

A surrogate marker of piperazine-resistant *Plasmodium falciparum* malaria: a phenotype–genotype association study

Benoit Witkowski*, Valentine Duru*, Nimol Khim, Leila S Ross, Benjamin Saintpierre, Johann Beghain, Sophy Chy, Saorin Kim, Sopheakvatey Ke, Nimol Kloeung, Rotha Eam, Chanra Khean, Malen Ken, Kaknika Loch, Anthony Bouillon, Anais Domergue, Laurence Ma, Christiane Bouchier, Rithea Leang, Rekol Huy, Grégory Nuel, Jean-Christophe Barale, Eric Legrand, Pascal Ringwald, David A Fidock, Odile Mercereau-Puijalon, Frédéric Arieu, Didier Ménard



Summary

Background Western Cambodia is the epicentre of *Plasmodium falciparum* multidrug resistance and is facing high rates of dihydroartemisinin–piperazine treatment failures. Genetic tools to detect the multidrug-resistant parasites are needed. Artemisinin resistance can be tracked using the K13 molecular marker, but no marker exists for piperazine resistance. We aimed to identify genetic markers of piperazine resistance and study their association with dihydroartemisinin–piperazine treatment failures.

Methods We obtained blood samples from Cambodian patients infected with *P falciparum* and treated with dihydroartemisinin–piperazine. Patients were followed up for 42 days during the years 2009–15. We established in-vitro and ex-vivo susceptibility profiles for a subset using piperazine survival assays. We determined whole-genome sequences by Illumina paired-reads sequencing, copy number variations by qPCR, RNA concentrations by qRT-PCR, and protein concentrations by immunoblotting. Fisher's exact and non-parametric Wilcoxon rank-sum tests were used to identify significant differences in single-nucleotide polymorphisms or copy number variants, respectively, for differential distribution between piperazine-resistant and piperazine-sensitive parasite lines.

Findings Whole-genome exon sequence analysis of 31 culture-adapted parasite lines associated amplification of the *plasmepsin 2–plasmepsin 3* gene cluster with in-vitro piperazine resistance. Ex-vivo piperazine survival assay profiles of 134 isolates correlated with *plasmepsin 2* gene copy number. In 725 patients treated with dihydroartemisinin–piperazine, multicopy *plasmepsin 2* in the sample collected before treatment was associated with an adjusted hazard ratio (aHR) for treatment failure of 20.4 (95% CI 9.1–45.5, $p < 0.0001$). Multicopy *plasmepsin 2* predicted dihydroartemisinin–piperazine failures with 0.94 (95% CI 0.88–0.98) sensitivity and 0.77 (0.74–0.81) specificity. Analysis of samples collected across the country from 2002 to 2015 showed that the geographical and temporal increase of the proportion of multicopy *plasmepsin 2* parasites was highly correlated with increasing dihydroartemisinin–piperazine treatment failure rates ($r = 0.89$ [95% CI 0.77–0.95], $p < 0.0001$, Spearman's coefficient of rank correlation). Dihydroartemisinin–piperazine efficacy at day 42 fell below 90% when the proportion of multicopy *plasmepsin 2* parasites exceeded 22%.

Interpretation Piperazine resistance in Cambodia is strongly associated with amplification of *plasmepsin 2–3*, encoding haemoglobin-digesting proteases, regardless of the location. Multicopy *plasmepsin 2* constitutes a surrogate molecular marker to track piperazine resistance. A molecular toolkit combining *plasmepsin 2* with *K13* and *mdr1* monitoring should provide timely information for antimalarial treatment and containment policies.

Funding Institut Pasteur in Cambodia, Institut Pasteur Paris, National Institutes of Health, WHO, Agence Nationale de la Recherche, Investissement d'Avenir programme, Laboratoire d'Excellence Integrative “Biology of Emerging Infectious Diseases”.

Copyright This is an Open Access article published under the CC BY 3.0 IGO license which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. In any use of this article, there should be no suggestion that WHO endorses any specific organisation, products or services. The use of the WHO logo is not permitted. This notice should be preserved along with the article's original URL.

Introduction

Antimalarial efficacy of artemisinin-based combination therapies, the first-line treatment for uncomplicated *Plasmodium falciparum* malaria, relies on both fast-acting artemisinin derivatives and long-lasting partner drugs. Resistance to artemisinin, which is now fixed in western Cambodia and observed across

southeast Asia, increases the proportion of parasites surviving a 3 day course of an artemisinin-based combination therapy. Resistance to the partner drug is a greater risk when more parasites survive artemisinin treatment. The reduced efficacy of artemisinin derivatives and partner drugs translates into late treatment failures and prolonged parasite carriage,

Lancet Infect Dis 2016

Published Online

November 3, 2016

[http://dx.doi.org/10.1016/S1473-3099\(16\)30415-7](http://dx.doi.org/10.1016/S1473-3099(16)30415-7)

See Online/Comment

[http://dx.doi.org/10.1016/S1473-3099\(16\)30414-5](http://dx.doi.org/10.1016/S1473-3099(16)30414-5)

See Online/Articles

[http://dx.doi.org/10.1016/S1473-3099\(16\)30409-1](http://dx.doi.org/10.1016/S1473-3099(16)30409-1)

*Contributed equally

Malaria Molecular Epidemiology Unit, Institut Pasteur in Cambodia, Phnom Penh, Cambodia

(B Witkowski PhD, V Duru MSc, N Khim PhD, S Chy BS, S Kim BS, S Ke BS, N Kloeung BS,

A Domergue MSc, R Eam BS, C Khean BS, M Ken BS, K Loch BS, D Ménard PhD); **Malaria Translational Research Unit, Institut Pasteur, Paris, France—Institut Pasteur in Cambodia, Phnom Penh, Cambodia** (B Witkowski,

N Khim, A Bouillon PhD, E Legrand PhD, J-C Barale PhD, D Ménard); **Department of Parasites and Insect Vectors, Institut Pasteur, Paris, France** (B Saintpierre MSc,

J Beghain MSc, E Legrand, O Mercereau-Puijalon PhD, F Arieu PhD); **Plate-forme Génomique, Département Génomes et Génétique, Institut Pasteur, Paris, France** (L Ma MSc, C Bouchier PhD);

Structural Microbiology Unit, Biology of Malaria Targets Group, Department of Structural Biology and Chemistry and CNRS, UMR3528, Institut Pasteur, Paris, France (A Bouillon,

J-C Barale); **Institut Cochin Inserm U1016, Université Paris-Descartes, Sorbonne Paris Cité, and Laboratoire de Parasitologie-Mycologie, Hôpital Cochin, Paris, France** (F Arieu); **National Center for Parasitology, Entomology and Malaria Control, Phnom Penh, Cambodia** (R Leang PhD,

R Huy MD); Laboratoire de Mathématiques Appliquées (MAP5) UMR CNRS 8145, Université Paris Descartes, Paris, France (Prof G Nuel PhD); Department of Microbiology and Immunology and Division of Infectious Diseases, Department of Medicine, Columbia University Medical Center, New York, NY, USA (L S Ross PhD, Prof D A Fidock PhD); and Global Malaria Programme, World Health Organization, Geneva, Switzerland (P Ringwald MD)

Correspondence to: Dr Didier Ménard, Malaria Molecular Epidemiology Unit, Institut Pasteur in Cambodia, 5 Boulevard Monivong, BP 983, Phnom Penh, Cambodia dmenard@pasteur-kh.org

Research in context

Evidence before this study

We searched PubMed for studies on piperazine resistance using the term “resistance” in combination with “falciparum” and “piperazine” on May 19, 2016, without any date or language restrictions, and identified 74 publications. These publications included clinical trials done in 11 countries evaluating the efficacy of dihydroartemisinin–piperazine for the treatment of uncomplicated *Plasmodium falciparum* malaria (26 reports) or asymptomatic infections (one report) and for intermittent preventive treatment of pregnant women (three reports) or infants (five reports). In all studies, cure rates were above 90%, except studies done in Cambodia after the year 2010, for which cure rates ranging from 85% to 40% were recorded. Overall, 26 publications reported susceptibility of parasites collected in 15 countries, studied using in-vitro or ex-vivo assays. Virtually all isolates tested by standard dose–response susceptibility assays (with parasite quantification based on isotopes, Sybr Green, or HRP2) were susceptible to piperazine (<100 nmol/L), except those collected in Cambodia after 2010 and samples collected in China before 1998 (when piperazine monotherapy was intensively used). Piperazine resistance at present appears confined to Cambodia. Resistance is a major concern because alternative therapeutic options are scarce and the reduced cure rates translate into prolonged parasite carriage and increased transmission potential of resistant parasites. To map the geographical extension of piperazine resistance and deploy containment measures to prevent its further spread, rapid detection tests are needed but are lacking at present. Potential molecular signatures associated with piperazine resistance were investigated in 11 studies. The only consistently recorded finding was an increased proportion of single copy *mdr1* parasites in piperazine-resistant areas. This marker is not informative for piperazine resistance because wild-type susceptible parasites can also have a single-copy *mdr1* locus.

Added value of this study

We identified amplification of the *plasmepsin 2–3* gene cluster encoding proteases involved in haemoglobin degradation as

the most significant molecular signature associated with in-vitro resistance to piperazine assessed using the piperazine survival assay. Using a large longitudinal collection of samples collected during clinical efficacy studies of dihydroartemisinin–piperazine done across Cambodia since 2009, we examined 725 *P. falciparum* isolates and found that an increased *plasmepsin 2* gene copy number was strongly associated with dihydroartemisinin–piperazine treatment failures. Patients harbouring multicopy *plasmepsin 2* parasites had a 20 times higher risk of recrudescence during the 42-day post-treatment follow-up (94% sensitivity and 77% specificity). Our retrospective analysis of samples collected in Cambodia during the last decade before and after introduction of dihydroartemisinin–piperazine as first-line treatment showed that the proportion of multicopy *plasmepsin 2* parasites correlated with the increase of dihydroartemisinin–piperazine treatment failure rates, from 2009 to 2015 in western Cambodia and during 2014–15 in eastern Cambodia. In areas of artemisinin resistance, the clinical efficacy of dihydroartemisinin–piperazine at day 42 fell under 90% when the local proportion of multicopy *plasmepsin 2* parasites rose above 22%.

Implications of all the available evidence

Dihydroartemisinin–piperazine failure rates have increased in western Cambodia since 2010 and in eastern Cambodia since 2014. They are caused by parasites that are resistant to both artemisinin and piperazine. Combined analysis of K13 polymorphisms and *plasmepsin 2* copy number represents the first informative molecular signature for dihydroartemisinin–piperazine failures. These molecular markers can now be used to track emergence and dissemination of resistance to artemisinin and piperazine in field populations, especially in areas where piperazine is being or will be recommended in combination with artemisinin derivatives as first-line treatment or in preventive treatment for infants or pregnant women, as developed in African settings.

thereby increasing the transmission potential of drug-resistant infections.

In Cambodia, artesunate–mefloquine was chosen as the first-line drug in 2001. By 2008, the high frequency of treatment failures in western provinces, the epicentre of *P. falciparum* multidrug resistance, led to its replacement with dihydroartemisinin–piperazine in those areas in 2008, and later throughout Cambodia in 2010. In recent years the spread of artemisinin-resistant *P. falciparum*, from western Cambodia to neighbouring provinces,^{1–5} has been followed by a substantial increase in dihydroartemisinin–piperazine failure rates. Failures are estimated to reach 60%,^{6–10} indicating a dramatic expansion of piperazine resistance. Until now, the

detection of piperazine resistance has been based on logistically demanding 42-day follow-up studies of patients treated with dihydroartemisinin–piperazine.¹¹ The in-vitro piperazine survival assay (PSA)⁷ has been shown with in-vitro culture-adapted parasites and freshly collected ex-vivo patient isolates to detect piperazine resistance and treatment failure more reliably than classic dose–response assays.⁷ The in-vitro PSA therefore provides a reliable tool to identify molecular signatures associated with resistance.

Here, we used the phenotypic information from the PSA to identify genetic marker(s) of piperazine resistance and study their association with dihydroartemisinin–piperazine treatment failures.

Methods

Overview

First, the exomes of culture-adapted artemisinin-resistant Cambodian *P. falciparum* lines defined as piperazine-susceptible or piperazine-resistant based on their PSA survival rates⁷ were compared for single-nucleotide polymorphisms (SNPs) and copy number variations (CNVs). This process identified an increased copy number of the *plasmepsin 2-plasmepsin 3* gene cluster as a putative genetic signature associated with in-vitro piperazine resistance. Increased *plasmepsin 2* gene copy number was then assessed as a candidate resistance marker in isolates with documented ex-vivo PSA survival rates and in blood samples collected during the years 2009–15 from Cambodian patients treated with dihydroartemisinin–piperazine and followed up for 42 days. Finally, we investigated the geographical and temporal distribution of multicopy *plasmepsin 2* parasites in the country from 2002 to 2015 and its correlation with dihydroartemisinin–piperazine treatment failures.

Study sites and patients

Patients with *P. falciparum* malaria were enrolled in clinical studies done at health centres located across Cambodia during the years 2009–15 (table 1, figure 1). After obtaining written informed consent, patients were treated with dihydroartemisinin–piperazine (Duo-Cotecxin [dihydroartemisinin 40 mg and piperazine 320 mg], Zhejiang Holley Nanhu Pharmaceutical Co Ltd, Jiaxing City, Zhejiang Province, China) and followed up for 42 days, as previously described.^{7,8,10} The endpoint to assess the efficacy of dihydroartemisinin–piperazine was the proportion of PCR-corrected recrudescence *P. falciparum* infections at day 42.¹¹ All studies were approved by the Ethical Committee for Health Research of the Cambodian Ministry of Health. Clinical trials were registered at the Australian New Zealand Clinical Trials Registry (numbers ACTRN 12615000793516, 12612000184875, 12612000183886, 12612000181808, and 12614000344695).

Procedures

Blood samples were collected into acid-citrate-dextrose tubes (Becton-Dickinson, Franklin Lakes, NJ, USA) before treatment and sent to Institut Pasteur in Cambodia within 24 h. A subset of freshly collected samples was used to do the ex-vivo PSA.⁷ All samples were cryopreserved in glycerol. Red cell pellets were stored at –20°C for molecular studies. Blood spots were prepared on day 0 and when applicable on the day of recrudescence.

Cryopreserved parasites were culture-adapted as described.¹² Susceptibility to piperazine was investigated using in-vitro PSA for culture-adapted parasites and ex-vivo PSA for fresh isolates. Survival rates were assessed microscopically and parasites with a survival

	Number of patients treated and followed up (n=725)	Number of patients classified as recrudescence (PCR-corrected; n=119)	Number of isolates with in-vitro PSA survival data (n=31)	Number of isolates with ex-vivo PSA survival data (n=134)
2009				
Pailin	32	3 (9.4%)	0	0
Preah Vihear	30	0	0	0
2010				
Pailin	21	4 (19.0%)	0	0
Pursat	32	3 (9.4%)	0	0
Rattanakiri	30	0	0	0
2011				
Kratié	51	2 (3.9%)	0	0
Preah Vihear	34	2 (5.9%)	0	0
Pursat	41	7 (17.1%)	0	0
2012				
Battambang	39	12 (30.8%)	19	0
Kampong Speu	22	0	4	0
Kampong Thom	38	0	2	0
Pursat	23	2 (8.7%)	6	0
2013				
Kampot	17	1 (5.9%)	0	0
Kratié	22	0	0	0
Preah Vihear	16	1 (6.3%)	0	0
Rattanakiri	31	1 (3.2%)	0	0
2014				
Mondulakiri	39	4 (10.3%)	0	0
Siemreap	40	25 (62.5%)	0	0
Stungtreng	33	11 (33.3%)	0	0
Rattanakiri	34	5 (14.7%)	0	34
2015				
Mondulakiri	16	4 (25.0%)	0	16
Rattanakiri	54	16 (29.6%)	0	54
Siemreap	17	10 (58.8%)	0	17
Stungtreng	13	6 (46.1%)	0	13

Site location and years of collection are provided for isolates with in-vitro and ex-vivo piperazine survival assay (PSA) profiles (see figure 1 for a map of the study site locations). Data are n or n (%).

Table 1: Proportion of PCR-corrected *Plasmodium falciparum* recrudescence recorded at day 42 in 2009–15 in 12 provinces across Cambodia in patients treated with a 3-day course of dihydroartemisinin–piperazine

rate of at least 10% were considered piperazine-resistant.⁷ *msp1*, *msp2*, and *glurp* polymorphisms were determined to distinguish recrudescence from new infections.¹³ Sequencing of the K13-propeller domain was used to screen for artemisinin resistance.¹ Whole-genome sequencing was done with Illumina paired-reads sequencing.¹ Data were integrated into the Whole-genome Data Manager database¹⁴ and exomes of piperazine-resistant and piperazine-sensitive lines were compared after excluding low-coverage positions (ie, lower than 25% of the genome-wide mean coverage). Genes from highly variable multigene families (*var*, *rifin*, *phist*, and *stevor*) were excluded.¹ SNPs and CNVs were investigated using PlasmocNVScan and the Phen2gen software (appendix).¹⁴ *Plasmepsin 2* and *mdr1*

See Online for appendix



Figure 1: Location of study sites (provinces) where dihydroartemisinin-piperaquine clinical efficacy studies were done in 2009–15 (42-day follow-up)

	Year	Site location	In-vitro PSA survival rate (%)	In-vitro susceptibility to PPQ*	DNA expansion type†
3D7	0.1%	Sensitive	No amplification
6273	2012	Kampong Speu	0.2%	Sensitive	No amplification
6337	2012	Kampong Speu	0.4%	Sensitive	No amplification
6403	2012	Pursat	0.5%	Sensitive	No amplification
6267	2012	Kampong Speu	0.5%	Sensitive	No amplification
6349	2012	Kampong Thom	0.6%	Sensitive	No amplification
6237	2012	Kampong Thom	0.8%	Sensitive	No amplification
6410	2012	Battambang	6.0%	Sensitive	No amplification
6369	2012	Pursat	6.4%	Sensitive	Type 2
6395	2012	Battambang	19.2%	Resistant	No amplification
6341	2012	Pursat	25.8%	Resistant	Type 2
6280	2012	Battambang	28.9%	Resistant	Type 2
6246	2012	Kampong Speu	36.9%	Resistant	No amplification
6293	2012	Battambang	39.3%	Resistant	Type 2
6391	2012	Battambang	39.4%	Resistant	Type 1
6272	2012	Battambang	40.0%	Resistant	Type 2
6218	2012	Battambang	40.8%	Resistant	Type 1
6302	2012	Battambang	42.5%	Resistant	Type 1
6229	2012	Battambang	46.6%	Resistant	Type 1
6443	2012	Battambang	49.6%	Resistant	Type 1
6430	2012	Battambang	51.3%	Resistant	Type 1
6429	2012	Pursat	51.8%	Resistant	Type 1
6365	2012	Battambang	51.8%	Resistant	Type 2
6394	2012	Battambang	56.7%	Resistant	Type 1
6219	2012	Battambang	58.6%	Resistant	Type 3
6408	2012	Battambang	58.7%	Resistant	Type 3

(Table 2 continues on next page)

copy number was determined by qPCR (appendix). Steady-state *plasmepsin 2* mRNA concentrations were measured by RT-qPCR (appendix) and *plasmepsin 2* protein expression by immunoblotting (appendix).

Statistical analysis

Data were analysed with MedCalc version 12 (Mariakerke, Belgium). Kruskal-Wallis or Mann-Whitney tests were used for non-parametric comparisons and Student's *t* test or one-way ANOVA were used for parametric comparisons. For proportions (expressed with percentages and 95% CIs), we used χ^2 or Fisher's exact tests. Manhattan plots were generated using the SNPEVG software.¹⁵ We did an SNP-wise analysis using a homemade script developed by FA and BS and used Fisher's exact test to identify significant SNP differences between piperaquine-resistant and piperaquine-sensitive parasite lines. We tested CNVs for differential distribution between piperaquine-resistant and piperaquine-sensitive parasite lines using a non-parametric Wilcoxon rank-sum test. The Bonferroni and the Benjamini-Hochberg corrections were used to assess genome-wide significance and adjust *p* values when statistical tests were done simultaneously on a single dataset (appendix). Relative risks were estimated using the Mantel-Haenszel test. Associations between a cumulative risk of failure at day 42 and molecular signatures associated with piperaquine resistance were assessed by survival analysis. Curves were compared with the Mantel-Haenszel log-rank test. The Cox proportional-hazards regression model was used to assess the association between parasite genotypes (*K13* mutations, *plasmepsin 2*, and *mdr1* copy number), sampling locations, and treatment responses. A linear regression analysis was used to assess the association between the efficacy of dihydroartemisinin-piperaquine and the proportion of parasites with multicopy *plasmepsin 2*. We deemed *p* values of less than 0.05 as significant.

