

An insight into the sialome of the adult female mosquito *Aedes albopictus*

Bruno Arcà^{a,b}, Fabrizio Lombardo^b, Ivo M.B. Francischetti^c, Van My Pham^c,
Montserrat Mestres-Simon^b, John F. Andersen^c, José M.C. Ribeiro^{c,*}

^aDepartment of Structural and Functional Biology, University 'Federico II', Naples, Italy

^bParasitology Section, Department of Public Health, University 'La Sapienza', Rome, Italy

^cSection of Vector Biology, Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases,
12735 Twinbrook Parkway, National Institutes of Health, Rockville, MD 20852, USA

Received 24 August 2006; received in revised form 20 October 2006; accepted 25 October 2006

Abstract

To gain insight into the molecular repertoire of the adult female salivary glands of the tiger mosquito *Aedes albopictus*, we performed transcriptome and proteome analysis. cDNA clones were sequenced and assembled in clusters of related sequences and the corresponding genes assigned to one of three categories: housekeeping (H; 31%), secreted (S; 34%), or unknown (U; 35%) function. Among the putative secreted factors are proteins known to be widely distributed in the saliva of blood-sucking Diptera, such as D7 and antigen 5 family members, as well as proteins that are mosquito- or culicine-specific, i.e., the 30-kDa allergen or the 62-kDa and 34-kDa families, respectively. Expression of 15 of these salivary proteins was confirmed by Edman degradation. Tissue and sex specificity of selected transcripts were evaluated by RT-PCR and identified at least 32 genes whose expression is restricted or enriched in the female salivary glands of *Ae. albopictus*, whereas 17 additional genes were expressed in female glands and adult males but not in other tissues of adult females. For approximately one third of the genes analyzed, involvement in blood-feeding, sugar digestion, immune response, or other more generic physiological roles can be postulated; however, no functions can be suggested for the remaining sequences, which therefore likely represent either novel functions or novel molecules recruited during the evolution of hematophagy. Supplemental spreadsheets with hyperlinks to all sequences used in this manuscript are hyperlinked throughout the text and can be found at <http://www.ncbi.nlm.nih.gov/projects/omes/#salivarytranscriptomes>.

Published by Elsevier Ltd.

Keywords: Saliva; Transcriptome; Hematophagy; Salivary proteins; Mosquito; *Aedes albopictus*

1. Introduction

The tiger mosquito *Aedes albopictus* is probably the fastest-spreading mosquito in the world. Historically original to tropical forests of Southeast Asia, it was already present in a large number of islands of the western Pacific and Indian Oceans in the middle of 20th century. Amazingly, in the last two decades, *Ae. albopictus* attained an almost global

distribution, being presently found in the United States, Central and South America (as far south as Argentina), Africa, Europe, and the Middle East. Several reasons contributed to the rapid spreading of this mosquito species; among the most important, certainly, are the resistance of its desiccated, dormant eggs and their efficient passive transport around the world, often in used tires (Tatem et al., 2006). Moreover, a significant contribution to the current widespread distribution came from the ability of *Ae. albopictus* populations to survive to relatively low temperatures (10 °C or lower) and to exploit for oviposition a large variety of breeding sites originated by human activities.

Ae. albopictus is an efficient laboratory vector for a large number of arboviruses (dengue, yellow fever, West Nile,

*Corresponding author. Section of Vector Biology, Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, 12735 Twinbrook Parkway, National Institutes of Health, Rockville, MD 20852, USA. Tel.: +1 301 496 9389; fax: +1 301 480 2571.

E-mail address: jribeiro@niaid.nih.gov (J.M.C. Ribeiro).

and several others), and its rapid global spreading has thus raised concern among public health officials; however, with the exception of some small outbreaks of dengue fever, its efficiency as a vector for arboviral diseases seems to be important in nature as suggested by laboratory data (Gratz, 2004; Gubler, 2003). After the first documented introduction in Genoa in 1990 (Sabatini et al., 1990), *Ae. albopictus* became well established in Italy. Thanks to some of its ecological characteristics (daylight biting habit, high density, aggressive and anthropophilic behavior), it became a major biting pest (Romi, 2001) and, as recently suggested, a possible natural vector of the filarial nematode *Dirofilaria immitis* (Cancrini et al., 2003).

The salivary glands of blood-feeding arthropods secrete a large number of bioactive molecules essential for efficient blood feeding; among these the most well known are certainly anti-hemostatics, but anti-inflammatory and immunomodulatory agents are also present (Ribeiro and Francischetti, 2003; Ribeiro, 1995). In the last few years, salivary gland transcriptome analyses have highlighted the complexity of mosquito salivary secretions, pointing out how little we know concerning targets and functional roles of salivary components (Arca et al., 2005; Calvo et al., 2004; Ribeiro et al., 2004; Valenzuela et al., 2002b, 2003). Salivary proteins injected into the host skin during feeding may also act as antigens and/or allergens, inducing host immune reactions. In this respect, *Ae. albopictus* may induce intense local cutaneous reactions and has been reported as the most common species associated with severe systemic allergic reaction to mosquito bites (Peng et al., 2004). Moreover, immunoblot analysis of salivary extracts with sera from individuals allergic to mosquitoes indicated that *Ae. albopictus* saliva carries at least 16 different allergens, the highest number among the 10 mosquito species examined (Peng and Simons, 2004). In view of its possible role as a disease vector and because of the unique allergenic potential of its salivary protein repertoire, we have begun transcriptome analysis of the salivary glands of the tiger mosquito *Ae. albopictus*, and we report here 69 novel cDNA's encoding putative salivary-secreted proteins.

2. Methods

2.1. Mosquitoes

The *Ae. albopictus* strain used in this study (kindly provided by R. Romi, Istituto Superiore di Sanità) was caught in Rome in 2000 and reared under standard laboratory conditions ($25 \pm 1^\circ\text{C}$, relative humidity $60 \pm 10\%$, light:dark photoperiod 14:10 h) in the insectary of the Department of Public Health at University La Sapienza, Rome. Adult females (1–5 days old) maintained on a 10% sucrose diet were used in this study. Salivary glands were dissected, collected in groups of 20 in 20 μl of PBS buffer (150 mM NaCl, 5 mM Na-phosphate, pH 7.4), frozen in liquid nitrogen, and stored at -80°C until needed.

2.2. Tissue preparation and cDNA library synthesis

Poly-A⁺ mRNA was isolated from 60 adult female salivary glands (never blood fed, 1–5 days post-emergence) with the MicroFast track mRNA isolation kit (Invitrogen) following the manufacturer's protocol. The cDNA library was constructed in the pTriplEx2 vector using the SMART cDNA library construction kit (BD Clontech). Random clones were sequenced from the 5' direction only, because successful sequencing from the 3' end was usually lower than 40%. Full-length sequences were obtained in selected cases by performing primer-based extension protocols. See (Francischetti et al., 2002) for more details.

2.3. Gel electrophoresis and Edman degradation studies

Sodium dodecyl sulfate polyacrylamide electrophoresis (SDS-PAGE) of homogenized salivary glands of *Ae. albopictus* adult females was done using NuPAGE 4–12% bis-tris gel, 1 mm thick (Invitrogen). Salivary gland homogenates were treated with NuPAGE LDS sample buffer (Invitrogen) in the presence of β -mercaptoethanol (5%, v/v). Twenty pairs of homogenized salivary glands ($\sim 50 \mu\text{g}$ protein) were applied/lane. The gel was run with MES buffer according to the manufacturer's instructions. To estimate the molecular weight of the samples, See-BlueTM markers from Invitrogen (myosin, BSA, glutamic dehydrogenase, alcohol dehydrogenase, carbonic anhydrase, myoglobin, lysozyme, aprotinin, and insulin, chain B) were used. The gel was stained with Coomassie blue (0.2%, g/v). For aminoterminal sequencing of the salivary proteins, the gel was transferred to polyvinylidene difluoride (PVDF) membrane using 10 mM CAPS, pH 11.0, 10% methanol as the transfer buffer on a blot module for the XCell II Mini-Cell (Invitrogen). The membrane was stained with Coomassie blue (0.02% in the absence of acetic acid). Stained bands were cut from the PVDF membrane and subjected to Edman degradation using a Procise sequencer (Perkin-Elmer Corp.). More details can be obtained in a previous publication (Francischetti et al., 2002). To find the cDNA sequences corresponding to the amino acid (aa) sequence—obtained by Edman degradation of the proteins transferred to PVDF membranes from PAGE gels—we wrote a search program (in Visual Basic) that evaluated these aa sequences against the three possible protein translations of each cDNA sequence obtained in the mass sequencing project or against the full-length sequences derived in this work. For details, see (Valenzuela et al., 2002b).

2.4. Reverse transcriptase-polymerase chain reaction (RT-PCR) expression analysis

Total RNA was extracted from dissected glands, carcasses (adult females from which salivary glands had been dissected), and adult males using the TRIZOL reagent (Invitrogen) and treated with RNase-free DNase I.

DNase-treated total RNA (50 ng) was used for amplification by RT-PCR using the Superscript one-step RT-PCR system (Invitrogen) according to manufacturer's instructions. Typically, reverse transcription (50 °C, 30 min) and heat inactivation of the reverse transcriptase (94 °C, 2 min) were followed by 30 PCR cycles (30 s at 94 °C, 30 s at 55 °C, 1 min at 72 °C); 25 cycles were used for amplification of mRNA encoding ribosomal protein L17 and S5 to keep the reaction below saturation levels and to allow a more reliable normalization. The sequence of the oligonucleotide primers used for *rpL17* and *rpS5* amplification were: for Aa_rpL17-F, 5'-TCA ATC CgT ACA TgT CgT Cg-3'; for Aa_rpL17-R, 5'-gCT ggA gCT TCT TCT Tgg AC-3'; for Aa_rpS5-F, 5'-ATT ACA TCG CCG TCA AGG AG-3'; and for Aa_rpS5-R, 5'-TCA TCA TCA GCG AGT TGG TC-3'. The sequence of the other oligonucleotide primers is provided in the Supplemental material. Amplification reactions were analyzed on 1.2% agarose gels stained with ethidium bromide. Each sample was analysed by RT-PCR at least twice and up to four times using independent batches of RNA.

2.5. Bioinformatic tools used

Expressed sequence tags (EST) were trimmed of primer and vector sequences, clusterized, and compared with other databases as described before (Valenzuela et al., 2003). The BLAST tool (Altschul and Gish, 1996) and CAP3 assembler (Huang and Madan, 1999) were used for the clusterization. First, sequences that had <5% N were blasted against each other and the output scanned to join in a single fasta file those that had minimum identity of 81 aa over a 90-aa stretch. These sequences were, in turn, input to the CAP3 assembler, which could further divide the blast cluster into additional clusters. The final output was piped into a tab-delimited file imported into an Excel spreadsheet (Microsoft Excel Analysis Tools). These operations were automated by a program written in Visual Basic named Cluster5. O-glycosylation sites on the predicted proteins were obtained with the program NetOGlyc (<http://www.cbs.dtu.dk/services/NetOGlyc/>) (Hansen et al., 1998). We submitted all translated sequences (starting with a Met) to the SignalP server (Nielsen et al., 1997) to detect signal peptides indicative of secretion. BLAST searches were done locally from executables obtained at the National Center of Biological Information (NCBI) FTP site (<ftp://ftp.ncbi.nih.gov/blast/executables/>) (Altschul et al., 1997) against the nonredundant (NR) protein database of the NCBI, the gene ontology fasta subset (Lewis et al., 2000), the conserved domains database of NCBI (Marchler-Bauer et al., 2002) containing the KOG (Tatusov et al., 2003), Pfam (Bateman et al., 2000) and Smart (Schultz et al., 2000) motifs, and to custom-downloaded databases containing mitochondrial and rRNA nucleotide sequences available at the NCBI. Phylogenetic analyses were done using the program ClustalX (Thompson et al., 1997) and the

MEGA3 package when there were too many divergent sequences (Kumar et al., 2004). Phylogenetic trees were constructed with the neighbor-joining algorithm and were bootstrapped 10,000 times. Gaps were taken into consideration by the pairwise deletion method; aa substitutions were considered by a Poisson model. The names of the sequences in the alignments and phylograms consist of a five-letter abbreviation of the species name (the first three from the genus, followed by the first two of the species name; thus, ANOGA for *Anopheles gambiae*) followed by the NCBI identifier number that follows the gi| for sequences deposited at GenBank at the time of manuscript writing; otherwise, the sequences reflect the arbitrary names of the clones or clusters of clones sequenced in this work or the *An. gambiae* sequences obtained from http://www.ncbi.nlm.nih.gov/projects/omes/An_gambiae_sialome-2005/Table4.xls.

2.6. Supplemental material

Supplemental Tables S1 and S2 and figures are hyperlinked throughout the paper to the NCBI page <http://www.ncbi.nlm.nih.gov/projects/omes/>, where this and other transcriptomes are located.

3. Results and discussion

3.1. cDNA library characteristics

A total of 1021 clones were sequenced and used to assemble a clusterized database (Supplemental Table S1), yielding 446 clusters of related sequences, 341 of which contained only one EST. The consensus sequence of each cluster is designated either a contig (deriving from two or more sequences) or a singleton (deriving from a single sequence); in this paper, for simplicity sake, we will use the denomination contig to address sequences deriving both from consensus sequences and from singletons. The 446 contigs were compared by the program blastx, blastn, or RPSBLAST (Altschul et al., 1997) to the NR protein database of the NCBI, to the gene ontology database (Ashburner et al., 2000), to the conserved domains database of the NCBI (Marchler-Bauer et al., 2002), and to a custom-prepared subset of the NCBI nucleotide database containing either mitochondrial or rRNA sequences. Because the libraries used are unidirectional, the three frame translations of the dataset were also derived, and open reading frames (ORF) starting with a methionine and longer than 40 aa residues were submitted to the SignalP server (Nielsen et al., 1997) to help identify putative secreted proteins. The EST assembly, BLAST, and signal peptide results were piped into an Excel spreadsheet for manual annotation.

