

Identification of putative innate immune related genes from a cell line of the mosquito *Aedes albopictus* following bacterial challenge

Rajnikant Dixit, Millind S. Patole, Yogesh S. Shouche

Molecular Biology Unit, National Center for Cell Science, Ganeshkhind, Pune, India

We report identification of putative innate immune related genes from a cell line of the mosquito *Aedes albopictus* challenged with heat-killed bacteria. Using a subtractive hybridization and sequencing approach, we analyzed a total 309 expressed sequence tags (ESTs) which clustered in 40 contigs. Thirty-five percent of genes yielded homology to known immune genes corresponding to antimicrobial peptides (AMPs), pathogen-associated molecular patterns, protease and immune signaling cascades. Interestingly, most of the genes have not been previously described from this mosquito and thus represent a class of novel immune genes. Further, 25% sequences did not match to any known species in the non-redundant databases, appear to be specific to the mosquito *A. albopictus* and merit further study.

Keywords: *Aedes albopictus*, innate immunity, subtractive hybridization, mosquito, ESTs

INTRODUCTION

Innate immunity is a crucial defense mechanism against microbial infection found in all animals. The ability of a host to distinguish between self and non-self remains a central hallmark of innate immunity.¹ Insects, including mosquitoes, also mount potential innate immune reactions in response to infection by viruses, bacteria, fungi, and macroparasites.² Although, in recent years much progress has been made in the field of insect immunity in different insect models including mosquitoes (viz *Anopheles gambiae* and *Aedes aegypti*), relatively little is known in the case of other mosquitoes.^{3,4}

The Asian tiger mosquito (*Aedes albopictus*), from the mosquito family *Culicidae*, is native to the tropical and subtropical areas of Southeast Asia; however, in the past couple of decades, this species has invaded many other countries in the world through the transport of goods and increasing international travel.⁵ This mosquito has become a significant pest in many communities because it closely associates with humans and, occasionally, an epidemic vector of dengue viruses in parts of Asia; it is also a competent vector of several other viruses under

experimental conditions. However, in comparison of *A. aegypti*, mosquito *A. albopictus* has been least studied in respect of the molecular immune response against microbial pathogens, though both species transmit similar diseases.

Currently, we are exploring the possibility of identifying immune-related genes from different vector mosquito systems (Sharma and Dixit, unpublished work).⁶ In the present context, we were interested in identifying immune-related genes that are differentially up-regulated during bacterial infection, using an *A. albopictus* cell line model system. Here, we report identification of several innate immunity and unknown genes, isolated from subtractive hybridization library of the immune-challenged mosquito cell line C6/36 (Igarashi).

MATERIALS AND METHODS

Mosquito cell line and immune challenge

The *A. albopictus* cell line C6/36 was grown in minimal essential medium (MEM) containing non-essential amino-acids and 10% fetal calf serum (FCS) at 28°C,

Received 27 June 2009; Revised 11 August 2009; Accepted 8 September 2009

Correspondence to: Rajnikant Dixit, Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, MD 20852, USA.

Tel: +1 301 435 3182; Fax: +1 301 480 1337; E-mail: dixit2k@yahoo.com

in a 5% CO₂ atmosphere. Next, these cells were plated, grown to form a nearly confluent monolayer in MEM in a 100-mm dish and washed in phosphate buffered saline. These cells were re-fed with MEM containing the heat-killed, Gram-negative bacterium *Pseudomonas aeruginosa* ($\sim 1 \times 10^9$ cells/ml). Control cells were treated under identical conditions, without bacterial exposure.

Total RNA preparation and cDNA synthesis

Total RNA was isolated by the standard Trizol method and mRNA was prepared using oligo(dT) magnetic beads. Isolated mRNA was treated with RNase-free DNase I to remove any genomic DNA contamination. DNase-treated mRNA (approximately 500 ng) was then reverse-transcribed, and PCR amplified to make double-stranded cDNA (ds-cDNA) through a PCR-based protocol using SMART cDNA synthesis kit (catalogue no. 634902; BD Clontech, Palo Alto, CA, USA).

Construction of the subtracted cDNA library

To generate a subtracted cDNA library, a suppression subtractive hybridization protocol was performed using the PCR Select cDNA subtraction kit (catalogue no. 637401; BD Clontech), following the manufacturer's protocol. Briefly, 500 ng mRNA from immune challenged (tester) and control (driver) were used to prepare ds-cDNAs as described above. Then ds-cDNA was purified and digested with *RSaI* enzyme. The tester cDNA was divided into two aliquots, each of which was ligated to one of two different adaptors – adaptor 1 or 2R. Both adaptor-ligated samples were denatured at 98°C for 90 s and then hybridized at 68°C for 8 h with an excess of driver cDNA. These two primary hybridized samples were mixed together without denaturing, and fresh denatured driver cDNA was added to enrich further differentially expressed genes by hybridizing at 68°C for 10 h. The entire population of cDNA molecules was then subjected to two rounds of suppression PCR for selective amplification of the differential transcripts. The resulting secondary PCR products were cloned into pGEM-T vector (Promega) and transformed to *Escherichia coli* (DH5a) competent cells. The cloned inserts were PCR-amplified using vector-specific M13 primers and sequenced with M13F primer in an ABI3730 sequencer.

Bioinformatic analysis of the databases

Each individual sequence chromatogram was examined manually. Vector- and end-specific low quality sequences were removed computationally from all sequences, using vector database-specific BLAST.

Then each individual sequence (FASTA FORMAT) was searched against non-redundant and other EST databases available at the National Center for Biotechnology and Information (NCBI) using BLASTN and BLASTX algorithms. The ESTs were assembled into contigs using CAP assembly program (BioEdit, v.7.0.8). Along with BLASTX analysis, different protein database (GO, KOG, PFAM, *etc.*) were used to determine the putative nature and function of the genes. Based on the predicted function of the homologous protein and significant BLASTX similarities, contigs were grouped into a relative family of proteins. ClustalX and GenDoc software were used for multiple sequence alignment and sequence identity analysis. All ESTs have been submitted to the NCBI GenBank (accession numbers from GR881567 to GR881877) for public domain use.

RESULTS AND DISCUSSION

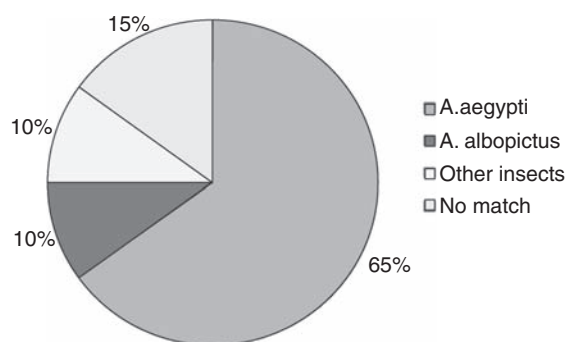
Recently, in *Drosophila*, it has been shown that the Toll pathway participates in resistance to infection by *P. aeruginosa*, an opportunistic virulent pathogen that exhibits an extremely broad host range including vertebrates, insects, nematodes and plants.^{7–10} However, in the mosquito, few studies have been done to examine the immune responses against this bacterial species, which appears to be independent of the other common bacterial species used in previous studies.¹¹

Therefore, to identify differentially expressed immune-responsive genes against this bacterial species, we first challenged an *A. albopictus* cell line with a heat-killed bacterial suspension of *P. aeruginosa*. Empirically, we decided to collect RNA samples at two time points – 2 h (early response) and 24 h (late response) – post challenge and generated a subtractive hybridization cDNA library as described above. To minimize the risk of PCR artifacts, we optimized the minimum PCR cycle numbers to amplify initial cDNA populations. A total of $\sim 10^6$ primary clones were obtained for the subtracted cDNA library. About 320 randomly selected positive recombinant clones (white colonies) were subjected to high throughput sequencing and 309 good quality sequences were recovered. These sequences were further analyzed to generate an EST catalogue for innate immune genes presumably activated in response to bacterial response. Table 1 represents the primary features of the subtraction cDNA library.

