



Published in final edited form as:

Nat Genet. 2014 November ; 46(11): 1197–1204. doi:10.1038/ng.3107.

Reappraisal of known malaria resistance loci in a large multi-centre study

A full list of authors and affiliations appears at the end of the article.

Abstract

Many human genetic associations with resistance to malaria have been reported but few have been reliably replicated. We collected data on 11,890 cases of severe malaria due to *Plasmodium falciparum* and 17,441 controls from 12 locations in Africa, Asia and Oceania. There was strong evidence of association with the *HBB*, *ABO*, *ATP2B4*, *G6PD* and *CD40LG* loci but previously reported associations at 22 other loci did not replicate in the multi-centre analysis. The large sample size made it possible to identify authentic genetic effects that are heterogeneous across populations or phenotypes, a striking example being the main African form of *G6PD* deficiency, which reduced the risk of cerebral malaria but increased the risk of severe malarial anaemia. The finding that *G6PD* deficiency has opposing effects on different fatal complications of *P. falciparum* infection indicates that the evolutionary origins of this common human genetic disorder are more complex than previously supposed.

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

*Corresponding Authors, kirk.rockett@well.ox.ac.uk and dominic.kwiatkowski@well.ox.ac.uk. Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK.

*A full list of authors appears at the end of this article. Full contribution is listed in Supplementary Note.

AUTHOR CONTRIBUTIONS.

All senior authors from each partner site have approved the final manuscript, their site authorship, acknowledgements and full contributors list.

Writing Group:

GMC, DPK, SM, KAR, CCAS

Project Management:

EA, TA, SA, AA, OA, KAB, DJC, VC, TMED, OD, CD, SJD, JF, TTH, KJJ, HK, AK, SK, KAK, DPK, KM, PM, DM, MM, IM, AN, NP, MP, BP, HR, ER, KAR, PS, SBS, GS, SS, TT, MAT, TNW, MDW

Sample Clinical Data Collection and Management:

SA, AA, LNA, OA, TA, KAB, ECB, GMC, DJC, SJD, AE, JE, KF, AG, LH, MJ, DK, HK, AK, SK, ML, AM, VDM, AM, LM, PM, SM, RM, AN, CMEN, AN, VN, SO, NP, NHP, MP, BP, NTNQ, HR, KAR, MS, GS, FS, SS, TT, CQT, MAT, OT, SU, SU, AV

Sample Processing Genotyping and Management:

AA, LNA, OA, TA, ECB, RC, AE, AG, AG, LH, CH, MJ, AEJ, DK, HK, ML, AM, VDM, AM, LM, SM, RM, AN, CMEN, AN, VN, SO, NHP, NTNQ, KAR, KR, FS, CQT, OT, SU, AV

Analysis:

GB, TGC, GMC, SQL, SM, MP, KAR, NS, CCAS

COMPETING FINANCIAL INTERESTS.

The authors declare no competing financial interests.

URLs.

MalariaGEN Partner Sites: <http://www.malariagen.net/projects/cp1>; MalariaGEN case-report form: <http://www.malariagen.net/resource/1>; The R-project: <http://www.r-project.org>;

Introduction

It was recognised over half a century ago that malaria has been a major force for evolutionary selection on the human genome, and that certain haematological disorders have risen to high frequency in malaria-endemic areas because they reduce the risk of death due to malaria.^{1–3} Sickle haemoglobin (HbS) and glucose-6-phosphate dehydrogenase (G6PD) deficiency are often-quoted examples of natural selection due to malaria, and many other genetic associations with resistance or susceptibility to malaria have been reported.^{2–9} However the current literature contains many conflicting lines of evidence based on relatively small studies that have not been independently replicated.

To address this problem we conducted a large multi-centre case-control study of severe malaria across 12 locations in Burkina Faso, Cameroon, The Gambia, Ghana, Kenya, Malawi, Mali, Nigeria, Tanzania, Vietnam and Papua New Guinea (Supplementary Table 1, Supplementary Figure 1). The structure of this consortial project has been described elsewhere¹⁰ and information about each of the partner studies can be found on the MalariaGEN website (see URLs). We used the World Health Organisation definition of severe malaria which comprises a broad spectrum of life-threatening clinical complications of *Plasmodium falciparum* infection.^{11–15} In this report we examine genetic associations with severe malaria in general and with two distinct clinical forms of severe malaria: cerebral malaria with a Blantyre coma score of less than 3; and severe malarial anaemia with a haemoglobin level of less than 5g/dl or a haematocrit of less than 15%.

RESULTS

Samples and Clinical Data

The first stage of work was to collect standardised clinical data on severe malaria from multiple locations (Supplementary Table 2). This presents many practical challenges as severe malaria is an acute illness that mainly occurs in resource-poor settings where laboratory facilities are limited and medical records can be unreliable. It was necessary to allow for variations in the design and implementation of the study in different settings, which depended on a range of local circumstances. Investigators at different sites agreed at the outset on principles for sharing data and on standardised clinical definitions, and they also worked together to define best ethical practices across different local settings including the development of guidelines for informed consent.^{10,16,17} A set of web tools was developed to enable investigators to curate data in their locally used format before transforming it to the standardised format necessary for data from different sites to be merged.

After data curation and quality control (see Online Methods) 11,890 cases of severe malaria and 17,441 controls were included for analysis (Table 1 and Supplementary Table 3). Controls were intended to be representative of the populations to which the cases belonged, i.e. a minority of controls may have subsequently gone on to develop severe malaria. Supplementary Table 4 shows the ethnic composition of cases and controls at each location. A total of 6,283 cases had cerebral malaria or severe malarial anaemia, of which 3,345 had cerebral malaria only, 2,196 had severe malarial anaemia only and 742 had both cerebral

malaria and severe malarial anaemia (Table 1). A further 5,607 cases did not have cerebral malaria or severe malarial anaemia according to the criteria used here but satisfied the WHO definition of severe malaria which includes a range of other clinical complications, such as acidosis, respiratory distress and hypoglycaemia, which are not explored in detail in the present analysis.¹¹

Most of the cases of severe malaria were young children, with median ages ranging from 1.3 to 3.8 years at different study sites, except in Vietnam where most were young adults with a median age of 29 years. The median age for severe malarial anaemia was 32 months and for cerebral malaria was 62 months (Supplementary Table 3). Cerebral malaria was more common than severe malarial anaemia in The Gambia, Kenya, Malawi and Vietnam, whereas severe malarial anaemia was more common than cerebral malaria in Burkina Faso, Cameroon, Ghana, Mali, Papua New Guinea and Tanzania. It has previously been observed that severe malarial anaemia particularly affects young African children exposed to extremely high levels of malaria transmission, whereas cerebral malaria particularly affects older African children exposed to lower levels of malaria transmission^{18,19}. The mean case fatality rate following treatment was 18% for cerebral malaria, 4% for severe malarial anaemia and 10% for severe malaria overall (Table 1).

Genetic loci analysed

We tested previously reported associations with severe malaria for 55 SNPs in 27 gene regions: *ABO*, *ADORA2B*, *ATP2B4*, *C6*, *CD36*, *CD40LG*, *CR1*, *DARC*, *G6PD*, *GNAS*, *HBB*, *ICAM1*, *IL10*, *IL13*, *IL1A*, *IL1B*, *IL22*, *IL4*, *IRF1*, *LTA*, *NOS2*, *SPTB*, *TLR1*, *TLR4*, *TLR6*, *TLR9*, and *TNF* (for references see Supplementary Table 5–7). Note that this report focuses on SNPs that we were able to genotype reliably using the Sequenom platform, i.e. we do not here consider structural variants such as those responsible for α -thalassaemia. Other SNPs were included for purposes of quality control (Supplementary Tables 5–7, Supplementary Figure 2). All SNPs were initially tested for association with severe malaria using a standard logistic regression method and assuming fixed effects across different populations (See Online Methods). We also tested for association with cerebral malaria and severe malarial anaemia individually, and for different models of inheritance. Results for all SNPs are shown in Supplementary Tables 8–10 and Supplementary Figure 3, and below we discuss those SNPs that showed strong evidence of association in the multi-centre analysis.

