

# Analysis of the *Plasmodium* and *Anopheles* Transcriptomes during Oocyst Differentiation\*<sup>§</sup>

Received for publication, July 14, 2003, and in revised form, October 13, 2003  
Published, JBC Papers in Press, November 19, 2003, DOI 10.1074/jbc.M307587200

Prakash Srinivasan<sup>‡</sup>§, Eappen G. Abraham<sup>‡</sup>§, Anil K. Ghosh<sup>‡</sup>§, Jesus Valenzuela<sup>¶</sup>,  
Jose M. C. Ribeiro<sup>¶</sup>, George Dimopoulos<sup>§</sup>||, Fotis C. Kafatos<sup>||</sup>, John H. Adams<sup>\*\*</sup>, Hisashi Fujioka<sup>‡</sup>,  
and Marcelo Jacobs-Lorena<sup>‡</sup>§<sup>‡‡</sup>

From the <sup>‡</sup>Department of Genetics, Case Western Reserve University, Cleveland, Ohio 44106, <sup>¶</sup>Medical Entomology Section, Laboratory of Parasitic Diseases, NIAID, National Institutes of Health, Bethesda, Maryland 20892-0425, <sup>||</sup>European Molecular Biology Laboratory, Heidelberg 69117, Germany, and <sup>\*\*</sup>Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana 46556

Understanding the life cycle of the malaria parasite in its mosquito vector is essential for developing new strategies to combat this disease. Subtractive hybridization cDNA libraries were constructed that are enriched for *Plasmodium berghei* and *Anopheles stephensi* genes expressed during oocyst differentiation on the midgut. Sequencing of 1485 random clones led to the identification of 1137 unique expressed sequence tags. Of the 608 expressed sequence tags with data base hits, 320 (53%) had significant matches to the non-redundant protein data base, whereas 288 (47%) with matches only to genomic data bases represent novel *Plasmodium* and *Anopheles* genes. Transcription of six novel parasite genes and two previously identified asexual stage genes was up-regulated during oocyst differentiation. In addition, the mRNA for an *Anopheles* fibrinogen domain gene was induced on day 2 after an infectious blood meal, at the time of ookinete to oocyst differentiation. The subcellular distribution of MAEBL, a sporozoite surface protein, is developmentally regulated from presumed storage organelles in day 15 oocysts to uniform distribution on the surface in day 22 oocysts. This redistribution may reflect a sporozoite maturation program in preparation for salivary gland invasion. Furthermore, apical membrane antigen 1, another parasite surface molecule, is translationally regulated late in sporozoite development, suggesting a role during infection of the vertebrate host. The present results and those of an accompanying report (Abraham, E. G., Islam, S., Srinivasan, P., Ghosh, A. K., Valenzuela, J., Ribeiro, J. M., Kafatos, F. C., Dimopoulos, G., & Jacobs-Lorena, M. (2003) *J. Biol. Chem.* 279, 5573–5580) provide the foundation for studies seeking to understand at the molecular level *Plasmodium* development and its interactions with the mosquito.

Malaria is the major vector-borne infectious disease. Nearly 40% of the world's population is at risk, and close to 2 million persons (mostly children under the age of five) die every year. Despite considerable efforts to contain the disease, the number of infected persons is raising (2). The lack of an effective vaccine and the emergence of drug-resistant parasites and insecticide-resistant mosquitoes warrant the search for alternate strategies to combat this disease.

The parasite undergoes a complex developmental program in the mosquito (3, 4). After traversing the midgut, the parasite lodges itself between the midgut epithelium and the basal lamina, giving rise to an oocyst. Upon maturation, sporozoites formed inside the oocyst are first released into the hemocoel and then invade the salivary glands. To date only a handful of parasite genes expressed specifically during differentiation in the mosquito have been identified. Oocyst differentiation, a process that leads to sporozoite formation, takes close to 2 weeks. Identifying genes that are required for this process and for the subsequent release of sporozoites could provide novel targets to interfere with parasite development. Thus far only two genes, namely circumsporozoite protein (CS)<sup>1</sup> and a scavenger receptor-like protein, have been shown to be required for sporozoite formation (5, 6). The mosquito responds to the presence of the parasite by mounting an immune response. A variety of immune genes are induced during ookinete penetration of the midgut and again at the time of sporozoite release into the hemocoel (7). The parasite elicitors of these responses are not known.

The lack of methods to fractionate oocysts from mosquito midgut tissues has hindered the molecular analysis of oocyst development. An additional limitation is that only a small proportion of the RNA extracted from infected mosquito guts originates from the parasite. In an attempt to overcome some of these limitations we used a subtractive hybridization strategy to enrich for *Plasmodium berghei* and *Anopheles stephensi* genes that are induced during oocyst development in the midgut. Numerous novel genes both of parasite and of mosquito origin were identified. A mosquito fibrinogen-like gene that is induced by the parasite and two *Plasmodium* surface proteins, MAEBL and apical membrane antigen 1 (AMA1), were further characterized.

## EXPERIMENTAL PROCEDURES

**Parasites**—*P. berghei* ANKA strain, clone 2.34, was maintained by passage in Swiss Webster mice. Mice with 15–20% parasitemia and 2–3 exflagellations per field (40× magnification) were used for mosquito infections.

<sup>1</sup> The abbreviations used are: CS, circumsporozoite protein; EST, expressed sequence tag; FBN, fibrinogen; AMA1, apical membrane antigen 1; AsFBN1, *A. stephensi* fibrinogen domain protein 1.

\* This work was supported by grants from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

<sup>§</sup> The on-line version of this article (available at <http://www.jbc.org>) contains Supplemental Fig. 1 and Table 1.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) CB603466 (L457), CB603120 (L1096), CB603146 (L1144), CB603148 (L1147), CB603277 (L2451), CB603415 (L3593), CB603107 (L1085), and CK320901 (M184).

<sup>‡</sup> Present address: Dept. Molecular Microbiology and Immunology and Malaria Research Institute, Johns Hopkins School of Public Health, Baltimore, MD 21205.

