Mol Gen Genomics (2003) 269: 753–764
DOI 10.1007/s00438-003-0882-7

ORIGINAL PAPER

I. Morlais · A. Mori · J. R. Schneider · D. W. Severson

A targeted approach to the identification of candidate genes determining susceptibility to *Plasmodium gallinaceum* in *Aedes aegypti*

Received: 15 January 2003 / Accepted: 6 June 2003 / Published online: 15 July 2003

Abstract The malaria parasite, *Plasmodium* , has evolved an intricate life cycle that includes stages specific to a mosquito vector and to the vertebrate host. The mosquito midgut represents the first barrier *Plasmodium* parasites encounter following their ingestion with a blood meal from an infected vertebrate. Elucidation of the molecular interaction between the parasite and the mosquito could help identify novel approaches to preventing parasite development and subsequent transmission to vertebrates. We have used an integrated Bulked Segregant Analysis-Differential Display (BSA-DD) approach to target genes expressed that are in the midgut and located within two genome regions involved in determining susceptibility to *P. gallinaceum* in the mosquito *Aedes aegypti*. A total of twenty-two genes were identified and characterized, including five genes with no homologues in public sequence databases. Eight of these genes were mapped genetically to intervals on chromosome 2 that contain two quantitative trait loci (QTLs) that determine susceptibility to infection by *P. gallinaceum*. Expression analysis revealed several expression patterns, and ten genes were specifically or preferentially expressed in the midgut of adult females. Real-time PCR quantification of expression with respect to the time of blood meal ingestion and infection status in mosquito strains permissive and refractory for malaria revealed a differential expression pattern for seven genes. These represent candidate genes that may influence the ability of the mosquito vector to support the development of *Plasmodium* parasites. Here we describe their isolation and discuss their putative roles in parasite-mosquito interactions and their use as potential targets in strategies designed to block transmission of malaria.

Communicated by D. Y. Thomas

I. Morlais · A. Mori · J. R. Schneider · D. W. Severson (✉)
Center for Tropical Disease Research and Training,
Department of Biological Sciences,
University of Notre Dame, Notre Dame, IN 46556, USA
E-mail: david.w.severson.1@nd.edu
Tel.: +1-574-6313826

Keywords Bulked segregant analysis · Differential display · Malaria · Quantitative trait loci (QTLs) · Vector competence

Introduction

Malaria, with AIDS and tuberculosis, remains one of the most devastating infectious diseases, affecting 300–500 million people and resulting in up to 2.7 million deaths per year, mainly in sub-Saharan Africa. Control strategies have been severely affected during the last four to five decades by the emergence of resistance to anti-malarial drugs in *Plasmodium* and resistance to insecticides in mosquito vectors. Ongoing research efforts have led to the successful germ line transformation of mosquito vectors (Jasinskiene et al. 1998; Catteruccia et al. 2000; Coates 2000), and promoters suitable for driving the expression of genes that could block parasite development have been identified (Kokoza et al. 2000; Moreira et al. 2000). However, the ability to engineer transgenic mosquitoes that are refractory to malaria parasite development requires the identification of genes capable of interfering with the development of *Plasmodium* in its vector.

Plasmodium parasites go through a complex life cycle in the mosquito (for review, see Ghosh et al. 2000). Parasite development in the mosquito vector begins when the mosquito ingests gametocytes present in the blood while feeding on an infected vertebrate host. In the lumen of the insect's midgut, gametocytes mature into gametes that fuse to produce a zygote. The zygote develops into a motile ookinete, which crosses the peritrophic matrix and the midgut epithelium, and then forms an oocyst beneath the basal lamina. Oocyst maturation leads to the release of sporozoites into the hemolymph. Sporozoites then invade the salivary glands and are injected with the saliva when the mosquito feeds on a new host. A significant reduction in parasite numbers occurs after a mosquito ingests infected blood:

only a few ookinetes are successful in reaching the basal lamina and, although thousands of sporozoites are released into the hemolymph, only a fraction succeeds in actually invading the salivary glands.

At the midgut level, parasites have to face two different barriers: the peritrophic matrix (PM) and the midgut epithelium. The PM forms a semi-permeable chitinous membrane surrounding the food bolus, and is associated with the digestive process, partitioning digested food components between the endo- and ectoperitrophic spaces (Richards and Richards 1977; Tellam et al. 1999). In addition, the PM provides a physical barrier against invasion by bacteria and parasites. To enable them to cross the PM, ookinetes secrete a chitinase that allows the local disruption of the PM during penetration (Huber et al. 1991; Vinetz et al. 2000), and it has been shown that chitinase inhibitors successfully block *Plasmodium* development (Sieber et al. 1991; Shahabuddin et al. 1995; Vinetz et al. 2000). Ookinetes that escape from the PM are then exposed to proteolytic enzymes in the ectoperitrophic space and need to penetrate the midgut epithelium. They bind to the microvillar surface of the epithelium and enter an epithelial cell. The mechanism that underlies the binding specificity is unknown, but adhesion requires an interaction with a carbohydrate ligand (Zieler et al. 1999), and mucin-like proteins that have been identified in the mosquito midgut could be involved in this process (Shen et al. 1999; Morlais and Severson 2001). Most ookinetes are lysed during their migration through the epithelium, but those that survive probably go on to develop into oocysts. Previous studies have shown that the mosquito midgut is capable of an active immune response following invasion by *Plasmodium*, and several defense molecules expressed in response to infection with *Plasmodium* have been described (Richman et al. 1997; Dimopoulos et al. 1998, 2001; Luckhart et al. 1998).

We have focused on the mosquito midgut as it represents the first barrier that parasites need to cross in order to complete their life cycle. Moreover, a laboratory strain of *Aedes aegypti* (Moyo-R) that is refractory to *P. gallinaceum* has previously been selected genetically (Thathy et al. 1994), with refractoriness being associated with failure to form oocysts. In combination with a permissive strain (Red), Moyo-R provides a suitable model for studying the genetic basis underlying refractoriness to *Plasmodium*. Indeed, two quantitative trait loci (QTLs) have been found to be associated with susceptibility to *Plasmodium*. The major one, *pgs1*, is located on chromosome 2, while *pgs2*, resides on chromosome 3 (Severson et al. 1995). More recently, four further QTLs have been identified on chromosome 2 (J. K. Meece and D. W. Severson, unpublished data).

Several approaches have been developed to identify genes expressed following infection of the insect vector by *Plasmodium* (Dimopoulos et al. 1996; Arca et al. 1999; Dessens et al. 1999; Oduol et al. 2000) and the number of putative anti-malarial immune defense molecules that have been identified has increased rapidly. However, none of these approaches considers aspects of the genetic basis for transmission or the fact that the expression of immune activated genes can reflect either a local or a systemic response. We recently described an integrated method, Bulked Segregant Analysis-Differential Display (BSA-DD), that takes these factors into account by targeting expressed genes located in genome areas known to be involved in determining susceptibility (Morlais and Severson 2001). The BSA-DD approach has already been used successfully to characterize a mucin-like protein precursor encoded by a gene (*AeIMUC1*) that maps within the *pgs1* interval and is over-expressed upon infection with *Plasmodium*.

We describe here genes identified and characterized using the BSA-DD method to target two genomic locations containing QTLs influencing susceptibility. Eighteen genes were mapped genetically: eight genes map in the QTL intervals *pgs1* and *pgs5* on chromosome 2. Expression analysis identified several categories of genes, including some that (1) are over-expressed in both susceptible and refractory mosquito strains at the time that ookinetes traverse the midgut, (2) exhibit differential regulation between the refractory and the susceptible strains, (3) exhibit sex-specific expression, or (4) are expressed in response to a blood meal.

Materials and methods

Mosquitoes and *P. gallinaceum* infections

A. aegypti mosquitoes were reared at 26°C and 85% relative humidity on a 12-h photoperiod. Adult mosquitoes were fed with a 10% sucrose solution for maintenance. The *A. aegypti* Red strain, which is susceptible to *P. gallinaceum*, the refractory Moyo-R strain and F_2 progenies obtained by pair-wise mating between Red females and Moyo-R males were used in this study. Infection with *P. gallinaceum* was carried out as previously described (Morlais and Severson 2001).

Dissection of mosquitoes

Midguts to be used for RNA extraction were dissected at 12 h, 24 h, 36 h, 48 h, 72 h and 7 days after blood feeding, and midguts from unfed females were also dissected. For each experiment, 15 midguts were homogenized in 120 μl of TriReagent (GibcoBRL) and stored at −80°C until RNA extraction. Each experiment was replicated three times. Control infected females were dissected 7 days after infection, midguts were stained with 2% mercurochrome and oocysts were counted.

