

# Serial Analysis of Gene Expression in *Plasmodium falciparum* Reveals the Global Expression Profile of Erythrocytic Stages and the Presence of Anti-Sense Transcripts in the Malarial Parasite

Swati Patankar,<sup>\*†</sup> Anusha Munasinghe,<sup>\*†</sup> Azadeh Shoaibi,<sup>‡</sup>  
Leda M. Cummings,<sup>‡</sup> and Dyann F. Wirth<sup>\*§</sup>

<sup>\*</sup>Department of Immunology and Infectious Diseases, Harvard School of Public Health, Harvard University, Boston, Massachusetts 02115; and <sup>‡</sup>The Institute for Genomic Research, Rockville, Maryland 20850

Submitted December 14, 2000; Revised May 24, 2001; Accepted July 20, 2001  
Monitoring Editor: David Botstein

Serial analysis of gene expression (SAGE) was applied to the malarial parasite *Plasmodium falciparum* to characterize the comprehensive transcriptional profile of erythrocytic stages. A SAGE library of ~8335 tags representing 4866 different genes was generated from 3D7 strain parasites. Basic local alignment search tool analysis of high abundance SAGE tags revealed that a majority (88%) corresponded to 3D7 sequence, and despite the low complexity of the genome, 70% of these highly abundant tags matched unique loci. Characterization of these suggested the major metabolic pathways that are used by the organism under normal culture conditions. Furthermore several tags expressed at high abundance (30% of tags matching to unique loci of the 3D7 genome) were derived from previously uncharacterized open reading frames, demonstrating the use of SAGE in genome annotation. The open platform “profiling” nature of SAGE also led to the important discovery of a novel transcriptional phenomenon in the malarial pathogen: a significant number of highly abundant tags that were derived from annotated genes (17%) corresponded to antisense transcripts. These SAGE data were validated by two independent means, strand specific reverse transcription-polymerase chain reaction and Northern analysis, where antisense messages were detected in both asexual and sexual stages. This finding has implications for transcriptional regulation of *Plasmodium* gene expression.

## INTRODUCTION

Malaria, an infectious disease caused by the protozoan parasite *Plasmodium falciparum*, affects 300–500 million people globally each year (WHO, 1997). Increasing drug-resistance in the parasite and insecticide-resistance in the *Anopheles* vector have exacerbated this substantial public health problem. Against this backdrop, effective strategies to combat the disease require a fundamental knowledge of the basic biology of *Plasmodium* to develop new pharmaceuticals and vaccines that target the parasite.

Most studies of *Plasmodium* biology have been directed at single genes thought to be important for pathogenesis. With

the advent of genomic technologies, however, new approaches to combat the disease, such as identifying entire repertoires of transcripts expressed under different conditions, have now become available. Genomic approaches were initiated with the sequencing of the *P. falciparum* (3D7 strain) genome, a collaborative project, undertaken by the Malaria Genome Consortium that is already close to completion (Butler, 1997; O’Brien, 1997; Craig *et al.*, 1999). Chromosomes 2 and 3 have been fully sequenced (Gardner *et al.*, 1998; Bowman *et al.*, 1999), whereas 80 to 90% of the estimated 6000 open reading frames (ORFs) in the 3D7 genome are now available as raw sequence data. The next challenge is to use this vast amount of data to study the functional relevance of various genes. For example, it is now possible to identify genes that are transcribed in different stages of the parasite’s development and also genes that are induced or repressed in response to various stimuli such as immune or drug pressure. For this reason, whole genome expression analyses with the use of high-density microarrays (Hayward

<sup>§</sup> Corresponding author. E-mail address: dfwirth@hsph.harvard.edu.

<sup>†</sup> These authors contributed equally to this work.

Abbreviations used: BLAST, basic local alignment search tool; SAGE, serial analysis of gene expression; RT-PCR, reverse transcription-polymerase chain reaction.

*et al.*, 2000) and serial analysis of gene expression (SAGE) (Munasinghe *et al.*, 2000) have been developed for *P. falciparum*. These new approaches will complement each other to generate data for the *Plasmodium* research community. Genome sequence will expedite the microarray and SAGE analyses; conversely, open platform profiling techniques such as SAGE will help the Malaria Genome Project with annotation of previously uncharacterized ORFs and with novel gene discovery.

SAGE provides a sensitive and highly quantitative description of the transcript profile of a given cell type (Velculescu *et al.*, 1995, 1997). The SAGE technology samples short sequence tags (14 bases) from mRNA transcripts in the population of interest. These tags contain sufficient sequence information to identify, by basic local alignment search tool (BLAST) analysis, the transcript from which each tag was derived (Munasinghe *et al.*, 2000). The frequency of each tag in the SAGE library is an accurate estimate of the abundance of its corresponding mRNA transcript. Numerous groups have used this technique successfully and described the SAGE protocol in detail (Velculescu *et al.*, 1995, 1997; Madden *et al.*, 1997; Polyak *et al.*, 1997; Matsumura *et al.*, 1999; Virlon *et al.*, 1999).

In this report, we show that SAGE can be used to study gene expression of the asexual stages of *P. falciparum*. Asexual parasites express many virulence factors and are the targets of antimalarials such as chloroquine; hence, an in-depth understanding of their transcriptional profiles will set the stage for future experiments addressing responses to immune or drug pressure.

SAGE was successfully applied to erythrocytic stage parasites (3D7 strain) of *P. falciparum* at baseline culturing conditions, and a SAGE library of ~8000 tags was generated. A majority of these corresponded to unique parasite genes, as demonstrated by BLAST analysis of a subset of tags. The SAGE data were validated by Northern and reverse transcription-polymerase chain reaction (RT-PCR) analysis of genes predicted to be highly expressed based on tag counts. BLAST analysis of highly abundant tags also provided insight into networks of major metabolic pathways that are used by the parasite under normal culture conditions. These pathways include mitochondrial, glucose, polyamine, and deoxy-D-xylulose 5-phosphate (DOXP) metabolism. Finally, SAGE also revealed the presence of antisense transcription in the malarial parasite, a phenomenon that has been previously missed by other methods of transcriptional analysis. These SAGE data were also validated by two independent methods, RT-PCR and Northern analysis; here antisense transcripts for genes expressed in asexual as well as sexual stage parasites were found. In summary, SAGE in *Plasmodium* has revealed many facets of the basic functioning of the parasite in culture, and it sets the stage for future comparisons of the transcriptional responses of *P. falciparum* to different stimuli.

## MATERIALS AND METHODS

### Parasite Culture and RNA Extraction

3D7 strain parasites were maintained under standard culturing conditions (Trager and Jensen, 1976) with modifications as previously described (Munasinghe *et al.*, 2000). Polyadenylated RNA was harvested from cultures at 8% parasitemia (1% rings, 5% trophozo-

ites, and 2% schizonts) and used in the SAGE procedure as previously described (Munasinghe *et al.*, 2000).

### Data Analysis

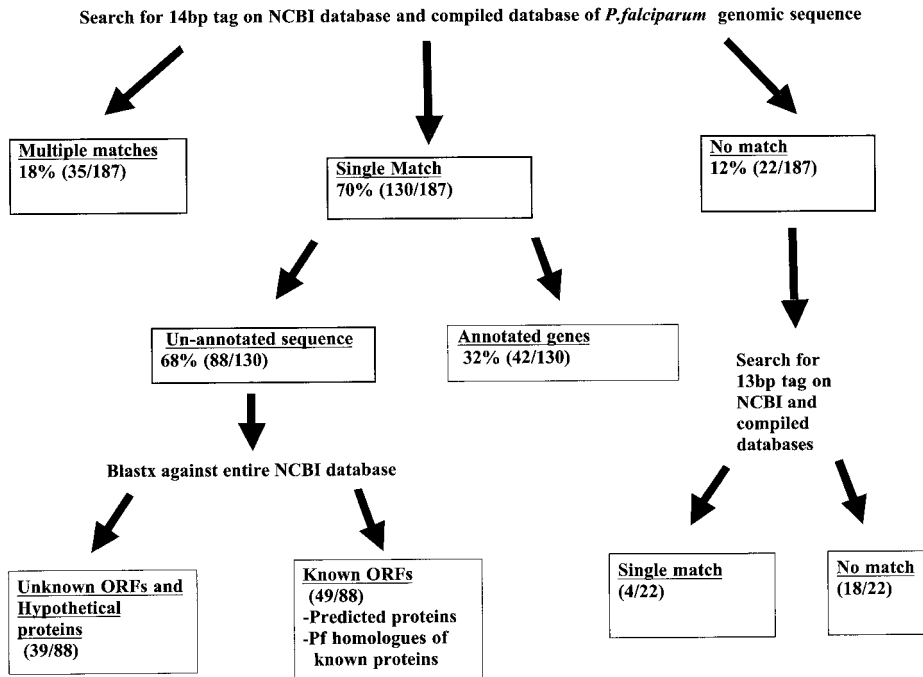
SAGE tags from 3D7 asexual stages were analyzed with the use of the SAGE software (Johns Hopkins University, Baltimore, MD, and Genzyme, Cambridge, MA), which extracts 14-bp tag counts from sequence files. To assign gene identity to each tag, the 3D7 experimental tag list was matched against a *P. falciparum* tag database. This database was created by extracting 14-bp tags from *P. falciparum* sequence deposited in GenBank (as of July 13, 2000), as well as from a compiled database of recently deposited 3D7 genome sequence (obtained from The Institute For Genomic Research (Rockville, MD), Sanger (Cambridge, UK), and Stanford (Palo Alto, CA), sequencing centers, and compiled at the University of Pennsylvania (Philadelphia, PA) as of July 26, 2000; kindly provided by Drs. Jessica Kissinger and David Roos). Because the *P. falciparum* genome is not fully annotated, all potential SAGE tags from both sense and antisense strands were extracted (i.e., tags were extracted from each database in the "genomic mode" rather than the "cDNA mode").

