Large-scale Genotyping and Genetic Mapping in *Plasmodium* Parasites

Xin-zhuan Su[1,*], Hongying Jiang[1], Ming Yi[2], Jianbing Mu[1] and Robert M. Stephens[2]

[1]*Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, USA;* [2]*Advanced Technology Program, SAIC-Frederick, Inc., NCI-Frederick, Frederick, Maryland 21702, USA*

Abstract: The completion of many malaria parasite genomes provides great opportunities for genomewide characterization of gene expression and high-throughput genotyping. Substantial progress in malaria genomics and genotyping has been made recently, particularly the development of various microarray platforms for large-scale characterization of the *Plasmodium falciparum* genome. Microarray has been used for gene expression analysis, detection of single nucleotide polymorphism (SNP) and copy number variation (CNV), characterization of chromatin modifications, and other applications. Here we discuss some recent advances in genetic mapping and genomic studies of malaria parasites, focusing on the use of high-throughput arrays for the detection of SNP and CNV in the *P. falciparum* genome. Strategies for genetic mapping of malaria traits are also discussed.

Key words: *Plasmodium*, malaria, phenotype, genotype, microsatellite, SNP

INTRODUCTION

Human malaria is a deadly disease that still kills more than a million people each year, mostly in tropical and subtropical regions [1,2]. The human disease is caused by infection with 1 of the 5 *Plasmodium* parasites, *Plasmodium falciparum, Plasmodium vivax, Plasmodium ovale, Plasmodium malariae,* and *Plasmodium knowlesi*. *P. falciparum* is the most deadly parasite, whereas *P. vivax* is the most wide-spread species. Although *P. vivax* is generally regarded as non-lethal, it can cause severe diseases in patients [3,4]. There are also malaria parasites that can infect many species of mammals, birds, and reptiles [5]. These animal malaria parasites can be excellent models for studying disease mechanisms and molecular pathways in the parasites.

The malaria parasite has a complex life cycle involving multiple rounds of sexual and asexual reproduction in its hosts. Sexual reproduction and genetic recombination of malaria parasites occur in mosquito vectors. The sexual stages are essential for transmission and genetic recombination that is critical for mapping studies. Most of the life cycle of a malaria parasite has a haploid genome. The genomes of the majority of malaria parasites are similar in size and contain 14 chromosomes with ~23-26 million nucleotides [6] and approximately 5,500 genes [7,8].

The nucleotide content (the percentage of G/A/T/C) can be very different: the *P. falciparum* genome has > 80% AT, whereas *P. vivax* has a much lower AT content (~63%) [7,9,10]. Another obvious difference is the species-specific gene families that encode variant antigens [6]. Unfortunately, the majority (60%) of the predicted genes in *P. falciparum* encode hypothetical proteins [7]. In this brief summary of malaria genetics and genomics, we shall focus on data from the *P. falciparum* parasite, although some information from other malaria parasites will also be briefly mentioned.

POLYMORPHISMS AND GENETIC MARKERS

A genetic mapping project generally consists of 3 parts: accurate measurement of phenotype, determination of genotype, and linkage or association analyses. Before a genetic mapping study can be performed, genetic markers for genotyping parasite DNA have to be developed. In the 1970s, polymorphisms in enzymes such as 6-replace with phosphogluconate hydrogenase and glucose-6-phosphate isomerase in starch gel electrophoresis were used to distinguish rodent malaria parasites and progeny from genetic crosses [11-13]. Restriction fragment-length polymorphisms (RFLP) became popular genetic markers for mapping malaria traits in the 1980s [14]. In the 1990s, microsatellites (MS) were introduced to fine map the chloroquine (CQ)-resistant gene [15, 16]. Now, in the 2000s, single-nucleotide polymorphism (SNP)

is becoming the marker of choice because of the development of high-throughput SNP genotyping methods [17,18]. SNP genotyping is actually based on the same polymorphism as RFLP (nucleotide substitution); the difference is in the method for detecting the polymorphism. Large numbers of MS and SNP have been developed for *P. falciparum* [19-22]; and genetic markers for other parasites such as *P. vivax* and rodent malaria parasites are also available or being developed [23-27].

MS markers have advantages for population studies because they are highly polymorphic and generally more selectively neutral than many SNP [28]. Compared with SNP, the high mutation rates and multiple alleles of MS markers provide better resolving power for studying closely related parasite populations or for fine mapping loci with candidate genes. MS markers are therefore useful for parasite identification that generally does not require typing large numbers of markers. Multiple alleles of a MS marker make it valuable for searching signatures of drug selection within and between parasite populations [29-31]. The distribution and frequency of MS among different malaria parasite genomes are very different. MS are abundant in *P. falciparum* [7,15], mostly likely due to its high-AT-content genome, but relatively few are present in *P. vivax* [24] and *P. yoelii* [26].

SNP are also quite abundant in both *P. falciparum* and *P. vivax* genomes, although the distribution of SNP can vary greatly among various chromosomal regions or different genes [20-22,24,32,33]. In an earlier study of 204 genes from chromosome 3 of *P. falciparum*, approximately 1 SNP was found from about 900 bp DNA sequences among 5 *P. falciparum* isolates (π = 4.9 $\times$ 10^{-4}) [32]. Recent large-scale sequencing suggested that the frequency of SNP in *P. falciparum* could be much higher [20,22]. High frequencies of SNP were found in genes encoding surface proteins [33] and putative transporters [34] that are likely under pressure from host immunity or antimalarial agents; but for housekeeping genes, the frequency of SNP can be much lower [35,36]. The uneven distribution of polymorphisms is one of the issues leading to the debate on the age of *P. falciparum* [32, 37,38]. In *P. vivax*, the frequency of SNP is also quite high, with approximately 1 SNP per 530 bp (π = 8.7 $\times$ 10^{-4}) [24]. The frequency of nucleotide substitution, in addition to the abundance of polymorphic MS, makes it feasible to develop a high-resolution genetic map for malaria parasites, particularly *P. falciparum*.