HBB

HBB encodes β -globin which has three well-known structural variants that have been associated with resistance to malaria: haemoglobin S (HbS), haemoglobin C (HbC) and haemoglobin E (HbE)^{2,7–9}. The SNP responsible for HbS, rs334, was present at all the African sites with heterozygote frequencies in controls ranging from 0.05 (Malawi) to 0.22 (Nigeria). Heterozygotes had a reduced risk of severe malaria (odds ratio (OR) 0.14, $P = 1.6 \times 10^{-225}$), cerebral malaria (OR 0.11, $P = 4.7 \times 10^{-88}$) and severe malarial anaemia (OR 0.11, $P = 9.3 \times 10^{-65}$) as shown in Figure 1, Table 2 and Supplementary Table 11. The SNP responsible for HbC, rs33930165, was present only in West Africa (Burkina Faso, The Gambia, Ghana, Mali, Cameroon and Nigeria) with frequencies of the derived (non-ancestral/non-reference) alleles ranging from 0.01 to 0.15 in controls. The strongest signal of

association with severe malaria was seen under an additive genetic model, i.e. the greatest protective effect was seen in homozygotes. As shown in Table 2, each copy of the derived allele reduced the risk of severe malaria by 29% (OR 0.71, $P = 6.9 \times 10^{-9}$), of cerebral malaria by 28% (OR 0.72, $P = 0.01$) and of severe malarial anaemia by 26% (OR 0.74, $P = 2.1 \times 10^{-3}$). The SNP responsible for HbE, rs33950507, was found at a derived allele frequency of 0.4 in the S'tieng ethnic group in Vietnam but was rare or absent in other ethnic groups, such that the sample was too small to estimate association with severe malaria.

ABO

ABO encodes the glycosyltransferase enzyme which determines ABO blood group. Individuals who are homozygous for a single nucleotide deletion (rs8176719) in *ABO* have an inactive form of the glycosyltransferase and can be classified as blood group O.²⁰ Estimated by this method, the frequency of blood group O in control samples ranged from 0.32 in Papua New Guinea to 0.62 in Nigeria (Supplementary Table 12). Combined analysis across all sites showed that blood group O was associated with decreased risk of severe malaria (OR 0.74, $P = 5.0 \times 10^{-32}$), of cerebral malaria (OR 0.73, $P = 8.9 \times 10^{-16}$) and of severe malarial anaemia (OR 0.68, $P = 7.9 \times 10^{-14}$) (Figure 1 and Table 2). We also analysed rs8176746, a non-synonymous coding SNP in *ABO* that is in linkage disequilibrium (LD) with rs8176719 and determines the production of B antigens, such that the majority of individuals carrying the derived allele express blood group B.²⁰⁻²² The derived allele was associated in a dominant gene model with increased risk of severe malaria (OR=1.25, $P = 2.0 \times 10^{-17}$) (Table 2).

G6PD

G6PD is an X-linked gene encoding glucose-6-phosphate dehydrogenase with many allelic variants²³. The major form of G6PD enzyme deficiency in Africa is encoded by the derived allele of rs1050828, commonly known as G6PD+202T.²⁴ Two other SNPs that cause G6PD deficiency are found in The Gambia but are rare in the other populations studied here.²⁵ In this study, the G6PD+202T allele was present at frequencies ranging from 0.03 in Gambia to 0.28 in Nigeria (Supplementary Table 13). Aggregated across all African sites, we found an increased risk of severe malarial anaemia in male hemizygotes (OR 1.49, $P = 3.6 \times 10^{-5}$) and in females homozygotes under a recessive model of association (OR=1.94, $P = 1.9 \times 10^{-3}$) (Table 3 and Supplementary Tables 14, 15). In contrast, there was a trend towards decreased risk of cerebral malaria in female heterozygotes (OR 0.87, $P = 0.06$) and male hemizygotes (OR 0.81, $P = 0.01$) (Supplementary Tables 14, 16, 17). Below we discuss this heterogeneity of effect in more detail. Similar but weaker trends were observed for rs1050829 which marks the ancestral lineage on which *G6PD*+202 originated²⁴.

ATP2B4

ATP2B4, encoding a calcium transporter found in the plasma membrane of erythrocytes, has been identified by genome-wide association study (GWAS) as a malaria resistance locus²⁶. We typed 4 SNPs in this gene which were found to be in LD such that the derived alleles of rs10900585 and rs55868763 were associated with increased risk of severe malaria, while the

derived alleles of rs4951074 and rs1541255 were associated with decreased risk (Table 2, Supplementary Tables 8–10 and 18). When averaged across all African sites, individuals carrying at least one copy of the derived allele of rs10900585 had an odds ratio of 1.32 for severe malaria ($P = 1.7 \times 10^{-9}$) while individuals homozygous for the derived allele at rs4951074 had an odds ratio of 0.77 ($P = 7.6 \times 10^{-7}$). In both cases the magnitude of genetic effect was similar for cerebral malaria and severe malarial anaemia (Figure 1).

CD40LG

CD40LG is a gene on the X chromosome encoding CD40 ligand that has previously been associated with severe malaria.²⁷ Homozygotes for the derived allele of a SNP in the 5' upstream region (rs3092945) showed reduced risk of severe malaria (OR=0.85, $P = 1.1 \times 10^{-6}$) with a similar trend of protection both in males (OR=0.90, $P = 0.01$) and in females (OR=0.78, $P = 8.9 \times 10^{-5}$) when the data were aggregated across sites (Table 3). However when sites were analysed individually the results were strikingly different between sites: homozygotes for the derived allele showed significantly reduced risk in of severe malaria in The Gambia but significantly increased risk in Kenya (Supplementary Table 19).

Other Loci

None of the other loci tested here showed consistent evidence of association with severe malaria in the multi-centre analysis with a significance of $P < 10^{-4}$. Supplementary Tables 8–10 and Supplementary Figure 3 show results for all variants tested, some of which are weak associations that merit further investigation. At the *CD36* locus, heterozygotes for codon variant G1439C tended to have reduced risk of severe malaria (OR 0.67, $P = 4.2 \times 10^{-4}$). Other weak signals of association (P values in the range of 0.05 to 0.001) were observed for *CD36*, *IL1A* and *IRF1* with severe malaria overall, for *CR1* and *IL4* with cerebral malaria and for *IL20RA* with severe malarial anaemia. While it is clear from these data that many genetic associations reported in the literature might have been false positives as has been observed for other common diseases²⁸, it is undoubtedly also the case that authentic genetic associations might be missed by multi-centre studies if the effect is weak and there is heterogeneity of effect across different study sites.

Epistasis Between Significantly Associated Loci

Epistasis between malaria resistance loci has been reported in previous studies^{29,30}. We therefore tested for pairwise interaction between all SNPs that showed significant associations at the *HBB*, *ABO*, *G6PD*, *ATP2B4* and *CD40LG* loci (Supplementary Table 20 and Supplementary Figure 4). This did not reveal any strong evidence of interaction but a marginally significant effect was observed between the *ATP2B4* locus (rs10900585) and HbC (rs33930165; $P = 1.3 \times 10^{-3}$) such that the ancestral allele of rs10900585, which is the minor allele in Africa, tends to reverse the protective effect of HbC. This warrants further investigation, since *ATP2B4* is the major erythrocyte calcium channel, and intracellular calcium levels have been noted to affect the clinical phenotype of sickling disorders³¹.

Heterogeneity of Effect

The large sample size of this study allowed us to investigate the heterogeneity of effect of malaria resistance loci in greater detail than has hitherto been possible. When associations with *HBB*, *ABO*, *G6PD*, *ATP2B4* and *CD40L* were broken down by population, and when cerebral malaria and severe malarial anaemia were treated as separate phenotypic entities, various different patterns were observed (Figure 1). The standard approach to genetic association analysis, such as the logistic regression methods used above, uses a fixed-effects model which assumes that true associations should be constant across different studies. However patterns of disease association could potentially vary due to a range of genetic, environmental and biological factors, and by understanding this variation we might gain important scientific insights. We therefore used a Bayesian statistical framework to evaluate different models of genetic association allowing for heterogeneity of effect (see Online Methods and ref³²). In essence this approach weighs up the evidence that a genetic effect is fixed or heterogeneous when it is compared across different populations and clinical phenotypes, i.e. cerebral malaria and severe malarial anaemia. Here we allow for two sorts of heterogeneous effect: correlated effects, that are not fixed but tend to behave in a similar way; and independent effects, where there is no tendency to behave in a similar way. We estimated the posterior probability of each model under the *a priori* assumption that all models are equally likely. Figure 2 shows the results for specific SNPs in *HBB*, *ABO*, *G6PD*, *ATP2B4* and *CD40L* and Supplementary Figure 5 shows the results for all SNPs.

For HbS and blood group O there was strong evidence for fixed or correlated effects across different populations and clinical phenotypes (Figures 1 and 2). HbS confers a remarkably fixed level of protection against cerebral malaria and severe malarial anaemia, with heterozygotes showing 89% reduced risk for both conditions when averaged across sites (Table 2). HbC and blood group O are less strongly protective than HbS but they too have very similar effects on cerebral malaria and severe malarial anaemia. There are many different theories about the molecular and cellular mechanisms by which HbS, HbC and blood group O act to protect against malaria.^{22,33–36} The fact that these variants protect equally against clinical complications as disparate as coma and severe anaemia implies that they act through some general mechanism, e.g. by suppressing parasite density, rather than through a specific effect on particular pathological processes such as cerebral malaria or severe malarial anaemia.

At the *CD40LG* and *ATP2B4* loci there was evidence for heterogeneity of effect across populations (Figures 1 and 2). For *CD40LG* rs3092945, the posterior probability of independent effects in different populations was greater than 90%, whereas the effects on clinical phenotype were relatively constant. In the case of *ATP2B4* rs10900585, the posterior probability was more evenly balanced between the various models, with some evidence for both fixed and independent effects in different populations. Heterogeneity of effect across populations might indicate some source of biological variation such as epistasis or gene-environment interactions, or it might be because the SNPs showing association are simply genetic markers showing variable patterns of linkage disequilibrium with true malaria resistance alleles. Variable patterns of linkage disequilibrium are particularly common in Africa, such that authentic genetic associations may fail to replicate in different

locations unless the causal variant is directly genotyped, but in the long term this form of population heterogeneity may prove extremely useful in genetic fine-mapping of causal variants.^{32,37,38}

The genetic effects observed for *G6PD* were strikingly different in character from those seen at the *HBB*, *ABO*, *ATP2B4* and *CD40LG* loci (Figures 1 and 2). There was strong heterogeneity of effect across different phenotypes, with more than 80% posterior probability that the *G6PD*+202T allele has independent effects on cerebral malaria and severe malarial anaemia, with relatively little heterogeneity of effect across different populations. Previous studies have concluded that G6PD deficiency is protective against *P. falciparum* although there has been debate about whether the protective effect is confined to female heterozygotes or is also present in male hemizygotes.^{4,5,25,39} With a much larger sample size than any previous study, these data reveal that male hemizygotes and female homozygotes have increased risk of severe malarial anaemia, whereas male hemizygotes and female heterozygotes have reduced risk of cerebral malaria. If males and females are combined in an additive model, the *G6PD*+202 T allele confers reduced risk of cerebral malaria (OR 0.91, $P=6.1 \times 10^{-3}$) but increased risk of severe malarial anaemia (OR 1.19, $P=2.6 \times 10^{-5}$) and the overall effect on severe malaria is close to neutral (OR 1.02, $P=0.15$) (Table 3, Supplementary Tables 13–17).

Discussion

Severe malarial anaemia is a complex pathological entity.^{40,41} Parasites invade and destroy erythrocytes as they replicate, and the host response to infection also leads to erythrocyte destruction and bone marrow suppression. In the developing world, the anaemia is often aggravated by chronic nutritional deficiency and helminthic infection.^{41,42} G6PD deficiency and the haemoglobinopathies are interesting examples of a biological trade-off between an inherent tendency to cause anaemia and the potential to protect against anaemia by protecting against malaria. The present study allows us to observe the outcome of this trade-off with greater resolution than has hitherto been possible, showing that the risk of severe malarial anaemia is significantly reduced for HbS heterozygotes, HbC homozygotes and HbC heterozygotes, while for G6PD-deficient male hemizygotes and female homozygotes the risk is significantly increased.

G6PD deficiency has excited much interest among evolutionary biologists because it is so common in the human population and displays such a remarkable diversity of allelic forms throughout the tropics and sub-tropics.^{3,6,9,23,39,43} The main force for evolutionary selection is widely assumed to be severe malaria due to *P. falciparum* but this is called into question by the present findings which show that G6PD deficiency has little effect on the overall risk of severe malaria. These data do not exclude the possibility that *P. falciparum* has played a role, as the present-day risk of cerebral malaria and severe malarial anaemia may not accurately reflect patterns of disease and fatality caused by *P. falciparum* in the past, particularly before anti-malarial drugs became widely used. However it is necessary to consider the possibility of some other evolutionary driving force and an obvious candidate is *P. vivax*, once regarded as benign but increasingly recognised to cause a significant burden of fatality and severe disease⁴⁴. In a region of Thailand where both species are endemic,

G6PD deficiency has been observed to suppress *P. vivax* more effectively than *P. falciparum* infection⁴⁵. An important biological difference from *P. falciparum* is that *P. vivax* preferentially infects reticulocytes (i.e. young erythrocytes) which have higher levels of G6PD enzyme activity than older erythrocytes, and it is conceivable that in these circumstances G6PD deficiency might exert a stronger protective effect. The geographic distribution of G6PD deficiency broadly coincides with the transmission range of *P. vivax*, with the notable exception of large parts of sub-Saharan Africa where G6PD deficiency is common but *P. vivax* is absent^{43,46}. It has recently been discovered that chimpanzee and gorillas in Africa carry parasites that are closely related to *P. vivax*⁴⁷. Evolutionary analysis of these ape parasite lineages indicates that human *P. vivax* is of African origin, suggesting that *P. vivax* may have been common in Africa before the selective sweep of the FY*O allele (Duffy negative blood group) led to its elimination from most of the continent^{48,49}.

Large multi-centre studies have transformed the field of human genetics over the past decade but there remain significant practical challenges in conducting such studies in the developing world. Obtaining reliable phenotypic data can be difficult in resource-poor settings, particularly for acute conditions such as severe malaria which rely on accurate clinical records at the time of illness. We have endeavoured to overcome these obstacles by establishing systems for standardising and sharing data from research groups in different countries. The extremely strong phenotypic associations observed for sickle cell trait and blood group O, and their remarkably consistent effects on different clinical forms of severe malaria, illustrate the effectiveness of this approach and provide a benchmark for evaluating other loci. A key finding is that most previously reported candidate gene associations fail to replicate, and an important use of the sample collections established by this project will be to discover authentic novel loci by conducting genome-wide association studies (GWAS) on a larger scale than previously possible. A critical consideration when conducting multi-centre GWAS is that true genetic effects can exhibit marked heterogeneity across different populations, and here we propose a statistical framework for dealing with this problem. There are many potential sources of heterogeneity including interactions with other infections and environmental variables. An important source of heterogeneity could be parasite genetic variation, whose impact on the clinical outcome of malaria remains very poorly understood, and this is an area of future investigation which could yield important biological insights. Previously this has been difficult to address in a systematic manner but with growing knowledge about the natural landscape of genome variation in both the host and the parasite there is the potential to discover novel genetic interactions, e.g. between parasite ligands and host receptors for erythrocyte invasion, that could be of great practical importance for vaccine development. This will require a framework for large-scale genetic epidemiology studies, and here we demonstrate the feasibility of integrating data across multiple locations to achieve scientific insights which could not be achieved by individual studies in isolation.

Online Methods

Cases of severe malaria were recruited on admission to hospital, usually as part of a larger programme of clinical research on malaria, designed and led by local investigators. A control group was recruited at each of the study sites to match the ethnic composition of the

cases (Supplementary Tables 1,2,4). The control group was intended to be representative of the general population, and cord blood samples were used as controls at several study sites. We describe elsewhere the details of study design at individual sites and local epidemiological conditions including malaria endemicity (refs^{21,50–53} and the MalariaGEN website (see URLs). Following consultation with the Severe Malaria in African Children network and other clinical experts, a standardised case-report form was developed to record the clinical features of severe malaria¹⁴ (The case-report form can be found on the MalariaGEN website [URLs]). This was not intended to replace local practice but to encourage uniformity in the core data collected across the different sites. A secure web application was developed to enable investigators to upload and curate their data, and to transform it into standardised units and format, before releasing it to the consortial database. A data fellow was appointed at each site with responsibility for the process of integrating local clinical data with the central database. As part of a capacity building program, data fellows received training in data management and analysis. Moreover, ethics advice, support and training/capacity-building in ethics were provided to Data Fellows and partners throughout the life of the study.

The normalised clinical data from each study site were combined to ascertain phenotypes in a standardised manner across the entire data set (Supplementary Table 21). A case of severe malaria was defined as an individual admitted to a hospital or clinic with *P. falciparum* parasites in the blood film and with clinical features of severe malaria as defined by WHO criteria^{11,12}. Severe malaria comprises a number of overlapping syndromes, the most commonly reported being cerebral malaria and severe malarial anaemia. In keeping with standard criteria cerebral malaria is defined here as a case of severe malaria with Blantyre Coma Score of <3 for a child, or a Glasgow Coma Score of <9 for an adult. Severe malarial anaemia is defined here as a case of severe malaria with a haemoglobin level of <5g/dl or a haematocrit level of <15%. In this report we do not attempt to classify other severe malaria syndromes, such as respiratory distress, that are more complicated to standardise among study sites, although they would be present in our data set. Control samples were either collected from cord bloods or if sampled from the local population, were microscopically negative for malaria.

Genotyping

We selected the Sequenom[®] iPLEX Mass-Array platform for genotyping because of its sample high-throughput capacity, adaptability for assay design and ability to genotype up to 40 SNPs in one reaction. We chose to design 2 iPLEX multiplexes as a compromise between maximising the SNPs we could type on any given sample and the time and cost to genotype all the samples submitted to the MalariaGEN Resource Centre for the various projects.

Altogether 89 SNPs were finally designed and tested in 2 rounds of multiplex design on 33,138 samples (Supplementary Tables 5–7, Supplementary Figure 2). Of these, 16 SNPs were excluded during the multiplex design and testing phase for a number of reasons: 9 were missing data in more than 20% of the samples; 1 showed a mismatch between the published sequence and the human reference genome; 1 was monomorphic; and 5 could not be

redesigned into the final multiplexes (Supplementary Table 7). Of the remaining 73 SNPs genotyped, 55 were included on the basis of a known genetic association with severe malaria (Supplementary Table 5), 3 were used to confirm or type gender and 15 were selected to aid sample quality control (Supplementary Table 6, Supplementary Note and Supplementary Tables 22–24). In the quality control phase, one of the 55 SNPs with a known association with severe malaria, rs1800750 (TNF-376), showed a large deviation in Hardy-Weinberg Equilibrium and was removed from further analysis (HWE see Statistical Analysis section below).

Samples

A total of 38,926 individual records comprising 16,433 cases of severe malaria and 22,492 controls were obtained from across the 12 study sites (Supplementary Table 2). Clinical data for gender were missing in 4% of records (we confirm or type sample genders by genotyping) and for ethnic group in 2% of records. A total of 33,138 samples were genotyped. Each sample was assessed for inclusion in the analysis if it was successfully genotyped in more than 90% of 65 analysis SNPs (excluding ATP2B4 SNPs): we excluded 789 samples on this basis (Supplementary Figure 6). The majority of sample failures were found to be due to blood storage and DNA extraction issues. After quality control of both phenotypic and genotypic data, 11,890 severe malaria cases and 17,441 controls were included for analysis (Table 1).

There were 213 different ethnic groups of which 41 comprised at least 5% of individuals at a study site; these included Mandinka, Jola, Wollof, Fula (Gambia); Bambara, Malinke, Peulh, Sarakole (Mali); Mossi (Burkina Faso); Akan, Frarra, Nankana, Kasem (Ghana); Yoruba (Nigeria); Bantu, Semi-Bantu (Cameroon); Chonyi, Giriama, Kauma (Kenya); Mzigua, Wasambaa, Wabondei (Tanzania); Chewa (Malawi); Madang, Sepik (Papua New Guinea); Kinh (Viet Nam). For purposes of analysis we classified ethnic groups with very small sample size (less than 5% of individuals at any study site) as 'other' (Supplementary Table 4).

Statistical Analysis

All statistical analyses were performed using the statistical software environment R (see URLs). As part of the genotyping QC process we identified SNPs with large deviations from HWE that might possibly signify assay failure. Overall HWE was assessed from the distribution of HWE P-values calculated for each SNP by country and ethnic group after discarding groups where the calculated allele frequency was less than $5/2N$ (where N is the number of individuals in the group)⁵⁴. An assay was marked for potential exclusion if the results deviated from HWE ($P < 10^{-4}$) in more than 4 ethnic groups (Supplementary Figure 7).

Single SNP tests of association, adjusted for HbS genotype, gender and ethnicity, for association with severe malaria and severe malaria subtypes cerebral malaria only and severe malarial anaemia only were performed for the 55 SNPs with a known association with severe malaria. Standard logistic regression models were used for tests of association at each autosomal SNP (Supplementary Table 25). Primary analyses comprised tests of

association between each SNP and severe malaria phenotypes across all individuals combined as well as separately by gender (X chromosome SNPs only) and site: genotypic, additive, dominant, recessive and heterozygote advantage genetic models were considered. For X-chromosome SNPs, males were treated as homozygous females. Therefore when analysing the males only for X chromosome SNPs, the genotypic, dominant, recessive and additive models are equivalent and the heterozygous model is redundant: in this case we present the results from the dominant model (referred to as the male hemizygote model for males at X chromosome SNPs) and note that the ORs correspond to the change in the odds of disease for males hemizygous for the derived allele compared to males hemizygous for the ancestral allele. For combined analyses of males and females at X chromosome SNPs, robust estimates of variance were used to account for the unequal variance⁵⁵ and all models are then appropriate. In secondary analyses we considered additional genetic models comparing effects between homozygotes and heterozygotes at selected SNPs. ORs and 95% confidence intervals were derived from Wald tests applied to regression coefficients. Significance was assessed using likelihood ratio tests of association except for combined analyses of males and females at X chromosome SNPs where Wald tests were applied using the robust variance estimates. Results are presented with respect to the association between the derived (non-ancestral allele) and the severe malaria phenotype in question.

The standard logistic regression analyses assume that effects are fixed across all sites. In order to investigate evidence for genetic heterogeneity across severe malaria subtypes both within and across African sites, we compared different models of association in a Bayesian statistical framework. The models we considered comprised fixed, independent or correlated effects between subtypes within a site crossed with fixed, independent or correlated effects of each subtype across all sites. See Supplementary Information for more details. For each SNP we assumed a normally distributed prior on the log odds ratio of association with mean zero and standard deviation σ where $\sigma = 1$ for rs334 to reflect the prior belief that the effect size is large consistent with the observed ORs of approximately 0.1; $\sigma = 0.4$ for SNPs found to be significant in fixed-effect analysis and; $\sigma = 0.2$ otherwise. To model fixed, independent or correlated effects either within or across sites, we set correlation parameters between subtypes to 1, 0.1 and 0.96 respectively. Multinomial regression was used to make independent maximum likelihood estimates of the effect of each SNP on these mutually exclusive subtypes for all individuals combined at each African site. Estimates were adjusted for sex, ethnicity and, with the exception of rs334, sickle-cell trait. Approximate Bayes Factors (ABFs) were then calculated for each SNP and model and used to estimate a posterior probability of each of the models for each SNP.

We also tested for interaction between all pairs of SNPs that were significant in the single SNP analysis; 25 pairs of markers were tested (Supplementary Table 20). We considered two different statistical models of interaction: (1) A 1 degree of freedom (df) “Best Model” test for the optimal genetic model for each SNP (as defined by association with SM for all individuals across all sites in a fixed effect model adjusted for ethnicity and gender) at each of the interacting loci; (2) A more general “Genotype” test using a model which allows for separate effects for heterozygous and homozygous genotypes at each of the interacting loci. At X chromosome SNPs, male individuals were treated as homozygous females and only additive effects were considered. For a pair of autosomal SNPs, the Genotype test is then a 4

degree of freedom (df) test of interaction; for a pair comprising an autosomal and an X chromosome SNP, it is a 3 df test and for a pair of X chromosome SNPs, a 1df test. Tests of interaction were performed by testing whether the regression coefficients that represent interaction terms in the corresponding logistic regression model were equal to zero or not. These tests are described in more detail in Cordell⁵⁶.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Malaria Genomic Epidemiology Network*, Kirk A. Rockett^{1,2,3}, Geraldine M. Clarke^{1,3}, Kathryn Fitzpatrick¹, Christina Hubbart¹, Anna E. Jeffreys¹, Kate Rowlands¹, Rachel Craik¹, Muminatou Jallow^{4,5}, David J. Conway^{4,6}, Kalifa A. Bojang⁴, Margaret Pinder⁴, Stanley Usen⁴, Fatoumatta Sisay-Joof⁴, Giorgio Sirugo⁴, Ousmane Toure⁷, Mahamadou A. Thera⁷, Salimata Konate⁷, Sibiry Sissoko⁷, Amadou Niangaly⁷, Belco Poudiougou⁷, Valentina D. Mangano⁸, Edith C. Bougouma⁹, Sodiomon B. Sirima⁹, David Modiano⁸, Lucas N. Amenga-Etego¹⁰, Anita Ghansah¹¹, Kwadwo A. Koram¹¹, Michael D. Wilson¹¹, Anthony Enimil¹², Jennifer Evans^{13,14}, Olukemi Amodu¹⁵, Subulade Olaniyan¹⁵, Tobias Apinjoh¹⁶, Regina Mugri¹⁷, Andre Ndi¹⁷, Carolyne M. Ndila¹⁸, Sophie Uyoga¹⁸, Alexander Macharia¹⁸, Norbert Peshu¹⁸, Thomas N. Williams^{18,19}, Alphaxard Manjurano^{20,21}, Eleanor Riley²¹, Chris Drakeley^{20,21}, Hugh Reyburn^{20,21}, Vysaul Nyirongo²², David Kachala²³, Malcolm Molyneux²⁴, Sarah J. Dunstan^{23,24,25}, Nguyen Hoan Phu^{23,26}, Nguyen Thi Ngoc Quyen²³, Cao Quang Thai^{23,26}, Tran Tinh Hien^{23,26}, Laurens Manning^{27,28}, Moses Laman²⁷, Peter Siba²⁷, Harin Karunajeewa²⁸, Steve Allen²⁹, Angela Allen³⁰, Timothy M. E. Davis²⁸, Pascal Michon^{27,31}, Ivo Mueller^{27,32,32}, Angie Green¹, Sile Molloy¹, Kimberly J. Johnson¹, Angeliki Kerasidou^{1,34}, Victoria Cornelius¹, Lee Hart¹, Aaron Vanderwal¹, Miguel SanJoaquin¹, Gavin Band¹, Si Quang Le¹, Matti Pirinen¹, Nuno Sepúlveda³⁵, Chris C.A. Spencer^{1,3}, Taane G. Clark³⁵, Tsiri Agbenyega^{12,36}, Eric Achidi¹⁷, Ogobara Doumbo⁷, Jeremy Farrar^{23,24}, Kevin Marsh¹⁸, Terrie Taylor³⁷, and Dominic P. Kwiatkowski^{1,2,3}

