Molecular Epidemiology of Malaria

David J. Conway*

Medical Research Council Laboratories, Fajara, P.O. Box 273, Banjul, The Gambia, and Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, United Kingdom

INTRODUCTION............................................................................................................................................................188
Definitions and Priorities in a Postgenomic Era........................................................................................................188
USE OF MOLECULAR METHODS TO DETECT MALARIA PARASITES......................................................................189
Correct Identification of Species.......................................................................................................................................189
Sensitive Detection............................................................................................................................................................190
COMPLEXITY OF INFECTIONS......................................................................................................................................190
Mixed Species....................................................................................................................................................................190
Multiple-Genotype Infections.........................................................................................................................................192
GENETIC STRUCTURES OF PARASITE POPULATIONS.................................................................................................193
Parasites in Populations in Areas of Endemicity.............................................................................................................193
Outbreaks, Epidemics, and Reintroductions of Malaria.................................................................................................194
SELECTION AND SPREAD OF DRUG RESISTANCE ALLELES......................................................................................195
Evolution...........................................................................................................................................................................195
Surveillance and Prevention..............................................................................................................................................195
ADAPTIVE ANTIGENIC POLYMORPHISMS......................................................................................................................197
Natural Immune Selection.................................................................................................................................................197
Vaccine Trials....................................................................................................................................................................197
MOLECULAR MECHANISMS OF VIRULENCE..................................................................................................................198
Severe Malaria..................................................................................................................................................................198
Placental Malaria..............................................................................................................................................................199
CURRENT CHALLENGES...............................................................................................................................................199
ACKNOWLEDGMENTS.......................................................................................................................................................199
REFERENCES.....................................................................................................................................................................199

INTRODUCTION

Malaria parasites are at the forefront of genomic approaches to biology. There is an unparalleled wealth of information on the comparative genomics of Plasmodium, as it is the most intensively sequenced eukaryotic genus, with genomes of seven different species complete or virtually so (29, 82, 87) (data at www.sanger.ac.uk/projects/protozoa and www.tigr.org/parasiteProjects.shtml). Each genome consists of ~23 megabases (Mb) distributed over 14 chromosomes, with separate mitochondrial and plastid genomes of ~6 and 35 kilobases, respectively. Parasites are haploid for most of their life cycle, and targeted gene manipulation approaches are being effectively used in several of the species that have been sequenced, including rodent and monkey malaria parasites that can be studied in models of infection (104, 192) as well as the most important parasite of humans, Plasmodium falciparum (46). The analysis of heritable variation that exists within P. falciparum has also rapidly advanced, with efforts for genomewide polymorphism discovery by sequencing and oligonucleotide hybridization mismatch strategies (30, 90, 130, 194) and performance of chromosome-wide as well as mitochondrial genome-based population genetic analyses (99, 131). Substantial description of nuclear and mitochondrial genomic variation has followed for the second most important malaria parasite, Plasmodium vivax (77, 94, 98, 132), which is more common than P. falciparum in many parts of South and Central America and Southeast Asia. The past several years have thereby yielded great advances in understanding these parasites.

Such new genomic information allows molecular epidemiology of malaria to be strongly built. A majority of previous studies and overviews were conducted when there was significantly less information or fewer technical resources (20, 86). Although many earlier studies gave understanding of lasting value, there is a current need to develop molecular epidemiology with genomic data and resources so that greater advances can be made in understanding and controlling this disease.

Definitions and Priorities in a Postgenomic Era

Molecular epidemiology is here defined as the study of pathogen genotypes and gene expression as it relates to the occurrence of infection and disease in human populations. In one sense, this is just an extension of normal epidemiological analysis to incorporate any molecular information on pathogens detected within individuals. However, molecular information is often so rich that it requires specialized analytical schemes, in the form of phylogenetic trees (99, 132), population genetic parameters (136), or gene expression profiles (49). Such schemes tend to determine the structure of the analysis that can be most usefully performed and generally yield more information than analyses that merely consider the presence or
absence of particular genes or genotypes. The goal is to see beyond a description of variability to detect the effects of molecular and cellular processes in infection, pathogenesis, immunity, or therapeutic responses.

Human and mosquito vector genetics are separate and well-defined fields that influence the epidemiology of malaria (110, 139) but fall outside the pathogen-oriented focus of molecular epidemiology here. A broader view would see them as part of the same system, with many potential points of ecological and evolutionary interaction. For example, interactive variation in human major histocompatibility complex and T-cell receptor genes can influence the type of selection operating among allelic forms of parasite antigens (113), while hemoglobin variants can have effects on display of variant surface antigens of the parasite on the erythrocyte membrane (71). Vector mosquitoes such as Anopheles gambiae in Africa show genetic variation in the ability to transmit malaria parasites (139) and have complex population structures that contain incipient species (181), features likely to affect the population biology of the parasite and to which it may respond adaptively. Essentially, many heritable variations in human and mosquito populations exist alongside other environmental variation that affects the molecular epidemiology of the parasite.

Some molecular processes lend themselves to mathematical modeling at a population level, particularly in genetics and immunology. Researchers who have modeled conceptual processes are increasingly keen to analyze the actual molecular variables that underlie these, and this helps to identify major knowledge gaps that can target research. Processes that have recently received modeling attention include the relationship between immunity and the evolution of virulence (81), the orchestration of antigenic variation by immune responses (160), and the emergence and spread of drug resistance (196).

The genome sequence of P. falciparum reveals the primary structure and arrangement of more than 5,000 protein coding genes (82), and the sequences of other Plasmodium species give comparable information (29, 87). This can be exploited by comparative genomic analyses that allow inferences to be drawn on molecular evolution (29, 87, 106) and microarray analyses to show profiles of transcription at all stages in the life cycle (including liver stages that are accessible in rodent malaria models but not in human malaria) (87). Data on gene evolution can lead to hypotheses about gene function, whereas studies of genome-wide patterns of gene transcription allow the potential role of otherwise nondescript genes to be highlighted (24, 87, 115). Either type of information is likely to ultimately relate to mechanistic hypotheses of clinical or epidemiological relevance.

USE OF MOLECULAR METHODS TO DETECT MALARIA PARASITES

Correct Identification of Species

Microscopic examination is the primary method of malaria parasite detection and species identification, although problems with this have been recognized for some time, as lamented by Knowles and Das Gupta when they grappled with the first malaria parasite samples from Asian monkeys, working in Calcutta in 1932.

“Now there is nothing easier—or in our opinion, more fallacious—than to make a general study of a blood film containing malaria parasites, and then to sit down and write a general account of the morphology of the parasite forms attached. In no branch of medicine is the well-worn tag of Hippocrates more appropriate—‘Experience is fallacious and judgement difficult’” (103).

Such problems have not diminished in recent years. Even the most skillful morphological analysis of stained parasites on blood films is not a very reliable basis for determining the identity of a malaria parasite species. Difficulties are compounded when infections contain more than one parasite species or when an unusual species is present.

