Chapter 8
Malaria Genomics and the Developing World

Christopher G. Jacob and Christopher V. Plowe

Essentials of Malaria

Malaria is a disease caused by parasites from genus *Plasmodium*, a member of the Apicomplexan family. Apicomplexans are unique in that they are the only fully parasitic large clade on the tree of life. It is thought that every type of mammal, bird, and reptile is parasitized by at least one species of *Plasmodium* (Morrison 2009). Five *Plasmodium* species infect humans: *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium malariae*, *Plasmodium ovale*, and *Plasmodium knowlesi*, which until recently was thought to infect only nonhuman primates (Cox-Singh et al. 2008). *P. falciparum* and *P. vivax* are the most prevalent species, and *P. falciparum* is responsible for most cases of severe malaria and death. A key feature of Apicomplexans is that these eukaryotic organisms exist mainly in a haploid state with most having only a brief obligatory diploid phase. Another difference is the apicoplast, a plastid believed to have originated from the phagocytosis of a chloroplast-containing microorganism (Waller and McFadden 2005).

Current estimates place over one-third of the world's population at risk of *Plasmodium spp*. infection. Malaria is transmitted throughout most of the tropics and subtropics, including sub-Saharan Africa, much of Asia, parts of Central and South America, and parts of the Middle East. However, most of the at-risk population lives in tropical Africa, where transmission is the highest, with an estimated 300–500 million cases of clinical malaria illness each year, of which slightly less than one million end in death, putting malaria in the top three major global killers along with HIV/AIDS and tuberculosis. Malaria genomic research to date has focused largely on *P. falciparum* as the leading cause of malaria-associated disease

C.G. Jacob • C.V. Plowe (✉)
Howard Hughes Medical Institute/Center for Vaccine Development,
University of Maryland School of Medicine, Baltimore, MD 21201, USA
e-mail: cplowe@medicine.umaryland.edu

K.E. Nelson and B. Jones-Nelson (eds.), *Genomics Applications for the Developing World*, Advances in Microbial Ecology,
DOI 10.1007/978-1-4614-2182-5_8,

and death, although the genomes of *P. vivax* (Carlton et al. 2008), the most common human malaria outside of Africa, and several animal *Plasmodia* used as model systems, have now been sequenced (Hall et al. 2005; Carlton et al. 2002; Pain et al. 2008), as shown in Table 8.1.

As a mosquito-borne infection, malaria follows the rains. Where transmission is intense (as many as 1,000 infected mosquito bites per person per year), as in much of sub-Saharan Africa, repeated exposure throughout early childhood results in naturally acquired immunity that protects older children and adults against disease but does not fully prevent infection. This immunity provides balancing selection pressure that drives extreme genetic diversity in the malaria antigens targeted by immune responses (Conway and Polley 2002). Immune protection is thought to require repeated exposure to diverse parasites to build up a full repertoire of allele-specific immune responses to parasites that infect a given human population (Hviid 2005). In the absence of this repeated exposure to malaria infection, immunity wanes over the course of a few years.

Infection with a *Plasmodium sp*. results in clinical symptoms ranging from asymptomatic infection to classical malarial fever paroxysms, to severe forms of *falciparum* malaria such as profound anemia, coma, seizures, and death, typically from respiratory failure. The greater pathogenic potential of *P. falciparum* is chiefly attributed to a large, diverse family of *var* genes encoding proteins expressed one at a time on the surface of the host red blood cell that mediate cytoadherence to host tissues. Gene switching among *vars* is thought to allow *P. falciparum* parasites to evade *var*-specific immune responses, and the interplay between allele-specific immune responses and expression of *vars* and other genes encoding variant antigens is likely to account for much of the variation in clinical manifestations of malaria (Su et al. 1995; Baruch et al. 1995; Smith et al. 1995).

The life cycle of malaria parasites is similar within the genus and begins in the female *Anopheles* mosquito. Wormlike sporozoites are injected into the bloodstream by a biting infected mosquito and quickly migrate to the liver where they enter hepatocytes. There, they develop into merozoites that multiply until the hepatocyte ruptures, releasing free merozoites that reenter the circulation where they invade red blood cells and either enter a continuous cycle of merozoite development and red cell invasion or develop into male and female gametocytes, which can be taken up during another mosquito blood meal. In the mosquito midgut, male gametes fertilize females to create the only diploid stage of the life cycle, providing an opportunity for genetic recombination when genetically nonidentical gametes mate. After mating, haploid sporozoites migrate to the mosquito salivary glands for subsequent infection.

Sequencing Human Plasmodia Genomes

Sequencing of the *P. falciparum* genome began in the mid-1990s. Parts of the genome were completed by 1998, and the entire draft genome was completed in 2002 (Gardner et al. 2002). This was accomplished in parallel with completion of

Table 8.1 Malaria sequencing status. Parasites selected for sequencing were based primarily on infectivity in humans, animal models of parasite biology, and phylogenetic relationships

