Title
Microarray analysis for identification of Plasmodium-refractoriness candidate genes in mosquitoes

Journal
Genome, 47(6)

ISSN
0831-2796

Authors
Chen, Haifeng
Wang, Jianxin
Liang, Ping
et al.

Publication Date
2004-12-01

DOI
10.1139/g04-056

Microarray analysis for identification of *Plasmodium*-refractoriness candidate genes in mosquitoes

Haifeng Chen, Jianxin Wang, Ping Liang, Monica Karsay-Klein, Anthony A. James, Daniel Brazeau, and Guiyun Yan

Abstract: The identification and cloning of genes conferring mosquito refractoriness to the malaria parasite is critical for understanding malaria transmission mechanisms and holds great promise for developing novel approaches to malaria control. The mosquito midgut is the first major site of interaction between the parasite and the mosquito. Failure of the parasite to negotiate this environment can be a barrier for development and is likely the main cause of mosquito refractoriness. This paper reports a study on *Aedes aegypti* midgut expressed sequence tag (EST) identification and the determination of genes differentially expressed in mosquito populations susceptible and refractory to the avian malaria parasite *Plasmodium gallinaceum*. We sequenced a total of 1200 cDNA clones and obtained 1183 high-quality mosquito midgut ESTs that were computationally collapsed into 105 contigs and 251 singlets. All 1200 midgut cDNA clones, together with an additional 102 genetically or physically mapped *Ae. aegypti* clones, were spotted on single arrays with 12 replicates. Of those interrogated microarray elements, 28 (2.3%) were differentially expressed between the susceptible and refractory mosquito populations. Twenty-seven elements showed at least a two-fold increase in expression in the susceptible population level relative to the refractory population and one clone showed reduced expression. Sequence analysis of these differentially expressed genes revealed that 10 showed no significant similarity to any known genes, 6 clones had matches with unannotated genes of *Anopheles gambiae*, and 12 clones exhibited significant similarity to known genes. Real-time quantitative RT-PCR of selected clones confirmed the mRNA expression profiles from the microarray analysis.

Key words: microarray, vector competence, real-time PCR, EST.

Résumé : L'identification et le clonage de gènes rendant les moustiques réfractaires à l'agent du paludisme constituent des étapes critiques dans la compréhension des mécanismes de transmission du paludisme et offrent de grandes promesses en vue du développement de nouvelles approches de lutte contre cette maladie. L'estomac du moustique constitue le premier site majeur d'interaction entre le parasite et le moustique. L'incapacité du parasite à s'adapter à cet environnement peut constituer une barrière à son développement et représente vraisemblablement la principale cause de la résistance chez le moustique. Dans ce travail, les auteurs ont identifié des EST (« expressed sequence tags ») exprimés dans l'estomac de l'*Aedes aegypti* et déterminé ceux dont l'expression différait chez des populations résistantes ou non à l'infection par le parasite du paludisme aviaire, le *Plasmodium gallinaceum*. Au total, 1200 clones d'ADNc provenant de l'estomac des moustiques ont été séquencés et 1183 EST de grande qualité ont été obtenus. L'analyse informatique de ces séquences a permis de former 105 contigs et 251 singlets. Les 1200 clones d'ADNc, ainsi qu'une collection de 102 clones de l'*Ae. aegypti* dont la cartographie génétique ou physique avait été faite, ont été déposés sur lame en 12 répétitions. Parmi les éléments présents sur la puce à ADN, 28 (2,3 %) montraient une expression différentielle chez les populations résistantes ou sensibles. Vingt-sept éléments montraient un accroissement de l'expression d'au moins deux fois chez la population sensible par rapport à la population résistante tandis qu'un clone montrait une expression réduite. Une analyse des séquences de ces gènes à expression différentielle a révélé que 10 d'entre eux ne présentaient aucune homologie significative avec des gènes connus, 6 clones s'avéraient semblables à des gènes non-annotés chez l'*Anopheles gambiae* et 12 clones montraient une homologie significative à des gènes connus. Une

Received 5 January 2004. Accepted 5 May 2004. Published on the NRC Research Press Web site at http://genome.nrc.ca on 24 November 2004.

Corresponding Editor: P.B. Moens.

H. Chen, M. Karsay-Klein, and G. Yan.[1] Department of Biological Sciences, State University of New York, Buffalo, NY 14260, USA.
J. Wang and P. Liang. Department of Cancer Genetics, Roswell Park Cancer Institute, Buffalo, NY 14263, USA.
A.A. James. Departments of Microbiology & Molecular Genetics and Molecular Biology and Biochemistry, University of California, Irvine, CA 92697, USA.
D. Brazeau. Department of Pharmaceutical Sciences, State University of New York, Buffalo, NY 14260, USA.

[1]Corresponding author (e-mail: gyan@buffalo.edu).

 doi: 10.1139/G04-056

analyse RT-PCR quantitative en temps réel effectuée sur certains clones a confirmé les profils d'expression révélés par l'analyse des puces à ADN.

Mots clés : puce à ADN, compétence face au vecteur, PCR en temps réel, EST.

[Traduit par la Rédaction]

Introduction

Transmission of malaria parasites to human hosts depends on the availability of competent mosquito vectors. *Plasmodium* parasite development in the vector begins when the mosquito ingests gametocytes present in the blood of an infected vertebrate host. Sexual fertilization occurs in the mosquito midgut and the resulting zygotes transform into motile ookinetes. Approximately 18–24 h after the bloodmeal, ookinetes begin to invade the epithelial cell layer of the mosquito midgut, and transform subsequently to oocysts. Oocyst maturation leads to the release of sporozoites into the hemolymph. When sporozoites invade the salivary glands, the mosquito becomes infectious. Failure to complete any of these steps inhibits parasite development and therefore prevents transmission to the vertebrate host. Identification and cloning of the parasite-inhibiting genes in the mosquito vector may lead to novel malaria control methods. For example, identification of genes involved in mosquito refractoriness may help in the design of drugs to block parasite transmission. In addition, sequence information of parasite-inhibiting genes may be used to develop molecular markers for diagnosis of mosquito susceptibility in nature and to assess malaria risk. Malaria transmission may also be affected by genetic disruption of mosquito vector competence through the release of genetically modified mosquitoes that do not transmit malaria parasites (James 2000).