Role of the funding source

The funders of this study had no role in study design, data collection, data analysis, data interpretation, writing of the report, and the decision to submit. The corresponding author had full access to all data in the study and final responsibility for the decision to submit for publication.

Results

From Sept 15, 2009, to Feb 23, 2015, 725 patients were enrolled in clinical studies to assess the efficacy of the standard 3-day dihydroartemisinin-piperaquine treatment. By 2015, the cumulative proportion of *P. falciparum* recrudescence at day 42 after PCR correction was 16.4% (119 of 725 patients), ranging from 0% to 62.5% depending on the site and the year of study (table 1, figure 1).

Whole-genome sequences were obtained from 31 artemisinin-resistant (*K13* C580Y mutant) culture-adapted parasite lines collected in Cambodia

in 2012, including 23 piperazine-resistant and eight piperazine-sensitive lines as defined by their in-vitro PSA survival rates (table 2). We recorded 120 691 exomic (coding sequence) SNPs. Genome-wide association analyses of SNPs identified significant differences between resistant and sensitive lines at two positions located in adjacent genes on chromosome 4: position 896588 of PF3D7_0420000 (encoding a putative zinc-finger protein; $p < 3.56 \times 10^{-7}$, Fisher's exact test; $p = 0.042$ after Bonferroni correction) and position 908385 of PF3D7_0420100 (encoding a Rio2 Ser–Thr protein kinase; $p < 3.56 \times 10^{-7}$, Fisher's exact test; $p = 0.042$ after Bonferroni correction). However, these positions (and indeed the sequences of both genes) were ambiguous with variable proportions of wild-type and mutant nucleotides, precluding identification of specific resistance-associated mutations, and were not studied further (appendix).

By contrast, signals of gene amplification were detected in the piperazine-resistant group for two adjacent genes from the cluster located on chromosome 14 that encode haemoglobin-digesting proteases known as plasmepsins ($p = 0.03795$ Wilcoxon test with Benjamini-Hochberg correction; figure 2, table 3). Irrespective of piperazine susceptibility, all *plasmepsin 3* sequences were wild type and all *plasmepsin 2* sequences had a Q442H *plasmepsin 2* polymorphism, which has been frequently recorded in reference laboratory lines or wild isolates. The correlation between in-vitro PSA survival rates and *plasmepsin 2–3* copy number was highly significant ($r = 0.83$ [95% CI 0.67–0.91], $p < 0.0001$ for *plasmepsin 2* copy number and $r = 0.85$ [0.71–0.93], $p < 0.0001$ for *plasmepsin 3* copy number). We recorded three different DNA expansion profiles (table 2, appendix). In-vitro PSA survival rates were significantly lower in parasites harbouring DNA expansion type 2 ($n = 6$, median PSA survival rate 34.1% [IQR 25.8–40.0]) compared with those harbouring DNA expansion type 1 ($n = 13$, median PSA survival rate 51.8% [IQR 45.6–61.7], $p = 0.006$, Mann-Whitney test) or type 3 ($n = 3$, median PSA survival rate 58.7%, $p = 0.02$, Mann-Whitney test).

Conversely, a cluster of five genes on chromosome 5 (PF3D7_0531700, PF3D7_0522900, PF3D7_0523000, PF3D7_0523100, and PF3D7_0523200), which included *mdr1*, had increased copy numbers in sensitive lines. *mdr1* was amplified in five of eight piperazine-sensitive lines but in none of the 23 piperazine-resistant lines ($p = 0.015$, Wilcoxon test; appendix).

To confirm the association between *plasmepsin* CNV and ex-vivo PSA survival rate, we used *plasmepsin 2* as an amplicon reporter. First, we optimised a qPCR method to assess *plasmepsin 2* gene copy number (appendix). *Plasmepsin 2* copy number detected by qPCR was 100% concordant with the whole-genome sequencing estimates for the 31 culture-adapted parasites ($p < 0.0001$, Fisher's test). From a set of 134 isolates with known ex-vivo PSA profiles, *plasmepsin 2* was amplified in 67 of

	Year	Site location	In-vitro PSA survival rate (%)	In-vitro susceptibility to PPQ*	DNA expansion type†
(Continued from previous page)					
6224	2012	Pursat	61.4%	Resistant	Type 1
6431	2012	Battambang	61.5%	Resistant	Type 1
6320	2012	Battambang	62.1%	Resistant	Type 1
6261	2012	Pursat	70.5%	Resistant	Type 1
6411	2012	Battambang	71.6%	Resistant	Type 1
6427	2012	Battambang	77.4%	Resistant	Type 3

The last column lists the DNA expansion types recorded in the region of chromosome 14 encoding the plasmepsin 1–4 haemoglobins. *Threshold used to define in-vitro susceptibility to PPQ: sensitive if survival rates were less than 10% and resistant if survival rates were 10% or more. †See appendix for details.

Table 2: Details of the 31 K13-C580Y mutant, piperazine (PPQ)-resistant and PPQ-sensitive culture-adapted parasites analysed by whole-genome sequencing and compared with the 3D7 reference line

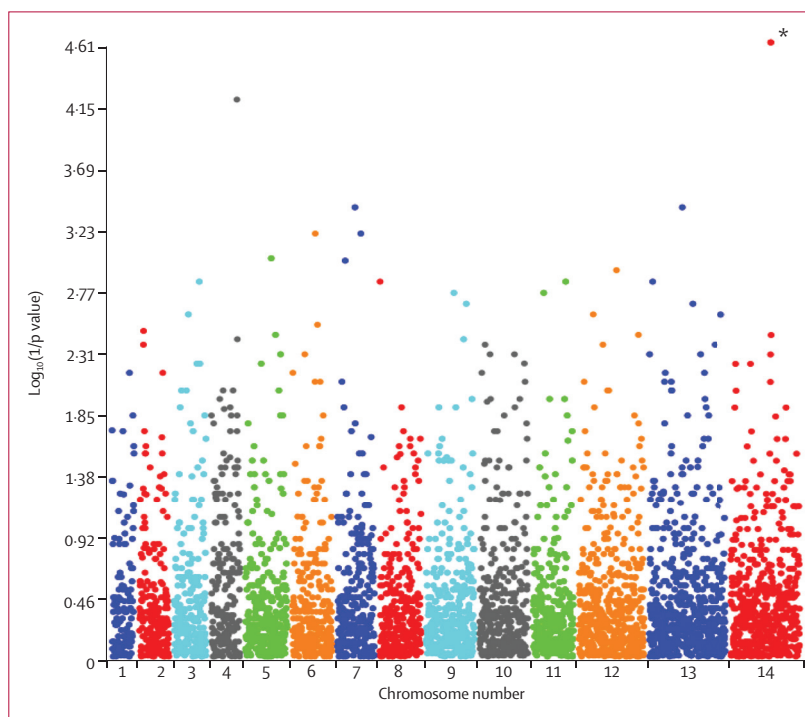


Figure 2: Manhattan plot showing the significance of copy number variations between whole-genome exome sequences of 23 piperazine-resistant and eight piperazine-sensitive culture-adapted lines collected in Cambodia in 2012 and phenotyped using the in-vitro piperazine survival assay Each dot represents a gene in the set of 31 culture-adapted parasites, according to chromosome. The x axis represents genomic location, and the y axis represents the \log_{10} transformed Wilcoxon test p values. *Wilcoxon test $p = 0.139$; after Benjamini-Hochberg correction, only two genes, PF3D7_1408000 (*plasmepsin 2*) and PF3D7_1408100 (*plasmepsin 3*) achieved genome-wide significance ($p = 0.03795$).

69 piperazine-resistant parasites (50, 15, and two isolates with two, three, or four *plasmepsin 2* copies, respectively), and zero of 65 piperazine-susceptible parasites (figure 3). The median ex-vivo PSA survival rate was significantly higher in isolates with at least two *plasmepsin 2* copies compared with those with unamplified *plasmepsin 2* (51.7% [IQR 29.7–75.1] vs 0.004% [0.003–0.39]; $p < 0.0001$, Mann-Whitney test). An increased *plasmepsin 2* copy number predicted ex-vivo piperazine resistance with a sensitivity of 0.97 (95% CI

For more on the Q442H PpPM2 polymorphism see <http://plasmodb.org/plasmo>

	Gene description	Chromosome number	Unadjusted Wilcoxon p value	Bonferroni*	Benjamini-Hochberg*
PF3D7_1408000	<i>Plasmepsin 2</i>	14	2.43×10^{-5}	0.1139	0.03795
PF3D7_1408100	<i>Plasmepsin 3</i> , histo-aspartic protease (HAP)	14	2.43×10^{-5}	0.1139	0.03795
PF3D7_0422000	Steroid dehydrogenase, putative	4	6.54×10^{-5}	0.306	0.0765
PF3D7_0700800	Pfmc-2TM Maurer's cleft two transmembrane protein (MC-2TM)	7	4.22×10^{-4}	1	0.2468
PF3D7_1353100	<i>Plasmodium</i> exported protein, unknown function	13	4.22×10^{-4}	1	0.2468
PF3D7_0713100	Pfmc-2TM Maurer's cleft two transmembrane protein (MC-2TM)	7	6.65×10^{-4}	1	0.3112
PF3D7_0605300	Ser-Thr protein kinase (ARK1)	6	6.65×10^{-4}	1	0.3112
PF3D7_0508400	Transcription factor IIb, putative	5	1.02×10^{-3}	1	0.4143
PF3D7_0715100	Conserved <i>Plasmodium</i> protein, unknown function	7	1.06×10^{-3}	1	0.4143
PF3D7_1211000	Kinesin-like protein, putative	12	1.25×10^{-3}	1	0.4208
PF3D7_1304500	Small heat shock protein, putative	13	1.52×10^{-3}	1	0.4208
PF3D7_1120100	Phosphoglycerate mutase, putative (PGM1)	11	1.52×10^{-3}	1	0.4208
PF3D7_0315600	Conserved <i>Plasmodium</i> protein, unknown function	3	1.52×10^{-3}	1	0.4208
PF3D7_0800700	Surface-associated interspersed gene 8,3 (SURFIN8,3) (SURF8,3)	8	1.52×10^{-3}	1	0.4208
PF3D7_1117700	GTP-binding nuclear protein ran/tc4 (RAN)	11	1.85×10^{-3}	1	0.4572
PF3D7_0909500	Subpellicular microtubule protein 1, putative (SPM1)	9	1.85×10^{-3}	1	0.4572
PF3D7_1310200	Conserved <i>Plasmodium</i> protein, unknown function	13	2.24×10^{-3}	1	0.4995
PF3D7_0925900	Conserved <i>Plasmodium</i> protein, unknown function	9	2.24×10^{-3}	1	0.4995
PF3D7_0322000	Peptidyl-prolyl <i>cis-trans</i> isomerase (CYP19A)	3	2.69×10^{-3}	1	0.5254

*Based on 4422 genes included in the analysis (total of 4678 screened genes; 256 genes with <500 bp were excluded from the final analysis).

Table 3: List of genes with copy number variation most strongly associated with in-vitro piperazine resistance expressed by the piperazine survival assay

0.90–0.99) and specificity of 1.00 (0.65–1.00). *K13* polymorphisms were detected in 65 piperazine-resistant and 17 piperazine-susceptible isolates (figure 3). Only four of 69 piperazine-resistant isolates harboured a wild-type *K13* sequence. In a multiple regression analysis, increased *plasmepsin 2* copy number was more strongly associated than *K13* mutations with in-vitro piperazine resistance ($r_{\text{partial}}=0.94$, $p<0.0001$ and $r_{\text{partial}}=0.25$, $p=0.004$, respectively).

Plasmepsin 2 transcript concentrations were 4.1–5.3 times higher in the piperazine-resistant line ID_6320 compared with the piperazine-sensitive line ID_6267 at all timepoints of the intra-erythrocytic cycle investigated. *Plasmepsin 2* protein concentrations were at least two times higher in piperazine-resistant parasites ID_6408 compared with the sensitive line ID_6267 (appendix). This finding is consistent with increased protein concentrations in the multicopy *plasmepsin 2* lines studied. However, further work is required to expand this analysis to additional lines.

We then explored the association between *plasmepsin 2* CNV and dihydroartemisinin–piperazine treatment outcome in the isolates from 725 patients collected before dihydroartemisinin–piperazine treatment, of whom 119 experienced recrudescence between day 12 and day 42 (figure 4). *Plasmepsin 2* was unamplified in 476 (65.7%) of 725 isolates, had two copies in

153 (21.1%) of 725 isolates, and three or more copies in 96 (13.2%) of 725 isolates. Only seven (1.5%) of 476 patients with unamplified *plasmepsin 2* parasites had recrudescence by day 42 compared with 112 (45.0%) of 249 patients infected with multicopy *plasmepsin 2* parasites (relative risk [RR] 22.8 [95% CI 10.7–48.6], $p<0.0001$). Recrudescence was more frequent for isolates with three or more *plasmepsin 2* copies compared with those with two copies (52 [54.2%] of 96 vs 60 [39.2%] of 153, $p=0.02$). The cumulative incidence of dihydroartemisinin–piperazine treatment failure increased with increasing *plasmepsin 2* gene copies: unamplified versus two copies, hazard ratio (HR) 32.2 (95% CI 17.9–58.0), $p<0.0001$; unamplified versus three copies, HR 49.0 (23.0–104.2), $p<0.0001$; or two copies versus three or more copies, HR 1.53 (1.04–2.25), $p=0.017$ (figure 5A). The mean time to recrudescence decreased with increasing *plasmepsin 2* copy number: 41.9 days (95% CI 41.8–42.0) for patients with unamplified *plasmepsin 2*, 36.0 days (34.6–37.4) for those with two copies, or 34.0 days (32.1–35.0) for those with three or more copies. Increased *plasmepsin 2* copy number predicted dihydroartemisinin–piperazine treatment failures with a sensitivity of 0.94 (95% CI 0.88–0.98) and a specificity of 0.77 (0.74–0.81). The AUC (area under the ROC curve) was 0.86 (95% CI 0.83–0.88),

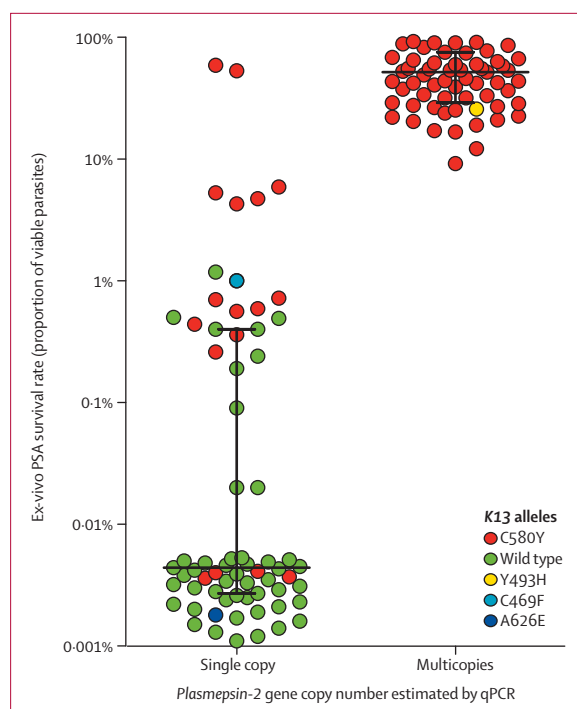


Figure 3: Ex-vivo piperazine survival assay (PSA) survival rates and single ($n=67$) and multicopy *plasmepsin 2* ($n=67$) as estimated by qPCR in isolates collected before dihydroartemisinin-piperazine (DHA-PPQ) treatment stratified by *K13* genotype

Patients were enrolled in clinical studies done in 2014–15 in Mondulakiri, Rattanakiri, Siem Reap, and Stungtreng provinces (see table 1). *K13* polymorphisms were detected in 65 of 69 piperazine-resistant isolates (64 C580Y, one Y493H) and 17 of 65 piperazine-susceptible isolates (15 C580Y, one C469F, and one A626E). Three parasite lines with discordant data were recorded: two resistant lines with non-amplified *plasmepsin 2* and *plasmepsin 3* loci (6246 and 6395) and one sensitive line with two *plasmepsin 2* copies (6369; table 2). The ex-vivo PSA survival rate (%) corresponds to the ratio of number of viable parasites in the PPQ-exposed cultures versus the number of viable parasites in the non-exposed culture.

significantly different from 0.5, the reference value of the null hypothesis ($p>0.0001$).

Among the 725 patients treated with dihydroartemisinin-piperazine, *K13* mutants were detected in 443 (61.1%) of 725 day 0 isolates (figure 5B). Of these, 116 (26.2%) of 443 were from patients who failed dihydroartemisinin-piperazine treatment by day 42 compared with three (1.1%) of 282 from patients harbouring *K13* wild-type parasites (RR 24.6 [95% CI 7.9–76.7], $p<0.0001$). A single *mdr1* gene copy was detected in 610 (84.1%) of 725 day 0 isolates. Dihydroartemisinin-piperazine failures were recorded in 112 (18.4%) of 610 patients infected with parasites harbouring a single *mdr1* copy and seven (6.1%) of 115 patients infected with multicopy *mdr1* parasites (RR 3.0 [95% CI 1.4–6.3], $p=0.003$). The cumulative incidence of dihydroartemisinin-piperazine treatment failure did not increase with increasing age (stratified in three classes: 0–15 years, 16–30 years, and >30 years; $p=0.1809$, log-rank test) or with increasing parasite

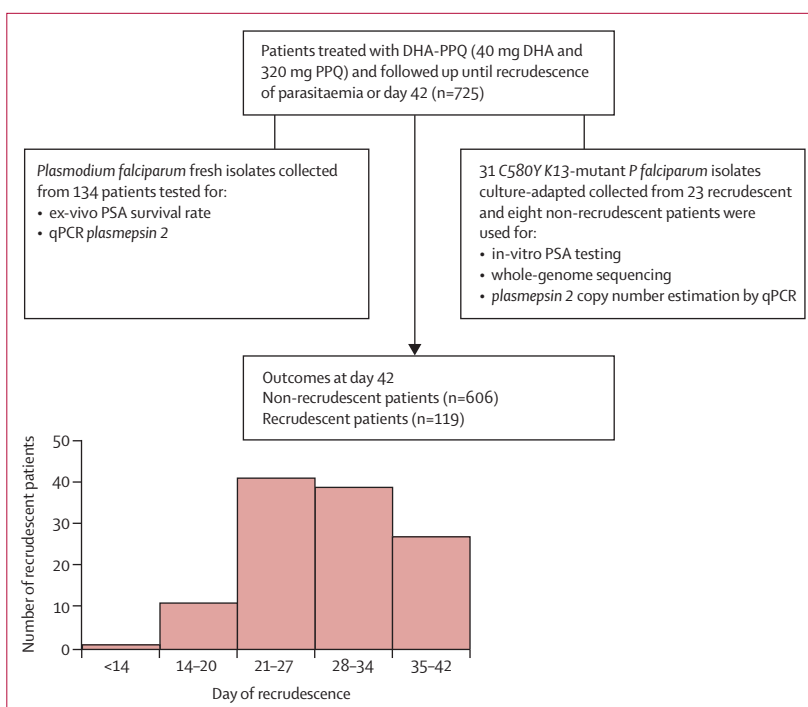


Figure 4: Patients enrolled in clinical studies done in 2009–15 in 12 provinces across Cambodia to assess the efficacy of the 3-day dihydroartemisinin-piperazine (DHA-PPQ) regimen, and isolates used to detect molecular signatures associated with in-vitro piperazine survival assay (PSA) resistance and DHA-PPQ clinical failure

Supervised DHA-PPQ was given once daily for 3 days (day 0, 24 h, 48 h). Dosing was based on bodyweight: less than 19 kg, 40 mg DHA–320 mg PPQ per day; 19–29 kg, 60 mg DHA–480 mg PPQ per day; 30–39 kg, 80 mg DHA–640 mg PPQ per day; greater than 40 kg, 20 mg DHA–960 mg PPQ per day. For children unable to swallow tablets, DHA-PPQ was dissolved in 5 mL of water. Patients were observed for 1 h post-dosing and were re-dosed with a full or half dose if vomiting occurred within 30 min or between 31 and 60 min, respectively. Those who vomited after the second dose were withdrawn from the study and were given parenteral rescue treatment (intramuscular artemether). Patients with axillary temperatures of 37.5°C were treated with paracetamol. Patients were seen daily to day 3 and then weekly for 6 weeks (day 42) for clinical examinations (axillary temperature, symptom check) and malaria blood films. Home visits were done if patients failed to come back for their follow-up appointments. Withdrawn patients, patients lost to follow-up, and patients classified as reinfect (based on *msp1*, *msp2*, and *glurp* genotypes) were excluded from the analysis.

numbers in isolates collected before dihydroartemisinin-piperazine treatment (stratified in four classes: <5000 parasites per μL , 5001–20000 parasites per μL , 20001–50000 parasites per μL , and >50000 parasites per μL ; $p=0.4612$, log-rank test).