Species	Strain/isolate/location	Coverage/status	Host	Citation
P. falciparum	3D7	Complete [published]	Human and nonhuman primates	(Gardner et al. 2002)
	>25 Isolates	In progress[a]		Plasmodium Writing Group
	>1,000 Isolates	Draft[b]		Wellcome Trust Sanger Institute
P. vivax	Salvador I	10×[Published]	Humans	(Carlton et al. 2008)
P. knowlesi	H	8×[Published]	Human and nonhuman primates	(Pain et al. 2008)
P. ovale	Nigeria I	8×[Incomplete]	Humans	Wellcome Trust Sanger Institute
P. coatneyi	Malaysia	In progress	Nonhuman primates	Plasmodium Writing Group
P. cynomolgi	Berok	In progress	Nonhuman primates	Plasmodium Writing Group
P. fragile	Sri Lanka	In progress	Nonhuman primates	Plasmodium Writing Group
P. inui	OS	In progress	Nonhuman primates	Plasmodium Writing Group
P. reichenowi	Oscar	In progress	Nonhuman primates	Wellcome Trust Sanger Institute
P. berghei	ANKA	3×[Published]	Rodents	(Hall et al. 2005)
	NK65	In progress		Plasmodium Writing Group
P. chabaudi	AS	8×[Published]	Rodents	(Hall et al. 2005)
P. vinckei	*P. v. vinckei*	In progress	Rodents	Plasmodium Writing Group
	P. v. petteri	In progress		Plasmodium Writing Group
P. yoelii	17XNL	5×[Published]	Rodents	+
	17XA	In progress		Plasmodium Writing Group
	17XYM	In progress		Plasmodium Writing Group
P. gallinaceum	A	3×[Incomplete]	Birds	Wellcome Trust Sanger Institute
P. relictum	K1	In progress	Birds	Plasmodium Writing Group
	KV115	In progress		Plasmodium Writing Group
P. mexicanum	U.S.A	In progress	Lizards	Plasmodium Writing Group

Plasmodium Writing Group, white paper listing approved projects sequencing malarial genomes, available at http://www.broadinstitute.org/files/shared/genomebio/Plasmodium100genomesWhitePaper.pdf
Wellcome Trust Sanger Institute, list of sequencing projects is available at http://www.sanger.ac.uk/resources/downloads/protozoa/

[a]More than 25 cloned *P. falciparum* isolates are currently in various stages of sequencing and assembly

[b]More than 1,000 field isolates of *P. falciparum* have been genotyped by genomic sequencing using next-generation technologies and mapping sequence reads to the reference genome, without plans for full assembly

the human genome in 2001 (Venter et al. 2001) and the sequencing of the *Anopheles gambiae* (the leading mosquito vector of malaria) genome also in 2002 (Holt et al. 2002). The completion of all three genomes marked the first time that genomes of a parasite, vector, and host for an infectious disease were available (Anonymous 2002).

The genome of *P. falciparum* posed challenges not previously encountered with sequencing eukaryote genomes. The ~23 Mb genome is extremely A+T rich with an average G+C content of only 19.4%, dropping in intergenic regions to 13.5%. The ~5,300 nuclear-encoded genes are distributed across 14 chromosomes varying in length from 650 Kb (chromosome 1) to 3.3 Mb (chromosome 14). Of the initially identified genes, over 60% lack significant homology to genes in other eukaryotes, leaving the majority of the *P. falciparum* genome not yet annotated. Also encoded by *P. falciparum* are the apicoplast and mitochondrial genomes consisting of 35 and 6 Kb, respectively (Gardner et al. 2002). Other basic insights gained from sequencing the *P. falciparum* genome included chromosomal location of genes involved in antigenic variation on the subtelomeric regions and the relative abundance of genes involved in immune evasion and host-parasite interactions compared to the genomes of free-living eukaryotes.

The second fully sequenced human malaria parasite was *P. vivax.* The genome for this organism was completed in 2008, and like all sequenced malaria parasites, it contains 14 chromosomes. Slightly larger than *P. falciparum*, the *P. vivax* genome is ~26.8 Mb and contains ~5,400 nuclear-encoded genes. A key difference between the *P. vivax* and *P. falciparum* genomes is their G+C content. While *P. falciparum* had the lowest G+C content of any sequenced genome, *P. vivax* has an average G+C content of 42.3% (Carlton et al. 2008), the highest of all sequenced *Plasmodia* (Pain and Hertz-Fowler 2009).

Post-Genomics: Fundamental Research

Since the completed *P. falciparum* genome was published in 2002, several more draft and complete genome sequences of parasites from various hosts have been completed (Table 8.1). The availability of multiple genome sequences within the *Plasmodium* clade allows researchers to compare functional differences in organisms with their genetic differences and provides an evolutionary view of genes, potentially highlighting genes under selection that could serve as vaccine or drug targets. A key example of such comparative genomics is the differential invasion machineries of *P. falciparum* and *P. vivax.* Phenotypic differences between the two species include the restricted invasion of *P. vivax* to reticulocytes expressing the Duffy receptor, an obligatory receptor for *vivax* invasion of host erythrocytes (Miller et al. 1976). This restriction led some to speculate that the *P. vivax* invasion machinery was less complex than that in *P. falciparum* (Carlton et al. 2008). Sequencing and analysis of the *P. vivax* genome revealed an expansion of the reticulocyte binding protein family that could provide diversity similar to *P. falciparum* (Galinski et al. 1992).

Comparative genomics is not limited to studying similarity of family members. Comparison of parasites and host genomes can provide insight into the feasibility of potential drug targets. Many essential enzymes that would be potential drug targets due to their necessity for parasite survival are too genetically similar between host and parasite, making it difficult to identify gene products that can be targeted by drugs without harming the host. Cellular pathways and processes that are currently considered to be good areas for locating drug targets include metabolism, DNA replication and transcription, and protein modification enzymes (De Azevedo and Soares 2009; Yeh and Altman 2006).

