Am. J. Trop. Med. Hyg., 77(2), 2007, pp. 201–208

Mechanisms of Gene Regulation in *Plasmodium*

Functional Genomics Workshop Group*
Summary of a Meeting Held September 15, 2006
The Board Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts†

EXECUTIVE SUMMARY

Summary. Recent investigations into gene expression in malarial systems suggest that locus-specific, promoter-based transcriptional control is not the dominant mode of regulation in *Plasmodium*. Although global transcript and protein profiles across the life cycle of *Plasmodium* imply significant control of developmental progression,[1,2] the exact mechanisms underlying this pattern are largely unknown, and appear to be unconventional compared with those of most model organisms. As such, our understanding of transcriptional, post-transcriptional, and epigenetic mechanisms in *Plasmodium*, as well as their relative contributions to gene regulation, lags far behind that of other eukaryotic systems. We seek a comprehensive characterization of the mechanisms that *Plasmodium* species have evolved to control gene expression during their complex life cycles.

To address these concerns, a panel of experts recently convened for an initial consultation at the Broad Institute (Cambridge MA) in September 2006. Both the integration of existing data in a manner that will be most useful to the malarial community and the types of information currently lacking in the field were discussed. Preliminary recommendations include further delineation of the parasite's genomic structure, as well as elucidation of chromatin structure and epigenetic features; a systematic comparison of global expression data from nuclear run-on, RNA expression microarrays, and proteomic based assays; and the re-evaluation and functional testing of putative regulatory nucleic acid motifs, including characterization of their binding proteins. Detailed examination of regulation on a whole-genome scale, as well as among specific loci, will ultimately make way for control measures aimed at interfering with processes crucial for parasite survival, such as differentiation, antigenic variation, and development of drug resistance.

Overview. *Plasmodium*, the causal agent of malaria, afflicts more than half a billion people worldwide. Despite global efforts to curb malaria, the spread of drug and insecticide resistance and the continuing lack of an effective vaccine contribute to its persistence as a major health burden. A fundamental understanding of how parasite genes governing transmission success, immune evasion, and drug resistance are regulated is critical to developing novel therapeutic strategies against these processes.

Although focused studies on candidate genes have provided some insights, we remain largely ignorant of the mechanisms underlying gene control, and the relative role of transcriptional and post-transcriptional/translational regulation in the parasite. For example, although genes appear to be monocistronically transcribed, no clear canonical promoter has been defined to date and the handful of functional *cis*-acting elements uncovered in *Plasmodium*[3–9] is unique to this system. The dearth of annotated transcription factors in the genome of *Plasmodium falciparum*[10,11] in conjunction with phased expression of stage-specific transcripts[1] suggests that post-transcriptional control may be a major means of regulating gene expression, as supported by a recent study documenting the significant role of translation repression in sexual differentiation.[12]

The successful application of high-throughput approaches in malarial systems now provides the opportunity to fundamentally transform the pace of post-genomic research on gene regulation in *Plasmodium*. For example, both microarray and proteomics-based assays document significant regulation of transcript and protein expression profiles across asexual and sexual stages of development.[1,2,13,14] Such technologies have opened the door to bioinformatically investigating the co-regulation of genes, with the potential to define sequence elements that may be behind this control. Variation in global rates of RNA decay during the intraerythrocytic developmental cycle contributes to mounting evidence supporting a more prominent role for post-transcriptional control in the specific regulation of transcript subsets, as do more recent experiments uncovering global control of transcriptional activity. This provocative combination follows a growing emphasis on chromatin remodeling[15,16] and gene-silencing,[4,15,17] and suggests a complex, multi-layered regulatory network in the parasite. Collectively, these studies underscore the need for a comprehensive, concerted effort to gain the same level of understanding in *Plasmodium* as has been achieved in the model eukaryotic systems for which these gene regulatory mechanisms were first elucidated.

Objectives/Conclusions. We first describe the state of the field in *Plasmodium,* with specific emphasis on what is known about the regulation of steady-state RNA levels during the life cycle and exposure to stress. Although microarray studies clearly demonstrate a "hard-wired" pattern of expression tied to developmental staging, perturbation with small molecules has rendered conflicting results, leading to difficulties in interpretation of this data. We next describe efforts to bioinformatically mine these steady-state RNA data sets to uncover *cis*-acting sequence motifs that may control expression of gene clusters. Preliminary evidence that more global mechanisms of regulating transcriptional activity may dominate this paradigm is also presented. The single exception to this model characterized to date, the antigenic *var* gene family, and its regulatory features are subsequently detailed. The growing body of evidence for post-transcriptional regulation in *Plasmodium* species is then discussed, including function-dependent changes in mRNA decay rate across the life cycle, and translational repression and proteomic regulation in *P.*

* Authors include: Kirk Deitsch, Manoj Duraisingh, Ron Dzikowski, Anusha Gunasekera, Shahid Khan, Karine Le Roch, Manuel Llinás, Gunnar Mair, Victoria McGovern, David Roos, Jennifer Shock, Jennifer Sims, Roger Wiegand, and Elizabeth Winzeler
† Consultation sponsored by the Burroughs Wellcome Fund.
Address correspondence to Dyann F. Wirth, Harvard School of Public Health, Boston MA 02115, E-mail: dfwirth@hsph.harvard.edu or Victoria McGovern, Burroughs Wellcome Fund, Research Triangle Park, NC 27709, E-mail: vmcgovern@bwfund.org

berghei gametocytes. Finally, we propose the following recommendations for the next steps to be taken in our multilateral investigation: 1) determination of chromatin organization and structure; 2) comparison and further characterization of global expression profiles from nuclear run-on, microarray, and proteomic-based assays; and 3) identification, functional testing, and re-evaluation of potentially regulatory nucleic acid binding proteins and their binding sites.