After controlling for *K13* and *mdr1* genotypes in a Cox proportional-hazards regression model, *plasmepsin 2* copy number (any increase compared with non-amplification) was the most significant molecular signature associated with dihydroartemisinin-piperazine treatment failure (adjusted HR [aHR] 20.4 [95% CI 9.1–45.5], $p<0.0001$), followed by *K13* mutation (aHR 5.5 [1.7–18.3], $p=0.005$), then *mdr1* single copy (aHR 2.05 [0.95–4.42], $p=0.06$). The cumulative incidence of dihydroartemisinin-piperazine treatment failure among patients harbouring artemisinin-resistant parasites (ie, an artemisinin resistance-associated *K13* mutation) increased significantly with *plasmepsin 2* copy number (unamplified vs two or more copies, seven [3.3%]

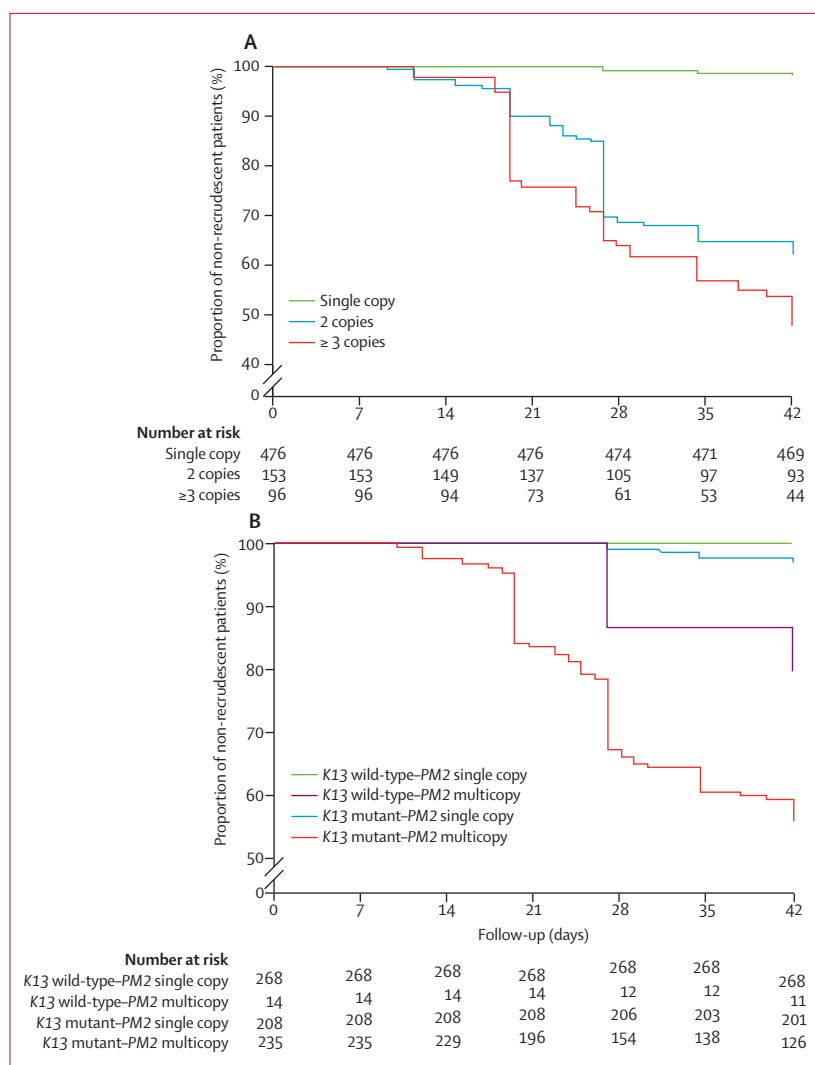


Figure 5: Cumulative proportion of non-recrudescent patients treated with a 3-day course of dihydroartemisinin-piperaquine

(A) *Plasmodium falciparum* (PM2) gene copy number. Log-rank test: $p < 0.0001$ overall; $p < 0.0001$ (hazard ratio [HR] 32.2 [95% CI 17.9–58.0]) for single copy vs two copies; $p < 0.0001$ (HR 49.0 [23.0–104.2]) for single copy vs three or more copies; $p = 0.017$ (HR 1.53 [1.04–2.25]) for two copies vs three or more copies. (B) PM2 gene copy number and K13 genotype detected in isolates collected at the time of enrolment, before treatment. Log-rank test: $p < 0.0001$ overall; $p < 0.0001$ for K13 wild-type-PM2 single copy vs K13 wild-type-PM2 multicopy; $p = 0.002$ for K13 wild-type-PM2 single copy vs K13 mutant-PM2 single copy; $p < 0.0001$ for K13 wild-type-PM2 single copy vs K13 mutant-PM2 multicopy; $p = 0.001$ (HR 6.9 [0.5–96.6]) for K13 wild-type-PM2 multicopy vs K13 mutant-PM2 single copy; $p = 0.07$ (HR 2.6 [1.3–5.5]) for K13 wild-type-PM2 multicopy vs K13 mutant-PM2 multicopy; $p < 0.0001$ (HR 17.5 [12.2–25.2]) for K13 mutant-PM2 single copy vs K13 mutant-PM2 multicopy.

of 208 vs 109 [46.4%] of 235; HR 17.5 [95% CI 12.2–25.2]).

CNVs in *plasmepsin 2* were investigated in 1252 samples collected across Cambodia from 2002 to 2015 (ie, before and after the introduction of dihydroartemisinin-piperaquine). This sample included 527 archived isolates in addition to the 725 studied above (appendix). A longitudinal sampling was done in Pailin (western Cambodia) and Rattanakiri (eastern Cambodia), where dihydroartemisinin-piperaquine was introduced in 2008

and 2010, respectively. In Pailin, the proportion of multicopy *plasmepsin 2* parasites increased from 27.9% (19 of 68) in 2008–09 to 91.2% (52 of 57) in 2014–15. In Rattanakiri, multicopy *plasmepsin 2* parasites were infrequent until 2012–13 (3.2% [one of 31]) but increased to 45.5% (40 of 88) in 2014–15 (appendix). A steady increase of multicopy *plasmepsin 2* parasites after introduction of dihydroartemisinin-piperaquine was recorded in other provinces as well (Preah Vihear, Pursat; appendix).

In the 12 sites where dihydroartemisinin-piperaquine efficacy studies were done in 2009–15, the proportion of multicopy *plasmepsin 2* isolates was negatively correlated with day 42 cure rates ($r = 0.89$ [95% CI 0.77–0.95], $p < 0.0001$; appendix). A Cox regression model showed that the risk of recrudescence following a dihydroartemisinin-piperaquine 3-day course was significantly associated ($p < 2 \times 10^{-16}$) with the presence of multicopy *plasmepsin 2* parasites on day 0 irrespective of the site of enrolment (appendix). A linear regression model showed that the clinical efficacy of dihydroartemisinin-piperaquine at day 42 fell below 90% when the proportion of multicopy *plasmepsin 2* parasites on K13-mutant genetic background rose above 22%.

Discussion

Following reports of increasing failure of artesunate-mefloquine in western Cambodia, dihydroartemisinin-piperaquine was adopted in 2008 in the western provinces and implemented nationwide in 2010. Resistance to this combination has recently accelerated to levels that render it widely ineffective.¹ The dearth of alternatives creates a perilous situation whereby these multi-drug-resistant infections might become untreatable and spread to other regions with endemic malaria.

The strategy used here to search for genetic associations with piperaquine resistance relied on genome-wide sequence comparisons of a set of artemisinin-resistant parasite lines collected in Cambodia in 2012, all harbouring the C580Y artemisinin resistance mutation and presenting in-vitro PSA survival rates indicative of piperaquine resistance or susceptibility. We reasoned that such a focused sampling in a geographically restricted population would reduce the genetic noise of artemisinin responses and population structure. Results identified amplification of the *plasmepsin 2–3* cluster as a putative genetic event associated with piperaquine resistance. To confirm this association across the country, we focused on *plasmepsin 2*, located in the centre of the amplicon. *Plasmepsin 2* amplification strongly correlated with ex-vivo PSA survival rates irrespective of artemisinin susceptibility and was highly predictive of dihydroartemisinin-piperaquine failures in all geographical areas of Cambodia. *Plasmepsin 2* amplification thus represents an informative marker for piperaquine resistance.

The strong association between K13 polymorphisms and *plasmepsin 2* amplification in the Cambodian parasites

studied herein most likely reflects the history of drug selection in Cambodia. The proportion of isolates with different *K13-plasmepsin 2* combinations (appendix) is consistent with a stepwise selection first for artemisinin resistance then for piperazine resistance. This scenario is in line with the delayed appearance of multicopy *plasmepsin 2* parasites in eastern provinces where the emergence of artemisinin resistance was delayed compared with western provinces (appendix). Dihydroartemisinin–piperazine treatment failures were rare in eastern Cambodia by 2013, confirming observations by others⁶ but increased steadily in 2014 to reach a high frequency by 2015. Most treatment failures had a single gene copy of *mdr1* (112 [94.1%] of 119), confirming earlier reports of failures.^{6–9} The presence of single copy *mdr1* is consistent with data reported for in-vitro-selected piperazine-resistant Dd2 parasites¹⁶ and analysis of field samples from Cambodia, suggesting opposing resistance mechanisms against these molecules.¹⁷ We did not observe the *crt* C101F mutation recorded in a piperazine-pressured parasite line selected in vitro. Thus, our data show that although the most informative marker for piperazine resistance is *plasmepsin 2* copy number, mutation of *K13* alongside a single *mdr1* gene copy contributes to the dihydroartemisinin–piperazine failure phenotype. This finding does not exclude the possibility that additional genes contribute to piperazine resistance. In particular, the significance of the mutations observed for PF3D7_0420000 and PF3D7_0420100 is unclear. Whether the notable sequence heterogeneity of both genes reflected ongoing purifying selection associated with piperazine resistance or loss of mefloquine resistance is uncertain. Analysis of a larger number of isolates with documented phenotypes for both mefloquine and piperazine is needed to address this question.

Drug-selected gene amplification is a well-known phenomenon in malaria parasites.^{18–21} The size of the amplicons on chromosome 14 varied depending on the isolate, as reported for *mdr1*.²² Gene amplification, which is more frequent than point mutation in *P. falciparum* parasites,²² is consistent with the remarkably rapid rise and spread of piperazine resistance in Cambodia. Conversely, *mdr1* de-amplification, consistent with regained susceptibility to mefloquine, occurred in Cambodia in recent years,^{6,8,23} and, as shown here, is associated with the emergence of piperazine-resistant strains.

Plasmepsins are expressed during the intra-erythrocytic asexual blood stage cycle and by sexual stage gametocytes that can be transmitted to the mosquito vector. All four plasmepsins are located in the digestive vacuole of intra-erythrocytic developmental forms where they engage in different steps of haemoglobin degradation. Studies of parasites disrupted in the *plasmepsin* genes pointed to redundancy in the haemoglobin degradation machinery.²⁴ To our knowledge, there are no reported studies about the consequences of overexpressing these proteases. We show here that *plasmepsin 2* amplification

is associated with a notable increase of steady-state mRNA and protein concentrations in two culture-adapted isolates. This observation needs to be confirmed with additional isolates. A reasonable hypothesis is that the amplification of plasmepsins overcomes the inhibitory effect of piperazine on haemoglobin degradation and haem detoxification, possibly by reducing concentrations of reactive haem species that are preferred substrates for piperazine binding. Piperazine-treated trophozoites have been shown to possess large digestive vacuoles containing membrane-bound packets of undigested haemoglobin.²⁵ The observation that piperazine-resistant parasites have a single *mdr1* copy is consistent with this scenario, since maintenance of a single *mdr1* copy (or reversion to a single copy) might avoid importing excessive amounts of piperazine into the digestive vacuole (appendix).^{26,27}

We note that the association of piperazine resistance with amplification of the *plasmepsin 2–3* cluster on chromosome 14 is not proof of causality. The structured populations of *P. falciparum* parasites in Cambodia²⁸ might confound the robustness of the association and additional loci might also contribute to piperazine resistance. The present findings should be complemented with laboratory investigations of the cellular consequences of this amplification on the parasite response to piperazine and on parasite fitness and transmissibility. Nonetheless, our data are timely in providing a molecular tool that predicts the appearance of piperazine resistance in endemic settings.

Piperazine is a well-tolerated partner drug used in combination with artemisinin derivatives or the ozonide compound arterolane (OZ277).²⁹ The mechanism of piperazine resistance in the specific context of Cambodia, where artemisinin resistance is nearly fixed and drug pressure is high, might not extrapolate to areas where artemisinin resistance has not yet been documented. Nevertheless, we propose to extend the assessment of *plasmepsin 2* gene copy number to areas where piperazine is being used in artemisinin-based combination therapies at a very large scale, and to combine this assay with *K13* sequencing to localise areas of parasite resistance to both components. In Cambodia, where the rapid failure of first-line artemisinin-based combination therapies is jeopardising elimination efforts and accelerating the emergence and spread of resistance, the opposing susceptibility between mefloquine and piperazine could be used to implement new strategies based on artemisinin-based combination drug rotation, sequential administration, or triple combinations including both artemisinin-based combination partner drugs. Although challenging to implement, these alternative strategies will help to ensure long-term efficacy of antimalarials to reach the elimination goal.

Contributors

BW, VD, LSR, FA, J-CB, DAF, OM-P, and DM contributed to study design. BS, JB, LM, CB, and FA analysed the whole-genome sequencing

data. NKH, SC, RE, CK, MK, KL, and EL did qPCR and RT-qPCR and LSR and AB did the antibody analysis. NKI, SKE, AD, VD, and BW did the in-vitro and ex-vivo drug assays. RL, RH, SKI, and PR gathered the clinical data. GN did the statistical analysis. VD, BW, FA, and DM analysed the data. OM-P, FA, LSR, DAF, and DM wrote the first manuscript. All authors read and approved the final manuscript.

Declaration of interests

All authors declare no competing interests. PR is a staff member of WHO. The authors are responsible for the views expressed in this publication, which do not necessarily represent the decisions, policy, or views of WHO.

Acknowledgments

We thank all patients enrolled in the therapeutic efficacy studies, the health centres' staff, and the members of the National Center for Parasitology, Entomology and Malaria Control in Cambodia for their support. We are grateful to Daniel E Goldberg (Washington University School of Medicine, St Louis, MI, USA) for providing anti-plasmeprin 2 antibody. This work was supported by the Institut Pasteur in Cambodia, Institut Pasteur Paris, and WHO. DAF gratefully acknowledges partial funding for this work from the National Institutes of Health (NIH; R01 AI109023 and AI124678). LSR is a recipient of an NIH NRSA fellowship F32 AI120578. OM-P acknowledges funding from the French Government's Investissement d'Avenir programme, Laboratoire d'Excellence "Integrative Biology of Emerging Infectious Diseases" (grant number ANR-10-LABX-62-IBED). VD was supported by a doctoral fellowship from the International Division, Institut Pasteur.

References

- 1 Arley F, Witkowski B, Amaratunga C, et al. A molecular marker of artemisinin-resistant *Plasmodium falciparum* malaria. *Nature* 2014; **505**: 50–55.
- 2 Ashley EA, Dhorda M, Fairhurst RM, et al. Spread of artemisinin resistance in *Plasmodium falciparum* malaria. *N Engl J Med* 2014; **371**: 411–23.
- 3 Takala-Harrison S, Jacob CG, Arze C, et al. Independent emergence of artemisinin resistance mutations among *Plasmodium falciparum* in Southeast Asia. *J Infect Dis* 2015; **211**: 670–79.
- 4 Dondorp AM, Nosten F, Yi P, et al. Artemisinin resistance in *Plasmodium falciparum* malaria. *N Engl J Med* 2009; **361**: 455–67.
- 5 Noedl H, Se Y, Schaefer K, et al. Evidence of artemisinin-resistant malaria in western Cambodia. *N Engl J Med* 2008; **359**: 2619–20.
- 6 Amaratunga C, Lim P, Suon S, et al. Dihydroartemisinin-piperaquine resistance in *Plasmodium falciparum* malaria in Cambodia: a multisite prospective cohort study. *Lancet Infect Dis* 2016; **16**: 357–65.
- 7 Duru V, Khim N, Leang R, et al. *Plasmodium falciparum* dihydroartemisinin-piperaquine failures in Cambodia are associated with mutant K13 parasites presenting high survival rates in novel piperaquine in vitro assays: retrospective and prospective investigations. *BMC Med* 2015; **13**: 305.
- 8 Leang R, Taylor WR, Bouth DM, et al. Evidence of *Plasmodium falciparum* malaria multidrug resistance to artemisinin and piperaquine in western Cambodia: dihydroartemisinin-piperaquine open-label multicenter clinical assessment. *Antimicrob Agents Chemother* 2015; **59**: 4719–26.
- 9 Spring MD, Lin JT, Manning JE, et al. Dihydroartemisinin-piperaquine failure associated with a triple mutant including *kelch13* C580Y in Cambodia: an observational cohort study. *Lancet Infect Dis* 2015; **15**: 683–91.
- 10 Leang R, Barrette A, Bouth DM, et al. Efficacy of dihydroartemisinin-piperaquine for treatment of uncomplicated *Plasmodium falciparum* and *Plasmodium vivax* in Cambodia, 2008 to 2010. *Antimicrob Agents Chemother* 2013; **57**: 818–26.
- 11 WHO. Methods for surveillance of antimalarial drug efficacy. Geneva: WHO Press; 2009.
- 12 Cranmer SL, Magowan C, Liang J, Coppel RL, Cooke BM. An alternative to serum for cultivation of *Plasmodium falciparum* in vitro. *Trans R Soc Trop Med Hyg* 1997; **91**: 363–65.
- 13 WHO. Methods and techniques for clinical trials on antimalarial drug efficacy: genotyping to identify parasite populations. Geneva: WHO Press; 2007.
- 14 Beghain J, Langlois AC, Legrand E, et al. *Plasmodium* copy number variation scan: gene copy numbers evaluation in haploid genomes. *Malaria J* 2016; **15**: 206.
- 15 Wang S, Dvorkin D, Da Y. SNPEVG: a graphical tool for GWAS graphing with mouse clicks. *BMC Bioinformatics* 2012; **13**: 319.
- 16 Eastman RT, Dharia NV, Winzeler EA, Fidock DA. 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.
- 17 Veiga MI, Dhinra SK, Henrich PP, et al. Globally prevalent PfMDR1 mutations modulate *Plasmodium falciparum* susceptibility to artemisinin-based combination therapies. *Nat Commun* 2016; **7**: 11553.
- 18 Cowman AF, Galatis D, Thompson JK. Selection for mefloquine resistance in *Plasmodium falciparum* is linked to amplification of the *pfmdr1* gene and cross-resistance to halofantrine and quinine. *Proc Natl Acad Sci USA* 1994; **91**: 1143–47.
- 19 Price RN, Uhlemann AC, Brockman A, et al. Mefloquine resistance in *Plasmodium falciparum* and increased *pfmdr1* gene copy number. *Lancet* 2004; **364**: 438–47.
- 20 Sidhu AB, Uhlemann AC, Valderramos SG, Valderramos JC, Krishna S, Fidock DA. Decreasing *pfmdr1* copy number in *Plasmodium falciparum* malaria heightens susceptibility to mefloquine, lumefantrine, halofantrine, quinine, and artemisinin. *J Infect Dis* 2006; **194**: 528–35.
- 21 Heinberg A, Siu E, Stern C, et al. Direct evidence for the adaptive role of copy number variation on antifolate susceptibility in *Plasmodium falciparum*. *Mol Microbiol* 2013; **88**: 702–12.
- 22 Nair S, Nash D, Sudimack D, et al. Recurrent gene amplification and soft selective sweeps during evolution of multidrug resistance in malaria parasites. *Mol Biol Evol* 2007; **24**: 562–73.
- 23 Lim P, Dek D, Try V, Sreng S, Suon S, Fairhurst RM. Decreasing *pfmdr1* copy number suggests that *Plasmodium falciparum* in western Cambodia is regaining in vitro susceptibility to mefloquine. *Antimicrob Agents Chemother* 2015; **59**: 2934–37.
- 24 Bonilla JA, Moura PA, Bonilla TD, Yowell CA, Fidock DA, Dame JB. Effects on growth, hemoglobin metabolism and paralogous gene expression resulting from disruption of genes encoding the digestive vacuole plasmepsins of *Plasmodium falciparum*. *Int J Parasitol* 2007; **37**: 317–27.
- 25 Sachanonta N, Chotivanich K, Chaisri U, et al. Ultrastructural and real-time microscopic changes in *P. falciparum*-infected red blood cells following treatment with antimalarial drugs. *Ultrastruct Pathol* 2011; **35**: 214–25.
- 26 Rohrbach P, Sanchez CP, Hayton K, et al. Genetic linkage of *pfmdr1* with food vacuolar solute import in *Plasmodium falciparum*. *EMBO J* 2006; **25**: 3000–11.
- 27 Sanchez CP, Rotmann A, Stein WD, Lanzer M. Polymorphisms within PfMDR1 alter the substrate specificity for anti-malarial drugs in *Plasmodium falciparum*. *Mol Microbiol* 2008; **70**: 786–98.
- 28 Miotto O, Almagro-Garcia J, Manske M, et al. Multiple populations of artemisinin-resistant *Plasmodium falciparum* in Cambodia. *Nat Genet* 2013; **45**: 648–55.
- 29 Valecha N, Krudsood S, Tangpukdee N, et al. Artesunate maleate plus piperaquine phosphate for treatment of uncomplicated *Plasmodium falciparum* malaria: a comparative, multicenter, randomized clinical trial. *Clin Infect Dis* 2012; **55**: 663–71.

