolic	34.7	0.67	14/21 (66%)	16/21 (76%)	GYWDIRG

Appendix IX. Detail lists of sequence homology matches for cat with query, subject, subject description, functional category, bit score, E-value and region of amino acid homology.

Query	Hit	Hit Description	Function	Bit-score	E-value	Identity	Positivity	6 amino acid homology
gi|1351908|sp|P49064|	gi|1351908|sp|P49064|	Serum albumin precursor (Allergen Fel d 2)	Homeostasis	1194	0	585/600 (97%)	585/600 (97%)	Highly homologous
gi|1363007|pir||S57632	gi|1351908|sp|P49064|	Serum albumin precursor (Allergen Fel d 2)	Homeostasis	1194	0	585/600 (97%)	585/600 (97%)	Highly homologous
gi|886485|emb|CAA59279.1|	gi|1351908|sp|P49064|	Serum albumin precursor (Allergen Fel d 2)	Homeostasis	1194	0	585/600 (97%)	585/600 (97%)	Highly homologous
gi|30962111|emb|CAD32275.1|	gi|1351908|sp|P49064|	Serum albumin precursor (Allergen Fel d 2)	Homeostasis	1139	0	555/576 (96%)	557/576 (96%)	Highly homologous
gi|6708502|gb|AAD09454.2|	gi|42559536|sp|Q9NJA9|	Paramyosin (Allergen Ani s 2)	Structural	593	e-168	317/828 (38%)	521/828 (62%)	RLQAEN
gi|38492848|pdb|1PUO|B	gi|38492847|pdb|1PUO|A[38492847]	Chain A, Crystal Structure Of Fel D 1	Unknown	325	2.00E-88	164/164 (100%)	164/164 (100%)	Highly homologous
gi|38492847|pdb|1PUO|A	gi|38492847|pdb|1PUO|A[38492847]	Chain A, Crystal Structure Of Fel D 1	Unknown	325	2.00E-88	164/164 (100%)	164/164 (100%)	Highly homologous
gi|25989117|gb|AAK33125.1|	gi|4138173|emb|CAA09884.1|	allergen [Malassezia sympodialis]	Unknown	229	2.00E-59	112/157 (71%)	122/157 (77%)	TGEKGFGY , QGGDFT , NGTGGKSIYG , LSMANAGPNTNGSQFFI , WLDGKHVVFG
gi|423188|pir||S33876	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	219	2.00E-56	101/160 (63%)	132/160 (82%)	QKVAGTW , AMAASDISLLD , APLRVYV , VLDTDY
gi|462475|sp|P33688|	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	219	2.00E-56	101/160 (63%)	132/160 (82%)	QKVAGTW , AMAASDISLLD , APLRVYV , VLDTDY
gi|448345|prf||1916447C	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	219	2.00E-56	101/160 (63%)	132/160 (82%)	QKVAGTW , AMAASDISLLD , APLRVYV , VLDTDY
gi|462473|sp|P33687|	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	218	4.00E-56	98/160 (61%)	133/160 (83%)	AMAASDISLLD , APLRVYV , VLDTDY
gi|448343|prf||1916447A	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	218	4.00E-56	98/160 (61%)	133/160 (83%)	AMAASDISLLD , APLRVYV , VLDTDY
gi|4322134|gb|AAD15975.1|	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	214	4.00E-55	97/154 (62%)	130/154 (84%)	QKVAGTW , AMAASDISLLD , APLRVYV , VLDTDY
gi|2119651|pir||S33875	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	211	3.00E-54	97/160 (60%)	132/160 (82%)	AMAASDISLLD , APLRVYV , VLDTDY
gi|4322130|gb|AAD15971.1|	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	211	4.00E-54	94/152 (61%)	128/152 (84%)	AMAASDISLLD , APLRVYV , VLDTDY
gi|4322131|gb|AAD15972.1|	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	207	6.00E-53	93/152 (61%)	127/152 (83%)	AMAASDISLLD , APLRVYV , VLDTDY
gi|232086|sp|P30440|	gi|232086|sp|P30440|	Major allergen I polypeptide chain 2 precursor	Unknown	184	5.00E-46	94/109 (86%)	94/109 (86%)	Highly homologous
gi|1082946|pir||C56413	gi|232086|sp|P30440|	Major allergen I polypeptide chain 2 precursor	Unknown	184	5.00E-46	94/109 (86%)	94/109 (86%)	Highly homologous
gi|163823|gb|AAC41616.1|	gi|232086|sp|P30440|	Major allergen I polypeptide chain 2 precursor	Unknown	184	5.00E-46	94/109 (86%)	94/109 (86%)	Highly homologous
gi|1169666|sp|P30439|	gi|1169666|sp|P30439|	Major allergen I polypeptide chain 1 minor form precursor	Unknown	181	4.00E-45	88/88 (100%)	88/88 (100%)	Highly homologous
