

32060

Sequence Tag Catalogs of Dust Mite-Expressed Genomes

Utility in Allergen and Acarologic Studies

Aaron Chen Angus, Seow Theng Ong and Fook Tim Chew

Department of Biological Sciences, National University of Singapore, Singapore

Contents

Abstract	357
1. The Expressed Sequence Tagging Approach	358
2. Rapid Identification of Allergen Homologs	359
3. Rapid Identification of Non-Allergen Homologs	363
4. Expression Profiling	363
5. Biologic Studies	363
6. Phylogenetic Studies	365
7. Assessment of Allergen Isoforms and Gene Families	365
8. Expressed Sequence Tags as a Means of Facilitating Proteomic Studies and Characterization	366
9. Conclusion	367

Abstract

Dust mites are a major source of indoor allergens. They contain a large number of components that react with immunoglobulin (Ig) E in individuals with allergies and are capable of inducing sensitization, and allergic respiratory and cutaneous diseases. With a significant proportion of the population affected in some way by mite allergies, it is essential that we improve our understanding of these organisms so that control strategies could be defined and its allergens better understood. Thus, we have initiated a project using the expressed sequence tagging (EST) strategy to study the major species of dust mites associated with allergic diseases, in particular, the American house dust mite, *Dermatophagoides farinae*, as well as *Blomia tropicalis*, the most prevalent mite in domestic tropical dwellings. The work has recently been expanded to include 'storage' mites such as *Tyrophagus putrescentiae*, *Acarus siro*, *Lepidoglyphus destructor*, *Glycyphagus domesticus*, *Suidasia medanensis*, and *Aleuroglyphus ovatus*.

More than 50% of the initial 3000 ESTs from the *D. farinae* and *B. tropicalis* dust mites showed significant matches to known genes and were categorized into eight functional groups (such as proteins involved in metabolism, gene expression, protein synthesis, cell signaling, etc.). Of specific interest, however, were the homologs to known mite allergens, in addition to a number of sequences bearing significant homology to allergens from non-mite sources previously not known to exist in mites. The availability of these allergen sequences has facilitated their expression and subsequent characterization in our laboratory in terms of their IgE-binding reactivity.

The wealth of sequence information, generated via the EST project, has also facilitated the identification of polymorphic forms of allergens, the investigation of differential gene expression under various environmental conditions via DNA microarrays, as well as the analysis of protein level expression profiling via the proteomics approach. Additionally, ESTs have also ameliorated the understanding of the phylogenetic relationships between

mites, and enabled the isolation of gene products crucial for life processes so that mite control strategies can be more effectively devised. Taken together, the utilization of the EST strategy has opened up numerous new avenues by which the allergist can engage more effectively in the study of dust mites with the ultimate aim of developing appropriate treatment regimens for mite-induced allergy.

House dust mites are the most important source of allergens in the indoor environment. Multiple epidemiological studies have shown strong causal links between sensitization to its allergens and the development of asthma.^[1] Several species of mites are predominant in the domestic indoor environment, the most common being the *Dermatophagoides pteronyssinus*, *D. farinae*, and *Euroglyphus maynei*, all of which have been shown to be ubiquitous in mite fauna worldwide.^[2] In tropical and subtropical regions, another species, *Blomia tropicalis*, has been shown to be prevalent.^[3,4] Several other mites have also been shown to be potent sources of allergens. These include *Tyrophagus putrescentiae*, *Acarus siro*, *Lepidoglyphus destructor*, *Glycyphagus domesticus*, *Suidasia medanensis*, *Aleuroglyphus ovatus*, and other storage mites. These mites are mainly present in farming communities and in stored products (hence their common name, storage mites) but have regularly been found in domestic dwellings as well.^[2]

Dust mite allergy constitutes a significant health problem worldwide. It is estimated that in the US alone, 15 million Americans suffer from mite allergies,^[5] while in Asian populations, such as in Singapore^[6] and Taiwan,^[7] up to one-third of the local population display allergic symptoms as a result of sensitization to mite allergens. With one in every three individuals in the population affected in some way by mite allergy, it is essential that a better understanding of the organism be achieved, such that its allergenic mechanisms can be delineated and control strategies subsequently defined.

To date, there are about 150 DNA sequences, sequence variants and fragments from various species of dust mites deposited in the public databases, out of an estimated 10 000 different expressed genes the organism may have (based on a genome size of approximately 10^8 – 10^9 bp),^[8] and few have been well characterized. A consequence of this is that efforts to understand the organism and search for control targets have mainly relied on empirical studies. However, despite the fact that this organism is so widespread, ecologically successful, and biomedically important, very little information is currently available; specifically only 1% of its total potential genomic information has been deposited in public databases.

1. The Expressed Sequence Tagging Approach

In the post-genomic era, numerous rapid and reliable methods of genome characterization and analysis have become accessible to the scientific community. In view of this, a pilot expressed sequence tagging (EST) approach was considered and applied to dust mite research to obtain a larger sampling and overview of the expressed genomes of several major species of dust mites.

ESTs constitute a useful tool for the identification of new genes. They have also been used to provide both quantitative and qualitative information regarding gene expression in a variety of tissues, cells, and organisms, or of various developmental and physiological (disease vs control) states.^[9–18] We have utilized this for the rapid identification and isolation of genes encoding potential allergenic components in dust mites. In our laboratory, we initiated the pilot characterization of two major species of mites, namely *D. farinae* and *B. tropicalis*, by EST analysis in the year 2000.^[19–22]

Our EST approach involved the partial sequencing of random clones selected from a λ -ZAP[®] cDNA libraries (Stratagene, La Jolla, CA, USA). With advances in automated sequencing procedures, we were able to read up to 700 base pairs (bp) efficiently and latest modifications have even achieved up to 1000bp without anomalies.^[23] As in many EST collections, any clone of interest would have been physically cataloged and stored, and, as such, longer sequence data from any of the tagged clones could easily be obtained by further sequencing.

These partial sequences were thereafter processed to remove sequences of poor quality as well as to trim off vector sequences, and the edited sequences were queried against a non-redundant protein database in the GenBank[®] database (<http://www.ncbi.nlm.nih.gov/>) using the BLASTX algorithm. The identification of homologs was generally based on high sequence identity over a relatively long sequence of bases. In general, search results with expectation values (E-values) <0.001 were considered

significant identities.¹ First, the optimal E-value to be used as a valid cut-off point was first evaluated (figure 1). Based on the results, E-value ≤ 0.001 was deemed to be optimal, as beyond this point a constant increment in the number of nonsignificant identities would be observed were a less stringent cut-off E-value chosen. For sequences with significant identities, a catalog was constructed to classify them according to their putative biological functions. The categories in this catalog were based on Mark Adams' 1991 classification of human brain cDNA libraries^[24] with minor modifications. These categories include: signaling and communication, metabolism, defense and homeostasis, structural, cell surface and motility, DNA synthesis and cell division, gene expression and protein synthesis, protease and protease inhibitor, hypothetical protein, unclassified protein and no homology.

Due to the fact that only a small proportion of the genome codes for proteins (e.g. 3% of the human genome), the EST strategy takes into account that the most important part of the genome is the protein coding region, and thus sequencing the expressed genome would give the largest return of biological information per base pair sequenced.^[25] The easiest way to access the coding regions would thus be to sequence cDNAs. Hence, the bulk isolation and identification of dust mite genes via EST would be a worthy investment for acarologists, allergists, and the medical community in general. Our preliminary results have already enabled us to identify numerous new allergens for subsequent characterization as well as several transcripts associated with a number of important life processes.^[8]

There are many possible uses for generated ESTs, such as probes for the isolation of full length clones and the characterization of promoters and other regulatory regions. ESTs can also be used as tissue-, cell-type-, developmental stage-, or even organism-specific molecular markers, as probes for the identification of new members of gene families, for physical chromosome mapping, analysis of gene expression patterns, and as a useful sequence resource for phylogenetic analysis and molecular evolution.