The software output files are organized in such a way that matches to a single locus, matches to multiple loci, and no matches to database sequences can be readily determined. For genomic sequence that is annotated, it is possible to assign gene identification to each tag in the manner outlined above; however, most of the available *P. falciparum* genome sequence is not annotated. Therefore, the 187 most abundant tags (abundance level of >4 hits) were characterized by manual BLASTx analysis; see flow chart in Figure 1. Here, for tags derived from unannotated reads, a 500-1000-bp sequence surrounding the tag was translated in all six reading frames and compared with the entire National Center for Biotechnology Information protein database. Fourteen bp tags that failed to match either database were analyzed with the use of only the first 13 bp of the tag sequence in the manner outlined above.

### RT-PCR

RT-PCR was performed with the use of the 3' RACE kit (Invitrogen, Carlsbad, CA) according to the manufacturer's protocol. First-strand cDNA synthesis was primed with oligo(dT)<sub>18</sub> (0.5 µg of mRNA was used per reaction), whereas PCR was performed with the gene specific primers described below. All primers anneal within coding regions of the genes and result in 350–790-bp PCR products. Calmodulin (sense): 5' GTCCATCACCATCAATAT-CAGC 3' calmodulin (antisense): 5' CTAAGGAGTTAGGAACGGT-CATG 3' *msp-3* (sense): 5' TTTTGTGTCTGGAAC-GCCTCCTCC 3' *msp-3* (antisense): 5' GCTTCGGAAGATGCTGAA-AAAGCTGC 3' *pfg27/25* (sense): 5' TCTTGTCTGTTATGATA-CGCTTC 3' *pfg27/25* (antisense): 5' GTACAAAAGGATAGT-GCCAAGCCCC 3' *rap-1* (sense): 5' CTTTGAAGAAATCTCTGAT-TTCAGC 3' *rap-1* (antisense): 5' GCTTTAGAAGGTGTCTGT-TCATATC 3' *hsp86* (sense): 5' CCGAATTACTCCGATTCCAA-ACCTC 3' *hsp86* (antisense): 5' CTCTTCCATTTTAGAAT-CGGTTGC 3' PCR reactions were carried out according to the manufacturer's protocol (3' RACE kit; Invitrogen). Initial denaturation of the template occurred at 94°C for 3 min. Amplification was performed for five cycles at 94°C for 45 s, 52–54°C for 45 s, and 72°C for 45 s, followed by 21–26 cycles of identical amplification where the annealing temperature was increased to 55–57°C. Finally, extension of partial PCR products was completed at 72°C for 6 min.

Strand-specific RT-PCR used 1 µg of total RNA per reaction and was performed with the express purpose of distinguishing sense RNA from antisense RNA (Yu *et al.*, 1995). RT-PCR was performed with the use of the 3' RACE kit (Invitrogen); however, first-strand cDNA was primed with gene-specific primers that hybridize to either sense or antisense messages, rather than with an oligo(dT)<sub>18</sub> primer. The gene-specific primers are identical to the primers listed above. A 10th of the cDNA sample was PCR amplified, with the



**Figure 1.** BLASTx analysis of highly abundant SAGE tags. SAGE tags from the 3D7 library were analyzed with the use of the SAGE software (as described in MATERIALS AND METHODS). SAGE tags at an abundance level of five counts or greater were further analyzed as depicted in the flow chart. The percentage of single matches, no matches, and multiple matches to the databases are indicated. Numbers in brackets correspond to the proportion of tag counts in each group. Single matches were divided into tags that matched annotated and unannotated sequence reads. Tags in the latter group were characterized by BLASTx. Fourteen bp tags that failed to match either database were analyzed with the use of only the first 13 bp of the tag sequence.

same set of gene-specific primers and amplification conditions described above.

PCR products were electrophoresed on 1.2% agarose gels. All resultant PCR products were cloned into the pCRII vector with the use of the TA cloning kit (Invitrogen) and sequenced to confirm the identity of the amplified cDNA.

### Northern Blots

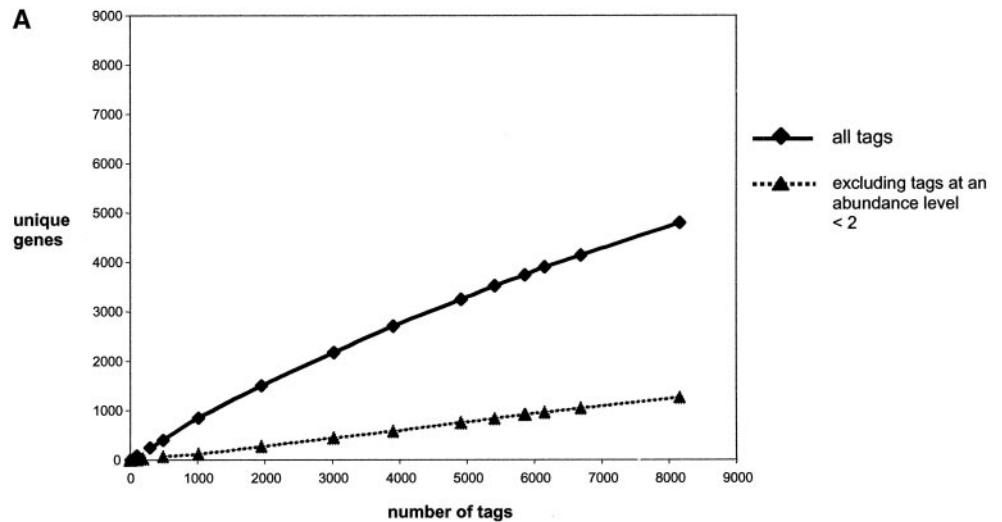
Northern analysis was performed according to standard protocols. Briefly, 1  $\mu$ g of mRNA from 3D7 cultures was gel electrophoresed, blotted onto BA85 nitrocellulose membranes (Schleicher & Schuell, Keene, NH), and probed with gene-specific DNA probes. All probes (calmodulin, *msp-3*, *rap-1*, and *pfg27/25*) were derived from the RT-PCR products described in the previous section. DNA probes were radiolabeled with [ $\alpha$ - $^{32}$ P]dATP with the use of random hexanucleotides and the Klenow fragment of DNA polymerase. Blots were visualized by autoradiography.

For strand-specific Northern blots, 20  $\mu$ g of total RNA were used per blot as described above. Synthetic RNAs corresponding to a sense or antisense fragment (~300 bp) of either calmodulin or *msp-3* were used as probes for strand-specific Northern analysis. The synthetic RNAs were generated in the following manner. Briefly, RT-PCR products of calmodulin and *msp-3* cDNAs (~300-bp fragment; see previous section) were cloned into pBluescript. The orientations of calmodulin and *msp-3* genes within pBluescript were determined by sequencing. Each plasmid (pBluescript-*msp-3* and pBluescript-calmodulin) was linearized with either *Bam*HI or *Xho*I. These plasmids were then used as DNA templates for in vitro transcription reactions with the use of T3 or T7 RNA polymerase to generate synthetic sense or antisense RNA fragments for each gene. Plasmids digested with *Bam*HI were incubated with T7 RNA polymerase (after standard protocols) to produce antisense RNAs for both genes. Similarly, plasmids digested with *Xho*I were incubated with T3 polymerase to produce sense RNAs. Strand-specific RNA probes were also obtained under the same conditions in the presence of [ $\alpha$ - $^{32}$ P]ATP.