gi|395407|emb|CAA44345.1|	gi|423192|pir||JC1127[423192]	major allergen chain 2 precursor, short form	Unknown	181	3.00E-45	92/107 (85%)	92/107 (85%)	Highly homologous
gi|423192|pir||JC1127	gi|423192|pir||JC1127[423192]	major allergen chain 2 precursor, short form	Unknown	181	3.00E-45	92/107 (85%)	92/107 (85%)	Highly homologous
gi|423191|pir||JC1126	gi|423191|pir||JC1126[423191]	major allergen chain 1 precursor B	Unknown	181	3.00E-45	88/88 (100%)	88/88 (100%)	Highly homologous
gi|108170|pir||S14719	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	181	7.00E-45	88/158 (55%)	120/158 (75%)	AMAASDISLLD , APLRVYV , VLDTDY
gi|125905|sp|P21664|	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	181	7.00E-45	88/158 (55%)	120/158 (75%)	AMAASDISLLD , APLRVYV , VLDTDY
gi|163827|gb|AAC41617.1|	gi|163827|gb|AAC41617.1|[163827]	major allergen I	Unknown	181	4.00E-45	88/88 (100%)	88/88 (100%)	Highly homologous
gi|4322133|gb|AAD15974.1|	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	177	6.00E-44	86/154 (55%)	118/154 (76%)	AMAASDISLLD , APLRVYV , VLDTDY
gi|448344|prf||1916447B	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	177	7.00E-44	87/158 (55%)	119/158 (75%)	AMAASDISLLD , APLRVYV , VLDTDY
gi|4322132|gb|AAD15973.1|	gi|125910|sp|P02754|	Beta-lactoglobulin precursor (Beta-LG)	Homeostasis	175	2.00E-43	85/154 (55%)	117/154 (75%)	AMAASDISLLD , APLRVYV , VLDTDY
gi|1082944|pir||A56413	gi|1082944|pir||A56413[1082944]	major allergen Fel dl chain 1 long form precursor	Unknown	165	3.00E-40	81/92 (88%)	81/92 (88%)	Highly homologous
gi|163825|gb|AAC37318.1|	gi|1082944|pir||A56413[1082944]	major allergen Fel dl chain 1 long form precursor	Unknown	165	3.00E-40	81/92 (88%)	81/92 (88%)	Highly homologous
gi|1169665|sp|P30438|	gi|1169665|sp|P30438|	Major allergen I polypeptide chain 1 major form precursor	Unknown	164	6.00E-40	81/92 (88%)	81/92 (88%)	Highly homologous
gi|423190|pir||JC1136	gi|423190|pir||JC1136[423190]	major allergen chain 1 precursor A	Unknown	164	4.00E-40	81/92 (88%)	81/92 (88%)	Highly homologous
gi|1364212|emb|CAA44343.1|	gi|423191|pir||JC1126[423191]	major allergen chain 1 precursor B	Unknown	150	8.00E-36	75/85 (88%)	77/85 (90%)	Highly homologous
gi|1364213|emb|CAA44344.1|	gi|423190|pir||JC1136[423190]	major allergen chain 1 precursor A	Unknown	149	1.00E-35	75/90 (83%)	77/90 (85%)	Highly homologous
gi|1708844|sp|P55031|	gi|548449|sp|Q06478|	Phospholipase A1 1 precursor (Allergen Dol m 1.01)	Metabolic	111	3.00E-23	77/276 (27%)	126/276 (45%)	GLDPAGP ,
gi|1150861|gb|AAB03848.1|	gi|548449|sp|Q06478|	Phospholipase A1 1 precursor (Allergen Dol m 1.01)	Metabolic	111	3.00E-23	77/276 (27%)	126/276 (45%)	GLDPAGP
gi|3914455|sp|Q28412|	gi|1314736|gb|AAA99805.1|[1314736]	Der f 3 mite allergen	Metabolic	94	1.00E-18	55/156 (35%)	80/156 (51%)	GDSGGP
gi|1304038|dbj|BAA07807.1|	gi|1314736|gb|AAA99805.1|[1314736]	Der f 3 mite allergen	Metabolic	94	1.00E-18	55/156 (35%)	80/156 (51%)	GDSGGP
gi|124848|sp|P08480|	gi|124757|sp|P01005|I	Ovomucoid precursor (Allergen Gal d 1)	Homeostasis	93.2	1.00E-18	46/119 (38%)	64/119 (53%)	PLCGSD
gi|89087|pir||A29654	gi|124757|sp|P01005|	Ovomucoid precursor (Allergen Gal d 1)	Homeostasis	93.2	1.00E-18	46/119 (38%)	64/119 (53%)	PLCGSD
gi|131098|sp|P80079|	gi|18281421|sp|Q91483|	Parvalbumin beta 2 (Major allergen Sal s 1)	Signal-transduction	92	2.00E-18	47/101 (46%)	62/101 (61%)	DKSGFIEEDEL
gi|7441490|pir||S27209	gi|18281421|sp|Q91483|	Parvalbumin beta 2 (Major allergen Sal s 1)	Signal-transduction	92	2.00E-18	47/101 (46%)	62/101 (61%)	DKSGFIEEDEL
gi|539716|pir||B53283	gi|539716|pir||B53283[539716]	major cat allergen Fel d I beta chain	Unknown	44.3	9.00E-04	20/20 (100%)	20/20 (100%)	VKMAETFPIFYDVFTAVANG

Appendix X. List of putative allergens predicted in pork using wavelet transform.