All ESTs could be assembled into 40 contigs, out of which 22 sequences remained unassembled and grouped as singleton. To determine the homology of the assembled contigs, sequences were further analyzed using the BLASTX search engine against NR database at NCBI. Only 10% of genes yielded homology to the

Table 1. Characteristic of subtraction cDNA library

Total number of positive recombinant clones sequenced	320
Good quality ESTs analyzed	309
Total number of clusters assembled	40
Average cluster length (bp)	428
Unique sequences (singleton)	55%

**Fig. 1.** EST homology to insect species in NR database.

mosquito *A. albopictus* while 65% of genes showed maximum homology to the mosquito *A. aegypti*. Another 10% sequences showed homology to other insect species while remaining 15% did not yield any significant homology to any species (Fig. 1).

To examine the putative nature and functions of the identified genes, first each cluster/contig was translated into six reading frames and then searched against available protein family databases (KOG, GO, PFAM, etc.). With this information, the assembled contigs were assigned putative functions based on close relatedness within a family of the known proteins in the available database (Table 2) and grouped in three major categories: (i) putative immune-related genes (35%); (ii) abundant genes (15% energy and metabolism, 20% protein synthesis and trafficking, 5% cell structure and organization related); and (iii) unknown/novel sequences (25%), which did not match any known species in the non-redundant databases but appear to be specific to the mosquito *A. albopictus* (Fig. 2).

Putative immunity-related genes

Mosquito innate immunity functions have been correlated in pattern recognition, activation of signaling cascades that initiate varied cytotoxic effector molecular responses to delimit the effect of foreign invaders. As a framework to examine mosquito *A. albopictus* general immune response, we identified 14 clusters (35%) coding for immunity-related proteins. These included

the antimicrobial peptide defensin, melanization cascade related family proteins viz. fibronectins, serine proteases and inhibitor, macroglobulins and clusters of signaling and cell adhesion/immunity proteins (Table 2). Interestingly, most of these genes have not been previously described from this mosquito and thus represent a class of novel immune genes. The evolutionary conserved nature of innate immune responsiveness to microbes makes this information pertinent to those interested in the biology of vector-pathogen-vertebrate host relationships.¹²

Antimicrobial peptides

The AMP-effector molecules that exhibit *in vitro* activities against microbes are considered a primary defense element in mosquito innate immunity. Transcriptional up-regulation of AMPs has been correlated not only with responses to bacteria and fungi but also with various stages of *Plasmodium* infection in the mosquitoes *Anopheles gambiae*¹³ and *Anopheles stephensi*.⁶

Defensin

Insect antimicrobial peptides can be assigned to three major classes – defensins, cecropins and proline/glycine-rich peptides. Although most of these antimicrobial peptides attack Gram-negative bacteria, defensins, characterized by strict conservation of six-cysteine residues, act primarily against Gram-positive bacteria.¹⁴ Interestingly, in the present study, we identified two identical cDNA sequences (466 bp) capable of encoding an unusual 123 amino-acid long defensin-like peptide which may be active against the Gram-negative bacterium *P. aeruginosa* similar to the previously reported human β -Def 2.¹⁵

Using SignalIP v.3.0 online software and comparative analysis, we predicted four distinct regions: 25 residues long unique extension upstream to signal peptides; 21 residues of signal peptide, 37 residues of prodefensin, and 40 residues of mature defensin from the deduced amino-acid sequence. Amino-acid composition analysis of this 123 amino-acid long defensin peptide showed relatively higher acidic properties (isoelectric point 6.07) compared to other insect basic defensins.

Multiple alignment analysis showed an overall maximum 73% similarity and strictly maximum identity of 97% to the mature peptide to the previously reported defensin D (gil3641324) from *A. albopictus*, whose expression was maximal at 12–36 h after Gram-positive bacterial challenge.¹⁶ Therefore, it would be interesting to examine the role and mechanism of this acidic defensin against bacterial infection. Phylogenetic analysis indicated an independent separation of *Aedine* mosquitoes from *Anopheline* defensins (Fig. 3)

Table 2. Molecular catalogue of *A. albopictus* genes identified from subtractive hybridization library

Assembled contig	Length	EST/contig	Best match to NR	e-Value	Sequence ID	% Identity	Species	Best match to PFAM	e-Value	Comments
Contig_1	839	105	Hypothetical protein AaeL_AAEL005331	3.00E-26	gbIEAT43232.11	86	<i>Aedes aegypti</i>			Function unknown
Contig_2	811	118	Unknown							Unknown
Contig_3	466	2	Defensin D	5.00E-47	gbIAAC36346.1	97	<i>Aedes albopictus</i>	Defensin_2	2.00E-11	Immunity
Contig_4	146	1	Unknown							Unknown
Contig_5	378	2	Thioredoxin	3.00E-32	gbIAAK70900.1	94	<i>Aedes aegypti</i>	Thioredoxin	2.00E-22	Metabolism
Contig_6	665	17	Unknown							Unknown
Contig_7	441	1	Macroglobulin/complement	4.00E-46	gbIEAT47037.1	60	<i>Aedes aegypti</i>	A2M	3.00E-04	Immunity
Contig_8	419	4	Macroglobulin/complement	9.00E-55	gbIEAT47037.1	75	<i>Aedes aegypti</i>	A2M	2.00E-07	Immunity
Contig_9	428	1	Elongation factor-1 alpha	8.00E-78	gbIABF18239.11	100	<i>Aedes aegypti</i>	GTP_EFTU	2.00E-19	Protein synthesis
Contig_10	459	1	GNBP-like protein	1.00E-77	gbIEAT40654.11	94	<i>Aedes aegypti</i>	Glyco-hydro-16	1.00E-04	Immunity
Contig_11	349	1	hypothetical protein AaeL_AAEL005638	2.00E-53	gbIEAT42879.1	92	<i>Aedes aegypti</i>	Ribosomal_L7		Protein synthesis
Contig_12	255	1	signal recognition particle	5.00E-32	gbIEAT45339.11	82	<i>Aedes aegypti</i>	SRP72		Signaling & immunity
Contig_13	366	1	Eukaryotic translation initiation factor	5.00E-48	gbIEAT34061.1	85	<i>Aedes aegypti</i>	S1	6.00E-12	Protein synthesis
Contig_14	315	1	14-3-3 protein sigma, gamma, zeta, beta/alpha	1.00E-39	gbIABF18291.1	100	<i>Aedes aegypti</i>	14-3-3	5.00E-49	Protein synthesis
Contig_15	737	1	Mitochondrial ATP synthase lipid binding protein precursor	2.00E-53	gbIAAV90735.1	98	<i>Aedes albopictus</i>	Ribosomal_L2	3.00E-11	Metabolism
Contig_16	513	10	Fibrinogen and fibronectin	3.00E-84	gbIEAT39569.1	98	<i>Aedes aegypti</i>	Fibrinogen_C	1.00E-11	Immunity
Contig_17	562	2	Trypsin, putative	1.00E-60	gbIEAT40249.1	67	<i>Aedes aegypti</i>	Trypsin	1.00E-17	Immunity
Contig_18	528	2	Hypothetical protein AaeL_AAEL004809	5.00E-58	gbIEAT43758.1	79	<i>Aedes aegypti</i>			Function unknown