At least four parasite-inhibiting mechanisms have been recognized in mosquitoes. These include parasite failure to exflagellate in the midgut of mosquitoes (Nijhout 1979), melanotic encapsulation of ookinetes by mosquito midgut cells (Collins et al. 1986; Paskewitz et al. 1988; Zheng et al. 1997), apparent physiological incompatibility between the mosquito and the parasite, termed "refractoriness" (Kilama and Craig 1969; Thathy et al. 1994; Severson et al. 1995), and a salivary-gland barrier that prevents sporozoite penetration or reduces survivorship of sporozoites (Rosenberg 1985). We focus here on the refractoriness phenotype and characterize the expression patterns of midgut-specific genes in *Aedes aegypti* populations susceptible and highly refractory to the avian malaria parasite, *Plasmodium gallinaceum*. In the susceptible population, *Plasmodium* oocysts form and mature in the midgut. Sporozoites then migrate to the salivary glands. However, oocysts do not normally develop in refractory mosquitoes, despite a sufficient number of healthy ookinetes in the midgut (Thathy et al. 1994, Kaplan et al. 2001). Refractoriness is a commonly observed phenomenon in natural mosquito populations. Kilama (1973) studied susceptibility of seven African *Ae. aegypti* populations to *P. gallinaceum*, and found that 10%–30% of the mosquitoes did not have oocysts. Toure et al. (1998) infected *Anopheles gambiae* F_1 populations from Mali with natural *Plasmodium falciparum* through direct feeding and found that an average of 40%–50% of the fully bloodfed mosquitoes were not infected with oocysts, though ookinetes form in all mosquitoes. A significant proportion of *An. gambiae* isofemale families screened for susceptibility to the *P. falciparum* parasite suggest that the frequency of refractory alleles is apparently high (Niare et al. 2002).

DNA microarrays have recently been used for high-throughput monitoring of gene expression (DeRisi et al. 1996; Carulli et al. 1998; Moch et al. 1999; Whitney et al. 1999; Ono et al. 2000; Watts et al. 2001). Dimopoulos et al. (2002) examined gene expression responses of *An. gambiae* to microbial and malarial infections using cDNA microarrays constructed from an expressed sequence tag (EST) clone collection. The present study focuses on mosquito innate immune response to malaria parasite infection, rather than on the parasite-induced immune responses. We used mosquito midgut-specific cDNAs to construct microarrays because we have previously determined that the mosquito midgut is a major barrier of ookinete penetration and likely the main cause of mosquito refractoriness (Kaplan et al. 2001). Genetic mapping experiments have demonstrated that at least two loci are responsible for mosquito refractoriness to malaria parasites (Severson et al. 1995).

The aims of this study were to characterize mosquito midgut ESTs and identify midgut genes differentially expressed between mosquitoes highly susceptible and refractory to malaria parasites through genome-wide gene expression analysis. Using cDNA microarrays to target differentially expressed genes between refractory and susceptible populations may allow for rapid and efficient identification of candidate genes conferring mosquito refractoriness to malaria parasites and determination of biochemical pathways involved in mosquito refractoriness.

Materials and methods

Mosquito strains and maintenance

Two *Ae. aegypti* mosquito strains (Moyo-R, Moyo-S) were used for this study. The two strains were selected from a common laboratory stock population, Moyo. The Moyo-R strain was selected for high refractoriness to *Plasmodium gallinaceum* by four generations of recurrent sib mating with a founder density of a single pair of mosquitoes each (Thathy et al. 1994), and has been maintained through mass mating for approximately 10 years. When challenged with a fully infectious bloodmeal, >85% of the Moyo-R mosquitoes have no *Plasmodium* oocysts, and only 1–10 oocysts were observed among those infected (Thathy et al. 1994; Yan et al. 1997). However, >90% of Moyo-S animals become infected, with intensities of infection exceeding 40 oocysts.

Unless otherwise stated, mosquitoes were reared as previously described (Yan et al. 1997). Because gene expression may be sensitive to environmental parameters, the two populations were reared under identical conditions. Thus, the dif-

ferent gene expression patterns between Moyo-R and Moyo-S populations are likely due to their genetic divergence or the interaction between genotype and environment, but not because of the environment itself. Three to four days after emergence, non-bloodfed female mosquitoes were collected for RNA extraction.

Mosquito midgut cDNA library preparation

Total RNA was extracted from 300 *Ae. aegypti* RED strain midguts (B. Beerntsen, D. Fidock and A.A. James, unpublished), half from non-bloodfed mosquitoes and half from those aged 10–12 days after a bloodmeal, using Trizol reagent (Invitrogen, Carlsbad, Calif.). Poly(A)+ RNA (1.79 μg) was selected for library construction using Oligotex (Qiagen, Valencia, Calif.). cDNA was directionally cloned with 5′ *Eco*RI and 3′ *Xho*I linkers in the λ Zap express vector (Stratagene, La Jolla, Calif.). The titer of the amplified library was 2.3×10^9 plaque forming units (PFU)/mL on *Escherichia coli* XL1-Blue MRF. The average insert size of the library was 1320 bp. The pBK-CMV phagemids were mass excised in vivo from the λ vector using ExAssist helper phage and plated using *E. coli* XLOLR.

Sequencing and analysis of ESTs

Minipreps for a total of 1200 randomly selected midgut cDNA colonies were prepared using the High Pure Plasmid Isolation Kit (Roche, Indianapolis, Ind.). The cDNAs were sequenced from the 5′ end using the BigDye™ terminator version 3.0 cycle sequencing kit (PE Applied Biosystems, Foster City, Calif.). The vector sequences were removed and low-quality sequences were trimmed. All sequences were compared against each other to determine sequence redundancy using a locally installed CAP3 program (Huang and Madan 1999). ESTs that share 95% or greater identity over 100 bp were grouped into one cluster. All EST clusters were further analyzed using BLASTX against the non-redundant SWISSPROT database (Altschul et al. 1997). DNA sequences that show marginal similarity to a known gene were considered as unknown genes. All EST sequences were deposited to GenBank dbEST.

