A principal target of human immunity to malaria identified by molecular population genetic and immunological analyses

DAVID J. CONWAY[1], DAVID R. CAVANAGH[2], KAZUYUKI TANABE[3], CALLY ROPER[1], ZSUZSANNA S. MIKES[1], NAOKO SAKIHAMA[3], KALIFA A. BOJANG [4], AYOADE M. J. ODUOLA[5], PETER G. KREMSNER[6], DAVID E. ARNOT[2], BRIAN M. GREENWOOD[1] & JANA S. MCBRIDE[2]

[1]*Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel St, London WC1E 7HT*
[2]*Institute of Cell, Animal, and Population Biology, University of Edinburgh, EH9 3JT*
[3]*Laboratory of Biology, Osaka Institute of Technology, Osaka 535-8535, Japan*
[4] *Medical Research Council Laboratories, Fajara, PO Box 273, Banjul, The Gambia*
[5]*Postgraduate Institute of Medical Research and Training, College of Medicine, University of Ibadan, Ibadan, Nigeria*
[6]*Department of Parasitology, Institute of Tropical Medicine, University of Tubingen, Germany*
Correspondence should be addressed to D.J.C.; email: david.conway@lshtm.ac.uk

New strategies are required to identify the most important targets of protective immunity in complex eukaryotic pathogens. Natural selection maintains allelic variation in some antigens of the malaria parasite *Plasmodium falciparum*[1–3]. Analysis of allele frequency distributions could identify the loci under most intense selection[4–7]. The merozoite surface protein 1 (Msp1) is the most-abundant surface component on the erythrocyte-invading stage of *P. falciparum*[8–10]. Immunization with whole Msp1 has protected monkeys completely against homologous[11] and partially against non-homologous[12] parasite strains. The single-copy *msp1* gene, of about 5 kilobases, has highly divergent alleles[13] with stable frequencies in endemic populations[14,15]. To identify the region of *msp1* under strongest selection to maintain alleles within populations, we studied multiple intragenic sequence loci in populations in different regions of Africa and Southeast Asia. On both continents, the locus with the lowest inter-population variance in allele frequencies was *block 2*, indicating selection in this part of the gene. To test the hypothesis of immune selection, we undertook a large prospective longitudinal cohort study. This demonstrated that serum IgG antibodies against each of the two most frequent allelic types of block 2 of the protein were strongly associated with protection from *P. falciparum* malaria.

We first determined allele frequency distributions throughout the merozoite surface protein 1 (*msp1*) gene by genotyping ten loci spanning the gene in 547 *Plasmodium falciparum* isolates from six populations in Africa (located in The Gambia, Nigeria, Gabon, Sudan, Tanzania and South Africa), using PCR amplification followed by allele sequence-specific oligonucleotide probing[16]. We determined the allele frequencies at each of the *msp1* intragenic loci in each population, together with the inter-population component of variance in allele frequencies (Wright's F_{ST}) (ref. 17) (Table 1). As recombination was frequent (linkage disequilibrium is generally less than half its maximal value over distances of more than 0.3 kilobases in most of the populations)[16], the loci had mostly independent allele frequency distributions. We sought the locus with the lowest F_{ST} value (that is, with the lowest inter-population divergence and the highest proportion of overall allelic diversity maintained within each population) as that under strongest selection to maintain alleles[6,7]. Nine of the loci had F_{ST} values that differed significantly from zero, indicating heterogeneity in allele frequencies among populations. In contrast, the *block 2* locus showed a non-significant value near zero (F_{ST} = 0.008), so almost all of the diversity at this locus existed within each population.

To determine the typical range of genetic variance among these *P. falciparum* populations at non-antigen loci for which there is no hypothesis of selection, we genotyped nine microsatellite tri-nucleotide TAA(n) repeat loci (*ta1*, *ta42*, *ta81*, *ta87*, *ta109*, *PfPK2*, *G377*, *Polya* and *ARA2*)[18,19] each in more than 50 isolates from each of the six populations. We discriminated alleles by size of products amplified by PCR (ref.19), with an observed number of distinct alleles per locus ranging from 7 (locus *G377*) to 22 (locus *ta1*). Overall, the microsatellite loci showed moderate inter-population variance in allele frequencies (mean F_{ST} value = 0.026), higher than that for *block 2* of *msp1*. For all except one (*G377*) of these nine loci individually, the F_{ST} values exceeded that of *block 2* and were highly significantly above zero ($P < 0.01$ for locus *ARA2*; $P < 0.001$ for each of the other seven loci). Thus, the *block 2* locus in *msp1* has a lower F_{ST} value than most putatively neutral loci.