(continued)

Table 2. Continued

Assembled contig	Length	EST/contig	Best match to NR	e-Value	Sequence ID	% Identity	Species	Best match to PFAM	e-Value	Comments
Contig_19	280	2	Serine protease	7.00E-44	gb EAT33372.1	86	<i>Aedes aegypti</i>	SP	1.00E-07	Immunity
Contig_20	595	2	Ribosomal protein S25, putative	1.00E-36	gb EAT48945.1	68	<i>Aedes aegypti</i>	Ribosomal_S25	6.00E-41	Protein synthesis
Contig_21	449	3	Transferrin precursor	1.00E-80	gb AAL58079.1	85	<i>Aedes aegypti</i>	Tyrosinase	8.30E-02	Metabolism
Contig_22	397	1	Unknown							
Contig_23	367	1	Putative carboxylesterase	2.00E-58	gb AAL76013.1	81	<i>Aedes aegypti</i>	COesterase	1.00E-04	Metabolism
Contig_24	261	1	Novel protein [Danio rerio]	4.00E-41	emb CAQ15541.1	97	Danio rerio	Ubiquitin	3.00E-24	Cell structure & organization
Contig_25	253	1	Peptidoglycan recognition protein-1c isoform	3.00E-27	gb EAT38640.1	76	<i>Aedes aegypti</i>	Amidase_2	6.00E-14	Immunity
Contig_26	315	5	Fibrinogen & fibronectin	1.00E-56	gb EAT39569.1	87	<i>Aedes aegypti</i>	Fibrinogen_C	6.00E-31	Immunity
Contig_27	455	1	Unknown							Unknown
Contig_28	166	1	60S ribosomal protein L18a	2.00E-19	gb AAV90708.1	100	<i>Aedes albopictus</i>	Ribosomal_L18a	4.00E-17	Protein synthesis
Contig_29	338	1	Coatomer subunit epsilon	3.00E-36	gb EDS43412.1	89	<i>Aedes aegypti</i>	Coatomer_E	7.00E-16	Protein trafficking
Contig_30	417	1	mrna turnover protein 4 mrt4	5.00E-70	gb EAT41471.1	94	<i>Aedes aegypti</i>			Protein synthesis
Contig_31	329	2	Serine protease inhibitor, serpin	5.00E-53	gb EAT42828.1	95	<i>Aedes aegypti</i>	Serpin	4.00E-17	Immunity
Contig_32	395	2	Serine protease	1.00E-40	gb EAT43471.1	63	<i>Aedes aegypti</i>	Trypsin	2.00E-11	Immunity
Contig_33	311	1	Unknown							Unknown
Contig_34	196	1	Midline fasciclin	2.00E-27	gb EAT39559.1	96	<i>Aedes aegypti</i>	Fasciclin	3.00E-04	Cell adhesion & immunity
Contig_35	249	1	Actin-5	2.00E-40	gb EDS37498.1	100	<i>Culex pipiens quinquefasciatus</i>	Actin	3.00E-46	Cell structure & organization
Contig_36	644	1	Unknown							Unknown
Contig_37	555	2	Unknown							Unknown
Contig_38	451	3	NADH dehydrogenase subunit 4	6.00E-46	gb ABY51631.1	65	<i>Aedes aegypti</i>	Oxidored_q1	1.00E-35	Metabolism
Contig_39	495	4	Cytochrome b	6.00E-76	gb AAL61970.1	86	<i>Aedes albopictus</i>	Cytochrom_B	5.00E-25	Metabolism
Contig_40	512	1	Conserved hypothetical protein	9E-51	XP_001658777.1	77	<i>Aedes aegypti</i>	Teneurin-1 with EGF-like repeats	8.00E-12	Signaling & immunity

suggesting an early duplication, followed by divergence under selective evolutionary pressure, specifically to combat the wide range of micro-organisms as well as parasites that they encounter in their body.¹⁷

Pattern recognition molecules

Recognition of pathogen-associated molecular patterns initiates innate immune responsiveness in both

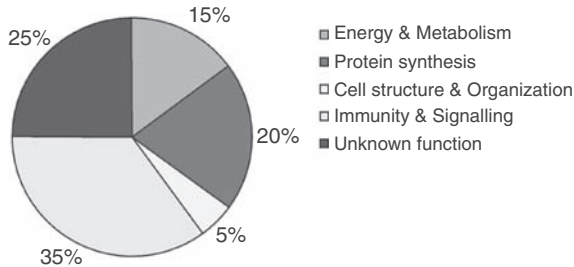


Fig. 2. Functional categories of genes identified.

vertebrates and invertebrates.¹⁰ The cascade of events begins when pathogen-associated molecular patterns (PAMPs) are bound by pattern recognition receptor (PRR) molecules. Different PRRs react with specific PAMPs, show distinct expression patterns, activate specific signaling pathways, and lead to distinct anti-pathogen responses.¹⁸ The majority of protein families of invertebrate PRRs include β -(1 \rightarrow 3)-glucan recognition proteins (β -GRPs) and Gram-negative bacterial-binding proteins (GNBPs), peptidoglycan recognition proteins (PGRPs), complement-like proteins and scavenger receptor proteins and interact with microbes to activate both cellular and humoral aspects of innate immunity.¹⁹ Here, we briefly describe the novel PRR-related genes identified from the mosquito *A. albopictus* cell line, likely induced due to bacterial challenge.

Peptidoglycan recognition proteins (PGRPs)

This family, distinguished by the PGRP domain (IPR002502), plays central and diverse roles in