In this review, we focus on the application of ESTs in allergen research. Specifically, we discuss their use in the rapid identification of putative allergens and their non-allergenic homologs (or even putative adjuvants), the elucidation of allergen polymorphisms and paralogs, comparative genomic studies, and also as a

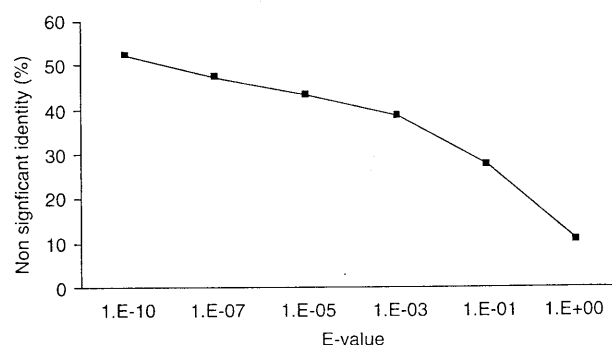


Fig. 1. Criteria for determination of expectation value (E-value) cut-off for significance of alignments. The percent of nonsignificant identity versus E-value cut-off is shown. E-value ≤ 0.001 was chosen as the optimal cut-off value.

means to facilitate proteomic and functional genomic studies on this organism.

2. Rapid Identification of Allergen Homologs

The first complete sequence of a dust mite allergen was published in 1988^[26] and since then cDNA- and PCR-based cloning have successfully defined the primary structures of a number of dust mite allergens. However, the majority of these are from *D. pteronyssinus*, and at the time we initiated our EST project, only three allergens of *B. tropicalis* were isolated and sequenced.

A glimpse into the expressed genomes of both *D. farinae* (Der f) and *B. tropicalis* (Blo t) revealed at least 30 distinct groups of putative allergens based on sequence homology across species and even across different types of organisms (including pollen and fungal allergen homologs). A summary of the various allergens identified from the EST collection is listed in table I. Through this collection, we were able to isolate the majority of the major allergens from both *D. farinae* and *B. tropicalis*.^[27,28] These include the groups 1–15 allergens from both species of dust mites. Many of these, such as Der f 4, Der f 5, Der f 8, Der f 9, and Der f 13, are newly identified homologues from *D. farinae*, while the corresponding allergens for *B. tropicalis* have not been reported previously. Of particular interest is the isolation of the Blo t 1,^[29] Blo t 2,^[29-31] Blo t 4,^[32] Blo t 9, and Der f 9,^[33] allergen cDNA clones. Many of these have now been full-length sequenced, cloned, and expressed as recombinant proteins, and have shown to be IgE binding.^[27-33] A preliminary profile of the IgE-binding responses of these allergens using sera of atopic patients from Singapore (where the mite fauna is predominantly dominated by

1 The expectation value is defined as the number of different alignments with scores equivalent to or better than the raw score *S* for an alignment (calculated by summing the scores for each aligned position and the scores for gaps) that are expected to occur in a database search by chance. The lower the E value, the more significant the score.

Table I. Summary of allergen homologs, paralogs, and isoforms identified from the *Dermatophagoides farinae* (Der f) and *Blomia tropicalis* (Blo t) expressed sequence tag catalogues

No.	Mite allergen	Identity	Der f	Blo t
Current allergen groups				
1	Group 1	Cysteine protease	X	X
	Group 1-like	Cathepsin B	Δ	Δ
	Group 1-like	Cathepsin C	Δ	NI
	Group 1-like	Cathepsin D	NI	Δ
	Group 1-like	Cathepsin L	Δ	Δ
2	Group 2	Unknown	X	Δ
3	Group 3	Trypsin	X	X
4	Group 4	Amylase	Δ	Δ
5	Group 5	Unknown	Δ	X
	Group 5 (paralog)	Unknown	Δ	Δ
6	Group 6	Chymotrypsin	X	X
7	Group 7	Unknown	X	Δ
	Group 7 (paralog)	Unknown	Δ	Δ
8	Group 8	Glutathione-S-transferase	Δ	Δ
9	Group 9	Collagenolytic serine protease	Δ	Δ
10	Group 10	Tropomyosin	X	X
11	Group 11	Paramyosin	X	X
12	Group 12	Unknown	NI	X
13	Group 13	Fatty acid binding protein	Δ	X
	Group 13 (paralog)	Fatty acid binding protein	Δ	Δ
14	Group 14	Apolipoprotein like protein/Mag 1 and Mag 3	X	Δ
15	Group 15	98kD chitinase	X	Δ
16	Group 16	Gelsolin/villin	X	NI
17	Group 17	Calcium-binding EF hand protein	X	NI
18	Group 18	60kD chitinase	X	Δ
19	Group 19	Anti-microbial peptide homologue	NI	X
20	Group 20	Arginine kinase	Δ	Δ
New allergen groups (mostly pan-allergens)				
21	Mag29 [Clah 4 homologue]	Heat shock protein 70	X	Δ
22	Asp f 12 homologue	Heat shock protein 90 (<i>Aspergillus fumigatus</i>)	Δ	Δ
23	Cop c 2 homologue	Thioredoxin (<i>Coprinus comatus</i>)	Δ	Δ
24	Act c 1 homologue	Actinidain (<i>Actinidia chinensis</i>)	Δ	Δ
25	Mal f 6 homologue	Cyclophilin (<i>Malassezia furfur</i>)	Δ	Δ
26	Ole e 3 homologue	Calbindin (<i>Olea europaea</i>)	Δ	Δ
27	Alt a 6 homologue	Acidic ribosomal protein P2 (<i>Alternaria alternata</i>)	Δ	Δ
28	Alt a 10/Clah 3 homologue	Aldehyde dehydrogenase	Δ	Δ
29	Gal d 1 homologue	Ovomucoid (<i>Gallus domesticus</i>)	Δ	Δ
30	Gal d 2 homologue	Ovalbumin (<i>G. domesticus</i>)	Δ	NI

Continued next page

Table I. Contd

No.	Mite allergen	Identity	Der f	Blo t
1	Heb v 8 homologue etc.	Profilin	Δ	Δ
2	Pen ch 20 homologue 68kD allergen	Hexosaminidase/N-acetyl-glucosaminidase (<i>Penicillium chrysogenum</i>)	Δ	Δ
3	Ves m 1 homologue	Phospholipase A1 (<i>Vespula maculifrons</i>)	Δ	Δ
4	Che a 3 homologue	Chenopodium album Polcalcin allergen	Δ NI	Δ NI
5	Sol i 3 homologue	Unknown (<i>Solenopsis invicta</i> , fire ant)	Δ	Δ
6	α-tubulin homologue	α-tubulins	Δ	Δ
7	Others	For example, lipid transfer proteins, Hevea allergen homologue, allergen V5/Tpx-1 related family, Api m 6 homologue, etc.	Δ	Δ

Δ = newly identified allergen (based on the current EST program); NI = no recombinant allergens available (not identified); X = previously identified allergen.

Blomia tropicalis and *Dermatophagoides* spp.) is shown in figure 2 and figure 3.

The EST-based gene discovery approach was also particularly valuable in the identification of allergens with homology to genes from more distantly related species or even across taxonomic kingdoms. Traditional DNA hybridization or PCR-based approaches often fail to find related genes because of nucleotide-level sequence divergence. This phenomenon has previously been demonstrated in many genes shared by the closely related nematodes, *Caenorhabditis elegans* and *C. briggsae*. Database searches, sometimes referred to as 'in silico hybridization', can provide a much more sensitive means of finding related genes, thereby allowing much larger inter-species leaps.^[34]

In our laboratory, we have also encountered this while trying to obtain the homologs to the group 1 allergens in *B. tropicalis*^[29] and the group 9 allergens.^[33] The traditional PCR-based cloning method failed to yield any clones showing homology to these allergens. In addition, we have previously shown that there was only low to moderate cross reactivity between the major allergenic components of *B. tropicalis* and the *Dermatophagoides* spp.,^[35] thus rendering antibody-based screening of cDNA libraries (using *Dermatophagoides*-specific antisera) difficult. With our pilot EST catalogs, however, we were able to identify homologs of these allergens fairly easily. In comparison to other group 1 dust mite allergens, for example, the Blo t 1 precursor protein has approximately only 30% amino acid identity with precursors of Der p 1, Der f 1, and Eur m 1.^[29]