Studies characterizing the transcriptome of *P. falciparum* have revealed novel parasite biology and validated gene models. Initial transcriptomic studies performed on intraerythrocytic parasites show patterns of gene expression unique to malaria. Using a large-scale parasite culture system, parasite transcripts were evaluated using an oligonucleotide array every hour post-invasion (Bozdech et al. 2003a). Results showed that transcripts are produced only during the point in the cell cycle where they are needed, and that *P. falciparum* nuclear genes are not polycistronic, whereas genes located in the apicoplast genome are coregulated (Bozdech et al. 2003b).

Malaria Genomics Helping the Developing World

Nearly every paper reporting results of genome sequencing projects, transcriptome studies, and other malaria genomics research endeavors claims that the reported advances will lead to the identification of new drug and vaccine targets. The first generation of malaria genomics studies focused on a small number of parasite strains that have been cultured in the laboratory for many generations, unexposed to human (or for that matter, mosquito) immunity and other environmental stimuli. Less widely appreciated is the notion that sequencing large numbers of wild parasite isolates from the field accompanied by demographic, clinical, and parasite phenotype information will directly inform vaccine development (Takala and Plowe 2009) and the discovery of mechanisms of drug action or markers of drug resistance. These kinds of genomic epidemiology studies, now underway, are likely to yield meaningful public health benefits for the developing world in the not-too-distant future.

Vaccine Target Discovery and Vaccine Development

Vaccine target discovery studies use sequencing and comparison of multiple isolates of the same species to locate highly polymorphic genes (Mu et al. 2007) and genes under diversifying selection (Mu et al. 2010) that likely encode immunogenic antigens. The approaches used to identify new antigens are verified by applying the same methods to known antigens, which give similar levels of selection and/or polymorphism.

The validity of this approach is constrained by the limited number of malaria antigens that have demonstrated meaningful efficacy as vaccines in humans—at present, just one such antigen, the circumsporozoite protein, is used in a vaccine formulation that prevents clinical malaria with modest but significant efficacy (Alonso et al. 2004). More highly polymorphic antigens such as the blood stage proteins merozoite surface protein 1 (MSP1) and apical membrane antigen 1 (AMA1) were thought to hold high promise as vaccine candidates based on in vitro and animal studies, but initial human trials of monovalent MSP1 (Ogutu et al. 2009) and bivalent AMA1 (Sagara et al. 2009) vaccines based on these antigens have shown no efficacy against clinical malaria. These disappointing results are likely to be in part if not chiefly due to insufficient cross-protection against malaria parasites with highly diverse forms of the vaccine antigens (Takala and Plowe 2009; Takala et al. 2007), although insufficient immunogenicity may also contribute to the lack of protective efficacy to date (Ouattara et al. 2010).

The extremely polymorphic blood stage antigen AMA1 demonstrates the pitfalls of using evidence of immune selection to pick vaccine candidates. Among 506 AMA1 sequences from parasites collected over 3 years in a single town in Mali, West Africa, 214 unique AMA1 haplotypes were detected. In the worst case, this might mean that a ~200-valent vaccine would be needed to provide complete protection against the diverse forms of malaria in a single rural town. In hopes of identifying a reduced number of variants that would be needed for a broadly protective polyvalent or chimeric AMA1 vaccine, in vitro and animal studies (Dutta et al. 2007) and molecular epidemiological approaches (Takala et al. 2009) have been used to try to pinpoint the polymorphic AMA1 codons that are most important in determining allele-specific immune responses. Based on these approaches, a single cluster of eight polymorphic codons was identified that could be used to define about ten haplotypes that might cover 80% of natural variants. Although a 10-valent malaria vaccine might still be infeasible, this complexity is within the range of currently licensed vaccines. Sequencing AMA1 from malaria episodes experienced by children immunized with a highly immunogenic AMA1 vaccine (Thera et al. 2010) in a recently completed field efficacy trial may permit further narrowing down of the number of variants that would be needed for a cross-protective AMA1 vaccine.

Recent serological profiling studies using a peptide array containing 1,200 recombinant proteins from the *P. falciparum* reference genome have suggested that protective humoral immune responses are directed against large numbers of malaria antigens (Crompton et al. 2010). This finding may help to explain the difficulty of achieving high and sustained efficacy with malaria vaccines based on just one or a few malaria antigens (Takala and Plowe 2009; Plowe et al. 2009). In the face of the failure of such subunit vaccines as well as DNA and viral-vectored vaccines to provide high-level protection, the concept of whole-organism vaccines has recently been revisited (Luke and Hoffman 2003). A radiation-attenuated, metabolically active, non-replicating whole sporozoite vaccine has been manufactured in and purified from aseptically raised mosquitoes (Hoffman et al. 2010) and was recently evaluated for safety and efficacy in an experimental sporozoite challenge trial in humans. If early clinical trials of sporozoite vaccines in the United States and Europe

demonstrate protection against homologous challenge with the same parasite clone used to make the vaccine, it will then be necessary to assess efficacy against natural heterologous challenge in high transmission settings such as Africa. It is hoped that immunizing with the very large number of antigens expressed by the whole organism will generate enough redundancy in protective immune responses to provide protection against diverse parasites. In the likely event that protection against diverse natural challenge afforded by a single-strain sporozoite vaccine is less than complete, comparing the genomes of breakthrough infections in vaccinated people and infections in unvaccinated controls will inform the design of multi-strain vaccines. This novel type of comparative genomics will identify genes encoding proteins that are under directional selection by vaccine-induced immunity, thus identifying the antigens most responsible for protective efficacy.