Identification of malaria parasites in peripheral blood samples can now be most reliably performed by analysis of DNA, and this, to some extent, transforms the possibilities for diagnosis and epidemiology of malaria. Plasmodium ovale has recently been shown to consist of two major clades that are as divergent as different species (199). This is based not only on a single locus sequence analysis that could give an unrepresentative result due to unusual gene lineage sorting or natural selection but on three different genes (the asexual stage expressed small subunit 18S rRNA, a nuclear gene encoding a cysteine protease, the mitochondrial cytb gene). The two major forms are indistinguishable morphologically but should be provisionally considered two different biological species, both of which exist in Africa and Southeast Asia. It would be relevant to study the epidemiology of each separately using molecular methods, although this is not yet a high priority, as the relative incidence of P. ovale overall is much lower than P. falciparum in Africa and P. vivax in Southeast Asia (37).

Parasite morphology has proved unsuitable for a systematic analysis of the relationships among the different species. Other parasitological features, such as data on the course of experimental infections (including periodicity of replication in the blood), also have little reliability for systematic purposes. For example, although most malaria parasites of macaques are more closely related to P. vivax than to other human malaria parasite species, Plasmodium inui was long considered to be more closely related to the human parasite Plasmodium malariae because of its “quartan” (72 h) periodicity. However, sequencing and phylogenetic analysis of the small subunit (SSU) rRNA genes from several P. inui isolates now clearly shows that they are all more related to other macaque malaria parasites than to P. malariae (102). Although it has yet to be confirmed by analysis of other genes, this result is likely to be robust, as the SSU rRNA sequence normally tracks the phylogeny of biological species accurately. In contrast, sequence analysis (SSU rRNA gene and part of the msp1 antigen gene) of the South American monkey malaria parasite Plasmodium brasilianum shows that it is not only superficially similar to P. malariae but apparently identical, comprising a single species that may be zoonotically transmitted in rainforest areas of the Amazon basin and Guianas (72).

The most vivid example of the epidemiological importance of molecular identification of malaria parasite species has emerged from Southeast Asia. Following a decline in the incidence of P. falciparum and P. vivax in the region, a surprisingly large number of cases of what was considered to be P. malariae remained in Malaysian Borneo. Molecular phyloge-
netic and epidemiological studies have shown these to be mostly due to the macaque malaria parasite species *Plasmodium knowlesi* (175). More than a hundred of these infections were detected in the Kapit District of Sarawak over a 2-year period, during which time no parasites were detected that were identified as *P. malariae* by PCR. The *csp* gene of 8 of these isolates is compared with *P. knowlesi* from monkeys and the other malaria parasite species in a phylogenetic analysis (Fig. 1A). They are all clearly indistinguishable from *P. knowlesi* and show a typical level of nucleotide polymorphism, indicating that they have not recently emerged from a single clonal source (Fig. 1B), a conclusion that is also clearly supported by sequence analysis of the SSU rRNA gene (175). This finding would have been as surprising to Knowles and colleagues who first saw these parasites in macaques 70 years previously as it has been to more recent investigators, since only two cases of natural human infections by this species were previously described in all that time. Today, the prospective screening of samples for *P. knowlesi* and normal human malaria parasite species or other previously unknown malaria parasite infections in humans is enabled by genus-specific and species-specific nested PCR assays such as those applied in this study (174). Shortly after the reporting of human cases of *P. knowlesi* in Borneo, a case was described in Thailand (97). It is likely that the infection is more widespread in humans in Southeast Asia, possibly throughout the long-tailed and pig-tailed macaque host ranges.

**Sensitive Detection**

The use of PCR-based methods to detect malaria parasites in blood samples increases the sensitivity of detection compared with microscopy. Qualitative PCR protocols that are robust, sensitive, and species specific have been available since the 1990s (174, 177), and there are now several quantitative PCR methods that allow estimation of parasitemia levels as well as positivity (9, 121, 167). Thick-film microscopy can allow the examination of ~0.1 to 1 µl of blood (50 to 500 high-power fields with ~0.002 µl per field) and, thus, the detection of more than ~10 parasites µl⁻¹. Most applications of PCR typically involve amplification of DNA template from the equivalent of 1 to 10 µl blood and are thus either slightly more or up to 100 times more sensitive than microscopy. DNA template can be prepared from larger volumes of blood to give even higher sensitivity, with detection of ~20 parasites ml⁻¹ being achieved (9), which is useful in clinical vaccine trials in which the time to first detectable blood-stage parasitemia is the endpoint (22).

Sensitive PCR methods for parasite detection have been used to good effect in epidemiological studies, revealing surprisingly high proportions of individuals who have persistent asymptomatic infections in some populations in areas of endemicity (162). Use of real-time quantitative PCR methods is being evaluated in clinical diagnostic laboratories in countries of endemicity with good resources (184) and reference laboratories in countries with substantial numbers of imported malaria cases (73). However, there is a limitation to the information provided by any method that samples peripheral blood for estimation of *P. falciparum* parasitemia, as sequestered mature asexual parasites may sometimes outnumber those in the peripheral blood (172). Therefore, methods to accurately estimate overall parasite loads need to be further developed. Circulating protein capture assays (such as those based on HRP2) are not yet able to exactly quantitate current infection levels (141), as protein persists in plasma for up to a few weeks after infections are resolved (122), although in some patient populations they may provide better estimates than peripheral blood parasitemia (58) when estimates of HRP2 production and clearance rates are incorporated in a model.

**COMPLEXITY OF INFECTIONS**

It has long been known that many malaria parasite infections contain a mixture of *Plasmodium* species, based on examination of morphological differences in stained blood films (although this has highly varying degrees of accuracy depending on the species and microscopist). By the 1970s and 1980s, different genotypes of common species such as *P. falciparum* could also be discriminated within infections by analyzing polymorphic protein products using electrophoretic (31) and serological (123) methods, although these were only performed in a few research laboratories. Intensive efforts to culture and clone different *P. falciparum* parasites from individual clinical isolates were successful in some cases, so that haploid clones within genetically complex infections could be clearly defined for combinations of several polymorphic characters as well as other genetically determined phenotypes, including antimalarial drug sensitivity in vitro. A classic example from this era was an isolate from a patient in Tak Province in Thailand from which seven different *P. falciparum* clones were derived (187). Once PCR methods became widely available, discrimination of parasite species and genotypes within blood samples could more easily be done.