In addition, we were also able to identify homologs to allergens from other non-mite organisms such as cockroaches, tree or pollen (*Juniperus ashei*, *Juniperus virginiana*, and *Prunus avium*, *Hevea brasiliensis*), yeast (*Malassezia* spp.), chicken (*Gallus gallus*),

yellow jacket venom, fungi (*Alternaria alternata*, *Cladosporium herbarum*, *Penicillium citrinum*, *Candida albicans*, and *Coprinus comatus*), and fruit (*Actinidia chinensis*). These allergens were identified as heat shock protein 70, profilin, enolase, ovomucoid precursor and ovalbumin, phospholipase A1, ribosomal protein P2, alcohol dehydrogenase, thioredoxin, hexosaminidase, cyclophilin and actinadain-like allergen homologs,^[21,22] and many were also shown to be IgE binding.^[36,37]

This pilot EST collection has proven to be a valuable resource for the rapid identification and isolation of new mite allergens. It has also provided us with the opportunity to perform multiple comparisons among these groups of allergens (e.g. in terms of cross-reactivity,^[38] and the degree of primary sensitization in different populations) as well as to test several novel hypotheses pertaining to the phylogenetic origins of several pan-allergens^[39] and functional relationship between these allergenic components.

With the large collection of ESTs we were also able to determine which among these allergens would be associated with atopic disease presentation.^[40] Although sensitization to crude allergenic extracts is a diagnostic hallmark of atopic diseases, a positive response (indicative of the presence of specific IgE) does not necessarily indicate disease manifestation. An allergic individual may respond to multiple components from a particular source, e.g. the dust mites, but which of the individual allergens are truly associated with disease presentation has never been evaluated as a whole. The EST collection, combined with our proteomic studies (see section 8), has enabled us to identify the majority, if not the full panel, of allergenic components from these dust mites, and clinical evaluation of their association with atopic disease manifestation is currently on-going in several populations. Ultimately, this would lead to a better understanding of the specific compo-

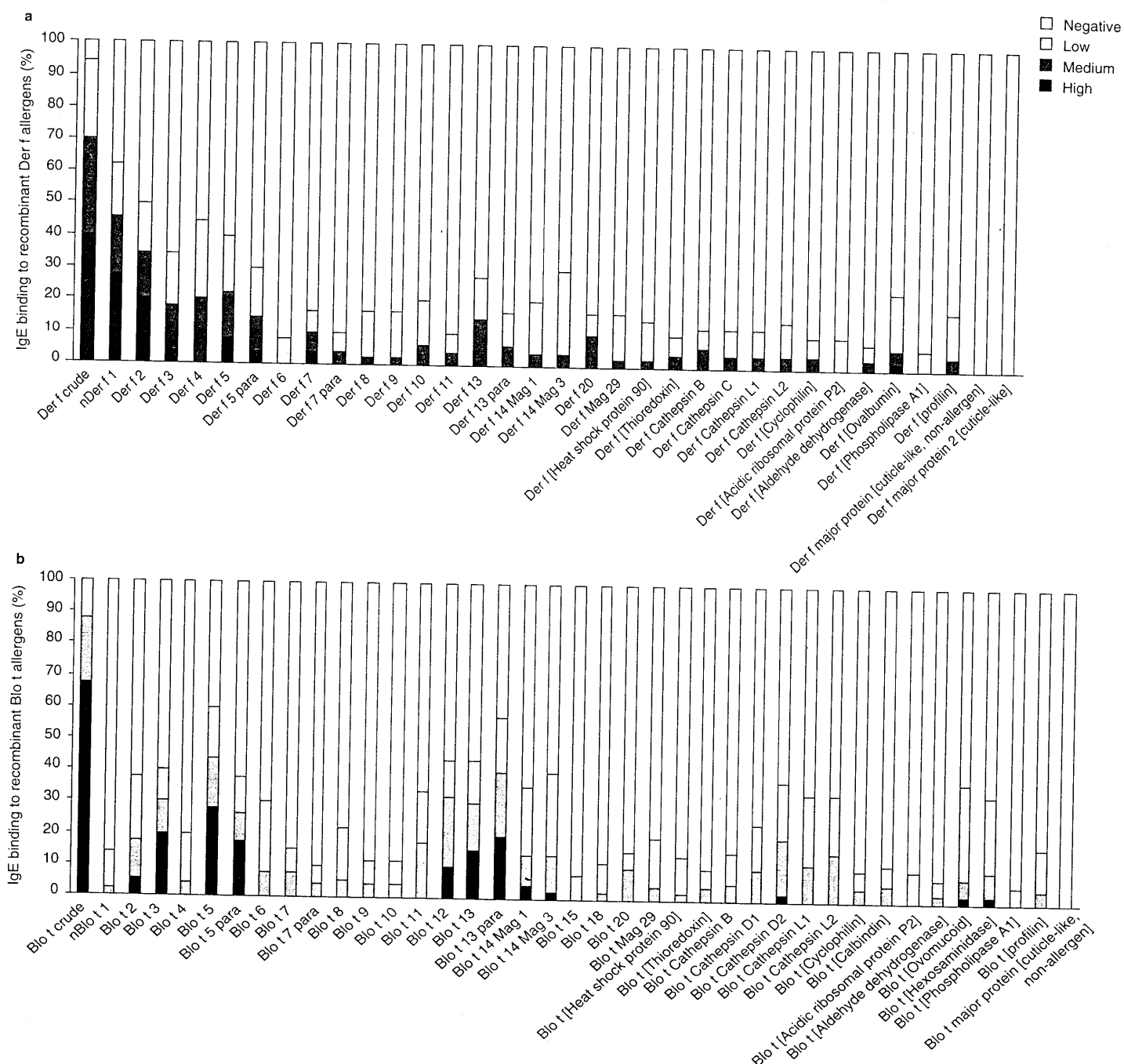


Fig. 2. Sera IgE-binding frequencies and intensities against (a) *Dermatophagoides* spp. and (b) *Blomia tropicalis* allergens in an evaluation of individuals sensitive to dust mites (preliminary results for Singapore). Included in this panel are the major Der f allergens and several other pan-allergens from *D. farinae* as well as the two major Der f protein components (cuticle like non-allergens). With the *Blomia tropicalis* panel, included are nBlo t 1, Blo t 2–Blo t 15, Blo t 18, Blo t 20, and several paralogous forms of the allergens together with the corresponding major cuticle-like protein (non-allergen) as a negative control. **Blo t** = *Blomia tropicalis*; **Der f** = *Dermatophagoides farinae*; **nBlo t** = natural purified *Blomia tropicalis*; **nDer f** = natural purified *Dermatophagoides farinae*.

nents that are associated with the disease processes and, thereafter, to the development of antigen-specific immunotherapy.^[41-43]

A further example of some of the work we are now able to pursue as a result of the wealth of new information gathered on

mite allergenic components is the functional relationship between some of the allergens present. Take for instance the dust mite group 14 allergen that was recently identified as a vitellogenin-apolipoprotein-like protein.^[44] Vitellogenins are the precursors of

egg-yolk proteins in most oviparous species from invertebrates to vertebrates. In oviparous vertebrates, vitellogenins are synthesized in the liver and transported through the blood to oocytes to provide nutrients in support of early embryogenesis. We have identified several other vitellogenin-like proteins from our EST collection. Additionally, we have also identified two allergens showing homology to ovomucoid precursors or ovalbumin, namely Gal d 1 and Gal d 2 homologs, the major chicken egg white allergens.^[36] Whether or not any relationship exists between these three allergens (group 14, Gal d 1, and Gal d 2 homologs) is an interesting point worthy of further experimentation. It would also be interesting to evaluate the possible cross reactivity between Gal d 1 and Gal d 2 with their corresponding homologs in the dust mites.

