

Plasmodium falciparum merozoite invasion ligands, linked antimalarial resistance loci and ex vivo responses to antimalarials in The Gambia

Haddijatou Mbye¹, Karim Mane¹, Mouhamadou Fadel Diop¹, Martha Anita Demba¹, Fatoumata Bojang¹,
Nuredin Ibrahim Mohammed¹, David Jeffries¹, Neils Ben Quashie², Umberto D'Alessandro¹
and Alfred Amambua-Ngwa ^{1*}

¹Medical Research Council at London School of Hygiene and Tropical Medicine, Fajara, The Gambia; ²University of Ghana, Legon, Accra, Ghana

*Corresponding author. E-mail: angwa@mrc.gm

Received 11 January 2022; accepted 23 June 2022

Background: Artemether/lumefantrine is the most commonly used artemisinin-based combination treatment (ACT) for malaria in sub-Saharan Africa. Drug resistance to ACT components is a major threat to malaria elimination efforts. Therefore, rigorous monitoring of drug efficacy is required for adequate management of malaria and to sustain the effectiveness of ACTs.

Objectives: This study identified and described genomic loci that correlate with differences in ex vivo responses of natural *Plasmodium falciparum* isolates from The Gambia to antimalarial drugs.

Methods: Natural *P. falciparum* isolates from The Gambia were assayed for IC₅₀ responses to four antimalarial drugs (artemether, dihydroartemisinin, amodiaquine and lumefantrine). Genome-wide SNPs from 56 of these *P. falciparum* isolates were applied to mixed-model regression and network analyses to determine linked loci correlating with drug responses. Genomic regions of shared haplotypes and positive selection within and between Gambian and Cambodian *P. falciparum* isolates were mapped by identity-by-descent (IBD) analysis of 209 genomes.

Results: SNPs in 71 genes, mostly involved in stress and drug resistance mechanisms correlated with drug responses. Additionally, erythrocyte invasion and permeability loci, including merozoite surface proteins (Pfdblmsp, Pfsurfin), and high-molecular-weight rhoptry protein 2 (Pfrhops2) were correlated with responses to multiple drugs. Haplotypes of *pfdblmsp2* and known drug resistance loci (*pfaat1*, *pfcr1* and *pfdhfr*) from The Gambia showed high IBD with those from Cambodia, indicating co-ancestry, with significant linkage disequilibrium between their alleles.

Conclusions: Multiple linked genic loci correlating with drug response phenotypes suggest a genomic backbone may be under selection by antimalarials. This calls for further analysis of molecular pathways to drug resistance in African *P. falciparum*.

Introduction

The malaria parasite *Plasmodium falciparum* has adapted to most previous antimalarial drugs.¹ Current treatment for uncomplicated malaria therefore requires artemisinin-based combination therapies (ACTs), combining a fast-acting artemisinin derivative, which is rapidly eliminated, with a long-acting partner drug.² However, resistance to artemisinin derivatives and to some partner drugs has been reported in South East Asia (SEA).³ In sub-Saharan Africa, ACTs remain highly efficacious, although parasites with mutations associated with artemisinin resistance have independently emerged in East Africa, and the

frequency of genetic variants associated with increased tolerance to partner drugs is also increasing.^{4–7} Therefore, monitoring the prevalence of known genetic resistance markers and the clinical efficacy of currently used ACTs is important to ensure national antimalarial treatment policies are updated to ascertain adequate responses to any change in treatment efficacy.

Genomic surveillance of antimalarial resistance markers has mainly focused on SEA, where mutations in the kelch 13 (*pfk13*) locus associated with artemisinin resistance and delayed parasite clearance were first identified.⁸ Nevertheless, these mutations remain rare in Africa, though the recent observation of the *pfk13* R561H mutant in Rwandan isolates and increased clearance time

for isolates with A675V and C469Y variants in Uganda suggest resistance may emerge independently in Africa.^{9,10} In addition, mutations in other genes such as *pfap2-mu* (I568T) have been identified in ACT-tolerant *P. falciparum*.¹¹ These mutations increase *in vitro* ring survival against artemisinin derivatives but are not associated with delayed parasite clearance in natural infections.¹² Moreover, copy number variation in *P. falciparum* plasmepsins II and III, associated with piperazine resistance, have been identified across Africa, possibly questioning the large-scale use of dihydroartemisinin/piperazine,¹³ although this combination drug is the second-line treatment in some countries like The Gambia. Like other malaria-endemic countries in Africa, The Gambia has previously used most approved antimalarial drugs, including quinine since the late nineteenth century until the 1920s, followed by chloroquine until it was used in combination with sulfadoxine/pyrimethamine from 1981 to 2008. From 2008 to date, The Gambia has relied on artemether/lumefantrine as the official first-line antimalarial ACT, though other combinations can be accessed via commercial vendors of pharmaceutical products. Chemoprevention with amodiaquine and sulfadoxine/pyrimethamine has been implemented for over 8 years in some parts of the country, overall increasing drug pressure on the local malaria parasite population.

In The Gambia, drug-tolerant *P. falciparum* isolates have been identified from *ex vivo* studies in the western region.¹⁴ While ACTs remain highly efficacious, strong selection signatures at drug resistance loci and recent directional selection in temporal populations have been reported from genome analysis of recent *P. falciparum* isolates.¹⁵ Indeed, between 2008 and 2014, the prevalence of *pfmdr1* WT variants increased, and mutations in cysteine desulfurase (*pfncs1* 65Q mutant) and chloroquine resistance transporter (*pfcr1*) differentiated in the low-transmission peri-urban setting of Brikama.¹⁵ Here we assess the correlation between genomic variants and *in vitro/ex vivo* responses of Gambian parasite isolates to current antimalarial drugs. Combining mixed-model regression, network and population genetics analyses, we identified several interacting genetic loci in pathways that could determine differences in susceptibility to antimalarials.

Methods

Samples and populations

P. falciparum-infected blood samples used in this study were a subsample of those collected for a different study from the same population.¹⁶ These were collected during the 2014 and 2015 malaria transmission seasons. Ethical approval was obtained from The Gambia Government/The Medical Research Council Unit The Gambia at the London School of Hygiene & Tropical Medicine (MRCG at LSHTM) Joint Ethics Committee. At Brikama Health Centre in western Gambia, patients with uncomplicated malaria and a parasite density of ≥ 1000 parasites/ μL were invited to participate in the study. Following written informed consent, a 2 mL venous blood sample was collected into EDTA tubes, and 50 μL aliquots spotted on Whatman filter paper for dried blood spots. Whole-genome sequences of *P. falciparum* isolates were generated in collaboration with the MalariaGEN consortium at the Wellcome Sanger Institute. Paired-end reads were aligned to the *P. falciparum* 3D7 reference genome (v3) and genomic variants (SNPs and small indels) retained as previously described.^{15,17} Freely available Gambian genome data from 2008 (1006-PF-GM-CONWAY) and those from Cambodian populations were obtained from previous publications and the recent Pf6 data release (<https://wellcomeopenresearch.org/articles/16-42/v2>).

