# Emergence of artemisinin-resistant *Plasmodium falciparum* with *kelch13* C580Y mutations on the island of New Guinea

Olivo Miotto[1,2,3,4,*#], Makoto Sekihara[5*], Shin-Ichiro Tachibana[5], Masato Yamauchi[5], Richard D Pearson[1,3], Roberto Amato[3], Sonia Gonçalves[3], Somya Mehra[6], Rintis Noviyanti[7], Jutta Marfurt[8], Sarah Auburn[2,4,8], Ric N Price[2,4,8], Ivo Mueller[6], Mie Ikeda[5], Toshiyuki Mori[5], Makoto Hirai[5], Livingstone Tavul[9], Manuel Hetzel[10,11], Moses Laman[9], Alyssa Barry[6,12,13,14], Pascal Ringwald[15], Jun Ohashi[16], Francis Hombhanje[17], Dominic P Kwiatkowski[1,3], Toshihiro Mita[5,#]

[1] MRC Centre for Genomics and Global Health, Big Data Institute, University of Oxford, Oxford, UK

[2] Mahidol-Oxford Tropical Medicine Research Unit, Mahidol University, Bangkok, Thailand

[3] Wellcome Sanger Institute, Hinxton, UK

[4] Centre for Tropical Medicine and Global Health, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK

[5] Department of Tropical Medicine and Parasitology, Juntendo University Faculty of Medicine, Tokyo, Japan

[6] Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia

[7] Eijkman Institute for Molecular Biology, Jakarta, Indonesia

[8] Global and Tropical Health Division, Menzies School of Health Research and Charles Darwin University, Darwin, Australia

[9] Papua New Guinea Institute of Medical Research, Madang, Papua New Guinea

[10] Swiss Tropical and Public Health Institute, Basel, Switzerland

[11] University of Basel, Basel, Switzerland

[12] University of Melbourne, Melbourne, Australia

[13] School of Medicine, Deakin University, Geelong, Australia

[14] Burnet Institute, Melbourne, Australia

[15] World Health Organization, Geneva, Switzerland

[16] Department of Biological Sciences, Graduate School of Science, University of Tokyo, Tokyo, Japan

[17] Centre for Health Research & Diagnostics, Divine Word University, Madang, Papua New Guinea

[*] Contributed equally

[#] Address correspondence to Toshihiro Mita (tmita@juntendo.ac.jp) or Olivo Miotto (olivo@tropmedres.ac)

Running title: Independent Emergence of *kelch13* C580Y in New Guinea

## Abstract

The rapid and aggressive spread of artemisinin-resistant Plasmodium falciparum carrying the kelch13 C580Y mutation is a growing threat to malaria elimination in Southeast Asia, but there is no evidence of their spread to other regions. We conducted cross-sectional surveys in 2016 and 2017 at two clinics in Wewak, Papua New Guinea (PNG) where we identified three infections caused by C580Y mutants among 239 genotyped clinical samples. One of these mutants exhibited the highest survival rate (6.8%) among all parasites surveyed in ring-stage survival assays (RSA) for artemisinin. Analyses of kelch13 flanking regions, and comparisons of deep sequencing data from 389 clinical samples from PNG, Indonesian Papua and Western Cambodia, suggested an independent origin of the Wewak C580Y mutation, showing that the mutants possess several distinctive genetic features. Identity by descent (IBD) showed that multiple portions of the mutants' genomes share a common origin with parasites found in Indonesian Papua, comprising several mutations within genes previously associated with drug resistance, such as mdr1, ferredoxin, atg18 and pnp. These findings suggest that a P. falciparum lineage circulating on the island of New Guinea has gradually acquired a complex ensemble of variants, including kelch13 C580Y, which have affected the parasites' drug sensitivity. This worrying development reinforces the need for increased surveillance of the evolving parasite populations on the island, to contain the spread of resistance.

