The Transcriptome of the Intraerythrocytic Developmental Cycle of *Plasmodium falciparum*

Zbynek Bozdech[1☯], Manuel Llinás[1☯], Brian Lee Pulliam[1], Edith D. Wong[1], Jingchun Zhu[2], Joseph L. DeRisi[1*]

1 Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, California, United States of America, **2** Department of Biological and Medical Informatics, University of California, San Francisco, San Francisco, California, United States of America

***Plasmodium falciparum* is the causative agent of the most burdensome form of human malaria, affecting 200–300 million individuals per year worldwide. The recently sequenced genome of *P. falciparum* revealed over 5,400 genes, of which 60% encode proteins of unknown function. Insights into the biochemical function and regulation of these genes will provide the foundation for future drug and vaccine development efforts toward eradication of this disease. By analyzing the complete asexual intraerythrocytic developmental cycle (IDC) transcriptome of the HB3 strain of *P. falciparum*, we demonstrate that at least 60% of the genome is transcriptionally active during this stage. Our data demonstrate that this parasite has evolved an extremely specialized mode of transcriptional regulation that produces a continuous cascade of gene expression, beginning with genes corresponding to general cellular processes, such as protein synthesis, and ending with *Plasmodium*-specific functionalities, such as genes involved in erythrocyte invasion. The data reveal that genes contiguous along the chromosomes are rarely coregulated, while transcription from the plastid genome is highly coregulated and likely polycistronic. Comparative genomic hybridization between HB3 and the reference genome strain (3D7) was used to distinguish between genes not expressed during the IDC and genes not detected because of possible sequence variations. Genomic differences between these strains were found almost exclusively in the highly antigenic subtelomeric regions of chromosomes. The simple cascade of gene regulation that directs the asexual development of *P. falciparum* is unprecedented in eukaryotic biology. The transcriptome of the IDC resembles a "just-in-time" manufacturing process whereby induction of any given gene occurs once per cycle and only at a time when it is required. These data provide to our knowledge the first comprehensive view of the timing of transcription throughout the intraerythrocytic development of *P. falciparum* and provide a resource for the identification of new chemotherapeutic and vaccine candidates.**

Introduction

Human malaria is caused by four species of the parasitic protozoan genus *Plasmodium*. Of these four species, *Plasmodium falciparum* is responsible for the vast majority of the 300–500 million episodes of malaria worldwide and accounts for 0.7–2.7 million annual deaths. In many endemic countries, malaria is responsible for economic stagnation, lowering the annual economic growth in some regions by up to 1.5% (Sachs and Malaney 2002). While isolated efforts to curb malaria with combinations of vector control, education, and drugs have proven successful, a global solution has not been reached. Currently, there are few antimalarial chemotherapeutics available that serve as both prophylaxis and treatment. Compounding this paucity of drugs is a worldwide increase in *P. falciparum* strains resistant to the mainstays of antimalarial treatment (Ridley 2002). In addition, the search for a malaria vaccine has thus far been unsuccessful. Given the genetic flexibility and the immunogenic complexity of *P. falciparum*, a comprehensive understanding of *Plasmodium* molecular biology will be essential for the development of new chemotherapeutic and vaccine strategies.

The 22.8 Mb genome of *P. falciparum* is comprised of 14 linear chromosomes, a circular plastid-like genome, and a linear mitochondrial genome. The malaria genome sequencing consortium estimates that more than 60% of the 5,409 predicted open reading frames (ORFs) lack sequence similarity to genes from any other known organism (Gardner et al. 2002). Although ascribing putative roles for these ORFs in the absence of sequence similarity remains challenging, their unique nature may be key to identifying *Plasmodium*-specific pathways as candidates for antimalarial strategies.

The complete *P. falciparum* lifecycle encompasses three major developmental stages: the mosquito, liver, and blood stages. It has long been a goal to understand the regulation of gene expression throughout each developmental stage. Previous attempts to apply functional genomics methods to address these questions used various approaches, including

Received June 12, 2003; Accepted July 25, 2003; Published August 18, 2003
DOI: 10.1371/journal.pbio.0000005

Abbreviations: ASL, adenylosuccinate lyase; CGH, comparative genomic hybridization; CLAG, cytoadherence-linked asexual gene; Clp, caseineolytic protease; DHFR-TS, dihydrofolate reductase–thymidylate synthetase; EBA, erythrocyte-binding antigen; EBL, erythrocyte-binding-like protein; FFT, fast Fourier transform; FP, falcipain; FV, food vacuole; gDNA, genomic DNA; HAP, histo-aspartyl protease; hpi, hours postinvasion; IDC, intraerythrocytic developmental cycle; MSP, merozoite surface protein; ORF, open reading frame; PM, plasmepsin; PV, parasitophorous vacuole; RBC, red blood cell; RBP, reticulocyte-binding protein; RESA, ring-infected surface antigen; SERA, serine repeat antigen; TCA, tricarboxylic acid

Academic Editor: Gary Ward, University of Vermont

* To whom correspondence should be addressed. E-mail: joe@derisilab.ucsf.edu

☯ These authors contributed equally to this work.