Ex vivo drug susceptibility assay

Parasite-infected erythrocytes were separated from leucocytes and assayed for 50% inhibitory concentration (IC₅₀) within 4 h of blood sample collection. Each isolate was tested against artemether, dihydroartemisinin, lumefantrine and amodiaquine in duplicate, at 0.5% parasitaemia and 2% haematocrit, as previously reported.¹⁶ Following a full cycle of growth with drugs, erythrocytes were counterstained for parasite DNA using 1:10 000 SYBR Green I in PBS. Stained cells were thereafter washed twice and resuspended with 100 μL of PBS followed by acquisition of 100 000 events on the Accuri flow cytometer (BD). Dd2 and 3D7 laboratory-adapted *P. falciparum* strains were used as controls. Data analysis was done using Prism (GraphPad Prism version 8.4.3).

Genome sequences

The genomes of 269 Gambian *P. falciparum* isolates (56 with high-quality IC₅₀ data) were obtained from the Pf6 MalariaGEN data release. An initial set of 16 390 PASS bi-allelic SNPs, supported by RMS mapping quality (MQ) ≥ 20 , read depth (DP) ≥ 5 and variant quality score log-odds (VQSLOD) ≥ 2 , were retained following previously described processing.^{15,18} Following filtration for missingness and variant call quality as above, the final dataset for analysis included 10 776 SNPs from 168 Gambian isolates, 96 sampled in 2008 and 72 sampled during 2014/15. The same SNPs were extracted for 87 monoclonal isolates (inbreeding coefficient, $F_{\text{ws}} > 0.95$) from Cambodia.

Genotype versus ex vivo drug susceptibility association analysis

For 56 Gambian isolates with IC₅₀ values obtained in 2014/15, we retained 9291 SNPs. Genotype–phenotype association was first carried out by applying the suite of mixed regression models in the Genetic Association and Prediction Integrated Tool (GAPIT) package (RStudio v1.2) (<https://zzlab.net/>). Secondly, we used parametric *t*-test to compare the mean IC₅₀ difference between SNP alleles. The derived *P* values were adjusted for multiple comparisons using the Bonferroni method. Thirdly, the R tool for network-based GWAS analysis, i.e. network-based Genome-Wide Association Studies (netGWAS), was used to build an SNP–drug phenotype network map to model lumefantrine and artemether responses using only SNP loci within regions of selective sweeps. netGWAS was used to reconstruct linkage disequilibrium (LD) networks from multilocus genotype data and to detect high-dimensional genotype–phenotype networks using graphical models. Network plots of relationships between phenotypes and genotypes were visualized using iGraph in R. To determine whether loci are associated with the response to multiple drugs, we ran MultiPhen in R, which jointly tested multiple phenotypes for association using ordinal regression in which SNPs are treated as the outcome and multiple phenotypes as the predictors. We applied a cut-off *P* value of 0.001 to identify outlier associated loci.

Positive selection and relatedness in *P. falciparum* populations

To determine genome-wide signatures of selection and shared haplotypes that may be affecting drug response, we scanned for regions of identity-by-descent (IBD) within and between populations, generating the *iR* index using the isoRelate package in R.¹⁹ Correction for multiple testing was applied on the resultant *P* values based on the Benjamin–Hochberg procedure with a 5% threshold for false discovery rate (FDR). Outlier SNP loci were extracted and their ontology inferred from PlasmoDB annotations. For selected loci of interest, multiple sequence alignments of haplotypes were exported from VCF files to fasta format using custom bash scripts and vk phylo. The haplotype distribution between populations was assessed using a TCS tree constructed with popART.

The genome data for isolates with *ex vivo* phenotypes were further reduced to SNPs around known drug resistance markers (*pfprt*, *pfmdr1*, *pfdhfr*, *pfdhps*, *pfpgch* and *pfk13*) and inter-chromosomal LD between drug resistance sweep SNPs calculated with VCFtools version 0.1.13 (<https://vcftools.github.io/documenta>

tion.html). These were compared with LD between randomly chosen loci outside drug resistance sweep regions for each pair of drug resistance loci on different chromosomes. All other scripts used in these analyses are freely available via <https://github.com/Karim-Mane/rSNPdata>.

Table 1. Summary of IC₅₀ (nM) of 56 samples assayed against each antimalarial tested in the 2014/15 transmission season

	Amodiaquine	Lumefantrine	Dihydroartemisinin	Artemether
First quartile	2.34	14.19	0.42	1.44
Third quartile	7.61	110.17	2.78	5
Median	4.58	44.94	1.3	2.76
Mean	10.02	88.77	2.87	4.4
Range	0.36–58.52	2.38–405.00	0.18–29.83	0.42–32.75
95% CI	6.16–13.90	61.10–116.45	1.57–4.12	3.03–5.76
3D7 control	45.38	65.12	2.37	8.76