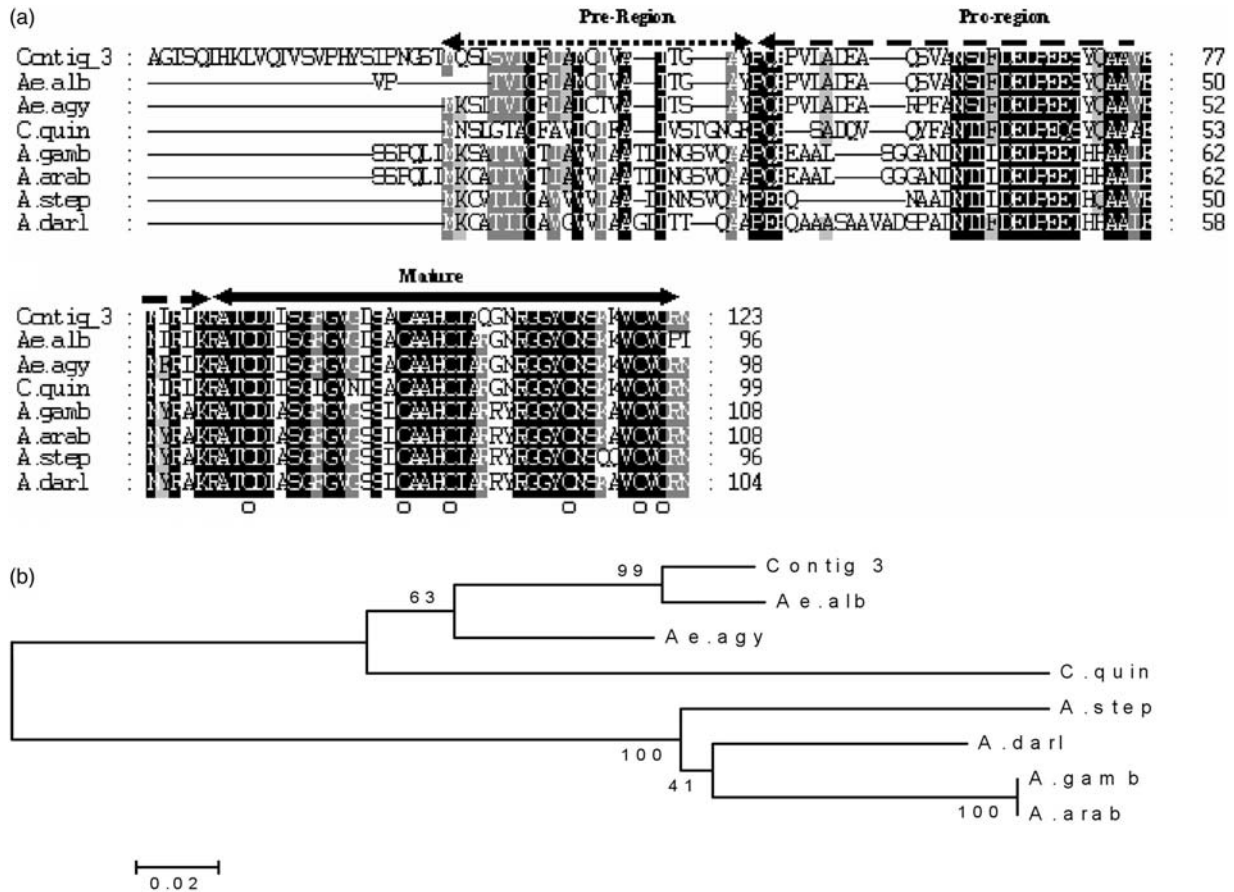


Fig. 3. Clustal alignment (a) and phylogenetic (b) analysis of known mosquito defensins. Highly conserved and similar residues are shown in black and gray, respectively. Pre and pro regions of defensin have been represented with double-headed arrows while mature region has been shown with single dark line. Cysteine residues have been marked as circle. Species names include: A.s., *Anopheles stephensi*; A.a., *Anopheles arabiensis*; A.g., *Anopheles gambiae*; A.e., *Aedes aegypti*; A.alb., *Aedes albopictus*.

activating insect immune reactions and hydrolyze peptidoglycans (PGNs) of bacterial cell walls. In the present study, we identified one cluster with a single cDNA (453 bp) capable of encoding an 84 amino-acid long peptide, which showed 76% identity to known PGRP in the mosquito *A. aegypti* (Gen bank accession gb|EAT38640.1). Primary BLASTP analysis showed the presence of putative PGRP and other domains related to amidase enzyme activity that cleaves the amide bond between *N*-acetylmuramoyl and L-amino-acids, preferentially D-lactyl-L-Ala, in bacterial cell walls (Fig. 4). Multiple amino-acid sequence alignment analysis with other known insect PGRPs clearly indicated highest homology to *Aedine* mosquitoes, while phylogenetic analysis clearly showed an early independent separation from other insects including the mosquito *Anopheles gambiae* (Fig. 4), suggesting a common functional conservation to combat different microbial infections.

Recently, it has been shown that a peptidoglycan recognition protein, PGRP-LC, is absolutely required for the induction of antibacterial peptide genes in response to infection in *Drosophila* and acts by controlling activation of the NF- κ B family transcription factor Relish.²⁰ However, so far, few studies have examined the role of PGRPs in mosquitoes. Therefore, it will be of

interest to identify and compare more PGRPs from the mosquito *A. albopictus* to understand better the adaptation values of PRR molecules towards differential responses to microbial infections.

Gram-negative bacteria binding proteins (GNBPs)

These proteins have sequence similarities to bacterial glucanases, and probably represent a case of either horizontal gene transfer or convergent evolution.²¹ Although they do not show enzymatic activity, various GNBPs are able to bind to fungal β -(1 \rightarrow 3)-glucan, bacterial lipopolysaccharides (LPSs), and/or bacterial lipoteichoic acid.²² During primary analysis of the immune-activated *A. albopictus* cell line transcriptional repertoire, we recognized a partial cDNA sequence (459 bp) capable of encoding a 152 amino-acid long peptide, which showed 94% identity to the predicted GGBP from the mosquito *A. aegypti* (gb|EAT40654.1).

Multiple alignment indicated higher conservation within the mosquito group relatively towards the C-terminus, a common feature of GNBPs, while phylogenetic analysis indicated a distant relationship with *Anopheles* mosquitoes and other insects (Fig. 5). The GGBP family includes members that are known to bind to Gram-negative bacteria, LPSs, and