Quantitative Northern analysis was carried out for calmodulin and *msp-3* to determine whether the ratio of their transcripts was comparable with that determined by SAGE. Northern blots and gene-specific DNA probes were prepared as described above. Known amounts of synthetic 300-bp RNA fragments (in the sense orientation) from each gene were run alongside the mRNA sample as markers for quantification. Blots were exposed to x-ray film (Kodak XO-MAT) such that the intensity of the signal was within the linear range of the film. Signal intensities for each of the transcripts in the mRNA sample were converted to molar amounts by reference to those of the synthetic RNAs. Signal intensities were measured by scanning the x-ray film into Adobe Photoshop (Adobe Systems, Mountain View, CA) and by using NIH Image software to quantify bands by pixel density.

### *Plasmodium gallinaceum* Gamete Preparation and RNA Isolation

*P. gallinaceum* parasites were propagated in White Leghorn chickens by serial injection into wing veins. At parasitemia of 50–70%, blood was withdrawn by heart puncture. Gametogenesis was induced as described previously (Goonewardene *et al.*, 1993), with the inclusion of xanthurenic acid (Sigma, St. Louis, MO) at a final concentration of 50  $\mu$ M in the exflagellation buffer. Gametes and zygotes were purified, also as described previously (Goonewardene *et al.*, 1993), and  $1 \times 10^7$  cells were incubated at 25°C in Medium 199 (Invitrogen) and harvested for analysis at 0, 24, and 48 h after isolation. Total RNA was isolated with the use of Tri reagent (Molecular Research Center, Cincinnati, OH) according to the manufacturer's protocol. Total RNA obtained from  $1 \times 10^7$  parasites was used for each RT-PCR reaction. Strand-specific RT-PCR was performed as described previously with the following primers: *pgs28* (sense): 5' CATCTAG-CATAGTCAGCACAAGGTTTATTTG 3' and *pgs28* (antisense): 5' CAAACGAAGATTATTTAGTCAAAC 3'.



**Figure 2.** SAGE 3D7 library analysis. (A) Cumulative total gene representation within the 3D7 SAGE library. Ascertained tags (from the 3D7 library) at increasing increments were plotted against the number of unique genes from which the tag subsets were derived. The solid line corresponds to all ascertained tags. The dotted line corresponds to the number of unique genes represented by tags at an abundance level of two counts or greater (some tags at an abundance level of one may be derived from sequencing errors). (B) SAGE abundance classes. SAGE tags (8335) are divided into abundance classes (percent abundance [first column] refers to the abundance of tags as a percentage of all 8335 tags in the SAGE library). The number of unique tags matching an entry in the *P. falciparum* National Center for Biotechnology Information database is listed per abundance class (last column), and the percentage of hits within the abundance class is given in brackets. ND, not determined.

Percent abundance	Total number of tags	Total number of genes	Matches to <i>P. falciparum</i> sequence databases
>1.1	96	1	1 (100%)
0.2-0.6	441	14	10 (71%)
0.05-0.2	1307	172	98 (57%)
0.025-0.05	2696	1092	ND
<0.025	3795	3587	ND
<b>Total</b>	<b>8335</b>	<b>4866</b>	<b>ND</b>

## RESULTS

### 3D7 SAGE Tag Library from Asexual Blood Stage Parasites

A total of 8335 SAGE tags was analyzed from the asexual blood stages of *P. falciparum*, 3D7 strain. A preliminary analysis showed that these 8335 tags corresponded to 4798 unique genes (Figure 2A). Of these, 1254 genes were present at an abundance of two hits (or counts) or greater. The 537 tags expressed at abundance levels  $\geq 20$  tags (percent abundance of 0.2) accounted for 6.4% of the total collection of tags but only 0.3% (15) of the total number of unique genes. As expected, these abundance groups had the highest percentage of matches to GenBank entries (Figure 2B), implying that many highly expressed messages have been readily cloned and studied. The lower abundance tags (abundance of <20 tags) accounted for 93.6% of the total collection of tags, and represented a vast majority of the unique genes expressed in the parasite. Moreover, these tags gave many fewer matches to GenBank; hence, SAGE in *P. falciparum* will aid in the discovery of novel malarial genes.

### BLAST Analysis of SAGE Tags

To assess whether 14-bp tags could uniquely identify genes in the highly A-T-rich *Plasmodium* genome, these SAGE tags were searched against 3D7 genome sequence. We decided that for an accurate estimate of the "tag to gene" mapping in *Plasmodium*, all available sequence data, both cDNA and genomic, would provide the most complete picture. Sequencing of the *P. falciparum* genome is close to completion; however, much of the newly available *P. falciparum* sequence data has yet to be annotated. Therefore, the 187 most abundant SAGE tags were analyzed in a more rigorous manner by BLASTx analysis. A schematic of the BLAST analysis is shown in Figure 1. This analysis revealed that a majority of the SAGE tags (88%) corresponded to *P. falciparum* genome sequence. Most of the tags that match to single loci (70%) lie within known genes; hence, SAGE tags can be used to uniquely identify genes in *Plasmodium*. The other 30% of tags that match single sites correspond to unknown genes and hypothetical open reading frames. Thus, SAGE data reveal not only predicted ORFs that are expressed but also previously uncharacterized transcripts; hence, SAGE in *Plasmodium* has the capacity to assist in annotation of the genome.



**Table 1.** Highly expressed genes in the 3D7 library

Tag	% Abundance	Gene description
TCAGGCGTTA	1.18	Mitochondrial 6-kb product
GTGGTGGTGC	0.70	No match to database
GAGCAAGCAG	0.58	Unknown protein
GAAGTCGAAA	0.44	5.8S ribosomal RNA
ATTTGAAGCA	0.36	Rhop H3
CTAAAGCACC	0.28	ras-related nuclear protein
TTGAAGCTGA	0.26	Heat shock protein-70
AACGACAAGA	0.25	Pfg27/25
CCAAATGATG	0.25	Polyubiquitin
TACAGCTGCT	0.18	Merozoite surface protein-3
GGAAATAAAG	0.17	Tartarate-resistant acid phosphatase
TGAGTCAAAC	0.17	No match to database
GGCACAACCTA	0.17	Thioredoxin
TAAAGTTTG	0.16	Rhoptry-associated protein-1
TTGTTTCATA	0.09	Rifin
CGAGTAAAAG	0.09	1-Deoxy-D-xylulose 5-phosphate (DOXP) synthase
CCAATAAGG	0.07	ATPase subunit B
TGATGGCTTG	0.07	Ornithine decarboxylase
TTCCGAACCT	0.07	Triose phosphate isomerase
AGAGATCCGC	0.06	Ubiquinol cytochrome-c reductase like protein
GATACTCTTG	0.06	26S proteasome beta subunit

Tag represents the 10-bp SAGE tag sequence adjacent to the NlaIII site. Abundance is listed as a percentage of all 8335 tags in the SAGE library.

Approximately 10% of the 187 most abundant SAGE tags did not match parasite sequence. We expect this number to decrease as the genome project nears completion. The percentage of SAGE tags that gave multiple matches within the *P. falciparum* genome was also calculated and found to be 18%. In the present study, the 35 tags that matched more than one loci were further investigated; of these tags, 21 (60%) matched two or three genes, whereas 14 (40%) matched greater than three genes. The latter set of tag sequences was of lower complexity in general. Northern blot analysis should help resolve whether tags that match multiple genes indeed represent multiple transcripts.

### Abundant Transcripts Expressed in *P. falciparum* Grown in Culture

The BLAST analysis described above enabled us to assign genes to highly abundant SAGE tags; examples of these are listed in Table 1. This analysis provided a snapshot of the major transcripts expressed by the parasite. A complete picture of metabolic pathways used by *P. falciparum* growing in culture will incorporate protein expression and stability; nevertheless, BLAST analysis of abundant SAGE tags provides the first global description of genes and hence, metabolic pathways that might be transcriptionally regulated at the level of expression. The most abundant transcripts were grouped into functional categories to reveal the transcriptional profile of 3D7 parasites grown in culture (Figure 3). Many tags represented housekeeping functions carried out

by all prokaryotic and eukaryotic cells (transcription, translation, chaperones, cytoskeleton), whereas some functional classes were highly specific for the unique life cycle of *Plasmodium* (membrane-associated proteins involved in invasion, DOXP pathway).

