Published in final edited form as:

Parasite Immunol. 2012 February ; 34(2-3): 50–60. doi:10.1111/j.1365-3024.2011.01340.x.

Genomics and Integrated Systems Biology in *Plasmodium falciparum*: A Path to Malaria Control and Eradication

Karine G. Le Roch, Duk-Won D. Chung, and Nadia Ponts

University of California Riverside, Department of Cell Biology and Neuroscience, Institute for Integrative Genome Biology, and Center for Disease Vector Research, 900 University Avenue, Riverside, CA 92521, USA.

Abstract

The first draft of the human malaria parasite's genome was released in 2002. Since then, the malaria scientific community has witnessed a steady embrace of new and powerful functional genomic studies. Over the years, these approaches have slowly revolutionized malaria research and enabled the comprehensive, unbiased investigation of various aspects of the parasite's biology. These genome-wide analyses delivered a refined annotation of the parasite's genome, a better knowledge of its RNA, proteins, and metabolite derivatives, and fostered the discovery of new vaccine and drug targets.

Despite the positive impacts of these genomic studies, most research and investment still focus on protein targets, drugs and vaccine candidates that were known before the publication of the parasite genome sequence. However, recent access to next-generation sequencing technologies, along with an increased number of genome-wide applications are expanding the impact of the parasite genome on biomedical research, contributing to a paradigm shift in research activities that may possibly lead to new optimized diagnosis and treatments. This review provides an update of *Plasmodium falciparum* genome sequences and an overview of the rapid development of genomics and system biology applications that have an immense potential of creating powerful tools for a successful malaria eradication campaign.

Keywords

Malaria; Genome; Systems Biology; Transcriptomics; Proteomics; Metabolomics; Epigenetics; Genomics; Interactomics; Immunomics; disease strategy

Introduction

Malaria is a mosquito-borne disease caused by a eukaryotic protozoan parasite of the genus *Plasmodium*. With up to one million deaths per year, malaria remains one of the deadliest infectious diseases in the world and has been recognized as one of the strongest forces driving evolutionary selection in the human genome. There are five different species of *Plasmodium* that can infect humans; *P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale*, and more recently *P. knowlesi*. *P. falciparum* is responsible for the most severe malignant malaria leading to death, especially in children under five years old in sub-Saharan African

Corresponding Author: Dr Le Roch, Karine 900 University Ave Riverside California United States 92521 karine.leroch@ucr.edu.

Disclosures:

The authors declare no potential conflicts or financial interests.

countries. In addition to its deleterious effects on human health, malaria has a significant impact on poverty and is a major impediment to economic development.

Despite the success of an eradication campaign after the Second World War in developed countries (Europe and North America) and a significant reduction of cases in developing parts of the world, malaria is still widespread in all tropical and subtropical areas and can still affect more than 40% of the world population. Recent advances in treatments – these include the development of new combinational therapies, the increased use of bed nets and improved insecticides – have contributed to the reduction of detected infections in select African countries and revived hope that malaria is a disease that can be eradicated.

While there is still no approved vaccine, malaria is a curable disease. Since ancient times, traditional medicinal plants have been used to treat malaria. Currently, quinine and artemisinin, isolated from the bark of the Peruvian *Cinchona succirubra* tree and the Chinese plant *Artemisia annua* respectively, remain the most effective available anti-malarials. Unfortunately, Artemisinin-based Combination Therapies (ACTs), recently adopted as our last resort in combating malaria infection, are already challenged by ACT-resistant strains detected in Southeast Asia. With the spread of parasite resistance to all current anti-malaria drugs, successful control and eradication will only be achieved if new efficient tools and cost-effective anti-malarial strategies are developed.

When the near completed sequence of the genome of the human malaria parasite *P. falciparum* was first published (1), the scientific community predicted that it would accelerate the discovery of new drug targets and vaccine strategies. Almost a decade later, this is still a work-in-progress. The genome sequence of the malaria parasite has nonetheless provided the foundation for modern biomedical research. The goal is now to transform our increasing knowledge of the parasite's biology into actual improvements of human health. Such achievement requires an integrated understanding of both the pathogen and the host's responses to infection. In this review, we present an overview of the *P. falciparum* genome as well as recent advances in genomics and systems biology that have led to major improvements in the understanding of the pathogen. We discuss the impact of these approaches on the development of new therapeutic strategies as well as exploring the long-term goal of global malaria eradication.

The malaria parasite genome

The first draft of the *P. falciparum* genome was published after seven years of international effort. The genome was sequenced using the Sanger method and chromosome shotgun strategy (1). The size of the genome was initially estimated at 22.8 Mb separated into 14 chromosomes and 5300 protein-encoding predicted genes. In addition to its nuclear genome, the parasite contains 6 kb and 35 kb circular DNA found in its mitochondria and plant-related apicoplast respectively. Today, the *P. falciparum* genome remains to be the most AT-rich genome. The overall (A+T)-composition is 80.6%, and can rise to 95% in introns and intergenic regions. After almost 9 years of coordinated genome curation efforts, the complete genome sequence is defined as haploid and 23.26 Mb in size. It contains 6372 genes and 5524 protein-coding genes (genome version: 06-01-2010, <http://plasmodb.org/plasmo/>). Approximately half of these genes have no detected sequence homology with any other model organism. Despite recent access to comparative and functional genomics studies and the completion of genome sequencing of more than eight *Plasmodium* species, the cellular function of most of the parasite genes remain obscure.

Over the past few years, extensive re-sequencing efforts have been successfully undertaken to identify genes and genetic traits associated with parasite's drug resistance and severity of the clinical outcomes. Initial sequencing surveys of genetic variation across the *P.*

falciparum genomes were published in 2007 (2-4). Volkman and colleagues sequenced high-quality draft genomes of three parasite laboratory clones (the reference sequenced as 3D7, HB3 and Dd2) isolated from different parts of the world (2). Their work alone identified 26845 single-nucleotide polymorphisms (SNPs) at a frequency of one SNP every 780 bases between the three clones and an additional 37039 insertion-deletions (indels) between 3D7 and HB3. They further extended their genotyping to 12 *P. falciparum* strains and 20 genomic regions from 54 worldwide *P. falciparum* isolates. Results were consistent with initial genetic diversity studies that were performed using whole-genome microarray analysis (5). All together they identified more than 46937 SNPs (one every 446 bases in average) across the whole genome. High levels of SNPs were detected in genes involved in antigenic variation as well as genes involved in drug-resistance. These data were further confirmed by the survey of ~60% of *P. falciparum* predicted genes (3) and a shotgun sequencing strategy of a Ghanaian clinical isolate (4). Taken together, these reports identified a high number of rare SNP variants and suggested that most SNPs have yet to be discovered. As a whole, these results underscore the importance of creating comprehensive maps of genetic diversity in *P. falciparum* field isolates. These SNPs are strongly suspected to be markers for various phenotypic traits such as virulence or resistance to drugs.

Recent advances in next-generation sequencing (NGS) technologies are enabling fast and affordable production of large amounts of genome sequence information. These technologies are already opening new perspectives of functional genomics in the field of primary, applied, and clinical malaria research.

High-throughput sequencing in malaria research

After 30 years of dominance of first generation “Sanger” dideoxy sequencing, the past five years have seen the explosion of NGS methods. NGS has transformed the field of whole-genome sequencing and analysis. Unlike Sanger sequencing, NGS avoids the need for bacterial cloning and therefore bypasses associated biases. For example, AT- or GC-rich regions are often toxic to bacteria and difficult to reliably read with cloning-based sequencing. This issue is of major importance in the case of the *P. falciparum*'s extremely AT-rich genome. The major leap forward from NGS is the ability to produce an enormous amount of data within small volumes; a tremendous number of DNA fragments, up to 2 billion short reads per instrument run, can be sequenced in parallel.