**Mixed Species**

One of the first questions to be considered using molecular detection methods in epidemiology is whether coinfection with one malaria parasite species (such as *P. vivax* or *P. malariae*) would modify the course of infection or risk of disease due to another (particularly *P. falciparum*). Epidemiological evidence from microscopically detected infections in Vanuatu suggests interaction (possibly cross-immunity) between *P. vivax* and *P. falciparum* that may be clinically important (120), and such interactions would potentially be important in many endemic areas. An early PCR study in Côte d’Ivoire gave preliminary data showing that *P. malariae* infections may reduce the risk of symptomatic *P. falciparum* infection (23), but little has been done to follow this up in larger studies or in other African populations. One analysis of mixed-species infections in Papua New Guinea indicated that there may be density-dependent regulation on all malaria parasites that coexist in the blood, so that a high parasitemia of any parasite clone would reduce the effective replication rate of any other parasite (of the same or a different species) (25). This is plausible, although the presence of each species in mixed infections was determined by nonquantitative PCR, whereas the density was estimated by microscopy only (with which it is difficult to reliably estimate
relative proportions of different species). Another study in Papua New Guinea has reported that the cooccurrence of different *Plasmodium* species is apparently random, as determined by microscopy and PCR with sequence-specific probing (125). The contrasting findings indicate that it may be interesting to conduct further studies of species-specific parasitemia levels in mixed- and single-species infections using quantitative PCR estimation.

![Diagram of the nonrepeat region of the csp gene of *P. knowlesi* parasites from 8 human infections (representative of 106 human *P. knowlesi* infections detected) in Kapit District Hospital (KH), Sarawak, Malaysia. (A) Phylogenetic analysis of the human isolates together with 2 *P. knowlesi* macaque isolates and other malaria parasite species. The large clade at the top contains *P. vivax* and the Asian macaque parasites, with *P. knowlesi* human and macaque sequences clearly distinct from all other species (numbers on the branches show bootstrap support, which is 100% in this case). (B) Sequence differences among the *P. knowlesi* isolates indicates a normal level of within-species polymorphism in the csp gene and strongly indicates that the human isolates have not arisen from a single recent source. Similar results are also obtained by studying the small subunit rRNA gene sequence in the same samples. (Reprinted from reference 175 with permission from Elsevier.)
Multiple-Genotype Infections

In contrast to the modest number of studies that have focused on analysis of mixed-species infections, there have been many studies that have analyzed mixed-genotype infections of *P. falciparum*. The high proportion of mixed-genotype infections in many areas of endemicity, along with the availability of simple protocols for PCR discrimination of different genotypes (75, 178), has encouraged many characterizations of mixed-genotype infections even in situations where there is no apparent idea of why this may be important. Some theoretical ideas and empirical data on possible interactions between different pathogen genotypes have been outlined (158), and the relevance to studies of malaria parasite genotypes in the blood of humans and experimental animals has been reviewed previously (159). Experiments in a murine model of *P. chabaudi* malaria have indicated that evolution of virulence may be driven by competition between coinfected malaria parasite clones (52). It is plausible that human malaria parasites in areas of low endemicity have evolved a lower level of virulence (due to less coinfection), whereas those in areas of high endemicity have higher virulence, although this has not been tested.

However, in controlled comparisons within populations in areas of endemicity, different numbers of genotypes in *P. falciparum* infections are not generally associated with different clinical symptoms or severity of malaria. Groups of severe and mild malaria cases have been compared independently in The Gambia (38), Senegal (161), and Gabon (109), with each study showing that the numbers of genotypes per infection were similar between the clinical groups. The study in Senegal indicated that, although means were similar, severe cases more often had only a single clone, but numbers of cases analyzed were small and they came from disparate areas that vary in endemicity (161). A remarkable postmortem analysis of parasites in multiple organs of Malawian children has indicated that cerebral malaria cases may be less genotypically mixed, supporting an idea that fatality may be caused by rapid replication or pathogenic cytoadhesion of particular virulent clones, although numbers of cases analyzed were necessarily small (129). There were also no differences in the number of genotypes per infection between mild symptomatic and asymptomatic *P. falciparum* infections in community-based studies in Tanzania (19), Kenya (111), or Papua New Guinea (68). Studies in Senegal and Sudan have shown very slightly higher mean numbers of genotypes in symptomatic infections than in asymptomatic infections but did not control for the level of parasitemia that is higher in clinical infections so that more parasites were sampled within them (163, 201). Differences have been noted between two sympatric ethnic groups in Mali, with a slightly higher mean number of genotypes detected in blood samples from the group that tends to suffer more from malaria and has higher average parasite densities, as may be expected (143).

Parasites detected in a single blood sample reveal only a part of the true picture, and a small number of studies have supplemented this by repeated sampling of individuals over a period of time. In areas of high endemicity, there are very frequent changes in the parasites detected in infected individuals that are asymptomatic. These generally have a pattern dominated by reappearance of particular genotypes every 2 days (due to the sequestration of parasites at mature stages of the 48-h developmental cycle), together with gradual changes in the genotypes detected over time (due to gain of genotypes by superinfection or recrudescence from low levels and loss by clearance) and apparently random sporadic appearance of genotypes only seen at one time point (due to parasites that are rapidly cleared or kept at low levels) (26, 74). Such frequent changes in the parasites detectable in a mixed-genotype infection, due to sequestration and fluctuation of parasite population densities, are an important feature that limits the possibility to derive a full genotypic profile of complex infections. A commonly made error is the assumption that a detected genotype must have arisen from a new infection if it was not detected in a blood sample taken at an earlier time (Fig. 2).

Clinical trials of antimalarials conducted in areas of endemicity often incorporate genotyping of highly polymorphic loci of parasites in a blood sample pretreatment (day 0) and on any occasion at which parasites have reappeared within the follow-up period and present the approach as being a robust means of distinguishing recrudescence from new infections, as applied in important recent studies (134, 155). The accuracy of this practice should be reconsidered and its potential value assessed more critically. A detected match between allelic profiles pre- and posttreatment is usually (but not always) correctly indicative of a recrudescence due to failure to clear the parasite completely, as the high level of polymorphism in most populations means that there is a low (although not zero) probability that a new parasite infection has a matched profile (195). However, detection of a parasite genotype in a follow-up sample that was not detected in a day 0 sample cannot be at all reliably ascribed as a new infection, as the genotype may have been present at an undetectable level in the patient on day 0 (at low level in the peripheral blood or entirely sequestered) cooccurring with a more abundant genotype. Thus, consideration of PCR-discordant cases as being inevitably due to new...

![Schematic diagram indicating fluctuation in peripheral blood parasitemia of *P. falciparum* relative to an arbitrary threshold of detectability (10^6 parasites in the body). Three coexisting genotypes (A to C) are considered. All genotypes show typical 2-day periodicity of detectable levels due to sequestration of mature asexual developmental stages that disappear from the peripheral blood. Genotype A declines over time, genotype B remains at a constant level throughout, while genotype C declines initially and then increases over time. The arrows show the genotypes that would be detected in peripheral blood samples taken on days 1, 6, 11, and 16, with all samples giving a different result. An incorrect interpretation would be that genotypes B and C appeared later as "new infections."](http://cmr.asm.org.onlinelibrary.wiley.com)
infections, and removal of them from being considered “treatment failures” in analysis can lead to undue inflation of reported parasitological cure rates. This is acknowledged in some trials (62), and even where this problem is not recognized, the general method can be seen to be problematic due to complexities in determining whether to score a match or mismatch when a mixed-genotype follow-up sample contains both (176).