<sup>‡‡</sup> To whom reprint requests should be addressed. Tel.: 443-287-0839; Fax: 410-955-0105; E-mail: [mlorena@jhsph.edu](mailto:mlorena@jhsph.edu).

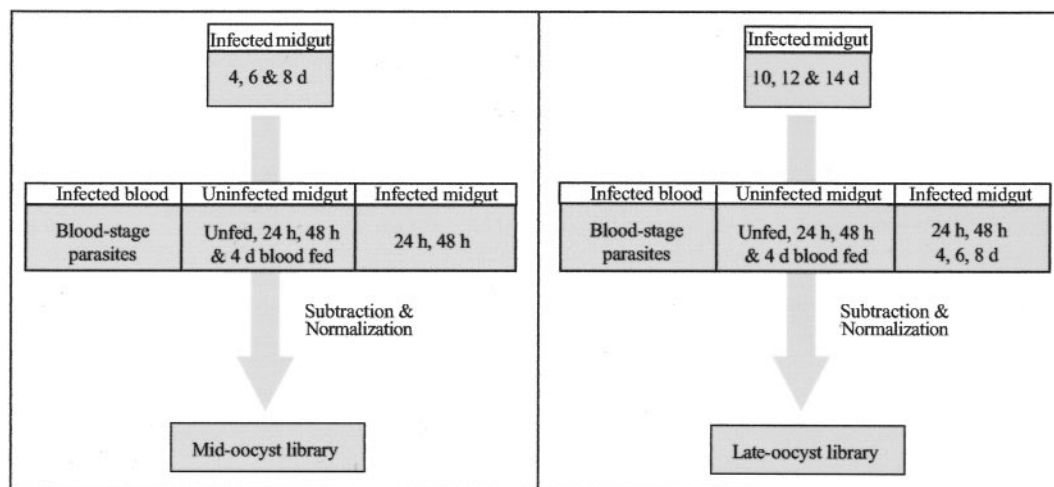


FIG. 1. **Construction of subtraction libraries.** For each library a mixture of equal amounts of cDNA from each of the samples listed in the middle set of three boxes was subtracted from a pool of equal amounts of cDNA from samples in the *Infected midgut* box at the top. *Infected blood*, cDNA from blood-stage parasites obtained from infected mice; *Uninfected* and *infected midgut*, cDNA prepared from mosquito midguts dissected at the indicated times (hours or days) after a non-infected or infected blood meal, respectively. *Unfed*, cDNA from midguts of mosquitoes that did not have a blood meal.

**Mosquito Infection**—*A. stephensi* mosquitoes were fed on infected mice in a 21 °C incubator. Mosquito infection was assessed by counting oocyst numbers of 10–15 guts dissected on day 15 after an infectious blood meal. For each experiment the average oocyst number per gut was 75 or higher.

**RNA Preparation and Subtraction Library Construction**—Total RNA was isolated with Trizol (Molecular Research Center) and polyadenylated RNA was prepared by chromatography on an oligo(dT) column. Polyadenylated RNA corresponding to 75 µg of total RNA of each sample was treated with DNase I (Invitrogen) to remove any remaining genomic DNA. The samples were then reverse-transcribed using the SMART cDNA synthesis primer (modified oligo(dT), Clontech), as described by the manufacturer. Equal amounts of cDNA from each sample was pooled as indicated in Fig. 1, and subtraction was carried out using the Clontech PCR Select cDNA subtraction kit (catalog number K1804-1). The subtraction products were cloned into pGEM-T easy vector (Promega) and transformed into *Escherichia coli* (DH5α) competent cells. Inserts were PCR amplified using T7 (5'-CTAATACGACTCACC-TATAGGGC-3') and SP6 (5'-GATTTAGGTGACACTATAG-3') primers and sequenced with a nested primer, PN1 (5'-TCGAGCGGCCGC-CCGGCAGGT-3') in an ABI377 sequencer.

**Clustering and Data Base Analysis**—Vector sequences were first removed computationally from all sequences followed by clustering of identical clones by internal BLAST analysis. Each cluster was searched for homology using the non-redundant data base and the Malaria Genetics and Genomics data bases at the National Center for Biotechnology Information (NCBI) using the BLASTN and BLASTX algorithm. In addition, sequences were also searched against the *Plasmodium* data base at www.Plasmdb.org, The Institute of Genomic Research (www.tigr.org), and Sanger (www.sanger.com). Sequences with significant BLASTX similarity ( $P(N) \leq 10^{-5}$ ) were grouped based on the predicted function of the homologous protein. Sequences with no significant BLASTX similarity were grouped based on their BLASTN similarity ( $P(N) \leq 10^{-10}$ ).

**Preparation of cDNA Blots**—After PCR amplification, subtracted and non-subtracted cDNAs were fractionated by electrophoresis on 1% agarose gels and transferred onto positively charged nylon membranes. A mosquito actin probe was generated by PCR with primers actin F (5'-TCAAGATGTGCGACGAAGAGG-3') and actin R (5'-CGAGAGATGTGGTGTGTTGTTCTC-3') and labeling with [<sup>32</sup>P]dCTP using random hexamers.

**Reverse Transcription-PCR Expression Analysis**—RNA was isolated from midguts dissected at various times after an infectious or a non-infected blood meal and from blood-stage parasites. Total RNA (1 µg) from each sample was reverse-transcribed with an oligo(dT) primer and PCR amplified (25 cycles) using the following gene-specific primers: PbRP.F (5'-GACTAACAAGAGCGGCAAGA-3') and PbRP.R (5'-GTACATAAAATCCCATTCCAT-3'); CS.F (5'-GTACCATTGTTAGTTGTAGCGTC-3') and CS.R (5'-CATCGCAAGTAATCTGTTG-3'); AMA1.F (5'-CCTTCAGGTAATGTCCAGT-3') and AMA1.R (5'-TTTCCAATC-ATCACGCA-3'); M184.F (5'-GCCTAAATGTGTTGCGAAGAGC-3') and

TABLE I  
Characteristics of the subtraction libraries

Singletons refer to ESTs found only once. Clusters with database similarities were based on BLASTN and/or BLASTX searches at NCBI, PlasmDB, TIGR, and Sanger with a probability cutoff of  $P(N) \leq 10^{-10}$  (BLASTN) or  $P(N) \leq 10^{-5}$  (BLASTX).

