



## Original Article

Gene regulation of the avian malaria parasite *Plasmodium relictum*, during the different stages within the mosquito vectorVaishnovi Sekar<sup>a,\*,1</sup>, Ana Rivero<sup>b,c</sup>, Romain Pigeault<sup>a,e</sup>, Sylvain Gandon<sup>d</sup>, Anna Drews<sup>f</sup>, Dag Ahren<sup>g</sup>, Olof Hellgren<sup>f,\*</sup><sup>a</sup> Department of Biology, Lund University, Sweden<sup>b</sup> MIVEGEC (CNRS – Université de Montpellier – IRD), 34394 Montpellier, France<sup>c</sup> CREES (Centre de Recherche en Ecologie et Evolution de la Santé), 34394 Montpellier, France<sup>d</sup> CEFE (CNRS - Université de Montpellier - Université Paul-Valéry – EPHE - IRD), Montpellier, France<sup>e</sup> Department of Ecology and Evolution, CH-1015 Lausanne, Switzerland<sup>f</sup> MEMEG, Department of Biology, Lund University, Sweden<sup>g</sup> National Bioinformatics Infrastructure Sweden (NBIS), SciLifeLab, Department of Biology, Lund, Sweden

## ARTICLE INFO

## Keywords:

Transmission

*Plasmodium relictum**Culex pipiens quinquefasciatus*

Extrinsic incubation

## ABSTRACT

The malaria parasite *Plasmodium relictum* is one of the most widespread species of avian malaria. As in the case of its human counterparts, bird *Plasmodium* undergoes a complex life cycle infecting two hosts: the arthropod vector and the vertebrate host. In this study, we examined transcriptomes of *P. relictum* (SGS1) during crucial timepoints within its vector, *Culex pipiens quinquefasciatus*. Differential gene-expression analyses identified genes linked to the parasites life-stages at: i) a few minutes after the blood meal is ingested, ii) during peak oocyst production phase, iii) during peak sporozoite phase and iv) during the late-stages of the infection. A large amount of genes coding for functions linked to host-immune invasion and multifunctional genes was active throughout the infection cycle. One gene associated with a conserved *Plasmodium* membrane protein with unknown function was upregulated throughout the parasite development in the vector, suggesting an important role in the successful completion of the sporogonic cycle. Gene expression analysis further identified genes, with unknown functions to be significantly differentially expressed during the infection in the vector as well as upregulation of reticulocyte-binding proteins, which raises the possibility of the multifunctionality of these RBPs. We establish the existence of highly stage-specific pathways being overexpressed during the infection. This first study of gene-expression of a non-human *Plasmodium* species in its vector provides a comprehensive insight into the molecular mechanisms of the common avian malaria parasite *P. relictum* and provides essential information on the evolutionary diversity in gene regulation of the *Plasmodium*'s vector stages.

## 1. Introduction

*Plasmodium* parasites are best known for the dramatic mortality and morbidity they cause in humans across the Southern hemisphere. This group of parasites can also be found infecting a diverse range of other hosts including non-human primates, bats, rodents, reptiles, and birds [1,2]. But for a few minor differences [3], all these species share a nearly identical life cycle, with an asexual replicative stage in the vertebrate host, and an sexual stage in a blood-sucking culicid mosquito (Diptera: Culicidae). The avian *Plasmodium* clade includes some of the world's

most genetically diverse [4] and virulent [5] of all malaria parasites known to date [6,7] and shows a large differentiation both in its geographical and host distribution [6,8].

Avian malaria parasites have played a key role in our comprehension of the prevalence, morbidity and epidemiology of the disease in natural populations [7,9]. Due to their high prevalence, widespread distributions and diverse host ranges, they have also been used to address several evolutionary issues such as host-parasite co-evolution [10,11], virulence evolution [12,13], and sexual selection [14]. To date there are more than 40 morphologically described species of avian *Plasmodium*

\* Correspondence to: Hellgren O, MEMEG, Department of Biology, Sölvegatan 37, 22362 Lund, Sweden.

\*\* Correspondence to: Sekar V, Tomtebodavägen 23a, 171 65 Solna, Sweden.

E-mail addresses: [vaishnovi.sekar@scilifelab.se](mailto:vaishnovi.sekar@scilifelab.se) (V. Sekar), [olof.hellgren@biol.lu.se](mailto:olof.hellgren@biol.lu.se) (O. Hellgren).<sup>1</sup> Current address: Science for Life Laboratory, Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, Sweden.

and over 1200 cytochrome *b* lineages, the vast majority of which are thought to be reproductively isolated entities [4,15]. Of the avian malaria parasites, *P. relictum* is the most prevalent and widespread morphospecies with a highly diverse host range [16], which places it amongst the top 100 most invasive species [17]. This parasite species has also been found to be associated with the decline and extinction of several bird species on the islands of Hawaii [18,19], with the mortality, in wild, endemic, and indigenous birds in New Zealand [20] and the mortality of penguins in zoos across the world [21]. Of the different mitochondrial cytochrome *b* lineages described to date within the morphologic species of *P. relictum*, SGS1 the most common in terms of geographic range, host range and host prevalence. This lineage has been found to infect 129 different bird species, (MalAvi, 2020-05-08, [4]) and may cause severe disease and even mortality in wild birds [22,23].

All *Plasmodium* parasites share a similar life cycle requiring two infection cycles: one in the arthropod vector and one in the vertebrate host. When competent mosquitoes take an infected blood meal, they ingest both male (*microgametocytes*) and female (*macrogametocytes*) parasites. Within the first few minutes, the gametocytes transform into *gametes* and fuse within the midgut to form a diploid *zygote*. The zygotes, in turn, undergo meiosis, become motile and elongated (*ookinetes*), traverse the wall of the midgut and start to develop into *oocysts*. Over the course of several days, the oocyst undergoes several rounds of mitosis to create a syncytial cell with thousands of nuclei. In a massive cytokinesis event, thousands of *sporozoites* erupt from each oocyst and migrate through the haemocoel towards the salivary glands [24]. Between 10 and 14 days after the initial infected blood meal, the parasite is in the salivary glands ready to be transmitted to a new host. When the mosquito takes a second blood meal, it injects sporozoites into the blood of the new host. These sporozoites multiply in various organs and blood cells of the host and a few days later end up releasing daughter parasites called *merozoites* into the bloodstream. Merozoites will continue the cycle of the parasite by invading other red cells and eventually produce the micro- and macrogametocytes which will restart the cycle in the mosquito [7].

Within the mosquito, the parasite faces several demographic bottlenecks that are crucial determinants of its onward transmission [25–28]: 1) the transition between the ingested micro (male) and macro (female) gametes to the formation of motile ookinetes, 2) the transition between the ookinete to the oocysts and 3) the transition between oocysts to the sporozoites in the salivary glands [7]. The molecular mechanisms underlying the different developmental stages of the parasite within the mosquito have to date only been studied in human malaria [29] and a few species of rodent malaria species using mosquito-parasite combinations not found in nature [30]. Therefore, in order to be able to find and study traits and genetic mechanisms that either have been conserved throughout the genera of *Plasmodium* or finding unique mechanisms linked to host or vector specificity, there is a need for knowledge of the parasite's molecular mechanisms in the mosquito outside the limited group of parasites that have been studied to-date [26,31–36].

Here, we set out to identify and study the first full transcriptome profile of a non-mammalian malaria *Plasmodium* as it goes through its different life stages in its mosquito vector. First, we carried out an experiment to establish the temporal dynamics of *P. relictum* (SGS1) development within its vector, the mosquito *Culex pipiens quinquefasciatus*. With this data in hand, we then used RNA sequencing to obtain the parasite's transcriptome profiles at key points of its development within the mosquito, namely: 1) in the first minutes after ingestion (30 min post infection on average, henceforth 30mpi), 2) during the peak oocyst production phase (8 days post infection, henceforth 8dpi), 3) during the peak sporozoite production phase (12dpi) and 4) during the latest stages of the mosquito infection when the sporozoites are mainly present in the salivary glands (22dpi). As a reference point, we also obtain the transcriptome of an infected blood sample from the host bird taken immediately before the mosquito feed. We discuss how our

observations will open up avenues of investigation into the pathways and molecular mechanisms that have been evolutionary conserved since the mammalian-avian malaria split.

## 2. Methods

### 2.1. Temporal dynamics of *Plasmodium* sporogony protocol

To establish the temporal dynamics of *P. relictum* (SGS1) development within its vector, the mosquito *Culex pipiens quinquefasciatus*, experiments were carried out using three birds (*Serinus canaria*) which were infected using standard infection protocols [37]. Ten days later, at the peak of the acute infection stage, each bird was placed overnight in a separate experimental cage with 180 7-day old female laboratory strains of *Culex pipiens quinquefasciatus* mosquitoes [38]. After the mosquitoes had taken a blood meal, the birds and unfed mosquitoes were removed from the cages and the blood fed mosquitoes were supplied with 10% ad libitum sugar solution until the end of the experiment. Every two days, starting on day 4 and finishing on day 26 post-blood meal, twelve mosquitoes were randomly sampled from each cage. Each mosquito was dissected to count the number of oocysts in the midgut, and its head-thorax was preserved at -20 °C for the quantification of the transmissible sporozoites. Developing oocysts were counted under the microscope [38]. Sporozoites were quantified using real-time quantitative PCR as the ratio of the parasite's *cytb* gene relative to the mosquito's *ace-2* gene [39].