Subject (hits)	Subject Description	Motif no.	E-values
P29700	Alpha-2-HS-glycoprotein precursor (Fetuin-A).	4	0.015
P18137	Alpha-lactalbumin precursor (Lactose synthase B protein).	17	9.00E-23
Q28944	Cathepsin L precursor.	6	0.03
P79401	Cytochrome P450 3A29 (CYPIIIA29).	13	0.00087
Q8WNV7	Dehydrogenase/reductase SDR family member 4	24	0.04
O02772	Fatty acid-binding protein, heart (H-FABP) (Heart-type fatty acid-binding protein).	8	0.0024
P10173	Fumarate hydratase, mitochondrial (Fumarase).	52	0.00053
P34930	HEAT SHOCK 70 KD PROTEIN 1 (HSP70.1).	15	6.70E-06
P34934	HEAT SHOCK 70 KD PROTEIN.	13	0.00031
Q04967	HEAT SHOCK 70 KDA PROTEIN 6 (HEAT SHOCK 70 KDA PROTEIN B').	15	0.00014
O02705	Heat shock protein HSP 90-alpha (HSP 86).	14	0.00049
Q29545	Inhibitor of carbonic anhydrase precursor.	27	0.0044
P14632	Lactotransferrin precursor (Lactoferrin).	28	0.0025
P00346	Malate dehydrogenase, mitochondrial precursor.	45	0.00016
Q95250	Membrane associated progesterone receptor component 1.	13	0.043
P79293	Myosin heavy chain, cardiac muscle beta isoform (MyHC-beta).	13	2.40E-05
Q29558	NADP-dependent malic enzyme (NADP-ME) (Malic enzyme 1).	34	0.00088
P00791	Pepsin A precursor.	25	0.0003
P09571	Serotransferrin (Transferrin) (Siderophilin) (Beta-1-metal binding globulin).	28	0.016
P08835	Serum albumin precursor.	34	1.80E-10
P42639	Tropomyosin 1 alpha chain (Alpha-tropomyosin).	26	1.50E-07
P53715	Von Ebner's gland protein precursor (VEG protein) (Tear prealbumin) (TP) (Tear lipocalin) (Lipocalin-	46	3.80E-05

Appendix XI. List of putative allergens predicted in chicken using wavelet transform.

Subject (hits)	Subject Description	Motif no.	E-values
P01005	Ovomucoid precursor (Allergen Gal d 1) (Gal d I).	55	5.50E-10
P79781	40S ribosomal protein S27a.	5	0.049
P27463	Aldehyde dehydrogenase 1A1 (Aldehyde dehydrogenase, cytosolic)	28	2.80E-05
O93344	Aldehyde dehydrogenase 1A2 (Retinaldehyde-specific dehydrogenase type 2) (RALDH(II)) (RALDH-2).	28	3.90E-05
P51913	Alpha enolase (2-phospho-D-glycerate hydro-lyase) (Phosphopyruvate hydratase).	42	4.20E-08
P07322	Beta enolase (2-phospho-D-glycerate hydro-lyase) (Phosphopyruvate hydratase).	42	4.20E-08
Q9DE13	Bromodomain adjacent to zinc finger domain 2B (Extracellular matrix protein F22).	33	0.0034
P07090	Calretinin (CR).	24	0.34
Q90640	Chromosome-associated kinesin KIF4A (Chromokinesin).	31	0.059
O42130	DNA topoisomerase II, alpha isozyme.	61	0.037
Q05423	Fatty acid-binding protein, retina (R-FABP).	8	1.30E-05
O57391	Gamma enolase (2-phospho-D-glycerate hydro-lyase) (Neural enolase) (NSE).	42	1.00E-08
P01013	GENE X PROTEIN (OVALBUMIN-RELATED).	34	1.30E-10
P01014	Gene Y protein (Ovalbumin-related).	34	3.30E-15
P08106	HEAT SHOCK 70 KD PROTEIN (HSP70).	15	4.20E-06
Q04619	Heat shock cognate protein HSP 90-beta.	14	0.022
P11501	Heat shock protein HSP 90-alpha.	14	0.00047
O93532	Keratin, type II cytoskeletal cochleal (Cytokeratin otokeratin).	55	5.00E-11
P29616	Myosin heavy chain, cardiac muscle isoform.	13	0.012
P02565	Myosin heavy chain, fast skeletal muscle, embryonic.	13	3.90E-07
P13538	Myosin heavy chain, skeletal muscle, adult.	13	0.005
P01012	Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II).	48	5.50E-10
P02789	Ovotransferrin precursor (Conalbumin) (Allergen Gal d 3) (Gal d III) (Serum transferrin).	27	1.90E-12
P80026	Parvalbumin, muscle.	45	0.016
P19753	Parvalbumin, thymic (Avian thymic hormone) (ATH) (Thymus-specific antigen T1).	45	7.60E-06
P43305	Parvalbumin, thymic CPV3 (Parvalbumin 3).	45	0.00021
P00793	Pepsin A precursor.	25	0.00043
P24367	Peptidyl-prolyl cis-trans isomerase B precursor (PPIase) (Rotamase) (Cyclophilin B) (S-cyclophilin) (SCYLP).	60	0.00028
O42184	Restin (Cytoplasmic linker protein-170) (CLIP-170).	13	0.057
P19121	Serum albumin precursor (Alpha-livetin) (Allergen Gal d 5).	9	3.80E-21
P00940	Triosephosphate isomerase (TIM) (Triose-phosphate isomerase).	7	0.0019
P04268	Tropomyosin 1 alpha chain (Alpha-tropomyosin).	26	1.10E-06
P19352	Tropomyosin beta chain (Tropomyosin 2) (Beta-tropomyosin).	26	4.10E-07
P02845	Vitellogenin II precursor (Major vitellogenin) [Contains: Lipovitellin I (LVI); Phosvitin (PV); Lipovitellin II (LVII);YG	57	5.30E-12

Appendix XII. List of putative allergens predicted in sheep using wavelet transform.