	Library	
	Mid-oocyst	Late-oocyst
ESTs sequenced and submitted	804	681
Singletons	623	334
Total number of clusters (unique sequences)	689 (100%)	448 (100%)
Clusters with database hits	367 (53%)	241 (55%)
BLASTX matches	191 (28%)	129 (29%)

M184.R (5'-CAGAAAGGGTGTATAATGATGG-3'); L1144.F (5'-AAGG-ATGGTTTTGACTGTTT-3') and L1144.R (5'-ATTTAATTTCTCCAAT-GGGT-3'); L1147.F (5'-TAAAAGGCGAGCAATAACAG-3') and L1147.R (5'-CATCAACAGATCCGACAC-3'); L457.F (5'-ATGCCTCTTTTATCT-GTCTA-3') and L457.R (5'-GATTACCAGCCGTATGTTGC-3'); L1085.F (5'-TATCATATTCGCGCTTTGC-3') and L1085.R (5'-TGTTGTCATAT-TTACATAATCCT-3'); L1096.F (5'-TGATGAAACAATCCAACTG-3') and L1096.R (5'-AATTTATCATGGCCCAACT-3'); L2451.F (5'-TGGA-GCAACTTGGTGTA-3') and L2451.R (5'-CCAGTATTGTTAGGGAAT-3'); L3593.F (5'-ATTCGGAAGTTTAATACCAA-3') and L3593.R (5'-TA-TGACCAGCTTTAGCAC-3'). Primers were designed such that their annealing temperature was 45–55 °C and the product size 100–250 bp. The PCR products were blotted and hybridized with a radioactive probe prepared from DNA amplified with the same primer using the cloned insert as the template.

**Indirect Immunofluorescence**—Guts were dissected on days 15 and 22, and salivary glands were dissected on day 25 after an infectious blood meal. Sporozoites obtained by gentle homogenization of guts or salivary glands were spotted on glass slides and fixed in ice-cold methanol. For double labeling, the fixed sporozoites were then incubated for 2 h in 5% bovine serum albumin in phosphate-buffered saline followed by incubation for 1 h with polyclonal antisera against the carboxyl cysteine-rich region of *P. berghei* MAEBL and a monoclonal antibody against CS (3D11, kindly provided by Dr. Victor Nussenzweig). Monoclonal antibody 28G2dc1 (kindly provided by Dr. Alan Thomas) that recognizes a highly conserved C-terminal region from several *Plasmodium* species was used to localize AMA1. The following secondary antibodies were used: rhodamine-conjugated goat anti-rabbit IgG (Sigma), fluorescein isothiocyanate-conjugated rabbit anti-mouse IgG (Sigma), and rhodamine-conjugated rabbit anti-mouse IgG (Sigma).

**Immuno-electron Microscopy**—*P. berghei* sporozoites isolated from salivary glands were fixed for 20 min at 4 °C with 1% formaldehyde and 0.2% glutaraldehyde in 0.1 M phosphate buffer, pH 7.4. Fixed samples were washed, dehydrated, and embedded in LR White resin (Poly-

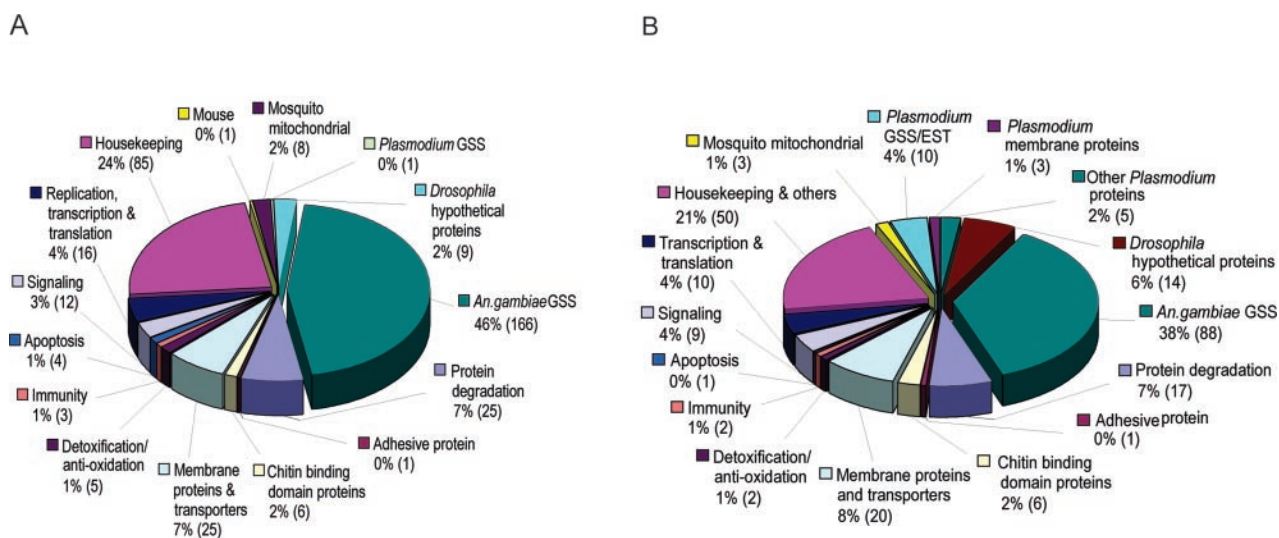


FIG. 2. Functional classification of ESTs based on BLASTX and BLASTN similarities. A total of 363 unique sequences from the mid-oocyst library and 245 unique sequences from the late-oocyst library had data base hits (Table I) and were grouped based on their predicted biological function. A, mid-oocyst library. B, late-oocyst library. Genomic Survey Sequence (GSS) is a sequence that has similarity with genomic DNA.