Three main NGS platforms have been commercialized over the past five years: the Roche 454, the Applied Biosystems SOLiD and finally the Illumina® (formally known as Solexa) Genome Analyzer and Hi-Seq platforms. The Illumina® sequencers are currently capable of producing up to 2 billion reads per run – this value is continually increasing due to constant improvements of reagents and consumables – with a recommended read length of 35-100 bp. This technique is by far the most successful NGS method to sequence the *P. falciparum* genome. Many variations of the technique were developed specifically for the sequencing of the (A+T)-rich genome of the malaria parasite (6-8) (Figure 1). Over the last couple of years only, many studies have used Illumina®'s NGS technology to identify SNPs and other mutations linked to drug resistance in the murine malaria parasite *P. chabaudi* (9,10) and the human malaria parasite *P. vivax* (11). Other analyses have contributed to the characterization of the *P. falciparum* transcriptome with the discovery of new splicing events (12-14) and transcription start sites (TSS) (15). Finally, Illumina®'s NGS technology was used to discover atypical features of *P. falciparum*'s chromatin (6,16) and various epigenetic events (7).

Currently, the future of high-throughput sequencing seems to be leaning towards single-cell sequencing applications. Going further, third-generation sequencing (TGS) technologies propose to use single molecules as direct templates for sequencing (techniques so far under

development at Helicos Biosciences and Pacific Biosciences). These TGS technologies should simplify the sample preparation procedure, avoid the bias introduced by DNA amplification and library preparation and be even more affordable than their predecessors. Nevertheless, the power of high-throughput sequencing also represents one of the major pitfalls for the analysts. The high-throughput and depth of quantitative measurements produced by NGS and TGS technologies come at the cost of producing sophisticated algorithms and software tools capable of accurately examining millions to billions of reads. The data generated by these methods are complex, novel, and abundant. The computational and statistical analysis of raw outputs is the tricky step where incorrect normalization and processing can yield misleading conclusions. Novel methods of quantitative analysis are constantly under development and testing. There is yet no consensus on which analytical approach is the most accurate, particularly for the *Plasmodium* genome.

The avalanche of whole-genome data over the past few years generated an immense source of knowledge that still requires maturing and processing. Nevertheless, in the near future, these powerful genomic approaches will certainly catalyze the transformation of this biological knowledge into viable therapeutic strategies. Single-cell sequencing will accelerate the genotyping of strains in patients' blood sample or other field isolates. Comparative genomics then will be an important source of information regarding the evolution and dynamics of malaria parasites' populations. Ultimately, such knowledge could be used for accurate diagnosis and targeted treatment of patients.

From genome to function

The ultimate goal of functional genomics and systems biology in the malaria field is to discover the properties of the parasite in order to pinpoint its weaknesses and identify long lasting anti-malarial strategies. The amalgamation of large-scale genome-wide analyses (microarrays, deep sequencing, quantitative mass spectrometry, epigenome mapping, computational modeling, *etc.*) has been used to mine *Plasmodium*'s genome in an unbiased manner and identify the genetic elements that may be targeted in the fight against malaria (Figure 2). Here, we present major contributions of the main “omics” to the malaria field.

ii) Transcriptomics—Microarray-based large scale analyses of *P. falciparum*'s transcripts led to the discovery of expressed genes, their functional association to the various stages of the parasite life cycle, and their involvement in particular biological processes with a high degree of accuracy (17-20). More recent sequencing-based studies, such as RNA-seq confirmed these initial microarray experiments and showed promising results on the prediction of new splicing events. These studies also allowed the identification of new open reading frames with their untranslated flanking regions (12-14,21). Moreover, transcriptome analyses in *P. falciparum* field isolates identified previously unknown factors involved in pathogenesis and immune evasion (22-26). Finally, analyses of transcription profiles of variant surface antigens identified patterns that are specific to the parasite's sexual stages and could be relevant for new vaccine interventions (27,28). In addition to mRNA-related transcriptomics, non-coding protein RNA (ncpRNA) transcriptome has been analyzed (29). In eukaryotes, structural ncpRNA are known to participate in the regulation of diverse biochemical pathways, *e.g.*, transcription, translation, epigenetic regulations, cell differentiation and proliferation. In *P. falciparum*, 604 putative ncpRNAs were detected (30-32) and were shown to form a complex regulatory network. All together, these latest analyses suggest that *P. falciparum* ncpRNAs may play a critical role in determining antigenic variation and virulence mechanisms (29).

ii) Proteomics—Previous proteomics (33-35) and interactomics (36) studies have confirmed and complemented the functional annotations proposed based on transcriptome

profiling. Numerous proteomics analyses surveyed stage-specific proteins, investigated as potential drug targets, including sex-specific proteins in male and female gametocytes that could be utilized for transmission blocking strategies (37). Parasite surface proteins (parasite proteins that are exported to the surface of the infected red blood cells) also represent new potential antigens for rational vaccine development (33-35,38,39). Genomics, cell biology, and proteomics studies identified a conserved protein export motif, the PEXEL motif, which has been reported in as many as 400 proteins. Most of these proteins are expressed during the erythrocytic stages. They are thought to be involved in creating knob structures on the surface of red blood cells, and are often associated with cytoadherence and antigenic variation. A complete understanding of their function and regulation will therefore be critical in order to disrupt one of the most pathological effects of *Plasmodium* infections.

iii) Metabolomics—In an effort to improve functional annotation and increase our understanding of the parasite's biology, a number of research groups have been leveraging biochemical metabolic profiling and metabolomics strategies (40). Metabolomics is the study of the entire repertoire of metabolites, *i.e.*, small molecules such as amino acids, sugars and fatty acids that are known to perform critical functions in various biological processes. Correlation analyses of transcriptomics, proteomics, and metabolomics data are a powerful way to identify new metabolic pathways as well as genes that encode for specific enzymatic functions (41,42). While the study of metabolomics in *Plasmodium* is still in its infancy, it has already uncovered important biological insights with possible implications in terms of adaptation, evolution, and host-pathogen interactions (43-45).

iv) Miscellaneous functional genomics—Functional genomics suffers from the lack of tools to analyze the malaria parasite's genome. For example, gene silencing using RNAi cannot be used in *Plasmodium* since the machinery does not exist in the parasite; gene knockout experiments are time consuming processes not compatible with large scale high-throughput analyses. However, in the past few years, a transposon-based mutagenesis approach in *Plasmodium* has been developed (46). A *Plasmodium*-specific selection cassette was added to the lepidopteran transposon *piggyBac* and transfected in parasites together with a transposase-containing helper plasmid (47). Random insertional mutants are obtained by multiple integrations of the transposon at TTAA recognition sites. Recent studies used *piggyBac*-based approaches to validate candidate parasite-specific secreted proteins (48) or identify genes that are essential for the parasite's proliferation (49). Used in combination with other genomics and proteomics analyses, *piggyBac*-based strategies could provide a better understanding of the parasite's biology and its interactions with its hosts.

v) Accessibility and usability of knowledge—The data of large-scale and functional genomic analyses must be accessible and intelligible for practical and efficient usage. The task belongs to the informatics and bioinformatics fields that can provide the necessary tools. Up to now, data depositary banks and web-based databases such as PlasmoDB (<http://plasmodb.org/plasmo/>) have greatly facilitated the access, the comprehensive visualization, and the analysis of large data sets. Gene predictions and annotations, new drug target identifications, and discoveries of vaccine candidates all resulted from various genome-wide analyses. However, it is critical that such resources remain well maintained and free for maximized accessibility. Indeed, a systemic view of the malaria parasite's biology can only be achieved with the successful integration and accessibility of the data from various origins.