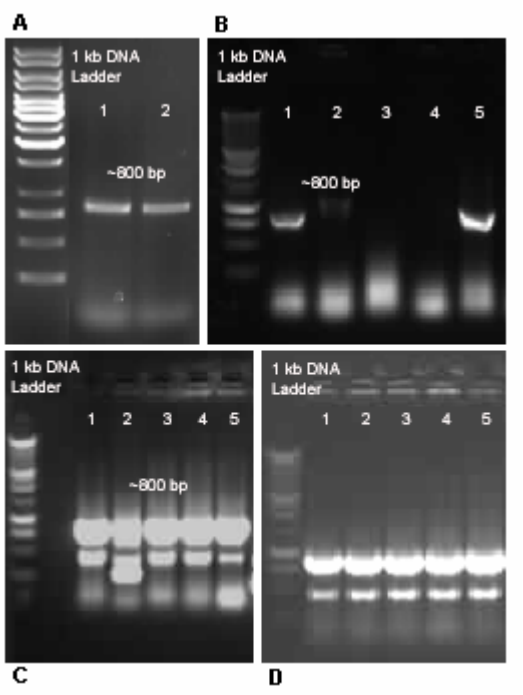
Subject (hits)	Subject Description	Motif no.	E-values
P51977	Aldehyde dehydrogenase 1A1 (Aldehyde dehydrogenase, cytosolic) (ALDH class 1) (ALHDII) (ALDH-E1).	28	0.0031
P09462	Alpha-lactalbumin precursor (Lactose synthase B protein).	17	4.20E-36
P04653	Alpha-S1 casein precursor.	2	0.005
P04654	Alpha-S2 casein precursor.	16	0.0011
P11839	Beta casein precursor.	12	0.013
P02757	Beta-lactoglobulin 1/B, 2/A, and 3/C precursor.	14	6.30E-06
Q06435	Interleukin-3 precursor (IL-3) (Multipotential colony-stimulating factor) (Hematopoietic growth factor) (P-cell stimu	29	0.041
P02539	Keratin, type II microfibrillar (Low-sulfur keratin).	55	4.30E-05
P25691	Keratin, type II microfibrillar, component 5.	55	6.50E-06
P15241	Keratin, type II microfibrillar, component 7C.	55	6.20E-06
P14639	Serum albumin precursor.	34	1.20E-09
P22793	Trichohyalin.	13	0.00089

Appendix XIII. List of putative allergens predicted in dog using wavelet transform.

Subject (hits)	Subject Description	Motif no.	E-values
Q9N2G9	Alpha-lactalbumin precursor (Lactose synthase B protein)	17	3.00E-25
P24643	Calnexin precursor (pp90)	41	0.026
O18873	Major allergen Can f 1 precursor (Allergen Dog 1)	46	8.30E-15
Q28269	Occludin	36	0.013
P49822	Serum albumin precursor (Allergen Can f 3)	34	1.20E-09
Q9XT60	Sex-determining region Y protein (Testis-determining factor)	36	0.079

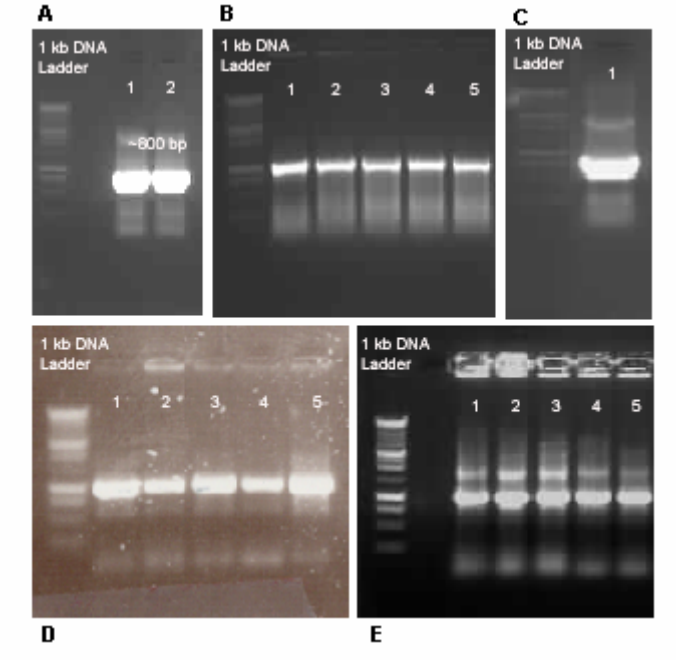
Appendix XIV. List of putative allergens predicted in cat using wavelet transform.