### 2.2. Parasite transcriptomic experimental protocol

This separate experiment was carried out using three new birds (*Serinus canaria*) which were infected using standard infection protocols [37]. Ten days later, at the peak of the acute infection stage, each bird was placed in a separate experimental cage with 150 7-day old female mosquitoes. On the day of the experiment (day 0), immediately before the bird was introduced into the cage, 20 µl blood was collected from the wing. A drop of this blood was used to quantify parasite load via a blood smear (as described in [7]), the rest was mixed with 500 µl of Trizol (Life Technologies, Carlsbad, CA, USA) and frozen at -80 °C for subsequent RNA extraction (henceforth 'bird' sample). Cages were visited 30 min after the bird's introduction into the cage and 20 fully-gorged resting mosquitoes were randomly sampled from each cage (this is our 30 min post infection (30mpi) sample). When the number of fully gorged mosquitoes found in a given cage after 30 min was <20, the cage was visited a second time 30 min later. Half of the mosquitoes sampled at 30mpi were immediately homogenised, mixed with 500 µl of Trizol and frozen at -80 °C for subsequent RNA extraction (1 tube per cage). The bird and unfed mosquitoes were taken out at the same time, and the blood fed mosquitoes were left in the cage with a source of sugar solution (10%) at our standard insectary conditions (25–27 °C, 70% RH). On day 8 after the start of the experiment, to coincide with the peak of oocyst production as estimated by the previous experiment, 30 further mosquitoes were randomly taken from each of the cages (henceforth '8 dpi' sample). For each cage, 10 of these mosquitoes were homogenised and conserved in 500 µl RNAlater (ThermoFisher Scientific/Ambion, Waltham, USA). The procedure was repeated on day 12 to coincide with the peak sporozoite production ('12 dpi' sample) and on day 22 during the late stage of the infection ('22 dpi' sample). The rationale behind that Trizol was used for the blood meal and the bird blood but not for the rest of the mosquito extractions stem from the fact that bird blood have nucleated erythrocytes and using Trizol in the initial step avoids clogging of the spin columns membranes during the initial stages of extraction.

### 2.3. RNA extraction

RNA from the bird blood samples was extracted with a combination

of TRizol LS and RNeasy Mini spin columns (Qiagen, Hilden, Germany). Homogenizing and phase separation were done according to the TRizol LS manufactures protocol resulting in an aqueous phase which was then mixed with one volume of 70% ethanol and placed in a RNeasy Mini spin column. From this point on the RNeasy Mini spin columns manufacturer's protocol was followed.

RNA from the 30mpi samples (collected in Trizol LS) were disrupted and homogenised using a TissueLyser (Qiagen, Hilden, Germany). Ten whole mosquitoes were collected in 500 µl of Trizol LS, the total volume of TRizol LS was adjusted to 750 µl and a 5 mm stainless steel bead was added. The TissueLyser was run for two cycles of 3 min at 30 Hz. Phase separation was done according to the TRizol LS manufactures protocol resulting in an aqueous phase which was then mixed with one volume of 70% ethanol and placed in a RNeasy Mini spin column. From this point on the RNeasy Mini spin columns manufactures protocol was followed.

RNA from the 8dpi, the 12dpi and 22dpi samples (collected in RNAlater) were extracted with RNeasy Mini spin columns following the manufactures protocol. Disruption and homogenization were done using a TissueLyser (Qiagen, Hilden, Germany). Ten whole mosquitoes were moved to a new tube and 600 µl of buffer RLT was added as well as a 5 mm stainless steel bead. The TissueLyser was run for two cycles of 3 min at 30 Hz.

The concentration of all RNA samples was measured on a Nanodrop 2000/2000c (Thermo Fisher Scientific, Wilmington, DE, USA). Dried samples were shipped on dry ice to Novogene (Hong Kong) for mRNA sequencing. 250–300 bp insert cDNA library where prepared by Novogene (Hong Kong) after removal of rRNA using ribo-zero kit. mRNA from each time point was sequenced using Illumina HiSeq platform at an average of 85 M reads per library. We obtained paired-end reads of 150 bp length for 15 sequenced transcriptomes which were used for further analysis.

#### 2.4. Data analysis

Raw RNA-Seq reads were assessed for quality, per-base sequence content, adapter content and overrepresented sequences using FastQC (Version 0.11.5) [40]. The raw sequence data has been submitted to the SRA database and can be accessed using BioProject id: PRJNA724741. Contamination due to other genomes such as human and bird was estimated using FastQ Screen (Version 0.11.1) [41]. FastQ Screen maps the raw-reads to the indexed reference genome using Bowtie 2 [42] or BWA [43] to calculate an estimate of the reference genome contamination in query sequences. Birds infected with the parasites in our study was *Serinus canaria* (GCF\_000534875.1\_SCA1) hence we used this genome for screening transcripts along with Zebra Finch (GCF\_000151805.1\_Taeniopygia\_guttata-3.2.4) genome which is widely used as bird model genome, along with mosquitoes (GCF\_000209185.1\_CulPip1.0) for malaria vector and *P. relictum* genome (GCA\_900005765.1\_PRELSG). The genomes were retrieved in the form of Fasta sequences from NCBI Genome browser (<https://www.ncbi.nlm.nih.gov/genome>). All the reads that did not map uniquely to bird and malaria vector reference genomes were filtered out using -filter function in FastQ Screen. Adapter sequences, ambiguous nucleotides, and low-quality sequence were trimmed off using Trimmomatic (Version 0.36) [44]. We used paired-end mode with options -phred33 for base quality encoding, a sliding window of 4:20, the minimum sequence length of 70, trailing 3 and leading 3. We used the standard Illumina adapter sequences (option: ILLUMINACLIP Truseq3-PE.fa:2:30:10) available in Trimmomatic to screen and trim the adapter sequences from the query. The choice of parameters was made after examining the reads for quality and presence of adapter sequences. Once trimmed, the sequences were again evaluated for quality (FastQC) before proceeding to the following steps. MultiQC [45], which summarizes results from various tools and generate a single report, was used for illustrative purpose.

The filtered and trimmed sequence files were aligned to *P. relictum*

published genome (GCA\_900005765.1\_PRELSG) using HISAT2 (Version 2.1.0) [46]. All the sequence files were aligned as paired end reads and with default parameter settings. HTseq-count [47] (parameters: -s no -t gene -i locus tag) was used to calculate the read count for each sample and custom bash scripts were written to post-process the files to generate gene count matrix. Gene count table is submitted in the supplementary files. Scripts and more detailed parameter explanation for this process can be provided upon request.

The differential gene expression analysis was performed in DESeq2 (Version 1.16.1) [48] package in R. Genes with no expression across all samples i.e. genes with zero read across all samples were filtered out from the gene count table before analysing the data further. To avoid having any bias in the analysis and to examine the data further, variance stabilized transformation of counts [49] was used to perform PCA. Plots and visualization were made using RColorBrewer, genefilter, dplyr, and ggplot2 packages in R.

Differential gene expression (DGE) was performed with the filtered gene count matrix. The bird samples were considered as a control to perform DGE analysis. The four cases were defined for the DGE analysis and genes were considered differentially expressed for adjusted p-value <0.01. The significantly differentially expressed upregulated genes (having positive log<sub>2</sub> fold change in our comparison) after foldchange and adjusted p-value sorting, from the four cases (30mpi vs bird, 8 days post-infection (dpi) vs bird, 12dpi vs bird, 22dpi vs bird) were considered for further gene ontology enrichment analysis. The genes from all the cases were compared cross-wise to identify genes common in all stages and the genes exclusive to a certain stage of infection.

We performed Gene Ontology Enrichment analysis to look for functions that were significantly overrepresented amongst the upregulated genes at the different stages. The significantly differentially expressed upregulated genes were analyzed to identify overrepresented Gene Ontology (GO) terms using the topGO package [50] in RStudio (Version 1.0.143). We used GO annotation for SGS1-like strain downloaded from PlasmoDB (<http://plasmodb.org/plasmo/>) as custom background for enrichment analysis. The background for analysis for each case is the list of genes and their p-adjusted valued as reported by DESeq2. All the significantly differentially expressed upregulated genes from each case was analyzed independently to identify the ontologies enriched specifically for each stage of the parasite's life cycle. The enrichment analysis is performed using classical algorithm and Fisher's exact test and the GO terms with p value <0.01 are considered enriched. The overrepresented GO were reported from all three domains: biological processes, molecular functions and cellular components.

### 3. Results

#### 3.1. Temporal dynamics of Plasmodium development in mosquitoes

The temporal dynamics of *Plasmodium* inside the mosquito were consistent between mosquito groups fed on the three infected birds (Fig. S1, Table S1). Oocysts first become visible in the mosquito midgut on days 4–6 after the blood meal and the peak oocyst burdens were reached on days 8–10 (Fig. 1, Fig. S1). Thereafter a rapid decrease in the number of oocysts was observed so that by day 12 post blood meal, only between 2 and 20% of the peak oocyst burden remains. The first sporozoites appear on the head-thorax homogenate on days 8–10, only 2–4 days after the first oocysts. Sporozoite concentrations reach the peak on day 12 (Fig. 1), except for the mosquitoes fed on the bird with the lowest parasite load (Bird\_3, Fig. S1) where the peak is reached much later (day 16, Fig. S1). Onwards the number of sporozoites decreases steadily but are still detectable, albeit at low numbers, three weeks after the initial infection (Fig. 1).