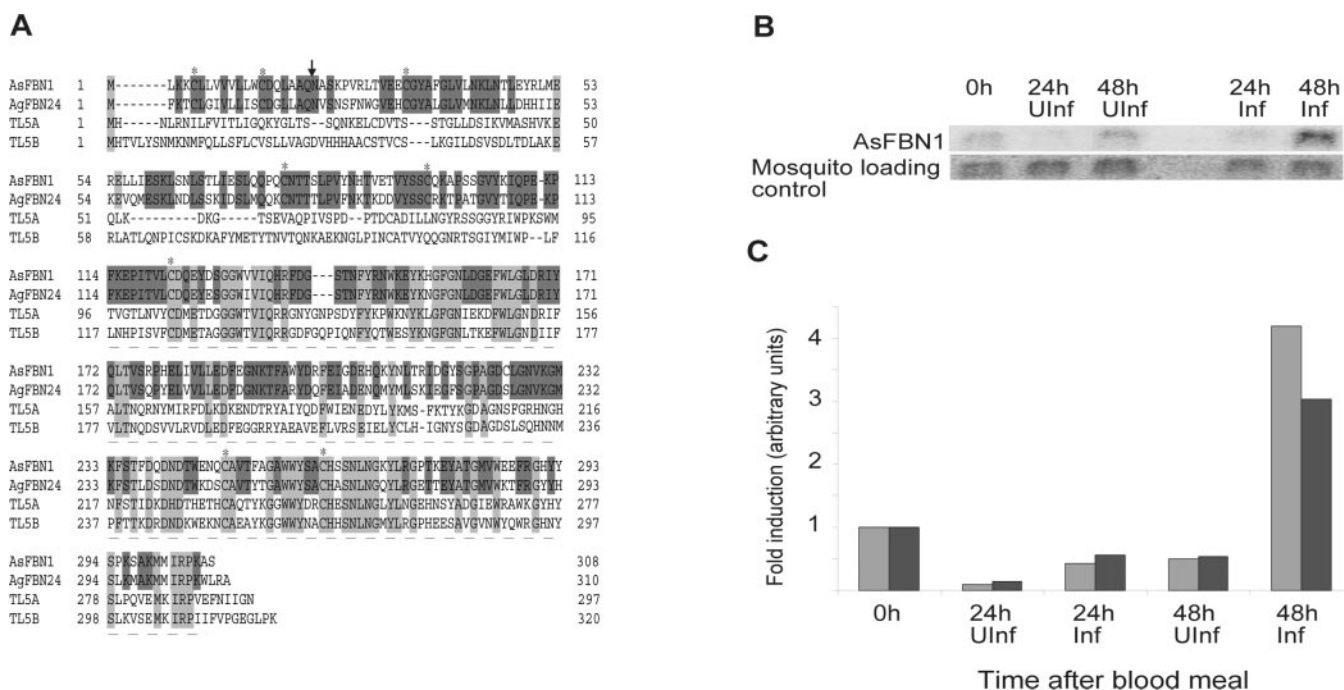


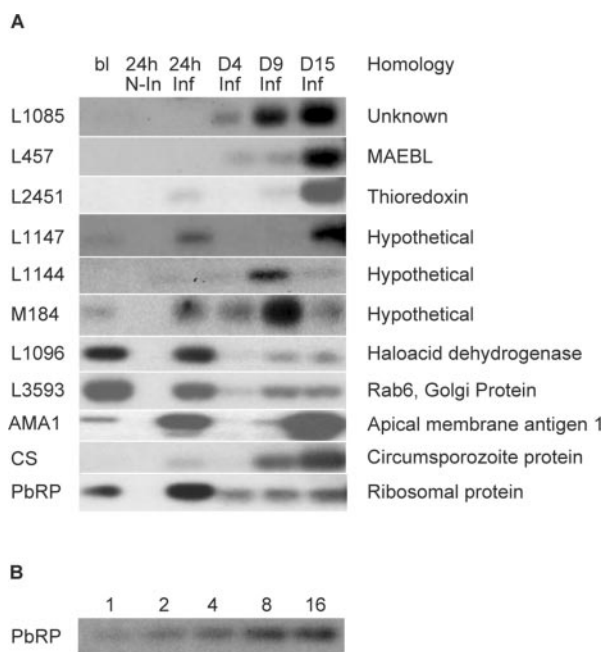
FIG. 3. Analysis of the AsFBN1. A, sequence comparison of the *A. stephensi* (As) and *A. gambiae* (Ag) fibrinogen domain protein genes with horseshoe crab tachylectins 5A and 5B. Amino acid residues that are conserved between all the four proteins are shaded light, and residues that conserved within the mosquito proteins are shaded dark. Arrow, putative signal peptide cleavage site; Dotted line, fibrinogen-like domain; \*, conserved cysteine residues. AgFBN24, *A. gambiae* fibrinogen domain protein 24; TL5A, tachylectin 5A; TL5B, tachylectin 5B. B, AsFBN1 is induced by *Plasmodium* in the mosquito midgut. Northern blot analysis shows that the gene is induced at 48 h after an infected but not after a non-infected, blood meal. The numbers refer to the time in hours after a blood meal at which the midguts were dissected. 0h, sugar-fed mosquitoes; UInf, uninfected blood meal; Inf, infected blood meal. C, Northern blots such as the one shown in B were quantified using phosphorimaging. The signals were normalized to a mosquito loading control (mitochondrial rRNA; Ref. 29) and are plotted relative to the value of sugar-fed controls. Gray and black bars represent expression profiles from two independent sets of RNA samples.

sciences). Thin sections were blocked for 30 min at room temperature in phosphate-buffered saline containing 5% w/v nonfat dry milk and 0.01% v/v Tween 20 (PBTM). Grids were then incubated with anti-AMA 1 primary antibody (diluted 1:50 in PBTM for 16 h at 4 °C. Control grids were incubated with normal mouse serum in PBTM at the same dilution. After washing grids were incubated for 1 h in 15-nm gold-conjugated goat anti-mouse IgG (Amersham Biosciences) diluted 1:20 in phosphate-buffered saline containing 1% w/v bovine serum albumin and 0.01% v/v Tween 20 (PBTB), rinsed with PBTB, and fixed with 2.5% glutaraldehyde to stabilize the gold particles. Samples were stained with uranyl acetate and lead citrate and then examined in a Zeiss CEM902 electron microscope (Oberkochen, Germany).