Preparation of cDNA microarrays

cDNA inserts for printing microarray slides were amplified using the universal T3 and T7 primers. For each cDNA clone, about 10 ng of plasmid DNA was applied in a 75-μL PCR containing 10 mM Tris–HCl (pH 8.3), 50 mM KCl, 2.5 mM $MgCl_2$, 0.2 μM T3 and T7 primers, 0.4 mM dNTPs, and 1.2 U *Taq* DNA polymerase (MBI Fermentas, Hanover, Md.). Cycling conditions in a DYAD Thermal Cycler (MJ Research, Waltham, Mass.) were 35 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 90 s. Amplification products were precipitated with 95% *v*/*v* ethanol, washed in 70% v/*v* ethanol twice, and resuspended in 20% dimethly sulfoxide (DMSO). Amplification products were run on 1% *w*/*v* agarose gels to estimate the yield from the reactions. Before mircoarray printing, the concentrations of all products were examined by spectrophotometer and each was diluted accordingly to ensure similar concentration for all clones. Each of the 1200 ESTs was spotted in a single array with 12 replicates using an Affymetrix 417 Arrayer. In addition to the midgut ESTs, we also printed 102 *Ae. aegypti* clones previously mapped by Severson et al. (1993, 2002), but the function and tissue specificity of expression are unknown for most of these mapped clones. *Aedes aegypti IMUC* (Morlais and Severson 2001), ribosomal protein S17, cot-1, and double-distilled water were used as negative controls.

Probe preparation and microarray hybridization

Total RNA was indirectly labeled using Atlas Powerscript labeling kits (BD Biosciences, Palo Alto, Calif.). Briefly, 2.5 μg of total RNAs extracted from Moyo-R and Moyo-S populations were labeled by incorporating either Cy3 or Cy5 (Amersham Pharmacia, Piscataway, N.J.) during random-primed reverse transcription. cDNA was synthesized at 42 °C for 1 h in 20 μL containing 4 μL 5× first-strand buffer, 2 μL 10× dNTP Mix, 2 μL DTT, 1 μL reverse transcriptase, and 2 μL random primer (Invitrogen, Carlsbad, Calif.). Ten microlitres of the resulting cDNA were mixed with 10 μL DMSO–Cy3 or –Cy5 reactive dye mixture in the dark at room temperature for 1 h. Labeled cDNA was purified using the Qiagenquick PCR purification kit (Qiagen, Valencia, Calif.). The labeled probe was eluted in 60 μL of elution buffer. Before hybridization, probe aliquots were denatured at 95 °C for 5 min and cooled on ice for 1 min, then spun at 13 000 rpm for 1 min. Arrays were hybridized overnight at 47 °C in a hybridization chamber. After hybridization, slides were washed on an orbital shaker as follows: 1× SSC, 0.2× SDS (42 °C) for 4 min; 0.1× SSC, 0.2 SDS for 4 min; 0.1× SSC for 2 min; and 0.1× SSC for 2 min. After washing, the slides were rinsed with 95% v/*v* ethanol and dried by spinning for 5 min at 500 rpm, then incubated for 2 min at 80 °C in an oven. The slides were scanned with the Affymetrix 428 Scanner. All of the above procedures were repeated by swapping the dye for each probe.

Microarray data analysis

The hybridized slides were scanned using an Affymetrix 428 Scanner to generate high-resolution images for both Cy3 and Cy5 channels. Image analysis was performed on the raw image files using ImaGene (version 4.1) (BioDiscovery; Los Angeles, Calif.). Each image and its background region were segmented using a proprietary optimized segmentation algorithm that excluded pixels that were not representative of the rest of the pixels in that region. The background corrected signal for each cDNA was the mean signal (of all the pixels in the region) minus the mean local background. Signals that were not significantly above background or had a poor coefficient of variance were excluded. For each signal a ratio was calculated from the background subtracted mean signal of the two channels. The ratios were then normalized on a log scale across the entire slide. For each clone on the slide, the expression ratio was the mean on the $\log_2$ scale of all of its replicates. The results from the two slides that made up the dye flip were averaged on the $\log_2$ scale and this became the final expression ratio of that clone. Because each clone was represented 12 times in one array, the average expression ratio of each was calculated. The expression ratio presented in the following section is the mean of the two arrays. The *t* test was used to determine whether the average log2-transformed expression ratio was significantly different from 0.

Real-time RT-PCR confirmation of microarray data

To validate selected data from the microarray experiments, real-time quantitative RT-PCR was performed for genes identified as differentially expressed from the above microarray analysis. Total RNA was extracted from 50 Moyo-R and Moyo-S mosquitoes, respectively. Samples, each containing 1 μg DNA-free RNA, were reverse transcribed to cDNA in 20-μL reaction volumes using SuperScript™ II RNase H- Reverse transcriptase with d(T) primer (GIBCO-BRL, Gaithersburg, Md.). Real-time PCR assays were conducted in triplicate using 5 μL of cDNA (1:10 dilution) as template with 4 μL of 5 μM primers and the SYBR Green Master Mix on an Mx4000 Multiplex Quantitative PCR System (Stratagene, La Jolla, Calif.) according to the instructions of the manufacturer. The PCR program was as follows: 95 °C for 10 min, followed by 40 cycles of 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s. The melting curve was from 55 to 95 °C. Amplification of mosquito 16S ribosomal RNA was used as an internal control for cDNA template quantity. The primers used for these experiments were designed by Primer 3.0 (http://www.broad.mit.edu/ cgi-bin/primer/primer3_www.cgi; Whitehead Institute, MIT Center for Genome Research, Boston, Mass). Real-time PCR data was evaluated for quantitative relative differences between the two mosquito populations. Correlation coefficients between expression data detected by the microarray method and by the quantitative real-time RT-PCR method were calculated.

Results

Aedes aegypti midgut EST sequencing

A total of 1200 randomly selected midgut cDNA clones were sequenced from the 5′ end, generating 1183 high-quality sequences. The sizes of the ESTs ranged from 68 to 2510 bp with an average of 799 bp. These ESTs correspond to 356 clone clusters (251 singlets and 105 contigs; Table 1), each potentially representing an individual gene. Thus, the overall redundancy of the library is about 70%. The most abundant ESTs were *Ae. aegypti* chymotrypsin-like proteins, trypsin-like protein precursors, serine proteases, and ribosomal proteins (Table 2). Three hundred clone clusters out of 356 (84.3%) showed significant BLASTX sequence identity (E value $< 1 \times 10^{-4}$) to proteins in the nonredundant SWISSPROT database, and they were grouped into eight distinct functional groups according to their putative functions (Fig. 1). In particular, 41 clusters showed significant similarity to previously identified or predicted *An. gambiae* genes and (or) proteins; 56 clusters had no significant match to known genes in the public database, and thus can be considered novel *Ae. aegypti* genes. All EST sequences were submitted to dbEST–GenBank under accession numbers CB250704–CB251886.