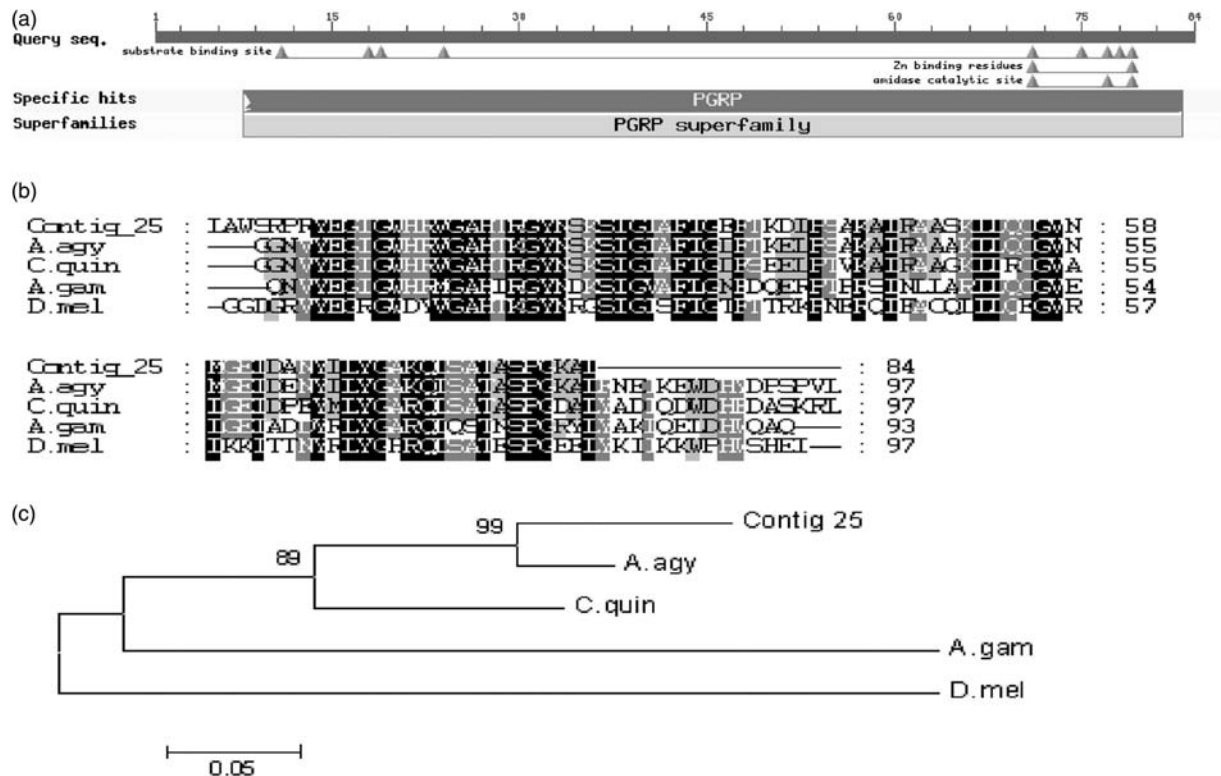


Fig. 4. Analysis of mosquito PGRPs. (a) Web-based prediction of putative PGRP and related domains in mosquito *A. albopictus* gene. Clustal alignment (b) and phylogenetic (c) analysis of mosquito PGRPs. Highly conserved and similar residues have been shown in black and gray, respectively. Species names include: Agam, *Anopheles gambiae*; Aeagy, *Aedes aegypti*; C.quin, *Culex quinquefasciatus*; D.mel, *Drosophila melanogaster*.

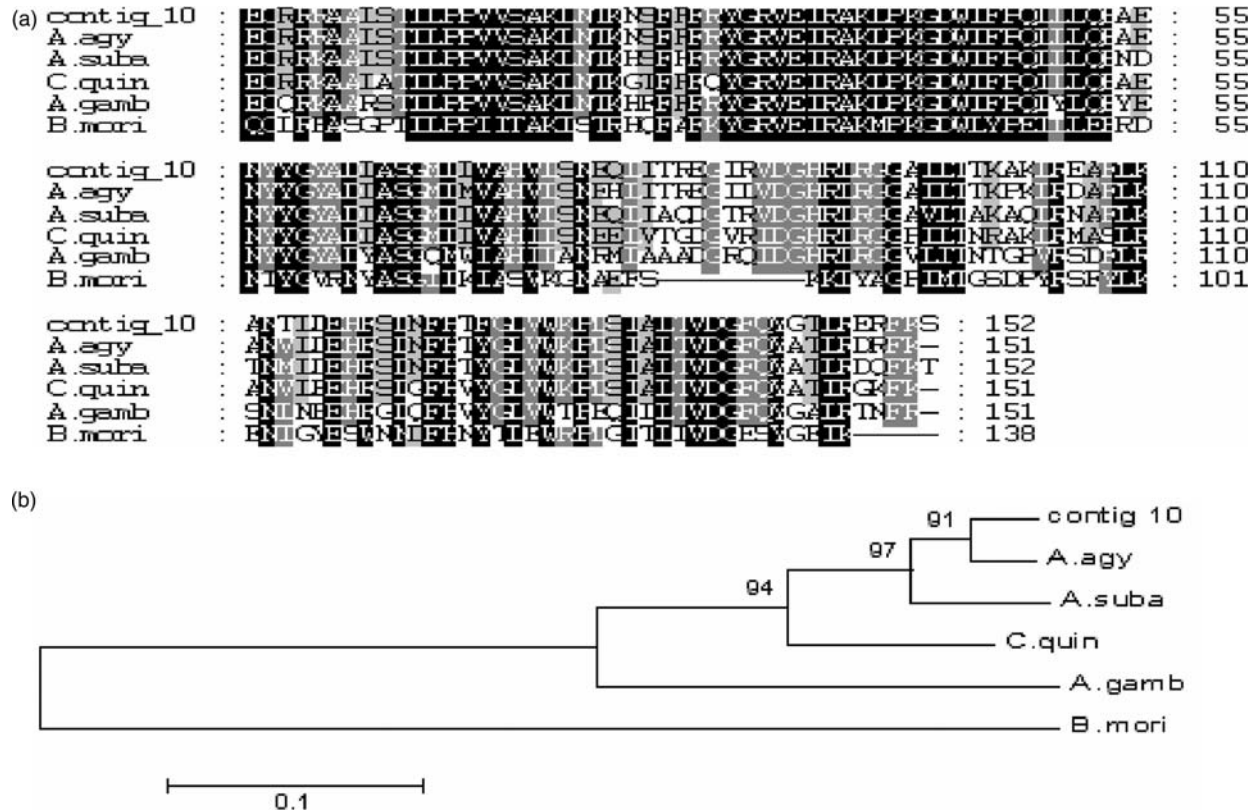


Fig. 5. Clustal alignment (a) and phylogenetic (b) analysis of known insect GNBP. Highly conserved and similar residues have been shown in black and gray, respectively. Species names include: Agam: *Anopheles gambiae*; Aeagy, *Aedes aegypti*; A.suba, *Aedes sublabatus*; C.quin, *Culex quinquefasciatus*; B.mori, *Bombyx mori*.

β -(1 \rightarrow 3)-glucan; and appears to be up-regulated by immune challenge in insects including mosquitoes.²² Recently, termite GNBP-2 (tGNBP-2) β -(1 \rightarrow 3)-glucanase effector activity has been tested as an antimicrobial effector protein and proposed as a strong molecular target to design new strategies for pest control.²³ Therefore, it would be important to determine the functional role of these family proteins in the mosquito *A. albopictus* to design vector-based control strategies of medically important diseases transmitted by mosquitoes.

Alpha-2 macroglobulin family proteins

The proteinase-binding alpha-macroglobulins (A2M) are large glycoproteins, found in the plasma of vertebrates and in the hemolymph of invertebrates, able to inhibit all four classes of proteinases by a 'trapping' mechanism.²⁴ Recent implication of the members of complement C3/ α_2 -macroglobulins superfamily proteins (TEPs) in strong anti-parasitic response, has been considered as the hallmark in the field of insect immunity.²⁵ Insect TEPs could be separated into two main groups: (i) the highly conserved orthologous trio (DmTep6/AgTep13/AaTep13); and (ii) much more divergent sequences mostly grouped in species-specific

expansions. Based on BLASTX analysis, we identified two novel partial cDNA clusters from five ESTs, encoding TEP-like proteins, which showed maximum identity of 60% and 75% homology to the TEP1 and TEP4 proteins of the mosquito *A. aegypti*, respectively (Fig. 6).

Usually, these proteins are glycosylated and secreted into the hemolymph by mosquito hemocytes and present either as a full-length form of approximately 165 kDa or as a cleaved smaller fragment of 80 kDa derived from the C-terminal region. Interestingly, TEP1 is also able to bind to Gram-positive and Gram-negative bacteria in a thioester-dependent manner and this opsonization is essential to promote the phagocytosis of bacteria by cultured mosquito cells.²⁶ Therefore, further knockout studies on the mosquito *A. albopictus* TEPs should provide an opportunity to understand better the conserved role of this family of proteins against other parasites and/or viruses transmitted by this species.