**Rapid Amplification of cDNA Ends**—Full-length cDNA for *A. stephensi* fibrinogen domain protein 1 was isolated by rapid amplification of cDNA ends using primers M2015F (5'-GTTCCGAAATTGGAGAC-GAGC-3') and M2015R (5'-CACCATGCTCCAGCAAACG-3'). Sequence analysis was performed at the Expaty Molecular Biology Server (www.Expaty.ch). Sequence alignment was done using ClustalW software.

RESULTS

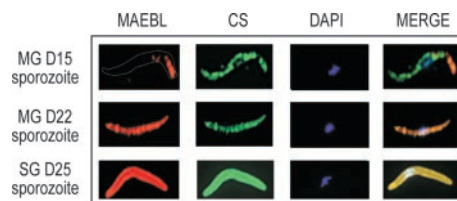
**Construction and Characterization of the Subtraction Libraries**—To construct libraries enriched for *P. berghei* and *A. stephensi* transcripts expressed during mid-oocyst (day 4–8) and



**FIG. 4. Temporal expression patterns of selected ESTs from the subtraction libraries.** *A*, reverse transcription-PCR products of samples indicated above each lane were amplified for 25 cycles with primers for genes specified to the left of each panel, fractionated by gel electrophoresis, and transferred onto nylon membranes. Radioactive probes generated by PCR amplification of the corresponding cloned EST were hybridized to the membranes. *Plasmodium* ribosomal protein gene (*PbRP*; accession number BF295783) was used as a loading control, and the circumsporozoite protein gene (accession number M14135) was used as a positive control. Similar profiles were obtained in at least four experiments with at least two independently isolated RNA samples. *M*, clones from the mid-oocyst library; *L*, clones from the late-oocyst library; *bl*, RNA from blood stage parasites; 24 h N-in, RNA from guts dissected 24 h after a non-infected blood meal; 24 h Inf, RNA from guts dissected 24 h after an infected blood meal; D4, D9, and D15, RNA from guts dissected at the indicated number of days after an infected blood meal; *Unknown*, EST having homology to the *Plasmodium* genomic data base that does not have a predicted open reading frame; *Hypothetical*, EST having homology to the *Plasmodium* genomic data base that has a predicted open reading frame. *B*, linearity of signal response. Increasing amounts of template (relative amounts indicated at the top of each lane) were amplified for 25 cycles and analyzed as in panel *A*. Note that the strength of the signal is proportional to the amount of RNA template.

late-oocyst (day 10–14) stages of parasite development, gut cDNAs of non-infected mosquitoes and of blood-stage parasites were subtracted from the cDNAs of guts harboring developing oocysts (Fig. 1). This procedure was expected to enrich for parasite transcripts expressed during oocyst differentiation and for mosquito transcripts induced in the gut by the parasite. Any common transcripts (e.g. genes expressed in both infected and non-infected guts or genes expressed in both blood-stage and mosquito-stage parasites) are depleted by the procedure. Subtraction efficiency was assessed by hybridizing a mosquito actin probe to Southern blots of cDNAs obtained before and after the subtraction procedure. The signal was much weaker in the “after subtraction” lanes, indicating effective depletion of common genes by the subtraction protocol (see Supplemental Fig. 1 and data not shown).

A total of 1485 clones from both libraries was sequenced and analyzed (Table I). The 804 sequences from the mid-oocyst library could be grouped into 689 unique clusters, of which 623 were singletons (sequences isolated only once). Of the 689 sequences, 367 (53%) had data base hits and were grouped according to their presumed function (Fig. 2A). Similarly, 681 sequences from the late-oocyst library could be grouped into 448 unique clusters, of which 334 were singletons. Of the 448



**FIG. 5. Differential localization of *P. berghei* MAEBL in oocyst and salivary gland sporozoites.** Anti-MAEBL and anti-CS antibodies were used to localize the corresponding proteins on sporozoites. The source of sporozoites was as follows: *MG D15*, midguts on day 15 after infection; *MG D22*, midguts on day 22 after infection; *SG D25*, salivary glands on day 25 after infection. The sporozoites were counterstained with 4,6-diamidino-2-phenylindole (*DAPI*) to show the location of the nucleus. The *MG D15* pattern was consistently observed in three independent preparations. No reproducible differences of protein distribution between day 22 *MG* and day 25 *SG* sporozoites were observed.

sequences, 241 (55%) had data base hits and were classified based on their predicted function (Fig. 2B). The high proportion of singletons in both libraries suggests that the subtraction procedure resulted in efficient normalization (removal of redundant sequences). As construction of the library involves digestion of cDNAs with a restriction enzyme, some of the clusters might represent non-overlapping regions of the same gene.

A number of mosquito genes involved in maintenance of the redox state of the cell were identified. These include ESTs with homology to thioredoxin, peroxiredoxin-1, metallothionein, glutathione transferase, and glutathione peroxidase (Fig. 2, *A* and *B*, and Supplemental Table I). In addition, several components of the electron transport chain, which is a major source of reactive oxygen species, were also identified (Supplemental Table I). Interestingly, transcripts from putative anti-apoptotic genes such as inhibitor of apoptosis 1 (*IAP1*) and Bax inhibitor 1 were found in mid- and late-oocyst libraries, respectively (Supplemental Table I). This is in contrast to the ookinete/early oocyst library, which contained ESTs corresponding to apoptosis-promoting rather than anti-apoptotic genes (1).