Affiliations

¹Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK

²The Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK

³MRC Centre for Genomics and Global Health, University of Oxford, Oxford, UK

⁴Medical Research Council Unit, Atlantic Boulevard, Serrekunda, The Gambia

⁵Royal Victoria Teaching Hospital, Independence Drive, Banjul, The Gambia

⁶Department of Pathogen Molecular Biology, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

- ⁷Malaria Research and Training Centre, Faculty of Medicine University of Bamako
Bamako Mali
- ⁸University of Rome La Sapienza, Italy
- ⁹Centre National de Recherche et de Formation sur le Paludisme (CNRFP),
Ouagadougou, Burkina Faso
- ¹⁰Navrongo Health Research Centre, Navrongo, Ghana
- ¹¹Noguchi Memorial Institute for Medical Research, Accra, Ghana
- ¹²Komfo Anokye Teaching Hospital, Ghana
- ¹³Department of Molecular Medicine, Bernhard Nocht Institute for Tropical Medicine,
Postfach 30 41 2, D-20324 Hamburg, Germany
- ¹⁴Kumasi Centre for Collaborative Research, Kumasi, Ghana
- ¹⁵University of Ibadan, Nigeria
- ¹⁶Department of Biochemistry & Molecular Biology, University of Buea, Buea, South
West Region, Cameroon
- ¹⁷Department of Medical Laboratory Sciences, University of Buea, Buea, South
West Region, Cameroon
- ¹⁸KEMRI-Wellcome Trust Research Programme, PO Box 230, Kilifi, Kenya
- ¹⁹Faculty of Medicine, Department of Medicine, Imperial College, Exhibition Road,
London SW7 2AZ, UK
- ²⁰Joint Malaria Programme, Kilimanjaro Christian Medical Centre, PO box 2228,
Moshi, Tanzania
- ²¹Faculty of Infectious and Tropical Diseases, London School of Hygiene and
Tropical Medicine, Keppel Street, London WC1E 7HT, UK
- ²²Malawi-Liverpool-Wellcome Trust Clinical Research Programme, College of
Medicine, University of Malawi, PO Box 30096, Blantyre, Malawi
- ²³Oxford University Clinical Research Unit, 764 Vo Van Kiet, District 5, Ho Chi Minh
City, Viet Nam
- ²⁴Centre for Tropical Medicine, Nuffield Department of Clinical Medicine, Oxford
University, OX3 7LJ, UK
- ²⁵Nossal Institute of Global Health, University of Melbourne, Australia
- ²⁶Hospital for Tropical Diseases, 764 Vo Van Kiet, District 5, Ho Chi Minh City, Viet
Nam
- ²⁷Papua New Guinea Institute of Medical Research, PO Box 60, Garoka, EHP441,
PNG
- ²⁸University of Western Australia, Perth, Australia
- ²⁹Swansea University, Swansea, Wales, UK

- ³⁰Weatherall Institute of Molecular Medicine, Oxford University, Oxford, UK
- ³¹Faculty of Health Sciences, Divine Word University, Madang, Papua New Guinea
- ³²Walter and Eliza Hall Institute of Medical Research, Australia
- ³³Barcelona Centre for International Health Research (CRESIB), Barcelona, Spain
- ³⁴The Ethox Centre, Nuffield Department of Population Health, University of Oxford, Old Road Campus, Oxford, OX3 7LF
- ³⁵London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK
- ³⁶Kwame Nkrumah University of Science and Technology, Ghana
- ³⁷Blantyre Malaria Project, College of Medicine, University of Malawi

Acknowledgements

The MalariaGEN Project is supported by the Wellcome Trust (WT077383/Z/05/Z) and the Bill & Melinda Gates Foundation through the Foundations of the National Institutes of Health (566) as part of the Grand Challenges in Global Health Initiative. The Resource Centre for Genomic Epidemiology of Malaria is supported by the Wellcome Trust (090770/Z/09/Z). This research was supported by the Medical Research Council (G0600718 ; G0600230), the Wellcome Trust Biomedical ethics Enhancement Award (087285) and Strategic Award (096527). Dominic Kwiatkowski receives support from the Medical Research Council (G19/9). Chris C A Spencer was supported by a Wellcome Trust Career Development Fellowship (097364/Z/11/Z). The Wellcome Trust also provides core awards to The Wellcome Trust Centre for Human Genetics (075491/Z/04; 090532/Z/09/Z) and the Wellcome Trust Sanger Institute (077012/Z/05/Z). The Mali MRTC – BMP group is supported by an ICDR grant of NIAID-NIH to UoMaryland and UoBamako and the Mali-NIAID/NIH ICER at USTTB, Mali. Contributions from Nigeria to the CP1 was supported financially by a grant within the BioMalPar European Network of Excellence (LSHP-CT-2004-503578). Eric Achidi received partial funding from the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement N° 242095 – EVIMalaR and the Central African Network for Tuberculosis, HIV/AIDS and Malaria (CANTAM) funded by the European and Developing Countries Clinical Trials Partnership (EDCTP). Thomas N Williams is funded by Senior Fellowships from the Wellcome Trust (076934/Z/05/Z and 091758/Z/10/Z) and through the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement N° 242095 – EVIMalaR. The KEMRI-Wellcome Trust Programme is funded through core support from the Wellcome Trust. This paper is published with the permission of the Director of KEMRI. Carolyne Ndila is supported through a strategic award to the KEMRI-Wellcome Trust Programme by the Wellcome Trust (084538). Tanzania/KCMC/JMP received funding from MRC grant number (G9901439). We would like to thank all the Vietnamese individuals who agreed to provide samples for this study. We acknowledge the work of the clinical staff from the Hospital of Tropical Diseases, HCMC and Phuoc Long and Dong Xoai District Hospitals in Binh Phuoc province, Viet Nam, who initially diagnosed and studied the patients with severe malaria. We would like to thank Dr. Nguyen Thi Hieu and his staff from Hung Vuong Obstetric Hospital for the collection of the cord blood controls. The clinical component of this study was funded through the Wellcome Trust Major Overseas Program in Vietnam (089276/Z/09/Z). Laurens Manning was supported through the Basser (Royal Australasian College of Physicians) and National Health and Medical Research Council (NHMRC) scholarships. Moses Laman was supported through a Fogarty Foundation Scholarship. Timothy ME Davis was supported through an NHMRC practitioner fellowship.

References

1. Haldane JBS. Disease and evolution. *Ricerca Sci Suppl.* 1949; 19:3–10.
2. Allison AC. Protection afforded by sickle-cell trait against subtertian malarial infection. *Br Med J.* 1954; 1:290–294. [PubMed: 13115700]
3. Allison AC. Genetic factors in resistance to malaria. *Ann N Y Acad Sci.* 1961; 91:710–729. [PubMed: 13682588]
4. Bienzle U, Ayeni O, Lucas AO, Luzzatto L. Glucose-6-phosphate dehydrogenase and malaria. Greater resistance of females heterozygous for enzyme deficiency and of males with non-deficient variant. *Lancet.* 1972; 1:107–110. [PubMed: 4108978]