Genes differentially expressed between mosquito populations highly refractory and susceptible to *P. gallinaceum*

Microarrays were used to examine differences in gene expression between *P. gallinaceum* refractory (Moyo-R) and susceptible mosquito populations (Moyo-S). Total RNA from these two samples was labeled by reverse transcriptase incorporation of Cy3 and Cy5, respectively. In a second set of labeling reactions, the fluorescent dyes were swapped for Moyo-S and Moyo-R. To assess assay reproducibility, we compared the expression intensity similarity between the two experiments using those spots with a signal intensity more than two-fold greater than the background. Data from a total of 316 clones (~26.3% of the total spots) were included in the analysis. The correlation coefficient (r) of Cy3:Cy5 for corresponding clones between the two independent experiments was remarkably similar ($r = 0.939$, $P = 7.2 \times 10^{-5}$; Fig. 2), indicating excellent reproducibility of clone expression. Comparative expression analysis of susceptible versus refractory mosquitoes in the two assays revealed 28 arrayed elements with either enhanced or reduced gene expression by at least two-fold (Table 3). The t tests found the expression level differed at $P < 0.05$ level for one gene (clone 58), $P < 0.001$ for three genes (GS1, ACHE, clone number 37), and $P < 0.0001$ for the remaining 24 clones (Table 3). The differentially expressed genes included 24 clones from the present midgut cDNA library, and four clones that were applied as genetic markers toward the identification of quantitative trait loci (QTL) influencing *Ae. aegypti* competence to the malaria parasite, *P. gallinaceum* (*MalI*, *GS1*, *AchE* and LF338; Severson et al. 1995, 2002). Of the 28 putative differentially expressed genes, 27 showed higher expression in the susceptible mosquito population; only one (*MalI*) displayed a two-fold expression reduction in the refractory mosquito population. The expression ratios of *Ae. aegypti IMUC* and ribosomal protein S17 genes between Moyo-S and Moyo-R were 1.26 and 1.14 fold, respectively; a result consistent with the finding of Morlais and Severson (2001).

To examine the similarity of these ESTs to known genes, cDNA sequences corresponding to each of the 28 array elements were subjected to a BLASTX program search against the non-redundant protein sequences, as well as TBLASTX to all available nucleotide sequences. Fourteen cDNA clones showed significant similarity to known genes (Table 3). Seven of them (374, 609, 542, 564, 847, 47, 210) exhibited significant sequence identity to *An. gambiae* genes. We also found that 11 differentially expressed elements had no matches in the public database entries. Of these 11 unknown genes, 3 (13, 14, 15) were assembled into a contig and the others were singlets. Clones 502 and 966 were noted for significant similarity to trypsin-like proteins. Respectively, they displayed 3.64 and 2.84-fold changes compared with the refractory population. Clone 307 encoded a putative carboxypeptide A. Previous studies indicated that *Ae. aegypti* carboxypeptidase A mRNA is expressed by posterior midgut epithelial cells and accumulates at high levels approximately 16–24 h after ingestion of a blood meal. The induction of this gene is similar to that of *Ae. aegypti* late trypsin (Edwards et al. 2000). Clone LF338 was related to the 65A cuticle protein gene of *Drosophila melanogaster.* Clone 464 displayed high similarity to heat-shock protein 90 from *D. melanogaster*, which is best known for its association with a variety of signal transduction systems (Felts et al. 2000). Clone GS1 encoded glutamine synthase, and clone 609 was associated with the SMC1 protein of *Drosophila*. Other clones, 27, 358, and 368 corresponded to animotransferase, a Lian-Ala retrotransposon protein and inosine triphosphate pyrophosphatse, respectively.

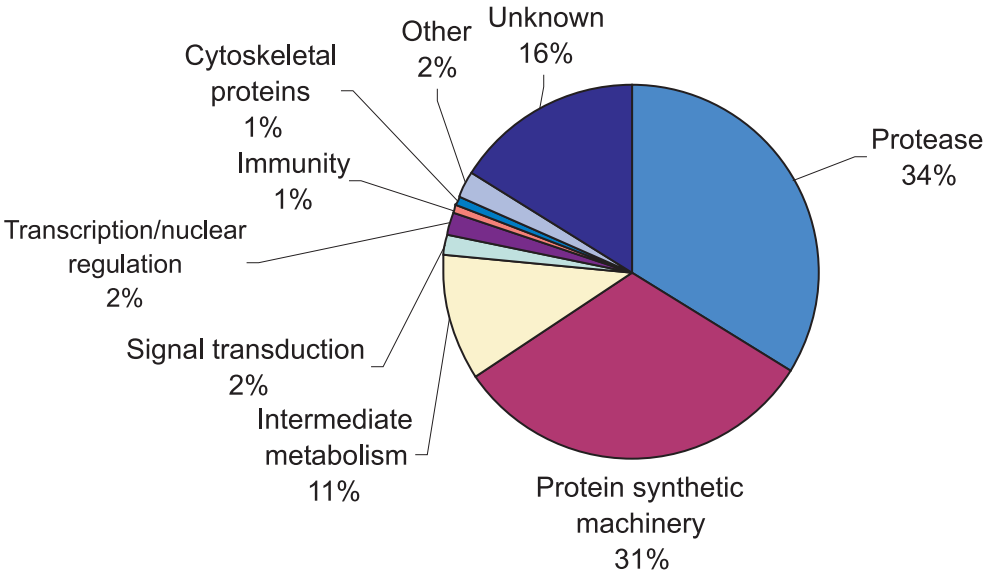
Table 1. Characterization of *Aedes aegypti* midgut ESTs.

Analyzed item	Total
Total sequence reads	1200
No. of high quality sequences	1183
Average high quality length (bp)	799
Singlets	251
Contigs	105
Total non-overlapping clone clusters	356
Homologous clone clusters	300
No. of genes homologous to all known and predicted *Anopheles gambiae* genes and (or) proteins	41
dbEST/GenBank accession Nos.	CB250704–CB251886

Table 2. Summary of the most abundant *Aedes aegypti* midgut ESTs.

