2003

Multiple Transporters Associated with Malaria Parasite Responses to Chloroquine and Quinine

Jianbing Mu

Michael T. Ferdig

Xiaorong Feng

Deirdre A. Joy

Junhui Duan

See next page for additional authors

Original Publication Citation

Mu, J. B., Ferdig, M. T., Feng, X. R., Joy, D. A., Duan, J. H., Furuya, T., . . . Su, X. Z. (2003). Multiple transporters associated with malaria parasite responses to chloroquine and quinine. *Molecular Microbiology, 49*(4), 977-989. doi:10.1046/j.1365-2958.2003.03627.x

Molecular Microbiology (2003) **49**(4), 977–989 doi:10.1046/j.1365-2958.2003.03627.x

Multiple transporters associated with malaria parasite responses to chloroquine and quinine

Jianbing Mu,[1‡] Michael T. Ferdig,[2‡] Xiaorong Feng,[1†] Deirdre A. Joy,[1] Junhui Duan,[1] Tetsuya Furuya,[1] G. Subramanian,[3] L. Aravind,[4] Roland A. Cooper,[5] John C. Wootton,[4] Momiao Xiong[6] and Xin-zhuan Su[1]*

[1]*Laboratory of Malaria and Vector Research, National Institutes of Health, Bethesda, MD, USA.*
[2]*Department of Biological Sciences, University of Notre Dame, Notre Dame, IN, USA.*
[3]*Human Genome Sciences, 9410 Key West Avenue, Rockville, MD, USA.*
[4]*National Center for Biotechnology Information, National Institutes of Health, Bethesda, MD, USA.*
[5]*Department of Biological Sciences, Old Dominion University, Norfolk, VA, USA.*
[6]*Human Genetics Center, University of Texas, Houston, TX, USA.*

Summary

Mutations and/or overexpression of various transporters are known to confer drug resistance in a variety of organisms. In the malaria parasite *Plasmodium falciparum*, a homologue of P-glycoprotein, PfMDR1, has been implicated in responses to chloroquine (CQ), quinine (QN) and other drugs, and a putative transporter, PfCRT, was recently demonstrated to be the key molecule in CQ resistance. However, other unknown molecules are probably involved, as different parasite clones carrying the same *pfcrt* and *pfmdr1* alleles show a wide range of quantitative responses to CQ and QN. Such molecules may contribute to increasing incidences of QN treatment failure, the molecular basis of which is not understood. To identify additional genes involved in parasite CQ and QN responses, we assayed the *in vitro* susceptibilities of 97 culture-adapted cloned isolates to CQ and QN and searched for single nucleotide polymorphisms (SNPs) in DNA encoding 49 putative transporters (total 113 kb) and in 39 housekeeping genes that acted as negative controls. SNPs in 11 of the putative transporter genes, including *pfcrt* and *pfmdr1*, showed significant associations with decreased sensitivity to CQ and/or QN in *P. falciparum*. Significant linkage disequilibria within and between these genes were also detected, suggesting interactions among the transporter genes. This study provides specific leads for better understanding of complex drug resistances in malaria parasites.

Accepted 19 May, 2003. *For correspondence. E-mail xsu@niaid.nih.gov; Tel. (+1) 301 402 0876; Fax (+1) 301 402 0079. †Present address: Department of Biochemistry and Molecular Biology, Pennsylvania State University College of Medicine, 500 University Drive, H171, Hershey, PA 17003, USA. ‡These authors contributed equally to this study.

Introduction

The human malaria parasite *Plasmodium falciparum* kills an estimated 1.1–2.7 million people each year (WHO, 2000), partly as a result of the emergence of parasites resistant to a wide array of antimalarial drugs (Peters, 1990; Thimasarn *et al*., 1997). Chloroquine (CQ), one of the most effective and affordable drugs over the past 50 years, has become useless in many parts of the world. Quinine (QN), another quinoline-based drug, has been effective against malaria parasites for centuries, but clinical failures have been reported from Asia and South America (Peters, 1987; Giboda and Denis, 1988; Pukrittayakamee *et al*., 1994). QN is structurally related to CQ, yet parasites resistant to CQ can be highly sensitive to QN and vice versa (Peters, 1987; Malin and Hall, 1990; Wongsrichanalai *et al*., 1997). Although mutations in two genes, *pfcrt* and *pfmdr1*, have been shown to confer resistance and modulate the response to CQ, respectively (Fidock *et al*., 2000; Reed *et al*., 2000; Djimde *et al*., 2001; Cooper *et al*., 2002; Sidhu *et al*., 2002), the mechanism of decreasing QN susceptibility is not clear. Five loci, however, have recently been implicated in QN response in a genetic cross in which inheritance of QN and CQ responses are correlated (M. T. Ferdig *et al*., unpublished).

It is well known that mutations and/or overexpression of certain transporters, especially transporters of the ABC superfamily, can confer drug resistance in organisms from bacteria to human cancer cell lines (Ouellette *et al*., 1994; Chakraborti *et al*., 1999; Allen *et al*., 2000). In human cancer cells, overexpression of P-glycoprotein and multidrug resistance-associated protein MRP1 are the principal causes of chemotherapy failure (Dean *et al*., 2001). Simultaneous changes in multiple glutamate chloride ion channels have been shown to be responsible for ivermectin resistance in nematodes (Blackhall *et al*., 1998; Dent

et al., 2000). In *P. falciparum*, *pfcrt* is a putative transporter with 10 transmembrane domains (Fidock *et al*., 2000). Two members of the ABC transporter superfamily from the malaria parasite *P. facilarum*, *pfmdr1* and *pfmdr2*, have also been implicated in responses to CQ and QN as well as other drugs (Wilson *et al*., 1989; 1993; Foote *et al*., 1990; Volkman and Wirth, 1998; Reed *et al*., 2000; Adagut and Warhurst, 2001; Babiker *et al*., 2001). Although the parasite has a limited number of transporters compared with yeast and other organisms, several additional ABC transporters are present in the genome (Gardner *et al*., 2002). Given the fact that most parasites carrying the same *pfcrt* and *pfmdr1* alleles often respond to CQ and QN differently (Fidock *et al*., 2000; Djimde *et al*., 2001; Cooper *et al*., 2002), it is clear that additional molecules, probably transporters or ion channels, play a role in modulating or contributing to drug responses.

We hypothesized that parasite transporters play an important role in the *P. falciparum* response to antimalarial drugs, including CQ and QN, and that the levels of the responses result from additive and/or interacting contributions of multiple proteins. Genetic variations and/or changes in gene expression in different parasite transporters should therefore affect the parasite response to antimalarial drugs. The *P. falciparum* genome sequencing project provides an excellent opportunity to study the potential contributions of transporters to drug resistance. The majority of the parasite genes are available in genome databases (Gardner *et al*., 2002). A comprehensive search of the parasite genome for genetic changes in putative transporters provides a unique approach to identify candidate transporters involved in drug transport or otherwise contributing to drug resistance.

To identify genes contributing to QN sensitivity as well as genes that may modulate parasite response to CQ, we collected single nucleotide polymorphisms (SNPs) from 49 genes that encode predicted or known transporters and transport regulatory proteins available from public databases. We then genotyped a total of 97 culture-adapted isolates (34 from Africa, 42 from Asia, 16 from the Americas and five from Papua New Guinea) and measured the dose–responses to CQ and QN *in vitro*. We show that SNPs from multiple transporters are associated with elevated levels of the parasite response to CQ and/or QN and provide evidence of co-selection of SNPs from the associated genes, supported by linkage disequilibra (LD) between genes on different chromosomes.

Results and discussion

SNPs from 49 putative transporter genes

We searched the *P. falciparum* genome databases at the websites of the genome sequencing consortium (websites of TIGR, Sanger Center and Stanford University) for protein motifs similar to known transporters. A total of 113 kb of DNA, 99 kb coding and 14 kb non-coding, containing 49 putative transporter sequences, was amplified from four parasite isolates (Hb3 of central America, Dd2 of south-east Asia, D10 of Papua New Guinea and 7G8 of South America) and sequenced. SNPs and polymorphic microsatellite (MS) sites were identified after alignment of the DNA sequences from five isolates (including 3D7 from the genome sequencing project). Two hundred and thirty-one polymorphic sites, including 67 MS and 164 SNPs, were obtained from 42 of the 49 genes (Table 1). Of the SNPs, 130 are in coding regions (cSNPs) and 34 are in non-coding regions. This gives an overall frequency of one SNP per 690 bp DNA (nucleotide polymorphism, $\theta = 7.3 \times 10^{-4}$), one SNP per 764 bp in coding regions ($\theta = 6.3 \times 10^{-4}$) and one SNP per 412 bp in non-coding regions ($\theta = 1.2 \times 10^{-3}$) respectively. Among the 130 cSNPs, 96 are non-synonymous substitutions (74%) and 34 are synonymous. MS are mostly in introns and are present at a frequency of one in ≈1.7 kb. These results confirm that the *P. falciparum* genome is highly polymorphic when parasites from around the world with different CQ selection histories are scored, with a polymorphic site occurring every 0.49 kb DNA in just five isolates, which is consistent with that reported for chromosome 3 (Mu *et al*., 2002).