PARASITE GENOTYPING METHODS

Over the years, various genetic typing methods have been developed. Polymorphisms in antigen-coding genes such as apical membrane antigen-1 (AMA-1), merozoite surface protein-1 (MSP-1), and circumsporozoite protein (CSP) have been widely used to type DNA samples collected from patients. More recently, MS markers are used because they are more likely to be neutral markers [15,39,40]. For typing and identification of parasite clones adapted to in vitro culture, genetic typing methods using multicopy genes can be very helpful [41-43]. Use of a selected number of MS markers (for example, 10 MS markers) or a set of SNP can be a very efficient and effective method for parasite 'fingerprinting' or parasite verification, as well [44-46].

MS products after PCR amplification can be labeled with radioactive or fluorescent materials and detected after separation in a DNA sequencing gel and an automatic DNA sequencer [44]. Because DNA sequencers are widely available, labeling PCR products with fluorescent dye is more convenient than labeling with isotopes; however, fluorescent-labeled primers can be expensive if large numbers of primers are required for a study; and a method using a single labeled universal primer for all the MS markers in a study can greatly reduce the primer cost [26]. Agarose gel can sometimes be used to resolve MS markers that have large size differences (4 bp or longer).

For high-throughput genotyping, it is clear that SNP is the marker of choice, because various high-throughput SNP typing methods are now available [47-52]. In *P. falciparum*, high-density arrays have been used to study genetic recombination in a progeny clone from the Dd2 $\times$ HB3 cross and to identify single-feature polymorphisms (SFPs) among field isolates [36,53]. Using an Affymetrix array with 298,782 25-mer *P. falciparum* probes, Kidgell et al. [36] identified 23,653 SFP and various copy-number variations (CNV) from the isolates. Similarly, genome-wide amplifications and deletions were also detected using oligonucleotides printed in glass slides [54]. More recently, another Affymetrix array containing ~2.5 million probes (PFSANGER array; designed at the Sanger Center, UK) became commercially available and has been used to study RNA transcripts [55].

In our laboratory, we have explored the possibility of using the Sanger tiling array for typing the *P. falciparum* genome. Biotin-labeled genomic DNA (10 μg) from a parasite was hybridized to the PFSANGER array at 45°C for 16 hr, and signals from field isolates were compared with those from 3D7. If there is a difference (substitution) in a probe sequence between 3D7 and another parasite, signal from the parasite will be reduced compared with that of 3D7. Because hybridization signals can be influenced by various factors such as the position of a base sub-

stitution in a probe and the GC content of a probe, criteria used to distinguish a real substitution in a probe from background signals require extensive evaluation and verification of hybridization signals from different parasites. Our data showed that the Sanger array could be successfully used to detect polymorphism within a probe (SFP) with an accuracy rate higher than 90% after comparison of more than 3,000 known SNP [21] with those detected using the array [53]. We also showed that the last 2 end positions in a probe did not significantly affect hybridization signals and that probes with GC contents < 16% should be excluded from SFP calling because of poor hybridization signals. SFP calls based on a single probe were not reliable either; and a conservative signal cutoff ratio of 3-5.0 and a vote among several adjacent probes (within 25 bp) with a majority of the probes calling for an SFP (or mSFP for SFP called by multiple probes with 75% of the probes calling for polymorphism) should be applied. To illustrate how well the array performed in calling nucleotide polymorphism, we applied the SFP calling parameters to identify 5 known substitutions within a gene encoding a homolog of the human P-glycoprotein (*pfmdr-1*) [21,56]. We were able to identify 3 of the 5 known mutations (NIAID SNP ID: PFE1150w-1, PFE1150w-3, and PFE1150w-5) in the *pfmdr-1* (Fig. 1A). The other 2 mutations (PFE1150w-2 and PFE1150w-4) were not detected because of absence of probes covering the substitution sites (Fig. 1B, C). Additionally, this array was a useful tool for detecting CNV in the *P. falciparum* genome. Genomic regions that were highly polymorphic, deleted, or amplified could be detected after comparison of hybridization signals from par-

asite isolates (Fig. 2). In particular, comparison of the chromosome 5 sequence from parasite 106/1 with that of 3D7 showed amplification of a ~100-kb segment containing *pfmdr-1* in the 106/1 parasite (Fig. 2A). In addition, potential amplifications on chromosome 1, 2, 4, 8, and 13 and highly polymorphic *var* clusters could also be detected (Fig. 2B). CNV has received more attention lately because it has been shown to be associated with parasite response to antimalarial drugs [57-60]. The results from this study showed that the Sanger array could be used to call both SFP and CNV in the *P. falciparum* genome.

A *P. falciparum*-specific SNP typing array using the molecular inversion probe (MIP) method [61] has also been developed for typing parasite SNP in our lab (J. Mu et al., unpublished). Because genomic DNA is amplified using a pair of primers specific for an SNP to be assayed, the MIP method can detect SNP using relatively small amount of DNA (~100 ng). This method therefore can potentially be used for genotyping DNA samples collected directly from patients. If a DNA sample is amplified using commercial kits such as REPLI-g® Whole Genome Amplification (QIAGEN) that can increase the amount of DNA in a sample 300-400 times, even smaller amounts of DNA will be required. Before genome-wide amplification, however, human DNA has to be removed before DNA purification, because the human genome is ~100 times that of a malaria parasite genome. Many other genotyping arrays such as Affymetrix 3 K and 75 K arrays (S. Volkman, D. Wirth et al., Harvard University), NimbleGen 60-mer oligo array (M. Ferdig, Notre Dame University), and another Affymetrix tiling with ~5 million probes (E. Winzeler,
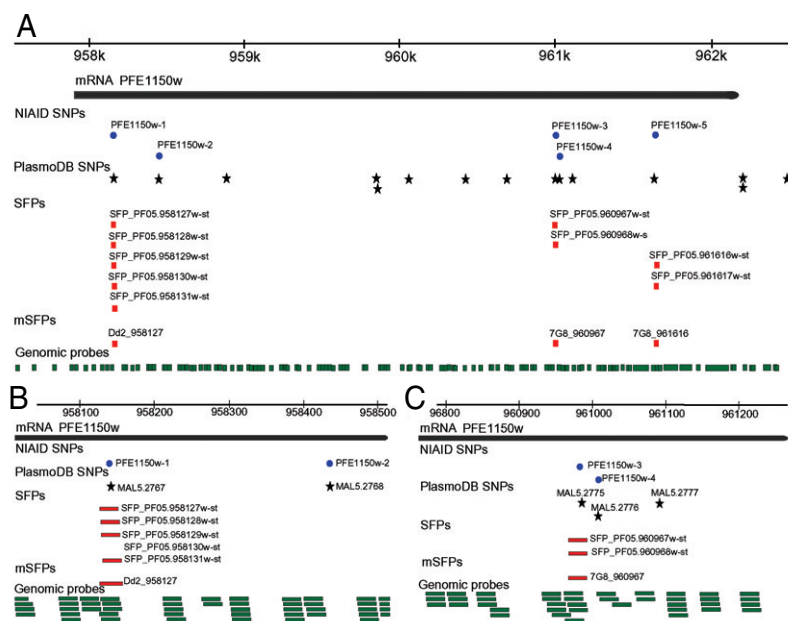