5. Ruwende C, et al. Natural selection of hemi- and heterozygotes for G6PD deficiency in Africa by resistance to severe malaria. *Nature*. 1995; 376:246–249. [PubMed: 7617034]
6. Tishkoff SA, et al. Haplotype diversity and linkage disequilibrium at human G6PD: recent origin of alleles that confer malarial resistance. *Science*. 2001; 293:455–462. [PubMed: 11423617]
7. Kwiatkowski DP. How malaria has affected the human genome and what human genetics can teach us about malaria. *Am J Hum Genet*. 2005; 77:171–192. [PubMed: 16001361]
8. Verra F, Mangano VD, Modiano D. Genetics of susceptibility to *Plasmodium falciparum*: from classical malaria resistance genes towards genome-wide association studies. *Parasite Immunol*. 2009; 31:234–253. [PubMed: 19388945]
9. Hedrick PW. Population genetics of malaria resistance in humans. *Heredity (Edinb)*. 2011; 107:283–304. [PubMed: 21427751]
10. Malaria_Genomic_Epidemiology_Network. A global network for investigating the genomic epidemiology of malaria. *Nature*. 2008; 456:732–737. [PubMed: 19079050]
11. World_Health_Organisation. Severe falciparum malaria. *Trans R Soc Trop Med Hyg*. 2000; 94(Suppl 1):S1–S90. [PubMed: 11103309]
12. World_Health_Organisation. second edn. World Health Organisation; 2010. Guidelines for the Treatment of Malaria; p. 194
13. Marsh K, et al. Indicators of life-threatening malaria in African children. *N Engl J Med*. 1995; 332:1399–1404. [PubMed: 7723795]
14. Taylor T, et al. Standardized data collection for multi-center clinical studies of severe malaria in African children: establishing the SMAC network. *Trans R Soc Trop Med Hyg*. 2006; 100:615–622. [PubMed: 16551469]
15. Jallow M, et al. Clinical features of severe malaria associated with death: a 13-year observational study in the Gambia. *PloS one*. 2012; 7:e45645. [PubMed: 23029157]
16. Chokshi DA, et al. Valid consent for genomic epidemiology in developing countries. *PLoS Med*. 2007; 4:e95. [PubMed: 17455985]
17. Parker M, et al. Ethical data release in genome-wide association studies in developing countries. *PLoS Med*. 2009; 6:e1000143. [PubMed: 19956792]
18. Snow RW, et al. Relation between severe malaria morbidity in children and level of *Plasmodium falciparum* transmission in Africa. *Lancet*. 1997; 349:1650–1654. [PubMed: 9186382]
19. Reyburn H, et al. Association of transmission intensity and age with clinical manifestations and case fatality of severe *Plasmodium falciparum* malaria. *JAMA : the journal of the American Medical Association*. 2005; 293:1461–1470. [PubMed: 15784869]
20. Yamamoto F, Clausen H, White T, Marken J, Hakomori S. Molecular genetic basis of the histo-blood group ABO system. *Nature*. 1990; 345:229–233. [PubMed: 2333095]
21. Ogasawara K, et al. Molecular genetic analysis of variant phenotypes of the ABO blood group system. *Blood*. 1996; 88:2732–2737. [PubMed: 8839869]
22. Fry AE, et al. Common variation in the ABO glycosyltransferase is associated with susceptibility to severe *Plasmodium falciparum* malaria. *Hum Mol Genet*. 2008; 17:567–576. [PubMed: 18003641]
23. Cappellini MD, Fiorelli G. Glucose-6-phosphate dehydrogenase deficiency. *Lancet*. 2008; 371:64–74. [PubMed: 18177777]
24. Hirono A, Beutler E. Molecular cloning and nucleotide sequence of cDNA for human glucose-6-phosphate dehydrogenase variant A(-). *Proc Natl Acad Sci U S A*. 1988; 85:3951–3954. [PubMed: 2836867]
25. Clark TG, et al. Allelic heterogeneity of G6PD deficiency in West Africa and severe malaria susceptibility. *European journal of human genetics : EJHG*. 2009; 17:1080–1085. [PubMed: 19223928]
26. Timmann C, et al. Genome-wide association study indicates two novel resistance loci for severe malaria. *Nature*. 2012; 489:443–446. [PubMed: 22895189]
27. Sabeti P, et al. CD40L association with protection from severe malaria. *Genes Immun*. 2002; 3:286–291. [PubMed: 12140747]

28. Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. A comprehensive review of genetic association studies. *Genet Med.* 2002; 4:45–61. [PubMed: 11882781]
29. Williams TN, et al. Negative epistasis between the malaria-protective effects of alpha+-thalassemia and the sickle cell trait. *Nat Genet.* 2005; 37:1253–1257. [PubMed: 16227994]
30. Atkinson SH, et al. Epistasis between the haptoglobin common variant and alpha+thalassemia influences risk of severe malaria in Kenyan children. *Blood.* 2014; 123:2008–2016. [PubMed: 24478401]
31. Steinberg MH, Eaton JW, Berger E, Coleman MB, Oelshlegel FJ. Erythrocyte calcium abnormalities and the clinical severity of sickling disorders. *Br J Haematol.* 1978; 40:533–539. [PubMed: 728370]
32. Band G, et al. Imputation-based meta-analysis of severe malaria in three African populations. *PLoS Genet.* 2013; 9:e1003509. [PubMed: 23717212]
33. Fairhurst RM, et al. Abnormal display of PfEMP-1 on erythrocytes carrying haemoglobin C may protect against malaria. *Nature.* 2005; 435:1117–1121. [PubMed: 15973412]
34. Bunn HF. The triumph of good over evil: protection by the sickle gene against malaria. *Blood.* 2012
35. Cserti CM, Dzik WH. The ABO blood group system and *Plasmodium falciparum* malaria. *Blood.* 2007; 110:2250–2258. [PubMed: 17502454]
36. Rowe JA, et al. Blood group O protects against severe *Plasmodium falciparum* malaria through the mechanism of reduced rosetting. *Proc Natl Acad Sci U S A.* 2007; 104:17471–17476. [PubMed: 17959777]
37. Teo YY, Small KS, Kwiatkowski DP. Methodological challenges of genome-wide association analysis in Africa. *Nat Rev Genet.* 2010; 11:149–160. [PubMed: 20084087]
38. Teo YY, et al. Genome-wide comparisons of variation in linkage disequilibrium. *Genome Res.* 2009; 19:1849–1860. [PubMed: 19541915]
39. Luzzatto L. G6PD deficiency and malaria selection. *Heredity.* 2012; 108:456. [PubMed: 22009270]
40. Menendez C, Fleming AF, Alonso PL. Malaria-related anaemia. *Parasitol Today.* 2000; 16:469–476. [PubMed: 11063857]
41. Weatherall DJ, Kwiatkowski D. Hematologic disorders of children in developing countries. *Pediatr Clin North Am.* 2002; 49:1149–1164. [PubMed: 12580360]
42. Calis JC, et al. Severe anemia in Malawian children. *N Engl J Med.* 2008; 358:888–899. [PubMed: 18305266]
43. Howes RE, Battle KE, Satyagraha AW, Baird JK, Hay SIG. 6PD deficiency: global distribution, genetic variants and primaquine therapy. *Adv Parasitol.* 2013; 81:133–201. [PubMed: 23384623]
44. Price RN, et al. Vivax malaria: neglected and not benign. *Am J Trop Med Hyg.* 2007; 77:79–87. [PubMed: 18165478]
45. Louicharoen C, et al. Positively selected G6PD-Mahidol mutation reduces *Plasmodium vivax* density in Southeast Asians. *Science.* 2009; 326:1546–1549. [PubMed: 20007901]
46. Guerra CA, et al. The international limits and population at risk of *Plasmodium vivax* transmission in 2009. *PLoS Negl Trop Dis.* 2010; 4:e774. [PubMed: 20689816]
47. Liu W, et al. African origin of the malaria parasite *Plasmodium vivax*. *Nat Commun.* 2014; 5:3346. [PubMed: 24557500]
48. Miller LH, Mason SJ, Clyde DF, McGinniss MH. The resistance factor to *Plasmodium vivax* in blacks. The Duffy-blood-group genotype, FyFy. *N Engl J Med.* 1976; 295:302–304. [PubMed: 778616]
49. Hamblin MT, Thompson EE, Di Rienzo A. Complex signatures of natural selection at the Duffy blood group locus. *Am J Hum Genet.* 2002; 70:369–383. [PubMed: 11753822]
50. Apinjoh TO, et al. Association of cytokine and Toll-like receptor gene polymorphisms with severe malaria in three regions of Cameroon. *PLoS One.* 2013; 8:e81071. [PubMed: 24312262]
51. Manjurano A, et al. Candidate human genetic polymorphisms and severe malaria in a Tanzanian population. *PLoS One.* 2012; 7:e47463. [PubMed: 23144702]

52. Dunstan SJ, et al. Variation in human genes encoding adhesion and proinflammatory molecules are associated with severe malaria in the Vietnamese. *Genes and immunity*. 2012
53. Toure O, et al. Candidate polymorphisms and severe malaria in a Malian population. *PLoS One*. 2012; 7:e43987. [PubMed: 22957039]
54. Vallone PM, Decker AE, Butler JM. Allele frequencies for 70 autosomal SNP loci with U.S. Caucasian, African-American, and Hispanic samples. *Forensic Sci Int*. 2005; 149:279–286. [PubMed: 15749374]
55. Clayton D. Testing for association on the X chromosome. *Biostatistics*. 2008; 9:593–600. [PubMed: 18441336]
56. Cordell HJ. Detecting gene-gene interactions that underlie human diseases. *Nat Rev Genet*. 2009; 10:392–404. [PubMed: 19434077]