The SNPs are not evenly distributed among the genes. There are genes that contain no SNPs, possibly reflecting functional constraints (Table 1). However, 14 genes are quite variable, having three or more non-synonymous substitutions, but few or no synonymous changes. This pattern suggests the possibility of positive selection, although the high frequency of non-synonymous substitutions partly reflects the paucity of synonymous sites (Mu *et al*., 2002).

Parasite responses to CQ and QN

To investigate whether SNPs from the transporter genes are associated with drug sensitivity, we determined *in vitro* IC_{50}s to both CQ and QN for 97 cloned isolates (Table 2). Plots of descending CQ and QN IC_{50} values showed continuous distributions (Fig. 1A and B) characteristic of multigenic, quantitative traits. One obvious gap in the distribution of CQ IC_{50} (dashed line in Fig. 1A) can be attributed to mutations in *pfcrt*: all the parasites below this interval carry a wild-type *pfcrt* allele, whereas parasites above this gap have mutant alleles (data not shown). Significantly, isolates carrying an identical *pfcrt* mutant allele displayed a wide range of IC_{50} values, indicating the participation of additional genes in modulating the level of response to CQ. A similar, but smoother continuous distribution of QN IC_{50}

Table 1. Summary of genes encoding 49 putative transporters and single nucleotide polymorphisms (SNPs).

Gene Code	Chrom. Location	PlasmoDB link (4.0)	Name	Coding (kb)	Non-coding (kb)	Total Seq (kb)	ncSNP	cSNP		Total SNP	Frequency (SNP/kb)	Nucleo. Poly. (θ)	MS	Total Poly.	Codon position			Nucleotide substitution	
								sSNP	nsSNP						1st	2nd	3rd	Transition	Transversion
G2	1[a]	PFA0590w	ABC transporter	3.6	0	3.6	0	0	3	3	0.8	0.00040	0	3	3	0	0	2	1
G3	14[a]	chr14.glm_514	ABC transporter (MDR2)	2.3	0	2.3	0	2	4	6	2.6	0.00125	1	7	1	4	1	5	1
G4	3	PFC0125w	ABC transporter	2.2	0	2.2	0	5	4	9	4.1	0.00196	1	10	3	1	5	5	4
G5	11	PF11–0225	ABC transporter	1.5	0	1.5	0	0	0	0	0	0.00000	0	0	0	0	0	0	0
G7	13[a]	PF13–0271	ABC transporter	2.0	0.2	2.2	1	1	1	3	1.4	0.00065	2	5	1	0	1	1	1
G8	11[a]	PF11–0466	ABC transporter	0.9	0	0.9	0	0	5	5	5.6	0.00267	0	5	3	1	1	1	4
G10	14	PF14–0321	ABC transporter	0.7	0	0.7	0	0	0	0	0	0.00000	1	1	0	0	0	0	0
G13	13	PF13–0019	NHE1	2.9	0	2.9	0	1	1	2	0.7	0.00033	2	4	1	0	1	1	1
G14	12[a]	PFL0685w	Multiple TM protein	1.5	1.2	2.7	0	0	1	1	0.4	0.00018	1	2	0	1	0	1	0
G15	3	PFC0160w	Ankyrin repeat protein	1.5	0	1.5	0	1	0	1	0.7	0.00032	1	2	0	0	1	0	1
G16	2[a]	PFB0210c	Monosaccharide transporter	1.5	0	1.5	0	3	0	3	2.0	0.00096	0	3	0	0	3	3	0
G17	2	PFB0465c	Putative transporter	1.2	0.8	2.0	0	0	0	0	0	0.00000	2	2	0	0	0	0	0
G18	2	PFB0535w	Multiple TM protein	0.6	0.7	1.3	1	1	1	3	2.3	0.00111	2	5	1	0	1	1	1
G20	3	PFC0875w	ATPase	2.2	0	2.2	0	1	3	4	1.8	0.00087	1	5	1	1	2	2	2
G21	14[a]	PF14–0622	Calcium–potassium channel	1.3	1.4	2.7	4	0	0	4	1.5	0.00071	1	5	0	0	0	0	0
G22	12[a]	PFL1315w	Potassium channel	2.5	0	2.5	0	0	3	3	1.2	0.00058	1	4	0	3	0	2	1
G23	11[a]	PF11–0338	Aquaporin	0.8	0.7	1.5	4	0	3	7	4.7	0.00224	2	9	1	2	0	3	0
G24	5	PFE1185w	Cation transporter	2.2	0.5	2.7	2	1	1	4	1.5	0.00071	3	7	1	0	1	1	1
G25	14[a]	PF14–0679	Sulphate permease	2.1	0.6	2.7	4	1	1	6	2.2	0.00107	5	11	0	1	1	1	1
G26	13[a]	PF13–0079	Multiple TM + TPR	3.9	0.3	4.2	3	1	2	6	1.4	0.00069	0	6	1	1	1	2	1
G28	2	PFB0770c	Multiple TM protein	2.2	0	2.2	2	0	0	2	0.9	0.00044	6	8	0	0	0	0	0
G29	2	PFB0580w	Multiple TM protein	2.0	0	2.0	1	1	0	2	1.0	0.00048	1	3	0	0	1	0	1
G30	14[a]	PF14–0292	GTPase	1.2	1	2.2	1	0	0	1	0.5	0.00022	0	1	0	0	0	0	0
G33	14[a]	PF14–0273	Multiple TM protein	2.6	0	2.6	3	1	1	5	1.9	0.00092	6	11	1	0	1	2	0
G34	12	PFL0765w	Multiple TM MFS family	1.4	0.5	1.9	2	1	1	4	2.1	0.00101	5	9	1	0	1	1	1
G41	9[a]	PFI0785c	Sugar transporter	2.3	0	2.3	1	0	1	2	0.9	0.00042	1	3	0	1	0	1	0
G42	12[a]	PFL0410w	ERLE domain protein	2.1	0.7	2.8	0	1	3	4	1.4	0.00069	1	5	2	1	1	2	2
G43	14[a]	PF14–0722	ERLE domain protein	0.8	1.5	2.3	4	0	0	4	1.7	0.00083	0	4	0	0	0	0	0

Table 1. *cont.*

Gene Code	Chrom. Location	PlasmoDB link (4.0)	Name	Coding (kb)	Non-coding (kb)	Total Seq (kb)	ncSNP	cSNP		Total SNP	Frequency (SNP/kb)	Nucleo. Poly. (θ)	MS	Total Poly.	Codon position			Nucleotide substitution	
								sSNP	nsSNP						1st	2nd	3rd	Transition	Transversion
G44	5[a]	PFE1525w	ERLE domain protein	0.8	2.7	3.5	0	0	0	0	0	0.00000	0	0	0	0	0	0	0
G45	8	PF08–0048	Multiple TM protein	0.3	0	0.3	0	0	1	1	3.3	0.00160	0	1	1	0	0	1	0
G46	13	MAL13P1.206	Phosphate transporter	2.4	0	2.4	0	0	2	2	0.8	0.00040	1	3	0	2	0	1	1
G47	5[a]	PFE0775c	Glycine transporter	1.4	0	1.4	0	0	3	3	2.1	0.00103	0	3	2	1	0	1	2
G48	9	PFI0765w	Multiple TM protein	0.7	0	0.7	0	0	0	0	0	0.00000	0	0	0	0	0	0	0
G49	8[a]	PF08–0078	ABC transporter	4.2	0	4.2	0	1	8	9	2.1	0.00103	2	11	4	3	2	1	8
G50	9	PFI1295c	Multiple TM protein	1.4	0	1.4	0	0	0	0	0	0.00000	0	0	0	0	0	0	0
G51	5	PFE0995c	DEAD box helicase	0.5	0	0.5	0	0	0	0	0	0.00000	0	0	0	0	0	0	0
G52	10[a]	PF10–0366	mt adenine transporter	0.9	0	0.9	0	0	2	2	2.2	0.00107	0	2	0	2	0	0	2
G53	14	PF14–0244	ABC transporter	2.0	0	2.0	0	0	0	0	0	0.00000	0	0	0	0	0	0	0
G54	14[a]	PF14–0260	Multiple TM protein	2.4	0	2.4	0	5	2	7	2.9	0.00140	1	8	3	0	4	6	1
G55	14	PF14–0133	ABC transporter	2.4	1.2	3.6	1	1	0	2	0.6	0.00027	2	4	0	0	1	1	0
G56	12[a]	PFL1410c	ABC transporter	6.3	0	6.3	0	2	6	8	1.3	0.00061	4	12	5	2	1	4	4
G57	12	PFL1125w	Cation transporter ATPase	0.8	0	0.8	0	0	0	0	0	0.00000	0	0	0	0	0	0	0
G58	12[a]	PFL2395c	Dimethyladenosine transferase	1.9	0	1.9	0	0	3	3	1.6	0.00076	3	6	0	2	1	1	2
G59	12[a]	PFL0420w	Multiple TM protein	4.8	0	4.8	0	0	0	0	0	0.00000	2	2	0	0	0	0	0
G60	12[a]	PFL1480w	Sec13 homologue	2.6	0	2.6	0	3	6	9	3.5	0.00166	5	14	3	2	4	6	3
G61	5[a]	PFE1150w	MDR1	4.5	0	4.5	0	0	5	5	1.1	0.00053	0	5	4	1	0	1	4
G62	7[a]	MAL7P1.27	pfCRT	1.6	0	1.6	0	0	14	14	8.8	0.00420	0	14	6	5	3	3	11
G69	12[a]	PFL0590c	P-type ATPase	3.6	0	3.6	0	0	4	4	1.1	0.00053	1	5	2	1	1	2	2
G70	12[a]	PFL0620c	Choline transporter	1.7	0	1.7	0	0	1	1	0.6	0.00028	0	1	1	0	0	1	0
Sum				99.2	14	113.2	34	34	96	164	1.4	0.00073[b]	67	231	52	38	40	66	64