Fibrinogen-related family proteins (FREPs)

Members of the fibrinogen-related proteins (FREPs) or fibrinogen-domain immunoelectin (FBN) family contain the evolutionary conserved fibrinogen domain that is

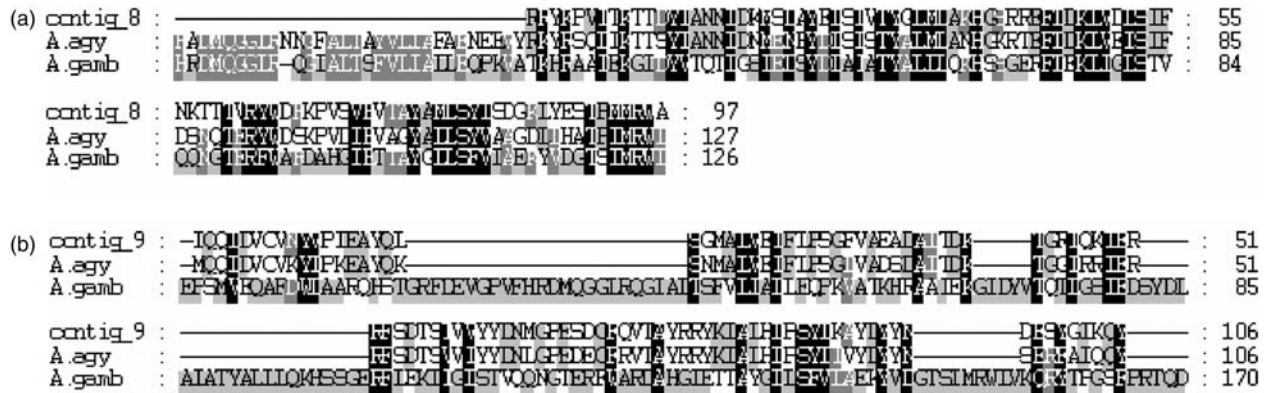


Fig. 6. Clustal alignment analysis of the mosquito TEPs, TEP-4 (a) and TEP-1 (b). Highly conserved and similar residues have been shown in black and gray, respectively. Please see text for details. Species names include: Agam: *Anopheles gambiae*; Aegy: *Aedes aegypti*.

also found in the mammalian ficolins.²⁷ The ability of FREPs to form multimers (dimers in Ag cell line supernatant) may enable an increase of the mosquito's pattern recognition receptor repertoire, since different combinations are likely to possess different binding specificities towards microbial products. Surprisingly, in the present study, we identified five contigs of FREP-related genes independently clustered from 15 ESTs. Thus exclusive expansion of this family protein prompted us to perform extensive bioinformatic analysis reported as a separate independent study.²⁸

Signal modulation and activation molecules

After recognition of infectious non-self particles, the extracellular cascades of activating serine proteases and their counter inhibitors, serpins, process the signal to amplify the immune responses involving AMP production, and melanotic encapsulation, phagocytosis, etc.

Serine protease

Among the large extracellular protease family, serine proteases that contain a 'paper clip'-like domain and are, therefore, referred as CLIP-domain serine protease (clip-SP), have been found to be involved in unique biological processes, such as immunity and development. Bioinformatic analysis and manual annotation of the *Anopheles gambiae* genome sequence defined two classes of CLIP. The CLIP-A class consists of 10 catalytically inactive putative SPH (CLIP-A1–10) and the CLIP-B class consists of 17 catalytically active enzymes (CLIP-B1–17).²⁹

In the present study, we recognized two novel cDNA contigs (each clustered from two ESTs) encoding proteins homologous to CLIP proteases of the mosquitoes *A. aegypti* and *Anopheles gambiae* (Table 2). Further bioinformatic and comparative analysis indicated that both belong to the CLIP-B class of serine proteases. Multiple alignment and phylogenetic

analysis of the longest read indicated that it is highly homologous to *Aedine* mosquitoes and distantly related to *Anopheline* mosquito species (Fig. 7), reflecting an evolutionary conserved function within mosquito species.

A recent report of an immune-responsive serine protease and its transcript abundance in naïve adult refractory mosquitoes, even in the absence of any immune challenge, from a field-collected refractory strain of *Anopheles culicifacies*, the main vector of malaria in India suggests the possibility of its contribution towards refractory phenotype to the *Anopheles culicifacies* population.³⁰ Therefore, it is of particular interest to examine further the role of these CLIP-B proteases in the immunity towards bacterial and other infectious agents like viruses in the mosquito *A. albopictus*.

Serine protease inhibitors (serpins)

Serpins are a very large and highly conserved family of serine protease inhibitors, found in all higher eukaryotes as well as viruses, with a wide range of biological functions such as blood-clotting and fibrinolytic cascades, inflammation and complement activation, tumor suppression, extracellular matrix maintenance or remodeling and apoptosis.³¹ They can be found intra- as well as extracellularly, and are usually 350–400 amino-acid residues long with a reactive center loop (RCL) that is located 30–40 residues from the C-terminal end.

In this study, we identified a novel partial cDNA sequence contig (with two ESTs), encoding a putative peptide homologous to predicted *A. aegypti* serine protease inhibitor. Through conserved domain analysis, we predicted a reactive center loop in the *A. albopictus* SRPN, a primary feature of this family of proteins (Fig. 8). This RCL binds to the active site of the specific target proteinase similar to the binding of a substrate. A high degree of conservation was observed with all