## Introduction

The global adoption of artemisinin-based combination treatments (ACTs), consisting of an artemisinin derivative and a partner drug, has played a key role in the worldwide reduction of malaria cases and deaths.[1] In the last decade, however, *Plasmodium falciparum* parasites with decreased sensitivity to artemisinin have emerged at multiple locations in the Greater Mekong Sub-Region (GMS).[2-4] These resistant parasites exhibit a reduced clearing rate during treatment,[2,3,5] such that treatment efficacy relies more heavily on the partner drug. As a result, resistance to the partner drug may emerge, as has recently been the case with the partner drug piperaquine, leading to ACT treatment failures.[4,6]

In the GMS, slow artemisinin clearance rates have been associated with genetic mutations in two domains of the parasite's *kelch13* gene.[7] Numerous independently arising *kelch13* mutations have been found in GMS resistant parasites,[3,8-10] and several of these mutations have been validated in relation with clinical and *in vitro* artemisinin resistance.[11] Increasingly, the genotyping of *kelch13* is being used for the surveillance of artemisinin resistance.[3,8-10] After an initial period in which many *kelch13* resistant alleles circulated concurrently, diversity has declined dramatically in the eastern part of the GMS (comprising Cambodia, northeast Thailand, southern Laos and Vietnam) where the C580Y allele has become the most prevalent, replacing other circulating mutations.[12] This allele is also associated with reduced clearing rate western Thailand,[5] where it has emerged independently.[13] The C580Y allele has also been detected in parasites outside the GMS, in Africa[8] and in Guyana, South America;[14] however, in these instances the artemisinin clearance rates were not known.

While the GMS has been at the epicenter of resistance to antimalarials multiple times over the last sixty years, resistance has also repeatedly arisen independently on the island of New Guinea, producing alternative resistant genetic variants.[15,16] This makes New Guinea a likely candidate for the emergence of novel forms of artemisinin resistance, and underlines the importance of monitoring locally circulating strains. The island of New Guinea is divided between the countries of Papua New Guinea (also known as "PNG") to the east, and Indonesia to the west (referred to as "Indonesian Papua" here). PNG has been using artemether-lumefantrine as the first-line treatment for malaria since late 2011,[17] and delayed parasite clearance has not been reported to date.[18] We previously reported the absence of *kelch13* mutations in PNG samples obtained in 2002-2003 in Wewak Town, East Sepik,[9] consistent with results from a later broader survey;[10] however, studies of genetic variants across the island have only been conducted at a limited number of sites. A recent report of an Australian citizen infected with a C580Y mutant after returning from PNG raised the prospect that artemisinin resistance may have emerged undetected.[19] To investigate this possibility, we conducted a cross-sectional study in 2016-2017 at the same sites as our previous study in Wewak. Here, we analyze clinical and genetic data from 257 cases surveyed in this study, to evaluate the potential emergence of artemisinin resistance. We assessed sensitivity to artemisinin using the ring-stage survival assay (RSA), a sensitive ex-vivo method,[20] and used next-generation sequencing data to reconstruct the epidemiology of *kelch13* mutants found.

## Results

### Identification and characterization of *kelch13* C580Y mutants

This study was performed in 2016 and 2017 at two clinics in Wewak Town, East Sepik in PNG (Figure 1, Supplementary Notes), enrolling 257 patients with uncomplicated malaria, 18 of whom (7.0%) had taken antimalarial drugs in the previous two weeks, and were therefore excluded from further analysis (Table 1). We were able to genotype the *kelch13* gene in 239 parasitized blood samples (Supplementary Table 1). None of the 2016 parasites carried *kelch13* mutations; in 2017, however, the C580Y allele was found in three samples (2.3%) labelled PNG-C580Y-1 to PNG-C580Y-3 (Supplementary Table 2). The patients carrying the mutant parasites were aged 7, 16 and 25 years, and presented with parasite densities of 0.35, 0.07 and 0.44%, respectively. They lived within 2 km of each other and had no history of foreign travel (Supplementary Figure 1).