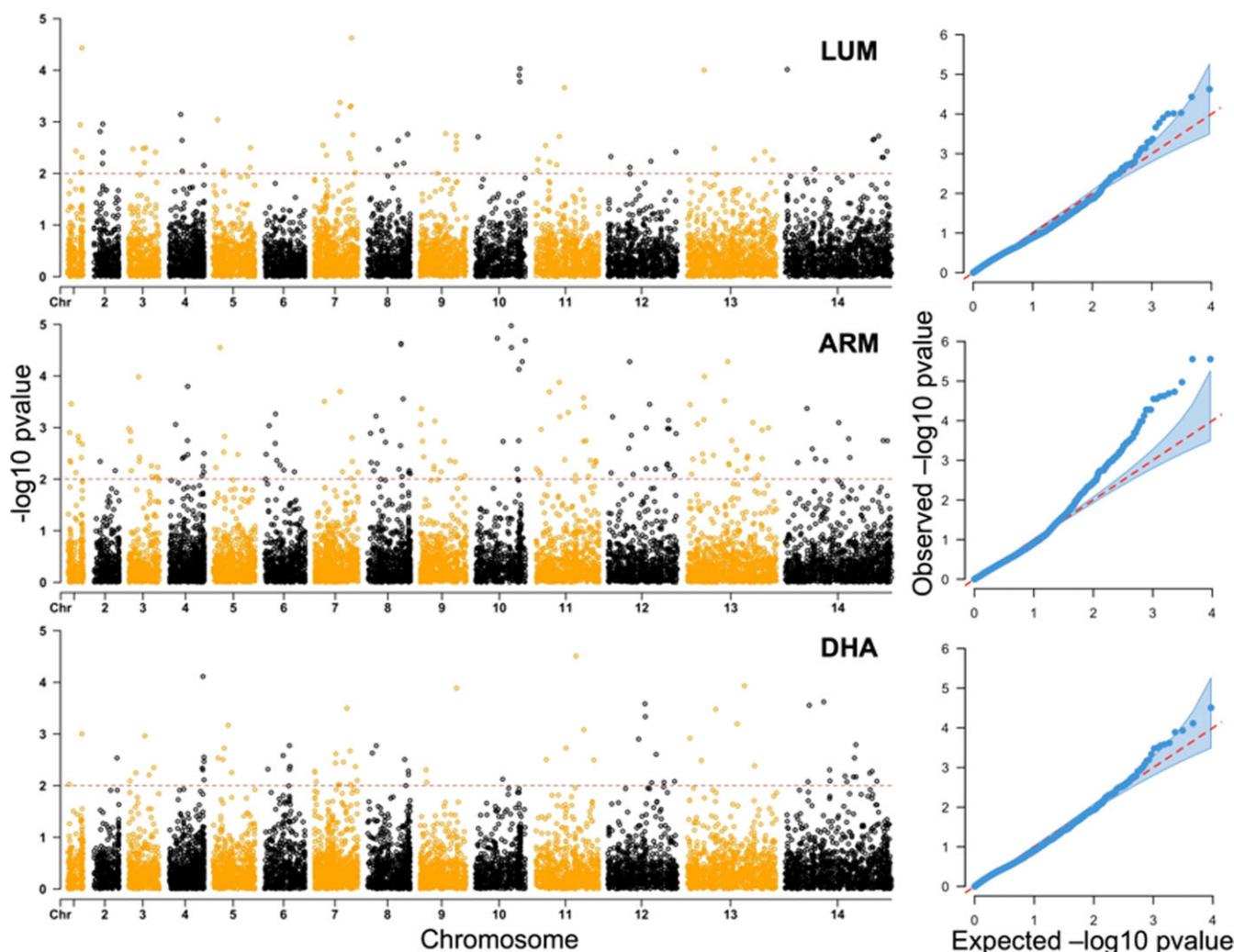


Figure 1. Association between genomic SNPs and Gambian *P. falciparum* IC₅₀ to antimalarial drugs from MLM. The left panel shows Manhattan plots of *P* values for association of each SNP with responses to lumefantrine (LUM) in the top row, artemether (ARM) in the middle row and dihydroartemisinin (DHA) in the bottom row. The right panel shows the corresponding QQ plot for each drug, indicating the deviation of the observed *P* values for SNPs against each drug as against random expectations. This figure appears in colour in the online version of *JAC* and in black and white in the print version of *JAC*.

Table 2. Effect sizes of top 30 genic loci associated with *ex vivo* responses of Gambian *P. falciparum* isolates to lumefantrine (LUM), artemether (ARM) and dihydroartemisinin (DHA)

Gene ID	Genomic location	Product description	Overall effect ^a	Drug
Pf3D7_1035700	Pf3D7_10_v3:1412641...1416363(+)	DBL merozoite surface protein (DBLMSP)	263.33	LUM
Pf3D7_0113800	Pf3D7_01_v3:526530...536562(+)	DBL-containing protein, unknown function	158.35	LUM
Pf3D7_1402800	Pf3D7_14_v3:100564...106014(+)	conserved <i>Plasmodium</i> protein, unknown function	149.48	LUM
Pf3D7_1037200	Pf3D7_10_v3:1473514...1476860(+)	conserved <i>Plasmodium</i> protein, unknown function	137.39	LUM
Pf3D7_0718800	Pf3D7_07_v3:833630...835051(-)	conserved protein, unknown function	130.31	LUM
Pf3D7_0719300	Pf3D7_07_v3:843566...848302(+)	actin-related protein (ARP6)	125.18	LUM
Pf3D7_0112700	Pf3D7_01_v3:477279...481382(+)	28S ribosomal RNA	103.86	LUM
Pf3D7_1345200	Pf3D7_13_v3:1811227...1813968(-)	rhomboid protease ROM6, putative	96.13	LUM
Pf3D7_0405000	Pf3D7_04_v3:268919...271537(-)	ATP-dependent RNA helicase DDX51, putative	95.07	LUM
Pf3D7_0409000	Pf3D7_04_v3:427283...435517(-)	conserved <i>Plasmodium</i> protein, unknown function	95.07	LUM
Pf3D7_1139300	Pf3D7_11_v3:1556447...1565469(-)	AP2 domain transcription factor, putative (ApiAP2)	89.79	LUM
Pf3D7_0728100	Pf3D7_07_v3:1193347...1210448(-)	anaphase-promoting complex subunit 1, putative	75.38	LUM
Pf3D7_0727000	Pf3D7_07_v3:1147264...1151617(-)	vacuolar protein sorting-associated protein 53, putative (VPS53)	73.90	LUM
Pf3D7_0717600	Pf3D7_07_v3:758490...763360(-)	inner membrane complex protein, putative (IMC32)	60.32	LUM
Pf3D7_0209000	Pf3D7_02_v3:369351...380156(+)	6-cysteine protein (P230)	59.14	LUM
Pf3D7_0504800	Pf3D7_05_v3:187586...205020(-)	conserved <i>Plasmodium</i> protein, unknown function	56.75	LUM
Pf3D7_1313800	Pf3D7_13_v3:588141...599917(-)	conserved <i>Plasmodium</i> membrane protein, unknown function	17.12	ARM
Pf3D7_0506100	Pf3D7_05_v3:256330...259207(-)	60S ribosomal subunit protein L24, putative	12.64	ARM
Pf3D7_1034500	Pf3D7_10_v3:1369717...1376708(+)	armadillo repeat protein, putative	12.64	ARM
Pf3D7_0824900	Pf3D7_08_v3:1079441...1083413(-)	conserved <i>Plasmodium</i> protein, unknown function	11.32	ARM
Pf3D7_1322300	Pf3D7_13_v3:944138...949434(-)	BRCT domain-containing protein, putative	10.33	ARM
Pf3D7_0826200	Pf3D7_08_v3:1142499...1144972(-)	alpha/beta hydrolase, putative	10.02	ARM
Pf3D7_1121300	Pf3D7_11_v3:801809...808198(-)	tyrosine kinase-like protein (TKL2)	9.43	ARM
Pf3D7_0308100	Pf3D7_03_v3:355867...361936(+)	zinc finger protein, putative	9.15	ARM
Pf3D7_1113000	Pf3D7_11_v3:491069...510781(+)	conserved <i>Plasmodium</i> protein, unknown function	8.89	DHA
Pf3D7_1018000	Pf3D7_10_v3:718371...720178(+)	tRNA pseudouridine synthase, putative	8.69	ARM
Pf3D7_1228900	Pf3D7_12_v3:1180404...1185971(-)	conserved protein, unknown function	8.57	ARM
Pf3D7_1303300	Pf3D7_13_v3:162566...164742(+)	conserved <i>Plasmodium</i> protein, unknown function	8.46	ARM
Pf3D7_1027300	Pf3D7_10_v3:1139669...1144013(-)	peroxiredoxin	8.42	ARM
Pf3D7_1039000	Pf3D7_10_v3:1567784...1572695(+)	serine/threonine protein kinase, FIKK family	8.36	ARM

^aEffect size projects the magnitude to which the phenotype is affected by the variant. Positive effects imply the minor allele is favourable to higher IC₅₀ values.

Table 3. Genes with SNPs correlating with multiple phenotypes from MultiPhen analysis

Gene ID	Genomic location (gene)	Product description	Cumulative log ₁₀ P value
Pf3D7_0424400	Pf3D7_04_v3:1099303...1107445(+)	surface-associated interspersed protein 4.2 (SURFIN 4.2)	3.98
Pf3D7_0709700	Pf3D7_07_v3:434748...436518(-)	prodrug activation and resistance esterase (PARE)	3.19
Pf3D7_0929400	Pf3D7_09_v3:1173579...1181380(+)	high molecular weight rhoptry protein 2 (RhopH2)	6.76
Pf3D7_0930300	Pf3D7_09_v3:1201305...1207576(+)	merozoite surface protein 1 (MSP1)	4.24
Pf3D7_0935900	Pf3D7_09_v3:1419960...1424840(-)	ring-exported protein 1 (Rex)	3.98

Results

SNPs in stress response and invasion genes associated with increased IC₅₀ for dihydroartemisinin and lumefantrine

The IC₅₀ values against amodiaquine, lumefantrine, dihydroartemisinin and artemether for the 56 isolates retained were mostly low. Relatively lower values were obtained for artemisinin derivatives and amodiaquine compared with lumefantrine (Table 1).