3. Rapid Identification of Non-Allergen Homologs

Allergens from dust mites correspond to proteins of diverse functions. A sizable group of these allergens, such as groups 1, 3, 4, 6, and 9, have been shown to possess enzymatic activity whereas the group 10 and 11 allergens shared homology with cytoskeletal proteins.^[45,46] Some other allergens also show high homology with glutathione *S*-transferase, fatty acid binding protein, and heat shock protein.^[47-49]

There is no unifying theory for why some proteins are allergenic and their biological counterparts are not. For instance, dust mite group 10 allergens shared high sequence similarity with muscle protein tropomyosin. This molecule belongs to a family of highly conserved proteins associated with the thin filaments in muscle and microfilaments in many non-muscle cells. However, tropomyosin has been shown by several laboratories to be a major allergen in shrimps,^[50-52] cockroaches,^[53,54] lobsters,^[55] and squid.^[56] Furthermore, Martinez and coworkers^[57] suggested that tropomyosin could be the molecule to which the cross-reactivity between cockroaches, mites, nematodes and crustaceans can be attributed.

Interestingly, the non-allergenic counterpart of tropomyosin is present in all living organisms and plays a functional role in the contractile activities in muscle cells together with actin and myosin, and is believed to participate in the regulation of cell morphology and motility in non-muscle cells. We have isolated different contigs of tropomyosin transcripts from the dust mites and it would be interesting to determine why one group can be allergenic and the other not. Similar work could also be done with proteases, as several groups of proteases from the dust mites were isolated in addition to the allergenic proteases. One example is cathepsin B, which shares some similar functional features to group 1 allergens.

Our EST database has allowed us to rapidly isolate not only the allergens but also their non-allergenic biological counterparts, and the data available will thus facilitate the study of their distinct properties. Full-length sequences of the clones of interest can be retrieved for some characterization studies such as sequence alignments, structural analysis, and cellular response, in the hope of locating clues as to why proteins so similar in identity can differ in their capacity to elicit allergic responses.

4. Expression Profiling

The analysis of cDNA libraries is a fundamental approach for the study of gene expression and has become an established method for the discovery and analysis of expressed genes. EST sampling can provide information on the spectrum of genes transcribed in the source tissue and the relative abundance of these transcripts, which serves as an effective and quantitative measure of steady-state mRNA levels. This 'electronic northern' approach has been used to develop expression profiles in a range of rat cells,^[58] invertebrate,^[59] and plant tissues.^[60]

These EST collections have enabled the generation of expression profiles useful for performing large-scale gene expression comparisons between two or more transcript populations, which allow the identification of differences in expression at the single gene level. When several cDNA libraries are used in a comparative study, the frequency of occurrence of the various clones can also be analyzed and compared. For instance, Ablett et al.^[60] found that in 5000 grape ESTs, 18% of the berry transcripts are specialized in the defense category as compared with 7% in the leaf, whereas transcripts involved in general housekeeping functions were similar in both tissues.^[60] We found that our *D. farinae* EST collection comprised more than 20% metabolic genes while *B. tropicalis* had only 14% of the ESTs linked to such functions (general housekeeping genes were similar in both as well). Also, the number of transcripts for cathepsins and its allergenic counterpart group 1 dust mite allergen is far higher in *D. farinae* than those identified in *B. tropicalis*. Efforts are currently underway to determine whether these differences have any relationship with the preferred niche environments of each dust mite.

5. Biologic Studies

Microarrays are best suited for determination of the change in relative abundance for transcripts by enabling the simultaneous monitoring of changes in gene expression at a global level. This is achieved by hybridization with two differentially labeled probes one corresponding to a reference or control condition and the other

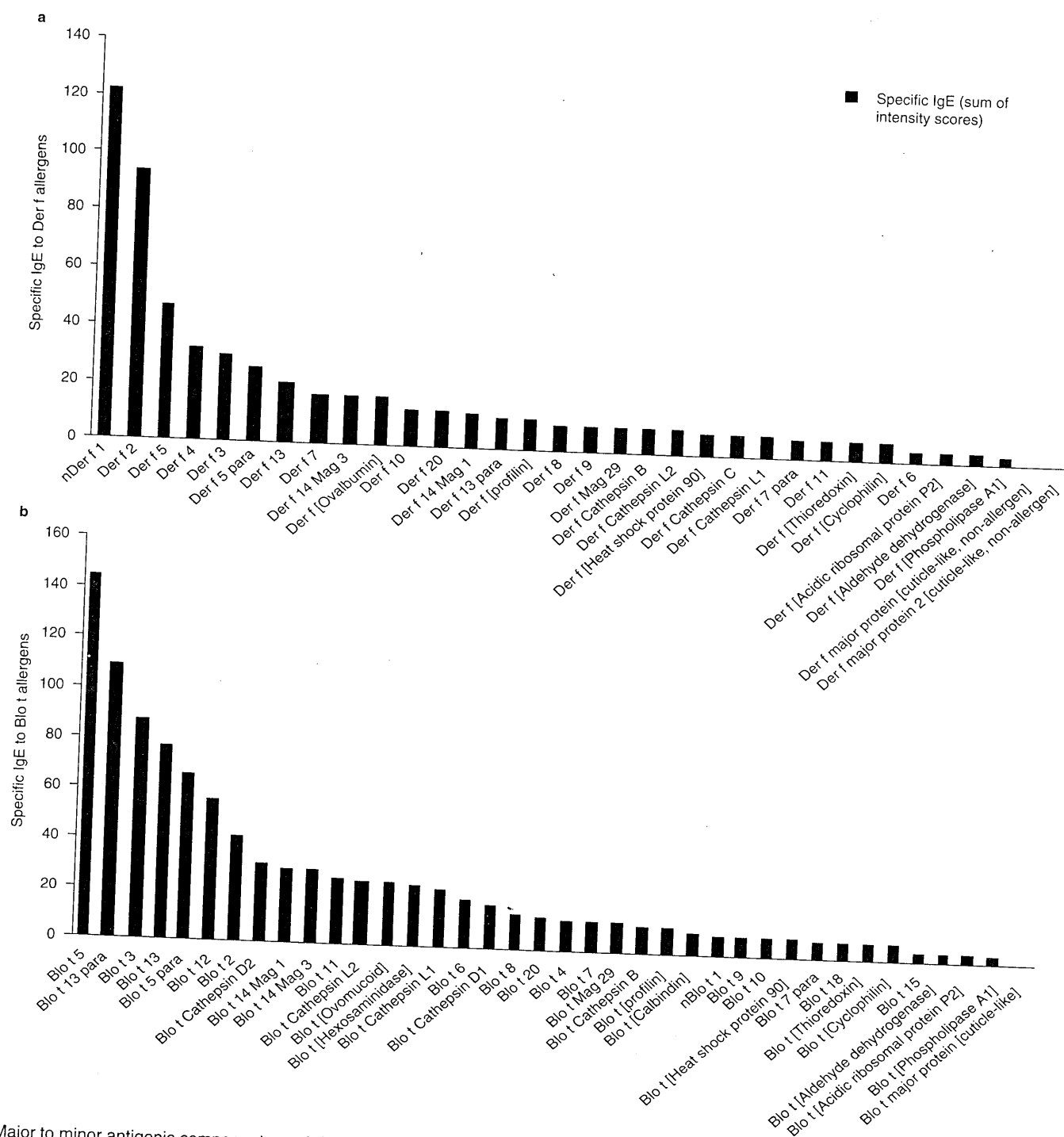


Fig. 3. Major to minor antigenic components as defined by the sum of sera specific immunoglobulin (Ig) E binding intensity scores of 181 house dust mite-sensitive patients screened with selected dust mite allergens; (a) *Dermatophagoides farinae* (Der f) and (b) *Blomia tropicalis* (Blo t). The units of the y-axis are the sum of intensity scores (arbitrary score units). para = paralog.

be representing the experimental condition. In this manner, the ratio of the hybridization signals determines the relative difference in abundance for a given transcript under different conditions. When the lack of availability of a complete genomic sequence for dust mites, one has to make use of the coding sequences available

or, alternatively, turn to ESTs to fulfill the requirements of a microarray study.

Having a large number of EST clones available allows the analysis of overall gene expression. cDNA clones can be gridded onto a glass slide and hybridized with various complex probes

corresponding to different development stages, different stimuli, or different genetic backgrounds. Alternatively, oligonucleotides derived from the ESTs can be spotted on DNA chips for other applications.