We used RSA[20] to measure artemisinin susceptibility in samples from cases with parasitemia >0.1% (n=117), including two of the C580Y mutants. The assays yielded interpretable results for 57 parasites (Supplementary Table 3), including PNG-C580Y-1. Undetermined results were mainly due to insufficient parasite growth (n=52) or low-quality blood smears (n=8). Survival rate could not be determined for PNG-C580Y-3, due to poor quality blood smears in non-exposure controls. Five samples (8.8%) exhibited a survival rate >1% (range: 1.6–6.8%) at 72 hours following dihydroartemisinin exposure, a threshold that correlates with parasite clearance half-life >5 hours in the GMS.[20] The PNG-C580Y-1 *kelch13* mutant produced the highest survival rate (6.8%), comparable to rates measured in Cambodian C580Y parasites,[7] suggesting that Wewak C580Y mutants possess levels of ring-stage artemisinin resistance similar to those in the GMS (Supplementary Figure 2).

### Assessment of *kelch13* mutants' genetic background

To investigate whether the Wewak C580Y mutants were imported from Southeast Asia, we first tested for the genetic background underpinning *kelch13* mutations in the GMS[21]. Of six characteristic markers tested, only one (*ferredoxin* D193Y) was found in Wewak C580Y mutants, compared to four or five in GMS artemisinin-resistant clones (Supplementary Table 4). The *ferredoxin* variant was detected in all three PNG kelch13 C580Y parasites, but was only found in one in ten randomly selected PNG parasites with wild-type *kelch13*. All 2017 samples were tested for amplifications of the *plasmepsin2-3* and *pfmdr1* genes, which are associated with resistance to the partner drugs piperaquine and mefloquine. All carried a single copy of *plasmepsin2-3*, while only a single *kelch13* wild-type parasite was found to possess an amplified *pfmdr1* (Supplementary Table 1).

### Genome-level relationship of *kelch13* mutants with other populations

To study in greater detail the provenance of the C580Y mutants, we compared whole genome sequencing (WGS) data from 389 high-quality samples with low within-sample diversity, including 73 from the present study, 83 from other PNG locations, 88 from Timika in Indonesian Papua, and 145 from western Cambodia (Table 2, sites in New Guinea shown in Figure 1). All three *kelch13* C580Y mutants were included in this analysis, and no other parasite from Indonesia or PNG carried *kelch13* mutations.

To cluster individuals by genetic similarity, we conducted a principal coordinates analysis (PCoA) based on genome-wide pairwise genetic distances. This showed a clear separation between Cambodian and New Guinean parasites, with no samples bridging the gap between these two populations (Supplementary Figure 3). To investigate the relationship between parasites within the island of New Guinea, we conducted a PCoA without Cambodian parasites (Figure 2A). In this analysis, the first component (PC1) clearly separates Papua Indonesia and PNG, showing that the parasite populations at the PNG sites surveyed are more similar to each other than they are to Indonesian populations. However, a number of individuals from both countries, including the Wewak C580Y parasites, occupy an intermediate position along PC1, consistent with a degree of admixture between the two populations. Furthermore, we observe that two Timika parasites cluster with the PNG population, suggesting recent importation from PNG into Timika.

Further evidence of admixture was provided by population structure analysis performed by the fastSTRUCTURE software (Supplementary Figure 4). This analysis indicated that K=6 putative ancestral populations best explained the underlying population structure (Figure 2B). Applying this model, we observe that New Guinea parasites mostly derive from two major ancestral populations, one in Papua Indonesia and one in PNG, with the exception of a small subpopulation in Wewak (gray bars). In both countries, however, several samples were reported to be significantly admixed (e.g. about 16% of samples from both PNG and Papua Indonesia shared >10% ancestry with the dominant group of the other region), suggesting frequent exchanges between the two populations. The Wewak C580Y mutants in particular were found to be amongst the most admixed in PNG, all three individuals estimated to have inherited 35-36% of their genome from the Indonesian ancestral population. Like all New Guinea samples, Wewak C580Y mutants appear to carry no contribution from GMS ancestral populations, ruling out the hypothesis that they are recent imports from that region.