More information on the genotypic composition of *P. falciparum* infections could be gained from multiple blood samples taken over a 48-h period from individuals prior to treatment, but a delay of treatment would be unethical for clinical cases. Genotyping of samples taken during the 48 h posttreatment gives limited information in addition to that available from a single pretreatment sample, as the parasitemia rapidly diminishes if the treatment is initially effective (76), and most of the mature sequestered parasites would not successfully develop a next generation of young ring-stage parasites that would be detectable in the peripheral blood. Such a reduction of parasitemia after treatment is common even if complete clearance is not achieved and recrudescence eventually occurs during the following weeks (possibly of a genotype that was initially at undetectable levels in the peripheral blood).

**GENETIC STRUCTURES OF PARASITE POPULATIONS**

**Parasites in Populations in Areas of Endemicity**

A consensus of data on nuclear and mitochondrial sequence diversity in populations indicates that modern *P. falciparum* populations are derived from a small original population in Africa that expanded sometime between approximately 10,000 and 100,000 years ago, as discussed fully elsewhere where different attempts at more subtle and precise estimation have been reviewed (90). It appears that *P. falciparum* and *P. vivax* populations at roughly the same time or slightly earlier (77, 94) but from a Southeast Asian source (70, 132). Both species were probably widely distributed in human populations, except for American populations, several thousands of years ago.

The most extensive data on malaria parasite population structure today pertain to *P. falciparum*. In populations in areas of high endemicity, *P. falciparum* has little or no linkage disequilibrium among alleles at loci in different parts of its genome (31, 39). This is not surprising given the frequent occurrence of infections with mixed genotypes that occur in areas of high endemicity, which allow mixing of gametes of different genotypes within the same mosquito blood meal. More remarkable is the fact that in areas of high endemicity strong linkage disequilibrium generally only occurs between very closely linked nucleotide sites (within less than 1 kb) and that it generally decays to very low levels beyond this (42, 149, 150). However, this is also consistent with the measured high meiotic recombination rate (*r*) in *P. falciparum*, with a 1% probability of crossover (1 centimorgan) every 17 kb along each of the chromosomes (183).

Populations with low endemicity have a lower proportion of mixed-genotype infections, due to a low rate of superinfection, and thus, the parasites have a relatively higher rate of self-fertilization between male and female gametes from the same clone taken up in a mosquito blood meal. The degree to which this constraint causes nonrandom mating in the parasite population overall is the inbreeding coefficient (*F*, with values on a scale from 0 to 1.0). This trend was initially illustrated by demonstration of a higher rate of inbreeding (measured from genotype frequencies of oocysts compared with Hardy-Weinberg expectations) in an area of moderate endemicity in Papua New Guinea (145) compared to an area of very high endemicity in Tanzania (13). The estimated inbreeding coefficients were ~0.9 and ~0.4 in the respective populations, but these should be regarded as approximate due to the technical possibility that allelic dropout in the genotyping data may have elevated the values (5). As noted by Dye and Williams (66), the effective recombination rate (*r’*) is reduced in proportion to the inbreeding coefficient, such that *r’* = *r*(1 − *F*), and the above data would therefore predict an effective recombination rate ~6 times lower in the Papua New Guinea population than that in Tanzania. Broadly consistent with such a difference between African and Southeast Asian or Pacific populations, the recombination parameter within the *ama1* gene has been estimated as ~3 times lower (i.e., linkage disequilibrium declines 3 times more slowly with molecular map distance) in a population in Thailand than in one in Nigeria (149, 150). An analysis of linkage disequilibrium patterns throughout parasite chromosome 3 shows a similar geographical difference and also indicates some likely variation in recombination rate between different parts of the chromosome (131).

There is a broad spectrum of *P. falciparum* population structures in different areas, dependent on the levels of transmission and endemicity of infection. Populations that have a high level of inbreeding sometimes show a significant degree of nonrandom association between alleles at loci that are on different chromosomes (loosely termed linkage disequilibrium although not due to physical linkage), as assessed by an analytical approach that combines data from multiple loci into a single “index of association.” The most extensive survey analyzed polymorphism of 12 microsatellite loci in 465 *P. falciparum* infections from nine geographical populations covering a worldwide distribution (4). Among the populations, the strength of multilocus linkage disequilibrium showed a strong negative correlation with the proportion of mixed-genotype infections, and there were major differences between continents that followed the expected pattern based on differences in endemicity. African populations had the highest proportion of mixed-genotype infections and the least linkage disequilibrium, South American populations had the fewest mixed-genotype infections and the strongest linkage disequilibrium, whereas Southeast Asian populations were intermediate for both (4). Populations in areas of low endemicity that show the most linkage disequilibrium and that are not influenced by a recent epidemic expansion are generally termed “clonal” in structure (4, 157) (the term does not imply no recombination but rather rare recombination in such situations, since most transmission is of self-fertilized parasites).

In regions where the transmission of malaria is now very patchy, as is the case in many countries of endemicity in Southeast Asia and South America, there may be extreme differences in population genetic structure between different areas within a country. In the Brazilian Amazon region (the largest area of endemicity in the Americas), geographical foci of *P. falciparum* exist with extremely different population genetic structures, with the amount of linkage disequilibrium in each
local area being inversely correlated with the proportion of mixed-genotype infections (118). Despite the existence of considerable microsatellite polymorphism in all 5 populations sampled in the Brazilian Amazon, those with the lowest endemicity show evidence of a “clonal” structure, and the geographical foci show a high level of genetic differentiation. A comparable study on 8 populations sampled in Malaysian Borneo (also an area of relatively low and focal \( P. falciparum \) endemicity where human populations are separated by extensive rainforest) also showed highly divergent local population patterns, with the most “clonal” structure appearing where mixed-genotype infections were rare (10). However, in contrast to the situation in Brazil, a strong relationship between genetic and geographical distances among the populations in Borneo supports a model of isolation by distance (Fig. 3). This is probably because \( P. falciparum \) malaria existed at a relatively stable level of endemicity in Borneo until it started to decline recently, with ongoing but low levels of gene flow between neighboring areas of endemicity. The Brazilian Amazon has a completely different history dominated by human colonization with sporadic introductions and population expansions of parasites. Analysis of data from a highly polymorphic antigen has also shown patchy and local differences in allele frequencies in parasite populations on islands of the Western Pacific, including Vanuatu and the Solomon Islands, likely due to strong founder effects and sporadic gene flow mediated by movements of infected humans between the islands, combined with random gene frequency drift on the smaller islands (117, 169).

Outbreaks, Epidemics, and Reintroductions of Malaria

There are some areas of the world from which malaria has been eradicated, but they remain at risk. When cases appear, genotyping of parasites can help evaluate whether local transmission has occurred and may also indicate the source of the introduction. Malaria had been eradicated from South Korea in the 1960s, but the numbers of indigenous cases of \( P. falciparum \) started to increase among soldiers near the border of North Korea in the early 1990s. The polymorphic antigen gene \( ama1 \) was sequenced from 30 of the isolates, and it shows very limited polymorphism, with only two alleles (closely related to each other) and virtual identity to sequences from Chinese isolates, suggesting a reintroduction from an adjacent regional source (no isolates were available from North Korea) (36).