No SNP attained an FDR-corrected significance threshold ($-\log_{10}P > 6$) for association with IC₅₀. However, several SNPs were identified as outliers from multilevel model regression (MLM) analysis, deviating from expectation (Figure 1). The least significant outcomes were against amodiaquine, with *P* values for SNPs mostly less significant than expected (Figure S1, available as [Supplementary data](#) at JAC Online). Considering an adjusted *P* value threshold of at least $-\log_{10}P > 3$, we identified 76 SNP loci in 71 genes having relatively higher correlation with *ex*

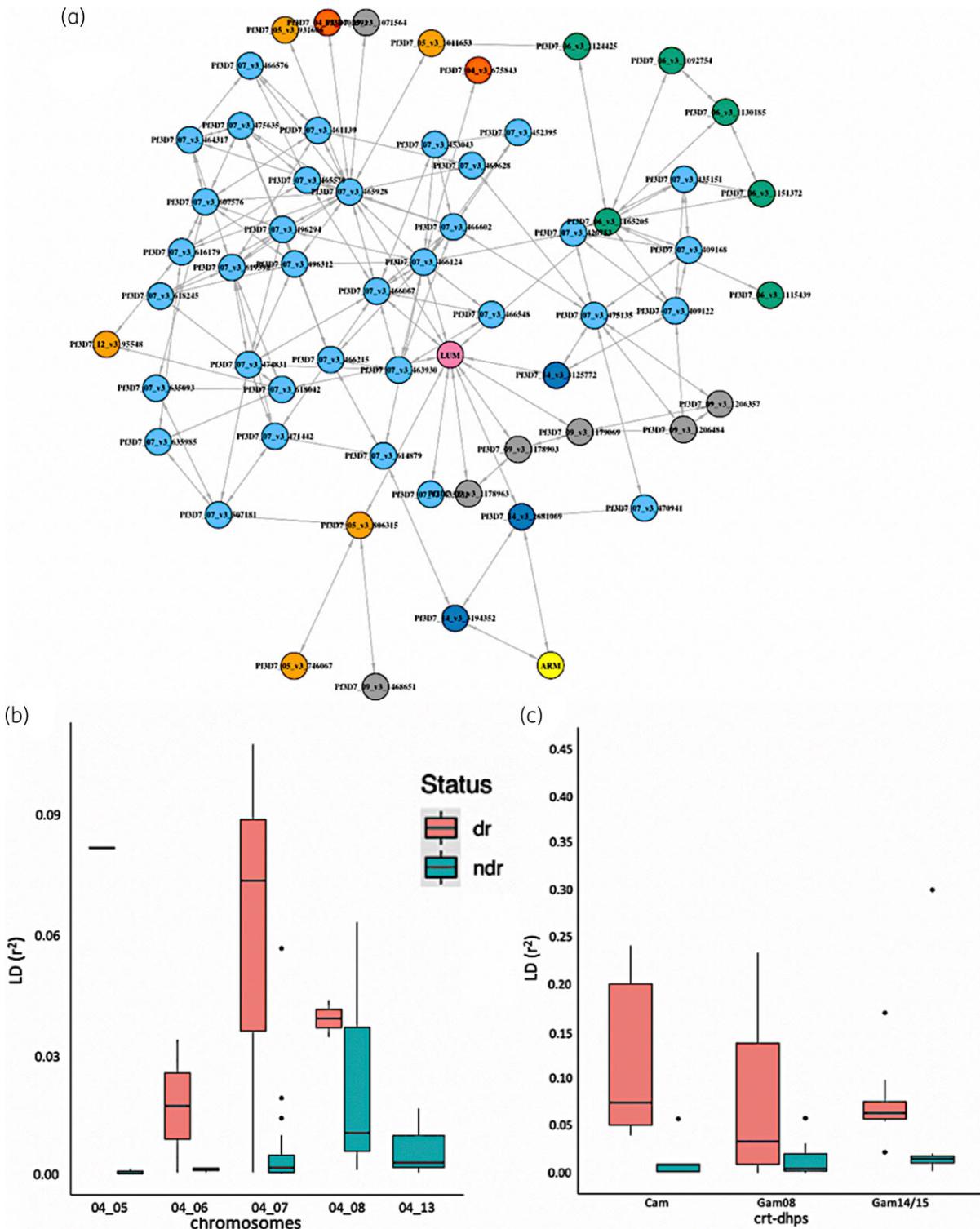


Figure 2. LD between SNPs and association with lumefantrine (LUM) and artemether (ARM) *ex vivo* responses. (a) Significant LD and genotype–phenotype interaction network for LUM (pink node) and ARM (yellow node). Dark orange, orange, green, blue, grey, mustard and dark-blue nodes represent chromosomes 4, 5, 6, 7, 9, 12 and 14, respectively. Each node is labelled with the SNP position. (b) Boxplots of LD between SNPs in or around drug loci (dr) and non-drug loci (ndr) across pairs of chromosomes as indicated on the x-axis. (c) LD between *pfert* and *pfdhps* variants for Cambodia (Cam), Gambia 2008 (Gam08) and 2014/2015 (Gam14/15). This figure appears in colour in the online version of JAC and in black and white in the print version of JAC.