Subject (hits)	Subject Description	Motif no.	E-values
P30439	Major allergen I polypeptide chain 1 minor form precursor (Allergen Fel d 1-A)	44	1.40E-09
P30438	Major allergen I polypeptide chain 1 precursor (Allergen Fel d 1-A)	44	1.40E-09
P30440	Major allergen I polypeptide chain 2 precursor (Allergen Fel d 1-B)	10	2.00E-06
P49064	Serum albumin precursor (Allergen Fel d 2)	34	3.80E-13



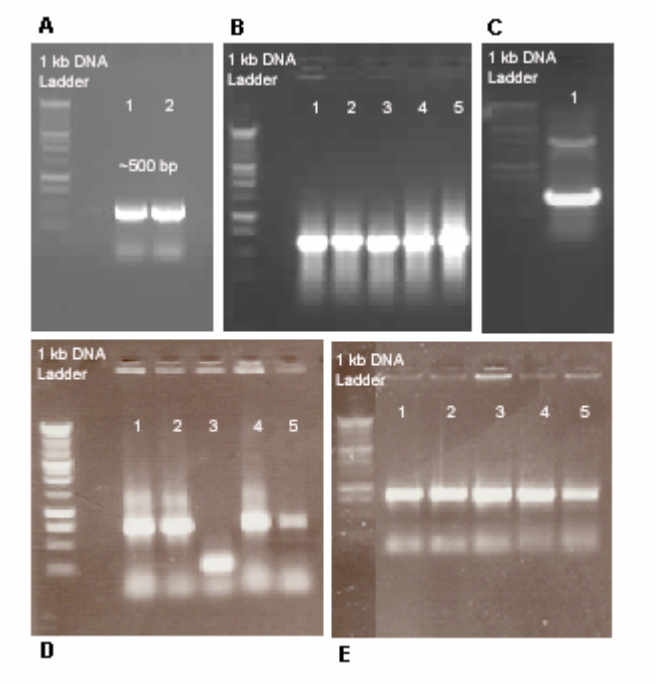
Appendix XV.

Molecular cloning of Tropo 3. (A) PCR amplification of target tropomyosin gene with gene specific primers using Expand long-template *Taq* DNA polymerase. Amplicons in Lane 1 and 2 corresponds to correct expected size of ~800 bp. PCR amplicons from both lanes were extracted and purified using QIA quick Gel extraction Kit (Qiagen) and ligated to pGEMT-Easy vector followed by transformation into XL-1 blue non-expression host. **(B) Colony screening of PCR inserts in pGEMT vector using SP6 and T7 primers:** A total of 10 colonies were screened for insert. Only five lanes were showed here (Lane: 1 to 5). Only 4 out of 10 clones showed the presence of insert with expected size of ~800 bp. **(C): Colony screening of pET32a ligated insert** in transformed XL 1-blue non-expression host strain using LIC primers. Lane 1 to 5 corresponds to 5 clones chosen with the correct size of insert. **(D): Sub cloning of ligated Pet Vector plasmid into BL-21 (DE3) (Novagen, USA) expression host**. Again, colony screening was performed (lane 1 to 5). The clones with the correct size insert were subsequently sequenced from both ends to check for correct reading frame. Clone 1 and 2 were selected for protein expression. Glycerol stocks were made from those clones that were used for expression.



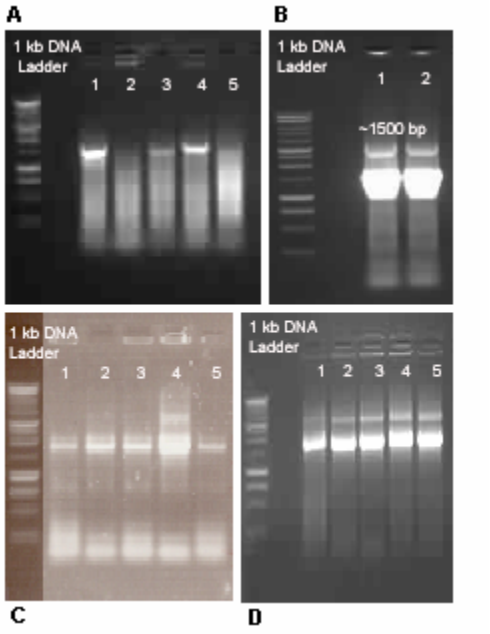
Appendix XVI.

Molecular cloning of TRNT. (A) PCR amplification of target troponin gene with gene specific primers using Expand long-template *Taq* DNA polymerase. Amplicons in Lane 1 and 2 corresponds to correct expected size of ~800 bp. PCR amplicons from both lanes were extracted and purified using QIA quick Gel extraction Kit (Qiagen) and ligated to pGEMT-Easy vector followed by transformation into XL-1 blue non-expression host. **(B) Colony screening of PCR inserts in pGEMT vector using SP6 and T7 primers:** All five colonies showed the presence of insert with expected size of ~800 bp. **(C): PCR amplification of target gene from pGEM-T plasmids with correct insert using designed LIC primer adaptors.** Purified pGEM-T plasmid from clone 1 (Lane 1 of Fig B) was used. The PCR amplified band was gel extracted, and purified. The final digested DNA fragment was ligated into pET-32a (+) expression vector (Novagen, USA) using T4 DNA ligase (Invitrogen, USA) and transformed into XL1-Blue non-expression host cell. **(D): Colony screening of pET32a ligated insert** in transformed XL 1-blue non-expression host strain using LIC primers. Lane 1 to 5 corresponds to 5 clones chosen with the correct size of insert. **(E): Sub cloning of ligated Pet Vector plasmid into BL-21 (DE3) (Novagen, USA) expression host**. Again, colony screening was performed (lane 1 to 5). The clones with the correct size insert were subsequently sequenced from both ends to check for correct reading frame. Clone 1 and 2 were selected for protein expression. Glycerol stocks were made from those clones that were used for expression.