In a separate study, we also determined whether *block 2* had an unusually low inter-population variance in allele frequencies in another continent. We genotyped *block 2* and three other polymorphic loci in the *msp1* gene in an overall sample of 655 alleles from two populations in Southeast Asia (located in Thailand and Vietnam), using a method of nested PCR with allele sequence-specific internal primers[15,20]. We determined the allele frequencies in each population, with F_{ST} indices (Table 2). Whereas the other loci have F_{ST} values significantly different from zero, the value for *block 2* was again non-significant and almost zero. Thus, results of African and Southeast Asian studies indicate that *block 2* is the region of the *msp1* gene under the strongest selection to maintain alleles in both continents.

Block 2 sequences of *msp1* belong to one or another of three highly divergent allelic types: *K1-like, MAD-like* and *RO33-like*. The *K1-like* and *MAD-like* types encode protein segments in the size ranges of 47–89 amino acids or 35–59 amino acids, respectively (each type incorporates many different subtypes that dif-

fer in sequence and length of tri-peptide repeats in the middle of the block, flanked by non-repetitive sequences that are highly conserved within each type but that differ between types). A mouse monoclonal antibody against the K1-like repeats inhibits *P. falciparum* growth *in vitro*[21]. The *RO33-like* type encodes a sequence of 46 amino acids without repeats and with little subtypic variation[22]. Recombinant proteins have been expressed in *Escherichia coli*, representing examples of subtypes from each of the three main allelic types of *block 2* (refs. 23,24). Type-specific epitopes on these block 2 proteins are commonly recognized by antibodies induced during natural *P. falciparum* infections, and antibodies against subtype-specific epitopes are detected occasionally[24], but no naturally acquired human antibodies cross-react among the three types of block 2 (refs. 23,24).

Table 1 Allele frequencies in six African populations and F_{ST} indices of inter-population variance at ten polymorphic loci within the *msp1* gene of *P. falciparum*

Locus	Allele	Population allele frequencies: Gambia (*n* = 91)	Nigeria (*n* = 107)	Gabon (*n* = 124)	Sudan (*n* = 66)	Tanzania (*n* = 86)	S.Africa (*n* = 73)	F_{ST}
aa44	*S*	0.88	0.79	0.60	0.68	0.42	0.32	0.170**
	G	0.12	0.21	0.40	0.32	0.58	0.68	
block 2	*MAD-like*	0.33	0.29	0.37	0.33	0.26	0.27	0.008
	K1-like	0.51	0.52	0.54	0.36	0.53	0.59	
	RO33-like	0.16	0.19	0.09	0.30	0.21	0.14	
aa160	*Q*	0.29	0.38	0.27	0.48	0.31	0.21	0.024*
	R	0.71	0.62	0.73	0.52	0.69	0.79	
aa222	*A*	0.22	0.22	0.12	0.00	0.02	0.01	0.086**
	V	0.78	0.78	0.88	1.00	0.98	0.99	
aa297	*D*	0.06	0.06	0.09	0.26	0.09	0.08	0.041**
	T	0.94	0.94	0.91	0.74	0.91	0.92	
aa320	*E*	0.73	0.57	0.68	0.44	0.76	0.64	0.040**
	Q	0.27	0.43	0.32	0.56	0.24	0.36	
block 4b	*MAD-like*	0.24	0.30	0.24	0.49	0.14	0.15	0.058**
	K1-like	0.76	0.70	0.76	0.51	0.86	0.85	
block 6-16	*MAD-like*	0.97	0.94	0.99	0.76	0.88	0.93	0.076**
	K1-like	0.03	0.06	0.01	0.24	0.12	0.07	
aa1644	*E*	0.42	0.43	0.50	0.13	0.42	0.24	0.064**
	Q	0.58	0.57	0.50	0.87	0.58	0.76	
aa1700	*S*	0.12	0.17	0.11	0.02	0.18	0.01	0.035**
	N	0.88	0.83	0.89	0.98	0.82	0.99	

Loci are indicated by codon (aa; amino-acid) position in a sequence alignment [22] or by block number (following a 17-block scheme of the gene)[13]. *, $P < 0.01$, and **, $P < 0.001$, compared with $F_{ST} = 0$.