a. Indicates genes that were mapped to chromosomes by typing polymorphic sites in or near the genes in 35 progeny from a genetic cross (Su et al., 1999). The chromosomal locations for other genes are according to the assignment of genome databases.
b. The average nucleotide polymorphism from five isolates.
The gene names are according to blast search hits or annotations from the genome sequencing centres.
nsSNP, non-synonymous SNP; cSNP, coding SNP; sSNP, synonymous SNP; and ncSNP, non-coding SNP.

Table 2. Parasite isolates, origins and responses to CQ and QN.

Isolates	Origin	CQ IC_{50}	QN IC_{50}	Isolates	Origin	CQ IC_{50}	QN IC_{50}
98-8a	Thailand	6.460	3.460	PNG2	PNG	0.559	0.110
98-5	Thailand	3.260	0.780	98-26b	Thailand	0.550	0.890
1088	Thailand	2.970	2.800	7G8	Brazil	0.546	0.390
Thai19	Thailand	2.163	0.824	DIV30	Brazil	0.538	0.278
123/5	Sudan	1.876	0.350	TM191c	Thailand	0.528	0.978
PC49	Peru	1.706	0.457	99-22a	Thailand	0.520	0.290
V1/S	Vietnam	1.631	1.263	98-9a	Thailand	0.510	0.680
98-8c	Thailand	1.580	0.290	TM92c-815	Thailand	0.500	0.410
PC09	Peru	1.518	1.147	T17-1	Thailand	0.500	0.800
Indo	Indochina	1.353	3.022	99-16b	Thailand	0.500	0.250
PC15	Peru	1.323	1.750	99-15	Thailand	0.500	0.650
JCK	Cambodia	1.290	1.446	PNG13	PNG	0.491	0.082
M97	The Gambia	1.281	0.550	98-13	Thailand	0.490	0.850
FCB	SE Asia	1.218	1.612	99-10a	Thailand	0.480	0.570
99-29a	Thailand	1.200	1.510	98-22b	Thailand	0.480	0.560
9020	Ghana	1.159	0.172	98-4	Thailand	0.440	0.560
P31	SE Asia	1.125	1.120	ECU	Ecuador	0.387	0.210
KMVII	Kenya	1.124	2.277	Fab6	S. Africa	0.173	0.416
98-11	Thailand	1.120	0.110	106/1	Sudan	0.112	0.852
DIV17	Brazil	1.092	0.670	K39	Kenya	0.107	0.351
124/8	Sudan	1.087	0.269	Hb3	Honduras	0.084	0.632
Thai16	Thailand	1.079	0.353	D10	PNG	0.075	0.131
102/1	Sudan	1.068	0.165	T2/c6	Thailand	0.073	0.166
DIV14	Brazil	1.064	0.573	Fab9	S. Africa	0.068	0.221
Dd2	Indochina	1.000	1.000	9021	Ghana	0.067	0.107
99-3	Thailand	0.950	0.990	SL/D6	Sierra Leone	0.064	0.136
128/4	Sudan	0.939	0.321	M24	Kenya	0.058	0.113
PC17	Peru	0.925	0.507	MR80	Vietnam	0.054	0.131
TM284	Thailand	0.899	0.940	601	The Gambia	0.054	0.080
D5	SE Asia	0.875	1.117	9016	Ghana	0.052	0.165
99-8	Thailand	0.860	2.040	M5	Mali	0.052	0.268
ICS	Brazil	0.859	0.551	Hu425	The Gambia	0.051	0.101
99-29d	Thailand	0.840	0.680	433	The Gambia	0.051	0.107
T18-1	Thailand	0.820	0.770	425	The Gambia	0.049	0.130
98-17a	Thailand	0.820	0.120	3D7	The Netherlands	0.048	0.253
M2	Mali	0.792	0.227	713	Guinea Bissau	0.047	0.128
PC26	Peru	0.783	0.710	M190	The Gambia	0.046	0.149
TM91c-40	Thailand	0.760	0.310	224	The Gambia	0.046	0.129
JAV	Colombia	0.757	0.499	418	The Gambia	0.042	0.120
9013	Ghana	0.756	0.313	449	The Gambia	0.041	0.109
T2-2	Thailand	0.750	0.880	4156	The Gambia	0.040	0.166
C2A	Thailand	0.746	1.139	REN	Sudan	0.040	0.173
ECP	Brazil	0.712	0.517	Haiti	Haiti	0.040	0.207
PAD	Brazil	0.666	0.280	MT/s-1	Thailand	0.039	0.105
PNG4	PNG	0.620	0.196	434	The Gambia	0.039	0.059
PNG3	PNG	0.601	0.082	LF4/1	Liberia	0.037	0.136
99-18	Thailand	0.570	1.480	P13	Mali	0.037	0.088
98-18	Thailand	0.570	1.600	Camp	Malaysia	0.031	0.051
S35	Mali	0.568	0.231				

The IC_{50} of each isolate for CQ and QN, representing median values from at least five independent assays, was normalized to that of control parasite Dd2 (mean IC_{50}: 404.1 nM for CQ and 315.9 nM for QN) included in each assay to account for day-to-day assay variation.

from the isolates was observed (Fig. 1B), suggesting the lack of a major genetic determinant. Although the majority of the CQ-sensitive (CQS) parasites carrying the wild-type *pfcrt* allele also have lower QN IC_{50} values, parasites carrying the mutant *pfcrt* alleles vary widely for QN IC_{50} (Fig. 1C). Best-fit curves of IC_{50} values (Fig. 1C) showed approximately parallel lines for CQ and QN, indicating the likely involvement of common genes, especially *pfcrt*, in levels of sensitivity to both drugs (Fig. 1C).