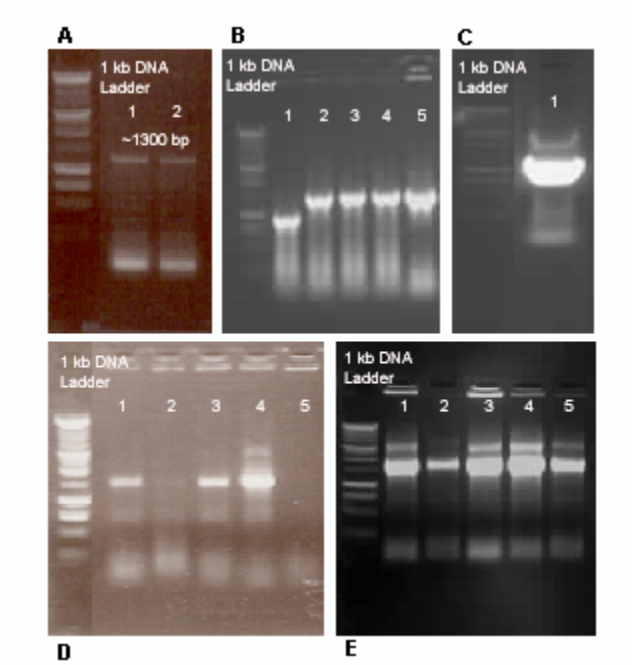
Appendix XVII.

Molecular cloning of Myo_L. (A) PCR amplification of target myosin light chain gene with gene specific primers using Expand long-template *Taq* DNA polymerase. Amplicons in Lane 1 and 2 corresponds to correct expected size of ~500 bp. PCR amplicons from both lanes were extracted and purified using QIA quick Gel extraction Kit (Qiagen) and ligated to pGEMT-Easy vector followed by transformation into XL-1 blue non-expression host. **(B) Colony screening of PCR inserts in pGEMT vector using SP6 and T7 primers:** All five colonies showed the presence of insert with expected size of ~500 bp. **(C): PCR amplification of target gene from pGEM-T plasmids with correct insert using designed LIC primer adaptors.** Purified pGEM-T plasmid from clone 1 (Lane 1 of Fig B) was used. The PCR amplified band was gel extracted, and purified. The final digested DNA fragment was ligated into pET-32a (+) expression vector (Novagen, USA) using T4 DNA ligase (Invitrogen, USA) and transformed into XL1-Blue non-expression host cell. **(D): Colony screening of pET32a ligated insert** in transformed XL 1-blue non-expression host strain using LIC primers. Four of the five clones chosen have the correct size insert. **(E): Sub cloning of ligated Pet Vector plasmid into BL-21 (DE3) (Novagen, USA) expression host**. Again, colony screening was performed (lane 1 to 5). The clones with the correct size insert were subsequently sequenced from both ends to check for correct reading frame. Clone 1 and 2 were selected for protein expression. Glycerol stocks were made from those clones that were used for expression.



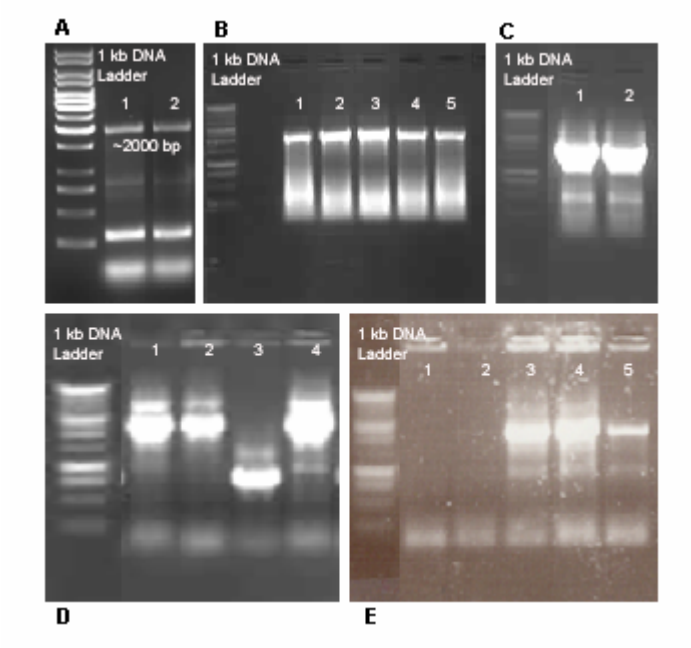
Appendix XVIII.