In conclusion, we articulate a way forward in the field of malarial gene expression, and propose experiments that are needed to pursue these goals systematically. This complex topic demands that experts from the malaria field as well as authorities on transcriptional, post-transcriptional, and epigenetic gene regulation in other systems be brought together. With this initial consultation, we have taken the first step for forming and funding such a consortium of scientists.

STATE OF THE FIELD

Regulation of transcription and steady-state RNA levels in *Plasmodium.* *Global regulation of steady-state RNA levels across the life cycle.* In recent years, RNA expression profiles of thousand of genes throughout the malarial life cycle have been analyzed; three main transcriptional analyses covering the entire genome of two *Plasmodium* species, *P. falciparum*[2,14] and *P. berghe*i,[18] have been published. Despite differences in technology, these approaches document similar expression profiles for all forms analyzed. A total of 88% of the parasite genome was expressed in at least one stage of the life cycle, with expression levels varying by five orders of magnitude throughout development.[2,14] A programmed cascade of cellular processes was observed across the erythrocytic cell cycle. Moreover, functionally related genes shared similar expression profiles.

The mechanisms underlying this regulation of mRNA levels are, however, still unclear. Only a few transcription factors have been identified so far in *Plasmodium*.[10] Comparison of mRNA and protein profiles throughout the parasite life cycle indicates a significant delay between the maximum detection of transcript and protein abundance for 45% of the genes examined.[1,18] This time shift effect appears to be marked in gene families. In summary, these analyses suggest a significant role for post-transcriptional events in the malaria parasite.

Global regulation of steady-state RNA levels in response to perturbagens. Although the approaches described above have consistently demonstrated pronounced variation in steady-state mRNA levels with parasite staging, results from perturbation studies are less clear. In contrast to other eukaryotic systems, many groups find that exposure to small molecules do not cause either profound fold-changes in *P. falciparum* gene expression or alterations in specific, functionally related cascades. One motivation behind this initial consultation was to compare the seemingly conflicting pieces of data obtained in similar assays by different investigators, and address the potential origins of such discrepancies.

In recent years, several studies have been undertaken to investigate environmental perturbations and their effect on transcript profiles in *P. falciparum*. However, despite assaying a large collection of treatments, few changes were successfully identified in the parasite. For example, only subtle alterations in parasite RNA levels were detected for genes involved in the folate biosynthesis pathway upon treatment with the antifolate WR99210.[19] Similarly, exposure to 10 different antimalarials suggested that the transcriptional program of *Plasmodium* does not have the capacity to respond to these compounds during either early or late phases of treatment.[20,21] Two different platforms, microarray and serial analysis of gene expression (SAGE), identified only low amplitude transcriptional responses to chloroquine in mixed stage cultures.[22] However, one drawback of the interpretation of these two studies is the inherent complexity of including mixed developmental stages (which might obscure subtle drug-induced changes) and, potentially, the poorly characterized states of arrest.

Studies using synchronized parasites have yielded mixed results. In one case, when synchronized cultures of *P. falciparum* were exposed to the toxic choline analog, T4 (Le Roch K, unpublished data), significant inductions (2–69-fold changes) of genes involved in sexual differentiation were observed after 24 to 36 hours of exposure. No specific changes were documented for enzymes involved in the targeted lipid biosynthesis pathway. Instead, a strong general arrest of genes involved in cell cycle progression was detected. These results highlight the fact that late responses to drug or stress may be confounded by effects of the cell cycle. Given these findings, it is tempting to speculate that, upon detection of a chemical stress, parasites arrest cell cycle progression to stimulate sexual development and escape the impending demise of its human host. Only one study published to date has shown a significant impact of drug on a metabolic pathway.[23] Transcriptional changes were documented after more than two 48-hour cycles of incubation with doxycycline; all notable alterations occurred in the expression profile of apicomplast genes, undoubtedly due to the loss of the apicomplast genome during drug treatment.

Conversely, other analyses have found evidence for regulation of RNA levels in response to environmental stress in the malaria system. Extensive microarray experiments have been carried out in which synchronized cultures were exposed to a variety of perturbed conditions including glucose starvation, serum depletion, hypoxanthine depletion, heat shock, chloroquine exposure, and treatment with methyl methane sulfonate (Llinás M and others, unpublished data). In all of the growth conditions tested, dramatic changes in transcript abundance were observed. As stated above, this may in some cases be due to altered growth rates confounding the measure of transcriptional changes in the parasite. In other cases, the cells may simply arrest. Nonetheless, the perturbations specifically assayed in these experiments do seem to affect the steady-state levels of transcripts. These changes appear to be non-random, which suggests that the parasite is able respond at the transcript level; whether this occurs through modulation of RNA stability or through transcriptional regulation remains to be seen. Finally, certain gene functions were enriched in specific gene clusters, whose transcript levels changed in concert across all the perturbations tested. For example, *P. falciparum* chloroquine resistance transporter and the V-type (H+) pyrophosphatase are reproducibly induced in chloroquine-treated cultures. In summary, these results have led Llinás and others to conclude that, regardless of mechanism, the parasite has some capacity to react at the transcript level to environmental perturbations.

Identification of putative regulatory motifs. Sequence motifs important for controlling transcription of *P. falciparum* have

been identified using promoter mutagenesis approaches as described below. The identities of some genes with life cycle-dependent expression patterns have been known for many years. A number of groups have used the flanking sequences/promoter regions of such genes to drive expression of reporter constructs in transient transfection assays, and thus probe for motifs associated with gene expression.[5,24] For example, the sequence AAGGAATA from the *pfs16* locus was required for reporter protein expression in gametes. This sequence is bound by a protein (PAF-1) in mosquito but not asexual parasite extracts;[5] however, the sequence of PAF-1 protein is not yet known. Linker-scanning mutagenesis has suggested that the sexual-stage promoter activity of *P. gallinaceum* Pgs28 is regulated by the sequence CAGACAGC.[3] None of the proteins binding to these experimentally validated motifs has yet been identified, and so far it has been difficult to identify protein binding activity using standard techniques such as electrophoretic mobility shift assays.