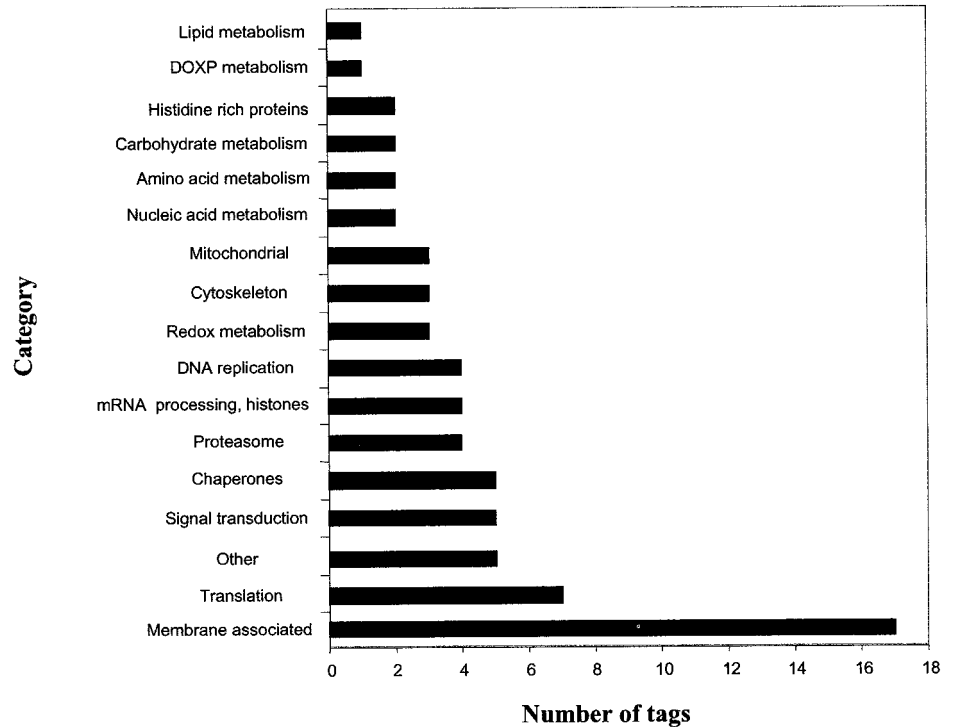
Interestingly, many of the highly abundant messages (5.3%) appear to be transcribed from the 6-kb mitochondrial genome, and another 2.1% (thioredoxin, vacuolar ATPase subunit B, ATPase transporter, ubiquinol cytochrome *c* reductase-like protein) are probably involved in oxidative metabolism. Therefore, a significant proportion of abundant transcripts encodes proteins that play a role in oxidative metabolism.

Stage-specific transcripts are highly represented in the list of abundant messages, reflecting the different developmental stages present in the culture. For example, mRNAs encoding cell surface proteins involved in merozoite invasion (Cowman *et al.*, 2000) comprise 8% of the most abundant transcripts. These include merozoite surface proteins 3 and 4 (MSP-3 and -4), rhoptry-associated protein-1 (RAP-1), and merozoite capping protein. Tags corresponding to serine repeat antigen, a soluble protein that is associated with the parasitophorous vacuole, were found at high abundance (0.32%). Surprisingly, a tag representing the gametocyte surface antigen Pfg27/25, shown to be essential for gametogenesis (Lobo *et al.*, 1999), was also present at high abundance (0.25%) in this SAGE library derived from asexual parasites.

Abundant SAGE tags represented major metabolic pathways of the malarial parasite. Because asexual blood stages of *Plasmodium* do not store energy reserves in the form of glycogen or lipids, glucose taken up from plasma is the primary source of energy (Sherman, 1991). Therefore, glucose metabolism is a prominent aspect of intracellular growth and not unexpectedly, proteins required for glucose metabolism were represented among the abundant tags (aldolase, phosphoenolpyruvate carboxykinase, and triose-phosphate isomerase).

Although lipids are not used as a major source of energy by *P. falciparum*, there is a significant increase in levels of phospholipids, diacylglycerol, and triacylglycerol, within the red blood cell upon merozoite invasion (Vial and Ancelin, 1998). This increase in the total lipid content is associated with a biosynthetic requirement for lipids during formation of the membranes surrounding the parasite (the parasitophorous vacuolar membrane and the tubovesicular membrane). *N*-Myristoyl transferase, an enzyme that plays a role in the formation of lipoproteins, was found among the 187 most abundant tags; however, tags representing proteins involved in lipid biosynthesis were not present.

Intraerythrocytic *P. falciparum* parasites are capable of de novo synthesis of pyrimidines from precursor molecules (Walsh and Sherman, 1968), with a requirement for para-aminobenzoic acid and folate cofactors. Unlike their hosts, malarial parasites do not use exogenous folate cofactors, but instead synthesize these de novo (Scheibel and Sherman, 1988). SAGE data revealed tags corresponding to ribonucleotide reductase, an enzyme of the pyrimidine biosynthetic pathway, and dihydrofolate synthase, an enzyme of the folate pathway. Polyamine biosynthetic enzymes were also represented among the SAGE tags (ornithine decarboxylase and ornithine aminotransferase).



**Figure 3.** Categories of highly expressed genes in the 3D7 control population. Highly abundant tags (tag counts of  $>0.05\%$  of the total number of tags) were examined for their matches in *P. falciparum* databases (see analysis described in Figures 1 and 2) and categorized by putative functions. The number of genes in each category is depicted.

The unique intracellular niche of malarial parasites results in the expression of many parasite-specific metabolic pathways. For example, growth of the asexual parasites within red blood cells is accompanied by degradation of hemoglobin and the subsequent detoxification of heme by-products (Foley and Tilley, 1998; Krogstad and De, 1998; Rosenthal and Meshnick, 1998). Tags representing proteins implicated in the detoxification of heme (histidine-rich proteins I and II, glutathione reductase) were found at high abundance in the SAGE library. Surprisingly, the plasmepsin and falcipain proteases that play a role in hemoglobin degradation were not found in the list of highly expressed genes. This may be due to the fact that their transcription occurs at an earlier stage in the parasite life cycle than the trophozoite stage, which was the predominant stage in the study population. Alternatively, these transcripts may be present at a very low abundance.

Finally, SAGE data revealed the expression of mRNA encoding DOXP synthase at high levels (0.09%). The DOXP pathway was recently identified as a parasite-specific metabolic pathway important for isoprenoid biosynthesis (Jomaa *et al.*, 1999). Because this pathway is localized in the apicoplast, a plant-derived organelle of *Plasmodium*, DOXP metabolism provides a novel target for antimalarial drug development.

#### Validation of SAGE Data

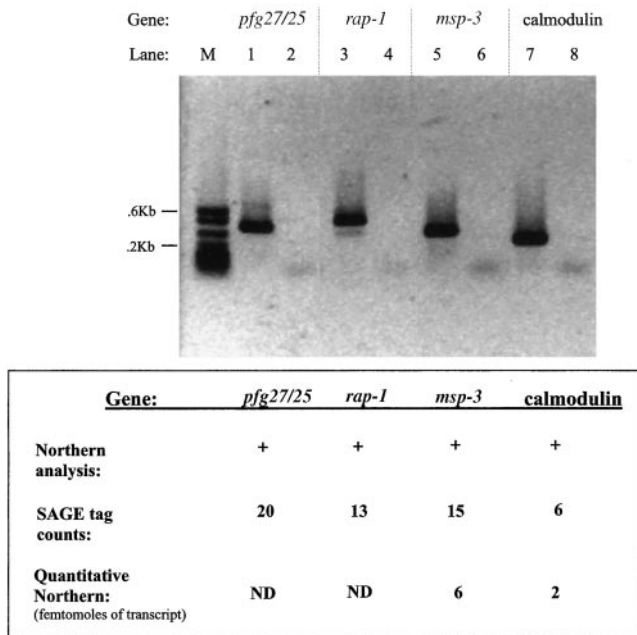
To confirm the expression data in asexual-stage parasites as determined by SAGE, RT-PCR and Northern analysis of several genes with highly abundant SAGE tag counts (calmodulin, *msp-3*, *rap-1*, and *pfg27/25*; Figure 4) were performed. *Pfg27/25* represents a gametocyte-specific antigen,

whereas the other three are predicted to be expressed in asexual stages. Because the SAGE library was derived from a culture that contained no detectable gametocytes, *pfg27/25* was specifically chosen for RT-PCR and Northern analysis. RT-PCR products for all four genes were generated from asexual-stage mRNA (Figure 4A). These were cloned, sequenced, and found to correspond to the expected gene. Transcripts at the predicted length for all four genes were also detected by Northern blotting (our unpublished results; Figure 4B). The presence of *pfg27/25* transcripts in the asexual stages of *P. falciparum* has been reported in another genome-wide expression analysis with the use of microarrays (Hayward *et al.*, 2000).

For a more quantitative estimate of gene expression, quantitative Northern analysis of two highly expressed genes (*msp-3* and calmodulin) was performed (our unpublished results; Figure 4B). Here, the molar ratio of *msp-3* to calmodulin was  $\sim 3:1$ , which is similar to the ratio of their SAGE tag counts (Figure 4B). Hence, SAGE tag data appear to correlate well with relative levels of mRNA within the cells.