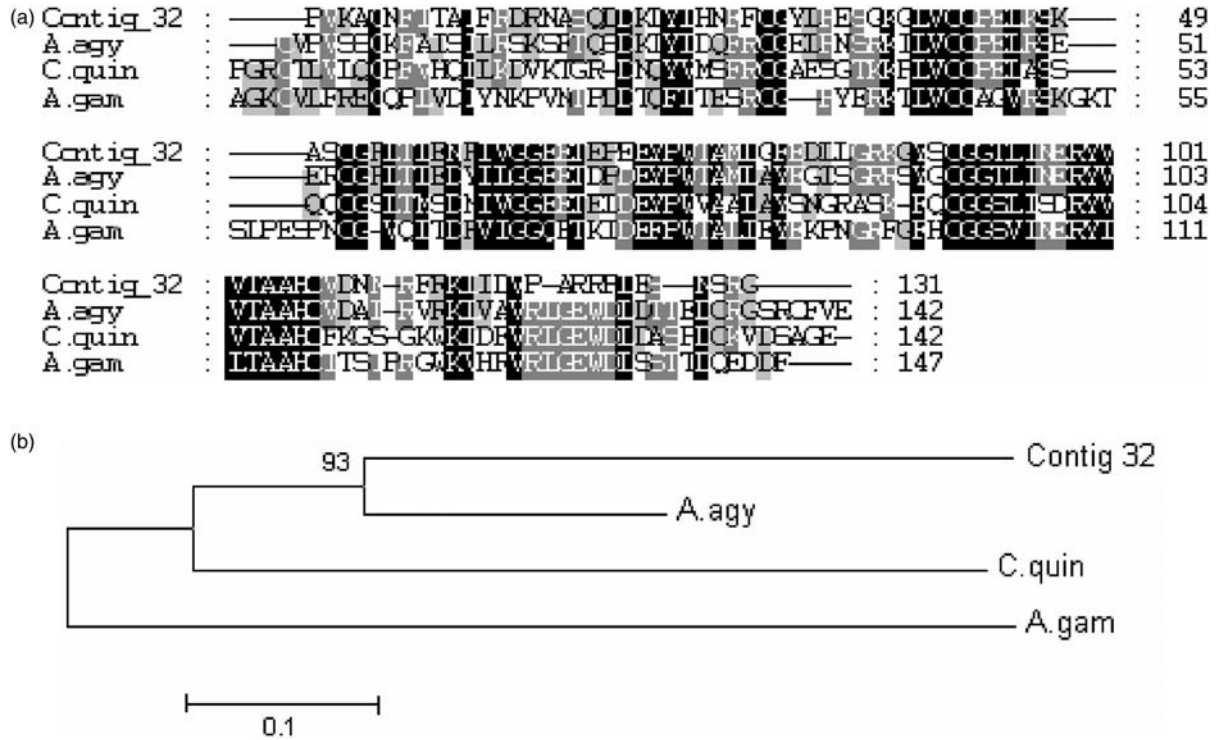


Fig. 7. Clustal alignment (a) and phylogenetic (b) analysis of known mosquito serine proteases (CLIP-B). Highly conserved and similar residues have been shown in black and gray, respectively. Species name includes: Agam, *Anopheles gambiae*; Aeagy, *Aedes aegypti*; A.suba, *Aedes sublabatus*; C.quin, *Culex quinquefasciatus*.

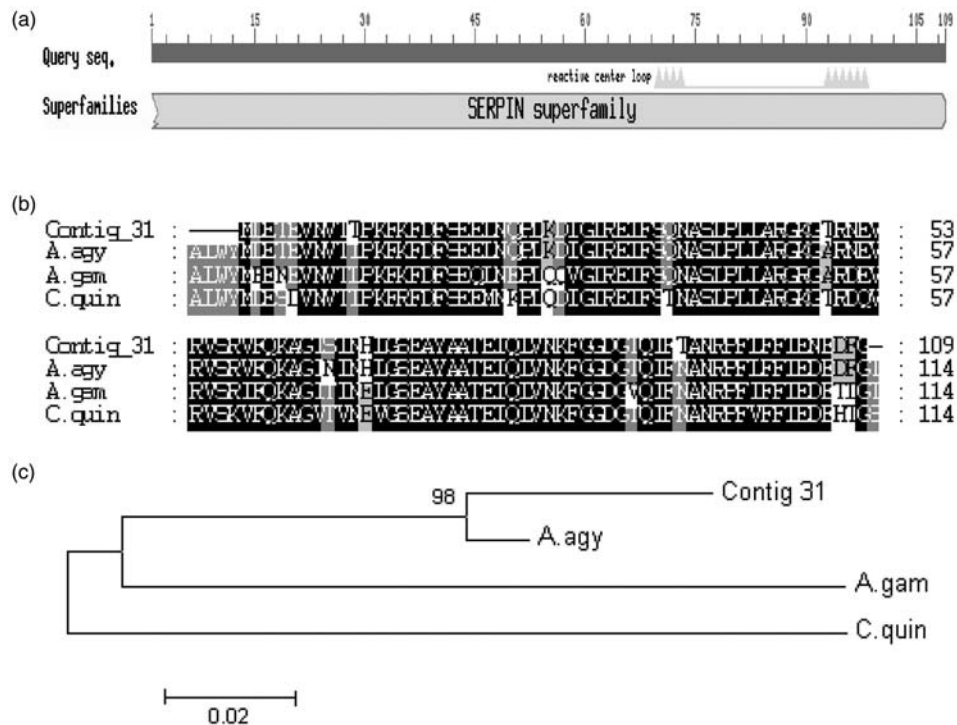


Fig. 8. Analysis of mosquito serpins. (a) Web-based prediction of putative SRPN and RCL loop in mosquito *A. albopictus* gene. Clustal alignment (b) and phylogenetic (c) analysis of mosquito serpins. Highly conserved and similar residues have been shown in black and gray, respectively. Species names include: Agam, *Anopheles gambiae*; Aeagy, *Aedes aegypti*; C.quin, *Culex quinquefasciatus*.

aligned mosquito SRPN sequences (92–95%) and phylogenetic analysis showed distant relationship with the *C. quinquefasciatus* mosquito SRPN, suggesting possible function in general regulation of immune cascade (Fig. 8). Recently, an ortholog (SRPN6) of mosquito SRPN family proteins has been strongly implicated in the innate immune response against *Plasmodium* and *E. coli* infection.³² Further detailed studies are required to elucidate the molecular mechanism of pathogen interaction in the mosquito *A. albopictus*.

Cellular function related genes

Sixteen clusters with 25 ESTs encode necessary products for cellular functions. Among these, 15% of genes included energy and metabolism related functions. The three large protein complexes (NADH dehydrogenase, cytochrome reductase, and cytochrome b oxidase) involved in the respiratory chain, in addition to ATPase (ATP synthase), are represented. These enzyme systems likely serve to provide the required energy for various cellular processes, or their up-regulation may be indicative of impending apoptosis, as has been observed in various vertebrate cells.³³ In addition to this, 25% of genes related to protein synthesis and cell structure organization are also illustrative of the concerted investment in immune responsiveness, employing multiple effector arms to combat pathogens, including phagocytosis via cytoskeletal re-arrangements, apoptosis, and the production of effector molecules (Fig. 2).^{34–37}

Mosquito-specific EST clusters

Interestingly, 10 clusters with 249 ESTs were observed that had no similarity to sequences in non-mosquito databases and to which no functional assignment could be made. In the protocol, we challenged mosquito cells with a heat-killed bacterial suspension rather live bacterial paste. Although the possibilities of contaminated sequences of microbial origin can not be ruled out, our specific BLAST analysis against bacterial genome sequences did not yield any significant homology, even if the e-value filter was kept low. Therefore, these intriguing sequences, especially those that are represented multiple times, may represent mosquito-specific gene products related to immune responsiveness and merit further investigation.

CONCLUSIONS

Genome sequence availability from different vector mosquito species and large-scale EST projects are underway to identify potential genes expressed in the different mosquito tissues/cells. Although, these EST

projects are descriptive in nature, they also generate hypotheses on the evolution of several mosquito-specific novel immune genes that may play an important role in disease transmission. Here, we have described transcriptome of immune challenged cell line in the mosquito *A. albopictus* and compared it with those of other mosquitoes and insects. Results indicate that genes are under intense selection and can be used as robust markers for closely related species. Together with ongoing high throughput proteomic and current genomic projects, this should provide further insight into the understanding of the molecular complexity of host–parasite interplay.

ACKNOWLEDGEMENTS

This work was financially supported by the Department of Biotechnology, New Delhi, Government of India. The authors are also thankful to Sarang Satoor for technical support during sequencing of ESTs.