Bioinformatic approaches to mining sequence motifs from the recent wealth of genomic and transcriptomic data have used gene families, functional relationships, and evolutionary conservation. Recently Militello and others identified a statistically significant putative regulatory sequence called the G box in *P. falciparum* by aligning the promoters of heat shock genes and then running the pattern finding algorithm AlignAce.[25] Important motifs tend to be conserved in evolution,[26,27] and thus the significance of a motif in *P. falciparum* may be evaluated by its conservation in the related parasites *P. yoelii*, *P. berghei*, *P. chabaudii*, and *P. vivax* for which shotgun genomic sequences are now available. Young and others used expression data, phylogenetic footprinting, and a method called GBSSR[28] to identify a motif upstream of genes associated with sexual stage development.[29] More recently, van Noort and Huynen combined genome-wide expression data and phylogenetic footprinting to find several candidate motifs, including AGAACAATC, which is associated with genes expressed during invasion.[30] These studies were limited in terms of the amount of expression data examined and the methods used, providing experimental support for motifs identified computationally only in some cases. Moreover, many of these functional assays measure a shift in reporter protein activity, leaving unresolved whether the observed effect is modulated at the transcriptional level, post-transcriptional level, or both.

Global regulation of transcriptional activity in P. falciparum. Data on the expression levels of specific genes obtained by microarray analysis reflect steady-state quantities of intracellular RNA, which are a product of the rates at which new transcripts are generated and the rate at which they are destroyed. In light of both direct and circumstantial evidence that post-transcriptional regulation plays a relatively dominant role in *Plasmodium*, nuclear run-on experiments have been conducted to distinguish regulatory patterns of transcriptional activity from downstream effects. A major finding discussed at the initial consultation was that transcriptional activity is not constant throughout the intraerythrocytic life cycle (Wirth D and others, unpublished data). In contrast, a sharp increase in global transcriptional activity during the late trophozoite and early schizont stages is observed by a nuclear run-on assay. Nuclear run-on experiments measuring the activity of the global transcriptome described a window of less than one-third of the life cycle during which most of the transcripts produced over the course of the life cycle were made. This bulk effect was corroborated on the level of individual genes. Gene-specific hybridization of the labeled transcripts showed that both the sense and antisense strands of several genes were transcribed, all peaking at a similar point in the life cycle as that observed in the global assay, which indicated that the event is not driven by classical promoters. The peak in transcriptional activity was specific to mRNA synthesized by RNA polymerase II. This pattern of life cycle-dependent global regulation of transcription sets *P. falciparum* apart from the model systems in which transcription has been studied directly and extensively; yeast and higher eukaryotes tend to turn on transcription in certain subsets of genes as their expression is needed by tightly controlled promoters. This recent observation suggests instead that most of the *Plasmodium* transcriptome is synthesized in a single, global activation event during the developmental cycle, which implicates post-transcriptional regulation as the dominant force in creating the differential transcript profiles observed by microarray. The single well-characterized exception to this model is the *var* gene family promoter that generates mutually exclusive expression at the level of transcriptional initiation (see below).

Transcriptional regulation of virulence gene expression in P. falciparum. Plasmodium falciparum virulence is in part attributed to the ability of the parasites to modify the erythrocyte surface, and undergo antigenic variation to avoid antibody recognition.[31] Central to both mechanisms is a protein known as *P. falciparum* erythrocyte membrane protein 1(PfEMP1), a highly variable molecule that is exported to the surface of infected erythrocytes and mediates adhesion. The parasite genome contains approximately 60 polymorphic *var* genes, each encoding a different form of PfEMP1. Immune evasion through antigenic variation depends on the ability of the parasite to exclusively express only a single *var* gene at a time, and to periodically switch expression to alternative *var* gene variants, thus altering the antigenic properties of the infected cell and avoiding recognition by antibodies directed against previously expressed forms of PfEMP1. Both *var* gene silencing and mutually exclusive expression are regulated at the level of transcription by mechanisms that seem to be linked.[32]

Mutually exclusive gene expression is a fundamental yet poorly understood aspect of gene regulation in many eukaryotic organisms. In recent years, several advances have been made in elucidating different aspects that contribute to the control of *var* gene expression, including chromatin modifications[15,16], promoter/promoter interactions,[4,33] the presence of sterile RNAs,[34,35] and changes in subnuclear localization.[15,36,37] In addition, recent work has identified the histone modifications that mark *var* genes for activation or silencing.[38] Current work is aimed toward understanding the role of homologs of the histone deacetylase inhibitor SIR2 on the chromatin structure surrounding *var* genes, as well as the function of non-coding RNAs that have been localized to the nucleus and appear to associate with silent *var* genes. An understanding of the molecular mechanisms that control antigenic gene transcription might support the development of drugs that interrupt expression of *var* genes, thus giving the human immune system an opportunity to clear the parasite infection and overcome the disease. Similar mechanisms may be responsible for the regulation of members of many other multigene families in *P. falciparum*.

Post-transcriptional control in *Plasmodium.* *Global regulation of mRNA stability in P. falciparum.* Variability in the rate of mRNA decay is an essential aspect of post-transcriptional regulation in all organisms. Globally, rates of mRNA decay increase dramatically during the asexual intraerythrocytic developmental cycle (Shock J, unpublished data). During the ring stage of the cycle, the average mRNA half-life is approximately 9.5 minutes, yet this is extended to an average of 65 minutes during the late schizont stage of development. This indicates that a major determinant of mRNA decay rate may be linked to the stage of intraerythrocytic development. Furthermore, specific variations in decay patterns are superimposed upon the dominant trend of progressive half-life lengthening. These variations in decay pattern are enriched for genes with specific cellular functions or processes. By complementing and extending previous mRNA abundance studies, elucidation of *Plasmodium* mRNA decay rates has provided a key element for deciphering mechanisms of genetic control.