Drug Resistance Mechanisms and Markers

The recent emergence in Southeast Asia of *P. falciparum* resistance to the leading class of antimalarial drugs (Noedl et al. 2008; Dondorp et al. 2009) represents a problem of urgent public health importance that malaria genomics can help to address through the identification of genetic loci associated with resistance that can be used as molecular markers for surveillance of resistance. The identification of such markers for other antimalarial drugs was accomplished in the pre-genomic era, but doing so without current genomic resources and technologies took so long that resistance had already spread globally by the time markers were identified and validated as surveillance tools (Plowe et al. 2007a). Genomic science has the potential to greatly accelerate this process, particularly as genome sequencing shifts from complete sequencing and assembly of a limited number of genomes to less comprehensive but higher throughput genome-wide genotyping of large numbers of samples from field studies.

Drug Resistance Mechanisms and Markers in the Pre-Genomic Era

Nearly a decade before the international effort to sequence the *P. falciparum* genome began in earnest in 1996, the gene encoding *P. falciparum* dihydrofolate reductase (*dhfr*) was cloned and sequenced using primers based on consensus in known *dhfr* sequences from other organisms (Bzik et al. 1987). Sequencing *dhfr* from *falciparum* strains sensitive and resistant to pyrimethamine and other antifolate drugs quickly identified a set of single nucleotide polymorphisms (SNPs) that caused resistance (Peterson et al. 1988; Peterson et al. 1990) and that had potential use as surveillance tools (Plowe et al. 1995).

Chloroquine-resistant *P. falciparum* emerged on the Thailand/Cambodia border in the late 1950s, spread throughout the region, and then disseminated globally, arriving in Africa in the late 1970s. The search for a chloroquine resistance marker

was less straightforward than that for antifolate resistance. With no known mechanism of resistance, searching for an orthologous candidate gene was not possible. Lacking an assembled genome that would permit modern approaches to identify the genetic locus of chloroquine resistance such as genome-wide association studies (GWAS), a genetic cross was completed in the mid-1980s between the chloroquine-sensitive HB3 clone of *P. falciparum* from Honduras and the chloroquine-resistant clone Dd2 from an Indochina lineage parasite (Walliker et al. 1987). The parental clones were mixed in culture and fed to mosquitoes where recombination occurred, and the mosquitoes were allowed to take a blood meal on splenectomized chimpanzees. Initial mapping of the resulting progeny was completed using 85 restriction length fragment polymorphisms across the 14 chromosomes. This cross showed that neither of two known *Plasmodium* multi-drug resistance-like candidate genes were associated with the resistance phenotype, but an ~400 Kb region on chromosome 7 was identified that did associate with the phenotype. This 400 Kb region was postulated to contain anywhere from 80 to 100 protein coding genes, so further mapping was needed to narrow the region (Wellems et al. 1991; Wellems et al. 1990).

To pinpoint the genetic determinant of resistance, a high-resolution linkage map was created using 342 microsatellites or simple sequence repeats (Su et al. 1997). Over several years and with some false alarms, this map was used to resolve the chloroquine resistance locus to 36 kb, and through directly sequencing this region in cross progeny and in geographically diverse isolates, the *P. falciparum* chloroquine resistance transporter (PfCRT) was identified as the primary determinant of chloroquine resistance. The central role of PfCRT in both in vitro and clinical chloroquine resistance was proven in genetic transformation studies (Fidock et al. 2000), and a single nucleotide polymorphism (SNP) in PfCRT was validated as a molecular marker for surveillance of chloroquine resistance in field studies (Djimde et al. 2001a). PCR-based protocols for detecting the marker in DNA extracted from filter paper blood spots collected from finger-pricks were disseminated through the World Health Organization even before research results were published, and these assays were widely deployed throughout the malaria-endemic world (Djimde et al. 2001b; Djimde et al. 2004; Plowe et al. 2007b) providing an example of the potential for genomics to improve the public health in developing countries. However, these results were published in 2000 and 2001, about 15 years after the effort to identify the genetic basis of chloroquine resistance began and after chloroquine efficacy was already compromised in many parts of the world by the global dissemination of resistant forms of PfCRT.

An Opportunity to Deter Drug Resistance

As resistant forms of PfCRT, DHFR, and dihydropteroate synthase (the target of sulfa drugs) spread globally (Plowe 2009), chloroquine and the antifolate combination sulfadoxine-pyrimethamine lost efficacy against malaria, resulting in large increases in malaria deaths (Trape et al. 1998). These older drugs were replaced as first-line therapies by artemisinins, a class of compounds derived from a Chinese

herb qinghausu or *Artemisia annua* (Hien and White 1993). Artemisinins are used in conjunction with one or more partner drugs as artemisinin-based combination therapy (ACT). ACTs are fast-acting, effective, and safe drugs that represent the last line of defense in areas with multi-drug resistance. While the partnering of drugs into combination therapies is meant to deter resistance (White and Olliaro 1996), it is probable that years of artemisinin monotherapy distribution along with use of substandard/counterfeit artemisinin and ACTs contributed to the emergence of resistance (Dondorp et al. 2010), which was recently reported on the Thailand-Cambodia border (Noedl et al. 2008; Dondorp et al. 2009).