Malaria had been virtually eradicated in the Cape Verde islands, hundreds of miles off the coast of West Africa, but an epidemic of \( P. falciparum \) occurred in a village in the main island of Santiago in 1995, resulting in at least 40% of the inhabitants becoming infected. Analysis of three polymorphic antigen genes showed that all of the cases were of an identical genotype, indicating a monoclonal epidemic parasite originating from a single source (11). It is likely that an individual imported the infection from an adjacent endemic country in West Africa from which there is frequent travel, such as Guinea-Bissau or Senegal, though there were no data to support this directly. The extensive genetic diversity of \( P. falciparum \) throughout that region prevents an exact identification of the source.

In Australia, an outbreak of \( P. vivax \) malaria affected 10 people who stayed at a campsite in Northern Queensland in 2002, caused by local transmission seeded by gametocytes from one of the campers who had recently been in Indonesia. Genotypic analysis showed that the index case was infected with at least two different \( P. vivax \) clones and that the secondary cases had a mixture of different reassorted genotypes (88).

In Europe, a particular outbreak risk is associated with “airport malaria,” where cases are acquired from infected mosquitoes that arrive on intercontinental flights. On average, there are a few such infections each year among people who live near airports but have never traveled to an area of malaria endemicity. Among these, occasionally more than one infection occurs in a given place at one time, and analyses in France have confirmed that a pair of cases in 1994 and a cluster of four cases in 1999 each resulted from single \( P. falciparum \) genotypes (51, 95). It is most likely that these clusters result from single mosquitoes infecting multiple people, as occurs frequently in areas of endemicity (40), although the possibility of secondary transmission locally (and thus the theoretical possibility of an outbreak) should be excluded only after careful consideration...
of the potential vectorial capacity of local mosquito populations.

SELECTION AND SPREAD OF DRUG RESISTANCE ALLELES

Evolution

Over the last few decades, the emergence and spread of many different drug resistance alleles has affected the epidemiology of malaria and options for its treatment (197). Many West African populations still use chloroquine as a first-line treatment for *P. falciparum* malaria, although parasite resistance in the region has recently risen to very high levels and it has long been very ineffective elsewhere. Resistance is mediated by a parasite food vacuole membrane transporter molecule, encoded by the gene *pfmdr1* (on chromosome 5 of the parasite) (156). In areas of endemicity, clones could have a higher mutation rate (potentially an adaptive “mutator” phenotype) (156). In endemic areas, resistance is mediated by a parasite food vacuole membrane transporter molecule, encoded by the gene *pfmdr1* (on chromosome 5 of the parasite). A large study in Mali was the first to confirm that treatment failure with chloroquine is determined by parasites having a *pfcr* allele encoding tyrosine at codon 76 and that this association is stronger than for polymorphisms in flanking genes on parasite chromosome 7 (54). An allele encoding tyrosine at codon 86 of the *pfmdr1* gene on chromosome 5 was also associated with treatment failure, consistent with earlier findings that this *pfmdr1* allele is associated with chloroquine resistance in African populations, and the role of both *pfcr* and *pfmdr1* alleles is clearly indicated in recent studies from different countries (101, 127).

Once resistance has evolved to make chloroquine ineffective in a country, the next first-line treatment to be adopted has commonly been the coformulated synergistic combination of sulfadoxine and pyrimethamine (SP), commonly referred to by the trade name Fansidar. These compounds target the dihydrofolate synthase and dihydrofolate reductase enzymes, respectively, of the folate biosynthesis pathway in the parasite. However, resistance has resulted from several structural mutations in each of the genes encoding these enzyme targets, *dhfr* (on chromosome 4) and *dhps* (on chromosome 8), and the increase in resistance has been even more rapid that that experienced with chloroquine.

Mutation in malaria parasites occurs at a fairly typical rate for a eukaryote (a point mutation rate of approximately $1 \times 10^{-9}$ per nucleotide site per mitotic division) (144), so new mutants are produced all the time in natural populations. One study even provides evidence suggesting that some *P. falciparum* clones could have a higher mutation rate (potentially an adaptive “mutator” phenotype) (156). In areas of endemicity, it is not uncommon for an infected person to be carrying more than $1 \times 10^{9}$ parasites (a child with an unremarkable parasite density of 1,000 µl$^{-1}$ in the blood would have approximately this number), and such infections would be likely to contain at least one parasite with a point mutation at almost any nucleotide position (ignoring strongly deleterious mutations that might prevent development or replication). It was considered that drug treatment could thus frequently select de novo mutants parasites arising within individual subjects and that resistance would arise everywhere due to the rate of background mutation and instantaneous selection pressure. However, chloroquine- and SP-resistant alleles detectable today appear to have only a small number of origins (135, 164, 200).

Resistance to chloroquine arose a few times independently in Southeast Asia and South America (78). Strong selection and an increase in frequency of resistance alleles generated absolute linkage disequilibrium over the *pfcr* locus and flanking genes on parasite chromosome 7 that initially made it difficult to tell exactly which was the gene under selection (182). However, once it was clarified that codon 76 and neighboring codons in *pfcr* encoded structural changes in the transporter that conferred the resistance phenotype (78), the linkage disequilibrium flanking the gene became useful as a signature of past selection and geographical spread of the resistance alleles (200). One major haplotype spread through populations in areas of endemicity in Asia and was introduced to East Africa in the late 1970s, from where it spread throughout the continent, rendering chloroquine increasingly ineffective. The observed extent of local linkage disequilibrium due to “hitchhiking” (covering approximately 200 kb flanking the gene on chromosome 7) has been compared with expectations from a theoretical model to derive the likely selection coefficient that has operated over time (200). It remains unknown why *pfcr* alleles conferring chloroquine resistance did not emerge de novo to become common in Africa.

Remarkably, the spread of pyrimethamine resistance due to mutant *dhfr* alleles shows a very similar pattern of origin and geographical colonization, with a highly resistant allele on an ancestral molecular haplotype flanking the *dhfr* gene on chromosome 4 that spread throughout Southeast Asia (135) and then into East Africa (165), with subsequent spread within Africa (164). Analysis of the strength of linkage disequilibrium flanking the highly resistant allele (causing a localized “valley” of reduced allelic diversity) shows that the intensity of selection has been strong in Southeast Asia (135) (Fig. 4) and also in Africa (146). Although it has been studied less intensively, there is apparently a similar effect of selection on the *dhps* allele (containing the codons 437G and 540E) that confers a high level of resistance to sulfadoxine, with spread of a single major haplotype in southern and east Africa (164). Thus, the geographical spread of resistance genes has been of primary importance, but rare resistance mutants that have been recently detected in Africa might also increase in frequency to become a significant problem in continent in the future (16, 124). Early mapping of new resistance genes in the *P. falciparum* genome in the future, even to antimalarials that are currently highly effective as they have not yet been widely deployed, might be achieved by prospective genome-wide scanning of polymorphisms to detect hitchhiked chromosomal haplotypes (7).