Genomics and gene regulation

Genome-wide transcriptomic and proteomic approaches have been extensively utilized to identify regulatory elements that control gene expression in the parasite's genome. For

example, one approach consisted of a DNA motif discovery framework based on the detection of dependencies between microarray-based transcriptomic data and the presence of DNA motifs within the 5' untranslated regions of genes (50). This approach identified *in silico* 21 potential motifs found in ~2700 genes expressed in *P. falciparum*. The method, however, may not perform very well on highly degenerated or atypical motifs. Another approach consists of identifying quantitative trait *loci* that are involved in gene expression variations (eQTLs) in various clones of *P. falciparum* (51). Using tiling arrays, Gonzales and colleagues identified hotspots of sequence polymorphisms spread throughout the entire genome that control the expression of nearly 18% of the genes from a distance. More recently, potential regulatory sequences found at nucleosome-free regions of DNA have been identified using formaldehyde-assisted isolation of regulatory elements (FAIRE) coupled with NGS at high resolution and large-scale (13). In addition, ChIP-on-chip experiments using histone H4-specific antibodies were used to discover nucleosome-bound sequences and also suggest the potential presence of nucleosome-free regulatory elements (52). These kinds of studies have provided a considerable amount of data in just a few years. The mechanisms that *P. falciparum* uses to regulate gene expression remain nonetheless elusive. Indeed, the remarkable changes in steady-state mRNA levels, with a tightly coordinated cascade of transcripts throughout the parasite life cycle, remain challenging to comprehend.

i) Transcriptional vs. post-transcriptional mechanisms of gene regulation—

The core transcriptional machinery that drives RNA polymerase II-dependent transcription (53), and 27 Apicomplexan AP2 (ApiAP2) plant-related transcription factors (54,55) have been identified as major regulators of parasite gene expression. All together, the proteins involved in the transcriptional machinery (including general transcription factors), along with ApiAP2-specific transcription factors, represent less than 2% of the total genome. Considering the *P. falciparum*'s genome size, twice this amount is required for a classical “transcription factor-mediated” model of gene regulation (53,56,57). Thus, either more atypical and elusive regulators remain to be discovered, or gene regulation in *Plasmodium* is not so classically based on the coordinated action of specific positive/negative regulators only.

The initial characterization of the ApiAP2 transcription factor family was a major step forward understanding key regulators in *Plasmodium* (58). However, their exact role in the parasite's biology remains to be determined. Furthermore, recent studies have started to underline that the malaria parasite may have adapted and optimized its mechanisms of transcriptional regulation for its lifestyle. Indeed, in contrast to what is observed in other eukaryotes, environmental perturbations and small molecule inhibitors do not cause specific transcriptional changes in *P. falciparum*, as revealed by genome-wide analyses of parasite expression profiles in response to stress (59-61). The concept of transcriptional rigidity in *Plasmodium* was recently conceived (59). Parasites subjected to chemical or environmental stresses do not specifically compensate for the stress-targeted pathways at the transcriptional level; instead, they exhibit a strong cell cycle arrest and an induction of genes involved in general (non-specific) stress responses and sexual differentiation. Taken together, these studies highlight an unusual method of transcriptional regulation with a limited capacity for positive or negative feedback mechanisms.

Additional analyses of mRNA *versus* protein profiles show significant varying time shifts between transcript and protein levels. These data enforce that extensive post-transcriptional mechanisms of gene regulation may have important roles during parasite development (38,62,63). Following these latest observations, the characterization of protein complexes involved in translational repression (64) and whole-genome analysis of mRNA decay rates

strongly supports the idea that post-transcriptional regulation may be an important mechanism for gene regulation in *P. falciparum* (65).

ii) Genome-wide chromatin structure—Recent studies highlight the importance of key chromatin components that regulate parasite development (66-68). A large number of chromatin modifying complexes have recently been identified (reviewed in (69)) leading to the hypothesis that malaria parasites may, in large part, be subject to epigenetic mechanisms that control gene expression. Epigenetic modifications involve reversible modifications to DNA or proteins that do not affect the genome sequence but are inheritable and modulate gene expression as well as other biological processes (70). In the human malaria parasite, heterochromatic silencing was shown to control mutually exclusive expression of antigenic variation genes in the parasite (67,68,71). More recently, several studies investigated the genome-wide distribution of various euchromatic/heterochromatic histone marks. Lopez-Rubio and colleagues used high-resolution ChIP-on-chip to map the positions of trimethylated lysine 9 histone H3 (H3K9me3), trimethylated lysine 4 histone H3 (H3K4me3), and acetylated lysine 9 histone H3 (H3K9ac) in *P. falciparum* (72). They showed that H3K9me3, a silencing mark, has an atypical distribution in the *P. falciparum* genome; H3K9me3 is indeed confined within the subtelomeric and limited chromosome internal regions that are closely associated with genes involved in antigenic variation. On the contrary, the active marks H3K4me3, and H3K9ac, display a broad distribution across the genome. Additional ChIP-on-chip profiling of histones H3, H3K4me3, H3K9me3, and H3K9ac confirmed that heterochromatin is restricted to the regions of the genome that contain the variant surface antigen families in *P. falciparum* (73). Further analyses also confirmed the co-localization of the heterochromatin protein 1 to H3K9me3, along with their association to regions of the genome that code for *Plasmodium* virulence factors (74,75). Global histone mass spectrometry analysis also confirmed the prevalence of active acetylated histone marks compared to inhibitory methylated ones (76). All together these results suggest an atypical euchromatin/heterochromatin structure in the malaria parasite; active chromatin is prevalent genome-wide whereas silencing marks are less frequent though they seem to play a significant role in transcriptional control of genes involved in phenotypic variation and pathogenesis.

iii) Chromatin remodeling across the parasite's cycle—Upon transcriptional activation, eukaryotic promoter nucleosomes are partially removed by sliding or disassembly, allowing DNA to become directly accessible to transcription factors (77,78) and other DNA binding proteins. Indeed, various genome-wide analyses provided evidence that active regulatory regions and gene promoters of highly expressed genes are, at least partially, nucleosome-depleted (79,80). Nucleosome positioning is typically driven by active remodeling complexes or dictated by the sequence of the binding DNA itself (81). In particular, poly(dA:dT) tracks are harder to bend around histones and nucleosomes have a lower affinity for such sequences (82,83). Considering the extremely high AT-content of *P. falciparum*'s genome, this latest observation may have important consequences for the parasite's biology.

Recently, the nucleosome landscape of *P. falciparum* was investigated in reference to gene regulation by using two genome-wide methods, both coupled to NGS: (i) Formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate protein-free DNA; and (ii) Micrococcal nuclease-Assisted Isolation of Mononucleosomal Elements (MAINE) to isolate DNA fragments associated with histones (13). The combined use of both methods provides a comprehensive view of the chromatin structure across *P. falciparum*'s genome. Complementary opposite results were obtained by both methods (nucleosome-bound regions were identified with MAINE, and interspersing nucleosome-free regions were identified with FAIRE) as reflected by a high negative correlation coefficient. Nucleosomes were

predominantly found within coding sequences, which have a higher GC-content relative to non-coding regions. Similar results were obtained using an anti-histone H4 ChIP-on-chip (52), and are consistent with three recent analyses of nucleosome distribution in human, worms, and flies demonstrating a marked preference of nucleosomes for exons (84-86). Moreover, Ponts and colleagues demonstrated the occurrence of massive and atypical genome-wide nucleosome depletion at the early trophozoite stage (“open” transcriptionally active state) before a progressive repacking of chromatin while the cycle progresses towards the schizont stage (“closed” transcriptionally silent stage) of the intra-erythrocytic cycle. These results are consistent with a dip in canonical histone abundance at trophozoite stage that was previously observed using mass spectrometry (33,38). Nonetheless, those changes in chromatin structure do not fully explain the changes of mRNA steady state levels across the intra-erythrocytic cycle, with the exception of ring stage- or exo-erythrocytic-specific genes (13-14). Such observations are consistent with recent data demonstrating that mRNA steady-state levels and transcription rate do not correlate for about half of the parasite's genes (87). In that case, genes could be massively transcribed at the trophozoite stage followed by major regulations at the post-transcriptional level. This hypothesis finds support in the fact that the parasite's pre-initiation complex interacts with both stage specific “active” and “inactive” promoters (88) and that mRNA decay rates are significantly lengthened during the intra-erythrocytic cycle suggesting major post-transcriptional regulations (65).