If artemisinin resistance follows the patterns established by chloroquine and antifolate resistance, which also originated along the Thailand-Cambodia border before disseminating globally, malaria deaths can be expected to sharply increase once again, reversing recent downward trends in malaria incidence and mortality and threatening to derail the renewed effort to eradicate malaria (Tanner and de Savigny 2008). The World Health Organization is coordinating an urgent effort to contain resistance in western Cambodia (World Health Organization 2011), but this initiative is hobbled by the lack of knowledge about whether and in what direction(s) resistance may be spreading from the site of origin. A molecular marker for resistance would greatly aid the containment effort and would provide a valuable tool for surveillance at sentinel sites where clinical resistance has not yet been observed.

Because they rely on dried blood spots that can be collected from finger-pricks and require no cold chain, molecular surveillance tools can be more readily standardized and widely and rapidly deployed than surveillance based on clinical protocols or on in vitro assays that require frozen venous blood (Plowe et al. 2007b). Research to identify artemisinin resistance markers has thus far mainly followed the same candidate gene approach that was used to identify antifolate resistance markers in the last century, focusing on genes known to play a role in resistance to other drugs or hypothesized to be involved in purported mechanisms of drug action. These approaches have not yet provided an understanding of the mode of action of artemisinins or artemisinin resistance, and no candidate gene has so far been associated with delayed parasite clearance (Imwong et al. 2010). A comprehensive genome-wide search for the molecular basis of resistance is therefore not only warranted but urgent.

Even with the malaria genome sequence in hand and rapidly improving next-generation sequencing platforms making it possible to sequence large numbers of parasite isolates, several challenges remain to using genomics to identify artemisinin resistance markers. First, a clearly defined and reproducible genetically inherited phenotype is needed. The reference genome along with SNP discovery studies has allowed for the creation of genome-wide diversity maps (Mu et al. 2007; Volkman et al. 2007), which have been used to create SNP arrays. These SNP arrays allow for the rapid and cost-effective genotyping of hundreds to thousands of parasites. The first GWAS in *P. falciparum* used an array with about 3,000 SNPs. This study identified genetic loci associated with in vitro susceptibility to artemisinins measured by culture-adapting field isolates and testing their ability to survive in the presence of different concentrations of artemisinins (Mu et al. 2010). However, the

relevance, if any, of these loci to clinical resistance remains unknown because the in vitro phenotype correlates poorly with delayed clearance of parasites following treatment with ACTs, the main clinical manifestation of the recently documented in vivo resistance.

Genetically identical parasite strains identified in multiple patients in western Cambodia share similarly fast or slow parasite clearance rates, indicating that these different phenotypes may be linked to unique heritable determinants and that clearance rate is a suitable phenotype for GWAS (Anderson et al. 2010). In addition to using standard regression methods for GWAS, powerful machine learning methods such as random forests (Cummings and Segal 2004) offer several advantages for analyzing large genetic datasets, including high accuracy and the ability to assess interactions. Complementary to GWAS, measures of signatures of selection locate changes in the genome such as regions of reduced heterozygosity and extended length haplotypes which could have resulted from recent selection, and are useful in that the methods are entirely phenotype independent, avoiding the need for costly clinical or in vitro investigations to measure phenotypes (Pickrell et al. 2009; Voight et al. 2006; Sabeti et al. 2007; Sabeti et al. 2002).

Several other impediments remain for genomic studies aimed at identifying genetic determinants of clinically important phenotypes such as drug resistance. Because resistant parasites can only be found in remote rural sites far from established cold chains, finding, collecting, preserving, and transporting samples from the field to genotyping and sequencing centers are fraught with challenges. Current genotyping and next-generation sequencing platforms require relatively high amounts of parasite DNA with little human DNA contamination. This requires either the costly and cumbersome filtration of leukocytes from blood samples obtained from sick patients at remote field sites or new, untested, and relatively expensive methods for separating human and parasite DNA post-extraction. Moreover, malaria control efforts targeting areas where resistance has emerged have reduced the incidence of malaria in some of these areas to the extent that it is increasingly difficult to enroll patients in the clinical trials that are needed to measure clinical resistance phenotypes and collect parasite samples. Finally, many residents of these areas are transient migrant workers or military personnel who may be reluctant or unable to participate in study activities, including follow-up to measure parasite clearance rates. Close collaboration between genomic and clinical scientists and public health officials, and multi-national, multi-site studies, are therefore required to design and conduct genomic epidemiology studies aimed at identifying artemisinin resistance markers.

Recently, the World Health Organization coordinated a set of such studies at four sites in Cambodia, Thailand, and Bangladesh, and a larger study at as many as 15 sites, including two sentinel sites in Africa, is now getting underway. Whole-genome sequencing of field isolates without culture-adaptation and cloning has been increasingly successful despite limitations of low parasite DNA and contaminating human DNA. At the Wellcome Trust Sanger Institute, more than 1,000 such field-collected samples, including hundreds from clinical trials of artemisinin efficacy, have been subjected to next-generation genomic sequencing. Although the quality and coverage

is not sufficient for complete de novo assembly, it is presently possible to genotype more than 80,000 SNPs in field samples with a high degree of confidence. The SNP calls resulting from these high-throughput genome sequencing efforts will be used for GWAS and studies of signatures of selection in hopes of identifying artemisinin resistance markers that can be used for surveillance and to guide containment efforts.

