Available online at www.sciencedirect.com
SCIENCE DIRECT
Molecular & Biochemical Parasitology 142 (2005) 184–192
MOLECULAR & BIOCHEMICAL PARASITOLOGY

Insights into the *P. y. yoelii* hepatic stage transcriptome reveal complex transcriptional patterns

Anne Charlotte Grüner [a,b,1], Stéphanie Hez-Deroubaix [a,1], Georges Snounou [c,2], Neil Hall [d,3], Christiane Bouchier [e], Frank Letourneur [f], Irène Landau [g], Pierre Druilhe [a,*]

[a] *Unité de Parasitologie Bio-Médicale, Institut Pasteur, 25 Rue du Dr Roux, 75731 Paris Cedex 15, France*
[b] *INSERM U567, CNRS UMR 8104, Département d'Immunologie, Université René Descartes, Hôpital Cochin, 27 Rue du Fbg St Jacques, 75014 Paris, France*
[c] *Unité de Parasitologie Bio-Médicale et CNRS URA 2581, Institut Pasteur, 25 Rue du Dr Roux, 75731 Paris Cedex 15, France*
[d] *The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK*
[e] *Pasteur Genopole Ile de France, Institut Pasteur, 25 Rue du Dr Roux, 75731 Paris Cedex 15, France*
[f] *Service Commun de Séquençage, Institut Cochin, Bâtiment Gustave Roussy, 27 Rue du Faubourg St Jacques, 75014 Paris, France*
[g] *Equipe Parasitologie Comparée et Modèles Expérimentaux USM 0307, Muséum National d'Histoire Naturelle, CP52, 61, Rue Buffon, 75005 Paris, France*

Received 21 December 2004; received in revised form 7 March 2005; accepted 23 March 2005
Available online 6 May 2005

Abstract

During their complex life cycle, malaria parasites adopt morphologically, biochemically and immunologically distinct forms. The intra-hepatic form is the least known, yet of established value in the induction of sterile immunity and as a target for chemoprophylaxis. Using *Plasmodium yoelii* as a model we present here a novel approach to the elucidation of the transcriptome of this poorly studied stage. Sequences from *Plasmodium* were obtained in 388 of the 3533 inserts (11%) isolated from liver stages cDNA obtained from optimized cultures with high yields. These corresponded to a total of 88 putative *P. yoelii* genes. The majority of the transcribed genes identified, code for predicted proteins of as yet unknown function. The RT-PCR analysis carried out for 29 of these genes, confirmed expression at the hepatic stage and provided evidence for complex patterns of genes transcription in the distinct stages found in the mosquito and vertebrate host. The results demonstrate the efficacy of the approach that can now be applied to further detailed analysis of the hepatic stage transcriptome of *Plasmodium*.

Keywords: Transcriptome; *Plasmodium yoelii*; Liver stages; Stage-specific transcripts

1. Introduction

The unicellular malaria parasites shuttle between the mosquito and the vertebrate hosts, adopting morphologically and biochemically distinct forms, implying complex regulation of the predicted 6000 genes to adapt to very distinct environments. Indeed, from a morphological and biological point of view, there are major differences between the gametes, the oocysts and sporozoites of the sexual cycle in the mosquito, or between these stages and the two distinct asexual rapidly dividing forms occurring first in the hepatocyte and thereafter in the red blood cell of the vertebrate. There are wide differences in parasite numbers depending on the stage, the least numerous being the hepatic forms, which might explain

Abbreviations: RTase, reverse transcriptase; MgS, midgut sporozoite; SgS, salivary gland sporozoite; LSc, liver stages parasites from cultures; LSv, liver stages parasites from in vivo; BSt, blood stage parasites; "s", sense transcription; "a", anti-sense transcription

* Corresponding author. Tel.: +33 1 45 68 85 78; fax: +33 1 45 68 86 40.
E-mail address: druilhe@pasteur.fr (P. Druilhe).

[1] Anne Charlotte Grüner and Stéphanie Hez-Deroubaix are equal first authors.

[2] Present address: Equipe Parasitologie Comparée et Modèles Expérimentaux USM 0307 et CNRS URA 2581, Muséum National d'Histoire Naturelle, CP52, 61, rue Buffon, 75005 Paris, France.

[3] Present address: Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, USA.

0166-6851/$ – see front matter © 2005 Published by Elsevier B.V.
doi:10.1016/j.molbiopara.2005.03.019

why they remained little studied for so long. Despite their scarcity, the characteristics of the intra-hepatic transcription are of considerable value from (a) a fundamental biological point of view, for e.g., the intriguing chronobiology of the dormant forms, (b) a malaria control point of view since it is the only stage against which sterile immunity can be induced in humans [1–4], and (c) a therapeutic point of view, as it is biochemically distinct from other stages rich in unique pathways that constitute ideal prophylactic targets as it precedes the pathogenic phase. Therefore, the elucidation of the pattern of gene expression during these hepatic stages is of considerable fundamental and practical value. Molecular studies of liver stages, however, have been difficult to undertake as severe technical and methodological limitations restricted access to the material [3].

In the present work we elaborated a strategy for the analysis of the hepatic stage transcriptome by capitalizing on four recent breakthroughs: the availability of the genome sequence of several *Plasmodium* species, that of the mouse genome, progress in high throughput sequencing, and the development of improved culture conditions of liver stages that lead to markedly higher rates of infection of hepatocytes. We selected *P. yoelii* as a model to examine the feasibility and the output of this approach, because this species is the most widely employed for vaccine and drug development in preclinical models and the life cycle can be easily handled in the laboratory. The first results presented here, demonstrate the validity of the approach.