Role of translational repression in sexual development of P. berghei. Translational repression (TR) in higher eukaryotes is an essential mechanism that controls protein expression during pre- and early post-fertilization development. Movement of certain mRNAs into cytoplasmic messenger ribonucleoprotein (mRNPs) complexes prevents their translation and generates both temporal and spatial expression control that is independent from transcriptional control and protein targeting signals. The assembly and maintenance of such complexes in *Plasmodium* depends on the evolutionarily highly conserved DEAD-box RNA helicase DOZI.[12]

Mutant *Plasmodium* parasites that lack DOZI fail to store certain transcripts (e.g., p25, p28) in translationally quiescent mRNPs. Instead, these mRNAs are targeted for degradation.[12] The effect on mRNA steady-state levels in blood stage gametocytes is profound and affects some 370 transcripts as shown by microarray analyses. Translationally repressed mRNAs (e.g., p25 and p28) co-purify with C-terminally tagged DOZI::GFP protein, as well as homologs of proteins known to function in TR complexes. Some 50 protein potential interactors were identified in mass spectrometric analysis of DOZI::GFP immunoprecipitates. The large number of proteins identified suggests that DOZI may be part of subpopulations of complexes that either target individual groups of mRNAs and/or that DOZI is a general regulator that is present during the entire life-span of such mRNPs, whereas other proteins take part only at specific times during assembly, transport and generation of final and conditionally stable complexes.

A loosely defined 47 nucleotide motif that resides in the 3′ untranslated regions (UTRs) of several transcriptionally silenced mRNAs in *P. berghei*[18]—among them p28—plays an important role in repression (Braks JA and Waters AP, unpublished data). This motif represses translation of a GFP transgene when placed immediately upstream of the silencing inactive dihydrofolate reductase/ts 3′ UTR and is also active in *P. gallinaeceum* and to a lesser degree in *P. falciparum*. Although the 3′ UTR of p28 (which contains this motif) is sufficient to mediate repression, mechanisms involving both 5′ and 3′ UTRs may be important in p25.

The challenges for the future will be to identify the function of individual proteins identified in immunoprecipitates (especially RNA binding proteins and their interactions with specific mRNAs), untangle individual mRNPs, elucidate mechanisms that allow the ordered de-repression and loading of mRNAs on to active ribosomes in post-fertilization development, and identify the role of TR during other *Plasmodium* life cycle stages.

Proteomic analysis of male and female gametocytes. In *P. berghei*, the young gametocytes of both sexes are morphologically indistinguishable from the trophozoite stage. However, the visible morphologic differentiation into either mature males or females occurs at this branching point but the mechanisms for this transition are poorly understood. It is conceivable that this critical event is controlled both by transcriptional and/or post-transcriptional regulation processes. The proteomic analysis of male and female gametocytes showed, not surprisingly, two distinct proteomes.[39] These proteomes showed that gametocytes store a number of proteins that will be used only upon activation inside the mosquito vector (data available from www.lumc.nl/1040/malariaprotocols/). Furthermore, female gametocytes make a distinction between the storage of gamete molecules as sequestered proteins, un-translated mRNA, or un-transcribed genes. There is an indication of active degradation of protein translation machinery in the male gametocyte, as shown by the absence of ribosomal proteins and a corresponding presence of proteosomal proteins. Conversely, female gametocytes/gametes have all the necessary translational machinery. Upon gamete activation, the male engages in the active process of rapid DNA division and flagella assembly, and becomes motile. The female also actively begins translating its stores of mRNA. The process of gamete activation and fertilization is rapid and requires a coordinated response in both male and female gametes.

The two parasite organelles with additional genomes, the mitochondrion and the apicoplast, are physically associated in the parasite and inherited through the female lineage. Consistent with this fact, gametocyte proteome analysis showed that mitochondrial proteins are clearly abundant in female gametocytes and underrepresented in male gametocytes compared with asexual parasites. However, apicomplast proteins appear to be equally abundant in asexual male and female gametocytes, which indicates a continued role for these proteins and this organelle. The distinction in different modes of protein translation and its significance, both immediately before and after fertilization, is intriguing and demands more attention.

NEXT STEPS

The next logical experimental aims and types of data needed to further our existing knowledge of malarial gene expression were discussed by the group. A major outcome of the consultation was consensus on new types of experiments required to investigate major aspects of gene regulation, which have hitherto received only a fraction of the attention they warrant. These include articulating the processes flanking transcription itself: regulation of the dynamics of chromatin organization and the regulation of RNA decay in the context of globally regulated transcriptional activity. Cutting-edge approaches to the discovery of putative regulatory motifs from microarray data sets, followed by a systematic effort to characterize their mechanisms of action, including

discovery of the proteins that bind them, are also outlined below.

Determining chromatin organization. Detailed studies of chromatin dynamics in recent years have illuminated their profound and intricate relationship with transcriptional regulation in model organisms. Although only a few specific transcription factors have been found in the *Plasmodium* genome, the basic components of eukaryotic chromatin and chromatin modification machinery are highly conserved.[40] *Plasmodium* species encode all of the core histones (H2A, H2B, H3, and H4) and variants thereof, although a histone H1 homolog has not yet been identified.[41–43] In addition, genes for histone modification enzymes including histone deacetylases,[44], a GCN5-like histone-acetyl transferase,[45] and methyltransferases[46] have been characterized. Nucleosome assembly proteins have also been identified and function as histone chaperones both within and outside of the nucleus.[47–49] Although the direct role of these many proteins in controlling the cascade of gene expression during development is unclear, nucleosome assemblies have been demonstrated to be essential for promoter function in *Plasmodium*. In particular, it is well established that control of sub-telomeric expression of the *Plasmodial var* gene family is tightly controlled by chromatin-mediated silencing. Whether and how other nucleosome assembly proteins govern the developmental cascade of gene expression remains unclear.