THE LANCET Infectious Diseases

Supplementary webappendix

This webappendix formed part of the original submission and has been peer reviewed.
We post it as supplied by the authors.

Supplement to: Witkowski B, Duru V, Khim N, et al. A surrogate marker of piperazine-resistant *Plasmodium falciparum* malaria: a phenotype-genotype association study. *Lancet Infect Dis* 2016; published online Nov 3. [http://dx.doi.org/10.1016/S1473-3099\(16\)30415-7](http://dx.doi.org/10.1016/S1473-3099(16)30415-7).

APPENDIX MATERIAL

A surrogate marker of piperaquine-resistant *Plasmodium falciparum* malaria

Benoit Witkowski*, Valentine Duru*, Nimol Khim, Leila S Ross, Benjamin Saintpierre, Johann Beghain, Sophy Chy, Saorin Kim, Sopheakvatey Ke, Nimol Kloeung, Rotha Eam, Chanra Khean, Malen Ken, Kaknika Loch, Anthony Bouillon, Anais Domergue, Laurence Ma, Christiane Bouchier, Rithea Leang, Rekol Huy, Grégory Nuel, Jean-Christophe Barale, Eric Legrand, Pascal Ringwald, David A Fidock, Odile Mercereau-Puijalon, Frédéric Arieu, Didier Ménar

*Contributed equally

CONTENTS

Appendix 1: Detailed methods: Piperaquine Survival Assays (in-vitro and ex-vivo), DNA, RNA and protein extraction, Whole-genome sequencing and immunoblotting.

Appendix 2: *PfPM2* and *Pfmdr1* copy number determination, with listing of primers, protocols and PCR amplification efficiencies.

Appendix 3: *PfPM2* mRNA expression profile, with listing of primers, protocols and RT-qPCR amplification efficiencies.

Appendix 4: Panel A. Manhattan plot showing the significance of single nucleotide polymorphisms (SNPs) between whole-genome exome sequences of 23 piperaquine-resistant and 8 piperaquine-sensitive culture-adapted lines phenotyped using in-vitro PSA. **Panel B.** List of the positions with variable proportions of wild type and mutant nucleotides of PF3D7_0420000 (encoding a putative zinc finger protein) and PF3D7_0420100 (encoding the Rio2 serine/threonine protein kinase) sequences of 23 piperaquine-resistant and 8 piperaquine-sensitive culture-adapted lines phenotyped using in-vitro PSA.

Appendix 5: Profiles of DNA expansion in the region of genes encoding proteins involved in hemoglobin-degrading activities positively associated with *in vitro* piperaquine resistance and methodology developed to confirm the four DNA expansion profiles.

Appendix 6: Distribution of the Wilcoxon Rank-Sum test p-values ranking the significance of CNVs of the 4,616 genes screened between whole-genome exome sequences of 23 piperaquine-resistant and 8 piperaquine-sensitive culture-adapted lines phenotyped using in-vitro PSA.

Appendix 7: Panel A. *PfPM2* mRNA transcript levels, relative to *Pfserine-tRNA ligase* mRNA in early ring stages, late ring stages, early trophozoite stages and mature trophozoites of in-vitro culture-adapted piperaquine-resistant and piperaquine-sensitive parasite lines.

Panel B. *PfPM2* expression in *Plasmodium falciparum* synchronized trophozoite-stage of culture-adapted piperaquine-resistant and piperaquine-sensitive parasite lines detected by western immunoblot.

Appendix 8: Spatio-temporal increase in frequency of parasites with multicopy *PfPM2* in Cambodia between 2002 and 2015. **Panel A.** Increase in frequency of parasites with multicopy *PfPM2* in Pailin (red histogram) and Rattanakiri (black histogram) between 2002 and 2015. **Panel B.** Dynamic of the spread of multicopy *PfPM2* parasites in 10 Cambodian provinces from 2002 to 2015

Appendix 9: Correlation between the proportion of parasites with multicopy *PfPM2* and DHA-PPQ treatment failure rates recorded in 12 sites across Cambodia from 2009 to 2015.

Appendix 10: Proportions of isolates with different genetic background (K13 wild-type/*PfPM2* single copy, K13 mutant/*PfPM2* single copy, K13 mutant/*PfPM2* multicopy and K13 wild-type/*PfPM2* multicopy).

Appendix 11: Hypothesis supporting the mechanisms of resistance of *P. falciparum* parasites to PPQ through the amplification of *PfPM2* and *PfPM3* genes, and the deamplification of the *Pfmdr1* gene in Cambodia.

Appendix 1: Detailed methods: Piperaquine Survival Assays (in-vitro and ex-vivo), DNA, RNA and protein extraction, Whole-genome sequencing and immunoblotting.

Piperaquine Survival assays (PSA). In-vitro susceptibility to PPQ was investigated using the in-vitro or the ex-vivo PSA, which is based on exposing very early ring-stage parasites to 200 nM PPQ for 48 hours, washing away the drug, and assessing parasite growth after a further 24 hours of culture. Survival rates at the 72 hour time point were assessed microscopically by counting the proportion of viable parasites in exposed and non-exposed cultures that developed into second-generation rings or trophozoites with normal morphology. Parasites with a survival rate $\geq 10\%$ were considered PPQ-resistant.⁵

DNA, RNA and protein extraction. Parasite DNA was extracted from blood spots with Instagen matrix (Bio-Rad, Marnes-la-Coquette, France), and from whole blood or cultured parasites with QIAamp DNA Blood MiniKit (Qiagen, Valencia, CA). Total RNA was isolated from cultured parasites using a Trizol reagent-based protocol (Life Technologies, Courtaboeuf, France) and purified with the RNeasy Mini Kit (Qiagen, Valencia, CA). Samples were DNase-treated (Life Technologies) to remove any contaminating genomic DNA. Proteins were extracted from cultured parasites that had been lysed with 0.15% saponin in PBS. The parasite pellet was washed four times with PBS, resuspended at 400,000 parasites per μL of PBS with 1x protease inhibitor cocktail (Sigma-Aldrich, St. Louis, MO, USA) and lysed with a BioRuptor Twin (10 cycles of 10 seconds each, low intensity).

Whole-genome sequencing. Image analysis, base calling and error estimation used the Illumina Analysis Pipeline version 1.7. Raw sequence files were filtered using Fqquality tool. Trimmed reads from controlled Fastq files were mapped onto the *P. falciparum* 3D7 reference genome with the Burrows-Wheeler Alignment (BWA), generating BAM files (binary files of tab-delimited format SAM). Samtools was used to prepare pileup files, which were formatted using in-house software to integrate the data into the Whole-genome Data Manager (WDM) database.¹⁵ Exomes of PPQ-resistant and -sensitive culture-adapted lines were compared after excluding positions with coverage lower than 25% of the genome-wide mean. SNPs were explored after excluding genes from highly variable multi-gene families (*var*, *rifin*, *phist* and *stevor*), as described.¹

Analysis of Single Nucleotide Polymorphisms (SNPs)

The analysis was based on the data generated by the samtools Pileup tools (based on sam file), which incorporates short indels information by correcting the effect of flanking tandem repeats. For each position, any information present in at least 20% of the sequences of that position ($\geq 20\%$ coverage) was considered. Sequence polymorphism was presented using uppercase IUPAC codes in case of base substitutions and a lowercase IUPAC code in case of short indels. This allowed included in a single analysis both types of mutations (30 different codes). The IUPAC Nucleotide code is: A: Adenine, C: Cytosine, G: Guanine, T: Thymine, R: A or G, Y: C or T, S: G or C, W: A or T, K: G or T, M: A or C, B: C or G or T, D: A or G or T, H: A or C or T, V: A or C or G, N: any base. If an indel is observed at a given position, letters are shown in lowercase.

For each nucleotide position, a contingency table was made: the number of columns was variable according to the number of variants observed whereas the line number was constant (sensitive and resistant according to the PSA value). 3D7 was included in the piperaquine-sensitive set. The Bonferroni correction was used to evaluate the significance of the observed polymorphism.

Analysis of Copy Number Variations (CNVs)

Copy Number Variations (CNVs) (after excluding indels) were investigated using PlasmocNVScan and Phen2gen software's.¹⁵ PlasmocNVScan is a C/C++ software that normalizes read depth (coverage) across the entire genome, thus by-passing the risk that the sequencing process is not uniform (i.e., the number of reads mapped to a region is assumed to follow a Poisson distribution and is proportional to the number of copies) and the need of reads mapped against a well-annotated reference genome. The underlying concept of ReadDepth-based methods is that the depth of coverage in a genomic segment is correlated with the copy number of the segment. Specifically, we first computed the average frequency for each motif (6-mer) across the whole exome, generating the genome coverage for each motif. We next recorded the local coverage for a motif at each position (extracted from the pileup file). Then, for each gene, we used a sliding window, computed the ratio between observed coverage and genome coverage for each position and calculated the average coverage for that gene. In the final analysis, genes with nucleotide sequence lengths less than 500 bp were excluded. We considered CNV as a continuous variable and used the Wilcoxon test to compare CNVs of PPQ-resistant and -sensitive parasite lines. CNVs were also classified according to their Wilcoxon rank-sum values. The Bonferroni and Benjamini-Hochberg corrections were as used to evaluate the significance of the CNVs.

Immunoblotting. Parasite lysates (synchronized trophozoite-stage cultures; 24-30 hours post invasion), were mixed with complete Laemmli buffer and boiled for 10 minutes at 95°C. Samples were run on a 10% Tris-Gly-SDS precast gel (BioRad) at 120V for 2 h with a Precision Kaleidoscope protein marker (Biorad). The gel contents were transferred to a nitrocellulose membrane (315 mA 90 minutes). Membranes were blocked with 2% nonfat dry milk and 1% BSA in TBS for 90 minutes at room temperature. Membranes were probed with antibody diluted in the blocking buffer at 4°C overnight. Dilutions used were 1:2,000 for anti-Plasmeprin2 (gift from Dan Goldberg) and 1:5000 for anti-beta actin (NovusBio). Membranes were washed in TBST, then probed with the appropriate 1:10,000 secondary antibody in blocking buffer for one hour at room temperature. Membranes were washed with TBS, then treated with ECL (Pierce) and exposed to film.

Appendix 2. *PfPM2* and *Pfmdr1* copy number determination, with listing of primers, protocols and PCR amplification efficiencies.

PfPM2 (PF3D7_1408000) and *Pfmdr1* (PF3D7_0523000) copy numbers were measured by qPCR using a CFX96 real-time PCR machine (Bio-Rad). As a reference, we used the single copy *β-tubulin* (PF3D7_1008700) gene. Listing of primers, protocols and PCR amplification efficiencies are provided in the table below.

qPCR	Primer Sequence	Sequences	T _m (°C)	Product size (bp)	Range of Melt T°C
<i>PfPM2</i>	PfPM2_CN_F	5'-TGGTGATGCAGAAGTTGGAG-3'	59.8	79	76.8 - 77.2
	PfPM2_CN_R	5'-TGGGACCCATAAATTAGCAGA-3'	59.4		
	Pf <i>β-tubulin</i> _CN_F	5'-TGATGTGCGCAAGTGATCC-3'	61.9	79	79.0 - 79.2
	Pf <i>β-tubulin</i> _CN_R	5'-TCCTTTGTGGACATTCTCCTC-3'	60.5		
<i>Pfmdr1</i>	Pfmdr1_CN_F	5'-TGCATCTATAAAACGATCAGACAAA-3'	60.0	87	77.8 - 78.0
	Pfmdr1_CN_R	5'-TCGTGTGTTCCATGTGACTGT-3'	60.0		
	Pf <i>β-tubulin</i> _CN_F	5'-TGATGTGCGCAAGTGATCC-3'	61.9	79	79.0 - 79.2
	Pf <i>β-tubulin</i> _CN_R	5'-TCCTTTGTGGACATTCTCCTC-3'	60.5		

PfPM2 copy number.

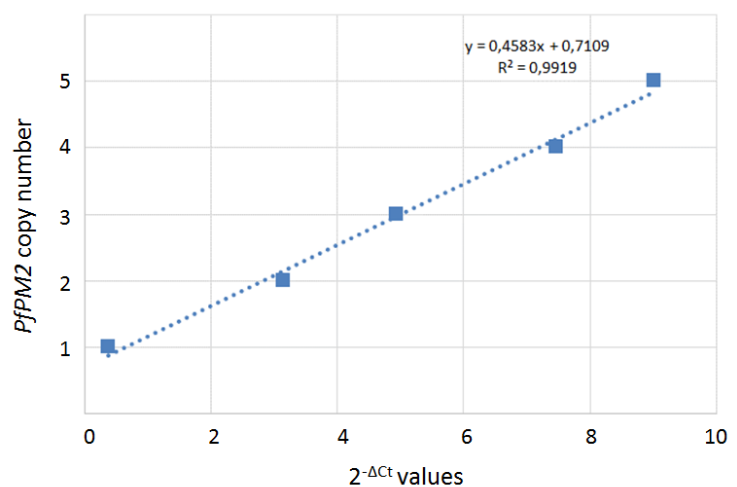
Quantitative PCR (qPCR) was carried out in 20 µl volumes in a 96-well plate containing 5X HOT FIREPol EvaGreen qPCR Mix ROX (Solis BioDyne, Estonia), 0.25 µM of each forward and reverse primer and 4 µl of template DNA. Final MgCl₂ concentrations were 2.5 mM and 4 mM for *PfPM2* and *Pfβ-tubulin*, respectively. Amplifications were performed under the following conditions: 95°C for 15min, followed by 45 cycles of 95°C for 15s, 58°C for 20s, and 72°C for 20s.

PfPM2 copy number of each sample was measured in triplicate relative to a standard curve using five standards of mixed synthetic gene fragments (Eurofins Genomics, Ebersberg, Germany). The lengths of the synthetic fragments for *PfPM2* and *Pfβ-tubulin*, including gene locations are for *PfPM2* (PF3D7_1408000, from position 367 to 560, 193 bp): 367-aggtagttcaaatgataatcgaattagtagatttccaaaatataatgtttta**tggtgatgcagaagttggagataa**

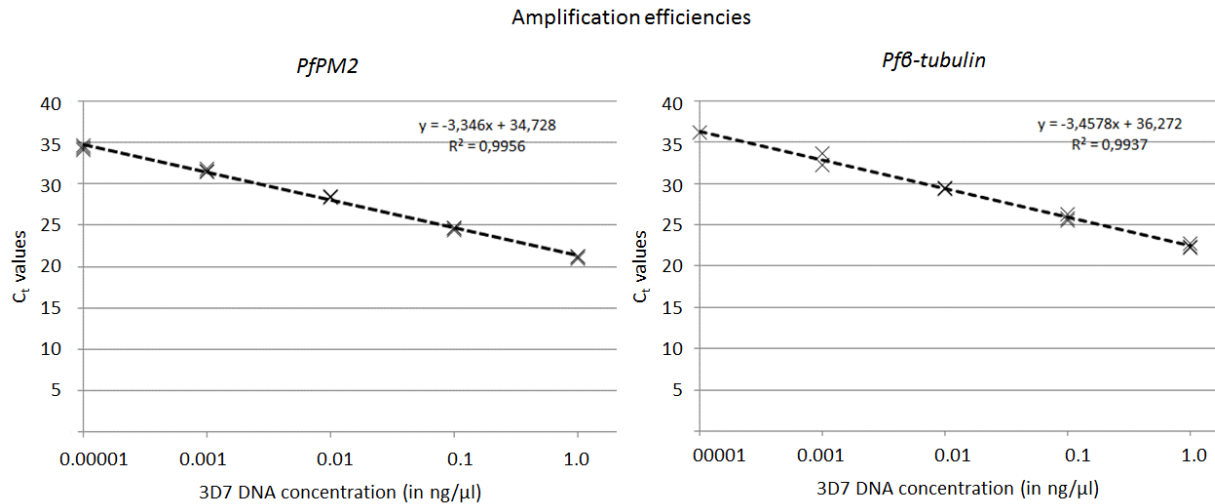
ccaacaaccatttacatttattcttgatacaggatctgctaatttatgggtcccaagtggttaaatgtacaactgcaggatgtttaactaa catctatatgattcatctaaatc-560; and for *Pfβ-tubulin* (PF3D7_1008700): from positions 1183 to 1391 (208 bp): 1183-tcaacaatacagagcctaactgtgccggagttaacacaacaaatgttcgacgcaaaaaata**tgatgtgccaag** **tgatccaagacatggaagatatttaacggcatgtgctatgttttagaggaagaatgtccacaaagga**agttgacgaacaaatgttaa acgttcaaaaataaaaactcatctattttgtcgaatggattcctcac-1391 (shown in bold font, the qPCR amplified portion).

The five standards of mixed synthetic gene fragments were: standard 1 (1:1 molar ratio of *PfPM2* & *β-tubulin*), standard 2 (2:1 molar ratio of *PfPM2* & *β-tubulin*), standard 3 (3:1 molar ratio of *PfPM2* & *β-tubulin*), standard 4 (4:1 molar ratio of *PfPM2* & *β-tubulin*) and standard 5 (5:1 molar ratio of *PfPM2* & *β-tubulin*).

PfPM2 standard curve



The 3D7 (Africa) line was included in each run as control (one copy of *PfPM2*). *PfPM2* copy number was calculated by the $2^{-\Delta C_t}$ method ($\Delta C_t = C_t \text{ PfPM2} - C_t \text{ Pf}\beta\text{-tubulin}$ where C_t is the threshold cycle) and deduced from the standard curve. A *PfPM2* copy number >1.6 was defined as an amplification of the gene. Amplification efficiencies of the *PfPM2* and the *Pf}\beta\text{-tubulin}* genes, measured using ten-fold dilutions of 3D7 DNA, were similar (99% and 95%, respectively).



***Pfmdr1* copy number.**

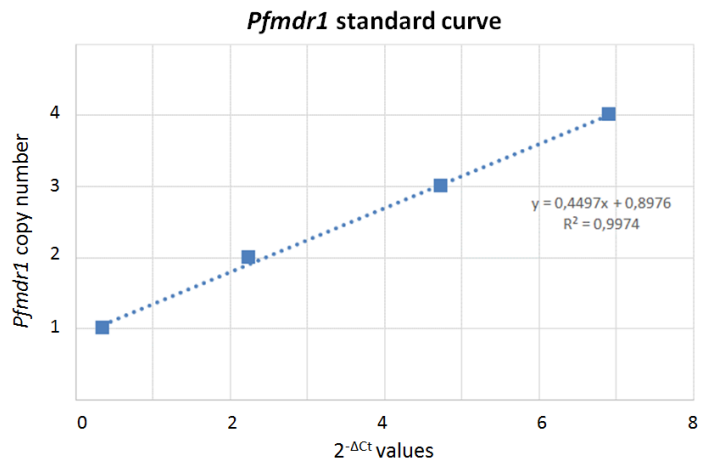
qPCRs were carried out in 25 μl volumes in a 96-well plate containing 5X HOT FIREPol EvaGreen qPCR Mix Plus (Solis BioDyne, Estonia), 0.3 μM concentrations of each forward and reverse primers, and 4 μl of template DNA.