To determine whether type-specific human antibodies against block 2 of Msp1 are protective against *P. falciparum* malaria, as the population genetic study would predict, we undertook a prospective longitudinal cohort study in The Gambia. In June 1996, immediately before the annual malaria transmission season, we collected sera from 337 children (3–7 years of age), to be tested later for antibody reactivity with recombinant antigens representing different allelic forms of block 2, or other blocks of the Msp1 protein. We monitored these children for five months (July to November), with active and passive case detection of malaria.

We determined the proportions of individuals who experienced clinical malaria at any time during the five months, among those with or without detectable serum IgG levels against particular Msp1 recombinant antigens at the start of the study (Table 3). Consistent with the deduced hypothesis, the presence of serum IgG against either of the two most common block 2 types, K1-like and MAD-like (with allele frequencies in The Gambia of 0.51 and 0.33, respectively; Table 1) was highly significantly and independently associated with a lower prospective risk of malaria. The relative risks (0.64–0.67 for IgG against the K1-like type, and 0.54–0.78 for IgG against the MAD-like type) gave protective efficacy estimates (1–relative risk) in the range of 22–46%. These effects remained statistically significant for five of the seven K1-like or MAD-like subtypes tested, after controlling for individuals' ages in years. Logistic regression analysis of relative antibody levels (absorbance levels) gave similar results to those for antibody positivity against antigens representing the K1-like and MAD-like types of block 2; that is, higher absorbance values were significantly associated with protection from clinical malaria ($P <$ 0.05 for each of the seven recombinant antigens representing these block 2 types).

The highly significant protective efficacy estimates of 22–46% for antibodies against either of the common K1-like or MAD-like block 2 types were not dissimilar to the maximum possible for type-specific immunity (given that the frequencies of these main allelic types of *block 2* in the Gambia were 0.51 and 0.33). If block 2 type-specific antibodies protect in a parasite allele-specific manner, their effects should be additive, and individuals with antibodies against both of the common block 2 types (K1-like and MAD-like) should be very well protected. This was indeed the case. Only 7 (21.2%) of the 33 individuals who had antibodies against representative antigens of both types (3D7 antigen of the K1-like type, and Wellcome antigen of the MAD-like type) subsequently had malaria, com-

Table 2 Allele frequencies in two Southeast Asian populations and F_{ST} indices of inter-population variance at four polymorphic loci within the *msp1* gene of *P. falciparum*

Locus	Allele	Population allele frequencies: Thailand (*n* = 287)	Vietnam (*n* = 368)	F_{ST}
block 2	*MAD-like*	0.65	0.64	0.000
	K1-like	0.20	0.23	
	RO33-like	0.15	0.13	
block 4a	*MAD-like*	0.51	0.41	0.016*
	K1-like	0.49	0.59	
block 4b	MAD-like	0.70	0.56	0.041**
	K1-like	0.30	0.44	
block 6-16	MAD-like	0.61	0.82	0.104**
	K1-like	0.39	0.18	

Loci are indicated by block number, following a 17-block scheme of the gene[13]. *, $P < 0.05$, and **, $P < 0.001$, compared with $F_{ST} = 0$.

pared with 145 (60.9%) of the 238 individuals who did not have antibodies against either (a relative risk of 0.35, with 95% confidence interval of 0.18–0.68). This gave an estimated protective efficacy of 65%, which is extremely significant (P = 0.0004 after correcting for individuals' ages in years; P = 0.00002 without correction).

Antibodies against the rarest of the block 2 types (RO33-like, with an allele frequency of 0.16) were not significantly associated with protection (relative risk, 0.94). This weaker non-significant effect is expected from its low allele frequency, even if such antibodies do offer type-specific protection. The protective effect of type-specific antibodies cannot exceed the allele frequency of that antigenic type of *P. falciparum* in the study population, so investigating effects of antibodies against this type would require a longitudinal cohort study in a population with a higher *RO33-like* allele frequency. There was no significant association between antibodies against block 1 and protection from malaria. Block 1 is very rarely recognized by naturally acquired antibodies in this as in another population[24], and immunization with a synthetic peptide vaccine containing part of this sequence failed to protect against malaria [25]. Antibody positivity against block 17 (that is, the C-terminal $Msp1_{19}$ fragment)[10] was also not associated with protection. However, antibody levels (absorbance values) against the $Msp1_{19}$ fragment were significantly associated with protection from malaria (P = 0.028 after correction for age by logistic regression). This indicates that differences in levels of antibody against $Msp1_{19}$ considerably higher than the cut-off value are important. A few other studies have indicated that serum IgG against $Msp1_{19}$ is associated with protection from malaria[26–28], but another recent study failed to show an association[29]. Detailed structural[30] and functional analyses[31] independently support the vaccine candidacy of $Msp1_{19}$.