Fig. 1. Genome browser displays (drawn in Canvas) showing SFP, mSFP and SNP, and probes at the *pfmdr-1* locus. (A) A genome browser window (~5 kb) showing chromosomal region covering *pfmdr-1* gene (top line) and the predicted *pfmdr-1* coding region (black bar); blue circle, known SNPs reported previously [21,56], black stars, SNPs in PlasmoDB; red squares, SFPs detected by individual probe and mSFP by a majority of the probes; green bars, genomic probe covering the *pfmdr-1* gene. Names of the SNPs and probes are as labeled. Note the mSFP matched well with those known SNPs. (B and C) An expanded view (500-bp window) showing 2 regions in the *pfmdr-1*, including 2 known substitutions (PFE1150w-1 and PFE1150w-4) that were not covered by any of the genomic probes. The labels are the same as in (A). For more information on genome-wide SNPs and SFPs, please visit: http://p-falcip.abcc.ncifcrf.gov/cgi-bin/gbrowse/malaria_sanger/.
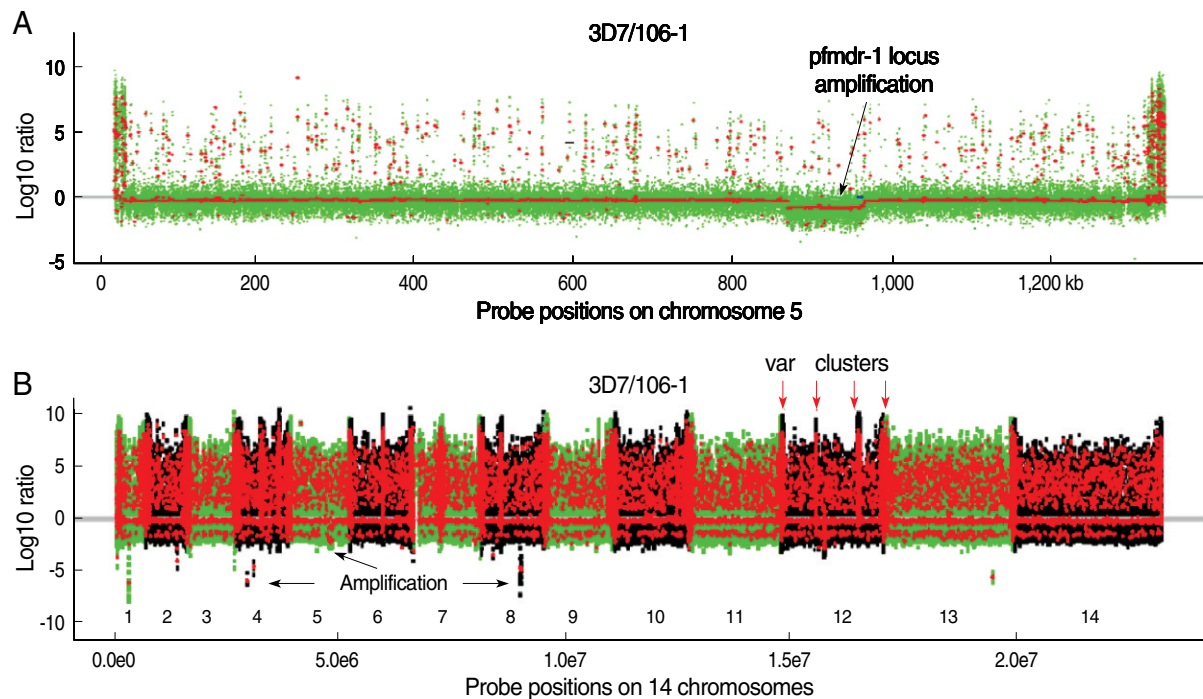
Fig. 2. Chromosome segmentation analysis showing genomic amplification and deletion/highly polymorphic regions on chromosomes of *Plasmodium falciparum*. Signals from DNA hybridizations to the Sanger tiling microarray were normalized as described [53]. Chromosomal segments (seen as dots) from circular binary segmentation (CBS) analysis [91] of the data for 106/1 strain using 3D7 strain as reference were plotted. (A) Probe signal ratios (3D7 over 106/1) showing an amplification of a locus containing *pfmdr-1* on chromosome 5 in the 106/1 parasite. Green dots are normalized log ratios of individual probe, and the red lines/dots are the mean values of normalized log ratios of all the probes in segments obtained after CBS analysis. Dots above the grey/red line at zero are polymorphic individual probes (green) or segments with multi-probes with similar reduced signals in 106/1 (red); and those under the line are potentially amplified regions (red). (B) CBS analysis applied to all chromosomes (1-14) of the 106/1 parasite using 3D7 as reference. The green dots are normalized log ratios of individual probes; and the red lines/dots are the mean values of normalized log ratios from all the probes in segments obtained after CBS analysis. The dots with positive values suggested polymorphisms between 106/1 and 3D7; and those with negative values were potential genomic amplification. Highly polymorphic chromosome ends and internal *var* gene clusters can be seen as intense red 'zones'. Red arrows point to 4 known *var* clusters on chromosome 12; and black arrows indicate some potential amplifications on chromosome 4, 5, and 8. The numbers 1-14 marks the parasite 14 chromosomes.