Global Collaboration

The sequencing of the malaria, mosquito, and human genomes was a monumental scientific accomplishment. Realizing the full potential public health benefit of this accomplishment will require collaboration among northern sequencing centers and clinical investigators and public health officials in developing countries. Malaria does not recognize national borders, and indeed drug resistance seems always to arise in international border regions. Efficient application of genomic science to vaccine design and development and especially to the identification of drug resistance markers that can be used to help contain emerging resistance requires not only collaboration among scientists of different stripes but close transnational collaboration, as well as international coordination.

In the case of artemisinin resistance, the coordinating role has been taken up by the World Health Organization's Global Malaria Programme. New programs like the Worldwide Antimalarial Resistance Network (WWARN) (Sibley et al. 2008; Sibley et al. 2007) and the Malaria Genomic Epidemiology Network (MalariaGEN) (Malaria Genomic Epidemiology Network 2008) are also working to bridge scientific and public health disciplines across international borders, creating internet-based tools that integrate clinical and genetic data and that are designed to be useful to both genomic and clinical scientists as well as, in the case of WWARN, public health officials. These networks and programs all have different mandates but share the goals of malaria control and eventual eradication, and continued collaboration is essential to the effective translation of advances in genomic science to benefiting public health in the developing world.

References

Anonymous (2002) Malaria after the genomes. Lancet 360:1107

Alonso PL, Sacarlal J, Aponte JJ, Leach A, Macete E et al (2004) Efficacy of the RTS, S/AS02A vaccine against *Plasmodium falciparum* infection and disease in young African children: randomised controlled trial. Lancet 364:1411–1420

Anderson TJ, Nair S, Nkhoma S, Williams JT, Imwong M et al (2010) High heritability of malaria parasite clearance rate indicates a genetic basis for artemisinin resistance in western Cambodia. J Infect Dis 201:1326–1330

Baruch DI, Pasloske BL, Singh HB, Bi X, Ma XC et al (1995) Cloning the *P. falciparum* gene encoding PfEMP1, a malarial variant antigen and adherence receptor on the surface of parasitized human erythrocytes. Cell 82:77–87

Bozdech Z, Zhu J, Joachimiak MP, Cohen FE, Pulliam B et al (2003a) Expression profiling of the schizont and trophozoite stages of *Plasmodium falciparum* with a long-oligonucleotide microarray. Genome Biol 4:R9

Bozdech Z, Llinas M, Pulliam BL, Wong ED, Zhu J et al (2003b) The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. PLoS Biol 1:E5. doi: 10.1371/journal.pbio.0000005 [doi]

Bzik DJ, Li WB, Horii T, Inselburg J (1987) Molecular cloning and sequence analysis of the *Plasmodium falciparum* dihydrofolate reductase-thymidylate synthase gene. Proc Natl Acad Sci USA 84:8360–8364

Carlton JM, Angiuoli SV, Suh BB, Kooij TW, Pertea M et al (2002) Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii yoelii*. Nature 419:512–519

Carlton JM, Adams JH, Silva JC, Bidwell SL, Lorenzi H et al (2008) Comparative genomics of the neglected human malaria parasite *Plasmodium vivax*. Nature 455:757–763

Conway DJ, Polley SD (2002) Measuring immune selection. Parasitology 125(Suppl):S3–S16

Cox-Singh J, Davis TM, Lee KS, Shamsul SS, Matusop A et al (2008) *Plasmodium knowlesi* malaria in humans is widely distributed and potentially life threatening. Clin Infect Dis 46:165–171

Crompton PD, Kayala MA, Traore B, Kayentao K, Ongoiba A et al (2010) A prospective analysis of the Ab response to *Plasmodium falciparum* before and after a malaria season by protein microarray. Proc Natl Acad Sci USA 107:6958–6963, 1001323107 [pii];10.1073/pnas.1001323107 [doi]

Cummings MP, Segal MR (2004) Few amino acid positions in rpoB are associated with most of the rifampin resistance in *Mycobacterium tuberculosis*. BMC Bioinformatics 5:137

De Azevedo WF Jr, Soares MB (2009) Selection of targets for drug development against protozoan parasites. Curr Drug Targets 10:193–201

Djimde A, Doumbo OK, Cortese JF, Kayentao K, Doumbo S et al (2001a) A molecular marker for chloroquine-resistant falciparum malaria. N Engl J Med 344:257–263

Djimde A, Doumbo OK, Steketee RW, Plowe CV (2001b) Application of a molecular marker for surveillance of chloroquine-resistant falciparum malaria. Lancet 358:890–891

Djimde AA, Dolo A, Ouattara A, Diakite S, Plowe CV et al (2004) Molecular diagnosis of resistance to antimalarial drugs during epidemics and in war zones. J Infect Dis 190:853–855

Dondorp AM, Nosten F, Yi P, Das D, Phyo AP et al (2009) Artemisinin resistance in *Plasmodium falciparum* malaria. N Engl J Med 361:455–467