Knowing where and when a gene is expressed often provides a strong clue as to its biologic role. Conversely, the patterns of gene expression in a cell can provide detailed information about its physiological state. For instance, ESTs derived from a *Brugia malayi* third-stage larvae cDNA library reveal a multitude of genes that are expressed in nematodes poised to invade the body.^[61] The ESTs generated from *Meloidogyne* and *Globodera* second-stage juveniles provide parallel information for plant parasitic nematodes. It has been said that knowledge of the biology and ecology of these mites is the starting point for the design of effective allergen control regimes.^[2] However, the available knowledge is rather scant, limited to only pictorial guides for taxonomic identification of mites found in homes,^[62] and some basic studies on the life history and environmental effects on the dynamics of dust mites.^[63] Hence, we now have the opportunity to embark on microarray-based studies to facilitate understanding of gene expression in the different life stages of dust mites, which might provide some valuable information on genes playing essential roles in development and maturation.

Virtually all changes in physiological status, regardless of their genetic or environmental bases, are mediated by changes in gene expression. Diversion from normal physiology, for instance, in the case of stress, is generally accompanied by a plethora of molecular and biochemical changes, including alterations in gene expression patterns. The regulation of genes is either the direct cause or result of such modulation in environment. For example, DeRisi et al.^[64] reported a comprehensive investigation of the temporal program of gene expression in yeast undergoing the metabolic shift from fermentation to respiration.

The life cycle of all mites consists of different stages: egg, larva, nymphal stages, and adulthood. The duration of development from egg to adult and population growth is dependent on both ambient relative humidity and temperature.^[65] The optimum growth conditions for dust mites are about 25–30°C and 80% relative humidity. However, the mite *A. siro* was reported to be able to complete its life cycle at 5°C with a humidity of 67.5%.^[66] We have observed in our own studies that we could hold entire cultures of *D. farinae* dust mites in an inanimate state for a long period of time at 4°C and revive them at room temperature (unpublished data). Additionally, we have noted that mites at certain stages of their life cycle seem to be resistant to extreme environmental pressures (unpublished data). Information on the

molecular mechanisms behind these phenomena would be useful in our understanding of dust mite biology and may ultimately lead to the development of novel control measures, which includes protocols aimed at the reduction of mite infestation.^[67] Nevertheless, the success of these measures is variable and relies heavily on empirical observations. Thus, investigating the change in gene expression profiles at different growth temperatures, humidity culture conditions, or in unique physiological phenomena as described above, might provide some valuable insights into mite control strategies.

6. Phylogenetic Studies

EST databases can be scanned for sequences coding for the ancient conserved regions (ACR) that are common to a number of proteins conserved in different phyla,^[68] more than two-thirds of the known ACR can be recognized in the Arabidopsis EST collection.^[69] Yan and Gong^[70] assembled a complete zebrafish mitochondrial 16S rRNA gene from overlapping ESTs. Mitochondrial 16S rRNA has been widely used to construct phylogenetic relationship throughout eukaryotic kingdoms because of its relatively rapid mutation rate and, thus, the high sensitivity for phylogenetic analysis among closely related species.

Evolutionary studies on mites have primarily surveyed the mitochondrial cytochrome oxidase subunit I (*COI*) gene, and work performed on *COI* sequences has revealed the population genetic structure of the species *T. urticae* throughout its distribution area.^[71] In our EST collection, multiple copies of the *COI* gene were found and, thus, a similar approach could be used to gather information on genetic relatedness of the different species of allergy-causing mites from different parts of the world so as to generate information useful for more comprehensive diagnostic and therapeutic approaches in allergy studies.

7. Assessment of Allergen Isoforms and Gene Families

A survey of the clones in our EST database has revealed the presence of a number of cDNA clones which matched sequences of known allergens, but which differed slightly with respect to their amino acid sequences. This is evidence of the presence of isoallergens – allergens that are very similar in molecular sizes and biological function but differ by a few residues in amino acid sequence. We have also seen several groups of allergens forming different contigs or gene clusters that were significantly more different than just a few nucleotide polymorphisms.^[72,73] Thus, the existence of polymorphisms and multi-gene families necessitate

their further characterization, as a diversity is introduced which may influence allergenicity.

Polymorphisms in dust mite allergens were first reported following the isolation of five cDNA clones of *D. pteronyssinus* group 1 allergen (Der p1) which differed from each other by one to three nucleotide residues.^[74] Though the changes were deemed minimal, they were significant on the basis that any change in amino acid sequence, including alterations in residues outside T cell epitopes, has the capacity to profoundly affect T cell epitopes. Moreover, three of the seven reported epitopes in Der p1 were found to possess polymorphic residues.

Screening of *D. pteronyssinus* cDNA libraries with a Der p2 cDNA clone as a probe as well as PCR amplification of genomic DNA with Der p2-specific primers from mites isolated from house dust resulted in the characterization of several distinct allelic variants.^[75] The translated polypeptides differed from each other by three to four amino acid residues and, interestingly, were also found in coding regions previously assigned to T cell epitopes. Affinity-purified natural Der p2 proteins were then subjected to two-dimensional (2D) gel electrophoresis and immunoblot analysis in order to address the question of these Der p2 variants at the protein level, and 10 isoforms with isoelectric point (pI) values ranging from 5.9 to 7.0 were observed. Genomic DNA from *D. farinae* was PCR-amplified with Der f2-specific primers, and the resulting amplified fragments compared found to correspond to four distinct polymorphic clones.^[76] It should be noted that the frequent polymorphic positions identified in Der p2 were at positions 40, 47, 111, and 117, while those reported for Der p2 were at positions 76, 88, 111, and 125. It is not known why the locations of polymorphisms do not coincide, despite the high level of sequence homology between both allergens, although this may be due to the fact that insufficient data exists for Der f2. Polymorphisms have also been identified in Der p3,^[77] Der f3,^[78] and Der p5.^[79] In contrast, two isoallergens of *L. destructor* group 2 allergen (Lep d2) were isolated, differing in 10% of their amino acid sequences,^[80] despite the demonstration that their IgE reactivity *in vivo* was shown to be comparable, as determined by skin prick testing.^[81,82]

Based on these studies, it can be observed that the frequency of polymorphisms in the allergens studied were, in many incidences, greater than the 1/10 000 error rate expected from replication during the construction of cDNAs. Moreover, the degree of sequence variation among the reported isoallergens was not usually found in DNA coding for other eukaryotic proteins, thus indicating that isoallergenic variation is an unusual occurrence that warrants further analysis. Another crucial point to note is that mites from

different environments and commercially cultured mites may differ with regards to the location and nature of the polymorphic residues, such as in the case of Der p1, where three amino acid substitutions were present in the environmental mites from different geographical locations in Australia and not in commercial mites obtained from the Commonwealth Serum Laboratories, Melbourne, Australia.^[83] Ultimately, this becomes a point of concern because commercial mites, in such cases, would not correspond to mites that are responsible for sensitization in the wider environment.

These observations have led us to analyze the nature and diversity of polymorphisms in dust mite allergens using the EST collection as a starting point. Advances in high-throughput mutation detection and genotyping technologies,^[84,85] largely previously applied to the analysis of single nucleotide polymorphisms in the human genome, may provide impetus for the elucidation of the spectrum of polymorphisms in dust mites from the environment. Examples include conformation-based mutation scanning and *in silico* single nucleotide polymorphism (SNP) mapping for polymorphism detection, and oligonucleotide ligation assay genotyping, mini-sequencing, homogenous solution hybridization using fluorescence resonance energy transfer detection, dye-labeled oligonucleotide ligation, and single-base extension with fluorescence detection, for genotyping. These assays seem to be both adequately sensitive and amenable even for studies using a single dust mite (unpublished data).

A full understanding of the degree of sequence polymorphism and the presence of multi-gene families of allergenic components would be essential for the development of therapeutic and diagnostic tools for mite allergy. The various naturally occurring isoforms would, thus, be useful in the elucidation of immunopathogenesis of mite allergy, in view of the observation that many of these polymorphisms lie within defined T cell epitopes.