SNPs from multiple transporter genes associated with CQ and QN responses

To identify genes that may contribute to responses to CQ and QN, we analysed the relationship between the SNPs and the quantitative parasite drug responses by direct genotype–phenotype association. When all 97 isolates were tested for associations (columns headed 'All', Table 3), 15 SNPs from six genes (*pfcrt*, G2, G7, G25, G30 and G49) were strongly associated with the quanti-

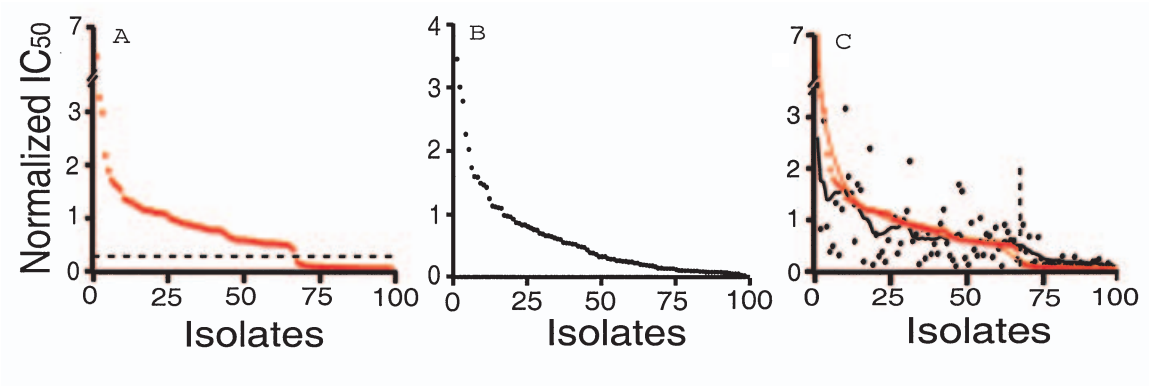

Fig. 1. Plots of *P. falciparum* isolate responses (IC_{50} normalized to Dd2 response in the same assay plate) to CQ and QN. Median IC_{50} from five independent assays were sorted in descending order and plotted. Each isolate shows a unique IC_{50} in responses to CQ (A) and QN (B). Parasites above the dashed line (A) have mutant *pfcrt* alleles, whereas those under the line carry the wild-type allele. Except for the five parasites with the highest IC_{50}, no obvious IC_{50} jump was found in the QN plot (B). Best-fit curves (C) of the IC_{50} values to CQ (red dots) and QN (black dots). Isolates were sorted according to CQ IC_{50} in descending order, and IC_{50} to both CQ and QN were plotted (Lowess regression using PRISM3 software). Note that CQS parasites (right side of the vertical dashed line) generally have low IC_{50} to QN.

tative CQ responses at $P < 0.001$, and four additional SNPs, from *pfmdr1*, G47 and G70, gave more marginal *P*-values (Table 3). Similar *P*-values were obtained for associations of these SNPs with the QN responses of the 97 isolates (Table 3). These *P*-values give an intriguing hint of multiple drug susceptibility determinants, as the strong CQ associations include most SNPs from *pfcrt*, and four of the associated genes encode putative ABC transporters (*pfmdr1*, G2, G7 and G49). However, the associations in the worldwide cohort of 97 parasites may reflect co-ancestry of lineages in different geographical regions in addition to drug response associations. Accordingly, we also analysed the parasites by different continental regions, where they have distinct CQ selection histories and CQR origins (Wootton *et al*., 2002).

The separate associations for Asia, Africa and the Americas (Table 3) show that 14 SNPs from six genes (*pfcrt*, *pfmdr1*, G2, G30, G49 and G55) are strongly associated with CQ responses in at least one geographical region ($P < 0.01$), and four additional SNPs from G7, G25 and G49 have *P*-values <0.022 (cut-off threshold from permutation analysis). Interestingly, different SNPs from *pfmdr1* are significantly associated with CQ responses among different parasite populations; the SNPs at amino acid position 86 and 1034 are significantly associated with CQ responses in parasites from Africa and South America respectively (Table 3), a finding consistent with results reported previously (Volkman and Wirth, 1998; Foote *et al*., 1990; Reed *et al*., 2000; Adagut and Warhurst, 2001; Babiker *et al*., 2001). For the South American parasites, which are all CQ resistant, most of the SNPs in *pfcrt* are not informative for quantitative drug response associations because of the absence of the ancestral *pfcrt* allele in this region. The majority of the *pfcrt* SNPs (except amino acid position 72 and 97) from African and Asian parasites were very significantly associated ($P < 0.0001$) with CQ response (Table 3).

Additionally, 12 SNPs from five genes showed evidence of association with the higher QN IC_{50} in the continental subpopulations ($P < 0.022$, Table 3). These included seven strongly significant ($P < 0.001$) *pfcrt* SNP associations in both Asian and African populations and, to a lesser degree, SNPs in G30 (Africa), G54 (Asia) and G70 (Asia). Also, two SNPs from G2 and one from G49, which have marginally non-significant *P*-values in the permutation tests, are significant by the regression analysis.

The strong association between *pfcrt* SNPs and responses to both CQ and QN is consistent with a scenario in which *pfcrt* may physically interact with both drugs, which is supported by various observations: (i) substitutions of K with I or N at amino acid position 76 of *pfcrt* changed the parasite response to both CQ and QN simultaneously (Cooper *et al*., 2002); (ii) the majority of CQS parasites carrying the wild-type *pfcrt* also have low QN IC_{50} (Fig. 1C); and (iii) the antimalarial effects of CQ and QN are antagonistic (Skinner-Adams and Davis, 1999). The results also agree with historical observations that no QN failures were reported before CQR (Peters, 1987) and are consistent with the proposal that the use of CQ may have led to background mutations contributing to steady decreases in QN potency (Knowles *et al*., 1984). The majority of the associations with both CQ and QN are corroborated by significant *P*-values ($P < 0.05$) from analysis of variance and linear regression (underlined in Table 3).

Table 3. SNPs from 11 *P. falciparum* putative transporter-encoding genes associated with CQ and QN responses.

Gene	Predicted products	Locus	AA change	Informativeness index[a] All	Asia	Africa	Americas	CQ (*P*-value) All	Asia	Africa	Americas	QN (*P*-value) All	Asia	Africa	Americas	Comments
Pfcrt	Putative transporter	72	C–S	0.61	0	0	0.94	0.087	NI	NI	0.19	0.25	NI	NI	0.17	No association
	10 TM segments	74	M–I	0.86	0.58	0.86	0	**9.6E–10**	**3.1E–7**	**2.5E–7**	NI	**7.6E–8**	**1.4E–4**	**0.001**	NI	CQ/QN Africa and Asia
	with CEGA motif	75	N–E	0.86	0.58	0.86	0.37	**5.6E–10**	**3.1E–7**	**2.5E–7**	NI	**5.1E–8**	**1.4E–4**	**0.001**	NI	CQ/QN Africa and Asia
		76	K–T	0.89	0.58	0.84	0.66	**1.7E–19**	**3.1E–7**	**1.1E–7**	**4.9E–4**	**5.3E–8**	**1.4E–4**	**0.0077**	0.28	CQ/all areas; QN/Asia and Africa
		220	A–S	0.85	0.58	0.86	0.66	**2.3E–18**	**3.1E–7**	**2.5E–7**	**4.9E–4**	**1.8E–9**	**1.4E–4**	**0.001**	0.28	CQ/all areas; QN/Asia and Africa
		271	F–Q	0.86	0.58	0.86	0	**9.6E–10**	**3.1E–7**	**2.5E–7**	NI	**7.6E–8**	**1.4E–4**	**0.001**	NI	CQ/QN Africa and Asia
		326	N–D/S	0.87	0.58	0.83	0.39	**1.2E–9**	**3.1E–7**	**6.7E–7**	NI	**2.3E–8**	**1.4E–4**	**0.0011**	0.26	CQ/QN Africa and Asia
		356	I–T/L	0.9	0.77	0.21	0.39	**3.7E–6**	**0.0012**	NI	NI	**1.8E–5**	0.044	NI	NI	CQ/Asia
		371	R–I	0.87	0.58	0.86	0.18	**9.5E–10**	**3.1E–7**	**2.5E–7**	NI	**6.2E–8**	**1.4E–4**	**0.001**	NI	CQ/QN Africa and Asia
Pfmdr1	ABC transporter	86	N–Y	0.77	0.74	0.82	0	0.065	**0.02**	**0.0048**	NI	0.38	0.29	0.091	NI	CQ/Africa and Asia
		1034	S–C[b]	0.63	0.48	0.1	0.88	**0.0066**	NI	NI	**0.0073**	**0.015**	NI	NI	0.12	CQ/Americas, may be geographical
		1042	N–D	0.8	0.79	0	0.37	**0.001**	0.19	NI	NI	**3.7E–4**	0.064	NI	NI	May be geographical
G2	ABC transporter	191	Y–H	0.86	0.7	0.36	0.99	**2.8E–7**	**0.0075**	NI	0.086	**6.6E–7**	0.067	NI	0.024	Weak CQ/Asia; weak QN/Americas
		437	A–S	0.85	0.68	0.47	0.97	**12E–7**	0.076	NI	0.031	**4.3E–6**	0.3	NI	0.029	Weak CQ/QN, Americas
G7	ABC transporter	1390	&1	0.74	0.4	0.87	0.92	**1.8E–6**	NI	**0.01**	0.32	**0.0032**	NI	0.18	0.37	CQ/Africa
G25	Sulphate permease	Intron	G–A	0.84	0.83	0.36	0	**7.6E–4**	**0.014**	NI	NI	**3.2E–5**	0.077	NI	NI	Weak CQ/Asia, may be geographical
G30	GTPase	Intron	C–G	0.65	0.22	0.94	0.37	**9.2E–8**	NI	**0.0048**	NI	**2.5E–5**	NI	**0.011**	NI	CQ/QN/Africa
G47	Glycine transporter	241	L–V	0.89	0.54	0.75	0.37	**0.0033**	0.075	0.4	NI	**7.1E–5**	**0.0056**	0.068	NI	QN/Asia, low, may be geographical
G49	ABC/ATPase	146	Q–E	0.82	0.88	0	0	**7.0E–5**	0.018	NI	NI	**9.3E–6**	0.025	NI	NI	Weak CQ/QN/Asia, may be geographical
		1046	L–I	0.82	0.78	0.36	0.92	0.076	0.23	NI	**0.021**	0.35	0.12	NI	0.14	Weak QN/Americas, may be geographical
		1116	L–I	0.88	0.5	0.73	0.8	**1.6E–7**	**0.0053**	0.24	0.16	**6.5E–6**	0.029	0.17	0.24	Weak CQ/QN/Asia, may be geographical
G54	Membrane protein	141	Y–Y	0.9	0.87	0.87	0.66	0.063	0.041	0.19	0.089	**0.0099**	**0.0047**	0.35	0.21	QN/Asia
		144	T–T	0.9	0.87	0.87	0.66	0.063	0.041	0.19	0.089	**0.0099**	**0.0047**	0.35	0.21	QN/Asia
G55	ABC transporter	–	&2	0.91	0.79	0.92	0	0.051	0.38	**0.0049**	NI	0.3	0.35	0.056	NI	CQ/Africa
G70	Choline transporter	105	E–K	0.85	0.95	0.55	0	**0.014**	0.059	0.32	NI	**1.5E–5**	**0.017**	0.16	NI	QN/Asia