Dondorp AM, Yeung S, White L, Nguon C, Day NP et al (2010) Artemisinin resistance: current status and scenarios for containment. Nat Rev Microbiol 8:272–280

Dutta S, Lee SY, Batchelor AH, Lanar DE (2007) Structural basis of antigenic escape of a malaria vaccine candidate. Proc Natl Acad Sci USA 104:12488–12493

Fidock DA, Nomura T, Talley AK, Cooper RA, Dzekunov SM et al (2000) Mutations in the *P. falciparum* digestive vacuole transmembrane protein PfCRT and evidence for their role in chloroquine resistance. Mol Cell 6:861–871

Galinski MR, Medina CC, Ingravallo P, Barnwell JW (1992) A reticulocyte-binding protein complex of *Plasmodium vivax* merozoites. Cell 69:1213–1226

Gardner MJ, Hall N, Fung E, White O, Berriman M et al (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature 419:498–511

Hall N, Karras M, Raine JD, Carlton JM, Kooij TW et al (2005) A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic, and proteomic analyses. Science 307:82–86

Hien TT, White NJ (1993) Qinghaosu. Lancet 341:603–608

Hoffman SL, Billingsley PF, James E, Richman A, Loyevsky M et al (2010) Development of a metabolically active, non-replicating sporozoite vaccine to prevent *Plasmodium falciparum* malaria. Hum Vaccin 6:97–106, 10396 [pii]

Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R et al (2002) The genome sequence of the malaria mosquito anopheles gambiae. Science 298:129–149

Hviid L (2005) Naturally acquired immunity to *Plasmodium falciparum* malaria in Africa. Acta Trop 95:270–275, S0001-706X(05)00165-8 [pii];10.1016/j.actatropica.2005.06.012 [doi]

Imwong M, Dondorp AM, Nosten F, Yi P, Mungthin M et al (2010) Exploring the contribution of candidate genes to artemisinin resistance in *Plasmodium falciparum*. Antimicrob Agents Chemother 54:2886–2892, AAC.00032-10 [pii];10.1128/AAC.00032-10 [doi]

Luke TC, Hoffman SL (2003) Rationale and plans for developing a non-replicating, metabolically active, radiation-attenuated *Plasmodium falciparum* sporozoite vaccine. J Exp Biol 206:3803–3808

Malaria Genomic Epidemiology Network (2008) A global network for investigating the genomic epidemiology of malaria. Nature 456:732–737, nature07632 [pii];10.1038/nature07632 [doi]

Miller LH, Mason SJ, Clyde DF, McGinniss MH (1976) The resistance factor to plasmodium vivax in blacks. The Duffy-blood-group genotype, FyFy. N Engl J Med 295:302–304

Morrison DA (2009) Evolution of the apicomplexa: where are we now? Trends Parasitol 25:375–382

Mu J, Awadalla P, Duan J, McGee KM, Keebler J et al (2007) Genome-wide variation and identification of vaccine targets in the *Plasmodium falciparum* genome. Nat Genet 39:126–130

Mu J, Myers RA, Jiang H, Liu S, Ricklefs S et al (2010) *Plasmodium falciparum* genome-wide scans for positive selection, recombination hot spots and resistance to antimalarial drugs. Nat Genet 42:268–271, ng.528 [pii];10.1038/ng.528 [doi]

Noedl H, Se Y, Schaecher K, Smith BL, Socheat D et al (2008) Evidence of artemisinin-resistant malaria in western Cambodia. N Engl J Med 359:2619–2620

Ogutu BR, Apollo OJ, McKinney D, Okoth W, Siangla J et al (2009) Blood stage malaria vaccine eliciting high antigen-specific antibody concentrations confers no protection to young children in Western Kenya. PLoS One 4:e4708

Ouattara A, Mu J, Takala-Harrison S, Saye R, Sagara I et al (2010) Lack of allele-specific efficacy of a bivalent AMA1 malaria vaccine. Malar J 9:175, 1475-2875-9-175 [pii];10.1186/1475-2875-9-175 [doi]

Pain A, Hertz-Fowler C (2009) Plasmodium genomics: latest milestone. Nat Rev Microbiol 7:180–181

Pain A, Bohme U, Berry AE, Mungall K, Finn RD et al (2008) The genome of the simian and human malaria parasite *Plasmodium knowlesi*. Nature 455:799–803, nature07306 [pii];10.1038/nature07306 [doi]

Peterson DS, Walliker D, Wellems TE (1988) Evidence that a point mutation in dihydrofolate reductase-thymidylate synthase confers resistance to pyrimethamine in falciparum malaria. Proc Natl Acad Sci USA 85:9114–9118

Peterson DS, Milhous WK, Wellems TE (1990) Molecular basis of differential resistance to cycloguanil and pyrimethamine in *Plasmodium falciparum* malaria. Proc Natl Acad Sci USA 87:3018–3022

Pickrell JK, Coop G, Novembre J, Kudaravalli S, Li JZ et al (2009) Signals of recent positive selection in a worldwide sample of human populations. Genome Res 19:826–837

Plowe CV (2009) The evolution of drug-resistant malaria. Trans R Soc Trop Med Hyg 103(Suppl 1):S11–S14

Plowe CV, Djimde A, Bouare M, Doumbo O, Wellems TE (1995) Pyrimethamine and proguanil resistance-conferring mutations in *Plasmodium falciparum* dihydrofolate reductase: polymerase chain reaction methods for surveillance in Africa. Am J Trop Med Hyg 52:565–568