8. Expressed Sequence Tags as a Means of Facilitating Proteomic Studies and Characterization

Very little is known about the major proteins in mites apart from the allergenic components that have been isolated and characterized via molecular biology. Knowledge of these other components, regardless of whether they are allergenic or not, may be useful in our comprehension of host allergic responses, mite biology, and control. Without the EST collection however, it would have been difficult to pursue larger scale proteomic-based studies, as the basis of protein identification via mass spectrometry is the availability of genomic or sequence data. Previous studies on

IgE-binding proteins screened via 2D-electrophoresis and immunoblotting, required Edman degradation based *N*-terminal amino acid sequencing to identify the proteins of interest.^[86] However, this process is slow, requires relatively large amounts of protein, and blocked *N*-termini are something that is not uncommon.

Nevertheless, to a certain degree the availability of an EST database facilitates the mass spectrometric-based identification of proteins. In our laboratory, we recently generated 2D protein profile maps of several species of dust mites. Without the use of the EST database, we observed that less than 5 (out of 100) of the major components isolated from each 2D map could be matched and identified via peptide mass fingerprinting. However, almost all peptide mass fingerprint profiles generated from these major proteins were able to match the dust mite EST sequences.^[87] Additionally, we have confirmed several of these matches using *de novo* tandem mass spectrometric sequencing.^[88]

Using a combination of peptide mass fingerprinting, tandem mass spectrometry, and contiguous EST alignments, we were able to identify, isolate, clone, and express, several major components of the house dust mite proteome for further studies. One example of these is an insect cuticle-like protein that was found to be one of the most abundant proteins in all dust mite species screened thus far.^[88] This protein was observed to be non-allergenic, and may even be a very weak antigen despite being a relatively large (40 kDa, pI 7.2), abundant, and soluble protein. Such proteins could then be used to contrast with dust mite allergens from several angles, e.g. environmental levels (dose) and exposure, host immunologic response, and their physicochemical properties. The philosophy behind this is that to understand what constitutes an allergen, one may also need to know what constitutes a non-allergen.

9. Conclusion

The EST database has revealed a tremendous amount of information on potential allergenic components from dust mites, in addition to facilitating biologic studies with the aim of understanding how to monitor and control this organism. We have extended our studies to include other mites, namely *T. putrescentiae*,^[89] *A. siro*,^[90] *G. domesticus*, *S. medanensis*, and *A. ovatus*, and it is hoped that these endeavors will accelerate our understanding of dust mite biology and reveal the full panel of its allergenic components.

Acknowledgments

This work was supported by grant BMRC/01/1/21/18/077 from the Biomedical Research Council (BMRC), Agency for Science, Technology & Research (A*STAR), Singapore.

The authors have no conflicts of interest that are directly relevant to the content of this review.

References

1. Arlian LG, Platts-Mills TA. The biology of dust mites and the remediation of mite allergens in allergic disease. *J Allergy Clin Immunol* 2001; 107: S406-13
2. Platts-Mills TA, Vervloet D, Thomas WR, et al. Indoor allergens and asthma: report of the Third International Workshop. *J Allergy Clin Immunol* 1997; 100: S2-24
3. Chew FT, Zhang L, Ho TM, et al. House dust mite fauna of tropical Singapore. *Clin Exp Allergy* 1999; 29: 201-6
4. Baratawidjaja IR, Baratawidjaja PP, Darwis A, et al. Mites in Jakarta homes. *Allergy* 1998; 53: 1226-7
5. Galant S, Berger W, Gillman S, et al. Prevalence of sensitization to aeroallergens in Californian patients with respiratory allergy. *Ann Allergy Asthma Immunol* 1998; 81: 203-10
6. Chew FT, Lim SH, Goh DY, et al. Sensitization to local dust-mite fauna in Singapore. *Allergy* 1999; 54: 1150-9
7. Tsai JJ, Wu HH, Shen HD, et al. Sensitization to *Blomia tropicalis* among asthmatic patients in Taiwan. *Int Arch Allergy Immunol* 1998; 115: 144-9
8. Chew FT, Ong ST, Wang WL, et al. Preliminary profile of the dust mite genomes: an expressed sequence tag (EST) analysis of *Dermatophagoides farinae* and *Blomia tropicalis* genes [abstract 192]. In: 57th American Academy of Allergy and Immunology Annual Meeting; 2001 Mar 16-21; New Orleans (LA). *J Allergy Clin Immunol* 2001; 107 Suppl.: 56
9. Ohlrogge J, Benning C. Unraveling plant metabolism by EST analysis. *Curr Opin Plant Biol* 2000; 3: 224-8
10. Rajkovic A, Yan MSC, Klysik M, et al. Discovery of germ cell-specific transcripts by expressed sequence tag database analysis. *Fertil Steril* 2001; 76: 550-4
11. Adams MD, Soares MB, Kerlavage AR, et al. Rapid cDNA sequencing (expressed sequence tags) from a directionally cloned human infant brain cDNA library. *Nat Genet* 1993; 4: 373-80
12. Adams MD, Kerlavage AR, Fleischmann RD, et al. Initial assessment of human gene diversity and expression patterns based on 83 million nucleotides of cDNA sequence. *Nature* 1995; 377: 3-17
13. Waterston R, Martin C, Craxton M, et al. A survey of expressed genes in *Caenorhabditis elegans*. *Nat Genet* 1992; 1: 114-23
14. Severson DW, Zhang Y. Generation of expressed sequence tags and sequence-tagged sites as physical landmarks in the mosquito, *Aedes aegypti*, genome. *Genome* 1996; 39: 224-9
15. Ewing RM, Ben Kahla A, Poirot O, et al. Large scale statistical analysis of rice ESTs reveal correlated patterns of gene expression. *Genome Res* 1999; 9 (10): 950-9
16. Gong Z, Yan T, Liao J, et al. Rapid identification and isolation of zebrafish cDNA clones. *Gene* 1997; 201: 87-98
17. Djikeng A, Agufa C, Donelson JE, et al. Generation of expressed sequence tags as physical landmarks in the genome of *Trypanosoma brucei*. *Gene* 1998; 221: 93-106
18. Santos TM, Johnston DA, Azevedo V, et al. Analysis of the gene expression profile of *Schistosoma mansoni cercariae* using the expressed sequence tag approach. *Mol Biochem Parasitol* 1999; 103: 79-97
19. Ong ST, Chew FT, Lee BW, et al. Identity of the initial 1500 expressed sequence tags (EST) of the *Dermatophagoides farinae* dust mite genome [abstract 639]. In: XVII International Congress of Allergology and Clinical Immunology; 2000 Oct 15-20; Sydney. *Allergy Clin Immunol International Suppl* 2000; 2: 176
20. Kuay KT, Chew FT, Wang WL, et al. Catalogue of the initial 1000 expressed sequence tags (EST) of the *Blomia tropicalis* dust mite genome [abstract 637].