Four categories of expressed genes derived from the manual annotation of the contigs. The putatively secreted (S) category contained 34% of the clusters and 61% of the sequences, with an average number of 4.2 sequences/

cluster. The housekeeping (H) category had 31% and 21% of the clusters and sequences, respectively, and an average of 1.5 sequences per cluster. Thirty-five percent of the clusters, containing 18% of all sequences, were classified as unknown (U) because no assignment for their function could be made; most of these consisted of singletons. A good proportion of these transcripts could have derived from 3' or 5' untranslated regions of genes of the above two categories, as was recently indicated for a sialotranscriptome of *An. gambiae* (Arcà et al., 2005). Possibly transposable elements originated two clusters. These could represent either active transposition or, more likely, expression of TE regulatory transcripts in *Ae. albopictus*.

3.2. Housekeeping (H) genes

The 138 gene clusters (comprising 212 EST) attributed to H genes expressed in the salivary glands of *Ae. albopictus* were further characterized into 16 subgroups according to function (Table 1). Not surprisingly for an organ specialized in secreting polypeptides, the two larger sets were associated with protein-synthesis machinery (95 EST in 44 clusters) and with energy metabolism (44 EST in 25 clusters). Also, the 17 EST (17 clusters) representing conserved proteins of unknown function and indicated as 'Conserved' in Table 1 are presumably associated with cellular metabolism. This group also includes a homologue of the membrane-associated *An. gambiae* SGS4 protein, which has been proposed to belong to a mosquito-specific protein family including candidate receptors for malaria sporozoite invasion of salivary glands (Korochkina et al., 2006) and that may have appeared in mosquito genomes by a lateral transfer event from a Wolbachia bacteria (Arcà et al., 2005). The group of clusters attributed to the protein modification function is of interest due to the specialized function of the salivary glands. This group with 10 clusters

includes chaperones and enzymes associated with disulfide bridge formation. Transporters were also identified, including those coding for two subunits of the V-ATPase complex necessary for mosquito salivation (Novak et al., 1995; Novak and Rowley, 1994). Additional information on each of the 138 gene clusters is available online (Supplemental Table S1).

3.3. Possibly secreted proteins and peptides

Supplemental Table S1 indicates the presence of several gene families previously described in the salivary glands of mosquitoes including the relatively abundantly expressed D7, 56-kDa, 30-kDa, angiopoietin, and antigen-5 (AG5), as well as enzymes associated with sugar (maltase, amylase) and blood feeding (apyrase, adenosine deaminase (ADA), and purine hydrolase). A summary of these transcripts organized by their abundance and protein family is shown in Table 2.

3.4. Preliminary characterization of the salivary transcriptome and proteome of the adult female *Ae. albopictus*

Several clusters of sequences coding for housekeeping and putative secreted polypeptides indicated in Supplemental Table S1 are abundant and complete enough to extract consensus sequences of novel cDNA. Additionally, primer extension studies on several clones allowed us to obtain full- or near full-length sequences of products of interest. A total of 107 novel sequences, 68 of which code for putative secreted proteins, are grouped in Supplemental Table S2.

To obtain information on the most abundant proteins in the salivary glands of *Ae. albopictus*, salivary gland homogenates of 20 pairs of glands were separated by

Table 1
Functional classification of the housekeeping genes expressed in adult female *Aedes albopictus* salivary glands

Types of transcripts	Clusters	Sequences	Sequences/cluster	% Cluster	% Sequence
Protein synthesis machinery	44	95	2.2	31.9	44.8
Metabolism, energy	25	44	1.8	18.1	20.8
Conserved	17	17	1.0	12.3	8.0
Protein modification	10	10	1.0	7.2	4.7
Transporters	8	9	1.1	5.8	4.2
Protein export machinery	8	8	1.0	5.8	3.8
Signal transduction	6	6	1.0	4.3	2.8
Cytoskeletal	4	5	1.3	2.9	2.4
Metabolism, lipid	3	3	1.0	2.2	1.4
Transcription machinery	3	3	1.0	2.2	1.4
Proteosome machinery	3	3	1.0	2.2	1.4
Metabolism, oxidant	2	3	1.5	1.4	1.4
Transcription factors	2	2	1.0	1.4	0.9
Metabolism, amino acid	1	2	2.0	0.7	0.9
Metabolism, carbohydrate	1	1	1.0	0.7	0.5
Metabolism, nucleotide	1	1	1.0	0.7	0.5
Total	138	212			

Table 2

Classification of transcripts coding for putative secreted proteins in adult female *Aedes albopictus* salivary glands

Types of transcripts	Clusters	Sequences	% Sequences
D7	24	105	16.8
30-kDa allergen	8	61	9.7
Maltase, glucosidase	10	57	9.1
56-kDa mosquito family	7	47	7.5
8.3-kDa secreted protein	1	44	7.0
poly His peptide	1	33	5.3
Antigen 5	4	26	4.2
Mucins	5	24	3.8
Apyrase	6	21	3.4
34-kDa mosquito family	3	17	2.7
C-type lectin	3	13	2.1
Serpins	9	13	2.1
Angiopoietins	6	11	1.8
Peptide 4.2k-1	1	11	1.8
58 kDa Aedes family	6	10	1.6
Serine proteases	6	19	3.0
Purine nucleosidase	2	8	1.3
Adenosine deaminase	1	4	0.6
Others	47	102	16.3
Total	150	626	

SDS-PAGE, the protein transferred to a PVDF membrane, and the stained bands submitted to Edman degradation. Fifteen bands yielded useful information, all of which could be assigned to predicted sequences from our clusterized database and to full-length sequences disclosed in this work. These results are summarized in Fig. 1. Notice that most of the X indicated in the Edman degradation correspond to Cys residues that do not yield Edman product. Descriptions of the proteins thus identified are provided below.

We also determined tissue and sex specificity of 59 transcripts, mostly encoding putative secretory products, because this information may be useful to direct future functional studies. To this end, oligonucleotide primers suitable for amplification of the corresponding mRNA were designed and used for RT-PCR on total RNA extracted from female salivary glands, female carcasses (i.e., adult females with salivary glands removed), and whole adult males. Primers amplifying the ribosomal proteins S5 and L17 were used for normalization and as a control for the amplification reactions. The results of this analysis are summarized in Table 3. A similar assay was previously used for analysis of the *An. gambiae* salivary transcriptome (Arcà et al., 2005), yielding results that matched well the information independently obtained by Marinotti et al. (1990) using the Affymetrix microarray chip. In this way, it is possible to distinguish three broad categories of genes. A first category is represented by genes with ubiquitous expression (UB in Table 3), i.e., detectable in the three tissues examined at approximately the same levels. Products encoded by these genes are most likely involved in housekeeping functions; ten of the transcripts analysed belong to this class. A second group is represented

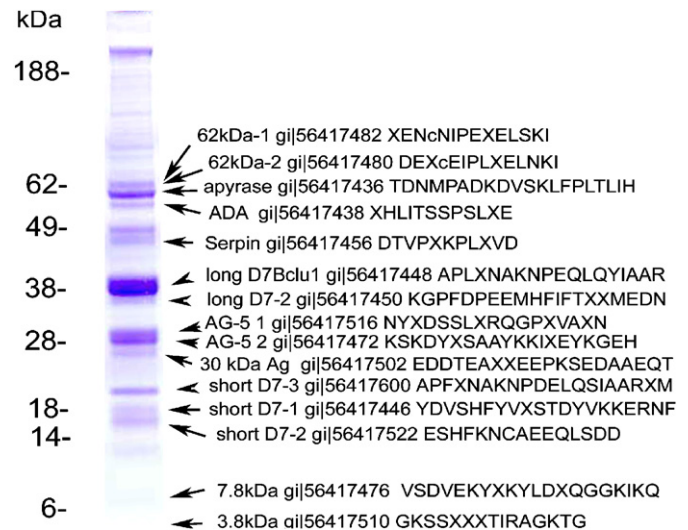


Fig. 1. Coomassie blue-stained PVDF membrane resulting from transfer of SDS-PAGE of 20 pairs of homogenized glands of adult female *Ae. albopictus*. Molecular weight markers are shown on the left. The right side lists the protein name, NCBI accession number, and Edman degradation product obtained from the band, respectively. An X indicates aa mismatch.

by genes that are female salivary gland specific (SG in Table 3) or whose expression appeared enriched in the female salivary glands compared with the other samples (ENR in Table 3); proteins encoded by these genes are likely to play some role in blood feeding, for example as anti-hemostatics or immunomodulators. Among the transcripts analysed, 22 were female gland specific and 10 were found enriched in female salivary glands; they include several genes known from previous studies on other mosquito species to be involved in blood-meal acquisition (see below). A third class of genes includes those expressed both in female glands and in adult males (SG/M in Table 3) but with no expression in female carcasses; many of these genes are supposed to be gland specific and expressed in both male and female glands. The corresponding products are likely involved in sugar feeding, antimicrobial activity, or other gland functions. Seventeen transcripts are part of this group. Accordingly, of the 59 transcripts shown in Table 3, 32 are either female specific (22) or enriched in female glands (10), whereas 17 are expressed both in female glands and in adult males. Ten transcripts were found ubiquitously expressed. Detailed descriptions of the full-length transcripts found in the salivary glands of adult female *Ae. albopictus* and of their profiles of expression are given below.

3.4.1. Putative salivary-secreted proteins for which a function is known or presumed

D7 protein family: The first member of the D7 protein family was described in the mosquito *Aedes aegypti* (James et al., 1991) and later found in virtually all mosquito sialotranscriptomes. Short (~15 kDa) and long (~30 kDa) forms are recognized. Long D7 forms also exist in sand flies

Table 3

Classification of selected genes encoding putative secretory products as determined by RT-PCR expression analysis on female salivary glands, carcasses of female mosquitoes, and whole males.

Family or Seq name	NCBI accession number	Comments	Edman	SG	ENR	SG/M	UB
<i>Secreted proteins</i>							
<i>D7 protein family</i>							
AL_7_90_90_8_CLU	56417450	D7L1	+	X			
Cluster-78	56417448	D7L2	+			X	
Cluster-92	56417442	D7 short	+	X			
AL-7-90-90-90-CLU	56417444	D7 short	+	X			
Cluster-91	56417440	D7 short	+	X			
AL_7_90_90_102_CLU	56417522	D7 short	+			X	
<i>Serpins</i>							
Cluster-71	56417462	Serpin			X		
Cluster-135	56417454	Serpin		X			
Cluster-174	56417456	Homologue of <i>Ae. aegypti</i> anticlotting	+	X			
<i>Mucins</i>							
AL_7_90_90_7_CLU	56417412	Putative secreted protein				X	
Cluster-4	56417410	Putative secreted protein				X	
Cluster-38	56417396	Mucin like perithrophin				X	
AL_7_90_90_139_CLU	56417398	Putative salivary mucin		X			
<i>Enzymes</i>							
Cluster-95	56417436	Apyrase	+			X	
Cluster-113	56417432	Salivary purine nucleosidase				X	
Cluster-144	56417438	Putative adenosine deaminase	+		X		
AL_7_90_90_24_CLU	56417434	Salivary chymotrypsin/elastase		X			
AL_7_90_90_31_CLU	56417392	Maltase				X	
<i>Immunity related products</i>							
<i>Peptides</i>							
AL_7_90_90_18_CLU	56417474	Putative 8.3 kDa secreted protein				X	
AL_7_90_90_12_CLU	56417428	Histidine rich peptide				X	
Cluster-278	56417406	Similar to <i>Ae. aegypti</i> i23 M peptide					X
Cluster-213	56417404	Llysozyme				X	
<i>Angiopoietin/fibrinogen-related proteins</i>							
AL-7-90-90-142-CLU	56417464	Angiopoietin-like protein-truncated?		X			
AL_7_90_90_136_CLU	56417466	Angiopoietin-like protein		X			
<i>Lectins</i>							
AL_7_90_90_120_CLU	56417402	Putative C-type lectin		X			
Cluster-125	56417400	Putative C-type lectin		X			
Cluster-324	56417408	Gram negative binding protein			X		
<i>Antigen 5 family</i>							
Cluster-36	56417516	Antigen-5 family member	+		X		
AL_7_90_90_138_CLU	56417468	Antigen 5-related salivary protein				X	
AL_7_90_90_121_CLU	56417470	Antigen 5-related protein		X			
AL_7_90_90_33_CLU	56417472	Antigen 5 precursor	+	X			
<i>Kazal and TIL domain-containing peptides</i>							
Cluster-380	56417460	Kazal domain containing peptide					X
Cluster-227	56417458	Til domain containing peptide				X	
<i>30 kDa family</i>							
AL_7_90_90_14_CLU	56417502	30 kDa allergen (shorter version)	+	X			
AL_7_90_90_19_CLU	56417508	30 kDa allergen (shorter version)		X			
AL_7_90_90_101_CLU	56417506	30 kDa salivary gland allergen		X			
<i>56 kDa family</i>							
Cluster-41	56417486	Putative 56.5 family allele 1				X	
<i>16 kDa family</i>							
Cluster-145	56417416	Similar to <i>An. gambiae</i> hyp16					X
<i>62 kDa family</i>							
Cluster-73	56417482	62 kDa protein 1	+	X			
Cluster-108	56417480	62 kDa protein 2	+	X			

Table 3 (continued)