2. Materials and methods

2.1. P. y. yoelii hepatic stage culture

Inbred 6–8 weeks female C57BL/6 (H-2^b) mice (Charles River, France) were housed and used with approval from the Pasteur Animal Welfare Committee. Hepatocytes were prepared by perfusion of mice liver following a two-step enzymatic protocol [5], with carefully pre-selected Collagenase H batches in a $CaCl_2$ buffer and purified on a 60% Percoll gradient at 2000 rpm for 2 min. Preparations with viability determined by blue trypan exclusion >90%, were suspended in Williams medium E (Gibco, Life Technologies, France) supplemented with 0.2% BSA (Sigma–Aldrich, France), 10% foetal calf serum, 1% penicillin–streptomycin (Gibco, Life Technologies, France), 1% spite (Sigma–Aldrich, France), and 1% L-glutamine. Cells were then seeded at a density of 7×10^4 per well, in 8-chamber LabTEK permanox slides (Nalge Nunc International, Naperville, Ill.). Sporozoites were obtained by dissection following Ozaki Method [6] from *P. y. yoelii* clone 1.1-infected *Anopheles stephensi* (14 days post-gametocyte infection), and kept in William's medium supplemented with 10% foetal calf serum and 1% penicillin–streptomycin, at 4 °C until use. Only batches with >15,000/20,000 sporozoites per mosquito were employed. Fifty thousand sporozoites in a volume of 200 μl William's medium (Gibco, Life Technologies, France) supplemented with dexamethasone (10^{-7} M) were added per well 24 h after cell plating. Infected hepatocyte cultures were centrifuged as described previously [7] and incubated at 37 °C in 5% CO_2 atmosphere for 3 h. The medium was carefully removed and 300 μl of supplemented William's with 3.5 g/l glucose was added to each well. Cultures were incubated at 37 °C in 5% CO_2 atmosphere and the medium changed daily. The growth of liver stages was monitored under a phase-contrast inverted microscope. Cells were harvested 48 h after infection only from a selected culture where the prevalence of high-infected hepatocytes was high (≥3%) and the parasitized cells were homogeneously 40–50 μm in size, and kept at −80 °C in RNA Later Buffer (Qiagen, Germany). Two control wells per slide were used to enumerate liver stage parasites using FITC-labelled anti-HSP70 antibody [8], and the residual sporozoites using FITC-labelled anti CS mAb NYS1. All experiments and procedures performed using animals conformed to the French Ministry of Agriculture Regulations for Animal Experimentation (1987).

2.2. RNA purification

Messenger RNA was purified from *P. yoelii yoelii* clone 1.1-infected primary mouse (C57Bl/6) hepatocyte cultures seeded in seven 8-well LabTEK culture plates. The cultures were harvested at 48 h post-sporozoites addition, when the parasites were at sub-mature schizont stage (Fig. 1). In the preparation selected, each well contained 70,000 hepatocytes

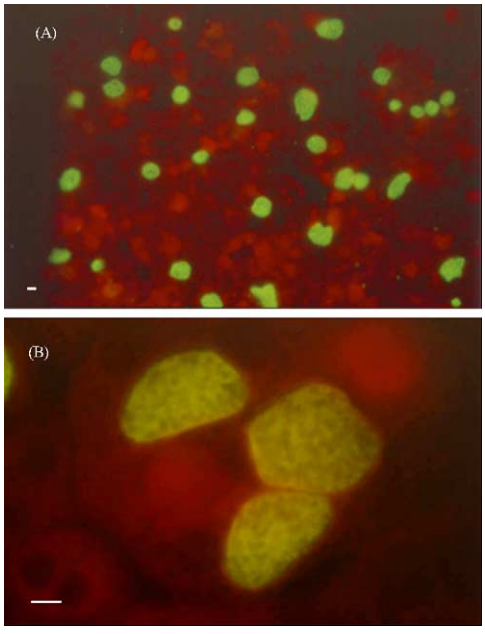

Fig. 1. *Plasmodium yoelii* liver stages employed in this study. IFAT of liver stages in primary hepatocytes culture, performed with antibodies directed against *Plasmodium* HSP70-2, showing the high proportion of parasites that were highly homogeneous in size and close to full maturation though still developing. Panel A: magnification 40×. Panel B: magnification 100×. Scale bar of both panels: 10 μm.

of which ca. 2500 were infected, i.e. the proportion of infected hepatocytes was 3.1%, and the rate of transformation of sporozoite into LS was ca. 5%. Thus the mRNA was purified from a total of 3.2 million hepatocytes of which 115,000 harboured *P. y. yoelii* infected hepatocytes.

Purification of the PolyA$^+$ RNA was carried out using the Micro Fast Track 2.0 Kit (Invitrogen, The Netherlands). The purified RNA was treated with RNase-free DNase I (Amersham, United Kingdom) in the presence of Rnasin (Promega, USA), before recovery by acid phenol extraction, and precipitation in isopropanol. The pellet obtained was washed in 70% Ethanol, dried and resuspended in 3 μl of H_2O. cDNA synthesis was performed using the SMART cDNA Synthesis kit (Clonetech, USA), following the manufacturer's instructions. Second strand synthesis was performed using 1/10 of the first strand cDNA, in a total volume of 100 μl, with either 24 or 27 cycles of amplification performed. The resulting PCR products were purified by the Qiagen PCR purification kit (Qiagen, Germany), and eluted in a total of 30 μl. Of this eluate, 4 μl were used for cloning in the TOPO pCR®4 vector (Invitrogen, The Netherlands), following the manufacturer's instructions. Top10 cells were transformed with 2 μl of the ligation mixture, and 20 μl were plated on LB agar plates supplemented with 100 μg/ml of Ampicillin. The resulting bacterial colonies were picked for plasmid purification and sequencing, and stored at −80 °C in the presence of 15% glycerol.

Total RNA was extracted from *P. y. yoelii* clone 1.1 infected C57/Bl6 mouse liver 44 h following intra-venous injection of 500,000 sporozoites, using RNeasy Mini Kit (Qiagen S.A., France) according to the manufacturer's instructions. The proportion of infected hepatocytes was ca. 0.03%, and liver schizonts were mature.

Total sporozoite RNA for use in the RT-PCR analyses was purified by the RNeasy Kit (Qiagen, Germany) from ca. 200,000 midgut sporozoites, or 540,000 salivary gland sporozoites dissected out of *P. y. yoelii* clone 1.1-infected *A. stephensi* mosquitoes 7 and 14 days post-feeding on infected mice, respectively. Total RNA from blood stage *P. y. yoelii* clone 1.1 parasites (150 μl of blood with 10% parasitaemia) was obtained in the same manner. All the purified RNA samples were exhaustively treated with RNase-free DNase, before phenol extraction and isopropanol precipitation before resuspension in 20 μl or 50 μl of RNase-free water for sporozoites and blood stage parasites preparations, respectively.

2.3. RT-PCR assays

Three oligonucleotide primers were designed for each of the 29 putative genes analyzed by RT-PCR, using the sequence obtained from the *P. y. yoelii* database [9]. The two outer primers were used separately for the RT step, and as a pair for the primary PCR amplification. The third inner primer was used in combination with one of the two outer primers for the semi-nested secondary amplification. For many of the putative genes it was not possible to deduce unambiguously the sense of transcription, especially when only a truncated sequence at the end of the contig was available. We have therefore arbitrarily assigned a "sense" ("s") and "antisense" ("as") direction of transcription. The sequences for the oligonucleotides used are provided in Supporting Information, Table 2, where it is indicated which primers were used for the "s" or "as" RT reaction, and which pairs were used for the primary (1) or secondary (2) PCR reactions.

The reverse transcription step was performed in a total volume of 20 μl with 1 μl of the purified total RNA, using the Superscript II RTase enzyme (Invitrogen, The Netherlands), according to the manufacturer's instructions. In order to optimize the use of the material, 8–10 different oligonucleotide primers were included in each reaction. Preliminary experiments conducted using the blood stage parasite total RNA, demonstrated that inclusion of multiple primers did not alter the sensitivity of the detection (data not shown). The oligonucleotides corresponding to the arbitrary "s" transcripts for each gene were included in one RT reaction, whereas those corresponding to the "a" transcripts were included in another separate RT reaction.

The primary and secondary PCR amplifications were carried out in a total volume of 20 and 30 μl, respectively. For the primary reaction 1 μl of reverse transcribed total RNA was used as template, and 1 μl of the product of this reaction was used as template for the secondary amplification reaction. Oligonucleotide primers were each used at a final concentration of 250 nM. The reaction mixture contained 10 mM Tris–HCl, pH 8.3, 50 mM KCl, 125 μM of each of the four dNTPs, and 0.4 units of AmpliTaq polymerase (Perkin Elmer Cetus, USA). Diverse Mg^{2+} concentrations were used for the various reactions as indicated (Supporting Information, Table 2). The amplification reactions were carried out in a PTC-200 thermocycler (M & J Research, USA) as follows: an initial denaturation step at 95 °C for 5 min was followed by 25 (primary reaction) or 30 cycles (secondary reaction) of annealing at 60 °C for 1 min, extension at 72 °C for 1 min, and denaturation at 94 °C for 30 s. After a final annealing step followed by 5 min of extension, the reaction was stopped. The presence of specific PCR products was established by electrophoresis on 3.0% MetaPhor agarose gels.

2.4. Sequencing and bioinformatics

Plasmid DNA purification were performed with Montage Plasmid Miniprep$_{96}$ Kit (Millipore, France**)** for one set of clones, the sequencing reactions were performed, from 5′ end of DNA plasmid, using ABI PRISM BigDye Terminator cycle sequencing ready reactions kit and run on a 3700 Genetic Analyzer (Applied Biosystems). The trace files were based-called using Phred [10]. Sequences not meeting production quality criteria (at least 100 bases called with a quality over 20) were discarded. A second set of clones was sequenced using a T3 or T7 primer using ABI BigDye terminator kits and run on ABI3730 DNA sequencing machines. The sequence was clipped for quality using Phred and vector us-

ing Cross Match (Phil Green, unpublished) and Svec_clip (Richard Mott, unpublished).

The sequences were searched against the mouse genome using BLASTN. Sequences with >90% identity over >80% of their length were considered to be of mouse origin. The remaining sequences were searched against the genome sequences and predicted genes from all sequences from *Plasmodium* species, including *P. berghei*, *P. chabaudi* (Hall et al. unpublished), *P. falciparum* [11] and *P. y. yoelii* [9], using BLASTN and BLASTX (cutoff E values of 10^{-10} and 10^{-5}, respectively). Alignments to genomic sequence were manually inspected with respect to the overlaid annotation using the Artemis genome browser [12].