To further complement these data, Bartfai and colleagues used a ChIP-seq approach to show that, unlike in other eukaryotes, the histone H2A variant H2A.z is a constant and ubiquitous feature of all intergenic regions throughout the parasite erythrocytic cycle (7). Since H2A.z is usually involved in chromatin destabilization and active transcription in eukaryotes (89-91), these results are consistent with a transcriptionally permissive state of *P.*

falciparum's chromatin during the asexual cycle. In addition, previous mass spectrometry studies showed that, unlike the abundant and more variable canonical histones, H2A.z is present at low and constant level throughout the parasite's cycle (33,38). This observation, combined with the high sensitivity of H2A.z to MNase digestion (89,90), is consistent with the relative nucleosome depletion that was observed by MAINE-seq and ChIP-on-chip in non-coding regions of the genome (6,52). Given the low levels of H2A.z and its extreme sensitivity to MNase digestion, H2A.z-containing nucleosomes can mostly be detected by targeted and specific immunoprecipitation-based sample enrichments. Quantitative measurements in such experiments, however, imply a careful normalization of histone variant levels vs. canonical histones.

All together these data confirm an unusual parasite chromatin structure and speculate an active transcriptional state during most of the erythrocytic cycle with a few exceptions such as clonally variant genes as well as genes known to be essential to early erythrocytic and sexual stage differentiation. It is therefore possible that part of transcriptional regulation in *P. falciparum* could occur during elongation rather than initiation. This hypothesis is supported by the recent observation that H2A.z seems to facilitate the passage of the RNA polymerase II (91).

As a whole, it is likely that a complex model that integrates major transcription factors, histone positioning and turnover, histone variants and modifications, and post-transcriptional regulations will be required to fully comprehend the *Plasmodium*'s unusual gene expression machinery. Given the major differences observed in parasite epigenetic features compared to all other eukaryotic organisms, inhibitors developed against *Plasmodium*-specific epigenetic enzymes have a strong potential for new therapeutic strategies against *P. falciparum*.

Genomics for the discovery of new antimalarial strategies

Many of the current drug therapies are based on chemically engineered variants of already known anti-malarial compounds (e.g., aminoquinolines and/or peroxides). Intensive exploration of the *P. falciparum* genome has led to the identification of parasite specific essential genes or metabolic pathways that could be targeted for rational drug designs (18,23,60,62,92-94). For example, a fosmidomycin-sensitive mevalonate-independent pathway of isoprenoid biosynthesis, absent from higher eukaryotes and located in the plant plastid-like parasite organelle namely the apicoplast, was identified in *P. falciparum* (95). Along with the discovery of new drug targets, the discovery of mechanisms of drug resistance has been significantly refined using genome wide analysis. Typically, mechanisms of drug resistance are determined by examining the genetic differences between sensitive and resistant strains. The best-studied case of drug resistance in *P. falciparum* is chloroquine resistance (CQR). CQR is mediated by a transporter gene (*Pfcr1*) and by the multi-drug resistance gene (*Pfmdr1*). The discovery of the genes associated with CQR took years of heavy molecular, epidemiology and genetic studies. Research is still ongoing to fully comprehend CQR in the parasite. Today, whole genome analytic tools provide the capability of analyzing rapidly the genetic changes that occur in the genome of a resistant strain. Whole-genome scanning using tiling microarrays has already been used for this purpose. For example initial analyses found relatively abundant copy number variations (CNV) in *P. falciparum* resistant strains (5). Point mutations in the apicoplast were recently associated with resistance to clindamycin, a drug used in combination with quinine for the treatment of malaria in pregnant women and infants (96). Another striking example of the power of genomics in drug discovery is the identification of a potent drug by cell proliferation-based compound screening (97) followed by the discovery of one of its targets using high-density microarrays and sequencing (98). Without the advent of genomics, such a process would have required many years. All together it is likely that these genome-wide approaches will soon uncover mechanism of drug resistance including emerging resistance of artemisinin. To further highlight the power of genomic studies for the discovery of new effective antimalarial strategies, a recent genome-wide SNP analysis identified regions of high and low recombination frequencies (*hotspots* and *coldspots*). The study found that *loci* associated with known drug resistance are likely found in hotspots (99). A practical consequence of these observations for a long-term antimalarial strategy is that drug targets, should be encoded by genes located in coldspots rather than hotspots.

Genome-wide proteomic analyses have generated a high number of potential new vaccine candidates. Several new parasite surface antigens have recently been discovered throughout the malaria parasite life cycle (33-35,38,39). The availability of the *P. falciparum* genome has also allowed the development of new genome-wide protein microarrays to probe human plasma from individuals before and after malaria season. These novel genome-wide methods have already delivered important insights into parasite proteins associated with immunoreactivity in an unbiased manner (100-102). It is highly probable that these studies will soon improve our understanding of the molecular basis of protective immunity and facilitates the discovery of new efficient vaccine strategies.

All together the increasing number and performances of genome-wide technologies is transforming the scientific field. Genomics and systems biological studies have already contributed significantly to a better understanding of the malaria parasite's biology. Most importantly, they have generated an exceptional pipeline of new drugs targets and vaccine candidates. The challenge today will be to bring these achievements to efficient and affordable anti-malarial products.

What future for malaria eradication?

Constantly diminishing costs of high-throughput genomics and DNA sequencing technologies have dramatically changed the way science is being done over the past few years. These changes should soon transform the way we assess genetic risk factors, and the way we think about medicine, treatments, and possible disease eradication in developing countries.

Genomics has already greatly contributed to our understanding of the malaria parasite and the human genetic factors that influence the susceptibility and the response to both malaria and anti-malarial drugs/vaccines. The full integration of the newly acquired knowledge to the disease strategy will undoubtedly provide bases to prevent the resurgence of malaria (*e.g.*, Peru (96)) and the arising and spread of resistances by analyzing parasites' population dynamics and evolution (*e.g.*, resistances to artemisinin in south-east Asia). The catalog of putative drugs and drug targets has already increased together with the panel of candidates for vaccination strategies. Beyond drug discovery, genomics was recently proven to be particularly efficient in the discovery of a drug mechanism of action within a two year- time span by coupling drug screening and genomics (98). Ultimately, diagnostic and curative treatment could be improved by genotyping both the host and the infecting parasite. Such optimized treatment would contribute to a better use of drugs and a better management of the spread of resistances.

Integrated genomics approaches, where all aspects of the disease management are considered, have the power to lead to better malaria strategy. However, it is only with free and open access to genome databases, continuing technology development, accurate identification of genetic and environmental factors, continuing financial investments, development of close private-public partnerships, collaboration of governmental and non-governmental organizations, academic institutions, individuals and good policy decisions, that the benefits of genomics and systems biological studies can be fully utilized and manifested to achieve new drug and vaccine targets that have emerged from genomic analyses and bring us closer to the eradication of malaria.