Amplifications were performed under the following conditions: 94°C for 15min, followed by 40 cycles of 94°C for 15s, 58°C for 20s, and 72°C for 20s. For each run, the *Pfmdr1* copy number of each sample was measured in triplicate relative to a standard curve using four standards of mixed synthetic gene fragments

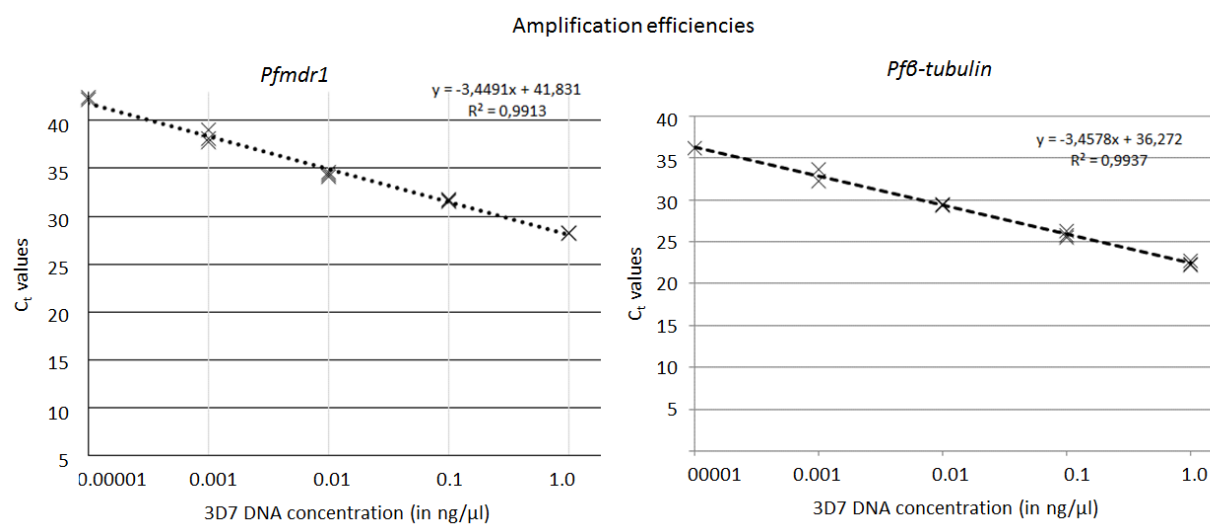
(Eurofins Genomics, Ebersberg, Germany). The lengths of the synthetic fragments for *Pfmdr1* (F3D7_0523000), including gene location are for *PfPM2* (PF3D7_1408000), from position 3981 to 4260 (204 bp): 3981- ctattgtagatattaaagataaagctgacaaaactattattactattgccacagaat**tgcatctataaaacgatcagacaaaattgtggtatttaataaccctgatcgaaatggaacctttgtacagtcacatggaacacacga**tgaattattatcagcacaagatggaatatataaaaaatatgtaaaattagctaaatga-4260 (shown in bold font, the qPCR amplified portion).

The four standards of mixed synthetic gene fragments were: from standard 1 (1:1 molar ratio of *Pfmdr1* and β-tubulin) to standard 4 (4:1 molar ratio of *Pfmdr1* and β-tubulin). The 3D7 Africa line (which has one copy of *Pfmdr1*) and the Dd2line (which has three copies of *Pfmdr1*) were included in each run as controls.

Pfmdr1 copy number was calculated by the $2^{-\Delta C_t}$ method ($\Delta C_t = C_t \text{ Pfmdr1} - C_t \text{ Pf}\beta\text{-tubulin}$ where C_t is the threshold cycle) and deduced from the standard curve. A *Pfmdr1* copy number >1.6 was defined as an amplification of the gene. Amplification efficiencies of the *Pfmdr1* and the *Pf}\beta\text{-tubulin}*



tubulin genes measured by using ten-fold dilutions of 3D7 DNA, were similar (95% and 95%, respectively).



Appendix 3: *PfPM2* mRNA expression profile, with listing of primers, protocols, RT-qPCR efficiencies.

One step reverse transcriptase (RT)-qPCRs were carried out using a CFX96 real-time PCR machine (Bio-Rad) in 25 µl volumes in a 96-well plate containing 2X SuperScript™ III Platinum One step qRT-PCR kit (Life Technologies, Courtaboeuf, France), 0.2 µM concentrations of each forward and reverse primers, 0.1 µM concentrations of specific probes (FAM-BHQ1) and 3 µl of DNase-treated RNA. Listing of primers, protocols and PCR amplification efficiencies are provided in the table below.

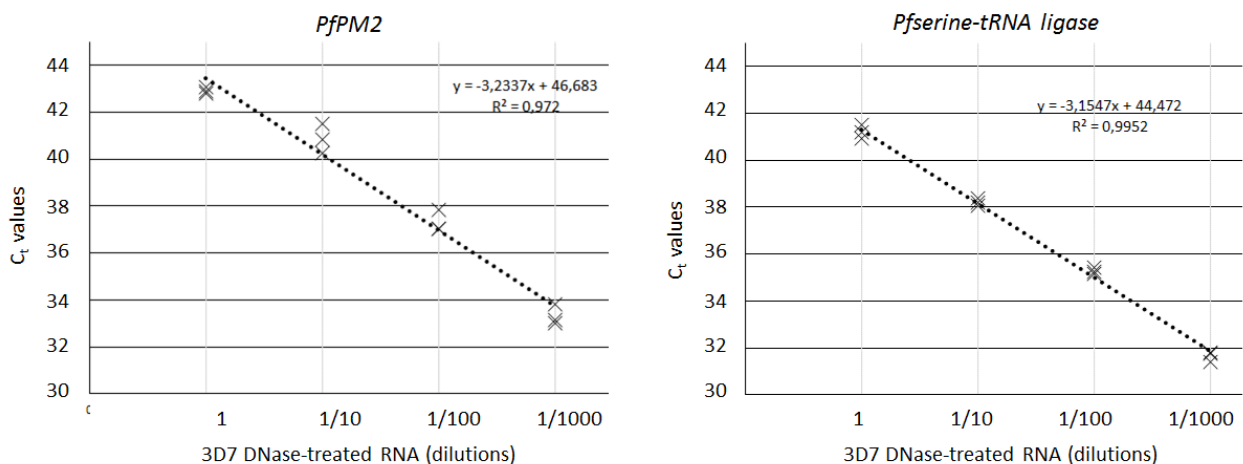
	Primer Sequence	Sequences	T _m (°C)	Product size (bp)
RT-PCR <i>PfPM2</i>	PfPM2_RTPCR_F	5'-GGATTCTGAACCACTTATACTGC-3'	59.1	90
	PfPM2_RTPCR_R	5'-AATTGGATCTACTGAACCTATTGATAA-3'	57.9	
	PfPM2_RTPCR_Probe	5'-FAM-CAACATTTGATGGTATCCTTGGTTTAGGATGGA-BHQ1-3'	71.3	
	Pfserine-tRNA ligase_RTPCR_F	5'-TGGAACAATGGTAGCTGCAC-3'	59.7	92
	Pfserine-tRNA ligase_RTPCR_R	5'-GGCGCAATTTTCAGGAACT-3'	61.5	
	Pfserine-tRNA ligase_RTPCR_Probe	5'-FAM-TGTCTTCTTGAAAATTATCAAAACGGCGAAGG-BHQ1-3'	71.6	

Amplifications were performed under the following conditions: 50°C for 15min, and 95°C for 2min, followed by 35 cycles of 95°C for 15s, 60°C for 30s and a final cycle at 35°C for 30s. Fluorescence data were collected during the 60°C annealing-extension steps.

For each run, *PfPM2* and *Pfserine-tRNA ligase* mRNAs expression were measured in triplicate for each sample. DNase-treated RNA from 3D7 parasites (collected at trophozoite stage, 24h post-invasion) was included in each run as control. *PfPM2* mRNA expression, normalized to *Pfserine-tRNA ligase* mRNA expression, was calculated by the $2^{-\Delta\Delta C_t}$ method, using the following formula: $\Delta\Delta C_t = [(C_t \text{ PfPM2} - C_t \text{ Pfserine-tRNA ligase})_{\text{sample}} - (C_t \text{ PfPM2} - C_t \text{ Pfserine-tRNA ligase})_{3D7}]$, where C_t is the threshold cycle.

RT-qPCR amplification efficiencies of *Pfplasmepsin2* and *Pfserine-tRNA ligase*, measured using ten-fold dilutions of DNase-treated 3D7 RNA, were similar (104% and 107%, respectively).

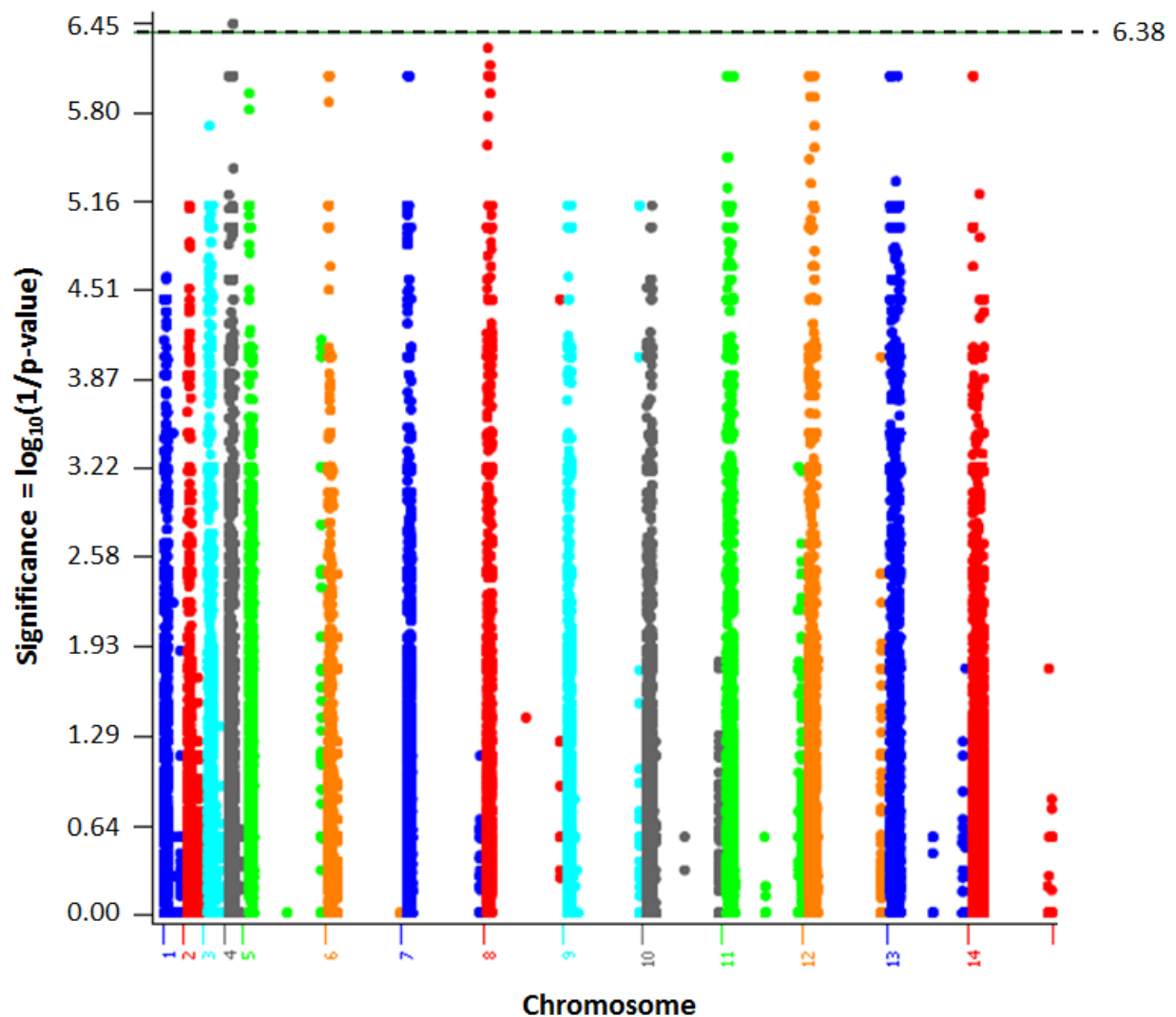
RT-qPCR efficiencies



Appendix 4:

Panel A. Manhattan plot showing the significance of single nucleotide polymorphisms (SNPs) between whole-genome exome sequences of 23 piperazine-resistant and 8 piperazine-sensitive culture-adapted lines phenotyped using in-vitro PSA.

Each dot represents a SNP in a set of 31 culture-adapted parasites, according to chromosome. The x axis represents genomic location, and the y axis represents the \log_{10} transformed Fischer exact test's p-values. After Bonferroni correction at level 5%, only 2 positions in 2 genes, position 896588 of PF3D7_0420000 (zinc finger protein, putative) and position 908385 of PF3D7_0420100 (serine/threonine protein kinase RIO2) achieved genome-wide significance between the resistant and sensitive lines ($p < 3.56 \times 10^{-7}$ for both SNP, Fisher's exact test) [$> 6.38 = \log_{10}(120691/0.05)$] (see Panel B for details).



Panel B. List of the positions with variable proportions of wild type and mutant nucleotides of PF3D7_0420000 (encoding a putative zinc finger protein) and PF3D7_0420100 (encoding the Rio2 serine/threonine protein kinase) sequences of 23 piperaquine-resistant (red background) and 8 piperaquine-sensitive (green background) culture-adapted lines phenotyped using in-vitro PSA (Multalin online software, <http://multalin.toulouse.inra.fr/multalin/>).

SNP (position 896588) in PF3D7_0420000 gene (encoding a putative zinc finger protein) ($p < 3.56 \times 10^{-7}$, Fisher's exact test) and SNP (position 908385) in PF3D7_0420100 gene (encoding the Rio2 serine/threonine protein kinase) ($p < 3.56 \times 10^{-7}$, Fisher's exact test) detected as significant between piperaquine-resistant and piperaquine-sensitive parasite lines are shown in red bold font

Nucleotide code (IUPAC nomenclature): A: Adenine, C: Cytosine, G: Guanine, T: Thymine, R: A or G, Y: C or T, S: G or C, W: A or T, K: G or T, M: A or C, B: C or G or T, D: A or G or T, H: A or C or T, V: A or C or G, N: any base. If an indel is observed (>20% of the genome-wide mean coverage), letters are shown in lowercase.

Analysis were done as follows:

For each position with a coverage >100, the variants with >20% frequency were included in the analysis (A, C, G, T or short INDEL). This results in a total of 30 possible values: 15 IUPAC code with upper case when no INDEL was detected and 15 with lower case when an INDEL was detected. For example, if a given position contains 17% A, 27% C, 32% G, 2% T, and 21% of reads with INDEL, the IUPAC code for this position and for this particular sample is « s ».

In particular, at position 896 588 of the gene PF3D7_0420000 (first table, line highlighted in red font), we observed 2 C and 7 Y in the sensitive lines (the 3D7 reference sequence was included because its sequence information is accurate for SNPs) and 3 C and 20 y in the resistant lines. Thus the contingency table and the Fisher exact probability test (3x2) results were:

	C	Y	y
Sensitive	2	7	0
Resistant	3	0	20

We found a significant p value of $3.56 \cdot 10^{-7}$, ($p=0.042$ after Bonferroni correction, for 120691 tests).

At position 908 385 of the gene PF3D7_0421000 (second table, line highlighted in red font), we observed 8 A and 1 m in the sensitive lines and 1 A and 22 a in the resistant lines. Thus the contingency table and the Fisher exact probability test (3x2) results were:

	A	m	a
Sensitive	8	1	0
Resistant	1	0	22

We also found a significant p value of $3.56 \cdot 10^{-7}$ ($p=0.042$ after Bonferroni correction, for 120691 tests).

PF3D7_0420000 (encoding a putative zinc finger protein)

Position on chromosome 4	Parasite line ID																															
	3D7	6273	6337	6267	6403	6349	6237	6410	6369	6395	6341	6280	6246	6293	6391	6272	6218	6302	6229	6443	6430	6365	6429	6394	6219	6408	6224	6431	6320	6261	6411	6427
	in-vitro PSA survival rate (%)																															
	0.1	0.2	0.4	0.5	0.5	0.6	0.8	6.0	6.4	19.2	25.8	28.9	36.9	39.3	39.4	40.0	40.8	42.5	46.6	49.6	51.3	51.8	51.8	56.7	58.6	58.7	61.4	61.5	62.1	70.5	71.6	77.4
	Mean Coverage (x)																															
435	242	200	134	278	231	202	243	132	257	369	200	125	194	243	237	198	139	188	166	210	223	179	273	224	58	263	157	138	294	199	216	
895028	A	A	A	M	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
895040	C	M	M	M	M	M	M	C	M	C	C	M	M	M	M	C	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
895055	G	R	R	R	R	R	R	G	R	G	G	R	R	R	R	G	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
895063	A	W	W	W	W	W	W	A	W	A	A	W	W	W	W	A	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W
895067	A	R	R	R	R	R	R	A	R	A	A	R	R	R	R	R	A	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
895099	A	W	W	W	W	W	W	A	W	A	A	W	W	W	W	W	A	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W
895111	T	W	W	W	W	W	W	T	W	T	T	W	W	W	W	W	T	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W
895131	G	G	G	G	G	G	G	G	G	G	G	A	G	A	A	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
895785	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	W	A	A	A	A	A	A
896254	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
896321	C	A	A	A	A	A	A	C	A	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
896552	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	Y	C	C	C	C	C	C	C	C	C	C	C	Y	Y	C	Y	C
896558	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	Y
896567	T	Y	Y	Y	Y	Y	Y	T	Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
896573	C	Y	Y	Y	Y	Y	Y	C	Y	C	C	C	Y	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
896585	T	T	T	T	T	T	T	T	T	T	T	t	T	t	t	t	T	T	T	t	t	t	t	t	t	t	t	t	t	t	t	t
896586	A	A	A	A	A	A	A	A	A	A	A	a	A	a	a	a	A	A	A	a	a	a	a	a	a	a	a	a	a	a	a	a
896587	T	T	T	T	T	T	T	T	T	T	T	t	T	t	t	t	T	T	T	t	t	t	t	t	t	t	t	t	t	t	t	t
896588	C	Y	Y	Y	Y	Y	Y	C	Y	C	C	y	y	y	y	y	C	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y
896804	C	C	C	C	C	C	C	C	C	C	C	T	T	T	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
897069	A	A	A	A	A	A	A	W	A	A	A	A	A	A	W	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
897087	A	A	A	A	A	W	W	A	A	A	A	W	W	A	W	W	A	A	A	A	A	A	A	A	A	A	W	A	A	A	W	A
897094	A	W	W	W	A	W	A	A	W	W	W	W	W	W	W	W	A	W	A	A	A	W	A	A	A	A	W	W	A	W	A	A
897095	G	G	K	G	G	K	G	K	G	G	K	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
897097	A	W	A	A	A	A	A	A	A	A	A	A	A	W	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
897114	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	K	G	G	G	G	G	
897121	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	K	G	G	G	G	G	G
897165	A	G	G	G	G	G	G	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
897266	C	G	G	G	G	G	G	C	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
897268	A	T	T	T	T	T	T	A	T	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
897275	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
897284	C	G	G	G	G	G	G	C	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
897286	A	T	T	T	T	T	T	A	T	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
897293	T	Y			C	Y		T		T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T