Surveillance and Prevention

Identification of the molecular determinants of resistance allows population surveillance to be performed, ideally to inform policy on first-line and second-line drug use within a country or region. As there are several alternative methods for efficiently genotyping polymorphisms in drug resistance genes,
FIG. 4. Microsatellite variability is reduced around the dhfr locus on P. falciparum chromosome 4 in a population sample from the Thai-Myanmar border. Diversity is plotted on the y axis as the expected heterozygosity index (with 1 standard error). Distances (in kb) of microsatellites from the dhfr gene are shown on the x axis. The observed reduction in diversity is due to a selective sweep that caused a drug resistance allele encoded by the dhfr gene to become very common in the population and fits expectations under population genetic models with realistic parameters, as shown by the black and dotted lines. (Reprinted from reference 135 by permission of Oxford University Press.)

for example, the dhfr and dhps genes that affect resistance to antifolate drugs (2, 6, 63, 147), this should be widely done. There are a few modestly funded regional networks that aim to provide surveillance and advice on changing patterns of drug resistance, and there is a need for an effective open-access global database that would give a better empirical basis and context to their work (171).

In East Africa, where failure rates after SP treatment have become very high over the last 10 years, dhps 540E shows the strongest allelic association with treatment failure, with dhfr 59R being next strongest (59, 107, 185). This is because these are present on haplotypes of dhps and dhfr that carry a combination of mutational changes conferring the highest resistance to sulfadoxine and pyrimethamine, respectively. Therefore, although full haplotypic typing of the genes would be ideal where resources permit, a lower cost screening or surveillance system could be based on a minimal genotyping system for dhps 540E (possibly together with dhfr 59R).

The geographical region in which surveillance of dhfr and dhps genotypes is now most important is West Africa, where resistance to SP is less common than in East Africa and where the treatment may be used (alone or with chloroquine or amodiaquine) for some time before other affordable alternatives are available. Of the 16 countries of malaria endemicity in the region (from Cameroon in the southeast to Mauritania in the northwest), genotype data on dhfr and dhps are publicly available for 10 countries, but samples taken since the year 2000 are available for only 7 countries (Table 1). Despite the relative paucity of data from such a large region of endemicity, it is clear that dhfr resistance alleles are very common (including 59R which is usually on the “triple” mutant allele), but frequencies can vary greatly among different sites within a country (e.g., in Senegal or Cameroon) (Table 1), as has been noted among different sites within countries in East Africa (1, 80, 147) and Southeast Asia (8). In contrast, dhps resistance alleles with 540E are absent or rare (detected only in Ghana and Nigeria so far). This would encourage the hope that many countries could have a few years during which SP drug use would be effective before dhps 540E alleles reach frequencies that cause high rates of treatment failure. However, there are already very high treatment failure rates in recently conflict-ridden Liberia and Sierra Leone (33, 34) for which there are no survey data for dhps 540, so it is possible that high frequencies of this important resistance allele are already established there. It is urgent to have more intensive surveillance for the dhps 540E allele while SP use continues in West Africa.

The tabulated data available (as represented in Table 1) are not allele frequencies in the parasite population, which would be ideal for evaluating the population genetic spread of resistance, but rather the proportions of infections that contain a given parasite allele (this will tend to be higher in populations that contain multiple-genotype infection). Individuals who are infected with a mix of different parasite genotypes are more likely to have at least one drug-resistant parasite genotype and thus to fail treatment, compared with individuals that are infected with only a single parasite genotype. A recent analysis of over 3,000 patients enrolled in clinical trials in over 7 sites in Uganda has clearly confirmed that this is so (114), with odds ratios of up to three for increased probability of failure in those with many genotypes prior to treatment compared to those with only 1 or 2 genotypes. Although this is an intuitively obvious effect, it has interesting implications. It would suggest that populations in areas of endemicity with more intense malaria transmission, and thus higher numbers of genotypes per infection, may have a higher treatment failure rate than populations that have low transmission (assuming the

<table>
<thead>
<tr>
<th>Country</th>
<th>Location</th>
<th>Yr of sample</th>
<th>% of isolates containing resistance allele:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mauritania</td>
<td>Aioun</td>
<td>1998</td>
<td>17 0 67</td>
</tr>
<tr>
<td>Senegal</td>
<td>Kobileni</td>
<td>1998</td>
<td>13 0 67</td>
</tr>
<tr>
<td>Senegal</td>
<td>Pikine</td>
<td>2003</td>
<td>61 0 137</td>
</tr>
<tr>
<td>Senegal</td>
<td>Thies</td>
<td>2003</td>
<td>40 0 137</td>
</tr>
<tr>
<td>Senegal</td>
<td>Tambacounda</td>
<td>2003</td>
<td>28 0 137</td>
</tr>
<tr>
<td>The Gambia</td>
<td>Farafenni</td>
<td>2001</td>
<td>75 0 62</td>
</tr>
<tr>
<td>Guinea-Bissau</td>
<td>Bissau</td>
<td>2001</td>
<td>42 0 105</td>
</tr>
<tr>
<td>Liberia</td>
<td>Harper</td>
<td>2000</td>
<td>84 ND* 33</td>
</tr>
<tr>
<td>Mali</td>
<td>Bamako</td>
<td>1995</td>
<td>13 0 53</td>
</tr>
<tr>
<td>Mali</td>
<td>Kidal</td>
<td>1999</td>
<td>35 0 56</td>
</tr>
<tr>
<td>Burkina Faso</td>
<td>Bogo-Dioullasso</td>
<td>2004</td>
<td>67 0 57</td>
</tr>
<tr>
<td>Ghana</td>
<td>Tamale</td>
<td>2002</td>
<td>65 1 128</td>
</tr>
<tr>
<td>Nigeria</td>
<td>Ibadan</td>
<td>2003–2004</td>
<td>32 24 89</td>
</tr>
<tr>
<td>Camerooon</td>
<td>Yaounde</td>
<td>1994–1995</td>
<td>ND 0 14</td>
</tr>
<tr>
<td>Camerooon</td>
<td>Yaounde</td>
<td>1999</td>
<td>71 ND 15</td>
</tr>
<tr>
<td>Douala</td>
<td>1999</td>
<td>86 ND 15</td>
<td></td>
</tr>
<tr>
<td>Eseka</td>
<td>1999</td>
<td>74 ND 15</td>
<td></td>
</tr>
<tr>
<td>Bertoua</td>
<td>1999</td>
<td>47 ND 15</td>
<td></td>
</tr>
</tbody>
</table>

* ND, not determined.
underlying allele frequencies are the same across the different populations). However, this relationship is not seen to hold, as people in higher transmission areas have more acquired immunity, and this appears to be a more important determinant of treatment success for a population than the prevalence of drug resistance alleles (80). Within local populations, individual treatment success with chloroquine is strongly age dependent in a manner that appears to clearly reflect acquired immunity (55) and correlates with acquired antibodies to merozoite antigens (148).