Acknowledgments

We would like to thank Randal Maile and Vance C. Huskins for their help with proofreading the manuscript. We apologize to the authors whose works were unable to be cited due to space limitations. This work is supported by the National Institute of Allergy and Infectious Diseases and the National Institutes of Health (#1R01AI085077-01A1).

REFERENCES

1. Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, et al. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature*. Oct 3; 2002 419(6906):498–511. [PubMed: 12368864]
2. Volkman SK, Sabeti PC, DeCaprio D, Neafsey DE, Schaffner SF, Milner DA, et al. A genome-wide map of diversity in *Plasmodium falciparum*. *Nat Genet*. Jan; 2007 39(1):113–9. [PubMed: 17159979]
3. Mu J, Awadalla P, Duan J, McGee KM, Keebler J, Seydel K, et al. Genome-wide variation and identification of vaccine targets in the *Plasmodium falciparum* genome. *Nat. Genet*. Jan; 2007 39(1):126–130. [PubMed: 17159981]
4. Jeffares DC, Pain A, Berry A, Cox AV, Stalker J, Ingle CE, et al. Genome variation and evolution of the malaria parasite *Plasmodium falciparum*. *Nat Genet*. Jan; 2007 39(1):120–5. [PubMed: 17159978]
5. Kidgell C, Volkman SK, Daily J, Borevitz JO, Plouffe D, Zhou Y, et al. A systematic map of genetic variation in *Plasmodium falciparum*. *PLoS Pathog*. Jun.2006 2(6):e57. [PubMed: 16789840]

6. Ponts N, Harris EY, Prudhomme J, Wick I, Eckhardt-Ludka C, Hicks GR, et al. Nucleosome landscape and control of transcription in the human malaria parasite. *Genome Res.* Feb; 2010 20(2): 228–238. [PubMed: 20054063]
7. Bártfai R, Hoeijmakers WAM, Salcedo-Amaya AM, Smits AH, Janssen-Megens E, Kaan A, et al. H2A.Z demarcates intergenic regions of the plasmodium falciparum epigenome that are dynamically marked by H3K9ac and H3K4me3. *PLoS Pathog.* 2010; 6(12):e1001223. [PubMed: 21187892]
8. Kozarewa I, Ning Z, Quail MA, Sanders MJ, Berriman M, Turner DJ. Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of (G+C)-biased genomes. *Nat. Methods.* Apr; 2009 6(4):291–295. [PubMed: 19287394]
9. Hunt P, Martinelli A, Modrzynska K, Borges S, Creasey A, Rodrigues L, et al. Experimental evolution, genetic analysis and genome re-sequencing reveal the mutation conferring artemisinin resistance in an isogenic lineage of malaria parasites. *BMC Genomics.* 2010; 11:499. [PubMed: 20846421]
10. Martinelli A, Henriques G, Cravo P, Hunt P. Whole genome re-sequencing identifies a mutation in an ABC transporter (mdr2) in a Plasmodium chabaudi clone with altered susceptibility to antifolate drugs. *Int. J. Parasitol.* Feb; 2011 41(2):165–171. [PubMed: 20858498]
11. Dharia NV, Bright AT, Westenberger SJ, Barnes SW, Batalov S, Kuhen K, et al. Whole-genome sequencing and microarray analysis of ex vivo Plasmodium vivax reveal selective pressure on putative drug resistance genes. *Proc. Natl. Acad. Sci. U.S.A.* Nov 16; 2010 107(46):20045–20050. [PubMed: 21037109]
12. Dimon MT, Sorber K, DeRisi JL. HMMSplicer: a tool for efficient and sensitive discovery of known and novel splice junctions in RNA-Seq data. *PLoS ONE.* 2010; 5(11):e13875. [PubMed: 21079731]
13. Otto TD, Wilinski D, Assefa S, Keane TM, Sarry LR, Böhme U, et al. New insights into the blood-stage transcriptome of Plasmodium falciparum using RNA-Seq. *Mol. Microbiol.* Apr; 2010 76(1): 12–24. [PubMed: 20141604]
14. Sorber K, Dimon MT, DeRisi JL. RNA-Seq analysis of splicing in Plasmodium falciparum uncovers new splice junctions, alternative splicing and splicing of antisense transcripts. *Nucleic Acids Res.* May; 2011 39(9):3820–3835. [PubMed: 21245033]
15. Tuda J, Mongan AE, Tolba MEM, Imada M, Yamagishi J, Xuan X, et al. Full-parasites: database of full-length cDNAs of apicomplexa parasites, 2010 update. *Nucleic Acids Res.* Jan; 2011 39(Database issue):D625–631. [PubMed: 21051343]
16. Ponts N, Harris EY, Lonardi S, Le Roch KG. Nucleosome occupancy at transcription start sites in the human malaria parasite: a hard-wired evolution of virulence? *Infect. Genet. Evol.* Jun; 2011 11(4):716–724. [PubMed: 20708104]
17. Le Roch KG, Zhou Y, Blair PL, Grainger M, Moch JK, Haynes JD, et al. Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science.* Sep 12; 2003 301(5639):1503–1508. [PubMed: 12893887]
18. Llinás M, Bozdech Z, Wong ED, Adai AT, DeRisi JL. Comparative whole genome transcriptome analysis of three Plasmodium falciparum strains. *Nucleic Acids Res.* 2006; 34(4):1166–1173. [PubMed: 16493140]
19. Young JA, Fivelman QL, Blair PL, de la Vega P, Le Roch KG, Zhou Y, et al. The Plasmodium falciparum sexual development transcriptome: a microarray analysis using ontology-based pattern identification. *Mol. Biochem. Parasitol.* Sep; 2005 143(1):67–79. [PubMed: 16005087]
20. Bozdech Z, Mok S, Hu G, Imwong M, Jaidee A, Russell B, et al. The transcriptome of Plasmodium vivax reveals divergence and diversity of transcriptional regulation in malaria parasites. *Proc. Natl. Acad. Sci. U.S.A.* Oct 21; 2008 105(42):16290–16295. [PubMed: 18852452]
21. Sorber K, Chiu C, Webster D, Dimon M, Ruby JG, Hekele A, et al. The long march: a sample preparation technique that enhances contig length and coverage by high-throughput short-read sequencing. *PLoS ONE.* 2008; 3(10):e3495. [PubMed: 18941527]
22. Siau A, Toure FS, Ouwe-Missi-Oukem-Boyer O, Ciceron L, Mahmoudi N, Vaquero C, et al. Whole-transcriptome analysis of Plasmodium falciparum field isolates: identification of new pathogenicity factors. *J. Infect. Dis.* Dec 1; 2007 196(11):1603–1612. [PubMed: 18008243]