## Origins of the Wewak *kelch13* C580Y haplotype

Even though the mutants were not recently imported from the GMS, it is still possible that a C580Y mutation originated in the GMS could have been acquired by New Guinea parasites through recombination, analogous to the *crt* haplotype of Asian origin that circulates in chloroquine-resistant African *P. falciparum*.[22] To investigate this possibility, we compared the regions flanking *kelch13*, to search for haplotypes matching those in the Wewak C580Y mutants. We tested 12 microsatellite loci and six SNPs flanking *kelch13*[14,23] in the 2017 parasites from Wewak, and eight other isolates for comparison (Supplementary Table 5), producing 64 unique haplotypes in 96 parasites. Upstream from *kelch13* (left flank), we found strong similarities between the haplotype sequence of the Wewak C580Y mutants, and that in both Cambodian resistant parasites (e.g. MRA-1236) and PNG wild type haplotypes (e.g. H5 and H13, found in 7 samples). Downstream from *kelch13*, however, the C580Y mutants bear greater similarity to the sequences of PNG haplotypes (e.g. H46 and H48) than to those in Cambodia. Overall, the H5 and H13 haplotypes from PNG were the most similar to those of the mutants, suggesting an independent local emergence rather than a GMS origin.

Similar results were obtained when comparing flanking haplotypes constructed by concatenating WGS genotypes at SNPs with minor allele frequency >= 0.01 in a region of 300kbp centered at *kelch13* (Supplementary Figure 5). Downstream of *kelch13*, the C580Y mutant flanking haplotype is

similar to those circulating in Cambodia, but is also found in *kelch13* wild-type parasites in New Guinea, particularly in Timika. In the upstream flank, considerably longer matches (~60kbp) are found in Indonesian and PNG samples than in Cambodian ones. Taken together, the above results provide no support for a Southeast Asian origin of the PNG *kelch13* C580Y haplotype. Since the mutant genomes show low similarity to GMS parasites, it is reasonable to conclude that the mutation is most likely to have emerged independently in New Guinea.

## Ancestry reconstruction of the PNG *kelch13* C580Y genomes

The level of admixture in the C580Y mutants suggest a genetic contribution from the Indonesian population, although C580Y mutations have not been reported in Indonesian Papua. One possibility is that the Wewak mutants originate from a geographically intermediate location, and are the result of long-term continuous recombination between Indonesian and PNG populations. A different hypothesis is that the mutants have acquired through recombination, and retained, portions of genomes circulating in Indonesian Papua which provide some selective advantage. The first scenario would result in a random distribution of shared alleles, while the second would likely produce a limited number of long shared haplotypes.

To evaluate these hypotheses, we performed a genome-wide estimation of *identity by descent* (IBD) between each pair of samples. IBD models recombination processes, comparing genotype sequences to identify genome regions that are likely to be identical because of common ancestry rather than by chance. We estimated the proportion of IBD between each pair of parasites genome, and for each of the sites we compared the IBD proportion in the Wewak C580Y mutants against that in the other (*kelch13* wild-type) Wewak parasites (Figure 3). The results indicate that mutants possess significantly higher IBD with Timika parasites than the Wewak wild-type parasites (medians IBD proportion: 14.4% vs 8.0%, $p<10^{-117}$ by Mann-Whitney test), and significantly less IBD with parasites from most PNG sites (Wewak medians: 8.9% vs 11%, $p<10^{-48}$; Madang medians: 9.5% vs 11.3%, $p<10^{-7}$; Maprik medians: 7.8% vs 10.7%, $p<10^{-37}$). An IBD network based on genome-wide data shows that parasites from Wewak, Timika and Cambodia form three large haplotype-sharing clusters, while the Wewak mutants constitute a separate group, confirming their genome composition is not typical of any of these populations (Figure 4A). Similarly, when restricting the analysis only to the regions flanking the *kelch13* gene, an IBD network shows that the haplotypes accompanying the C580Y mutations are distinct from those circulating elsewhere and, in particular, they are different from those in both Cambodia and Papua Indonesia (Figure 4B).