Family or Seq name	NCBI accession number	Comments	Edman	SG	ENR	SG/M	UB
<i>41.9 kDa family</i>							
AL_7_90_90_124_CLU	56417494	41 kDa protein				X	
<i>34 kDa family</i>							
Cluster-148	56417496	Putative 34 kDa secreted protein			X		
AL_7_90_90_28_CLU	56417498	34 kDa secreted protein	+	X			
<i>27 kDa family</i>							
AL_7_90_90_122_CLU	56417514	Putative 27 kDa secreted protein					X
Cluster-356	56417512	Putative 28 kDa secreted protein				X	
<i>23.4 kDa family</i>							
AL_7_90_90_152_CLU	56417518	Putative 21.5 kDa salivary protein			X		
<i>WPCWW 8.9 kDa peptide</i>							
AL_7_90_90_123_CLU	56417390	W rich similar to <i>Culex</i>				X	
<i>7.6 kDa basic family</i>							
AL_7_90_90_105_CLU	56417476	Putative 7.6 kDa secreted protein	+	X			
AL_7_90_90_106_CLU	56417478	Putative 7.6 kDa secreted protein	+	X			
<i>Basic 3.8 kDa peptide</i>							
AL_7_90_90_26_CLU	56417510	Putative 3.8 kDa basic salivary peptide	+		X		
<i>Other peptides of unknown function</i>							
Cluster-316	56417452	Putative cys rich salivary peptide					X
Cluster-273	56417414	Putative ecreted peptide					X
Cluster-425	56417426	QG rich Putative secreted peptide					X
AL_7_90_90_137_CLU	56417524	Putative salivary peptide					X
<i>13 kDa family</i>							
Cluster-155	56417422	Putative salivary peptide			X		
Cluster-154	56417420	Putative salivary peptide			X		
Cluster-153	56417418	Putative salivary peptide			X		
Cluster-335	56417424	Putative salivary peptide					X
<i>Housekeeping products</i>							
Cluster-51	56417394	Small secreted conserved peptide					X

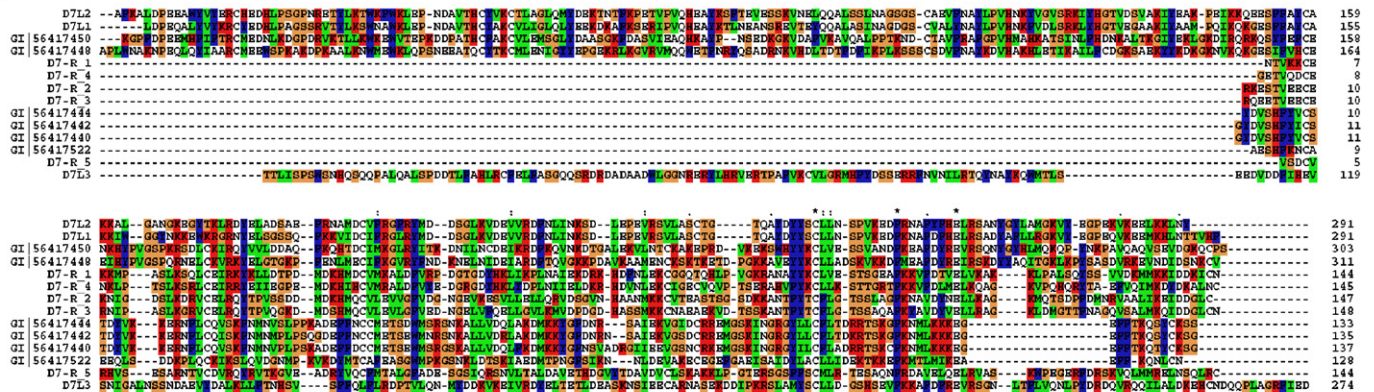
For additional data, see [Supplemental Table S2](#).

*Edman, polypeptides whose identity and secretory nature was confirmed by Edman degradation; SG, signal found only in female salivary glands; ENR, enriched in adult female salivary glands; SG/M, found in female salivary glands and adult males, but not in female carcasses lacking salivary glands; UB, expressed in all tissue examined.

(Valenzuela et al., 2002a) and Culicoides (Campbell et al., 2005). Some of these proteins have been associated with binding of biogenic amines such as serotonin, histamine, and norepinephrine (Calvo et al., 2006). Additionally, one short D7 protein from *An. stephensi*, named hamadarin, was shown to prevent kallikrein activation by Factor XIIa (Isawa et al., 2002). In *An. gambiae*, five short and three long D7 proteins were shown to occur as tandem duplications in chromosome 3R (Arca et al., 2005). In *Ae. aegypti*, two long and four short D7 proteins were found with five of the corresponding genes grouped in a cluster spanning ~80 kb (Ribeiro, J.M.C., et al., manuscript submitted). In *Ae. albopictus*, we describe two long and four short D7 proteins in addition to alleles of some of these proteins (Table 3 and Supplemental Table S2). Evidence for expression for all but one of these D7 proteins was found by Edman degradation of gel bands from SDS-PAGE (Fig. 1). The long D7 proteins contributed to the strongest staining bands in SDS-PAGE

(Fig. 1) and yielded strong signals on Edman degradation. Alignment of the six *Ae. albopictus* D7 sequences with the eight well-characterized D7 proteins from *An. gambiae* (Fig. 2A) resulted in a phylogram (Fig. 2B) in which four major groups are resolved. Group I contains the last genes in the short and long *An. gambiae* D7 cassette, which are poorly expressed and share lower sequence similarity with the other D7 proteins (Calvo et al., 2006). Group II contains all other D7 long proteins, both from *Aedes* and *Anopheles*. Group III contains all short D7 proteins from *Ae. albopictus*, and Group IV contains all remaining *An. gambiae* short proteins. The relative large similarity distance among different D7 proteins indicates the rapid pace of evolution for this family, which appears to have been co-opted very early for an exclusive blood-feeding role in hematophagous Nematocera (Calvo et al., 2005). Accordingly, *Ae. albopictus* D7 appeared to be expressed only or predominantly in female salivary glands (Table 3 and Fig. 3). This pattern of expression is also in line with

A



B

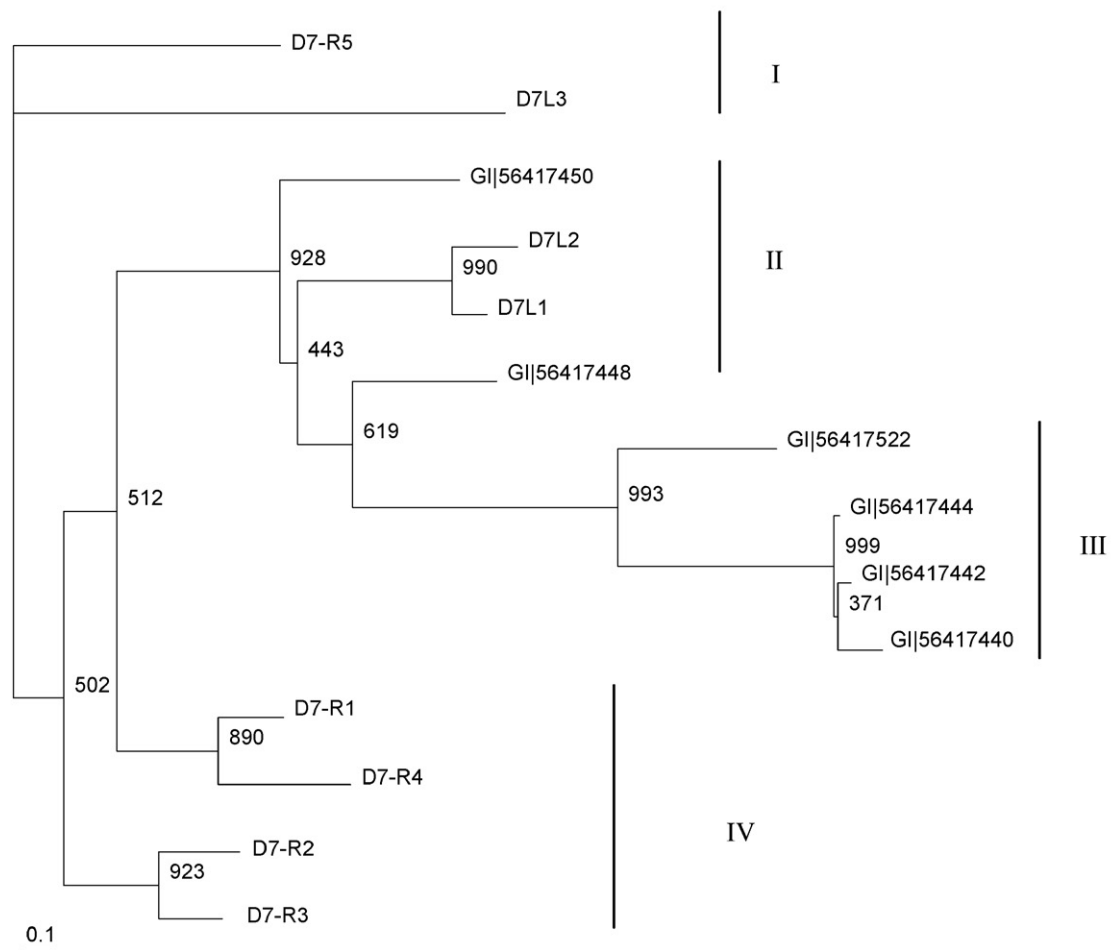


Fig. 2. Alignment of the mature forms of *Ae. albopictus* D7 salivary proteins with the eight known D7 proteins from *An. gambiae*. (A) Clustal alignment. (B) Phylogram. Numbers on branches indicate bootstrap value for 1000 trials. Bar indicates 10% distance in aa sequence. *Ae. albopictus* sequences are indicated by their NCBI accession number. Other proteins are from *An. gambiae* as annotated in <http://www.ncbi.nlm.nih.gov/projects/omes/Angambiasialome-2005/Table4.xls>.

that previously observed in *An. gambiae*, where expression of the D7 genes is restricted to female salivary glands with the only exceptions of D7r5 and D7L3, the most divergent members of the family that also cluster independently in Group I (Arcà et al., 2005) (Fig. 2B). A very similar pattern

of expression is also found in the yellow fever vector *Ae. aegypti* (Ribeiro, J.M.C., et al., manuscript submitted).

Serpins: The serine protease inhibitor (serpin) protein family is ubiquitous in animals and plants; in *Ae. aegypti*, one member is responsible for the main salivary

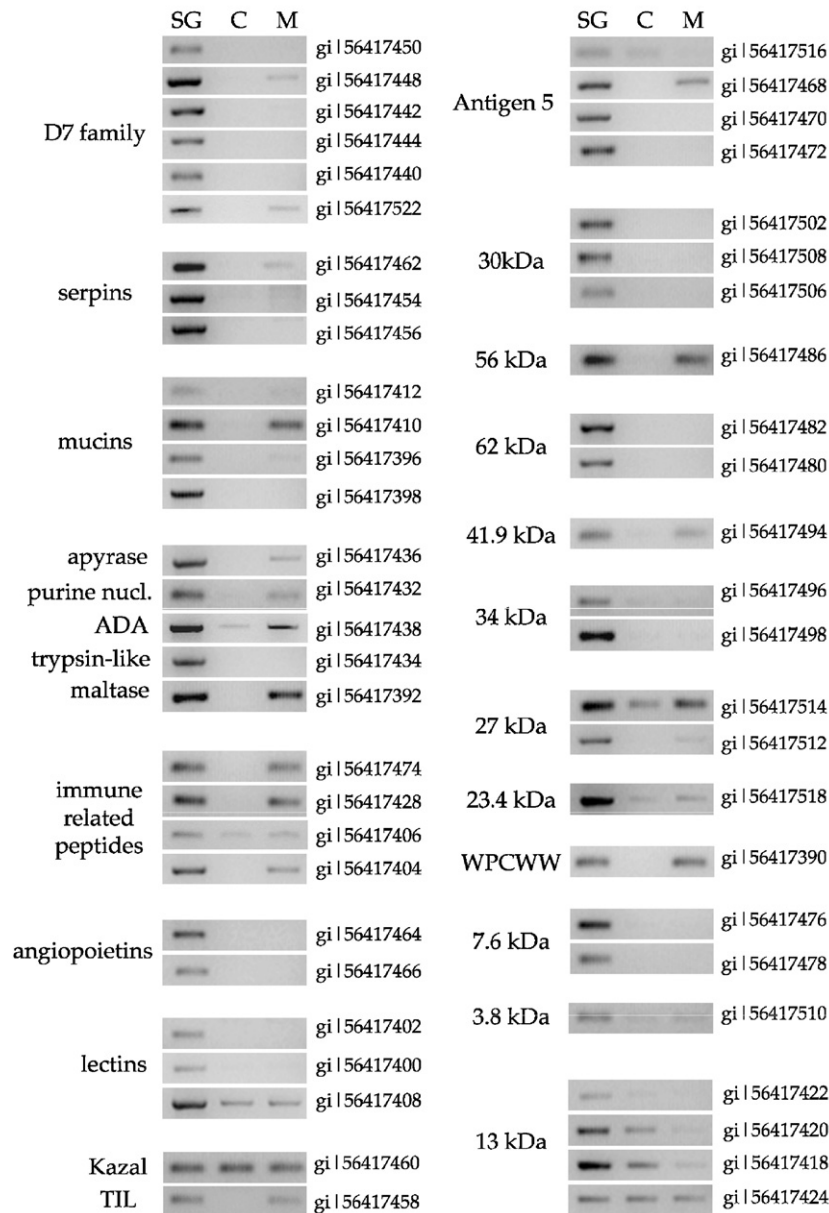


Fig. 3. Tissue specificity of genes expressed in the salivary glands of adult female *Ae. albopictus* as evidenced by RT-PCR. Total RNA from female salivary glands (SG), carcasses (C) (adult females with salivary glands removed), or adult males (M) was amplified using primers for the genes coding for the proteins indicated on the right by their NCBI accession numbers.

anti-blood-clotting activity (gi|3411116), having specificity for blood-clotting Factor Xa (Stark and James, 1995; Stark and James, 1998). Three members of this protein family were found transcribed in the salivary glands of *Ae. albopictus*, one of which was detected in the gel reported in Fig. 1. Alignment of these three proteins with the three previously reported serpins from *Ae. aegypti* (Fig. 4) indicates that *Ae. albopictus* gi|56417456 is the homologue of *Ae. aegypti* FXa inhibitor, with 62% identity at the aa level. The three serpins of each species cluster more closely between than within species, with strong bootstrap support (Fig. 4B), indicating that the ancestral *Aedes* mosquito genome already had these three proteins. The aa sequence DTVPXKPLXVD was found in the expected region of the

experiment reported in Fig. 1, indicating abundant expression of gi|56417456, the putative *Ae. albopictus* anti-Factor Xa salivary serpin. The target of the other serpins is unknown, but their function is also likely connected to blood feeding, as expression analysis indicates that transcripts encoding these three serpins are found only or mostly in adult female salivary glands (Table 3 and Fig. 3). It should be noted that serpins have not been described in the sialotranscriptomes of *Anopheles* (Calvo et al., 2004; Francischetti et al., 2002; Valenzuela et al., 2003), perhaps because anophelines abundantly express transcripts coding for peptides of the anophelin family, which was shown to be a potent antithrombin in *An. albimanus* (Francischetti et al., 1999; Valenzuela et al., 1999).