23. Daily JP, Scanfeld D, Pochet N, Le Roch K, Plouffe D, Kamal M, et al. Distinct physiological states of *Plasmodium falciparum* in malaria-infected patients. *Nature*. Dec 13; 2007 450(7172): 1091–1095. [PubMed: 18046333]
24. Daily JP, Le Roch KG, Sarr O, Fang X, Zhou Y, Ndir O, et al. In vivo transcriptional profiling of *Plasmodium falciparum*. *Malar. J.* Aug 5.2004 3:30. [PubMed: 15296511]
25. Daily JP, Le Roch KG, Sarr O, Ndiaye D, Lukens A, Zhou Y, et al. In vivo transcriptome of *Plasmodium falciparum* reveals overexpression of transcripts that encode surface proteins. *J. Infect. Dis.* Apr 1; 2005 191(7):1196–1203. [PubMed: 15747257]
26. Mackinnon MJ, Li J, Mok S, Kortok MM, Marsh K, Preiser PR, et al. Comparative transcriptional and genomic analysis of *Plasmodium falciparum* field isolates. *PLoS Pathog.* Oct.2009 5(10):e1000644. [PubMed: 19898609]
27. Petter M, Bonow I, Klinkert M-Q. Diverse expression patterns of subgroups of the rif multigene family during *Plasmodium falciparum* gametocytogenesis. *PLoS ONE*. 2008; 3(11):e3779. [PubMed: 19020666]
28. Wang CW, Mwakalinga SB, Sutherland CJ, Schwank S, Sharp S, Hermesen CC, et al. Identification of a major rif transcript common to gametocytes and sporozoites of *Plasmodium falciparum*. *Malar. J.* 2010; 9:147. [PubMed: 20509952]
29. Raabe CA, Sanchez CP, Randau G, Robeck T, Skryabin BV, Chinni SV, et al. A global view of the nonprotein-coding transcriptome in *Plasmodium falciparum*. *Nucleic Acids Res.* Jan; 2010 38(2): 608–617. [PubMed: 19864253]
30. Chakrabarti K, Pearson M, Grate L, Sterne-Weiler T, Deans J, Donohue JP, et al. Structural RNAs of known and unknown function identified in malaria parasites by comparative genomics and RNA analysis. *RNA*. Nov; 2007 13(11):1923–1939. [PubMed: 17901154]
31. Li F, Sonbuchner L, Kyes SA, Epp C, Deitsch KW. Nuclear non-coding RNAs are transcribed from the centromeres of *Plasmodium falciparum* and are associated with centromeric chromatin. *J. Biol. Chem.* Feb 29; 2008 283(9):5692–5698. [PubMed: 18165241]
32. Mourier T, Carret C, Kyes S, Christodoulou Z, Gardner PP, Jeffares DC, et al. Genome-wide discovery and verification of novel structured RNAs in *Plasmodium falciparum*. *Genome Res.* Feb; 2008 18(2):281–292. [PubMed: 18096748]
33. Florens L, Washburn MP, Raine JD, Anthony RM, Grainger M, Haynes JD, et al. A proteomic view of the *Plasmodium falciparum* life cycle. *Nature*. Oct 3; 2002 419(6906):520–526. [PubMed: 12368866]
34. Florens L, Liu X, Wang Y, Yang S, Schwartz O, Peglar M, et al. Proteomics approach reveals novel proteins on the surface of malaria-infected erythrocytes. *Mol. Biochem. Parasitol.* May; 2004 135(1):1–11. [PubMed: 15287581]
35. Lasonder E, Ishihama Y, Andersen JS, Vermunt AMW, Pain A, Sauerwein RW, et al. Analysis of the *Plasmodium falciparum* proteome by high-accuracy mass spectrometry. *Nature*. Oct 3; 2002 419(6906):537–542. [PubMed: 12368870]
36. LaCount DJ, Vignali M, Chettier R, Phansalkar A, Bell R, Hesselberth JR, et al. A protein interaction network of the malaria parasite *Plasmodium falciparum*. *Nature*. Nov 3; 2005 438(7064):103–107. [PubMed: 16267556]
37. Khan SM, Franke-Fayard B, Mair GR, Lasonder E, Janse CJ, Mann M, et al. Proteome analysis of separated male and female gametocytes reveals novel sex-specific *Plasmodium* biology. *Cell*. Jun 3; 2005 121(5):675–687. [PubMed: 15935755]
38. Le Roch KG, Johnson JR, Florens L, Zhou Y, Santrosyan A, Grainger M, et al. Global analysis of transcript and protein levels across the *Plasmodium falciparum* life cycle. *Genome Res.* Nov; 2004 14(11):2308–2318. [PubMed: 15520293]
39. Sam-Yellowe TY, Florens L, Wang T, Raine JD, Carucci DJ, Sinden R, et al. Proteome analysis of rhoptry-enriched fractions isolated from *Plasmodium merozoites*. *J. Proteome Res.* Oct; 2004 3(5): 995–1001. [PubMed: 15473688]
40. Lakshmanan V, Rhee KY, Daily JP. Metabolomics and malaria biology. *Mol. Biochem. Parasitol.* Feb; 2011 175(2):104–111. [PubMed: 20970461]

41. Fridman E, Pichersky E. Metabolomics, genomics, proteomics, and the identification of enzymes and their substrates and products. *Curr. Opin. Plant Biol.* Jun; 2005 8(3):242–248. [PubMed: 15860420]
42. van Brummelen AC, Olszewski KL, Wilinski D, Llinás M, Louw AI, Birkholtz L-M. Co-inhibition of *Plasmodium falciparum* S-adenosylmethionine decarboxylase/ornithine decarboxylase reveals perturbation-specific compensatory mechanisms by transcriptome, proteome, and metabolome analyses. *J. Biol. Chem.* Feb 13; 2009 284(7):4635–4646. [PubMed: 19073607]
43. Lian L-Y, Al-Helal M, Roslani AM, Fisher N, Bray PG, Ward SA, et al. Glycerol: an unexpected major metabolite of energy metabolism by the human malaria parasite. *Malar. J.* 2009; 8:38. [PubMed: 19267910]
44. Olszewski KL, Morrissey JM, Wilinski D, Burns JM, Vaidya AB, Rabinowitz JD, et al. Host-parasite interactions revealed by *Plasmodium falciparum* metabolomics. *Cell Host Microbe.* Feb 19; 2009 5(2):191–199. [PubMed: 19218089]
45. Teng R, Junankar PR, Bubba WA, Rae C, Mercier P, Kirk K. Metabolite profiling of the intraerythrocytic malaria parasite *Plasmodium falciparum* by (1)H NMR spectroscopy. *NMR Biomed.* Apr; 2009 22(3):292–302. [PubMed: 19021153]
46. Balu B, Chauhan C, Maher SP, Shoue DA, Kissinger JC, Fraser MJ, et al. piggyBac is an effective tool for functional analysis of the *Plasmodium falciparum* genome. *BMC Microbiol.* 2009; 9:83. [PubMed: 19422698]
47. Balu B, Shoue DA, Fraser MJ, Adams JH. High-efficiency transformation of *Plasmodium falciparum* by the lepidopteran transposable element piggyBac. *Proc. Natl. Acad. Sci. U.S.A.* Nov 8; 2005 102(45):16391–16396. [PubMed: 16260745]
48. van Ooij C, Tamez P, Bhattacharjee S, Hiller NL, Harrison T, Liolios K, et al. The malaria secretome: from algorithms to essential function in blood stage infection. *PLoS Pathog.* Jun.2008 4(6):e1000084. [PubMed: 18551176]
49. Balu B, Singh N, Maher SP, Adams JH. A genetic screen for attenuated growth identifies genes crucial for intraerythrocytic development of *Plasmodium falciparum*. *PLoS ONE.* 2010; 5(10):e13282. [PubMed: 20949012]
50. Elemento O, Slonim N, Tavazoie S. A universal framework for regulatory element discovery across all genomes and data types. *Mol. Cell.* Oct 26; 2007 28(2):337–350. [PubMed: 17964271]
51. Gonzales JM, Patel JJ, Ponmee N, Jiang L, Tan A, Maher SP, et al. Regulatory hotspots in the malaria parasite genome dictate transcriptional variation. *PLoS Biol.* Sep 30.2008 6(9):e238. [PubMed: 18828674]
52. Westenberger SJ, Cui L, Dharia N, Winzeler E, Cui L. Genome-wide nucleosome mapping of *Plasmodium falciparum* reveals histone-rich coding and histone-poor intergenic regions and chromatin remodeling of core and subtelomeric genes. *BMC Genomics.* 2009; 10:610. [PubMed: 20015349]
53. Coulson RM, Hall N, Ouzounis CA. Comparative genomics of transcriptional control in the human malaria parasite *Plasmodium falciparum*. *Genome Res.* Aug; 2004 14(8):1548–54. [PubMed: 15256513]
54. Balaji S, Babu MM, Iyer LM, Aravind L. Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic Acids Res.* 2005; 33(13):3994–4006. [PubMed: 16040597]
55. Painter HJ, Campbell TL, Llinás M. The Apicomplexan AP2 family: integral factors regulating *Plasmodium* development. *Mol. Biochem. Parasitol.* Mar; 2011 176(1):1–7. [PubMed: 21126543]
56. Templeton TJ. The varieties of gene amplification, diversification and hypervariability in the human malaria parasite, *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* Aug; 2009 166(2):109–116. [PubMed: 19375460]
57. Aravind L, Iyer LM, Wellem TE, Miller LH. *Plasmodium* Biology: Genomic Gleanings. *Cell.* Décembre; 2003 115(7):771–785. [PubMed: 14697197]
58. Campbell TL, De Silva EK, Olszewski KL, Elemento O, Llinás M. Identification and genome-wide prediction of DNA binding specificities for the ApiAP2 family of regulators from the malaria parasite. *PLoS Pathog.* 2010; 6(10):e1001165. [PubMed: 21060817]

59. Ganesan K, Ponmee N, Jiang L, Fowble JW, White J, Kamchonwongpaisan S, et al. A genetically hard-wired metabolic transcriptome in *Plasmodium falciparum* fails to mount protective responses to lethal antifolates. *PLoS Pathog.* Nov.2008 4(11):e1000214. [PubMed: 19023412]
60. Le Roch KG, Johnson JR, Ahiboh H, Chung DW, Prudhomme J, Plouffe D, et al. A systematic approach to understand the mechanism of action of the bithiazolium compound T4 on the human malaria parasite, *Plasmodium falciparum*. *BMC Genomics.* 2008; 9:513. [PubMed: 18973684]
61. Hu G, Cabrera A, Kono M, Mok S, Chaal BK, Haase S, et al. Transcriptional profiling of growth perturbations of the human malaria parasite *Plasmodium falciparum*. *Nat. Biotechnol.* Jan; 2010 28(1):91–98. [PubMed: 20037583]
62. Hall N, Karras M, Raine JD, Carlton JM, Kooij TWA, Berriman M, et al. A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic, and proteomic analyses. *Science.* Jan 7; 2005 307(5706):82–86. [PubMed: 15637271]
63. Foth BJ, Ralph SA, Tonkin CJ, Struck NS, Fraunholz M, Roos DS, et al. Dissecting apicoplast targeting in the malaria parasite *Plasmodium falciparum*. *Science.* Jan 31; 2003 299(5607):705–708. [PubMed: 12560551]
64. Mair GR, Braks JAM, Garver LS, Wiegant JCAG, Hall N, Dirks RW, et al. Regulation of sexual development of *Plasmodium* by translational repression. *Science.* Aug 4; 2006 313(5787):667–669. [PubMed: 16888139]
65. Shock JL, Fischer KF, DeRisi JL. Whole-genome analysis of mRNA decay in *Plasmodium falciparum* reveals a global lengthening of mRNA half-life during the intra-erythrocytic development cycle. *Genome Biol.* 2007; 8(7):R134. [PubMed: 17612404]
66. Coulson RMR, Hall N, Ouzounis CA. Comparative genomics of transcriptional control in the human malaria parasite *Plasmodium falciparum*. *Genome Res.* Aug; 2004 14(8):1548–1554. [PubMed: 15256513]
67. Duraisingh MT, Voss TS, Marty AJ, Duffy MF, Good RT, Thompson JK, et al. Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in *Plasmodium falciparum*. *Cell.* Apr 8; 2005 121(1):13–24. [PubMed: 15820675]
68. Freitas-Junior LH, Hernandez-Rivas R, Ralph SA, Montiel-Condado D, Ruvalcaba-Salazar OK, Rojas-Meza AP, et al. Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. *Cell.* Apr 8; 2005 121(1):25–36. [PubMed: 15820676]
69. Cui L, Miao J. Chromatin-mediated epigenetic regulation in the malaria parasite *Plasmodium falciparum*. *Eukaryotic Cell.* Aug; 2010 9(8):1138–1149. [PubMed: 20453074]
70. Workman JL, Kingston RE. Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annu. Rev. Biochem.* 1998; 67:545–579. [PubMed: 9759497]
71. Chookajorn T, Dzikowski R, Frank M, Li F, Jiwani AZ, Hartl DL, et al. Epigenetic memory at malaria virulence genes. *Proc. Natl. Acad. Sci. U.S.A.* Jan 16; 2007 104(3):899–902. [PubMed: 17209011]
72. Lopez-Rubio J-J, Mancio-Silva L, Scherf A. Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. *Cell Host Microbe.* Feb 19; 2009 5(2):179–190. [PubMed: 19218088]
73. Salcedo-Amaya AM, van Driel MA, Alako BT, Trelle MB, van den Elzen AMG, Cohen AM, et al. Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. U.S.A.* Jun 16; 2009 106(24):9655–9660. [PubMed: 19497874]
74. Flueck C, Bartfai R, Volz J, Niederwieser I, Salcedo-Amaya AM, Alako BTF, et al. *Plasmodium falciparum* heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors. *PLoS Pathog.* Sep.2009 5(9):e1000569. [PubMed: 19730695]
75. Pérez-Toledo K, Rojas-Meza AP, Mancio-Silva L, Hernández-Cuevas NA, Delgadillo DM, Vargas M, et al. *Plasmodium falciparum* heterochromatin protein 1 binds to tri-methylated histone 3 lysine 9 and is linked to mutually exclusive expression of var genes. *Nucleic Acids Res.* May; 2009 37(8):2596–2606. [PubMed: 19270070]
76. Trelle MB, Salcedo-Amaya AM, Cohen AM, Stunnenberg HG, Jensen ON. Global histone analysis by mass spectrometry reveals a high content of acetylated lysine residues in the malaria