The observation that the Wewak C580Y mutants share a higher proportion of IBD with Timika than with local wild-type parasites might appear to contradict the finding that the proportion of PNG ancestry in these mutants is greater than that of Indonesia, as was determined by fastSTRUCTURE. However, these results could be reconciled if the mutants' genomes are the product of long-term mixing of PNG parasites, resulting in a random distribution of alleles common in PNG, with recent segments acquired from the Indonesian population, which are identified as IBD, but occupy a smaller portion of the genome. To test this hypothesis, we selected SNPs that are highly differentiated between the Wewak and Timika populations ($F_{ST} \geq 0.3$), and labelled them according to which of the two populations has the allele present in the C580Y mutants as its dominant allele. Of the 467

selected SNPs, 57.6% (n=269) carried the Wewak allele, consistent with a predominantly PNG-like genetic background. When mapped across the genome, the majority of Timika-like alleles do not appear randomly distributed, but rather clustered in regions where the C580Y mutants are in IBD with a high proportion of the Timika populations (Figure 5). For example, 63% of variants where the mutants carry the Timika-like allele are in regions where there is IBD with ≥ 10% of the Timika population (30% of all variants); outside of these regions, 77% of SNPs exhibit Wewak-like alleles.

Although some stretches of the C580Y genomes share IBD with both Timika and Wewak parasites, possibly indicating earlier gene flow between these populations, there are a number of loci where IBD haplotypes are only observed in Indonesian parasites- notably in chromosomes 3, 5, 10 and 13. The sizes of these haplotypes (generally > 100 kbp) suggests they were recently acquired from a donor population- putatively, one circulating in Indonesian Papua- and possibly under evolutionary selection, raising the question of a possible relation to the C580Y mutation.

### Genetic variants association with *kelch13* C580Y

If the long shared IBD segments were acquired from Indonesian parasites because they contain beneficial alleles, we may expect these alleles to cause amino acid changes, to circulate widely in Indonesian Papua, and to be less common in PNG. Therefore, from the list of highly differentiated SNPs between the Wewak and Timika populations ($F_{ST} \geq 0.3$), we extracted those non-synonymous variants where the C580Y mutants carries the Timika-like allele and ordered them by decreasing $F_{ST}$ (n=144, Supplementary Table 6). Several of these variants, particularly those in regions of IBD with a large proportion of the Timika parasites, are implicated in drug response, or located in genes associated with drug resistance.

The most differentiated SNP was *ferredoxin* (PF3D7_1318100) D193Y on chromosome 13, previously identified as a component of the genetic background associated with artemisinin resistance in the GMS.[21] In the large IBD region on chromosome 10 we found three noteworthy hits: the T38I mutation in the autophagy-related protein 18 gene (*atg18*, PF3D7_1012900), associated with artemisinin response and widespread in SE Asia; and two SNPs (N280S and N659S) in the NLI interacting factor-like phosphatase (*nif4*, PF3D7_1012700), implicated in conjunction with *atg18*[24] and identified as associated with artemisinin response.[21] On chromosome 5, another extensive IBD region contains mutations Y184F and N1042D in the *mdr1* gene (PF3D7_0523000), which has been implicated in resistance to multiple antimalarials[25,26] including artemisinin; the Y184F variant is common in the GMS, where it has been under selection.[27] In a different region of chromosome 5, we found the Q225H mutations in the purine nucleoside phosphorylase gene (*pnp*, PF3D7_0513300), which circulates in the GMS and has also been associated with drug resistance.[28] