Epidemiological and Population Genetic Studies on Polymorphic Antigens of *Plasmodium falciparum*

A thesis by

David Joseph Conway

Submitted for the degree of Doctor of Philosophy

University of Edinburgh
Institute of Cell, Animal, and Population Biology.

Dedicated with love to my Mum,

and in loving memory of my Dad.
Praise the Lord, all you nations; extol him, all you peoples.
For great is his love towards us, and the faithfulness of the Lord endures for ever.

Praise the Lord.

(Psalm 117)
Table of Contents

Preface

Acknowledgements

List of Abbreviations

Abstract

Chapter 1: Introduction

1.1. A human disease and its control
   1.1.1. Malaria
   1.1.2. Strategic Control

1.2. Epidemiological insight and after-thought
   1.2.1. A priori modelling
   1.2.2. Modification and revision of models
   1.2.3. Current need for model development

1.3. The point of population genetics
   1.3.1. The study of processes
   1.3.2. Genetic polymorphism and the epidemiology of parasitic infections
   1.3.3. Human and mosquito polymorphisms

1.4. Immune responses and immunity to malaria
   1.4.1. Observations from human infections
   1.4.2. Antibodies
   1.4.3. Cells and cytokines

1.5. Polymorphic proteins
   1.5.1. Characterisation of protein polymorphisms
   1.5.2. Antigenic polymorphisms
1.5.3. The precursor to the major merozoite surface proteins - MSP1 (or PMMSA, MSA-1, PSA, gp195, p190)
1.5.4. The second merozoite surface protein - MSP2 (or GYMSSA, MSA-2, gp35-56)
1.5.5. An exported protein - Exp1 (or CRA, p23)
1.5.6. The significance of merozoite surface protein polymorphism

1.6. Aims and outline of the thesis

Chapter 2: Materials and methods

2.1. Isolation and culture of *Plasmodium falciparum*

2.1.1. Study areas, clinics, and laboratories
2.1.2. Diagnostic identification of *P.falciparum*
2.1.3. Blood collection from patients
2.1.4. Washing erythrocytes
2.1.5. Malaria culture medium
2.1.6. Freezing and thawing of culture-adapted *P.falciparum* clones
2.1.7. *In vitro* culture
2.1.8. Parasite concentration for DNA preparation

2.2. Amplification, cloning, and sequencing of parasite DNA

2.2.1. Parasite genomic DNA preparation
2.2.2. DNA amplification
2.2.3. DNA fractionation by agarose gel electrophoresis
2.2.4. Purification of DNA
2.2.5. Kinasing insert DNA
2.2.6. Digestion of bacteriophage M13
2.2.7. Ligation of insert with vector
2.2.8. Transformation of host cells
2.2.9. Identification of recombinants
2.2.10. Preparation of single-stranded DNA templates
2.2.11. Sequencing reactions
2.2.12. Sequencing gel electrophoresis

2.3. Characterising Plasmodium falciparum antigen polymorphisms

2.3.1. Preparation of multipot schizont slides
2.3.2. Monoclonal antibodies
2.3.3. Single-colour indirect immunofluorescence
2.3.4. Two-colour indirect immunofluorescence
2.3.5. Photographing microscopic fields

2.4. Blood group typing

2.4.1. Cells and reagents
2.4.2. Direct agglutination of erythrocytes

2.5. Computer methods

2.5.1. Statistical tests using commercial software
2.5.2. A program for analysis of genotype diversity

Chapter 3: Are Plasmodium falciparum populations 'clonal'?

3.1. Introduction
3.2. Methods
  3.2.1. Practical methods
  3.2.2 Statistical methods
3.3. Results
3.3.1. Allelic serotype frequencies at the MSP1, MSP2, and Exp-1 antigen loci
3.3.2. Testing for linkage disequilibrium
3.3.3. Expected and observed 3-locus genotype diversity within the urban/periurban region
3.3.4. 3-locus genotype diversity in villages

3.4. Discussion
3.5. Summary

Chapter 4: The epidemiology of multiple-clone *Plasmodium falciparum* infections in malaria patients

4.1. Introduction
4.2. Methods
   4.2.1. Practical methods
   4.2.2. Statistical methods
4.3. Results
   4.3.1. Clonal diversity within natural *P. falciparum* infections
   4.3.2. Mean number of clones per isolate
   4.3.3. Origin of multiple-clone infections
4.4. Discussion
4.5. Summary

Chapter 5: Genetic evidence for the importance of interrupted feeding by mosquitoes in the transmission of malaria

5.1. Introduction
5.2. Methods
   5.2.1. Practical methods
   5.2.2. Statistical methods
5.3. Results
5.4. Discussion
5.5. Summary
Chapter 6: Intragenic Recombination and non-random associations between polymorphic domains of the precursor to the major merozoite surface antigens

6.1. Introduction
6.2. Methods
6.3. Results
   6.3.1. Nucleotide sequences and putative intragenic recombination sites within a region of the MSP1 gene
   6.3.2. Epitope locations deduced from correlations between amino acid sequences and reactivities with MAbs
   6.3.3. Allelic serotypes of MSP1 in endemic populations of P. falciparum
   6.3.4. Non-random associations between polymorphic epitopes of MSP1
6.4. Discussion
6.5. Summary

Chapter 7: A longitudinal study of defined Plasmodium falciparum antigen polymorphisms in a Gambian population

7.1. Introduction
7.2. Methods
7.3. Results
   7.3.1. Scoring of P. falciparum isolates for expression of antigen epitopes
   7.3.2. Comparison of polymorphic epitope frequencies during a single transmission season
   7.3.3. Comparison of polymorphic epitope frequencies during the period 1982-1989
7.4. Discussion
7.5. Summary
Chapter 8: Testing for associations between *Plasmodium falciparum* merozoite surface protein polymorphisms and blood groups in malaria patients

8.1. Introduction
8.2. Methods
8.3. Results
8.4. Discussion
8.5. Summary

Chapter 9: Conclusions and implications of the thesis

References
Preface

This thesis has been composed by myself, and I did all the work described, except where explicitly stated.
Acknowledgements

Particular thanks are due to my project supervisor, Dr. Jana McBride for her active support and discerning criticism of this work.

Drs. Brian Greenwood, Dominic Kwiatkowski, Adrian Hill, Steve Allen, Hilton Whittle, Harry Campbell, Andrew Wilkins, Mr. Chris Grummitt, and Mr. Momadou Sanyang, provided support in sample collection and in facilitating laboratory work at MRC in The Gambia. Professor Lateef Salako, Drs. Ayoade Oduola, O. Walker, and A. Sowunmi, kindly provided facilities for the study at University College Hospital, in Ibadan, Nigeria. Dr. Virgilio do Rosario provided isolates from Brazil, with support from Drs. Alvaro Couto and Salma Oliveira.

Anti-P.falciparum monoclonal antibodies were kindly provided by Drs. Anthony Holder, Jeffrey Lyon, Richard Pink, Robert Reese, and Alan Saul, in addition to those produced by Dr. Jana McBride. Antiserum and monoclonal antibodies for blood group typing were provided by the Scottish National Blood Transfusion Service, with technical advice from Mr. Charles Green and Dr. Robin Fraser.

I received advice on statistical analyses from Dr. Peter Keighley, Dr. Steve Bennett, Dr. Andrew Leigh-Brown, Mr. Adam Eyre-Walker, Ms. Jo Armstrong, and Professor Bill Hill. Many thanks are due to those who patiently
encouraged me in laboratory techniques, particularly Mrs. Jane Robinson, Dr. Anjam Khan, Dr. John Clark, Mr. Richard Fawcett, Miss Alison Creasey, and Professor David Walliker. Thanks to Graham and Frank for their photographic services, to Professor Bob Sinden and Dr. Philip Ashmole for helpful discussion, and to members of the Institute of Cell, Animal, and Population Biology who have helped me directly or indirectly.

I am grateful to the Medical Research Council, for providing support for me on a research studentship, and travel awards for the studies in West Africa.
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>BTS</td>
<td>Scottish National Blood Transfusion Service</td>
</tr>
<tr>
<td>cDNA</td>
<td>DNA copy of RNA</td>
</tr>
<tr>
<td>CSP</td>
<td>circumsporozoite protein</td>
</tr>
<tr>
<td>DAPI</td>
<td>4'-6-diamino-2-phenylindole</td>
</tr>
<tr>
<td>dATP</td>
<td>deoxyadenosine triphosphate</td>
</tr>
<tr>
<td>dCTP</td>
<td>deoxycytidine triphosphate</td>
</tr>
<tr>
<td>ddH$_2$O</td>
<td>double distilled water</td>
</tr>
<tr>
<td>dGTP</td>
<td>deoxyguanosine triphosphate</td>
</tr>
<tr>
<td>dH$_2$O</td>
<td>distilled water</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>dTTP</td>
<td>deoxythymidine triphosphate</td>
</tr>
<tr>
<td>DTT</td>
<td>dithioerythriol</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetra-acetic acid</td>
</tr>
<tr>
<td>Exp-1</td>
<td>an exported malaria protein (CRA, p23)</td>
</tr>
<tr>
<td>FITC</td>
<td>fluorescein isothiocyanate</td>
</tr>
<tr>
<td>HLA</td>
<td>human leucocyte antigen</td>
</tr>
<tr>
<td>HEPES</td>
<td>N-2-hydroxyethylpiperazine-N'-2-ethane sulphonic acid</td>
</tr>
<tr>
<td>IFA</td>
<td>immunofluorescence assay</td>
</tr>
<tr>
<td>IFN-γ</td>
<td>interferon gamma</td>
</tr>
<tr>
<td>kD</td>
<td>kilodalton</td>
</tr>
<tr>
<td>MAb</td>
<td>monoclonal antibody</td>
</tr>
<tr>
<td>MHC</td>
<td>major histocompatibility complex</td>
</tr>
<tr>
<td>MRC</td>
<td>Medical Research Council</td>
</tr>
<tr>
<td>MSP1</td>
<td>the precursor to the major merozoite surface proteins (PMMSA/MSA-1/p190/gp195)</td>
</tr>
<tr>
<td>MSP2</td>
<td>a second merozoite surface protein (GYMSSA, MSA-2, gp35-56)</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline (pH 7.3)</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PEG</td>
<td>polyethylene glycol</td>
</tr>
<tr>
<td>RESA</td>
<td>ring-infected erythrocyte surface antigen (pf155)</td>
</tr>
</tbody>
</table>
Single Letter Amino Acid Code

<table>
<thead>
<tr>
<th>Single Letter</th>
<th>Amino Acid</th>
<th>Single Letter</th>
<th>Amino Acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>alanine</td>
<td>M</td>
<td>methionine</td>
</tr>
<tr>
<td>C</td>
<td>cysteine</td>
<td>N</td>
<td>asparagine</td>
</tr>
<tr>
<td>D</td>
<td>aspartic acid</td>
<td>P</td>
<td>proline</td>
</tr>
<tr>
<td>E</td>
<td>glutamic acid</td>
<td>Q</td>
<td>glutamine</td>
</tr>
<tr>
<td>F</td>
<td>phenylalanine</td>
<td>R</td>
<td>arginine</td>
</tr>
<tr>
<td>G</td>
<td>glycine</td>
<td>S</td>
<td>serine</td>
</tr>
<tr>
<td>H</td>
<td>histidine</td>
<td>T</td>
<td>threonine</td>
</tr>
<tr>
<td>I</td>
<td>isoleucine</td>
<td>V</td>
<td>valine</td>
</tr>
<tr>
<td>K</td>
<td>lysine</td>
<td>W</td>
<td>tryptophan</td>
</tr>
<tr>
<td>L</td>
<td>leucine</td>
<td>Y</td>
<td>tyrosine</td>
</tr>
</tbody>
</table>
Allelic polymorphism at three unlinked loci coding for blood stage proteins of *Plasmodium falciparum* was studied serologically, using a panel of 27 monoclonal antibodies. The proteins, MSP1, MSP2, and Exp-1, exhibited 39, 8 and 2 serotypes respectively, among 567 Gambian, Nigerian, and Brazilian clinical isolates. In each of two years, within an urban/periurban study area in The Gambia, the observed number of 3-locus combinations was in accordance with expectations assuming random assortment between the loci, i.e. panmixia in the *P. falciparum* population. A mean of 2.0 *P. falciparum* clones was detected in patients from this area. Two clones within an individual were more frequently identical at the MSP1 locus than two clones picked randomly from the local population, indicating that non-identical sibling parasites are sometimes acquired from a single mosquito bite. Parasites isolated from children who sleep in the same room are very frequently identical at all three loci, suggesting that a single mosquito may inoculate more than one human on a given occasion.

Intragenic recombination in the MSP1 gene accounts for much of the extensive allelic polymorphism detected serologically. Putative epitopes for several monoclonal antibodies are mapped on the basis of allelic sequence-serology correlations. Strong non-random associations
between epitopes at different domains of MSP1 are broadly similar in three countries, which could be a result of differential selection on recombinant alleles. The frequencies of polymorphic epitopes of MSP1, MSP2, and Exp-1 remained stable over seven years in the Gambian study area, rare epitopes remaining at a low frequency compared to common alternatives, suggesting that some of these polymorphisms are not maintained by frequency-dependent selection. Little evidence was obtained for association between patients' blood groups and merozoite surface protein polymorphisms, although one statistical association between blood group 0 and epitope 8F6/49 on MSP2 should be tested further.
Chapter 1. Introduction

1.1. A Human Disease and its Control

1.1.1. Malaria

Malaria is caused by intra-erythrocytic parasitic protozoa of the genus \textit{Plasmodium}, which complete their life cycle in \textit{Anopheles} mosquitoes (Figure 1). The point of human concern is that infection with \textit{Plasmodium falciparum}, \textit{P.vivax}, \textit{P.malariae} or \textit{P.ovale} can cause acute fever, involving painful and debilitating symptoms. A person infected with \textit{P.falciparum} may develop a life threatening illness, involving complications such as severe anaemia, or hypoglycaemia and coma, although most clinical cases are milder and self cure within a few weeks. Perhaps the most dangerous feature of a \textit{P.falciparum} infection is the unpredictability of its effects (Kitchen, 1949).

The severity of disease is loosely correlated with parasitaemia (Kitchen, 1949), and adherence of large numbers of infected red cells to the endothelium of cerebral capillary venules (Macpherson \textit{et al.} 1985; Berendt \textit{et al.} 1990), although immunopathological processes are also involved (Clark, 1987; Kwiatowski \textit{et al.} 1990). The number of fatalities due to \textit{P.falciparum} malaria is of the order of one million per year, a large
Figure 1. Life cycle of *Plasmodium falciparum* (order Eucoccidiorida), the parasite causing malignant tertian malaria. Parasites are haploid at all stages of the cycle, except immediately after fertilisation in the mosquito midgut. Diploid zygote formation is followed by a meiotic division, and asexual proliferation of haploid parasites occurs within the oocyst on the mosquito gut wall, within human hepatocytes, and within human erythrocytes (Walliker, 1989). During the erythrocytic cycle, intracellular parasites develop from ring stages, to trophozoites, to schizonts which rupture to release 16-48 extracellular merozoites which may invade other erythrocytes and continue the cycle. (Figure reproduced from Trager, 1986).
**Incubation period**

- Hepatic schizont
- Invasion of hepatic cell

**Sporozoites injected to bloodstream with saliva of mosquito**

**IN ANOPHELES**

- Sporozoites in salivary glands
- Sporogony
- Motile zygote. Ookinetes at 24 hr

**IN HUMAN**

- Asexual cycle of schizogony (48 hr) in erythrocytes
- Small proportion of erythrocytic merozoites differentiates to gametocytes (8-12 days)
- Gametocytes
- Taken into midgut of mosquito with blood meal
- Gametes
- Gamete formation and fertilization in midgut

**Mature oocyst at 10-14 days**

- Young oocyst on gut wall protruding into hemocoel

8th-10th day: merozoites from hepatic schizonts initiate erythrocytic cycle
It is more difficult to estimate the annual number of uncomplicated clinical cases caused by all four *Plasmodium* species, as they are mostly unreported, although an estimate of 96 million was given by the World Health Organisation Expert Committee on Malaria (WHO, 1986). The number of asymptomatic infections is much larger.

History reveals nothing concerning the origins of human malaria, except that the earliest writers were already familiar with its symptoms. The different species infecting humans show less genetic similarity to each other than to species infecting other mammals, and may have become adapted independently from previous zoonotic associations (Bruce-Chwatt, 1988). Apart from history and host specificity, human genetics provides compelling evidence for a relatively long association with *P. falciparum* (section 1.3.3.).

### 1.1.2. Strategic Control

Malaria parasites were first identified in human blood by Laveran in 1880, and in an *Anopheles* mosquito by Ross in 1897. Prior to these important scientific developments, a simple and effective cure had been discovered. In the 17th century, an extract of the bark of the Peruvian Cinchona tree was found to be effective
in the treatment of 'intermittent fevers'. The active component, quinine, was characterised in the 19th century, and from the 1840s had widespread prophylactic use among the British colonial services (Bruce-Chwatt, 1988).

The impact of the subsequent parasitological developments was chiefly to introduce the concept of eliminating transmission by reducing mosquito populations, which was shown to be a theoretical possibility (section 1.2.1.). In practice, larval breeding sites were identified and drained, treated with petroleum oil, Paris green insecticide, or by the addition of larvivorous Gambusia fish. Methods of personal protection from mosquito bites were actively encouraged. Where control methods were vigorously employed, results were often encouraging. However, some of the variability in success was recognised to be due to differences in the ecology and bionomics of local mosquito vector populations, and this stimulated entomological research, which became the most active field of malariology in the first half of the present century (see Boyd, 1949a).

A notable success was achieved in the elimination of the efficient vector Anopheles gambiae from Brazil in 1940, ten years after its accidental introduction from West Africa. Then, during the 1940s residual insecticides became available for indoor house spraying against adult
mosquitoes, which was theoretically (section 1.2.2.) and practically a more powerful approach towards reducing transmission. The goal of malaria eradication appeared more feasible than ever, and was officially adopted by the Eighth World Health Assembly in 1955.

Elimination of malaria was achieved on some islands, notably Mauritius, Réunion, and several Caribbean islands, and within some continental areas such as North Africa, where the mosquito populations were smaller and more focalised. However, within most highly endemic areas the impact of vector control was minimal. Between 1969 and 1976, a large WHO co-ordinated study was undertaken at Garki in northern Nigeria, in order to assess the effects of intensive house spraying and mass administration of chloroquine to the local rural population. Thorough epidemiological analysis revealed that, although the transmission rate may have been reduced by as much as 90%, the prevalence of *P. falciparum* had decreased by only about 25%, and returned to its previous level once the intensive control had ceased (Molineaux & Gramiccia, 1980).

Without attempting to reduce transmission, targetted chemoprophylaxis (with chloroquine, proguanil or pyrimethamine) has been used to lower the incidence of clinical malaria in risk groups such as the under-fives, pregnant women, or primary schoolchildren. Possible disadvantages of chemoprophylaxis are the increased risk
of selection for drug resistance, and an altered acquired immunity in the targeted group after terminating chemoprophylaxis (Greenwood, 1984; Otoo et al. 1989). Other disadvantages include the cost and the operational difficulty of obtaining a constant supply of any drug for the targeted group.

Since the Alma-Ata declaration in 1978 (WHO, 1978), concerning the priority of primary health care, much emphasis has been placed on the use of presumptive treatment of malaria cases by village health workers, and the development of support facilities at higher levels in the health services. However, *P. falciparum* parasites resistant to each of the widely used drugs have emerged in different parts of the world, and particular concern has arisen as cases of chloroquine resistance have been identified throughout Africa, a continent which relies heavily on this compound. As a result, much of the research on malaria within the last ten years has aimed to identify and develop new anti-malarial compounds. Mefloquine and Halofantrine have become commercially available, and Artemesin is likely to follow.

Research on molecular and genetic aspects of *Plasmodium falciparum* has been stimulated in part by the desire to understand the observed plasticity of the parasite, in terms of mutation, recombination and adaptive spread of resistant traits in natural populations (sections 1.3. and 1.5.). Proteins which
stimulate immune responses account for the majority of those characterised, as acquired immunity to the parasite plays an important role in regulating natural infections (sections 1.4. and 1.5.). Some targets of immunity are partially known, but it is unknown whether immunisation with any defined polypeptides will induce a clinically useful anti-parasite or anti-disease response.
1.2. Epidemiological Insight and After-Thought

1.2.1. A priori modelling

Models of malaria transmission have played a critical role in epidemiological thinking, and in directing attempts at control or eradication of malaria.

Ross began working on a quantitative model within a few years of his discovery that Anopheles mosquitoes could transmit the disease. Health administrators at the time argued that since it was impossible to eliminate every single mosquito, it would also be impossible to interrupt transmission. Ross was concerned to demonstrate, against the prevailing scepticism, that it was feasible to eradicate the disease by lowering the mosquito density, without eliminating every mosquito.

In its developed form, his was 'a theory of happenings', which predicted rates of occurrence of an event (e.g. infection) from known quantifiable determinants (e.g. mosquito density, biting rate, etc.). He described it as an a priori approach, to be distinguished from inference of determinants by curve fitting to known epidemic data, which he termed the a posteriori approach. It is worth quoting Ross' classic distinction between the two fundamentally different approaches to epidemiology:
'The whole subject is capable of study by two distinct methods which are used in other branches of science, which are complementary of each other, and which should converge towards the same results - the a posteriori and the a priori methods. In the former we commence with observed statistics, endeavour to fit analytical laws to them, and so work backwards to the underlying cause; and in the latter we assume a knowledge of the causes, construct our differential equations on that supposition, follow up the logical consequences, and finally test the calculated results by comparing them with the observed statistics.'


The a priori approach has been adopted as the logical method for developing modern epidemiological models. Practically speaking, the conclusion of Ross' differential equations was that there is indeed a threshold level of mosquito density below which transmission will be interrupted. This validated subsequent attempts at mosquito control and much research on the bionomics of vector species.

1.2.2. Modification and Revision of Models

It was to be 40 years before another useful model was developed, stimulated by new insights into the biology of malaria and new possibilities of control. During the 1950s, Macdonald developed a model which was very similar in structure to the Ross model, but which advanced it in specific ways. Firstly, he assumed that, 'the existence of infection is no barrier to
superinfection, so that two or more broods of organisms may flourish side by side, the duration of infection due to one being unaltered by others' (Macdonald, 1950). This assumption was included in the model, despite a mathematical error in its formulation (Fine, 1975b). Secondly, Macdonald realised that the probability of an infected mosquito surviving long enough to transmit the infection was better expressed as an inverse function of the average mosquito longevity, rather than as a constant as Ross had defined it. The practical importance of this was to illustrate that a reduction of adult mosquito longevity would have a disproportionately large effect on lowering the transmission rate, compared with a similar reduction of the larval mosquito population. This gave validity to the widespread use of adulticides rather than larvicides during the WHO Malaria Eradication Programme.

Theoretically, the major advance of Macdonald's model was the incorporation of the 'basic reproductive rate' of the parasite, defined as the number of secondary cases of infection generated by a single infected individual in a population of susceptibles. The argument followed that if this particular number could be reduced to less than 1, eradication of the disease would occur (Macdonald, 1957).

Macdonald demonstrated that his formula could produce a curve which fitted the observed age vs. prevalence rates among infants in Tanzania and Uganda
Macdonald, 1955). He also recognised acquired immunity as a regulating factor, although it could not be quantified, and so was not included into his model. Although the model has contributed greatly to the understanding of malaria epidemiology, experience during the WHO Malaria Eradication Programme has shown that it is insufficient to describe reality. A large eradication project in Katsina state, northern Nigeria, was planned and evaluated using Macdonald's model, and illustrated well its inadequacies (Najera, 1974). In particular, the predicted effects on human parasite rates of residual insecticide spraying in a hyperendemic area were far too high.

In the light of these failures, a new mathematical model was designed (Deitz, Molineaux & Thomas, 1974; Molineaux & Gramiccia, 1980), and tested during the WHO Garki Project in Kano state, northern Nigeria. This model included some of Macdonald's thinking, but incorporated some presumed effects of acquired immunity. Interestingly, the development of the model and its parameters used the a posteriori as well as the a priori approach. In other words, it was modified interactively with data emerging from two villages in the Garki study area, theoretical curves being fitted to the observed data by an iterative computer program designed to minimise differences. The final model was then tested against data from other villages. It was concluded that,
'the fit was quite good on the whole', and, 'the model simulates the epidemiology of *P.falciparum* infections with acceptable realism'  
(Molineaux & Gramiccia, 1980).

However, this 'acceptable realism' led to the conclusion that malaria control in hyperendemic areas by residual insecticide spraying and mass drug administration is not feasible. This conclusion has resulted in research into alternative control strategies.

1.2.3. Current Need for Model Development

The model of Deitz *et al.* (1974) has been considered, 'to date, the most practical and realistic one' (Aron & May, 1982). However, it is fundamentally similar to all previous ones in that it has a compartmental structure, i.e. individuals are assumed to change from one defined state to another (susceptible, infectious, immune non-infectious, etc.). Aron & May (1982) considered that a theoretical advance could be achieved by constructing a continuum model, in which an individual's infection and immune status is quantified rather than compartmentalised. Such an approach has been considered appropriate for modelling helminthic ('macroparasite') infections, which are relatively long lived and to which immunity is partial (May & Anderson,
In generalised terms, bacterial and viral ('microparasite') infections are relatively short lived and confer a strong immunity to re-infection, and these can be more adequately described by compartmental models (Anderson & May, 1979). For modelling purposes, malaria shows more of the characteristics of a 'macroparasite' than a 'microparasite' (Aron & May, 1982).

Of more immediate need are suitable models of individual components of the malaria system, to roughly predict effects of new control strategies. Curtis & Otoo (1986) predicted the theoretical optimum coverage of a population by chemoprophylaxis with a combination of two drugs for which parasite resistance genes are unlinked, in order to minimise the risk of emergence of drug resistance. Anderson, May & Gupta (1989) predicted the required coverage of a population with either a sporozoite or a gametocyte vaccine, to reduce parasite prevalence, depending on the efficiency and the duration of protection afforded by the vaccine. The predictions of Anderson et al. (1989) do not give an optimistic view of the usefulness of such vaccines, and although these predictions are based on rudimentary assumptions it should be noted that no malaria model has yet erred on the side of pessimism.
1.3. The Point of Population Genetics

1.3.1. The Study of Processes

Population genetic studies involve sampling a finite number of individuals from a population in order to quantify a particular genetic variable or a process underlying such a variable in the population. A study can be designed to test either an evolutionary or an ecological hypothesis.

(i) Evolutionary population genetics is concerned with processes causing divergence between individuals of a population, between populations, or between species. Such processes may be essentially constant (e.g. the mutation rate), or subject to deterministic or stochastic changes in environmental factors or population distribution (e.g. incorporation and spread of a mutation in a population). The neutral theory of molecular evolution (Kimura, 1983) argues that most molecular changes are not driven by natural selection, although some changes do have adaptive significance.

(ii) Ecological population genetics involves analysis of gene (allele) frequencies, diploid genotype (zygote) frequencies, and haploid multi-locus genotype (gamete) frequencies, as follows:

Gene (allele) frequencies - detectable alleles are
enumerated, and their frequencies analysed, to test whether the allelic composition in a local population may be partly a result of selective processes. Spatial analysis may involve testing for correlations of allele frequency clines (Sokal & Oden, 1978) with environmental gradients (Dillon, 1984). Temporal analysis may test for deterministic changes in frequency due to effects of currently operating selection (e.g. frequency-dependant selection: Haldane & Jayakar, 1963; Clarke, B. & O'Donald, 1964; or environmental changes: Clarke, C.A., Mani & Wynne, 1985), on the background of randomly expected stochastic changes (e.g. genetic drift: Kimura, 1955).

Diploid genotype (zygote) frequencies - these can be calculated from allelic frequencies, assuming the Hardy-Weinberg equilibrium. However, selection may act on the diploid genotype in such a way as to promote 'heterosis' (heterozygote advantage). This may maintain allelic polymorphism (i.e. heterozygosity) in the population at higher levels than otherwise expected (Kimura & Crow, 1964). Alternatively, dominance of one allele over another can also influence selection at the zygotic level (Clarke & O'Donald, 1964).

Haploid multi-locus genotype (gamete) frequencies - expected haploid multi-locus genotype frequencies are calculated as the product of the component allelic frequencies, assuming unlinked loci. Significant deviations between observed and expected multi-locus
Genotype frequencies indicate either low rates of recombination (genetic drift will lead to lower genotype diversity and an observed 'clonal' population structure: Selander & Whittam, 1983), ecologically subdivided populations (Ohta, 1982), or epistatic interactions between loci (causing differences in fitness of particular combinations: Lewontin & White, 1960).

1.3.2. Genetic Polymorphism and the Epidemiology of Parasitic Infections

Most advances in population genetics have been derived from studies of free-living organisms, particularly species amenable to laboratory breeding and colonisation. In contrast, protozan and helminth parasites have been studied very inadequately. Most data are contained in tables of isoenzyme frequencies, with sample sizes which may be adequate for interspecific differentiation, but are less useful for considerations of ecological population genetics. Nevertheless, widespread intraspecific variation has been documented for most parasites studied.

Genetic characterisation of parasites may contribute to epidemiology in either of two ways. Firstly, if genetic polymorphisms within a species are considered adaptively neutral, they can be used to analyse the breeding structure of parasite populations and to
identify non-random spatial distribution (e.g. household clustering of genotypes) which may reflect focal or non-random mechanisms of transmission. Alternatively, characterised genetic polymorphisms may determine, or be linked to genes which determine, intraspecific variations in pathogenicity, drug resistance, or host specificity. Such polymorphisms might be useful markers for clinical diagnosis or studies to elucidate zoonotic transmission cycles.

Different population structures are evident among helminth species. For example, the genotypes of *Ascaris suum* sampled by Leslie *et al.* (1982) showed no deviation from Hardy-Weinberg equilibrium, suggesting a considerable amount of cross fertilisation between adults of different genotype in the intestinal lumen. In contrast, no allelic polymorphisms have been detected among *Strongyloides ratti* isolates from different geographical regions (Mark Viney, personal communication). Nadler (1987) has discussed the implications of genetic variation in parasitic helminths from an evolutionary perspective.

In several helminth species, 'strain' differences have behavioral or ecological significance. Within both *Trichinella spiralis* and *Echinococcus granulosus*, genetic characters appear to correlate with the different alternative transmission cycles (Bryant & Flockhart, 1987), and in *T. spiralis* may also correlate with disease
severity (Nelson, 1982). The different 'strains' may reflect sub-species. An analysis of enzyme polymorphisms in *Onchocerca volvulus* from different African countries revealed a small degree of geographic divergence, but no systematic differences between the ecologically defined 'forest' and 'savannah' forms (Flockhart, Cibulskis & Albiez, 1986).

Apart from *Plasmodium* species, the protozoan parasites which have been most intensively studied by genetic techniques are the Kinetoplastida, including *Trypanosoma* and *Leishmania* species. Cluster analysis of iso-enzyme data from *Trypanosoma brucei* isolates has revealed systematic genetic differences between ecologically defined sub-species (Gibson, Marshall & Godfrey, 1980). An attempt has also been made to relate the population genetics of *T. brucei* in Kenya to the local epidemiology of sleeping sickness (Mihok, Otieno & Darji, 1990). Although genetic exchange can occur between *T. brucei* parasites within the tse-tse fly vector (Turner et al., 1990), it appears that in the field, sympatric populations may exist with a degree of reproductive isolation (Mihok et al., 1990).

An analysis of Tibayrenc et al. (1986) suggests that a clonal population structure exists in *T. cruzi*, as most clones are genetically very different from each other, implying a very low rate of genetic recombination. Similarly, genetic recombination is probably rare in Old
World *Leishmania* species, which exhibit a very low frequency of heterozygosity (Tibayrenc, Kjellberg & Ayala, 1990).

It is evident that studies of genetic diversity within parasite species have different aims, whether to determine the population breeding structure, or to identify functional polymorphisms. It is important that study aims are defined in order to avoid confusion, or bias, in interpretations. To quote from a textbook,

'We should not merely be interested in the detection of differences *per se* between isolates, which may often be possible with the extreme resolution of some sophisticated DNA techniques. There should always be a sound basis to our search for variation, and once detected the significance of such variation should be elucidated.'

(Thompson, 1988).

Genetic polymorphisms within *P.falciparum* have been intensively studied (section 1.5.). The haploid parasite has 14 chromosomes, upon which certain loci are identified (Kemp *et al.* 1987). Although there are geographic allele frequency differences (Creasey *et al.* 1990), evidence suggests that *P.falciparum* is a globally panmictic species (Walliker, 1985), and sexual reproduction within the mosquito vector allows mendelian recombination between different genotypes (Walliker *et al.* 1987). In contrast, Tibayrenc *et al.* (1990) suggested that *P.falciparum* populations consist of multiple
discrete clones, implying a low rate of cross-fertilisation. The question of the genetic structure of *P. falciparum* populations is important for any consideration of its genetics, and is thoroughly addressed in this thesis (Chapter 3). Once the breeding structure has been established, it is possible to use genetic characterisation to study the parasite epidemiology in detail (Chapters 4 & 5).

Particular polymorphisms of *P. falciparum* may be of functional significance. *In vitro* drug resistance properties are clearly genetically determined (Walliker *et al.* 1987), and efforts have been made to determine the genetic polymorphisms underlying these traits in nature (Foote *et al.* 1990; Zolg *et al.* 1989). The possible significance of polymorphisms in proteins on the merozoite surface is considered in this thesis (section 1.5., and Chapters 6-8).

1.3.3. Human and Mosquito Polymorphisms

With respect to malaria, important genetic polymorphisms have already been identified in the human host and the mosquito vector.

The geographic distributions of certain human red blood cell polymorphisms, particularly those related to mutations of the beta-globin gene, correlate closely with the distribution of endemic *P. falciparum* malaria. The
The most thoroughly studied polymorphism is that involving the haemoglobin S (sickle-cell) gene. Within Africa, Allison (1964) demonstrated that children with sickle-cell trait (heterozygote genotype AS) had a lower incidence of clinical malaria than normal children (homozygote AA), and that adults with the trait were more resistant to an experimental *P. falciparum* infection than normal adults. In Uganda, Raper (1956) showed that mortality from cerebral malaria was lower in children with sickle-cell trait. Similar studies were repeated by many investigators (reviewed by Livingstone, 1971), which on the whole support these original findings. Despite the high mortality among individuals with sickle-cell disease (homozygote SS), the heterosis has enabled the S allele to be maintained at frequencies of 10% or above in many malaria endemic areas of Africa.

The distributions of other haemoglobinopathies, caused by haemoglobin C, E, and beta-thalassaemia, are also restricted to populations in which malaria is or was recently endemic, suggesting that heterozygote advantage is responsible for maintaining these mutations in these populations (Livingstone, 1971). The frequency of alpha-thalassaemia, caused by deletion of alpha-globin loci, correlates closely with historical levels of malaria endemicity in the Melanesian archipelago (Flint et al., 1986). A direct demonstration of the protective effect of heterozygous genotypes has been more difficult to obtain.
in these cases, however. Apart from the haemoglobinopathies, there is conclusive evidence that individuals heterozygous for glucose-6-phosphate dehydrogenase (G6PD) deficiency are protected against the severe effects of malaria infection, and G6PD deficiency frequencies are highest in malaria endemic areas (Livingstone, 1971; Luzzatto, 1979). In contrast, common blood group polymorphisms do not apparently influence susceptibility to \textit{P.falciparum} malaria (Martin et al., 1979; Molineaux & Gramiccia, 1980).

A large number of \textit{Anopheles} species are responsible for malaria transmission. The geographic range of each species is limited, however, and at a given location there is usually only one or a few species involved (Boyd, 1949a). Differences between sympatric species are not always discernable morphologically, especially among members of a species complex such as \textit{A.gambiae} sensu lato. Cytogenetic techniques can differentiate between the six species of the \textit{A.gambiae} s.l. complex, each of which differs in vectorial capacity (Coluzzi, 1984). In addition, there is much intraspecific diversity within members of the complex, particularly \textit{A.gambiae} sensu stricto and \textit{A.arabiensis}. Intraspecific differences in biting behaviour (anthropophily/zoophily) and resting behaviour (endophily/exophily) complicates vector control strategies aimed at reducing malaria transmission (Molineaux & Gramiccia, 1980).
These traits are genetically determined, and may correlate with inversion polymorphisms on chromosome 2 (Coluzzi et al., 1977). Seasonal and geographic variations in the relative frequencies of different inversion polymorphisms have been documented (White, 1974; Coluzzi et al., 1979), and it appears that population genetic studies using these chromosomal markers would allow the feasibility and effectiveness of vector control measures to be better evaluated (WHO, 1984).
1.4. Immune Responses and Immunity Against Malaria

1.4.1. Observations from Human Infections

The existence of acquired immunity to malaria among inhabitants of endemic areas has been recognised for a long time. The age profiles of prevalence and intensity of parasitaemia in hyperendemic areas consistently show lower levels in older individuals which cannot be explained as a lower level of exposure to infection (Figure 2). From these profiles, two characteristics of anti-parasite immunity to malaria can be deduced: (i) immunity is acquired slowly and gradually after repeated exposure, and (ii) immunity is non-sterilising but maintains parasitaemia at lower levels after re-infection. Studies on induced malaria in volunteers and neurosyphalitic patients confirmed the characteristics of immunity which could be deduced from the age profiles noted above (Jeffery, 1966; Powell, McNamara & Rieckmann, 1972), although they did not address questions of specific immune mechanisms. More recent investigations have attempted this, leading to the current conclusions:

1.4.2. Antibodies

Antibodies are very important. The strongest
Figure 2. Malaria survey data from a hyperendemic area of Tanzania. The parasite rate is the percentage of the age group with blood samples positive for malaria. The spleen rate is the percentage of the age group with enlarged spleens. The parasite density is the number of parasites per cubic millimeter (in thousands) in those blood samples which were positive for the parasites. (Figure reproduced from Aron & May, 1982, p. 165).
evidence comes from the demonstration that passive transfer of purified gamma-globulin from Gambian adult immune serum strongly reduced parasitaemia and alleviated clinical symptoms in children with malaria (Cohen, McGregor & Carrington, 1961; McGregor, Carrington & Cohen, 1963). Human immune gamma-globulin also protected chimpanzees against experimental infection (Sadun et al., 1966). The evidence from sero-epidemiology has been less clear. The Garki Project (Molineaux & Gramiccia, 1980), demonstrated that serum activity to total *P. falciparum* extracted antigens (measured by Ouchterlony double diffusion and indirect haemagglutination assays) showed a positive correlation with parasite abundance in the youngest age groups, and a negative correlation in the older age groups. This was interpreted as indicating that the activity of antibodies in young children reflected the level of recent or current infection, while in older individuals reflected 'acquired immune status' (Molineaux & Gramiccia, 1980). This conclusion was speculative, because, as noted by Marsh et al. (1989),

'The use of methods employing crude blood-stage antigens to measure the humoral anti-malarial response does not allow the differentiation of protective responses from those that merely reflect cumulative exposure.'

Indeed, in a cross-sectional survey and a
longitudinal cohort study in The Gambia, there were no correlations between parasitaemia or clinical illness and antibody titres to crude antigen preparations (Marsh et al. 1989).

In vitro assays have had limited use in determining the role of protective antibodies. Opsonisation of infected erythrocytes by monocytes is enhanced by the addition of immune serum in vitro (Celada, Cruchaud & Perrin, 1982), but it is unclear whether this assay correlates with clinical immunity among individuals. Immune serum has been shown to inhibit merozoite invasion of erythrocytes in vitro (Green et al. 1981; Chulay et al. 1981), but not invariably (Brown et al. 1982), and Fandeur et al. (1984) demonstrated with Saimiri monkeys that in vitro inhibition did not correlate with in vivo protection. Marsh et al. (1989) also noted no correlation between in vitro invasion inhibition and protective immunity, and it has been argued on theoretical grounds that antibody mediated inhibition of invasion may not be an effective mechanism for anti-parasite immunity (Saul, 1987).

1.4.3. Cells and Cytokines

Antibody responses depend on the clonal proliferation of T helper lymphocytes specific for parasite epitopes presented in conjunction with major
histocompatibility complex (MHC) class II molecules on the surface of B cells or macrophages. Only proliferation and release of lymphokines from T cells can stimulate the transformation and clonal expansion of B cells for production of large amounts of specific antibodies. Cells may also function in parasiticidal immune mechanisms, independent of antibodies (Clark, 1987). However, the importance of particular cell populations cannot be confirmed by adoptive transfer experiments in humans, due to the incompatibility of MHC molecules between donor and recipient. Therefore, the cellular role in immunity has been studied by the use of animal malaria models, or by in vitro assays.

Neonatally thymectomised rats, in which T cells cannot mature, have an enhanced susceptibility to *P. berghei* (Brown *et al.*, 1968). Murine cytotoxic T cells are involved in killing intracellular liver-stage parasites in an antigen-specific manner, as evidenced by the effects of adoptive transfer of a cytotoxic T cell clone specific for an epitope on the circumsporozoite protein of *P. berghei* (Romero *et al.*, 1989).

In vitro proliferation of lymphocytes is considered to be a useful means of identifying the presence of T cell epitopes on defined malaria proteins. Such epitopes have been identified on the circumsporozoite protein (CSP: Good *et al.*, 1988), the ring-infected erythrocyte surface antigen (RESA: Kabilan *et al.*, 1988), and the
precursor to the major merozoite surface antigens (PMMSA or MSP1: Sinigaglia et al., 1988), among others. Longitudinal epidemiological studies have attempted to find correlations between lymphoproliferative responses to particular domains of these antigens and subsequent clinical and parasitological experience of individuals (Eleanor Riley, personal communication)

Cytokines (inter-cellular chemical messengers) are closely involved in the functioning of cellular immunity, from the central interaction between T helper cells and antigen-presenting cells to the lesser understood involvements of other leucocyte populations. The study of cytokine production by defined leucocyte populations upon stimulation with defined polypeptide antigens in vitro is the most precise way of studying components of an immune mechanism. Such an approach has been used to study interleukin-4, and interferon-gamma (IFN-γ) secretion in response to synthetic peptides of RESA/pf155 (Kabilan et al., 1988; Kabilan, 1990).

Some cytokines are closely involved in parasiticidal mechanisms, e.g. tumor necrosis factor (TNF: Clark, 1987; Kwiatkowski et al., 1990), and IFN-γ (Schofield et al., 1987). Reactive oxygen species may be released by macrophages and neutrophils to kill intra-erythrocytic parasites, possibly in the spleen (Dockrell, Alavi & Playfair, 1986). There also remain elusive 'factors' which may occasionally be important, such as 'crisis form
factor' (Jensen et al. 1983).
1.5. Polymorphic Proteins

1.5.1. Characterisation of Protein Polymorphisms

There is a range of biochemical and immunological techniques available for the characterisation of proteins, and the determination of allelic polymorphism in malaria parasites. Once the gene for a protein has been cloned and sequenced, it is also possible to study allelic polymorphism with a variety of molecular techniques such as oligonucleotide probe hybridisation and DNA sequencing.

The broadest view of protein polymorphism may be obtained by comparing two-dimensional polyacrylamide gel electrophoresis patterns, of different parasite clones at a given stage in the life cycle (for convenience, the asexual blood stage). For sensitivity to allelic differences, the technique surpasses that of one-dimensional iso-enzyme electrophoresis. Nevertheless, it has been important to painstakingly establish the genetic basis of the observed two-dimensional polymorphisms (Fenton, 1987; Fenton & Walliker, 1990). The identity of some proteins has also been revealed by correlation with other methods of characterisation (Fenton et al., 1989; Fenton & Walliker, 1990).