3. Results

We opted to focus this first analysis on sub-mature dividing schizonts. *P. yoelii* liver schizont maturation was continuously monitored microscopically. Success of cultivation depends on a number of difficult to standardise factors, amongst which the quality of the hepatocytes preparation, the quality of the mosquito colony and the infectivity of the sporozoites are the most critical factors. Thus, from a series of cultures performed on a weekly basis, we selected one batch where the following criteria were met: high invasion rate, homogeneity of the maturation amongst parasites, and finally large size of individual schizonts (40–50 μm) (Fig. 1).

PolyA$^+$ RNA was purified from a culture containing 115,000 infected hepatocytes, representing 3% of the total number of hepatocytes, which is high by pre-erythrocytic culture standards [13]. Harvesting was performed 48 h post-infection with sporozoites. For this culture, hardly any of the 50,000 sporozoites added to each well were detectable at this time, as determined by microscopic observation after staining with anti-CS antibodies (data not shown). The cDNA library was derived from the whole batch of RNA through a PCR-based protocol. Using a moderate number of amplification cycles to minimize bias, a potential library of 1,800,000 colonies could be generated, and 5952 colonies (62×96-well plates) were picked for further analysis so as to assess the validity of the approach.

A total of 3533 meaningful sequences were obtained (59% of the total), and these were then assembled to yield 171 clusters (representing 3363 sequences) and 169 singletons. BLASTN analysis against mouse and *Plasmodium* genomes revealed that although 3144 of the 3533 sequences (89%) were of mouse origin (117 clusters and 122 singletons), 388 (11%) were ascribed to the *P. y. yoelii* or other plasmodial genomes (54 clusters and 46 singletons), with one singleton that could not be ascribed to either organisms. The redundancy for the 117 clusters corresponding to mouse sequences was low (geometric mean of 5.8) and included 27 clusters that were assembled from more than 10 sequences with only four of these comprising more than 100 members (112, 301, 449 and 1090 sequences, respectively). The redundancy for the 54 *P. y. yoelii* clusters was also low (geometric mean of 3.2), and these included only two that were assembled from more than 10 sequences (22 and 155, respectively). These observations indicated that preferential amplification of a subset of transcripts did not occur as a result of the amplification step included for the construction of the cDNA library we generated.

Of the 100 sequences (388 reads), 96 (384 reads) were assigned through BLASTN analysis to a total of 88 different contigs of the *P. y. yoelii* 5X coverage database (Supporting Information, Table 3). In five cases two distinct contigs were found linked with a cDNA sequence overlapping their ends. Alignment placed the cDNAs within or close to annotated ORFs of >50 aa in 87 of the 96 distinct sequences, identifying a total of 79 predicted genes. No ORFs were predicted for the remaining nine, but the cDNA insert lay close to the contig end in six of these, and the present *P. y. yoelii* genome is known to be incompletely covered [9]. All 100 sequences identified were also sought in the *P. y. yoelii* EST database derived from sporozoites [14] and that from blood stage parasites (both available through http://www.tigr.org/tdb/tgi/pygi/). A total of 47 hits were obtained, and these 13 were to sporozoite ESTs uniquely, 31 were to blood stage parasite ESTs uniquely, and three were to both sets. The remaining 53 cDNA inserts were negative in these BLASTN analyses. Homologues of the *P. y. yoelii* cDNA sequences were also sought in the genomes of *P. berghei* and that of *P. falciparum*. Hits against *P. berghei* DNA or blood stage EST sequences were obtained for 89 of the 100 sequences, while BLASTX hits against annotated *P. falciparum* proteins were obtained for 44 of the 100 sequences (Supporting Information Supplementary data, Table S3). These analyses confirmed that all 100 sequences were derived from a *Plasmodium* genome.

In order to confirm that the sequences identified were indeed transcribed by the parasite, we selected 29 of the putative genes for RT-PCR analysis, three of which had no associated predicted ORF. Eleven of these putative genes were identified with a singleton cDNA sequence, while the remainder were represented by clusters with varying redundancies. We further wished to determine or confirm transcription of these genes throughout the parasite life cycle. Thus, RT-PCR analysis was performed on total RNA purified from two insect stages, midgut sporozoites (MgS) and salivary gland sporozoites (SgS); and from two mammalian stages, liver stage parasites both from cultures (LSc) and from an in vivo infection (LSv), and blood stage parasites (BSt). Moreover, for each of the 29 identified sequences and for each type of RNA, the RT step was performed independently with the "sense" ("s") and the "anti-sense" ("a") primers. Finally, we adopted a semi-nested PCR strategy where the sensitivity of detection was of the order of 100 target copies. The relative quantity of the different RNA aliquots used in the assays was ascertained by real-time quantitative PCR based on the ribosomal RNA genes [15]. Since we wished to determine whether the putative genes were specifically transcribed in

the pre-erythrocytic stages, the quantity of RNA included in the RT reactions for the BSt was a 1000-fold larger than that added for the other stages.

Evidence for transcription was obtained from RT-PCR analysis for all the 29 putative genes tested, and transcription patterns at the different stages were found to be complex (Table 1). All 29 putative genes were found to be transcribed in blood stage parasites, but only one of these (Table 1, primers A16) was not found transcribed in one or more of the other stages. Evidence of transcription at the hepatic stage was obtained for 26 of the putative genes, of which seven were not transcribed in the insect stages. Evidence for hepatocytes stage transcription was obtained for 25 of the putative genes using RNA purified from LSc, but only for 9 of these genes when the RNA used was purified from LSv. For gene A03 (Table 1) evidence for hepatic stage transcription was only obtained with the LSv RNA. Transcription in sporozoites was detected for 21 of the putative genes that included two for which no RT-PCR signal was obtained from liver stage parasites. Transcripts from three of the 29 putative genes (Table 1, primers A12, A16 and A33) were not detected in LS, either because these genes were transcribed during a very narrow time frame (different batches were used for the cDNA construction and RNA purification for RT-PCR assays), or because the corresponding RNA was particularly labile. These transcripts might also correspond to genes expressed by residual sporozoites at the time of harvesting, though this is unlikely as very few such sporozoites were present, furthermore no transcript was found in sporozoite RNA for one of these three putative genes.

The 29 putative genes included five with hits to sporozoite ESTs and 11 to blood stage ESTs, one to both and 12 to neither. This was indeed confirmed for all except one of the sporozoite-expressed genes. However, the RT-PCR analysis revealed that nine of the 11 putative genes found in the blood stage ESTs only, were also found transcribed in sporozoites,

Table 1
RT-PCR derived transcription patterns during the life cycle

Primers	cDNA	Red.	Gene	MgS		SaS		LSc		LSv		BSt	
				s	a	s	a	s	a	s	a	s	a
A01	NC102	4	NONE	s		s		s		s		s	s
A03	GS0016E06.b	1	PY01114										
A04	NC074	3	PY00417										
A06	AGY012d12.p1k	1	PY02417	s		s		s				s	
A07	NC061	2	PY02936		w								
A12	NC077	3	PY00348			w							
A13	NC021	2	PY01282	w	s		s	s	s		s	s	s
A14	NC045	2	PY01616										
A15	NC049	2	NONE	w	w						w		
A16	GS0030B01.b	1	PY00082									gs	
A17	GS0004G04.b	1	PY00448										
A18	GS0027G12.b	1	PY00628										
A19	NC171	155	PY06160	gs				gs		gs		gs	
A20	GS0020E04.b	1	PY06945	w	w								
A21	GS0028F09.b	1	PY00355										
A22	NC01000	1	PY08617						w	w			
A23	GS0005D11.b	1	PY00898										
A24	NC001	2	PY01761										
A27	NC067	3	PY03599										
A28	NC012	2	PY05236	s	s			s	s			s	s
A29	NC097	3	PY05881										
A31	NC070	3	PY00230										
A33	GS0013C05.b	1	PY02102										
A36	GS0025H05.b	1	PY03841										
A37	NC115	5	PY04223			gs							gs
A40	NC092	3	PY04707						w				
A41	NC073	3	PY04834										
A42	NC128	6	PY05285					s				s	gs
A46	NC023 & NC159	24	NONE	s	s		s	s	s			s	s

Oligonucleotide primers (A numbers) were designed for each of 29 putative genes for which the corresponding contig or singleton number, EST redundancy in the library (Red.), and gene number in the *P. y. yoelii* 1.1 data base are indicated. Independent RT reactions were performed using the sense "s" (filled cells) or anti-sense "a" (stippled cells) primers, prior to PCR amplification to detect the presence of transcripts in total RNA purified from different stages of the life cycle: MgS, midgut sporozoites (light blue), SgS, salivary gland sporozoites (dark blue), hepatic stages parasites obtained from in vitro cultures, LSc, or from sporozoite-infected mouse livers, LSv (pink for both types), and finally BST, blood stage parasites (red). Detection of an RT-PCR product is indicated by a filled cell (color-coded as above), and the letter within the cells denotes whether the size of the product was as expected (g, genomic) smaller than expected (s, evidence of splicing), for some both were observed (gs). For primers A23, sequencing of the RT-PCR products amplified from the sporozoite RNA samples were shown to be derived from *Anopheles* (grey cells).

and four of the five solely found in the sporozoite ESTs were also found transcribed in the hepatic and blood stages (Fig. 2).

Splicing was predicted to occur for 11 of the 79 predicted genes identified by the cDNA inserts sequenced. Evidence for splicing was obtained for four of these putative genes that were included among those selected for RT-PCR analysis. Evidence for splicing was obtained for further five genes from the selection. Sequencing of the amplified products was carried out for seven of these nine putative genes and confirmed that the fragments did represent a spliced version of their respective genomic DNA sequences. Evidence of splicing was also obtained by RT-PCR for three genes where splicing was not predicted. Finally, two of the three sequences for which no ORF had been assigned were also found to be spliced, highlighting that gene finder predictions are only indicative.

Two noteworthy features of the transcripts of the putative genes were revealed by RT-PCR, where the RT step was undertaken independently with "sense" and anti-sense" primers. First, transcripts in both senses were observed for all 29 putative genes in the blood stage parasites, and for 14 of these genes in the other parasite stages, namely sporozoite and hepatic parasites (Fig. 2). The fact that the quantity of RNA, as ascertained by real-time quantitative PCR using the small subunit RNA genes, used in each assay was 1000-fold higher for the blood stages as compared to the other stages might account for this observation. However, equivalent quantities of RNA were used for the sporozoite and hepatic stage RT reactions, thus the results for these assays were comparable. When the data from the blood stages were excluded, it was interesting to note that "s" and "a" transcripts were more often observed for MgS (9/13) than for SgS (3/13), and they were detected for all the 13 putative genes using RNA from LSc but for none from LSv (Fig. 2).