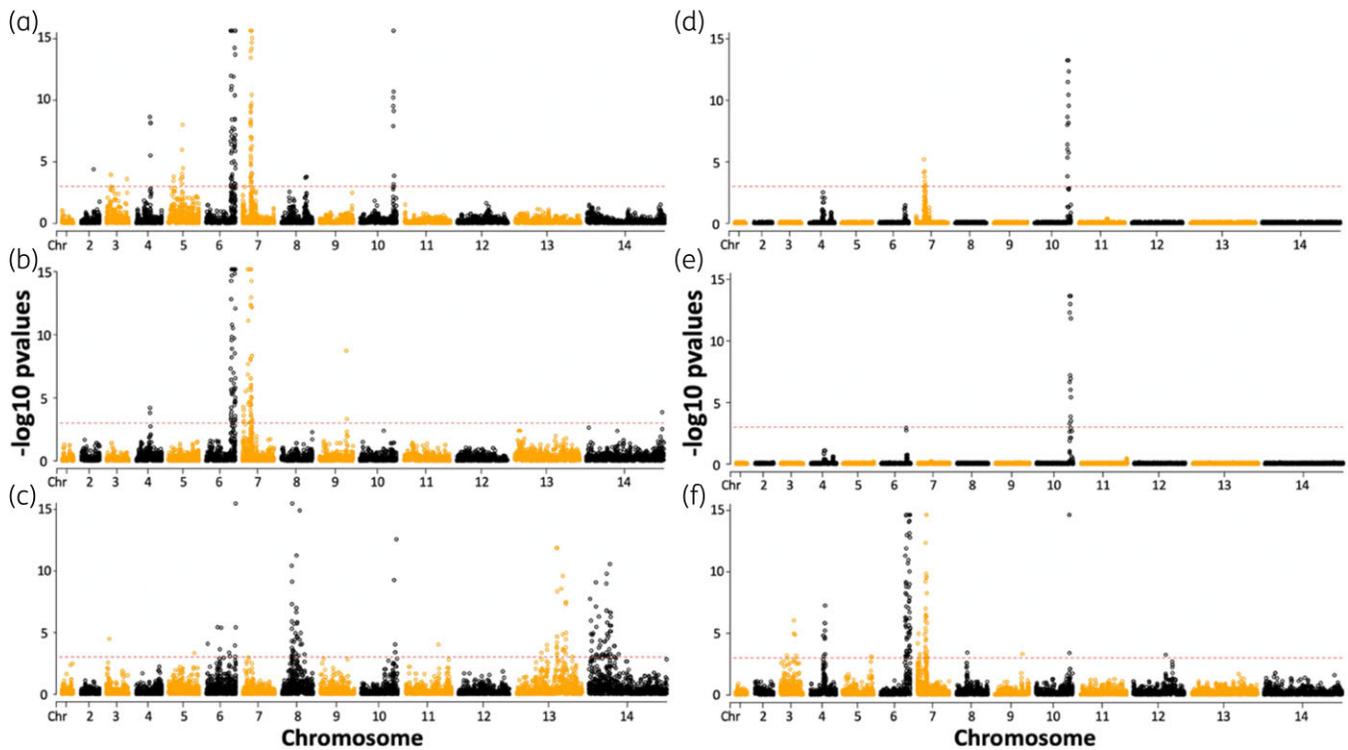


Figure 3. Manhattan plots of IBD-based signatures of positive selection across the genome of Gambian and Cambodian *P. falciparum* isolates. The y-axis of each plot shows the $-\log_{10}$ of corrected *P* values of iR for each SNP. The left panel shows IBD for pairs of isolates within populations: (a) Gambia 2008; (b) Gambia 2014/15; and (c) Cambodia. The right panel shows IBD between isolates from different populations: (d) Gambia 2008 against Cambodia; (e) Gambia 2014/15 against Cambodia; and (f) Gambia 2008 against Gambia 2014/15. This figure appears in colour in the online version of *JAC* and in black and white in the print version of *JAC*.

vivo responses to drugs (Table S1). These loci correlating with relatively higher tolerance had minor allele frequencies ranging from 4% to 31%. The highest effect sizes were for responses to lumefantrine, though the strongest deviation of expected *P* values was seen for artemether (Table 2).

The strongest effect was for an SNP locus in *pfdblmsp* located within the *pfmsp* gene cluster on chromosome 10, which includes *pfdblmsp2*. Associated SNPs were also found in several conserved proteins of unknown function, with the most significant semantic clusters for the biological functions of the top 71 genes being ‘response to chemical stimuli’, ‘response to stress’ and ‘protein ubiquitination’ (Figure S2).

At a $-\log_{10}P$ of at least 3, MultiPhen analysis identified alleles in six non-synonymous SNPs in five genes on chromosomes 4, 7 and 9 as predicting multiple drug responses (Table 3). This included: *pfrrhoph2*, which is part of the cell permeability complex; *pfmsp1* and *pfsurfin4.2*, which are involved in erythrocyte invasion mechanisms; *pfrex1*, which is associated with the Maurer’s cleft; and prodrug activation esterase (*pfpare*), an esterase with a drug activation function.

From parametric *t*-test comparing the mean differences in responses between the alleles at each SNP locus, a total of 16, 27, 59 and 13 SNPs were significantly associated with parasite IC_{50} levels to artemether, dihydroartemisinin, lumefantrine and amodiaquine, respectively, with a higher significance threshold of $-\log_{10}P$ of at least 5 (Figure S3). Of these, 56 loci had a minor

allele frequency of over 5%, 26 of which affected responses to lumefantrine (Table S2). The 56 genic loci identified include antigens (Hyp, EBA-175, schizont egress protein, Surfin14.2) and several metabolic and housekeeping genes. The most significant gene ontologies include ‘chromosome segregation’, ‘sos response’, ‘cell cycle’ and ‘growth’ (Figure S4).

Linked loci across drug resistance sweeps associated with *ex vivo* artemether/lumefantrine responses

Genotype–phenotype network analysis identified several SNPs clustered together with drug susceptibility phenotypes to artemether and lumefantrine, the components of artemether/lumefantrine. Lumefantrine had direct interactions with 12 SNPs found on chromosomes 4, 5, 7, 9 and 14. Half of these loci were SNPs around the chloroquine resistance transporter (*pfCRT*), while three were from *rrhoph2* (PF3D7_0929400) (Figure 2a). Artemether was directly associated with two SNPs in PF3D7_1465800 (dynein beta chain, putative) and PF3D7_1477600 [surface-associated interspersed protein 14.1 (*surfin14.1*)] coded for by chromosome 14. Most of these drug response network SNPs were in LD across chromosomes even when not directly associated with drug phenotypes (Figure S5). In the Gambian population, LD was strong within known drug resistance genes but also between them, with *pfDHPS* and *pfGCH* on chromosomes 8 and 12, respectively, having the highest number of linked SNPs (Figure S6). SNPs around

Table 4. Top genes with high IBD (iR index >4) between population pairs within The Gambia (Gam14_15/Gam08) and between Gambian and Cambodian *P. falciparum* isolates (Gam08/Cam and Gam14_15/Cam)

Gene ID	Product description	Gam14_15/Gam08	Gam08/Cam	Gam14_15/Cam
PF3D7_0414000	structural maintenance of chromosomes protein 3	4.8	1	0.2
PF3D7_0415100	conserved protein, unknown function	5.8	1.7	0.2
PF3D7_0415200	conserved <i>Plasmodium</i> protein, unknown function	15.7	0.9	0.3
PF3D7_0627700	transportin	12.4	1.5	3.4
PF3D7_0627800	acetyl-CoA synthetase, putative (ACS)	38.5	1	3.8
PF3D7_0628100	HECT domain-containing protein 1, putative (HECT1)	126.8	6.2	4
PF3D7_0629700	SET domain protein, putative	106.9	1.2	0.6
PF3D7_0630300	DNA polymerase epsilon catalytic subunit A, putative	13.1	0	0
PF3D7_0630700	bifunctional methylenetetrahydrofolate dehydrogenase/cyclohydrolase	18.3	0.2	0.1
PF3D7_0709100	Cg1 protein	14.7	1.8	0.2
PF3D7_0709300	mediator of RNA polymerase II transcription subunit 14	55.2	5.3	0.4
PF3D7_0709700	prodrug activation and resistance esterase (PARE)	4.5	2.2	0.1
PF3D7_0709900	hydrolase, putative	27.4	6.5	0.1
PF3D7_0710000	conserved <i>Plasmodium</i> protein, unknown function	21	0.3	0.2
PF3D7_0710200	conserved <i>Plasmodium</i> protein, unknown function	15.6	9.9	0.2
PF3D7_1036000	merozoite surface protein 11 (MSP11)	2	21.2	27.3
PF3D7_1036300	DBL merozoite surface protein 2 (DBLMSP2)	20.7	340.9	280.7
PF3D7_1036500	conserved <i>Plasmodium</i> protein, unknown function	29.3	26.5	27.3
PF3D7_1036800	acetyl-CoA transporter, putative (ACT)	37.4	64.5	38.8
PF3D7_1036900	conserved <i>Plasmodium</i> protein, unknown function	1.9	26.5	27.3
PF3D7_1037000	DNA polymerase zeta catalytic subunit, putative (REV3)	2.1	13.2	6
PF3D7_1037100	pyruvate kinase 2	1.3	10.4	6.9
PF3D7_1037400	conserved <i>Plasmodium</i> protein, unknown function	2.7	21.9	7.6

drug resistance genes were also in relatively stronger LD, even when located in different chromosomes (Figure 2b and 2c, Figure S7). Across all populations, strong LD was seen between *pfcr* and *pfdhps* on chromosomes 7 and 8, respectively (Figure 2c).