However, if a range of monoclonal antibodies (MAbs) against polymorphic domains of a protein is available, a
serological method of characterisation may be more sensitive to allelic differences than electrophoretic methods. Additional advantages of a serological method are that it may be quicker or require less parasite material. A large number of MAb s have been raised against *P. falciparum* proteins, in the attempt to characterise those which may subsequently be credited as 'vaccine candidate antigens'.

1.5.2. Antigenic Polymorphisms

The observed serological and molecular polymorphism in many cloned *P. falciparum* antigens has frequently been interpreted as an 'immune evasion strategy' by the parasite (Kemp *et al.*, 1986). It is suggested that an immune response against one allelic type of a parasite protein may not recognise an alternative allelic type encountered during a subsequent infection, thereby enabling the parasite population to avoid an acquired immune response to the protein. This is a popular concept, although there is less popular consideration given to the role of neutral processes (section 1.3.1.), or of selection by mechanisms other than protective immunity (section 1.5.6.), in maintaining allelic polymorphisms.

The following observations have made the 'immune evasion' idea very attractive, and possibly led to the
over-popularisation of the concept. Firstly, it has been observed that protective immunity to *P. falciparum* in non-human primates apparently involves a 'strain-specific' component, immunisation conferring a lower level of protection to challenge with a 'heterologous strain' than with the 'immunising strain' (Sadun *et al.*, 1966; Cadigan & Chiacumpa, 1969; Voller, Green & Richards, 1973). The evidence of such a phenomenon with *P. falciparum* infections of humans is less consistent, however (James, Nicol & Shute, 1932; Boyd, Stratman-Thomas & Kitchen, 1936; Jeffery, 1966; Powell *et al.*, 1972). Secondly, individuals in a hyperendemic or holoendemic area are infected several (and probably many) times before acquiring a significant level of anti-parasite immunity. For a long time it has been considered that such a phenomenon is partly due to gradual exposure to different 'strains' of malaria circulating in a locality (Bray *et al.*, 1962). Thirdly, there is the relatively recent historical description of an exquisite molecular mechanism which allows a single blood-stage clone of *Trypanosoma brucei* to sequentially express variant forms of a major surface glycoprotein, thereby avoiding variant-specific immunity by means of antigenic variation (Turner, 1984). Although *T. brucei* and *P. falciparum* are members of different phyla, and differential gene expression is not the same as allelic polymorphism, parallels have been drawn. Fourthly, a general argument
is that the degree of sequence polymorphism in surface antigens is so great that it must have some adaptive significance, possibly immune evasion. More specifically, it is claimed that the ratio of non-synonymous (resulting in alteration of amino acid sequence) to synonymous nucleotide polymorphisms is unusually high in antigen genes, and this is suggested as evidence that amino acid sequence polymorphism \textit{per se} has been selected for (McCutchan, Good & Miller, 1989). This is a reasonable argument, although there is a problem in knowing what to compare the data with. Obviously, internal 'housekeeping' proteins are unsuitable for comparison, as these are conserved even among widely divergent taxa.

Hypotheses based on the idea of immune evasion must be considered and tested with respect to specific polymorphisms of particular antigens. Even in general consideration it should be noted that there are different types of polymorphisms, resulting from different molecular mechanisms (e.g. nucleotide substitution, deletion, unequal crossing over within a repetitive sequence at meiosis, intragenic recombination within a non-repetitive sequence). At the protein level, it is possible to grossly differentiate between:

(a) Regions of 'microheterogeneity' (i.e. a clustering of substitutions within a short sequence of amino acids). These are not spectacular at the molecular level, but may be adequate to cause a high degree of
epitope polymorphism, as described at the Th2R and Th3R epitopes of CSP (the circumsporozoite protein: Good et al. 1988; Lockyer, Marsh & Newbold, 1989). At such regions, there may be a large number of allelic variants, each differing by only one or a few amino acids.

(b) Large domains or 'blocks' of sequence accounting for a sizeable proportion of a gene, of which there are only two, or a few allelic variants, each differing at most residues throughout the sequence (e.g. domains flanking the repeat sequence in the S-antigen: Brown et al. 1987; and dimorphic sequences in the precursor to the major merozoite surface proteins: section 1.5.3.). These polymorphisms are spectacular at the molecular level, but even if they contain important epitopes, the number of variants may be too low to confer great advantage to the parasite in a hyperendemic or holoendemic area, in which individuals may be infected several times each year. It has to be considered that, even if these polymorphisms have adaptive significance, this may be unrelated to immune responses.

Three polymorphic proteins have been studied in this thesis, and a outline of their known characteristics is given here.

1.5.3. The Precursor to the Major Merozoite Surface Proteins - MSP1 (or PMMSA, MSA-1, PSA, gp195, p190)
This protein is synthesised by blood-stage and liver-stage schizonts (Holder & Freeman, 1982; Szarfman et al., 1988), and is attached to the outer face of the parasite membrane by a glycosylphosphatidylinositol anchor (Haldar, Ferguson & Cross, 1985). It is coded by a single gene locus, on chromosome 9 (numbering from the smallest upwards) of the 14 chromosomes separated by pulse field gel electrophoresis (Kemp et al., 1987). Each haploid parasite therefore produces one allelic type of the protein, in the size range 185-205 kD (McBride, Newbold & Anand, 1985; Howard, R.J. et al., 1986; McBride & Heidrich, 1987).

The protein is cleaved into at least 4 smaller fragments (approximate sizes 80-83, 28-31, 36-45, and 40-45 kD) shortly before merozoite release from schizonts (Figure 3). These fragments are located at the surface of free merozoites (Holder et al., 1987), where at least two of them are involved in a non-covalently associated complex (McBride & Heidrich, 1987). There is evidence that the C-terminal 40-45 kD fragment is further cleaved into 15-18 and 16-19 kD fragments on the merozoite surface (Holder et al., 1987; McBride & Heidrich, 1987). The 16-19 kD fragment is retained by the merozoite after invasion, and epitopes on this fragment are recognised by indirect immunofluorescence on early ring stages (Blackman et al., 1990). The N-terminal 80-83 kD fragment is apparently shed from the merozoite surface around the
Figure 3. Schematic representation of the natural processing of the MSP1 molecule into smaller fragments, which are incorporated on the merozoite surface (Holder, 1988). Polymorphism in MSP1 is also shown schematically by dividing the sequence into 17 domains or 'blocks', each of which is 'conserved' (unshaded), 'semi-conserved' (diagonal lines), or 'polymorphic' (fully shaded), according to Tanabe et al. (1987).
time of invasion, and may be recovered from in vitro culture supernatant (Holder et al. 1985). The fate of the other fragments is unknown as they are not uniquely recognised by any MAb (Holder, 1988). Although the processing scheme has been described independently by different workers (Lyon et al. 1986; Holder et al. 1987), there is a report from one laboratory claiming that the 200 kD precursor remains unprocessed on the surface of merozoites (Pirson & Perkins, 1985).

The function of MSP1 is unknown, although the location of the processed fragments at the surface of merozoites suggests a possible role in the attachment to or invasion of erythrocytes. One study suggests that MSP1 binds to glycophorin molecules on the surface of human erythrocytes, although it implied that MSP1 was on the merozoite surface in an unprocessed form (Perkins & Rocco, 1988).

Immunisation with purified MSP1 has induced partial protection to challenge infection in Saimiri sciurus monkeys (Perrin et al. 1984; Hall et al. 1984) and apparently complete protection in Aotus lemurinus monkeys (Siddiqui et al. 1987). The homologue of MSP1 in P. yoelii, a 230 kD molecule, induces a species-specific immunity in mice (Holder & Freeman, 1981). Taken overall, these results have encouraged research on MSP1, in the hope that it may contain epitopes inducing protection in humans. It has been shown that T and B cell epitopes are
present at particular domains of the molecule (Sinigaglia et al. 1988 & 1990), and are recognised by individuals in malaria endemic areas (Riley et al. 1991; Früh et al. 1991).

Allelic polymorphism of MSP1 has been studied by comparing nucleotide sequences from different P.falciparum isolates and clones (Tanabe et al. 1987). In a comparison of 4 allelic sequences it was determined that particular domains of the gene are 'polymorphic' (less than 40% homology between sequences at the amino acid level), others are 'conserved' (over 87% homology between sequences), and others are intermediate or 'semi-conserved'. This variation in the degree of polymorphism at different domains has been schematically outlined (Tanabe et al. 1987) by dividing the sequence into 17 domains or 'blocks' (Figure 3).

Additional allelic sequences have demonstrated that sequence polymorphism in domain 2 is very high, with a different sequence described from almost every allele investigated (Mackay et al. 1985; Holder et al. 1985; Weber et al. 1986; Cheung et al. 1986; Howard, R.F., Ardeshir & Reese, 1986; Tanabe et al. 1987; Certa et al. 1987; Peterson et al. 1988 a,b; Chang et al. 1988; Scherf, Barbot & Langsley, 1989). Domain 4 has at least 3 alternative sequences (Tanabe et al. 1987; Certa et al. 1987). All the polymorphic and 'semi-conserved' domains downstream of domain 5 have essentially only 2
alternative sequences, very few sites having a third alternative amino acid, and have therefore been termed 'dimorphic' (Tanabe et al. 1987). A study of several isolates with sequence-specific oligonucleotide probes (Peterson et al. 1988a) has confirmed this pattern of polymorphism, and supports the hypothesis that intragenic recombination events have occurred at sites within domains 2-5 producing a range of observed recombinant alleles (Tanabe et al. 1987). Further evidence of intragenic recombination is documented in Chapter 6.

Polymorphism of MSP1 was originally characterised using a panel of monoclonal antibodies (MAbs), some of which recognise polymorphic epitopes of the molecule (McBride et al. 1985). A large number of MAbs have subsequently been produced in different laboratories (Holder, 1988). Many of these MAbs recognise naturally processed products of MSP1, allowing the epitope location to be approximately determined (Lyon et al. 1986; Holder et al. 1985; McBride & Heidrich, 1987). The epitopes of some MAbs are approximately known as they recognise recombinant fusion proteins representing a known portion of the total sequence (Lyon et al. 1987; Gentz et al. 1988). The probable site of some epitopes can be precisely determined by a correlation of serological and sequence polymorphisms among different parasite isolates or clones (Chapter 6).
1.5.4. The Second Merozoite Surface Protein - MSP2 (or GYMSSA, MSA-2, gp35-56)

MSP2 is synthesised in blood-stage schizonts and incorporated on the schizont surface and the surface of free merozoites (Stanley, Howard & Reese, 1985; Miettinen-Baumann et al. 1988), probably via a glycosylphosphatidylinositol membrane anchor (Smythe et al. 1988). Its size varies among parasite clones, in the range 35-56 kD (Stanley, Howard & Reese, 1985; Clark et al. 1989; Fenton et al. 1989), and is considerably larger than that predicted on the basis of deduced amino acid sequences (262-347 amino acid residues: Smythe et al. 1990; Thomas et al. 1990; Fenton et al. 1991; Smythe et al. 1991) presumably due to glycosylation (Stanley et al. 1985; Smythe et al. 1988; Clark et al. 1989).

The function of MSP2 is not known. No challenge immunisation experiments have been attempted, and no homologues of the molecule are yet described from other Plasmodium species. However, it is recognised by sera, and T lymphocytes, from adults in endemic areas (Smythe et al. 1990; Rzepczyk et al. 1990), and is immunogenic in rabbits (Miettinen-Baumann et al. 1988), and monkeys (Stanley et al. 1985). Several murine monoclonal antibodies have been produced against MSP2, at least 3 of which have been demonstrated to inhibit merozoite invasion of erythrocytes in vitro (Epping et al. 1988;
Figure 4. Schematic representation of allelic polymorphism in MSP2. The regions are numbered according to Thomas et al. (1990). The sequences in regions 1 and 4 are very highly conserved. Region 2 consists of polymorphic repeat sequences, and region 3 consists of polymorphic or 'dimorphic' non-repetitive sequences.
Clark et al. 1989; Ramasamy et al. 1990).

All the available MAbs appear to recognise polymorphic epitopes, and genetic analyses have suggested that the protein is coded by a single locus at which there is considerable allelic polymorphism (Fenton et al. 1989). The locus has been mapped to chromosome 2 (locus Ag 513: Kemp et al. 1987). Comparisons of gene sequences from different isolates and clones have revealed the structural basis of the allelic polymorphisms (Smythe et al. 1990, 1991; Thomas et al. 1990; Fenton et al. 1991). The 43 N-terminal and the 74 C-terminal amino acid residues are very highly conserved, whereas there is extensive sequence diversity within the central region of the gene (Figure 4). The epitopes of two MAbs have been precisely determined by synthetic peptide binding assays (Epping et al. 1988; Saul et al. 1989; Ramasamy et al. 1990), and the probable locations of other epitopes have been deduced by correlations between serology and sequence polymorphisms (Fenton et al. 1991).

1.5.5. An Exported Protein – Exp-1 (or CRA, p23)

Exp-1 is a protein of 23 kD in all asexual erythrocytic stage parasites. Indirect immunofluorescence studies, using MAb 5.1 (Hall et al. 1983), or immune human antibodies affinity purified on extracts of an expression clone containing Exp-1 (Coppel et al. 1985),
demonstrate that in ring stages the protein is localised at the parasite surface. At the trophozoite stage, the protein is also localised in the parasitophorous vacuole membrane and in the membranes of vesicles within the erythrocyte cytoplasm, as demonstrated by immuno-electron microscopic labelling (Simmons et al. 1987). The location suggests that the molecule may be involved in intracellular transport mechanisms.

The deduced amino acid sequence from cDNA clones contains a region typical of transmembrane sequences of integral membrane proteins (Hope et al. 1985; Coppel et al. 1985). Sequences from genomic DNA show that the gene contains two introns (Simmons et al. 1987), and hybridisation studies have located the gene on chromosome 11 (Kemp et al. 1987).

The sequence is very conserved, as demonstrated by a comparison of five allelic sequences from different isolates or clones. Only two sites of amino acid substitution were revealed, one of which is within the epitope recognised by MAb 5.1 (Simmons et al. 1987). MAb 5.1 cross reacts with the sporozoite surface by indirect immunofluorescence (Hope et al. 1984), as a result of sequence homology between the 5.1. epitope of Exp-1 and the CSP repeats (Hope et al. 1985; Coppel et al. 1985). The amino acid substitution within the 5.1 epitope explains the previously described isolate specificity of the MAb (McBride, Walliker & Morgan, 1982; Hall et al.
Parasites with aspartic acid at position 136 of Exp-1 are recognised by MAb 5.1, whereas parasites with glycine at that position are not recognised (Simmons et al. 1987).

No homologues of Exp-1 have been described from other Plasmodium species, and no challenge immunisation experiments have been undertaken to evaluate whether an immune response against the protein has a protective effect.

1.5.6. The Significance of Merozoite Surface Protein Polymorphism

As discussed in sections 1.3.1. and 1.5.2., allelic polymorphisms of proteins are not always maintained by selection, and even polymorphisms of merozoite surface components may be adaptively neutral. However, the merozoite surface interacts with the host in at least two very important ways, and it has been suggested that merozoite surface protein polymorphism may be involved in either of these interactions:

(i) Merozoite surface proteins induce a degree of acquired immunity to reinfection (section 1.5.3.). Antibodies against the merozoite surface could theoretically inhibit parasite multiplication by either, (i) cross-linking and agglutinating merozoites before they have dispersed from a segmenting schizont when the
erythrocyte membrane ruptures, (ii) agglutinating free merozoites, or (iii) specifically blocking a merozoite surface receptor molecule so that it cannot interact with an uninfected erythrocyte. The third mechanism might be achieved with lower antibody concentration, as cross-linking and agglutination would not need to occur.

The hypothesis that polymorphisms in target antigens are maintained in parasite populations to avoid the acquired immune responses of the local host population has been considered in general in section 1.5.2. MSP1 is certainly a target of protective immunity (section 1.5.3.), while MSP2 awaits further study (section 1.5.4.), and at present the immune evasion hypothesis is widely advocated to explain at least some of the extensive sequence polymorphisms in these two proteins. Testing this hypothesis experimentally would be expensive, involving immunisation of monkeys or human volunteers with a defined antigen, and challenge with a range of different P. falciparum clones with defined allelic types of that antigen. Such controlled experiments have not been undertaken. An alternative, indirect approach has been taken in this thesis. Parasites are sampled from a human population over a period of time, to test whether rare antigenic variants increase in frequency relative to alternative common variants, due to frequency-dependent selection by acquired variant-specific immunity (chapter 7).
Merozoites interact with erythrocytes in a complex invasion process. Merozoites attach to the erythrocyte surface in any orientation, and then re-orient so that the apex (containing the rhoptry organelles) is in contact with the erythrocyte. The vesicular contents of the rhoptries are discharged, fusing with the erythrocyte membrane, which undergoes a wave of deformation. The membrane invaginates around the merozoite, which is thereby internalised into the erythrocyte, surrounded by this membrane (termed the parasitophorous vacuole membrane).

Merozoite surface components are involved in a specific receptor-ligand interaction with the erythrocyte surface during this complex invasion process. In vitro studies have demonstrated that sialic acid residues of erythrocyte-surface glycophorin A, and possibly also glycophorin B, are required for optimal invasion of erythrocytes by merozoites (Pasvol et al. 1982 a,b; Facer, 1983; Hermentin et al. 1987). However, some isolates of *P. falciparum* are more dependent than others on the presence of this sialic acid. For example, Thai-Tn parasites invaded neuraminidase-treated erythrocytes (sialic acid removed), and Cad erythrocytes (altered sialic acid conformation), with a greater relative efficiency than Camp parasites (Mitchell et al. 1986; Hadley & Miller, 1988). In addition, 7G8 parasites invaded neuraminidase-treated erythrocytes, En(a-)
erythrocytes (lacking glycophorin A), and MkMk erythrocytes (lacking glycophorin A and B), with a greater relative efficiency than Camp parasites (Hadley et al., 1987; Hadley & Miller, 1988). It is thought that invasion of sialic acid-deficient erythrocytes depends on the specific recognition of another erythrocyte surface ligand. Treatment of MkMk or En(a−) cells with trypsin reduced the residual invasion of all *P. falciparum* isolates tested to approximately 5% of that of normal control cells (Mitchell et al., 1986; Hadley & Miller, 1988). Hence, it was proposed that there are two alternative invasion pathways, one involving sialic acid on glycophorin A (and possibly glycophorin B), and another involving a trypsin-sensitive non-sialic acid residue, and parasites differ in the extent to which they rely on each pathway.

Polymorphism of the merozoite surface proteins MSP1 and MSP2 might explain the observed differences among parasite clones in dependence on particular erythrocyte ligands. This was suggested by Tanabe et al. (1987) in their paper describing the 'dimorphism' in allelic sequence throughout most of the MSP1 gene, but the hypothesis has not been tested.

Erythrocyte sialic-acid deficiency is extremely rare among humans, and has never been detected in malaria endemic areas, although individuals lacking glycophorin B are not uncommon in some parts of Africa (the S-s-U-
phenotype: Mourant, Kopec & Domaniewska-Sobczak, 1976; Issitt, 1981). Therefore, it is difficult to understand how some parasites could have an advantage in recognising an alternative ligand, sufficient to maintain polymorphism in a receptor molecule. In the absence of further evidence on the receptor-ligand interaction(s), an epidemiological approach has been taken to test for correlations between erythrocyte surface polymorphisms (ABO, MN, and Ss blood groups) and merozoite surface polymorphisms (MSP1 and MSP2) in malaria patients (Chapter 8).
1.6. Aims and Outline of the Thesis

The thesis begins by addressing the population genetic structure and epidemiology of *P. falciparum*. The three polymorphic proteins (MSP1, MSP2, and Exp-1) are used as genetic markers for examining the *P. falciparum* population structure within communities (Chapter 3), individuals (Chapter 4), and households (Chapter 5). The second main aim of the thesis is to analyse particular allelic polymorphisms themselves, to consider whether there are adaptive constraints on intragenic recombination (Chapter 6), relative frequencies of different variants in a population (Chapter 7), or associations with patients of different blood group phenotype (Chapter 8).

From the outset, it is important to establish whether *Plasmodium falciparum* can be considered to be generally an outbreeding species, as this affects the interpretation of all other results. Previously, limited data from characterisation of wild isolates was consistent with an outbreeding population structure, although a few anomalous published results had caused confusion leading to a hypothesis that populations were predominantly clonal. In Chapter 3, extensive new data is analysed, to determine whether there is random assortment between three unlinked loci.
Multiple-clone infections were previously noted to be common in endemic areas. Taking a classical view, this is a prerequisite for genetic recombination, as a mosquito may thereby acquire two or more gametocyte clones. In Chapter 4 a technique is described for distinguishing up to four clones in a particular isolate without in vitro parasite cloning. Effects of season, parasitaemia, patient's age, and disease severity are tested for. A statistical analysis on gene diversity within isolates is performed, in an attempt to explain whether multiple-clone infections arise by a mechanism of superinfection alone, or whether there is a significant incidence of non-identical sibling parasite clones acquired from single mosquitoes.

Chapter 5 examines the importance of a previously unquantified mode of malaria transmission. Mosquito pre-feeding probing activity, and interrupted blood feeding, had been described as means whereby a single mosquito may infect more than one person on a single occasion. The effect of these activities on infection is studied indirectly by measuring the occurrence of identical *P. falciparum* genotypes in children presenting simultaneously with malaria from the same household, compared with children from different households.

The structural basis for serological polymorphism of MSP1 is studied in Chapter 6. Sequence data is presented in support of the hypothesis of intragenic recombination.
towards the 5' end of the gene. The probable epitope determinants of particular monoclonal antibodies are assigned on the basis of sequence-serology correlations. Associations between epitopes at different domains of MSP1 are examined statistically, for parasites in each of three different countries, and are discussed with reference to the intragenic recombination hypothesis.

Allelic polymorphisms at particular protein domains might be maintained in a population if an acquired immune response against one variant of a protein is ineffective against an alternative variant. In such a situation frequency-dependent selection may favour rare variants in the population. Frequencies of allelic variants of three antigenic proteins (MSP1, MSP2, and Exp-1) were recorded at different times over a period of 7 years in a hyperendemic area. The results are presented in Chapter 7, and discussed with reference to the hypothesis of frequency-dependent selection.

Merozoite surface proteins may be involved in the receptor-ligand interaction at the erythrocyte surface prior to invasion. It has been suggested that different allelic types of a merozoite surface protein may recognise different molecules or variants of a molecule on the erythrocyte surface. Different erythrocyte phenotypes may be preferentially invaded by different parasites. Blood group phenotypes (ABO, MN, and S) and merozoite surface antigen polymorphisms (MSP1 and MSP2)
from the same patients are analysed in Chapter 8, in order to test for statistical associations which may reveal dependent relationships.

Each of the chapters has discrete conclusions, but it is useful to briefly consider them together. Chapter 9 therefore summarises the main conclusions of the thesis, and considers its contribution to malariology.
Chapter 2. Materials and Methods

2.1. Isolation and Culture of Plasmodium falciparum

2.1.1. Study Areas, Clinics, and Laboratories

The Gambia, on the coast of West Africa, is within the 'Northern Guinea Savannah' ecological zone. The country experiences approximately 500-1000 mm of rain annually between May and October, resulting in seasonal hyperendemic malaria transmission between July and November, during which the point prevalence of *P. falciparum* in the human population is typically above 50 % (Greenwood *et al.*, 1987).

Patients presented with malaria to the outpatients departments of the Medical Research Council (MRC), Fajara, and the Royal Victoria Hospital, Banjul. The majority of patients lived within an urban/periurban area of approximately 7 km radius centred on Serekunda, 5 km south of Fajara (Figure 5). Approximately 50-100,000 people live in this area. Patients from Sarakunda village on the north bank of the river, and from Brefet village on the south bank of the river (Figure 5), presented to village clinics run by MRC staff. Parasites from all isolates were cultured within 6 hours of sample collection, in the MRC laboratories at Fajara or Farafenni.
Figure 5. A map of the western part of The Gambia showing the locations of the urban/periurban region and the two villages from which isolates were collected.
Nigeria incorporates several ecological zones, and hence areas of different malaria endemicity. Ibadan, a city of approximately 2 million people in Oyo State, is within the 'Forest' zone in the southwest of the country. Ibadan typically experiences 1000-1500 mm of rain throughout the year, mainly during April-July and September-October. Although malaria transmission is at a peak during and following these rainy seasons, malaria is holoendemic, as transmission persists during the rest of the year (Salako et al. 1990).

Patients presented to the Outpatients Department of the University College Hospital, Ibadan. All of the patients recruited for this study lived within Ibadan. Parasites from all isolates were cultured within 6 hours of blood collection, in the laboratories assigned to the Department of Pharmacology and Therapeutics, in the Institute of Postgraduate Research of University College Hospital.

Isolates from Brazil were collected from different locations within the Amazon basin by workers at Instituto Evandro Chagas da Fundação Serviços de Saúde Pública, Belém, in the research programme led by Dr. Virgilio Do Rosario. The *in vitro* culture of these isolates was performed by Dr. Rosario and colleagues in Brazil.
2.1.2. Diagnostic Identification of Plasmodium falciparum

Air-dried thick blood films were stained in Field's or Giemsa staining solutions.

(i) Field's stain: The blood film was dipped in a staining jar containing Solution A (0.16 % medical methylene blue, 0.1 % Azure 1, 1.0 % anhydrous Na$_2$HPO$_4$, 1.25 % KH$_2$PO$_4$, in distilled water - dH$_2$O) for 3 seconds, and rinsed in a jar of dH O for 3 seconds. The film was then dipped in a jar containing Solution B (0.2 % eosin, 1.0 % anhydrous Na$_2$HPO$_4$, 1.25 % KH$_2$PO$_4$, in dH$_2$O) for 1 second, and rinsed in a jar of dH$_2$O for 5 seconds, before air drying.

(ii) Giemsa Stain: Giemsa stock solution was prepared by partially dissolving 3.8g Giemsa powder (Gurr, BDH Ltd.) in 250 ml methanol, then adding 250 ml glycerol and approximately 50 glass beads. The mixture was kept in a dark bottle, and shaken vigorously for a few minutes each day for the following 3 days to ensure that the Giemsa powder had completely dissolved. For staining, an aliquot of this stock solution was diluted 1:9 with buffer solution pH 7.2 (0.2 % Na$_2$HPO$_4$, 0.14 % KH$_2$PO$_4$), and 5 ml poured onto each blood film, and left for 30 minutes. The stain was then washed off with buffer solution pH 7.2, and dried.

*Plasmodium falciparum* parasites were morphologically
distinguished from other \textit{Plasmodium} species, using the WHO Bench Aids for the Diagnosis of Malaria (WHO, 1985). The prevalence of \textit{P.malariae} and \textit{P.ovale} in the study area was very low, and isolates were only collected from patients with \textit{P.falciparum} parasitaemia unmixed with either of these species.

2.1.3. Blood Collection from Patients

A 0.3 ml blood sample was obtained from each patient by either of two methods:

(i) The middle finger was cleaned with a cotton wool swab impregnated in 70 \% ethanol, and wiped with dry cotton wool. A sterile lancet was used to puncture the skin, slightly to the lateral side of the fingertip, and the first drop of blood was removed with dry cotton wool. Drops of blood were collected in a microtainer containing ammonium heparin (Becton Dickinson), which was numbered with a water-insoluble marker to identify which individual the sample was from.

(ii) The sample was aliquoted from a 5-10 ml venous sample, which was collected by syringe and transferred to a 15 ml polystyrene tube (Falcon, Becton Dickinson) containing ammonium heparin (100 units / 1 ml blood).

2.1.4. Washing Erythrocytes
Each blood sample was transferred to a 15 ml polystyrene centrifuge tube (Falcon), and spun at 2000 rpm for 5 minutes. Plasma was removed, and the cells were resuspended in at least 5 ml sterile phosphate buffered saline pH 7.3 (PBS). The tube was spun at 2000 rpm for 5 minutes and the supernatant removed. Cells were resuspended in PBS and centrifuged as before, and the resulting cell pellet was resuspended in complete RPMI culture medium (see 2.1.5.) at approximately 3 % haematocrit.

2.1.5. Malaria Culture Medium

5.94g HEPES powder (Sigma), 2g NaHCO₃, and one sachet of RPMI 1640 powder (Flow Laboratories) were dissolved in double distilled water (ddH₂O) to make a 1 litre solution. The solution was adjusted to pH 7.2 by the addition of 1 M NaOH, and filtered at 0.23 μ into 200 ml sterile bottles or culture flasks. 10 % volume human serum (blood group AB) was added.

2.1.6. Freezing and Thawing of Culture-adapted P. falciparum Clones

Clones T9/96 and T9/101 were originally obtained by limiting dilution in vitro cloning of parasites isolated from a Thai patient (Thaithong et al., 1984). Frozen
stabilates of the clones were obtained from the WHO Registry of Standard Malaria Strains, University of Edinburgh. Each ampoule was removed from liquid N₂, and thawed to room temperature. The infected blood was transferred to a 15 ml graduated polystyrene tube. The volume (V) was measured, and 0.2 x V of 12 % NaCl solution was added dropwise using a syringe, while agitating the cells by tapping and shaking the tube. Then 10 x V of 1.6 % NaCl solution was added dropwise, being mixed with the blood cells by shaking the tube. This mixture was spun at 2000 rpm for 5 minutes, supernatant was removed, and 10 x V of 0.9 % NaCl, 0.2 % dextrose was added in a similar manner as in the previous steps. The mixture was spun at 2000 rpm for 5 minutes, supernatant was removed, and the cells were resuspended in complete RPMI medium at approximately 3 % haematocrit for culture.

Ring-stage cultured parasites were cryopreserved for future use by the following method: Parasitised erythrocytes were centrifuged, and culture medium removed. The volume of packed cells was measured, and an equal volume of cryopreservative (28 % glycerol, 3 % sorbitol, 0.65 % NaCl) was mixed with the cells by adding dropwise in a similar manner as described during the thawing process. 0.3-0.5 ml aliquots of this mixture were transferred to polypropylene ampoules (Nunc), and placed directly into liquid N₂ for cryopreservation.
2.1.7. In Vitro Culture

At the beginning of culture most P. falciparum parasites were ring stages (< 24 h old trophozoites), either because (i) more mature parasites are not normally isolated from peripheral blood of a patient, as they sequester in blood vessels of internal organs, or (ii) more mature parasites do not survive the cryopreservation procedure used on in vitro culture-adapted clones (section 2.1.6.).

Parasites freshly isolated from patients were grown in approximately 3 ml culture volumes at 3 % haematocrit, in covered petri dishes (Falcon, 35 x 10 mm) under reduced O₂ tension in a candle jar at 37°C (Trager & Jensen, 1976). They were cultured for 24-48 hours, until the majority had matured to schizonts. After 24 hours culture, the maturity of the parasites was assessed morphologically on a thin blood film, stained with Giemsa (as described for thick blood films in 2.1.2.). The schizont-stage cultures were harvested, and those not yet matured were returned to culture and harvested at the appropriate later time.

When required (for DNA preparation from wild isolates), larger volumes (50-100 ml) were cultured in 175 cm² culture flasks (Nunc), which were placed in a candle jar for culture at 37°C.

Culture-adapted parasite clones were grown for
several asexual cycles (approximately 2-3 weeks continuous culture after thawing from liquid N\textsubscript{2}), to yield a parasitaemia of at least 5 % (which was desirable for DNA preparation). Approximately 5 ml volumes were grown in 25 cm culture flasks, filled with a cylinder gas mixture (1 \% O\textsubscript{2}, 3 \% CO\textsubscript{2}, 96 \% N\textsubscript{2}) which allowed more efficient parasite growth than the candle jar method. The flasks were incubated at 37\textdegree C for culture.

2.1.8. Parasite Concentration for DNA Extraction

Concentrated preparations of parasites were prepared from \textit{in vitro} culture by either of two methods:

(i) A method of enrichment of schizont-infected erythrocytes \cite{Pasvol et al. 1978}. Parasitised erythrocytes were pelleted by centrifugation at 2000 rpm for 5 minutes, washed in complete RPMI 1640 medium and pelleted again at 1300 rpm for 10 minutes in a 15 ml or a 50 ml graduated tube (Falcon). The volume of the pellet (V) was measured from the tube gradations, and the pellet was resuspended in 1.4 x V of complete RPMI 1640 medium (to achieve a total volume of 2.4 x V).

An equal volume of Plasmagel (Laboratoire Roger Bellon, Neuilly Sur Seine, France) was added to this suspension, and the tube was incubated at 37\textdegree C for 30 minutes without disturbing. After incubation, the more
mature parasite stages were suspended in the Plasmagel mixture, and the ring-infected and uninfected erythrocytes had sedimented to the bottom.

The suspension was transferred to another centrifuge tube, and erythrocytes (approximately 80% of which were schizont-infected) were pelleted by centrifugation at 2000 rpm for 5 minutes. The cells were resuspended in < 1.0 ml PBS and microcentrifuged in a 1.5 ml polypropylene tube for 30 seconds. Supernatant was removed and the cell pellet frozen immediately at -70°C.

(ii) A method of releasing parasites from lysed erythrocytes. Cultured parasite-infected erythrocytes were pelleted by centrifugation at 2000 rpm for 5 minutes. The volume of the pellet was measured, and the pellet was resuspended in an equal volume of RPMI 1640 medium. Saponin was added (to a final concentration of 0.1%), and the mixture was incubated on ice for 5 minutes. The erythrocytes lysed, and the released parasites were pelleted by centrifugation at 2400 rpm for 10 minutes, resuspended in < 1.0 ml PBS and microcentrifuged in a 1.5 ml tube for 30 seconds. The supernatant was removed, and the pellet frozen at -70°C until required for DNA extraction.
2.2. Amplification, Cloning, and Sequencing of Parasite DNA

2.2.1. Parasite Genomic DNA Preparation

A 5-20 μl parasite pellet which had been stored at -70°C was allowed to thaw to room temperature, and resuspended in 400 μl of 150 mM NaCl, 25 mM EDTA pH 8.0 solution. 10 μl of Proteinase K (Sigma) solution (10 mg/ml), and 10 μl of 10% sodium dodecyl sulphate (SDS) solution were added. The contents of the tube were gently mixed, and the tube was sealed and incubated at 37°C overnight.

200 μl phenol and 200 μl chloroform (both previously equilibrated with TE buffer: 10 mM Tris base, 1mM EDTA), were added to the parasite lysate. The tube was vortexed for 15 seconds and microcentrifuged for 5 minutes. After spinning, the aqueous phase containing nucleic acids was on the top, and the organic phase containing proteins and lipids was on the bottom. The aqueous phase was removed to another tube, its high viscosity requiring care not to transfer the interface components.

This process of phenol-chloroform extraction was repeated twice. Then extraction was performed 3 more times using 400 μl chloroform, to remove traces of phenol. Finally, 400 μl of ether was added to the aqueous
DNA solution, and the mixture was vortexed for 15 seconds. After spinning in a microcentrifuge for 30 seconds, the aqueous phase was on the bottom. Most of the ether was removed by pipette, and traces were allowed to evaporate by leaving the tube standing open for 10 minutes. Approximately 400 µl of aqueous DNA solution remained, depending on the efficiency of transfer during the extraction stages.

To precipitate the DNA, 800 µl absolute ethanol (AnalR) and 120 µl 3M sodium acetate solution pH 5.0 were added, and the mixture was inverted a few times before placing at -70°C for 30 minutes. The tube was microfuged in a 4°C cold room for 20 minutes to pellet the DNA. The supernatant was carefully removed, 1.2 ml 70% ethanol was added to wash the DNA, and the tube was microfuged in the cold room for 10 minutes. All of the supernatant was removed, using a glass micropipette to remove the visible traces, and the open tube was vacuum dried to allow invisible traces of ethanol to evaporate.

The DNA pellet was resuspended in 40 µl of TE buffer. 1 µl of this solution was run on one lane of an agarose mini gel (as described in 2.2.3.), to estimate the DNA content by comparison with a control lane containing 500 ng Hind III digested phage lambda.

2.2.2. DNA Amplification
A short DNA sequence (approximately 350 base pairs) from the MSP1 gene was amplified from parasite genomic DNA by the polymerase chain reaction (PCR) method first described by Saiki et al. (1985). Amplification was primed by two synthetic oligonucleotides (a 24-mer: 5'-TGAAGGAAGTAAGAAAACAATTGA-3', and a 28-mer: 5'-TCTAATTCAAGTGGATCAGTAAATAAAC-3', synthesised by the Oswel DNA Service, Department of Chemistry, University of Edinburgh) designed to hybridise to the 5' ends of the coding and non-coding DNA strands at each end of the sequence. 10 μl of a 1 μM solution of each amplification primer, 2 μl of a 10 mM solution of each dNTP, 100ng of parasite genomic DNA, and 10 μl of 10x Cetus Taq buffer (500 mM KCl, 100mM Tris-HCl pH 8.3, 15 mM MgCl2, 0.1 % w/v gelatin), were added to a 1.5 ml tube, and made up to a volume of 98 μl with sterile double distilled water. A 1 unit/μl solution of Taq DNA polymerase (Perkin Elmer Cetus) was thawed from - 20°C storage, and 2 μl added to the amplification mixture in the tube. The mixture was overlaid with 100 μl sterile mineral oil (Sigma). The tube was sealed, and a small hole was peirced in the lid using a sterile needle.

The tube was placed in a well of an intelligent heating block (Hybaid), which was programmed for the following routine: First, 95°C for 5 minutes (for genomic DNA denaturation). Then, 10 seconds at 45°C (for annealing of primers), 2 minutes at 70°C (for DNA
polymerisation), and 10 seconds at 92°C (for denaturation of double-stranded amplified product), with this cycle repeated twice. Then 27 more cycles, with the annealing step at 40°C instead of 45°C, to maximise the yield of product. The final PCR product, cooled to room temperature, was double stranded.

2.2.3. DNA Fractionation by Agarose Gel Electrophoresis

Agarose gels for fractionation and purification of DNA were made in TAE buffer (40 mM Tris base, 95 mM glacial acetic acid, 1 mM EDTA pH 8.0). An appropriate weight (0.8 - 1.0 g) of agarose powder (Sigma) was added to 100 ml TAE, stirred and heated to boiling point until the agarose melted. The mixture was cooled to approximately 55°C in a water bath. 20 ml of the mixture was carefully poured onto a flat 60 x 90 mm perspex plate. A perspex mould with 4 mm wide teeth was positioned to form the sample wells. The mixture solidified within 20 minutes, forming a gel at room temperature. The gel was placed in a horizontal electrophoresis tank containing 150 ml TAE with 15 μl of ethidium bromide (5 mg/ml).

DNA samples for electrophoresis were made up to 10 μl or 20 μl in double distilled water, with 10 % volume formamide dye mix (0.025 % bromophenol blue, 10 % Ficoll, in dH₂O), for loading in each well. The samples were
electrophoresed at a constant 150 volts, until the DNA had fractionated sufficiently for purification (approximately 40 minutes). Under ultraviolet incident light, the required DNA fraction was visualised and excised using a sterile razor blade.

Agarose gels for estimation of DNA fragment size and/or concentration were prepared similarly, except that they were made in TBE buffer (90 mM Tris base, 90mM boric acid, 2mM EDTA pH 8.0). and samples were electrophoresed in TBE buffer.

The following molecular weight markers were used (Boehringer Mannheim, 500ng per track): lambda DNA.HindIII (fragment sizes, 23130, 9416, 6682, 4361, 2322, 2027, 564 base pairs), or pBR328 DNA.BglI + pBR328 DNA.HinfI (fragment sizes 2176, 1766, 1230, 1033, 653, 517, 453, 394, 298, 298, 234, 234, 220, 154, 154 base pairs).

2.2.4. Purification of DNA

Excised agarose containing DNA (from 2.2.3.) was cut into pieces approximately 1 mm square, using a sterile razor blade. The pieces were placed in a 1.5 ml tube and spun in a microfuge for 5 seconds, to determine the approximate volume of agarose. To dissolve the agarose, 2.5 to 3 volumes of NaI stock solution (Geneclean, Bio 101 Inc.) was added, and the tube was transferred to a 55°
C waterbath incubator for 5 minutes, being removed every one or two minutes for mixing of the contents by inversion.

A tube of insoluble silica matrix stock, termed 'Glassmilk' (Geneclean TM), was vortexed vigorously for 1 minute to resuspend the silica particles. 5 μl of this suspension was mixed into the solution containing DNA, which was placed on ice for at least 5 minutes to allow binding of the DNA to the silica matrix, mixing every 1 - 2 minutes to ensure that the silica particles remained suspended.

The silica matrix with bound DNA was pelleted by spinning the tube in a microfuge for 5 seconds, and the supernatant was discarded. The pellet was resuspended in 500 μl ice-cold 'New Wash' (Geneclean), by digging into the pellet with a pipette tip, while pipetting up and down. After resuspension, the silica matrix with bound DNA was pelleted by spinning the tube in a microcentrifuge for 5 seconds. The supernatant was discarded and the wash procedure was repeated twice. The last traces of 'New Wash' were removed from the tube with a glass micropipette.

The DNA was eluted from the silica matrix by resuspending the pellet in 10 μl TE buffer, and incubating at 55°C for 3 minutes. The tube was then centrifuged for 30 seconds to form a solid pellet, and the supernatant containing the eluted DNA was carefully
removed and placed in a new tube. The pellet was resuspended in 5 μl TE buffer to elute any remaining DNA, and the tube was incubated and spun as before. The supernatant was added to that obtained from the first elution, to give approximately 15 μl of purified DNA solution.

2.2.5. Kinasing Insert DNA

Prior to ligation into the Smal site of the bacteriophage vector, the purified DNA fragment (PCR product) was kinased, to phosphorylate the 5' ends. The kinasing mix consisted of 7 μl purified DNA solution, 1 μl of 10 mM ATP, 1 μl of 10x kinase buffer (500 mM Tris-HCl pH 7.6, 100 mM MgCl₂, 50 mM DTT [dithioerythriol], 1 mM spermidine, 1mM EDTA), and 1 μl T4 polynucleotide kinase (Boehringer Mannheim, 1 unit/μl). The mixture was incubated in a 37°C waterbath for 1 hour. The enzyme was then inactivated by incubating the tube at 65°C for 10 minutes.

2.2.6. Digestion of Bacteriophage M13

Bacteriophage M13 (mp18 or mp19: Yanisch-Perron, Vieira & Messing, 1985) was cut at a single Smal site, to produce blunt ends for ligation. The digestion mixture consisted of 2 μl purified mp18 or mp19 phage
(approximately 1 µg), 2 µl 10x Smal buffer (15 mM Tris-HCl, 15 mM KCl, 6 mM 2-mercaptoethanol, pH 8.5), 15.5 µl sterile ddH₂O, and 0.5 µl Smal enzyme (BCL, 2 units/µl). The mixture was incubated at 25°C for 4 hours. The restriction enzyme was then inactivated by incubating the tube at 65°C for 10 minutes.

2.2.7. Ligation of Insert with Vector

1 µl of the kinased insert DNA and 1 µl of the Smal restricted bacteriophage M13 were run on adjacent lanes of an agarose mini gel to compare molar concentrations of the solutions. The ligation was performed at an insert:vector molar ratio of 3:1 in a final volume of 10 µl. This volume included 1 µl of 10x ligation buffer (500 mM Tris-HCl pH 7.5, 100 mM MgCl₂, 100 mM DTT, 100 mM ATP), and 0.5 µl T4 DNA ligase (BCL, 1 unit/µl). The tube was placed at 15°C overnight. Control mixtures included ligations with a lower insert:vector ratio, and without any insert. The ligated phage mixtures were stored at -20°C until required for transformation of host cells.

2.2.8. Transformation of Host Cells

The ligated M13 vector was introduced into E.coli TG-1 (genotype K12, lac-pro, SupE, thi, hsdD5, F'traD36, pro A+B+, lacI ,lacZ, M15), as follows: A colony of TG-1
was picked and expanded in 5 ml L-broth (10g Difco Bacto Tryptone, 5g Difco Bacto yeast extract, 5g NaCl, per litre pH 7.2), while shaken overnight at 37°C. 1 ml of overnight culture (in stationary growth phase) was added to 50 ml L-broth and expanded for 90-120 minutes at 37°C in a 500 ml conical flask. These cells (in log growth phase) were then chilled on ice for 30 minutes, and centrifuged at 1500 rpm for 15 minutes.

The cells were suspended in half the previous volume (i.e. 25 ml) in ice-cold 100 mM MgCl₂, and harvested again by centrifugation at 1500 rpm for 15 minutes. The cells were resuspended in one twentieth the original volume (i.e. 2.5 ml) in ice-cold 100 mM CaCl₂, and kept on ice for at least 30 minutes. These cells were then competent for transformation.

200 µl of competent cell suspension were transferred to each of 5 tubes. 5 µl of each ligation mixture was added to a separate tube. 1 µl uncut vector DNA (1/50 dilution) and 0.5 µl cut vector DNA were each added to a tube as positive and negative transformation controls, respectively. The mixtures of cells and phage were kept on ice for 30 minutes, and then heat-shocked at 42°C for 2 minutes, before being returned to ice for another 30 minutes.

10 µl and 100 µl aliquots of each transformed cell suspension were added to 12 mm diameter glass bacteriological tubes containing 200 µl untransformed
competent cell suspension. To each of these tubes 10 μl IPTG (Isopropyl-beta-D-thiogalactosidase, 100 mM in ddH₂O), 20 μl X-Gal indicator (5-bromo-4-chloro-3-indoyl-beta galactosidase, 20 mg/ml in diethyl formamide), and 3.0 ml molten BBL top agar (10g Baltimore Biological Laboratories trypticase, 5g NaCl, 6.5g Difco agar, per litre) were added, and the mixture was rapidly spread on plates containing minimal agar medium. The plates were left for 10 minutes until the BBL-broth had set, and were inverted for incubation at 37°C overnight.

The minimal plates had been prepared as follows: 6 g of Bacto Agar (Difco) was added to 300 mls double distilled water, and autoclaved. After cooling to 55°C, 4 mls of a 20% glucose solution, 80 mls of Spizizen salt solution (2g \([\text{NH}_4]_2\text{SO}_4\), 14g K₂HPO₄, 6g KH₂PO₄, 1g trisodium citrate, 0.2g MgSO₄, per litre), and 200 μl of 100 mg/ml vitamin B1 solution were added, and mixed by inverting a few times. Approximately 20 ml was poured into each of 20 plates.

2.2.9. Identification of Recombinants

The insertion of a DNA fragment into the Smal site of the M13mp vectors prevents alpha-complementation and generation of active beta-galactosidase. This allows recombinant phage to be identified by colour development on the plates (white colonies are recombinants, blue are
non-recombinants). White colonies were picked and expanded for phage preparation (section 2.2.10.), and the size of the phage was compared with that from a non-recombinant control, on an agarose mini gel, to confirm the presence of an insert.

2.2.10. Preparation of Single Stranded DNA Templates

Each white colony was picked and transferred to 0.5 ml phage buffer (3g KH$_2$PO$_4$, 7g anhydrous Na$_2$HPO$_4$, 5g NaCl, 1mM MgSO$_4$, 0.1mM CaCl$_2$, 1ml of 1 % gelatin solution, per litre), for storage at 4°C. 10 µl of this phage suspension was added to a 1.5 ml aliquot of a 1:50 dilution of an overnight culture of TG-1 cells. The cells and phage were incubated for 6 hours at 37°C with constant shaking. The cells were spun in a microcentrifuge for 5 minutes, and the supernatant containing the single stranded phage was decanted into a fresh microcentrifuge tube. The supernatant was centrifuged again, to remove any residual cells. To precipitate the phage, 1.2 ml of supernatant was added to 300 µl of 20 % PEG (polyethylene glycol) 6000, 2.5 M NaCl, and stood for at least 15 minutes at room temperature. The mixture was spun in a microfuge for 10 minutes to harvest the phage precipitate. The supernatant was discarded, the tube re-centrifuged for a few seconds, and the residual supernatant removed with a glass
micropipette. The pellet was resuspended in 100 μl TE buffer with 0.1 M NaCl. An equal volume of phenol (saturated with TE) was added. The tube was vortexed for 20 seconds, and stood for 2 minutes, this was repeated twice, and then microcentrifuged for 5 minutes. The aqueous layer was added to an equal volume of chloroform (saturated with TE). This tube was vortexed for 20 seconds, and microcentrifuged for 5 minutes. 80 μl of the aqueous solution was removed, and DNA was precipitated by adding 9 μl of 3M sodium acetate pH 5.0, and 200 μl absolute ethanol, and placing the tube at -70°C for 30 minutes. The DNA was pelleted, washed, and resuspended in 10 μl TE (as in 2.2.1.).

2.2.11. Sequencing Reactions

Dideoxy sequencing was performed using the Sequenase kit (United States Biochemical). The Universal Smal M13 sequencing primer (5'-GTAAAACGACGGCCAGT-3') was annealed to the DNA template, by adding the following to a tube: 1 μl primer (0.5 pmol), 2 μl of 5x sequencing buffer (200mM Tris-HCl pH 7.5, 100 mM MgCl2, 250 mM NaCl), 3 μl DNA template solution (approximately 1 mg, from 2.2.10.), 4 μl sterile double distilled water. The tube was heated to 65°C for 2 minutes, then cooled slowly to <37°C. While this mixture was cooling, 2.5 μl of each termination mixture (containing 8 μM ddGTP, ddATP, ddTTP,
or ddCTP in 80 μM dGTP, 80 μM dATP, 80 μM dTTP, 80 μM dCTP, 50 mM NaCl) was aliquoted into a separate tube, labelled G, A, T, or C, and placed in a 37°C waterbath.

The sequencing reaction was initiated by adding the following to the tube containing 10 μl of the annealed DNA mixture: 1 μl DTT (0.1M), 2 μl labelling mix (diluted 1/5 from stock, in ddH₂O), 0.5 μl 35S-dATP (Amersham, 10 μci / μl), 2 μl Sequenase™ enzyme (diluted 1/8 from stock, in 10 mM Tris-HCl pH 7.5, 5 mM DTT, 0.5 mg/ml BSA). This mixture was incubated at room temperature for 5 minutes. Then, a 3.5 μl aliquot was added to each of the four tubes labelled G, A, T, and C, mixed and incubated at 37°C for 5 minutes. The sequencing reactions were stopped by adding 4 μl of formamide dye solution. These tubes contained the samples ready for electrophoresis.

2.2.12. Sequencing Gel Electrophoresis

20 x 50 cm glass plates were thoroughly cleaned with a sponge in hot soapy water. After cleaning, the plates were rinsed in distilled water, wiped with tissue paper impregnated with ethanol, and dried with tissue paper. The inner surface of one of the plates was siliconised by wiping thoroughly with Sigmacote™ (Sigma), and the plates were taped together along the sides and bottom, using yellow electrical tape to ensure a complete seal.
A 6% polyacrylamide/urea sequencing gel was prepared as follows: 17 g urea was added to 19 ml double distilled water, 2 ml 10x TBE, and 6 ml acrylamide stock solution (38% acrylamide, 2% bis-acrylamide). After the urea had dissolved, the solution was filtered. 240 μl of 10% ammonium persulphate solution, and 35 μl of Temed (N,N,N',N'-tetramethyl-ethylene diamine: Sigma) were added immediately before pouring the gel.

The gel was poured by transferring the solution to a 50 ml syringe, which was clamped in position so that the solution could slowly run through. The plates were held at a low angle so that the gel solution could run slowly down one side of the space between the plates. Small bubbles forming while pouring were removed by tapping the plates. After pouring the gel, the flat edge of a plastic sharkstooth comb was inserted between the plates to a depth of 5-8 mm.

Once the gel had set, the top of the gel was washed with distilled water, and the comb was removed leaving a straight edge at the top. The tape from the bottom of the gel plates was removed, and the gel sandwich was clamped into the electrophoresis chamber, using vaseline grease to ensure a tight seal against the upper chamber. The upper chamber was filled with 0.5x TBE buffer, and checked for leakages. The comb was placed between the glass plates, with the teeth down, so that the teeth were resting on the top edge of the gel, thereby forming
sample wells. The sample wells were individually rinsed with a small syringe, to remove urea which tends to diffuse into the wells and would otherwise lower the resolution of electrophoresis. 0.5x TBE buffer was placed in the lower chamber. Immediately before loading, the DNA was denatured by heating the samples to 75°C for 2 minutes, and placing them on ice. The samples were loaded in the order G-A-T-C into adjacent sample wells (1-2 µl per well). Electrophoresis was performed at 30 V for 40 minutes - 3.5 hours, depending on the portion of the sequence to be obtained.

After electrophoresis, the apparatus was dismantled so that the gel remained intact and attached to the non-siliconised plate. The plate with the gel was transferred to a tray of fixing solution (10 % glacial acetic acid, 10 % methanol, in distilled water). The gel was transferred to Whatman 3 mm blotting paper, covered in Saran Wrap, and dried under vacuum on a gel drier at 80°C for 45 minutes. The dried gel was placed against X-ray film (Curix, RPI) for 1-4 days exposure.
2.3. Characterising Plasmodium falciparum antigen polymorphisms

2.3.1. Preparation of Multispot Schizont Slides

Schizont-infected erythrocytes (from 2.1.7.) were washed 3 times and resuspended at approximately 3% haematocrit in PBS. Multispot slides (C.A. Hendley Ltd., Essex) of schizonts were prepared with 20 µl of the cell suspension per spot. The slides were dried in a desiccator or a well air-conditioned room, and stored under desiccation at -20°C, in sealed polythene bags containing calcium chloride crystals (McBride, Welsby & Walliker, 1984).

2.3.2. Monoclonal Antibodies

A panel of 27 murine monoclonal antibodies (MAbs), recognising allelic variants of three *P. falciparum* antigens coded by unlinked loci were used to type parasites in each isolate. The antigens included MSP1, the precursor to several merozoite surface antigens (also known as PMMSA, PSA, p190, gp195 or MSA-1; Section 1.5.3.), another merozoite surface protein MSP2 (also called gp35-53, GP3 or MSA-2; Section 1.5.4.), and an exported protein Exp-1 (also called p23 or CRA; Section
1.5.5.). Details on the specificity, isotype, working dilution and source of the MAbs are given in Table 1.

2.3.3. Single-colour Indirect Immunofluorescence

Each *P. falciparum* isolate was analysed in a series of immunofluorescence assays (IFA) using the above MAbs (McBride, Walliker & Morgan, 1982). Working dilutions of antibodies were made in PBS containing 1% bovine serum albumin, and 0.01% sodium azide. 20 µl of each MAb was incubated on a well of an acetone-fixed multispot slide preparation of schizonts at room temperature for 30 minutes, carefully removed by pasteur pipette, and slides were washed 3 times (1, 5, and 5 minutes) in PBS. After drying the slides at approximately 60°C for 10 minutes, 20 µl of a 1:100 dilution of FITC-conjugated polyvalent rabbit anti-mouse Ig antibody (ICN Immunobiologicals, Lisle, Israel), was added to each spot, and incubated for 30 minutes. After 2 washes (1 and 5 minutes) in PBS, parasite nuclei were stained with DAPI (4',6-diamino-2-phenylindole, Sigma Ltd.: 1 x 10⁻⁶ w/v in PBS) for 1 minute. The slides were washed twice (1 and 5 minutes) in PBS, and mounted in Citifluor (City University, London) or a 4:1 glycerol:PBS mixture under coverslips. Parasites were visualised by FITC-fluorescence (MAb specific) and DAPI-fluorescence (DNA specific), with incident light of 450-490 nm and 390-440 nm respectively, at magnification...
Table 1. (opposite)

Summary of specificities of monoclonal antibodies against MSP1, MSP2, and Exp-1.

Epitopes marked by identical symbols exhibit identical allelic distributions.

<table>
<thead>
<tr>
<th>Reciprocal dilution for IFA</th>
<th>Antibody</th>
<th>Isotype</th>
<th>Epitope and location</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Anti-MSP1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.8-4-4-1</td>
<td>1000</td>
<td>IgG1</td>
<td>* conserved conformational</td>
<td>2,3</td>
</tr>
<tr>
<td>12.4-3-4</td>
<td>500</td>
<td>IgG1</td>
<td>* conserved conformational</td>
<td>2,3</td>
</tr>
<tr>
<td>12.8-2</td>
<td>1000</td>
<td>IgG2b</td>
<td>* conserved block16-17,16K fragment</td>
<td>3</td>
</tr>
<tr>
<td>12.2-1-1</td>
<td>2000</td>
<td>IgG1</td>
<td>polymorphic block 2 repeats</td>
<td>1-3</td>
</tr>
<tr>
<td>3D3.10</td>
<td>1000</td>
<td>IgG2b</td>
<td>polymorphic block 2 repeats</td>
<td>6</td>
</tr>
<tr>
<td>9.5-1-5-1</td>
<td>500</td>
<td>IgG2b</td>
<td>polymorphic 80K fragment</td>
<td>1-3</td>
</tr>
<tr>
<td>13.2-3</td>
<td>2000</td>
<td>IgG1</td>
<td>polymorphic 80K fragment</td>
<td>3</td>
</tr>
<tr>
<td>12.1-5-4</td>
<td>2000</td>
<td>IgG1</td>
<td>polymorphic block 4, 80K fragment</td>
<td>1-3,8</td>
</tr>
<tr>
<td>10-2B</td>
<td>2000</td>
<td>IgG2a</td>
<td>polymorphic block 4, 80K fragment</td>
<td>4</td>
</tr>
<tr>
<td>9.2-6-2</td>
<td>2000</td>
<td>IgG1</td>
<td>* dimorphic block 12-14</td>
<td>1-3,8</td>
</tr>
<tr>
<td>9.7-1</td>
<td>500</td>
<td>IgG1</td>
<td>* dimorphic conformational</td>
<td>1-3</td>
</tr>
<tr>
<td>10.3-2</td>
<td>500</td>
<td>IgG1</td>
<td>* dimorphic conformational</td>
<td>1-3</td>
</tr>
<tr>
<td>1-1C</td>
<td>500</td>
<td>IgG1</td>
<td>§ dimorphic 80K fragment</td>
<td>4</td>
</tr>
<tr>
<td>7.3-7</td>
<td>1000</td>
<td>IgG2a</td>
<td>§ dimorphic conformational</td>
<td>1-3</td>
</tr>
<tr>
<td>34-5</td>
<td>undiluted</td>
<td>IgG1</td>
<td>§ dimorphic 36K fragment</td>
<td>5</td>
</tr>
<tr>
<td>6.1-1-3</td>
<td>500</td>
<td>IgG1</td>
<td>§ dimorphic block 16, 40K fragment</td>
<td>1-3,8</td>
</tr>
<tr>
<td>13.1-2</td>
<td>2000</td>
<td>IgG1</td>
<td>§ dimorphic block 16, 40K fragment</td>
<td>1-3,8</td>
</tr>
<tr>
<td>17.1-3</td>
<td>2000</td>
<td>IgG1</td>
<td>§ dimorphic block 16, 40K fragment</td>
<td>1-3,8</td>
</tr>
<tr>
<td>111.4</td>
<td>1000</td>
<td>IgG1</td>
<td>polymorphic block16-17,16K fragment</td>
<td>7</td>
</tr>
<tr>
<td><strong>Anti-MSP2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.3-2-2</td>
<td>500</td>
<td>IgG1</td>
<td># polymorphic</td>
<td>9</td>
</tr>
<tr>
<td>12.5-1-2</td>
<td>500</td>
<td>IgG1</td>
<td># polymorphic</td>
<td>9</td>
</tr>
<tr>
<td>13.4-2-1</td>
<td>500</td>
<td>IgG1</td>
<td>polymorphic</td>
<td>9,10</td>
</tr>
<tr>
<td>4-4F</td>
<td>300</td>
<td>IgM</td>
<td>* polymorphic</td>
<td>11</td>
</tr>
<tr>
<td>8-5D</td>
<td>300</td>
<td>IgM</td>
<td>* polymorphic</td>
<td>11</td>
</tr>
<tr>
<td>8G10/48</td>
<td>300</td>
<td>IgG2b</td>
<td>polymorphic</td>
<td>12</td>
</tr>
<tr>
<td>8F6/49</td>
<td>40</td>
<td>IgG3</td>
<td>polymorphic</td>
<td>12</td>
</tr>
<tr>
<td><strong>Anti-Exp-1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.1-4</td>
<td>500</td>
<td>IgG1</td>
<td>polymorphic</td>
<td>13</td>
</tr>
</tbody>
</table>
Isolates within which any polymorphic epitopes were expressed by only a proportion of schizonts, suggesting the presence of heterogeneous parasite sub-populations (Chapter 4), were tested further using a double-labelled (two-colour) IFA.

2.3.4. Two-colour Indirect Immunofluorescence

Combinations of two MAbs with different isotypes, and different epitope specificities, were used to test the homogeneity of parasites within particular isolates. Each pair of MAbs was incubated together on a slide for 30 minutes. An RITC-conjugated and an FITC-conjugated antibody (Southern Biotechnology Associates Inc., Birmingham, Alabama; dilution 1:50), each specific for the different isotypes of the two MAbs, were then incubated together on the slide for the second stage of 30 minutes, and carefully removed with a pasteur pipette. All other steps were performed using exactly the same method as described above for the individual MAb IFA.

In addition to FITC-fluorescence (green) and DAPI fluorescence (blue), as described above, RITC-fluorescence (red) was visualised using incident light of 515-560 nm. The proportion of schizonts showing (i) green (and blue) fluorescence only, (ii) red (and blue) fluorescence only, (iii) red and green (and blue)
fluorescence, and (iv) neither red nor green (only blue) fluorescence, was recorded for each pair of MAbs tested. Combined results, obtained with a series of different pairs of MAbs, resolved the number of distinct parasite sub-populations within each isolate (Chapter 4). The antigenic profile of the majority sub-population within each isolate was also determined (Chapter 3).

2.3.5. Photographing Microscopic Fields

Fluorescent parasites were viewed with a x60 objective lens, under water immersion, on a Leitz Dialux 20 fluorescence microscope. Photographs were taken with an attached Wild MPS 515 camera on Kodak 800 print film, exposures controlled by a Wild Photoautomat MPS 45.
2.4. Blood Group Typing

2.4.1. Cells and reagents

Erythrocytes from patients were washed (section 2.1.4.), and 15 µl packed cells were resuspended in 500 µl PBS (3 % haematocrit). These cells were kept at 4°C, and blood typing was performed within the following 7 days. Erythrocytes of control phenotypes were provided by the Scottish National Blood Transfusion Service (BTS). As the control cells had a shelf life of 1 month, controls were also identified among laboratory workers. Anti-A, anti-B, anti-M, and anti-N monoclonal antibodies were provided in working dilution by the BTS, and stored at 4°C. The monoclonal antibody cell lines were, LM103107 / ES9 (anti-A), ES4 (anti-B), LM110140 (anti-M), and LM1719 (anti-N). Anti-S rabbit antiserum was also provided by the BTS, and stored at -20°C.

2.4.2. Direct agglutination of erythrocytes

For typing ABO and MN groups, two drops (60-100 µl) of the appropriate monoclonal antibody were added to a 12 x 75 mm glass bacteriological tube, and one drop (30-50 µl) of blood cells (at 3% haematocrit) also added. The tube was gently shaken, stood at room temperature for 5 minutes, and centrifuged at 1000 g (2400 rpm) for 15
seconds. Agglutination of erythrocytes was determined macroscopically by holding the tube at an angle and gently tapping the bottom to dislodge the pellet. Erythrocytes were tested similarly with the anti-S antiserum, except that incubation was at 37°C for 15 minutes.
2.5. Computer Methods

2.5.1. Statistical tests using commercial software

Data was coded and organised in files, using the EMAS-3 operating system (University of Edinburgh) on a mainframe computer. The MINITAB statistical package was used for the following procedures: comparison of sample means by paired t-tests, calculation of standard errors of sample means, chi-square tests on 2 x 2 contingency tables, and generation of probability distributions.

2.5.2. A program for analysis of genotype diversity

The VARPAS program was written in PASCAL by Mr. Adam Eyre-Walker (see Appendix), and run on the EMAS-3 operating system on the mainframe computer. The program employed a procedure of random sampling with replacement, to calculate the expected number of different multi-locus genotypes within a sample of given size, from the observed allelic frequencies at each locus, assuming random recombination between loci. The application of the program is described in Chapter 3.
Malaria parasites reproduce asexually and are haploid within vertebrate hosts, but produce gametes and undergo fertilisation, zygote formation and sexual recombination at meiosis within mosquitoes (Figure 1, section 1.1.1.; Sinden, Hartley & Winger, 1985; Walliker et al., 1987). The sexual reproduction in the mosquito vector is an obligatory event in the parasite life-cycle, and without it natural transmission of malaria cannot proceed. In natural populations sexuality thus provides opportunities for generation of recombinant genotypes by chromosomal reassortment.

In *Plasmodium falciparum* allelic polymorphisms exist at many loci, including antigens, enzymes and other proteins (section 1.5.). Many combinations of alleles at different loci have been observed among culture-adapted parasites (Fenton, Walker & Walliker, 1985; Creasey et al., 1990), consistent with the hypothesis that genetic recombination and segregation of alleles occurs in natural populations (reviewed by Walliker, 1985, & 1989). An experimental genetic cross between two defined clones of *P. falciparum* directly demonstrated that, during transmission of a mixed infection, recombinant genotypes were generated (Walliker et al., 1987).
Despite the evidence that genetic exchange occurs, probably at meiosis, it has been suggested that this may be a relatively infrequent occurrence in natural populations. Tibayrenc, Kjellberg & Ayala (1990), reviewed data on genetic polymorphisms in several species of parasitic protozoa, including *P. falciparum*, and proposed that a predominantly 'clonal' population structure may be characteristic of parasitic protozoa generally. By the term 'clonal', with respect to a sexually reproducing species such as *P. falciparum*, they referred to the possibility that a high degree of inbreeding occurs, due to almost exclusively self-fertilising events. Such a predominance of self-fertilisation could allow the development of linkage disequilibria, ultimately resulting in some multi-locus genotypes being over-represented in the finite parasite population, due to random genetic drift (Hedrick, 1985).

To determine whether alleles from unlinked polymorphic loci are assorted randomly in a population, or whether linkage disequilibria exist, a technique is required to identify alleles, or allelic products, of different loci in the same parasite, and to characterise a large sample of the population. In this study, indirect immunofluorescence analyses (IFA) with monoclonal antibodies (MAbs) are used to identify in situ allelic variants of three polymorphic blood stage antigens whose diversity is well understood at the DNA level. The antigens, MSP1, MSP2, and Exp-1, are each encoded by a single locus within the haploid genome,
and the loci have been mapped to different chromosomes (Kemp et al. 1987, 1990). Each antigen exhibits allelic polymorphism determined by DNA sequence and detectable by MAbs as epitope differences among parasite clones (sections 1.5.3., 1.5.4., 1.5.5., and Chapter 6).

When appropriate sample data are available, population genetic analyses which may be used to test for inbreeding, i.e. effects of self-fertilisation, within a population of organisms could include the following:

(a) The coefficient of inbreeding, \( f \), may be calculated from genotype frequencies at a single locus with two alleles, in a population of diploid organisms,

\[
f = 1 - \frac{H}{2pq}
\]

where \( H \) is the frequency of heterozygotes, \( p \) is the frequency of one allele and \( q \) is the frequency of the other allele. \( 2pq \) is the 'Hardy-Weinberg' heterozygosity, i.e. the frequency of heterozygotes expected if the population is randomly mating. If inbreeding is occurring, the frequency of heterozygotes, \( H \), will be less than expected, and hence the value of \( f \) will be more than zero (Hedrick, 1985). This analysis may not, of course, be used on data from haploid organisms, such as the blood stages of \( P. falciparum \).

(b) Linkage disequilibrium, \( D \), between two loci, may be calculated if frequencies of the two-locus genotypes (i.e. gamete types) are known. However, the existence of linkage disequilibrium may be either due to inbreeding and
random genetic drift, or to interactions between genes at two or more polymorphic loci influencing the fitness of the phenotype in a given environment (Hedrick, 1985). If inbreeding is the cause of linkage disequilibria, then such disequilibria should be detectable among many loci at which polymorphisms are neutral.

The most simple measure of linkage disequilibrium involves a calculation of the non-random association between alleles at two loci, by sampling haploid individuals such as gametes from a population. Calculation of the measures $D$, $D'$, and the statistic $Q$, for this purpose are described in the methods section 3.2.2.

(c) If allelic polymorphism at the loci studied is high, and hence the number of two-locus genotypes is very high, it is possible that a sample from the population may not contain sufficient numbers of each genotype for conventional analyses of linkage disequilibrium to be valid statistically. In such a case, it is possible to measure genotype diversity within a sample and test whether this is lower than expected under assumption of panmixia, i.e. in the absence of inbreeding.

The genotype diversity (single- or multi-locus) within a population may be calculated from a sample of the population, using the G-test (Stoddart & Taylor, 1988). In the case of multi-locus genotype diversity, an expected value of $G$ may be calculated, using the allelic frequencies in the sample, under the assumption of panmixia. A test
value of G can be derived from the observed numbers of each multi-locus genotype, and compared with the expected value of G, using a t-test or a goodness of fit test. The equations involved are quite complex, and there is particular difficulty in obtaining a measure of the variance of G for the t-test (Stoddart & Taylor, 1988). Because of these complexities, the G-test was discarded for the purposes of this study, in favour of the following method of analysis of genotype diversity.

(d) Assuming panmixia, an expected number of multi-locus genotypes within a sample of given size may be calculated from maximum-likelihood estimates of allelic frequencies, using a computer program to sample genotypes randomly with replacement. A large number, say 1000, of repeated random samples will yield a frequency distribution of generated numbers of expected multi-locus genotypes, with which an observed number can be compared. This is conventionally termed a 'bootstrap' process (Weir, 1990). Unlike the G-test, this is not a parametric statistical test, but does allow a sensitive comparison between observed and expected genotype diversity within a sample. The procedure followed in this study is described in the methods section 3.2.2.
3.2. Methods

3.2.1. Practical Methods

A *P. falciparum* isolate is defined as a sample of parasites taken from a patient on a single occasion. Isolates were collected from patients in three areas of The Gambia (Figure 5, section 2.1.1.):

(a) An urban/periurban region within a 7km radius of Serekunda, from Bakau in the north to the Yundum International Airport in the south. 344 isolates were obtained from patients presenting to the Outpatients Departments of the Medical Research Council, Fajara, and the Royal Victoria Hospital, Banjul, during July-December 1988 (N=225), and October-December 1989 (N=119). None of the patients had travelled outside the region during three weeks prior to clinical presentation.

(b) 9 isolates were obtained in Sarakunda village, near Farafenni on the north bank of the Gambia river, on October 5th & 6th 1988.

(c) 17 isolates were obtained in Brefet village, on the south bank in the Lower River Division, during October-December 1989.

Heparinised blood samples were obtained from each patient (section 2.1.3.), erythrocytes were washed (section 2.1.4.), and the parasites were cultured for 24-48 hours to obtain mature schizonts (section 2.1.7.). Multispot slides
of schizonts were prepared (section 2.3.1.) for immunofluorescence typing with MAbs against MSP1, MSP2, and Exp-1 (Table 1, section 2.3.2.).

All MAbs were first used individually in indirect immunofluorescence assays (section 2.3.3.) to determine MSP1, MSP2 and Exp-1 variants expressed by schizonts of each isolate. Since many isolates contained more than one 
P. falciparum clone (Chapter 4), the antigen phenotype was determined for the majority clone within each isolate. This was achieved using pairs of MAbs with different isotypes and epitope specificities in a double-labelled (2-colour) IFA. Combined results from different pairs of MAbs allowed the antigen phenotype of the majority clone within each isolate to be accurately resolved (section 2.3.4.).

Allelic serotypes for each of the antigens were distinguished by their different profiles of reactivity with typing MAbs. Each allelic serotype was assigned a number, according to a classification scheme based on combinations of individual MAb specificities.

3-locus genotypes were determined as different combinations of allelic serotypes of the 3 antigens. Many isolates contained more than one parasite clone (Chapter 4), although the genotypes of minority clones could not be accurately determined in every case. As the present analysis simply requires a random sample of genotypes from the local parasite population, it was decided that only the genotype of the majority parasite clone within each isolate...
would be used for analysis, as this was accurately determined for each isolate.

3.2.2. Statistical Methods

3.2.2.1. Measurement of linkage disequilibrium

If there is random association between alleles at two loci, the expected frequency of each 2-locus gamete type, $x_1$, $x_2$, $x_3$, or $x_4$, is equal to the product of its component allele frequencies. Thus,

$$
x_1 = p_1q_1
$$

$$
x_2 = p_1q_2
$$

$$
x_3 = p_2q_1
$$

$$
x_4 = p_2q_2
$$

where $p_1$ and $p_2$ are the frequencies of the two alleles at the first locus, and $q_1$ and $q_2$ are the frequencies of the two alleles at the second locus. Deviation from random association (i.e. disequilibrium) between the two loci may be measured by the value, $D$, where

$$
D = x_1x_4 - x_2x_3 \quad \text{(Hedrick, 1985)}
$$

$D$ has a range from $-0.25$ to $0.25$. However, its value is not only sensitive to the degree of non-random association between alleles at the different loci, but also to the allele frequencies themselves. For example, $D$ can only possibly attain its maximum (or minimum) value when allelic frequencies are all equal, i.e. $p_1 = p_2 = q_1 = q_2 = 0.5$. When allelic frequencies are unequal, there are
restrictive limits to the value D, even if the degree of non-random association is absolute.

A helpful modification, D', is calculated by dividing the value D by the maximum attainable value (Dmax) given the component allelic frequencies. If D is positive, Dmax is equal to the lesser of \( p_1 q_2 \) or \( p_2 q_1 \). If D is negative, Dmax is equal to the lesser of \( p_1 q_1 \) or \( p_2 q_2 \). An advantage of using the measure D' (which has a range from -1.0 to 1.0) is that it is not biased by the component allelic frequencies at each locus.

To test whether an observed linkage disequilibrium is significant, a statistic, Q, may be calculated,

\[
Q = n \frac{D^2}{p_1 p_2 q_1 q_2}
\]

(Hedrick, 1985)

where \( n \) = the number of haplotypes sampled, and \( p_1, p_2, q_1, \) and \( q_2 \) are the observed allelic frequencies within the sample. Q approximates to the chi-square value obtained for a 2 x 2 contingency table of the numbers of each 2-locus haplotype sampled, and the statistical significance of the disequilibrium can be accordingly read from a statistical table for chi-square values (Hill, 1974).

3.2.2.2. Measurement of multi-locus genotype diversity

Assuming panmixia, expected 3-locus genotype frequencies were firstly calculated from the observed allelic frequencies of MSP1, MSP2, and Exp-1, taken from a large sample size (N = 225 in 1988, and N = 119 in 1989) and stored in a data file. Secondly, a PASCAL program
(section 2.5.2., and Appendix) was used to calculate probability distribution histograms of the expected no. of different 3-locus haploid genotypes within samples of given size, using a 'bootstrap' method of sampling with replacement (Weir, 1990). The program performed the following functions:

(i) Converted the expected 3-locus genotype frequencies into cumulative values, between 0 and 1.

(ii) Randomly picked N values between 0 and 1 in order to obtain N(d) different cumulative values from (i). This corresponds to N(d) different genotypes from a random sample of size N.

(iii) Repeated step (ii) 1000 times to obtain a distribution of values of N(d), which was plotted as a probability distribution histogram, corresponding to the approximate probability of obtaining a given number of different genotypes within a particular sample size, N.

Thirdly, using the generated probability distributions, observed 3-locus genotype diversities within samples were compared with those expected, in the following cases:

(a) Within the urban/periurban region, observed numbers of different 3-locus genotypes were compared to expected numbers, N(d), for the whole sample in each year. The test was applied to 1988 and 1989 data separately.

(b) Within Sarakunda and Brefet villages, observed numbers of different genotypes were compared to numbers
expected for identical sample sizes taken randomly from the urban/periurban region during the same year. This comparison could only be considered as a fair test for inbreeding within the villages if it is assumed that the allelic frequencies within villages are distributed similarly to the urban/periurban region.
3.3. Results

3.3.1. Allelic Serotype Frequencies at the MSP1, MSP2 and Exp-1 Antigen Loci

Serotypes at the three loci were determined by IFA tests with a panel of MAbs, on parasites present within 344 isolates collected in the urban/periurban region during 1988 (July-December) and 1989 (October-December). Less than half of the isolates appeared to consist of a single parasite clone whose reactions with all MAbs were homogeneous, while 2 or more clones could be clearly differentiated in other isolates (Chapter 4). In the present analysis, each of such mixed isolates is represented once, by its most predominant clone whose serotype was determined accurately by two-colour IFA.

MSP1 exhibited a very high degree of allelic polymorphism, with 36 different serotypes detected (type examples are given in Table 2). MSP1 serotypes and their frequencies are shown in Figure 6. 35 different serotypes were detected among 225 isolates in 1988. In 1989, 24 serotypes were represented among 119 isolates at frequencies similar to the previous year.

Allelic polymorphism of MSP2 was also high, with 8 different serotypes detected (type examples in Table 3). MSP2 serotypes and their frequencies are illustrated by Figure 7. The 8 serotypes were detected in 1988 (N=221).
Table 2. Type examples of each of the MSP1 serotypes as defined by combinations of reactivities with individual MAbs

<table>
<thead>
<tr>
<th>Isolate</th>
<th>MSP1 MAb Epitopes</th>
<th>Allelic Serotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>9.8</td>
<td></td>
</tr>
<tr>
<td>GF88-200</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-210</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-213</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-230</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-39</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GB88-8</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GB88-10</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-23</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF89-42</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-32</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-249</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-201</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GB88-15</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-233</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-175</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-185</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-137</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-159</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-182</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-168</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF89-16</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-172</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-9</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-67</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GB88-7</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF89-68</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GB88-36</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF89-96</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF89-101</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF89-133</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF89-65</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF89-7</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-160</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-177</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF88-74</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF89-64</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
<tr>
<td>GF89-116</td>
<td>+ + + + + + +</td>
<td>+ + + + + + +</td>
</tr>
</tbody>
</table>

Serotypes are numbered according to the scheme in Figure 6.
Table 3. Type examples of each of the MSP2 serotypes as defined by combinations of reactivities with individual MAbs

<table>
<thead>
<tr>
<th>Isolate</th>
<th>MSP2 MAb Epitopes</th>
<th>MSP2 Allelic Serotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12.3 12.5 8G10/48 8-5D 4-4F 8F6/49</td>
<td></td>
</tr>
<tr>
<td>GB88-7</td>
<td>+ + - + + + -</td>
<td>1</td>
</tr>
<tr>
<td>GF88-53</td>
<td>+ + - + + - -</td>
<td>2</td>
</tr>
<tr>
<td>GB88-17</td>
<td>+ + - - - - -</td>
<td>4</td>
</tr>
<tr>
<td>GF88-71</td>
<td>- - + + + - +</td>
<td>5</td>
</tr>
<tr>
<td>GF88-130</td>
<td>- - + + + - -</td>
<td>6</td>
</tr>
<tr>
<td>GF88-156</td>
<td>- - + - - - +</td>
<td>7</td>
</tr>
<tr>
<td>GF88-171</td>
<td>- - + - - - -</td>
<td>8</td>
</tr>
<tr>
<td>GF88-255</td>
<td>- - - + + - -</td>
<td>10</td>
</tr>
</tbody>
</table>

Serotypes are numbered according to the scheme in Figure 7.
Figure 6. MSP1 allelic serotype frequencies in the urban/periurban region in The Gambia. Each serotype is defined by a unique profile of reactivity with the monoclonal antibodies, shown at the bottom of the figure. Reactivities with MAbs 12.2 and 3D3 are symbolised as follows: \( \alpha = 3D3+,12.2+; \ \beta = 3D3+,12.2-; \ \gamma = 3D3-,12.2+; \ \delta = 3D3-,12.2-. \) Certain combinations have never been observed, e.g. types 17, 18, 19, 21 etc. (MAbs 7.3, 13.1, 17.1, 1-1C, and 34-5, showed identical serotype distribution to MAb 6.1, and are omitted from the figure for the sake of clarity).
Figure 7. MSP2 allelic serotype frequencies in the urban/periurban region in The Gambia. Each serotype is defined by a unique profile of reactivity with the monoclonal antibodies, shown at the bottom of the figure. Types 3 and 9 have never been observed. (MAb 12.5 showed an identical serotype distribution to MAb 12.3, and MAb 4-4F showed an identical serotype distribution to MAb 8-5D, and are omitted from the figure for the sake of clarity).
The frequencies of 7 serotypes represented also in 1989 were similar to those found in the previous year (N=117).

Serotype polymorphism in Exp-1 is limited to a single epitope change recognised by MAb 5.1 (section 1.5.5.). In 1988, 30.4% of the majority clones within isolates typed positive for this epitope (N=224). 30.3% were positive in 1989 (N=119).

3.3.2. Testing for linkage disequilibrium

Calculation of linkage disequilibrium measures, D, D', and the statistic Q, is cumbersome when there are more than 2 alleles at a locus, and the method may also be insensitive due to low numbers of each allele within a sample. However, an option for the analysis of such data is to consider the most common allele as 'allele A1', and to lump the remaining alleles for analysis as 'allele A2'.

Here, the most common allele of MSP2 had a frequency of greater than 0.3 in 1988 and 1989 (allelic serotype 2, Figure 7). To test for linkage disequilibrium between MSP2 and Exp-1, this allele is termed 'allele A1', and the remaining alleles of MSP2 were lumped and treated as a single allele, 'allele A2'. In the case of Exp-1, for which only two serotypes were identified, no lumping is necessary, and the two serotypes are considered as 'allele B1' (epitope 5.1+) and 'allele B2' (epitope 5.1-). The genotype frequencies and calculations of D, D', and Q for 1988 and 1989 samples are given in Table 4. In neither year
Table 4.

a) Number of 2-locus genotypes (MSP2 and Exp-1)

<table>
<thead>
<tr>
<th></th>
<th>A1B1</th>
<th>A1B2</th>
<th>A2B1</th>
<th>A2B2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1988</td>
<td>20</td>
<td>60</td>
<td>47</td>
<td>94</td>
<td>221</td>
</tr>
<tr>
<td>1989</td>
<td>12</td>
<td>29</td>
<td>24</td>
<td>52</td>
<td>117</td>
</tr>
</tbody>
</table>

See text for designation of alleles A1, A2, B1, and B2

b) Test for linkage disequilibrium between MSP2 and Exp-1

<table>
<thead>
<tr>
<th></th>
<th>D</th>
<th>Dmax</th>
<th>D'</th>
<th>$\chi^2 \approx Q$</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1988</td>
<td>-0.0192</td>
<td>0.0905</td>
<td>-0.212</td>
<td>1.669</td>
<td>p&gt;0.1</td>
</tr>
<tr>
<td>1989</td>
<td>-0.0053</td>
<td>0.103</td>
<td>-0.051</td>
<td>0.067</td>
<td>p&gt;0.5</td>
</tr>
</tbody>
</table>
was any significant linkage disequilibrium detected.

Linkage disequilibrium between MSP1 and either of the other loci was not tested for, because each of the MSP1 allelic serotypes was present at a frequency of less than 0.15 (Figure 6), which renders the statistic Q very insensitive to detecting actual linkage disequilibrium, due to the low numbers in each 2-locus genotype class. For this reason, the more sensitive 'bootstrap' method for the analysis of 3-locus genotype diversity (described in section 3.2.2.) was used to test for evidence of 'clonality' in the population genetic structure.

3.3.3. Expected and Observed 3-Locus Genotype Diversity Within the Urban/periurban Region

Each 3-locus genotype was defined as a different combination of MSP1, MSP2, and Exp-1 allelic serotypes. In 1988 among 220 parasite isolates 126 different genotypes were observed (Figure 8). In 1989 among 117 parasite isolates 75 genotypes were observed (Figure 9). These observed values were then compared to probability distribution histograms of genotype diversity generated by the 'bootstrap' method described in section 3.2.2.2.

For the 1988 data (sample size N=220), the probability distribution predicted a median value of 126 and a lower 95% point of 117 different genotypes. The observed value of 126 was thus exactly as expected (Figure 10a). For the 1989
Figure 8. 3-locus genotypes of the majority clone within each of 220 isolates from the urban/peri-urban region in 1988. The MSP1 types (columns) and MSP2 types (rows) are numbered as defined in Figures 6 and 7, respectively. Exp-1 types are positive or negative with MAb 5.1, shown by closed and open circles respectively. 126 different genotypes are represented in this sample.
Figure 9. 3-locus genotypes of the majority clone within each of 117 isolates from the urban/periurban region in 1989. The MSP1 types (columns) and MSP2 types (rows) are numbered as defined in Figures 6 and 7, respectively. Exp-1 types are positive or negative with MAb 5.1, shown by closed and open circles respectively. 75 different genotypes are represented in this sample.
Figure 10. Probability distribution histograms (generated by a bootstrap method, section 3.2.2.) of the expected number of different 3-locus genotypes within samples of (a) 220 isolates in 1988, and (b) 117 isolates in 1989, on the assumption of random assortment of alleles at different loci. The observed number of different 3-locus genotypes within each sample is shown by an arrow.
sample (N=117), the probability distribution generated an expected median value of 80, and a lower 95% point of 73. The observed value of 75 was within the expected range (Figure 10b).

Therefore, for both 1988 and 1989 transmission seasons, the observed 3-locus genotype diversity values were within the range expected from the hypothesis of random assortment of alleles.

3.3.4. 3-Locus Genotype Diversity in Villages

Samples were taken from two villages distant from the urban region (Figure 5, section 2.1.1.). The samples from both villages were small, so allelic frequency distributions could not be accurately estimated. For this reason, making an assumption that these allelic frequency distributions were similar to those observed in the urban/periurban region, numbers of 3-locus genotypes observed in the villages were compared to expected values generated for identical sample sizes taken from the urban/periurban region in the appropriate year.

In a sample of 9 isolates from Sarakunda village, 7 different genotypes were observed (three isolates contained an identical genotype). The median expected value was 9 and the lower 95% point was 8. The observed value was therefore lower than expected.

In a sample of 17 isolates from Brefet village 13
different genotypes were observed (in 1989). The median expected value was 16 and the lower 95% point was 14. The observed value was therefore lower than expected.
3.4. Discussion

Within the Gambian urban/periurban region, the observed extensive diversity of *P. falciparum* 3-locus genotypes was in accordance with expectations assuming panmixia, using the 'bootstrap' procedure of random sampling with replacement from the observed allele frequencies at each of the three loci. In concordance with this, no linkage disequilibrium was detected between the MSP2 and Exp-1 loci, using the more conventional but less sensitive test for disequilibrium.

These results do not support a hypothesis of Tibayrenc *et al.* (1990), who propose that 'clonal' (i.e. self-fertilising) reproduction is more predominant than cross-fertilising reproduction, within populations of *P. falciparum*. Although the clonal hypothesis may apply to other genera of parasitic protozoa, there is no data to support the hypothesis for any plasmodia. Neither of the sources cited by Tibayrenc *et al.* (1990) contains data suitable for analysis of population genetic structure. The unusual observation of identical multi-locus genotypes in two culture established isolates from different parts of Thailand (Fenton *et al.*, 1985), has already been explained by Fenton (1988) as probably due to isolate cross-contamination. The cited isoenzyme data on 15 geographically disparate cultured isolates (Sanderson *et al.*, 1981), include two identical clones from the same
patient. More importantly, in that study the local allelic frequencies were not known. In contrast, 2-locus analyses of clinical samples from The Gambia and Tanzania revealed no deviations from local expected genotype frequencies (Carter & McGregor, 1973; Carter & Voller, 1975).

Information on other *Plasmodium* species is also consistent with frequent genetic exchange in natural populations. Joshi et al. (1989) observed random combinations of alleles at 3 iso-enzyme loci among 76 *P. vivax* isolates from northern India. An earlier study of isoenzyme polymorphism among *P. chabaudi* clones from central Africa did not demonstrate any non-random associations of alleles at different loci, either (Beale, Carter & Walliker, 1978). Although there is no evidence of 'clonality' in any natural plasmodial populations, it has been rightly pointed out that a degree of inbreeding may occur which is insufficient to cause detectable linkage disequilibrium (Dye, 1991).

In the present study, 3-locus genotype diversity was significantly lower in each of two villages than in the urban/periurban region. A possible confounding variable with respect to Sarakunda village is that the isolates were collected on two consecutive days, and their diversity compared with expected values calculated from a sample collected in the urban/periurban region during six months. However, the Brefet village sample was collected during the same period as the urban/peri-urban sample, with which the
genotype diversity was compared.

It must be noted that the lower 3-locus genotype diversity within villages might not be due to inbreeding, but could be due to a lower allelic diversity at each of the three loci, although the local allelic frequency distributions could not be determined due to the small sample sizes. Lower allelic diversity could arise if there is a restricted parasite population size within villages. The mean number of detectable *P. falciparum* clones per acute infection in The Gambia is 2.0 (Chapter 4). Assuming that the mean number of clones within all infections is about 2, it follows that in an area of 50% *P. falciparum* prevalence for example, the number of blood stage clones will not exceed the number of people. If some of the blood stage clones do not contribute to the next generation (Burkot et al., 1984), the effective *P. falciparum* population size will be even lower (Kimura & Ohta, 1971; Hedrick, 1985). Within a village of only a few hundred people for example, the effective *P. falciparum* population size may be small enough to result in a reduction of genetic diversity, due to local extinction of rare alleles by random genetic drift (Kimura & Crow, 1964). Even though *P. falciparum* populations in different villages are not completely isolated from one another, the introduction of new allelic variants may occur at a slower rate than the local extinction of alleles, resulting in a lower genetic diversity than within a larger population.
The above argument has implications for the interpretation of epidemiological studies on *P. falciparum*. Forsyth *et al.* (1989) recorded differences in the prevalence of an allelic serotype of the S-antigen between villages in Papua New Guinea, and between samples taken on different dates from particular villages. It was suggested (Forsyth *et al.* 1988) that frequency-dependant selection due to variant-specific immunity was the most likely explanation for changes in the prevalence of the serotype. However, changes in allelic frequency are explainable by non-selective genetic drift in any population of limited size (Kimura, 1955). In the Gambian urban/periurban study region, the MSP1, MSP2, and Exp-1 serotype frequencies were similar in 1988 and 1989, and the frequencies of individual epitopes were stable over the 1982-1989 period (Chapter 7).

The interpretation of studies on allelic polymorphism is often controversial, due to the difficulty in resolving the contribution of neutral and selective processes. With respect to *P. falciparum*, many studies are undertaken on genes for which there is an implicit assumption of selection. To be informative, such studies require a knowledge of the local *P. falciparum* population structure, and a recognition of the importance of neutral processes.
3.5. Summary

Serotyping with monoclonal antibodies was used to estimate the number and frequencies of allelic types of two merozoite surface proteins, MSP1 and MSP2, and an exported protein Exp-1, in a sample of 344 clinical isolates of *P. falciparum* from an urban/periurban region of The Gambia. Represented among the isolates were 36, 8 and 2 alleles of the MSP1, MSP2 and Exp-1 loci respectively. Relative frequencies of these alleles remained stable in the parasite population over two years of the study. No linkage disequilibrium was detected. A computer program was used to calculate, from the frequencies of individual alleles at the three loci, the probable number of different 3-locus genotypes in samples from the population, assuming random assortment among the unlinked loci. No significant difference was found between the expected and the observed 3-locus genotype diversity. It is concluded that recombination among unlinked loci is a common consequence of sexual reproduction of *P. falciparum* in The Gambia. Slightly lower genotype diversity was observed in each of two villages, which may be a consequence of smaller parasite population size compared with the urban/periurban region.
Chapter 4: The Epidemiology of Multiple-Clone 
Plasmodium falciparum Infections in Malaria Patients

4.1. Introduction

*Plasmodium falciparum* infections containing more than one genotype were first analysed *in vitro* by Rosario (1981), using limiting dilution cloning to demonstrate heterogeneity of parasites with respect to electrophoretic forms of glucose phosphate isomerase (GPI). Thaithong *et al.* (1984) used isoenzyme electrophoresis, monoclonal antibody immunofluorescence, 2-dimensional gel electrophoresis, and drug susceptibility assays to characterise clones derived from a Thai patient, and obtained seven distinct types, each presumed to represent a different haploid clone present in the infection.

Clonal heterogeneity of *P. falciparum* within *in vitro* cultured isolates complicates the interpretation of parasite characterisation. For example, different DNA sequences for the blood stage antigen MSP1 (section 1.5.3.) were independently obtained from the Palo Alto isolate (Chang *et al.* 1988; Scherf, Barbot & Langsley, 1989). Although the existence of clonal heterogeneity within *P. falciparum* isolates *in vitro* is well established, the epidemiology of multiple-clone infections in patients within endemic areas has not
been investigated systematically.

The number of *P.falciparum* clones in infected individuals may determine the genetic structure of the local *P.falciparum* population (Chapter 3). The rate of genetic recombination in populations of plasmodia depends on the frequency at which vector mosquitoes acquire a blood meal containing gametocytes of more than one genotype. A successful fertilisation event in the mosquito mid-gut is followed by a brief diploid stage and meiosis, producing recombinant haploid sporozoite progeny which may be inoculated at a later blood feed.

Carter & McGregor (1973) studied isoenzyme polymorphism of *P.falciparum* in The Gambia, and calculated from the allelic frequencies of lactate dehydrogenase (LDH) and GPI, and the proportion of isolates showing more than one allelic type of either enzyme, that wild isolates contained approximately two clones each.

The present study, also in the Gambia, involved the characterisation of allelic serotypes of three polymorphic blood stage antigens using monoclonal antibodies (MAbs). Different parasite clones were characterised by particular combinations of allelic serotypes for the three antigens. The study aimed to determine numbers of different clones within individual isolates, and test for effects of patients' age,
parasitaemia, stage of transmission season, and correlations with disease severity. An analysis of gene identity was also used to test whether multiple-clone infections can be explained by superinfection alone, or by inoculation of sibling recombinant parasites from the same parent zygote from a single mosquito.
4.2. Methods

4.2.1. Practical Methods

During 1988 (July-December) and 1989 (October-December), 0.3 ml blood samples were collected from 355 *P. falciparum* malaria patients at the outpatients departments of the Medical Research Council, Fajara, and the Royal Victoria Hospital, Banjul (section 2.1.1.). 153 of these patients were included in a case-control study to determine risk factors for severe malaria, 35 of whom had severe disease according to the modified Wellcome criteria (here mostly cerebral malaria: Kwiatkowski et al. 1990).

Erythrocytes were washed (section 2.1.4.), and parasites cultured for 24-48 hours until schizonts had matured (section 2.1.7.). Multispot slides of schizonts were prepared from each isolate (section 2.3.1.). A panel of 27 murine monoclonal antibodies (MAbs), specific for allelic variants of three *P. falciparum* antigens coded by unlinked loci were used to type parasites in each isolate by IFA. The antigens included MSP1, the precursor to several merozoite surface proteins (section 1.5.3.), another merozoite surface protein MSP2 (section 1.5.4.), and an exported protein Exp-1 (section 1.5.5.). Details on the specificity, isotype, working dilution and source of the MAbs are
given in Table 1 (section 2.3.2.). Each *P. falciparum* isolate was analysed in a series of IFA tests using the above MAbs (section 2.3.3.). Within some isolates, all parasites appeared to have an identical antigenic phenotype, and these isolates were considered to contain only one clone. Other isolates within which only a proportion of schizonts expressed certain polymorphic epitopes, suggesting the presence of heterogeneous parasite sub-populations, were tested further using a double-labelled (two-colour) IFA. Combinations of two MAbs with different isotypes, and different epitope specificities, were used to test the homogeneity of parasites within each isolate (section 2.3.4.). Combined results, obtained with a series of different pairs of MAbs, resolved the number of distinct parasite sub-populations within each isolate, within the limitations outlined below. The antigenic profile of the majority clone within each isolate was also determined (Chapter 3).

The number of schizonts (visualised by DAPI fluorescence) varied considerably among isolates, from less than 1 to over 500 per field. To avoid bias against detection of minority clones in isolates with low parasitaemia, an arbitrary limit of sensitivity was set at one percent of the total parasite number within any isolate. Therefore, no isolate was included in the analysis if parasite numbers were too low to score at
least 200 schizonts per test, and detected minority clones of less than one percent were also discarded from the analysis.

Despite the high resolution of double-labelled IFA analyses, it is not possible to reliably determine whether there are more than 4 clones within an isolate. Therefore, each isolate was scored as containing 1, 2, 3 or 4 clones.

4.2.2. Statistical Methods

The hypothesis that each parasite clone within a human infection is acquired from a separate mosquito inoculation (i.e. random superinfection) was tested by a statistical analysis of gene identity (i.e. theoretical 'homozygosity', \( jx \)) at the MSP1 locus between pairs of different clones within multiple-clone infections.

To test the hypothesis, an assumption is that different mosquitoes carry \( P.falciparum \) genotypes randomly acquired from the gene pool in the urban/periurban study region, i.e. that the \( P.falciparum \) population in this area is fully panmictic. Following this assumption, an expectation of the hypothesis is that each \( P.falciparum \) genotype within an infection is randomly acquired from the gene pool in the urban/periurban study region. Therefore,
the level of gene identity \((jx)\) of MSP1 among clones within an infection should be equal to the level of gene identity of MSP1 within the study region.

It is straightforward to estimate the gene identity \((jx)\) of MSP1 within the study region, as it is equal to the sum of the squares of the frequencies of each MSP1 allele \(\sum x_i^2\) in the population.

\[ j(x) = \sum x_i^2 \quad \text{(Hedrick, 1985)} \]

However, as we have only sampled a finite number of MSP1 alleles from the total population, it is important to also calculate the sampling variance \(V(jx)\) of this estimate of gene identity, using an equation given by Nei & Roychoudhury (1974):

\[ V(jx) = \frac{2(n-1)}{n^3} \left\{ (3-2n)jx^2 + 2(n-2)\sum x_i^3 + jx \right\} \]

Once \(jx\) and \(V(jx)\) have been obtained, the observed gene identity among pairs of different clones within multiple-clone infections may be calculated and compared.

In the present study, clones were first differentiated at the MSP2 locus, before pairs of clones were sampled for the analysis. If the two most predominant clones within an infection were shown to be different at the MSP2 locus, this pair of clones was then sampled to determine whether they were identical or different at the MSP1 locus. A sample of 105 of such pairs of clones was obtained. The observed level of gene identity of MSP1 within infections was therefore
calculated as the proportion of pairs which had an identical MSP1 type, among all the 105 pairs sampled.
4.3. Results

4.3.1. Clonal Diversity Within Natural *P. falciparum* Infections

Infections containing more than one clone of *P. falciparum* were frequently encountered in Gambian patients. Figure 11 shows an example of *in situ* resolution of different clones within an isolate (GF88-96), using double-labelled IFA. The different combinations of anti-MSP1 MAbs showed that there were at least three different MSP1 serotypes within isolate GF88-96 (e.g. 10-2B+, 111.4+; 10-2B+, 111.4-; 10-2B-, 111.4+). Additional tests with combinations of other MAbs against MSP1, as well as MSP2 and Exp-1, were also used to detect minority clones, and resolve the antigenic phenotypes of the majority clones within this and other multiple-clone isolates. For example, Table 5 shows the resolved antigenic phenotypes of clones within three multiple-clone isolates.

4.3.2. Mean Number of Clones per Isolate

Among the 355 infections studied, the mean number (*s.e.*) of detectable clones per isolate was 2.02 ± 0.06. Isolates from patients with severe and mild clinical malaria contained similar numbers of clones,
Figure 11. In situ resolution of different antigenic phenotypes of *P. falciparum* in isolate GF88-96, using DNA staining and differential (two-colour) labelling of monoclonal antibodies against MSP1. (a) to (c) show one microscopic field, stained with: (a) DAPI (DNA-specific), (b) MAb 9.5 (isotype IgG2b) + FITC-conjugated anti-IgG2b, (c) MAb 10-2B (isotype IgG2a) + RITC-conjugated anti-IgG2a. In this field, each of the schizonts is positive with both MAbs. (d) to (f) show a second field, stained with: (d) DAPI, (e) MAb 12.1 (isotype IgG1) + FITC-conjugated anti-IgG1, (f) MAb 10-2B + RITC-conjugated anti-IgG2a. In this field, each schizont is positive with either one MAb or the other. (g) to (j) show a third field, stained with: (g) DAPI, (h) MAb 111.4 (isotype IgG1) + FITC-conjugated anti-IgG1, (j) MAb 10-2B + RITC-conjugated anti-IgG2a. In this field one of the schizonts is positive with both MAbs, and the others are positive with one or other of the MAbs.
Table 5. Examples of the resolution of MSP1, MSP2, and Exp-1 serotypes in multiple-clone *P. falciparum* infections

<table>
<thead>
<tr>
<th>MAb Epitopes</th>
<th>MSP1</th>
<th>MSP2</th>
<th>Exp-1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>9.8, 12.4, 12.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isolate</td>
<td>6.1, 9.2, 13.2</td>
<td>10.1</td>
<td>12.2, 12.3, 11.4</td>
</tr>
<tr>
<td></td>
<td>8G10/48, 4-4F</td>
<td>8-5D, 13.4</td>
<td>8F6/49</td>
</tr>
<tr>
<td>GF88-96</td>
<td>(i) 45%</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>(ii) 35%</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>(iii) 20%</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>(iv) 8%</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>84%</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>8%</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>8%</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>GF89-23</td>
<td>(i) 60%</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>(ii) 20%</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>(iii) 12%</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>(iv) 8%</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Blank spaces indicate where the presence or absence of an epitope on a particular clone could not be determined.

Epitopes 13.1, 17.1, 7.3, 1-1C, and 34-5, omitted from the table to save space, exhibited an identical serotype distribution to epitope 6.1.
with a mean of $2.11 \pm 0.20$ (N=35) and $2.04 \pm 0.10$ (N=118), respectively. Figure 12 shows the frequencies of isolates containing 1, 2, 3, or 4 detectable clones. There were no significant differences in the mean number of parasite clones according to the age of patients (Table 6) or parasitaemia (Table 7).

The mean number varied slightly between seasons of *P. falciparum* transmission (Table 8). Isolates collected during October-November 1989 contained a significantly higher mean number of clones compared with those obtained during July-August 1988 (paired t-test, p<0.002) and September-October 1988 (paired t-test, p<0.05). A slight increase observed through the 1988 transmission season was not significant between periods.

4.3.3. Origin of Multiple-Clone Infections

The hypothesis that each parasite clone within a human infection is acquired from a separate mosquito inoculation, and hence that all multiple-clone blood stage infections are due to superinfection, was tested using an analysis of gene identity at the MSP1 locus between pairs of clones from 105 multiple-clone isolates. Each of these isolates contained two parasite clones distinguished by their different MSP2 types, whose MSP1 types were also rigorously determined to be
Figure 12. The percentage of *P. falciparum* isolates containing 1 - 4 clones among severe (N=35) and mild (N=118) clinical cases, and among the total analysed in 1988-1989 (N=355).
Table 6. Mean Number of Clones per Isolate According to Age of Patients

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Mean ± s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 2</td>
<td>2.05 ± 0.13  (n=77)</td>
</tr>
<tr>
<td>2-4</td>
<td>2.02 ± 0.10  (n=120)</td>
</tr>
<tr>
<td>5 - 9</td>
<td>1.90 ± 0.10  (n=111)</td>
</tr>
<tr>
<td>&gt; 10</td>
<td>2.30 ± 0.20  (n=40)</td>
</tr>
</tbody>
</table>

(No significant differences between means; paired t-tests, p > 0.05 for each comparison)
Table 7. Mean Number of Clones per Isolate According to Parasite Density

<table>
<thead>
<tr>
<th>Parasite Density (Number/high power field)</th>
<th>Mean ± s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 50</td>
<td>2.11 ± 0.14 (n=64)</td>
</tr>
<tr>
<td>50 - 199</td>
<td>2.08 ± 0.10 (n=131)</td>
</tr>
<tr>
<td>&gt; 200</td>
<td>1.87 ± 0.09 (n=128)</td>
</tr>
</tbody>
</table>

(No significant differences between means; paired t-tests, p > 0.1 for each comparison)
Table 8. Mean Number of Clones per Isolate During Different Periods

<table>
<thead>
<tr>
<th>Sampling Period</th>
<th>Mean ± s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>July - August 1988</td>
<td>1.67 ± 0.14 (n=45)</td>
</tr>
<tr>
<td>September - October 1988</td>
<td>1.94 ± 0.09 (n=123)</td>
</tr>
<tr>
<td>November - December 1988</td>
<td>2.05 ± 0.18 (n=39)</td>
</tr>
<tr>
<td>October - November 1989</td>
<td>2.25 ± 0.12 (n=112)</td>
</tr>
</tbody>
</table>

(Mean for Oct-Nov 1989 is significantly higher than that for Jul-Aug 1988 and Sept-Oct 1988, p < 0.002 and p < 0.05 for respective paired t-tests. Other comparisons showed no significant differences)
either identical or different to each other. Once the MSP1 alleles were determined to be different, the allelic serotype of the second most predominant clone was not necessarily determined completely.

Following the method of calculation outlined in section 4.2.2., the gene identity of MSP1 among all the isolates from the urban/periurban region, $jx = 0.0650$, with a sampling variance $Vs(jx) = 0.00001856$, and therefore a standard deviation $= 0.00432$.

In 21 of the 105 multiple-clone isolates, suitable for the present analysis, from among the total number of isolates examined, the two clones which were different with respect to MSP2 were identical for MSP1. The clones within the other 84 isolates differed in both MSP1 and MSP2 serotypes. Therefore the observed gene identity of MSP1 between clones in the same isolate $= 21/105 = 0.020$, much higher than randomly expected (the variance of gene identity in this sample could not be calculated using the equation in section 4.2.2., since the allelic serotypes of MSP1 were not always fully characterised in both of the non-identical paired clones). Therefore the random superinfection hypothesis is not sufficient to explain the observed number of isolates in which both clones were identical with respect to MSP1. The likelihood of some pairs of clones being sibling recombinant progeny from the same oocyst is sufficient to explain this.
4.4. Discussion

The double-labelled IFA described here determines the number of *P. falciparum* clones per patient to a high degree of resolution. As the level of detectable MSP1, MSP2, and Exp-1 polymorphism within the parasite population is very high, there is a low probability of randomly encountering 2 different clones which have an identical genotype with respect to the three antigen loci (less than 1% probability for any 2 clones selected randomly from the urban/periurban population: Chapter 3). Therefore, the number of clones missed due to undetected genetic polymorphism is unlikely to be high. However, in low level infections, clones accounting for less than 1% of the total parasite number cannot reliably be detected, and such minor sub-populations were thus excluded from the analysis to standardise the sensitivity of the method between isolates.

The visualisation of individual parasites by double-labelled IFA allows a sensitive determination of the number and relative proportions of different clones within an uncloned isolate. This technique offers certain advantages over methods requiring cellular extracts (e.g. isoenzyme electrophoresis, 2-dimensional gel electrophoresis, restriction fragment length polymorphism analysis, or DNA amplification and allele-
specific oligonucleotide probe analysis). Using extracted protein or DNA material, it may be impossible to resolve multi-locus genotype combinations present within a mixed isolate, without prior \textit{in vitro} cloning of the parasites. This is important, as long term \textit{in vitro} culture may result in clonal selection due to different \textit{in vitro} growth rates (Oduola \textit{et al.}, 1988), and the procedure carries a risk of isolate cross-contamination (Fenton, 1988). An additional advantage of the IFA technique is that it requires only a small volume of parasitised blood, which can be obtained from a fingerprick sample.

The mean value of 2.0 \textit{P. falciparum} clones per patient is similar to the value calculated by Carter \& McGregor (1973) from iso-enzyme frequencies in the Gambia. Other workers have reported the detection of 1 - 4 MSP1 alleles within clinical isolates from Senegal (Scherf, Mattei \& Sarthou, 1991). The isolation of 7 distinct clones from a single Thai isolate, T9 (Thaithong \textit{et al.} 1984), and 6 distinct clones from another Thai patient before and after mefloquine treatment (Pinswasdi \textit{et al.} 1987) lies well outside the range of 1 to 4 clones detectable in this study. The technique of \textit{in vitro} cloning followed by long term culture, as used on the Thai isolates, may be necessary to select clones originally present at undetectable levels. The number of clones per patient may differ
between areas with different levels of endemicity (Creasey et al. 1990), although this has yet to be investigated in a systematic way. It is possible that the Thai isolates described are atypical examples of mixed infection.

The number of clones per infection appears to have no relevance to the development of severe clinical disease. The number was neither related to the age nor the parasitaemia of the patient. The slightly higher mean number in 1989 probably reflects the fact that malaria transmission was particularly high in The Gambia that year. The comparison between different sample periods in 1988 did not reveal significant seasonal variation. Collectively, these results indicate that the small mean number of \textit{P. falciparum} clones per infection is relatively stable within the Gambian study area. Multiple-clone infections clearly occur at frequencies which would allow frequent cross fertilisation and genetic recombination within vector mosquitoes, consistent with the local \textit{P. falciparum} population structure (Chapter 3).

Mathematical models of malaria epidemiology have allowed for superinfection of humans by parasites from different mosquitoes (section 1.2.; MacDonald, 1950; Deitz, Molyneaux & Thomas, 1974; Deitz, 1988). An attempt was made to address the question of the origin of multiple-clone infections using the analysis of MSP1
identity, testing for the identity at this locus between clones within such infections. Multiple-clone infections cannot be explained as a result of superinfection alone. The analysis showed that the level of observed identity at the MSP1 locus was too high to be explained by a random acquisition of genotypes from the parasite population.

Therefore, to explain the higher than expected identity at the MSP1 locus within infections, it is concluded that at least some multiple-clone infections are the result of multiple-clone inoculation from single mosquitoes carrying sibling recombinant parasites. Within a hyperendemic area, where recombination between \textit{P.falciparum} genotypes occurs frequently, a vector mosquito may frequently contain sporozoites of more than one \textit{P.falciparum} genotype.

The frequency and genetic similarity of multiple-clone \textit{P.falciparum} infections in single mosquitoes could be investigated directly by analysis of DNA polymorphism in oocysts and sporozoites isolated from individual wild-caught mosquitoes.
The occurrence of multiple-clone *P. falciparum* haploid blood-stage infections is a prerequisite for cross fertilisation and genetic exchange at the diploid mosquito stage. A mean of two *P. falciparum* clones was detected in the blood of each malaria patient in The Gambia. Using monoclonal antibodies against three polymorphic blood stage antigens, a method of two-colour differential immunofluorescence allowed the resolution of between one and four clones per isolate. The mean number per patient showed no correlation with age, parasitaemia, or disease severity. There was a slight difference in mean number between sample periods, probably reflecting temporal differences in transmission intensity. An analysis of identity at the MSP1 locus, between different clones within isolates, suggests that some multiple-clone infections result from inoculation of sibling recombinant parasites by a single mosquito, and others result from superinfection.

4.5. Summary
5.1. Introduction

Models of malaria transmission (section 1.2.) assume that a malaria vector mosquito can infect only a single person per blood meal. The mosquito inoculates infective sporozoites from its salivary glands, while probing the skin in an attempt to locate a blood capillary. Sporozoites inoculated into the skin, not directly into a blood capillary, were sufficient to induce malaria in a human volunteer (Boyd, 1949b). If a mosquito fails to locate a blood capillary, it may attempt to feed on another individual nearby, and thereby infect more than one individual (Rossignol & Mackay-Rossignol, 1988). Moreover, some blood meals are obtained from more than one individual, due to interrupted feeding (Boreham & Garrett-Jones, 1973). Early experiments on human volunteers demonstrated that a single mosquito could infect more than one person due to interrupted feeding (Boyd, 1949b).

The effects of repeated probing and/or interrupted feeding on the natural transmission of malaria have never been demonstrated. There are inherent difficulties in correlating inoculation and incidence rates by conventional means, in endemic areas (Pull & Grab, 1974). A new approach
to the problem is the characterisation of *Plasmodium falciparum* parasite genotypes from individuals with malaria. A single mosquito infecting two people is expected to inoculate identical parasite genotypes into both of them. If this phenomenon occurs frequently, in an area where malaria transmission occurs at night, children from the same house presenting with malaria at the same time are expected to share identical *P.falciparum* genotypes more often than those living in different houses in the same area.
5.2. Methods

5.2.1. Practical Methods

The study was conducted in an urban/periurban area of The Gambia (section 2.1.1.), where *Plasmodium falciparum* malaria is transmitted by *Anopheles gambiae* sensu lato. *Anopheles gambiae* s.l. bites mainly between 10 pm and 4 am (Holstein, 1954). During July-December 1988, and October-December 1989, *P. falciparum* isolates were collected from 339 malaria patients presenting to the outpatients departments of the Medical Research Council, Fajara, and the Royal Victoria Hospital, Banjul. Among these patients were 17 pairs of siblings presenting together. The parents of 11 of the pairs were questioned about where their children slept in the house. In all cases, both children slept in the same room, and in all except one case, in the same bed. All of the patients lived in the same urban/periurban area.

Heparinised samples of parasitised blood were cultured for 24-48 hours until mature schizonts were obtained (section 2.1.). Multispot slides of schizonts were prepared (section 2.3.1.), for analysis of *Plasmodium falciparum* genotypes by indirect immunofluorescence using a panel of monoclonal antibodies recognising allelic variants of three polymorphic blood stage proteins: MSP1, MSP2, and Exp-1 (Table 1, section 2.3.2.). The proteins are coded by genes
on different chromosomes (Sections 5.5., 5.4., and 5.5.), which segregate at meiosis in the mosquito vector (Figure 1, and section 1.3.2.).

At least 200 schizonts in each isolate were typed with each monoclonal antibody by indirect immunofluorescence (section 2.3.3.). A method of two-colour differential labelling was used to resolve the majority parasite clone in any isolate containing more than one recognisable clone (section 2.3.4.). Only the majority clone in each isolate was included in the statistical analyses.

5.2.2. Statistical Methods

Among the 339 isolates, there were 36 distinguishable allelic serotypes of MSP1, 8 types of MSP2, and 2 types of Exp-1 (Chapter 3). From the observed allelic frequencies at each locus, the expected frequencies of 3-locus genotypes (i.e. combinations of different MSP1, MSP2, and Exp-1 alleles) were calculated for 1988 and 1989 independently, assuming random assortment among these unlinked loci. The observed numbers of different 3-locus genotypes (126 out of 220 isolates sampled in 1988, and 75 out of 117 isolates sampled in 1989) were in accordance with expectations of random recombination in the population (Chapter 3).

The probability (p) that a randomly selected pair of isolates would contain an identical 3-locus genotype was calculated as the sum of the squares of the individual
expected 3-locus genotype frequencies (p=0.00802 for 1988, and 0.00898 for 1989: data in Chapter 3). Assuming a random distribution of genotypes in the local area, the expected number of household pairs of isolates containing an identical *P. falciparum* 3-locus genotype was calculated by the binomial distribution, as follows.

Within each year's sample, the probability of the observed number occurring by chance alone was calculated by the equation:

\[
Pr(i) = \left( \frac{N!}{i!j!} \right) p^i q^j
\]

where, \( p \) = the probability of a randomly selected pair of genotypes being identical, \( q \) = the probability of a randomly selected pair of genotypes being non-identical (i.e. 1-\( p \)), \( N \) = the number of household pairs sampled, \( i \) = the observed number of identical genotype pairs among the household pairs, \( j \) = the number of non-identical genotype pairs among the household pairs, the symbol '!!' denoting the factorial of the number.
5.3. Results

The 3-locus genotype of the majority \textit{P.falciparum} clone within each of the household paired children is given in Table 9. Figure 13 shows the expected numbers of pairs containing identical genotypes as probability distribution histograms for each year. In both years, the expected number was 0 or 1. In 1988, 3 out of 8 pairs contained identical genotypes. The probability of this particular number occurring randomly, \( \text{Pr}(i) = 2.77 \times 10^{-5} \) (Figure 13a). In 1989, 6 out of 9 pairs contained identical genotypes. The probability of this particular number occurring randomly, \( \text{Pr}(i) = 2.55 \times 10^{-10} \) (Figure 13b). The cumulative density function of the probability distributions showed that the observed numbers of identical genotype pairs were higher than randomly expected (\( p<0.00005 \), for each year).
Table 9. 3-locus genotypes of *P. falciparum* isolated from pairs of children who sleep in the same house

(a) 1988

<table>
<thead>
<tr>
<th>Pair</th>
<th>Name</th>
<th>Age/Sex</th>
<th>3-locus Genotype</th>
<th>Sleeping Arrangement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A.N.</td>
<td>3 / F</td>
<td>15-8-2</td>
<td>n.i.</td>
</tr>
<tr>
<td></td>
<td>M.N.</td>
<td>1 / M</td>
<td>28-5-1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>B.J.</td>
<td>5 / M</td>
<td>*10-8-1</td>
<td>n.i.</td>
</tr>
<tr>
<td></td>
<td>F.J.</td>
<td>3 / F</td>
<td>10-8-1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>M.S.</td>
<td>5 / F</td>
<td>15-4-2</td>
<td>n.i.</td>
</tr>
<tr>
<td></td>
<td>A.S.</td>
<td>2 / M</td>
<td>52-2-2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>A.B.</td>
<td>4 / M</td>
<td>*52-2-2</td>
<td>n.i.</td>
</tr>
<tr>
<td></td>
<td>M.B.</td>
<td>6 / F</td>
<td>52-2-2</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>S.J.</td>
<td>3 / F</td>
<td>28-2-2</td>
<td>same bed</td>
</tr>
<tr>
<td></td>
<td>A.J.</td>
<td>1 / F</td>
<td>32-2-2</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>A.T.</td>
<td>5 / F</td>
<td>20-2-1</td>
<td>same bed</td>
</tr>
<tr>
<td></td>
<td>S.T.</td>
<td>2 / M</td>
<td>11-1-2</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>F.C.</td>
<td>4 / M</td>
<td>11-7-2</td>
<td>n.i.</td>
</tr>
<tr>
<td></td>
<td>P.C.</td>
<td>&lt;1 / M</td>
<td>12-2-2</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>B.M.</td>
<td>&lt;1 / F</td>
<td>*12-2-2</td>
<td>same bed</td>
</tr>
<tr>
<td></td>
<td>K.M.</td>
<td>3 / M</td>
<td>12-2-2</td>
<td></td>
</tr>
</tbody>
</table>
3-locus genotype numbers refer to allelic serotypes of MSP1 (Figure 6), MSP2 (Figure 7), and Exp-1 (1 = positive for epitope 5.1, 2 = negative for epitope 5.1), respectively.

Asterisks denote the pairs which share an identical *P. falciparum* genotype

n.i. = no information
Figure 13. Observed (arrows) versus expected (histograms) numbers of identical \textit{P.falciparum} 3-locus genotype pairs, among pairs of children sleeping in the same house. (a) 8 household pairs in 1988 (expectation: binomial distribution of $P=0.00802$ and $N=8$), (b) 9 household pairs in 1989 (expectation: binomial distribution of $P=0.00898$ and $N=9$). Details on statistics are given in section 5.2.2.
(a) 8 Household Pairs Sampled (1988)

(b) 9 Household Pairs Sampled (1989)
5.4. Discussion

Approximately 50% of pairs of children from the same household shared an identical 3-locus *P. falciparum* genotype, compared with approximately 1% of pairs randomly chosen from the study area as a whole.

This exceptionally high frequency of genetically identical parasite pairs among patients from the same household can be explained in one of two ways: (a) single mosquitoes frequently infect two or more people on the same occasion, due to interrupted feeding and/or repeated probing, or (b) extreme clustering of mosquitoes carrying identical *P. falciparum* genotypes exists within households, due to a lack of dispersal during the whole period of development of the parasite in the mosquito.

There is no evidence for the second explanation. *Anopheles gambiae* s.l. mosquitoes are unlikely to remain clustered in a household during the 10-14 day developmental cycle of the parasite, as they lay eggs at outdoor breeding sites approximately every 4 days (Holstein, 1954).

Therefore, single mosquitoes must frequently be responsible for infecting more than one individual with malaria. Haptoglobin typing of blood meals revealed that 3-10% of human blood meals taken by *Anopheles gambiae* s.l. in Africa are acquired from two or more individuals (Boreham *et al.* 1979; Port, Boreham & Bryan, 1980). The observed frequency of genetic identity between household paired
isolates was approximately 50%, suggesting that mosquito probing activity prior to feeding is also important.

Field studies of pre-feeding probing behaviour by malaria vectors have not been attempted, and until now the effects on malaria transmission of vectors probing more than one individual have been unconfirmed. Interestingly, experimental studies have shown that sporozoite-infected mosquitoes spend more time than non-infected ones in probing the skin (Rossignol, Ribeiro & Spielman, 1984). There is probably great diversity in both mosquito probing behaviour and interrupted blood feeding, depending on factors such as host irritability (Lenahan & Boreham, 1976) as well as salivary gland pathology (Rossignol et al., 1984), and therefore such behaviour may be difficult to quantify.

These aspects of mosquito behaviour will tend to increase malaria transmission, as a larger number of people receive infective bites, increasing the basic reproductive rate of the parasite. The results presented in this chapter strongly suggest that the effects of single mosquitoes biting two or more individuals are important, and that present models of malaria transmission therefore underestimate the vectorial capacity of vector populations.
5.5. Summary

*Plasmodium falciparum* isolates were obtained from 17 pairs of Gambian children, each pair living in the same house, presenting with malaria at the same time. Frequencies of allelic serotypes of three polymorphic blood stage proteins (MSP1, MSP2, and Exp-1) were previously determined in the local area, and the probability of a random pair of isolates containing an identical genotype was calculated to be less than 0.01. However, three out of eight household pairs in one year, and six out of nine pairs in the next year, contained identical *P.falciparum* genotypes, a much higher frequency than expected randomly (p<0.00005, for each year). This finding is discussed in terms of the probable contribution of single mosquitoes infecting more than one person.
Chapter 6: Intragenic Recombination and Non-random Associations Between Polymorphic Domains of the Precursor to the Major Merozoite Surface Antigens

6.1. Introduction

The precursor to the major merozoite surface antigens of Plasmodium falciparum (MSP1, PMMSA, or p190) is polymorphic in natural populations (section 5.2.). Allelic polymorphism of the MSP1 gene has been analysed by comparisons of DNA sequences from different parasite isolates and clones (Weber et al. 1986; Tanabe et al. 1987; Peterson et al. 1988). Although certain domains are highly conserved, long sections of the gene exist as one or the other of two essentially dimorphic sequences, and additional allelic polymorphism is a result of intragenic recombination events at sites near the 5' end of the sequence (Tanabe et al. 1987). Many different sequences could theoretically have been produced by recombination events between two parental alleles. Additional allelic sequences presented in this chapter support this conclusion.

At the protein level, antigenic polymorphism of MSP1 has been demonstrated using a panel of variant-specific MAbs (Chapter 3). Approximate locations of certain polymorphic epitopes have been determined, as some of the MAbs recognise fusion proteins encoded by
fragments of the MSP1 gene (Lyon et al. 1986; Gentz et al. 1988; Früh et al. 1991), or naturally occurring processing fragments of the native protein (Howard et al. 1986; Holder et al. 1987; McBride & Heidrich, 1987; Holder, 1988). An alternative approach is to identify sequence polymorphisms which correlate with serological differences among allelic types of the protein. In this chapter, putative locations of different polymorphic epitopes are proposed on the basis of correlation with sequence polymorphisms.

Finally, MSP1 variants of *P. falciparum* from a large number of clinical isolates from The Gambia, Nigeria, and Brazil were serotyped for various polymorphic epitopes. Associations among the epitopes were then analysed statistically, and are discussed in view of the intragenic recombination hypothesis. Non-random associations between epitopes on different domains of MSP1 suggest that, though intragenic recombination may occur, strong intragenic disequilibria exist within MSP1 in natural populations of *P. falciparum*. 
6.2. Methods

*P. falciparum* clones T9/96 and T9/101, expressing different MSP1 types (Thaithong et al. 1984; McBride et al. 1985), were obtained from the WHO Registry of Standard Malaria Strains at Edinburgh University. Frozen stabilates were thawed and cultured *in vitro* (sections 2.1.6., 2.1.7., and 2.1.8.) for preparation of parasite DNA. A wild *P. falciparum* isolate GF88-160 was obtained from 10 ml of venous blood from a Gambian patient with 5% parasitaemia, and cultured *in vitro* for 48 hours until mature schizonts developed. Schizont-infected erythrocytes were concentrated by Plasmagel (section 2.1.8.).

Parasite genomic DNA was extracted (section 2.2.1.) and the polymerase chain reaction (PCR: Saiki et al. 1985) was used to amplify a polymorphic region of the MSP1 gene corresponding to base pairs 774 - 1087 of the coding sequence of the MAD 20 allele (Tanabe et al. 1987). Synthetic 24-mer (5'-TGAAGGAGTAAAGAAAACAATTGA-3') and 28-mer (5'-TCTAATTCAAGTGATCGTAGTAAATAAC-3') amplification primers corresponded to conserved flanking sequences (section 2.2.2.).

Amplified DNA was purified (sections 2.2.3. and 2.2.4.), and kinased (section 2.2.5.) for insertion into the Smal site of bacteriophage lambda M13.
(sections 2.2.6. and 2.2.7.). The recombinant M13 was used to transfect \textit{E. coli} TG-1 host cells (sections 2.2.8. and 2.2.9.), which were grown to produce single-stranded DNA (section 2.2.10.). The single-stranded DNA was used as template for dideoxy sequencing of the cloned insert (sections 2.2.11. and 2.2.12.).

\textit{P. falciparum} isolates were obtained from 445 malaria patients in the Gambia in 1988 and 1989, 60 patients in Ibadan, Nigeria, in 1989, and 62 patients in the Amazon basin, Brazil during 1983-1989 (section 2.1.1.). Parasites were cultured to schizonts (sections 2.1.3. and 2.1.7.), and multispot slides of schizonts prepared for immunofluorescence typing (section 2.3.1.). Slides were also prepared from culture-adapted clones and isolates (T9/96, T9/101, MAD20, FC27, RO-33, Camp, Palo Alto, K1, and Wellcome), obtained from the WHO Registry of Standard Malaria Strains, University of Edinburgh.

Each isolate was tested with 19 anti-MSP1 MAbs (Table 1, section 2.3.2.). Three of the MAbs recognised conserved epitopes, and 16 recognised epitopes on polymorphic domains of MSP1. Reactivities of at least 200 schizonts in each isolate were scored for by indirect immunofluorescence microscopy (section 2.3.3.). Two-colour fluorescence analyses, using pairs of MAbs with different isotypes, and isotype-specific FITC- and RITC-conjugated second antibodies (Southern
Biotechnology Associates Inc.), were performed to resolve the MSP1 profile of the majority parasite clones within certain mixed isolates (section 2.3.4.).

Statistical associations between polymorphic epitopes on different domains of MSP1 in each country were determined by chi-square tests on 2 x 2 contingency tables, which contained the numbers with both, neither, or only one of a given pair of epitopes. Disequilibrium measures, D, and D', between each pair of epitopes, were calculated as for disequilibrium between gene loci (section 3.2.2.).
6.3. Results

6.3.1. Nucleotide Sequences and Putative Intragenic Recombination Sites Within a Region of the MSP1 gene

Figure 14 compares nucleotide sequences of a PCR-amplified region of the MSP1 gene from parasite clones T9/96 and T9/101, and an uncloned isolate GF88-160 (containing only one MSP1 serotype), with the previously published MAD20 and Camp sequences (Tanabe et al. 1987).

Each of the three newly presented sequences differs from those previously published, but each can be explained by intragenic recombination events between parental sequences identical to the Camp and MAD20 sequences, except for two base pair differences. One of these differences, thymidine at position 977 in the T9/96 sequence, is previously undescribed in other MSP1 alleles, and might conceivably represent a PCR misincorporation (four independent clones, two forward strand and two reverse strand, were obtained and sequenced from one PCR amplification mix). The T9/96 sequence is identical to Camp upstream of bp 860 and to MAD20 downstream of bp 926 (bp 860-926 marked as recombination site 2 in Figure 14). The T9/101 sequence is identical to MAD20 upstream of bp 993 and to Camp downstream of bp 999 (bp 993-999 marked as site 3). The
Figure 14. Nucleotide sequences of MSP1 from clones T9/96 and T9/101 and an uncloned isolate GF88-160 show previously undescribed intragenic recombination within a region corresponding to base pairs 774-1087 of the MAD 20 coding sequence (Tanabe et al. 1987). Nucleotide differences in the Camp (Weber et al. 1986) and MAD 20 sequences are shown for comparison. The T9/96 sequence is shown in full, together with differences in the other sequences: identities are shown as (.) and deletions as (-). Horizontal bars 1 to 3 indicate regions containing putative recombination sites. An asterisk marks a position at which the T9/96 sequence contains a previously undescribed nucleotide substitution.
<table>
<thead>
<tr>
<th>Site 1</th>
<th>Site 2</th>
<th>Site 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>T9-96</strong></td>
<td><strong>TC</strong>AATAGATGCAGATAGTGRGARAGARAGAAAAAAATATACCGTCTATATGAT</td>
<td><strong>CG</strong>ATTGACACCTTTAAAARATGAAACATAATTTACTGAGATATAGATAAA</td>
</tr>
<tr>
<td><strong>T9-101</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>GF88-160</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>MAD20</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>CAMP</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>SITE 1</strong></td>
<td><strong>TT</strong>TCTATTCAATACATACATAGAGCAGACATATTTAATAAGC6TTTTAGAAHAA</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>SITE 2</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>SITE 3</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>T9-96</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>T9-101</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>GF88-160</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>MAD20</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>CAMP</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>T9-96</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>T9-101</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>GF88-160</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>MAD20</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>CAMP</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>T9-96</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>T9-101</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>GF88-160</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>MAD20</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>CAMP</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>T9-96</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>T9-101</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>GF88-160</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>MAD20</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>CAMP</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>T9-96</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>T9-101</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>GF88-160</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>MAD20</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>CAMP</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>T9-96</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>T9-101</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>GF88-160</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>MAD20</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
<tr>
<td><strong>CAMP</strong></td>
<td>..................................................</td>
<td>..................................................</td>
</tr>
</tbody>
</table>

**T9-96**

**T9-101**

**GF88-160**

**MAD20**

**CAMP**
GF88-160 sequence is identical to Camp upstream of bp 832 and to T9/101 downstream of bp 838 (bp 832-838 marked as site 1). Sites 1, 2 and 3 therefore mark putative sites of intragenic recombination.

6.3.2. Epitope locations deduced from correlations between amino acid sequences and reactivities with MAbs

Figure 15 compares deduced amino acid sequences of the T9/96, T9/101, and GF88-160 alleles to other published sequences of MSP1 (covering residues 263-362 of the MAD20 sequence, Tanabe et al. 1987). The sequences are correlated to serological reactions of the allelic products with MAbs 9.5 or 13.2, and 12.1 or 10-2B.

MSP1 variants positive for epitope 12.1 all share a unique 12 amino acid sequence (PLPENKKKEVEG), while those positive for the alternative epitope 10-2B share an alternative unique 11 amino acid sequence (TLLDKNKKIEE). The putative linear epitopes 12.1 and 10-2B are shown boxed in Figure 15. Of more than 550 P. falciparum isolates tested, no parasites have been seen to be positive for both alternative epitopes. Parasites negative for both MAbs are rare, and the single example which has been sequenced, RO-33, has a third alternative sequence in the corresponding region.

Epitopes 13.2 and 9.5 are also mutually exclusive.
flAb SEROLOGY
9.5
13.2

9.5 (Q)
MAO 20
FCQ27
11033
CAM P
PALO ALTO
Ki
UELLCO1E

113
OSrITKTYADL
OSHTKTYAOL
OSORKSYAOL
OSORKSYAOL
OSOAKSVHOL
OSOAKSYADL
OSOAKSYAUL

KHRUQFIYLFT
KHRUQHYLFT
KHAUQUYLFT
KHRUR1YLFT
KHRU A HYLFT
KHRLJAHYLLT
KHRUAHYLLT

IKELKYPELF
IKELKYPELF
IKELKYPELF
IKELKYPELF
IKELKYPELF
(KELKYPQLF
IKELKYPQLF

OLTNHrILTLS
OLTHKrILTLS
OLTHHI1LTLC
OLTHHf1LTLC
OLTNHI1LTLC
OLTUHULTLC
OLTHHI1LTLC

162
KHUEJGFKYLI
KHUOGFKYLI
()HIHGFKYLI
DHIHGFKYLI
IJNI HGFKYLI
OMIHGFKYLI
DHIHGFKYLI

+

-

+
+

-

-

+

-

-

+

-

+

13.2 (R)
1111020
FCO27
11033
CAN P
PALO ALTO
K
UELLCOI1E

163
OGVEEINELL
OGVEEUIELL
DGYEEIHELL
OGYEEIIIELL
OGYEEINELL
OGYEEIHELL
EJGYEEIHELL

YKLHFYYOLL
YKLHFYYDLL
YKLHFYFOLL
YKLtIFYFDLL
YKLHFYFDLL
YKLHFYFOLL
YKLHFYFOLL

RAKLHORCAH
RAKLNOACAH
RAKLIIOUCAH
RAKLMOUCAN
RAKLHOUCAU
RAKLHNUCAH
RAKLtIOUCAH

SYCQIPFHLK
SYCQIPFHLK
OYCQIPRILK
OYCQIPFIILK
OYCQIPFHLK
OVCQIPFHLK
OYCQIPFHLK

212
IRAMELDULK
IRAMELOULK
IRAHELOULK
IRAHELOULK
IRAHELDULK
IARHELDULK
IRAHELDULK

MA020
FCO27
11033
CAMP
PALO ALTO
K
UELLCOI1E

213
KIUFGYRKPL
KIUFGYRKPL
KLUFGYRKPL
KLUFGYAKPL
KLUFGYRKPL
KLUFGYRKPL
KLUFGYRKPL

DNIKOHUGKI1
OHIKONtJGKII
OFIKDH(JGKI1
OHIKOHUGKI1
DtlIKOH(JGK11
DNIKOHUGKII
DHI KOHtJGKI1

EDYIKKHKTT
EOYIKKHKTT
EOYIKKHKTT
EOYIKKHKTT
EDYIKKHKTT
EOYIKKHKKT
EDY IKKHKKT

IANHiELIEC
IRtIIHELIEG
IRHIHELIEG
IRHIHELIEG
IAHINELIEG
IEHINELIEE
lEN IHELI EE

262
SKKTIDQIIKH
SKKTIOQHKH
SKKTIDQHKH
SKKTIOQHKN
SKKTIOQHKH
SKKTIDKHKN
SKKTIOKNKH

9,5
I1A020
FCO27
T9-96
11033
T9-101
6F80-160
CAMP
PALO ALTO
KI
UELLCOI1E

263
AOHEEGKKKL
ROHEEGKKKL
ROHEEGKKKL
AOHEEGKKKL
RCHEEGKKKL
AOHEEGKKKL
ROtIEEGKKKL
AOHEE6KKKL
RTKEEEKKKI
RTKEEEKKKL

YQAQYHLFIV
YQRQVIILFIV
'i'QRQYOLSIY
YQRQYOLFIV
YQAQYHLFIY
YQRQYOLFIY
YQAQYDLSIY
YQRQYOLSIY
YQRQYOLFIY
YQAQYOLSI?

HKQL
HKQL
tIKQL
NKQL
IiKQL
HKOL
IIKQL
HKQL
tIKQL
tIKQL

Q
0
E
Q
Q
Q
E
E
E
E

lThb SEROLOGY
9.5
13.2

(0)
ERHNL
ERHHL
ERHUL
EAHHL
ERHHL
EAHHL
ERHHL
ERHHL
ERHHL
EAHHL

ISULEKRIOT
ISULEKRIOT
ISULEKRIOT
ISULEKRIOT
ISIJLEKRIOT
ISULEKRIDT
ISULEKRIOT
ISULEKRIOT
ISIJIEKRICT
ISULEKRIOT

312
LKKHEHIKKL
LKKNEHIKKL
LKKMEHIKKL
LKKHEHIKKL
LKKHEHIKKL
LKKHEHIKKL
LKKHEHIKEL
LKKHEHIKEL
IKKHEHIKEL
LKKHEHIKEL

+

-

+

-

-

+

+

-

-

+

-

+
+

-

1:3.2 (E)
IlRb SEROLOGY
12.1
10-20

12.1
1111020
FCO27
T9-96
R033
19-101
GF68-160
CAf1P
PALO ALTO
KI
UELLCOIIE

313
LEDIDKIKTD
LEDIOKIKTD
LED IOKIKTD
LEDIOKIKIO
LEOIOKIKTO
LEDIOKIKID
LOKIHEIK.H
LDKIHEIK.H
LOKIHEIK.H
LDKItiEIK.tI

REPIPTTGSKP
HEHPTTGSKP
REKLTTGSKP
REKPTTGUt1Q
REKPTTGSKP
REKPTTGSKP
PPPRHSGHTP
PPPRHSGHTP
PPPAHSGHTP
PPPRHSGHTP

H
H
H
I
H
H
H
H
H
11

PLPEHKKKEUEGjHEEKIKE
PLPEHKKKEUEGIHEEKIKE
PLPEHKKKEUEGJHEEKIKE
ISLRLEKESR.. HEEKIKE
TLLOKHKK.IEEIHEEKIKE
TU..OKHKK.IEE HEEKIKE
TLLOKHKK.IEE HEEKIKE
TIIOKHKK.•IEE HEEKIKE
TLIOKHKK.IEE HEEKIKE
TLLDKHKK.IEE HEEKIKE
10-26

362
IRKTIKFHIO
IAKTIKFHID
IRKTIKFHIO
IRKTIKFHID
IRKTIKFHIO
IRKTIKFtIIO
IAKTIKFtIIO
IRKTIKFHIO
IRKTIKFHIO
IRKTIKFHIO

+

-

+

-

-

-

-

+

-

+
I


Parasites positive for epitope 13.2 have glutamate at residue 287 (numbered according to the MAD 20 sequence), while those positive for the alternative epitope 9.5 have glutamine at residue 287. However, the sequence from isolate GF88-160, which is negative for both 13.2 and 9.5, also has glutamine at this residue. Therefore, the 9.5 epitope cannot be explained by this residue alone. The 9.5 epitope may be conformational, requiring glutamine in this position as a part of the epitope. Residue 127, upstream of the region cloned and sequenced here, could also be important for the epitope, since parasites positive for 9.5 have glutamine and those positive for 13.2 have arginine (Figure 15).

Figure 15 also summarises antigenic consequences of the recombination events within the region (Figure 14). The different epitope combinations of e.g. the T9/96 and T9/101 proteins (13.2+ 12.1+, and 9.5+ 10-2B+ respectively) result from different recombinations between the MAD20 (9.5+ 12.1+) and Camp (13.2+ 10-2B+) sequences.

Figure 16 illustrates the approximate locations of the above epitopes, relative to other epitopes which are used below to characterise allelic serotypes of MSP1 on parasites from clinical isolates.
Figure 16. Alternative epitope specificities at different domains of MSP1 are shown boxed. MSP1 sequence polymorphism is shown schematically divided into 17 domains or 'blocks', according to Tanabe et al. (1987). Unshaded blocks are least polymorphic, hatched blocks are more so, and fully shaded blocks are most polymorphic (section 1.5.3.). The locations of epitopes 12.1, 10-2B, 13.2, and 9.5 are discussed in the text. The specificities of the other MAbs are listed in Table 1.
6.3.3. Allelic Serotypes of MSP1 in Endemic Populations of *P. falciparum*

MSP1 serotypes of parasites in clinical isolates of *P. falciparum* were defined as different combinations of polymorphic epitopes recognised by MAbs described in Table 1 and Figure 16. Many isolates contained more than one *P. falciparum* clone, and in such isolates the MSP1 serotype of the majority clone was resolved by two-colour immunofluorescence, as described in Chapter 3. The minority clones (Chapter 4) were not included in the present analysis.

Figure 17 shows the frequencies of 39 different MSP1 serotypes identified in the Gambia, Nigeria, and Brazil. It is assumed that each different serotype represents a different allelic form of MSP1, since the antigen is encoded by a single locus (Tanabe *et al.* 1987), and haploid parasite clones retain their serotype identity during long-term *in vitro* culture (McBride *et al.* 1985). All schizonts in every isolate carried the conserved epitopes recognised by MAbs 9.8, 12.4, and 12.8 (not shown on Figure 17).
Figure 17. MSP1 allelic serotype frequencies in The Gambia, Nigeria, and Brazil. Each serotype is defined by a unique combination of polymorphic epitopes defined by monoclonal antibodies (MAbs), shown at the bottom of the Figure. The serotype numbers (1-52) are as defined in Chapter 3. Reactivities with MAbs recognising domain 2 are summarised as follows: $\alpha = 3D3^+, 12.2^+$; $\beta = 3D3^+, 12.2^-$; $\gamma = 3D3^-, 12.2^+$; $\delta = 3D3^-, 12.2^-$. MAbs 1-1C, 34-5, 7.3, 13.1, and 17.1, which were tested on all the Nigerian and Gambian isolates, showed identical serotype distribution to MAb 6.1, and are omitted from the Figure for the sake of clarity. Two isolates in Nigeria contained a rare phenotype 12.1-, 10-2B-, otherwise similar to serotype 20 or 28, and one Gambian isolate contained the phenotype 12.1-, 10-2B-, otherwise similar to serotype 5 or 13 (phenotypes not included in the Figure). The Brazilian isolates were not tested with MAbs 9.7 or 10.3, and only 28 of the 62 Brazilian isolates were tested with MAb 10-2B, although all those negative for MAb 12.1 are shown here as '10-2B+.'
6.3.4. Non-random Associations Between Polymorphic Epitopes of MSP1

Certain pairs of epitopes are mutually exclusive, never occurring together on the same parasite. This is expected for epitopes determined by alternative variant sequences at the same domain, e.g. 9.5 and 13.2 (domain 3) or 12.1 and 10-2B (domain 4). However, epitopes 12.2 and 3D3 (both on domain 2) are not mutually exclusive, being detected either singly or together on the same parasite. The different combinations of the two epitopes within domain 2 suggest that intragenic recombination events within domain 2 gave rise to recombinant proteins. Similarly, different combinations among epitopes on domain 3 (13.2 or 9.5), and domain 4 (12.1 or 10-2B) are interpreted as being due to intragenic recombination between these domains (as detailed in sections 6.3.1. and 6.3.2.). Interestingly, the MAbs recognising epitopes on domains 6-16 (Figure 16) grouped into 2 distinct alternative specificities, suggesting that there is no intragenic recombination in the corresponding portion of the sequence (6.1, 13.1, 17.1, 7.3, 1-1C and 34-5 recognising the KL/Wellcome type, 9.2, 9.7 and 10.3 recognising the MAD20 type).

Significant non-random associations were detected between epitopes on different domains separated by known sites of intragenic recombination. For example,
13.2 occurred more frequently together with 10-2B, while 9.5 occurred more frequently with 12.1. Tables 10 - 12 show the measures of disequilibrium, D, and D', between each pair of epitopes, together with the chi-square test for significant non-random association, within the above-mentioned samples from The Gambia, Nigeria, and Brazil. Figure 18 summarises the significant associations between epitopes. The large number of strong associations between epitopes on different domains indicates that, despite the occurrence of intragenic recombination between these domains, recombinant alleles do not exist at equilibrium frequencies in natural populations. Mechanisms to explain this are considered in the Discussion (section 6.4.).

Most of the statistical associations were similar in the three countries, with no association being reversed (from positive to negative, or vice versa) in one country compared to another. However, a strong positive association between epitopes 12.1 (domain 4) and 111.4 (domain 16/17) was observed only in Brazil. The absence of statistical association between epitopes 13.2 and 10-2B in Brazil (in contrast with The Gambia and Nigeria) may be due to the fact that only 28 of the Brazilian isolates were tested for 10-2B, so the statistical power of the chi-square test was lower.

Epitopes on domains 6-16 were excluded from the
<table>
<thead>
<tr>
<th>Epitope pair</th>
<th>D</th>
<th>Dmax</th>
<th>D'</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.2 vs. 12.1</td>
<td>-0.0652</td>
<td>0.1731</td>
<td>-0.377</td>
<td>33.5 ***</td>
</tr>
<tr>
<td>13.2 vs. 10-2B</td>
<td>0.0663</td>
<td>0.1731</td>
<td>0.383</td>
<td>34.6 ***</td>
</tr>
<tr>
<td>13.2 vs. 12.2</td>
<td>0.1113</td>
<td>0.1562</td>
<td>0.712</td>
<td>102.3 ***</td>
</tr>
<tr>
<td>13.2 vs. 3D3</td>
<td>0.0521</td>
<td>0.0993</td>
<td>0.525</td>
<td>30.2 ***</td>
</tr>
<tr>
<td>13.2 vs. 111.4</td>
<td>0.0176</td>
<td>0.2154</td>
<td>0.082</td>
<td>2.26</td>
</tr>
<tr>
<td>9.5 vs. 12.1</td>
<td>0.0521</td>
<td>0.2093</td>
<td>0.254</td>
<td>22.5 ***</td>
</tr>
<tr>
<td>9.5 vs. 10-2B</td>
<td>-0.0534</td>
<td>0.2108</td>
<td>-0.253</td>
<td>23.6 ***</td>
</tr>
<tr>
<td>9.5 vs. 12.2</td>
<td>-0.1208</td>
<td>0.1230</td>
<td>-0.982</td>
<td>126.4 ***</td>
</tr>
<tr>
<td>9.5 vs. 3D3</td>
<td>-0.0782</td>
<td>0.0782</td>
<td>-1.000</td>
<td>71.1 ***</td>
</tr>
<tr>
<td>9.5 vs. 111.4</td>
<td>-0.0213</td>
<td>0.1695</td>
<td>-0.126</td>
<td>3.44</td>
</tr>
<tr>
<td>12.1 vs. 12.2</td>
<td>-0.0520</td>
<td>0.1082</td>
<td>-0.481</td>
<td>24.7 ***</td>
</tr>
<tr>
<td>12.1 vs. 3D3</td>
<td>-0.0148</td>
<td>0.0688</td>
<td>-0.216</td>
<td>2.71</td>
</tr>
<tr>
<td>12.1 vs. 111.4</td>
<td>-0.0210</td>
<td>0.1646</td>
<td>-0.128</td>
<td>3.55</td>
</tr>
<tr>
<td>10-2B vs. 12.2</td>
<td>0.0504</td>
<td>0.1089</td>
<td>0.483</td>
<td>23.2 ***</td>
</tr>
<tr>
<td>10-2B vs. 3D3</td>
<td>0.0153</td>
<td>0.0692</td>
<td>0.221</td>
<td>2.87</td>
</tr>
<tr>
<td>10-2B vs. 111.4</td>
<td>0.0198</td>
<td>0.1501</td>
<td>0.132</td>
<td>3.13</td>
</tr>
<tr>
<td>12.2 vs. 3D3</td>
<td>0.0382</td>
<td>0.0629</td>
<td>0.607</td>
<td>18.8 ***</td>
</tr>
<tr>
<td>12.2 vs. 111.4</td>
<td>0.0119</td>
<td>0.1782</td>
<td>0.067</td>
<td>1.18</td>
</tr>
<tr>
<td>3D3 vs. 111.4</td>
<td>-0.0036</td>
<td>0.0867</td>
<td>-0.042</td>
<td>0.146</td>
</tr>
</tbody>
</table>

***, p<0.001    **, p<0.01    *, p<0.05
### Table 11. Intragenic disequilibria between epitopes on MSP1 in Nigeria

<table>
<thead>
<tr>
<th>Epitope pair</th>
<th>D</th>
<th>Dmax</th>
<th>D'</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.2 vs. 12.1</td>
<td>-0.0894</td>
<td>0.1228</td>
<td>-0.728</td>
<td>9.63 ***</td>
</tr>
<tr>
<td>13.2 vs. 10-2B</td>
<td>0.0989</td>
<td>0.1322</td>
<td>0.748</td>
<td>11.6 ***</td>
</tr>
<tr>
<td>13.2 vs. 12.2</td>
<td>0.1317</td>
<td>0.1983</td>
<td>0.664</td>
<td>24.4 ***</td>
</tr>
<tr>
<td>13.2 vs. 3D3</td>
<td>0.0314</td>
<td>0.1313</td>
<td>0.239</td>
<td>1.94</td>
</tr>
<tr>
<td>13.2 vs. 111.4</td>
<td>-0.0205</td>
<td>0.1039</td>
<td>-0.197</td>
<td>0.538</td>
</tr>
<tr>
<td>9.5 vs. 12.1</td>
<td>0.0775</td>
<td>0.2167</td>
<td>0.358</td>
<td>4.34 *</td>
</tr>
<tr>
<td>9.5 vs. 10-2B</td>
<td>-0.0833</td>
<td>0.2333</td>
<td>-0.357</td>
<td>6.70 **</td>
</tr>
<tr>
<td>9.5 vs. 12.2</td>
<td>-0.1500</td>
<td>0.1500</td>
<td>-1.000</td>
<td>25.7 ***</td>
</tr>
<tr>
<td>9.5 vs. 3D3</td>
<td>-0.0917</td>
<td>0.0917</td>
<td>-1.000</td>
<td>13.5 ***</td>
</tr>
<tr>
<td>9.5 vs. 111.4</td>
<td>0.0167</td>
<td>0.1833</td>
<td>0.091</td>
<td>0.287</td>
</tr>
<tr>
<td>12.1 vs. 12.2</td>
<td>0.0800</td>
<td>0.1606</td>
<td>0.498</td>
<td>7.45 **</td>
</tr>
<tr>
<td>12.1 vs. 3D3</td>
<td>0.0128</td>
<td>0.0794</td>
<td>-0.161</td>
<td>0.266</td>
</tr>
<tr>
<td>12.1 vs. 111.4</td>
<td>0.0578</td>
<td>0.2078</td>
<td>0.278</td>
<td>3.52</td>
</tr>
<tr>
<td>10-2B vs. 12.2</td>
<td>0.0900</td>
<td>0.1322</td>
<td>0.681</td>
<td>9.30 **</td>
</tr>
<tr>
<td>10-2B vs. 3D3</td>
<td>0.0189</td>
<td>0.0856</td>
<td>0.221</td>
<td>0.574</td>
</tr>
<tr>
<td>10-2B vs. 111.4</td>
<td>-0.0789</td>
<td>0.1956</td>
<td>-0.403</td>
<td>6.46 **</td>
</tr>
<tr>
<td>12.2 vs. 3D3</td>
<td>0.0117</td>
<td>0.1283</td>
<td>0.091</td>
<td>0.260</td>
</tr>
<tr>
<td>12.2 vs. 111.4</td>
<td>0.0067</td>
<td>0.1800</td>
<td>0.037</td>
<td>0.055</td>
</tr>
<tr>
<td>3D3 vs. 111.4</td>
<td>-0.0006</td>
<td>0.0672</td>
<td>-0.009</td>
<td>0.001</td>
</tr>
</tbody>
</table>

***, p<0.001   **, p<0.01   *, p<0.05
<table>
<thead>
<tr>
<th>Epitope pair</th>
<th>D</th>
<th>Dmax</th>
<th>D'</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.2 vs. 12.1</td>
<td>-0.0413</td>
<td>0.1048</td>
<td>-0.394</td>
<td>2.342</td>
</tr>
<tr>
<td>13.2 vs. 10-2B</td>
<td>0.0357</td>
<td>0.1071</td>
<td>0.333</td>
<td>not valid</td>
</tr>
<tr>
<td>13.2 vs. 12.2</td>
<td>0.0156</td>
<td>0.0474</td>
<td>0.329</td>
<td>not valid</td>
</tr>
<tr>
<td>13.2 vs. 3D3</td>
<td>0.1300</td>
<td>0.1776</td>
<td>0.732</td>
<td>40.10 ***</td>
</tr>
<tr>
<td>13.2 vs. 111.4</td>
<td>-0.0098</td>
<td>0.1209</td>
<td>0.081</td>
<td>0.098</td>
</tr>
<tr>
<td>9.5 vs 10-2B</td>
<td>-0.0714</td>
<td>0.1429</td>
<td>0.500</td>
<td>not valid</td>
</tr>
<tr>
<td>9.5 vs. 12.1</td>
<td>0.0635</td>
<td>0.1260</td>
<td>0.504</td>
<td>5.067 *</td>
</tr>
<tr>
<td>9.5 vs. 12.2</td>
<td>-0.0433</td>
<td>0.0433</td>
<td>-1.000</td>
<td>not valid</td>
</tr>
<tr>
<td>9.5 vs. 3D3</td>
<td>-0.1625</td>
<td>0.1625</td>
<td>-1.000</td>
<td>50.6 ***</td>
</tr>
<tr>
<td>9.5 vs. 111.4</td>
<td>-0.0242</td>
<td>0.1512</td>
<td>0.160</td>
<td>0.082</td>
</tr>
<tr>
<td>12.1 vs. 12.2</td>
<td>-0.0262</td>
<td>0.0262</td>
<td>-1.000</td>
<td>not valid</td>
</tr>
<tr>
<td>12.1 vs. 3D3</td>
<td>-0.0348</td>
<td>0.0983</td>
<td>0.354</td>
<td>2.262</td>
</tr>
<tr>
<td>12.1 vs. 111.4</td>
<td>0.1807</td>
<td>0.1965</td>
<td>0.919</td>
<td>39.10 ***</td>
</tr>
<tr>
<td>10-2B vs. 12.2</td>
<td>0.0536</td>
<td>0.0536</td>
<td>1.000</td>
<td>not valid</td>
</tr>
<tr>
<td>10-2B vs. 3D3</td>
<td>0.0714</td>
<td>0.1429</td>
<td>0.500</td>
<td>not valid</td>
</tr>
<tr>
<td>10-2B vs. 111.4</td>
<td>-0.1964</td>
<td>0.1964</td>
<td>1.000</td>
<td>18.12 ***</td>
</tr>
<tr>
<td>12.2 vs. 3D3</td>
<td>0.0484</td>
<td>0.0484</td>
<td>1.000</td>
<td>not valid</td>
</tr>
<tr>
<td>12.2 vs. 111.4</td>
<td>-0.0015</td>
<td>0.0333</td>
<td>0.045</td>
<td>not valid</td>
</tr>
<tr>
<td>3D3 vs. 111.4</td>
<td>0.0181</td>
<td>0.1134</td>
<td>0.160</td>
<td>0.80</td>
</tr>
</tbody>
</table>

***, $p<0.001$  **, $p<0.01$  *, $p<0.05$
Figure 18. Statistical associations between epitopes at different domains of MSP1, among parasites sampled from The Gambia, Nigeria, and Brazil. A solid line indicates significant positive association, and a broken line significant negative association. The chi-square p values were high, p<0.005 for all significant values, except for 12.2 vs. 12.1, and 10-2B vs. 9.5, in Nigeria (p<0.01), 10-2B vs. 111.4, and 9.5 vs. 12.1 in Nigeria (p<0.05), and 9.5 vs. 12.1 in Brazil (p<0.05). Epitope 12.2 was rare in Brazil, and therefore not included in statistical analyses.
MSP1 domain

2 3 4 16/17

The Gambia
N = 445

Nigeria
N = 60

Brazil
N = 62
chi-square analyses, since one of the two alternative variants (recognised by MAbs 6.1, 13.1, 17.1, 7.3, 1-1C, and 34-5) was present at such a low frequency in each of the countries that the analysis would be invalid. However, the 21 isolates containing such parasites all had an identical profile of epitopes at other domains (i.e. serotype 52 in Figure 17). This was in marked contrast with parasites with the other form of domain 6-16 (recognised by MAbs 9.2, 9.7 and 10.3), which exhibited a very high diversity of epitope combinations at other domains (the remaining 38 different serotypes, Figure 17).
6.4 Discussion

Three sites of intragenic recombination within the MSP1 gene are proposed to explain the newly described allelic sequences, in addition to sites proposed by Tanabe et al. (1987). Identification of the sites confirms and extends the evidence for a clustering of recombination sites within a relatively short 5' region of the gene (covering less than 20% of the total sequence), discussed by Peterson et al. (1988a).

Location of epitope determinants by sequence-serology correlation has been previously undertaken for chicken lysozyme c (Smith-Gill et al. 1982), murine MHC antigens (Landais et al. 1985), and the HLA-DR, -DQ, and -DP loci (Bugawan et al. 1988; Horn et al. 1988; Marsh & Bodmer, 1989). The approach has proved useful, although a potential complicating factor is that a critical conformational change in an epitope is sometimes determined by an amino acid substitution remote from the epitope site (White, Ibrahimi & Wilson, 1978).

Here, allelic sequences within a polymorphic region of MSP1 explain the observed alternate specificities of MAbs 12.1 and 10.2B, and the likely position of the epitopes has been deduced (amino acids 334-345 in domain 4, according to the scheme of Tanabe et al. 1987). Similarly, amino acids 127 and/or 287 in
domain 3 may determine another pair of alternate epitopes, 13.2 and 9.5. The present analysis has also shown that intragenic recombinations can be identified serologically as combinations of the respective epitopes.

If recombination events occur frequently within the MSP1 gene, and the recombinant proteins are not under differential selection, then polymorphic epitopes at different domains of the gene would be expected to assort randomly and their combinations exist at equilibrium frequencies in natural populations. However, the strong non-random associations observed between epitopes on different domains is evidence that intragenic recombinants are not at equilibrium frequencies. Most notably, parasites with one of the variants at domains 6-16 (recognised by MAb 6.1, 13.1, 17.1, 7.3, 1-1C, and 34-5) all had an identical combination of MAb epitopes at other domains. Such non-random associations could arise by either of the following mechanisms.

Firstly, if intragenic recombination is a rare event, random allelic frequency changes in natural populations could explain disequilibria between polymorphic domains of MSP1. Although there are several recombination sites within a short region of the MSP1 gene, it does not follow that recombination at these sites has occurred frequently. Reassortment between
unlinked loci occurs frequently at meiosis (Walliker et al. 1987, and Chapter 3), but it is not possible to estimate the rate of intragenic recombination from present data.

Alternatively, it is conceivable that recombinant alleles differ in 'fitness', so that particular recombinants may be selected even if intragenic recombination is a frequent event. The observed similarities between the epitope associations in three separate parasite populations suggest that unknown selective factors may consistently operate in favour of particular recombinant alleles. The function of MSP1 is unknown, so mechanisms whereby particular alleles have a selective advantage must be purely conjectural.

Polymorphism of MSP1 may be related to recognition of different erythrocyte surface receptors, as suggested by Tanabe et al. (1987). *P. falciparum* isolates differ in their ability to invade glycophorin-deficient or sialic acid-deficient red blood cells in vitro (Mitchell et al. 1986; Hadley et al. 1987; Perkins & Holt, 1988), although it is not known whether this is linked to MSP1 polymorphism (Perkins, 1989). MSP1 can induce a protective anti-parasite immune response (Perrin et al. 1984; Hall et al. 1984; Siddiqui et al. 1987), although there is no clear evidence that polymorphisms are maintained by acquired variant-specific immunity (Chapter 7).
Further studies are therefore required to demonstrate whether the clustering of intragenic recombination sites, and the observed disequilibria between polymorphic domains of MSP1, are a result of selective constraints on the protein.
6.5. Summary

Extensive allelic polymorphism in the *Plasmodium falciparum* precursor to the major merozoite surface antigens (MSP1) is partly due to intragenic recombination events within a short region near the 5' end of the gene. Newly described allelic sequences from this region of the gene are compared to those previously published, revealing additional sites of intragenic recombination. Epitope sites on the protein have been assigned on the basis of correlations between serology and sequence polymorphisms among different allelic types of MSP1. Certain pairs of epitopes, although sited on MSP1 domains separated by known sites of intragenic recombination, are highly significantly associated on parasites in endemic populations. Analyses on 567 wild isolates from The Gambia, Nigeria, and Brazil, reveal that most associations are similar in the three countries. These associations are discussed with respect to the intragenic recombination hypothesis.
Chapter 7: A Longitudinal Study of Defined Plasmodium falciparum Antigen Polymorphisms in a Gambian Population

7.1. Introduction

The existence of antigenic diversity among 'strains' of Plasmodium falciparum was originally suggested to explain the slow rate of acquisition of immunity to malaria by inhabitants of endemic areas, and the apparent susceptibility of 'immune' adults who travelled from one endemic area to another. In experimental infections of non-human primates, acquired immunity appeared to involve a 'strain-specific' component (Sadun et al., 1966; Cadigan & Chiacumpa, 1969; Voller, Green & Richards, 1973), lending support to the importance of antigenic diversity. However, results from human infections were less consistent in this respect (James et al., 1932; Boyd et al., 1936; Bray et al., 1962; Jeffery, 1966; Powell et al., 1972).

Serological characterisation of P. falciparum isolates subsequently demonstrated heterogeneity of parasite soluble products (Wilson et al., 1969; Wilson, 1980), internal and surface antigens (McBride, Walliker & Morgan, 1982) and infected erythrocyte surface components (Hommel, David & Oligino, 1983; Udeinya et al., 1983; Marsh & Howard, 1986). In only one study, on the schizont-infected erythrocyte surface in squirrel monkeys
(Hommel et al., 1983), has serological diversity been linked with variant-specific immunity.

DNA sequencing has elucidated the structural basis for serological polymorphism of some proteins, including the precursor to the major merozoite surface antigens (MSP1: section 1.5.3.), a second merozoite surface antigen (MSP2: section 1.5.4.), and an exported protein (Exp-1: section 1.5.5.). Immunisation with MSP1 elicited a degree of immunity in monkeys (Perrin et al. 1984; Hall et al. 1984; Siddiqui et al. 1987). Immunisation experiments with MSP2 or Exp-1 have not been attempted. However, these antigens may be immunologically important. Exp-1 cross reacts with the NANP(n) repeat epitope of the circumsporozoite protein (Hope et al. 1985; Simmons et al. 1987). Monoclonal antibodies against MSP2 partially inhibit in vitro invasion of erythrocytes (Clark et al. 1989; Saul et al. 1989).

It is not known whether polymorphisms in any of these antigens enable the parasite to evade acquired immune responses. If so, such polymorphisms might be maintained in a parasite population by frequency-dependent selection, rare types having a selective advantage due to lower levels of acquired immunity against them (section 1.5.2.).

In this study, frequencies of variant epitopes on different domains of MSP1, MSP2 and Exp-1 have been determined in P. falciparum isolates collected from an
urban/periurban area of The Gambia, over the period 1982 to 1989. The results are discussed with reference to the hypothesis of frequency-dependant selection by variant-specific immunity.
7.2. Methods

*P. falciparum* isolates were collected from 424 malaria patients presenting to the Outpatients Departments of the Medical Research Council (MRC) Fajara, or the Royal Victoria Hospital, Banjul, during October-December 1982 (N=46), Oct-Dec 1983 (N=31), July-Dec 1988 (N=228), and Oct-Dec 1989 (N=119). All lived within a radius of 7 km in an urban/periurban area (section 2.1.1.). The collection and characterisation of the 1982 and 1983 isolates was performed by Dr. J.S. McBride.

Blood samples were obtained (section 2.1.3.), parasites cultured for 24-48 hours (section 2.1.7.), and multispot schizont slides prepared (section 2.2.1.). Indirect immunofluorescence (section 2.3.3.) was performed on schizonts of each isolate using murine MAbs against MSP1 (MAbs 6.1, 7.3, 13.1, 17.1, 1-1C, 34-5, 9.2, 9.7, 10.3, 13.2, 9.5, 12.1, 10-2B, 12.2, 3D3, 111.4), MSP2 (MAbs 12.3, 12.5, 13.4), and Exp-1 (MAb 5.1). The specificities of these MAbs are listed in Table 1 (section 2.3.2.). The epitope sites of MAbs against MSP1 are shown in Figures 15 and 16 (Chapter 6).
7.3. Results

7.3.1. Scoring of P. falciparum Isolates for Expression of Antigen Epitopes Recognised by Monoclonal Antibodies

At least 200 schizonts from each isolate were scored for reactions with each of 23 MAbs by indirect immunofluorescence. All schizonts in every isolate gave specific reactions with control MAbs 9.8, 12.4, and 12.8. Each of the other MAbs identified some isolates within which all schizonts were positive, some isolates within which all parasites were negative, and some isolates within which only a proportion of schizonts gave specific fluorescence (i.e. multiple-clone infections). Multiple-clone infections are common in the Gambia (Chapter 4). The statistical analyses included as positive all isolates within which the majority (50-100%) of schizonts expressed a given epitope. Although not included in statistical analyses, data on the percentages of isolates within which more than 1% of schizonts expressed a given epitope are also presented.

7.3.2. Comparison of Polymorphic Epitope Frequencies During a Single Transmission Season

Table 13 compares frequencies of different MSP1,
Table 13. Percentage Frequency of Epitopes Marking Polymorphic Domains of MSP1, MSP2, and Exp-1 at Different Periods of the 1988 Transmission Season

Open figures denote the percentage of isolates in which ≥ 50% of schizonts are positive for a given epitope (bracketed figures denote the percentage in which > 1% are positive).

The one significant difference between periods (p<0.05) is underlined.

Omitted from the table are: MAb 13.1, 7.3, 17.1, 34.5, and 1-1C which showed identical specificity to MAb 6.1; MAb 9.7 and 10.3 which showed identical specificity to MAb 9.2; and MAb 12.5, which showed identical specificity to MAb 12.3.
### (a) MSP1

<table>
<thead>
<tr>
<th>Domain</th>
<th>Monoclonal Antibody</th>
<th>July-August (N=51)</th>
<th>Sept-Oct (N=135)</th>
<th>Nov-Dec (N=42)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>12.2</td>
<td>25.5 (37.3)</td>
<td>29.6 (45.2)</td>
<td>26.2 (40.5)</td>
</tr>
<tr>
<td>2</td>
<td>3D3</td>
<td>27.5 (31.4)</td>
<td>20.0 (28.9)</td>
<td>11.9 (26.2)</td>
</tr>
<tr>
<td>3</td>
<td>13.2</td>
<td>52.9 (66.6)</td>
<td>48.2 (63.0)</td>
<td>51.2 (63.4)</td>
</tr>
<tr>
<td>3</td>
<td>9.5</td>
<td>39.2 (45.1)</td>
<td>38.5 (55.0)</td>
<td>43.9 (53.7)</td>
</tr>
<tr>
<td>4</td>
<td>12.1</td>
<td>33.3 (45.1)</td>
<td>32.6 (50.4)</td>
<td>45.2 (59.5)</td>
</tr>
<tr>
<td>4</td>
<td>10-2B</td>
<td>66.7 (72.8)</td>
<td>68.2 (75.6)</td>
<td>54.8 (64.3)</td>
</tr>
<tr>
<td>6-16</td>
<td>6.1</td>
<td>2.0 (2.0)</td>
<td>5.2 (5.2)</td>
<td>2.4 (2.4)</td>
</tr>
<tr>
<td>6-16</td>
<td>9.2</td>
<td>98.0 (98.0)</td>
<td>94.8 (97.0)</td>
<td>97.6 (97.6)</td>
</tr>
<tr>
<td>16-17</td>
<td>111.4</td>
<td>45.1 (58.8)</td>
<td>41.5 (53.4)</td>
<td>52.4 (64.3)</td>
</tr>
</tbody>
</table>

### (b) MSP2

<table>
<thead>
<tr>
<th>Monoclonal Antibody</th>
<th>July-August (N=50)</th>
<th>Sept-Oct (N=133)</th>
<th>Nov-Dec (N=41)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.3</td>
<td>50.0 (62.0)</td>
<td>69.2 (80.4)</td>
<td>78.1 (90.3)</td>
</tr>
<tr>
<td>13.4</td>
<td>14.0 (16.0)</td>
<td>8.3 (13.6)</td>
<td>7.1 (14.3)</td>
</tr>
</tbody>
</table>

### (c) Exp-1

<table>
<thead>
<tr>
<th>Monoclonal Antibody</th>
<th>July-August (N=49)</th>
<th>Sept-Oct (N=133)</th>
<th>Nov-Dec (N=42)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>30.6 (51.0)</td>
<td>30.1 (45.9)</td>
<td>28.6 (50.0)</td>
</tr>
</tbody>
</table>
MSP2, and Exp-1 epitopes among Gambian clinical isolates at the beginning (July-August), middle (September-October), and end (November-December) of the 1988 transmission season. Only epitope 12.3 on MSP2 occurred slightly less frequently during July-August than during September-October or November-December (chi-square, p<0.05, for each comparison). There were no significant differences between periods for frequencies of any other epitopes.
7.3.3. Comparison of Polymorphic Epitope Frequencies During the Period 1982-1989

Figure 19 compares frequencies of the same epitopes as above among *P. falciparum* isolates collected in the same area during the October-December period of 1982 (N=46), 1983 (N=31), 1988 (N=101), and 1989 (N=119). There were no significant changes in the frequencies of most epitopes. The epitope 12.2 on the repeat domain 2 of MSP1 was more frequent in 1983 than in each other year (chi-square, p<0.05), and MAb epitope 12.1 on domain 4 of MSP1 was more frequent in 1982 than in each other year (chi-square, p<0.05).
Figure 19. Percentage of isolates containing parasites expressing different polymorphic epitopes of MSP1, MSP2, and Exp-1 during October–December 1982 (N=46), 1983 (N=31), 1988 (N=101), and 1989 (N=119). Shaded bars indicate the proportion of isolates in which 50–100% of parasites expressed a given epitope, and open bars show the proportion of isolates in which 1–100% of parasites had each epitope. Asterisks mark the two values which were higher in one year (chi-square, p<0.05).
7.4. Discussion

The main finding of the comparison was that the relative frequencies of MSP1, MSP2 and Exp-1 variant epitopes among isolates from Gambian patients remained stable during the 1982-1989 period. The MSP1 epitopes recognised by MAbs 12.1 and 12.2 had a significantly higher frequency in 1982 and 1983 respectively, although these differences are no more than expected by chance (due to the large number of chi-square comparisons made in total). The frequencies were also stable throughout a single transmission season (with only one significant difference among all the chi-square comparisons).

Therefore, the more common variants remained at a higher frequency over several years, and the rare variants remained at a lower frequency. This was particularly notable with respect to the two alternative variants at domain 6-16 of MSP1 (Figure 16, Chapter 6). This 'dimorphic' region of MSP1 extends over most of the sequence, and is represented in the sequence divergence between laboratory isolates Mad20 and K1/Wellcome (Tanabe et al. 1987). The K1/Wellcome-like variant (6.1+, 7.3+, 13.1+, 17.1+, 34-5+, 1-1C+) remained at a very low frequency (< 5%), and the Mad20-like variant (9.2+, 9.7+, 10.3+) maintained a high frequency (> 95%).

Theoretically, frequency-dependent selection can
maintain genetic polymorphisms with either stable or fluctuating variant frequencies (Haldane & Jayakar, 1963; Clarke & O'Donald, 1964). However, if acquired variant-specific immunity is the cause of such selection, it is difficult to explain how this could lead to a stable polymorphism in which two alternative MSP1 variants are maintained at such greatly differing frequencies. Moreover, it is improbable that the parasite population would maintain what could be described as a 'conserved dimorphism' in response to a specific immune response. Extensive sequence polymorphism within T or B cell epitopes would be a more likely result of immune selection (Good et al., 1988). Hypotheses other than frequency dependant selection by acquired variant-specific immunity are therefore required to explain the sequence 'dimorphism' throughout most of the MSP1 gene.

The alternative variants at other domains of MSP1, e.g. 13.2 and 9.5 on domain 3, were present in the population at more similar frequencies, so it is conceivable that these could represent stable equilibrium frequencies under frequency-dependant selection.

By contrast, temporal variation in the prevalence of an S-antigen epitope was recorded in villages of Papua New Guinea (Forsyth et al., 1989). These variations could possibly result from frequency-dependant immune selection (Forsyth et al., 1988). However, as pointed out elsewhere (Chapter 3), random fluctuations in allelic frequencies
are probably inevitable in *P. falciparum* populations in villages, due to genetic drift imposed by a restricted population size. The S-antigen data do not reveal any directional frequency changes, as opposed to random fluctuations.

In this chapter, the frequencies of each epitope have been analysed separately, independently of one another. However, different epitopes on the same protein are likely to be strongly associated genetically or structurally. For example, even though intragenic recombination can occur in the MSP1 gene, strong associations exist in local populations between polymorphic epitopes located on different domains of the protein (Chapter 6). Therefore, allelic serotype frequencies should be considered, in addition to individual polymorphic epitope frequencies. Allelic serotypes have been resolved for the 1988 and 1989 samples, over which period the frequencies remained stable (Chapter 3, Figure 6 and 7).

Three MAbs recognised MSP1 epitopes which were invariant on all schizonts of the 424 isolates studied here, and also among 122 isolates from Brazil and Nigeria (Chapter 6). There is evidence that some conserved regions of MSP1 induce human immune responses. Sinigaglia et al. (1988 & 1990) have described human T and B cell epitopes within domain 3 of MSP1, a relatively conserved region of the sequence. Also, a synthetic peptide vaccine
including a conserved sequence from domain 1 of MSP1 appears to induce a partial protective immunity to malaria in humans (Pattaroyo et al. 1988).

To conclude, the results presented in this chapter suggest that some major antigen polymorphisms are not maintained by frequency-dependant immune selection, particularly the 'dimorphism' which includes most of the sequence of MSP1. Other polymorphisms in MSP1, MSP2, and Exp-1 might be balanced by the effects of variant-specific immunity, although immunological studies using defined recombinant or synthetic peptide antigens are required to characterise such variant-specific immune responses.
7.5. Summary

*Plasmodium falciparum* merozoite surface antigens MSP1 and MSP2, and an exported antigen Exp-1, exhibit allelic polymorphism in natural populations. A hypothesis is that antigen polymorphisms are maintained by frequency-dependent immune selection, rare variants having an advantage over common variants due to a lower level of acquired immunity against them. The frequencies of polymorphic epitopes of MSP1, MSP2, and Exp-1, were determined among isolates from malaria patients in an urban area of The Gambia, during different stages of a transmission season (1988), and different years (1982, 1983, 1988, and 1989). The frequencies remained very stable throughout the period of study, alternative epitope variants remaining either rare or common, without shifts in relative frequencies. The implications of these results on the immune-selection hypothesis are discussed.
Chapter 8: Testing for Associations Between Plasmodium falciparum Merozoite Surface Protein Polymorphisms and Blood Groups in Malaria Patients

8.1. Introduction

*Plasmodium falciparum* merozoite invasion of human erythrocytes is dependent on glycophorin molecules on the erythrocyte surface (Pasvol *et al.* 1982a,b; Facer, 1983), particularly on the O-linked oligosaccharide side chains bearing sialic acid residues (Hermentin *et al.* 1987). However, the degree of this dependence differs among cultured lines of *P. falciparum*, suggesting that there is also a sialic acid-independent pathway of invasion (Mitchell *et al.* 1986; Hadley *et al.* 1987; Perkins & Holt, 1988; Hadley & Miller, 1988). The molecules important in the sialic acid-independent pathway are unidentified, although a trypsin-sensitive component is involved (Hadley & Miller 1988).

Merozoite surface components are probably important in the specific recognition of erythrocyte surface molecules. MSP1 and MSP2 are coded by gene loci at which there are extensive allelic polymorphisms (sections 1.5.3. and 1.5.4.). The significance of these polymorphisms is unclear. However, both proteins contain domains which are 'dimorphic' (at which there are two alternative allelic sequences), and it has been suggested
that such alternative sequences may form receptor structures which recognise different erythrocyte surface molecules (Tanabe et al. 1987). It is possible that parasites with different allelic forms of MSP1 or MSP2 have adapted to invade erythrocytes of different surface phenotype (section 1.5.6.).

Here, this hypothesis was tested with respect to the ABO, MN, and SS blood group systems. The ABO determinants are carbohydrate moieties attached to band 3 (though not exclusively), an erythrocyte transmembrane component reported to be important for *P. falciparum* invasion into erythrocytes (Okoye & Bennett, 1985). The MN determinants are on the peptide chain of glycophorin A, and the SS determinants on the peptide chain of glycophorin B (Issitt, 1976, and 1981).

Blood groups, and *P. falciparum* MSP1 and MSP2 types, were characterised in a large sample of malaria patients, to test for statistical associations which may reveal relationships between erythrocyte and parasite polymorphisms. There was no a priori expectation that any particular epitope of MSP1 or MSP2 would be associated with any particular blood group.
8.2. Methods

Blood samples were obtained from 264 malaria patients presenting to the outpatients departments of the Medical Research Council, Fajara, and the Royal Victoria Hospital, Banjul, in The Gambia (section 2.1.1.), during July-December 1988 and October-December 1989. Erythrocytes were typed for ABO, MN, and S blood group antigens (section 2.4.). There was insufficient time during the study to type erythrocytes for the s determinant by indirect agglutination. Parasites were cultured for 24-48 hours (section 2.1.7.), and multispot schizont slides prepared (section 2.3.1.). Indirect immunofluorescence (section 2.3.3.) was performed on schizonts of each isolate, using murine MAbs against MSP1 and MSP2 (Table 1, section 2.3.2.). The epitope locations of MAbs against MSP1 are shown in Figures 15 and 16 (Chapter 6).

Many isolates contained more than one P.falciparum clone (Chapter 4), so in the present analysis an isolate was considered positive for a given MSP1 or MSP2 epitope if the majority (> 50%) of schizonts in the isolate expressed the epitope. Differences in the proportions of patients of different blood groups, containing a majority of parasites expressing a given epitope, were tested by chi-square analyses on 2 x 2 contingency tables.
8.3. Results

Of 264 patients studied, 28.8% were blood group A, 28.0% group B, 2.3% group AB, and 40.9% group O. Of 258 patients, 14.0% were blood group M, 37.6% group N, and 48.4% group MN. Of 241 patients, 22.4% were positive and 77.6% were negative for blood group determinant S. These frequencies are similar to those reported among adults from The Gambia and Senegal (Mourant et al., 1976), as expected from previous studies showing no associations between these blood groups and malaria per se (Martin et al., 1979; Molineaux & Gramiccia, 1980).

Table 14 shows the proportion of patients of different ABO blood groups, containing parasites of which the majority (> 50%) of schizonts had a given epitope. The proportions were compared between each blood group by chi-square tests, with the exception of blood group AB for which the numbers were too small for statistical comparison. Values for MSP1 epitopes 6.1 and 9.2 could not be analysed by chi-square tests, as too few isolates were positive and negative for these epitopes, respectively.

The MSP2 epitope 8F6/49 was found more frequently on parasites from patients of blood group O than groups A (p<0.001) or B (p<0.005). All other polymorphic MSP2 and
Table 14. Proportions of isolates containing parasites of which > 50 % expressed particular MSP1 and MSP2 epitopes, from patients of different ABO blood groups

<table>
<thead>
<tr>
<th>MSP1</th>
<th>A</th>
<th>B</th>
<th>O</th>
<th>AB</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>4/76</td>
<td>0/74</td>
<td>3/108</td>
<td>0/6</td>
</tr>
<tr>
<td>9.2</td>
<td>72/76</td>
<td>74/74</td>
<td>105/108</td>
<td>6/6</td>
</tr>
<tr>
<td>13.2</td>
<td>38/76</td>
<td>38/74</td>
<td>55/108</td>
<td>4/6</td>
</tr>
<tr>
<td>9.5</td>
<td>25/76</td>
<td>27/74</td>
<td>45/108</td>
<td>1/6</td>
</tr>
<tr>
<td>12.1</td>
<td>23/76</td>
<td>32/74</td>
<td>40/108</td>
<td>4/6</td>
</tr>
<tr>
<td>10-2B</td>
<td>53/76</td>
<td>42/74</td>
<td>68/108</td>
<td>2/6</td>
</tr>
<tr>
<td>111.4</td>
<td>36/76</td>
<td>34/74</td>
<td>48/108</td>
<td>3/6</td>
</tr>
<tr>
<td>12.2</td>
<td>26/76</td>
<td>25/74</td>
<td>27/108</td>
<td>2/6</td>
</tr>
<tr>
<td>3D3</td>
<td>14/76</td>
<td>15/74</td>
<td>22/108</td>
<td>2/6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MSP2</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>12.3</td>
<td>54/75</td>
<td>53/74</td>
<td>60/104</td>
</tr>
<tr>
<td>8G10/48</td>
<td>21/75</td>
<td>21/74</td>
<td>42/100</td>
</tr>
<tr>
<td>8-5D</td>
<td>39/75</td>
<td>41/72</td>
<td>52/103</td>
</tr>
<tr>
<td>13.4</td>
<td>4/75</td>
<td>7/74</td>
<td>10/104</td>
</tr>
<tr>
<td>8F6/49</td>
<td>4/73</td>
<td>6/70</td>
<td>26/101</td>
</tr>
</tbody>
</table>

Values with the same superscript differed significantly by chi-square comparison: a, p<0.001; b, p<0.005.
MSP1 epitopes were present at similar frequencies within patients of different ABO blood groups. To test whether the observed strong association between ABO blood groups and MSP2 epitope 8F6/49 was consistent between years, the data from 1988 and 1989 were analysed separately (Table 15). In 1988, parasites with epitope 8F6/49 were more frequently isolated from patients of blood group O than blood group B (chi-square, p<0.05), but not than blood group A (chi-square, 0.1> p >0.05). In 1989, parasites with epitope 8F6/49 were more frequently isolated from patients of blood group O than blood group A (chi-square, p<0.01), but not than blood group B (chi-square, p>0.1). Thus there were differences between years in the observed statistical associations.

Table 16 shows several differences in the proportions of patients of different MN blood groups, in which the majority of parasites had a given epitope. MSP1 epitope 9.5 was more frequent on parasites from patients of blood group MN than M (p<0.025). MSP1 epitope 111.4 was more frequent on parasites from patients of blood group MN than N (p<0.05). MSP1 epitope 12.2 was more frequent on parasites from patients of blood group M than N (p<0.05), while epitope 3D3 was more frequent on parasites from patients of blood group M than MN (p<0.025). MSP2 epitope 8-5D was more frequent on parasites from patients of blood group MN than M (p<0.05) or N (p<0.05).
Table 15. Proportions of isolates containing parasites of which > 50 % expressed the MSP2 epitope 8F6/49, from patients of different ABO blood groups, analysed separately for 1988 and 1989.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>O</th>
<th>AB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1988</td>
<td>3/37</td>
<td>3/47</td>
<td>12/51</td>
<td>1/4</td>
</tr>
<tr>
<td>1989</td>
<td>1/36</td>
<td>3/23</td>
<td>14/50</td>
<td>1/2</td>
</tr>
</tbody>
</table>

Values with the same superscript differed significantly by chi-square comparison: a, p<0.05; b, p<0.01.
Table 16. Proportions of isolates containing parasites of which > 50% expressed particular MSP1 and MSP2 epitopes, in patients of different MN and S blood groups.

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>N</th>
<th>MN</th>
<th>S+</th>
<th>S-</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MSP1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.1</td>
<td>1/36</td>
<td>2/97</td>
<td>4/125</td>
<td>2/54</td>
<td>6/187</td>
</tr>
<tr>
<td>9.2</td>
<td>35/36</td>
<td>95/97</td>
<td>121/125</td>
<td>52/54</td>
<td>181/187</td>
</tr>
<tr>
<td>13.2</td>
<td>20/36</td>
<td>52/97</td>
<td>61/125</td>
<td>25/54</td>
<td>91/187</td>
</tr>
<tr>
<td>9.5</td>
<td>7/36</td>
<td>36/97</td>
<td>52/125</td>
<td>20/54</td>
<td>70/187</td>
</tr>
<tr>
<td>12.1</td>
<td>16/36</td>
<td>37/97</td>
<td>44/125</td>
<td>22/54</td>
<td>68/187</td>
</tr>
<tr>
<td>10-2B</td>
<td>20/36</td>
<td>60/97</td>
<td>81/125</td>
<td>32/54</td>
<td>119/187</td>
</tr>
<tr>
<td>111.4</td>
<td>16/36</td>
<td>38/97</td>
<td>66/125</td>
<td>24/54</td>
<td>85/187</td>
</tr>
<tr>
<td>12.2</td>
<td>15/36</td>
<td>23/97</td>
<td>41/125</td>
<td>17/54</td>
<td>54/187</td>
</tr>
<tr>
<td>3D3</td>
<td>11/36</td>
<td>23/97</td>
<td>17/125</td>
<td>16/54</td>
<td>33/187</td>
</tr>
<tr>
<td><strong>MSP2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.3</td>
<td>26/36</td>
<td>54/95</td>
<td>85/122</td>
<td>32/54</td>
<td>121/183</td>
</tr>
<tr>
<td>8G10/48</td>
<td>10/35</td>
<td>40/94</td>
<td>36/120</td>
<td>21/53</td>
<td>61/180</td>
</tr>
<tr>
<td>8-5D</td>
<td>15/36</td>
<td>44/94</td>
<td>73/120</td>
<td>23/54</td>
<td>102/180</td>
</tr>
<tr>
<td>13.4</td>
<td>3/36</td>
<td>8/95</td>
<td>11/122</td>
<td>3/54</td>
<td>17/183</td>
</tr>
<tr>
<td>8F6/49</td>
<td>4/34</td>
<td>17/91</td>
<td>16/119</td>
<td>8/51</td>
<td>27/177</td>
</tr>
</tbody>
</table>

Values with the same superscript differed significantly by chisquare comparison: a, p<0.025; b, p<0.05; c, p<0.05; d, p<0.025; e, p<0.05; f, p<0.05.
There were no differences in the frequency of any MSP1 or MSP2 epitopes between parasites from patients with or without blood group determinant S (Table 16).
8.4. Discussion

The number of significant associations between parasite epitopes and particular MN blood groups was greater than expected. If there were no underlying associations, one out of 20 chi-square tests would be expected to show a significance level of \( p<0.05 \). However, 6 out of 33 tests were significant at this level. This is difficult to interpret, as the direction of the association is different in different cases, and no single chi-square value is highly significant. Although there may be some real underlying associations, the data does not provide a strong case for any particular one being important. No previous studies have implicated the MN determinants, on the peptide chain of glycophorin A, as important sites of parasite recognition (Pasvol et al. 1982a).

In contrast, the only significant association between a parasite epitope and ABO blood groups had a high level of significance. Epitope 8F6/49 on MSP2 was more commonly detected on parasites in patients of blood group 0 than groups A (\( p<0.001 \)) or B (\( p<0.005 \)). There were differences in the strength of these associations in 1988 and 1989, however.

The epitope 8F6/49 has been identified as the linear sequence DTPTATE, present in some allelic types of MSP2.

The difference between the A and B polysaccharide determinants, and H (the determinant responsible for blood group 0) consists in the terminal sugar residues (Issitt, 1976). The A substance has N-acetyl-D-galactosamine linked (alpha 1,3 bond) to the penultimate D-galactose, the B substance has D-galactose linked (alpha 1,3) to D-galactose, whereas the H substance has no alpha 1,3 linked sugar on D-galactose. It is conceivable that the sequence DTPTATE may preferentially recognise the terminal of the H substance, in a receptor-ligand interaction involved in one of the alternative invasion pathways. There is no experimental evidence available to support this suggestion. Differences among parasite clones in ability to invade erythrocytes of different ABO phenotypes have not been tested for.

Studies on in vitro invasion of erythrocytes of different phenotype, by P.falciparum clones of different defined MSP1 and MSP2 types, would test more directly whether particular variants of these proteins preferentially recognise particular erythrocyte surface components. This approach could test the relevance of the highly statistically significant observed association between MSP2 epitope 8F6/49 and blood group 0.
The possibility that different polymorphic types of merozoite surface proteins recognise different erythrocyte surface ligands was investigated epidemiologically. Allelic polymorphisms of two proteins (MSP1 and MSP2) were determined serologically in isolates collected from 264 Gambian malaria patients, who were typed for blood group antigens ABO, MN, and S. The frequency of monoclonal antibody epitope 8F6/49 on MSP2 was much higher in patients of blood group O, than group A (p<0.001), or B (p<0.005). Several associations which were not highly statistically significant were noted between other parasite epitopes and MN blood groups.
Chapter 9: Conclusions and Implications of the Thesis

The genetic structure of *Plasmodium falciparum* in The Gambia is consistent with frequent recombination between unlinked loci (Chapter 3). This has implications for theoretical considerations on the emergence and spread of drug resistance in *P. falciparum*. For example, a model by Curtis & Otoo (1986) predicted the spread of resistance to two drugs (for which single resistance genes are on different loci), during selection on a proportion of a hypothetical parasite population by an administered mixture of the two compounds. Depending on whether recombination between loci was assumed, the predictions differed greatly, with a slower emergence of resistance when frequent recombination was assumed.

This outbreeding population structure is to be expected from the observation that most malaria patients contain more than one *P. falciparum* clone (Chapter 4), thereby allowing frequent cross-fertilisation of *P. falciparum* genotypes at the obligate sexual stage during infection in the mosquito (Walliker et al. 1987). The mean number of asexual blood stage clones detected per patient was 2.0, and does not correlate with the patient's age, parasitaemia, or the severity of clinical disease.

The possibility of acquiring more than one 'brood' of *P. falciparum* by superinfection has already been
incorporated into models of malaria epidemiology (Macdonald, 1950; Deitz, Molineaux & Thomas, 1974). Multiple-clone \textit{P.falciparum} infections in malaria patients can only partly be explained by superinfection, however. A statistical analysis reveals that it is also necessary to assume that a single mosquito may inoculate non-identical sibling recombinant clones of \textit{P.falciparum} into a human. Multiple-clone inoculation by single mosquitoes may require incorporation into future epidemiological models, if such models are to take account of parasite genetics.

A single mosquito may infect two individuals with \textit{P.falciparum} on the same occasion, due to interrupted blood feeding, or repeated skin probing prior to blood feeding. This had previously been established by entomological studies (Boyd, 1949b), and is here taken to explain the high frequency of genetic similarity between \textit{P.falciparum} clones isolated from children who sleep in the same house, presenting together with malaria (Chapter 5). Models of malaria transmission have assumed that an infectious mosquito can infect only one individual per blood feeding occasion, and this is probably one reason why such models tend to underestimate the predicted vectorial capacity of mosquito populations (Molineaux et al., 1988). It is practically impossible to accurately estimate vectorial capacity in a given epidemiological situation, due to factors such as this, which either have
not or cannot be measured. A more realistic, and relevant goal than an estimation of vectorial capacity, is to estimate the relative reduction in vectorial capacity expected under different mosquito control strategies (Dye, 1990).

Structurally and antigenically, allelic polymorphism in the precursor to the major merozoite surface proteins, MSP1, is very extensive. Thirty-nine different allelic serotypes of the protein are described in Chapter 6, and this number would increase if additional monoclonal antibodies were used in the serological characterisation. Much of this allelic polymorphism is due to intragenic recombination between domains of the gene at which polymorphism is more limited, and recombinant allelic proteins are distinguished serologically by different combinations of monoclonal antibody epitopes.

All the known sites of intragenic recombination in MSP1 are clustered within 20% of the gene sequence, towards the 5' end. Within three countries, The Gambia, Nigeria, and Brazil, there were strong non-random associations between epitopes whose determinants are on different domains of the sequence, separated by sites of intragenic recombination (Chapter 6). Such disequilibria can no longer be explained by postulating a 'clonal' parasite population structure (e.g. Weber, 1988). Neutral effects could give rise to disequilibria between polymorphic domains, if intragenic recombination events
were infrequent. However, most of the associations between polymorphic domains were similar in each of the countries. These observations suggest that there are constraints on certain recombinant alleles of MSP1, and that neutral effects may not be adequate to explain the observed patterns of structural and antigenic polymorphism of the protein.

Any mechanism of selection on MSP1 is probably related to its function, which remains unknown. Polymorphisms at certain domains, particularly the major sequence dimorphism throughout most of the gene, are probably not maintained by immune selection. In the Gambian study area over a period of seven years, rare polymorphic types remained rare while common alternative types remained common (Chapter 7). This does not support the hypothesis of frequency-dependent selection under the effects of acquired immunity, by which rare types should have a selective advantage.

There were no strong statistical correlations between ABO, MN, and S blood group antigens of patients, and the MSP1 types of parasites isolated from them (Chapter 8). It follows that, if different MSP1 types recognise different erythrocyte ligands prior to invasion, these ligands are not the ABO, MN, and S blood group determinants. The possible role of MSP1 in merozoite invasion of erythrocytes remains unclear.

Compared with MSP1, there is less knowledge of the
structural basis for serological polymorphism in the second merozoite surface protein, MSP2. In this thesis, serological polymorphisms of the protein have been studied, with little information on the molecular polymorphisms underlying these. However, more information on the allelic sequence polymorphisms is becoming available (Smythe et al. 1990, 1991; Thomas et al. 1990; Fenton et al. 1991).

There was a highly significant observed association between patients of blood group 0 and parasites with MSP2 epitope 8F6/49, determined by the linear sequence DTPTATE on MSP2 (Chapter 8). In the absence of other highly significant associations, this may merit further investigation, to determine whether the observed association reflects a specific receptor-ligand interaction.

It is unknown whether allelic polymorphism of the merozoite surface proteins MSP1 or MSP2 will be an obstacle to vaccine development. Polymorphism is a potential problem only if protective immune responses are directed against polymorphic domains of either protein, rather than the conserved domains. Even so, at certain domains of the MSP1 and MSP2 genes, polymorphism is limited to only two or three alternative sequences, so if any of these domains is a target of protective immunity it may be feasible to incorporate the different polymorphic types in a multivalent vaccine.
Investigations on human immune responses to defined portions of MSP1 and MSP2 are necessary, to determine firstly, whether immune responses against either protein can reliably protect against a challenge infection, and secondly, to which domains of the proteins such responses are targetted.
REFERENCES


Appendix

The 'VARPAS' computer program written in PASCAL by Mr. Adam Eyre-Walker, used in an analysis described in Chapter 3.
program resample_frequencies (input, output, data, results);

const max_no_freq = 2000;

var cum_prob : array [0..max_no_freq] of real;
filename : array [1..20] of char;
no_of_freq, no_of_individs, no_of_trials : integer;
trialno : integer;
data, results : text;

procedure g05ccc; extern;
{NAG routine to initialise random number generator}

function g05cac : real; extern;
{NAG routine to generate random numbers between 0 and 1 from a uniform distribution}

procedure read_data;

var i : integer;
{Reads in the expected frequencies from a user specified file}

begin
  cum_prob[0] := 0;
  writeln('Filename : ');
  readln(filename);
  reset(data, filename);
  i := 0;
  while not eof(data) do
    begin
      i := i + 1;
      read(data, cum_prob[i]);
      cum_prob[i] := cum_prob[i] + cum_prob[i-1];
    end;
  no_of_freq := i;
  for i := 1 to no_of_freq do
    chosen[i] := false;
  for i := 1 to no_of_individs do
    begin
      rnd := g05cac;
      j := 1;
      while rnd > cum_prob[j] do j := j + 1;
      if not chosen[j] then
        begin
          chosen[j] := true;
          no_of_genotypes := no_of_genotypes + 1;
        end;
    end;
end;
begin
g05ccc;
writeln('Filename for results : ');
readln(filename);
rewrite(results, filename);
writeln('No of trials : ');
readln(no_of_trials);
writeln('No of individuals in a sample : ');
readln(no_of_individs);
read_data;
for trialno:=1 to no_of_trials do
  writeln(results, resample);
end.
The epidemiology of multiple-clone Plasmodium falciparum infections in Gambian patients

D. J. Conway*, B. M. Greenwood and J. S. McBride

Institute of Cell, Animal and Population Biology, University of Edinburgh, Kings Buildings, West Mains Road, Edinburgh EH9 3JT, Scotland, U.K.

MRC Laboratories, Fajara, P.O. Box 273, Banjul, The Gambia, West Africa

(Received 27 September 1990; revised 14 February 1991; accepted 14 February 1991)

SUMMARY

The occurrence of multiple-clone Plasmodium falciparum haploid blood-stage infections is a pre-requisite for cross-fertilization and genetic exchange at the diploid stage in the mosquito. Using monoclonal antibodies against 3 polymorphic blood-stage antigens, a method of two-colour differential immunofluorescence allowed the resolution of between 1 and 4 clones/isolate. A mean of 2 P. falciparum clones was detected in the blood of malaria patients in The Gambia. The mean number of clones/patient showed no correlation with age, parasitaemia, or disease severity. There was a slight difference in mean number between sample periods, probably reflecting temporal differences in transmission intensity. A statistical analysis of 2-locus genetic diversity of clones within isolates concludes that not all multiple-clone infections result from superinfection, but that some are due to single multiple-clone inoculations.

Key words: Plasmodium, malaria, genetics, clone, epidemiology, immunofluorescence.

INTRODUCTION

The number of Plasmodium falciparum clones in infected human individuals may determine the genetic structure of the local P. falciparum population. The rate of genetic assortment and recombination in a population of plasmodia depends on the frequency at which vector mosquitoes acquire a bloodmeal containing gametocytes of more than one genotype. A successful cross-fertilization in the mosquito mid-gut is followed by a brief diploid stage and then by meiosis, producing recombinant haploid sporozoite progeny which can be inoculated at a later blood feed into the next human host (Sinden, Hartley & Winger, 1985; Walliker et al. 1987).

Plasmodium falciparum infections containing more than one genotype were first analysed in vitro by Rosario (1981), using limiting dilution cloning to demonstrate heterogeneity of parasites with respect to electrophoretic forms of glucose phosphate isomerase (GPI). Thaithong et al. (1984) used isoenzyme electrophoresis, monoclonal antibody immunofluorescence, 2-dimensional gel electrophoresis, and drug susceptibility assays to characterize clones derived from a Thai patient, and obtained 7 distinct types, each presumed to represent a different haploid clone present in the infection. Clonal heterogeneity of P. falciparum within in vitro-cultured isolates complicates the interpretation of parasite characterization. For example, different DNA sequences for the blood-stage antigen MSP1 (PMMMsa/gp195) were independently obtained from the widely cultured Palo Alto isolate (Chang et al. 1988; Scherf, Barbot & Langsley, 1989).

Although the existence of clonal heterogeneity within P. falciparum isolates in vitro is well established, the epidemiology of multiple-clone infections in patients within endemic areas has not been investigated systematically. Carter & McGregor (1973) studied isoenzyme polymorphism of P. falciparum in The Gambia, and calculated from the allelic frequencies of lactate dehydrogenase and glucose phosphate isomerase, and the proportion of isolates showing more than one allelic type of either enzyme, that wild isolates contained approximately 2 clones each.

The present study, also in The Gambia, involved the characterization of allelic serotypes of 3 polymorphic blood-stage antigens using monoclonal antibodies (MAbs). Different parasite clones were characterized by particular combinations of allelic serotypes for the 3 antigens. The study aimed to determine numbers of different clones within individual isolates, and to test for effects of patients' age, parasitaemia, stage of transmission season, and for correlations with disease severity. In addition, a statistical analysis was used to test whether multiple-clone infections can be adequately explained by superinfection events, or whether some multiple-clone infections are due to single multiple-clone inoculations.

MATERIALS AND METHODS

Study area, patients and P. falciparum isolates

The study was carried out in an urban/periurban region of The Gambia, West Africa, where the
incidence of clinical malaria is highly seasonal with the majority of cases occurring between August and November each year. During 1988 (July–December) and 1989 (October–December), 355 patients were enrolled at the outpatient departments of the Medical Research Council, Fajara, and the Royal Victoria Hospital, Banjul. They all presented with symptoms of malaria, and a *P. falciparum* density of at least 5 parasites per high-power microscopic field (approximately 0.1% parasitaemia). None of the patients contained parasites of other *Plasmodium* species. In all, 153 of these patients were included in a case-control study to determine risk factors for severe malaria, 35 of whom had severe disease according to the modified Wellcome criteria (Kwiatowski *et al.* 1990: here mostly cerebral malaria).

A vol. of 0.3 ml of heparinized blood was obtained from each patient, by fingerprick or as part of a larger volume obtained for other studies, under approval of the Scientific and Ethical Committees of MRC and the Gambian government. After removal of plasma, the blood cells were washed 3 times in 10 ml of sterile phosphate-buffered saline, pH 7.3 (PBS). Parasites were cultured (Trager & Jensen, 1976) for 24–48 h, until schizonts had matured (McBride, Welsby & Walliker, 1984). After washing 3 times and resuspending at approximately 3% haematocrit in PBS, multispot slides of schizonts were prepared with 20 μl of the cell suspension/spot. The slides were dried in a well air-conditioned room, and stored under desiccation at −20 °C.

**Monoclonal antibodies**

A panel of 27 murine monoclonal antibodies (MAbs), specific for allelic variants of 3 *P. falciparum* antigens coded by unlinked loci were used to type parasites in each isolate. The antigens included MSP1, the precursor to several merozoite surface antigens (also known as PMMSA, PSA, p190, gp195 or MSA-1; reviewed by Holder, 1988), another merozoite surface protein MSP2 (also called gp35-53, GP3 or MSA-2; reviewed by Kemp, Cowman & Walliker, 1990), and an exported protein Exp-1 (also called p23 or CRA; Simmons *et al.* 1987). Details on the specificity, isotype, working dilution and source of the MAbs are given in an accompanying paper (Conway & McBride, 1991).

**Indirect immunofluorescence assay (IFA)**

Each *P. falciparum* isolate was analysed in a series of IFA tests using the above MAbs (Voller & O'Neill, 1971; McBride, Walliker & Morgan, 1982). Working dilutions of antibodies were made in PBS containing 1% bovine serum albumin, and 0.01% sodium azide. A vol. 20 μl of each MAb was incubated on a well of an acetone-fixed multipot slide preparation of schizonts at room temperature for 30 min, carefully removed by Pasteur pipette, and slides were washed 3 times (1, 5, and 5 min) in PBS. After drying the slides at approximately 60 °C for 10 min, 20 μl of a 1:100 dilution of FITC (fluoresceine isothiocyanate)-conjugated polyvalent rabbit anti-mouse Ig antibody (ICN Immunobiologicals, Lisle, Israel), was added to each spot, and incubated for 30 min. After 2 washes (1 and 5 min) in PBS, parasite nuclei were stained with DAPI (4',6-diamino-2-phenylindole, Sigma Ltd: 1 x 10−6 (w/v) in PBS) for 1 min. The slides were washed twice (1 and 5 min) in PBS, and mounted in Citifluor (City University, London) under cover-slips. Parasites were visualized by FITC-fluorescence (MAb specific) and DAPI-fluorescence (DNA specific), with incident light of 450–490 nm and 390–440 nm respectively, at magnification ×360 or ×600.

Isolates within which only a proportion of schizonts expressed certain polymorphic epitopes, indicating the presence of more than one parasite clone, were analysed further using a double-labelled (two-colour) IFA.

**Double-labelled IFA**

Combinations of two MAbs with different isotypes, and different epitope specificities, were used to test the homogeneity of parasites within each isolate. Each pair of MAbs was incubated together on a well for 30 min. An RITC (rhodamine isothiocyanate)-conjugated and an FITC-conjugated antibody (Southern Biotechnology Associates Inc., Birmingham, Alabama; dilution 1:50), each specific for the different isotypes of the two MAbs, were then incubated together on the well for the second stage of 30 min, and carefully removed with a Pasteur pipette. All other steps were performed using exactly the same method as described above for the individual MAb IFA.

In addition to FITC-fluorescence (green) and DAPI fluorescence (blue), as described above, RITC-fluorescence (red) was visualized using incident light of 515–560 nm. The proportion of schizonts showing (i) green (and blue) fluorescence only, (ii) red (and blue) fluorescence only, (iii) red and green (and blue) fluorescence, and (iv) neither red nor green (only blue) fluorescence, was recorded for each pair of MAbs tested. Combined results, obtained with a series of different pairs of MAbs, resolved the number of distinct parasite clones within each isolate. The antigenic profile of the majority clone within each isolate was also determined (Conway & McBride, 1991).

**Limits on the resolution of the number of clones per isolate**

The number of schizonts (visualized by DAPI fluorescence) varied considerably among isolates,
from less than 1 to over 500/field. To avoid bias against detection of minority clones in isolates with low parasitaemia, an arbitrary limit of sensitivity was set at 1% of the total parasite number within any isolate. Therefore, no isolate was included in the analysis if parasite numbers were too low to score at least 200 schizonts/test, and detected minority clones representing less than 1% of parasites in certain isolates were also discarded from the analysis.

Despite the high resolution of double-labelled IFA analyses, it is not possible to reliably determine whether there are more than 4 clones within an isolate. Therefore, each isolate was scored as containing 1, 2, 3 or 4 clones.

Hypothesis testing by statistical methods

The hypothesis that each parasite clone within a human infection is acquired from a separate mosquito inoculation (i.e. superinfection) was tested by a statistical analysis of 2-locus genetic diversity of parasites within isolates.

The frequencies of allelic serotypes of MSP1 within the urban/periurban study area had been determined (Conway & McBride, 1991), and these frequencies were used to generate predictions based on the superinfection hypothesis. The hypothesis depends on the fact that the MSP1 and MSP2 loci are on different chromosomes (Kemp et al., 1987), and on the observation that alleles at the two loci are randomly assorted in the parasite population (Conway & McBride, 1991). Hence, if two clones within an isolate differ at one locus, MSP2, the probability that they are identical at the other unlinked locus, MSP1, equals the probability of randomly selecting two identical MSP1 alleles from the population. This probability equals the sum of the squares of individual MSP1 allele frequencies \(P = 0.063\) for MSP1 in the study area in 1988: data from Conway & McBride, 1991).

Among 105 isolates, each of which contains two clones distinguished by alleles of MSP2, the hypothesis predicts the number in which the two clones have an identical MSP1 type. This prediction is in the form of a probability distribution curve (binomial distribution \(N = 105, P = 0.063\), against which the observed number can be compared. If the observed number is higher than expected, it is concluded that a significant proportion of multiple-clone human infections result from a mechanism other than superinfection.

RESULTS

Clonal diversity within natural P. falciparum infections

Infections containing more than one clone of P. falciparum were frequently encountered in Gambian patients. Fig. 1 shows an example of in situ resolution of different clones within an isolate (GF88-96), using double-labelled IFA. The illustrated different combinations of anti-MSP1 MAbs showed that there were at least three different MSP1 serotypes within isolate GF88-96 (Fig. 1G-I). Additional tests with combinations of other MAbs against MSP1, as well as MSP2 and Exp-1, were also used to detect minority clones and resolve the antigenic phenotypes of the individual clones.

Mean number of clones per isolate

Among the 355 infections studied, the mean number (± s.e.) of detectable clones per isolate was 2.02 ± 0.06. Isolates from patients with severe and mild clinical malaria contained similar numbers of clones, with a mean of 2.11 ± 0.20 (\(N = 35\)) and 2.04 ± 0.10 (\(N = 118\)), respectively. Fig. 2 shows the frequencies of isolates containing 1, 2, 3, or 4 detectable clones. There were no significant differences in the mean number of parasite clones according to the age of patients (Table 1) or parasitaemia (Table 2).

The mean number varied slightly between seasons of P. falciparum transmission (Table 3). Isolates collected during October-November 1989 contained a significantly higher mean number of clones compared with those obtained during July-August 1988 (paired t-test, \(P < 0.002\)) and September-October 1988 (paired t-test, \(P < 0.05\)). A slight increase observed through the 1988 transmission season was not significant between periods.

Origin of multiple-clone infections

The hypothesis that each parasite clone within a human infection is acquired from a separate mosquito inoculation, and hence that all multiple-clone

Fig. 1. In situ resolution of different MSP1 phenotypes of Plasmodium falciparum in isolate GF88-96, using DNA staining and differential two-colour labelling of monoclonal antibodies. (A)-(C) show one microscopic field stained with: (A) DAPI (DNA-specific), (B) MAb 9.5 (isotype IgG2b) + FITC-conjugated anti-IgG2b, (C) MAb 10-2B (isotype IgG2a) + RITC-conjugated anti-IgG2a. In this field, each of the schizonts is positive with both MAbs. (D)-(F) show a second field, stained with: (D) DAPI, (E) MAb 12.1 (isotype IgG1) + FITC-conjugated anti-IgG1, (F) MAb 10-2B + RITC-conjugated anti-IgG2a. In this field, each schizont is positive with either one or the other MAb. (G)-(I) show a third field, stained with: (G) DAPI, (H) MAb 111.4 (isotype IgG1) + FITC-conjugated anti-IgG1, (I) MAb 10-2B + RITC-conjugated anti-IgG2a. In this field one of the schizonts is positive with both MAbs, and the others are positive with one or other of the MAbs (i.e. there are three distinct parasite phenotypes).
The percentage of Plasmodium falciparum isolates containing 1, 2, 3, or 4 clones among severe ( ), N = 35) and mild ( , N = 118) clinical cases, and among the total analysed in 1988–1989 ( , N = 355).

Table 1. Mean number of clones/isolate according to age of patients
(No significant differences between means; paired t-tests, P > 0.05 for each comparison.)

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Mean ± s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 2</td>
<td>2.05 ± 0.13 (n = 77)</td>
</tr>
<tr>
<td>2–4</td>
<td>2.02 ± 0.10 (n = 120)</td>
</tr>
<tr>
<td>5–9</td>
<td>1.90 ± 0.10 (n = 111)</td>
</tr>
<tr>
<td>&gt; 10</td>
<td>2.30 ± 0.20 (n = 40)</td>
</tr>
</tbody>
</table>

Table 2. Mean number of clones/isolate according to parasite density
(No significant differences between means; paired t-tests, P > 0.1 for each comparison.)

<table>
<thead>
<tr>
<th>Parasite density (number/high-power field)</th>
<th>Mean ± s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 50</td>
<td>2.11 ± 0.14 (n = 64)</td>
</tr>
<tr>
<td>50–199</td>
<td>2.08 ± 0.10 (n = 131)</td>
</tr>
<tr>
<td>&gt; 200</td>
<td>1.87 ± 0.09 (n = 128)</td>
</tr>
</tbody>
</table>

Table 3. Mean number of clones/isolate during different periods
(Mean for Oct–Nov 1989 is significantly higher than that for Jul–Aug 1988 and Sept–Oct 1988, P < 0.002 and P < 0.05 for respective paired t-tests. Other comparisons showed no significant differences.)

<table>
<thead>
<tr>
<th>Sampling period</th>
<th>Mean ± s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>July–August 1988</td>
<td>1.67 ± 0.14 (n = 45)</td>
</tr>
<tr>
<td>September–October 1988</td>
<td>1.94 ± 0.09 (n = 123)</td>
</tr>
<tr>
<td>November–December 1988</td>
<td>2.05 ± 0.18 (n = 39)</td>
</tr>
<tr>
<td>October–November 1989</td>
<td>2.25 ± 0.12 (n = 112)</td>
</tr>
</tbody>
</table>

Fig. 3. The number of isolates containing two clones with an identical MSPI type, out of 105 isolates in which the two clones were differentiated with respect to MSP2 types. The arrow shows the observed number. The probability distribution curve shows the expectations of the hypothesis that all multiple-clone infections are due to random superinfections (binomial distribution N = 105, P = 0.063). For details of the hypothesis see the Materials and Methods section.

The double-labelled IFA analyses described here determined the number of P. falciparum clones per patient to a high degree of resolution. As the level of detectable MSPI, MSP2, and Exp-1 polymorphism within the parasite population is very high, there is a low probability of randomly encountering 2 clones which have an identical genotype with respect to the three antigen loci (Conway & McBride, 1991). However, a limitation of the technique is that, in low level infections, clones accounting for less than 1% of the total parasite number cannot reliably be detected. Such minor sub-populations were excluded from our analysis in order to standardize the sensitivity of the method between isolates.
The visualization of individual parasites by double-labelled IFA allows determination of the number and relative proportions of different clones within an uncloned isolate. This offers certain advantages over methods requiring cellular extracts (e.g. isoenzyme electrophoresis, 2-dimensional gel electrophoresis, restriction fragment length polymorphism analysis, or DNA amplification and allelespecific oligonucleotide probe analysis). Using extracted protein or DNA material, it may be impossible to resolve multi-locus genotype combinations present within a mixed isolate, without prior in vitro cloning of the parasites. This is important, as long-term in vitro culture may result in clonal selection due to different in vitro growth rates (Oduola et al. 1988), and the procedure carries a risk of isolate cross-contamination (Fenton, 1988). An additional advantage of the IFA technique is that it requires only a small volume of parasitized blood, which can be obtained from a fingerprick sample.

The mean value of 2-0 P. falciparum clones/patient is similar to the value calculated by Carter & McGregor (1973) from isoenzyme frequencies in The Gambia. It is also comparable to the finding of 1-4 MSP1 alleles in clinical isolates from Senegal (Scherf, Mattei & Sarthou, 1991). The isolation of 7 distinct clones from a single Thai isolate, T9 (Thaithong et al. 1984), and 6 distinct clones from another Thai patient before and after mefloquine treatment (Pinswasdi et al. 1987) lies well outside the range of 1 to 4 clones detectable in the present study. The technique of in vitro cloning followed by long-term culture, as used on the Thai isolates, may be necessary to select clones originally present at undetectable levels. The number of clones/patient may differ between areas with different levels of endemicity (Creasey et al. 1990), although this has yet to be investigated in a systematic way. It is also possible that the two Thai isolates described are atypical examples of mixed infection.

The number of clones/infection appears to have no relevance to the development of severe clinical disease. The number was neither related to the age nor the parasitaemia of the patient. The slightly higher mean number in 1989 probably reflects the fact that malaria transmission was particularly high that year, but the comparison between different sample periods in 1988 did not reveal significant seasonal variation. Collectively, these results indicate that the mean number of P. falciparum clones/infection is relatively stable within the Gambian study area. The mixed infections clearly occur at a frequency which would allow frequent cross-fertilization and genetic recombination within vector mosquitoes, consistent with the observed local P. falciparum population structure (Conway & McBride, 1991).

Mathematical models of malaria epidemiology have allowed for superinfection of humans by parasites from different mosquitoes (Macdonald, 1950; Deitz, Molineaux & Thomas, 1974; Deitz, 1988). We have attempted to address the question of the origin of multiple-clone infections using a 2-locus statistical analysis. Multiple-clone infections cannot be explained as a result of superinfection alone, since the 2-locus analysis showed that the observed identity at the second locus (MSP1) was too high to be explained by a random acquisition of genotypes from the local parasite population. Spatial clustering of parasite genotypes would confound the analysis, but there was no evidence of such clustering within the urban/perurban study area (Conway & McBride, 1991).

Therefore, to explain the higher than expected identity at the second locus within individual infections, it is proposed that at least some multiple-clone infections are the result of multiple-clone inoculations from single mosquitoes carrying sibling recombinant parasites. Within a hyperendemic area, where recombination between P. falciparum genotypes occurs frequently (Conway & McBride, 1991), a vector mosquito may frequently contain sporozoites of more than one P. falciparum genotype. The frequency and genetic similarity of multiple-clone P. falciparum infections in single mosquitoes could be investigated directly by analysis of DNA polymorphism in oocysts and sporozoites isolated from individual wild-caught mosquitoes.

We thank all the patients for willing participation. The study depended on logistic support from Drs Dominic Kwiatowski, Adrian Hill, and the staff at the Medical Research Council Laboratories in The Gambia. We are grateful to Drs Robert Reese, Jeffrey Lyon, Anthony Holder, Alan Saul and Richard Pink for provision of their monoclonal antibodies, and to Richard Carter for comments on the manuscript. The work was supported by grants from the Medical Research Council, the Wellcome Trust, and the UNDP/World Bank/WHO Special Programme for Training and Research in Tropical Diseases.

REFERENCES


D. J. Conway, B. M. Greenwood and J. S. McBride


Population genetics of *Plasmodium falciparum* within a malaria hyperendemic area

D. J. CONWAY* and J. S. McBRIDE†

_Institute of Cell, Animal and Population Biology, University of Edinburgh, Kings Buildings, West Mains Road, Edinburgh EH9 3JT, Scotland, U.K._

(Received 27 September 1990; revised 14 February 1991; accepted 14 February 1991)

**SUMMARY**

Serotyping with monoclonal antibodies was used to estimate the number and frequencies of allelic variants of two merozoite surface proteins, MSPI and MSP2, and an exported protein Exp-1, in a sample of 344 clinical isolates of *Plasmodium falciparum* from an urban region of The Gambia. Represented among the isolates were 36, 8 and 2 alleles of the MSPI, MSP2 and Exp-1 loci respectively. Relative frequencies of these alleles remained stable in the parasite population over the 2 years of the study. A computer program was used to calculate from the frequencies of individual alleles at the three loci, the probable number of different genotypes in samples from the population, assuming random assortment among the loci. No significant difference was found between the expected and the observed genotype diversity. It is concluded that recombination among unlinked loci is a common consequence of sexual reproduction of *P. falciparum* in The Gambia. Slightly lower genotype diversity was observed in each of two villages, which may be a consequence of smaller population size compared with the urban region.