In conclusion, we have predicted here which polymorphic region of a malaria parasite antigen is most likely to be a target of acquired immunity in endemic populations, and then showed that antibody responses to this region are strongly associated with protection from malaria. This demonstrates that molecular population genetic analysis can identify regions of pathogen genes under selection, and is an effective approach to discover targets of protective immunity to endemic pathogens. Other population genetic approaches to identifying malaria parasite loci under immune selection have been suggested, but not yet tested[32,33]. We have identified block 2 of Msp1 as a principal candidate for a *P. falciparum* vaccine. The vaccine should incorporate the main allelic types of block 2 to induce broad, though essentially type-specific, immunity against merozoites. Studies are underway to identify epitopes on the different block 2 types recognized by human antibodies, and to develop recombinant proteins suitable for testing the efficacy of vaccination in primate models and human volunteers.

Table 3 Proportions of Gambian children acquiring clinical malaria (July–November) in groups with or without serum IgG against the indicated regions of Msp1 in June

Percentages and numbers acquiring clinical malaria of the cohort of 337 children grouped according to IgG reactivities in June against: Msp1 recombinant antigen	IgG positive	IgG negative	RR (95% CI)[a] ¶	*P* value
K1-like Block 2				
3D7	37.8% (28/74)	59.3% (156/263)	0.64 (0.47–0.87)	0.0010***
Palo Alto	40.0% (34/85)	59.9% (151/252)	0.67 (0.51–0.89)	0.0018**
DWK1#1	37.3% (22/59)	58.3% (162/278)	0.64 (0.45–0.90)	0.0033**
DWK1#2	37.3% (19/51)	57.7% (165/286)	0.65 (0.45–0.93)	0.0069**
MAD-like Block 2				
Wellcome	32.7% (19/58)	60.2% (168/279)	0.54 (0.36–0.79)	0.0001***
MAD20	45.1% (37/82)	57.6% (147/255)	0.78 (0.60–1.02)	0.048*
DWMAD20	34.7% (17/49)	58.0% (167/288)	0.60 (0.40–0.89)	0.0025**
RO33-like Block 2				
RO33	51.9% (27/52)	55.1% (157/285)	0.94 (0.71–1.25)	0.67
DWRO33	52.2% (48/92)	55.5% (136/245)	0.94 (0.75–1.18)	0.58
Block 1	54.5% (6/11)	54.6% (178/326)	1.00 (0.58–1.73)	1.0
Block 17	51.7% (109/211)	59.5% (75/126)	0.87 (0.71–1.05)	0.16

*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$: All of these protective associations except two (for DWK1#2 and MAD20) remain statistically significant after adjusting for individuals' ages in years by multiple logistic regression analysis. [a]Relative risks (with 95% confidence intervals).

Methods

***P. falciparum* populations and DNA samples.** Human subjects with *P. falciparum* infections were identified in studies on the epidemiology of malaria in six geographical locations in Africa and in two locations in Southeast Asia. The African *P. falciparum* isolates were those previously collected[16] from Basse in Upper River District, The Gambia, 1995 (n (number of isolates) = 91); Ibadan in Oyo state, south-western Nigeria, 1996 (n = 107); Lambarene District in Gabon, 1996 (n = 124); Daraweesh in Gedaref state, eastern Sudan, 1992-1995 (n = 66); Muheza District in Tanzania, 1995 (n = 86); and KwaZulu-Natal in South Africa, 1996 (n = 73). The Southeast Asian isolates were collected from Mae Sot in Thailand, 1996 (n = 184) (ref. 20); and Bao Loc in Vietnam, 1994 and 1996 (n = 210) (ref. 15). Peripheral blood samples were collected with permission and under guidelines of the relevant government and institutional ethical committees, and DNA was isolated using standard proteinase K digestion and phenol:chloroform extraction methods.