- In: XVII International Congress of Allergology and Clinical Immunology; 2000 Oct 15-20; Sydney. Allergy Clin Immunol International Suppl 2000; 2: 175
21. Lim SH, Chew FT, Ong ST, et al. Identification of new Dermatophagoides allergens and isoforms from an expressed sequence tag database showing homology to other known allergens [abstract 638]. In: XVII International Congress of Allergology and Clinical Immunology; 2000 Oct 15-20; Sydney. Allergy Clin Immunol International Suppl 2000; 2: 175-6
 22. Wang WL, Chew FT, Kuay KT, et al. Identification of putative *Blomia tropicalis* allergens with sequence homology to other known allergens by expressed sequence tagging [abstract 110]. In: XVII International Congress of Allergology and Clinical Immunology; 2000 Oct 15-20; Sydney. Allergy Clin Immunol International Suppl 2000; 2: 32
 23. Huang GM. High-throughput DNA sequencing: a genomic data manufacturing process. DNA Seq 1999; 10: 149-53
 24. Adams MD, Kelly JM, Gocayne JD, et al. Complimentary DNA sequencing: expressed sequence tags and Human Genome Project. Science 1991; 252: 1651-6
 25. Marra MA, Kucaba TA, Hillier LW, et al. High-throughput plasmid DNA purification for 3 cents per sample. Nucleic Acids Res 1999; 27: e37
 26. Chua KY, Stewart GA, Thomas WR, et al. Sequence analysis of cDNA coding for a major house dust mite allergen, Der p 1: homology with cysteine proteases. J Exp Med 1988; 167: 175-82
 27. Ong ST, Shang HS, Lua BL, et al. Specific IgE binding reactivity to nine recombinant allergens from dust mite, *Dermatophagoides farinae* [abstract 479]. In: 58th American Academy of Allergy and Immunology Annual Meeting; 2002 Mar 1-6; New York. J Allergy Clin Immunol 2002; 109 Suppl.: 479
 28. Wang WI, Kuay KT, Shang HS, et al. Specific IgE reactivity to fourteen recombinant allergens from house dust mite, *Blomia tropicalis* [abstract 619]. In: 58th American Academy of Allergy and Immunology Annual Meeting; 2002 Mar 1-6; New York. J Allergy Clin Immunol 2002; 109 Suppl.: 619
 29. Kuay KT, Wang WL, Shang HS, et al. Molecular cloning and characterization of group 1 and 2 allergens from dust mite, *Blomia tropicalis* [abstract 478]. In: 58th American Academy of Allergy and Immunology Annual Meeting; 2002 Mar 1-6; New York. J Allergy Clin Immunol 2002; 109 Suppl.: 478
 30. Christopel JJ, Hozic N, Arruda LK, et al. Expression of the recombinant group 2 allergen from *Blomia tropicalis*: comparison of the immunoreactivity of Blo t 2 and Blo t 5 [abstract 393]. In: 58th American Academy of Allergy and Immunology Annual Meeting; 2002 Mar 1-6; New York. J Allergy Clin Immunol 2002; 109 Suppl.: 393
 31. Angus AC, Lalitha K, Kuay KT, et al. Characterization of the Group 2 Allergens from Dust Mite, *Blomia tropicalis* [abstract P17-6]. In: XVIII International Congress of Allergology and Clinical Immunology; 2003 Sep 7-12; Vancouver. Allergy Clin Immunol International Suppl 2003; 1: 188
 32. Kuay KT, Wang WL, Shang HS, et al. Molecular cloning and characterization of a Group 4 alpha-amylase *Blomia tropicalis* allergen [abstract 545]. In: 59th American Academy of Allergy and Immunology Annual Meeting; 2003 Mar 7-12; Denver (CO). J Allergy Clin Immunol 2003; 111 Suppl.: 204
 33. Shang HS, Ong ST, Bi XZ, et al. Cloning and immuno-characterization of complete Group 9 allergens from 3 major dust mites, *Blomia tropicalis*, *Dermatophagoides farinae* and *Dermatophagoides pteronyssinus* [abstract 1022]. In: 59th American Academy of Allergy and Immunology Annual Meeting; 2003 Mar 7-12; Denver (CO). J Allergy Clin Immunol 2003; 111 Suppl.: 324
 34. Streit A, Li W, Robertson B, et al. Homologs of the Caenorhabditis elegans masculinizing gene her-1 in *C. briggsae* and the filarial parasite *Brugia malayi*. Genetics 1999; 152: 1573-84
 35. Chew FT, Yi FC, Chua KY, et al. Allergenic differences between the domestic mites *Blomia tropicalis* and *Dermatophagoides pteronyssinus*. Clin Exp Allergy 1999; 29: 982-9
 36. Ong ST, Angus AC, Tsai LC, et al. Molecular cloning and characterization of an ovalbumin (Gal d 2) homologue in house dust mite, *Dermatophagoides farinae* [abstract 1025]. In: 59th American Academy of Allergy and Immunology Annual Meeting; 2003 Mar 7-12; Denver (CO). J Allergy Clin Immunol 2003; 111 Suppl.: 324-5
 37. Shang HS, Kuay KT, Wang WL, et al. Allergenicity of homologs to fungal, pollen, food and insect allergens in dust mite, *Blomia tropicalis*: identification of new classes of pan-allergens [abstract O-16-5]. In: XVIII International Congress of Allergology and Clinical Immunology; 2003 Sep 7-12; Vancouver, Canada. Allergy Clin Immunol International Suppl 2003; 1: 32
 38. Angus AC, Chua SC, Wun ST, et al. Patterns of allergic sensitisation and cross-reactivity between *Blomia tropicalis* and *Dermatophagoides farinae*: a comparative study [abstract 624]. In: 58th American Academy of Allergy and Immunology Annual Meeting; 2002 Mar 1-6; New York City, USA. J Allergy Clin Immunol 2002; 109 Suppl.: 624
 39. Loo AHB, Tan SPL, Angus AC, et al. Genetic relationship between allergy-causing dust mites: phylogenetic inference from random amplified polymorphic DNA (RAPD) markers, housekeeping gene (18S rDNA) and group 2 allergens [abstract 373]. In: 59th American Academy of Allergy and Immunology Annual Meeting; 2003 Mar 7-12; Denver, Colorado, USA. J Allergy Clin Immunol 2003; 111 Suppl.: 162
 40. Angus AC, Wang XS, Mari A, et al. Association between sensitization to specific dust mites allergens and manifestation of allergic diseases in four different populations [abstract 1023]. In: 59th American Academy of Allergy and Immunology Annual Meeting; 2003 Mar 7-12; Denver (CO). J Allergy Clin Immunol 2003; 111 Suppl.: 324
 41. Valenta R. The future of antigen-specific immunotherapy of allergy. Nat Rev Immunol 2002; 2: 446-53
 42. Westritschnig K, Valenta R. Can we genetically engineer safer and more effective immunotherapy reagents? Curr Opin Allergy Clin Immunol 2003; 3: 495-500
 43. Valenta R, Kraft D. From allergen structure to new forms of allergen-specific immunotherapy. Curr Opin Immunol 2002; 14: 718-27
 44. Epton MJ, Dilworth RJ, Smith W, et al. Sensitisation to the lipid-binding apolipoprotein allergen Der p 14 and the peptide Mag-1. Int Arch Allergy Immunol 2001; 124: 57-60
 45. Aki T, Kodama T, Fujikawa A, et al. Immunochemical characterization of recombinant and native tropomyosins as a new allergen from the house dust mite, *Dermatophagoides farinae*. J Allergy Clin Immunol 1995; 96: 74-83
 46. Tsai LC, Chao PL, Shen HD, et al. Isolation and characterization of a novel 98-kd *Dermatophagoides farinae* mite allergen. J Allergy Clin Immunol 1998; 102: 295-303
 47. O'Neill GM, Donovan GR, Baldo BA. Cloning and characterization of a major allergen of the house dust mite, *Dermatophagoides pteronyssinus*, homologous with glutathione S-transferase. Biochim Biophys Acta 1994; 1219: 521-8
 48. Caraballo L, Puerta L, Jimenez S, et al. Cloning and IgE binding of a recombinant allergen from the mite *Blomia tropicalis*, homologous with fatty acid-binding proteins. Int Arch Allergy Immunol 1997; 112: 341-7
 49. Aki T, Fujikawa A, Wada T, et al. Cloning and expression of cDNA coding for a new allergen from the house dust mite, *Dermatophagoides farinae*: homology with human heat shock cognate proteins in the heat shock protein 70 family. J Biochem 1994; 115: 435-40
 50. Daul CB, Slattery M, Reese G, et al. Identification of the major brown shrimp (*Penaeus aztecus*) as the muscle protein tropomyosin. Int Arch Allergy Clin Immunol 1994; 105: 49-55
 51. Shanti KN, Martin BN, Nagpal S, et al. Identification of tropomyosin as the major shrimp allergen and characterization of its IgE binding epitopes. J Immunol 1993; 151: 5354-63
 52. Leung PSC, Chu KH, Chow WK, et al. Cloning, expression and primary structure of *Metapenaeus ensis* tropomyosin, the major heat-stable shrimp allergen. J Allergy Clin Immunol 1994; 92: 837-45
 53. Santos ABR, Tobias KR, Ferriani VPL, et al. Identification of tropomyosin from *Periplaneta americana* as a major cockroach allergen. J Allergy Clin Immunol 1999; 103 (1 Pt 2): S122
 54. Asturias JA, Arilla MC, Gomez-Bayon N, et al. Molecular characterization of American cockroach tropomyosin (Per a 7), a cross-reactive allergen. J Immunol 1999; 162: 4342-8
 55. Leung PSC, Chen YC, Mykles DL, et al. Molecular identification of the lobster muscle protein tropomyosin as a seafood allergen. Mol Mar Biol Biotechnol 1998; 7: 12-20