Rank	No. of ESTs	Description of homologous subject sequence
1	186	*Aedes aegypti* chymotrypsin-like protein
2	56	*Aedes aegypti* early trypsin-like protein precursor
3	27	*Stomoxys calcitrans* serine protease Ssp3
4	14	Goldfish mitochondrion hypothetical 18K protein
5	13	*Drosophila pseudoobscura* ribosomal protein S2
6	11	*Mus musculus* ribosomal protein L17
7	10	*Spodoptera frugiperda* ribosomal protein S3A
8	8	*Anopheles gambiae* agCP1749
	8	Unknown
9	7	*Xenopus laevis* ribosomal protein S3
	7	*Chironomus tentans* gene for yeast ribosomal YL10 protein homologue
	7	*Anopheles gambiae* agCP1641
	7	*Anopheles gambiae* agCP14063
10	6	Unknown
	6	*Anopheles gambiae* agCP14553
	6	*Anopheles gambiae* agCP12861

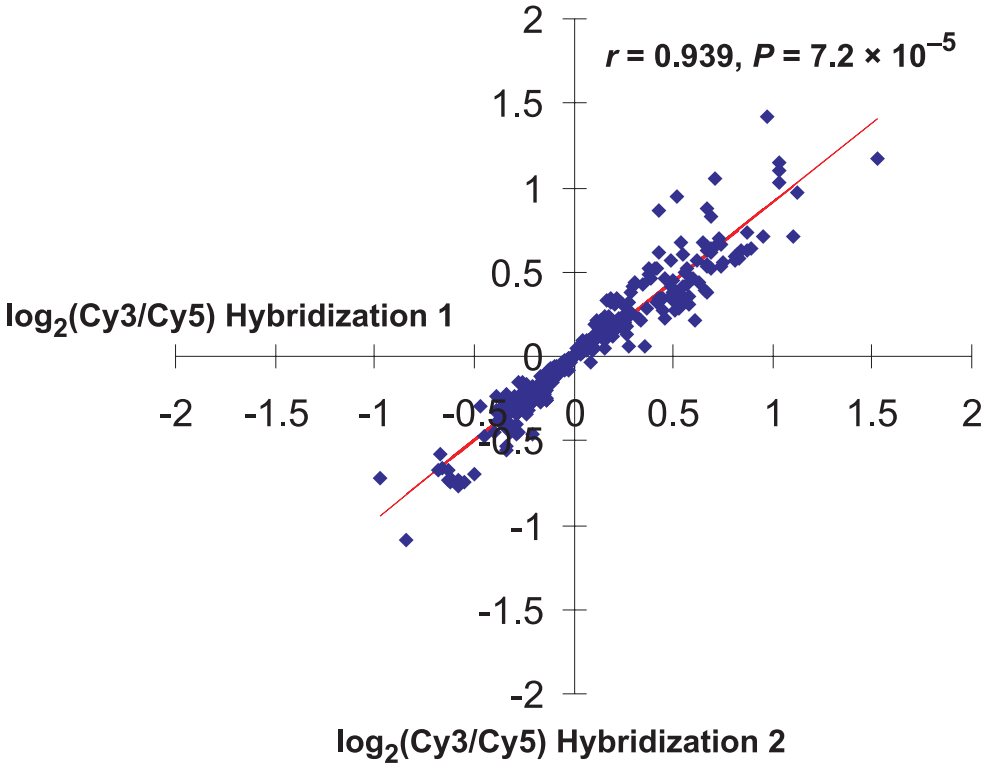
Fig. 1. Distribution of *Aedes aegypti* midgut EST clusters based on putative functions.



Concordance between microarray and quantitative real-time RT-PCR data

To confirm the data from cDNA arrays and to further characterize the mRNA expression profiles, 16 selected genes were analyzed using real-time quantitative RT-PCR. In this study, total RNAs were extracted from 5-day old adult mosquitoes from the Moyo-R and Moyo-S populations. Extracted RNAs were treated with RNase-free DNase to remove genomic DNA contamination prior to reverse transcription. Specific primers were designed according to

Fig. 2. Scatter plot of the Cy3:Cy5 fluorescence intensity ratio for corresponding signals between the two hybridization experiments using *Plasmodium* refractory and susceptible *Aedes aegypti* populations to show the reproducibility of microarray assays. The expression ratio data were log transformed. Solid line represents a perfect correlation (r = 1.0).



the sequences of identified genes. For both Moyo-R and Moyo-S mosquito populations (16 comparisons), there was a high degree of concordance (r = 0.896, P < 1.1E-06) between data from microarray and real-time RT-PCR (Fig. 3).

Discussion

A total of 1183 high-quality *Aedes aegypti* midgut EST sequences were obtained. Those ESTs were collapsed into 105 contigs and 251 singlets, each of which represented putatively different transcripts. Among the contig sequences that allowed interpretation of coding potential, we identified classes of abundant transcripts and found that chymotrypsin and trypsin-like proteins were the most numerous, accounting for 186 and 56 overlapping sequences, respectively. Their frequencies were much higher than the average abundance of inserts in the cDNA library. This result is not surprising because the ESTs were generated from a cDNA library made from mosquito midgut 24 h to 12 days after blood feeding. Theoretically, ESTs generated from a cDNA library should represent the entire set of expressed genes in the tissue from which the library was constructed. However, because the expression patterns of different genes in a given tissue yield mRNAs in different abundances, it is difficult to capture rare mRNAs from primary cDNA libraries. This problem also leads to redundant sequences of clones representing the same highly expressed genes, thereby affecting the efficiency of the ESTs approach and increasing the cost of novel gene discovery (Bonaldo et al. 1996).

The availability of large numbers of ESTs provides a rich source of mosquito cDNA clones for genome-wide investigation using microarray analysis. Using the microarrays containing 1200 randomly selected midgut ESTs, we have identified 27 midgut ESTs differentially expressed between the two *Ae. aegypti* mosquito populations that are highly refractory and highly susceptible to the *P. gallinaceum* parasite. Dimopoulos et al. (2002) examined the gene expression profile of *An. gambiae* in responses to injury, bacterial challenge, and malarial infection. The differentially expressed genes identified in the present study showed no overlap with theirs. Thus, pathogen infection or injury likely induces the immediate expression of the same classes of genes in both refractory and susceptible lines, and those are distinct from the ones specific to either of the two selected lines. The two mosquito populations (Moyo-R and Moyo-S) were selected from the same stock population carried through four generations of selective inbreeding (Thathy et al. 1994). Therefore, these differentially expressed genes should not be associated with the genetic background of the mosquito populations. In addition, we used two mosquito strains of equivalent age, reared under the same conditions to prepare hybridization probes for the microarray analyses, and thus we minimized the effects of environmental conditions on gene expression differences between the two strains.