#### Antisense Transcripts

A surprising observation of SAGE in *P. falciparum* was the large proportion of tags corresponding to antisense transcripts. Unlike microarrays, SAGE is able to detect antisense transcription because the orientation of the SAGE tag on the mRNA can be readily determined. A SAGE tag consists of the 4-bp recognition sequence (CATG) of the restriction enzyme *NlaIII* (this enzyme defines the position of each tag in an mRNA transcript) and 10 bp of adjacent sequence in the direction of the 3' poly(A) tail of the RNA molecule. Among 45 annotated genes whose 5' and 3' ends are clearly



**Figure 4.** Validation of SAGE data by RT-PCR and Northern analysis (A) RT-PCR of genes represented in the 3D7 SAGE library. RT-PCR products generated by specific primers for *pfg27/25* (lanes 1 and 2), *rap-1* (lanes 3 and 4), *msp-3* (lanes 5 and 6), and calmodulin (lanes 7 and 8) are shown. RT minus samples were also PCR amplified as negative controls (lanes 2, 4, 6, and 8). pBR322 *MspI* digest (lane M). (B) Summary of expression data. For Northern analysis 1  $\mu$ g of mRNA from 3D7 cultures was gel electrophoresed, blotted onto a nylon membrane, and probed with gene specific  $^{32}$ P-labeled DNA probes. + indicates that a specific signal corresponding to each transcript was detected in the Northern analysis. Tag counts from the 3D7 library (8335 tags) for all four genes are listed. Quantitative Northern analysis was carried out for calmodulin and *msp-3*. Northern blots and gene-specific probes were prepared as described above. Known amounts of in vitro synthesized RNA fragments were run alongside the mRNA sample as markers for quantification. Signal intensities for each of the transcripts in the mRNA sample were converted to molar amounts by reference to those of the synthetic RNAs.

denoted, 17% of the tags consisted of a CATG and the 3' adjacent 10 bp, in the direction of the 5' end of the transcript, on the noncoding strand of cDNA. This result was unexpected; hence, we wanted independent confirmation of the SAGE data. This was accomplished by strand-specific RT-PCR analysis of asexual as well as sexual blood stages, and strand-specific Northern analysis in erythrocytic stage parasites.

We confirmed the presence of antisense transcripts from erythrocytic stages by strand-specific RT-PCR analysis of the three genes calmodulin, *rap-1*, and *msp-3*, and subsequent sequencing of the RT-PCR products to establish gene identity. Based on SAGE data, we expected all three transcripts to be present in both the sense and antisense orientations, a prediction that was confirmed by RT-PCR (Figure 5A, lanes 1–12) and sequence analysis. On the other hand, a PCR product for *hsp-86* was only detected for sense RNA (Figure

5A, lane 15), consistent with the absence of an antisense SAGE tag for this gene. Importantly, control experiments that excluded reverse transcriptase (lanes 2, 4, 6, 8, 10, 12, 14, and 16) indicated a lack of contaminating genomic DNA, showing that the PCR products obtained during strand-specific RT-PCR were indeed derived from RNA. These data validate the antisense transcripts predicted by SAGE.

The presence of antisense transcripts was also confirmed by strand-specific Northern analysis for calmodulin and *msp-3*. To control for the specificity of the strand-specific RNA probes, synthetic RNA corresponding to the sense or antisense strands of each gene was included in the experiment. This synthetic RNA consisted of short transcripts (250–300 bp within the coding regions) derived from each gene in vitro. Figure 5B shows that strand-specific probes can specifically detect synthetic antisense RNA (lanes 1 and 2 for calmodulin; lanes 7 and 8 for *msp-3*) or synthetic sense RNA (lanes 4 and 5 for calmodulin; lanes 10 and 11 for *msp-3*). With the use of these strand-specific probes, total RNA isolated from asexual stage parasites was shown to contain both antisense (~1 kb) and sense (~1.2 kb) transcripts for both calmodulin (lanes 3 and 6) and *msp-3* (~2 kb) (lanes 9 and 12). Therefore, as confirmed by two independent techniques, the presence of antisense tags in the SAGE library reflects antisense transcription in asexual stages of the malarial parasite.

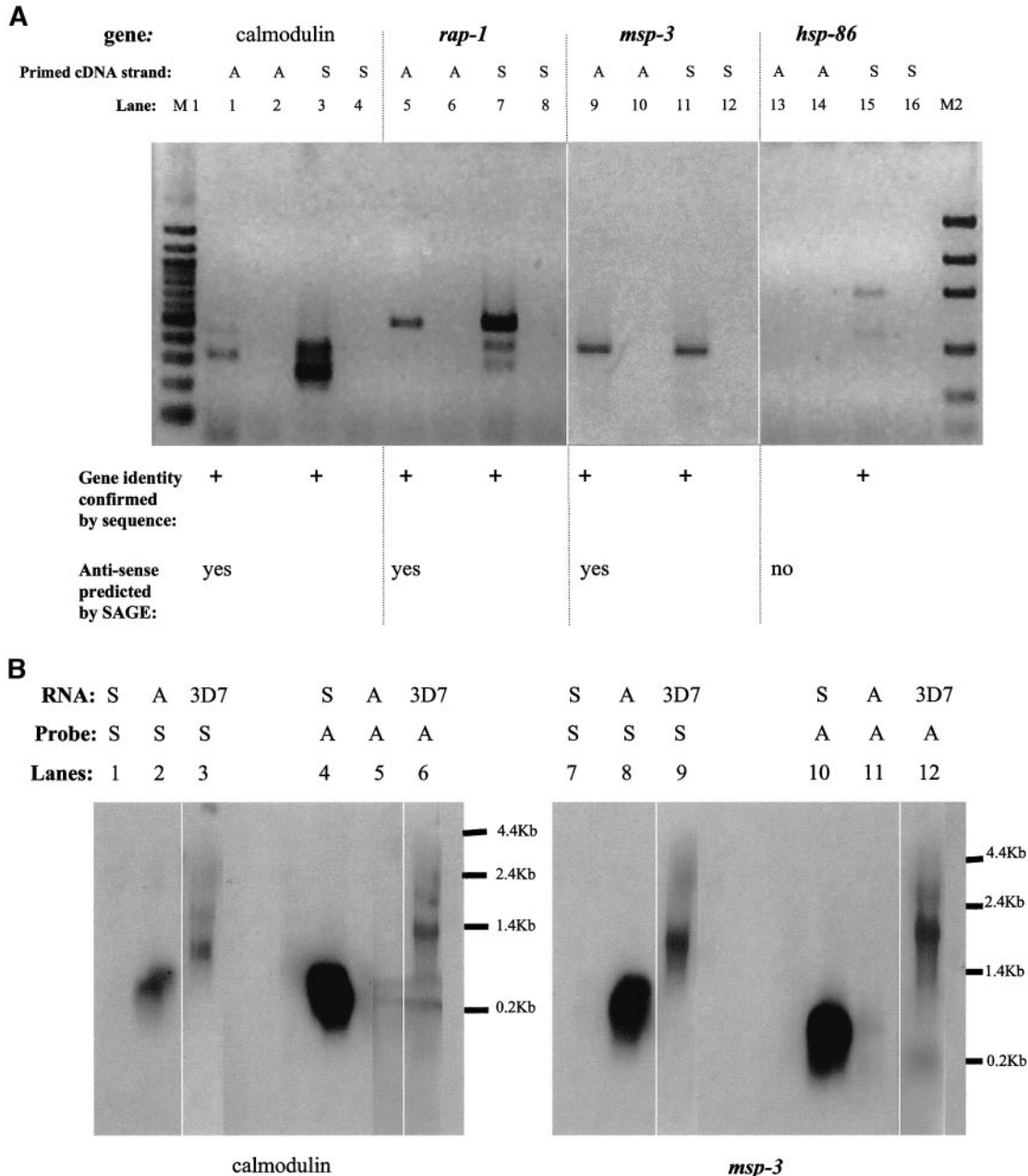
We wondered whether genes expressed in other stages of the *Plasmodium* life cycle also exhibited antisense transcription. To address this, the sexual stages (zygotes and ookinetes) of the chicken malarial parasite *P. gallinaceum* were tested for the presence of antisense RNAs. *Pgs28* is a major surface antigen of *P. gallinaceum* sexual stages (Duffy *et al.*, 1993), and transcription of the *pgs28* gene has been studied previously. Strand-specific RT-PCR of total RNA from zygotes (0 h) and mature ookinetes (48 h) showed that the *pgs28* gene expressed both sense and antisense transcripts (Figure 6) at different stages of in vitro development (lanes 1, 5, and 9 show antisense PCR product).

## DISCUSSION

This report demonstrates the application of SAGE in *P. falciparum*. Despite the low complexity of the genome, SAGE tags as short as 14 bp can uniquely identify a majority of genes in *P. falciparum*. This observation has been exploited to study transcription in the asexual stages of the parasite, resulting in new insights into the biology of the pathogen. First, we provide a description of the transcriptional profile of the 3D7 strain of *P. falciparum* that builds upon the extensive data generated by the Malaria Genome Project. Second, the major metabolic pathways present in blood stage parasites are delineated; modulation of these pathways in response to stimuli such as drug and immune pressure can now be studied. Finally, this report shows that *Plasmodium* parasites express antisense RNAs at multiple stages during the developmental cycle, a finding that has implications for transcriptional regulation of *Plasmodium* gene expression.