*A. stephensi Fibrinogen Domain Protein 1 (AsFBN1) Is Induced by the Parasite*—Two overlapping ESTs (accession numbers CB602443 and CB602444) with similarity to fibrinogen domain proteins were isolated from the mid-oocyst library. There is evidence that in mosquitoes and horseshoe crabs, members of this gene family are induced in response to foreign organisms (8, 9). The full-length cDNA was cloned by 5'- and 3'-rapid amplification of cDNA ends. The resulting 921-bp sequence predicts a protein of 306 amino acids, including a 19-amino acid secretion signal sequence at the N terminus and a 200-amino acid fibrinogen domain at the C terminus (Fig. 3A). *AsFBN1* is related to horseshoe crab tachylectins 5A and 5B (25% identity, 26% similarity), suggesting that it could act as a broad specificity immune molecule by binding carbohydrate molecules on glycoproteins (8). *AsFBN1* and the *Anopheles gambiae* homologue are 72% identical, including six conserved cysteine residues in the secreted portion of the protein and two in the signal peptide (Fig. 3A). Analysis of *AsFBN1* gene expression revealed that it is induced at 48 h by a parasite-containing blood meal but not by a non-infected blood meal (Fig. 3, *B* and *C*). This suggests that induction is likely to be triggered by the presence of the differentiating oocyst.

*Temporal Patterns of Plasmodium Gene Expression*—Reverse transcription-PCR was used to assess changes in parasite mRNA abundance as a function of time after mosquito infection (Fig 4A). In initial experiments, the assay was run by resolving PCR products on agarose gels followed by ethidium bromide staining. Quantification by this protocol was not reliable because many of the target mRNAs were rare in the samples and

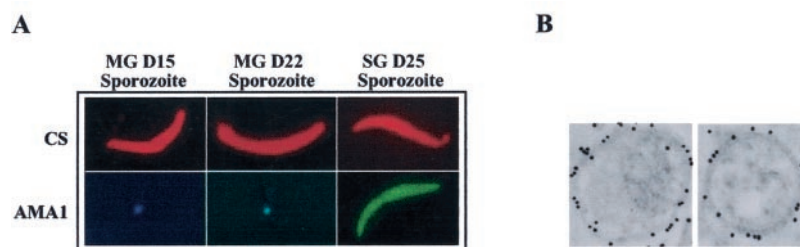


FIG. 6. Regulation of AMA1 expression in sporozoites. A, anti-AMA1 and anti-CS antibodies were used to localize the corresponding proteins on sporozoites. The source of sporozoites was as follows: *MG D15*, midguts on day 15 after infection; *MG D22*, midguts on day 22 after infection; *SG D25*, salivary glands on day 25 after infection. Sporozoites with no AMA1 expression were stained with 4,6-diamidino-2-phenylindole, which labels the nucleus, to show the presence of the parasite. B, cross-sectional view of salivary gland sporozoites labeled with anti-AMA1 antibody. Immuno-electron microscopy shows the surface localization of AMA1 on salivary gland-invaded sporozoites. No staining was observed with midgut sporozoites (data not shown).

required large number of PCR amplification cycles for detection. The radioactive hybridization approach (*cf.* "Materials and Methods") proved to be much more sensitive, linear, and reproducible. Fig. 4B shows that under the conditions used the signal is proportional to RNA concentration. To partially compensate for the presence of large numbers of blood-stage parasites in the 24h *inf* sample, more template was used for this than later time points (Fig. 4A, *PbRP*). Note that all blood-stage parasites eventually die and are eliminated from the mosquito guts. CS, which is expressed during oocyst development (5), was used as a positive control (Fig. 4A).

To analyze temporal expression patterns, eight ESTs that had hits to the *Plasmodium* data bases were chosen (Fig. 4A). Of these, six (L1085, L2451, L1147, L1144, M184, L1096) are novel, and two (L457 and L3593) had been isolated previously. All but L1096 and L3593 appear to be induced specifically during the mosquito stages and have an expression profile consistent with the library (mid-oocyst or late-oocyst) from which they were derived (Fig. 4A). L1085, L2451, and L1147 start being expressed at different times but all have peak expression during the terminal stages of oocyst differentiation (day 15). These gene products could function in final stages of oocyst differentiation and/or be stored for use during sporozoite invasion of the mosquito salivary gland or vertebrate liver. L2451 encodes a thioredoxin domain protein. Thioredoxins play a crucial role in defense against free radicals (10, 11) and may be involved in protection against reactive oxygen species generated by the mosquito immune system. L1147 may also be expressed in ookinetes (24 h).

M184 and L1144 had peak expression at mid-oocyst stages (day 9) followed by a sharp decrease of transcript abundance, indicating that these gene products may be required for intermediate steps in sporozoite differentiation. L1096 is expressed at all stages of parasite development. The predicted protein has homology to halo-acid dehalogenase family proteins. This family includes L-2-haloacid dehalogenase (12) and epoxide hydrolases (13) that function in halo-carbon metabolism and detoxification of epoxide substrates, respectively. This family also includes certain phosphatases.

The expression of two ESTs (L457 and L3593) with similarity to previously identified *Plasmodium* genes were also analyzed (Fig. 4A). L3593 encodes a homologue of Rab6, a protein associated with the Golgi apparatus of blood-stage *P. falciparum* (14). This is of interest because there has been no morphological evidence of a functional secretory pathway in the mosquito stages of parasite development. The finding that L3593 is expressed during sporozoite differentiation raises the possibility that the oocyst employs a Golgi-like pathway for protein trafficking. MAEBL (L457) and AMA1, two proteins originally thought to be expressed only in blood stages, were found to be strongly activated at late stages of

oocyst differentiation (Fig. 4A). Further studies on these two genes are described below.