P-values are emphasized if significant association was determined by permutation analysis (bold) or linear regression/ANOVA analysis ($P < 0.05$, underlined): see *Experimental procedures* for these two independent statistical strategies. Some strong *P*-values for 'All' regions may reflect geographical subdivision (indicated in the comments column) rather than CQ or QN associations, as confirmed by direct geographical association tests (data not shown): in these cases, only weak or non-significant *P*-values were obtained for isolates from individual continents.

a. Informativeness index (*I*, see *Experimental procedures*) was used to exclude SNPs with low minor allele frequency: loci with $I < 0.5$ are denoted non-informative (NI). Cases with marginal *I*-values of 0.5–0.7 should also be viewed with care because of relatively low minor allele frequency or parasite sample sizes. &1 represents a trinucleotide insertion and &2 is also a microsatellite polymorphism 700 bp downstream of the stop codon.

b. Indicates a singleton third allele.

Negative controls and SNPs from 39 putative housekeeping genes

Of the 39 putative transporter genes that have SNPs, 28 show no significant associations with CQ or QN responses. These SNPs/genes can serve as controls for false associations due to population structures or other unknown factors. Additionally, we also searched for SNPs from 39 putative housekeeping genes (or partial genes) on chromosome 3 among five isolates (Dd2, 3D7, Hb3, D10 and 7G8), totalling 33.6 kb coding and 3.9 kb non-coding sequences (Mu *et al.*, 2002) (Table 4). These SNPs were then assayed in all 97 isolates. Only 13 of these 39 putative housekeeping genes (compared with 39 of 49 putative transporter genes) showed nucleotide substitutions, a significant difference between the transporter and housekeeping classes in the number of polymorphic genes (χ^2 test, $P < 0.001$). Nucleotide substitution rates within these 13 housekeeping genes are also significantly lower (χ^2 test, $P < 0.0001$), suggesting purifying selection acting on the housekeeping genes and/or positive selection on the transporters ($\theta = 2.0 \times 10^{-4}$ for the housekeeping genes compared with $\theta = 7.3 \times 10^{-4}$ for the transporter genes). Of the 13 genes with nucleotide substitutions, only two satisfied our informativeness index criterion of $I \geq 0.5$ among the 97 isolates (guanine nucleotide-binding protein $I = 0.59$ and 40S ribosomal protein S12 $I = 0.66$). The SNPs in these two genes are not significantly associated with either CQ or QN responses according to our *P*-value thresholds from permutation analysis (data not shown).

Linkage disequilibrium among associated genes

Co-selection of *pfcrt* amino acid T76 and *pfmdr1* Y86 positions by CQ treatment has been reported recently (Adagut and Warhurst, 2001; Babiker *et al.*, 2001; Djimde *et al.*, 2001), suggesting either that the two genes may work in concert in determining CQ resistance levels or that the mutation at *pfmdr*1Y86 may compensate deleterious *pfcrt* mutations. To detect potential co-selection, we evaluated LD (*D*) between pairs of SNPs within a geographical region, including Africa (A), Asia (B) and South America (C) (Fig. 2). In addition to strong LD within *pfcrt*, the most notable finding was the strong LD detected between *pfcrt* and *pfmdr1* (position 86), G7, G30 and G55 in African parasites, and between *pfcrt* and G2, G7, G47 and G49 in Asian parasites ($P < 0.00001$, red, Fig. 2A and B). Additionally, strong LD is present in African isolates within G2 and G54 and between G25 and G49 (Fig. 2A), and in Asian parasites within G2, G49 and G54 and between the following pairs: G7/G47, G7/G49, G47/G49 and G54/G70 (Fig. 2B). For parasites from South America, strong LD is detected between *pfcrt* and *pfmdr1* (position 1034 only), *pfmdr1*/G47 and G49/54 in addition to LD within *pfcrt*, G2

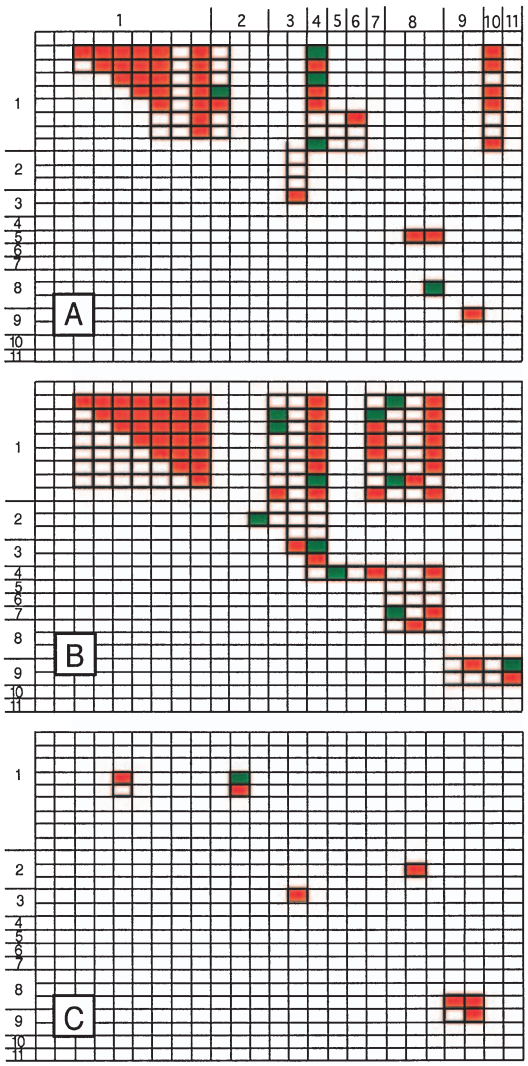

Fig. 2. LD among SNPs within and between 11 putative transporter genes associated with CQ and QN responses. Pairwise LD parameter *D* was calculated as described in the *Experimental procedures*. *P*-values and confidence intervals were determined from permutation tests, giving thresholds for unadjusted *P*-values $P < 0.001$ (marginal, green) and $P < 0.00001$ (highly significant, red). More genes show LD among parasites from Africa (A) and Asia (B) than from South America (C), probably because of a more uniform genetic background in the South American population: for example, most *pfcrt* positions are not informative for LD detection in South American isolates because there is only one allele. The numbers represent individual genes that have one or more SNPs (small rectangles) significantly associated with drug responses in Table 3: 1, *pfcrt*; 2, *pfmdr1*; 3, G2; 4, G7; 5, G25; 6, G30; 7, G47; 8, G49; 9, G54; 10, G55; and 11, G70. All the LDs in this figure are also significant ($P < 0.01$) by tests performed using an R^2 estimate implemented in DNASP (data not shown). The permutation test criterion was also used to evaluate LD between all SNP pairs from the two informative housekeeping genes and the 28 transporter-encoding genes that showed no significant drug associations. These were negative for LD: only a few of these SNP pairs showed near-marginal *P*-values, as expected by chance (data not shown).

Table 4. Single nucleotide polymorphisms in 39 putative housekeeping genes.