#### 3.2. Data preprocessing and sequence alignment

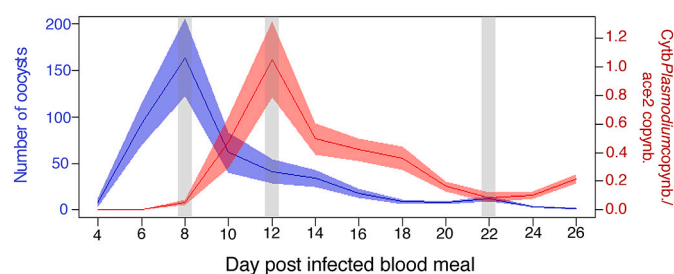
We assessed the sequence data for quality and contamination.

FastQC did not report any sequences having an overall Phred Score lower than 24, however, six samples had a drop in the quality for last 8–18 bases (Fig. S2, Table S2), which were trimmed using Trimmomatic. FastQ Screen mapped the reads with an average of 13% to the Zebra finch, 20% to Canary and 91% to the malaria vector (mosquito) reference sequences. Majority of the reads from the blood samples obtained from infected bird immediately before the mosquito feed mapped to the bird reference sequences (Fig. S3A; Table S3). After removal of reads that uniquely mapped to the birds and vector genomes, an average of 2% of the remaining reads mapped to the parasite genome (around 0.65 Million reads) and were retained for further analysis (Fig. S3B; Table S4). On average 71% of paired-end reads were retained after trimming for low-quality reads and adapter content (Table S5). GC content-based filtering for the transcriptomic data was not performed as it has been estimated that the GC percentage of coding sequences (CDSs) are greater than the genomic GC percentage in the *Plasmodium* spp. [51]. The average GC content and average length of sequences processed by FastQ Screen and Trimmomatic were 37% and 138 bp, respectively (Table S6). HISAT2 produced an average of 77% overall alignment for the transcriptome data aligned to *P. relictum* genome (Table S7).

### 3.3. Transcript assembly and differential gene expression analysis

Transcript assembly and abundance calculation reported 5286 genes (that we found a higher number of genes than present in the annotated genome is due to that alternative splice variants is interpreted as different genes) in the gene count table, which include mitochondrial and apicoplast genes (Table S8). Of these, 160 of the genes found in the annotated genome had no expression data in any of the samples and were dropped from further analysis as they did not contribute to any information to the analysis. The remaining 5126 genes were analyzed for differential gene expression using the DESeq2 package in the R statistical suite. A principal component analysis of expression clustered parasite samples belonging to the same time point together (Fig. 2). The different timepoints fall on four different non overlapping PC1 (53% variance) intervals, where the expression in the bird, 30mpi, 8dpi each have their own interval whereas 12 and 22dpi share overlapping PC1 values, which explained the time from the start of infection and is indicative of parasite development (30mpi -> 8dpi -> 12dpi -> 22dpi). However, when also including the PC2 values, each time point occupies a unique non-overlapping area on the PC1/PC2 space.

Four time points (30mpi, 8dpi, 12dpi, and 22dpi) were compared for *P. relictum* genes differentially expressed against the bird samples. As baseline control, we used the transcriptomic profile of the parasite in the vertebrate host immediately before the blood meal (bird samples). The analysis reported 71, 311, 605 and 421 significantly differentially



**Fig. 1.** Temporal dynamics of *Plasmodium* development in mosquitoes. Average oocyst burden (blue) and sporozoite (red) counts at each dissection day. Blue and red shadows represent standard error. The left axis represents the average number of oocysts counted per female. The right axis represents the amount of sporozoites quantified by qPCR as the ratio of the parasite's *cytb* gene relative to the mosquito's *ace-2* gene. Light grey areas represent three of the five sampling points used to study the *Plasmodium* transcriptome. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

expressed genes for each of the time points, out of which 8, 96, 202 and 166 genes were upregulated in 30mpi, 8dpi, 12dpi and 22dpi vs bird respectively (Tables S9–S12). Over 60% of the genes were found to be downregulated (Fig. 3; Fig. S4). Genes that were significantly downregulated with respect to the baseline, were interpreted as being highly expressed genes that are linked to the development within the vertebrate host (Fig. 4). For this study, we therefore only considered the genes that were significantly upregulated within the mosquito as compared to our baseline controls.

Two genes were reported to be upregulated at all time points versus bird. One of these genes had a known function associated with 28S ribosomal RNA and the other one was a conserved *Plasmodium* membrane protein with unknown function. The genes involved in sporozoite invasion, TRAP-like protein, early transcribed membrane protein, cysteine repeat modular protein, oocyst capsule protein Cas380, p25-alpha family protein, Circumsporozoite protein (CSP) were upregulated in all stages except the 30mpi stage. Two reticulocyte binding proteins were upregulated during all stages except the 30mpi stage. Several other conserved *Plasmodium* proteins with unknown function were also reported to be upregulated across timepoints. These results are summarized in Supplementary File Table S13, and detailed information about the corresponding genes is included in the Supplementary File Table S14.

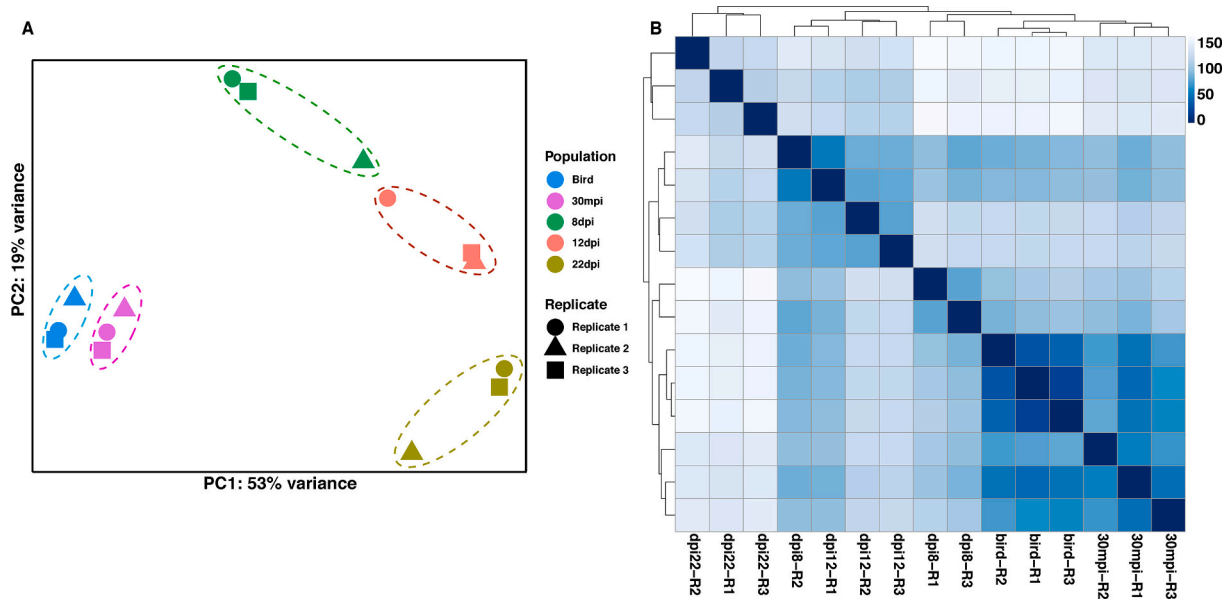
### 3.4. Gene ontology enrichment

Annotations for 2583 *P. relictum* genes were retrieved from PlasmoDB database. Gene ontology (GO) enrichment analysis revealed distinct GO enriched at different time points (Fig. 5; Fig. S5; Fig. S6). GO terms associated with pseudouridine synthesis were enriched for the 30mpi time point, several GO terms related to oxidoreductase activity were enriched for 8dpi, GO terms related to rhoptry neck, entry into host cell (vector in our study) and membrane part were enriched for 12dpi, and GO terms related to locomotion, regulation of RNA metabolic process and lipid metabolic processes were enriched for 22dpi. GO terms related to DNA-binding transcription factor was enriched from 8dpi to 22dpi. GO terms related to exit from host, movement in host environment (vector in our study) and protein kinase activity were enriched during the late stages of infection (12dpi and 22dpi). Most of the significantly overexpressed GO terms were supported by only a few genes (Tables S15–S18).

## 4. Discussion

*Plasmodium* undergoes complex molecular processes during its life-cycle to successfully invade the vector and overcome its defence mechanisms to continue the infection cycle. However, our knowledge about the processes, pathways and enzymes involved during these stages in non-human malaria parasites is limited to studies of rodent malaria parasites using vector-parasite combinations not found in nature [30]. Given the well-documented differences in parasite performance between natural and non-natural *Plasmodium*-mosquito combinations [52–54], these latter results have to be interpreted with caution. In this study, we study the transcriptome of *P. relictum* in its vector: the *Culex pipiens* mosquito, which is the closest to the tropical form found in nature.

The average transcriptome GC content for the reads retained after quality control and that mapping to the *P. relictum* genome was calculated to be higher than that of *P. falciparum* and the avian malaria genome of *P. ashfordi* [32] (Fig. S7). This could be due to the variation in codon usage bias and gene regulation mechanisms between the species as explained by Yadav et al. [51]. The PC analysis (Fig. 2) shows how samples occupies discrete areas on the PC1/PC2 space based on time-points, as expected when same time points share similar expression profiles as a result of similar biological responses. Two samples, one from 8dpi and one from 12dpi, corresponding to the peak oocyst and peak sporozoite formation (Fig. 1), clustered intermediate between



**Fig. 2.** A. PCA plot with variable stabilizing transformation. The samples cluster together around each the time point (different stages of parasite life cycle in the mosquito). However, one 8dpi sample (from replicate 2) and one 12dpi sample (from replicate 1) cluster close to each other and away from their corresponding time point. These results are supported by the heatmap (B). 2B: Heatmap portraying Euclidian distance measured between the different samples. Lighter colour indicates greater distance.

these two time points, indicating an active transition stage. The fact that these two samples originate from different individual birds, suggests that this is due either to a delay (in one of the 12 dpi sample) and an early onset (of the 8 dpi sample) of the infection in the mosquitoes due to differences between the birds resulting in variation in the timing of transitions to the different life stages. These differences could stem from differences in bird parasitemia which are known to influence gene expression of the parasites in the host [32]. We note that both of these intermediate samples come from the two birds with the varying parasitemia levels (bird #2: 4.4% and birds #1: 9.5%, bird #3: 9.3%).

#### 4.1. Invasion genes are expressed throughout the sporogonic cycle

The majority of the genes in the parasite genome, apart from 160 genes, are expressed at one stage or other in the parasite's life cycle within the mosquito. The differential gene expression analysis revealed that less than 40% of the total genes were upregulated at a given time point during the infection cycle in the vector as compared to our baseline control. The transition from gametocyte activation in the mosquito midgut (30mpi) to the salivary gland invasion by sporozoites (8dpi to 22dpi, Fig. 1) is associated with the expression of several important genes specific to mobility, entry into the host, DNA transcription and pathogenesis (Table S13). These include a putative sporozoite invasion associated protein, which is reported to participate in host-pathogen interactions during cell-traversal [55], a Thrombospondin-related protein 1 and a Cysteine repeat modular protein 4 (CRMP4) which facilitate the salivary gland invasion. These genes have been found in previous human malaria [56,57] and rodent malaria studies [57–59] indicating that these genes are evolutionary conserved. The oocyst growth and the sporozoites release are concomitant during a long period of time (Fig. 1). We observe several genes which are associated with sporozoite development to be upregulated throughout the 8dpi-22dpi transition. Consistent with previous reports [60–64], our analysis revealed Circumsporozoite protein (CSP) to be upregulated during these stages. CSP is a key gene having a multifunctional role in oocyst development, formation of sporoblast and sporozoites, salivary gland infection, onset of sporozoites invasion, sporozoite mobility, salivary gland invasion and hepatocyte infection [25,61,62,65]. Interestingly, three genes associated TRAP-like proteins (TLP) were also found to be upregulated during

these stages, which is in line with studies that have reported TLP to an important player in sporozoite cell traversal [66]. Oocyst capsule protein Cap380, reported to be active during the oocyst development, sporozoite differentiation and disruption of this genes may affect parasite's ability to invade host or vector cells [67], was also in our study upregulated during this transition highlighting its essential role in the parasite growth within the vector. In contrast to previously reported, two Reticulocyte-Binding Proteins (RBPs) which are involved in red blood cell invasion [32], were upregulated during 8dpi-22dpi transition. Due to sequence divergence, we cannot know with certainty which of the RBPs is the orthologous gene to the RBP we identified in this study. This suggests, either that these genes, which play a crucial role in the blood stages of the parasite, have pleiotropic effects in the mosquito, or that they are RBPs-like genes that have acquired another function in the avian malaria system.

#### 4.2. Genes upregulated during the different stages in vector

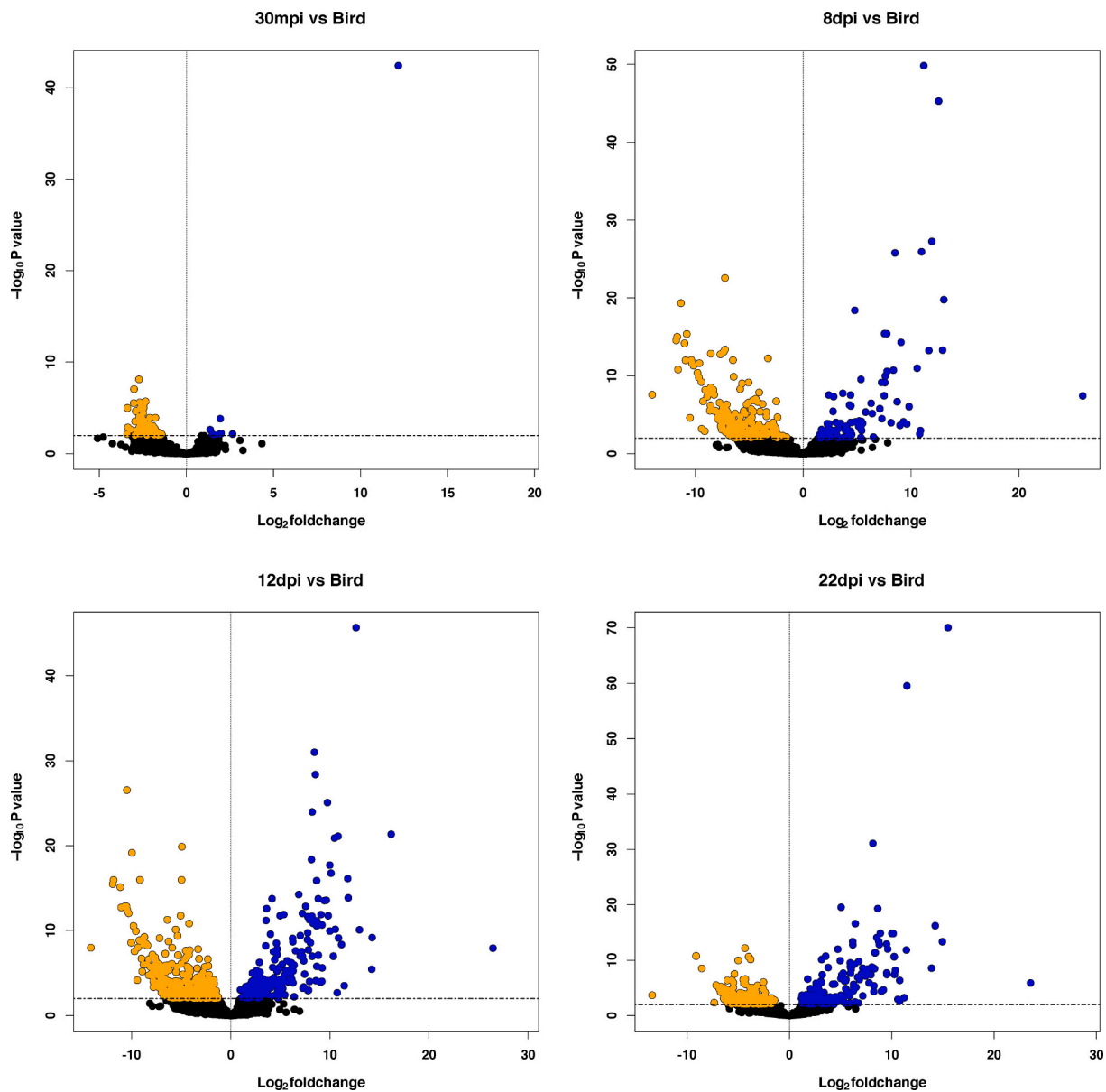
Along with the number of invasive genes, several genes were specifically upregulated during a certain timepoint (Table S13).

##### a. During 30mpi

Our analysis also reported genes which are exclusive to each time point (Table S13). We identified 4 genes upregulated exclusively during the 30mpi stage. One such gene, Putative schizont egress antigen-1, is exclusively upregulated and is amongst genes upregulated only at this time point. Studies with *Plasmodium falciparum* have reported that inhibiting this gene prevents the schizonts from leaving the infect RBCs thereby affecting infection [68]. Whether this affects the zygote production by reducing differentiation of male and female gametocytes is an interesting question to explore. Three other conserved *Plasmodium* protein with unknown functions were also upregulated during this stage suggesting a direct or indirect role in successful parasite infection and/or zygote development.