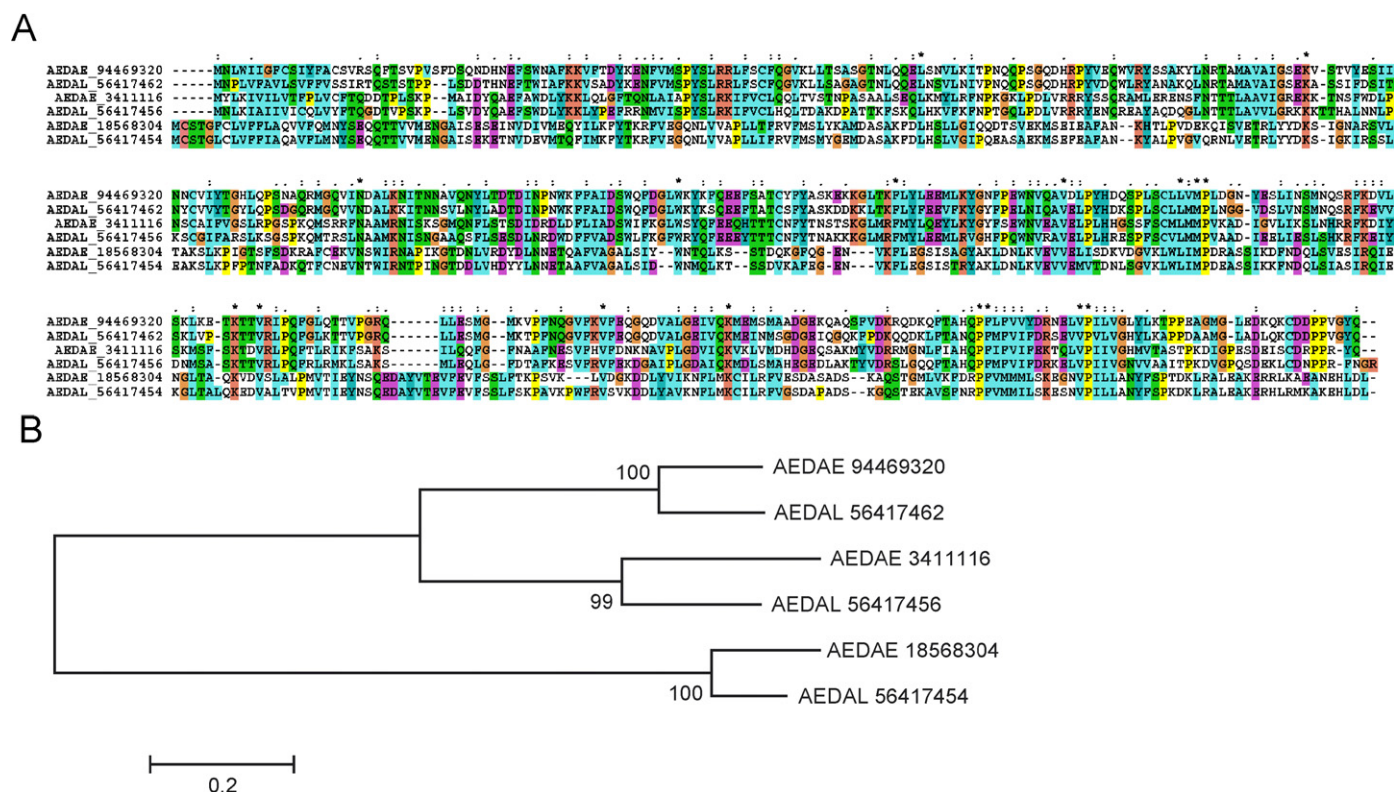


Fig. 4. (A) Clustal alignment of the three *Ae. albopictus* members of the serpin family with the two known *Ae. aegypti* salivary serpins. (B) Phylogram of the alignment. Bar indicates 20% distance in aa sequence. The numbers on the phylogram represent percentage bootstrap support for 10,000 trials.

Mucins: Due to their putative high numbers of serine and threonine and their high probability of having a substantial percentage of their backbone with *O*-linked *N*-acetylgalactosamine, four proteins are identified as mucins. Two of these code for mature peptides of less than 8 kDa; one is similar to a previously described *Ae. aegypti* peptide, and the other has no significant matches to eukaryotes. The remaining mucins are similar to previously described mosquito peritrophins or mucins. These proteins might function in the lubrication of the mosquito mouthparts. One of them (gi|56417396) has an indication of a chitin-binding site domain, and it is possible it binds to the chitinous linings of the salivary ducts and mouthparts to aid food canal lubrication. Transcripts encoding these proteins were found either in adult female salivary glands (gi|56417398) or in female salivary glands and in adult males but never in female carcasses (Table 3 and Fig. 3), a pattern compatible with expression in both male and female salivary glands and with the proposed physiological function.

Enzymes: The translation products for the enzymes apyrase, ADA, and purine nucleosidase are described. These three enzymes were found highly expressed in female (but not male) salivary glands of *Ae. aegypti*, where they may play a role in blood feeding by degrading purinergic mediators of hemostasis and inflammation (Champagne et al., 1995; Ribeiro et al., 2001; Ribeiro and Valenzuela, 2003). Edman degradation products matching the predicted mature amino terminal sequences of *Ae. albopictus*

apyrase and ADA were found in protein bands located at their expected positions (Fig. 1), indicating their abundant expression. In agreement with this observation, RT-PCR analysis also showed significant expression in female salivary glands; however, mRNA corresponding to these enzymes were additionally detected in males and, in the case of ADA, also in female carcasses, although at lower levels (Table 3 and Fig. 3). An identical expression pattern was found in the other culicine mosquito, *Ae. aegypti* (Ribeiro, J.M.C., et al., manuscript submitted). Mosquito apyrases evolved from the 5'-nucleotidase family by gene duplication and divergent evolution that involved the loss of a C-terminal membrane-anchoring portion, allowing the shift from a membrane-bound to a secreted protein and its recruitment to a blood-feeding role (Champagne et al., 1995). Two secreted members of the 5'-nucleotidase family were found in the saliva of *An. gambiae*, where they may both play a blood-feeding role (Lombardo et al., 2000). The first was found expressed specifically in female salivary glands and, for this reason, it was designated the putative apyrase of this mosquito. The second, expressed at significantly higher levels but not exclusively in female glands, was designated salivary 5'-nucleotidase. Also in the culicine mosquito *Ae. aegypti*, more than one member of the 5'-nucleotidase family is expressed in female salivary glands (Valenzuela et al., 2002b). In *Ae. albopictus* sialotranscriptome, only one 5'-nucleotidase family member was found, and its expression is not restricted to female salivary glands; it is possible that a second salivary member

of the 5'-nucleotidase family is present also in *Ae. albopictus* but not found in this transcriptome analysis due to a lower representation of the corresponding transcript. We also report the [homologue of *Ae. aegypti* salivary maltase](#), an enzyme that might help sugar digestion in these mosquitoes. In *Ae. aegypti*, this enzyme is found in both male and female salivary glands (James et al., 1989); the same holds true for *Ae. albopictus*, where the corresponding mRNA was detected both in adult female salivary glands and in adult males (Table 3 and Fig. 3). Finally, we report the sequence of a trypsin-like salivary serine protease. With the sole exception of an *Ae. aegypti* homologue (Ribeiro, J.M.C., et al., manuscript submitted), no significant matches were found in the NR protein database. In *Ae. albopictus* this enzyme appeared to be expressed only in female salivary glands whereas in *Ae. aegypti*, the corresponding transcript was also found in adult males, a pattern compatible with a gland-specific role. The function of this enzyme is unknown, confirming our ignorance regarding the role of other mosquito salivary serine proteases (Arca et al., 2005; Valenzuela et al., 2002b); however, its restricted distribution within mosquitoes suggests it may represent a trypsin-like serine protease specifically recruited to salivary-gland functions in *Aedes* mosquitoes.

Immune-related proteins and peptides: Possible antimicrobial polypeptides, lectins, a [Gram-negative binding protein](#), and fibrinogen-related proteins are included in this group, where their function may be related to preventing bacterial growth in the sugar meal and/or blood meal. Antimicrobial polypeptides here described include the *Ae. albopictus* salivary lysozyme, which is [87% identical to *Ae. aegypti* salivary lysozyme](#); three members of a histidine-rich peptide family; and one [peptide similar to *Ae. aegypti* i23M](#), which is annotated as being upregulated under *Plasmodium gallinaceum* challenge. One histidine-rich peptide of *Ae. albopictus* also qualifies as a mucin, having [19 putative galactosylation sites](#), and is [similar to *Aedes aegypti* and *Cx. quinquefasciatus* salivary polypeptides](#). Two other histidine-rich peptides are characterized by a [His repeat region followed by a proline-rich carboxyterminal region](#) and may be splice variants, as indicated by their alignment. Their presumed antimicrobial function derives from the histidine-rich stretch found in some antibacterial peptides (Lai et al., 2004), and their effect may derive from chelating essential Zn or Ni ions necessary for bacterial growth. Lysozyme and His-rich peptides mRNA were found both in female salivary glands and in adult males (Table 3 and Fig. 3), a pattern well matching with expression in salivary glands of both sexes and with the presumed role of limiting bacterial growth in the sugar- or blood-meal. Three C-type lectins and two proteins with fibrinogen/ficolin/angiopoietin domains or similarities are also reported. These proteins may assist immune functions in saliva and have been described before in the sialotranscriptomes of *Ae. aegypti* and *An. gambiae*. Notably, this group of immune-related secreted peptides described above is found only in adult female

salivary glands with the sole exception of a gram-negative binding protein whose expression is enriched in but not restricted to female glands (Table 3 and Fig. 3). It may thus be possible that salivary lectins are playing a role in inflammation or hemostasis, as is the case of C-type lectins found in snake venoms that work as anticlotting or platelet aggregation inhibitors (Jennings et al., 2005; Morita, 2005; Wang et al., 1999).

3.4.2. Putative secreted salivary proteins of unknown function

3.4.2.1. Members of ubiquitous protein families, unknown function. AG5 protein family: Four different genes coding for members of the AG5 family were identified by matching salivary transcripts. Notably, the same number of genes of this family was found expressed in *An. gambiae* adult female glands (Arca et al., 2005). This protein family, widely found in the venom of vespids (King and Spangfort, 2000) and in the salivary glands of many blood-sucking insects (Francischetti et al., 2002; Li et al., 2001; Valenzuela et al., 2002b), is found as a multigene family in most animal genomes such as *Drosophila* (Megraw et al., 1998). These proteins belong to the larger family of cysteine-rich extracellular proteins ubiquitously found in animals and plants (Schreiber et al., 1997). Most proteins have no known function, except in a few cases: one *Conus* protein was recently shown to have proteolytic activity (Milne et al., 2003), snake venom proteins of this same family have been shown to contain smooth-muscle-relaxing activity (Yamazaki et al., 2002; Yamazaki and Morita, 2004), and the salivary neurotoxin of the venomous lizard *Heloderma horridum* is also a member of this protein family (Nobile et al., 1996). Alignment of the four *Ae. albopictus* salivary AG5-related sequences with those of other Diptera (obtained by collecting a NR set of proteins representing the best 25 matches of each of the four *Ae. albopictus* sequences when compared by blastp to the NR NCBI protein database) followed by phylogeny reconstruction indicates the presence of four robust branches (Fig. 5).

- Branch I contains salivary proteins from diverse hematophagous Diptera including sand flies (*Phlebotomus papatasi*, *Phlebotomus ariasi*, and *Lutzomyia longipalpis*), culicine mosquitoes (*Ae. aegypti*, *Ae. albopictus*, and *Culex pipiens*), and *Culicoides sonorensis* found in previously described salivary transcriptomes (Campbell et al., 2005; Oliveira et al., 2006; Ribeiro et al., 2004; Valenzuela et al., 2002b). Assuming the AG5 members evolved within organisms by gene duplication, this branch may point to the original AG5 gene that was co-opted for salivary function early in the evolution of hematophagy among Diptera. Female salivary gland-specific expression of the two *Ae. albopictus* members included in this branch, gi|56417472 and gi|56417470, fits perfectly with this hypothesis (Table 3 and Fig. 3). Notably, members of any anopheline-derived protein are absent among this group.