Strong IBD at drug resistance-associated loci in Gambia and Cambodia

At 5% FDR, 139, 115 and 142 loci showed significant IBD in Gambian 2008, 2014/15 and Cambodian populations, respectively (Table S3). Pairwise population analysis also identified 124 SNPs with excess IBD between temporal Gambian populations, 58 SNPs between Gambia 2008 and Cambodia, and 52 loci between Gambia 2014/15 and Cambodia. The IBD index, iR, located these SNPs in segments of positive selection that were stronger within and between Gambian temporal populations than against Cambodia (Figure 3). In the Gambian populations, these were prominent around drug resistance-associated loci, including *pfdhfr* on chromosome 4, *pfmdr1* on chromosome 5, *pfat1* on chromosome 6, *pfcr* on chromosome 7, *pfdhps* on chromosome 8 and a cluster on chromosome 10 that encodes the Duffy binding-like (DBL) domain containing proteins including Pfdblmsp and Pfdblmsp2. IBD between Gambian and Cambodian isolates was prominent around *pfdhfr*, *pfat1*, *pfcr* and *pfdblmsp*. The region around *pfdblmsp2* on chromosome 10 was the dominant region of shared IBD and positive selection between Gambian and Cambodian isolates, spanning several genes (Table 4). The *pfdblmsp2* region had 88 SNPs with minor allele frequencies of at least 5%. Haplotype clustering with these loci

across Gambian and Cambodian populations showed 32 haplotypes with at least two identical sequences. Half of these shared haplotypes were present only in the Gambian population, while one was unique to Cambodia, with six isolates (Figure 4). There were four major haplotype clades shared between Gambian and Cambodian populations. A fifth clade (clade 5) was predominantly Gambian, with only a single isolate from Cambodia.

Discussion

To further understand the relationship between SNPs in the *P. falciparum* genome and the parasite's responses to current antimalarials, we determined the *ex vivo* susceptibility of Gambian natural *P. falciparum* isolates to artemisinin derivatives (artemether and dihydroartemisinin) and two partner drugs, lumefantrine and amodiaquine. Artemether/lumefantrine is the first-line antimalarial ACT in The Gambia and has been used for over a decade. Overall, IC₅₀ levels to these drugs were associated with SNPs in many genes, which included known drug response, resistance, stress response and antigenic invasion loci, as has been previously described in similar studies.²⁰⁻²² The results of the current study were obtained by combining mixed-model regression methods, simple parametric *t*-test and population genetic analyses, as the small number of samples did not allow for power for strong genome-wide association approaches.

The main antigenic invasion locus associated with differences in antimalarial susceptibility was the DBL merozoite surface protein (*pfdblmsp2*) on chromosome 10. We considered three aspects as indicators of its possible correlation with antimalarial

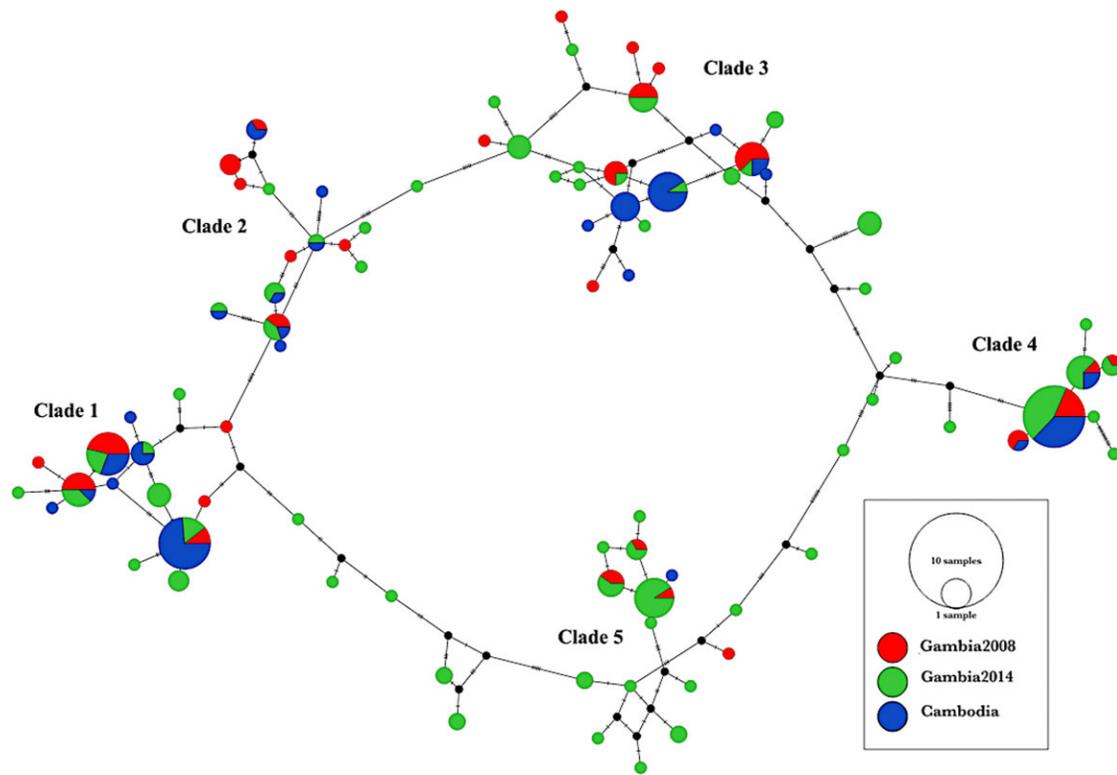


Figure 4. TCS haplotype network of DBLMSP2 amongst Gambian and Cambodian isolates. Each pie chart shows the proportion of haplotypes from each population: Gambia 2008 (red); Gambia 2014/15 (green); and Cambodia (blue). The black edges are unsampled putative haplotypes not represented in the dataset. This figure appears in colour in the online version of JAC and in black and white in the print version of JAC.