There are effective antimalarial drugs that have not yet been widely used in populations in areas of endemicity, generally because they are too expensive. Atovaquone targets the parasite mitochondrial cytochrome b and is generally very effective in its use together with the antifolate proguanil (in the single-tablet Malarone formulation) as prophylaxis and treatment for travelers from areas of nonendemicity. Rare cases of prophylaxis or treatment failure are caused by new point mutations in the parasite cytb gene, at key positions such as codon 268 where different mutational changes have occurred (Tyr268Ser and Tyr268Asn) (79, 108, 198). In many organisms, the mitochondrial genome has a higher mutation rate than the nuclear genome, so it is possible that mutations in the parasite cytochrome b are more common than in other drug targets encoded by nuclear genes. Together with the fact that different mutational changes in the same codon can cause resistance, this would allow resistant mutants to arise particularly frequently. However, as the cases of atovaquone failure are usually treated with effective rescue therapy and many of them present after travelers return to areas of nonendemicity, the resistance alleles are not transmitted, so there has apparently been no selection on natural parasite populations (133). Thus, treatment failure due to parasite resistance to atovaquone has been restricted to an absolute minimum caused only by new mutations. This could probably not be maintained if the drug were widely used in areas of endemicity.

Effective resistance management is possible in areas of endemicity, however. In Thailand, mefloquine resistance emerged rapidly in the late 1980s after it was introduced in combination with SP, to which there was already substantial resistance (191). Mefloquine is commonly used as prophylaxis in travelers and nonimmune individuals living in areas of endemicity for considerable periods, but (as for atovaquone) rare cases of treatment failure in these relatively privileged individuals are normally dealt with by effective rescue therapy. Despite an early emergence of resistance in Thailand, mefloquine has now been preserved for long-term use but is not available for most people with malaria.

**ADAPTIVE ANTIGENIC POLYMORPHISMS**

**Natural Immune Selection**

There are a number of challenges that need to be overcome to develop a malaria vaccine. Protective immune responses are slowly acquired after repeated infections, and there is considerable genetic polymorphism of many target antigens within local populations and phenotypic variation due to differential expression of genes within parasite genomes. Antigenic polymorphisms can be maintained in populations due to the frequency-dependent effect of acquired immune responses, with alleles having a selective advantage when rare (44). A general way of looking for evidence of selection that maintains alleles within populations (balancing selection) is to analyze allele frequency distributions and compare these with the patterns that are likely to occur in the absence of any selection (under a neutral model of molecular evolution that incorporates mutation and random gene frequency drift).

Among the blood stage antigens of *P. falciparum*, analysis of allele frequencies yields the strongest evidence of such selection on the block 2 region of *msp1* (43) and on domains I and III of *ama1* (149, 150), with some evidence also for *eva1* (18), *msp2* (41), and *msp3* (153). Studies on naturally acquired antibodies against these particular polymorphic antigens of *P. falciparum* support a hypothesis that they are targets of responses that protect against clinical malaria (32, 43, 126, 151–153, 186). However, in some cases (such as for *AMA1* and* EBA175*), the data do not resolve whether the most effective immune responses are against polymorphic epitopes or whether immunity is directed mainly to conserved epitopes and only partially to polymorphic epitopes (to an extent sufficient to exert some frequency-dependent selection). To address this, there is scope for further design of antigenic reagents for these and other antigens that contain polymorphisms and appear to be targets of protective immunity. There is also a need for immunoepidemiological studies that use outcome measures incorporating information on parasite alleles present in clinical infections so that allele-specific effects of antibodies can be assessed.

**Vaccine Trials**

Trials of the efficacy of candidate malaria vaccines against natural infections present an opportunity to study whether vaccine-induced immune responses are selectively effective against the vaccine-type allele. If this is so, deployment of such a vaccine might cause the non-vaccine-type alleles in a population to replace the vaccine-type allele, and the overall efficacy of the vaccine would progressively decline. Allele-specific data for two different vaccines have been analyzed in this way in large phase 2 and 3 trials. The synthetic peptide combination SPI66 that contains a short sequence from near the N terminus of the *P. falciparum* merozoite surface protein 1 (MSP1) was tested in The Gambia (91), and the recombinant protein-based RTS,S/AS02A that contains a large portion of the *P. falcipa*
circumsporozoite protein was tested in The Gambia (3) and Mozambique (69). The RTS,S/AS02A vaccine had a significantly protective effect and the SPf66 vaccine did not, but neither affected the allele frequencies of their respective antigens in the infections that occurred in the vaccine group compared to the control group (3, 69, 91). In contrast, a smaller trial in Papua New Guinea of a combination vaccine that included one allelic form of MSP2 showed that subjects who were vaccinated had subsequent P. falciparum infections with disproportionately higher relative frequencies of the alternative MSP2 allelic type, suggesting that this vaccine had an allele-specific effect on MSP2 (84).

MOLECULAR MECHANISMS OF VIRULENCE

Severe Malaria

It is not yet clear whether there are polymorphisms that cause some P. falciparum parasites to be inherently more virulent than others. Polymorphisms in a small number of antigen genes have been typed in a wide variety of population samples, so it is not surprising that there have been occasional reports of statistically significant differences in the proportions of alleles at such loci as msp1 (12, 109), msp2 (68, 161), and eba175 (47) between severe and mild malaria samples. Leaving aside the possibility of confounding, which is a particular problem for case control comparisons, one would expect some of these associations to be due to chance, but causal differences are not ruled out. It is probable that most common virulence polymorphisms that can be studied are targets of immunity, as balancing selection from the frequency-dependent acquired immune response would slow down or arrest fixation of optimally fit alleles. Otherwise, evolution toward intermediate virulence is expected due to a trade-off between transmission rate and avoidance of host mortality, and polymorphism would only be transient before the alleles that cause optimal intermediate virulence are fixed. Possible disease associations may need to be tested in further studies by analyzing parasite alleles alongside measurements of allele-specific immune responses in the same subjects.

There is evidence to suggest differences in the infected erythrocyte variable surface antigen (VSA) repertoire expressed by parasites isolated from patients with severe and mild malaria and also differences between infections from naive and immune individuals (140). It is proposed that VSA subsets causing most-efficient cytoadhesion, rapid parasite replication, and higher virulence predominate in the naive host, but immunity to these can be gained after only a few infections (if the host survives), after which the VSAs expressed are associated with a lower parasite replication rate and, thus, milder clinical infections. If this is so, it encourages the possibility of developing a vaccine that reduces the risk of severe malaria, based on a restricted subset of var gene products (PlEMP1 antigenic variants) that may largely determine the VSA phenotypes.