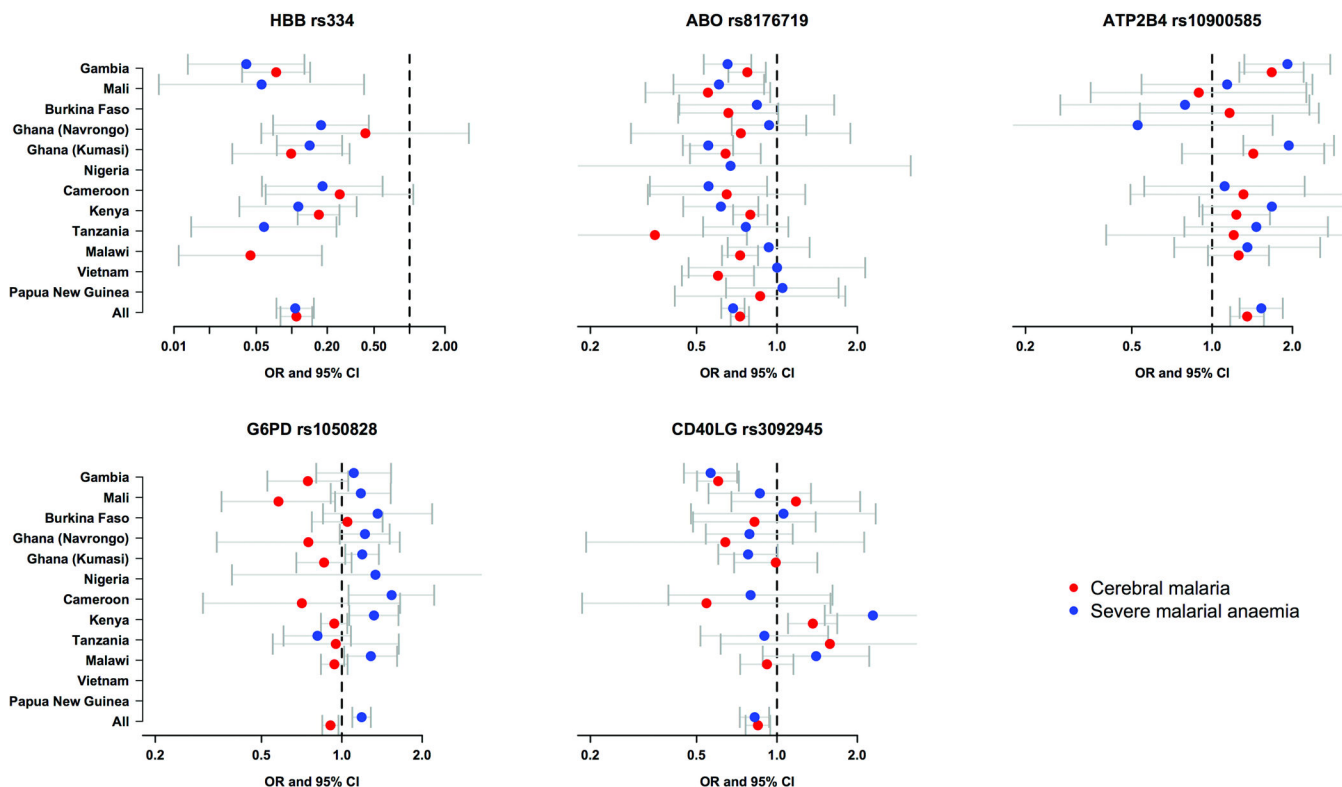


Figure 1. Forest plots for association with severe malaria and sub phenotypes
 Odds ratios and 95% confidence intervals (CI) are shown for Sickle-cell trait (rs334, heterozygote model), Blood group O (rs8176719, recessive model), ATP2B4 (rs10900585, dominant model), G6PD deficiency (rs1050828, additive model) and (e) CD40LG (rs3092945, recessive model) for association with cerebral malaria (lines with red dots) and severe malarial anaemia (lines with blue dots) in all individuals combined. Results are adjusted for gender, ethnicity and (with the exception of rs334) for sickle cell trait. Results are not presented when the sample size is too small (fewer than 5 cases or controls with relevant genotype) or for locations where the derived allele is absent. Further details are available in Supplementary Tables 11–19. The OR=1 for no effect is highlighted by the vertical dashed line.

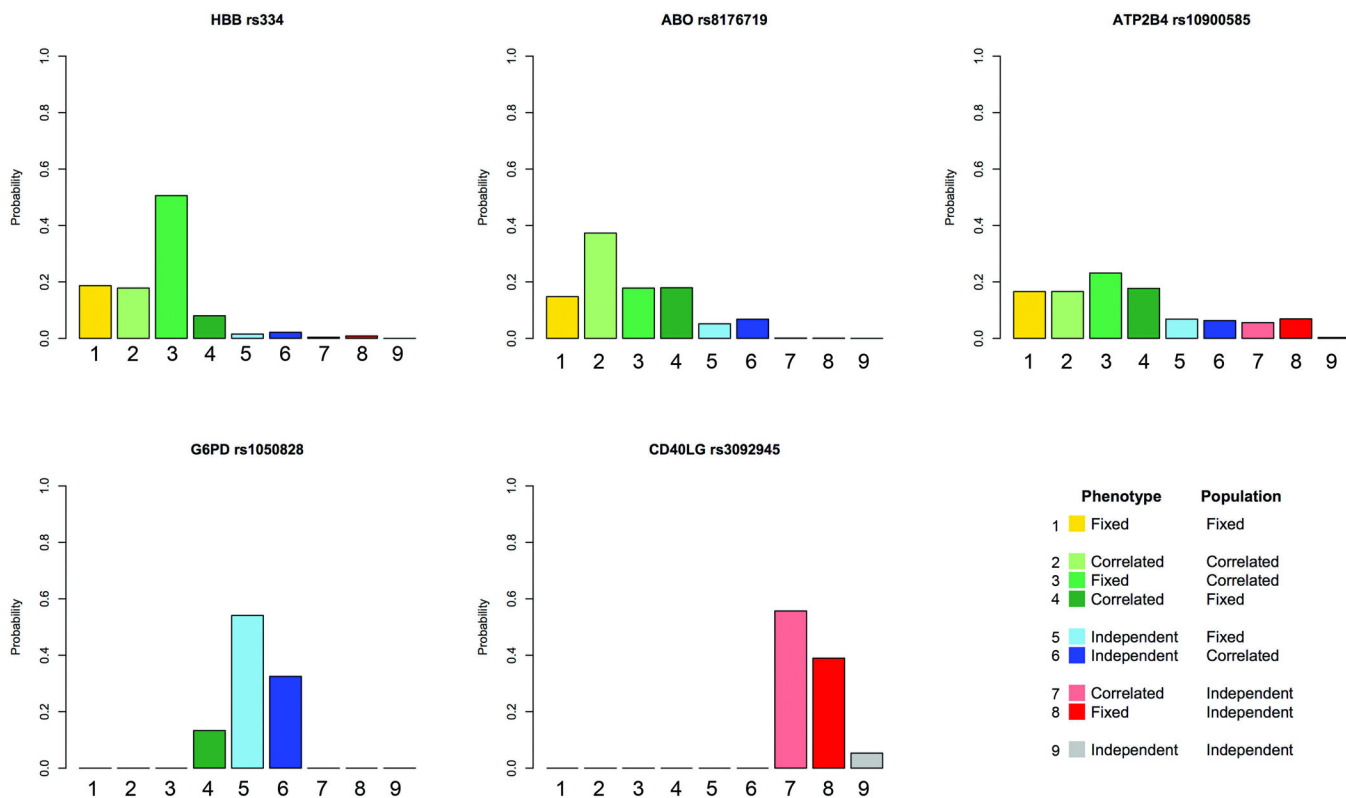


Figure 2. Genetic heterogeneity of severe malaria subtypes cerebral malaria only and severe malarial anaemia only within and across African sites for significant loci
 Bar plots show the distribution of probability between each of nine models of association where the effects on each phenotype are fixed, independent or correlated *within* a site combined with being fixed independent or correlated across all sites (see legend). Models are *a priori* assumed to be equally likely (See Online Methods and Supplementary Note for details). Results are shown for SNPs rs334 (HbS, heterozygote model) in HBB, rs8176719 homozygotes (Blood Group O) in ABO, rs10900585 in ATP2B4, G6PD+202 (rs1050828) in G6PD and rs3092945 in CD40LG.

Table 1

Clinical phenotype case counts and percentage of case fatalities.