Gene ID	Products	Coding (bp)	Non-coding (bp)	Total seq. (bp)	ncSNP	cSNP		Total SNP	Frequency (SNP/kb)	Nucleotide Poly. (θ)	MS	Total
						sSNP	nsSNP					
PFC0060c	Ser/Thr protein kinase	1200	90	1290	0	0	2	2	1.5504	0.0007	1	3
PFC0065c	Alpha/beta hydrolase	1314	40	1354	0	0	0	0	0.0000	0	1	1
PFC0100c	Guanine nucleotide-binding protein	1120	0	1120	0	0	1	1	0.8929	0.0004	1	2
PFC0135c	Chromosome region maintenance protein 1	1200	0	1200	0	1	0	1	0.8333	0.0004	0	1
PFC0155c	DNA-directed RNA polymerase, 14 kDa subunit	492	0	492	0	0	0	0	0.0000	0	0	0
PFC0190c	Homologue of *D. melanogaster* PAST protein	1219	0	1219	0	0	0	0	0.0000	0	0	0
PFC0200w	60 ribosomal protein L44	315	234	549	0	0	0	0	0.0000	0	1	1
PFC0205c	Glutaredoxin	509	240	749	0	0	0	0	0.0000	0	1	1
PFC0225c	Elongation factor TS	1173	104	1277	1	0	0	1	0.7831	0.0004	1	2
PFC0255c	Ubiquitin conjugating enzyme	429	182	611	0	1	0	1	1.6367	0.0008	1	2
PFC0275w	Glycerol-3-phosphate dehydrogenase	1203	0	1203	0	0	0	0	0.0000	0	0	0
PFC0295c	40S ribosomal protein S12	426	166	592	1	1	0	2	3.3784	0.0016	1	3
PFC0300c	60 ribosomal protein L7	711	245	956	0	0	0	0	0.0000	0	1	1
PFC0305w	Putative homologue of human EB1 protein	1183	0	1183	0	0	0	0	0.0000	0	1	1
PFC0310c	ATP-dependent CLP protease	1100	0	1100	0	1	1	2	1.8182	0.0009	0	2
PFC0375c	Homologue of *C. elegans* T08A11.2 protein	981	0	981	0	0	0	0	0.0000	0	1	1
PFC0385c	Serine/threonine protein kinase	1220	0	1220	0	0	0	0	0.0000	0	2	2
PFC0390w	Homologue of *C. elegans* Y48E1c.2 protein	1239	0	1239	0	0	0	0	0.0000	0	0	0
PFC0395w	Asparagine synthetase	1019	273	1292	1	0	0	1	0.7740	0.0004	1	2
PFC0400w	60S Acidic ribosomal protein P2	339	0	339	0	0	0	0	0.0000	0	0	0
PFC0410w	Putative homologue of rat brain protein YT51	567	552	1119	1	0	0	1	0.8937	0.0004	2	3
PFC0420w	Calcium-dependent protein kinase	978	20	998	0	1	0	1	1.0020	0.0005	0	1
PFC0470w	Valyl-tRNA synthethase	1095	195	1290	0	0	0	0	0.0000	0	1	1
PFC0495w	Aspartyl protease	1249	162	1411	0	1	0	1	0.7087	0.0003	1	2
PFC0520w	26S Proteasome regulatory subunit S14	915	0	915	0	0	0	0	0.0000	0	0	0
PFC0525c	Glycogen synthase kinase	495	0	495	0	0	0	0	0.0000	0	0	0
PFC0535w	60S ribosomal protein L26	371	209	580	0	0	0	0	0.0000	0	1	1
PFC0595c	Serine/threonine protein phosphatase	927	0	927	0	0	0	0	0.0000	0	0	0
PFC0710w	Inorganic pyrophosphatase	1080	0	1080	0	0	0	0	0.0000	0	0	0
PFC0735w	40S ribosomal protein S15A	387	240	627	1	0	0	1	1.5949	0.0008	1	2
PFC0775w	40S ribosomal protein S11	480	287	767	1	0	0	1	1.3038	0.0006	2	3
PFC0800w	Putative band 7 protein	1124	0	1124	0	0	0	0	0.0000	0	0	0
PFC0825c	Polyadenylation specificity factor	1220	0	1220	0	0	0	0	0.0000	0	0	0
PFC0860w	Kinesin-related protein	1240	0	1240	0	0	0	0	0.0000	0	0	0
PFC0865w	RNA-binding protein	665	100	765	0	0	0	0	0.0000	0	2	2
PFC0920w	Histone H21A variant	144	384	528	0	0	0	0	0.0000	0	1	1
PFC0950c	Acylaminoacyl-peptidase	1018	0	1018	0	0	0	0	0.0000	0	0	0
PFC0975c	Cyclophilin	516	0	516	0	0	0	0	0.0000	0	0	0
PFC1020c	40S ribosomal protein S3A	738	161	899	0	0	0	0	0.0000	0	1	1
Sum		33601	3884	37485	6	6	4	16	0.4403[a]	0.0002[a]	26	42

a. Indicates averaged frequency and nucleotide polymorphism respectively.
The gene names are according to annotations in PlasmoDB (4.0).
nsSNP, non-synonymous SNP; cSNP, coding SNP; sSNP, synonymous SNP; and ncSNP, non-coding SNP.

and G54 (Fig. 2C) and, to a lesser extent ($P = 0.001–0.00001$), between *pfcrt* and *pfmdr1* (position 1034). Some SNPs do not show LD in the South American population because many of the positions (for example in *pfcrt*) have only one allele. These LDs were corroborated using estimates of R^2 implemented in DNASP (Rozas and Rozas, 1999; data not shown).

The LD between *pfcrt* and *pfmdr1* in parasites from Africa (*pfmdr1* position 86) and South America (*pfmdr1* position 1034) supports previous reports suggesting co-selection or involvement of both *pfcrt* and *pfmdr1* in CQR (Adagut and Warhurst, 2001; Babiker *et al*., 2001; Djimde *et al*., 2001; Chen *et al*., 2002). It is interesting that *pfcrt* is linked to different SNPs in *pfmdr1* in parasites from Africa and South America. This implies that specific genetic backgrounds are associated with various *pfcrt* mutant alleles and that patterns in LD reflect distinct drug selection histories in Africa and South America (Wootton *et al*., 2002). Lack of LD between *pfcrt* and *pfmdr1* in the Asian population could also be a result of extensive mefloquine use that may counter the selective effect of CQ (Duraisingh *et al*., 1997). Strong LD between the genes located on different chromosomes (Table 1) provides indirect evidence that some genes work in concert with *pfcrt* in CQ and QN responses.

Of course, many unknown factors may also contribute to the observed LD, including other antimalarial agents not tested in this study; therefore, strong LD between putative transporters does not constitute formal proof that they are linked as a result of CQ and QN selection. Similarly, the strong associations between SNPs and drug response phenotypes do not formally prove that drug selection acted historically on these SNPs *per se* rather than on closely linked loci. However, our results as a whole provide support for the hypothesis that multiple drug response determinants have acted in different combinations on different continents.

All the proteins encoded by the genes associated with CQ and QN responses are predicted to be membrane-spanning transporters or transport regulators that function in either the plasma or the organellar membrane (Table 1). Four of the genes (G2, G7, G49 and G55) encode ABC transporter/ATPases similar to *pfmdr1*, consistent with the hypothesis that these genes are generally involved in drug transport. G30 encodes a GTPase predicted to be a translation factor (Leipe *et al*., 2002) that may affect protein synthesis. The majority of substitutions observed are unlikely, however, to cause drastic distortions in the structures of the encoded proteins. It is possible that some of the genes may affect gene expression levels caused by polymorphisms in non-coding regions such as those in G7, G30 and G54.

Results from this study show that, in addition to *pfcrt* and *pfmdr1*, SNPs from several putative transporters are significantly associated with elevated responses to CQ and/or QN, providing evidence that the level of CQ and QN response is a multigenic phenomenon and that mutations in different transporter genes may impact the response to antimalarial compounds. Additionally, we provide strong evidence that *pfcrt* is also involved in responses to QN. There appear to be shared genes underlying CQ and QN responses, and mutations in *pfcrt* are probably necessary, but not sufficient, to confer QN resistance. With the near saturation of mutant *pfcrt* alleles in south-east Asia and South America and an increasing prevalence in Africa, the genetic background for QN treatment failure may exist in current parasite populations. Overlapping, but not identical sets of genes, including many encoding unknown proteins, may provide an explanation for why a parasite can be resistant to CQ but highly sensitive to QN. Although the number of parasite isolates tested in this study is relatively small, identification of these genes implicated in drug responses provides a foundation for their further functional characterization and allows for a better understanding of the genetic basis of drug resistance in malaria parasites. The molecular roles of these candidate genes and transporter proteins in CQ and QN responses can now be tested rigorously using transfection and gene knock-out experiments.