##### b. During 8dpi

At the 8dpi time point, we observed an upregulation of several genes



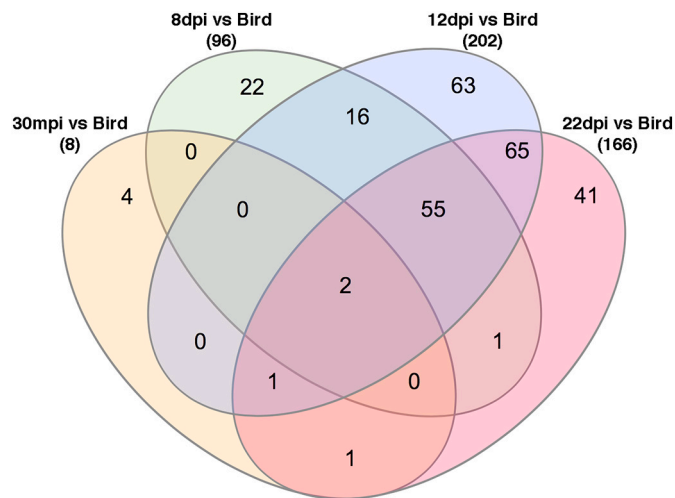
**Fig. 3.** Volcano plots showing  $\log_2$  fold change in expression on the x-axis and P-adjusted values on the y-axis for each of the 4 time points. Each dot represents a different gene. Differentially expressed genes with an adjusted p-value  $< 0.01$  and absolute fold change  $> 0$  are coloured blue (upregulated) and differentially expressed genes with an adjusted p-value  $< 0.01$  and absolute fold change  $< 0$  are coloured orange (downregulated). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

involved in translation initiation: ATP dependent RNA helicase DBP1, and eukaryotic translation initiation factor 4E. We also found several genes associated with carbon and TCA cycle. These included NADP-specific glutamate dehydrogenase, succinyl-CoA ligase, phosphoenolpyruvate carboxyl kinase. Previous studies in *P. falciparum* have reported that carbon and TCA cycle are critical for oocyst production and maturation [31]. We also identified a putative gene, coproporphyrinogen-III oxidase with GO term association to oxidation-reduction process. Previous studies in *P. falciparum* have demonstrated the role of this genes in a hybrid Heme biosynthetic pathway [69] and the importance of this pathway in parasite survival in mosquito stages [70]. We suspect that this gene could be conserved between human and avian malaria and involved in critical processes for parasite growth under stress condition such as against the host defence mechanism in mosquito vector. As observed in the temporal dynamics of the parasite development in the vector, our analysis supports the hypothesis that this stage is marked by the peak oocyst production in the infection cycle and

suggests that these genes might have a critical role in cell regulation metabolism in *Plasmodium* species.

### c. During 12dpi

Several genes involved in oocyst rupture were reported to be significantly upregulated exclusively during the stage which is marked by the peak of oocysts burst and sporozoite release (12dpi). These included oocyst rupture protein 1 and oocyst rupture protein 2 [71]. Genes involved in host cell invasion, such as putative rhostry neck protein (RON6), merozoite surface protein 1 paralog and rhomboid protease ROM1 were also found upregulated during this stage. Other upregulated genes included lysine-specific histone demethylase, which is associated with oxidoreductase activity, putative inorganic anion exchange, involved in sulphate transmembrane transporter activity, and cGMP-specific phosphodiesterase. A previously reported protein involved in mobility and infectivity of the ookinetes, CDPK3 [72], and



**Fig. 4.** Venn diagram representing a crosswise comparison of upregulated genes in each of the 4 time points. Orange (30mpi), Green (8dpi), Blue (12dpi) and Red (22dpi). 2 genes are upregulated during all stages of infection versus bird whereas 4, 22, 63 and 41 genes are exclusively upregulated during the 30mpi, 8dpi, 12dpi and 22dpi timepoints versus bird, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

one conserved *Plasmodium* protein associated with microtubule-based movement was also upregulated along with several other conserved *Plasmodium* proteins with unknown functions, suggesting their role in oocyst maturation and rupture, sporozoite formation, release and transportation.

#### d. During 22dpi

The last time point (22dpi) characterised by the low density of sporozoites. During this time, the sporozoites are mainly present in the salivary glands. Putative sporozoite and liver stage asparagine-rich protein (SLARP) and CDPK5 are upregulated during this timepoint. These proteins which are reported to have a role in the regulation of transcription, have been previously described in *Plasmodium falciparum* [73,74]. Upregulation of these genes and other genes involved in kinase activity – serine/threonine protein kinase, inositol polyphosphate multikinase, suggests an essential role in invasion mechanism during this transition. Some of the previously reported stage specific markers identified in other *Plasmodium* species was not observed in our study, such as UOS3 [75–77], MAEBL [78], CelTOS [79]. If this pattern is due to methodological reasons or a true biological difference, we can't determine at this point but is important to disentangle in future studies in order to understand how the gene functions have evolved across the different species.

#### e. During 8dpi and 12dpi

Secreted ookinete protein 25 was upregulated at both these time point (8dpi and 12dpi) which are characterised by peak oocyst formation and peak sporozoite concentration, respectively. Secreted ookinete protein 25 is reported to affect ookinete formation and the formation of midgut oocysts [80]. Other genes that are upregulated during this transition are genes involved in motility, *Plasmodium* exported protein (PHIST), asparagine-rich antigen, neurotransmitter, oxidation-reduction process and surface related antigen SRA.

#### f. During 12dpi and 22dpi

Genes involved in sporozoite maturation and cell invasion, copies of

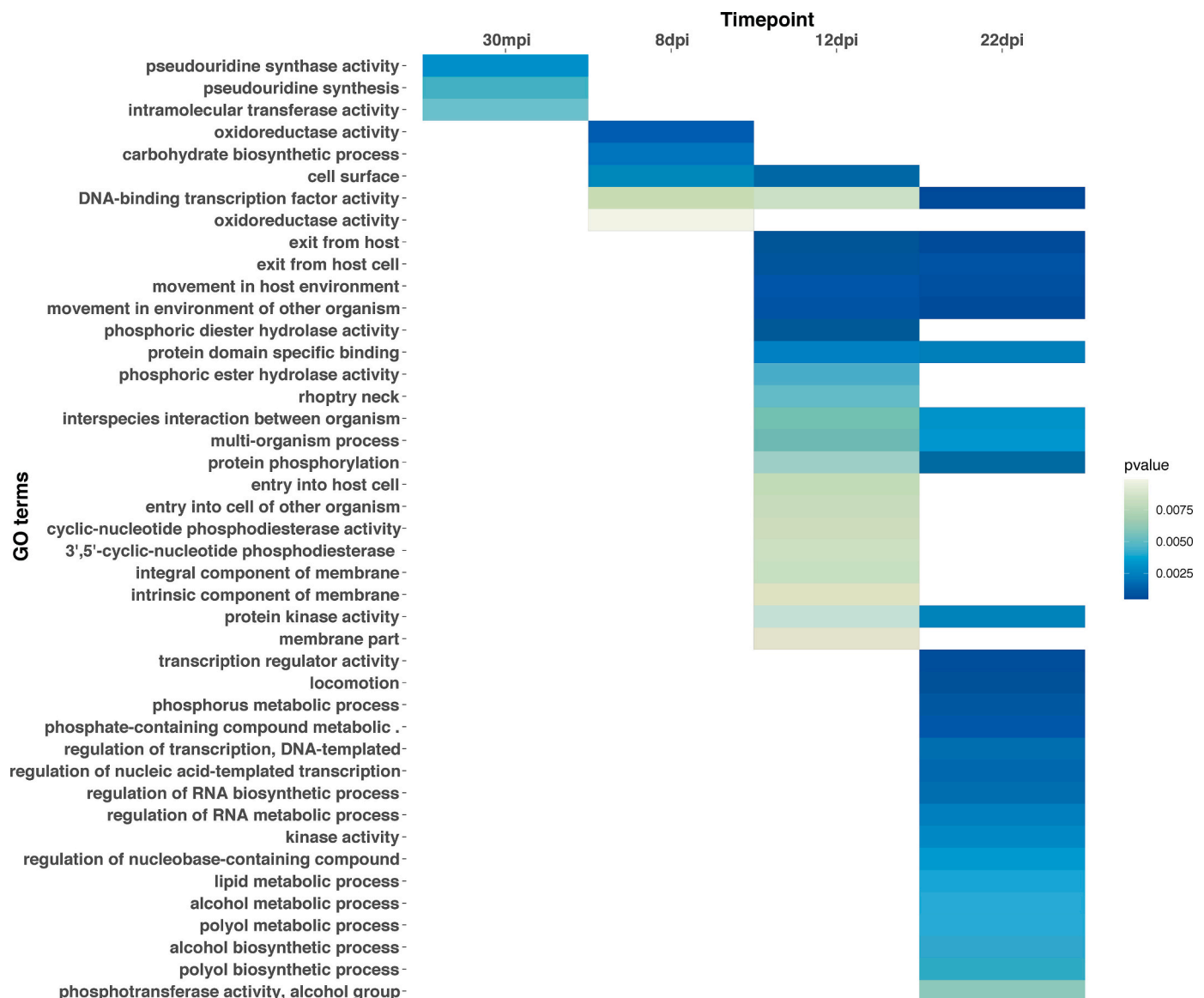
erythrocyte membrane-associated antigen and reticulocyte binding proteins were also upregulated during the later stages of infection (12dpi to 22dpi). These include sporozoite surface protein 3, Thrombospondin-related anonymous protein (TRAP), CRMP1, CRMP2 and Calcium dependent protein kinase 6 (CDPK6). RBPs are essentially involved in erythrocyte invasion [32] and along with Sporozoite surface protein, CRMP1 and CDPK6, effectively invade the salivary gland during the later stages of the infection [25,81]. Sporozoite micronemal protein essential for cell traversal (SPECT1), a key gene involved in host cell traversal [82], are also upregulated during this transition. TRAP is a key player in the attachment and parasite invasion of the salivary glands [58]. Other interesting genes upregulated during this transition are putative plasmepsin X, which is involved in entry and exit into the host cell, and several copies of serine/threonine protein kinase and eukaryotic translation initiation factor 2-alpha kinase 2, which are involved in ATP binding and protein kinase activity and 6-cysteine protein which affects the sporozoites' ability to invade host cells.