Study site	Cases					Controls
	As Phenotypes					
	All Severe Malaria	Cerebral Malaria	Severe Malarial Anaemia	Cerebral Malaria and Severe Malarial Anaemia	Other Severe Malaria	
Gambia	2425 (14)	785 (26)	458 (4)	126 (26)	1056 (6)	3342
Mali	453 (14)	86 (26)	185 (8)	74 (20)	108 (25)	344
Burkina Faso	865 (5)	107 (20)	399 (10)	20 (20)	339 (4)	729
Ghana (Navrongo)	682 (4)	22 (27)	248 (2)	14 (36)	398 (3)	489
Ghana (Kumasi)	1496 (4)	230 (10)	551 (2)	75 (13)	640 (2)	2042
Nigeria	77 (4)	6 (17)	8 (12)	0 (0)	63 (2)	40
Cameroon	621 (5)	39 (19)	82 (7)	8 (50)	492 (4)	578
Kenya	2268 (11)	909 (14)	160 (8)	214 (16)	985 (7)	3949
Tanzania	429 (11)	34 (41)	182 (4)	28 (29)	185 (11)	453
Malawi	1388 (17)	873 (15)	132 (7)	166 (19)	217 (2)	2697
Vietnam	794 (11)	211 (18)	31 (6)	8 (12)	544 (9)	2538
Papua New Guinea	392 (2)	43 (5)	120 (2)	9 (11)	220 (0)	240
Total	11890(10)	3345(18)	2196(4)	742(20)	5607(5)	17441

Numbers are given for cases of all-severe-malaria by study site with percentage of case-fatalities in brackets. Cases are further divided into those with cerebral malaria only, severe malarial anaemia only, both cerebral malaria and severe malarial anaemia and other severe malaria.

Table 2
Autosomal single nucleotide polymorphisms (SNPs) with strong association signals.

Gene SNP	Alleles Ancestral /Derived	Frequency of the Derived Allele in Controls (Ancestral Hom/Het/ Derived Hom)	Case Phenotype	Frequency of the Derived Allele in Cases (Ancestral Hom/Het/ Derived Hom)	Model ^a	Model OR(95% CI)	Model P
ABO	rs8176719 I/D	0.69(1700/7215/8238)	Severe malaria	0.64(1506/5533/4750)	R	0.74(0.7–0.78)	4.99 × 10 ⁻³³
			Cerebral malaria only	0.64(415/1537/1373)	R	0.73(0.67–0.79)	8.85 × 10 ⁻¹⁶
			Severe malarial anaemia only	0.62(302/1054/816)	R	0.68(0.62–0.76)	7.97 × 10 ⁻¹⁴
rs8176746 C/A	0.17(11943/4852/565)	Severe malaria	0.2(7512/3817/507)	D	1.25(1.19–1.32)	2.01 × 10 ⁻¹⁷	
		Cerebral malaria only	0.2(2148/1049/138)	A	1.27(1.18–1.36)	2.00 × 10 ⁻¹¹	
		Severe malarial anaemia only	0.22(1342/746/97)	D	1.28(1.16–1.42)	1.71 × 10 ⁻⁰⁶	
A1P2B4	rs10900585 G/T	0.66(1644/5722/5737)	Severe malaria	0.68(868/4056/4203)	D	1.32(1.21–1.45)	1.69 × 10 ⁻⁰⁹
			Cerebral malaria only	0.69(245/1168/1236)	D	1.35(1.17–1.57)	3.06 × 10 ⁻⁰⁵
			Severe malarial anaemia only	0.69(159/816/851)	D	1.53(1.27–1.84)	3.68 × 10 ⁻⁰⁶
			Severe malaria	0.29(4558/3922/743)	R	0.75(0.68–0.83)	4.87 × 10 ⁻⁰⁹
			Cerebral malaria only	0.29(1334/1122/200)	R	0.7(0.59–0.82)	4.03 × 10 ⁻⁰⁶
			Severe malarial anaemia only	0.29(932/788/137)	R	0.67(0.55–0.82)	3.96 × 10 ⁻⁰⁵
HBB	rs334 A/T	0.07(12773/1791/77)	Severe malaria	0.02(10388/213/84)	H	0.14(0.12–0.16)	1.62 × 10 ⁻²²⁵
			Cerebral malaria only	0.01(3041/42/5)	H	0.11(0.08–0.15)	4.67 × 10 ⁻⁸⁸
			Severe malarial anaemia only	0.03(1965/31/45)	H	0.11(0.07–0.15)	9.25 × 10 ⁻⁶⁵
			Severe malaria	0.04(6866/445/46)	A	0.71(0.63–0.8)	6.87 × 10 ⁻⁰⁹
			Cerebral malaria only	0.02(1412/56/5)	A	0.72(0.56–0.94)	0.01
			Severe malarial anaemia only	0.04(1465/116/10)	A	0.74(0.6–0.9)	2.11 × 10 ⁻⁰³

Odds Ratios (OR), 95% Confidence Intervals (95% CI) and P-values (P) for the optimal genotypic model (Model) are presented for males and females combined. Results are adjusted for gender, ethnicity and (with the exception of rs334) for sickle cell trait. Study sites at which a SNP was monomorphic were excluded from analysis. Further models are shown in Tables S8–S10 and S11–S13.

Association P-values for all SNPs can be found in Supplementary Figure 3.

Hom, Homozygote; Het, Heterozygote;

^aModels are A, Additive; D, Dominant; H Heterozygote Advantage; R, Recessive.

Table 3

X-chromosome single nucleotide polymorphisms (SNPs) with strong association signals.

Gene	SNP	Alleles Ancestral /Derived	Sample	Frequency of the Derived Allele in Controls (Ancestral Hom/Het/ Derived Hom)	Case Phenotype	Frequency of the Derived Allele in Cases (Ancestral Hom/Het/ Derived Hom)	Model ^a	Model OR(95% CI)	Model P
CD40LG	rs3092945	T/C	All	0.27(11030/2621/3197)	Severe malaria	0.29(7123/2035/2250)	R	0.85(0.79–0.91)	1.11 × 10 ⁻⁰⁶
					Cerebral malaria only	0.27(2102/576/596)	R	0.85(0.76–0.94)	2.45 × 10 ⁻⁰³
					Severe malarial anaemia only	0.32(1197/411/452)	R	0.82(0.73–0.94)	2.97 × 10 ⁻⁰³
			F	0.27(4581/2621/849)	Severe malaria	0.3(2636/2035/513)	R	0.78(0.69–0.88)	8.93 × 10 ⁻⁰⁵
					Cerebral malaria only	0.29(813/576/157)	R	0.9(0.74–1.09)	0.27
					Severe malarial anaemia only	0.33(424/411/103)	R	0.71(0.56–0.9)	3.49 × 10 ⁻⁰³
			M	0.27(6449/0/2348)	Severe malaria	0.28(4487/0/1737)	M	0.9(0.83–0.98)	0.01
					Cerebral malaria only	0.25(1289/0/439)	M	0.85(0.75–0.97)	0.01
					Severe malarial anaemia only	0.31(773/0/349)	M	0.87(0.75–1.01)	0.07
G6PD	rs1050828	C/T	All	0.15(11552/1770/1279)	Severe malaria	0.15(8516/1152/1000)	A	1.02(0.97–1.06)	0.15
					Cerebral malaria only	0.13(2513/338/230)	A	0.91(0.85–0.97)	6.08 × 10 ⁻⁰³
					Severe malarial anaemia only	0.18(1563/215/263)	A	1.19(1.1–1.28)	2.62 × 10 ⁻⁰⁵
			F	0.15(5069/1770/174)	Severe malaria	0.14(3705/1152/134)	A	0.95(0.88–1.03)	0.23
							R	1.15(0.9–1.46)	0.27
							H	0.9(0.82–0.99)	0.02
					Cerebral malaria only	0.14(1129/338/39)	A	0.92(0.82–1.04)	0.19
							R	1.09(0.76–1.57)	0.65
							H	0.87(0.76–1.01)	0.06
			M	0.15(6483/0/1105)	Severe malarial anaemia only	0.16(669/215/41)	A	1.12(0.96–1.3)	0.15
							R	1.94(1.3–2.89)	1.92 × 10 ⁻⁰³
							H	0.93(0.77–1.11)	0.42
		Severe malaria	0.15(4811/0/866)	M	1.1(0.99–1.22)	0.07			
		Cerebral malaria only	0.12(1384/0/191)	M	0.81(0.68–0.96)	0.01			
		Severe malarial anaemia only	0.2(894/0/222)	M	1.49(1.24–1.79)	3.55 × 10 ⁻⁰⁵			

Odds Ratios (OR), 95% Confidence Intervals (95% CI) and P-values (P) for the optimal genetic model (Model). Results are adjusted for gender (all individuals combined), ethnicity and sickle cell trait. Study sites at which a SNP was monomorphic were excluded from analysis. Further models are shown in Tables S8–S10 and S14–S17.

Association P-values for all SNPs can be found in Supplementary Figure 3.

Hom, Homozygote; Het, Heterozygote.

^aModels are A, Additive; D, Dominant; H Heterozygote Advantage; M, Male Hemizygote; R, Recessive.