Molecular cloning of ADH. (A) Colony screening of PCR inserts in pGEMT vector using SP6 and T7 primers: Three out of five clones showed the presence of insert with expected size of ~1500 bp. **(B): PCR amplification of target gene from pGEM-T plasmids with correct insert using designed LIC primer adaptors.** Purified pGEM-T plasmid from clone 1 (Lane 1 of Fig A) was used. The PCR amplified band was gel extracted, and purified. The final digested DNA fragment was ligated into pET-32a (+) expression vector (Novagen, USA) using T4 DNA ligase (Invitrogen, USA) and transformed into XL1-Blue non-expression host cell. **(C): Colony screening of pET32a ligated insert** in transformed XL 1-blue non-expression host strain using LIC primers. Lane 1 to 5 corresponds to 5 clones chosen with the correct size of insert. **(D): Sub cloning of ligated Pet Vector plasmid into BL-21 (DE3) (Novagen, USA) expression host**. Again, colony screening was performed (lane 1 to 5). The clones with the correct size insert were subsequently sequenced from both ends to check for correct reading frame. Clone 1 and 2 were selected for protein expression. Glycerol stocks were made from those clones that were used for expression.



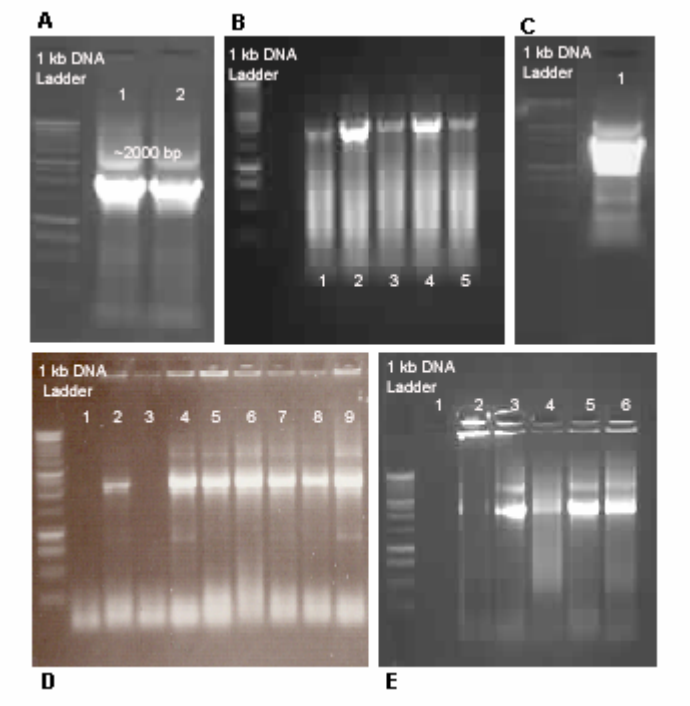
Appendix XIX.

Molecular cloning of ENO 1. (A) PCR amplification of target enolase 1 gene with gene specific primers using Expand long-template *Taq* DNA polymerase. Amplicons in Lane 1 and 2 corresponds to correct expected size of ~1300 bp. PCR amplicons from both lanes were extracted and purified using QIA quick Gel extraction Kit (Qiagen) and ligated to pGEMT-Easy vector followed by transformation into XL-1 blue non-expression host. **(B) Colony screening of PCR inserts in pGEMT vector using SP6 and T7 primers:** Four out of five clones showed the presence of insert with expected size of ~900 bp. **(C): PCR amplification of target gene from pGEM-T plasmids with correct insert using designed LIC primer adaptors.** Purified pGEM-T plasmid from clone 1 (Lane 2 of Fig B) was used. The PCR amplified band was gel extracted, and purified. The final digested DNA fragment was ligated into pET-32a (+) expression vector (Novagen, USA) using T4 DNA ligase (Invitrogen, USA) and transformed into XL1-Blue non-expression host cell. **(D): Colony screening of pET32a ligated insert** in transformed XL 1-blue non-expression host strain using LIC primers. Three of the five clones chosen have the correct size insert. **(E): Sub cloning of ligated Pet Vector plasmid into BL-21 (DE3) (Novagen, USA) expression host**. Again, colony screening was performed (lane 1 to 5). The clones with the correct size insert were subsequently sequenced from both ends to check for correct reading frame. Clone 1 and 3 were selected for protein expression. Glycerol stocks were made from those clones that were used for expression.



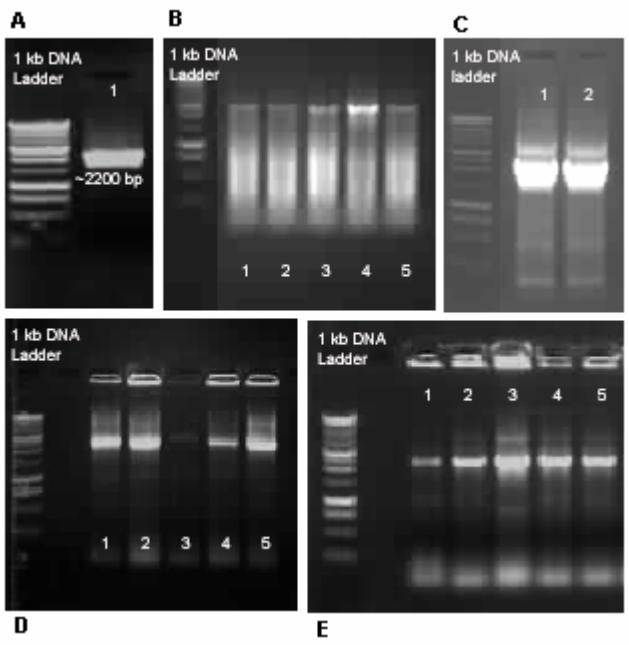
Appendix XX.