As a first step, global nucleosome positions can reveal transcriptionally active regions of the genome. This could be accomplished through micrococcal nuclease footprinting assays coupled with high-resolution DNA microarrays to readout nucleosome distribution along the chromosomes. Recent results in other organisms using chromatin immunoprecipitation (ChIP) against specific nucleosome modifications have demonstrated positional patterns that correlate well with active regions of the genome, as well as transcriptional start sites.[50–53] Trimethylated H3K4 in higher eukaryotes for example is enriched at the 5′ end of active genes.[54] Similar experiments in *Plasmodium* would define the genome-wide relevance of methylation and acetylation patterns on histones.

Comparison of the rate of transcription and mRNA decay. Of equal importance as deciphering events upstream of transcriptional regulation is synthesizing the datasets on transcript production and degradation. Steady-state mRNA abundance is a common and useful metric for the expression of specific genes, especially given recent advances in the availability and standardization of high-throughput microarrays for *Plasmodium*. However, the steady-state RNA level of any gene is a function of both the rate of transcription and the rate of RNA degradation, and it is yet unclear whether this pattern of expression of individual genes is dominated by life cycle-dependent changes in the rate of transcription, rate of mRNA decay, or both. The apparent paucity of recognizable promoter elements, along with evidence from nuclear run-on that transcriptional activity is regulated globally (discussed above), suggests that gene-specific, promoter-based transcriptional control is not the dominant mode of regulating expression in *P. falciparum*. Furthermore, the rate of mRNA decay appears to be life cycle-dependent: the average half-life of the transcriptome increases approximately seven-fold during the asexual erythrocyte cycle, with different functional clusters of genes varying in their exact pattern of decay rate.

To discern which mechanisms govern the life cycle-dependent expression of individual genes, it is necessary to synthesize data sets pertaining to both the rate of transcription and the rate of degradation. Models of their relative impacts on each member of the transcriptome can then be tested against previously acquired steady-state data. Current research (Wirth D and DeRisi JL, unpublished data) promises to enable this much-needed analysis of the origins of steady-state mRNA abundance.

Additional transcriptome analyses. There are several computational approaches to find regulatory elements that control transcriptional initiation. These approaches make use of available *Plasmodium* genome sequences and expression profiles. The first is to search for motifs identified in other organisms and listed in TRANSFAC. Various laboratories have looked for known motifs in the *P. falciparum* genome[55] and have been unable to find much of interest, which is probably due to divergence from the other organisms in the database and the unusually high AT content of this genome. AT-rich motifs, a telltale characteristic of regulatory regions in other organisms, occur randomly throughout the *P. falciparum* genome. An alternate approach involves grouping genes based on their common functions or expression patterns and search for overrepresented elements upstream of the ATG start codon.[56] Several different algorithms are available for this purpose and a comprehensive review and assessment of these methods has recently been published.[57]

The use of global expression data in combination with phylogenetic footprinting to understand regulation is promising, but additional data are needed. First, expression data from all life stages would undoubtedly improve the quality of putative regulatory motifs that are identified after cluster analysis. Comprehensive expression data (both protein and RNA) from liver, oocysts, male and female gametocytes, and ookinete stages would complement existing data. Expression data from related species such as *T. gondii*, *P. berghei*, *P. yoelii*, and *P. vivax* would also help to refine predictions and validate the presumption of evolutionarily conserved function. Finally, additional data about transcription start and stop sites in the parasite genome that could be obtained from tiling arrays would increase the quality of motif predictions produced by search algorithms.

Functional testing of putative transcriptional regulatory factors. Despite the wealth of biologic information generated by genome-wide analyses of *Plasmodium* species using transcriptomic and/or proteomic techniques, the key regulatory elements that drive the observed changes in the malaria parasite, are still unknown. A computational study using the profile-hidden Markov model (HMMs) identified 156 putative transcription factors in the *P. falciparum genome.*[10] Although the identification of these 156 proteins represents only one-third of the transcriptional control elements expected in a genome the size of *Plasmodium,* understanding their role in regulation of parasite transcription will be of a considerable interest and value to the malaria community.

In recent years, ChIP arrays have emerged as an elegant method to probe the protein-DNA interface on a genome wide scale and rapidly identify the precise binding sites of specific DNA-binding proteins, such as transcription factors and polymerases. By combining chromatin immunoprecipitation with high-density tiling or promoter region arrays of a set of genes, we can screen possible targets of critical transcription factors.

To overcome the difficulties of obtaining a specific antibody, a modified version of the ChIP protocol, the tandem affinity purification (Tap-tag) strategy, can be systematically adopted. The Tap-tag method enables efficient purification of protein and protein complexes present at low levels under native condition. The development of such protocols will enable purification of potential transcription factor or DNA binding protein using commercially available products. A ChIP experiment combined with the Tap-tag strategy should provide interesting insights regarding transcriptional regulation in the malaria parasite. To complete the functional analysis of any potential transcription factors or DNA binding proteins, a knockout of the targeted gene should also be engineered.

Identification of proteins involved in gene regulatory mechanisms. Given the small number of transcriptional regulators identified by homology to well-characterized regulators in other genomes, we are faced with the challenge of identifying putative regulators in *P. falciparum*. Regulators of interest may be involved in transcriptional regulation by enhancing or repressing transcription directly or may have downstream roles through interactions with transcribed RNAs. Post-transcriptional RNA processing may contain several levels of regulation including splicing (potentially alternative splicing), transcript stability, or transcript sequestration.