RNA extraction

Total RNA was extracted from dissected female tissues (midguts and carcasses) and intact animals (pupae, larvae and adult males) in TriReagent according to the manufacturer's instructions.

BSA-DD procedure

The Bulked Segregant Analysis-Differential Display (BSA-DD) method was performed as previously described (Morlais and

Severson 2001). We first applied the BSA-DD method to target genes located inside the *pgs1* QTL, which has been mapped on chromosome 2 (Severson et al. 1995). More recently, fine-scale mapping has shown that four more QTLs (*pgs3*– *pgs6*) are located on chromosome 2 and are therefore linked to *pgs1* (J. K. Meece and D. W. Severson, unpublished data). We also applied the BSA-DD technique to the *pgs5* region. RFLP markers used for genetic mapping were: LF181, LF282 and LF338, LF272 for *pgs1* and LF272, LF173 and LF282, LF223 for *pgs5* (in each case, the first two flank the QTL and the others are further away; Fig. 1).

Genetic mapping

Selective genotyping was performed using the cloned DD-PCR products as probes on Southern blots of genomic DNAs obtained from the F_2 females. *Eco* RI polymorphism was first verified on Southern blots of six individuals each from both parental strains. Procedures for Southern blotting and hybridization were performed as described elsewhere (Severson 1997). Southern blots were exposed to phosphorimaging screens and scanned using a Storm 840 phosphorimaging system (Molecular Dynamics). Multipoint linkage analysis was conducted using the MAPMAKER computer package (Lander et al. 1987; Lincoln and Lander 1990). Recombination frequencies were converted into map distances (cM) using the Kosambi function (Kosambi 1944).

SNP genotyping

When no *Eco* RI polymorphism could be detected between the parental strains on Southern blots, or when sequence data showed strain-specific polymorphisms, single nucleotide polymorphism (SNP) genotyping was investigated using a melting curve assay (McSNP; Akey et al. 2001). The assays were monitored with the DASH machine (Hybaid). The same F_2 female population was used for all mapping efforts, as a fraction (1/5) of the genomic DNA from individual mosquitoes was retained before Southern blots were prepared. SNPs were identified by sequence alignments of at least two clones from both Red and Moyo-R strains, and substitutions occurring at a restriction enzyme site were chosen for the SNP assays. Individual genotypes were characterized using the automatic scoring feature of the DASH software by adjusting the matching zones to the parental patterns. Linkage analysis of individual SNPs was performed as described above.

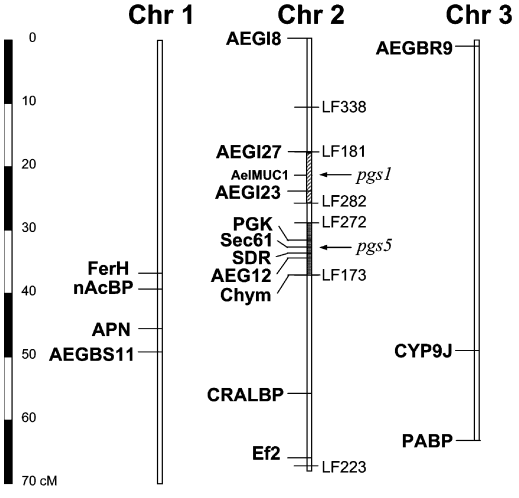


Fig. 1 Locations of genes identified by BSA-DD on the *Aedes aegypti* genetic map. RFLP markers used for the BSA procedure are indicated to the right of chromosome 2. Targeted QTL intervals are indicated by the *arrows*

Virtual Northern analysis

Virtual Northern blots were prepared to assess RNA expression of the DD inserts in midguts from the parental Red and Moyo-R strains. cDNAs were first obtained from 1-μg aliquots of total RNA using the SMART PCR cDNA synthesis kit (Clontech). Virtual blots were performed as previously described using 400 ng of cDNA (Morlais and Severson 2001).

Full-length genes

To obtain full-length genes, an internal gene-specific reverse primer, designed from the DD insert sequence, was used in combination with the SMO primer from the SMART cDNA kit to screen a midgut cDNA library. PCRs were performed using standard conditions with an annealing temperature of 62°C. The PCR products, corresponding to the 5′ends of the respective mRNAs, were cloned into the TopoTA pCRII vector (Invitrogen) and end-sequenced using the ABI Prism 310 genetic analyzer (PerkinElmer). Specific primers were then designed at the 5′ and 3′-ends. The primer set was used on the parental cDNA libraries to obtain the full-length gene for both Red and Moyo-R strains. At least two clones from each strain were sequenced. We also amplified the full-length products from genomic DNA extracted from individual mosquitoes. For instance, where the gDNA product was larger than the cDNA, the PCR products from at least two different mosquitoes for each strain were sequenced, and introns were identified by sequence comparison between cDNAs and gDNAs.

Sequence analysis

Sequence analysis was performed using the GCG Sequence Analysis package (Genetics Computer Group, Madison, Wis.) for sequence assembly and the ExPASy proteonomic tools (http://www.expasy.org/) for general analyses including PSORT II for intracellular sorting signals, and ProfileScan for pattern searches. Sequence homology searches in the non-redundant databases were carried out with the BLAST programs (Altschul et al. 1997). When no homologue was found, sequences were also used to search the EST database. Amino-acid alignments were constructed using ClustalW (Thompson et al. 1994). Trees were constructed with NjPlot from Clustal alignments using the neighbor joining method with bootstrapping.

Expression analysis

Reverse transcription was done using 500-ng aliquots of total RNA in a 20-μl reaction mixture containing 50 mM TRIS-HCl pH 8.3, 3 mM $MgCl_2$, 75 mM KCl, 10 mM DTT, 500 μM of each dNTP, 20 pmol of oligo-$(dT)_{18}$ (Clontech) and 200 U of Superscript II RNaseH- (Gibco-BRL). Reactions were performed at 42°C for 1 h and then heated to 70°C for 15 min to inactivate the enzyme. The cDNAs were resuspended in a final volume of 100 μl, and a 2-μl sample was used as template for PCR. All cDNA templates were subjected to a standard PCR using gene-specific primers for a ribosomal gene (*RpS17*) to check the quality and efficiency of the RT reaction.

Developmental and tissue specific expression

PCRs were performed under our standard conditions (see below), using cDNAs from larvae, pupae, adult males, female carcasses minus midguts or female midguts as template. For each primer set, the cycle number was defined empirically to avoid saturation and ranged from 16 to 22. The *RpS17* gene was used as an internal control. PCR products were electrophoresed on 2% agarose gels, stained with ethidium bromide, and photographed.

Real-time PCR

Real-time PCR was performed using the SYBR Green I PCR kit (PE Applied Biosystems) according to the manufacturer's instructions. The threshold cycle or Ct value represents the cycle at which a significant increase is detected in the intensity of SybrGreen fluoresence; the Ct occurs when the increase in signal is associated with an exponential increase in the amount of PCR product. Relative quantification was normalized to the *RpS17* gene as the internal standard, according to the manufacturer's recommendations. Reactions were conducted in the GeneAmp 7700 sequence detection system (Perkin Elmer). Real-time PCRs were carried out on samples from three independent dissection experiments. Six different time points were investigated, e.g., 12, 24, 36, 48 and 72 h and 7 days after feeding of the two strains with infected or uninfected blood, and quantifications were also performed for midguts from unfed females. Sequences of gene-specific primer sets are given in Table 1.

Statistical methods

Expression levels for each gene were analyzed using a three-way ANOVA (Proc GLM; SAS Institute, Cary, N.C., 2000), and included time, strain and infection status as the independent variables. Multiple comparisons of the least-square means were performed using Tukey's studentized range test.

Results and discussion

BSA-DD fragments

A total of 52 candidate fragments were excised from the differential display gels, cloned and sequenced. Twenty-four fragments were isolated upon comparing bulks targeting the *pgs1* QTL and twenty-eight when the *pgs5* QTL was targeted. Nucleotide sequences were subjected to similarity searches using the BLAST program. Seven sequences were isolated at least twice with different primer combinations, but in all such cases only the oligo(dT) used in the primer set was different. To confirm gene expression and estimate the size of the corresponding transcripts, the cloned DD fragments were used to probe virtual Northern blots. Eleven fragments failed to produce a clear expression pattern, shared no sequence identity with known nucleotide sequences, and were not analyzed further. A total of 22 genes were identified and characterized from our BSA-DD assay. Putative identities for 21 of these genes are summarized in Table 2. Note that the *AeIMUC1* gene, also isolated during this screen, has been described previously (Morlais and Severson 2002) and was not included here.