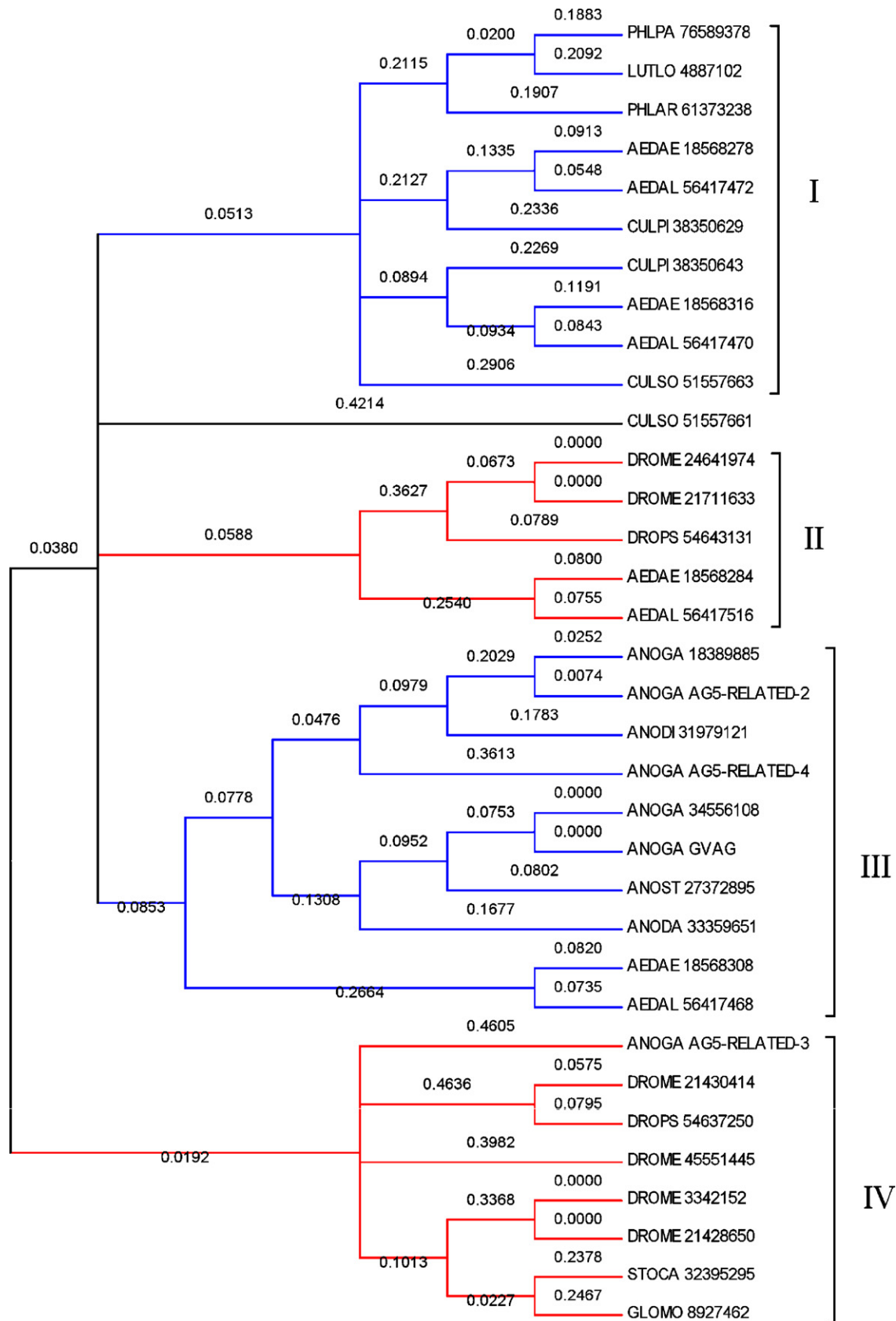


Fig. 5. Phylogenetic tree of selected members of the antigen-5 family of proteins from Diptera, obtained by the NJ method. The numbers in the tree branches indicate the percentage of time that the bootstrap procedure generated a different grouping. The four groups are marked with Roman numerals. For details, see text.

- Branch II contains both *Drosophila* and *Aedes* sequences, with the *Ae. albopictus* gi|56417516 that appears only slightly enriched in female salivary glands (Table 3 and Fig. 3).
- Branch III contains all known salivary-expressed anopheline sequences (except for one) from *An. gambiae* (Arca et al., 2005), *An. dirus*, *An. stephensi* (Valenzuela et al., 2003), and *An. darlingi* (Calvo et al., 2004) as well as proteins from *Ae. albopictus* and *Ae. aegypti*. This branch shows three distinct branches, two of which are exclusive to anophelines and the third exclusively ædine. The *Ae. albopictus* AG5 that is part of this branch (gi|56417516) was found expressed both in female salivary glands and in adult males (Table 3 and Fig. 3).
- Branch IV contains many *Drosophila* sequences but also one *An. gambiae* sequence and salivary AG5 molecules from the hematophagous higher Diptera, *Stomoxys calcitrans* and *Glossina morsitans* (Li et al., 2001). Considering that higher and lower Diptera evolved independently to hematophagy, this branch appears to contain the gene co-opted for the abundant expression of AG5 proteins found in the saliva of the higher flies: stable fly and tsetse.

Overall, it appears that mosquitoes have a set of four different AG5 proteins expressed in the salivary glands, if we can generalize from the observations above. Members of Branches I and III appear to be uniquely differentiated to exclude non-blood-feeding members. Accordingly, Branch I contains the blood-feeding Nematocera including culicines, sand flies, and biting midges; this branch may point to the oldest AG5 gene that was co-opted for salivary function early in the evolution of hematophagy among Diptera. It is tempting to speculate that the ancestral AG5 gene found in Branch I has gone extinct in anophelines; alternatively, it has yet to be discovered in further transcriptomes or proteome work. AG5 genes grouped in Branch III are instead exclusive to mosquitoes. Evidence for expression for two members of the *Ae. albopictus* AG5 proteins was confirmed by Edman degradation products for gi|56417516 (NYcDSsLcRQGPhVAcN [lower-case letters were not found in the Edman reaction]), a member of branch II, and gi|56417472 (KSKDYcSAaYKKIcEYKGEH), a member of Branch I (Fig. 1). These proteins are associated with two strong bands when separated by SDS-PAGE (which appear fused in Fig. 1), suggesting their strong expression in salivary glands. The enrichment in female salivary glands and the strong expression of gi|56417516 suggest a possible role in blood-feeding, even if this sequence is grouped in Branch II with *Drosophila* proteins. Perhaps it represents an AG5 gene independently recruited to blood-feeding roles in *Aedes* mosquitoes.

Kazal and TIL-domain containing peptides: Two peptides with mature molecular weight between 7.3 and 7.5 kDa and containing, respectively, a Kazal and a TIL domain are predicted. Although these domains are associated with protease inhibitors, they are also found in other extra-

cellular proteins without this function or containing antimicrobial activity (Fogaca et al., 2005). Kazal-type serine protease inhibitors usually contain more than one Kazal domain as tandem repeats. The *Ae. albopictus* Kazal peptide gi|56417460 has a single domain, but this aligns well with one of the Kazal domains of several proteins annotated as thrombin inhibitors; however, the transcript encoding this polypeptide was found in all tissue examined (Table 3 and Fig. 3), a pattern of expression not expected in an anti-hemostatic factor. A homologue has also been found in the sialotranscriptome of *Ae. aegypti*, where it exhibits the same ubiquitous expression profile (Ribeiro, J.M.C., et al., manuscript submitted). The *Ae. albopictus* TIL domain peptide is similar to several related salivary peptides previously reported in *An. gambiae* male and female transcriptomes. Identification of these TIL-containing peptides in male salivary glands (Calvo et al., 2006) points to their action as antimicrobial agents; expression of the corresponding gene would be compatible with such a role, as it was found expressed both in female salivary glands and in adult males (Table 3 and Fig. 3). The *Ae. aegypti* homologue shares 75% similarity with this *Ae. albopictus* TIL-domain peptide as well as an identical expression pattern (Ribeiro, J.M.C., et al., manuscript submitted).

3.4.2.2. Members of mosquito-specific protein families. Found in both culicines and anophelines

- **30-kDa allergen:** This family of proteins is ubiquitously found exclusively in the salivary glands of adult female mosquitoes, but its function is unknown to date. Recombinant members of this family from *Ae. aegypti* are known to be antigenic to humans (Simons and Peng, 2001). While only one member of this protein family is known in *An. gambiae*, it appears to have expanded in the genome of *Ae. albopictus*, where at least three genes code for the five different proteins reported in this transcriptome. Two of the proteins may derive from the same gene as alternative splice transcripts (Fig. 6), and two are probable alleles. All predicted proteins have an acidic pI below 4.4 and are very rich in Asp and Glu. All *Ae. albopictus* members of this family have 3–14 mucin-type glycosylations in their amino-terminal region. Of the three putative genes coding for this protein family in *Ae. albopictus*, one codes for the normally long protein of 30kDa having homologues in Anopheles, but the other two code for a smaller version lacking a strand of 60 aa residues rich in Gly, Asp, and Glu (shown in Fig. 6A in reverse turquoise; acidic aa in red). These proteins have three Cys residues in the mature form. Evidence for expression of gi|56417502 was obtained by Edman degradation, and the product EDDTEAXTEXP (Fig. 1) is a few aa beyond the predicted cleavage of the signal peptide. The position of this protein (predicted mature molecular weight of 20.7kDa) in the SDS gel (Fig. 1) suggests it may indeed be glycosylated. The three

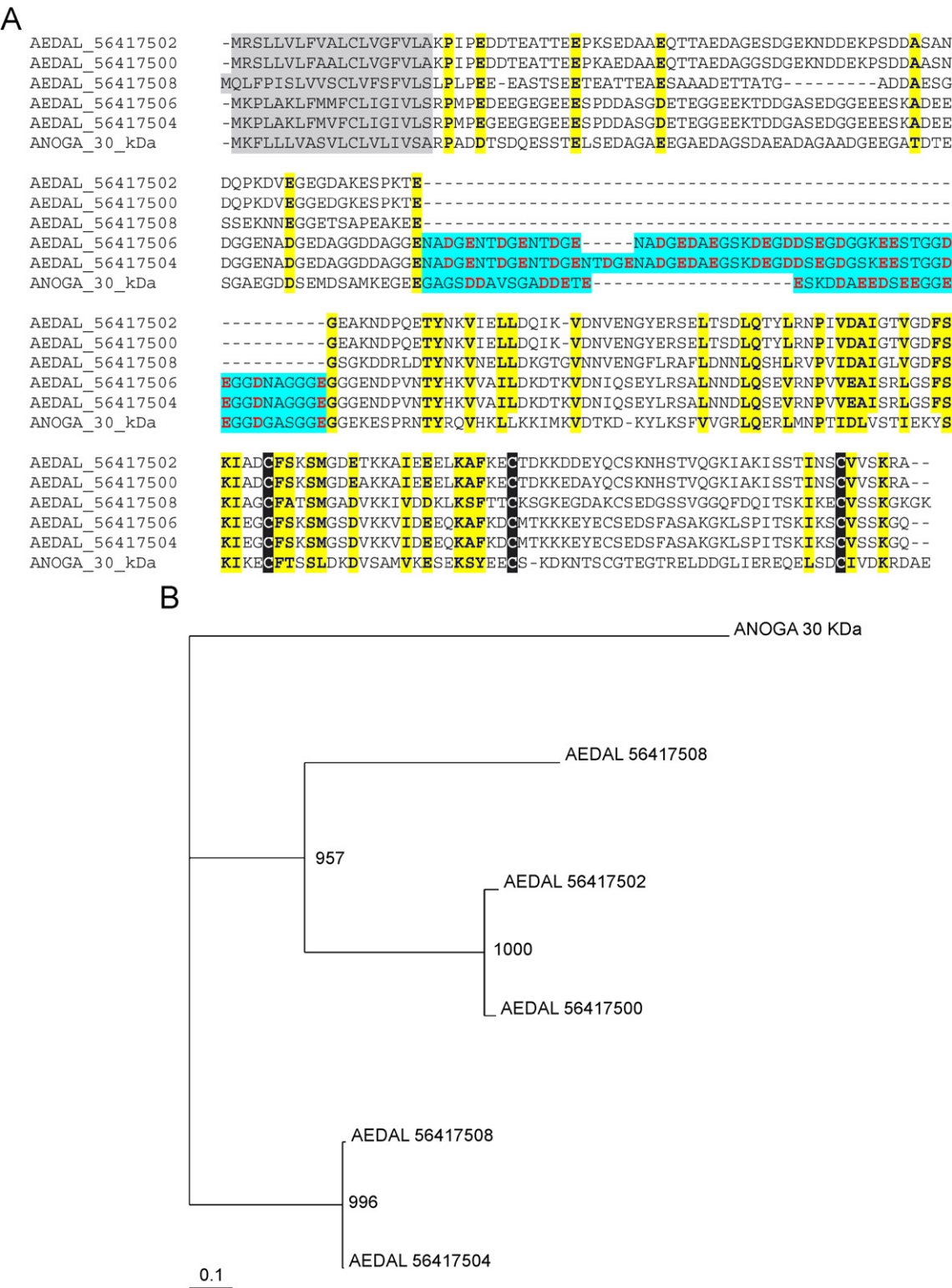


Fig. 6. (A) Clustal alignment of the four *Ae. albopictus* members of the 30-kDa antigen family with the single *An. gambiae* protein of the same family. (B) Phylogram of the alignment. The signal peptide region is shown in reverse gray. The numbers on the phylogram represent bootstrap support for 1000 trials. For details see text.

genes encoding for this protein family in *Ae. albopictus* are exclusively expressed in female salivary glands, as shown in Table 3 and Fig. 3. The 30-kDa family members also are female-gland specific in the yellow

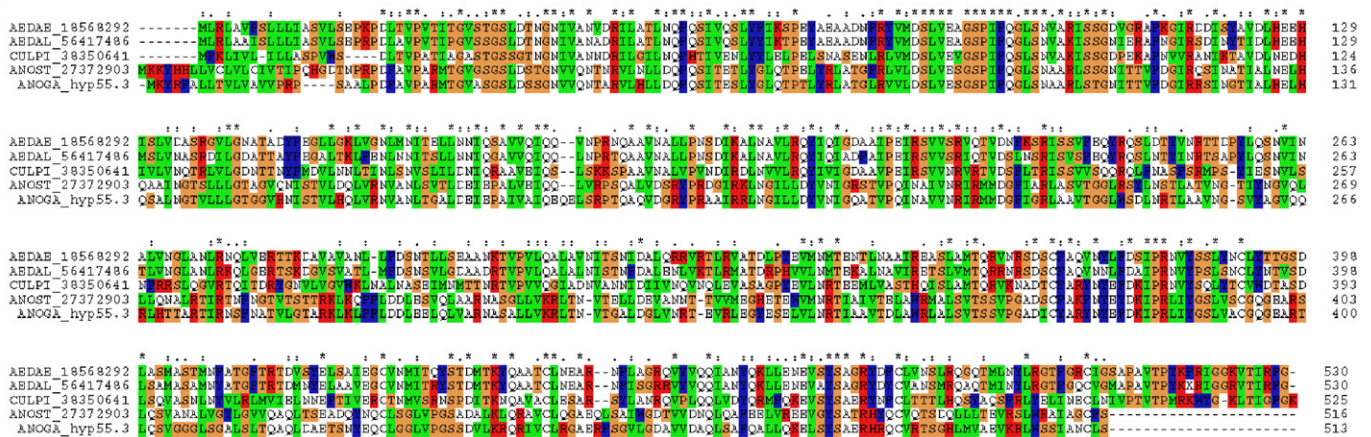
fever mosquito *Ae. aegypti*, where, however, only two genes encoding different splice variants and alleles have been reported (Ribeiro, J.M.C., et al., manuscript submitted).