Although candidate approaches may shed some insight into putative regulators (e.g., characterization of the Myb1 homolog), real strides could be made using a screen for genes necessary for specific transcription. For a screen it would be ideal to generate a knockout library, which could potentially be generated using the piggy-BAC system.[58] The library could then be screened for the transcription of specific reporter constructs pertaining to different stages of *Plasmodium* development in erythrocytes. Conversely, a more biochemical approach would be to use large-batch culturing to immunoprecipitate proteins associated with transcription to find associated factors. This could be accomplished with antibodies to RNA polymerase II or to other key components of the basal transcription machinery, such as TATA binding protein and the other core transcription-associated factors.

A POSSIBLE WAY FORWARD

Two primary functions of this initial consultation were to establish experimental aims critical for advancing the field of gene expression in *Plasmodium* and to discuss the possibility of building consortia dedicated to achieving these aims. Although there is a clear abundance of data on certain aspects of malarial gene regulation, such as developmental control of steady-state transcript levels, and post-transcriptional regulation of sexual differentiation, other areas such as chromatin modification in *Plasmodium*, the balance between transcriptional activity and RNA degradation, and identification and functional evaluation of regulatory proteins and their binding sites, remain unresolved.

The group reached consensus regarding short-term experimental goals. This consensus involves forming consortia to undertake an in-depth investigation of chromatin-dependent gene regulation; coordinated analysis of data on transcriptional activity and mRNA degradation across the life cycle; and application of innovative approaches to mine relevant data from the extant microarray technology. We find it critical in the long-term to initiate multiple facets of discussion both within the broader malaria community and with experts on gene regulation in other eukaryotic fields, especially those in which both technology and dogma are already well-established. Also essential to the future progress of this field is the engagement of funders concerned with the ongoing burden of malaria to determine the level of interest in and feasibility of supporting these goals.

Received February 27, 2007. Accepted for publication February 27, 2007.

LIST OF CONTRIBUTORS

Kirk Deitsch, Cornell University, Ithaca, New York; Manoj Duraisingh, Harvard School of Public Health, Boston, Massachusetts; Ron Dzikowski, Cornell University, Ithaca, New York; Anusha Gunasekera, Harvard School of Public Health, Boston, Massachusetts; Shahid Khan, Leiden University Medical Center, Leiden, The Netherlands; Karine Le Roch, University of California, Riverside, California; Manuel Llinás, Princeton University, Princeton, New Jersey; Gunnar Mair, Leiden University Medical Center, Leiden, The Netherlands; Victoria McGovern, Burroughs Wellcome Fund, Research Triangle Park, North Carolina; David Roos, University of Pennsylvania; Philadelphia, Pennsylvania; Jennifer Shock, University of California, San Francisco, California; Jennifer Sims, Harvard School of Public Health, Boston, Massachusetts; Roger Wiegand, Broad Institute of Harvard, Cambridge, Massachusetts and Massachusetts Institute of Technology, Cambridge, Massachusetts; Elizabeth Winzeler, Scripps Research Institute, La Jolla, California

REFERENCES

1. Le Roch KG, Johnson JR, Florens L, Zhou Y, Santrosyan A, Grainger M, Yan SF, Williamson KC, Holder AA, Carucci DJ, Yates JR III, Winzeler EA, 2004. Global analysis of transcript and protein levels across the *Plasmodium falciparum* life cycle. *Genome Res 14:* 2308–2318.
2. Bozdech Z, Llinas M, Pulliam BL, Wong ED, Zhu J, DeRisi JL, 2003. The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol 1:* E5.
3. Chow CS, Wirth DF, 2003. Linker scanning mutagenesis of the *Plasmodium gallinaceum* sexual stage specific gene pgs28 reveals a novel downstream cis-control element. *Mol Biochem Parasitol 129:* 199–208.
4. Deitsch KW, Calderwood MS, Wellems TE, 2001. Malaria. Cooperative silencing elements in var genes. *Nature 412:* 875–876.
5. Dechering KJ, Kaan AM, Mbacham W, Wirth DF, Eling W, Konings RN, Stunnenberg HG, 1999. Isolation and functional characterization of two distinct sexual-stage-specific promoters of the human malaria parasite *Plasmodium falciparum*. *Mol Cell Biol 19:* 967–978.
6. Horrocks P, Lanzer M, 1999. Mutational analysis identifies a five base pair cis-acting sequence essential for GBP130 promoter activity in *Plasmodium falciparum*. *Mol Biochem Parasitol 99:* 77–87.
7. Voss TS, Kaestli M, Vogel D, Bopp S, Beck HP, 2003. Identification of nuclear proteins that interact differentially with *Plasmodium falciparum* var gene promoters. *Mol Microbiol 48:* 1593–1607.
8. Osta M, Gannoun-Zaki L, Bonnefoy S, Roy C, Vial HJ, 2002. A 24 bp cis-acting element essential for the transcriptional activ-