#### 4.3. Role of AP2 transcription factors in the sporogonic cycle

Several copies of putative AP2 domain transcription factors (TFs) are upregulated during the infection period, and our analysis also captures a few specific AP2 TFs associated with the various stages of parasite development. AP2 TF family constitutes a major and important TF class in *Plasmodium* species [83]. Our analysis identifies several interesting AP2 TFs which could be possible targets for malaria transmission control applications. Previous studies have reported the functional role of AP2-SP2 during the sporozoite formation [84]. In our analysis, we identified this transcription factor to be upregulated from 8dpi to 22dpi. Previous finding report that AP2-O3 is associated with ookinete formation, gliding and invasion [85]. However, the direct target of this transcription factor is unknown. We observe AP2-O3 (putative) to be upregulated at 8dpi, which might suggest that AP2-O3 regulates genes during the peak oocyst stage. Putative AP2-EXP which, in *P. falciparum*, AP2-EXP is reported to regulate virulence genes [86], is upregulated exclusively at 12dpi. Our results are consistent with previous studies in *P. berghei* [84] where AP2-EXP has been shown to be expressed specifically during the sporozoite stage. AP2-L, an AP2 transcription factor associated with liver exoerythrocytic stages [87] and sporozoites [84] was also found to be upregulated during the 22dpi stage of infection. We also identified upregulation of AP2-O5 during this stage. Even though the role of AP2-O in activation of genes associated with ookinete mobility and oocyst development has been previously established [88], the exact function of AP2-O5 during the sporozoite stage is unclear. Furthermore, several copies of putative AP2 transcription factors were upregulated throughout the duration of infection suggesting that these transcription factors might be responsible for regulation of specific genes in the respective stages of infection. This could be explained by the fact the experiment is conducted *in vivo* and that the different parasite development stages are not completely separate in time. Many copies of AP2 genes indicate that there are several transcriptional factors that are associated and active at different time points, indicating the strong importance of these transcription factors during the entire infection process. This also suggests that AP2 genes might have very selective roles in the parasite development processes and are tightly regulated. The presence of these essential genes in distantly related mammalian and avian *Plasmodium* also suggest that they share some essentially conserved genes [25].

#### 4.4. GO enrichment revealed highly stages specific pathways and metabolism

The different *Plasmodium* species have diverged over time and several genes might have been gained and lost during the evolution. Many of the genes in the *P. relictum* genome are uncharacterized [89]. This potentially limits our ability to pin down all the essential



**Fig. 5.** Gene ontology enrichment analysis for upregulated DEGs. Simplified figure shows the GO terms significantly enriched ( $p$  value  $< 0.01$ ) for different biological processes, molecular functions and cellular components for each time point as compared to the bird baseline. Dark colour indicates higher  $p$  value.

mechanism during the lifecycle of the parasite. To strengthen our understanding of the biological functions behind some of the transcribed genes, we make use of Gene Ontology (GO) system of classification for the genes. GO enrichment allows to establish if the genes of interest are associated with certain biological based on statistical testing. For this purpose, we used the up-to-date annotation for *P. relictum* from PlasmoDB which includes both experimentally validated and computationally predicted GO terms. We observed several GO terms enriched in a very time specific manner suggesting the coordinated regulation of these processes during the development of the parasite (Table S14-S17).

#### 4.5. 30mpi

Only a few GO terms were significantly enriched for the upregulated genes in 30mpi stage. These included pseudouridine synthesis, pseudouridine synthase activity and intermolecular transferase activity. This is indicative of an active progress during this stage of gamete maturation and zygote formation. The successful production of zygotes and the ookinete formation is supported by the overexpressed pseudouridine synthesis which is directly associated with macromolecule modification.

Surprisingly, we also saw the GO term associated with reproduction (GO:0000003) being reported amongst the top 20 GO terms enriched for biological processes. This term was, however, statistically not significant in our analysis (Fig. 5; Table S14), probably due to the limitation in the available knowledge regarding the genes annotated to this term.

#### 4.6. 8dpi

At 8dpi the repeated nuclear division (endomitosis) in oocysts is almost at its peak. This stage is enriched for GO terms associated with oxidoreductase activity, carbohydrate biosynthesis process, DNA binding transcription factor and cell surface (Fig. 5; Table S15). Consistent with DGE analysis, transcription and metabolic processes are active and essential for energy demand during parasite growth and transition [31]. This analysis also reveals the overexpression of oxidation-reduction process, supported by genes such as coproporphyrinogen-III oxidase, which supports regulation of redox balance in the parasite. During infection, the parasites can fight the host's immune oxidative stress mechanism by altering the redox balance and antioxidant defence system [90,91].



#### 4.7. 12dpi

The 12dpi stage is associated with overexpressed GO terms associated with the exit from the host cell, movement in the host environment, interspecies interactions, multi-organism processes (meaning interaction between organisms), protein phosphorylation, and entry into the host cells. The overexpression of these GO terms along with overexpression of GO terms associated with cell surface, membrane parts and integral component of membrane, reflects the crucial role of these pathways in oocyst maturation, sporozoite release and invasion [92] (Fig. 5; Table S16). Other GO terms relating to the interaction with the host, the formation of the parasite's cellular and its mobility within the host seem to be critical for the parasite entry and transmission [25]. Enrichment of the GO term rhoptry neck indicates a host cell invasion mechanism active during this transition. Several molecular functions related to protein domain specific binding, DNA binding transcription factor activities and a number of protein kinase activities are also enriched during this stage, which are indicative of active transition from oocyst to sporozoite formation.

#### 4.8. 22dpi

A wide range of biological processes are associated to the genes upregulated during the later stage of the infection (22dpi). The overexpressed GO terms are associated with the exit from host cell (vector cells in our study), movement in the host environment (vector cells in our study), protein phosphorylation, regulation of RNA metabolic processes, interaction between organisms, regulation of transcription, interspecies interaction and regulation of RNA metabolic process. The molecular functions that are overexpressed are associated with DNA-binding transcription factors, and transcription regulator, protein kinase and phosphotransferase activities (Fig. 5; Table S17). Along with the roles of protein kinases identified in previous studies [93], our analysis also suggests an essential role of this pathways in gene regulation in sporozoite invasion. It is suggested that these pathways have a key role in the stage specific development of the parasite. However, several of the interesting GO terms were not significant according to our cut-off ( $p < 0.01$ ) (Fig. S8). Our analysis was also heavily influenced by the existing GO annotation of the genes. Hence these observations have to be considered with caution. Since not a lot is known about the metabolism and enzymes in these stages, this work provides first-hand insight into the mechanism of the parasite in the vector system.

### 5. Conclusion

This is the first study to provide a comprehensive insight into the molecular mechanism of one of the most harmful avian malaria parasite *P. relictum* in *Culex pipiens quinquefasciatus*, its vector thereby providing valuable knowledge about the genes involved in critical transitions in the lifecycle of the parasite. We have captured a snapshot of genes associated with host immunity, even the ones that were lowly expressed, during all stages during infection within the vector. We identified several known genes associated with cell invasion along with several gene with unknown functions specific to different infection stages of the parasite life cycle, which can be potential candidates for functional studies. We also identified copies of reticulocyte binding proteins upregulated during different stages of the infection, either suggesting a that these proteins have a wider function than previously thought or that they have evolved a different function in *P. relictum*. Another interesting protein Ookinete protein 25 was identified to be upregulated in 8dpi and 12dpi. Even though it has established role in ookinete and oocyst formation, this protein might have other functions of which we at this stage are unaware of. Like many other uncharacterized proteins identified in this study, it will benefit to explore the roles of Ookinete protein 25 and reticulocyte binding proteins in parasite survival in its vector. We also identified many significant genes specific to different stages of parasite

development in this analysis. Although only upregulated genes were considered for this study; we acknowledge that downregulated genes could have a potential cascading effect on the parasite growth which may have been missed in our analyses. Some of the genes have been studied using genetic knockout mutants and their functions have been validated experimentally. However, a large portion of the parasite genome remains uncharacterized *Plasmodium* genes, which limits our ability to list all active genes during the sporogonic cycle. Future research will help determine the function and features of these genes. This work also contributes to improve the genetic resources for *P. relictum*. The knowledge gathered from this study could contribute to our understanding of the critical stages in *Plasmodium* life cycle and could be used as an active model to further conduct specific studies on targeted genes related to invasion. As a result, this work provides an insight into the functional relatedness and mechanism of the development of different *Plasmodium* species within the vector. This can inform further research aimed at devising broader ecological solutions regarding the disease control and conservation programs.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ygeno.2021.05.021>.

#### Ethics statement

Bird manipulations were carried out in strict accordance with the “National Charter on the Ethics of Animal Experimentation” of the French Government. Experiments were approved by the Ethical Committee for Animal Experimentation established by the authors' institution (CNRS) under the auspices of the French Ministry of Education and Research (permit number CEEA- LR-1051).

#### Funding

This study was funded by the Swedish Research Council (grant 2016-03419) and Nilsson-Ehle foundation to O.H. A.R. was funded through the ANR-16-CE35-0001-01 ('EVODRUG').

#### Data availability

The raw sequence data can be obtained from SRA, BioProject id: PRJNA724741. Gene count table is submitted in the supplementary files.