Scripps) are being developed and will be available soon. These arrays and methods will greatly facilitate our ability to detect large numbers of genome polymorphisms in *P. falciparum* parasites.

MAPPING APPROACHES

Genetic mapping in malaria parasites can be generally classified into 3 different approaches. The first is linkage mapping using genetic crosses. The first genetic cross in malaria parasites was performed by Dr. David Walliker and his colleagues at the University of Edinburgh, UK [62]. Following the initial cross, many more genetic crosses have been performed in the human malaria parasite *P. falciparum* and in rodent malaria parasites [12-14,63-65]. One successful story of mapping malaria traits using progeny from a genetic cross was the identification of *pfcrt*, the gene that plays a key role in CQ resistance in *P. falciparum* [14,16,66,67]. The advantage of mapping using progeny from genetic crosses is that the genetic backgrounds in the progeny derive from the 2 parents, reducing genomic background noises seen in field isolates; however, performing a genetic cross can be expensive and laborious, not to mention the ethical concerns of using nonhuman primates for crosses of human malaria parasites. In this regard, use of progeny from crosses of rodent malaria parasites can be good alternatives. Although phenotypes in rodent malaria do not always exactly reflect those seen in human malaria parasites, what we learn from rodent malaria parasites will provide important information for studying human malaria parasites. Importantly, disease phenotypes such as virulence can be dissected more easily using rodent malaria parasites,

because host genetic background variation can be controlled. The second approach is to investigate the association of candidate genes such as *pfmdr-1*, *pfdhfr*, and *pfcrt* with drug resistance in field isolates [34,68], but candidate genes have to be available or identified before a study can be conducted. The third approach is genome-wide association, which is becoming popular with the development of various high throughput genotyping methods [18,31]. The majority of genetic studies in malaria parasites have been performed using rodent malaria parasites (*Plasmodium chabaudi*, *Plasmodium yoelii*, *Plasmodium berghei*) and *P. falciparum* and have focused on mapping genes contributing to drug resistance [14,18,69-75]. Differences in parasite growth and development, virulence, transmission, immunogenicity, variation in gene expression, and a 'mutator' phenotype are some of the traits that have also been studied [76-83].

Genetic recombination (or crosses) occurs in the field when parasites are ingested by a mosquito; however, a recombination event will be detected only if 2 or more parasites with different genotypes are present. It is quite common for a mosquito to carry parasites with different genotypes, particularly in regions with relatively high transmission rates, such as Africa and Asia [39,84,85]. Genetic recombination occurring in parasite populations can be explored for mapping genes contributing to drug resistance and other parasite traits. Under drug pressure, genetic changes will occur in some parasites, leading to parasites that can survive at a high drug concentration. Some wild type parasites may acquire the changes through genetic recombination and survive. This process will select for parasites carrying drug-resistant mutations, leading to a drug 'selective sweep' that can be detected through genotyping. Signature of drug selection has been reported for CQ and sulfadoxine-pyrimethamine resistance [29,30,86-89].

We are interested in mapping genes that may affect parasite responses to various antimalarial drugs including quinine, mefloquine, and others. While developing high-throughput methods to genotype *P. falciparum* parasites collected from different regions of the world, we have also been testing the parasite responses to different drugs. Hundreds of parasites have been collected, adapted to in vitro culture, and tested. We insisted on testing parasite drug responses after adapting parasites from patients into in vitro culture and genotyping them to make sure they were clonal, because mixed infections and many other factors might influence drug test results [90]. Accurate measurement of drug response is particularly important when a phenotype being studied is controlled by multiple genes. DNA sam-

ples from the parasites are being typed using the MIP and the Sanger tiling microarrays. We hope to associate some polymorphisms in the parasite genome to elevated parasite responses (increased IC_{50}) to various antimalarial drugs. After identification of candidate genes, molecular and biochemical approaches will be employed to evaluate the functions of the candidate genes.

CONCLUSION

High-throughput genotyping methods are now available for typing DNA from *P. falciparum* and for mapping parasite traits; and many more typing methods are under development, including those for other malaria species. Unfortunately, the malaria parasite is a single-cell organism, and it is thus challenging to detect or measure reproducible phenotypic variation between individual parasites. The future direction for mapping malaria traits should focus on developing methods to accurately characterize and measure phenotypic variation among individual parasites.

ACKNOWLEDGEMENTS

REFERENCES

1. Snow RW, Guerra CA, Noor AM, Myint HY, Hay SI. The global distribution of clinical episodes of *Plasmodium falciparum* malaria. Nature 2005; 434: 214-217.

2. Cox J, Hay SI, Abeku TA, Checchi F, Snow RW. The uncertain burden of *Plasmodium falciparum* epidemics in Africa. Trends Parasitol 2007; 23: 142-148.

3. Mendis K, Sina BJ, Marchesini P, Carter R. The neglected burden of *Plasmodium vivax* malaria. Am J Trop Med Hyg 2001; 64 (1-2 Suppl): 97-106.

4. Genton B, D'Acremont V, Rare L, Baea K, Reeder JC, Alpers MP, Müller I. *Plasmodium vivax* and mixed infections are associated with severe malaria in children: a prospective cohort study from Papua New Guinea. PLoS Med 2008; 5: e127.

5. Perkins SL, Schall JJ. A molecular phylogeny of malarial parasites recovered from cytochrome b gene sequences. J Parasitol 2002; 88: 972-978.

6. Carlton J, Silva J, Hall N. The genome of model malaria parasites, and comparative genomics. Curr Issues Mol Biol 2005; 7: 23-37.

7. Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan MS, Nene V, Shallom SJ, Suh B, Peterson J, Angiuoli S, Pertea M, Allen J, Selengut J, Haft D, Mather MW, Vaidya AB, Martin DM, Fairlamb AH, Fraunholz MJ, Roos DS, Ralph SA, McFadden GI, Cummings LM, Subramanian GM, Mungall C, Venter JC, Carucci DJ, Hoffman SL, Newbold C, Davis RW, Fraser CM, Barrell B. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature 2002; 419: 498-511.

8. Hall N, Carlton J. Comparative genomics of malaria parasites. Curr Opin Genet Dev 2005; 15: 609-613.

9. Pollack Y, Katzen AL, Spira DT, Golenser J. The genome of *Plasmodium falciparum*. I. DNA base composition. Nucleic Acids Res 1982; 10: 539-546.

10. Carlton JM, Adams JH, Silva JC, Bidwell SL, Lorenzi H, Caler E, Crabtree J, Angiuoli SV, Merino EF, Amedeo P, Cheng Q, Coulson RM, Crabb BS, Del Portillo HA, Essien K, Feldblyum TV, Fernandez-Becerra C, Gilson PR, Gueye AH, Guo X, Kang'a S, Kooij TW, Korsinczky M, Meyer EV, Nene V, Paulsen I, White O, Ralph SA, Ren Q, Sargeant TJ, Salzberg SL, Stoeckert CJ, Sullivan SA, Yamamoto MM, Hoffman SL, Wortman JR, Gardner MJ, Galinski MR, Barnwell JW, Fraser-Liggett CM. Comparative genomics of the neglected human malaria parasite *Plasmodium vivax*. Nature 2008; 455: 757-763.

11. Carter R, Voller A. Enzyme variants in primate malaria parasites. Trans R Soc Trop Med Hyg 1973; 67: 14-15.

12. Walliker D, Sanderson A, Yoeli M, Hargreaves BJ. A genetic investigation of virulence in a rodent malaria parasite. Parasitology 1976; 72: 183-194.

13. Walliker D, Quakyi IA, Wellems TE, McCutchan TF, Szarfman A, London WT, Corcoran LM, Burkot TR, Carter R. Genetic analysis of the human malaria parasite *Plasmodium falciparum*. Science 1987; 236: 1661-1666.

14. Wellems TE, Panton LJ, Gluzman IY, do Rosario VE, Gwadz RW, Walker-Jonah A, Krogstad DJ. Chloroquine resistance not linked to mdr-like genes in a *Plasmodium falciparum* cross. Nature 1990; 345: 253-255.

15. Su XZ, Wellems TE. Toward a high-resolution *Plasmodium falciparum* linkage map: polymorphic markers from hundreds of simple sequence repeats. Genomics 1996; 33: 430-444.

16. Su XZ, Kirkman LA, Fujioka H, Wellems TE. Complex polymorphisms in an approximately 330 kDa protein are linked to chloroquine-resistant *P. falciparum* in Southeast Asia and Africa. Cell 1997; 91: 593-603.

17. Maresso K, Broeckel U. Genotyping platforms for mass-throughput genotyping with SNPs, including human genome-wide scans. Adv Genet 2008; 60: 107-139.

18. Su XZ, Hayton K, Wellems TE. Genetic linkage and association analyses for trait mapping in *Plasmodium falciparum*. Nat Rev Genet 2007; 8: 497-506.

19. Su XZ, Ferdig MT, Huang Y, Huynh CQ, Liu A, You J, Wootton JC, Wellems TE. A genetic map and recombination parameters of the human malaria parasite *Plasmodium falciparum*. Science 1999; 286: 1351-1353.

20. Jeffares DC, Pain A, Berry A, Cox AV, Stalker J, Ingle CE, Thomas A, Quail MA, Siebenthall K, Uhlemann AC, Kyes S, Krishna S, Newbold C, Dermitzakis ET, Berriman M. Genome variation and evolution of the malaria parasite *Plasmodium falciparum*. Nat Genet 2007; 39: 120-125.

21. Mu J, Awadalla P, Duan J, McGee KM, Keebler J, Seydel K, McVean GA, Su XZ. Genome-wide variation and identification of vaccine targets in the *Plasmodium falciparum* genome. Nat Genet 2007; 39: 126-130.

22. Volkman SK, Sabeti PC, DeCaprio D, Neafsey DE, Schaffner SF, Milner DA Jr, Daily JP, Sarr O, Ndiaye D, Ndir O, Mboup S, Duraisingh MT, Lukens A, Derr A, Stange-Thomann N, Waggoner S, Onofrio R, Ziaugra L, Mauceli E, Gnerre S, Jaffe DB, Zainoun J, Wiegand RC, Birren BW, Hartl DL, Galagan JE, Lander ES, Wirth DF. A genome-wide map of diversity in *Plasmodium falciparum*. Nat Genet 2007; 39: 113-119.

23. Grech K, Martinelli A, Pathirana S, Walliker D, Hunt P, Carter R. Numerous, robust genetic markers for *Plasmodium chabaudi* by the method of amplified fragment length polymorphism. Mol Biochem Parasitol 2002; 123: 95-104.

24. Feng X, Carlton JM, Joy DA, Mu J, Furuya T, Suh BB, Wang Y, Barnwell JW, Su XZ. Single-nucleotide polymorphisms and genome diversity in *Plasmodium vivax*. Proc Natl Acad Sci USA 2003; 100: 8502-8507.

25. Martinelli A, Hunt P, Cheesman SJ, Carter R. Amplified fragment length polymorphism measures proportions of malaria parasites carrying specific alleles in complex genetic mixtures. Mol Biochem Parasitol 2004; 136: 117-122.

26. Li J, Zhang Y, Sullivan M, Hong L, Huang L, Lu F, McCutchan TF, Su XZ. Typing *Plasmodium yoelii* microsatellites using a simple and affordable fluorescent labeling method. Mol Biochem Parasitol

2007; 155: 94-102.