Tissue- and stage-specific expression

The tissue and stage specificity of expression was assessed by RT-PCR using specific primers designed from the DD insert sequences and RNAs extracted from the Moyo-R strain. Expression patterns are shown in Fig. 2. Three genes (*Chym*, *CYP9J* and *AEG12*) are specifically expressed in adult female midguts; three are midgut-specific in adult females, but are also expressed in adult males (*SDR*) or at all developmental stages (*CRALBP* and *APN*). Four genes (*AEGBS11*, *FerH*, *AEGI27*, *ODC-AZ*) are preferentially expressed in adult female midguts, while five (*AEGI23*, *PABP*, *PGK*, *AEGI8* and *RpS11*) are repressed. Expression of *PGK* is enhanced in adult males and *AEGI23* is only expressed during adult stages. Finally, six genes (*Sec61*, *nAcBP*, *RNABP*, *PAPS*, *Ef2* and *AEGBR9*) do not exhibit differential expression, and are therefore constitutively expressed.

Genetic mapping

The cloned DD fragments were used to probe the Southern blots initially employed to genotype the F_2 population. In addition, SNP genotyping was investigated with several genes (*AEG12*, *APN* , *Chym* , *FerH* , *PGK* and *CYP9J*) for which SNPs were identified between the two strains in the full-length sequences. We were unable to map four fragments that did not show any strain-specific *Eco* RI polymorphisms and did not contain SNPs suitable for SNP genotyping. A total of 17 genes were mapped (Fig. 1). Four are located on chromosome 1, ten on chromosome 2 and three on chromosome 3. Genes on chromosome 2 are distributed across the entire length of the linkage group, but most importantly, seven are located within the targeted QTL areas.

Table 1 Primer sequences used for real-time PCR analysis with SYBR Green

Gene	Primer orientation	Primer sequence (5′ to 3′)	Amplicon size (bp)
APN	Forward	GGATGTTCGGAAAACACTGC	206
	Reverse	CCGAAACTTCCTCCGTTGTA	
CYP9J	Forward	TTTCAACCTGCAAGCGTC	154
	Reverse	TCAGAGCCAGTAGATAAACCC	
Chym	Forward	GGTGAAACAGTCCGGTGAAG	174
	Reverse	AAATAACATCACATCCGAGTCC	
FerH	Forward	GTTCCTGTTCGACAAGAACC	150
	Reverse	GCATTTTCCGGTGGTCTTC	
AEG12	Forward	CGGTTGCTTTATTGCCTCTG	84
	Reverse	CGAAGAAAGCCTTGAACTCG	
SDR	Forward	AAGCACAGCGAAAAAACCCA	173
	Reverse	TCCACTAAATCAAAACGCACT	
CRALBP	Forward	TATTCCGGCAAGGATTGG	112
	Reverse	TCGCTTTCTCTTTCCCTCTC	

Table 2 Characterization of cDNAs isolated by the BSA-DD procedure

Gene ID	GenBank Accession No.	Best homology (*E* value)	Polypeptide length	Map location[a]	Expression[b]
CHYM	AY038039, AY038040	*A. gambiae* serine protease (7e−64)	268	2	a♀G
AEG12	AY038041, AY038042	*A. gambiae* ANG12 precursor (3e−46)	207	2	a♀G
CRALBP	AF329893, AF390101	*D. melanogaster* CG5958 gene product (2e−89)	290	2	G
SDR	AY033621, AY033626	*D. melanogaster* CG12171 gene product (6e−64)	254	2	aG
AEGI23	AY033624, AY033625	None	78	2	−G, a
PGK	AY043171, AY043172	*D. melanogaster* phosphoglycerate kinase (1e−98)	415	2	−G, +a♂
EF2	AF331798, AY040342	*D. melanogaster* Ef2b gene product (0.0)	844	2	c-ex
AEGI8	AF326339, AF326340	*D. melanogaster* unknown (e−124)	335	2	−G
Sec61	AF326338, AF392805	Sec61 alpha isoform 2 (0.0)	476	2	c-ex
APN	AF378117, AF390100	*D. melanogaster* CG5839 gene product (e−167)	955	1	G
FerH	AF326341, AF326342	*D. melanogaster* Fer1HCH gene product (9e−35)	209	1	+G
AEGBS11	AY033622, AY033623	None	140	1	+G
CYP9J	AF329892, AF390099	*D. mettleri* CYP9 cytochrome P450 (e−100)	536	3	a♀G
ODC-AZ	AF396870, AF396871	*D. melanogaster* Gutfeeling protein (2e−40)	240	ND	+G
RNABP	AY033620	*D. melanogaster* CG11844 gene product (5e−23)	419	ND	c-ex
RpS11	AF315552	*A. gambiae* 40S ribosomal protein S11 (5e−58)	152	ND	−G
PAPS	AY038044	*D. melanogaster Paps* gene product (e−151)	–	ND	c-ex
PABP	AY038043	*D. melanogaster Pabp* gene product (3e−94)	–	3	−G
nAcBP	AY040341	*D. melanogaster* CG16751 gene product (3e−19)	–	1	c-ex
AEGBR9	BG937403	*A. aegypti* AEMTAI57 EST (7e−34)	–	3	c-ex
AEGI27	BG937399-402, BI074240-3	*A. aegypti* AEMTBG08 EST (0.0)	–	2	+G

[a]ND, not determined
[b]a, adult; ♀, female; ♂,male; G, gut; c-ex, constitutively expressed

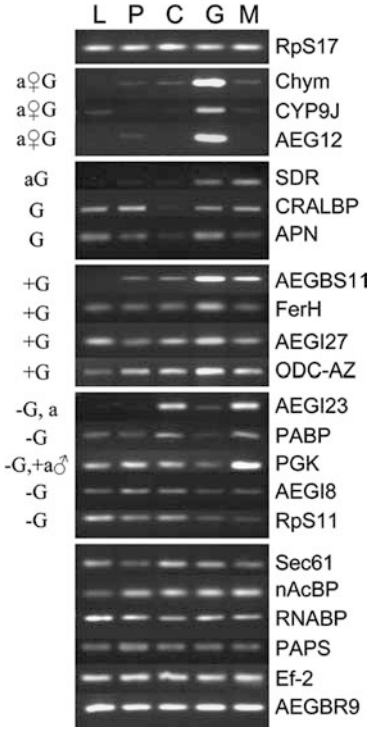


Fig. 2 Analysis of developmental and tissue-specific expression by RT-PCR. Expression profiles of BSA-DD genes in larvae (L), pupae (P), adult female carcasses minus midguts (C), adult female midguts (G) and adult males (M) are depicted. Expression levels are compared relative to ribosomal protein S17. The expression patterns are classified on the *left*: a, adult; ♀, female; ♂, male; G, gut

AEIG27 and *AEI23* map to the *pgs1* QTL region, and *AEG12*, *SDR* , *Sec61* , *PGK* and *Chym* were assigned to the *pgs5* QTL region. Note that the *AeIMUCI* gene also maps to the *pgs1* region (Morlais and Severson 2001). Therefore, 8 of 18 genes (44%) isolated using BSA-DD map within the targeted QTL regions.

Genes that are expressed differentially depending on the time after blood feeding and the infection status of the blood meal

Detailed expression analyses were performed on seven genes (*APN*, *AEG12*, *Chym*, *CRALBP*, *CYP9J*, *FerH*, *SDR*) that were identified in initial assays as being midgut-specific or differentially expressed in the midgut (Fig. 2). For these analyses, midgut expression was compared using real-time PCR at seven time points (before, and 12, 24, 36, 48 and 72 h and 7 days after blood feeding) in both the susceptible Red and the refractory Moyo-R strain, following feeding with *Plasmodium* -infected or uninfected blood meals. The observed results confirmed a differential expression pattern for these genes. Each gene showed a specific temporal pattern of expression after blood feeding, and gene expression varied depending on whether the blood meal contained the parasite or not, both among and within strains (Table 3 and Fig. 3).

Overall expression patterns for *APN*, *AEG12*, *CRALBP* and *FerH* were similar: expression was low in the female gut prior to blood feeding, and increased significantly in both strains by 12 h after feeding with either uninfected or infected blood. Thereafter, transcript levels decreased to the levels observed before blood feeding by 48–72 h.

APN was significantly up-regulated at 12 h following feeding with an infected compared to an uninfected

Table 3 Statistical analysis of expression data

Gene	*P* value[a]			
	Time	Strain	Infection status	Strain*Infection status
APN	< 0.0001	0.0193	0.0090	0.2418
CYP9J	< 0.0001	< 0.0001	0.0018	0.3069
Chym	< 0.0001	0.5514	0.0454	0.9269
FerH	< 0.0001	0.0299	0.0246	0.8900
AEG12	< 0.0001	< 0.0001	< 0.0001	< 0.0001
SDR	< 0.0001	< 0.0001	0.0263	0.2239
CRALBP	< 0.0001	< 0.0001	0.0164	0.0099

[a]Expression levels for each gene were analyzed using a three-way ANOVA Global Model (Proc GLM; SAS) and included time, strain and infection status as the independent factors. The table shows the *P* values for each factor

blood meal in both the Red and Moyo-R strains (Red, $P < 0.0018$, Moyo-R, $P < 0.0002$), but was quickly down-regulated thereafter. *APN* shows strong sequence identity with aminopeptidase genes from other insects, including *Bombyx mori* (GenBank Accession No. BAA32140, 52%), *Manduca sexta* (P91885, 52%) and *Drosophila melanogaster* (the CG5839 gene, AAF55912). The *Drosophila* gene contains two tandem repeats of the aminopeptidase motif; sequence identities with *APN* were 57% with the 3′-end repeat and 55% with the 5′-end sequence. Aminopeptidases are proteolytic enzymes that are involved in a wide range of functions, including digestion. Several reports have shown that aminopeptidases are also involved in defense responses (Delmas et al. 1992; Pautot et al. 1993; Giugni et al. 1996), and aminopeptidase N from insect midguts is known to be involved in insecticide resistance. APNs also contain a putative *Cry1A* toxin-binding domain that is likely to be involved in determining affinity for Cry toxins (Nakanishi et al. 2002). Alignment of our *A. aegypti* APN with domains of other insect APNs reveals that this domain is highly conserved in the mosquito (Fig. 4A). Interestingly, aminopeptidases from *Anopheles* mosquitoes were previously shown to play a role in parasite-vector interactions. In *A. stephensi*, aminopeptidase activity is higher in a strain selected for refractoriness to P. falciparum (Feldmann et al. 1990). In a *Plasmodium* -resistant strain of *A. gambiae* (L35), a midgut aminopeptidase was up-regulated 24 h after a *P. berghei* -infected blood-meal, and no up-regulation was observed in the *Plasmodium* -susceptible G3 strain (Rosenfeld and Vanderberg 1999). Our results for *A. aegypti* and the cited reports for *Anopheles* spp. provide complementary evidence that *APN* induction is likely to be related to the mosquito immune response against malaria parasites.

AEG12 mRNA levels were also significantly up-regulated at 12 h following feeding with infected compared to uninfected blood in both the Red and Moyo-R strains (Red, $P < 0.0001$, Moyo-R, $P < 0.0001$). This parasite-induced up-regulation was observed only at 12 h for the Moyo-R strain, but the gene remained up-regulated in the Red strain at 24 h ($P < 0.0001$) and 36 h ($P < 0.0001$). *AEG12* shows 70% sequence identity to the *ANG12* precursor from *A. gambiae* (Q17040), 60% identity with *Blattella germanica* major allergens Bla g 1.0101 (AAD13530) and Bla g 1.02 (AAD13531) and 56% with *Periplaneta americana* Cr-PII (AAB82404). A sequence alignment of *AEG12* with *ANG12* is shown in Fig. 5A. The high degree of sequence identity and the similarity in expression pattern between *AEG12* and *ANG12* suggest that both genes serve the same function. The deduced amino acid sequence of AEG12 indicates that the protein follows a secretory pathway, and given its female midgut specific expression the protein may have a function in the digestion of blood. The AEG12 protein has recently been described as a microvillar membrane protein (L. Shao and M. Jacobs-Lorena, unpublished; GenBank sequence AY050565). The high levels of identity of AEG12 and ANG12 to cockroach allergens suggest a possible role in the immune response, but in contrast to AEG12 and ANG12, the cockroach allergens contain multiple tandem amino acid repeats (Pomes et al. 1998; Wu et al. 1998). In *A. aegypti*, the rapid up-regulation of *AEG12* following an infected blood meal supports this suggestion.

CRALBP exhibited strain-specific up-regulation in response to an infected blood meal; its expression level was significantly higher at 24 h ($P < 0.0057$) in Red strain females that had fed on infected blood than in those that had been fed with an uninfected blood meal. In the Moyo-R strain, expression levels were enhanced by blood feeding, but the infection status of the blood meal itself had no effect ($P > 0.05$). *CRALBP* shows significant similarity to cellular retinaldehyde-binding proteins from other organisms and tocopherol-associated proteins (TAP), which are involved in vitamin metabolism. The protein exhibits the CRAL_TRIO domain, a common ligand-binding domain, required for binding and as a substrate carrier in humans (Crabb et al. 1998; Zimmer et al. 2000). Two distinct, but not- *A. aegypti* strain-specific, alleles of *CRALBP* were identified, which differ by a 48-bp deletion in the 3′-UTR sequence.

FerH expression levels in Red strain females receiving either an uninfected or infected blood meal were slightly higher than those observed in Moyo-R strain females; however, most comparisons were not significantly different. The only exception was that transcription was significantly higher ($P < 0.0243$) in Red strain females at 12 h following an infected blood meal compared to Moyo-R females fed with uninfected blood. The complete ORF for the *A. aegypti* ferritin heavy chain, *FerH* , was contained in the DD fragment; however, sequence alignment with nucleotide sequences from the *A. aegypti* Rockefeller strain (AF126431) and Bahamas strain (L37082) revealed that the 5′-UTR and 3′-UTR regions were incomplete. The FerH sequences differ between the Red and the Moyo-R strain by one amino-acid change (position 145, R↔K) and, the Rockefeller strain shares the Moyo-R allele, whereas