Genotyping of allelic sequences in *msp1*. Ten intragenic loci in the *msp1* gene were genotyped in the African samples, using PCR and sequence-specific oligonucleotide probing as described[16]. Natural *P. falciparum* infections commonly have multiple clones and may therefore contain more than one haploid parasite genotype[34], so for each locus the single or the most abundant allele in each isolate (that giving the strongest hybridization signal) was included in a sample of allele frequency in the local population. On the rare occasions on which different alleles were apparently equally abundant in an isolate, a random number table was used to determine which allele was to be sampled statistically. Thus, the sample size of alleles was equal to the number of isolates (n = 547). Four intragenic loci in the *msp1* gene were genotyped in the Southeast Asian samples, using PCR with sequence-specific internal primers[15,20,35]. In accordance with the previous use of this method, every allele found in an isolate was included in overall frequencies in each population, so the number of alleles sampled (n = 655) in Southeast Asia was greater than the number of isolates collected.

Genotyping of microsatellites. Nine *P. falciparum* microsatellite loci were genotyped (*ta1*, *ta42*, *ta81*, *ta87*, *ta109*, *PfPK2*, *G377*, *Polya* and *ARA2*)[18] using a semi-nested PCR method[19], with some modifications in the choice of fluorescent dye used to label the internal primer for some loci. Alleles were identified by size after electrophoresis on an ABI 377 using GENESCAN and GENOTYPER software (Applied Biosystems, Warrington, Cheshire, UK). For each locus in each isolate, a single allele (that is, the

only allele detected, or the most abundant allele in isolates containing more than one genotype of *P. falciparum*) was counted for statistical analysis of allele frequencies.

Immuno-epidemiology study cohort. Children (n = 337) 3–7 years old (42% female), living in a rural community near Basse in the Upper River Division of The Gambia (with intense, very seasonal malaria transmission), were recruited for a study of immune responses and surveillance of malaria incidence. After informed consent was obtained, 10 ml venous blood was collected from each child at the end of June 1996, before the malaria transmission season. The children were then actively monitored throughout the transmission season (July–November) for incidence of clinical malaria. The main measured outcome of clinical malaria in this population was a parasitemia of more than 5,000 *P. falciparum* parasites per µl blood, coincident with an axillary temperature of more than 37.5 °C.

Recombinant Msp1 antigens and antibody assays. Block 1 (MAD20 sequence), block 2 (3D7, Palo Alto, DW K1#1, DWK1#2, MAD20, Wellcome, DWMAD20, RO33 and DWRO33) and block 17 ($Msp1_{19}$ Wellcome sequence) proteins were expressed in *E. coli* as recombinant proteins fused to the C terminus of glutathione S-transferase of *Schistosoma japonicum* using pGEX vectors, as described[23,24,36,37]. Human sera were tested by enzyme-linked immunosorbent assay (ELISA) at a dilution of 1:500 for the presence of IgG antibodies recognizing these recombinant proteins, as described[23,24]. Cut-off values above which binding of antibody was regarded as positive were calculated as the mean + 3 s.d. of readings obtained with negative control sera from 15 healthy Scottish blood donors.

Statistical analyses. Encoded genotypic data were imported into EXCEL and SPSS 8.0 for analysis and tabulation of allele frequencies in each of the geographical populations. For each locus, the inter-population variance in allele frequencies, F_{ST}, among the six populations in Africa or between the two populations in Southeast Asia, was calculated using Weir & Cockerham's θ estimator[17], with the FSTAT program (version 1.2) . Antibody ELISA results were imported into an EXCEL worksheet by investigators 'blinded' to the epidemiological data (kept separately on a STATA file). Data files were merged, and were analyzed independently by two investigators using EPI-INFO version 6 and SPSS 8.0 for univariate and multivariate (logistic regression) analyses of the predictive effects of specific antibodies on malaria incidence.

Acknowledgments

We thank O. Ogundahunsi, M. Grobusch, K. Hirayama, K. Na-Bangchang, C. Curtis, J. Trigg, J. Curtis, G. Warnasuriya, and D. Warhurst for support with sample collection and processing; H. Weiss for help with the cohort study database; T. Anderson for advice on microsatellites; and R. Gregory for technical assistance with the ABI 377. This work was supported by The Wellcome Trust (grant references 013163/Z/94/B,C and 0471191/Z/96), the UK Medical Research Council (grant reference G9803180), and The Ministry of Education, Science and Culture, Japan.

RECEIVED 9 SEPTEMBER 1999; ACCEPTED 10 APRIL 2000

1. Hughes, M. K. & Hughes, A. L. Natural selection on *Plasmodium* surface proteins. *Mol. Biochem. Parasitol.* **71**, 99–113 (1995).
2. Conway, D.J. Natural selection on polymorphic malaria antigens and the search for a vaccine. *Parasitol. Today* **13**, 26–29 (1997).
3. Escalante, A.A., Lal, A.A. & Ayala, F.J. Genetic polymorphism and natural selection in the malaria parasite *Plasmodium falciparum*. *Genetics* **149**, 189–202 (1998).
4. Berry, A. & Kreitman, M. Molecular analysis of an allozyme cline: Alcohol dehydrogenase in *Drosophila melanogaster* on the east coast of North America. *Genetics* **134**, 869–893 (1993).
5. Taylor, M.F.J., Shen, Y. & Kreitman, M.E.A population genetic test of selection at the molecular level. *Science* **270**, 1497–1499 (1995).
6. McDonald, J.H. in *Non-Neutral Evolution: Theories and Molecular Data.* (ed. Golding, B.) 88-100 (Chapman–Hall, London, UK, 1994).
7. Beaumont, M.A. & Nichols, R.A. Evaluating loci for use in the genetic analysis of population structure. *Proc. R. Soc. Lond. B* **263**, 1619–1626 (1996).
8. Holder, A.A. *et al.* Primary structure of the precursor to the three major surface antigens of *Plasmodium falciparum* merozoites. *Nature* **317**, 270–273 (1985).
9. McBride, J.S. & Heidrich, H.-G. Fragments of the polymorphic Mr 185 000 glycoprotein from the surface of isolated *Plasmodium falciparum* merozoites form an antigenic complex. *Mol. Biochem. Parasitol.* **23**, 71–84 (1987).
10. Blackman, M.J., Heidrich, H.-G., Donachie, S., McBride, J.S. & Holder, A.A. A single fragment of a malaria merozoite surface protein remains on the parasite during red cell invasion and is the target of invasion-inhibiting antibodies. *J. Exp. Med.* **172**, 379–382 (1990).
11. Siddiqui, W.A. *et al.* Merozoite surface coat precursor protein completely protects *Aotus* monkeys against *Plasmodium falciparum* malaria. *Proc. Natl. Acad. Sci., USA* **84**, 3014–3018 (1987).
12. Hall, R. *et al.* Major surface antigen gene of a human malaria parasite cloned and expressed in bacteria. *Nature* **311**, 379–382 (1984).
13. Tanabe, K., Mackay, M., Goman, M. & Scaife, J.G. Allelic dimorphism in a surface antigen gene of the malaria parasite *Plasmodium falciparum*. *J. Mol. Biol.* **195**, 273–287 (1987).
14. Conway, D.J., Greenwood, B.M. & McBride, J.S. Longitudinal study of *Plasmodium falciparum* polymorphic antigens in a malaria endemic population. *Infect. Immun.* **60**, 1122–1127 (1992).
15. Ferreira, M.U. *et al.* Stable patterns of allelic diversity at the merozoite surface protein-1 locus of *Plasmodium falciparum* in clinical isolates from southern Vietnam. *J. Euk. Microbiol.* **45**, 131–136 (1998).
16. Conway, D.J. *et al.* High recombination rate in natural populations of *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. USA* **96**, 4506–4511 (1999).
17. Weir, B.S. & Cockerham, C.C. Estimating F statistics for the analysis of population structure. *Evolution* **38**, 1358–1370 (1984).
18. Su, X.-Z. & Wellems, T.E. Towards a high resolution *Plasmodium falciparum* linkage map: Polymorphic markers from hundreds of simple sequence repeats. *Genomics* **33**, 430–444 (1996).
19. Anderson, T.J.C., Su, X.-Z., Bockaire, M., Lagog, M. & Day, K.P. Twelve microsatellite markers for characterisation of *Plasmodium falciparum* from finger prick blood samples. *Parasitology* **119**, 113–125 (1999).
20. Sakihama, N. *et al.* Allelic recombination and linkage disequilibrium within *Msp-1* of *Plasmodium falciparum*, the malignant human malaria parasite. *Gene* **230**, 47–54 (1999).