27. Karunaweera ND, Ferreira MU, Munasinghe A, Barnwell JW, Collins WE, King CL, Kawamoto F, Hartl DL, Wirth DF. Extensive microsatellite diversity in the human malaria parasite *Plasmodium vivax*. Gene 2008; 410: 105-112.

28. Anderson TJ, Su XZ, Roddam A, Day KP. Complex mutations in a high proportion of microsatellite loci from the protozoan parasite *Plasmodium falciparum*. Mol Ecol 2000; 9: 1599-1608.

29. Wootton JC, Feng X, Ferdig MT, Cooper RA, Mu J, Baruch DI, Magill AJ, Su XZ. Genetic diversity and chloroquine selective sweeps in *Plasmodium falciparum*. Nature 2002; 418: 320-323.

30. Roper C, Pearce R, Nair S, Sharp B, Nosten F, Anderson T. Intercontinental spread of pyrimethamine-resistant malaria. Science 2004; 305: 1124.

31. Anderson TJ. Mapping drug resistance genes in *Plasmodium falciparum* by genome-wide association. Curr Drug Targets Infect Disord 2004; 4: 65-78.

32. Mu J, Duan J, Makova K, Joy DA, Huynh CQ, Branch OH, Li WH, Su XZ. Chromosome-wide SNPs reveal an ancient origin for *Plasmodium falciparum*. Nature 2002; 418: 323-326.

33. Volkman SK, Hartl DL, Wirth DF, Nielsen KM, Choi M, Batalov S, Zhou Y, Plouffe D, Le Roch KG, Abagyan R, Winzeler EA. Excess polymorphisms in genes for membrane proteins in *Plasmodium falciparum*. Science 2002; 298: 216-218.

34. Mu J, Ferdig MT, Feng X, Joy DA, Duan J, Furuya T, Subramanian G, Aravind L, Cooper RA, Wootton JC, Xiong M, Su XZ. Multiple transporters associated with malaria parasite responses to chloroquine and quinine. Mol Microbiol 2003; 49: 977-989.

35. Volkman SK, Barry AE, Lyons EJ, Nielsen KM, Thomas SM, Choi M, Thakore SS, Day KP, Wirth DF, Hartl DL. Recent origin of *Plasmodium falciparum* from a single progenitor. Science 2001; 293: 482-484.

36. Kidgell C, Volkman SK, Daily J, Borevitz JO, Plouffe D, Zhou Y, Johnson JR, Le Roch K, Sarr O, Ndir O, Mboup S, Batalov S, Wirth DF, Winzeler EA. A systematic map of genetic variation in *Plasmodium falciparum*. PLoS Pathog 2006; 2: e57.

37. Rich SM, Licht MC, Hudson RR, Ayala FJ. Malaria's Eve: evidence of a recent population bottleneck throughout the world populations of *Plasmodium falciparum*. Proc Natl Acad Sci USA 1998; 95: 4425-4430.

38. Hughes AL, Verra F. Very large long-term effective population size in the virulent human malaria parasite *Plasmodium falciparum*. Proc R Soc Lond B Biol Sci 2001; 268: 1855-1860.

39. Anderson TJ, Haubold B, Williams JT, Estrada-Franco JG, Richardson L, Mollinedo R, Bockarie M, Mokili J, Mharakurwa S, French N, Whitworth J, Velez ID, Brockman AH, Nosten F, Ferreira MU, Day KP. Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. Mol Biol Evol 2000; 17: 1467-1482.

40. Greenhouse B, Myrick A, Dokomajilar C, Woo JM, Carlson EJ, Rosenthal PJ, Dorsey G. Validation of microsatellite markers for use in genotyping polyclonal *Plasmodium falciparum* infections. Am J Trop Med Hyg 2006; 75: 836-842.

41. Dolan SA, Herrfeldt JA, Wellems TE. Restriction polymorphisms and fingerprint patterns from an interspersed repetitive element of *Plasmodium falciparum* DNA. Mol Biochem Parasitol 1993; 61: 137-142.

42. Su XZ, Wellems TE. *Plasmodium falciparum*: a rapid DNA fingerprinting method using microsatellite sequences within var clusters. Exp Parasitol 1997; 86: 235-236.

43. Su XZ, Carucci DJ, Wellems TE. *Plasmodium falciparum*: parasite typing by using a multicopy microsatellite marker, PfRRM. Exp Parasitol 1998; 89: 262-265.

44. Ferdig MT, Su XZ. Microsatellite markers and genetic mapping in *Plasmodium falciparum*. Parasitol Today 2000; 16: 307-312.

45. Crameri A, Marfurt J, Mugittu K, Maire N, Regös A, Coppee JY, Sismeiro O, Burki R, Huber E, Laubscher D, Puijalon O, Genton B, Felger I, Beck HP. Rapid microarray-based method for monitoring of all currently known single-nucleotide polymorphisms associated with parasite resistance to antimalaria drugs. J Clin Microbiol 2007; 45: 3685-3691.

46. Daniels R, Volkman SK, Milner DA, Mahesh N, Neafsey DE, Park DJ, Rosen D, Angelino E, Sabeti PC, Wirth DF, Wiegand RC. A general SNP-based molecular barcode for *Plasmodium falciparum* identification and tracking. Malar J 2008; 7: 223.

47. Lindblad-Toh K, Winchester E, Daly MJ, Wang DG, Hirschhorn JN, Laviolette JP, Ardlie K, Reich DE, Robinson E, Sklar P, Shah N, Thomas D, Fan JB, Gingeras T, Warrington J, Patil N, Hudson TJ, Lander ES. Large-scale discovery and genotyping of single-nucleotide polymorphisms in the mouse. Nat Genet 2000; 24: 381-386.

48. Kennedy GC, Matsuzaki H, Dong S, Liu WM, Huang J, Liu G, Su X, Cao M, Chen W, Zhang J, Liu W, Yang G, Di X, Ryder T, He Z, Surti U, Phillips MS, Boyce-Jacino MT, Fodor SP, Jones KW. Large-scale genotyping of complex DNA. Nat Biotechnol 2003; 21: 1233-1237.