Plowe CV, Roper C, Barnwell JW, Happi CT, Joshi HH et al (2007a) World Antimalarial Resistance Network (WARN) III: molecular markers for drug resistant malaria. Malar J 6:121

Plowe CV, Roper C, Barnwell JW, Happi CT, Joshi HH et al (2007b) World Antimalarial Resistance Network (WARN) III: molecular markers for drug resistant malaria. Malar J 6:121

Plowe CV, Alonso P, Hoffman SL (2009) The potential role of vaccines in the elimination of falciparum malaria and the eventual eradication of malaria. J Infect Dis 200:1646–1649. doi: 10.1086/646613 [doi]

Sabeti PC, Reich DE, Higgins JM, Levine HZ, Richter DJ et al (2002) Detecting recent positive selection in the human genome from haplotype structure. Nature 419:832–837

Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E et al (2007) Genome-wide detection and characterization of positive selection in human populations. Nature 449:913–918
Sagara I, Dicko A, Ellis RD, Fay MP, Diawara SI et al (2009) A randomized controlled phase 2 trial of the blood stage AMA1-C1/Alhydrogel malaria vaccine in children in Mali. Vaccine 27:3090–3098
Sibley CH, Barnes KI, Plowe CV (2007) The rationale and plan for creating a World Antimalarial Resistance Network (WARN). Malar J 6:118
Sibley CH, Barnes KI, Watkins WM, Plowe CV (2008) A network to monitor antimalarial drug resistance: a plan for moving forward. Trends Parasitol 24:43–48
Smith JD, Chitnis CE, Craig AG, Roberts DJ, Hudson-Taylor DE et al (1995) Switches in expression of *Plasmodium falciparum* var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. Cell 82:101–110
Su XZ, Heatwole VM, Wertheimer SP, Guinet F, Herrfeldt JA et al (1995) The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of *Plasmodium falciparum*-infected erythrocytes. Cell 82:89–100
Su X, Kirkman LA, Fujioka H, Wellems TE (1997) Complex polymorphisms in an approximately 330 kDa protein are linked to chloroquine-resistant *P. falciparum* in Southeast Asia and Africa. Cell 91:593–603
Takala SL, Plowe CV (2009) Genetic diversity and malaria vaccine design, testing and efficacy: preventing and overcoming 'vaccine resistant malaria'. Parasite Immunol 31:560–573, PIM1138 [pii];10.1111/j.1365-3024.2009.01138.x [doi]
Takala SL, Coulibaly D, Thera MA, Dicko A, Smith DL et al (2007) Dynamics of polymorphism in a malaria vaccine antigen at a vaccine-testing site in Mali. PLoS Med 4:e93
Takala SL, Coulibaly D, Thera MA, Batchelor AH, Cummings MP et al (2009) Extreme polymorphism in a vaccine antigen and risk of clinical malaria: implications for vaccine development. Sci Transl Med 1:2ra5
Tanner M, de Savigny D (2008) Malaria eradication back on the table. Bull World Health Organ 86:82, S0042-96862008000200002 [pii]
Thera MA, Doumbo OK, Coulibaly D, Laurens MB, Kone AK et al (2010) Safety and immunogenicity of an AMA1 malaria vaccine in Malian children: results of a phase 1 randomized controlled trial. PLoS One 5:e9041. doi:10.1371/journal.pone.0009041 [doi]
Trape JF, Pison G, Preziosi MP, Enel C, Dulou AD et al (1998) Impact of chloroquine resistance on malaria mortality. Comptes Rendus de l'Academie des Sciences de Paris/Life Sciences 321:689–697
Venter JC, Adams MD, Myers EW, Li PW, Mural RJ et al (2001) The sequence of the human genome. Science 291:1304–1351
Voight BF, Kudaravalli S, Wen X, Pritchard JK (2006) A map of recent positive selection in the human genome. PLoS Biol 4:e72
Volkman SK, Sabeti PC, DeCaprio D, Neafsey DE, Schaffner SF et al (2007) A genome-wide map of diversity in *Plasmodium falciparum*. Nat Genet 39:113–119
Waller RF, McFadden GI (2005) The apicoplast: a review of the derived plastid of apicomplexan parasites. Curr Issues Mol Biol 7:57–79
Walliker D, Quakyi IA, Wellems TE, McCutchan TF, Szarfman A et al (1987) Genetic analysis of the human malaria parasite *Plasmodium falciparum*. Science 236:1661–1666
Wellems TE, Panton LJ, Gluzman IY, do Rosario VE, Gwadz RW et al (1990) Chloroquine resistance not linked to mdr-like genes in a *Plasmodium falciparum* cross. Nature 345:253–255
Wellems TE, Walker-Jonah A, Panton LJ (1991) Genetic mapping of the chloroquine-resistance locus on *Plasmodium falciparum* chromosome 7. Proc Natl Acad Sci USA 88:3382–3386
White NJ, Olliaro PL (1996) Strategies for the prevention of antimalarial drug resistance: rationale for combination chemotherapy for malaria. Parasitol Today 12:399–401
World Health Organization (2011) Global plan for artemisinin resistance containment (GPARC)
Yeh I, Altman RB (2006) Drug targets for *Plasmodium falciparum*: a post-genomic review/survey. Mini Rev Med Chem 6:177–202