Consistent with this, the pattern of switching of var gene expression is ordered and differs between genotypically distinct clones of parasites in vitro (92), so it is likely that a highly nonrandom program of var gene expression operates that underlies the VSA distribution observed in infections (27, 140). However, it has not yet been possible to define VSA patterns that are reproducibly associated with severity of disease (27, 28). The var gene family is very diverse (82), but the gene structures can be clustered into different groups, with one classification system identifying major groups A to E based on the arrangement and number of Duffy binding-like (DBL) domains (112) and another based on comparative sequence analysis of a region of the N-terminal DBL domain (DBL-1α) that is contained in most var genes (27). Most detailed studies of expression have been conducted with laboratory clones such as 3D7 and FCRC3/A4, and the validation of these is challenging, with differences in the performance of quantitative PCR and mini-array methods (60). Analysis of very early blood-stage parasites in a sporozoite-infected volunteer indicates that group A var genes may be associated with high replication rate in nonimmune individuals (96). Recent studies on expression of different var gene groups in field isolates from Tanzania and Papua New Guinea show varying associations, with differences between severe and mild malaria (and broadly between symptomatic and asymptomatic) samples that need to be tested further (100, 166) and an indication that Papua New Guinea isolates with a rosetting phenotype tended to have higher levels of group A var transcripts than nonrosetting isolates (100).

There is a need to develop assays for var gene transcripts that can be applied in an unbiased manner to genetically heterogeneous isolates, which will be helped by sequencing projects on additional P. falciparum genomes (http://www.sanger.ac.uk/Projects/P_falciparum) as it requires comprehensive data on the var gene repertoire.

Variation among P. falciparum genotypes in intrinsic replication rate in the blood (173) is likely to be due in part to efficiency of erythrocyte invasion mediated by alternative receptor usage. A study in Thailand indicated that the initial replication rate of parasites in culture was higher for those isolated from severe malaria versus mild malaria cases (35), although such a difference was not seen in Mali or Kenya (50). Isolates of P. falciparum show a broad range of erythrocyte invasion phenotypes (defined as reliance on different erythrocyte receptors that have distinct susceptibility to proteolytic enzymes and neuraminidase) in culture invasion assays, with those sampled from single areas of endemicity showing as great a range of phenotypes as those from diverse geographical sources (17, 116, 142). Studies using transfection-mediated targeted gene knockout of laboratory-adapted cultured lines have shown that the presence (and different levels of expression) of Rh genes (particularly Rh1, Rh2b, and Rh4) and eba genes (particularly eba175, eba140, and eba181) in different parasite lines affect the use of alternative invasion pathways (64, 65, 83, 85, 119, 180, 188). Studies on clinical isolates in Kenya have shown distinct expression profiles of these genes that are consistent with those shown in laboratory-adapted isolates (138). Larger studies of expression profiles in clinical isolates, possibly combined with analysis of antibodies to these parasite proteins in the same individuals, should reveal if they determine invasion phenotypes or risk of severe disease.

The whole transcript profile of isolates may be assayed using microarray methods, for which the technology and informatics base has continued to expand since the development of the first useful arrays (24, 115). Technical methods for analysis of clinical isolates have been established that are sensitive enough
Placental Malaria

Placental *P. falciparum* infection during pregnancy causes a substantial risk of miscarriage or poor birth outcome. The enhanced infection of the placenta is largely due to a variant ligand on the *P. falciparum*-infected erythrocyte that binds to chondroitin sulfate A on the placental capillary endothelia (93), with some evidence that a secondary ligand may bind to hyaluronic acid (21). Substantial evidence indicates that the product of a particular *P. falciparum var* gene (var2csa) that encodes a variant of PIEMP1 is responsible (170, 193) and is expressed at high levels in most infections of pregnant women (61, 190) but only occasionally in infections of others (190). Initial evidence suggesting the involvement of two other candidate *var* genes (var1csa and CS2var) has not been similarly supported (168, 170). The var2csa gene encodes a unique arrangement of 6 DBL domains, the first N-terminal three of which are unusual types, “DBLx,” while the C-terminal three are the DBLe type. It exists in an inverted subtelomeric orientation compared with other subtelomeric *var* genes, and this together with its distinct gene structure may help to protect its integrity from ectopic recombination with other *var* genes.

Infected pregnant women can make antibodies, pregnancy-associated malaria (PAM)-specific parasite molecules, of which the response to the var2csa-encoded PIEMP1 appears to be the most important, as it can block adhesion of these parasites to chondroitin sulfate A. The presence of such antibodies is associated with better pregnancy outcome, such as higher birth weight (179). The specificity of the PAM-associated antigenic targets is emphasized by the observation that the presence of antibodies to other (non-PAM-associated) VSAs does not correlate with improved birth outcome (179). A recent study has suggested the existence of serological polymorphism in the chondroitin sulfate A-binding pregnancy-associated malaria antigen at a population level (45), indicating that the adaptive and immunological significance of polymorphisms in var2csa should be investigated. An initial analysis of divergent allelic sequences from laboratory isolates, and a partial sequence from the homologous *var* gene in the chimpanzee parasite *Plasmodium reichenowi*, indicates that var2csa is under diversifying selection in *P. falciparum* (189).

CURRENT CHALLENGES

Prospects for research and for a rapidly improved understanding of malaria are good. On the basis of information now available, there should be few limits to the scope for molecular biological and genetic investigations to be combined effectively into epidemiological and clinical studies. Genome projects not only transformed understanding of parasites but also facilitated a cultural change among researchers, so there is now greater willingness to contribute to shared priorities and projects. Potential pitfalls ahead would seem to be those that are the inevitable side products of an immense development of this subject.

Current work appears to be driven by technology and by the massive amount of information in genomes. Molecular evolutionary analysis, as it forms the backbone of comparative genomics, has now become a high-throughput industry. Population genetics is also moving up to a genomic scale as multiple isolates of a species are sequenced, and single-nucleotide polymorphism discovery allows broad genotyping approaches to be used to identify signatures of selection and associations with phenotypes. This could distract from primary questions about disease causation and epidemiology. Newer and more sophisticated assays and technical platforms are constantly becoming available, while methods and items of equipment that were only recently established are rapidly considered obsolete.

Thus, the gap between resources available to leading research groups and those in most countries where malaria is endemic is wider than ever, despite the freely available genome information.

Solutions to this will be costly and need to go deeper than technology transfer or the development of circumscribed scientific network schemes. The understandable desire for training and “capacity building” that exists in resource-poor settings, and is shared by well-intended donors, needs to be more effectively channeled. This requires stronger long-term leadership by scientists in countries of endemicity and well-targeted international funding to support institutes and research programs. This is a multilateral challenge, which if successful, will ensure that future work on the molecular epidemiology of malaria will be even more interesting and relevant.

ACKNOWLEDGMENTS

I am grateful to Bal Singh and colleagues for helpful discussions and to anonymous reviewers for comments and suggestions.

REFERENCES