**Differential Localization of MAEBL during Sporozoite Maturation**—Two non-overlapping ESTs (accession numbers CB603466 and CB603325) corresponding to the C-terminal region of the erythrocyte-binding protein MAEBL were recovered from the late-oocyst library. MAEBL was initially identified as a blood-stage gene whose protein localizes to the apical organelles of merozoites (15, 16). Recently, MAEBL was reported to be expressed also in sporozoites (17, 18). We found that the gene is sharply induced during late stages of oocyst maturation (Fig. 4A). Indirect immunofluorescence revealed that MAEBL subcellular localization changes as the sporozoite matures. In immature sporozoites (day 15) the protein was restricted to the apical end, presumably confined to secretory organelles (Fig. 5). In contrast, CS, which is another major sporozoite surface protein, was uniformly distributed in the same sporozoites. In mature midgut sporozoites (day 22) and in sporozoites that had invaded the salivary glands both MAEBL and CS were uniformly distributed on the sporozoite surface. The deployment of the protein to the sporozoite surface upon maturation is consistent with the observation that MAEBL is essential for salivary gland invasion (19). The mechanism that regulates MAEBL translocation from the apical end to the surface is unknown. In blood stage parasites, MAEBL has been shown to transit through the ER-Golgi secretory pathway (20). The sporozoites could employ a similar secretory system. Reverse transcription-PCR performed using primers representing both N- and C-terminal regions of the coding sequence indicates that MAEBL is also expressed in liver-stage parasites (results not shown). Thus, MAEBL may be required both for salivary gland and liver invasion, as is the case for CS (21, 22).

**AMA1 Expression Is Post-transcriptionally Regulated during Oocyst Development**—The expression of MAEBL during oocyst development prompted the analysis of other blood-stage parasite surface proteins, including MSP1 and AMA1. MSP1 expression seems to be restricted to blood forms of the parasite (not shown). However, AMA1 is strongly induced during the final stages of oocyst maturation (Fig. 4A). In contrast to CS and MAEBL, AMA1 protein expression could not be detected in sporozoites isolated from either day 15 or day 22 oocysts (Fig. 6A). Interestingly, protein can only be detected after sporozoite invasion of salivary glands (day 25, Fig. 6A). Moreover, immuno-electron microscopy revealed that AMA1 is localized on the surface of salivary gland sporozoites (Fig. 6B). In agreement with the immunofluorescence results (Fig. 6A), no immunogold staining was observed with midgut sporozoites (not shown). Thus, the presence of the transcript (Fig. 4A) but not of the protein (Fig. 6A) in midgut sporozoites indicates that AMA1 is translationally regulated during sporozoite maturation. The presence of the AMA1 protein only on the surface of

salivary gland sporozoites (Fig. 6B) suggests that AMA1 might be required for invasion of hepatocytes.

#### DISCUSSION

In this study we report the identification of candidate genes of both mosquito and parasite origin that are induced during oocyst development on the mosquito midgut. The lack of suitable techniques to purify oocysts from mosquito midgut tissues and the small proportion of parasite RNA in total RNA from infected midguts have seriously hampered the molecular analysis of *Plasmodium* development in the mosquito. The recent demonstration of sexual stage parasite development in an *in vitro* system (23) has yet to be adapted for large scale cultures. The present study has partially overcome these limitations by the construction of stage-specific subtraction libraries. By reverse transcription-PCR we demonstrate that these libraries indeed contain genes induced during oocyst development on the mosquito midgut. Ongoing microarray analysis of clones from the two subtraction libraries promises to provide a more detailed survey of mosquito and parasite genes differentially regulated during oocyst differentiation.

A large proportion of the *Plasmodium* life cycle in the mosquito is dedicated to the differentiation of the oocyst, resulting in the formation of thousands of sporozoites. Parasite genes expressed at specific stages of the differentiation process are likely to participate in these events. For instance, M184 and L1144 may be required for intermediate steps in sporozoite differentiation because they have peak expression at mid-oocyst stages (day 9, Fig. 4) and L1085, L2451, and L1147 may be required in terminal stages, as indicated by their activation late during oocyst differentiation (day 15, Fig. 4). It is well established that the mosquito responds to the presence of the invading parasite (ookinete) by mounting an innate immune response (7). Less well known are the responses against the developing oocyst. This study provides the first insights on the response of the mosquito to this stage of the parasite (Supplemental Table 1, A and B). Some of the transcripts identified include genes with homology to anti-oxidants and key anti-apoptotic proteins such as inhibitor of apoptosis 1 and Bax inhibitor 1.

Annotation of the recently completed *A. gambiae* genome revealed the presence of a large family of fibrinogen-related proteins (FBN; Refs. 24 and 25). Some FBN family genes are induced either immediately after a bacterial challenge or 24 h after a *Plasmodium*-infected blood meal (9). In contrast, *AsFBN1* is induced on day 2 after an infectious blood meal, at which time ookinete-to-oocyst transformation is completed. It is likely that this induction is triggered by the developing oocyst. Because fibrinogen can interact with thrombospondin I (26), the possibility arises that *AsFBN1* recognizes thrombospondin I domains of *Plasmodium* surface proteins, such as circumsporozoite protein and thrombospondin-related adhesive protein.

Surprisingly, MAEBL localization was found to be developmentally regulated in sporozoites, from apical localization (presumably in secretory organelles) in immature sporozoites to uniform distribution on the entire surface in mature sporozoites. After invasion of the salivary glands, MAEBL remains on the sporozoite surface, indicating that the protein might be required for vertebrate stage (liver) infection. The mechanism that regulates MAEBL translocation from the apical end to the surface is unknown.

AMA1 has been shown in various *Plasmodium* species to be required for invasion of erythrocytes (27) and as such is a candidate for a blood-stage vaccine in both rodent and primate models (28, 29). AMA1, which was thought to be a surface

protein specific to the asexual stages of the parasite life cycle, was found to be strongly induced in day 15 oocysts. The mRNA is translationally regulated, since the protein appears on the sporozoite surface only after invasion of the salivary gland. This raises the intriguing possibility that AMA1 is required during invasion of the vertebrate liver.

The ESTs generated by this and the accompanying study (1) will contribute to the annotation of the recently completed *Plasmodium yoelii* and *A. gambiae* genomes (30, 31). For instance L1085 has homology to the *Plasmodium* nucleotide data base but has no predicted open reading frame. Many of the genes identified in this study are novel, and elucidation of their function will provide a stronger foundation for understanding parasite development and *Plasmodium* interactions with the mosquito.

*Acknowledgments*—We are indebted to Marilyn Donnelly-Doman for expert technical assistance. We thank Dr. Victor Nussenzweig for the 3D11 and Dr. Alan Thomas for the 28G2dc1 antibodies.