**Key words:** *Plasmodium*, genetics, recombination, epidemiology, antigens, immunofluorescence.

**INTRODUCTION**

Malaria parasites reproduce asexually and are haploid within vertebrate hosts, but produce gametes and undergo fertilization, zygote formation and sexual recombination at meiosis within mosquitoes (Sinden, Hartley & Winger, 1985; Walliker _et al._ 1987). The sexual reproduction in the mosquito vector is an obligatory event in the parasite life-cycle and, without it, natural transmission of malaria cannot proceed. In natural populations sexuality thus provides opportunities for generation of recombinant parasite genotypes by chromosomal re-assortment.

In *Plasmodium falciparum* allelic polymorphisms exist at many loci including antigens, enzymes and other proteins (reviewed by Kemp, Cowman & Walliker, 1990). Many combinations of alleles at different loci have been observed in culture-adapted parasites (Fenton, Walker & Walliker, 1985; Creasey _et al._ 1990), consistent with the hypothesis that genetic recombination and segregation of alleles occurs in natural populations (reviewed by Walliker, 1985, 1989). An experimental genetic cross between two defined clones of *P. falciparum* directly demonstrated that, during transmission of a mixed infection, recombinant genotypes were generated (Walliker _et al._ 1987).

However, in wild populations the frequency of recombination events remains unknown, and a variety of recombinant genotypes could persist even if heterozygous recombination is infrequent. Evidence that heterozygous recombination may be rare or lacking in a given population, would include the over- or under-representation of particular multilocus genotypes, or a lower diversity of genotypes than expected on the assumption of random assortment of alleles. Tibayrenc, Kjellberg & Ayala (1990) reviewed data from several species of parasitic protozoa, including *P. falciparum*, and proposed that a predominantly clonal population structure may be characteristic of parasitic protozoa generally.

The genetic diversity of *P. falciparum* may be limited by two variables in a natural population: (1) the number and frequencies of different alleles present in the population, and (2) the frequency of genetically mixed infections since only these offer the possibility of cross-fertilization in the mosquito vector. The two variables are examined respectively here and in an accompanying study (Conway, Greenwood & McBride, 1991), for a parasite population in The Gambia.

To determine by statistical methods, whether alleles from unlinked polymorphic loci are assorted randomly in a population, a technique is required to identify the allelic products in the same parasite, and to characterize a large sample of the population. In this study, indirect immunofluorescence analyses (IFA) with monoclonal antibodies (MAbs) were used to identify _in situ_ allelic variants of three polymorphic bloodstage antigens whose diversity is well understood at the DNA level. The antigens, MSP1 (also known as PMMSA, PSA, p190, gp195
or MSA-1), MSP2 (also called MSA-2, gp35-56, or GP3), and Exp-1 (CRA or p23), are each encoded by a single locus within the haploid genome, and the loci have been mapped to different chromosomes (Kemp et al. 1987, 1990). Each antigen exhibits allelic polymorphism determined by DNA sequence and detectable by MAbs as epitope differences among parasite clones (MSP1: McBride, Newbold & Anand, 1985; Lyon et al. 1987; Tanabe et al. 1987; Gentz et al. 1988. MSP2: Clark et al. 1989; Fenton et al. 1989, 1991; Smythe et al. 1990; Exp-1: Simmons et al. 1987).

The present study involved the determination of allelic frequencies at each antigen locus, and statistical analyses of 3-locus genotype diversity among *P. falciparum* isolates from malaria patients within an urban/periurban region, and two villages, in The Gambia. The results show that the observed diversity of genotypes is compatible with random assortment of alleles and outbreeding in this population.

**MATERIALS AND METHODS**

**Study region and *P. falciparum* isolates**

The Gambia, within the Northern Guinea Savannah climatic zone of West Africa, experiences highly seasonal rainfall (500–1000 mm annually, between May and October). This results in hyperendemic *P. falciparum* transmission with a seasonal prevalence which may exceed 50%, and a peak of clinical cases occurring between August and November (Marsh et al. 1989).

A *P. falciparum* isolate is defined as a sample of parasites taken from a patient on a single occasion. Isolates were collected from patients in three areas (Fig. 1). (a) An urban/periurban region within a 7 km radius of Serekunda, from Bakau in the north to the Yundum International Airport in the south. A total of 344 isolates were obtained from patients presenting to the Outpatients Departments of the Medical Research Council, Fajara, and the Royal Victoria Hospital, Banjul, during July–December 1988 (N = 225), and October–December 1989 (N = 119). None of the patients had travelled outside the region during 3 weeks prior to clinical presentation. (b) Nine isolates were obtained in Sarakunda village, near Farafenni on the north bank of the Gambia river, on 5 and 6 October 1988. (c) Seventeen isolates were obtained in Brefet village, on the south bank in the Lower River Division, during October–December 1989.

Samples of 0.3 ml of heparinized blood were obtained from each patient after consent, under the approval of the Ethical Committee of the Medical Research Council and the Gambian Government. The blood was washed 3 times in sterile phosphate-buffered saline, and the parasites were cultured for 24–48 h to obtain mature schizonts (McBride, Welsby & Walliker, 1984). Schizont-infected cells were air dried on multipot slides, and stored desiccated at —20°C until use in immunofluorescence typing (Conway et al. 1991).

**Antigens and monoclonal antibodies (MAbs)**

Characteristics of murine MAbs specific for epitopes on polymorphic domains of *P. falciparum* antigens MSP1, MSP2, and Exp-1 are described in Table 1. Correlations between deduced amino-acid sequences and epitope polymorphism of the antigens are detailed elsewhere, for MSP1 (Tanabe et al. 1987; Gentz et al. 1988; Conway & McBride, unpublished results), MSP2 (Saul et al. 1989; Smythe et al. 1990; Fenton et al. 1991), and Exp-1 (Simmons et al. 1987).

**Individual (1-colour) and double-labelled (2-colour) MAb immunofluorescence**

All MAbs were first used individually in indirect immunofluorescence assays (IFA) to identify MSP1, MSP2 and Exp-1 variants expressed by schizonts of each isolate. Since many isolates contained more than one *P. falciparum* clone (Conway et al. 1991), the antigen phenotype was determined for the majority clone within each isolate. This was achieved by using pairs of MAbs with different isotypes and epitope specificities in a double-labelled (two-colour) IFA. Combined results from different pairs of MAbs allowed the phenotype of the majority clone within each isolate to be accurately resolved (Conway et al. 1991).

**Allelic serotypes and 3-locus genotype classification**

Allelic serotypes of each of the antigens were distinguished by their different profiles of reactivity with the typing MAbs. Each allelic serotype was assigned a number, according to a classification scheme based on combinations of individual MAb specificities.

The 3-locus genotypes were determined as different combinations of allelic serotypes of the 3 antigens. Only the genotype of the majority parasite clone within each isolate was used for analysis, as this was accurately determined for each isolate by the double-labelled IFA method.

**Statistical analyses**

**Genotype diversity.** Assuming random assortment of alleles at the unlinked loci, expected genotype frequencies were first calculated from the known allelic frequencies of each of the three antigens, and stored in a data file. Secondly, a PASCAL program was used to calculate probability distribution curves of genotype diversity (i.e. the number of different 3-locus combinations) within samples of given size. The program performed the following functions. (i)
Converted the expected 3-locus genotype frequencies into cumulative values, between 0 and 1. (ii) Randomly picked \( N \) values between 0 and 1 in order to obtain \( N(d) \) different cumulative values from (i). This corresponds to \( N(d) \) different genotypes from a random sample of size \( N \). (iii) Repeated step (ii) 1000 times to obtain a distribution of values of \( N(d) \), which was plotted as a probability distribution curve, corresponding to the probability of obtaining a given number of different genotypes within a particular sample size, \( N \).

Thirdly, using the generated probability distributions, the following tests were performed. (a) Within the urban/periurban region, observed numbers of different 3-locus genotypes were compared to expected numbers, \( N(d) \), to test whether the observed genotype diversity was lower than expected on the hypothesis of random assortment of alleles at the three different loci. The test was applied to 1988 and 1989 data separately. (b) Within Sarakunda and Brefet villages, and also within small areas of the urban/periurban region, observed numbers of different genotypes were compared with numbers expected for identical sample sizes taken randomly from the whole urban/periurban region.

**Pairs of identical genotypes.** The expected number of pairs of identical genotypes, \( N(i) \), within a sample of given size, \( N \), was calculated on the basis of expected 3-locus genotype frequencies within the urban/

### Table 1. Specificities of typing monoclonal antibodies

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Antibody†</th>
<th>Reciprocal dilution for IFA</th>
<th>Isotype</th>
<th>Reference‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSP1</td>
<td>*9. 8-4-4-1</td>
<td>1000</td>
<td>IgG1</td>
<td>2, 3</td>
</tr>
<tr>
<td></td>
<td>*12. 4-3-4</td>
<td>500</td>
<td>IgG1</td>
<td>2, 3</td>
</tr>
<tr>
<td></td>
<td>*12. 8-2</td>
<td>1000</td>
<td>IgG2b</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>*9. 2-6-2</td>
<td>2000</td>
<td>IgG1</td>
<td>1-3</td>
</tr>
<tr>
<td></td>
<td>*9. 7-1</td>
<td>500</td>
<td>IgG1</td>
<td>1-3</td>
</tr>
<tr>
<td></td>
<td>*10. 3-2</td>
<td>500</td>
<td>IgG1</td>
<td>1-3</td>
</tr>
<tr>
<td></td>
<td>*1-1C</td>
<td>500</td>
<td>IgG1</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>*7. 3-7</td>
<td>1000</td>
<td>IgG2a</td>
<td>1-3</td>
</tr>
<tr>
<td></td>
<td>*34. 5</td>
<td>Undiluted</td>
<td>IgG1</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>*6. 1-1-3</td>
<td>500</td>
<td>IgG1</td>
<td>1-3</td>
</tr>
<tr>
<td></td>
<td>*13. 1-2</td>
<td>2000</td>
<td>IgG1</td>
<td>1-3</td>
</tr>
<tr>
<td></td>
<td>*17. 1-3</td>
<td>2000</td>
<td>IgG1</td>
<td>1-3</td>
</tr>
<tr>
<td></td>
<td>12. 2-1-1</td>
<td>2000</td>
<td>IgG1</td>
<td>1-3</td>
</tr>
<tr>
<td></td>
<td>3D3. 10</td>
<td>1000</td>
<td>IgG2b</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>9. 5-1-5-1</td>
<td>500</td>
<td>IgG2b</td>
<td>1-3</td>
</tr>
<tr>
<td></td>
<td>12. 1-5-4</td>
<td>2000</td>
<td>IgG1</td>
<td>1-3</td>
</tr>
<tr>
<td></td>
<td>13. 2-3</td>
<td>2000</td>
<td>IgG1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>10-2B</td>
<td>2000</td>
<td>IgG2a</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>111. 4</td>
<td>1000</td>
<td>IgG1</td>
<td>7</td>
</tr>
<tr>
<td>MSP2</td>
<td>*12. 3-2-2</td>
<td>500</td>
<td>IgG1</td>
<td>1, 8-10</td>
</tr>
<tr>
<td></td>
<td>*12. 5-1-2</td>
<td>500</td>
<td>IgG1</td>
<td>1, 10</td>
</tr>
<tr>
<td></td>
<td>13. 4-2-1</td>
<td>500</td>
<td>IgG1</td>
<td>8-10</td>
</tr>
<tr>
<td></td>
<td>*8-5D</td>
<td>200</td>
<td>IgM</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>*4-4F</td>
<td>200</td>
<td>IgM</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>8G10/48</td>
<td>300</td>
<td>IgG2b</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>8F6/49</td>
<td>40</td>
<td>IgG3</td>
<td>12</td>
</tr>
<tr>
<td>Exp-1</td>
<td>5. 1-4</td>
<td>500</td>
<td>IgG1</td>
<td>13</td>
</tr>
</tbody>
</table>

† MAb epitopes marked * are conserved among all isolates tested. MAb epitopes marked by identical symbols (°, ‡, §) exhibit identical allelic distributions.
perurban region. The probability, \( P(i) \), of a random pair of genotypes being identical is equal to the sum of the squares of all individual genotype frequencies. The number of all possible pairwise comparisons between genotypes within a sample of size \( N \), is equal to \( N(N-1)/2 \). The probability distribution of the expected number of pairs of identical genotypes, \( N(i) \), assuming random genotype distribution, is given by the binomial distribution of \( P(i) \) for \( N(N-1)/2 \).

Numbers of pairs of identical genotypes observed in two villages and within small areas of the urban/periurban region, were tested against the binomial distribution of \( P(i) \), \( N(N-1)/2 \) for equivalent sample sizes from the urban/periurban region as a whole.

RESULTS

The extent of polymorphism and allelic serotype frequencies at the MSP1, MSP2 and Exp-1 antigen loci

Serotypes at the three loci were determined by IFA tests with a panel of MAbs, on parasites present within 344 isolates collected in the urban/periurban region during 1988 (July–December) and 1989 (October–December). Less than half of the isolates contained only a single parasite clone, while two or more clones could be clearly differentiated in other isolates (Conway et al. 1991). In the present analysis, each of such mixed isolates is represented once, by its most predominant clone whose serotype was determined accurately by two-colour IFA.

Diversity of MSP1 serotypes and their frequencies are shown in Fig. 2. MSP1 exhibited a very high degree of allelic polymorphism, with 35 different serotypes detected among 225 isolates in 1988. In 1989, 24 serotypes were represented among 119 isolates at frequencies similar to the previous year.

MSP2 serotypes and their frequencies are illustrated in Fig. 3. Allelic polymorphism of MSP2 was also high, with 8 different serotypes detected in 1988 (\( N = 221 \)). The frequencies of 7 serotypes also represented in 1989 were similar to those found in the previous year (\( N = 117 \)).

Polymorphism in Exp-1 is limited to a single epitope change recognized by MAb 5.1. In 1988, 30-4% of the majority clones within isolates typed positive for this epitope (\( N = 224 \)). In 1989 30-3% were positive (\( N = 119 \)).
Fig. 3. MSP2 allelic serotype frequencies in the urban/periurban region. Each numbered serotype is defined by a unique combination of epitopes recognized by the monoclonal antibodies, shown at the bottom of the figure. Types 3 and 9 have never been observed. (■) 1988 (N = 221); (▲) 1989 (N = 117).

3-Locus genotype diversity within the urban/periurban area

Each 3-locus genotype was defined as a different combination of MSP1, MSP2, and Exp-1 allelic serotypes. For example, in 1988, among 220 parasite isolates, 126 different genotypes were observed (Fig. 4). The observed value was then compared to a probability distribution curve of genotype diversity generated by a statistical method described in the Materials and Methods section.

For the 1988 data (sample size N = 220), the probability distribution predicted a median value of 126 and a lower 95% confidence limit of 117 different genotypes. The observed value of 126 was thus exactly as expected (Fig. 5A). For the 1989 sample (N = 117), the probability distribution generated an expected median value of 80, and a lower 95% confidence value of 73. The observed value was 75 (Fig. 5B).

Therefore, for both 1988 and 1989 transmission seasons, the observed 3-locus genotype diversity values were within the range expected from the hypothesis of random assortment of alleles.

3-Locus genotype diversity in small areas within the urban/periurban region

To investigate the possibility of spatial clustering of parasite genotypes within the urban/periurban region, the observed 3-locus genotype diversities within small areas were compared to those expected for samples of the same sizes taken from the region as a whole, using data collected in 1988. According to where the patient lived, each isolate was assigned to 1 of 8 areas (each of radius approximately 1.5 km). In each area, the observed number of different genotypes was within expected statistical limits (Table 2).
Fig. 5. Probability distribution curves of the expected number of different 3-locus genotypes within samples of (A) 220 isolates in 1988, and (B) 117 isolates in 1989, on the assumption of random assortment of alleles at different loci. The area outside the lower 95% confidence limit under each curve is shaded. The observed number of different 3-locus genotypes within each sample is shown by an arrow. (The generation of the probability distributions is described in the Materials and Methods section.)

The number of identical genotype pairs, \( N(i) \), also fitted randomly expected values (Table 3), as determined by the binomial distribution described in the Materials and Methods section.

3-Locus genotype diversity in villages

To investigate whether genotype diversity might be lower in small village populations than in the larger urban region, samples were analysed from two villages distant from the urban region (Fig. 1). The samples from both villages were small, so 3-locus genotype diversities observed in the villages were compared to expected values generated for identical sample sizes taken from the urban region in the appropriate year.

In a sample of 9 isolates from Sarakunda village, 7 different genotypes were observed (3 isolates contained an identical genotype). The median expected value was 9 and the lower 95% confidence limit was 8. The observed value was therefore lower than expected. Out of a total of \( N(N-1)/2 = 36 \) pairwise comparisons between isolates, there were 3 pairs of identical genotypes. This is higher than expected \((P < 0.01)\), according to the binomial distribution for \( P(i) = 0.00802 \) (the sum of the squares of expected genotype frequencies in the urban/periurban region in 1988), which gave a median expected value of 0. The analysis is explained in the Materials and Methods section.

In a sample of 17 isolates from Brefet village 13 different genotypes were observed (in 1989). The median expected value was 16 and the lower 95% confidence limit was 14. The observed value was therefore lower than expected. Out of a total of \( N(N-1)/2 = 136 \) pairwise comparisons between isolates, there were 5 pairs of identical genotypes. This is a higher number than expected \((P < 0.01)\) according to the binomial distribution for \( P(i) = 0.00898 \) (the sum of the squares of expected genotype frequencies in the urban/periurban region in 1989), which gave a median expected value of 1.

In summary, the observed \( P. falciparum \) genotype diversity was lower in the village samples than in the urban/periurban region.

Discussion

The PASCAL program employed a simple randomization procedure, based on expected genotype frequencies calculated from known allelic frequencies. In order to obtain accurate allelic frequencies initially, large sample sizes are required. It is considered that this sort of randomization method is the most appropriate means of analysing genotype diversity when allelic polymorphism is high at more than one locus, and the individual genotype frequencies are therefore expected to be very low.

Within the Gambian urban/periurban population, the observed extensive diversity of \( P. falciparum \) 3-locus genotypes was in accordance with expectations assuming random mating and assortment of alleles. The results do not support a hypothesis of Tibayrenc et al. (1990) who proposed that clonal reproduction involving self-fertilization is more predominant than cross-fertilizing reproduction, within natural populations of \( P. falciparum \). Although the clonal hypothesis may apply to other genera of parasitic protozoa, there are no data to support the hypothesis for any plasmodia. Neither of the sources cited by Tibayrenc et al. (1990) contains data suitable for analysis of population genetic structure. The unusual observation of identical multi-locus genotypes in two culture-established isolates presumed to originate from different parts of Thailand (Fenton et al. 1985), has already been explained by Fenton (1988) as probably due to a cross-contamination in the laboratory. The cited isoenzyme data on 15 geographically disparate cultured isolates (Sanderson, Walliker & Molez, 1981), include two identical clones from the same patient. More importantly, in that study the local allelic frequencies were not known. In contrast,
Table 2. Comparison of expected and observed numbers of different genotypes, $N(d)$, in small areas within the urban/periurban region in 1988

<table>
<thead>
<tr>
<th>Area</th>
<th>Address</th>
<th>No. of isolates</th>
<th>Expected $N(d)$</th>
<th>Lower 95% CL*</th>
<th>Observed $N(d)$†</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bakau</td>
<td>26</td>
<td>24</td>
<td>21</td>
<td>25</td>
<td>&gt; 0.5</td>
</tr>
<tr>
<td>2</td>
<td>Dippakunda</td>
<td>46</td>
<td>39</td>
<td>35</td>
<td>40</td>
<td>&gt; 0.5</td>
</tr>
<tr>
<td></td>
<td>Latrikunda</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Jeshwang</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Bundung</td>
<td>31</td>
<td>28</td>
<td>25</td>
<td>26</td>
<td>&gt; 0.1</td>
</tr>
<tr>
<td></td>
<td>Ebotown</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>London Corner</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Serekunda</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Talinding</td>
<td>29</td>
<td>26</td>
<td>23</td>
<td>27</td>
<td>&gt; 0.5</td>
</tr>
<tr>
<td></td>
<td>Fajikunda</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wellingara</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nemakunku</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Lamin</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>9</td>
<td>&gt; 0.5</td>
</tr>
<tr>
<td></td>
<td>Banjulinding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Sukuta</td>
<td>36</td>
<td>32</td>
<td>28</td>
<td>30</td>
<td>&gt; 0.1</td>
</tr>
<tr>
<td></td>
<td>Bakoteh</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Brufut</td>
<td>12</td>
<td>12</td>
<td>10</td>
<td>12</td>
<td>&gt; 0.5</td>
</tr>
<tr>
<td></td>
<td>Kotu</td>
<td>14</td>
<td>13</td>
<td>12</td>
<td>14</td>
<td>&gt; 0.5</td>
</tr>
<tr>
<td></td>
<td>Manjaikunda</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* CL, confidence limit.
† The observed numbers did not differ from those expected assuming random distribution of genotypes in the urban/periurban region.

Table 3. Comparison of expected and observed numbers of identical genotype pairs, $N(i)$, in small areas within the urban/periurban region in 1988

<table>
<thead>
<tr>
<th>Area*</th>
<th>No. of isolates</th>
<th>No. of pairwise comparisons $N(N-1)/2$</th>
<th>Expected (median) $N(i)$</th>
<th>Observed $N(i)$</th>
<th>P value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>26</td>
<td>325</td>
<td>2</td>
<td>1</td>
<td>&gt; 0.5</td>
</tr>
<tr>
<td>2</td>
<td>46</td>
<td>1035</td>
<td>8</td>
<td>7</td>
<td>&gt; 0.5</td>
</tr>
<tr>
<td>3</td>
<td>31</td>
<td>465</td>
<td>4</td>
<td>6</td>
<td>&gt; 0.1</td>
</tr>
<tr>
<td>4</td>
<td>29</td>
<td>406</td>
<td>3</td>
<td>2</td>
<td>&gt; 0.5</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>36</td>
<td>0</td>
<td>0</td>
<td>&gt; 0.1</td>
</tr>
<tr>
<td>6</td>
<td>36</td>
<td>630</td>
<td>5</td>
<td>7</td>
<td>&gt; 0.1</td>
</tr>
<tr>
<td>7</td>
<td>12</td>
<td>66</td>
<td>0</td>
<td>0</td>
<td>&gt; 0.5</td>
</tr>
<tr>
<td>8</td>
<td>14</td>
<td>91</td>
<td>1</td>
<td>0</td>
<td>&gt; 0.5</td>
</tr>
</tbody>
</table>

* Areas were as described in Table 2.
† The P values were obtained by comparing the observed values of $N(i)$ with the expected values calculated by the binomial distribution of $P(i) = 0.00802$, $N(N-1)/2$ (described in the Materials and Methods section). There were no significant differences between observed and expected values.

2-locus isoenzyme analyses of clinical samples from The Gambia and Tanzania revealed no deviations from local expected genotype frequencies (Carter & McGregor, 1973; Carter & Voller, 1975).

Information on other *Plasmodium* species is also consistent with frequent genetic exchange in natural populations. Joshi *et al.* (1989) observed random combinations of alleles at 3 isoenzyme loci among 76 *P. vivax* isolates from northern India. An earlier study of isoenzyme polymorphism among clones of the rodent malaria parasite *P. chabaudi* from central Africa did not demonstrate any non-random associations of alleles at different loci (Beale, Carter & Walliker, 1978).
In the present study, genotype diversity was significantly lower in each of two villages than in the urban/periorurban region. A possible confounding variable with respect to Sarakunda village was that the isolates were collected on two consecutive days, and their diversity compared with expected values calculated from a sample collected in the urban/periorurban region over 6 months. However, the Brefet village sample was collected during the same period as the urban/periorurban sample, with which its genotype diversity was compared. Although the 3-locus genotype diversity was lower within the villages, it was not possible to determine whether the allelic diversity at each locus was also significantly lower, as the small sample sizes did not allow sufficient statistical power for these tests.

Small areas of the urban/periorurban region did not show a lower genotype diversity than the region as a whole. These areas were necessarily larger than the villages (in order to obtain sufficient sample sizes), so a direct comparison cannot be made.

The lower genotype diversity observed in the villages is consistent with the idea of a limited population size. The mean number of detectable \textit{P. falciparum} clones per acute infection in The Gambia is 2-0 (Conway \textit{et al.} 1991). Assuming that the mean number of clones within all infections is about 2, it follows that in an area of 50 \% \textit{P. falciparum} prevalence for example, the number of bloodstage clones will not exceed the number of people. If some of the bloodstage clones do not contribute to the next generation (Burkot \textit{et al.} 1984), the effective \textit{P. falciparum} population size will be even lower (Kimura & Ohta, 1971). Within a village of 500 people for example, the effective \textit{P. falciparum} population size will probably be small enough to result in a reduction of genetic diversity, due to local extinction of rare alleles by random genetic drift (Kimura & Crow, 1964). Even though \textit{P. falciparum} populations in different villages are not completely isolated from one another, the introduction of new allelic variants may occur at a slower rate than the local extinction of alleles, resulting in a lower genetic diversity than within a larger population.

The above argument has implications for the interpretation of epidemiological studies on \textit{P. falciparum}. Forsyth \textit{et al.} (1989) recorded differences in the prevalence of an allelic serotype of the S-antigen between villages in Papua New Guinea, and between samples taken on different dates from particular villages. It was suggested (Forsyth \textit{et al.} 1988) that frequency-dependent selection due to serotype-specific immunity was the most likely explanation for changes in the prevalence of the serotype. However, changes in allelic frequency are explainable by non-selective genetic drift in any population of limited size (Kimura, 1955).

The interpretation of studies on allelic polymorphism is often controversial, due to the difficulty in resolving the contribution of neutral and selective processes. With respect to \textit{P. falciparum}, many studies are undertaken on genes for which there is an implicit assumption of selection. To be informative, such studies require a knowledge of the population structure, and a recognition of the importance of neutral processes.

We thank the patients for their willing co-operation in the study. We are indebted to Drs Brian Greenwood, Steve Allen, Dominic Kwiatowski, Adrian Hill, other workers at the Medical Research Council Laboratories in The Gambia, and Mr Momadou Sanyang of Brefet village, for their clinical and organizational help. Mr Adam Eyre-Walker kindly wrote the \textit{pascal} program for us, while Drs Steve Bennett, Peter Keightley and Professor Bill Hill gave valuable advice on statistics. Drs Anthony Holder, Jeffrey Lyon, Richard Pink, Robert Reese and Alan Saul, provided some of the monoclonal antibodies. The study was supported by grants from the Medical Research Council, the Wellcome Trust, and the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases.

\textbf{REFERENCES}


Population structure of malaria parasites


Genetic evidence for the importance of interrupted feeding by mosquitoes in the transmission of malaria

David J. Conway* and Jana S. McBride Department of Cell, Animal, and Population Biology, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh, EH9 3JF, UK

Abstract

Plasmodium falciparum isolates were obtained from 17 pairs of Gambian children, each pair living in the same house and presenting with malaria at the same time. Frequencies of allelic serotypes of 3 polymorphic blood stage proteins (MSP1, MSP2, and Exp-1) were previously determined from a large number of isolates from patients in the local area, and the probability of a random pair of isolates containing an identical genotype was calculated to be less than 0.01. However, 3 of 8 household pairs in one year, and 6 of 9 in the next year, contained identical P. falciparum genotypes, a much higher frequency than expected randomly (P<0.0005, for each year). This finding is discussed in terms of the probable contribution of single mosquitoes infecting more than one person.

Introduction

Models of malaria transmission (reviewed by Aron & May, 1982) assume that a malaria vector mosquito can infect only a single person per blood meal. A vector mosquito inoculates infective sporozoites from its salivary glands while probing the skin in an attempt to locate a blood capillary. Sporozoites inoculated into the skin, not directly into a blood capillary, were sufficient to induce malaria in a human volunteer (Boyd, 1949). Experiments also showed that a single mosquito could infect more than one person due to interrupted blood feeding (Boyd, 1949).

In nature, some blood meals are obtained from more than one individual, due to interrupted feeding (Boreham & Garrett-Jones, 1973). Moreover, if a mosquito fails to locate a blood capillary initially, it may attempt to feed on another individual nearby, and thereby infect more than one individual (Rossignol & Mackay-Rossignol, 1988). However, the effects of repeated probing and/or interrupted feeding on the natural transmission of malaria have not been investigated.

There are inherent difficulties in correlating inoculation and incubation rates by conventional means in endemic areas (Pull & Grab, 1974). Therefore, it would be very difficult to estimate directly the effects on malaria transmission of repeated mosquito probing and/or interrupted blood feeding. A new approach to the problem involves the characterization of Plasmodium falciparum parasite genotypes from individuals with malaria. A single mosquito infecting 2 people would be expected to inoculate identical parasite genotypes into both of them. If this occurred frequently, in an area where the malaria vector bites indoors at night, patients from the same house presenting with malaria at the same time would be expected to share identical P. falciparum genotypes more often than those living in different houses in the same area.

Materials and Methods

Practical methods

The study was conducted in an urban/perurban area of The Gambia, West Africa, where P. falciparum malaria is transmitted by Anopheles gambiae sensu lato, which bites mainly between 2200 and 0400 h (Holstein, 1954). Malaria transmission is seasonal and hyperendemic, reaching a peak between August and November.

During July–December 1988 and October–December 1989 P. falciparum isolates were collected from 337 malaria patients presenting to the out-patients departments of the Medical Research Council, Fajara, and the Royal Victoria Hospital, Banjul (220 isolates in 1988, 117 in 1989). Among these patients were 17 pairs of siblings presenting together. The parents of 11 of the pairs were questioned about where their children slept in the house. In all cases, both children slept in the same room and, in all except one case, in the same bed. All the patients lived in the same urban/perurban area (Conway & McBride, 1991).

Heparinized samples of parasitized blood from each patient were cultured from 24–48 h until mature schizonts were obtained. Multi-spot slides of schizonts were prepared (McBride et al., 1984), for analysis of P. falciparum genotypes by indirect immunofluorescence using a panel of 27 monoclonal antibodies (as listed by Conway & McBride, 1991), recognizing allelic variants of 3 polymorphic blood stage proteins: (i) the precursor to the major merozoite surface proteins (MSP1, PMMSA, or gp195: Holder, 1988), (ii) a second merozoite surface protein (MSP2, MSA-2, or gp35–56: Fenton et al., 1991), and (iii) an exported protein (Exp-1, CRA, or p23: Simmons et al., 1987). These proteins are encoded by genes on different chromosomes (Kemp et al., 1987), which segregate at meiosis in the mosquito vector (Walliker et al., 1987).

At least 200 schizonts of each isolate were typed with each monoclonal antibody by indirect immunofluorescence (Conway & McBride, 1991). A method of two-colour differential labelling was used to resolve the majority parasite clone in any isolate containing more than one recognizable clone (Conway et al., 1991). Only the majority clone in each isolate was included in the statistical analyses.

Statistical methods

Among the 337 isolates, there were 36 distinguishable allelic serotypes of MSP1, 8 types of MSP2, and 2
types of Exp-1. From observed allelic frequencies at each locus, the expected frequencies of different three-locus genotypes (i.e., combinations of different

MSP1, MSP2, and Exp-1 types) were calculated for 1988 and 1989 independently, assuming random outbreeding in the parasite population. The observed numbers of different three-locus genotypes (126 of 220 sampled in 1988, and 75 of 117 sampled in 1989) were in accordance with the expectations of random assortment among the loci (CONWAY & MCBRIDE, 1991).

The probability ($P$) that a randomly selected pair of isolates would contain an identical three-locus genotype was calculated as the sum of the squares of the individual expected three-locus genotype frequencies ($P=0.00802$ for 1988, and $P=0.00898$ for 1989: data from CONWAY & MCBRIDE, 1991). The expected number of household pairs of isolates containing an identical $P. falciparum$ three-locus genotype was calculated as the binomial distribution of the above probability values for the total number of household pairs (8 in 1988, 9 in 1989).

Results

The three-locus genotype of the majority $P. falciparum$ clone within each of the household-paired children is given in the Table. The Figure shows the expected numbers of pairs containing identical genotypes as probability distribution histograms for Table. Three-locus genotypes of $P. falciparum$ isolated from pairs of children who slept in the same house.

<table>
<thead>
<tr>
<th>Pair no.</th>
<th>Initials</th>
<th>Age (years)</th>
<th>Sex</th>
<th>Genotype</th>
<th>Sleeping arrangement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1988</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>AN</td>
<td>3</td>
<td>F</td>
<td>15-8-2</td>
<td>NI</td>
</tr>
<tr>
<td></td>
<td>MN</td>
<td>1</td>
<td>M</td>
<td>28-5-1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>BJ</td>
<td>5</td>
<td>M</td>
<td>10-8-1*</td>
<td>NI</td>
</tr>
<tr>
<td></td>
<td>FJ</td>
<td>3</td>
<td>F</td>
<td>10-8-1*</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>MS</td>
<td>5</td>
<td>F</td>
<td>15-4-2</td>
<td>NI</td>
</tr>
<tr>
<td></td>
<td>AS</td>
<td>2</td>
<td>M</td>
<td>52-2-2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>AB</td>
<td>4</td>
<td>M</td>
<td>52-2-2*</td>
<td>NI</td>
</tr>
<tr>
<td></td>
<td>MB</td>
<td>6</td>
<td>F</td>
<td>52-2-2*</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>SJ</td>
<td>3</td>
<td>F</td>
<td>28-2-2</td>
<td>Same</td>
</tr>
<tr>
<td></td>
<td>AJ</td>
<td>1</td>
<td>F</td>
<td>32-2-2</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>AT</td>
<td>5</td>
<td>F</td>
<td>20-2-1</td>
<td>Same</td>
</tr>
<tr>
<td></td>
<td>ST</td>
<td>2</td>
<td>M</td>
<td>11-1-2</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>FC</td>
<td>4</td>
<td>F</td>
<td>11-7-2</td>
<td>NI</td>
</tr>
<tr>
<td></td>
<td>PC</td>
<td>&lt;1</td>
<td>M</td>
<td>12-2-2*</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>BM</td>
<td>&lt;1</td>
<td>F</td>
<td>12-2-2*</td>
<td>Same</td>
</tr>
<tr>
<td></td>
<td>KM</td>
<td>3</td>
<td>M</td>
<td>12-2-2</td>
<td></td>
</tr>
<tr>
<td>1989</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>SF</td>
<td>8</td>
<td>M</td>
<td>20-2-2</td>
<td>Different</td>
</tr>
<tr>
<td></td>
<td>MF</td>
<td>8</td>
<td>M</td>
<td>11-7-2</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>SJ</td>
<td>15</td>
<td>M</td>
<td>15-2-2</td>
<td>Same</td>
</tr>
<tr>
<td></td>
<td>EJ</td>
<td>12</td>
<td>M</td>
<td>15-2-2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>BM</td>
<td>5</td>
<td>F</td>
<td>7-1-1*</td>
<td>Same</td>
</tr>
<tr>
<td></td>
<td>TM</td>
<td>3</td>
<td>M</td>
<td>7-1-1*</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>NN</td>
<td>5</td>
<td>F</td>
<td>10-7-1</td>
<td>Same</td>
</tr>
<tr>
<td></td>
<td>SN</td>
<td>1</td>
<td>M</td>
<td>4-2-1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>SD</td>
<td>4</td>
<td>F</td>
<td>28-2-2</td>
<td>Same</td>
</tr>
<tr>
<td></td>
<td>JD</td>
<td>6</td>
<td>F</td>
<td>28-2-2</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>XS</td>
<td>9</td>
<td>M</td>
<td>52-7-2</td>
<td>Same</td>
</tr>
<tr>
<td></td>
<td>MS</td>
<td>12</td>
<td>M</td>
<td>41-7-2</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>AC</td>
<td>6</td>
<td>M</td>
<td>8-8-1*</td>
<td>Same</td>
</tr>
<tr>
<td></td>
<td>JC</td>
<td>12</td>
<td>F</td>
<td>8-8-1</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>KB</td>
<td>15</td>
<td>M</td>
<td>40-2-2</td>
<td>Same</td>
</tr>
<tr>
<td></td>
<td>OB</td>
<td>14</td>
<td>M</td>
<td>40-2-2</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>SC</td>
<td>4</td>
<td>F</td>
<td>16-6-2</td>
<td>NI</td>
</tr>
<tr>
<td></td>
<td>AC</td>
<td>6</td>
<td>F</td>
<td>16-6-2</td>
<td></td>
</tr>
</tbody>
</table>

F=female, M= male.

Discussion

Approximately 50% of pairs of children with malaria from the same household contained an identical three-locus $P. falciparum$ genotype, compared with approximately 1% of pairs randomly chosen from the study area as a whole.

This exceptionally high frequency of genetically
identical parasite pairs among patients from the same household could be explained in one of 2 ways: (i) single mosquitoes frequently infect 2 or more people on the same occasion, due to interrupted feeding and/or repeated probing, or (ii) extreme clustering of mosquitoes carrying identical P. falciparum genotypes exists within households, due to a lack of dispersal during the whole period of development of the parasite in the mosquito.

There is no evidence for the second explanation. An. gambiae s.l. are not likely to remain clustered in a household during the 10-14 d development period of the parasite, as they lay eggs at outdoor breeding sites approximately every 4 d (HOLSTEIN, 1954). Accordingly, no spatial clustering of P. falciparum genotypes was observed in the study area (CONWAY & McBRIEDE, 1991).

Therefore, it is very likely that single mosquitoes are frequently responsible for infecting more than one individual with malaria. Haptoglobin typing of blood meals revealed that 3-10% of human blood meals taken by An. gambiae s.l. in Africa were acquired from 2 or more individuals (BOREHAM et al., 1979; PORT et al., 1980). Here, the observed frequency of genetic identity between household paired isolates was approximately 50%, suggesting that mosquito probing activity before feeding also contributed significantly to malaria transmission.

Field studies of pre-feeding probing behaviour by malaria-vectors have not been attempted, and until now the effects on malaria transmission have not been demonstrated. There is probably great diversity in both mosquito probing behaviour and interrupted blood feeding, depending on factors such as host irritability (LENAHAN & BOREHAM, 1976) as well as salivary gland pathology (ROSSIGNOL et al., 1984), and therefore such behaviour might be difficult to assess precisely.

Mosquito pre-feeding probing behaviour will tend to increase malaria transmission, as a larger number of people receive infective bites. The results presented in this paper strongly suggest that the effects of such behaviour are important, and therefore that present models of malaria transmission underestimate the vectorial capacity of mosquito populations. Since a precise estimation of vectorial capacity may not be possible, a more realistic goal of mathematical models is the prediction of relative reduction in vectorial capacity under a given control effort (DYE, 1990).

Acknowledgements

We are grateful to workers at the Medical Research Council laboratories in the Gambia, and particularly to Drs Brian Greenwood, Hilton Whittle, Andrew Wilkins, Dominic Kwiatkowski, Harry Campbell and Adrian Hill, for making the work possible. We thank Dr Peter Keighley and Professor Bill Hill for statistical advice. This investigation received financial support from the Medical Research Council, The Wellcome Trust, and the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases.

References


Received 25 February 1991; accepted for publication 22 March 1991
**Plasmodium falciparum**: Intragenic Recombination and Nonrandom Associations between Polymorphic Domains of the Precursor to the Major Merozoite Surface Antigens


*Institute of Cell, Animal, and Population Biology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT Scotland, United Kingdom; †Department of Pharmacology and Therapeutics, University College Hospital, University of Ibadan, Ibadan, Nigeria; and ‡Medical Research Council, Fajara, P.O. Box 273 Banjul, The Gambia, West Africa

Conway, D. J., Rosario, V., Oduola, A. M. J., Salako, L. A., Greenwood, B. M., and McBride, J. S. 1991. *Plasmodium falciparum*: Intragenic recombination and nonrandom associations between polymorphic domains of the precursor to the major merozoite surface antigens. Experimental Parasitology 73, 469-480. Extensive allelic polymorphism in the *Plasmodium falciparum* major merozoite antigen precursor (MSP1/PMMSA) is partly due to intragenic recombination events within a short region near the 5' end of the gene. Newly described allelic sequences from this region of the gene are compared to those previously published, revealing additional sites of intragenic recombination. Epitopes recognised by monoclonal antibodies on the protein have been assigned on the basis of correlations between serology and amino acid sequence polymorphisms among different allelic types of MSP1. Serological analyses of MSP1 from 567 wild isolates from The Gambia, Nigeria, and Brazil reveal that certain pairs of epitopes, although sited on MSP1 domains separated by known sites of intragenic recombination, are highly significantly associated on parasites in endemic populations. Most associations are similar in the three countries. These associations are discussed with respect to the intragenic recombination hypothesis.

INDEX DESCRIPTORS AND ABBREVIATIONS: *Plasmodium falciparum*; Malaria antigen; Allelic polymorphism; Recombination; The precursor to the major merozoite surface proteins (MSP1, also termed PMMSA, MSA-1, PSA, p190, or gp195); Base pair (bp); Fluoresceine isothiocyanate (FITC); Rhodamine isothiocyanate (RITC); Monoclonal antibody (MAb); Phosphate-buffered saline (PBS); Polymerase chain reaction (PCR).

INTRODUCTION

The precursor to the major merozoite surface antigens of *Plasmodium falciparum* (MSP1, PMMSA, or p190) is polymorphic in natural populations of the parasite (Howard et al. 1986; Conway and McBride, 1991), a fact which may compromise its potential as a vaccine against malaria (Holder, 1988). Allelic polymorphism of the MSP1 gene has been analysed by comparisons of DNA sequences from different parasite isolates and clones (Weber et al. 1986; Tanabe et al. 1987; Peterson et al. 1988). Although certain domains of the gene are highly conserved, other domains exist as one or the other of two essentially dimorphic sequences, and most of the additional allelic polymorphism is a result of intragenic recombination events at sites near the 5' end of the sequence (Tanabe et al. 1987). Many different sequences could theoretically have been produced by recombination events between two parental alleles. Additional allelic sequences presented in this paper support this conclusion.

At the protein level, antigenic polymorphism of MSP1 has been demonstrated using a panel of variant-specific MAb
Bride et al. 1985; Conway and McBride, 1991). Approximate locations of certain polymorphic epitopes have been determined, as some of the MAbs recognise fusion proteins encoded by fragments of the MSP1 gene (Gentz et al. 1988; Früh et al. 1991), or naturally occurring processing fragments of the native protein (Howard et al. 1985; Lyon et al. 1987; Holder et al. 1987; McBride and Heidrich, 1987; Holder, 1988). An alternative approach is to identify sequence polymorphisms which correlate with serological differences among allelic types of the protein. In the present study, putative locations of different polymorphic epitopes are proposed on the basis of correlation with sequence polymorphisms.

Finally, to determine the extent of recombinational polymorphism in natural populations of *P. falciparum*, MSP1 variants from a large number of clinical isolates from The Gambia, Nigeria, and Brazil were serotyped for various polymorphic epitopes. A total of 39 different MSP1 serotypes were identified among the wild parasites. Nonrandom associations between epitopes on different domains of MSP1 were revealed by statistical analyses. The associations suggest that, though intragenic recombinations do occur, strong intragenic disequilibria exist within MSP1 in natural populations of *P. falciparum*.