The number of annotated genes identified (79) through this preliminary analysis of the *P. y. yoelii* hepatic stage transcriptome, is too limited to allow meaningful comparison with transcripts found in the more accessible parasite stages. Nonetheless, when the possible function of these genes is considered, half (39/79) represented hypothetical proteins, and the remainder were either broadly involved in metabolism (22/79) or in nucleic acid interactions (18/79). It is interesting to note that the putative gene with the highest redundancy in the cDNA sequences (155, accounting for 40% of all the reads of *Plasmodium* origin) was predicted to code

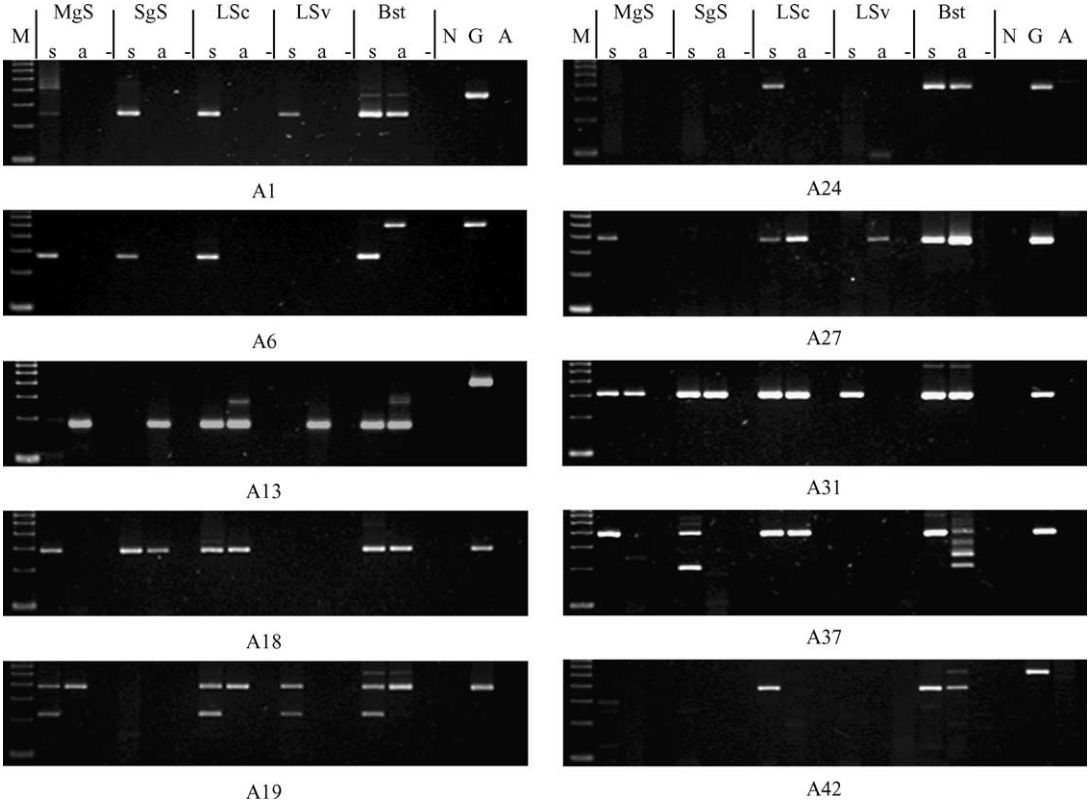

Fig. 2. RT-PCR amplification performed on total RNA purified from different parasite forms. RT-PCR results for 10 selected putative genes (corresponding primer numbers are derived from Table 1, and are indicated below each panel), obtained using total RNA purified from different stages of the parasite's life cycle, abbreviated as follows above the top panels: MgS, midgut sporozoites, SgS, salivary gland sporozoites, hepatic stages parasites obtained from in vitro cultures, LSc, or from sporozoite-infected mouse livers, LSv, and finally BST, blood stage parasites. The initial RT reaction was performed in the presence of the sense "s" or the "anti-sense" primer, or in the absence of the reverse transcriptase "–", as indicated above the corresponding lanes. Amplified products were separated by electrophoresis on 3% MetaPhor agarose gels, and visualized by UV-transillumination following ethidium bromide staining. Lane "N" represents the "no template" control for the PCR, lane "G" represents amplification from purified *P. y. yoelii* 1.1 genomic DNA, and lane "A" represents amplification from purifies *Anopheles stephensi* genomic DNA. Molecular weight markers "M", consisted of a 100 bp ladder, with the lowest band visible being 100 bp in size.

for a translation-elongation factor (Supporting Information Supplementary data, Table S3).

4. Discussion

In this work we demonstrate that the transcriptome of the hepatic stages is accessible to classical genomic analysis despite the difficulty to access large numbers of infected cells, provided a large number of sequences are screened. We chose *P. y. yoelii*, a rodent malaria parasite, as a model because it is extensively used for fundamental and applied research, because its genome is sufficiently characterized [9], and because a similar approach can be applied to *P. falciparum*. We opted for in vitro cultured material, because it reduces the number of contaminating host cells, large numbers of sporozoites are brought into contact with a relatively small number of hepatocytes, because hepatocytes do not replicate in vitro whereas the parasites do so actively and finally because the degree of parasite maturation can be continuously monitored microscopically.

The efficiency of the approach was remarkably high in that although ca. 3% of the hepatocytes were infected, 11% of the cDNA clones corresponded to *P. y. yoelii* sequences. There are three possible non-exclusive explanations underlying this observation. First selection of PolyA$^+$ RNA might be biased towards AT-rich sequences characteristic of *Plasmodium* genomes, indeed some of the cDNA corresponded to segments in the middle of an ORF. Second, it is known that the metabolic activity of in vitro-maintained primary hepatocytes diminishes with time, their cellular division is rarely observed, whereas that of the liver schizonts is remarkably active. Third, the sub-mature dividing schizonts from which we opted to purify RNA, contain multiple parasite nuclei (Fig. 1) and represent a stage where transcription was most likely very active.