REFERENCES

1. Medzhitov R, Janeway CA. Innate immunity: the virtues of a nonclonal system of recognition. *Cell* 1997; **91**: 295–298.
2. Kim C, Park J, Ha N, Kang HJ, Lee BL. Innate immune response in insects: recognition of bacterial peptidoglycan and amplification of its recognition signal. *BMB Reports* 2008; **41**: 93–101.
3. Barillas-Mury C, Wikel B, Han YS. Mosquito immune responses and malaria transmission: lessons from insect model systems and implications for vertebrate innate immunity and vaccine development. *Insect Biochem Mol Biol* 2000; **30**: 429–442.
4. Lowenberger C. Innate immune response of *Aedes aegypti*. *Insect Biochem Mol Biol* 2001; **31**: 219–229.
5. Scholte JE, Schaffner F. Waiting for the tiger: establishment and spread of the *Aedes albopictus* mosquito in Europe. In: von Takken W, Knols BGJ, (eds) *Emerging pests and vector-borne diseases in Europe*, Vol. 1. Wageningen: Academic Press, 2007; 24–260.
6. Dixit R, Sharma A, Mourya DT, Kamaraju R, Patole MS, Shouche YS. Salivary gland transcriptome analysis during *Plasmodium* infection in malaria vector *Anopheles stephensi*. *Int J Infect Dis* 2009; **13**: 636–646.
7. Fick Jr RB. (ed) *Pseudomonas aeruginosa the Opportunist: Pathogenesis and Disease*. Boca Raton, FL: CRC Press, 1993.
8. D'Argenio DA, Gallagher LA, Berg CA, Manoil C. *Drosophila* as a model host for *Pseudomonas aeruginosa* infection. *J Bacteriol* 2001; **183**: 1466–1471.
9. Mahajan-Miklos S, Tan MW, Rahme LG, Ausubel FM. Molecular mechanisms of bacterial virulence elucidated using a *Pseudomonas aeruginosa*–*Caenorhabditis elegans* pathogenesis model. *Cell* 1999; **96**: 47–56.
10. Rahme LG, Stevens EJ, Wolfort SF, Shao J, Tompkins RG, Ausubel FM. Common virulence factors for bacterial pathogenicity in plants and animals. *Science* 1995; **268**: 1899–1902.
11. Hillyer JF, Schmidt SL, Christensen BM. The antibacterial innate immune response by the mosquito *Aedes aegypti* is mediated by hemocytes and independent of Gram type and pathogenicity. *Microbes Infect* 2004; **6**: 448–459.

12. Lehmann T, Hume JC, Licht M *et al.* Molecular evolution of immune genes in the malaria mosquito *Anopheles gambiae*. *PLoS One* 2009; **4**: e4549.
13. Dimopoulos G, Seeley D, Wolf A, Kafatos FC. Malaria infection of the mosquito *Anopheles gambiae* activates immune-responsive genes during the critical stages of the parasite life cycle. *EMBO J* 1998; **17**: 6115–6123.
14. Cociancich S, Ghazi A, Hoffmann JA, Hetrus C, Letellier C. Insect defensin, an inducible antibacterial peptide, forms voltage dependent channels in *Micrococcus luteus*. *J Biol Chem* 1993; **268**: 19239–19245.
15. Huang LC, Redfern RL, Narayanan S, Reins RY, McDermott AM. *In vitro* activity of human β -defensin 2 against *Pseudomonas aeruginosa* in the presence of tear fluid. *Antimicrob Agents Chemother* 2007; **51**: 3853–3860.
16. Gao Y, Hernandez VP, Fallon AM. Immunity proteins from mosquito cell lines include three defensin A isoforms from *Aedes aegypti* and a defensin D from *Aedes albopictus*. *Insect Mol Biol* 1999; **8**: 311–318.
17. Simard F, Licht M, Besansky NJ, Lehmann T. Polymorphism at the defensin gene in the *Anopheles gambiae* complex: testing different selection hypotheses. *Infect Genet Evol* 2007; **7**: 285–292.
18. Akira S, Uematsu S, Takeuchi O. Pathogen recognition and innate immunity. *Cell* 2006; **124**: 783–801.
19. Fabrick JA, Baker JE, Kanost MR. Innate immunity in a *Pyrallid* moth: functional evaluation of domains from a beta-1,3-glucan recognition protein. *J Biol Chem* 2004; **279**: 26605–26611.
20. Choe K, Werner T, Stöven S, Hultmark D, Anderson KV. Requirement for a peptidoglycan recognition protein (PGRP) in Relish activation and antibacterial immune responses in *Drosophila*. *Science* 2002; **296**: 359–362.
21. Jiggins FM, Kim K. Contrasting evolutionary patterns in *Drosophila* immune receptors. *J Mol Evol* 2006; **63**: 769–780.
22. Dimopoulos G, Richman A, Müller HM, Kafatos FC. Molecular immune responses of the mosquito *Anopheles gambiae* to bacteria and malaria parasites. *Proc Natl Acad Sci USA* 1997; **94**: 11508–11513.
23. Bulmer MS, Bachelet I, Raman R, Rosengaus RB, Sasisekharan R. Targeting an antimicrobial effector function in insect immunity as a pest control strategy. *Proc Natl Acad Sci USA* 2009; **106**: 12652–12657.
24. Sottrup-Jensen L. Alpha-macroglobulins: structure, shape, and mechanism of proteinase complex formation. *J Biol Chem* 1989; **264**: 11539–11542.
25. Blandin SA, Marois E, Levashina EA. Antimalarial responses in *Anopheles gambiae*: from a complement-like protein to a complement-like pathway. *Cell Host Microbe* 2008; **3**: 364–374.
26. Levashina EA, Moita LF, Blandin S, Vriend G, Lagueux M, Kafatos FC. Conserved role of a complement-like protein in phagocytosis revealed by dsRNA knockout in cultured cells of the mosquito, *Anopheles gambiae*. *Cell* 2001; **104**: 709–718.
27. Wang X, Zhao Q, Christensen BM. Identification and characterization of the fibrinogen-like domain of fibrinogen-related proteins in the mosquito, *Anopheles gambiae*, and the fruitfly, *Drosophila melanogaster*, genomes. *BMC Genomics* 2005; **6**: 114.
28. Dixit R, Roy U, Patole MS, Shouche YS. Molecular and phylogenetic analysis of a novel family of fibrinogen-related proteins from mosquito *Aedes albopictus* cell line. *Comput Biol Chem* 2008; **32**: 382–386.
29. Christophides GK, Zdobnov E, Barillas-Mury C *et al.* Immunity-related genes and gene families in *Anopheles gambiae*. *Science* 2002; **298**: 159–165.
30. Rodrigues J, Agrawal N, Sharma A *et al.* Transcriptional analysis of an immune-responsive serine protease from Indian malarial vector, *Anopheles culicifacies*. *BMC Mol Biol* 2007; **15**: 8–33.
31. Ragg H. The role of serpins in the surveillance of the secretory pathway. *Cell Mol Life Sci* 2007; **64**: 2763–2770.
32. Abraham EG, Pinto SB, Ghosh A *et al.* An immune-responsive serpin, SRPN6, mediates mosquito defense against malaria parasites. *Proc Natl Acad Sci USA* 2005; **102**: 16327–16332.
33. Chandra D, Liu JW, Tang DG. Early mitochondrial activation and cytochrome c up-regulation during apoptosis. *J Biol Chem* 2002; **277**: 50842–50854.
34. Opferman JT, Korsmeyer SJ. Apoptosis in the development and maintenance of the immune system. *Nat Immunol* 2003; **4**: 410–415.
35. Nappi AJ, Ottaviani E. Cytotoxicity and cytotoxic molecules in invertebrates. *Bioessays* 2000; **22**: 469–480.
36. Greenberg S, Grinstein S. Phagocytosis and innate immunity. *Curr Opin Immunol* 2002; **14**: 136–145.
37. Boutros M, Agaisse H, Perrimon N. Sequential activation of signaling pathways during innate immune responses in *Drosophila*. *Dev Cell* 2002; **3**: 711–722.