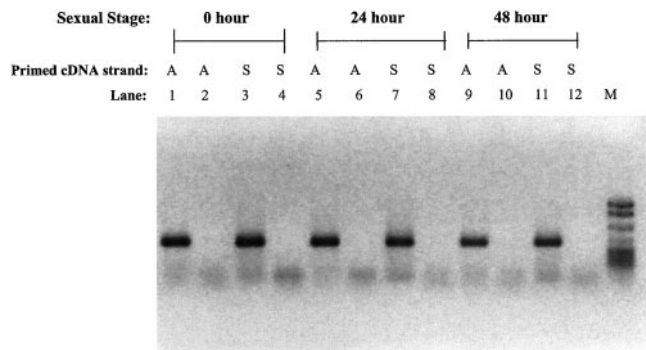
### Analysis of SAGE Tags

Of the tags that matched to single loci, 70% matched to known genes, whereas 30% matched to unknown genes or



**Figure 5.** Validation of antisense SAGE data by strand-specific RT-PCR and strand-specific Northern analysis in asexual stage parasites. (A) Strand-specific RT-PCR analysis. First-strand cDNA from asexual stages was synthesized with a primer that specifically hybridizes to either antisense message (A) for calmodulin (lanes 1 and 2), *rap-1* (lanes 5 and 6), *msp-3* (lanes 9 and 10), and *hsp-86* (lanes 13 and 14); or a primer that binds to sense message (S) for calmodulin (lanes 3 and 4), *rap-1* (lanes 7 and 8), *msp-3* (lanes 11 and 12), and *hsp-86* (lanes 15 and 16). RT-PCR was performed on these cDNA samples with the use of the same primer pair for calmodulin (lanes 1–4), *rap-1* (lanes 5–8), *msp-3* (lanes 9–12), and *hsp-86* (lanes 13–16). Products in lanes 1, 3, 5, 7, 9, 11, and 15 were cloned to confirm gene identity. + indicates that the gel product corresponded to the specific gene under investigation. In some cases, nonspecific PCR products were also generated, in addition to the specific product for the particular gene under study. RT minus samples were also PCR amplified as negative controls (lanes 2, 4, 6, 8, 10, 12, 14, and 16). 100-bp ladder (NEB), lane M1; low DNA mass ladder (Invitrogen), lane M2. Genes for which antisense transcripts were found in the SAGE library are indicated by yes or no (last row). Products for *msp-3* (lanes 9–12) were generated in a separate experiment. (B) Strand-specific Northern analysis. Synthetic sense (S) RNA fragments (lanes 1 and 4 for calmodulin; lanes 7 and 10 for *msp-3*), synthetic antisense (A) RNA fragments (lanes 2 and 5 for calmodulin; lanes 8 and 11 for *msp-3*), or 20  $\mu$ g of total RNA from 3D7 cultures (lanes 3, 6, 9, and 12) were gel electrophoresed. Blots were transferred onto nitrocellulose membranes and probed with  $^{32}$ P-labeled sense (S) RNA probes (calmodulin, lanes 1–3; *msp-3*, lanes 7–9) or antisense (A) RNA probes (calmodulin, lanes 4–6; *msp-3*, lanes 10–12). Northern blots were exposed to x-ray film; optimal exposure times for signals from synthetic RNA fragments were 5–10 s, whereas those for total RNA were 5–10 min. Adobe Photoshop was used to make a composite image of these x-ray films.





**Figure 6.** Strand-specific RT-PCR analysis of *pgs28* in sexual-stage parasites. First-strand cDNA from sexual stages was synthesized with a primer that specifically hybridizes to either antisense message (A) for *pgs28* (lanes 1, 2, 5, 6, 9, and 10), or a primer that binds to sense message (S) for *pgs28* (lanes 3, 4, 7, 8, 11, and 12). cDNA was synthesized from total RNA obtained from purified gametes and zygotes immediately (lanes 1–4), 24 h (lanes 5–8), and 48 h (lanes 9–12) after isolation. RT-PCR was performed on all samples with the use of the same primer pair. RT minus samples were also PCR amplified as negative controls (lanes 2, 4, 6, 8, 10, and 12). pBR322 *MspI* digest (lane M).

hypothetical ORFs. This distribution is in stark contrast to genome sequencing data, where 60% of the putative ORFs were of unknown function, whereas 40% were genes encoding proteins of known functions (Gardner *et al.*, 1998). This discrepancy could be explained by the fact that the asexual blood stages are more amenable to cultivation and experimental manipulation in the laboratory than other stages; hence, many of the transcripts expressed in these stages at high abundance have been previously studied and are of known functions. It is also likely that a majority of the transcripts expressed during laboratory culture of asexual blood stages encode proteins that serve housekeeping functions conserved within organisms widely separated on the phylogenetic tree. The genes of unknown function identified by the Malaria Genome Project may turn out to be of importance in host–parasite interactions and disease; however, under culturing conditions only relatively few may be expressed at high levels. Alternatively, the higher percentage of uncharacterized, putative ORFs in the sequence data might be due to overprediction of genes. Because SAGE data reveals genes that are actually expressed in asexual stage parasites, identification of tags that correspond to unknown genes and hypothetical ORFs will be of tremendous use in annotation of the *P. falciparum* genome.

Some tags (10%) did not match to the *Plasmodium* databases. Because the *P. falciparum* genome is 80–90% complete, these tags should prove to be informative as the genome project proceeds to completion. Alternatively, tags that do not match genome sequence may turn out to span splice junctions. These questions should be resolved as more genome sequence becomes available. Nevertheless, SAGE in *P. falciparum* is comparable with other studies where tags with no matches to the genome were as high as 20% (Matsumura *et al.*, 1999) and 23% (Yamashita *et al.*, 2000) of the total tags.

Finally, of the 8335 tags, 18% gave multiple matches to *Plasmodium* databases, a number that is fourfold higher than

that obtained from human pancreatic SAGE libraries, where ~5% of tags gave multiple matches (Velculescu *et al.*, 1995). However, pancreatic SAGE tags were only searched against RNA sequence databases, in contrast to our more extensive analysis that surveyed all available *Plasmodium* genome sequence. Hence, the higher percentage of multiple matches to the genome may reflect the method of analysis rather than any limitation of the technique when applied to the A-T rich genome of *Plasmodium*. Alternatively, the higher percentage of tags giving multiple matches may be a consequence of the lower complexity of the *Plasmodium* genome. Ambiguous tags of interest can be investigated further on an individual basis by Northern analysis.

### Metabolic Pathways Defined by SAGE

Other reports on SAGE have revealed metabolic profiles that are highly specific to the organism or tissue under study. For example, SAGE of mouse kidney revealed a preponderance of ion channels and mitochondrial enzymes, consistent with the role of the kidney in filtration and solute transport and the high-energy requirement for the same (El-Meanawy *et al.*, 2000). Transcriptional profiling of the 100 most abundant SAGE tags derived from seedlings of the rice plant, *Oryza sativa* L., demonstrated a prevalence of prolamin, a storage protein expressed in seeds (Matsumura *et al.*, 1999). As expected, other highly abundant transcripts included those encoding water channels and respiratory metabolism enzymes.

SAGE data from *P. falciparum* shed light on the transcriptional profile of blood stage parasites and hence reveal the classes of proteins and metabolic pathways that are probably used during asexual growth. For example, membrane-associated proteins form the most abundant category of expressed proteins. This is not surprising in light of the fact that the parasite is separated from its extracellular environment by three separate membranes: the host red blood cell membrane, the parasitophorous vacuole membrane, and the parasite plasma membrane (Torii and Masamichi, 1998). Many of these highly expressed proteins are stage specific and have been previously shown to be important in invasion of the red blood cell (MSP-3 and -4) (Barnwell and Galinski, 1998); others are transporters that may import nutrients into the parasite cell (importin  $\beta$ -subunit). Hence, the unique niche of the malarial parasite within the red blood cell requires the high expression of specific surface proteins.

A significant proportion (7.4%) of the most abundant tags was derived either from transcripts encoded on the 6-kb mitochondrial genome or from nuclear encoded transcripts involved in oxidative metabolism. High levels of RNA synthesis from the 6-kb element may reflect the fact that this episomally replicating molecule is present at ~20 copies per cell (Preiser *et al.*, 1996). However, a high demand for mitochondrial function and oxidative metabolism is suggested by the abundance of nuclear transcripts encoding proteins (thioredoxin, ubiquinol cytochrome *c* reductase-like protein) probably involved in the maintenance of intracellular oxidative homeostasis. Moreover, SAGE data show that transcripts encoding the molecular chaperones *hsp-60* and *-70*, which may be involved in import of nuclear-encoded proteins into the mitochondria (Das *et al.*, 1997), are also expressed at high levels. Hence, mitochondrial functions are most highly represented in the abundant classes of SAGE

tags, probably reflecting the microaerophilic lifestyle of the parasite within the red blood cell. The robust expression of genes involved in mitochondrial physiology may explain why mitochondrial pathways have been excellent targets for antimalarial drugs.