ity of *Plasmodium falciparum* CDP-diacylglycerol synthase gene promoter. *Mol Biochem Parasitol 121:* 87–98.
9. Porter ME, 2002. Positive and negative effects of deletions and mutations within the 5′ flanking sequences of *Plasmodium falciparum* DNA polymerase delta. *Mol Biochem Parasitol 122:* 9–19.
10. Coulson RM, Hall N, Ouzounis CA, 2004. Comparative genomics of transcriptional control in the human malaria parasite *Plasmodium falciparu*m. *Genome Res 14:* 1548–1554.
11. Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan MS, Nene V, Shallom SJ, Suh B, Peterson J, Angiuoli S, Pertea M, Allen J, Selengut J, Haft D, Mather MW, Vaidya AB, Martin DM, Fairlamb AH, Fraunholz MJ, Roos DS, Ralph SA, McFadden GI, Cummings LM, Subramanian GM, Mungall C, Venter JC, Carucci DL, Hoffman SL, Newbold C, Davis RW, Fraser CM, Barrell B, 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature 419:* 498–511.
12. Mair GR, Braks JA, Garver LS, Wiegant JC, Hall N, Dirks RW, Khan SM, Dimopoulos G, Janse CJ, Waters AP, 2006. Regulation of sexual development of *Plasmodium* by translational repression. *Science 313:* 667–669.
13. Lasonder E, Ishihama Y, Andersen JS, Vermunt AM, Pain A, Sauerwein RW, Eling WM, Hall N, Waters AP, Stunnengerg HG, Mann M, 2002. Analysis of the *Plasmodium falciparum* proteome by high-accuracy mass spectrometry. *Nature 419:* 537–542.
14. Le Roch KG, Zhou Y, Blair PL, Grainger M, Moch JK, Haynes JD, de la Vega P, Holder AA, Batalov S, Carucci DJ, Winzeler EA, 2003. Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science 301:* 1503–1508.
15. Duraisingh MT, Voss TS, Marty AJ, Duffy MF, Good RT, Thompson JK, Freitas-Junior LH, Scherf A, Crabb BS, Cowman AF, 2005. Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in *Plasmodium falciparum*. *Cell 121:* 13–24.
16. Freitas-Junior LH, Hernandez-Rivas R, Ralph SA, Montiel-Condado D, Ruvalcaba-Salazar OK, Rojas-Meza AP, Mancio-Silva L, Leal-Silvestre RJ, Gontijo AM, Shorte S, Scherf A, 2005. Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. *Cell 121:* 25–36.
17. Deitsch KW, 2001. Gene silencing and antigenic variation in malaria parasites. *ScientificWorldJournal 1:* 650–652.
18. Hall N, Karras M, Raine JD, Carlton JM, Kooij TW, Berriman M, Florens L, Janssen CS, Pain A, Christophides GK, James K, Rutherford K, Harris B, Harris D, Churcher C, Quail MA, Ormond D, Doggett J, Trueman HE, Mendoza J, Bidwell SL, Rajandream MA, Carucci DJ, Yates JR III, Kafatos FC, Janse CJ, Barrell B, Turner CM, Waters AP, Sinden RE, 2005. A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic, and proteomic analyses. *Science 307:* 82–86.
19. Ganesan K, Jiang L, White J, Rathod P, 2003. Rigidity of the *Plasmodium* trancriptome revealed by a lethal antifolate. Molecular Parasitology Meeting, Woods Hole, MA.
20. Painter H, Morrisey J, Stumhofer J, Vaidya A, 2004. Early response to antimalarial drug treatment examined by whole genome transcription profiling. Molecular Parasitology Meeting. Woods Hole, MA.
21. Vaidya A, Morrisey J, Painter H, Mather M, 2006. Adaptive changes in transgenic *P. falciparum* expressing type 1A dihydroorotate dehydrogenase continuously exposed to mitochondrial electron transport inhibitors. Molecular Parasitology Meeting. Woods Hole, MA.
22. Gunasekera AM, Patankar S, Schug J, Eisen G, Wirth DF, 2003. Drug-induced alterations in gene expression of the asexual blood forms of *Plasmodium falciparum*. *Mol Microbiol 50:* 1229–1239.
23. Dahl EL, Shock JL, Shenai BR, Gut J, DeRisi JL, Rosenthal PJ, 2006. Tetracyclines specifically target the apicoplast of the malaria parasite *Plasmodium falci*parum. *Antimicrob Agents Chemother 50:* 3124–3131.
24. Mbacham WF, Chow CS, Daily J, Golightly LM, Wirth DF, 2001. Deletion analysis of the 5′ flanking sequence of the *Plasmodium gallinaceum* sexual stage specific gene pgs28 suggests a bipartite arrangement of cis-control elements. *Mol Biochem Parasitol 113:* 183–187.
25. Militello KT, Dodge M, Bethke L, Wirth DF, 2004. Identification of regulatory elements in the *Plasmodium falciparum* genome. *Mol Biochem Parasitol 134:* 75–88.
26. Kellis M, Patterson N, Endrizzi M, Birren B, Lander ES, 2003. Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature 423:* 241–254.
27. Cliften P, Sudarsanam P, Desikan A, Fulton L, Fulton B, Majors J, Waterston R, Cohen BA, Johnston M, 2003. Finding functional features in *Saccharomyces* genomes by phylogenetic footprinting. *Science 301:* 71–76.
28. Kreps J, Budworth P, Goff S, Wang R, 2003. Identification of putative plant cold responsive regulatory elements by gene expression profiling and a pattern enumeration algorithm. *Plant Biotechnol J 1:* 345–352.
29. Young JA, Fivelman QL, Blair PL, de la Vega P, Le Roch KG, Zhou Y, Carucci DJ, Baker DA, Winzeler EA, 2005. The *Plasmodium falciparum* sexual development transcriptome: a microarray analysis using ontology-based pattern identification. *Mol Biochem Parasitol 143:* 67–79.
30. van Noort V, Huynen MA, 2006. Combinatorial gene regulation in *Plasmodium falciparum*. *Trends Genet 22:* 73–78.
31. Dzikowski R, Templeton TJ, Deitsch K, 2006. Variant antigen gene expression in malaria. *Cell Microbiol 8:* 1371–1381.
32. Dzikowski R, Frank M, Deitsch K, 2006. Mutually exclusive expression of virulence genes by malaria parasites is regulated independently of antigen production. *PLoS Pathog 2:* e22.
33. Frank M, Dzikowski R, Costantini D, Amulic B, Berdougo E, Deitsch K, 2006. Strict pairing of var promoters and introns is required for var gene silencing in the malaria parasite *Plasmodium falciparum*. *J Biol Chem 281:* 9942–9952.
34. Calderwood MS, Gannoun-Zaki L, Wellems TE, Deitsch KW, 2003. *Plasmodiu*m *falciparum* var genes are regulated by two regions with separate promoters, one upstream of the coding region and a second within the intron. *J Biol Chem 278:* 34125–34132.
35. Kyes SA, Christodoulou Z, Raza A, Horrocks P, Pinches R, Rowe JA, Newbold CI, 2003. A well-conserved *Plasmodium falciparum* var gene shows an unusual stage-specific transcript pattern. *Mol Microbiol 48:* 1339–1348.
36. Ralph SA, Scheidig-Benatar C, Scherf A, 2005. Antigenic variation in *Plasmodium falciparum* is associated with movement of var loci between subnuclear locations. *Proc Natl Acad Sci USA 102:* 5414–5419.
37. Voss TS, Healer J, Marty AJ, Duffy MF, Thompson JK, Beeson JG, Reeder JC, Crabb BS, Cowman AF, 2006. A var gene promoter controls allelic exclusion of virulence genes in *Plasmodium falciparum* malaria. *Nature 439:* 1004–1008.
38. Chookajorn T, Dzikowski R, Frank M, Li F, Jiwani AZ, Hartl DL, Deitsch KW, 2007. Epigenetic memory at malaria virulence genes. *Proc Natl Acad Sci USA 104:* 899–902.
39. Khan SM, Franke-Fayard B, Mair GR, Lasonder E, Janse CJ, Mann M, Waters AP, 2005. Proteome analysis of separated male and female gametocytes reveals novel sex-specific *Plasmodium* biology. *Cell 121:* 675–687.
40. Templeton TJ, Iyer LM, Anantharaman V, Enomoto S, Abrahante JE, Subramanian GM, Hoffman SL, Abrahamsen MS, Aravind L, 2004. Comparative analysis of apicomplexa and genomic diversity in eukaryotes. *Genome Res 14:* 1686–1695.
41. Bennett BJ, Thompson J, Coppel RL, 1995. Identification of *Plasmodium falciparum* histone 2B and histone 3 genes. *Mol Biochem Parasitol 70:* 231–233.
42. Creedon KA, Kaslow DC, Rathod PK, Wellems TE, 1992. Identification of a *Plasmodium falciparum* histone 2A gene. *Mol Biochem Parasitol 54:* 113–115.
43. Longhurst HJ, Holder AA, 1997. The histones of *Plasmodium falciparum*: identification, purification and a possible role in the pathology of malaria. *Parasitology 114:* 413–419.
44. Joshi MB, Lin DT, Chiang PH, Goldman ND, Fujioka H, Aikawa M, Syin C, 1999. Molecular cloning and nuclear localization of