Position on chromosome 4	Parasite line ID																																
	3D7	6273	6337	6267	6403	6349	6237	6410	6369	6395	6341	6280	6246	6293	6391	6272	6218	6302	6229	6443	6430	6365	6429	6394	6219	6408	6224	6431	6320	6261	6411	6427	
	in-vitro PSA survival rate (%)																																
	0.1	0.2	0.4	0.5	0.6	0.8	6.0	6.4	19.2	25.8	28.9	36.9	39.3	39.4	40.0	40.8	42.5	46.6	49.6	51.3	51.8	51.8	56.7	58.6	58.7	61.4	61.5	62.1	70.5	71.6	77.4		
	Mean Coverage (x)																																
435	242	200	134	278	231	202	243	132	257	369	200	125	194	243	237	198	139	188	166	210	223	179	273	224	58	263	157	138	294	199	216		
897296	C	C			T	T		C		C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
897297	C	C						C		C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
897302	C	C						C		C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
897304	A	W						W		W	W	W	W	W	T	W	W	W	W	W	W	W	T	W	W	W	T	W	W	W	W	T	
897305	C	C						C		C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
897306	C	C						C		C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
897314	C	Y	C			T		C		C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
897315	C	Y	C			T		C		C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
897320	C	C	C	C	C	C	C	Y	C	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
897329	T	G	G	G	G	G	G	K	G	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	
897338	C	T	T	T	T	T	T	Y	T	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
897347	C	G	K	G	G	G	G	K	K	K	K	G	K	K	K	G	G	G	G	G	G	G	K	K	G	G	G	K	G	K	G	K	
897358	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	
897360	G	G	G	G	G	G	G	G	G	G	G	G	K	G	G	G	G	G	G	G	G	K	G	G	G	G	G	G	G	G	G	G	
897363	A	A	A	A	A	A	A	A	A	W	W	A	W	A	A	A	A	A	A	A	A	A	W	A	A	A	A	A	A	A	A	A	W
897371	A	A	A	A	A	A	A	A	A	A	A	A	W	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
897378	A	A	A	A	A	A	A	A	A	A	A	A	W	A	A	W	A	A	A	A	A	A	A	A	A	A	W	A	A	A	A	A	A
897390	G	G	G	G	G	G	G	G	G	G	G	G	G	G	K	G	G	G	G	G	K	G	G	G	G	G	G	K	G	G	G	G	
897393	A	A	A	A	A	A	A	A	A	A	W	W	W	W	W	W	W	W	W	W	A	W	W	W	W	W	W	A	W	W	W	A	W
897419	T	C	C	C	C	C	C	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	
897462	C	A	A	A	A	A	A	C	A	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
897468	C							C		C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
898387	G	G	G	G	G	G	G	A	G	A	A	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
898407	A	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
898504	C	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c
898522	C	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
898528	T	Y	T	Y	Y	Y	Y	T	Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	Y	Y	T	T	T
898537	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	Y	Y	C
898538	C	C	C	C	C	C	C	C	C	C	C	C	Y	C	C	Y	C	Y	C	C	C	C	C	C	C	C	C	Y	C	C	Y	Y	C
898540	T	T	T	T	T	T	T	T	T	T	T	Y	Y	Y	Y	Y	Y	T	Y	T	T	T	T	T	T	Y	Y	Y	T	Y	Y	Y	Y
898543	A	W	W	W	W	W	W	A	W	A	A	W	W	W	W	W	A	W	W	W	W	W	W	A	W	W	W	W	W	W	W	W	W
898555	C	c	c	c	c	c	c	C	c	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
898556	C	c	c	c	c	c	c	C	c	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
898557	A	a	a	a	a	a	a	A	a	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
898558	T	T	t	y	t	y	t	T	t	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
898561	A	W	W	W	W	W	W	A	W	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
898976	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
899306	T	T	T	T	T	T	T	T	T	T	T	Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
899339	G	G	G	G	C	G	G	G	C	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
899359	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T

Position on chromosome 4	Parasite line ID																																
	3D7	6273	6337	6267	6403	6349	6237	6410	6369	6395	6341	6280	6246	6293	6391	6272	6218	6302	6229	6443	6430	6365	6429	6394	6219	6408	6224	6431	6320	6261	6411	6427	
	in-vitro PSA survival rate (%)																																
	0.1	0.2	0.4	0.5	0.5	0.6	0.8	6.0	6.4	19.2	25.8	28.9	36.9	39.3	39.4	40.0	40.8	42.5	46.6	49.6	51.3	51.8	51.8	56.7	58.6	58.7	61.4	61.5	62.1	70.5	71.6	77.4	
	Mean Coverage (x)																																
435	242	200	134	278	231	202	243	132	257	369	200	125	194	243	237	198	139	188	166	210	223	179	273	224	58	263	157	138	294	199	216		
899379	A	T	T	T	T	T	T	A	T	A	A	T	T	T	T	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	
899385	A	a	a	a	A	a	a	A	A	A	A	a	a	a	a	A	a	a	a	a	A	A	a	a	a	a	a	a	a	a	a	a	
899386	T	Y	Y	Y		Y	Y	T	C	T	T	Y	Y	Y	Y	T	Y	Y	Y	Y	C	C	Y	Y	T	Y	T	Y	T	Y	Y	Y	
899387	T	W	W	W		A	W	T	T	T	T	A	A	W	A	A	T	W	A	W	W	W	T	W	A	A	W	A	W	W	W	W	
899388	T	M	M	M		C	M	T		T	T	C	C	M	C	C	T	H	C	M	M	M	M	C	C	M	C	C	Y	M	M	M	
899389	T	Y	Y	Y		T	Y	T		T	T	T	T	T	Y	T	T	Y	T	Y	Y	C	Y	Y	T	T	Y	T	T	T	Y	Y	
899390	G	T	Y	Y		Y	C	G		G	G	C	Y	C	C	C	G	K	C	T	C	T	T	Y	Y	C	Y	C	C	S	Y	Y	
899391	A	A	A	A		A	A	A		A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
899392	T	T	T	T		T	T	T		T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	
899393	C							C		C	C						C	C													C		
899394	A							A		A	A						A	A													A		
899399	G				C			G	C	G	G						G	G													G		
899400	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		A			A	A	A	A	A	A	
899401	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	W	T	T	T	T	T		T	T	T	T	T	T	T	
899402	C	G	G	G	C	G	G	C	C	C	C	G	G	G	G	C	G	G	S	G	G	G	G	G		G	G	G	S	G	G	G	
899403	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		A	A	A	A	A	A	A
899404	C	T	T	T	T	T	T	Y	T	Y	Y	T	T	T	T	Y	T	T	T	T	T	T	T	T		T	T	T	Y	T	T	T	
899405	T	A	A	A	A	A	A	T	A	T	T	A	A	A	A	A	T	A	A	A	A	A	A	A	A	A	A	A	A	W	A	A	
899406	A	T	T	T	T	T	T	A	T	A	A	T	T	T	T	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	
899408	G	C	C	C	C	C	C	G	C	G	G	C	C	C	C	C	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
899411	C	A	A	A	A	A	A	C	A	C	C	A	A	A	A	A	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
899915	T	T	T	T	T	T	T	T	T	T	T	T	Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	
899970	T	T	T	T	T	T	T	T	T	T	T	T	Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	
900145	A	T	T	T	A	T	T	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
900189	G	G	G	G	T	G	G	T	T	T	T	G	G	G	G	G	T	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	
900191	C	C	C	C	T	C	C	T	T	T	T	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
900206	T	T	T	T	C	T	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	
900208	A	A	A	A	T	A	A	A	T	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
900210	G	G	G	G	A	G	G	A	A	A	A	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	
900213	A	A	A	A	T	A	A	T	T	T	T	A	A	A	A	A	T	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
900214	T	T	T	T	A	T	T	A	A	A	A	T	T	T	T	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	
900218	G	G	G	G	T	G	G	T	T	T	T	G	G	G	G	T	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	
900225	C	C	C	C	C	C	C	T	C	T	T	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
900229	A	A	A	A	T	A	A	W	T	W	W	A	A	A	A	A	W	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
900231	G	G	G	G	A	G	G	R	A	R	R	G	G	G	G	R	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	
900242	C	C	C	C	C	C	C	S	C	S	S	C	C	C	C	C	S	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
900246	A	A	A	A	C	A	A	C	C	C	C	A	A	A	A	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
900252	G	G	G	G	A	G	G	A	A	A	A	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	
900254	C	C	C	C	T	C	C	T	T	T	T	C	Y	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
900255	A	A	A	A	T	A	A	T	T	T	T	A	A	A	A	A	T	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	

Position on chromosome 4	Parasite line ID																															
	3D7	6273	6337	6267	6403	6349	6237	6410	6369	6395	6341	6280	6246	6293	6391	6272	6218	6302	6229	6443	6430	6365	6429	6394	6219	6408	6224	6431	6320	6261	6411	6427
	in-vitro PSA survival rate (%)																															
	0.1	0.2	0.4	0.5	0.5	0.6	0.8	6.0	6.4	19.2	25.8	28.9	36.9	39.3	39.4	40.0	40.8	42.5	46.6	49.6	51.3	51.8	51.8	56.7	58.6	58.7	61.4	61.5	62.1	70.5	71.6	77.4
	Mean Coverage (x)																															
435	242	200	134	278	231	202	243	132	257	369	200	125	194	243	237	198	139	188	166	210	223	179	273	224	58	263	157	138	294	199	216	
900256	T	T	T	T	A	T	T	A	A	A	A	T	T	T	T	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
900260	G	G	G	G	T	G	G	T	T	T	T	G	G	G	G	T	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
900267	A	A	A	A	A	A	A	A	A	A	A	A	M	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
900273	G	G	G	G	K	G	G	T	K	T	T	G	G	G	G	T	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
900275	C	C	C	C	T	C	C	T	T	T	T	C	Y	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
900281	G	G	G	G	G	G	G	G	G	G	G	G	K	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
900284	C	C	C	C	S	C	C	C	S	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
900290	T	T	T	T	Y	T	T	T	Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
900292	A	A	A	A	W	A	A	A	W	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
900298	A	A	A	A	A	A	A	A	A	A	A	A	W	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
900309	T	T	T	T	M	T	T	T	M	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
900313	T	T	T	T	W	T	T	T	W	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
900315	A	A	A	A	W	A	A	A	W	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
900318	A	A	A	A	A	A	A	A	A	A	A	W	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
900323	G	G	G	K	g	K	G	G	g	G	G	K	G	K	G	G	G	G	G	G	G	G	G	K	G	G	G	G	G	G	K	G
900324	T	T	T	T	t	T	T	T	t	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
900325	T	T	T	T	t	T	T	T	t	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
900326	C	C	C	C	c	C	C	C	c	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
900327	G	G	G	G	g	G	G	G	g	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
900328	T	T	T	T	t	T	T	T	t	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
900329	C	C	C	C	c	C	C	C	c	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
900330	C	C	C	C	c	C	C	C	c	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
900331	A	A	A	A	a	A	A	A	a	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
900332	c	C	C	C	c	C	C	C	c	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
900333	T	T	T	T	a	T	T	A	a	A	A	T	T	T	T	T	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
900334	T	T	T	T	t	T	T	T	t	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
900335	T	T	T	T	t	T	T	T	t	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
900336	T	T	T	T	w	T	T	W	w	W	W	T	T	T	T	T	W	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
900337	T	T	T	T	t	T	T	T	t	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
900338	T	T	T	T	t	T	T	T	t	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
900339	T	T	T	T	w	T	T	W	w	W	W	T	T	T	T	T	W	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
900340	T	T	T	T	w	T	T	W	w	W	W	T	T	T	T	T	W	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
900341	T	T	T	T	t	T	T	T	t	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
900342	T	T	T	T	t	T	T	T	t	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
900343	T	T	T	T	t	T	T	T	t	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
900344	T	T	T	T	k	T	T	K	k	K	K	T	T	T	T	T	K	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
900348	G	G	G	K	G	G	G	G	G	G	G	G	G	G	G	G	G	K	G	G	G	G	G	G	G	G	G	G	G	G	G	G
900353	C	T	T	T	C	T	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
900355	T	W	W	W	T	W	W	W	T	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W
900357	A	W	A	A	A	A	A	A	A	A	A	A	W	A	W	A	A	A	A	A	A	W	W	A	A	A	W	W	A	W	A	W

Position on chromosome 4	Parasite line ID																																			
	3D7	6273	6337	6267	6403	6349	6237	6410	6369	6395	6341	6280	6246	6293	6391	6272	6218	6302	6229	6443	6430	6365	6429	6394	6219	6408	6224	6431	6320	6261	6411	6427				
	in-vitro PSA survival rate (%)																																			
	0.1	0.2	0.4	0.5	0.5	0.6	0.8	6.0	6.4	19.2	25.8	28.9	36.9	39.3	39.4	40.0	40.8	42.5	46.6	49.6	51.3	51.8	51.8	56.7	58.6	58.7	61.4	61.5	62.1	70.5	71.6	77.4				
	Mean Coverage (x)																																			
	435	242	200	134	278	231	202	243	132	257	369	200	125	194	243	237	198	139	188	166	210	223	179	273	224	58	263	157	138	294	199	216				
900361	A	W	W	W	A	W	W	A	A	A	A	A	W	A	A	W	A	A	W	A	A	W	W	A	A	A	A	W	W	W	W	W				
900363	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	W	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T				
900372	C	T	T	T	C	T	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T				
900381	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	W	A	A	A	A				
900386	G	G	G	K	G	G	K	G	G	G	G	G	G	G	G	G	K	K	G	K	G	G	G	K	G	G	K	K	K	K	K	G				
900389	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	Y	C			
900390	G	G	G	G	G	G	K	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G			
900395	C	c	c	c	C	c	c	C	C	C	C	c	c	c	c	c	C	c	C	c	c	c	c	c	c	c	c	c	c	c	c	c	c			
900396	A	t	t	t	A	t	t	A	A	A	A	t	T	T	t	t	A	T	T	t	t	t	t	t	T	t	t	t	t	t	t	t	t			
900399	A	T	T	T	A	T	T	A	A	A	A	T	T	T	T	T	A	W	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T			
900403	A	T	T	T	A	T	T	A	A	A	A	T	T	T	T	T	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T			
900457	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	K	G	G	G	G	G	G			
900665	A	A	A	A	G	A	A	A	G	A	A	G	R	G	G	G	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G			
900780	G	G	G	G	G	G	G	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A			
900957	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C			
900988	A	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C			
901096	G	G	G	G	G	G	G	A	G	A	A	G	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G			
901097	G	G	G	G	G	G	G	C	G	C	C	G	G	G	G	G	C	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G			
901197	A	A	A	A	A	A	A	G	A	G	G	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A			
901321	A	A	A	A	a	A	A	A	a	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A			
901322	T	T	T	T	t	T	T	T	t	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T			
901323	A	A	A	A	a	A	A	A	a	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A			
901324	T	T	T	T	t	T	T	T	t	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T			
901325	T	T	T	T	t	T	T	T	t	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T			
901326	A	A	A	A	a	A	A	A	a	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A			
901327	T	T	T	T	t	T	T	T	t	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T			
901328	T	T	T	T	t	T	T	T	t	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T			
901329	A	A	A	A	t	A	A	A	t	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A			
901330	T	T	T	T	a	T	T	T	a	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T			
901345	T	T	T	T	T	T	T	t	T	t	t	T	T	T	T	T	t	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T			
901408	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T			
901416	C	A	A	A	C	A	A	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C			
901429	T	T	T	T	C	T	T	T	C	T	T	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C			
901771	T	T	T	T	T	T	T	C	T	C	C	T	T	T	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T			
901913	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A			
902548	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T			
902591	G	C	C	C	C	C	C	G	C	G	G	C	C	C	C	C	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C			
902609	C	c	c	c	c	c	c	C	c	c	C	c	c	c	c	y	c	C	y	c	y	c	c	C	c	Y	c	c	c	c	c	c	y			
902610	A	a	a	a	a	a	a	A	a	a	A	a	a	a	a	a	A	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a			
902611	T	t	t	t	t	t	t	T	t	t	T	t	t	t	t	t	T	t	t	t	t	t	t	t	t	T	t	t	t	t	t	t	t			

Position on chromosome 4	Parasite line ID																																	
	3D7	6273	6337	6267	6403	6349	6237	6410	6369	6395	6341	6280	6246	6293	6391	6272	6218	6302	6229	6443	6430	6365	6429	6394	6219	6408	6224	6431	6320	6261	6411	6427		
	in-vitro PSA survival rate (%)																																	
	0.1	0.2	0.4	0.5	0.5	0.6	0.8	6.0	6.4	19.2	25.8	28.9	36.9	39.3	39.4	40.0	40.8	42.5	46.6	49.6	51.3	51.8	51.8	56.7	58.6	58.7	61.4	61.5	62.1	70.5	71.6	77.4		
	Mean Coverage (x)																																	
435	242	200	134	278	231	202	243	132	257	369	200	125	194	243	237	198	139	188	166	210	223	179	273	224	58	263	157	138	294	199	216			
902612	T	t	t	t	t	t	t	T	t	t	T	t	t	t	y	y	T	y	t	y	y	t	y	T	t	Y	t	t	t	t	t	y		
902613	A	A	A	A	A	A	A	A	A	a	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		
902615	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	K	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T		
902624	T	Y	Y	Y	Y	Y	Y	T	T	T	T	Y	Y	Y	T	Y	T	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y		
902627	T	Y	Y	Y	Y	Y	Y	T	Y	T	T	Y	Y	Y	Y	Y	T	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y		
902633	T	T	T	T	T	T	T	Y	T	Y	Y	T	T	T	T	Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T		
902642	C	C	C	C	C	C	C	T	C	Y	T	C	C	Y	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C		
902663	C	T	T	T	T	T	T	T	T	Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T		
903129	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	W	A	A	A	A	A	A	A		
903141	A	A	A	A	W	A	A	A	W	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		
903155	A	A	A	A	W	A	A	A	A	A	A	A	A	A	A	A	A	A	W	A	A	A	A	W	A	A	A	A	A	A	A	A		
903161	C	C	C	C	C	C	C	Y	C	C	Y	C	Y	C	C	C	C	Y	Y	Y	C	Y	C	C	Y	C	Y	C	Y	C	Y			
903176	G	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	G	G	K	K	K	K	K	K	K	K	K	K	K	K		
903181	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	W	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		
903182	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	W	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		
903183	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	W	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		
903186	G	G	K	G	G	G	G	G	G	G	G	G	K	G	G	G	K	G	G	G	G	G	G	G	K	G	G	G	G	G	G	G		
903187	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	K	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G		
903188	A	A	W	A	A	A	A	A	A	A	A	A	A	A	A	A	W	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		
903207	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		
903414	A	A	A	A	A	A	A	A	A	A	A	G	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G		
903891	G	G	G	G	G	G	G	T	G	T	T	G	G	G	G	T	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G		
903950	A	A	A	A	A	A	A	C	A	C	C	A	A	A	A	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		
904088	G	G	G	G	G	G	G	G	G	G	G	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		
904159	C	C	C	C	C	C	C	T	C	T	T	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C		
904629	G	G	G	G	G	K	G	G	G	K	G	G	K	G	G	G	G	G	G	G	G	G	G	G	G	G	G	K	G	G	G	G		

PF3D7_0420100 (encoding the Rio2 serine/threonine protein kinase)

Position on chromosome 4	Parasite line ID																																
	3D7	6273	6337	6267	6403	6349	6237	6410	6369	6395	6341	6280	6246	6293	6391	6272	6218	6302	6229	6443	6430	6365	6429	6394	6219	6408	6224	6431	6320	6261	6411	6427	
	in-vitro PSA survival rate (%)																																
	0.1	0.2	0.4	0.5	0.5	0.6	0.8	6.0	6.4	19.2	25.8	28.9	36.9	39.3	39.4	40.0	40.8	42.5	46.6	49.6	51.3	51.8	51.8	56.7	58.6	58.7	61.4	61.5	62.1	70.5	71.6	77.4	
	Mean Coverage (x)																																
	272	232	214	132	325	242	203	198	131	203	237	121	143	132	179	169	127	88	125	117	109	167	146	152	150	31	209	109	67	250	104	172	
907571	T	C	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
908347	T	A	A	A	A	A	A	A	A	A	A	T	A	T	T	T	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	
908382	G	G	G	G	G	G	G	R	G	R	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	
908385	A	A	A	A	A	A	A	m	A	a	a	a	A	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	
908388	A	A	A	A	A	A	A	A	A	A	A	R	A	A	A	A	A	R	A	A	R	A	A	R	A	R	A	R	R	A	R	A	
908391	C	C	C	C	C	C	C	C	C	C	C	M	C	C	C	C	M	M	C	C	M	C	C	M	C	M	C	M	C	M	C	M	
908394	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	R	A	R	A	A	R	A	R	A	
908458	G	g	g	g	g	g	g	G	g	G	G	G	g	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	
908459	T	t	t	t	t	t	t	T	t	T	T	T	t	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	
908460	G	a	a	a	a	a	a	G	a	G	G	g	a	g	g	g	G	g	g	g	g	g	g	g	g	g	g	g	g	g	g	g	g
908461	A	a	a	a	a	a	a	A	a	A	A	a	a	a	a	a	A	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a
908462	T	T	T	T	T	T	T	T	T	T	T	t	T	t	t	t	T	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t
908463	A	A	A	A	A	A	A	A	A	A	A	a	A	a	a	a	A	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a
908464	A	A	A	A	A	A	A	A	A	A	A	a	A	a	a	a	A	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a
908465	T	T	T	T	T	T	T	T	T	T	T	t	T	t	t	t	T	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t
908466	A	A	A	A	A	A	A	A	A	A	A	a	A	a	a	a	A	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a
908467	A	A	A	A	A	A	A	A	A	A	A	a	A	a	a	a	A	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a
908468	T	T	T	T	T	T	T	T	T	T	T	t	T	t	t	t	T	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t
908469	A	A	A	A	A	A	A	A	A	A	A	a	A	a	a	a	A	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a
908470	A	A	A	A	A	A	A	A	A	A	A	a	A	a	a		A	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a
908471	T	T	T	T	T	T	T	T	T	T	T	t	T	t	t	t	T	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t	t
908472	A	A	A	A	A	A	A	A	A	A	A	a	A	a	a	a	A	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a
908685	A	G	G	G	G	G	G	A	G	A	A	G	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
908711	T	T	T	T	W	T	T	W	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
908713	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	R	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
908725	G	G	G	G	G	G	G	G	G	G	G	G	R	G	G	G	G	R	G	G	G	G	G	G	G	G	G	G	G	G	G	G	R
908726	T	T	T	T	T	T	T	W	T	W	W	T	W	T	W	T	T	T	T	T	W	T	W	T	T	T	W	W	T	W	T	W	W
908729	G	G	G	G	G	G	G	G	G	G	G	G	R	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	R
908730	G	G	G	G	G	G	G	R	G	R	G	G	R	G	G	G	G	G	R	G	G	G	R	G	G	G	G	G	G	G	G	G	G
908736	T	T	T	T	T	T	T	W	T	W	W	T	W	W	W	W	T	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W
908738	T	T	T	T	W	T	T	W	W	W	W	T	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	T	W	W	W	W
908742	G	G	R	R	R	G	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
908747	T	T	T	W	W	T	T	T	W	W	W	W	W	W	W	W	T	T	T	W	T	T	W	W	W	W	W	W	T	W	W	W	W
908748	T	T	T	W	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	W	W	T	T	T	T	T	T	T	T	T