21. Locher, C.P., Tam, L.Q., Chang, S.P., McBride, J.S. & Siddiqui, W.A. *Plasmodium falciparum*: gp195 tripeptide repeat-specific monoclonal antibody inhibits parasite growth *in vitro*. *Exp. Parasitol.* **84**, 74–83 (1996).
22. Miller, L.H., Roberts, T., Shahabuddin, M. & McCutchan, T.F. Analysis of sequence diversity in the *Plasmodium falciparum* merozoite surface protein-1 (MSP-1). *Mol. Biochem. Parasitol.* **59**, 1–14 (1993).
23. Cavanagh, D.R. & McBride, J.S. Antigenicity of recombinant proteins derived from *Plasmodium falciparum* merozoite surface protein 1. *Mol. Biochem. Parasitol.* **85**, 197–211 (1997).
24. Cavanagh, D.R. *et al.* A longitudinal study of type-specific antibody responses to *Plasmodium falciparum* merozoite surface protein-1 in an area of unstable malaria in Sudan. *J. Immunol.* **161**, 347–359 (1998).
25. D'Alessandro, U. *et al.* Efficacy trial of the malaria vaccine SPf66 in Gambian infants. *Lancet* **346**, 462–467 (1995).
26. Riley, E.M. *et al.* Naturally acquired cellular and humoral immune responses to the major merozoite surface antigen (PfMSP1) of *Plasmodium falciparum* are associated with reduced malaria morbidity. *Parasite Immunol.* **14**, 321–337 (1992).
27. Al-Yaman, F. *et al.* Assessment of the role of naturally acquired antibody levels to *Plasmodium falciparum* merozoite surface protein-1 in protecting Papua New Guinean children from malaria mortality. *Am. J. Trop. Med. Hyg.* **54**, 443–448 (1996).
28. Egan, A.F. *et al.* Clinical immunity to *Plasmodium falciparum* malaria is associated with serum antibodies to the 19kDa C-terminal fragment of the merozoite surface antigen, PfMSP1. *J. Infect. Dis.* **173**, 765–769 (1996).
29. Dodoo, D. *et al.* Levels of antibody to conserved parts of *Plasmodium falciparum* merozoite surface protein 1 in Ghanaian children are not associated with protection from clinical malaria. *Infect. Immun.* **67**, 2131–2137 (1999).
30. Morgan, W.D. *et al.* Solution structure of an EGF module pair from the *Plasmodium falciparum* merozoite surface protein 1. *J. Mol. Biol.* **289**, 113–122 (1999).
31. Patino, J.A., Holder, A.A., McBride, J.S. & Blackman, M.J. Antibodies that inhibit malaria merozoite surface protein-1 processing and erythrocyte invasion are blocked by naturally acquired human antibodies. *J. Exp. Med.* **186**, 1689–1699 (1997).
32. Gupta, S. *et al.* The maintenance of strain structure in populations of recombining infectious agents. *Nature Med.* **2**, 437–442 (1996).
33. Hastings, I. Population genetics and the detection of immunogenic and drug-resistant loci in *Plasmodium*. *Parasitol.* **112**, 155–164 (1996).
34. Conway, D.J., Greenwood, B.M. & McBride, J.S. The epidemiology of multiple-clone *Plasmodium falciparum* infections in Gambian patients. *Parasitology* **103**, 1–6 (1991).
35. Kaneko, O., Kimura, M., Kawamoto, F., Ferreira, M.U. & Tanabe, K. *Plasmodium falciparum*: Allelic variation in the merozoite surface protein 1 gene in wild isolates from southern Vietnam. *Exp. Parasitol.* **86**, 45–57 (1997).
36. Burghaus, P.A. & Holder, A.A. Expression of the 19-kilodalton carboxy-terminal fragment of the *Plasmodium falciparum* merozoite surface protein-1 in *Escherichia coli* as a correctly folded protein. *Mol. Biochem. Parasitol.* **64**, 165–169 (1994).
37. Egan, A.F. *et al.* Serum antibodies from malaria-exposed people recognize conserved epitopes formed by the two epidermal growth factor motifs of MSP1(19), the carboxy-terminal fragment of the major merozoite surface protein of *Plasmodium falciparum*. *Infect. Immun.* **63**, 456–466 (1995).