49. Gunderson KL, Steemers FJ, Lee G, Mendoza LG, Chee MS. A genome-wide scalable SNP genotyping assay using microarray technology. Nat Genet 2005; 37: 549-554.

50. Hardenbol P, Yu F, Belmont J, Mackenzie J, Bruckner C, Brundage T, Boudreau A, Chow S, Eberle J, Erbilgin A, Falkowski M, Fitzgerald R, Ghose S, Iartchouk O, Jain M, Karlin-Neumann G, Lu X, Miao X, Moore B, Moorhead M, Namsaraev E, Pasternak S, Prakash E, Tran K, Wang Z, Jones HB, Davis RW, Willis TD, Gibbs RA. Highly multiplexed molecular inversion probe genotyping: over 10,000 targeted SNPs genotyped in a single tube assay. Genome Res 2005; 15: 269-275.

51. Steemers FJ, Gunderson KL. Whole genome genotyping technologies on the BeadArray platform. Biotechnol J 2007; 2: 41-49.

52. Hagiwara H, Sawakami-Kobayashi K, Yamamoto M, Iwasaki S, Sugiura M, Abe H, Kunihiro-Ohashi S, Takase K, Yamane N, Kato K, Son R, Nakamura M, Segawa O, Yoshida M, Yohda M, Tajima H, Kobori M, Takahama Y, Itakura M, Machida M. Development of an automated SNP analysis method using a paramagnetic beads handling robot. Biotechnol Bioeng 2007; 98: 420-428.

53. Jiang H, Yi M, Mu J, Zhang L, Ivens A, Klimczak LJ, Huyen Y, Stephens RM, Su XZ. Detection of genome wide polymorphisms in

the AT rich *Plasmodium falciparum* genome using a high density microarray. BMC Genomics 2008; 9: 398.

54. Ribacke U, Mok BW, Wirta V, Normark J, Lundeberg J, Kironde F, Egwang TG, Nilsson P, Wahlgren M. Genome wide gene amplifications and deletions in *Plasmodium falciparum*. Mol Biochem Parasitol 2007; 155: 33-44.

55. Mourier T, Carret C, Kyes S, Christodoulou Z, Gardner PP, Jeffares DC, Pinches R, Barrell B, Berriman M, Griffiths-Jones S, Ivens A, Newbold C, Pain A. Genome-wide discovery and verification of novel structured RNAs in *Plasmodium falciparum*. Genome Res 2008; 18: 281-292.

56. Foote SJ, Kyle DE, Martin RK, Oduola AM, Forsyth K, Kemp DJ, Cowman AF. Several alleles of the multidrug-resistance gene are closely linked to chloroquine resistance in *Plasmodium falciparum*. Nature 1990; 345: 255-258.

57. Cowman AF, Galatis D, Thompson JK. Selection for mefloquine resistance in *Plasmodium falciparum* is linked to amplification of the pfmdr1 gene and cross-resistance to halofantrine and quinine. Proc Natl Acad Sci USA 1994; 9: 1143-1147.

58. Price RN, Uhlemann AC, Brockman A, McGready R, Ashley E, Phaipun L, Patel R, Laing K, Looareesuwan S, White NJ, Nosten F, Krishna S. Mefloquine resistance in *Plasmodium falciparum* and increased pfmdr1 gene copy number. Lancet 2004; 364: 438-447.

59. Nair S, Nash D, Sudimack D, Jaidee A, Barends M, Uhlemann AC, Krishna S, Nosten F, Anderson TJ. Recurrent gene amplification and soft selective sweeps during evolution of multidrug resistance in malaria parasites. Mol Biol Evol 2007; 24: 562-573.

60. Jiang H, Patel JJ, Yi M, Mu J, Ding J, Stephens R, Cooper RA, Ferdig MT, Su XZ. Genome-wide compensatory changes accompany drug-selected mutations in the *Plasmodium falciparum* crt gene. PLoS ONE 2008; 3: e2484.

61. Hardenbol P, Baner J, Jain M, Nilsson M, Namsaraev EA, Karlin-Neumann GA, Fakhrai-Rad H, Ronaghi M, Willis TD, Landegren U, Davis RW. Multiplexed genotyping with sequence-tagged molecular inversion probes. Nat Biotechnol 2003; 21: 673-678.

62. Walliker D, Carter R, Morgan S. Genetic recombination in malaria parasites. Nature 1971; 232: 561-562.

63. Walliker D, Carter R, Morgan S. Genetic recombination in *Plasmodium berghei*. Parasitology 1973; 66: 309-320.

64. Walliker D, Carter R, Sanderson A. Genetic studies on *Plasmodium chabaudi*: recombination between enzyme markers. Parasitology 1975; 70: 19-24.

65. Hayton K, Gaur D, Liu A, Takahashi J, Henschen B, Singh S, Lambert L, Furuya T, Bouttenot R, Doll M, Nawaz F, Mu J, Jiang L, Miller LH, Wellems TE. Erythrocyte binding protein PfRH5 polymorphisms determine species-specific pathways of *Plasmodium falciparum* invasion. Cell Host Microbe 2008; 4: 40-51.

66. Wellems TE, Walker-Jonah A, Panton LJ. Genetic mapping of the chloroquine-resistance locus on *Plasmodium falciparum* chromosome 7. Proc Natl Acad Sci USA 1991; 88: 3382-3386.

67. Fidock DA, Nomura T, Talley AK, Cooper RA, Dzekunov SM, Ferdig MT, Ursos LM, Sidhu AB, Naudé B, Deitsch KW, Su XZ, Wootton JC, Roepe PD, Wellems TE. Mutations in the *P. falci-*

parum digestive vacuole transmembrane protein PfCRT and evidence for their role in chloroquine resistance. Mol Cell 2000; 6: 861-871.