resistance. Firstly, like other drug resistance loci (*pfprt*, *pfdfhr*, *pfdhps* and *pfmdr1*), the genomic region around this gene had strong IBD and shared haplotypes between Gambian and Cambodian populations. Drug resistance is generally thought to have migrated from SEA to Africa, with introgression of drug resistance loci from SEA. Secondly, *pfdblmsp2* showed significant signatures of positive selection, which could be due to drug selection as seen for other drug resistance loci. Thirdly, our regression and network analyses showed a correlation between the locus and *ex vivo* responses of isolates. This association between *pfdblmsp2* and drug response has been previously reported, with increased gene copies associated with reduced *in vitro* susceptibility to halofantrine, a phenanthrene compound like lumefantrine, mefloquine and quinine.²³ *pfdblmsp2* and other DBL genes are involved in the cell invasion process,²⁴ and the results here raise questions on the role invasion mechanisms and pathways play in antimalarial drug resistance. Erythrocyte invasion by *P. falciparum* merozoites is a complex process and the ligands involved are also targeted by balancing selection from host immune responses. Indeed, other invasion/surface-associated loci (*surfin4.2*, *rex* and *msp1*) were associated with combined responses to multiple drugs. Merozoite invasion has been considered as a viable target for the development of new antimalarial drugs.^{25,26} Association between drug response and variants at invasion ligands, Surfin1 and Clag3, have been reported before.²⁰ These antigens bind to IgM antibodies and this may interfere with effective IgG immunoreactivity and immune responses

that accompany parasite clearance by drugs *in vivo*.^{27,28} It is therefore possible that some alleles in these antigens could be affecting invasion and immunity, thereby limiting the effect of drugs.

Across Africa, loci associated with antimalarial resistance have been linked with importation from SEA, with shared haplotypes.²⁹ An extended region of LD with reduced recombination has been previously reported around *pfdblmsp2*, with two highly divergent haplotypes.²⁸ Here, we found five major haplotype clades of *pfdblmsp2*, four of which were common between Cambodia and Gambia. This aligned with increased relatedness and LD around this locus and at other drug resistance-associated loci. Hence, the background of drug resistance imported from SEA could have included haplotypes of *pfdblmsp2* as well as chloroquine resistance markers (*pfprt*) and antifolate resistance markers (*pfdfhr*), which show signatures of positive selection and stronger inter-chromosomal LD compared with the genome background. This LD between drug resistance loci across the genome has been shown before for isolates from the Thailand–Cambodia border, defining the backbone of artemisinin and multidrug resistance.¹⁷ Alongside the standard use of ACTs, malaria chemoprevention combining antifolates and quinolines is ongoing in regions of seasonal malaria chemoprevention and antifolates are used for prevention of malaria in pregnancy. These could be sustaining the high frequencies of mutant variants at antifolate markers (*pfdfhr* and *pfdhps*), which are almost fixed in our population.¹⁴ Therefore, most other antimalarial resistance markers

will be evolving on the background of these antifolate markers. A small number of multi-locus LD across drug resistance loci was also recently reported in Ghana.³⁰ Like in SEA, response to drugs seems to be driven by multiple loci, presenting the risk of multi-drug resistance, which needs systematic and rigorous monitoring as multiple antimalarial prophylaxes and treatments are implemented.

MultiPhen also identified a PARE, where gene mutations have been reported to cause resistance to a lead antimalarial drug candidate, benzoxaborole (AN13762), due to blocked compound activation and altered ubiquitination or sumoylation.³¹ Ubiquitination has been implicated in the pathway of resistance to artemisinin derivatives.³² Cell death from artemisinin derivatives follows perturbation in protein folding, inhibition of the proteasome stress response process and accumulation of polyubiquitinated proteins.³³ Alteration of ubiquitination and proteasome processes via PARE could alter drug response. Other candidate-associated loci, *pfrex2* and *pfroph2*, are part of the protein exportome. With the exception of significant association between rho-try-associated membrane antigen (RAMA) and the resistant variant of the chloroquine resistance transporter, Pfcrt, rho-try proteins have not been implicated in drug response.³⁴ However, Rho-try is part of the cell permeability complex and is functional in nutrient transport with effects on cell metabolism and growth.³⁵

From this small set of *ex vivo* drug susceptibility phenotypes and genomes, we identified putative loci that may be associated with *P. falciparum* responses to antimalarials. These included several SNPs in antigenic loci and exported proteins, which could be directly implicated or linked to causal variants, modulating the invasion process or interactions with drugs. With relatively higher IBD around these loci and higher LD across drug resistance loci, the evolution and functional role of linked loci and invasion ligands in drug resistance development needs to be further investigated. Despite these findings, this study remains limited by the small samples phenotyped and seasonal variation. While the IC₅₀ metric has been used for both efficacy screening and drug discovery, its high variability and poor reproducibility could confound phenotype–genotype association analyses. Hence, validation of the identified SNPs in drug resistance pathways will require larger samples with innovative approaches to normalize drug response phenotypes. As current and new drug combinations are deployed for elimination of malaria, well-powered genotype–phenotype analysis combining both *in vitro/ex vivo* and *in vivo* responses could further our understanding of drug resistance and the identification of markers for the prediction and monitoring of drug interventions against malaria.

Acknowledgements

We thank the participants for consenting to this study. We acknowledge all the members of the malaria population biology team at MRCG at LSHTM and Mr Simon Corea for facilitating sampling and maintenance of the malaria culture labs. Genomic sequences were freely accessed online and through long-standing collaboration with MalariaGEN.

Funding

This work was part of the PhD studies of Mrs Haddijatou Mbye, funded partially by the West African Centre for Cell Biology of Infectious

Pathogens (WACCBIP). WACCBIP is partly supported through the Developing Excellence and Leadership, Training and Science (DELTA) Africa programme (grant number 107755/Z/15/Z). The views expressed in this publication are those of the author(s) and not necessarily those of The African Academy of Sciences (AAS), the AUDA-NEPAD Agency, Wellcome Trust or the UK government. Dr Amambua-Ngwa's group supported the studies with funds from the Human Hereditary and Health in Africa (H3Africa) project and the Pan-African Malaria Genetic Epidemiology Network (PAMGEN, grant number H3A/18/002) funded through the AAS.

Transparency declarations

None to declare.

Supplementary data

Figures S1 to S7 and Tables S1 to S3 are available as [Supplementary data](#) at JAC Online.