The major transcriptional pathways in the parasite as revealed by SAGE will help to identify potential drug targets and lead compounds. For example, atovaquone inhibits erythrocytic growth by targeting the mitochondrial cytochrome *bc<sub>1</sub>* complex (Fry and Pudney, 1992). Further evidence that other highly expressed metabolic pathways could also serve as drug targets is found in the following studies: the antimalarial drug fosmidomycin has been shown to target DOXP metabolism (Jomaa *et al.*, 1999); the ornithine decarboxylase inhibitor, difluoro-methylornithine, inhibits erythrocytic growth of *P. falciparum* in culture (Assaraf *et al.*, 1984); and folate antagonists such as pyrimethamine and cycloguanil target dihydrofolate reductase (Ferre *et al.*, 1969). Other major transcriptional patterns uncovered by SAGE in the parasite (proteasome, chaperones, unknown ORFs) may provide new targets for antimalarial drug development.

### **SAGE Reveals Novel Transcriptional Phenomena in *P. falciparum***

Most techniques for global analysis of gene expression are unable to distinguish sense and antisense transcripts. Due to the directional nature of SAGE tags (4 bp representing the *NlaIII* site that is closest to the 3' end of each transcript, and 10 bp downstream on the coding strand), we were able to identify numerous antisense transcripts in the transcriptional repertoire of *P. falciparum* asexual-stage parasites. Strand-specific RT-PCR and Northern analysis confirmed this observation for three of the genes (*msp-3*, *rap-1*, and calmodulin) predicted to transcribe antisense messages. The fact that antisense transcription can be detected in *Plasmodium* by three independent methods suggests that this is a bona fide biological phenomenon and not an artifact of the SAGE procedure.

Whether antisense RNAs transcribed by the malarial parasite have poly(A) tails is unclear. However, their representation in the SAGE library suggests that (similar to our observations with mitochondrial transcripts in the study by Munasinghe *et al.*, 2000) due to the presence of long poly(A) tracts in the genome of *P. falciparum*, SAGE results in the sampling of both polyadenylated RNAs as well as those lacking poly(A) tails.

Our data also demonstrate that antisense transcripts are expressed in other stages of *Plasmodium* development. The *pgs28* gene that encodes a major surface antigen of *P. gallinaceum* sexual stages has been studied extensively (Duffy *et al.*, 1993). Transcription of *pgs28* is restricted to the zygotes and ookinetes. Strand-specific RT-PCR shows that *pgs28* expresses both sense and antisense transcripts in both stages. Hence, the presence of antisense transcripts may be a widespread phenomenon in multiple stages of *Plasmodium* development and should be tested further. For example, a family of genes (*var*) encodes variable surface proteins involved in host-parasite interactions; *var* genes are transcribed during erythrocytic growth, resulting in the expression of the PfEMP-1 protein (Wahlgren *et al.*, 1999). Several *var* genes

are transcribed in the ring stages, whereas a single *var* gene is transcribed in trophozoites (Chen *et al.*, 1998; Scherf *et al.*, 1998). It would be interesting to test whether any of the ring stage *var* transcripts are antisense.

What is the biological significance of antisense transcription in *P. falciparum*? Antisense transcripts may reflect mechanisms of transcriptional initiation in a parasite with a highly A-T-rich genome (86% A-T in noncoding regions and 76% in coding sequence) (Bowman *et al.*, 1999). Numerous studies have shown that transcription in *P. falciparum* is initiated from the A-T-rich 5' upstream region of genes, resulting in sense transcripts (Horrocks *et al.*, 1998; Dechering *et al.*, 1999; Horrocks and Lanzer, 1999). The presence of antisense transcripts for 17% of annotated genes implies novel mechanisms of transcriptional initiation and termination, including potential roles in posttranscriptional control of protein expression.

In conclusion, we have shown that SAGE can be readily adapted for the study of global transcription in *P. falciparum*. SAGE of 3D7 asexual parasites sheds light on the prominent metabolic pathways used in these stages. Because blood stages are the targets of both antimalarial drugs and the host immune system, this comprehensive transcriptional profile generated by SAGE will form the basis for future comparisons of gene expression under drug or immune pressure. Finally, the unique nature of SAGE reveals a novel phenomenon, that of antisense transcription that has previously been missed.

### **ACKNOWLEDGMENTS**

We thank Drs. J. Kissinger and D. Roos (University of Pennsylvania, Philadelphia, PA; <http://www.plasmoDB.org>) for providing access to assembled *P. falciparum* genomic sequences. We also thank Dr. Connie Chow for insightful comments about this manuscript, and Dr. Kevin Militello for providing the hsp86 primers and valuable suggestions. We acknowledge the invaluable data provided by the Malaria Genome Consortium: Sequence data for *P. falciparum* chromosome (1, 3, 4, 5, 6, 7, 8, 9, and 13) was obtained from The Sanger Center Web site at [http://www.sanger.ac.uk/Projects/P\\_falciparum/](http://www.sanger.ac.uk/Projects/P_falciparum/). Sequencing of *P. falciparum* chromosome (1, 3, 4, 5, 6, 7, 8, 9, and 13) was accomplished as part of the Malaria Genome Project with support by The Wellcome Trust. Sequence data for *P. falciparum* chromosome 12 was obtained from the Stanford DNA Sequencing and Technology Center Web site at <http://www-sequence.stanford.edu/group/malaria>. Sequencing of *P. falciparum* chromosome 12 was accomplished as part of the Malaria Genome Project with support by the Burroughs Wellcome Fund. Preliminary sequence data for *P. falciparum* chromosome (2, 10, 11, and 14) was obtained from The Institute for Genomic Research Web site ([www.tigr.org](http://www.tigr.org)). Sequencing of chromosome (2, 10, 11, and 14) was part of the International Malaria Genome Sequencing Project and was supported by awards from the Burroughs Wellcome Fund and the U.S. Department of Defense. The Chromosome 2 Sequencing Project was a collaborative effort by The Institute for Genomic Research (TIGR) and the Naval Medical Research Center. This work was supported by the Burroughs Wellcome Fund and by the U.S. Department of Defense.

### **REFERENCES**

Assaraf, Y.G., Golenser, J., Spira, D.T., and Bachrach, U. (1984). Polyamine levels and the activity of their biosynthetic enzymes in