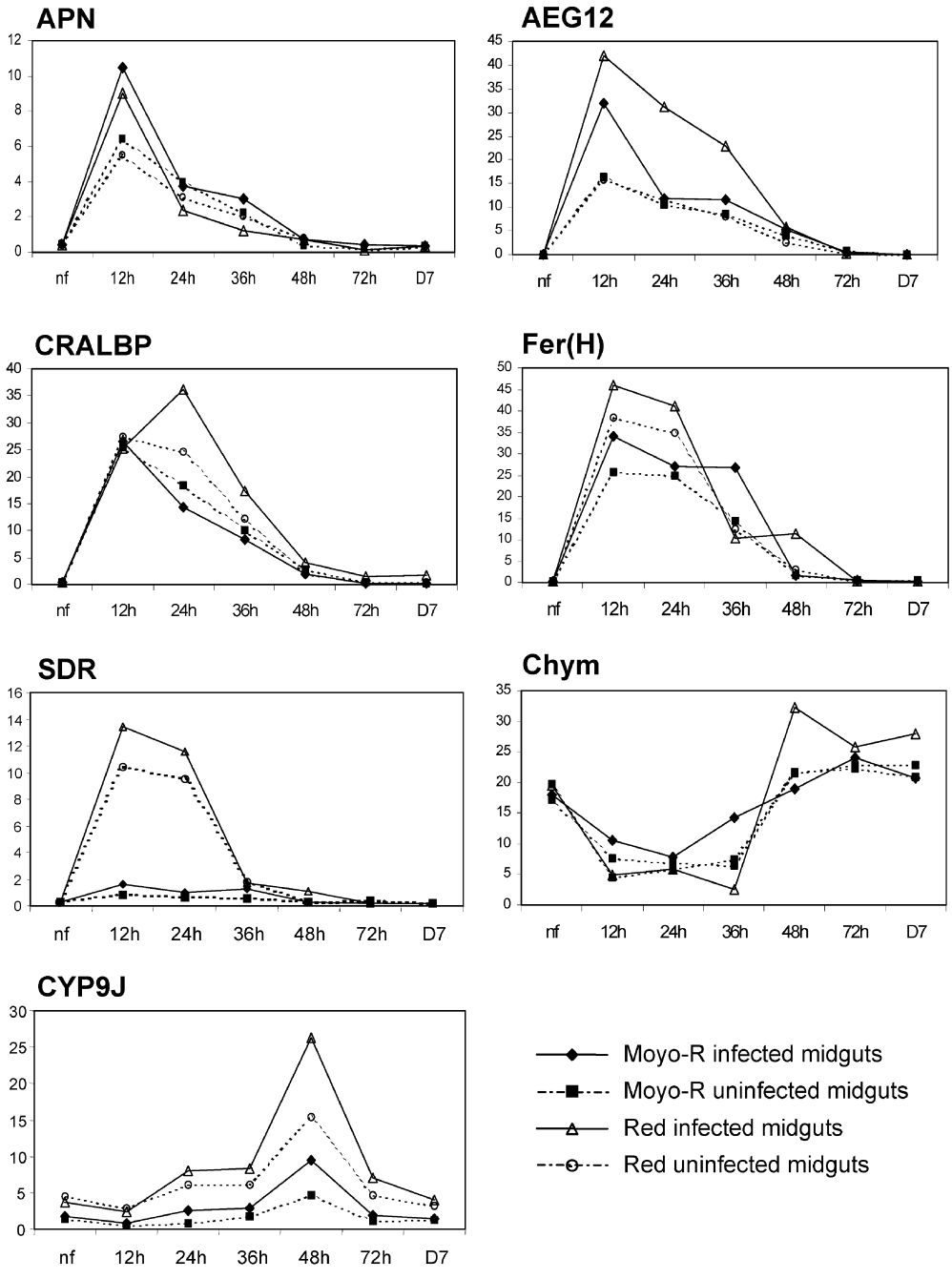
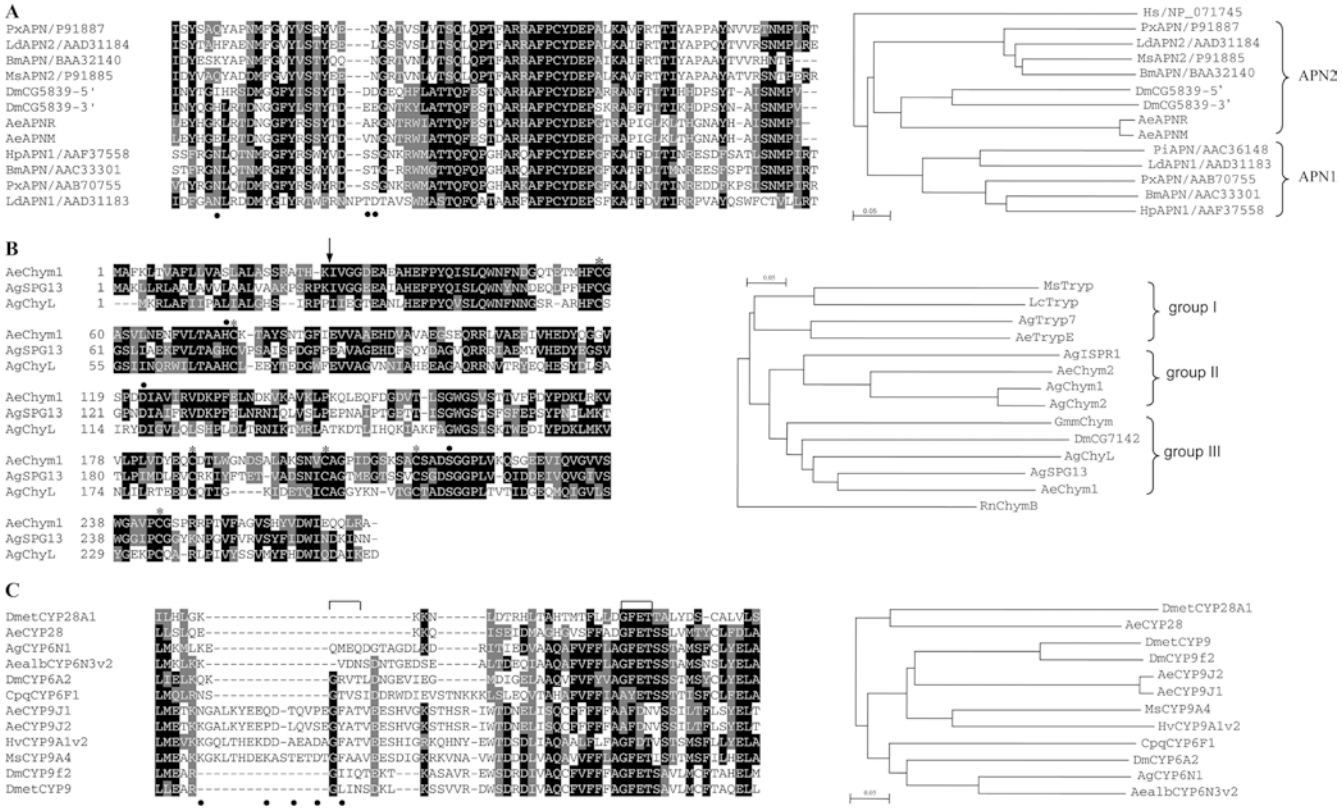


Fig. 3 Time course of BSA-DD transcript expression in the midguts of female mosquitoes. Real-time PCR analysis was performed using SYBR Green with RNA prepared from midguts at six different time points (12, 24, 36, 48, 72 h, and 7 days) of non-fed (nf) females, and from *P. gallinaceum* - infected and uninfected females of both the Moyo-R and Red strains of *A. aegypti*. All quantifications were normalized with respect to values for the gene for ribosomal protein S17. Data are the means of three independent experiments

the Bahamas strain has the Red strain genotype. Ferritins are iron-storage proteins that have been extensively studied in vertebrates, particularly because of their role in immune processes. In insects, ferritin proteins have been characterized from only a few species, and it has been postulated that they may play a role in intracellular iron transport (Dunkov et al. 1995).

SDR expression showed strong strain-specific up-regulation; it increased by 10- to 12-fold ($P < 0.0001$) in the Red strain at 12 h and 24 h following either an uninfected or infected blood meal. Interestingly, expression levels in the Moyo-R strain remained at the pre-blood feeding levels at all time points and, therefore, were not affected by blood feeding. SDR shows strong sequence similarity to other members of the short-chain dehydrogenase/reductase enzyme family. Its amino-acid sequence clearly exhibits the characteristic features of this family: the putative consensus sequence $G(X)_3$ GXG at Gly^{12} for the cofactor binding site, the consensus sequence $Y(X)_3$ K at Tyr^{152} for the active site and the three additional residues Gly^{132}, Ser^{139} and Pro^{182} (Fig. 5B). In humans, several members of the SDR family are involved in enzymatic metabolism and others participate in pathophysiological processes (Oppermann et al. 2001).

Chym mRNA levels were high before blood feeding, decreased rapidly after a blood meal and remained at



low levels until 36 h after blood feeding, returning to the pre-blood feeding levels by 48 h. *Chym* expression patterns showed no dependence on strain or blood meal infection status at any time point except at 48 h post-blood feeding; at that point expression was higher ($P < 0.0335$) in Red females that had obtained an infected blood meal. *A. aegypti Chym* shows 67% similarity to *A. gambiae* SPG13 (CAA93818) and 56% similarity to *A. gambiae* ChyL (AAC02700). Chymotrypsins are serine proteases associated with digestion in the mosquito midgut (Jiang et al. 1997; Muller et al. 1995; Shen et al. 2000). It is noteworthy that the expression profile we observed for *A. aegypti Chym* is similar to that observed in *A. gambiae* (Shen et al. 2000). Insect serine proteases also participate in the immune process (Hoffmann 1995; Hultmark 1993), and chymotrypsins have been reported to be involved in *Bt* resistance in lepidopteran larvae (Shao et al. 1998). Interestingly, the gene most closely related to *A. aegypti Chym*, AgSPG13, is strongly induced in larvae of *A. gambiae* challenged with bacteria (Dimopoulos et al. 1996). To assess the phylogenetic relationship of the *A. aegypti* chymotrypsin to the three groups of insect serine proteases (Shen et al. 2000), a consensus dendrogram was constructed from an alignment with other insect serine proteases. Figure 4B shows that the *A. aegypti Chym* sequence clusters within group III. Group III is characterized by a short activation peptide and high similarity to mammalian serine proteases known to participate in the immune response (Dimopoulos et al. 1996; Jiang et al. 1997; Shen et al. 2000).