#### REFERENCES

- Abraham, E. G., Islam, S., Srinivasan, P., Ghosh, A. K., Valenzuela, J., Ribeiro, J. M., Kafatos, F. C., Dimopoulos, G. & Jacobs-Lorena, M. (2003) *J. Biol. Chem.* **279**, 5573–5580
- Breman, J. G., Egan A. & Keusch G. T. (2001) *Am. J. Trop. Med. Hyg.* **64**, Suppl. 1, 4–7
- Ghosh, A., Edwards M. J. & Jacobs-Lorena, M. (2000) *Parasitol. Today* **16**, 196–201
- Sinden, R. E. (1998) in *Malaria: Parasite Biology, Pathogenesis, and Protection* (Sherman, I. W., ed) pp. 25–49, American Society for Microbiology, Washington, D. C.
- Menard, R., Sultan, A. A., Cortes, C., Altszuler, R., van Dijk, M. R., Janse, C. J., Water, A. P., Nussenzweig, R. S. & Nussenzweig, V. (1997) *Nature* **385**, 336–340
- Claudianos, C., Dessens, J. T., Trueman, H. E., Arai, M., Mendoza, J., Butcher, G. A., Crompton, T. & Sinden, R. E. (2002) *Mol. Microbiol.* **45**, 1473–1484
- Dimopoulos, G., Richman, A., Muller, H. M. & Kafatos, F. C. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 11508–11513
- Gokudan, S., Muta, T., Tsuda, R., Koori, K., Kawahara, T., Seki, N., Mizunoe, Y., Wai, S. N., Iwanaga, S. & Kawabata, S. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 10086–10091
- Dimopoulos, G., Christophides, G. K., Meister, S., Schultz, J., White, K. P., Barillas-Mury, C. & Kafatos, F. C. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99**, 8814–8819
- Rahlfis, S., Schirmer, R. H. & Becker, K. (2002) *Cell. Mol. Life Sci.* **59**, 1024–1041
- Aslund, F. & Beckwith, J. (1999) *Cell* **96**, 751–753
- Tsang, J. S. & Sam, L. (1999) *J. Bacteriol.* **181**, 6003–6009
- Argiriadi, M. A., Morisseau, C., Hammock, B. D. & Christianson, D. W. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 10637–10642
- de Castro, F. A., Ward, G. E., Jambou, R., Attal, G., Mayau, V., Jaureguiberry, G., Braun-Breton, C., Chakrabarti, D. & Langsley, G. (1996) *Mol. Biochem. Parasitol.* **80**, 77–88
- Kappe, S. H. I., Curley, G. P., Noe, A. R., Dalton, J. P. & Adams, J. H. (1997) *Mol. Biochem. Parasitol.* **89**, 137–148
- Kappe, S. H. I., Noe, A. R., Fraser, T. S., Blair, P. L. & Adams, J. H. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 1230–1235
- Kappe, S. H. I., Gardner, M. J., Brown, S. M., Ross, J., Matuschewski, K., Ribeiro, J. M., Adams, J. H., Quackenbush, J., Cho, J., Carucci, D. J., Hoffman, S. L. & Nussenzweig, V. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98**, 9895–9900
- Ghai, M., Dutta, S., Hall, T., Freilich, D. & Ockenhouse, C. F. (2002) *Mol. Biochem. Parasitol.* **123**, 35–45
- Kariu, T., Yuda, M., Yano, K. & Chinzei, Y. (2002) *J. Exp. Med.* **195**, 1317–1323
- Noe, A. R., Fishkind, D. J. & Adams, J. H. (2000) *Mol. Biochem. Parasitol.* **108**, 169–185
- Warburg, A., Touray, M., Krettli, A. U. & Miller, L. H. (1992) *Exp. Parasitol.* **75**, 303–307
- Cerami, C., Frevert, U., Sinnis, P., Takacs, B., Clavijo, P., Santos, M. J. & Nussenzweig, V. (1992) *Cell* **70**, 1021–1033
- Al-Olayan, E. M., Beetsma, A. L., Butcher, G. A., Sinden, R. E. & Hurd, H. (2002) *Science* **295**, 677–679
- Christophides, G. K., Zdobnov, E., Barillas-Mury, C., Birney, E., Blandin, S., Blass, C., Brey, P. T., Collins, F. C., Danielli, A., Dimopoulos, G., Hetru, C., Hoa, N. T., Hoffman, J. A., Kanzok, S. M., Letunic, I., et al. (2002) *Science* **298**, 159–165
- Zdobnov, E. M., von Mering, C., Letunic, I., Torrents, D., Suyama, M., Copley, R. R., Christophides, G. K., Thomasova, D., Holt, R. A., Subramanian, G. M., Mueller, H., Dimopoulos, G., Law, J. H., Wells, M. A., Birney, E., et al. (2002) *Science* **298**, 149–159
- Bonnefoy, A., Hantgan, R., Legrand, C. & Frojmovic, M. M. (2001) *J. Biol. Chem.* **276**, 5605–5612
- Triglia, T., Healer, J., Caruana, S. R., Hodder, A. N., Anders, R. F., Crabb, B. S. & Cowman, A. F. (2000) *Mol. Microbiol.* **38**, 706–718

28. Narum, D. L., Ogun, S. A., Thomas, A. W. & Holder, A. A. (2000) *Infect. Immun.* **68**, 2899–2906
29. Stowers, A. W., Kennedy, M. C., Keegan, B. P., Saul, A., Long, C. A. & Miller, L. H. (2002) *Infect. Immun.* **70**, 6961–6967
30. Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., Carlton, J. M., Pain, A., Nelson, K. E., Bowman, S., Paulson, I. T., James, K., Eisen, J. A., Rutherford, K., Salzberg, S. L., *et al.* (2002) *Nature* **419**, 498–511
31. Holt, R. A., Subramanian, G. M., Halpern, A., Sutton, G. G., Charlab, R., Nusskern, D. R., Wincker, P., Clark, A. G., Ribeiro, J. M., Wides, R., Salzberg, S. L., Loftus, B., Yandell, M., Majoros, W. H., Rusch, D. B., *et al.* (2002) *Science* **298**, 129–149