**MATERIALS AND METHODS**

*P. falciparum* Genomic DNA Preparations

*P. falciparum* clones T9/96 and T9/101, expressing different MSP1 serotypes (Thaithong et al. 1984; McBride et al. 1985), were obtained from the WHO Registry of Standard Malaria Strains at Edinburgh University. Frozen stabiles were thawed and cultured in vitro (Trager and Jensen, 1976) until schizonts were obtained at approximately 5% parasitaemia (in 5 ml of culture at 3% haematocrit). Erythrocytes were lysed in a 0.15% saponin solution in phosphate-buffered saline (PBS), and parasites pelleted and washed twice in PBS before being frozen at −70°C. A wild *P. falciparum* isolate GF88-160 was obtained from 10 ml of venous blood from a Gambian patient with 5% parasitaemia and cultured in vitro for 48 hr until mature schizonts developed. Schizont-infected erythrocytes were concentrated by Plasmagel (Pasvol et al. 1978), washed twice in PBS, and frozen at −70°C.

Each parasite pellet was thawed and suspended in 400 μl of 150 mM NaCl, 25 mM EDTA (pH 8.0) solution; 10 μl of 10% SDS solution and 10 μl of 10 mg/ml Proteinase K (Sigma) solution were added before incubation at 37°C overnight. Proteins were removed using phenol–chloroform (1:1), and DNA was ethanol precipitated before being redisolved in 10 mM Tris, 1 mM EDTA (pH 8.0) solution (Fenton et al. 1991).

**DNA Amplification, Cloning, and Sequencing**

The polymerase chain reaction (PCR: Saiki et al. 1985) was used to amplify a polymorphic region of the MSP1 gene corresponding to base pairs 774–1087 of the coding sequence of the MAD20 allele (Tanabe et al. 1987). Synthetic 24-mer (5’-TGAAGGAAGTAA-GAAACAATTGA-3’) and 28-mer (5’-TCTAATTCAAGTGGATCAGTAAATAAAC-3’) amplification primers corresponded to conserved flanking sequences. The amplification mixture contained 2 units of Taq DNA polymerase (Perkin Elmer Cetus), 0.1 ng of parasite genomic DNA, 10 μl of a 1 μM solution of each amplification primer, 2 μl of a 10 mM solution of each dNTP, in a total volume of 100 μl of 50 mM KCl, 10 mM Tris–HCl (pH 8.3), 1.5 mM MgCl2, 0.01% gelatin (w/v). After initial DNA denaturation at 95°C for 5 min the first 3 cycles of DNA amplification were performed using an annealing temperature of 45°C, polymerisation at 70°C, and denaturation at 92°C. The annealing temperature was set at 40°C for the remaining 27 cycles in order to maximise the yield of product.

The PCR product was fractionated in 1% agarose gel (in 40 mM Tris base, 19.7 mM acetic acid, 1 mM EDTA pH 8.0), and purified using the Gene clean kit (Bio 101 Inc.). The purified product was kinased and cloned into the Smal site of phage M13 (mp18 or mp19). Recombinant phages were grown to produce single-stranded DNA template for sequencing by the dideoxy chain termination method (Sanger et al. 1977), using the Sequenase II kit (United States Biochemicals) with the universal primer.

**P. falciparum Schizont Antigen Slides for Immunofluorescence Typing**

*P. falciparum* isolates were obtained from 445 malaria patients in the Gambia in 1988 and 1989; 60 patients in Ibadan, Nigeria in 1989; and 62 patients in the Amazon basin, Brazil, during 1983–1989. Blood, 0.3 ml, was obtained from each patient by fingerprick or as part of a venous sample obtained for other studies, after consent, under approval from the appropriate government, scientific, and ethical committees. The heparinised blood was washed 3 times in sterile PBS,
and parasites were cultured (Trager and Jensen, 1976) for 24–48 hr until schizonts had matured. After washing the cultured cells three times and resuspending at 3% haematocrit in PBS, multispot slides were prepared with 20-µl cell suspension per spot, were air dried (using a desiccator or a well air-conditioned room), and stored under desiccation at −20°C (McBride et al. 1984). Slides were also prepared from culture-adapted clones and isolates (T9/96, T9/101, MAD20, FC27, RO-33, Camp, Palo Alto, K1 and Wellcome), obtained from the WHO Registry of Standard Malaria Strains, University of Edinburgh.

**Indirect Immunofluorescence Typing with Monoclonal Antibodies**

Table 1 describes 19 MSP1 specific murine monoclonal antibodies (MAbs). Three MAbs recognised conserved epitopes and 16 recognised epitopes on polymorphic domains. Acetone-fixed schizonts of each isolate were first tested with the individual MAbs, using a FITC-conjugated rabbit anti-mouse IgG (ICN Immunobiologicals) to allow visualisation of positively reacting parasites microscopically (McBride et al. 1985). Two-colour fluorescence analyses, using pairs of MAbs of different isotypes, and isotype-specific FITC- and RITC-conjugated second antibodies (Southern Biotechnology Associates Inc.) were performed to resolve the MSP1 profile of the majority parasite clones within certain mixed isolated (Conway and McBride, 1991; Conway et al. 1991).

**Statistical Tests**

Associations between polymorphic epitopes on different domains of MSP1 were determined by χ² tests on 2 × 2 contingency tables, which contained the numbers of parasite clones with both, neither, or either one of a given pair of epitopes. The analyses were performed separately for collections of parasites obtained from populations in The Gambia, Nigeria, and Brazil.

**RESULTS**

**Nucleotide Sequences and Putative Intragenic Recombination Sites within a Region of the MSP1 Gene**

Figure 1 compares nucleotide sequences of a PCR-amplified region of the MSP1 gene from Thai parasite clones T9/96 and T9/101, and an uncloned Gambian isolate GF88-160 (containing only one MSP1 serotype), with the previously published MAD20 and Camp sequences (Weber et al. 1986; Tanabe et al. 1987).

Each of the three newly presented sequences differs from those previously published, but each can be explained as a result of intragenic recombination events between parental sequences identical to the Camp and MAD20 sequences, except for two base pair differences. One of these differences, thymidine at position 977 in the T9/96 sequence, is previously undescribed in other MSP1 alleles, and might conceivably represent a PCR misincorporation (four independent clones, two forward strand and two reverse strand, were obtained and sequenced from one PCR amplification mix). The T9/96 sequence is identical to Camp upstream of bp 860 and to MAD20 downstream of bp 926 (bp 860–926 marked as recombination site 2 in Fig. 1). The T9/101 sequence is identical to MAD20 upstream of bp 993 and to Camp downstream of bp 999 (bp 993–999 marked as site 3). The GF88-160 sequence is identical to Camp upstream of bp 832 and to T9/101 downstream of bp 838 (bp 832–838 marked as site 1). Sites 1, 2, and 3 therefore mark putative sites of intragenic recombination. Site 3 is located in the variable domain 4 (Tanabe et al. 1987) which is of particular interest because it shows for the first time that recombinations can occur within variable domains rather than only between them (Peterson et al. 1988; Jongwutiwes et al. 1991).

**Epitope Locations Deduced from Correlations between Amino Acid Sequences and Reactivities with MAbs**

Figure 2 compares deduced amino acid sequences of the T9/96, T9/101, and GF88-160 alleles to other published sequences of MSP1 (covering a region corresponding to residues 113–362 of the MAD20 sequence, Tanabe et al. 1987). The sequences are correlated to serological reactions of the allelic products with MAbs 9.5 or 13.2, and 12.1 or 10-2B.

MSP1 variants positive for epitope 12.1 all share a 12 amino acid sequence (PLPENKKKEVEG), while those positive for
the alternative epitope 10-2B share an alternative 11 amino acid sequence (TLLDKNKKIEE). The putative linear epitopes 12.1 and 10-2B are shown boxed in Fig. 2. Of more than 550 *P. falciparum* isolates tested, no parasites have been seen to be positive for both these alternative epitopes. Parasites negative for both epitopes are rare, and the single example which has been sequenced, RO-33, has a third alternative sequence in the corresponding region.

Epitopes 13.2 and 9.5 are also mutually exclusive. Parasites positive for epitope
Fig. 2. Correlations between amino acid sequences and monoclonal antibody serology for a polymorphic region of MSP1 in 10 *P. falciparum* clones and isolates. Putative epitope determinants are shown boxed. References to published sequences: MAD20, Tanabe et al. 1987; FC27, Peterson et al. 1988; RO-33, Certa et al. 1987; Camp, Weber et al. 1986; Palo Alto, Chang et al. 1988; K1, Mackay et al. 1985; Wellcome, Holder et al. 1985.
13.2 have glutamate at residue 287 (numbered according to the MAD20 sequence), while those positive for the alternative epitope 9.5 have glutamine at residue 287. However, the sequence from isolate GF88-160, which is negative for both 13.2 and 9.5, also has glutamine at this residue. Therefore, the 9.5 epitope cannot be explained by this residue alone. The 9.5 epitope may be conformational, requiring glutamine in this position as a part of the epitope. Residue 127, upstream of the region cloned and sequenced here, could also be important for the epitope, since parasites positive for 9.5 have glutamine and those positive for 13.2 have arginine (Fig. 2).

Figure 2 also summarises antigenic consequences of the recombination events within the region (Fig. 1). For example, the different epitope combinations of the T9/96 and T9/101 proteins (13.2 + 12.1+, and 9.5 + 10-2B+, respectively) can be seen as the results of recombination events between the MAD20 (9.5 + 12.1+) and Camp (13.2 + 10-2B+) sequences.

Figure 3 schematically illustrates the approximate locations of the above epitopes, as well as other epitopes which are used below to characterise allelic serotypes of MSP1 on parasites from clinical isolates.

### Allelic Serotypes of MSP1 in Endemic Populations of P. falciparum

MSP1 serotypes of parasites in clinical isolates of *P. falciparum* were defined as different combinations of polymorphic epitopes described in Table 1 and Fig. 3. Many isolates contained more than one *P. falciparum* clone and in such isolates the MSP1 serotype of the majority clone was

---

### TABLE 1

**Summary of Specificities of Monoclonal Antibodies against MSP1**

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Reciprocal dilution for IFA</th>
<th>Isotype</th>
<th>Epitope and location</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.8-4-4-1</td>
<td>1000</td>
<td>IgG1</td>
<td>*conserved conformational</td>
<td>2,3</td>
</tr>
<tr>
<td>12.4-3-4</td>
<td>500</td>
<td>IgG1</td>
<td>*conserved conformational</td>
<td>2,3</td>
</tr>
<tr>
<td>12.8-2</td>
<td>1000</td>
<td>IgG1</td>
<td>*conserved block 16-17, 16 K fragment</td>
<td>3</td>
</tr>
<tr>
<td>12.2-1-1</td>
<td>2000</td>
<td>IgG1</td>
<td>polymorphic block 2 repeats</td>
<td>1-3</td>
</tr>
<tr>
<td>3D3.10</td>
<td>1000</td>
<td>IgG1</td>
<td>polymorphic block 2 repeats</td>
<td>6</td>
</tr>
<tr>
<td>9.5-1-5-1</td>
<td>500</td>
<td>IgG2b</td>
<td>polymorphic 80 K fragment</td>
<td>1-3</td>
</tr>
<tr>
<td>13.2-3</td>
<td>2000</td>
<td>IgG1</td>
<td>polymorphic 80 K fragment</td>
<td>3</td>
</tr>
<tr>
<td>12.1-5-4</td>
<td>2000</td>
<td>IgG1</td>
<td>polymorphic block 4, 80 K fragment</td>
<td>1-3, 8, Fig. 2</td>
</tr>
<tr>
<td>10-2B</td>
<td>2000</td>
<td>IgG2a</td>
<td>polymorphic block 4, 80 K fragment</td>
<td>4, Fig. 2</td>
</tr>
<tr>
<td>9.2-6-2</td>
<td>2000</td>
<td>IgG1</td>
<td>♂dimorphic block 12-14</td>
<td>1-3, 8</td>
</tr>
<tr>
<td>9.7-1</td>
<td>500</td>
<td>IgG1</td>
<td>♂dimorphic conformational</td>
<td>1-3</td>
</tr>
<tr>
<td>10.3-2</td>
<td>500</td>
<td>IgG1</td>
<td>♂dimorphic conformational</td>
<td>1-3</td>
</tr>
<tr>
<td>1-1C</td>
<td>500</td>
<td>IgG1</td>
<td>♂dimorphic 80 K fragment</td>
<td>4</td>
</tr>
<tr>
<td>7.3-7</td>
<td>1000</td>
<td>IgG2a</td>
<td>♂dimorphic conformational</td>
<td>1-3</td>
</tr>
<tr>
<td>34-5</td>
<td>undiluted</td>
<td>IgG1</td>
<td>♂dimorphic 36 K fragment</td>
<td>5</td>
</tr>
<tr>
<td>6.1-1-3</td>
<td>500</td>
<td>IgG1</td>
<td>♂dimorphic block 16, 40 K fragment</td>
<td>1-3, 8</td>
</tr>
<tr>
<td>13.1-2</td>
<td>2000</td>
<td>IgG1</td>
<td>♂dimorphic block 16, 40 K fragment</td>
<td>1-3, 8</td>
</tr>
<tr>
<td>17.1-3</td>
<td>2000</td>
<td>IgG1</td>
<td>♂dimorphic block 16, 40 K fragment</td>
<td>1-3, 8</td>
</tr>
<tr>
<td>111.4</td>
<td>1000</td>
<td>IgG1</td>
<td>polymorphic block 16–17, 16 K fragment</td>
<td>7</td>
</tr>
</tbody>
</table>

**Note.** In this study, serological epitopes of MSP1 are referred to by the same code numbers as the mAbs which identify them. Epitopes marked * are conserved among all isolates. Polymorphic epitopes marked by identical symbols (♂ or ♀) exhibit identical allelic distributions. Sequence block numbers are as proposed by Tanabe et al. (1987). Locations of MSP1 epitopes are shown schematically in Fig. 1. References: 1. McBride et al. 1984. 2. McBride et al. 1985. 3. McBride and Heidrich 1987, and unpublished. 4. Howard et al. 1985. 5. Gentz et al. 1988. 6. Lyon et al. 1987. 7. Holder et al. 1985. 8. Früh et al. 1991.
Fig. 3. Alternative epitope specificities at different domains of MSP1 are shown boxed. MSP1 sequence polymorphism is shown schematically divided into 17 domains or "blocks," according to Tanabe et al. (1987). Unshaded blocks are the least polymorphic, hatched blocks are more so, and solid blocks are most polymorphic. The locations of epitopes 12.1, 10-2B, 13.2, and 9.5 are discussed in the text. The approximate locations of the other epitopes are listed in Table I.

Nonrandom Associations between Polymorphic Epitopes of MSP1

Certain pairs of epitopes are mutually exclusive, never occurring together on the same parasite. This is expected for epitopes determined by alternative variant sequences, e.g., 12.1 and 10-2B (domain 4). Different combinations among epitopes on domain 3 (13.2 or 9.5) and domain 4 (12.1 or 10-2B) have been shown to be due to intragenic recombination between these domains (as detailed above). Similarly, epitopes 12.2 and 3D3 (both on domain 2) are not mutually exclusive, being detected either singly or together on the same parasite (Fig. 4). The different combinations of epitopes within domain 2 suggest that intragenic recombination events within domain 2 gave rise to recombinant proteins. In contrast, all epitopes on domains 6–16 (Figure 3) grouped into only two distinct alternative specificities, suggesting that there has been no intragenic recombination in the corresponding portion of the sequence (epitopes 6.1, 13.1, 17.1, 7.3, 1-IC, and 34-5 expressed by the Ki/Wellcome type, and epitopes 9.2, 9.7, and 10.3 expressed by the MAD20 type).

$\chi^2$ tests revealed nonrandom associations between epitopes on different domains separated by known sites of intragenic recombination. Figure 5 summarises the significant positive and negative associations between epitopes on different domains of MSP1 in The Gambia, Nigeria, and Brazil. For example, epitope 13.2 occurred more frequently together with epitope 10-2B, while epitope 9.5 occurred more frequently with epitope 12.1. The absence of a significant association between epitopes 13.2 and 10-2B in Brazil (in contrast with The Gambia and Nigeria) may be due to the fact that only 28 of the Brazilian isolates were tested for 10-2B, so the statistical power of the $\chi^2$
P. falciparum: POLYMORPHISM IN MEROZOITE SURFACE ANTIGEN

The Gambia

N = 445

Nigeria

N = 60

Brazil

N = 62

Fig. 4. MSP1 allelic serotype frequencies in The Gambia, Nigeria, and Brazil. Each serotype is defined by a unique combination of polymorphic epitopes shown at the bottom of the figure. Certain epitope combinations have never been observed (e.g., types 17, 18, 19, and 21, etc.). Combinations of known epitopes on domain 2 are summarized as follows: a = 3D3 +, 12.2; β = 3D3 +, 12.2; γ = 3D3 −, 12.2; δ = 3D3 −, 12.2−. Epitopes 1–IC, 7.3, 13.1, and 17.1, which were tested for on all Nigerian and Gambian isolates, and epitope 34-5 tested for on only Gambian isolates, showed identical serotype specificity to epitope 6.1 and are omitted from the figure for the sake of clarity. The Brazilian isolates were not tested with MAbs 9.7 or 10.3, and only 28 of the 62 Brazilian isolates were tested with MAb 10-2B, although all those negative for MAb 12.1 are shown here as "10-2B +." Two isolates in Nigeria contained a rare phenotype 12.1-, 10-2B-, otherwise similar to serotype 20 or 28, and one Gambian isolate contained the phenotype 12.1-, 10-2B-, otherwise similar to serotype 5 or 13 (not included in the figure).

Fig. 5. Summary of statistical associations between epitopes at different domains of MSP1, among parasites sampled from The Gambia, Nigeria, and Brazil. Epitopes depicted as boxes are joined by lines indicating a statistically significant nonrandom association. A solid line indicates a positive association, and a broken line a negative association. χ² p values were all <0.005, except for 12.2 vs 12.1 and 10-2B vs 9.5 in Nigeria (P < 0.01); 10-2B vs 111.4 and 9.5 vs 12.1 in Nigeria; and 9.5 vs 12.1 in Brazil (P < 0.05). Epitope 12.2 was rare in Brazil and therefore not included in statistical analyses for that country.

Test was lower. Most of the statistical associations were similar in the three countries, with no association being reversed (from positive to negative, or vice versa) in one country compared to another. Interestingly, a positive association between epitopes 12.1 (domain 4) and 111.4 (domain 16/17) was observed only in Brazil.

The large number of significant associations between epitopes on different domains indicates that, despite the occurrence of intragenic recombination between these domains, recombinant alleles do not exist at equilibrium frequencies in natural populations. Mechanisms to explain this are considered under Discussion.

Epitopes on domains 6–16 were excluded from the χ² analyses, since one of the two alternative variants (K1/Wellcome recognised by MAbs 6.1, 13.1, 17.1, 7.3, 1-IC, and 34-5) was present at such a low frequency in each of the countries that the analysis would be invalid. However, the 21 isolates containing such parasites all had an identical profile of epitopes at the other domains (serotype 52 in Fig. 4). This was in marked contrast with parasites with the other form of domains 6–16 (MAD20 recognised by MAbs 9.2, 9.7, and 10.3), which exhibited a very high diversity of epitope combinations at other domains (the remaining 38 different serotypes, Fig. 4).

DISCUSSION

Three sites of intragenic recombination toward the 5′ end of the MSP1 gene are proposed here to explain the newly described allelic sequences, in addition to sites proposed by Tanabe et al. (1987). This confirms and extends the evidence for a
clustering of recombination sites within a relatively short 5' region of the gene (covering less than 20% of the total sequence), discussed by Peterson et al. (1988).

Location of epitope determinants by sequence-serology correlations has been previously undertaken for chicken lysozyme c (Smith-Gill et al. 1982), murine MHC antigens (Landais et al. 1985), and the HLA-DR, -DQ, and -DP loci (Bugawan et al. 1988; Horn et al. 1988; Marsh and Bodmer, 1989). The approach has proved useful for linear epitopes, although there is a potential complication if a critical conformational change in an epitope is determined by an amino acid substitution remote from the epitope site (White et al. 1978).

Here, comparisons of allelic sequences within a polymorphic region of MSP1 explain the observed alternate reactivities of MAbs 12.1 and 10.2B, and the likely position of the epitopes has been deduced (amino acids 334-345 in domain 4, according to the scheme of Tanabe et al. 1987). Similarly, amino acids 127 and/or 287 in domain 3 may determine another pair of alternate epitopes, 13.2 and 9.5. The present analysis has also shown clearly that intragenic recombinations can be identified serologically as combinations of the respective epitopes.

If recombination events occur frequently within the MSP1 gene, and the recombinant proteins are not under differential selection, then polymorphic epitopes at different domains of the gene would be expected to assort randomly and the combinations exist at equilibrium frequencies in natural populations. However, the existence of strong nonrandom associations between epitopes on different domains shows that intragenic recombinants are not at equilibrium frequencies. Most notably, parasites with one of the two types of domains 6–16 (K1/Wellcome) all had an identical combination of epitopes at the other domains. Such nonrandom associations could arise by either of the following mechanisms.

First, if intragenic recombination is a rare event, random allelic frequency changes in natural populations could explain the observed disequilibria between polymorphic domains of MSP1. Although there are several recombination sites within the MSP1 gene, it does not follow that recombination at these sites has occurred frequently. It is known that in *P. falciparum* populations, intergenic recombinations between unlinked loci occur frequently at meiosis (Walliker et al. 1987; Conway and McBride 1991), but intragenic recombination events are likely to be much less frequent. To obtain an estimate of the frequency of such events would require extensive analyses of the MSP1 gene from progeny of a number of experimental crosses of the parasite.

Second, it is conceivable that recombinant alleles differ in “fitness,” and thus may differ in frequency, generating intragenic disequilibria even if intragenic recombination is a frequent event. The observed similarities between the epitope associations in three geographically separate parasite populations suggest that unknown selective factors might consistently operate in favour of particular recombinant alleles. Since the function of MSP1 is unknown, mechanisms whereby particular alleles might have a selective advantage must remain conjectural at this time.

Tanabe et al. (1987) speculated that polymorphism of MSP1 may be related to recognition of different erythrocyte surface receptors. *P. falciparum* isolates differ in their ability to invade glycophorin-deficient or sialic acid-deficient red blood cells *in vitro* (Mitchell et al. 1986; Hadley et al. 1987; Perkins and Holt 1988), but whether this ability is linked to MSP1 polymorphism has not been investigated (Perkins 1989). Alternatively, since MSP1 can induce a protective antiparasite immunity (Perrin et al. 1984; Hall et al. 1984; Siddiqui et al. 1987), it is possible that some of the allelic polymorphism is selectively maintained by acquired variant-specific immunity.

Future studies are required to indicate whether the clustering of intragenic recom-
bination sites, and the observed disequilibria between polymorphic domains of MSP1, are a result of selective constraints on the protein.

ACKNOWLEDGMENTS

We thank the patients for their willing cooperation in this study, and Drs. A. Couto, S. Oliveira (Instituto Evandro Chagas, Belém, Brazil), D. Kwiatkowski, A. V. S. Hill (MRC Laboratories, The Gambia), O. Walker, and A. Sowunmi (University College Hospital, Ibadan, Nigeria) for valuable help in sample collection. We also thank Dr. A. Khan and Mrs. J. Robinson for technical guidance; Drs. J. A. Lyon, R. Pink, R. T. Reese, and A. A. Holder for providing some of the monoclonal antibodies; and Dr. E. M. Riley for comments on the manuscript. The investigation received financial support from the Medical Research Council, the Wellcome Trust, and the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases.

REFERENCES


Received 29 March 1991; accepted with revision 12 July 1991.
Longitudinal Study of *Plasmodium falciparum* Polymorphic Antigens in a Malaria-Endemic Population

DAVID J. CONWAY,† BRIAN M. GREENWOOD,² AND JANA S. MCBRIDE¹*

Institute of Cell, Animal and Population Biology, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh EH9 3JT, United Kingdom, and Medical Research Council Laboratories, Fajara, Banjul, The Gambia²

Received 2 July 1991/Accepted 26 December 1991

*Corresponding author.
† Present address: Department of Biology, Imperial College, London SW7 2BB, United Kingdom.

**Plasmodium falciparum** merozoite surface antigens MSP1 and MSP2 and an exported antigen, Exp-1, exhibit allelic polymorphism in natural populations. To explain this, one hypothesis is that antigen polymorphisms are maintained by frequency-dependent immune selection. An expectation of the hypothesis is that rare variants have an advantage over common variants because of a lower level of acquired immunity against them and thus increase in frequency until an equilibrium is attained. To test this hypothesis, the frequencies of polymorphic epitopes of MSP1, MSP2, and Exp-1 were determined among isolates from malaria patients in an urban area of The Gambia, during different periods of one transmission season (1988) and in different years (1982, 1983, 1988, and 1989). The frequencies remained very stable throughout the period of study, alternative epitope variants remaining either rare or common, without shifts in relative frequencies. These results are discussed with reference to the immune-selection hypothesis, with the conclusion that frequencies of the major dimorphic serological classes of MSP1 are probably not maintained by immune selection.

The existence of antigenic diversity among "strains" of *Plasmodium falciparum* was suggested originally to explain the low rate of acquisition of immunity to malaria by inhabitants of endemic areas and the apparent recurrence of susceptibility among "immune" adults who travelled from one endemic area to another. In experimental infections of nonhuman primates, acquired immunity appeared to involve a "strain-specific" component (3, 33, 43), lending support to the importance of antigenic diversity. However, data from human infections were less consistent in this respect (1, 2, 22, 23).

Serological characterization of *P. falciparum* isolates subsequently demonstrated heterogeneity of soluble parasite antigens (44), internal and surface antigens (10, 27, 28), and infected erythrocyte surface components (20, 25, 42).

DNA sequencing has elucidated the genetic basis for the serological polymorphism of certain proteins, including the precursor to the major merozoite surface antigen (PMMSA or MSP1) (9, 19, 32, 39), a second merozoite surface antigen (MSA-2 or MSP2) (10, 34, 37, 40), and an exported protein (Exp-1) (36). MSP1 and MSP2 are among antigens considered for the development of a malaria vaccine. Immunization with MSP1 elicited a degree of immunity in monkeys (18, 31, 35). Immunization experiments with MSP2 have not been reported, but monoclonal antibodies (MAbs) against MSP2 partially inhibit in vitro invasion of erythrocytes (5, 34). Exp-1 may also be immunologically important, since it cross-reacts with the NANP(n) repeat epitope of the circumsporozoite protein (36).

It is not known whether polymorphisms in any of these antigens enable the parasite to evade variant-specific acquired immunity. If so, such polymorphisms might be subject to frequency-dependent selection, rare variants having a selective advantage because of lower levels of acquired immunity against them.

In this study, frequencies of variant epitopes on different domains of MSP1, MSP2, and Exp-1 have been determined for *P. falciparum* isolates collected from a small area of The Gambia over the period 1982 to 1989. The results are discussed with reference to the hypothesis of frequency-dependent selection by variant-specific immunity.

**MATERIALS AND METHODS**

**Study area and patients.** In The Gambia, malaria transmission is seasonal, with most infections with *P. falciparum* occurring from July through November (16). Blood samples were obtained from 424 malaria patients with *P. falciparum* infection at a density of ≥5 parasites per high-power field (x1,000 magnification), equivalent to a parasite density of ≥2,500 per μl (15). None of the patients was positive for other *Plasmodium* species. The patients presented to the Outpatients Departments of the Medical Research Council (MRC), Fajara, or the Royal Victoria Hospital, Banjul, from October through December 1982 (n = 46), October through December 1983 (n = 31), July through December 1988 (n = 228), and October through December 1989 (n = 119), and all lived within a radius of 7 km in an urban/peri-urban area described previously (7, 8).

**Blood collection and parasite culture.** A 0.3-ml sample of blood was obtained from each patient by finger prick or as part of a venous sample obtained for other studies, after consent, with approval from the Gambian government and the MRC Scientific and Ethical Committees. The heparinized blood was washed three times in sterile phosphate-buffered saline, pH 7.3 (PBS), and parasites were cultured (41) for 24 to 48 h, until schizonts were obtained (29). After the cells were washed three times and resuspended at approximately 3% hematocrit in PBS, multipot schizont slides were prepared with approximately 20 μl of cell suspension per spot. The slides were dried in a well-air-conditioned room and stored under desiccation at −20°C.

Copyright © 1992, American Society for Microbiology
TABLE 1. Summary of MAbs specific for epitopes of MSP1, MSP2, and Exp-1 antigens of *P. falciparum*

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Antibody</th>
<th>Reciprocal dilution for IFA</th>
<th>Isotype</th>
<th>Epitope and location</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSP1</td>
<td>9.8-4-4-1</td>
<td>1,000</td>
<td>IgG1</td>
<td>Conserved conformational</td>
<td>26, 27</td>
</tr>
<tr>
<td></td>
<td>12.4-3-4</td>
<td>500</td>
<td>IgG1</td>
<td>Conserved conformational</td>
<td>26, 27</td>
</tr>
<tr>
<td></td>
<td>12.8-2</td>
<td>1,000</td>
<td>IgG2b</td>
<td>Conserved domain 16–17, 16K fragment</td>
<td>13, 26</td>
</tr>
<tr>
<td></td>
<td>12.2-1-1</td>
<td>2,000</td>
<td>IgG1</td>
<td>Polymorphic domain 2 repeats</td>
<td>26, 27, 29</td>
</tr>
<tr>
<td></td>
<td>12 3D3.10</td>
<td>1,000</td>
<td>IgG2b</td>
<td>Polymorphic domain 2 repeats</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>9.5-1-5-1</td>
<td>500</td>
<td>IgG2b</td>
<td>Polymorphic domain 3, 80K fragment</td>
<td>26, 27, 29</td>
</tr>
<tr>
<td></td>
<td>13.2-3</td>
<td>2,000</td>
<td>IgG1</td>
<td>Polymorphic domain 3, 80K fragment</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>12.1-5-4</td>
<td>2,000</td>
<td>IgG1</td>
<td>Polymorphic domain 4, 80K fragment</td>
<td>9, 13, 26, 27</td>
</tr>
<tr>
<td></td>
<td>10-2B</td>
<td>2,000</td>
<td>IgG2a</td>
<td>Dimorphic conformational*</td>
<td>13, 26, 27</td>
</tr>
<tr>
<td></td>
<td>9.2-6-2</td>
<td>2,000</td>
<td>IgG1</td>
<td>Dimorphic conformational*</td>
<td>27, 29</td>
</tr>
<tr>
<td></td>
<td>9.7-1</td>
<td>500</td>
<td>IgG1</td>
<td>Dimorphic conformational*</td>
<td>27, 29</td>
</tr>
<tr>
<td></td>
<td>10.3-2</td>
<td>500</td>
<td>IgG1</td>
<td>Dimorphic conformational*</td>
<td>27, 29</td>
</tr>
<tr>
<td></td>
<td>1-IC</td>
<td>500</td>
<td>IgG1</td>
<td>Dimorphic, 80K fragment†</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>7.3-7</td>
<td>1,000</td>
<td>IgG2a</td>
<td>Dimorphic conformational†</td>
<td>26, 27, 29</td>
</tr>
<tr>
<td></td>
<td>34-5</td>
<td>500</td>
<td>IgG1</td>
<td>Dimorphic, 36K fragment†</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>6.1-1-3</td>
<td>500</td>
<td>IgG1</td>
<td>Dimorphic domain 16, 40K fragment†</td>
<td>13, 26, 27, 29</td>
</tr>
<tr>
<td></td>
<td>13.1-2</td>
<td>2,000</td>
<td>IgG1</td>
<td>Dimorphic domain 16, 40K fragment†</td>
<td>13, 26, 27, 29</td>
</tr>
<tr>
<td></td>
<td>17.1-3</td>
<td>2,000</td>
<td>IgG1</td>
<td>Dimorphic domain 16, 40K fragment†</td>
<td>13, 26, 27, 29</td>
</tr>
<tr>
<td></td>
<td>111.4</td>
<td>1,000</td>
<td>IgG1</td>
<td>Dimorphic domain 16–17, 16K fragment</td>
<td>19</td>
</tr>
</tbody>
</table>

| MSP2    | 12.3-2-2 | 500                         | IgG1    | Dimorphic group A-specific region† | 10, 29 |
|         | 12.5-1-2 | 500                         | IgG1    | Dimorphic group A-specific region† | 10, 29 |
|         | 13.4-2-1 | 500                         | IgG1    | Group A polymorphic repeat (GSAG) | 5, 10 |
|         | 8-5D     | 200                         | IgM     | Polymorphic§ | 10, 38 |
|         | 4-4F     | 200                         | IgM     | Polymorphic§ | 10, 38 |
|         | 8G10/48  | 300                         | IgG2b   | Dimorphic group B-specific region (STNS) | 10, 34 |
|         | 8F6/49   | 50                          | IgG3    | Group B polymorphic (DTPTATE) | 10, 34 |
| Exp-1   | 5-1-4    | 500                         | IgG1    | Dimorphic NADP | 36 |

*In this study, serological epitopes are referred to by the same code numbers as the MAbs which identify them. Conserved MSP1 epitopes were detectable on schizonts of all isolates. Polymorphic epitopes marked by identical symbols (*, †, ‡, or §) exhibit identical allelic distributions. Locations of MSP1 epitopes are shown schematically in Fig. 1.*

**IFA**, immunofluorescence assay.

1. IgG1, immunoglobulin G1.
2. **—**, undiluted.

MABS. The working dilution, source, and specificity of each of the 27 murine MAbs against MSP1, MSP2, and Exp-1 are listed in Table 1. Three of the MAbs against MSP1 recognize epitopes which are conserved among all isolates previously studied, and 16 recognize MSP1 variant epitopes. Locations of the variant epitopes on different domains of MSP1 are shown schematically in Fig. 1. MAbs against MSP1, MSP2, and Exp-1 all recognize variant epitopes.

**Indirect immunofluorescence assay.** Indirect immunofluorescence typing of each *P. falciparum* isolate was performed with individual MAbs on separate wells of acetone-fixed schizont slides (7, 29). A working dilution of each MAb (Table 1) was incubated on schizonts for ≥30 min. After careful removal of MAbs by Pasteur pipette, the slides were washed three times (1, 5, and 30 min) in PBS and dried. A 0.20 µl volume of a 1:100 dilution (in PBS, 1% bovine serum albumin, and 0.01% sodium azide) of fluorescein isothiocyanate (FITC)-conjugated rabbit anti-mouse immunoglobulin antibody (ICN Immunobiologicals) was added to each spot, and the slides were incubated for ≥30 min. After two washes (1 and 5 min) in PBS, slides were stained with DAPI (4',6-diamidino-2-phenylindole [Sigma Ltd.]; 10−5 wt/vol) in PBS for 1 min and again washed twice (1 and 5 min in PBS). The slides were dried and mounted under coverslips with Citifluor (Citifluor Ltd., City University, London) or glycerol-PBS (4:1). Parasites were visualized by DNA-specific DAPI fluorescence (incident light of 390 to 440 nm) at a magnification of ×360 or ×600. For each isolate, the percentage of schizonts giving MAb-specific positive FITC fluorescence (incident light of 450 to 490 nm) was recorded for each MAb. Many isolates contained two or more *P. falciparum* clones, as resolved by double-labelled immunofluorescence using different pairs of MAbs (7). In the analyses described below, isolates in which a majority of parasites expressed a given epitope are distinguished from those in which only a minority expressed the epitope.

**RESULTS**

**Scoring of *P. falciparum* isolates for expression of antigen epitopes recognized by MAbs.** At least 200 schizonts from each isolate were scored for reactions with each of 27 MAbs by indirect immunofluorescence. All mature schizonts in every isolate gave specific parasite surface fluorescence with control MAbs 9.8, 12.4, and 12.8 against conserved epitopes of MSP1. Each of the other MAbs identified some isolates within which all schizonts were positive and other isolates within which all schizonts were negative. Such apparently homogeneous isolates were thus scored as being either positive or negative for a given epitope. However, genetically heterogeneous multiple-clone *P. falciparum* infections are common in The Gambia (4, 7), and thus there were isolates within which only a proportion of total schizonts reacted positively with one or more of the MAbs against polymorphic epitopes. The statistical analyses considered...
positive all isolates within which the majority (50 to 100%) of schizonts expressed a given epitope. Although not included within the statistical analyses, data on the percentages of isolates within which more than 1% of schizonts expressed a given epitope are also presented.

**Comparison of polymorphic epitope frequencies during a single transmission season.** Table 2 compares frequencies of different MSP1, MSP2, and Exp-1 epitopes among clinical isolates collected at the beginning (July and August), middle (September and October), and end (November and December) of the 1988 transmission season. The MSP2 epitope 12.3 occurred less frequently (and hence the alternative epitope 8G10/48 more frequently) during July and August than during September and October or November and December (chi-square, \( P < 0.05 \) for each comparison). There were no significant differences in frequencies of any other epitopes between the periods.

**Comparison of polymorphic epitope frequencies during the period 1982 to 1989.** To investigate whether slow antigenic drift may occur over a period of several years, frequencies of polymorphic epitopes were recorded between 1982 and 1989. Figure 2 compares epitope frequencies among *P. falciparum* isolates collected in the same area as above during the October to December periods of 1982 (10 = 46), 1983 (n = 31), 1988 (n = 101), and 1989 (n = 119). Only two statistically significant frequency changes were recorded. The epitope 12.2 on the repeat domain of MSP1 was more frequent in 1983 than in each other year (chi-square, \( P < 0.05 \) for each comparison), and the epitope 12.1 on domain 4 of MSP1 was more frequent in 1982 than in each other year (chi-square, \( P < 0.05 \) for each comparison).

Data presented in Table 2 and Fig. 2 illustrate a very striking difference between the frequencies of the two alternative major serological classes of MSP1, here identified by epitopes on domains 6 to 16 (Fig. 1 and Table 1). The frequency of the “MAD20-like” class (identified by epitopes 9.2, 9.7, and 10.3, with 9.2 shown as an example in Table 2 and Fig. 2) was high, >95%. In contrast, the frequency of the “K1/Wellcome-like” class (identified by epitopes 6.1, 7.3, 13.1, 17.1, 1-1C, and 34-5, with 6.1 shown as an example in Table 2 and Fig. 2) was less than 5%. This difference remained a stable characteristic of the parasite population over the period of the study.

**DISCUSSION**

The main finding of the study was that the relative frequencies of MSP1, MSP2, and Exp-1 variant epitopes among *P. falciparum* isolates from Gambian patients remained stable during the period 1982 to 1989. Only the MSP1 epitopes 12.1 and 12.2 had significantly higher frequencies in 1982 and 1983, respectively, although these differences are no more than expected by chance (because of the large total number of chi-square comparisons made). The frequencies were also stable throughout a single transmission season (with only two significant differences among all the chi-square comparisons, as would again have been expected by chance).

Therefore, the more common variants remained at a higher frequency over several years, and the rare variants remained at a lower frequency. This was particularly notable with respect to the two alternative types of MSP1 domains. This “dimorphic” region of MSP1 extends over most of the sequence (32, 39) and is represented in the sequence divergence between laboratory isolates MAD20 and K1 Wellcome (39). All MAbs recognizing this dimorphic region have one of two alternative specificities and are useful...
FIG. 2. Percentage of isolates containing parasites expressing different polymorphic epitopes of MSP1, MSP2, and Exp-1 from October through December 1982 (n = 46), 1983 (n = 31), 1988 (n = 101), and 1989 (n = 119). Shaded bars indicate the proportion of isolates in which 50 to 100% of parasites expressed a given epitope, and open bars show the proportion of isolates in which 1 to 100% of parasites had each epitope. Asterisks mark the two values which were higher in one year (chi-square, \( P < 0.05 \)). Omitted from the figure are MSP1 epitopes 7.3, 13.1, 17.1, 1-1C, and 34-5, which showed parasite distribution identical to that of epitope 6.1 (epitopes 1-1C and 34-5 were tested for only in 1988 and 1989 isolates); MSP1 epitopes 9.7 and 10.3, which showed specificity identical to that of the epitope 9.2; and MSP2 epitope 12.5, with specificity identical to that of 12.3.

markers of the dimorphism even though naturally acquired human antibodies may be directed against different epitopes. The K1/Wellcome-like type (here distinguished by epitopes 6.1, 7.3, 13.1, 17.1, 1-1C, and 34-5) remained at a very low frequency (<5%), while the MAD20-like type (distinguished by epitopes 9.2, 9.7, and 10.3) maintained a high frequency (>95%). During the intervening years 1984 to 1987, the population was not systematically sampled. However, in nine isolates available from 1986, all parasites had the common MAD20-like dimorphic type, confirming that there
was no significant frequency shift during this time (unpublished data; isolates supplied by M. J. Blackman).

Theoretically, a frequency-dependent selection can maintain genetic polymorphisms with either stable or fluctuating variant frequencies, depending on whether equilibrium frequencies have been or are being reached (6, 17). If acquired variant-specific immunity operated on the dimorphic regions of MSP1, it is difficult to explain how the two alternative frequencies have been or are being reached (6, 17). If acquired variant-specific immunity are therefore required to explain the sequence dimorphism at domains 6 to 16 in MSP1. Since MSP1 may be a parasite receptor involved in the recognition of the host erythrocyte (30), one such mechanism may be a corresponding dimorphism in a ligand of the erythrocyte surface in the human population.

However, more extensive polymorphisms exist in domains of tandem repeats in MSP1 (32, 39) and MSP2 (10, 37, 40), and these might be subject to immune selection. In a hyperendemic area, such as The Gambia, it may be that equilibrium frequencies of such variants have been reached, which could explain the stable frequencies of, e.g., epitopes 12.2 and 3D3 on MSP1 and 13.4 on MSP2. Alternatively, the polymorphisms may be neutral with respect to immunity, and the lack of change in these frequencies may simply result from an absence of selection.

Temporal variations in the prevalence of an S-antigen epitope were recorded in villages of Papua New Guinea (11). It was argued that these variations could possibly result from frequency-dependent immune selection (12). However, as pointed out elsewhere (8), random fluctuations in allelic frequencies are likely to occur in P. falciparum populations in villages because of genetic drift in small populations, so selective mechanisms are not required to explain the data.

In the present study, the frequencies of each epitope have been analyzed separately, independently of one another. However, different epitopes on the same protein are likely to be strongly associated genetically or structurally. For example, even though intragenic recombinations occur in the MSP1 gene (9, 32, 39), strong associations between variant epitopes located on different domains of the protein exist in local populations (9). Therefore, a consideration of allelic serotype frequencies, in addition to individual variant epitope frequencies, is necessary. In The Gambia, MSP1 and MSP2 allelic serotype frequencies in 1988 and 1989 were resolved, and they remained similar in both years (8).

In conclusion, the present results suggest that some major antigen polymorphisms, particularly the dimorphism which includes most of the sequence of MSP1, are not maintained by frequency-dependent immune selection. It is not excluded that other polymorphisms in MSP1 and MSP2 might be balanced by effects of frequency-dependent immune selection, although immunological studies are now required to demonstrate such variant-specific immune responses.

ACKNOWLEDGMENTS

We thank the patients for cooperation in the study, and we also thank Dominic Kwiatkowski, Adrian Hill, and all the staff at the MRC laboratories in The Gambia for clinical and technical support. Richard Pink, Anthony Holder, Jeffery Lyon, Alan Saul, and Robert Reese kindly provided some of the MAbs.

This investigation received financial support from the MRC, the Wellcome Trust, and the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases.

REFERENCES