Of all the differentially expressed elements, we found that several genes related to bloodmeal digestion (e.g., trypsin and carboxypeptidase A) exhibited increased mRNA abundance in susceptible mosquitoes. We observed previously that susceptible mosquitoes exhibited higher enzymatic ac-

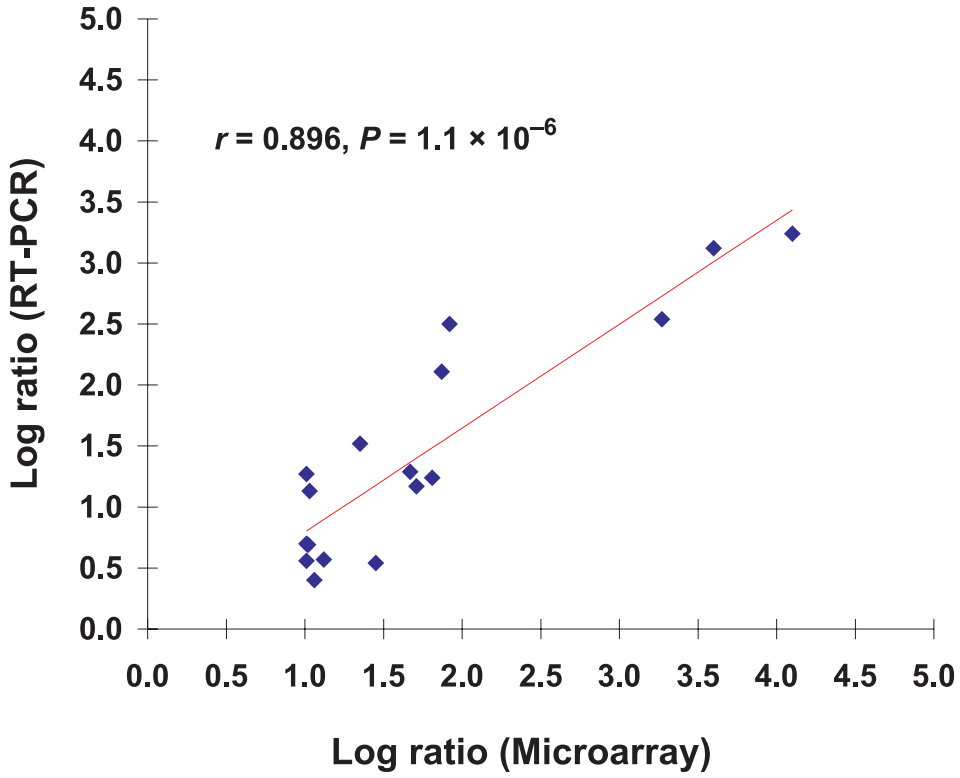
Table 3. cDNAs that are differentially expressed between *Aedes aegypti* mosquitoes highly refractory and susceptible to *Plasmodium gallinaceum* parasite and best BLASTX identity matches.

Clone ID	Expression difference (folds)	*P**	Best match (GenBank acc. No.)	Putative function	Organism	Amino acid identity†	BLASTX *E* value
Mal 1	0.41	<0.0001	M30442	Maltase-like enzyme	*Aedes aegypti*		
569	2	<0.0001	None				
374	2.01	<0.0001	AAAB01008960	Short chain dehydrogenase/reductase	*Anopheles gambiae*	165/235 (70)	2.0×10^{-86}
GS1	2.01	<0.001	AF004351	Glutamine synthase	*Aedes aegypti*		
609	2.01	<0.0001	CAD59403	SMC1 protein	*Anopheles gambiae*	156/212 (73)	3.0×10^{-82}
58	2.02	<0.05	None				
14	2.04	<0.0001	None				
542	2.07	<0.0001	EAA09736	agCP4376	*Anopheles gambiae*	44/133 (33)	1.0×10^{-9}
15	2.1	<0.0001	None				
464	2.11	<0.0001	AAD29307	Hsp90 protein	*Drosophila melanogaster*	169/251 (67)	5.0×10^{-92}
AchE	2.12	<0.001	AAB35001	acetylcholinesterase	*Aedes aegypti*		
13	2.12	<0.0001	None				
564	2.16	<0.0001	EAA05224	Fibrinogen	*Anopheles gambiae*	71/165 (43)	4.0×10^{-36}
847	2.16	<0.0001	EAA00339	Glycosyl transferase	*Anopheles gambiae*	101/162 (62)	3.0×10^{-58}
27	2.17	<0.0001	None				0.002
LF338	2.29	<0.0001	U84745	Cuticle protein LCP65Ac	*Drosophila melanogaster*	40/58 (68)	5.0×10^{-19}
47	2.38	<0.0001	EAA06892	agCP7421	*Anopheles gambiae*	54/62 (87)	1.0×10^{-24}
307	2.41	<0.001	XP_317083	Carboxypeptidase A	*Anopheles gambiae*	64/116(55)	3.0×10^{-15}
89	2.55	<0.0001	None				
210	2.6	<0.0001	EAA14591	agCP8048	*Anopheles gambiae*	106/159 (66)	3.0×10^{-53}
358	2.73	<0.0001	AAB65093	Lian-Aa1 retrotransposon protein	*Aedes aegypti*	100/320 (31)	1.0×10^{-34}
367	2.83	<0.0001	None				
966	2.84	<0.0001	TRWV5Y	Trypsin	*Aedes aegypti*	185/213 (86)	1.0×10^{-105}
986	2.86	<0.0001	None				
368	3.19	<0.0001	AAK21848	Inosine triphosphate pyrophosphatase	*Homo sapiens*	83/112 (74)	6.0×10^{-42}
247	3.64	<0.0001	None				
502	3.64	<0.0001	T13598	Trypsin homolog	*Drosophila*	80/213 (37)	2.0×10^{-25}
172	4.17	<0.0001	None				

*P value refers to t test significance for the null hypothesis that there was no difference in the expression level between *Plasmodium* refractory and susceptible *Aedes aegypti* populations for the clone.

†Percent identity in parentheses.