Analysis of the 388 reads unambiguously assigned to *Plasmodium*, uncovered 88 putative genes in the transcriptome of *P. y. yoelii* liver stage parasites. This number compares favourably with that obtained for the transcriptome of a more accessible parasite stage, the sporozoite. Using a similar cDNA library construction method from a starting material of half a million *P. y. yoelii* sporozoites, the 1900 ESTs obtained assembled to 160 putative genes only [14]. This relatively lower yield might reflect the quiescent metabolic state of the mature salivary gland sporozoites used. These yields should not be compared to those obtained from cDNA libraries of *P. yoelii* blood stage parasites, since these libraries did not comprise the amplification step included for the construction of sporozoite and liver stage cDNA libraries.

The bioinformatics analyses, and the associated detailed RT-PCR assays conducted for 29 of the 88 putative genes, uncovered a number of novel and unexpected features of transcription in *Plasmodium*. First, none were found to be transcribed solely in hepatic stage parasites, and transcripts for 21 genes were detected in both insect and mammalian stages and for seven genes in mammalian stages only. Liver stage transcription was more frequently observed in LSc RNA derived from in vitro cultures (25/26 of the putative genes) as compared to LSv RNA derives from *in vivo* mouse infection (10/26). This is most likely due to the difference in maturity of the parasites from which the RNA was purified. Maturation of liver is slower than in vivo, moreover, full maturation of merozoites and their release from primary mouse hepatocytes culture does not occur in vitro for *P. yoelii*. The differences observed for these two hepatic stages RNA sampled suggest that the transcript levels for many putative genes have declined in the fully mature forms. The fact that transcripts from all 29 genes were detected in the blood stages might be due to the higher amount of RNA included. Since similar amounts of RNA from the other stages were used for RT-PCR assays, the analysis suggested that 70% (18/26) of the genes transcribed in the hepatic parasite were also transcribed in sporozoites. Second, the RT-PCR analysis confirmed the widespread occurrence of sense and anti-sense transcription ("s"/"a"), a phenomenon, hitherto described only for blood stages [16–18], that has now been extended to other parasite stages in the mammal and in the insect hosts. The pattern of "s"/"a" transcription differed between the stages, and was most frequently associated with the actively dividing forms (midgut sporozoites, sub-mature cultured liver schizonts) than with the quiescent salivary gland sporozoite or fully mature liver schizonts in vivo. We are currently undertaking more detailed analyses to confirm and better define these patterns of transcription. A functional significance of anti-sense transcripts in *Plasmodium*, is still a matter for speculation, although recent studies showing RNA polymerase II is specifically associated with this transcription, and that the relative levers of these transcripts vary with different genes [19] suggests that anti-sense transcripts might play a role in the life cycle of the parasite. Further investigations of transcriptional patterns in malaria parasites are clearly warranted.

Attempts to access the transcriptome of hepatic stage *Plasmodium* parasites have been recently initiated. A differential display approach was attempted for *P. yoelii*-infected hepatocytes [20], but the results obtained were disappointing in that the five sequences identified were negative against *Plasmodium* genomes by BLAST analyses, and one actually corresponded to a mouse sequence. The potential of a second approach based on fastidious laser microdissection of individual infected hepatocytes, was demonstrated using both *P. falciparum* [21] and *P. yoelii* [22] parasites. For this latter parasite, the EST set derived from this approach has become recently available (http://www.tigr.org/tdb/e2k1/pya1/) but interestingly none of the EST's or the corresponding genes was found to overlap with the transcripts we describe here. The reason may lie in the fact that we chose mature forms that had not yet reached liver merozoite formation from the pseudo-cytomeres, in order to increase the likelihood to obtain LS transcripts, whereas laser micro-dissection was performed on the largest schizonts in order to obtain

sufficient material, i.e. that had most likely neared their full development. This clearly demonstrates that the two approaches are complementary, and further indicates that major switches in transcriptional patterns occur within relatively short periods during hepatic stage development, in a manner similar to that observed in blood stage parasites [23]. Finally, in a very recently published study, transcripts were obtained from sporozoites that have rounded-up in axenic cultures [24]. Though cultured outside hepatocytes, the forms obtained are reminiscent of early liver stage trophozoites. This EST dataset, derived from the opposite spectrum of hepatic stage development, was also compared to our dataset. Only 15 of the 652 unique transcripts identified in the axenic cultures were observed in our dataset.

The approach we have validated by the present work offers advantages, over alternative strategies, that are of importance for future investigations of *P. falciparum* and further analysis of *P. yoelii*. First, our strategy would allow access to all developmental stages of the complex liver schizogony, whereas micro-dissection is mostly restricted to mature forms. Liver forms at various steps of maturation are of fundamental interest. They are of practical interest because they are both at the origin of protective immunity [13,25] and the main target of immune effectors [13]. Monitoring of the synchronous cultures can lead to isolating parasites with a precise degree of maturation. Second, parasite numbers accessible through in vitro culture exceed those obtained by micro-dissection by several orders of magnitude. High numbers could be recovered from liver biopsies, though at the cost of massively increased contamination by uninfected cells (ca. 100 fold), an aspect that will be critical not only for genomic but also for proteomic analysis of the hepatic stages. Moreover, these in vivo approaches would be far more difficult to apply to *P. falciparum* [26]. Third, the in vitro system is better suited to undertake further research, e.g. to study gene expression regulation, to dissect the influence of immune effectors, or to investigate the influence of drugs on gene expression.