- human erythrocytes infected with the malarial parasite *Plasmodium falciparum*. *Biochem. J.* 222, 815–819.
- Barnwell, J.W., and Galinski, M.R. (1998). Invasion of vertebrate cells: erythrocytes. In: *Malaria: Parasite Biology, Pathogenesis, and Protection*, ed. I.W. Sherman, Washington, DC: ASM Press, 93–120.
- Bowman, S., Lawson, D., Basham, D., Brown, D., Chillingworth, T., Churcher, C.M., Craig, A., Davies, R.M., Devlin, K., Feltwell, T., Gentles, S., Gwilliam, R., Hamlin, N., Harris, D., Holroyd, S., Hornsby, T., Horrocks, P., Jagels, K., Jassal, B., Kyes, S., McLean, J., Moule, S., Mungall, K., Murphy, L., Barrell, B.G. and, et al. (1999). The complete nucleotide sequence of chromosome 3 of *Plasmodium falciparum*. *Nature* 400, 532–538.
- Butler, D. (1997). Funding assured for international malaria sequencing project. *Nature* 388, 701.
- Chen, Q., Fernandez, V., Sundstrom, A., Schlichtherle, M., Datta, S., Hagblom, P., and Wahlgren, M. (1998). Developmental selection of var gene expression in *Plasmodium falciparum*. *Nature* 394, 392–395.
- Cowman, A.F., Baldi, D.L., Healer, J., Mills, K.E., O'Donnell, R.A., Reed, M.B., Triglia, T., Wickham, M.E., and Crabb, B.S. (2000). Functional analysis of proteins involved in *Plasmodium falciparum* merozoite invasion of red blood cells. *FEBS Lett.* 476, 84–88.
- Craig, A.G., Waters, A.P., and Ridley, R.G. (1999). Malaria genome project task force: a post-genomic agenda for functional analysis. *Parasitol. Today* 15, 211–214.
- Das, A., Syin, C., Fujioka, H., Zheng, H., Goldman, N., Aikawa, M., and Kumar, N. (1997). Molecular characterization and ultrastructural localization of *Plasmodium falciparum* Hsp 60. *Mol. Biochem. Parasitol.* 88, 95–104.
- Dechering, K.J., Kaan, A.M., Mbacham, W., Wirth, D.F., Eling, W., Konings, R.N., and Stunnenberg, H.G. (1999). Isolation and functional characterization of two distinct sexual-stage-specific promoters of the human malaria parasite *Plasmodium falciparum*. *Mol. Cell Biol.* 19, 967–978.
- Duffy, P.E., Pimenta, P., and Kaslow, D.C. (1993). Pgs28 belongs to a family of epidermal growth factor-like antigens that are targets of malaria transmission-blocking antibodies. *J. Exp. Med.* 177, 505–510.
- El-Meanawy, M.A., Schelling, J.R., Pozuelo, F., Churpek, M.M., Ficker, E.K., Iyengar, S., and Sedor, J.R. (2000). Use of serial analysis of gene expression to generate kidney expression libraries. *Am. J. Physiol. Renal Physiol.* 279, F383–F392.
- Ferone, R., Burchall, J.J., and Hitchings, G.H. (1969). *Plasmodium berghei* dihydrofolate reductase. Isolation, properties, and inhibition by antifolates. *Mol. Pharmacol.* 5, 49–59.
- Foley, M., and Tilley, L. (1998). Quinoline antimalarials: mechanisms of action and resistance and prospects for new agents. *Pharmacol. Ther.* 79, 55–87.
- Fry, M., and Pudney, M. (1992). Site of action of the antimalarial hydroxynaphthoquinone, 2-[*trans*-4-(4'-chlorophyll) cyclohexyl]-3-hydroxy-1,4-naphthoquinone (566C80). *Biochem. Pharmacol.* 43, 1545–1553.
- Gardner, M.J., Tettelin, H., Carucci, D.J., Cummings, L.M., Aravind, L., Koonin, E.V., Shallom, S., Mason, T., Yu, K., Fujii, C., Pederson, J., Shen, K., Jing, J., Aston, C., Lai, Z., Schwartz, D.C., Perlea, M., Salzberg, S., Zhou, L., Sutton, G.G., Clayton, R., White, O., Smith, H.O., Fraser, C.M., Hoffman, S.L. et al. (1998). Chromosome 2 sequence of the human malaria parasite *Plasmodium falciparum* *Science* 282, 1126–1132.
- Goonewardene, R., Daily, J., Kaslow, D., Sullivan, T.J., Duffy, P., Carter, R., Mendis, K., and Wirth, D. (1993). Transfection of the malaria parasite and expression of firefly luciferase. *Proc. Natl. Acad. Sci. USA* 90, 5234–5236.
- Hayward, R.E., Derisi, J.L., Alfidhli, S., Kaslow, D.C., Brown, P.O., and Rathod, P.K. (2000). Shotgun DNA microarrays and stage-specific gene expression in *Plasmodium falciparum* malaria. *Mol. Microbiol.* 35, 6–14.
- Horrocks, P., Dechering, K., and Lanzer, M. (1998). Control of gene expression in *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* 95, 171–181.
- Horrocks, P., and Lanzer, M. (1999). Mutational analysis identifies a five base pair cis-acting sequence essential for GBP130 promoter activity in *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* 99, 77–87.
- Jomaa, H., Wiesner, J., Sanderbrand, S., Altincicek, B., Weidemeyer, C., Hintz, M., Turbachova, I., Eberl, M., Zeidler, J., Lichtenthaler, H.K., Soldati, D., and Beck, E. (1999). Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. *Science* 285, 1573–1576.
- Krogstad, D., and De, D. (1998). Chloroquine: modes of action and resistance and the activity of chloroquine analogs. In: *Malaria: Parasite Biology, Pathogenesis, and Protection*, ed. I.W. Sherman, Washington, DC: ASM Press, 331–339.
- Lobo, C.A., Fujioka, H., Aikawa, M., and Kumar, N. (1999). Disruption of the Pfg27 locus by homologous recombination leads to loss of the sexual phenotype in *P. falciparum*. *Mol. Cell* 3, 793–798.
- Madden, S.L., Galella, E.A., Zhu, J., Bertelsen, A.H., and Beaudry, G.A. (1997). SAGE transcript profiles for p53-dependent growth regulation. *Oncogene* 15, 1079–1085.
- Matsumura, H., Nirasawa, S., and Terauchi, R. (1999). Technical advance: transcript profiling in rice (*Oryza sativa* L.) seedlings using serial analysis of gene expression (SAGE). *Plant J.* 20, 719–726.
- Munasinghe, A., Patankar, S., Cook, B.P., Madden, S.L., Martin, R.K., Kyle, D.E., Shoaibi, A., Cummings, L.M., and Wirth, D.F. (2000). Serial analysis of gene expression (SAGE) in *Plasmodium falciparum*: application of the technique to A-T rich genomes. *Mol. Biochem. Parasitol.* 113, 23–34.
- O'Brien, C. (1997). Malaria genome project gets a funding boost. *Mol. Med. Today* 3, 3.
- Polyak, K., Xia, Y., Zweier, J.L., Kinzler, K.W., and Vogelstein, B. (1997). A model for p53-induced apoptosis. *Nature* 389, 300–305.
- Preiser, P.R., Wilson, R.J., Moore, P.W., McCready, S., Hajibagheri, M.A., Blight, K.J., Strath, M., and Williamson, D.H. (1996). Recombination associated with replication of malarial mitochondrial DNA. *EMBO J.* 15, 684–693.
- Rosenthal, P., and Meshnick, S. (1998). Hemoglobin processing and the metabolism of amino acids, heme and iron. In: *Malaria: Parasite Biology, Pathogenesis, and Protection*, ed. I.W. Sherman, Washington, DC: ASM Press, 145–158.
- Scheibel, L., and Sherman, I.W. (1988). Plasmodial metabolism and related organellar function during various stages of the life cycle: proteins, lipids, nucleic acids and vitamins. In: *Malaria, Principles and Practice of Malariology*, ed. W. Wernsdorfer and I. McGregor, Edinburgh, UK: Churchill Livingstone, 219–252.
- Scherf, A., Hernandez-Rivas, R., Buffet, P., Bottius, E., Benatar, C., Pouvelle, B., Gysin, J., and Lanzer, M. (1998). Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of var genes during intra-erythrocytic development in *Plasmodium falciparum*. *EMBO J.* 17, 5418–5426.
- Sherman, I.W. (1991). The biochemistry of malaria: an overview. In: *Biochemical Protozoology*, ed. G. Coombs and M. North, London, England: Taylor and Francis, 6–34.
- Torii, M., and Masamichi, A. (1998). Ultrastructure of asexual stages. In: *Malaria: Parasite Biology, Pathogenesis, and Protection*, ed. I.W. Sherman, Washington, DC: ASM Press, 123–134.

- Trager, W., and Jensen, J.B. (1976). Human malaria parasites in continuous culture. *Science* 193, 673–675.
- Velculescu, V.E., Zhang, L., Vogelstein, B., and Kinzler, K.W. (1995). Serial analysis of gene expression. *Science* 270, 484–487.
- Velculescu, V.E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M.A., Bassett, D.E., Jr., Hieter, P., Vogelstein, B., and Kinzler, K.W. (1997). Characterization of the yeast transcriptome. *Cell* 88, 243–251.
- Vial, H.J., and Ancelin, M.L. (1998). Malarial lipids. In: *Malaria: Parasite Biology, Pathogenesis, and Protection*, ed. I.W. Sherman, I.W., Washington, DC: ASM Press, 159–175.
- Virlon, B., Cheval, L., Buhler, J.M., Billon, E., Doucet, A., and Elalouf, J.M. (1999). Serial microanalysis of renal transcriptomes. *Proc. Natl. Acad. Sci. USA* 96, 15286–15291.
- Wahlgren, M., Fernandez, V., Chen, Q., Svard, S., and Hagblom, P. (1999). Waves of malarial variations. *Cell* 96, 603–606.
- Walsh, C.J., and Sherman, I.W. (1968). Purine and pyrimidine synthesis by the avian malaria parasite, *Plasmodium lophurae*. *J. Protozool.* 15, 763–770.
- World Health Organization (WHO) (1997). *The World Health Report. Conquering, Suffering, Enriching Humanity*. Geneva, Switzerland: WHO Publishers.
- Yamashita, T., Hashimoto, S., Kaneko, S., Nagai, S., Toyoda, N., Suzuki, T., Kobayashi, K., and Matsushima, K. (2000). Comprehensive gene expression profile of a normal human liver. *Biochem. Biophys. Res. Commun.* 269, 110–116.
- Yu, D.C., Wang, A.L., Wu, C.H., and Wang, C.C. (1995). Virus-mediated expression of firefly luciferase in the parasitic protozoan *Giardia lamblia*. *Mol. Cell. Biol.* 15, 4867–4872.