Position on chromosome 4	Parasite line ID																																
	3D7	6273	6337	6267	6403	6349	6237	6410	6369	6395	6341	6280	6246	6293	6391	6272	6218	6302	6229	6443	6430	6365	6429	6394	6219	6408	6224	6431	6320	6261	6411	6427	
	in-vitro PSA survival rate (%)																																
	0.1	0.2	0.4	0.5	0.5	0.6	0.8	6.0	6.4	19.2	25.8	28.9	36.9	39.3	39.4	40.0	40.8	42.5	46.6	49.6	51.3	51.8	51.8	56.7	58.6	58.7	61.4	61.5	62.1	70.5	71.6	77.7	
	Mean Coverage (x)																																
	272	232	214	132	325	242	203	198	131	203	237	121	143	132	179	169	127	88	125	117	109	167	146	152	150	31	209	109	67	250	104	172	
908749	T	T	T	W	T	T	W	W	W	W	W	T	W	W	W	W	W	W	W	T	W	W	W	T	W	W	W	T	W	W	T	W	
908755	C	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	
908762	T	W	W	W	W	W	W	W	W	A	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	
908772	G	r	R	R	R	R	G	R	G	R	R	R	G	R	R	R	G	R	G	R	G	R	R	G	R	R	R	R	R	R	R	G	R
908773	T	t	T	T	W	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	W	T	T	T	T	T	T	T	T	T	T	T
908774	T	t	T	T	W	T	T	T	T	T	T	T	T	T	T	T	T	W	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
908776	T	T	T	W	W	T	T	T	T	T	T	T	T	T	W	T	T	T	T	T	T	T	T	T	T	W	T	T	T	T	T	T	T
908778	G	G	R	G	R	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
908780	G	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
908785	T	W	W	W	W	T	T	W	T	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W
908786	A	W	W	W	W	W	W	A	W	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
908792	T	T	W	T	W	W	W	W	T	T	T	W	T	T	T	W	T	T	T	T	T	T	T	T	T	W	T	T	T	W	T	T	T
908793	T	W	W	T	W	T	T	T	W	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
908794	A	A	A	M	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
908796	T	H	Y	H	H	M	y	T	Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
908797	T	W	W	T	W	Y	W	T	W	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
908799	T	W	W	W	W	T	W	W	W	W	W	W	W	T	W	W	W	T	W	T	W	W	W	W	W	W	W	W	W	W	W	T	W

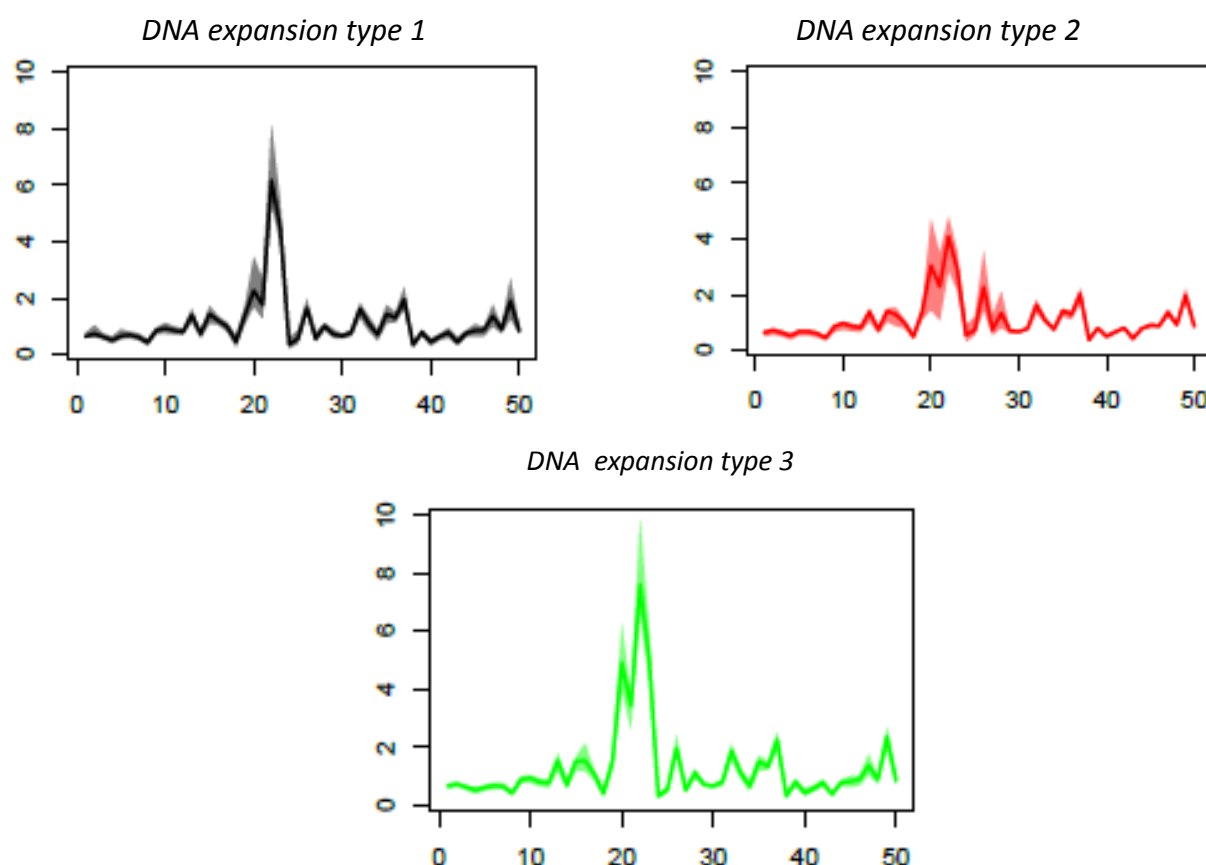
Appendix 5: Profiles of DNA expansion in the region of genes encoding proteins involved in hemoglobin-degrading activities positively associated with in-vitro piperaquine resistance and methodology developed to confirm the four DNA expansion profiles.

We performed an unsupervised classification of the amplification profiles (50 genes flanking PfPM2-3) using a Gaussian mixture. Each class has its own prior proportion, a class-specific mean for each gene, and a global residual variance, which is shared across genes and classes (this unrealistic homoscedastic assumption was necessary to allow a good fitting of the mixture). Fitting was performed using the classical Expectation-Maximization algorithm (Dempster *et al.*, Journal of the Royal Statistical Society. Series B (Methodological), Vol. 39, No. 1. (1977), pp. 1-38) and in order to avoid sub-optimal solution we replicated the algorithm 50 times only retaining the best fitting. The mixture model was fitted for k=2, 3, 4, 5, 6 classes. Model selection was performed using the BIC (Bayesian Information Criterion). The selected model has k=3 classes, as shown below:

Model selection

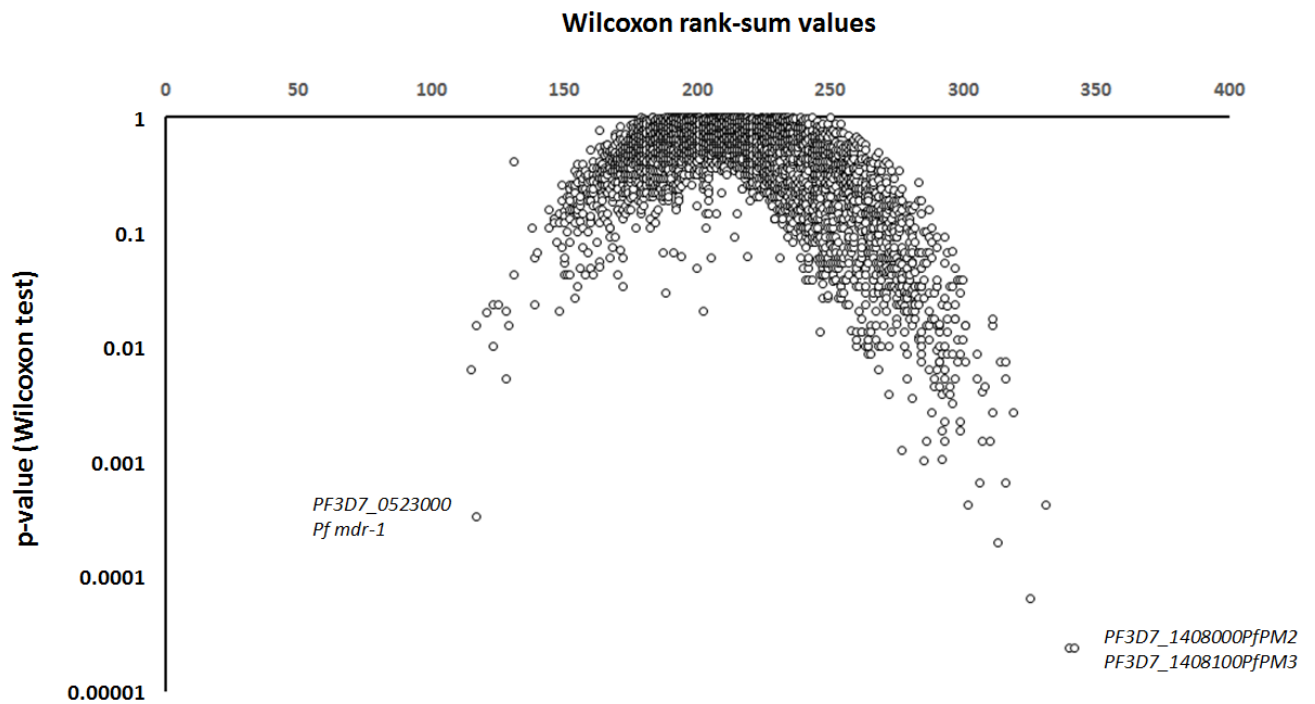
	k=2	k=3	k=4	k=5	k=6
loglik	-250.102	-41.52158	132.9167	307.3057	430.4599
dim	103.000	157.0000	213.000	271.0000	331.0000
bic	1221.520	1182.524444	1225.8195	1283.2192	1457.0948

Specific signals of the three types of DNA expansion (types by parasite lines are given in Table 2)
(the x-axis represents the gene position, and the y-axis represents the amplification signal).



Appendix 6: Distribution of the Wilcoxon Rank-Sum test p-values ranking the significance of CNVs of the 4,616 genes screened between whole-genome exome sequences of 23 piperazine-resistant and 8 piperazine-sensitive culture-adapted lines phenotyped using in-vitro PSA.

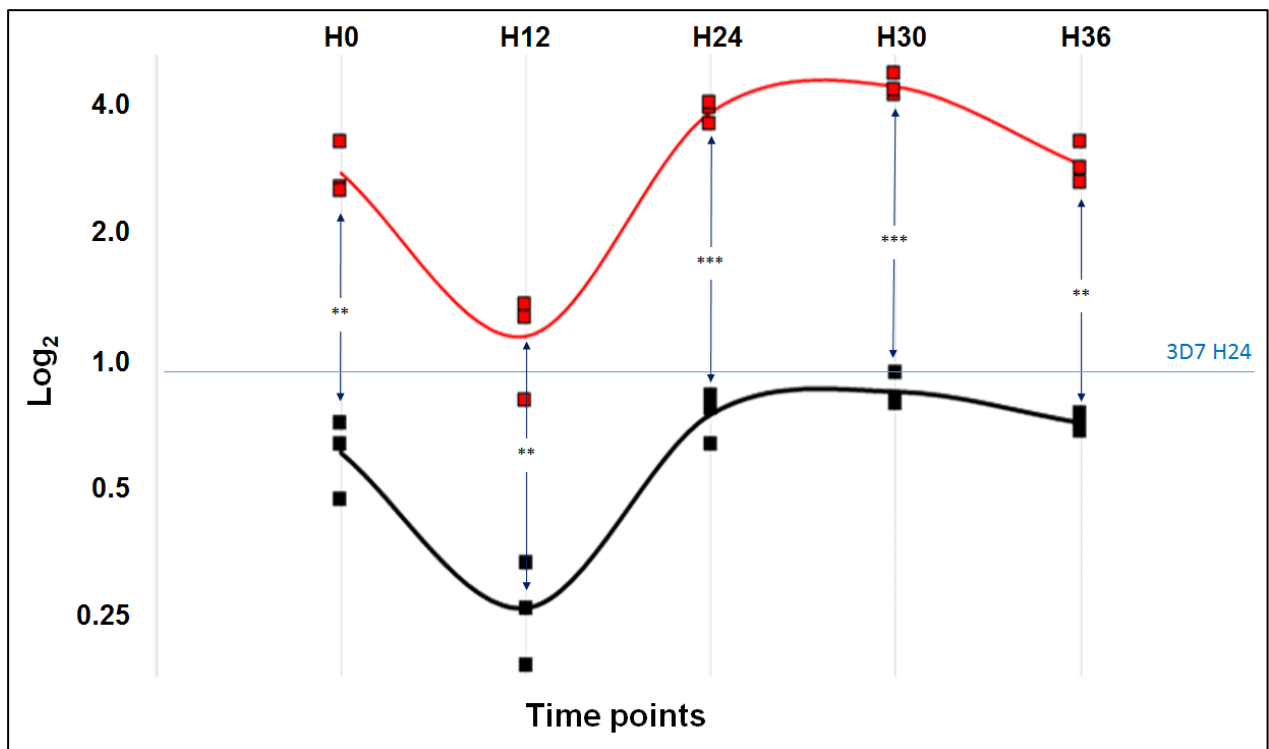
Each dot represents a CNV. The x-axis represents the Wilcoxon rank-sum values, and the y-axis represents the p-values (Wilcoxon test). PF3D7_1408000 (*PfPM2*) and PF3D7_1408100 (*PfPM3*) ranked in the two first positions ($p=2.43 \times 10^{-5}$). PF3D7_0523000 (*Pfmdr1*) was classified at the 4676th/4678 position ($p=0.015$).



Appendix 7:

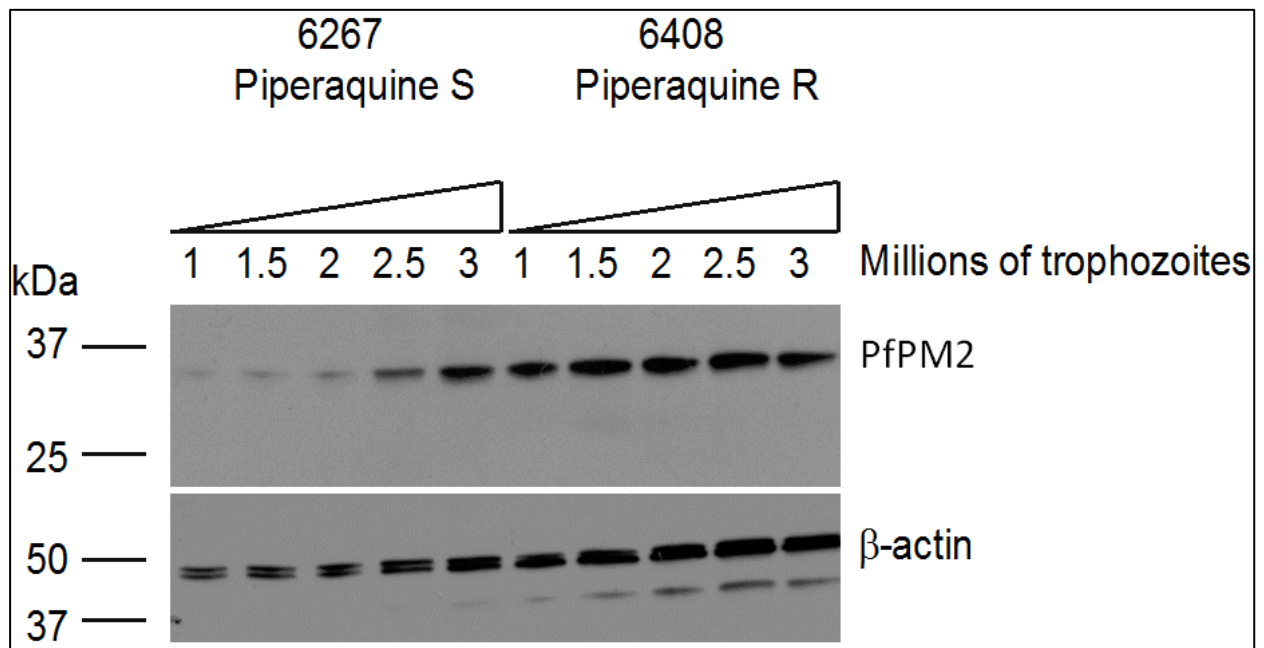
Panel A. *PfPM2* mRNA transcript levels, relative to *Pfserine-tRNA ligase* mRNA in early ring stages (H0: 0-3h post invasion), late ring stages (H12: 12-15h post invasion), early trophozoite stages (H24: 24-27h post invasion) and mature trophozoites (H36: 36-39h post invasion) of the in-vitro culture-adapted piperazine-resistant ID_6320 line (PSA survival rate=62.1%, 2 copies *PfPM2*, C580Y K13 allele, solid red line) and piperazine-sensitive ID_6267 line (PSA survival rate=0.5%, single copy *PfPM2*, C580Y K13 allele, solid black line) (see table 2 for details).

The x axis represents different time points post invasion, and the y axis represents \log_2 transformed $2^{-\Delta\Delta C_t}$ values (see appendix 4 for details). The horizontal solid blue line corresponds to the *PfPM2* mRNA level of 3D7 trophozoites (24h post invasion), used as a control. Black arrows indicate significant differences in *PfPM2* mRNA level (* $p<0.05$ -0.011, ** $p<0.01$ -0.0011 and *** $p<0.001$) between the 2 strains at each time point.