Experimental procedures

Gene sequences and parasite DNA

Predicted coding sequences of the 3D7 parasite were downloaded from the websites of the genome sequencing centres and PlasmoDB (http://www.sanger.ac.uk/Projects/Protozoa/; http://www.stanford.edu/group/malaria/index.html; http://www.tigr.org/tdb/edb2/pfa1/htmls/; http://www.plasmodb.org/). Text search using key word 'transporter' and motif searches for membrane proteins and ion channels were performed to identify sequences with homologies to various transporters. Coding sequences are according to annotation in the database. Parasite isolates were genotyped with MS markers (Wootton *et al*., 2002) to verify clonality, and isolates with mixed genotype were cloned by limiting dilution before drug tests. Parasite culture, DNA extraction and *in vitro* drug assays were performed as described previously (Su *et al*., 1997).

Amplification of parasite DNA and DNA sequencing

Oligonucleotide primers were designed to amplify and sequence DNA of the transporter genes from four isolates first (Dd2, Hb3, D10 and 7G8). These isolates have been genotyped previously with 342 MS markers and shown to have diverse genetic backgrounds (Wootton *et al*., 2002). Primers, 20–25 bp long, were synthesized in a DNA synthesizer (Applied Biosystems) or obtained through a commercial supplier (Invitrogen). Polymerase chain reaction (PCR) set-

ups include 4 μl of DNA (≈ 5 ng), 0.5 μl of each primer (50 pM) and 45 μl of PCR mix containing 5 μl of 10× PCR buffer, 1.0 μl of dNTPs (10 mM) and 0.1 μl (5 U μl^{-1}) of *Taq* polymerase (Invitrogen). All the amplifications were performed with one cycling condition: 94°C for 2 min, 35 cycles of 94°C for 20 s, 52°C 10 s, 48°C for 10 s and 60°C for 1–4 min, and 60°C for 5 min. Five microlitres of the PCR products were run on 1% agarose gel to check for quality of amplification. If there was a single band, and no obvious 'primer–dimer' was present, the PCR product was treated with 1 μl of ExoSAP-IT (USB) at 37°C for 15 min and 80°C for another 15 min. The PCR product (2–5 μl) was used in a sequencing reaction using dichlororhodamine or BigDye terminator chemistry on ABI377 or ABI3100 (Applied Biosystems).

SNP discovery and verification

DNA sequences were aligned using SEQUENCHER 3.1 (Gene Codes Corporation) or ASSEMBLYLIGN software (Oxford Molecular). All potential SNPs and each of the ambiguities were verified by visual inspection. Polymorphic MS sequences were aligned with the software first, then with visual assistance to minimize artificial SNPs.

SNP association and linkage disequilibrium analyses

Genetic analysis was adapted to the haploid inheritance system of *P. falciparum*. We used two independent methods, namely permutation analysis and quantitative trait locus (QTL) regression analysis, to evaluate association between the multilocus alleles and the drug susceptibility phenotypes. Permutation analysis used refinements of methods described by Churchill and Doerge (1994), Long and Langley (1999) and Zhao *et al*. (2000). The association statistic, A_i for the *i*th polymorphic site, used the discrete small-sample equivalent of the *t*-test statistic based on the direct distance between the binary-encoded genotype vector and the scaled (0,1) phenotype vector of log IC_{50} values. The significance of A_i was determined from the distribution of simulated values obtained from 1000 random permutations of the scaled phenotype vector on the fixed indices of the genotypes. If A_i was within the range of values from permutations, an exact, distribution-independent *P*-value was obtained from the proportion of permuted samples more extreme than A_i. If A_i was more extreme than the computed permutation distribution, its *P*-value was estimated by extrapolation using Fisher's Z-score approximation. These unadjusted *P*-values are presented in Table 3; however, as long recognized (Good, 1953), such estimates do not translate directly into conventional significance levels. Accordingly, empirical significance thresholds were determined using second-order nested permutations of the permuted phenotype vectors. Intervals estimated by this test were: 'marginal', $P < 0.022$ (one false positive expected by chance in the entire multilocus association analysis); and 'highly significant', $P < 0.001$. Such thresholds are sample size independent and account for both the non-normality of the permutation distribution and the multiple tests of the multilocus analysis, thus enabling us to minimize the false discovery rate of associations.

We also used the same 1000 permutations to make an empirical estimate of the 'informativeness index', I_i, of each *i*th SNP locus for the four geographical sets of isolates analysed ('All regions', 'Asia', Africa' and 'Americas'). I_i is the ratio T_i/T_{max}, where T_i is the range (A_{upper}–A_{lower}) of the association statistic obtained from the 1000 permutations for locus *i*, and T_{max} is the corresponding range from 1000 random permutations for a maximally informative model locus with equal allele frequencies. I_i is independent of sample size and provides a consistent criterion to exclude the less informative SNPs from significance tests: loci with $I < 0.5$ are denoted 'NI' (not informative) in Table 3.

Population-based quantitative trait loci (QTL) regression association was performed as described previously (Zhao *et al*., 2001). Briefly, each locus tested is assumed to have two alleles, *A* (wild-type) and a (mutant), with frequencies P_A and P_a respectively. Let y_i be the IC_{50} value of the *i*th parasite response to CQ and QN and x_i be an indicator variable of the allele at the locus, defined as $x_i = 1$ if the allele is a and $x_i = -1$ if the allele is *A*. QTL analysis was modelled by the following regression:

$$y_i = \mu + x_i\alpha + e_i$$

where μ is overall population mean, α is genetic additive effect and e_i is an error with $E[e_i] = 0$ and Var $(e_i) = \sigma^2_e$. The generalized likelihood ratio test statistic, which follows an $F_{1,n-2}$ distribution, was used to test the null hypothesis H_0: $\alpha = 0$. *P*-values for the *F* statistics were assessed using regression ANOVA tests.

In view of caveats about inferences from single LD measures (Hedrick and Kumar, 2001), we used two independent methods to calculate the amount of LD between alleles (nucleotides) at different polymorphic sites. The basic LD measure, *D* (Lewontin and Kojima, 1960), was computed for all pairs of loci from the numerically encoded allele vectors of loci *A* and *B*, and its significance was evaluated by permutation analysis. The observed *D*-value was compared with the distribution computed from 400 random permutations of the allele values of locus *B* on the fixed indices of the locus *A* vector. *P*-values and confidence intervals were determined from these permutation tests, as described above for genotype–phenotype associations: the thresholds for unadjusted *P*-values (Fig. 2) were $P < 0.001$ (marginal) and $P < 0.00001$ (highly significant). LD was also calculated using a pairwise R^2 estimate implemented in the DNASP package (Rozas and Rozas, 1999). A concatenated sequence (haplotype) of the associated genes for each isolate was created and aligned according to geographical origins and imported into DNASP. The significance levels of R^2 were evaluated by the χ^2-test option of this package. Nucleotide polymorphism (θ) was also calculated using DNASP.

Acknowledgements

We greatly appreciate the valuable discussions with Drs T. J. C. Anderson and Y.-F. Wang. We also thank Ms Brenda R. Marshall for editorial assistance. This study would not have been possible without the efforts of the *P. falciparum* genome sequencing consortium that includes The Institute of Genome Research, Sanger Center and Stanford University.

References

Adagut, I.S., and Warhurst, D.C. (2001) *Plasmodium falciparum*: linkage disequilibrium between loci in chromosomes 7 and 5 and chloroquine selective pressure in Northern Nigeria. *Parasitology* **123:** 219–224.

Allen, J.D., Brinkhuis, R.F., van Deemter, L., Wijnholds, J., and Schinkel, A.H. (2000) Extensive contribution of the multidrug transporters P-glycoprotein and Mrp1 to basal drug resistance. *Cancer Res* **60:** 5761–5766.

Babiker, H.A., Pringle, S.J., Abdel-Muhsin, A., Mackinnon, M., Hunt, P., and Walliker, D. (2001) High-level chloroquine resistance in Sudanese isolates of *Plasmodium falciparum* is associated with mutations in the chloroquine resistance transporter gene *pfcrt* and the multidrug resistance gene *pfmdr1*. *J Infect Dis* **183:** 1535–1538.

Blackhall, W.J., Pouliot, J.F., Prichard, R.K., and Beech, R.N. (1998) *Haemonchus contortus*: selection at a glutamate-gated chloride channel gene in ivermectin- and moxidectin-selected strains. *Exp Parasitol* **90:** 42–48.