CYP9J expression was low in unfed mosquitoes and increased only slightly in the 36 h following a blood meal; although not significantly different, expression levels were consistently higher in the Red strain during this period. Expression levels at 48 h increased sharply in the Red strain, compared to the Moyo-R strain, fed with either uninfected ($P < 0.0011$) or infected ($P < 0.0001$) blood; expression levels at 48 h in the Moyo-R individuals that had received an infected blood meal showed a modest but significant increase ($P < 0.0208$) over unfed levels. Expression levels dropped rapidly in both strains to reach pre-blood feeding levels by 72 h. *CYP9J* is a member of the cytochrome P450 family and has high similarity with other insect CYP9s. The sequences from the Red and Moyo-R strains were designated *CYP9J* 1 and *CYP9J* 2, respectively, following the recommendations for P450 nomenclature (http://drnelson.utmem.edu/Cytochrome P450.html). *CYP9J* shows motifs common to all cytochrome P450 genes, such as a membrane-anchoring signal (up to residue 25), the highly conserved heme-iron ligand signature FXXGXXXCXG (starting at position 471) and the EXXR motif within helix K (position 393–396). However, the GXXT motif associated with formation of the substrate-binding pocket of helix I is not present in *A. aegypti*, as shown in the alignment with other insect cytochrome P450s (Fig. 4C). Instead, a putative GXXT motif is found 33 amino acids upstream; the sequence is GFAT in the Red sequence and GYAT in the Moyo-R sequence. Cytochrome P450s (CYPs) represent a superfamily of

Fig. 4A–C Sequence alignment and phylogenetic analyses of selected RNAs expressed in the female midgut. **A** *Cry1A* toxin-binding domain alignment and phylogenetic tree of insect APNs. Accession Nos. are listed in the Figure. DmCG5839-3' and DmCG5839-5' represent the two tandem repeats of the *D. melanogaster* aminopeptidase, AAF55912. *Dots* indicate amino-acid changes between the *A. aegypti* sequences. The *Homo sapiens* sequence (NP_071745) was used as an outgroup to construct the dendrogram. The dendrogram places the *A. aegypti APN* sequences within the APN2 cluster. **B** Alignment of the *A. aegypti* chymotrypsin with the *A. gambiae* chymotrypsin-like protease AgChyL (AAC02700) and the *A. gambiae* serine protease Ag-SPG13 (CAA93818), and phylogenetic tree of insect serine proteases. *Dots* mark the catalytic triad (His^{73} Asp^{122} and Ser^{218}) typical of serine proteases, while the six cysteines residues that form three disulfide bonds in chymotrypsin proteins are each marked with a *star*, and the peptide cleavage site is indicated by an *arrow*. The dendrogram shows the three groups of insect serine proteases (Shen et al. 2000) and the *A. aegypti Chym* (*AeChym1*) sequence clusters within group III. *GmmChym*, *Glossina morsitans* chymotrypsin-like serine protease (AAF91345); *AeChym2*, *A. aegypti* chymotrypsin II-like protein (AAF43707); *AgChym2*, *A. gambiae* chymotrypsinogen-like protease (CAA83567); *AeTrypE*, *A. aegypti* early trypsin (AAM34268); *AgChym1*, *A. gambiae* chymotrypsinogen-like protease (CAA83568); *DmCG7142*, *D. melanogaster* CG7142 gene product (AAF55527); *AgTryp7*, *A. gambiae* TRYP7 gene (P35041); *LcTryp*, *Lucilia cuprina* chymotrypsinogen (AAA68986); *MsTryp*, *Manduca sexta* chymotrypsinogen (AAA58743); *AgISPR1*, *A. gambiae* immune-responsive chymotrypsin-like serine protease (AAF19831); *RnChymB*, *Rattus norvegicus* chymotrypsinogen (NP_036668). This last was used as the outgroup. **C** Alignment of the insect CYP substrate-binding pocket region and dendrogram showing the location of the *A. aegypti CYP9J* sequence within the CYP9 group. *AeCYP28*, *A. aegypti* cytochrome P450 (AAG36958); *HvCYP9A1v2*, *Heliothis virescens* cytochrome P450 (AAD25104); *DmetCYP9*, *Drosophila mettleri* CYP9 cytochrome P450 (AAC33299); *DmetCYP28A1*, *D. mettleri* cytochrome P450 (AAB95198); *DmCYP9f2*, *D. melanogaster* cytochrome P450 9f2 (Q9VG82); *AealbCYP6N3v2*, *A. albopictus* cytochrome P450 (AAF97937); *AgCYP6N1*, *A. gambiae* cytochrome P450 (AAK32960); *DmCYP6A2*, *D. melanogaster* cytochrome P450 6a2 (P33270); *CpqCYP6F1*: *Culex pipiens quinquefasciatus* cytochrome P450 (BAA92152); *MsCYP9A4*: *M. sexta* cytochrome P450 (AAD51037). The putative GXXT motifs are indicated by the *bracket*

diverse and widely distributed enzymes involved in many metabolic systems. In insects, their functions include the metabolism of steroids and fatty acids, although cytochrome P450s are particularly well known for their role in both insecticide resistance and tolerance to plant toxins (Feyereisen 1999; Scott et al. 1998). *CYP9J* sequences were compared with insect cytochrome P450s of three families (CYP28, CYP6 and CYP9) and the phylogenetic tree obtained from the sequence alignment (Fig. 4C) is consistent with the classification of *A. aegypti CYP9J* as a member of the CYP9 family. The CYP9 family has only recently been created, and the first member of this family, CYP9A1 from *Heliothis virescens*, has been associated with insecticide resistance (Rose et al. 1997). Several isoforms of CYP9A from *M. sexta* have been identified and they are induced in the lepidoptera midgut by xenobiotics and allelochemicals (Stevens et al. 2000).

Genes with unknown functions isolated by the BSA-DD procedure

As eukaryotic genome projects multiply, the number of genes isolated that have no similarity with any known proteins has increased, and their functions remain unknown (Schmid and Tautz 1997; Tautz and Schmid 1998). Although most of the genes that we isolated showed similarity to previously described genes with known function, we also identified five novel genes (AEGBS11, AEGBR9, AEGI23, AEGI27 and AEGI8) with no known homologues or functionally related genes in the public databases. Northern analysis did, however, indicate that these sequences correspond to expressed genes, and primary structure analysis of their deduced amino acid sequences identified several functional domains. These could represent uncharacterized eukaryotic protein classes or they may represent rapidly evolving sequences. Because genes involved in specific adaptations such as parasite defense evolve very fast, it seems likely that some of the sequences identified as unknown proteins may play a role in these adaptations (Schmid and Tautz 1997).

The AEGI27 fragment is 962 bp long, and a BLAST search in the EST database revealed over 93% identity with several sequences from the *A. aegypti* MT pSPORT library (AI619038, AI618909, AI621631, AI637417, AI629432, AI612185) and with a random cDNA sequence from *A. aegypti*, LF181 (T58316, T58317), that had previously been placed on the genetic linkage map (Severson et al. 2002). The transcript was confirmed as expressed by Virtual Northern analysis and the transcript size was estimated to be ~1.5 kb. Because of its map location, within the *pgs1* QTL interval on chromosome 2, we attempted to obtain the full-length sequence. Fragments of 1.5 kb were obtained from our cDNA libraries and several clones were sequenced for each strain. Unfortunately, none of these sequences contained the expected complete ORF. Strain-specific polymorphisms were found, with a total of 11 SNPs and a 2-bp deletion. A Southern blot bearing total genomic DNA digested with four different restriction enzymes (*Eco*RI, *Hinc* II, *Hind* III and *Bam* HI) was hybridized with the cloned AEGI27 insert. We observed multiple restriction fragments in each digest, indicating that AEGI27 is likely to represent a multicopy gene (data not shown).

The AEGBS11 gene is predicted to encode a protein of 140 amino acids that contains two hydrophobic segments (Fig. 5C). The N terminal hydrophobic sequence consists of a 16-residue signal peptide, and a signal cleavage site is predicted at Ser^{18}. The C terminal hydrophobic sequence of 22 amino acids represents a GPI-anchored signal. The presence of these two hydrophobic segments indicates that the protein is secreted and exported.

AEGI23 consists of 309 bp and contains a putative ORF encoding a 78-amino acid peptide. Virtual Northern blotting confirmed transcript expression;

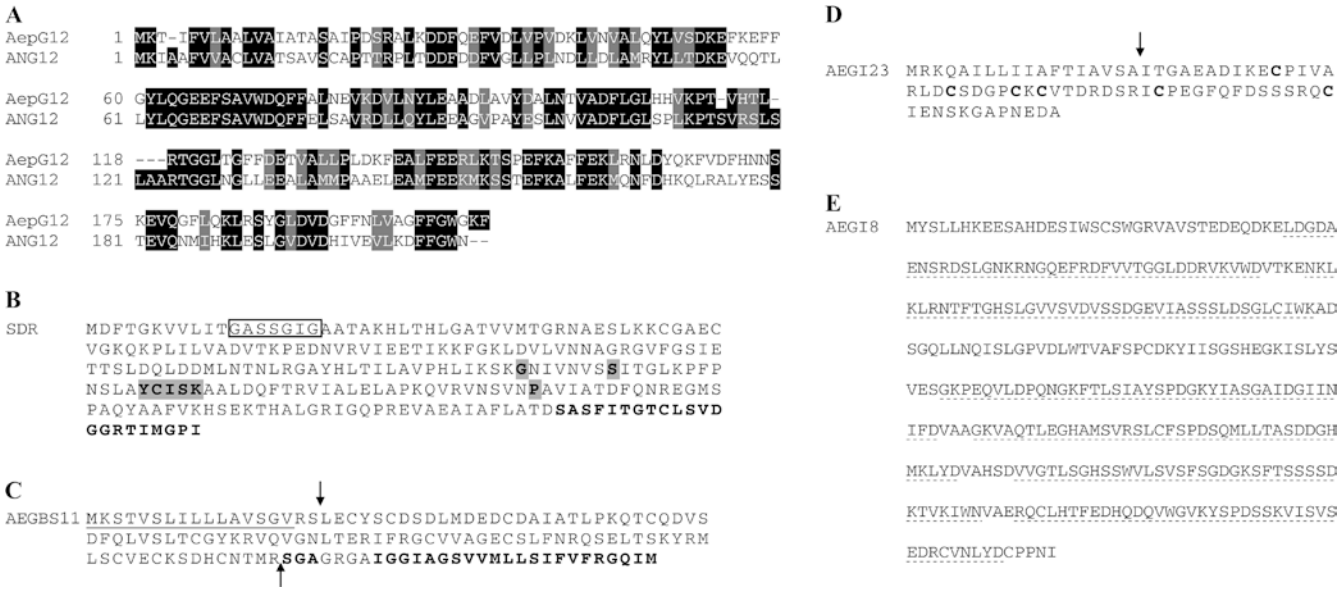


Fig. 5A–E Predicted sequences of the protein products of genes isolated by the BSA-DD procedure. **A** Alignment of the *AEG12* amino-acid sequence with the homologous *ANG12* precursor (Q17040). **B** Amino-acid sequence of the putative product of the *SDR* gene. The $G(X)_3$ GXG motif is *boxed*, the active site including the $Y(X)_3$ K motif at Tyr^{152} and the three additional residues Gly^{132}, Ser^{139} and Pro^{182} are *shaded*. **C** Amino-acid sequence of the *AEGBS11* gene product. The N terminal hydrophobic sequence is *underlined* and the C terminal GPI-anchored signal is shown in *bold*. The *arrows* indicate potential cleavage sites for each hydrophobic segment. **D** Amino acid sequence of the *AEGI23* gene product. The N-terminal putative signal peptide is *underlined* and the *arrow* indicates the putative cleavage site. The six cysteine residues are shown in *bold*. **E** Amino-acid sequence of the predicted product of the *AEGI8* gene. The seven WD40 repeats are *underlined*

transcript size is consistent with an estimated product size of ~80 amino acids. The eighteen-residue N-terminal segment is predicted to be a putative cleavable signal peptide by von Heijne's rules (Nielsen et al. 1997) (Fig. 5D). In addition, the peptide sequence contains six cysteine residues that could be involved in protein tertiary structure.

The *A. aegypti* AEGI8 sequence is 1120 bp long, and codes for a 335-amino acid polypeptide. The protein shows 79% similarity to a protein of unknown function from *D. melanogaster* (AAF54480), 70% similarity to the recombination protein REC14 from *Homo sapiens* (NP_079510) and 59% similarity to a G-protein beta-subunit from *Schistosoma mansoni* (AAC46896). The peptide sequence exhibits seven putative WD40 repeat motifs (residues 34–73, 79–118, 122–160, 164–203, 207–245, 248–287 and 291– 329; see Fig. 5E), spanning the entire length of the sequence. WD-40 repeats are found in β-transducin, the β-subunit of the heterodimeric G protein complexes that participate in the transduction of signals generated by transmembrane receptors.

AEGBR9 is a 659-bp fragment, which shows no match to known genes in the databases, but the last 99 amino acids of the sequence are 93% identical to the 5′-end of the AEMTAI57 sequence (AI637429) from the *A. aegypti* MT pSPORT library. The transcript has an estimated size of 2.8 kb and is expressed constitutively during mosquito development.

Conclusions

The BSA-DD procedure allowed us to compare transcriptional activity in the midguts of both susceptible and refractory female *A. aegypti* mosquitoes following a normal blood meal and a *P. gallinaceum* -infected blood meal, and to focus on transcripts derived from regions (QTLs) known to contain genes determining susceptibility. To target genes located within these QTL intervals, we bulked midgut RNAs isolated from F_2 females that had previously been genotyped at two linked QTLs on chromosome 2, *pgs1* and *pgs5*, with each RNA bulk representing females homozygous for one or the other parental genotype at the loci flanking the individual QTL. For each QTL interval, the bulked RNA from females with the susceptible genotype was compared by the differential display procedure (Liang and Pardee 1992) to the bulked RNA from those with the refractory genotype. The RNA bulks were prepared from midguts dissected from females 48 h after feeding with infected blood. By that time, ookinetes should have developed and crossed the midgut epithelium, and initiated development to the oocyst stage under the basal lamina in susceptible individuals.

Detailed sequence information and expression profiles were determined for 21 sequences derived from our BSA-DD screens. Results from the analysis of an additional BSA-DD isolate, *AeIMUCI* , are reported elsewhere (Morlais and Severson 2001). It should be mentioned here that the differential display method is well known for generating a high proportion of false positives and also showing a bias for isolation of abundant transcripts (Bertioli et al. 1995; Ledakis et al.

1998). However, sixteen of the 22 BSA-DD sequences we characterized (73%) showed some level of differential expression pattern, when expression in the female midgut was compared to that in larvae, pupae, female carcass minus midgut, or adult males. Furthermore, genetic linkage analyses with 18 of these sequences indicated that 11 map to the chromosome targeted by BSA (Severson et al. 2002), and eight map directly within the two targeted QTL regions on chromosome 2. Eight of the 22 BSA-DD sequences (36%) represent genes that have been shown, or are likely, to be involved in the mosquito immune response.

In conclusion, the integrated BSA-DD approach that we used to target genes involved in determining susceptibility to infection by *P. gallinaceum* in *A. aegypti* allowed us to identify several genes with a putative role in the host-parasite interactions in the mosquito midgut. Some of these genes may represent valuable candidate genes for use in the development of strategies to block parasite transmission, although further studies are needed to determine their specific function and their possible roles in vector competence.

Acknowledgements This work was supported by Grant No. AI33127 from the National Institutes of Health.

References

Akey JM, Sosnoski D, Parra E, Dios S, Hiester K, Su B, Bonilla C, Jin L, Shriver MD (2001) Melting curve analysis of SNPs (McSNP): a gel-free and inexpensive approach for SNP genotyping. Biotechniques 30:358–367

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Arca B, Lombardo F, de Lara Capurro M, della Torre A, Dimopoulos G, James AA, Coluzzi M (1999) Trapping cDNAs encoding secreted proteins from the salivary glands of the malaria vector *Anopheles gambiae*. Proc Natl Acad Sci USA 96:1516–1521

Bertioli DJ, Schlichter UHA, Adams MJ, Burrows PK, Steinbiss HH, Antoniw JF (1995) An analysis of differential display shows a strong bias towards high copy number messenger RNAs. Nucleic Acids Res 23:4520–4523

Catteruccia F, Nolan T, Blass C, Muller HM, Crisanti A, Kafatos FC, Loukeris TG (2000) Toward *Anopheles* transformation: *Minos* element activity in anopheline cells and embryos. Proc Natl Acad Sci USA 97:2157–2162

Coates CJ (2000) Malaria. A mosquito transformed. Nature 405:900–901

Crabb JW, Carlson A, Chen Y, Goldflam S, Intres R, West KA, Hulmes JD, Kapron JT, Luck LA, Horwitz J, Bok D (1998) Structural and functional characterization of recombinant human cellular retinaldehyde-binding protein. Protein Sci 7:746–757

Delmas B, Gelfi J, L'Haridon R, Vogel LK, Sjostrom H, Noren O, Laude H (1992) Aminopeptidase N is a major receptor for the entero-pathogenic coronavirus TGEV. Nature 357:417–420

Dessens JT, Beetsma AL, Dimopoulos G, Wengelnik K, Crisanti A, Kafatos FC, Sinden RE (1999) CTRP is essential for mosquito infection by malaria ookinetes. EMBO J 18:6221–6227

Dimopoulos G, Richman A, della Torre A, Kafatos FC, Louis C (1996) Identification and characterization of differentially expressed cDNAs of the vector mosquito, *Anopheles gambiae*. Proc Natl Acad Sci USA 93:13066–13071

Dimopoulos G, Seeley D, Wolf A, Kafatos FC (1998) Malaria infection of the mosquito *Anopheles gambiae* activates immune-responsive genes during critical transition stages of the parasite life cycle. EMBO J 17:6115–6123

Dimopoulos G, Muller HM, Levashina EA, Kafatos FC (2001) Innate immune defense against malaria infection in the mosquito. Curr Opin Immunol 13:79–88

Dunkov BC, Zhang D, Choumarov K, Winzerling JJ, Law JH (1995) Isolation and characterization of mosquito ferritin and cloning of a cDNA that encodes one subunit. Arch Insect Biochem Physiol 29:293–307

Feldmann AM, Billingsley PF, Savelkoul E (1990) Bloodmeal digestion by strains of *Anopheles stephensi* Liston (Diptera: Culicidae) of differing susceptibility to *Plasmodium falciparum*. Parasitology 101:193–200

Feyereisen R (1999) Insect P450 enzymes. Annu Rev Entomol 44:507–533

Ghosh A, Edwards MJ, Jacobs-Lorena M (2000) The journey of the malaria parasite in the mosquito: hopes for the new century. Parasitol Today 16:196–201

Giugni TD, Soderberg C, Ham DJ, Bautista RM, Hedlund KO, Moller E, Zaia JA (1996) Neutralization of human cytomegalovirus by human CD13-specific antibodies. J Infect Dis 173:1062–1071

Hoffmann JA (1995) Innate immunity of insects. Curr Opin Immunol 7:4–10

Huber M, Cabib E, Miller LH (1991) Malaria parasite chitinase and penetration of the mosquito peritrophic membrane. Proc Natl Acad Sci USA 88:2807–2810

Hultmark D (1993) Immune reactions in Drosophila and other insects: a model for innate immunity. Trends Genet 9:178–83.