- **56-kDa family:** This protein family is found exclusively in the salivary glands of culicines and anophelines. In *An. gambiae*, a single gene is apparent for this protein (Arca et al., 2005), and this appears to be the case for *Ae. aegypti* and most likely also for other mosquito species. In *Ae. albopictus*, the gene appears to be polymorphic; we report four alleles from the 56-kDa locus of *Ae. albopictus* (Supplemental Table S2). Alignment of proteins from five different species and three different genera indicates that, besides a very well-conserved 20 aa region at the amino-terminus, they are divergent enough between genera to preclude a clear phylogeny, indicating their rapid pace of change (Fig. 7). They may be good markers for closely related mosquito species, because they appear to be from single genes and under intense evolutionary change. In both *An. gambiae* and *Ae. aegypti*, the gene coding for the 56-kDa protein was found expressed in female salivary glands and in male mosquitoes (Arca et al., 2005), a pattern of expression identical to that observed in *Ae. albopictus* (Fig. 3). The presence of this protein family only in mosquitoes, the single-exon structure, and the retrieval of only mosquito and bacterial proteins after four reiterations of PSIblast led to the hypothesis that this gene family may have originated by lateral transfer from a bacterial genome (Ribeiro, J.M.C., et al., manuscript submitted). It is possible that the 56-kDa protein family plays a role as an antimicrobial agent in the crop-stored sugar meal.
- **16-kDa family:** One transcript from *Ae. albopictus* is similar to a previously reported salivary protein from *An. stephensi* (Valenzuela et al., 2003) and to a possible antennal protein from *An. gambiae* (Justice et al., 2003). No other significant matches to this protein sequence are found in the NR protein database. The gene encoding this peptide was found to have an ubiquitous profile of expression (not shown).

Proteins found so far only in culicine salivary gland transcriptomes, unknown function:

- **62-kDa family:** Members of this protein family have been found exclusively in *Aedes* mosquitoes, where they are abundantly transcribed. We report here the products of two genes and one allele present in the sialotranscriptome of *Ae. albopictus*. Transcripts corresponding to these genes were found only in female salivary glands (Table 3 and Fig. 3). The two genes code for proteins with 43% sequence identity and 65% similarity among themselves. Both *Ae. albopictus* genes have homologues in *Ae. aegypti*, where they share a larger degree of similarity than among proteins from the same species, indicating that the gene duplication event occurred well before speciation. Indication of expression of both proteins was obtained by Edman degradation of protein bands separated from salivary gland homogenate submitted to SDS-PAGE (Fig. 1). Due to the uniqueness among metazoan and also to the single-exon structure for this gene family, acquisition by horizontal transfer has been postulated. The abundant protein expression and apparent selectivity to the genus *Aedes* may make this protein family a good marker of *Aedes* exposure.
- **41.9-kDa basic protein family:** Members of this family are found in *Cx. quinquefasciatus* (38% identity) and in *Ae. aegypti*. No other significant similarities occur when these sequences are compared with the NR protein database. RT-PCR expression analysis indicated expression of this gene both in female glands and in adult males (Table 3 and Fig. 3). The function of members of this family is presently unknown.
- **34-kDa family:** Similar to the 62-kDa family, the 34-kDa family appears to be exclusive to *Aedes*. Two gene products are described here for *Ae. albopictus*, whereas several alleles and at least three different genes have been found in *Ae. aegypti*. In both species, members of this family are either specific or highly enriched in female glands (Table 3 and Fig. 3). The two *Ae. albopictus* translation products share only 33% amino acid sequence identity among themselves but significantly higher identities with their *Ae. aegypti* homologues, as is the case with the 62-kDa family members. The weak signal for the Edman degradation product NPTXKXXTVXE corresponding to the expected mature protein amino terminus was found at the expected location by SDS-PAGE (Fig. 1), indicating synthesis of this protein by the salivary glands. Also, this protein family in *Ae. aegypti* exhibits a single exon structure and, therefore, may have been introduced into the genome of culicine mosquitoes by horizontal transfer.
- **27-kDa family:** Similar to the 34 and 62 kDa families, the 27-kDa family appears to be also exclusive to *Aedes*. Two gene products are known for *Ae. aegypti*, and two are described here for *Ae. albopictus*. The two *Ae. albopictus* translation products share only 57% amino acid sequence identity among themselves, but they share a higher identity of 77% and 58% with their *Ae. aegypti* homologues, as is the case with the 62 and 34-kDa family members, all of which appear to be *Aedes*-specific, salivary gland-duplicated genes.
- **23.4-kDa family:** This unique protein family has one member in the *Culex pipiens quinquefasciatus* sialome that is 38% identical to the reported *Ae. albopictus* homologue. The gene encoding the 23.4-kDa protein showed female salivary gland-enriched expression (Table 3 and Fig. 3).
- **WPCWW 8.9-kDa peptide:** A putative salivary-secreted peptide with 8.9 kDa mass containing an unusual cluster of tryptophane residues has been previously described in the sialome of *Ae. aegypti* (Valenzuela et al., 2002b). One member of this family was found in the *Ae. Albopictus* sialotranscriptome, where it was found expressed in female salivary glands and in adult males (Table 3 and Fig. 3). This peptide is distantly related to a *C. p. quinquefasciatus* peptide of the same size, where the WPCWW motif is changed to WPFWF.

A



B

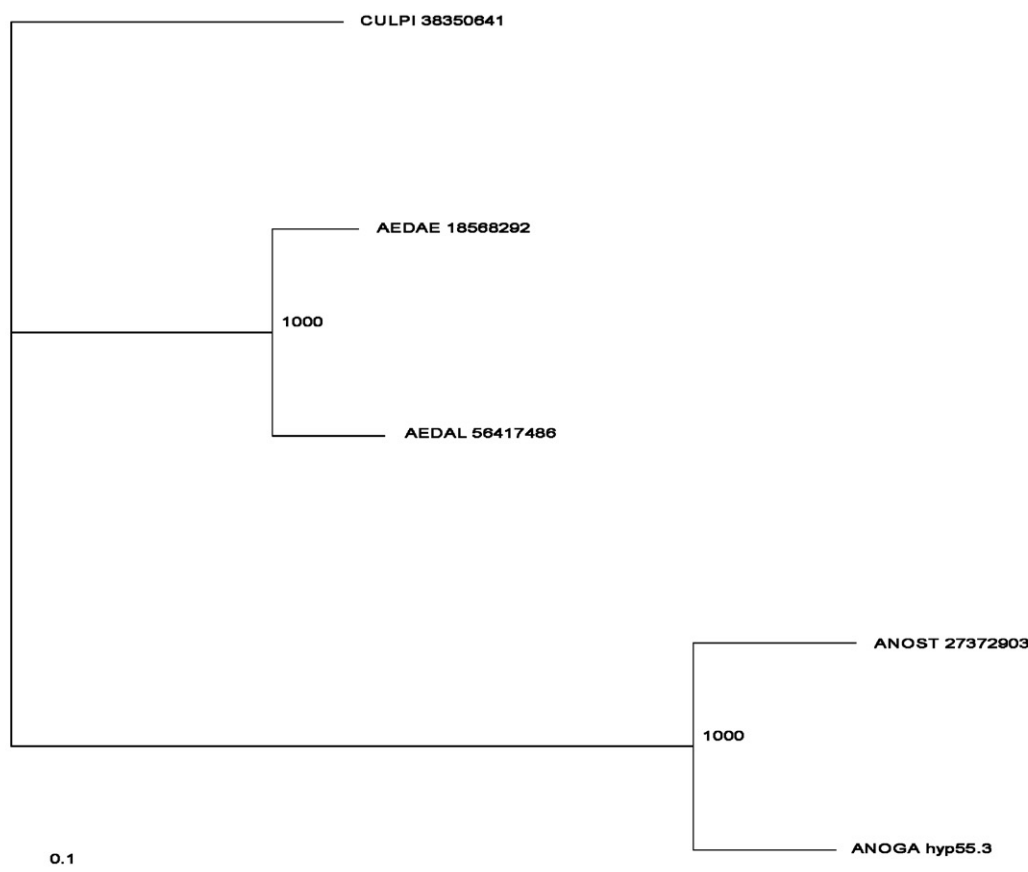


Fig. 7. The 56-kDa family of mosquito salivary proteins. (A) Clustal alignment. (B) Phylogram. The sequences derive from *An. gambiae*, *An. stephensi*, *Ae. albopictus*, *Ae. aegypti*, and *C. p. quinquefasciatus*. The numbers represent NCBI accession numbers, except for the *An. gambiae* protein, which was as annotated in <http://www.ncbi.nlm.nih.gov/projects/omes/Angambiaesialome-2005/Table4.xls>. The numbers on the phylogram represent bootstrap support for 1000 trials.

- **Basic 7.6-kDa peptide:** Two predicted peptides with mature mass of 7.6 kDa were found in the *Ae. albopictus* transcriptome. They share 85% identity at the amino acid sequence level, are specifically expressed in female salivary glands (Table 3 and Fig. 3), and most probably represent the effect of a gene duplication event. They share 68% and 62% homology to one

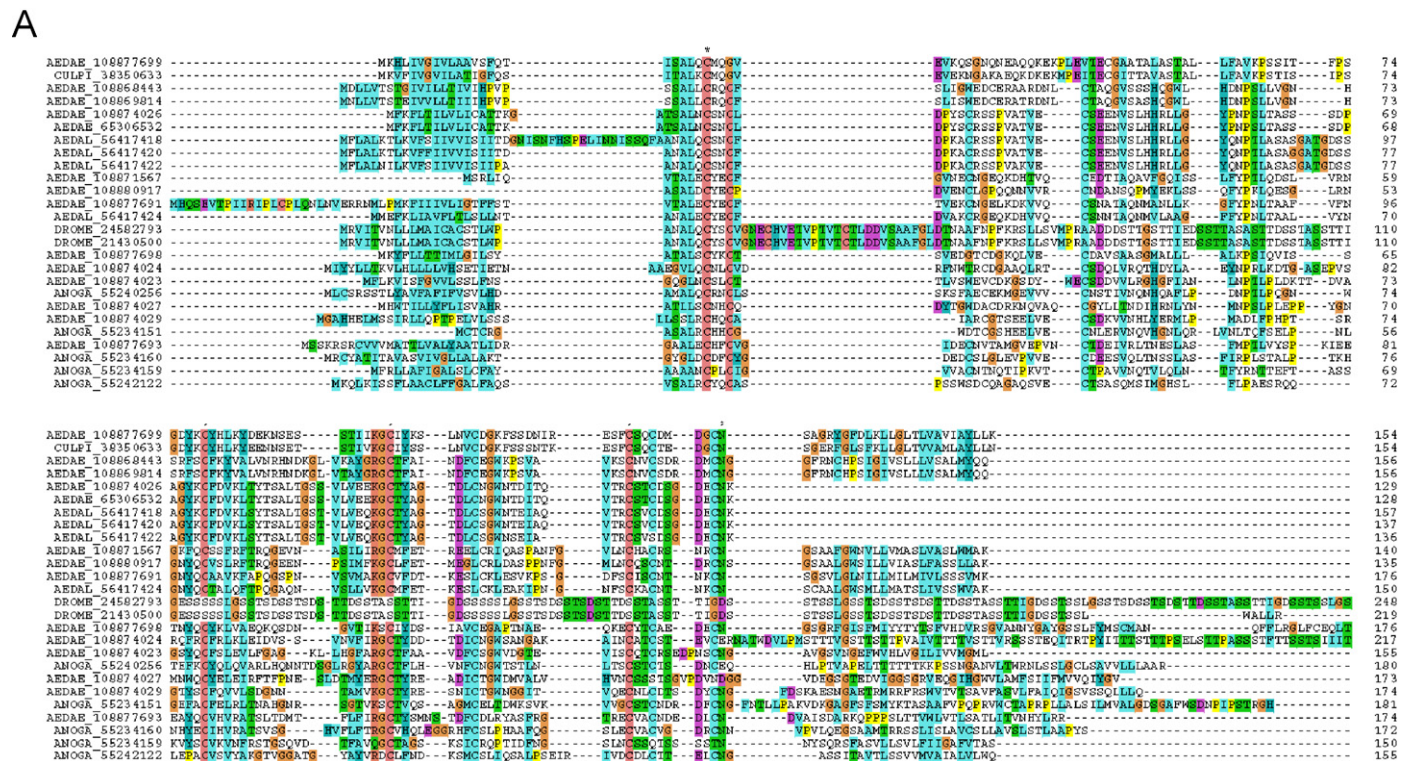
similar peptide reported for *Ae. aegypti*. Strong expression of at least one family member is indicated by Edman degradation sequence VSDVEKYX-KYLDXQGGK IKQ from protein bands of SDS-PAGE (Fig. 1), where the two X were predicted cysteines that do not yield a signal in the Edman procedure.