68. Hayton K, Su XZ. Drug resistance and genetic mapping in *Plasmodium falciparum*. Curr Genet 2008; 54: 223-239.

69. Wang P, Read M, Sims PF, Hyde JE. Sulfadoxine resistance in the human malaria parasite *Plasmodium falciparum* is determined by mutations in dihydropteroate synthetase and an additional factor associated with folate utilization. Mol Microbiol 1997; 23: 979-986.

70. Su XZ, Wootton JC. Genetic mapping in the human malaria parasite *Plasmodium falciparum*. Mol Microbiol 2004; 53: 1573-1582.

71. Ferdig MT, Cooper RA, Mu J, Deng B, Joy DA, Su XZ, Wellems TE. Dissecting the loci of low-level quinine resistance in malaria parasites. Mol Microbiol 2004; 52: 985-997.

72. Hunt P, Cravo PV, Donleavy P, Carlton JM, Walliker D. Chloroquine resistance in *Plasmodium chabaudi*: are chloroquine-resistance transporter (crt) and multi-drug resistance (mdr1) orthologues involved? Mol Biochem Parasitol 2004; 133: 27-35.

73. Carlton J, Mackinnon M, Walliker D. A chloroquine resistance locus in the rodent malaria parasite *Plasmodium chabaudi*. Mol Biochem Parasitol 1998; 93: 57-72.

74. Cravo PV, Carlton JM, Hunt P, Bisoni L, Padua RA, Walliker D. Genetics of mefloquine resistance in the rodent malaria parasite *Plasmodium chabaudi*. Antimicrob Agents Chemother 2003; 47: 709-718.

75. Culleton R, Martinelli A, Hunt P, Carter R. Linkage group selection: rapid gene discovery in malaria parasites. Genome Res 2005; 15: 92-97.

76. Walliker D. Genetic factors in malaria parasites and their effect on host-parasite relationships. In Taylor AE, Muller R eds, Genetic Aspects of Host-Parasite Relationships. Oxford, UK. Blackwell. 1976, p 25-44.

77. Martinelli A, Cheesman S, Hunt P, Culleton R, Raza A, Mackinnon M, Carter R. A genetic approach to the de novo identification of targets of strain-specific immunity in malaria parasites. Proc Natl Acad Sci USA 2005; 102: 814-819.

78. Vaidya AB, Muratova O, Guinet F, Keister D, Wellems TE, Kaslow DC. A genetic locus on *Plasmodium falciparum* chromosome 12 linked to a defect in mosquito-infectivity and male gametogenesis. Mol Biochem Parasitol 1995; 69: 65-71.

79. Furuya T, Mu J, Hayton K, Liu A, Duan J, Nkrumah L, Joy DA, Fidock DA, Fujioka H, Vaidya AB, Wellems TE, Su XZ. Disruption of a *Plasmodium falciparum* gene linked to male sexual development causes early arrest in gametocytogenesis. Proc Natl Acad Sci USA 2005; 102: 16813-16818.

80. Klotz FW, Chulay JD, Daniel W, Miller LH. Invasion of mouse erythrocytes by the human malaria parasite, *Plasmodium falciparum*. J Exp Med 1987; 165: 1713-1718.

81. Mackinnon MJ, Gaffney DJ, Read AF. Virulence in rodent malaria: host genotype by parasite genotype interactions. Infect Genet Evol 2002; 1: 287-296.

82. Rathod PK, McErlean T, Lee PC. Variations in frequencies of drug

resistance in *Plasmodium falciparum*. Proc Natl Acad Sci USA 1997; 94: 9389-9393.

83. Gonzales JM, Patel JJ, Ponmee N, Jiang L, Tan A, Maher SP, Wuchty S, Rathod PK, Ferdig MT. Regulatory hotspots in the malaria parasite genome dictate transcriptional variation. PLoS Biol 2008; 6: e238.

84. Lee SA, Yeka A, Nsobya SL, Dokomajilar C, Rosenthal PJ, Talisuna A, Dorsey G. Complexity of *Plasmodium falciparum* infections and antimalarial drug efficacy at 7 sites in Uganda. J Infect Dis 2006; 193: 1160-1163.

85. Kobbe R, Neuhoff R, Marks F, Adjei S, Langefeld I, von Reden C, Adjei O, Meyer CG, May J. Seasonal variation and high multiplicity of first *Plasmodium falciparum* infections in children from a holoendemic area in Ghana, West Africa. Trop Med Int Health 2006; 11: 613-619.

86. Nair S, Williams JT, Brockman A, Paiphun L, Mayxay M, Newton PN, Guthmann JP, Smithuis FM, Hien TT, White NJ, Nosten F, Anderson TJ. A selective sweep driven by pyrimethamine treatment in southeast asian malaria parasites. Mol Biol Evol 2003; 20: 1526-1536.

87. Roper C, Pearce R, Bredenkamp B, Gumede J, Drakeley C, Mosha F, Chandramohan D, Sharp B. Antifolate antimalarial resistance in southeast Africa: a population-based analysis. Lancet 2003; 361: 1174-1181.

88. McCollum AM, Poe AC, Hamel M, Huber C, Zhou Z, Shi YP, Ouma P, Vulule J, Bloland P, Slutsker L, Barnwell JW, Udhayakumar V, Escalante AA. Antifolate resistance in *Plasmodium falciparum*: multiple origins and identification of novel dhfr alleles. J Infect Dis 2006; 194: 189-197.

89. McCollum AM, Mueller K, Villegas L, Udhayakumar V, Escalante AA. Common origin and fixation of *Plasmodium falciparum* dhfr and dhps mutations associated with sulfadoxine-pyrimethamine resistance in a low-transmission area in South America. Antimicrob Agents Chemother 2007; 51: 2085-2091.

90. Liu S, Mu J, Jiang H, Su XZ. Effects of *Plasmodium falciparum* mixed infections on in vitro antimalarial drug tests and genotyping. Am J Trop Med Hyg 2008; 79: 178-184.

91. Olshen AB, Venkatraman ES, Lucito R, Wigler M. Circular binary segmentation for the analysis of array-based DNA copy number data. Biostatistics 2004; 5: 557-572.