Jasinskiene N, Coates CJ, Benedict MQ, Cornel AJ, Rafferty CS, James AA, Collins FH (1998) Stable transformation of the yellow fever mosquito, *Aedes aegypti* , with the *Hermes* element from the housefly. Proc Natl Acad Sci USA 95:3743–3747

Jiang Q, Hall M, Noriega FG, Wells M (1997) cDNA cloning and pattern of expression of an adult, female-specific chymotrypsin from *Aedes aegypti* midgut. Insect Biochem Mol Biol 27:283–289

Kokoza V, Ahmed A, Cho WL, Jasinskiene N, James AA, Raikhel A (2000) Engineering blood meal-activated systemic immunity in the yellow fever mosquito, *Aedes aegypti*. Proc Natl Acad Sci USA 97:9144–9149

Kosambi DD (1944) The estimation of map distance from recombination values. Ann Eugen 12:172–175

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Ledakis P, Tanimura H, Fojo T (1998) Limitations of differential display. Biochim Biophys Res Comm 251:653–656

Liang P, Pardee AB (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257:967–971

Lincoln SE, Lander ES (1990) Mapping genes controlling quantitative traits using MAPMAKER/QTL. Whitehead Institute for Biomedical Research, Cambridge, Mass.

Luckhart S, Vodovotz Y, Cui L, Rosenberg R (1998) The mosquito *Anopheles stephensi* limits malaria parasite development with inducible synthesis of nitric oxide. Proc Natl Acad Sci USA 95:5700–5705

Moreira LA, Edwards MJ, Adhami F, Jasinskiene N, James AA, Jacobs-Lorena M (2000) Robust gut-specific gene expression in transgenic *Aedes aegypti* mosquitoes. Proc Natl Acad Sci USA 97:10895–10898

Morlais I, Severson DW (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

Muller HM, Catteruccia F, Vizioli J, della Torre A, Crisanti A (1995) Constitutive and blood meal-induced trypsin genes in *Anopheles gambiae*. Exp Parasitol 81:371–385

Nakanishi K, Yaoi K, Nagino Y, Hara H, Kitami M, Atsumi S, Miura N, Sato R (2002) Aminopeptidase N isoforms from the midgut of *Bombyx mori* and *Plutella xylostella* —their classification and the factors that detrmine their binding specificity to *Bacillus thuringiensis* CryIA toxin. FEBS Lett 519:215–220

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

Oduol F, Xu J, Niare O, Natarajan R, Vernick KD (2000) Genes identified by an expression screen of the vector mosquito *Anopheles gambiae* display differential molecular immune response to malaria parasites and bacteria. Proc Natl Acad Sci USA 97:11397–11402

Oppermann UC, Filling C, Jornvall H (2001) Forms and functions of human SDR enzymes. Chem Biol Interact 130–132:699–705

Pautot V, Holzer FM, Reisch B, Walling LL (1993) Leucine aminopeptidase—an inducible component of the defense response in *Lycopersicon esculentum* (tomato). Proc Natl Acad Sci USA 90:9906–9910

Pomes A, Melen E, Vailes LD, Retief JD, Arruda LK, Chapman MD (1998) Novel allergen structures with tandem amino acid repeats derived from German and American cockroach. J Biol Chem 273:30801–30807

Richards AG, Richards PA (1977) The peritrophic membranes of insects. Annu Rev Entomol 22:219–240

Richman AM, Dimopoulos G, Seeley D, Kafatos FC (1997) *Plasmodium* activates the innate immune response of *Anopheles gambiae* mosquitoes. EMBO J 16:6114–6119

Rose RL, Goh D, Thompson DM, Verma KD, Heckel DG, Gahan LJ, Roe RM, Hodgson E (1997) Cytochrome P450 (CYP)9A1 in *Heliothis virescens*: the first member of a new CYP family. Insect Biochem Mol Biol 27:605–615

Rosenfeld A, Vanderberg JP (1999) *Plasmodium berghei*: induction of aminopeptidase in malaria-resistant strain of *Anopheles gambiae*. Exp Parasitol 93:101–104

Schmid KJ, Tautz D (1997) A screen for fast evolving genes from Drosophila. Proc Natl Acad Sci USA 94:9746–9750

Scott JG, Liu N, Wen Z (1998) Insect cytochromes P450: diversity, insecticide resistance and tolerance to plant toxins. Comp Biochem Physiol C Pharmacol Toxicol Endocrinol 121:147–155

Severson DW (1997) RFLP analysis of insect genomes. In: Crampton JM, Beard CB, Louis C (eds) The Molecular biology of insect disease vectors. Chapman and Hall, London, pp 309–320

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

Severson DW, Meece JK, Lovin DD, Saha G, 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, Criscio M, Kaslow DC (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

Shao Z, Cui Y, Liu X, Yi H, Ji J, Yu Z (1998) Processing of delta-endotoxin of *Bacillus thuringiensis* subsp. kurstaki HD-1 in *Heliothis armigera* midgut juice and the effects of protease inhibitors. J Invertebr Pathol 72:73–81

Shen Z, Dimopoulos G, Kafatos FC, Jacobs-Lorena M (1999) A cell surface mucin specifically expressed in the midgut of the malaria mosquito *Anopheles gambiae*. Proc Natl Acad Sci USA 96:5610–5615

Shen Z, Edwards MJ, Jacobs-Lorena M (2000) A gut-specific serine protease from the malaria vector *Anopheles gambiae* is down-regulated after blood ingestion. Insect Mol Biol 9:223–229

Sieber KP, Huber M, Kaslow D, Banks SM, Torii M, Aikawa M, Miller LH (1991) The peritrophic membrane as a barrier: its penetration by *Plasmodium gallinaceum* and the effect of a monoclonal antibody to ookinetes. Exp Parasitol 72:145–156

Stevens JL, Snyder MJ, Koener JF, Feyereisen R (2000) Inducible P450s of the CYP9 family from larval *Manduca sexta* midgut. Insect Biochem Mol Biol 30:559–568

Tautz D, Schmid KJ (1998) From genes to individuals: developmental genes and the generation of the phenotype. Philos Trans R Soc Lond B Biol Sci 353:231–240

Tellam RL, Wijffels G, Willadsen P (1999) Peritrophic matrix proteins. Insect Biochem Mol Biol 29:87–101

Thathy V, Severson DW, Christensen BM (1994) Reinterpretation of the genetics of susceptibility of *Aedes aegypti* to *Plasmodium gallinaceum*. J Parasitol 80:705–712

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Vinetz JM, Valenzuela JG, Specht CA, Aravind L, Langer RC, Ribeiro JM, Kaslow DC (2000) Chitinases of the avian malaria parasite *Plasmodium gallinaceum*, a class of enzymes necessary for parasite invasion of the mosquito midgut. J Biol Chem 275:10331–10341

Wu CH, Wang NM, Lee MF, Kao CY, Luo SF (1998) Cloning of the American cockroach Cr-PII allergens: evidence for the existence of cross-reactive allergens between species. J Allergy Clin Immunol 101:832–840

Zieler H, Nawrocki JP, Shahabuddin M (1999) *Plasmodium gallinaceum* ookinetes adhere specifically to the midgut epithelium of *Aedes aegypti* by interaction with a carbohydrate ligand. J Exp Biol 202:485–495

Zimmer S, Stocker A, Sarbolouki MN, Spycher SE, Sassoon J, Azzi A (2000) A novel human tocopherol-associated protein: cloning, in vitro expression, and characterization. J Biol Chem 275:25672–25680