56. Miyazawa H, Fukamachi H, Inagaki Y, et al. Identification of the first major allergen of a squid (*Todarodes pacificus*). *J Allergy Clin Immunol* 1996; 98: 948-53
57. Martinez A, Martinez J, Palacios R, et al. Importance of tropomyosin in the allergy to household arthropods: cross reactivity with their invertebrate extracts. *Allergol Immunopathol (Madr)* 1997; 25: 118-26
58. Lee NH, Weinstock KG, Kirkness EF, et al. Comparative expressed sequence tag analysis of differential gene expression profiles in PC-12 cells before and after nerve growth factor treatment. *Proc Natl Acad Sci U S A* 1995; 92: 8303-7
59. Santos TM, Johnston DA, Azevedo V, et al. Analysis of the gene expression profile of *Schistosoma mansoni cercariae* using the expressed sequence tag approach. *Mol Biochem Parasitol* 1999; 103: 79-97
60. Ablett E, Seaton G, Scott K, et al. Analysis of grape ESTs: global gene expression patterns in leaf and berry. *Plant Sci* 2000; 159: 87-95
61. Blaxter M, Aslett M, Guiliano D, et al. The Filarial Genome Project: parasitic helminth genomic. *Parasitology* 1999; 118: S39-51
62. Colloff MJ, Spiekma FT. Pictorial keys for the identification of domestic mites. *Clin Exp Allergy* 1992; 22: 823-30
63. Arlian LG. Water balance and humidity requirements of house dust mites. *Exp Appl Acarol* 1992; 16: 15-35
64. DeRisi JL, Iyer VR, Brown PO. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 1997; 278: 680-6
65. Hart BJ. Life cycle and reproduction of house-dust mites: environmental factors influencing mite populations. *Allergy* 1998; 53 (48 Suppl.): 13-7
66. Cunnington AM. The effect of physical conditions on the development and increase of some important storage mites. *Ann Appl Biol* 1976; 82: 175-8
67. Tovey ER, Marks G. Methods and effectiveness of environmental control. *J Allergy Clin Immunol* 1999; 103: 179-91
68. Green P, Lipman D, Hittier L, et al. Ancient conserved regions in new gene-sequences and the protein databases. *Science* 1993; 259: 1711-6
69. Delseny M, Cooke R, Raynal M, et al. The Arabidopsis thaliana cDNA sequencing project. *FEBS Lett* 1997; 405: 129-32
70. Yan T, Gong ZY. Assembly of a complete zebrafish mitochondrial 16S rRNA gene from overlapping expressed sequence tags. *DNA Seq* 1998; 9: 145-8
71. Navajas MJ, Lagnel J, Gutierrez J, et al. Species wide homogeneity of nuclear ribosomal ITS2 sequences in the spider mite *Tetranychus urticae* contrasts with extensive mitochondrial COI polymorphism. *Heredity* 1998; 80: 742-52
72. Shang HS, Ong ST, Gulzar Mohd R, et al. Comparison of the Group 7 allergens from two major dust mites, *Blomia tropicalis* and *Dermatophagoides farinae* [abstract P17-11]. In: XVIII International Congress of Allergology and Clinical Immunology; 2003 Sep 7-12; Vancouver. *Allergy Clin Immunol International Suppl* 2003; 1: 190
73. Ong ST, Gulzar Mohd R, Lua BL, et al. Differential IgE binding patterns in five polymorphic Group 13 allergens identified from dust mites, *Dermatophagoides farinae* and *Blomia tropicalis* [abstract P17-12]. In: XVIII International Congress of Allergology and Clinical Immunology; 2003 Sep 7-12; Vancouver, Canada. *Allergy Clin Immunol International Suppl* 2003; 1: 190
74. Chua KY, Kehal PK, Thomas WR. Sequence polymorphisms of cDNA clones encoding the mite allergen Der p 1. *Int Arch Allergy Immunol* 1993; 101: 364-8
75. Chua KY, Huang CH, Shen HD, et al. Analysis of sequence polymorphisms of a major mite allergen Der p 2. *Clin Exp Allergy* 1996; 26: 829-37
76. Yuuki T, Okumura Y, Okudaira H. Genomic organization and polymorphisms of the major house dust mite allergen Der f 2. *Int Arch Allergy Immunol* 1997; 112: 44-8
77. Smith WA, Thomas WR. Sequence polymorphisms of the Der p 3 house dust mite allergen. *Clin Exp Allergy* 1996; 26: 571-9
78. Smith WA, Thomas WR. Comparative analysis of the genes encoding group 3 allergen from *Dermatophagoides pteronyssinus* and *Dermatophagoides farinae*. *Int Arch Allergy Immunol* 1996; 109: 133-40
79. Lin KL, Chua KY, Thomas WR, et al. Characterisation of Der p 5 allergen: cDNA analysis and IgE-mediated reactivity to the recombinant protein. *J Allergy Clin Immunol* 1994; 94: 989-96
80. Schmidt M, Olsson S, van der Ploeg I, et al. cDNA analysis of the mite allergen Lep d 1 identifies two different isoallergens and variants. *FEBS Lett* 1995; 370: 11-4
81. Kronqvist M, Johansson E, Magnusson CGM, et al. Skin prick test and serological analysis with recombinant group 2 allergens of the dust mites *L. destructor* and *T. putrescentiae*. *Clin Exp Allergy* 2000; 30: 670-6
82. Kaiser L, Gafvelin G, Johansson E, et al. Lep d 2 polymorphisms in wild and cultured *Lepidoglyphus destructor* mites. *Eur J Biochem* 2003; 270: 646-53
83. Thomas WR, Smith W, Hales BJ, et al. Functional effects of polymorphism in house dust mite allergens. *Int Arch Allergy Immunol* 1997; 112: 96-8
84. Bowcock A, Cavalli-Sforza L. The study of variations in the human genome. *Genomics* 1997; 11: 491-8
85. Shi MM. Enabling large-scale pharmacogenetic studies by high-throughput mutation detection and genotyping technologies. *Clin Chem* 2001; 47: 164-72
86. Le Mao J, Mayer CE, Peltre G, et al. Mapping of dermatophagoides farinae mite allergens by two-dimensional immunoblotting. *J Allergy Clin Immunol* 1998; 102: 631-6
87. Bi XZ, Angus AC, Ong ST, et al. The dust mite proteome: what are the major components of extracts apart from the allergens? [abstract O-1-5]. In: XVIII International Congress of Allergology and Clinical Immunology; 2003 Sep 7-12; Vancouver. *Allergy Clin Immunol International Suppl* 2003; 1: 2
88. Bi XZ, Ong ST, Chew FT. Identification, cloning and expression of a major component of the house dust mite proteome, a putative cuticle-like protein, via peptide mass fingerprinting, tandem mass spectrometry and contiguous expressed sequence tag alignments [abstract 387]. In: 59th American Academy of Allergy and Immunology Annual Meeting; 2003 Mar 7-12; Denver (CO). *J Allergy Clin Immunol* 2003; 111 Suppl.: 165
89. Reginald K, Haroon-Rashid L, Sew YS, et al. Identification of putative *Tyrophagus putrescentiae* allergens with sequence homology to other known allergens by expressed sequence tagging. *Allergy* 2002; 57 Suppl. 73: 286-7
90. Gao YF, Chew FT. Large scale screening of putative allergen genes from *Acarus siro* [abstract 1030]. In: 59th American Academy of Allergy and Immunology Annual Meeting; 2003 Mar 7-12; Denver (CO). *J Allergy Clin Immunol* 2003; 111 Suppl.: 326

Correspondence and offprints: Dr Fook Tim Chew, Department of Biological Sciences, National University of Singapore, Functional Genomic Laboratories, Block S3, Level 2, Science Drive 4, off Lower Kent Ridge Road, 14, 117543, Singapore.
E-mail: dbscft@nus.edu.sg