• **Basic 3.8-kDa peptide:** A putative secreted peptide of pI = 11 and mature molecular weight of 3.8 kDa is predicted in the sialotranscriptome of *Ae. albopictus*. This peptide shares 46% identity and 68% similarity with previously described peptides, respectively, from the salivary transcriptomes of *C. p. quinquefasciatus* (Ribeiro et al., 2004) and *Ae. aegypti* (Ribeiro, J.M.C., et al., manuscript submitted). Expression of this peptide is confirmed by Edman product GKSSXXXTIR-AGKTGGAAER, matching the expected amino terminus of the predicted mature peptide. RT-PCR expression analysis indicated highly enriched expression in female glands (Table 3 and Fig. 3). Members of this peptide family have no additional similarities to other proteins in the NR database and may represent a culicine-specific family possibly functioning as antimicrobial agents.

3.4.2.3. Other predicted salivary polypeptides of unknown function. In this group, we describe eight additional gene products expressed in the salivary glands of *Ae. albopictus* either found uniquely in the salivary transcriptome of *Ae. albopictus* or distant members of other Diptera families.

Four of these proteins represent a gene family producing mature peptides in the range of 11–14 kDa that for this

reason are indicated as the 13-kDa family. Among these, three are probably the result of gene duplications and one appears to be a **longer splice variant** of the same gene. To obtain more information on this protein family, sequence [gi|56417418](#) was submitted to PSIblast against the NR database. Convergence was obtained after seven iterations, identifying 26 related proteins (results not shown), including all *Ae. albopictus* sequences reported in this group. Except for two *Drosophila* proteins, all others were from mosquitoes. Three of these sequences were reported previously in sialotranscriptomes of *Ae. aegypti*, *Culex pipiens quinquefasciatus*, and in male sialotranscriptomes of *An. gambiae* (see hyperlinks on Fig. 8B). Alignment of these sequences indicates that the main similarities reside on a framework of 10 cysteines (Fig. 8A). The *Drosophila* sequences, as well as two *Ae. aegypti* sequences, are over 200 aa in length, while all other proteins have fewer than 177 residues. Realignment of the smaller set of 22 proteins clarifies the cysteine framework (Fig. 8B). Note that two proteins lack cysteines 2 and 3, and another two lack cysteines 9 and 10. One protein lacks Cys 9. Additionally, six proteins have an extra cysteine residue in their carboxyterminal region (* in Fig. 8B). The finding of a member of this family in male *Anopheles* sialotranscriptomes points to a function either related to sugar feeding or as an antimicrobial agent, suggesting that members of this



B

	1	2	3	4	5	6
AEDAE_108877699	--TISALQ	CMQGV	EVKQSGNQ	EAQKQ	KEPLEVTE	CGAAT
CULPI_38350633	--SITALK	CMQGV	EVKNGAKA	EQKDK	EMPEITE	CGITTA
AEDAE_108871567	--QVTALE	CYEC	FGVNE	---CNGE	---QKDHT	VQCED
AEDAE_108880917	---ASALD	CYEC	PDVEN	---CLGP	---QONNV	VRCDANS
AEDAE_108877691	--TVTALE	CYEC	FTVEK	---CNGE	---LKD	KVVQ
AEDAL_56417424	--TANALE	CYEC	FDVAK	---CRGE	---QKDH	VVQ
AEDAE_108869814	--VPSSALL	CRQC	FLISWEDC	---ERATRD	NLCTA	QGV
AEDAE_108877693	--RGAAL	CHFC	VGIDE	---CNVT	---AMG	VEPVN
ANOGA_55234160	---YGLD	CDFO	YGDED	---CSLG	---LEV	PVVE
ANOGA_55240256	---AMALQ	CRNCL	SSSFABC	---EKMGE	VVVCNST	IVNQ
AEDAE_108874023	---GQGLN	CSLCT	TLVSWEVC	---DKGSD	YWECSDD	VVLRG
ANOGA_55242122	---VSALRC	YCCAS	PSWSDC	---QAGA	QSVECTS	ASQMS
AEDAE_108874029	---LLSLR	CHQCA	IAR	---CGTS	---EELVE	---CSDK
ANOGA_55234151	---ASALR	CHHCG	WDT	---CGSH	---EELVE	---CNLER
AEDAE_108868443	--VPSSALL	CRQC	FLISWEDC	---ERAARD	NLCTA	QGV
ANOGA_55234159	---AAAAN	CPLCT	GVVA	---CNTN	---QTI	PIKVT
AEDAE_108874027	---ATLLS	CHNC	QDYTGWDAC	---DRKN	QVACQ	GYLLT
AEDAL_56417418	QFAANAL	QCSN	CFDPKA	---CRSS	---PVATVE	---CSEEN
AEDAL_56417420	---ANALQ	CSNCF	DPKA	---CRSS	---PVATVE	---CSEEN
AEDAL_56417422	---ANALQ	CSNCF	DPKA	---CRSS	---PVAKVE	---CSEEN
AEDAE_108874026	--ATSALN	CSNCL	DPYS	---CRSS	---PVATVE	---CSEEN
AEDAE_65306532	--TSALN	CSNCL	DPYS	---CRSS	---PVATVE	---CSEEN
	7	8	9	10	*	*
AEDAE_108877699	LN--VCDG	KFSSD	--NIRE	SFSCQ	CDM--DGCN	S--
CULPI_38350633	LN--VCDG	KFSSS	--NTKE	SFSCQ	CTE--DGCN	S--
AEDAE_108871567	REE--LCRI	QASPA	--NFGV	LNCHACRS	---NRON	G--
AEDAE_108880917	MEG--LCRL	DASPP	--NFGM	LNQSCNT	---DRON	SG--
AEDAE_108877691	KES--ICKLE	SVKP	--S-GDFS	CSICNT	---NKON	S--
AEDAL_56417424	KES--ICKLE	AKIP	--N-GNFS	CKACNT	---NKON	S--
AEDAE_108869814	ND--FCGE	WKPSV	--A-VKSON	VCSDR	---DMONG	G--
AEDAE_108877693	NST--DFED	LDYASF	--R-GTREC	VACNDE	---DLOND	V--
ANOGA_55234160	EGGRH	FCSL	PHAAF	--QG-SLEC	VACVG	---DRON
ANOGA_55240256	VN--FCNG	WTSTL	--N-LTSC	STCTS	---DNCE	QHLPT
AEDAE_108877693	VD--FCSG	VVDGT	--E-VISC	QTCRSED	PNSCN	---
ANOGA_55242122	KS--MCSL	IQALP	SEIRI	VDCDLC	TTT--ELCN	---
AEDAE_108874029	SN--ICTG	WNGGIT	---VQEO	NLCOTS	---DYCN	GFDS
ANOGA_55234151	AG--ICEL	DLRWKS	---VKVVG	CSTCND	---DFCNG	FNTLL
AEDAE_108868443	ND--FCGE	WKPSV	--A-VKSON	VCSDR	---DMONG	G--
ANOGA_55234159	KS--ICRQ	PTIDF	--N-GSLN	CSSTSS	---SSTNN	---
AEDAE_108874027	AD--ICTG	WDMVA	--LV-HVNC	SSSTSG	VPDV	---NDGG
AEDAL_56417418	TD--ICSG	WNTEIA	---QVTR	CSVCD	SG--DECN	K--
AEDAL_56417420	TD--ICSG	WNTEIA	---QVTR	CSVCD	SG--DECN	K--
AEDAL_56417422	TD--ICSG	WNTEIA	---QVTR	CSVCD	SG--DECN	K--
AEDAE_108874026	TD--ICNG	WNTEIT	---QVTR	CSVCD	SG--DECN	K--
AEDAE_65306532	TD--ICNG	WNTEIT	---QVTR	CSVCD	SG--DECN	K--

Fig. 8. (Continued)

family have been co-opted to be transcribed in the salivary glands of mosquitoes. Accordingly, all the corresponding transcripts but one were found enriched in female salivary glands (Table 3 and Fig. 3).

The remaining four gene products of this group, which can be browsed in Supplemental Table S2, showed a ubiquitous profile of expression (Table 3).

3.4.3. Housekeeping genes

Supplemental Table S2 displays 39 proteins associated with housekeeping functions of the salivary glands. They include a member of the V-ATPase complex, which has been shown to be important for salivary secretion in insects (Zimmermann et al., 2003), several ribosomal proteins, ubiquitin and ubiquitin-conjugating enzymes, cytochromes, and enzymes associated with mitochondrial ATP synthesis. These are available for browsing in Supplemental Table S2.

3.5. Comparisons on the rate of evolution of housekeeping and salivary gland genes between *Ae. aegypti* and *Ae. albopictus*

It has been proposed that salivary proteins from adult female American sand flies are under strong selection due to the deleterious effect that host immunity has on feeding (Lanzaro et al., 1999; Milleron et al., 2004). In support of this observation, when comparing the sialotranscriptomes of *An. stephensi* with that of *An. gambiae* (Valenzuela et al., 2003), it was noted that the putative secreted salivary proteins were $62 \pm 15\%$ identical (average $\pm$ SD), while the housekeeping protein products were $93 \pm 6\%$ identical, suggesting a faster pace of evolution for salivary gland proteins of these two anthropophilic mosquitoes belonging to the same subgenus. To compare our data set on *Ae. albopictus* with that of the closely related *Ae. aegypti*, we used the data from Supplemental Table S2, which presents

68 proteins from *Ae. albopictus* which are attributed with a salivary secretory function, and 39 proteins attributed to housekeeping function. These proteins were compared to all proteins from *Ae. aegypti* deposited in Genbank (which includes gene release 1.1 and all our recently deposited salivary sequences). Because the *Ae. albopictus* set contains alleles of the same genes, these were removed from the data set by excluding any pair of *Ae. albopictus* sequences matching the same *Ae. aegypti* sequence. We thus obtained a set of 55 putative salivary secreted proteins, and 38 housekeeping proteins matching *Ae. aegypti* proteins. The housekeeping group were $94 \pm 13\%$ (Mean $\pm$ SD) identical to the best matched *Ae. aegypti* proteins, while the salivary proteins had significantly lower identity of $71.5 \pm 11\%$. These results support the hypothesis that salivary gland genes are at a faster pace of evolution than housekeeping genes. It should be reminded in this context that while blood feeding *Aedes* mosquitoes also secrete components associated with sugar feeding, such as the maltase enzyme, although the reverse is not true (Marinotti et al., 1990). Accordingly, salivary components associated with sugar feeding could contribute to allergic reactions at the site of the bite as well as components adapted to disrupt host physiological processes.

3.6. Concluding remarks

For several years, our knowledge of mosquito salivary glands proceeded through classical biochemical analysis: typically, the initial identification of a given activity in the saliva was followed by protein purification, peptide sequencing, and gene cloning. In the last few years, development of techniques for high-throughput analysis allowed rapid acquisition of large amount of information on the sialotranscriptomes of arthropod disease vectors. Detailed lists of secreted salivary proteins are presently available for two very important mosquito vectors: the African malaria mosquito *An. gambiae* (Arca et al., 2005) and the yellow fever and dengue vector *Ae. aegypti* (Ribeiro, J.M.C., et al., manuscript submitted). Moreover, insights into the transcriptomes and proteomes of the Asian and South American malaria vectors *An. stephensi* and *An. darlingi* (Valenzuela et al., 2003; Calvo et al., 2004) and of the cosmopolitan arboviral vector *C. p. quinquefasciatus* (Ribeiro et al., 2004) have been obtained. This report adds information on the sialotranscriptome of an additional *Aedes* species, *Ae. albopictus*. This is probably the most readily spreading mosquito in the world, a potential vector of several arboviruses, and one of the primary mosquitoes responsible of the phenomenon of hypersensitivity to mosquito bites. These analyses allowed us to discover several mosquito salivary protein families that are unique among metazoans and whose functions are still unknown, highlighting how much we still have to learn about mosquito salivary components. The salivary glands are placed at the interface between the pathogens that mosquitoes transmit and the human host. Increasing

experimental evidence suggests that salivary proteins are not only crucial for blood feeding but also play important roles in pathogen transmission and in host immune and inflammatory responses. We believe that deeper knowledge of the salivary secretions of arthropod disease vectors will allow in the near future a better understanding of the complex relationships between pathogens, vectors, and hosts and perhaps, will be of help in devising strategies to better control some important human and/or animal diseases.

Acknowledgements

This work was supported by the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health (J.M.C.R.) and in part by grants from the European Union (M. Coluzzi and BA; BioMalPar N 503578). We thank Mark Garfield from the NIAID Research Technologies Branch for performing the Edman degradation reactions, the NIAID intramural editor Brenda Rae Marshall for assistance, and Chuong Huynh (NCBI) for help with posting the data. We also thank Roberto Romi for providing *Ae. albopictus* eggs and Maria Calzetta for technical assistance.