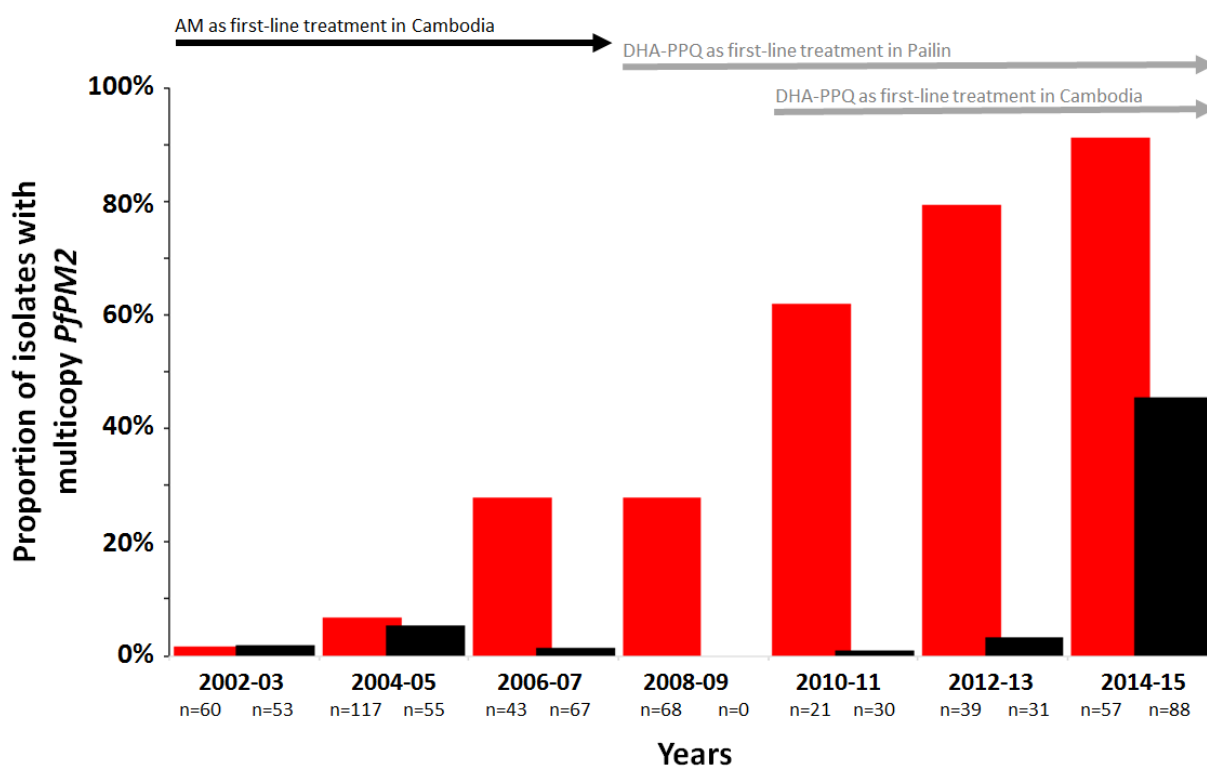
Panel B. PfPM2 expression in *Plasmodium falciparum* synchronized trophozoite-stage of piperazine-resistant (6408, PSA survival rate=58.7%, multicopy *PfPM2*) and piperazine-sensitive parasite lines (6267, PSA survival rate=0.5%, single copy *PfPM2*) detected by western immunoblot.

The piperazine-resistant parasite line 6408 has higher PfPM2 protein levels than the sensitive line 6267. Synchronized trophozoite-stage cultures (24-30 hours post invasion) were probed with anti-PfPM2 (gift of Daniel Goldberg) and anti-beta-actin (NovusBio) antibodies. The resistant line has approximately twice as much PfPM2 as the sensitive line.

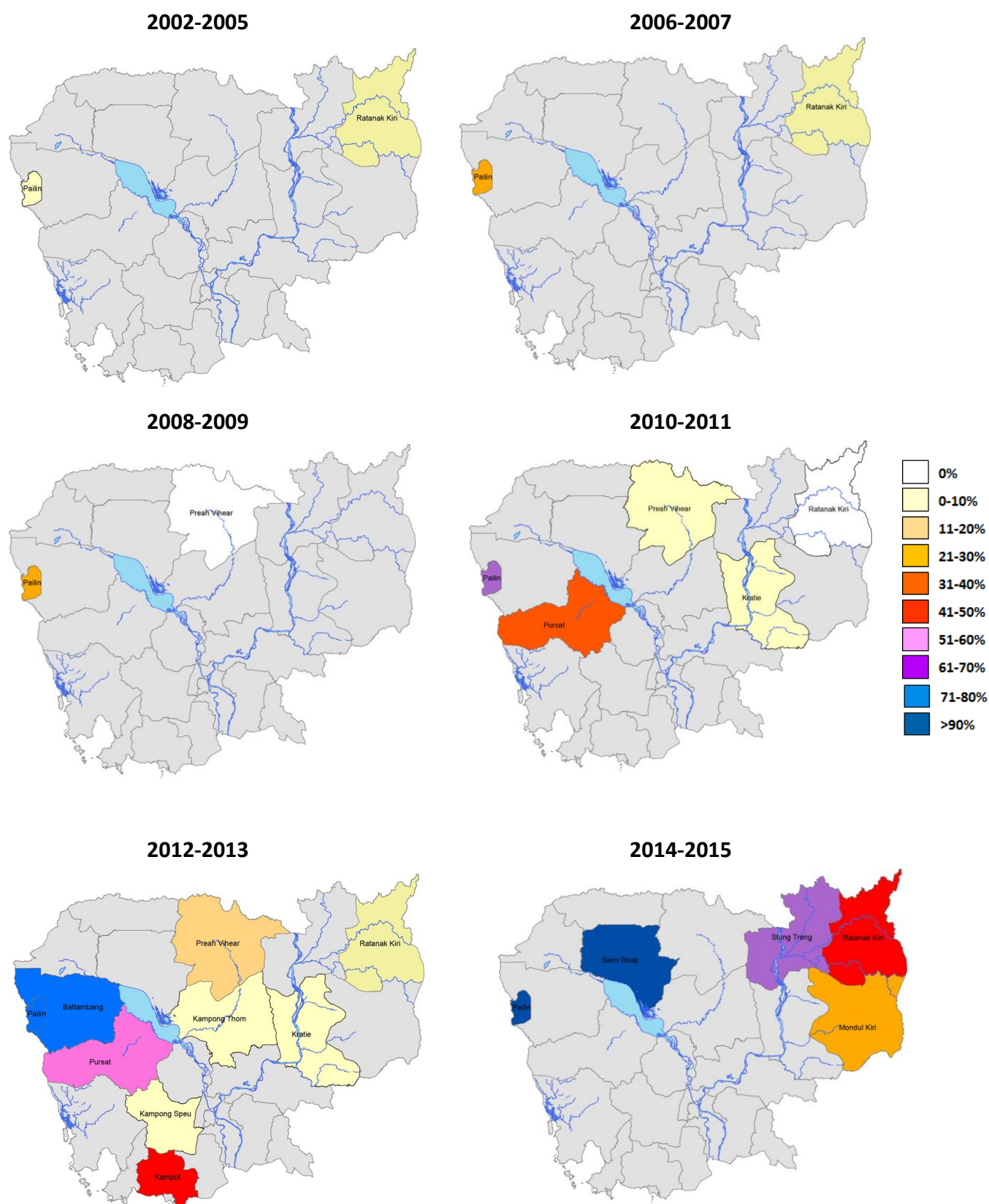


Appendix 8: Spatio-temporal increase in frequency of parasites with multicopy *PfPM2* in Cambodia from 2002 to 2015.

Panel A. Increase in frequency of parasites with multicopy *PfPM2* in Pailin (red histogram) and in Rattanakiri (black histogram) between 2002 and 2015. On the x-axis, sample sizes are given by site and year.



Panel B. Dynamic of the spread of multicopy *PfPM2* parasites in 10 Cambodian provinces from 2002 to 2015.



Legend: Color-codes are the proportion of parasites with multicopy *PfPM2*

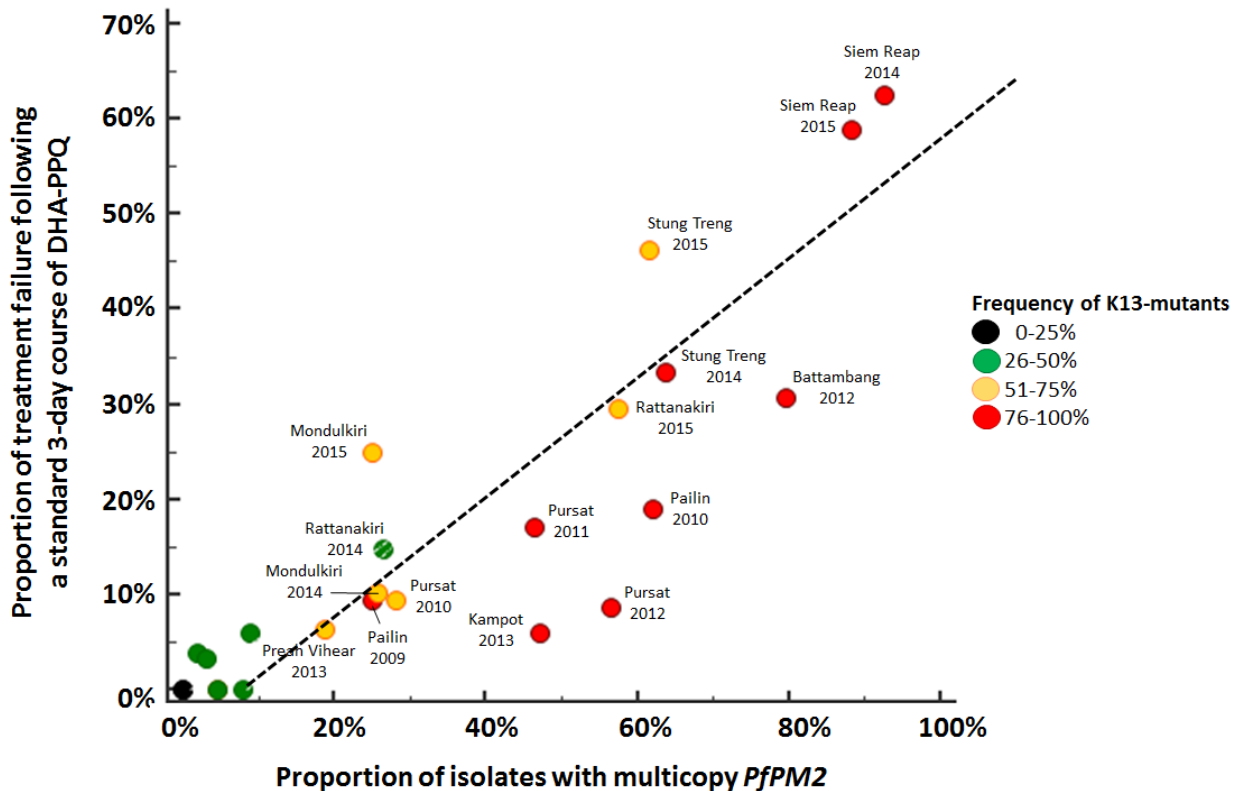
Period (year)	Province	No isolates studied*	% isolates with multicopy <i>PfPM2</i>
2002-2003	Pailin	60	1.7
	Rattanakiri	53	1.9
2004-2005	Pailin	117	6.8
	Rattanakiri	55	5.4
2006-2007	Pailin	43	27.9
	Rattanakiri	67	1.5
2008-2009	Pailin	68	27.9
	Preah Vihear	30	0,0
2010-2011	Pailin	21	61.9
	Rattanakiri	30	0,0
	Pursat	73	38,4
	Kratié	51	2,0
	Preah Vihear	34	8,8
2012-2013	Pailin	39	79.5
	Rattanakiri	31	3.22
	Pursat	23	56.5
	Battambang	39	79.5
	Kampong Speu	22	4.5
	Kampot	17	47.1
	Kratié	22	4.5
	Preah Vihear	16	18.7
2014-2015	Pailin	57	91.2
	Rattanakiri	88	45.5
	Mondulkiri	55	25.4
	Siemreap	57	91.2
	Stungtreng	46	63.0

* Details regarding the samples collected from 2002 to 2008 are given in Arie *et al.*, (Nature, 505, 50-55, 2014, A molecular marker of artemisinin-resistant *Plasmodium falciparum* malaria. (reference 1 of the main text).

Appendix 9:

Panel A. Correlation between the proportion of parasites with multicopy *PfPM2* and DHA-PPQ treatment failure rates recorded in 12 sites across Cambodia from 2009 to 2015.

Results from each clinical study (site and year) are represented by a colored dot. The position of the dot corresponds to the proportion of parasites with multicopy *PfPM2* (x axis) and the DHA-PPQ treatment failure rate (y axis). The color code refers to the proportion of K13 mutant parasites in each site by year.



Panel B. Cox regression model: association of *PfPM2* copy number, sampling sites and treatment response.

Variables	coef	exp(coef)	se(coef)	z	p value
Kampong Speu	$-1.57 \cdot 10^1$	$1.49 \cdot 10^{-7}$	$2.95 \cdot 10^3$	-0.01	0.9957
Kampong Thom	$-1.59 \cdot 10^1$	$1.21 \cdot 10^{-7}$	$2.08 \cdot 10^3$	-0.01	0.9939
Kampot	-1.32	$2.68 \cdot 10^{-1}$	1.04	-1.27	0.2059
Kratie	$1.30 \cdot 10^{-2}$	1.01	$7.94 \cdot 10^{-1}$	0.02	0.9870
Mondulhiri	$3.37 \cdot 10^{-1}$	1.40	$4.58 \cdot 10^{-1}$	0.73	0.4626
Pailin	$-8.79 \cdot 10^{-2}$	$9.16 \cdot 10^{-1}$	$4.76 \cdot 10^{-1}$	-0.18	0.8536
Preah Vihear	$-2.54 \cdot 10^{-1}$	$7.76 \cdot 10^{-1}$	$6.55 \cdot 10^{-1}$	-0.39	0.6988
Pursat	$-3.61 \cdot 10^{-1}$	$6.97 \cdot 10^{-1}$	$4.09 \cdot 10^{-1}$	-0.88	0.3771
Rattanakiri	$4.33 \cdot 10^{-1}$	1.54	$3.62 \cdot 10^{-1}$	1.20	0.2311
Siem Reap	$9.18 \cdot 10^{-1}$	2.51	$3.35 \cdot 10^{-1}$	2.74	0.0061
Stung Treng	$5.77 \cdot 10^{-1}$	1.78	$3.78 \cdot 10^{-1}$	1.53	0.1268
<i>PfPM2</i> amplification	3.44	$3.12 \cdot 10^1$	$4.17 \cdot 10^{-1}$	8.26	$<2.10^{-16}$

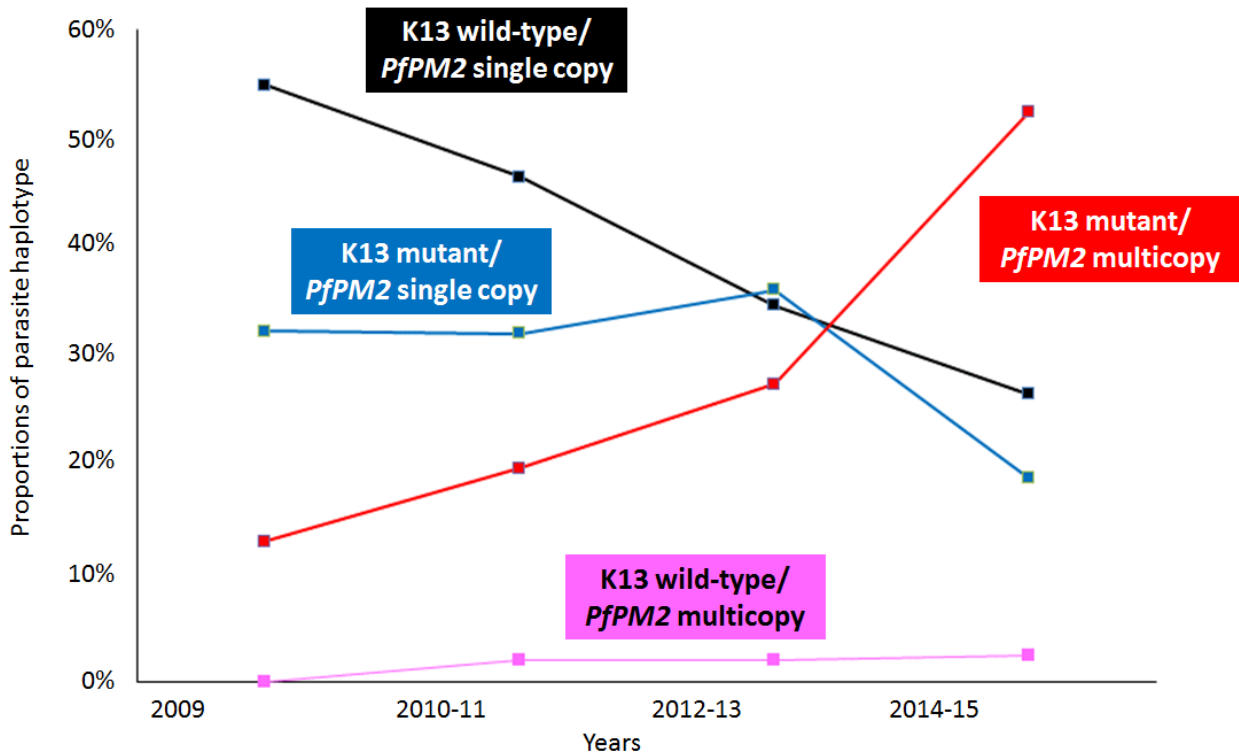
Likelihood ratio test=267 on 12 df, p=0

n= 725, number of events= 119

Appendix 10:

Panel A. Over-time trends of the proportions of isolates with different genetic background (K13 wild-type/*PfPM2* single copy, K13 mutant/*PfPM2* single copy, K13 mutant/*PfPM2* multicopy and K13 wild-type/*PfPM2* multicopy) observed in Cambodia from 2009 to 2015.

The color code is the same as in Figure 4B: K13 wild-type/*PfPM2* single copy (black), K13 wild-type/*PfPM2* multicopy (light purple), K13 mutant/*PfPM2* single copy (blue) and K13 mutant/*PfPM2* multicopy (red).



Panel B. Proposed scenario of the stepwise selection process for the emergence DHA-PPQ resistant parasites in Cambodia.

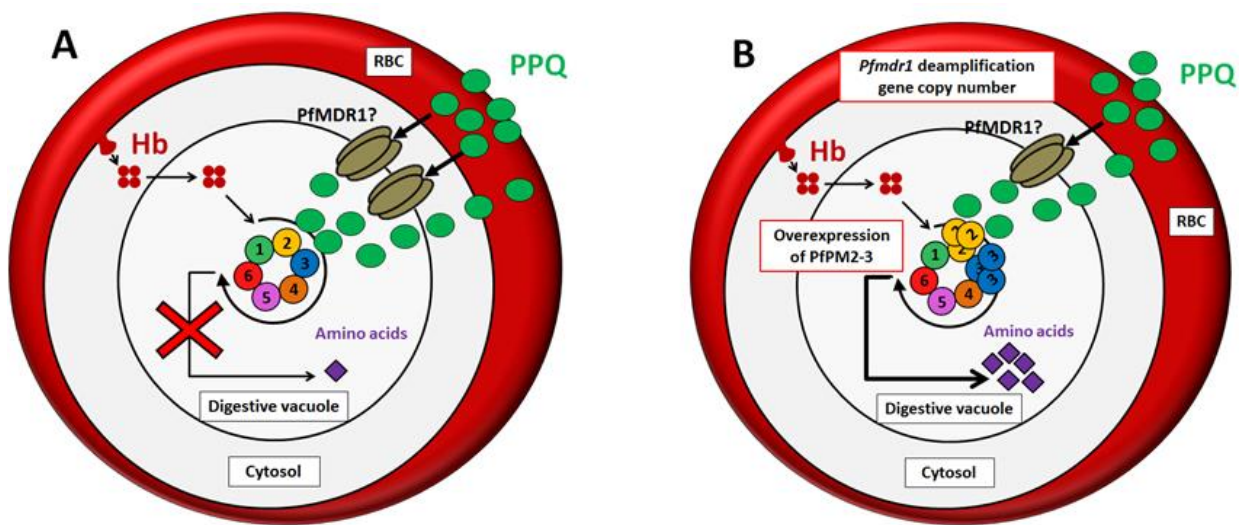


The thickness of the arrow is proportional to the probability of the selection process.

Appendix 11. Hypothesis supporting the mechanisms of resistance of *P. falciparum* parasites to PPQ through the amplification of *PfPM2* and *PfPM3* genes, and the deamplification of *Pfmdr1* gene in Cambodia.

Panel A. PPQ-sensitive parasite. PPQ accumulates in the food digestive vacuole via its weak-base properties. PfMDR1 transporter might help concentrate PPQ into the digestive vacuole, explaining the selection against multicopy *Pfmdr1*. PPQ inhibits hemoglobin degradation leading to the disruption of amino acid production. Parasite death is provoked.

Panel B: PPQ-resistant parasite. Amplification of the *PfPM2* and *PfPM3* genes and increased production of the PfPM2 and PfPM3 proteases is proposed to compensate for the PPQ inhibition of hemoglobin catabolism, restoring normal globin-derived peptide levels used for amino acid production, adequate intracellular osmotic pressure and promoting parasite survival. PfCRT might play a possible role in the efflux of PPQ from the digestive vacuole in some PPQ-resistant parasites.



Hb: hemoglobin; 1: PfPM1; 2: PfPM2; 3: PfPM4; 4: PfFalcipain; 5: PfPM3; 6: PfFalcilysin.