References

- Huijben S, Paaijmans KP. Putting evolution in elimination: winning our ongoing battle with evolving malaria mosquitoes and parasites. *Evol Appl* 2018; **11**: 415–30.
- Nosten F, White NJ. Artemisinin-based combination treatment of falciparum malaria. 2007. <https://www.ncbi.nlm.nih.gov/books/NBK1713/>.
- Hamilton WL, Amato R, van der Pluijm RW *et al*. Evolution and expansion of multidrug-resistant malaria in southeast Asia: a genomic epidemiology study. *Lancet Infect Dis* 2019; **19**: 943–51.
- Baliraine FN, Rosenthal PJ. Prolonged selection of *pfmdr1* polymorphisms after treatment of falciparum malaria with artemether-lumefantrine in Uganda. *J Infect Dis* 2011; **204**: 1120–4.
- Eastman RT, Dharia NV, Winzeler EA *et al*. Piperaquine resistance is associated with a copy number variation on chromosome 5 in drug-pressured *Plasmodium falciparum* parasites. *Antimicrob Agents Chemother* 2011; **55**: 3908–16.
- Dama S, Niangaly H, Ouattara A *et al*. Reduced *ex vivo* susceptibility of *Plasmodium falciparum* after oral artemether–lumefantrine treatment in Mali. *Malar J* 2017; **16**: 59.
- Ebohon O, Irabor F, Ebohon LO *et al*. Therapeutic failure after regimen with artemether-lumefantrine combination therapy: a report of three cases in Benin City, Nigeria. *Rev Soc Bras Med Trop* 2019; **52**: e20190163.
- Ariey F, Witkowski B, Amaratunga C *et al*. A molecular marker of artemisinin-resistant *Plasmodium falciparum* malaria. *Nature* 2014; **505**: 50–5.
- Uwimana A, Legrand E, Stokes BH *et al*. Emergence and clonal expansion of *in vitro* artemisinin-resistant *Plasmodium falciparum* kelch13 R561H mutant parasites in Rwanda. *Nat Med* 2020; **26**: 1602–8.
- Balikagala B, Fukuda N, Ikeda M *et al*. Evidence of artemisinin-resistant malaria in Africa. *N Engl J Med* 2021; **385**: 1163–71.
- Henriques G, Hallett RL, Beshir KB *et al*. Directional selection at the *pfmdr1*, *pfprt*, *pfubp1*, and *pfap2mu* loci of *Plasmodium falciparum* in Kenyan children treated with ACT. *J Infect Dis* 2014; **210**: 2001.
- Henrici RC, van Schalkwyk DA, Sutherland CJ. Modification of *pfap2mu* and *pfubp1* markedly reduces ring-stage susceptibility of *Plasmodium falciparum* to artemisinin *in vitro*. *Antimicrob Agents Chemother* 2019; **64**: e01542-19.

- 13** Leroy D, Macintyre F, Adoke Y et al. African isolates show a high proportion of multiple copies of the *Plasmodium falciparum* plasmepsin-2 gene, a piperaquine resistance marker. *Malar J* 2019; **18**: 126.
- 14** Mbye H, Bojang F, Jawara AS et al. Tolerance of Gambian *Plasmodium falciparum* to dihydroartemisinin and lumefantrine detected by ex vivo parasite survival rate assay. *Antimicrob Agents Chemother* 2020; **65**: e00720-20.
- 15** Amambua-Ngwa A, Jeffries D, Amato R et al. Consistent signatures of selection from genomic analysis of pairs of temporal and spatial *Plasmodium falciparum* populations from The Gambia. *Sci Rep* 2018; **8**: 9687.
- 16** Amambua-Ngwa A, Okebe J, Mbye H et al. Sustained ex vivo susceptibility of *Plasmodium falciparum* to artemisinin derivatives but increasing tolerance to artemisinin combination therapy partner quinolines in The Gambia. *Antimicrob Agents Chemother* 2017; **61**: e00759-17.
- 17** Miotto O, Amato R, Ashley EA et al. Genetic architecture of artemisinin-resistant *Plasmodium falciparum*. *Nat Genet* 2015; **47**: 226-34.
- 18** Amambua-Ngwa A, Amenga-Etego L, Kamau E et al. Major subpopulations of *Plasmodium falciparum* in sub-Saharan Africa. *Science* 2019; **365**: 813-6.
- 19** Henden L, Lee S, Mueller I et al. Identity-by-descent analyses for measuring population dynamics and selection in recombining pathogens. *PLoS Genet* 2018; **14**: e1007279.
- 20** Mu J, Myers RA, Jiang H et al. *Plasmodium falciparum* genome-wide scans for positive selection, recombination hot spots and resistance to antimalarial drugs. *Nat Genet* 2010; **42**: 268.
- 21** Park DJ, Lukens AK, Neafsey DE et al. Sequence-based association and selection scans identify drug resistance loci in the *Plasmodium falciparum* malaria parasite. *Proc Natl Acad Sci U S A* 2012; **109**: 13052-7.
- 22** Wang Z, Cabrera M, Yang J et al. Genome-wide association analysis identifies genetic loci associated with resistance to multiple antimalarials in *Plasmodium falciparum* from China-Myanmar border. *Sci Rep* 2016; **6**: ytac083.
- 23** Van Tyne D, Uboldi AD, Healer J et al. Modulation of PF10-0355 (MSPDL2) alters *Plasmodium falciparum* response to antimalarial drugs. *Antimicrob Agents Chemother* 2013; **57**: 2937-41.
- 24** Hodder AN, Czabotar PE, Uboldi AD et al. Insights into Duffy binding-like domains through the crystal structure and function of the merozoite surface protein MSPDL2 from *Plasmodium falciparum*. *J Biol Chem* 2012; **287**: 32922-39.
- 25** Dans MG, Weiss GE, Wilson DW et al. Screening the Medicines for Malaria Venture Pathogen Box for invasion and egress inhibitors of the blood stage of *Plasmodium falciparum* reveals several inhibitory compounds. *Int J Parasitol* 2020; **50**: 235-52.
- 26** Wilson DW, Goodman CD, Sleebbs BE et al. Macrolides rapidly inhibit red blood cell invasion by the human malaria parasite, *Plasmodium falciparum*. *BMC Biol* 2015; **13**: 52.
- 27** Amambua-Ngwa A, Tetteh KKA, Manske M et al. Population genomic scan for candidate signatures of balancing selection to guide antigen characterization in malaria parasites. *PLoS Genet* 2012; **8**: e1002992.
- 28** Crosnier C, Iqbal Z, Knuepfer E et al. Binding of *Plasmodium falciparum* merozoite surface proteins DBLMSP and DBLMSP2 to human immunoglobulin M is conserved among broadly diverged sequence variants. *J Biol Chem* 2016; **291**: 14285-99.
- 29** Pearce RJ, Pota H, Evehe M-SB et al. Multiple origins and regional dispersal of resistant *dhps* in African *Plasmodium falciparum* malaria. *PLoS Med* 2009; **6**: e1000055.
- 30** Narh CA, Ghansah A, Duffy MF et al. Evolution of antimalarial drug resistance markers in the reservoir of *Plasmodium falciparum* infections in the Upper East Region of Ghana. *J Infect Dis* 2020; **222**: 1692-701.
- 31** Sindhe KMV, Wu W, Legac J et al. *Plasmodium falciparum* resistance to a lead benzoxaborole due to blocked compound activation and altered ubiquitination or sumoylation. *mBio* 2020; **11**: e02640-19.
- 32** Tilley L, Straimer J, Gnädig NF et al. Artemisinin action and resistance in *Plasmodium falciparum*. *Trends Parasitol* 2016; **32**: 682.
- 33** Bridgford JL, Xie SC, Cobbold SA et al. Artemisinin kills malaria parasites by damaging proteins and inhibiting the proteasome. *Nat Commun* 2018; **9**: 3801.
- 34** Sanchez CP, Liu CH, Mayer S et al. A HECT ubiquitin-protein ligase as a novel candidate gene for altered quinine and quinidine responses in *Plasmodium falciparum*. *PLoS Genet* 2014; **10**: e1004382.
- 35** Counihan NA, Chisholm SA, Bullen HE et al. *Plasmodium falciparum* parasites deploy RhopH2 into the host erythrocyte to obtain nutrients, grow and replicate. *Elife* 2017; **6**: e23217.