a histone deacetylase homologue in *Plasmodium falciparum*. *Mol Biochem Parasitol 99:* 11–19.
45. Fan Q, An L, Cui L, 2004. *Plasmodium falciparum* histone acetyltransferase, a yeast GCN5 homologue involved in chromatin remodeling. *Eukaryot Cell 3:* 264–276.
46. Pollack Y, Kogan N, Golenser J, 1991. *Plasmodium falciparum*: evidence for a DNA methylation pattern. *Exp Parasitol 72:* 339–344.
47. Chandra BR, Olivieri A, Silvestrini F, Alano P, Sharma A, 2005. Biochemical characterization of the two nucleosome assembly proteins from *Plasmodium falciparum*. *Mol Biochem Parasitol 142:* 237–247.
48. Navadgi VM, Chandra BR, Mishra PC, Sharma A, 2006. The two *Plasmodium falciparum* nucleosome assembly proteins play distinct roles in histone transport and chromatin assembly. *J Biol Chem 281:* 16978–16984.
49. Pace T, Olivieri A, Sanchez M, Albanesi V, Picci L, Siden Kiamos I, Janse CJ, Waters AP, Pizzi E, Ponzi M, 2006. Set regulation in asexual and sexual Plasmodium parasites reveals a novel mechanism of stage-specific expression. *Mol Microbiol 60:* 870–882.
50. Bernstein BE, Liu CL, Humphrey EL, Perlstein EO, Schreiber SL, 2004. Global nucleosome occupancy in yeast. *Genome Biol 5:* R62.
51. Lee CK, Shibata Y, Rao B, Strahl BD, Lieb JD, 2004. Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nat Genet 36:* 900–905.
52. Liu CL, Kaplan T, Kim M, Buratowski S, Schreiber SL, Friedman N, Rando OJ, 2005. Single-nucleosome mapping of histone modifications in *S. cerevisiae*. *PLoS Biol 3:* e328.
53. Yuan GC, Liu YJ, Dion MF, Slack MD, Wu LF, Altschuler SJ, Rando OJ, 2005. Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science 309:* 626–630.
54. Millar CB, Grunstein M, 2006. Genome-wide patterns of histone modifications in yeast. *Nat Rev Mol Cell Biol 7:* 657–666.
55. Kumar N, Cha G, Pineda F, Maciel J, Haddad D, Bhattacharyya M, Nagayasu E, 2004. Molecular complexity of sexual development and gene regulation in *Plasmodium falciparum*. *Int J Parasitol 34:* 1451–1458.
56. Cho RJ, Campbell MJ, Winzeler EA, Steinmetz L, Conway A, Wodicka L, Wolfsberg TG, Gabrielian AE, Landsman D, Lockhart DJ, Davis RW, 1998. A genome-wide transcriptional analysis of the mitotic cell cycle. *Mol Cell 2:* 65–73.
57. Tompa M, Li N, Bailey TL, Church GM, de Moor B, Eskin E, Favorov AV, Frith MC, Fu Y, Kent WJ, Makeev VJ, Mironov AA, Pavesi G, Pesole G, Regnier M, Simonis N, Sinha S, Thijs G, van Helden J, Vandenbogaert M, Wenz Z, Workman C, Ye C, Zhu Z, 2005. Assessing computational tools for the discovery of transcription factor binding sites. *Nat Biotechnol 23:* 137–144.
58. Balu B, Shoue DA, Fraser MJ Jr, Adams JH, 2005. High-efficiency transformation of *Plasmodium falciparum* by the lepidopteran transposable element piggyBac. *Proc Natl Acad Sci USA 102:* 16391–16396.