Fig. 3. Correlation between log-transformed expression ratios obtained from microarray and real-time quantitative RT-PCR experiments. A total of 16 differentially expressed genes between *Plasmodium* refractory and susceptible *Aedes aegypti* populations identified by microarray analysis were used in the correlation analysis. Solid line represents a perfect correlation (r = 1.0).



tivity for trypsin and aminopeptidase than refractory populations in blood meal digestion (Kaplan et al. 2001). Midgut digestive proteases are vital to the success of *Plasmodium* development in mosquitoes. For example, trypsin is essential for activating chitinase, an enzyme required for ookinete penetration of the peritrophic membrane (Richards and Richards 1977; Shahabuddin et al. 1993). Inhibition of trypsin-like activity reduces *Ae. aegypti* susceptibility (Shahabuddin et al. 1995). On the other hand, trypsin can damage the early developing ookinetes (Gass 1977; Gass and Yeates 1979), but fully developed *Plasmodium* ookinetes are relatively protected from midgut proteases (Yeates and Steiger 1981). Interestingly, trypsin in *An. gambiae* was reported to be down-regulated by *Plasmodium*-infected blood 18 h after blood feeding (Bonnet et al. 2001). The increased level of midgut proteases in susceptible populations may enhance ookinete penetration of the peritrophic membrane.

None of the genes identified in this report share sequence similarity with the leucine-rich repeat protein or the C-type lectins that act as an antagonist or protective agonists, respectively, to *P. berghei* infection in *An. gambiae* (Osta et al. 2004). This may reflect the specificity of responses of each of the mosquito species to the respective parasites and the nature of the resistance or refractory phenotype. *Anopheles gambiae* resists infection of *P. berghei* by encapsulating the developed oocysts in melanin, whereas *Ae. aegypti* refractory to *P. gallinaceum* show few or no developed oocysts.

We observed one cDNA element corresponding to the gene, *Maltase-like I* (*M*alI) that was up regulated in the refractory mosquitoes. *MalI* is specifically expressed in the adult salivary glands of both female and male (James et al. 1989). This gene was one of the QTL markers for *P. gallinaceum* susceptibility identified in *An. aegypti* (Severson et al. 1995) and included in this array. Differential expression of *MalI* gene between the susceptible and refractory populations indicates that salivary gland gene expression may also affect sporozoite development success.

In summary, we have characterized *Ae. aegypti* mosquito midgut ESTs and identify 27 midgut cDNAs and 1 salivary gland gene differentially expressed between mosquitoes highly susceptible and refractory to *P. gallinaceum* parasites using microarray analyses. Some of the differentially expressed genes may not be directly related to mosquito refractoriness, however the determination of mosquito refractoriness QTLs (Severson et al. 1995) can be used to screen candidate genes. For example, the candidate genes can be genetically mapped with the available markers (Severson et al. 1993) to determine whether they are in the general chromosomal regions where refractory QTLs reside. This approach, coupled with detailed biochemical analysis, holds promise for yielding useful information on genes involved in the biomedical pathway leading to mosquito refractoriness to malaria parasites.

Acknowledgements

We thank anonymous reviewers for helpful comments on the manuscript. This work was supported by the interdisciplinary research and creative activities fund of the State Uni-

versity of New York, Buffalo, NY, to G.Y. A.A.J. is grateful to the Burroughs–Wellcome Fund.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.L. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**: 3389–3402.

Bonaldo, M.F., Lennon, G., and Soares, M.B. 1996. Normalization and subtraction: two approaches facilitate gene discovery. Genome Res. **6**: 791–806.

Bonnet, S., Prevot, G., Jacques, J.C., Boudin, C., and Bourgouin, C. 2001. Transcription of the malaria vector *Anopheles gambiae* that are differentially regulated in the midgut upon exposure to invasive stages of plasmodium falciparum. Cell Microbiol. **3**: 449–58.

Carulli, J.P., Artinger, M., Swain, P.M., Root, C.D., Chee, L., Tulig, C., Guerin, J., Osborne, M., Stein, G., Lian, J., and Lomedico, P.T. 1998. High throughput analysis of differential gene expression. J. Cell. Biochem. (Suppl.) **30–31**: 286–296.

Collins, F.H., Sakai, R.K., Vernick, K.D., Paskewitz, S., Seeley, D.C., Miller, L.H., Collins, W.E., Campbell, C.C., and Gwadz, R.W. 1986. Genetic selection of a substrate of a *Plasmodium*-refractory strain of malaria vector *Anopheles gambiae*. Science (Washington, D.C.), **234**: 607–610.

DeRisi, J., Penland, L., Brown, P.O., Bittner, M.L., Meltzer, P.S., Ray, M., Chen, Y., Su, Y.A., and Trent, J.M. 1996. Use of a cDNA microarray to analyse gene expression patterns in human cancer. Nat. Genet. **14**: 457–460.

Dimopoulos, G., Christophides, G.K., Meister, S., Schultz, J., White, K.P., Barillas-Mury, C., and Kafatos, F.C. 2002. Genome expression analysis of *Anopheles gambiae*: responses to injury, bacterial challenge, and malarial infection. Proc. Natl. Sci. Acad. USA. **99**: 8814–8819.

Edwards, M.J., Moskalyk, L.A., Donelly-Doman, M., Vlaskova, M., Noriega, F.G., Walker, V.K., and Jacobs-Lorena, M. 2000. Characterization of a carboxypeptidase A gene from mosquito, *Aedes aegypti*. Insect Mol. Bio. **9**: 33–38.

Felts, S.J., Owen, B.A., Nguyen, P., Trepel, J., Donner, D.B., and Toft, D.O. 2000. The hsp90-related protein TRAP1 is a mitochondrial protein with distinct functional properties. J. Biol. Chem. **275**: 3305–3312.

Gass, R.F. 1977. Influences of blood digestion on the development of *Plasmodium gallinaceum* (Brumpt) in the midgut of *Aedes aegypti* (L.). Acta Tropica, **34**: 127–140.

Gass, R.F., and Yeates, R.A. 1979. In vitro damage of cultured ookinetes of *Plasmodium gallinaceum* by digestive proteinases from susceptible *Aedes aegypti*. Acta Tropica, **36**: 243–252.

Huang, X., and Madan, A. 1999. CAP3: A DNA sequence assembly program. Genome Res. **9**: 868–877.

James, A.A. 2000. Control of disease transmission through genetic modification of mosquitoes. *In* Insect transgenesis. *Edited by* A.M. Handler and A.A. James. CRC Press, Boca Raton, Fla. 319–332.