References

- Altschul, S.F., Gish, W., 1996. Local alignment statistics. *Methods Enzymol.* 266, 460–480.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402.
- Arca, B., Lombardo, F., Valenzuela, J.G., Francischetti, I.M., Marinotti, O., Coluzzi, M., Ribeiro, J.M., 2005. An updated catalogue of salivary gland transcripts in the adult female mosquito, *Anopheles gambiae*. *J. Exp. Biol.* 208, 3971–3986.
- Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al., 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25, 25–29.
- Bateman, A., Birney, E., Durbin, R., Eddy, S.R., Howe, K.L., Sonnhammer, E.L., 2000. The Pfam protein families database. *Nucleic Acids Res.* 28, 263–266.
- Calvo, E., Andersen, J., Francischetti, I.M., de, L.C.M., deBianchi, A.G., James, A.A., Ribeiro, J.M., Marinotti, O., 2004. The transcriptome of adult female *Anopheles darlingi* salivary glands. *Insect Mol. Biol.* 13, 73–88.
- Calvo, E., Mans, B.J., Andersen, J.F., Ribeiro, J.M., 2005. Function and evolution of a mosquito salivary protein family. *J. Biol. Chem.*
- Calvo, E., Mans, B.J., Andersen, J.F., Ribeiro, J.M., 2006. Function and evolution of a mosquito salivary protein family. *J. Biol. Chem.* 281, 1935–1942.
- Campbell, C.L., Vandyke, K.A., Letchworth, G.J., Drolet, B.S., Hanekamp, T., Wilson, W.C., 2005. Midgut and salivary gland transcriptomes of the arbovirus vector *Culicoides sonorensis* (Diptera: Ceratopogonidae). *Insect Mol. Biol.* 14, 121–136.
- Cancrini, G., Frangipane di Regalbano, A., Ricci, I., Tessarin, C., Gabrielli, S., Pietrobelli, M., 2003. *Aedes albopictus* is a natural vector of *Dirofilaria immitis* in Italy. *Vet. Parasitol.* 118, 195–202.
- Champagne, D.E., Smartt, C.T., Ribeiro, J.M., James, A.A., 1995. The salivary gland-specific apyrase of the mosquito *Aedes aegypti* is a

- member of the 5'-nucleotidase family. *Proc. Natl. Acad. Sci. USA* 92, 694–698.
- Fogaca, A.C., Almeida, I.C., Eberlin, M.N., Tanaka, A.S., Bulet, P., Daffre, S., 2005. Ixodidin, a novel antimicrobial peptide from the hemocytes of the cattle tick *Boophilus microplus* with inhibitory activity against serine proteinases. *Peptides*.
- Francischetti, I.M., Valenzuela, J.G., Ribeiro, J.M., 1999. Anophelin: kinetics and mechanism of thrombin inhibition. *Biochemistry* 38, 16678–16685.
- Francischetti, I.M., Valenzuela, J.G., Pham, V.M., Garfield, M.K., Ribeiro, J.M., 2002. Toward a catalog for the transcripts and proteins (sialome) from the salivary gland of the malaria vector *Anopheles gambiae*. *J. Exp. Biol.* 205, 2429–2451.
- Gratz, N.G., 2004. Critical review of the vector status of *Aedes albopictus*. *Med. Vet. Entomol.* 18, 215–227.
- Gubler, D.J., 2003. *Aedes albopictus* in Africa. *Lancet Infect. Dis.* 3, 751–752.
- Hansen, J.E., Lund, O., Tolstrup, N., Gooley, A.A., Williams, K.L., Brunak, S., 1998. NetOglyc: prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility. *Glycoconj. J.* 15, 115–130.
- Huang, X., Madan, A., 1999. CAP3: A DNA sequence assembly program. *Genome Res.* 9, 868–877.
- Isawa, H., Yuda, M., Orito, Y., Chinzei, Y., 2002. A mosquito salivary protein inhibits activation of the plasma contact system by binding to factor XII and high molecular weight kininogen. *J. Biol. Chem.* 277, 13–13.
- James, A.A., Blackmer, K., Racioppi, J.V., 1989. A salivary gland-specific, maltase-like gene of the vector mosquito, *Aedes aegypti*. *Gene* 75, 73–83.
- James, A.A., Blackmer, K., Marinotti, O., Ghosn, C.R., Racioppi, J.V., 1991. Isolation and characterization of the gene expressing the major salivary gland protein of the female mosquito, *Aedes aegypti*. *Mol. Biochem. Parasitol.* 44, 245–253.
- Jennings, B., Spearman, W., Shephard, E., 2005. A novel 25 kDa protein from the venom of *Bitis arietans* with similarity to C-type lectins causes fibrinogen-dependent platelet agglutination. *Toxicon* 46, 687–698.
- Justice, R.W., Dimitratos, S., Walter, M.F., Woods, D.F., Biessmann, H., 2003. Sexual dimorphic expression of putative antennal carrier protein genes in the malaria vector *Anopheles gambiae*. *Insect Mol. Biol.* 12, 581–594.
- King, T.P., Spangfort, M.D., 2000. Structure and biology of stinging insect venom allergens. *Int. Arch. Allergy Immunol.* 123, 99–106.
- Korochkina, S., Barreau, C., Pradel, G., Jeffery, E., Li, J., Natarajan, R., Shabanowitz, J., Hunt, D., Frevert, U., Vernick, K.D., 2006. A mosquito-specific protein family includes candidate receptors for malaria sporozoite invasion of salivary glands. *Cell. Microbiol.* 8, 163–175.
- Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform.* 5, 150–163.
- Lai, R., Takeuchi, H., Lomas, L.O., Jonczyk, J., Rigden, D.J., Rees, H.H., Turner, P.C., 2004. A new type of antimicrobial protein with multiple histidines from the hard tick, *Amblyomma hebraeum*. *Faseb J.* 18, 1447–1449.
- Lanzaro, G.C., Lopes, A.H., Ribeiro, J.M., Shoemaker, C.B., Warburg, A., Soares, M., Titus, R.G., 1999. Variation in the salivary peptide, maxadilan, from species in the *Lutzomyia longipalpis* complex. *Insect Mol. Biol.* 8, 267–275.
- Lewis, S., Ashburner, M., Reese, M.G., 2000. Annotating eukaryote genomes. *Curr. Opin. Struct. Biol.* 10, 349–354.
- Li, S., Kwon, J., Aksoy, S., 2001. Characterization of genes expressed in the salivary glands of the tsetse fly, *Glossina morsitans morsitans*. *Insect Mol. Biol.* 10, 69–76.
- Lombardo, F., Di Cristina, M., Spanos, L., Louis, C., Coluzzi, M., Arca, B., 2000. Promoter sequences of the putative *Anopheles gambiae* apyrase confer salivary gland expression in *Drosophila melanogaster*. *J. Biol. Chem.* 275, 23861–23868.
- Marchler-Bauer, A., Panchenko, A.R., Shoemaker, B.A., Thiessen, P.A., Geer, L.Y., Bryant, S.H., 2002. CDD: a database of conserved domain alignments with links to domain three-dimensional structure. *Nucleic Acids Res.* 30, 281–283.
- Marinotti, O., James, A., Ribeiro, J.M.C., 1990. Diet and salivation in female *Aedes aegypti* mosquitoes. *J. Insect Physiol.* 36, 545–548.
- Megraw, T., Kaufman, T.C., Kovalick, G.E., 1998. Sequence and expression of *Drosophila* Antigen 5-related 2, a new member of the CAP gene family. *Gene* 222, 297–304.
- Milleron, R.S., Ribeiro, J.M., Elnaime, D., Soong, L., Lanzaro, G.C., 2004. Negative effect of antibodies against maxadilan on the fitness of the sand fly vector of American visceral leishmaniasis. *Am. J. Trop. Med. Hyg.* 70, 278–285.
- Milne, T.J., Abbenante, G., Tyndall, J.D., Halliday, J., Lewis, R.J., 2003. Isolation and characterization of a cone snail protease with homology to CRISP proteins of the pathogenesis-related protein superfamily. *J. Biol. Chem.* 278, 31105–31110.
- Morita, T., 2005. Structures and functions of snake venom CLPs (C-type lectin-like proteins) with anticoagulant-, procoagulant-, and platelet-modulating activities. *Toxicon* 45, 1099–1114.
- Nielsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* 10, 1–6.
- Nobile, M., Noceti, F., Prestipino, G., Possani, L.D., 1996. Helothermine, a lizard venom toxin, inhibits calcium current in cerebellar granules. *Exp. Brain Res.* 110, 15–20.
- Novak, M.G., Rowley, W.A., 1994. Serotonin depletion affects blood-feeding but not host-seeking in *Aedes triseriatus* (Diptera: Culicidae). *J. Med. Entomol.* 31, 600–606.
- Novak, M.G., Ribeiro, J.M.C., Hildebrand, J.G., 1995. 5-Hydroxytryptamine in the salivary glands of adult female *Aedes aegypti* and its role in regulation of salivation. *J. Exp. Biol.* 198, 167–174.
- Oliveira, F., Kamhawi, S., Seitz, A.E., Pham, V.M., Guigal, P.M., Fischer, L., Ward, J., Valenzuela, J.G., 2006. From transcriptome to immunome: identification of DTH inducing proteins from a *Phlebotomus ariasi* salivary gland cDNA library. *Vaccine* 24, 374–390.
- Peng, Z., Simons, F.E., 2004. Mosquito allergy: immune mechanisms and recombinant salivary allergens. *Int. Arch. Allergy Immunol.* 133, 198–209.
- Peng, Z., Beckett, A.N., Engler, R.J., Hoffman, D.R., Ott, N.L., Simons, F.E., 2004. Immune responses to mosquito saliva in 14 individuals with acute systemic allergic reactions to mosquito bites. *J. Allergy Clin. Immunol.* 114, 1189–1194.
- Ribeiro, J.M.C., 1995. Blood-feeding arthropods: live syringes or invertebrate pharmacologists? *Infect. Agents Dis.* 4, 143–152.
- Ribeiro, J.M., Francischetti, I.M., 2003. Role of arthropod saliva in blood feeding: sialome and post-sialome perspectives. *Annu. Rev. Entomol.* 48, 73–88.
- Ribeiro, J.M., Valenzuela, J.G., 2003. The salivary purine nucleosidase of the mosquito, *Aedes aegypti*. *Insect Biochem. Mol. Biol.* 33, 13–22.
- Ribeiro, J.M., Charlab, R., Valenzuela, J.G., 2001. The salivary adenosine deaminase activity of the mosquitoes *Culex quinquefasciatus* and *Aedes aegypti*. *J. Exp. Biol.* 204, 2001–2010.
- Ribeiro, J.M., Charlab, R., Pham, V.M., Garfield, M., Valenzuela, J.G., 2004. An insight into the salivary transcriptome and proteome of the adult female mosquito *Culex pipiens quinquefasciatus*. *Insect Biochem. Mol. Biol.* 34, 543–563.
- Romi, R., 2001. *Aedes albopictus* in Italy: an underestimated health problem. *Ann. Ist Super Sanita* 37, 241–247.
- Sabatini, A., Raineri, V., Trovato, G., Coluzzi, M., 1990. *Aedes albopictus* in Italy and possible diffusion of the species into the Mediterranean area. *Parassitologia* 32, 301–304.
- Schreiber, M.C., Karlo, J.C., Kovalick, G.E., 1997. A novel cDNA from *Drosophila* encoding a protein with similarity to mammalian cysteine-rich secretory proteins, wasp venom antigen 5, and plant group I pathogenesis-related proteins. *Gene* 191, 135–141.

- Schultz, J., Copley, R.R., Doerks, T., Ponting, C.P., Bork, P., 2000. SMART: a web-based tool for the study of genetically mobile domains. *Nucleic Acids Res.* 28, 231–234.
- Simons, F.E., Peng, Z., 2001. Mosquito allergy: recombinant mosquito salivary antigens for new diagnostic tests. *Int. Arch. Allergy Immunol.* 124, 403–405.
- Stark, K.R., James, A.A., 1995. A factor Xa-directed anticoagulant from the salivary glands of the yellow fever mosquito *Aedes aegypti*. *Exp. Parasitol.* 81, 321–331.
- Stark, K.R., James, A.A., 1998. Isolation and characterization of the gene encoding a novel factor Xa-directed anticoagulant from the yellow fever mosquito, *Aedes aegypti*. *J. Biol. Chem.* 273, 20802–20809.
- Tatem, A.J., Hay, S.I., Rogers, D.J., 2006. Global traffic and disease vector dispersal. *Proc. Natl. Acad. Sci. USA* 103, 6242–6247.
- Tatusov, R.L., Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Kiryutin, B., Koonin, E.V., Krylov, D.M., Mazumder, R., Mekhedov, S.L., Nikolskaya, A.N., et al., 2003. The COG database: an updated version includes eukaryotes. *BMC Bioinform.* 4, 41.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25, 4876–4882.
- Valenzuela, J.G., Francischetti, I.M., Ribeiro, J.M., 1999. Purification, cloning, and synthesis of a novel salivary anti-thrombin from the mosquito *Anopheles albimanus*. *Biochemistry* 38, 11209–11215.
- Valenzuela, J.G., Charlab, R., Gonzalez, E.C., Miranda-Santos, I.K.F., Marinotti, O., Francischetti, I.M., Ribeiro, J.M.C., 2002a. The D7 family of salivary proteins in blood sucking Diptera. *Insect Mol. Biol.* 11, 149–155.
- Valenzuela, J.G., Pham, V.M., Garfield, M.K., Francischetti, I.M., Ribeiro, J.M.C., 2002b. Toward a description of the sialome of the adult female mosquito *Aedes aegypti*. *Insect Biochem. Mol. Biol.* 32, 1101–1122.
- Valenzuela, J.G., Francischetti, I.M., Pham, V.M., Garfield, M.K., Ribeiro, J.M., 2003. Exploring the salivary gland transcriptome and proteome of the *Anopheles stephensi* mosquito. *Insect Biochem. Mol. Biol.* 33, 717–732.
- Wang, R., Kini, R.M., Chung, M.C., 1999. Rhodocetin, a novel platelet aggregation inhibitor from the venom of *Calloselasma rhodostoma* (Malayan pit viper): synergistic and noncovalent interaction between its subunits. *Biochemistry* 38, 7584–7593.
- Yamazaki, Y., Morita, T., 2004. Structure and function of snake venom cysteine-rich secretory proteins. *Toxicon* 44, 227–231.
- Yamazaki, Y., Koike, H., Sugiyama, Y., Motoyoshi, K., Wada, T., Hishinuma, S., Mita, M., Morita, T., 2002. Cloning and characterization of novel snake venom proteins that block smooth muscle contraction. *Eur. J. Biochem.* 269, 2708–2715.
- Zimmermann, B., Dames, P., Walz, B., Baumann, O., 2003. Distribution and serotonin-induced activation of vacuolar-type H⁺-ATPase in the salivary glands of the blowfly *Calliphora vicina*. *J. Exp. Biol.* 206, 1867–1876.