#### Author contributions

O.H. planned and designed the study with input from A.R and S.G. A. R and S.G. designed and performed the mosquito experiment for the RNA sequencing. V.S. performed the bioinformatic and statistical analyses with help and advice from D.A. and O.H. R.P. performed and analyzed the experiment that determined the timing of the life stages in the mosquito. A.D. performed the molecular lab work. V.S. wrote the article with input from all authors.

#### Author statement

All authors have participated in the revision and approved the final version of the manuscript being submitted and declare no conflicting interest.

#### References

- [1] Ott, K.J., Malaria Parasites and Other Haemosporidia. P. C. C. Garnham. Blackwell: Oxford, England; Davis, Philadelphia, 1966. 1132 pp., illus. \$35. 1967.
- [2] N.D. Levine, *The protozoan phylum Apicomplexa. Volume I vol. II*, CRC Press, Inc., Boca Raton, FL 33431 USA, 1988, p. 154.
- [3] J.J. Schall, Malarial parasites of lizards: diversity and ecology, *Adv. Parasitol.* 37 (1996) 255–333.

- [4] S. Bensch, O. Hellgren, J. Pérez-Tris, MalAvi: a public database of malaria parasites and related haemosporidians in avian hosts based on mitochondrial cytochrome b lineages, *Mol. Ecol. Resour.* 9 (2009).
- [5] R.E. Warner, The role of introduced diseases in the extinction of the endemic Hawaiian avifauna, *Condor* 70 (2) (1968) 101–120.
- [6] C.R.F. Chagas, et al., Diversity and distribution of avian malaria and related haemosporidian parasites in captive birds from a Brazilian megalopolis, *Malar. J.* 16 (1) (2017) 83.
- [7] G. Valkiūnas, *Avian Malaria Parasites and Other Haemosporidia*, CRC, Boca Raton, 2005.
- [8] M.C.I. Medeiros, G.L. Hamer, R.E. Ricklefs, *Host compatibility rather than vector-host-encounter rate determines the host range of avian Plasmodium parasites*. *Proceedings, Biol. Sci.* 280 (1760) (2013) 20122947.
- [9] Sylvia M. Fallon, Eldredge Bermingham, Robert E. Ricklefs, Host specialization and geographic localization of avian malaria parasites: a regional analysis in the lesser antilles, *Am. Nat.* 165 (4) (2005) 466–480.
- [10] Charleston, M. And S. Perkins, *Tangled Trees: Phylogeny, Cospeciation, and Coevolution*. 2003.
- [11] J. Mu, et al., Host switch leads to emergence of *Plasmodium vivax* malaria in humans, *Mol. Biol. Evol.* 22 (8) (2005) 1686–1693.
- [12] A.S. Bell, et al., Within-host competition in genetically diverse malaria infections: parasite virulence and competitive success, *Evolution* 60 (7) (2006) 1358–1371.
- [13] J.J. Schall, Parasite virulence, in: E.E. Lewis, J.F. Cambell, M.V.K. Sukhdeo (Eds.), *The Behavioural Ecology of Parasites*, CABI Publishing, 2002, pp. 283–313.
- [14] K.A. Spencer, et al., *Parasites affect song complexity and neural development in a songbird*. *Proceedings, Biol. Sci.* 272 (1576) (2005) 2037–2043.
- [15] S. Bensch, et al., Linkage between nuclear and mitochondrial dna sequences in avian malaria parasites: multiple cases of cryptic speciation? *Evolution* 58 (7) (2004) 1617–1621.
- [16] O. Hellgren, et al., Global phylogeography of the avian malaria pathogen *Plasmodium relictum* based on MSP1 allelic diversity, *Ecography* 38 (8) (2015) 842–850.
- [17] S. Boudjelas, et al., 100 of the world's worst invasive alien species: a selection from the Global Invasive Species Database, in: *The Invasive Species Specialist Group (ISSG) a specialist group of the Species Survival Commission (SSC) of the World Conservation Union (IUCN)*, 2000.
- [18] C. van Riper, et al., The epizootiology and ecological significance of malaria in Hawaiian land birds, *Ecol. Monogr.* 56 (4) (1986) 327–344.
- [19] C.T. Atkinson, D.A. LaPointe, Introduced avian diseases, climate change, and the future of Hawaiian honeycreepers, *J. Avian Med. Surg.* 23 (1) (2009) 53–63 (11).
- [20] D.A. Lapointe, C.T. Atkinson, M.D. Samuel, Ecology and conservation biology of avian malaria, *Ann. N. Y. Acad. Sci.* 1249 (2012) 211–226.
- [21] R.E.T. Vanstreels, et al., Epidemiology and pathology of avian malaria in penguins undergoing rehabilitation in Brazil, *Vet. Res.* 46 (2015) 30.
- [22] V. Palinauskas, et al., *Plasmodium relictum* (lineage P-SGS1): effects on experimentally infected passerine birds, *Exp. Parasitol.* 120 (2008).
- [23] V. Palinauskas, et al., *Plasmodium relictum* (lineage SGS1) and *Plasmodium ashfordi* (lineage GRW2): the effects of the co-infection on experimentally infected passerine birds, *Exp. Parasitol.* 127 (2) (2011) 527–533.
- [24] N. Gerald, B. Mahajan, S. Kumar, Mitosis in the human malaria parasite *Plasmodium falciparum*, *Eukaryot. Cell* 10 (4) (2011) 474–482.
- [25] A.S.I. Aly, A.M. Vaughan, S.H.I. Kappe, Malaria parasite development in the mosquito and infection of the mammalian host, *Annu. Rev. Microbiol.* 63 (2009) 195–221.
- [26] K.A. Akinosoglou, et al., Characterization of *Plasmodium* developmental transcriptomes in *Anopheles gambiae* midgut reveals novel regulators of malaria transmission, *Cell. Microbiol.* 17 (2) (2015) 254–268.
- [27] J.A. Vaughan, Population dynamics of *Plasmodium* sporogony, *Trends Parasitol.* 23 (2) (2007) 63–70.
- [28] R.E. Sinden, et al., Progression of *Plasmodium berghei* through *Anopheles stephensi* is density-dependent, *PLoS Pathog.* 3 (12) (2007), e195.
- [29] S.E. Lindner, et al., Transcriptomics and proteomics reveal two waves of translational repression during the maturation of malaria parasite sporozoites, *Nat. Commun.* 10 (1) (2019) 4964.
- [30] X. Xu, et al., Transcriptome analysis of *Anopheles stephensi*-*Plasmodium berghei* interactions, *Mol. Biochem. Parasitol.* 142 (1) (2005) 76–87.
- [31] A. Srivastava, et al., Stage-specific changes in *Plasmodium* metabolism required for differentiation and adaptation to different host and vector environments, *PLoS Pathog.* 12 (12) (2016) e1006094.
- [32] E. Videvall, et al., The transcriptome of the avian malaria parasite *Plasmodium ashfordi* displays host-specific gene expression, *Mol. Ecol.* 26 (11) (2017) 2939–2958.
- [33] E. Videvall, et al., The avian transcriptome response to malaria infection, *Mol. Biol. Evol.* 32 (5) (2015) 1255–1267.
- [34] Z. Bozdech, et al., The Transcriptome of the Intraerythrocytic developmental cycle of *Plasmodium falciparum*, *PLoS Biol.* 1 (1) (2003), e5.
- [35] T.D. Otto, et al., New insights into the blood-stage transcriptome of *Plasmodium falciparum* using RNA-Seq, *Mol. Microbiol.* 76 (1) (2010) 12–24.
- [36] T.N. Siegel, et al., Strand-specific RNA-Seq reveals widespread and developmentally regulated transcription of natural antisense transcripts in *Plasmodium falciparum*, *BMC Genom.* 15 (1) (2014) 150.
- [37] R. Pigeault, et al., Avian malaria: a new lease of life for an old experimental model to study the evolutionary ecology of *Plasmodium*, *Phil. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 370 (1675) (2015) 20140300.
- [38] J. Vézilier, et al., Insecticide resistance and malaria transmission: infection rate and oocyst burden in *Culex pipiens* mosquitoes infected with *Plasmodium relictum*, *Malar. J.* 9 (1) (2010) 379.
- [39] F. Zélé, et al., *Wolbachia* increases susceptibility to *Plasmodium* infection in a natural system, *Proc. Biol. Sci.* 281 (1779) (2014) 20132837.
- [40] Andrews, S., et al., **FastQC: a quality control tool for high throughput sequence data**. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- [41] S.W. Wingett, S. Andrews, **FastQ Screen: a tool for multi-genome mapping and quality control**, *F1000Res.* 7 (2018) 1338.
- [42] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2, *Nat. Methods* 9 (4) (2012) 357–359.
- [43] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform, *Bioinformatics* (Oxford, England) 25 (14) (2009) 1754–1760.
- [44] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, *Bioinformatics* (Oxford, England) 30 (15) (2014) 2114–2120.
- [45] P. Ewels, et al., MultiQC: summarize analysis results for multiple tools and samples in a single report, *Bioinformatics* (Oxford, England) 32 (19) (2016) 3047–3048.
- [46] D. Kim, B. Langmead, S.L. Salzberg, HISAT: a fast spliced aligner with low memory requirements, *Nat. Methods* 12 (4) (2015) 357–360.
- [47] S. Anders, P.T. Pyl, W. Huber, HTSeq—a Python framework to work with high-throughput sequencing data, *Bioinformatics* (Oxford, England) 31 (2) (2015) 166–169.
- [48] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, *Genome Biol.* 15 (12) (2014) 550.
- [49] S. Anders, W. Huber, Differential expression analysis for sequence count data, *Genome Biol.* 11 (10) (2010) R106.
- [50] A. Alexa, J. Rahnenfuhrer, topGO: Enrichment Analysis for Gene Ontology. R package version 2.38.1, 2019.
- [51] M.K. Yadav, D. Swati, Comparative genome analysis of six malarial parasites using codon usage bias based tools, *Bioinformatics* 8 (24) (2012) 1230–1239.
- [52] A. Cohuet, et al., *Anopheles* and *Plasmodium*: from laboratory models to natural systems in the field, *EMBO Rep.* 7 (12) (2006) 1285–1289.
- [53] G. Jaramillo-Gutierrez, et al., Mosquito immune responses and compatibility between *Plasmodium* parasites and anopheline mosquitoes, *BMC Microbiol.* 9 (1) (2009) 154.
- [54] R. Aguilar, et al., *Anopheles* infection responses; laboratory models versus field malaria transmission systems, *Acta Trop.* 95 (3) (2005) 285–291.
- [55] G. Arévalo-Pinzón, et al., Synthetic peptides from two *Pf* sporozoite invasion-associated proteins specifically interact with HeLa and HepG2 cells, *Peptides* 32 (9) (2011) 1902–1908.
- [56] K. Wengelnik, et al., The A-domain and the thrombospondin-related motif of *Plasmodium falciparum* TRAP are implicated in the invasion process of mosquito salivary glands, *EMBO J.* 18 (19) (1999) 5195–5204.
- [57] J. Thompson, et al., *Plasmodium* cysteine repeat modular proteins 1–4: complex proteins with roles throughout the malaria parasite life cycle, *Cell. Microbiol.* 9 (6) (2007) 1466–1480.
- [58] A.A. Sultan, et al., TRAP is necessary for gliding motility and infectivity of *Plasmodium* Sporozoites, *Cell* 90 (3) (1997) 511–522.
- [59] B. Douradinha, et al., *Plasmodium* cysteine repeat modular proteins 3 and 4 are essential for malaria parasite transmission from the mosquito to the host, *Malar. J.* 10 (2011) 71.
- [60] R. Ménard, et al., Circumsporozoite protein is required for development of malaria sporozoites in mosquitoes, *Nature* 385 (6614) (1997) 336–340.
- [61] Q. Wang, H. Fujioka, V. Nussenzweig, Mutational analysis of the GPI-anchor addition sequence from the circumsporozoite protein of *Plasmodium*, *Cell. Microbiol.* 7 (11) (2005) 1616–1626.
- [62] A. Warburg, et al., *Plasmodium gallinaceum*: antibodies to circumsporozoite protein prevent sporozoites from invading the salivary glands of *Aedes aegypti*, *Exp. Parasitol.* 75 (3) (1992) 303–307.
- [63] A. Coppi, et al., The malaria circumsporozoite protein has two functional domains, each with distinct roles as sporozoites journey from mosquito to mammalian host, *J. Exp. Med.* 208 (2) (2011) 341–356.
- [64] C. Aldrich, et al., Roles of the amino terminal region and repeat region of the *Plasmodium berghei* Circumsporozoite protein in parasite infectivity, *PLoS One* 7 (2) (2012), e32524.
- [65] R. Ménard, et al., Circumsporozoite protein is required for development of malaria sporozoites in mosquitoes, *Nature* 385 (6614) (1997) 336–340.
- [66] C.K. Moreira, et al., The *Plasmodium* TRAP/MIC2 family member, TRAP-Like Protein (TLP), is involved in tissue traversal by sporozoites, *Cell. Microbiol.* 10 (7) (2008) 1505–1516.
- [67] P. Srinivasan, H. Fujioka, M. Jacobs-Lorena, PbCap380, a novel oocyst capsule protein, is essential for malaria parasite survival in the mosquito, *Cell. Microbiol.* 10 (6) (2008) 1304–1312.
- [68] D.K. Raj, et al., Antibodies to PfSEA-1 block parasite egress from RBCs and protect against malaria infection, *Science* 344 (6186) (2014) 871.
- [69] V.A. Nagaraj, et al., Characterization of coproporphyrinogen III oxidase in *Plasmodium falciparum* cytosol, *Parasitol. Int.* 59 (2) (2010) 121–127.
- [70] H. Ke, et al., The Heme biosynthesis pathway is essential for *Plasmodium falciparum* development in mosquito stage but not in blood stages \*, *J. Biol. Chem.* 289 (50) (2014) 34827–34837.
- [71] I. Siden-Kiamos, et al., Identification of *Plasmodium berghei* Oocyst Rupture Protein 2 (ORP2) domains involved in sporozoite egress from the oocyst, *Int. J. Parasitol.* 48 (14) (2018) 1127–1136.
- [72] I. Siden-Kiamos, et al., *Plasmodium berghei* calcium-dependent protein kinase 3 is required for ookinete gliding motility and mosquito midgut invasion, *Mol. Microbiol.* 60 (6) (2006) 1355–1363.