Chakraborti, P.K., Bhatt, K., Banerjee, S.K., and Misra, P. (1999) Role of an ABC importer in mycobacterial drug resistance. *Biosci Rep* **19:** 293–300.

Chen, N., Russell, B., Fowler, E., Peters, J., and Cheng, Q. (2002) Levels of chloroquine resistance in *Plasmodium falciparum* are determined by loci other than *pfcrt and pfmdr1*. *J Infect Dis* **185:** 405–407.

Churchill, G.A., and Doerge, R.W. (1994) Empirical threshold values for quantitative trait mapping. *Genetics* **138:** 963–971.

Cooper, R.A., Ferdig, M.T., Su, X.Z., Ursos, L.M., Mu, J., Nomura, T., *et al.* (2002) Alternative mutations at position 76 of the vacuolar transmembrane protein PfCRT are associated with chloroquine resistance and unique stereospecific quinine and quinidine responses in *Plasmodium falciparum*. *Mol Pharmacol* **61:** 35–42.

Dean, M., Rzhetsky, A., and Allikmets, R. (2001) The human atp-binding cassette (abc) transporter superfamily. *Genome Res* **11:** 1156–1166.

Dent, J.A., Smith, M.M., Vassilatis, D.K., and Avery, L. (2000) The genetics of ivermectin resistance in *Caenorhabditis elegans*. *Proc Natl Acad Sci USA* **97:** 2674–2679.

Djimde, A., Doumbo, O.K., Cortese, J.F., Kayentao, K., Doumbo, S., Diourte, Y., *et al.* (2001) A molecular marker for chloroquine-resistant falciparum malaria. *N Engl J Med* **344:** 257–263.

Duraisingh, M.T., Drakeley, C.J., Muller, O., Bailey, R., Snounou, G., Targett, G.A., *et al.* (1997) Evidence for selection for the tyrosine-86 allele of the *pfmdr1* gene of *Plasmodium falciparum* by chloroquine and amodiaquine. *Parasitology* **114:** 205–211.

Fidock, D.A., Nomura, T., Talley, A.K., Cooper, R.A., Dzekunov, S.M., Ferdig, M.T., *et al.* (2000) Mutations in the *P. falciparum* digestive vacuole transmembrane protein PfCRT and evidence for their role in chloroquine resistance. *Mol Cell* **6:** 861–871.

Foote, S.J., Kyle, D.E., Martin, R.K., Oduola, A.M., Forsyth, K., Kemp, D.J., *et al.* (1990) Several alleles of the multidrug-resistance gene are closely linked to chloroquine resistance in *Plasmodium falciparum*. *Nature* **345:** 255–258.

Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., *et al.* (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419:** 498–511.

Giboda, M., and Denis, M.B. (1988) Response of Kampuchean strains of *Plasmodium falciparum* to antimalarials: *in vivo* assessment of quinine and quinine plus tetracycline; multiple drug resistance *in vitro*. *J Trop Med Hyg* **91:** 205–211.

Good, I.J. (1953) The population frequencies of species and the estimation of population parameters. *Biometrika* **40:** 237–264.

Hedrick, P., and Kumar, S. (2001) Mutation and linkage disequilibrium in human mtDNA. *Eur J Hum Genet* **9:** 969–972.

Knowles, G., Davidson, W.L., Jolley, D., and Alpers, M.P. (1984) The relationship between the in vitro response of *Plasmodium falciparum* to chloroquine, quinine and mefloquine. *Trans R Soc Trop Med Hyg* **78:** 146–150.

Leipe, D.D., Wolf, Y.I., Koonin, E.V., and Aravind, L. (2002) Classification and evolution of P-loop GTPases and related ATPases. *J Mol Biol* **317:** 41–72.

Lewontin, R.C., and Kojima, K. (1960) The evolutionary dynamics of complex polymorphisms. *Evolution* **14:** 450–472.

Long, A.D., and Langley, C.H. (1999) The power of association studies to detect the contribution of candidate genetic loci to variation in complex traits. *Genome Res* **9:** 720–731.

Malin, A.S., and Hall, A.P. (1990) Falciparum malaria resistant to quinine and pyrimethamine-sulfadoxine successfully treated with mefloquine. *Br Med J* **300:** 1175.

Mu, J., Duan, J., Makova, K., Joy, D.A., Huynh, C.Q., Branch, O.H., *et al.* (2002) Chromosome-wide SNPs reveal an ancient origin for *Plasmodium falciparum*. *Nature* **418:** 323–326.

Ouellette, M., Legare, D., and Papadopoulou, B. (1994) Microbial multidrug-resistance ABC transporters. *Trends Microbiol* **2:** 407–411.

Peters, W. (1987) *Chemotherapy and Drug Resistance in Malaria*. London: Academic Press.

Peters, W. (1990) Drug resistance in malaria. *Recent Prog Med* **81:** 749–753.

Pukrittayakamee, S., Supanaranond, W., Looareesuwan, S., Vanijanonta, S., and White, N.J. (1994) Quinine in severe *falciparum* malaria: evidence of declining efficacy in Thailand. *Trans R Soc Trop Med Hyg* **88:** 324–327.

Reed, M.B., Saliba, K.J., Caruana, S.R., Kirk, K., and Cowman, A.F. (2000) Pgh1 modulates sensitivity and resistance to multiple antimalarials in *Plasmodium falciparum*. *Nature* **403:** 906–909.

Rozas, J., and Rozas, R. (1999) DnaSP, Version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* **15:** 174–175.

Sidhu, A.B., Verdier-Pinard, D., and Fidock, D.A. (2002) Chloroquine resistance in *Plasmodium falciparum* malaria parasites conferred by *pfcrt* mutations. *Science* **298:** 210–213.

Skinner-Adams, T., and Davis, T.M. (1999) Synergistic *in vitro* antimalarial activity of omeprazole and quinine. *Antimicrob Agents Chemother* **43:** 1304–1306.

Su, X., Kirkman, L.A., Fujioka, H., and Wellems, T.E. (1997) Complex polymorphisms in an approximately 330 kDa pro-

tein are linked to chloroquine-resistant *P. falciparum* in Southeast Asia and Africa. *Cell* **91:** 593–603.

Su, X., Ferdig, M.T., Huang, Y., Huynh, C.Q., Liu, A., You, J., *et al.* (1999) A genetic map and recombination parameters of the human malaria parasite *Plasmodium falciparum*. *Science* **286:** 1351–1353.

Thimasarn, K., Sirichaisinthop, J., Chanyakhun, P., Palananth, C., and Rooney, W. (1997) A comparative study of artesunate and artemether in combination with mefloquine on multidrug resistant falciparum malaria in eastern Thailand. *SE Asian J Trop Med Public Health* **28:** 465–471.

Volkman, S., and Wirth, D. (1998) Functional analysis of *pfmdr1* gene of *Plasmodium falciparum*. *Methods Enzymol* **292:** 174–181.

WHO (2000) WHO Expert Committee on Malaria. *WHO Tech Rep Series* **892:** 1–74.

Wilson, C.M., Serrano, A.E., Wasley, A., Bogenschutz, M.P., Shankar, A.H., and Wirth, D.F. (1989) Amplification of a gene related to mammalian *mdr* genes in drug-resistant *Plasmodium falciparum*. *Science* **244:** 1184–1186.

Wilson, C.M., Volkman, S.K., Thaithong, S., Martin, R.K., Kyle, D.E., Milhous, W.K., *et al.* (1993) Amplification of *pfmdr 1* associated with mefloquine and halofantrine resistance in *Plasmodium falciparum* from Thailand. *Mol Biochem Parasitol* **57:** 151–160.

Wongsrichanalai, C., Nguyen, T.D., Trieu, N.T., Wimonwattrawatee, T., Sookto, P., Heppner, D.G., *et al.* (1997) *In vitro* susceptibility of *Plasmodium falciparum* isolates in Vietnam to artemisinin derivatives and other antimalarials. *Acta Trop* **63:** 151–158.

Wootton, J.C., Feng, X., Ferdig, M.T., Cooper, R.A., Mu, J., Baruch, D.I., *et al.* (2002) Genetic diversity and chloroquine selective sweeps in *Plasmodium falciparum*. *Nature* **418:** 320–323.

Zhao, J.H., Curtis, D., and Sham, P.C. (2000) Model-free analysis and permutation tests for allelic associations. *Hum Hered* **50:** 133–139.

Zhao, J., Li, W., and Xiong, M.M. (2001) Population based linkage disequilibrium mapping of QTL: An application to simulated data in an isolated population. *Genet Epidemiol* **21** (Suppl. 1)**:** s655–s659.