James, A.A., Blackmer, K., and Racioppi, J.V. 1989. A salivary gland-specific, maltase-like gene of the vector mosquito, *Aedes aegypti*. Gene, **75**: 73–83.

Kaplan, R.A., Zwiers, S.H., and Yan, G. 2001. *Plasmodium gallinaceum*: ookinete formation and proteolytic enzyme dynamics in highly refractory *Aedes aegypti* populations. Exp. Parasitol. **98**: 115–122.

Kilama, W.L. 1973. Distribution of a gene for susceptibility to *Plasmodium gallinaceum* in populations of *Aedes aegypti* (L.). J. Parasitol. **59**: 920–924.

Kilama, W.L., and Craig, G.B. 1969. Monofactorial inheritance of susceptibility to *Plasmodium gallinaceum* in *Aedes aegypti*. Ann. Trop. Med. Parasitol. **63**: 419–432.

Moch, H., Schraml, P., Bubendorf, L., Mirlacher, M., Kononen, J., Gasser, T., Mihatsch, M.J., Kallioniemi, O.P., and Sauter, G. 1999. High-throughput tissue microarray analysis to evaluate genes uncovered by cDNA microarray screening in renal cell carcinoma. Am. J. Pathol. **154**: 981–986.

Morlais, I., and Severson, D.W. 2001. Identification of a polymorphic mucin-like gene expressed in the midgut of the mosquito, *Aedes aegypti*, using an integrated bulked segregant and differential display analysis. Genetics, **158**: 1125–1136.

Niare, O., Markianos, K., Volz, J., Odoul, F., Toure, A., et al. 2002. Genetic loci affecting resistance to human malaria parasites in a West African mosquito vector population. Science (Washington, D.C.), **298**: 213–216.

Nijhout, M.M. 1979. *Plasmodium gallinaceum*: Exflagellation stimulated by a mosquito factor. Exp. Parasitol. **48**: 75–80.

Ono, K., Tanaka, T., Tsunoda, T., Kitahara, O., Kihara, C., Okamoto, A., Ochiai, K., Takagi, T., and Nakamura, Y. 2000. Identification by cDNA microarray of genes involved in ovarian carcinogenesis. Cancer Res. **60**: 5007–5011.

Osta, M.A., Christophides, G.K., and Kafatos, F.C. 2004. Effects of mosquito genes on *Plasmodium* development. Science (Washington, D.C.), **203**: 2030–2032.

Paskewitz, S.M., Brown, M.R., Lea, A.O., and Collins, F.H. 1988. Ultrastructure of the encapsulation of *Plasmodium cynomolgi* (B strain) on the midgut of refractory stain of *Anopheles gambiae*. J. Parasitol. **74**: 432–439.

Richards, A.G., and Richards, P.A. 1977. The peritrophic membranes of insects. Annu. Rev. Entomol. **22**: 219–240.

Rosenberg, R. 1985. Inability of *Plasmodium knowlesi* sporozoites to invade *Anopheles freeborni* salivary glands. Am. J. Trop. Med. Hyg. **34**: 687–691.

Severson, D.W., Mori, A., Zhang, Y., and Christensen, B.M. 1993. Linkage map for *Aedes aegypti* using restriction fragment length polymorphisms. J. Hered. **84**: 241–247.

Severson, D.W., Thathy, V., Mori, A., Zhang, Y., and Christensen, B.M. 1995. Restriction fragment length polymorphism mapping of quantitative trait loci for malaria parasite susceptibility in the mosquito *Aedes aegypti*. Genetics, **139**: 1711–1717.

Severson, D.W., Meece, J.K., Lovin, D.D., Saha, G., and Morlais, I. 2002. Linkage map organization of expressed sequence tags and sequence tagged sites in the mosquito, *Aedes aegypti*. Insect Mol. Biol. **11**: 371–378.

Shahabuddin, M., Toyoshima, T., Aikawa, M., and Kaslow, D.C. 1993. Transmission-blocking activity of a chitinase inhibitor and activation of malaria parasite chitinase by mosquito protease. Proc. Natl. Sci. Acad. USA. **90**: 4266–4270.

Shahabuddin, M., Criscio, M., and Kaslow, D.C. 1995. Unique specificity of in vitro inhibition of mosquito midgut trypsin-like activity correlates with in vivo inhibition of malaria parasite infectivity. Exp. Parasitol. **80**: 212–219.

Thathy, V., Severson, D.W., and Christensen, B.M. 1994. Reinterpretation of genetics of susceptibility of *Aedes aegypti* to *Plasmodium gallinaceum*. J. Parasitol. **80**: 705–712.

Toure, Y.T., Doumbo, O., Toure, A., Bagayoko, M., Diallo, M., et al. 1998. Gametocyte infectivity by direct mosquito feeds in an area of seasonal malaria transmission: implications for Bancoumana, Mali as a transmission-blocking vaccine site. Am. J. Trop. Med. Hyg. **59**: 481–486.

Watts, G.S., Futscher, B.W., Isett, R., Gleason-Guzman, M., Kunkel, M.W., and Salmon, S.E. 2001. cDNA microarray analysis of multidrug resistance: doxorubicin selection produces multiple defects in apoptosis signaling pathways. J. Pharmacol. Exp. Ther. **299**: 434–41.

Whitney, L.W., Becker, K.G., Tresser, N.J., Caballero-Ramos, C.I., Munson, P.J., Prabhu, V.V., Trent, J.M., McFarland, H.F., and Biddison, W.E. 1999. Analysis of gene expression in multiple sclerosis lesions using cDNA microarrays. Ann. Neurol. **46**: 425–428.

Yan, G., Severson, D.W., and Christensen, B.M. 1997. Cost and benefits of mosquito refractoriness to malaria parasites: Implication for genetic diversity of mosquitoes and genetic control of malaria. Evolution, **51**: 441–450.

Yeates, R.A., and Steiger, S. 1981. Ultrastructural damage of in vitro cultured ookinetes of *Plasmodium gallinaceum* (Brumpt) by purified proteinases of susceptible *Aedes aegypti* (L.). Z Parasitenkd. **66**: 93–7.

Zheng, L., Cornel, A.J., Wang, R., Erfle, H., Voss, H., Ansorge, W., Kafatos, F.C., and Collins, F.H. 1997. Quantitative trait loci for refractoriness of *Anopheles gambiae* to *Plasmodium cynomolgi* B. Science (Washington, D.C.), **276**: 425–428.