- [73] O. Silvie, K. Goetz, K. Matuschewski, A sporozoite asparagine-rich protein controls initiation of *Plasmodium* liver stage development, *PLoS Pathog.* 4 (6) (2008), e1000086.
- [74] A.S.I. Aly, et al., SAP1 is a critical post-transcriptional regulator of infectivity in malaria parasite sporozoite stages, *Mol. Microbiol.* 79 (4) (2011) 929–939.
- [75] M. Steinbuechel, K. Matuschewski, Role for the *Plasmodium* sporozoite-specific transmembrane protein S6 in parasite motility and efficient malaria transmission, *Cell. Microbiol.* 11 (2) (2009) 279–288.
- [76] A. Combe, et al., TREP, a novel protein necessary for gliding motility of the malaria sporozoite, *Int. J. Parasitol.* 39 (4) (2009) 489–496.
- [77] S.A. Mikolajczak, et al., Distinct malaria parasite sporozoites reveal transcriptional changes that cause differential tissue infection competence in the mosquito vector and mammalian host, *Mol. Cell. Biol.* 28 (20) (2008) 6196–6207.
- [78] T. Kariu, et al., MAEBL is essential for malarial sporozoite infection of the mosquito salivary gland, *J. Exp. Med.* 195 (10) (2002) 1317–1323.
- [79] T. Kariu, et al., CelTOS, a novel malarial protein that mediates transmission to mosquito and vertebrate hosts, *Mol. Microbiol.* 59 (5) (2006) 1369–1379.
- [80] W. Zheng, et al., Functional characterization of *Plasmodium berghei* PSOP25 during ookinete development and as a malaria transmission-blocking vaccine candidate, *Parasit. Vectors* 10 (1) (2017) 8.
- [81] M.M. Mota, J.C.R. Hafalla, A. Rodriguez, Migration through host cells activates *Plasmodium* sporozoites for infection, *Nat. Med.* 8 (11) (2002) 1318–1322.
- [82] M.M. Mota, et al., Migration of *Plasmodium* Sporozoites through cells before infection, *Science* 291 (5501) (2001) 141.
- [83] H.J. Painter, T.L. Campbell, M. Llinás, The Apicomplexan AP2 family: integral factors regulating *Plasmodium* development, *Mol. Biochem. Parasitol.* 176 (1) (2011) 1–7.
- [84] M. Yuda, et al., Transcription factor AP2-Sp and its target genes in malarial sporozoites, *Mol. Microbiol.* 75 (4) (2010) 854–863.
- [85] K. Modrzynska, et al., A knockout screen of ApiAP2 genes reveals networks of interacting transcriptional regulators controlling the *Plasmodium* life cycle, *Cell Host Microbe* 21 (1) (2017) 11–22.
- [86] R.M. Martins, et al., An ApiAP2 member regulates expression of clonally variant genes of the human malaria parasite *Plasmodium falciparum*, *Sci. Rep.* 7 (1) (2017) 14042.
- [87] S. Iwanaga, et al., Identification of an AP2-family protein that is critical for malaria liver stage development, *PLoS One* 7 (11) (2012), e47557.
- [88] M. Yuda, et al., Identification of a transcription factor in the mosquito-invasive stage of malaria parasites, *Mol. Microbiol.* 71 (6) (2009) 1402–1414.
- [89] U. Böhme, et al., Complete avian malaria parasite genomes reveal features associated with lineage-specific evolution in birds and mammals, *Genome Res.* 28 (4) (2018) 547–560.
- [90] K. Becker, et al., Oxidative stress in malaria parasite-infected erythrocytes: host–parasite interactions, *Int. J. Parasitol.* 34 (2) (2004) 163–189.
- [91] S. Müller, Redox and antioxidant systems of the malaria parasite *Plasmodium falciparum*, *Mol. Microbiol.* 53 (5) (2004) 1291–1305.
- [92] E.-M. Patzewitz, et al., An ancient protein phosphatase, SHLP1, is critical to microneme development in *Plasmodium* ookinetes and parasite transmission, *Cell Rep.* 3 (3) (2013) 622–629.
- [93] C. Doerig, et al., Protein kinases of malaria parasites: an update, *Trends Parasitol.* 24 (12) (2008) 570–577.