- parasite *Plasmodium falciparum*. *J. Proteome Res.* Jul; 2009 8(7):3439–3450. [PubMed: 19351122]
77. Saha A, Wittmeyer J, Cairns BR. Chromatin remodelling: the industrial revolution of DNA around histones. *Nat. Rev. Mol. Cell Biol.* Jun; 2006 7(6):437–447. [PubMed: 16723979]
 78. Boeger H, Griesenbeck J, Strattan JS, Kornberg RD. Removal of promoter nucleosomes by disassembly rather than sliding in vivo. *Mol. Cell.* Jun 4; 2004 14(5):667–673. [PubMed: 15175161]
 79. Lee C-K, Shibata Y, Rao B, Strahl BD, Lieb JD. Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nat. Genet.* Aug; 2004 36(8):900–905. [PubMed: 15247917]
 80. Hogan GJ, Lee C-K, Lieb JD. Cell cycle-specified fluctuation of nucleosome occupancy at gene promoters. *PLoS Genet.* Sep 22; 2006 2(9):e158. [PubMed: 17002501]
 81. Segal E, Fondufe-Mittendorf Y, Chen L, Thåström A, Field Y, Moore IK, et al. A genomic code for nucleosome positioning. *Nature.* Aug 17; 2006 442(7104):772–778. [PubMed: 16862119]
 82. Segal E, Widom J. Poly(dA:dT) tracts: major determinants of nucleosome organization. *Curr. Opin. Struct. Biol.* Feb; 2009 19(1):65–71. [PubMed: 19208466]
 83. Kaplan N, Moore I, Fondufe-Mittendorf Y, Gossett AJ, Tillo D, Field Y, et al. Nucleosome sequence preferences influence in vivo nucleosome organization. *Nat. Struct. Mol. Biol.* Aug; 2010 17(8):918–920. author reply 920–922. [PubMed: 20683473]
 84. Andersson R, Enroth S, Rada-Iglesias A, Wadelius C, Komorowski J. Nucleosomes are well positioned in exons and carry characteristic histone modifications. *Genome Res.* Oct; 2009 19(10):1732–1741. [PubMed: 19687145]
 85. Schwartz S, Meshorer E, Ast G. Chromatin organization marks exon-intron structure. *Nat. Struct. Mol. Biol.* Sep; 2009 16(9):990–995. [PubMed: 19684600]
 86. Tilgner H, Nikolaou C, Althammer S, Sammeth M, Beato M, Valcárcel J, et al. Nucleosome positioning as a determinant of exon recognition. *Nat. Struct. Mol. Biol.* Sep; 2009 16(9):996–1001. [PubMed: 19684599]
 87. Sims JS, Militello KT, Sims PA, Patel VP, Kasper JM, Wirth DF. Patterns of gene-specific and total transcriptional activity during the *Plasmodium falciparum* intraerythrocytic developmental cycle. *Eukaryotic Cell.* Mar; 2009 8(3):327–338. [PubMed: 19151330]
 88. Gopalakrishnan AM, Nyindodo LA, Ross Fergus M, López-Estraño C. *Plasmodium falciparum*: Preinitiation complex occupancy of active and inactive promoters during erythrocytic stage. *Exp. Parasitol.* Jan; 2009 121(1):46–54. [PubMed: 18951895]
 89. Jin C, Felsenfeld G. Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes Dev.* Jun 15; 2007 21(12):1519–1529. [PubMed: 17575053]
 90. Fan JY, Gordon F, Luger K, Hansen JC, Tremethick DJ. The essential histone variant H2A.Z regulates the equilibrium between different chromatin conformational states. *Nat. Struct. Biol.* Mar; 2002 9(3):172–176. [PubMed: 11850638]
 91. Santisteban MS, Hang M, Smith MM. Histone variant H2A.Z and RNA polymerase II transcription elongation. *Mol. Cell. Biol.* May; 2011 31(9):1848–1860. [PubMed: 21357739]
 92. Mok BW, Ribacke U, Rasti N, Kironde F, Chen Q, Nilsson P, et al. Default Pathway of var2csa switching and translational repression in *Plasmodium falciparum*. *PLoS ONE.* 2008; 3(4):e1982. [PubMed: 18431472]
 93. Natalang O, Bischoff E, Deplaine G, Proux C, Dillies M-A, Sismeiro O, et al. Dynamic RNA profiling in *Plasmodium falciparum* synchronized blood stages exposed to lethal doses of artesunate. *BMC Genomics.* 2008; 9:388. [PubMed: 18706115]
 94. Jiang H, Yi M, Mu J, Zhang L, Ivens A, Klimczak LJ, et al. Detection of genome-wide polymorphisms in the AT-rich *Plasmodium falciparum* genome using a high-density microarray. *BMC Genomics.* 2008; 9:398. [PubMed: 18724869]
 95. Jomaa H, Wiesner J, Sanderbrand S, Altincicek B, Weidemeyer C, Hintz M, et al. Inhibitors of the Nonmevalonate Pathway of Isoprenoid Biosynthesis as Antimalarial Drugs. *Science.* Sep 3; 1999 285(5433):1573–1576. [PubMed: 10477522]
 96. Dharia NV, Plouffe D, Bopp SER, González-Pérez GE, Lucas C, Salas C, et al. Genome scanning of Amazonian *Plasmodium falciparum* shows subtelomeric instability and clindamycin-resistant parasites. *Genome Res.* Nov; 2010 20(11):1534–1544. [PubMed: 20829224]

97. Plouffe D, Brinker A, McNamara C, Henson K, Kato N, Kuhen K, et al. In silico activity profiling reveals the mechanism of action of antimalarials discovered in a high-throughput screen. *Proc. Natl. Acad. Sci. U.S.A.* Jul 1; 2008 105(26):9059–9064. [PubMed: 18579783]
98. Rottmann M, McNamara C, Yeung BKS, Lee MCS, Zou B, Russell B, et al. Spiroindolones, a potent compound class for the treatment of malaria. *Science*. Sep 3; 2010 329(5996):1175–1180. [PubMed: 20813948]
99. Mu J, Myers RA, Jiang H, Liu S, Ricklefs S, Waisberg M, et al. *Plasmodium falciparum* genome-wide scans for positive selection, recombination hot spots and resistance to antimalarial drugs. *Nat Genet.* Mar; 2010 42(3):268–271. [PubMed: 20101240]
100. Crompton PD, Kayala MA, Traore B, Kayentao K, Ongoiba A, Weiss GE, et al. A prospective analysis of the Ab response to *Plasmodium falciparum* before and after a malaria season by protein microarray. *Proc. Natl. Acad. Sci. U.S.A.* Apr 13; 2010 107(15):6958–6963. [PubMed: 20351286]
101. Doolan DL, Mu Y, Unal B, Sundaresh S, Hirst S, Valdez C, et al. Profiling humoral immune responses to *P. falciparum* infection with protein microarrays. *Proteomics*. Nov; 2008 8(22):4680–4694. [PubMed: 18937256]
102. Sundaresh S, Doolan DL, Hirst S, Mu Y, Unal B, Davies DH, et al. Identification of humoral immune responses in protein microarrays using DNA microarray data analysis techniques. *Bioinformatics*. Jul 15; 2006 22(14):1760–1766. [PubMed: 16644788]

Figure 1. Methods for next-generation sequencing of the human malaria parasite's genome

Library preparation for sequencing on Illumina® NGS platforms usually involves the fragmentation of the dsDNA by a method of choice (*e.g.*, sonication), end repair (fill in), and the addition of a single A tail on the 3' end of each DNA strand before specific and oriented double-stranded adapters can be ligated. These chimera fragments are then PCR-amplified using specific primers that contain sequences compatible with the oligos attached to the Illumina® sequencing flowcell. In the amplification-free protocol (8), the PCR step is entirely skipped. The adapters are thus customized in order to contain the flowcell-compatible sequences in the absence of PCR. After adapter ligation, the library can be eventually size selected and purified prior to hybridization on the flowcell and sequencing. In both the optimized amplification (6) and the linear amplification protocols (7) the PCR step is performed with a polymerase that is different from the original Illumina® protocol. The linear amplification protocol uses the T7 polymerase and therefore requires that one of the adapters contains the sequence corresponding to the T7 promoter. In the case of the optimized amplification protocol, a mix of enzymes is used to perform the PCR (for example, a combination of the NEB Phusion and the Takara Ex Taq have been successfully used in the past (6,16)). The parameters for PCR cycling have also been adapted to the very high AT-content of the genome.

Figure 2. A system biology approach to understand the parasite biology for the design of new drug and vaccines strategies

Genomes to identify genetic diversity and drug resistances, *Epigenomes* to understand transcriptional regulation, *Transcriptomes* to evaluate mRNA steady state, *Proteomes* to evaluate protein levels, *Immunomes* to discovered protective antigens, *Interactomes* to comprehend protein-protein interactions, *Metabolomes* to evaluate metabolites.