In conclusion, our results demonstrated the feasibility of the approach, thus opening the way to extend studies to *P. falciparum*. The strategy employed provides a new and systematic means to investigate transcriptional patterns throughout the developmental stages of hepatic parasites. This is particularly desirable as our results indicated that complex patterns of gene expression occur at this pivotal yet paradoxically the least understood stage of the malaria life cycle.

Acknowledgments

This work was supported by the World Health Organisation, the Genopole of the Institut Pasteur, and by the Wellcome Trust. Anne Charlotte Grüner was partly supported by a grant from the Carlsberg Foundation (Denmark), and Stéphanie Hez-Deroubaix by a fellowship from CANAM (France).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.molbiopara.2005.03.019.

References

[1] Clyde DF, Most H, McCarthy VC, Vanderberg JP. Immunization of man against sporozite-induced falciparum malaria. Am J Med Sci 1973;266:169–77.
[2] Clyde DF. Immunization of man against falciparum and vivax malaria by use of attenuated sporozoites. Am J Trop Med Hyg 1975;24:397–401.
[3] Druilhe P, Marchand C. From sporozoite to liver stages: the saga of the irradiated sporozoite vaccine. New strategies in parasitology. Churchill Livingstone: K.P.W.J. McAdam; 1989. pp. 39–48.
[4] Clyde DF. Immunity to falciparum and vivax malaria induced by irradiated sporozoites: a review of the University of Maryland studies, 1971–75. Bull World Health Organ 1990;68:9–12.
[5] Guguen-Guillouzo C, Campion JP, Brissot P, et al. High yield preparation of isolated human adult hepatocytes by enzymatic perfusion of the liver. Cell Biol Int Rep 1982;6:625–8.
[6] Ozaki LS, Gwadz RW, Godson GN. Simple centrifugation method for rapid separation of sporozoites from mosquitoes. J Parasitol 1984;70:831–3.
[7] Millet P, Collins WE. Enhancement of in vitro infectivity of simian malaria sporozoites to hepatocytes by centrifugation. J Parasitol 1989;75:992–4.
[8] Rénia L, Mattei DM, Goma J, et al. A malaria heat-shock-like determinant expressed on the infected hepatocyte surface is the target of antibody-dependent cell-mediated cytotoxic mechanisms by nonparenchymal liver cells. Eur J Immunol 1990;20:1445–9.
[9] Carlton JM-R, Angiuoli SV, Suh BB, et al. Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii yoelii*. Nature 2002;419:512–9.
[10] Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 1998;8:175–85.
[11] Gardner MJ, Hall N, Fung E, et al. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature 2002;419:498–511.
[12] Rutherford K, Parkhill J, Crook J, et al. Artemis: sequence visualization and annotation. Bioinformatics 2000;16:944–5.
[13] Druilhe P, Rénia L, Fidock DA. Immunity to liver stages. In: Malaria: parasite biology, pathogenesis, and protection. I.W. Sherman, ASM Press; 1998. pp. 513–543.
[14] Kappe SHI, Gardner MJ, Brown SM, et al. Exploring the transcriptome of the malaria sporozoite stage. Proc Natl Acad Sci USA 2001;98:9895–900.
[15] Bruña-Romero O, Hafalla JCR, González-Aseguinolaza G, et al. Detection of malaria liver-stages in mice infected through the bite of a single Anopheles mosquito using a highly sensitive real-time PCR. Int J Parasitol 2001;31:1499–502.
[16] Patankar S, Munasinghe A, Shoaibi A, et al. 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. Mol Biol Cell 2001;12:3114–25.
[17] Kyes SA, Christodoulou Z, Pinches R, Newbold CI. Stage-specific merozoite surface protein 2 antisense transcripts in *Plasmodium falciparum*. Mol Biochem Parasitol 2002;123:79–83.
[18] Gunasekera AM, Patankar S, Schug J, et al. Widespread distribution of antisense transcripts in the *Plasmodium falciparum* genome. Mol Biochem Parasitol 2004;136:35–42.

[19] Militello KT, Patel V, Chessler AD, et al. RNA polymerase II synthesizes antisense RNA in *Plasmodium falciparum*. RNA 2005;11:1–6.
[20] Lau AOT, Sacci Jr JB, Azad AF. Retrieving parasite specific liver stage gene products in *Plasmodium yoelii* infected livers using differential display. Mol Biochem Parasitol 2000;111:143–51.
[21] Semblat J-P, Silvie O, Franetich J-F, et al. Laser capture microdissection of *Plasmodium falciparum* liver stages for mRNA analysis. Mol Biochem Parasitol 2002;121:179–83.
[22] Sacci Jr JB, Aguiar JC, Lau AOT, Hoffman SL. Laser capture microdissection and molecular analysis of *Plasmodium yoelii* liver-stage parasites. Mol Biochem Parasitol 2002;119:285–9.
[23] Bozdech Z, Llinas M, Pulliam BL, et al. The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. PLoS Biol 2003;1:85–100.
[24] Wang Q, Brown S, Roos DS, et al. Transcriptome of axenic liver stages of *Plasmodium yoelii*. Mol Biochem Parasitol 2004: 137.
[25] Mellouk S, Lunel F, Sedegah M, et al. Protection against malaria induced by irradiated sporozoites. Lancet 1990;335:721.
[26] Meis JF, Ponnudurai T, Mons B, et al. *Plasmodium falciparum*: studies on mature exoerythrocytic forms in the liver of the chimpanzee, Pan troglodytes. Exp Parasitol 1990;70:1–11.