Molecular cloning of pHSP70. (A) PCR amplification of target heat shock protein 70 gene with gene specific primers using Expand long-template *Taq* DNA polymerase. Amplicons in Lane 1 and 2 corresponds to correct expected size of ~2000 bp. PCR amplicons from both lanes were extracted and purified using QIA quick Gel extraction Kit (Qiagen) and ligated to pGEMT-Easy vector followed by transformation into XL-1 blue non-expression host. **(B) Colony screening of PCR inserts in pGEMT vector using SP6 and T7 primers:** All five colonies showed the presence of insert with expected size of ~2000 bp. **(C) PCR amplification of target gene from pGEM-T plasmids with correct insert using designed LIC primer adaptors.** Purified pGEM-T plasmid from clone 1 (Lane 1 of Fig B) was used. The PCR amplified band was gel extracted, and purified. The final digested DNA fragment was ligated into pET-32a (+) expression vector (Novagen, USA) using T4 DNA ligase (Invitrogen, USA) and transformed into XL1-Blue non-expression host cell. **(D): Colony screening of pET32a ligated insert** in transformed XL 1-blue non-expression host strain using LIC primers. Three of the four clones chosen have the correct size insert. **(D): Sub cloning of ligated Pet Vector plasmid into BL-21 (DE3) (Novagen, USA) expression host**. Again, colony screening was performed (lane 1 to 5). The clones with the correct size insert were subsequently sequenced from both ends to check for correct reading frame. Clone 3 and 4 were selected for protein expression. Glycerol stocks were made from those clones that were used for expression.



Appendix XXI.

Molecular cloning of bHSP70. (A) PCR amplification of target heat shock protein 70 gene with gene specific primers using Expand long-template *Taq* DNA polymerase. Amplicons in Lane 1 and 2 corresponds to correct expected size of ~2000 bp. PCR amplicons from both lanes were extracted and purified using QIA quick Gel extraction Kit (Qiagen) and ligated to pGEMT-Easy vector followed by transformation into XL-1 blue non-expression host. **(B) Colony screening of PCR inserts in pGEMT vector using SP6 and T7 primers:** All five colonies showed the presence of insert with expected size of ~2000 bp. **(C) PCR amplification of target gene from pGEM-T plasmids with correct insert using designed LIC primer adaptors.** Purified pGEM-T plasmid from clone 1 (Lane 2 of Fig B) was used. The PCR amplified band was gel extracted, and purified. The final digested DNA fragment was ligated into pET-32a (+) expression vector (Novagen, USA) using T4 DNA ligase (Invitrogen, USA) and transformed into XL1-Blue non-expression host cell. **(D): Colony screening of pET32a ligated insert** in transformed XL 1-blue non-expression host strain using LIC primers. Seven of the nine clones chosen have the correct size insert. **(D): Sub cloning of ligated Pet Vector plasmid into BL-21 (DE3) (Novagen, USA) expression host**. Again, colony screening was performed (lane 1 to 5). The clones with the correct size insert were subsequently sequenced from both ends to check for correct reading frame. Clone 5 and 6 were selected for protein expression. Glycerol stocks were made from those clones that were used for expression.



Appendix XXII.

Molecular cloning of bHSP90. (A) PCR amplification of target heat shock protein 90 gene with gene specific primers using Expand long-template *Taq* DNA polymerase. Amplicons in Lane 1 and 2 corresponds to correct expected size of ~2200 bp. PCR amplicons from both lanes were extracted and purified using QIA quick Gel extraction Kit (Qiagen) and ligated to pGEMT-Easy vector followed by transformation into XL-1 blue non-expression host. **(B) Colony screening of PCR inserts in pGEMT vector using SP6 and T7 primers:** All five colonies showed the presence of insert with expected size of ~2000 bp. **(C) PCR amplification of target gene from pGEM-T plasmids with correct insert using designed LIC primer adaptors.** Purified pGEM-T plasmid from clone 1 (Lane 4 of Fig B) was used. The PCR amplified band was gel extracted, and purified. The final digested DNA fragment was ligated into pET-32a (+) expression vector (Novagen, USA) using T4 DNA ligase (Invitrogen, USA) and transformed into XL1-Blue non-expression host cell. **(D): Colony screening of pET32a ligated insert** in transformed XL 1-blue non-expression host strain using LIC primers. Four of the five clones chosen have the correct size insert. **(D): Sub cloning of ligated Pet Vector plasmid into BL-21 (DE3) (Novagen, USA) expression host**. Again, colony screening was performed (lane 1 to 5). The clones with the correct size insert were subsequently sequenced from both ends to check for correct reading frame. Clone 2 and 3 were selected for protein expression. Glycerol stocks were made from those clones that were used for expression.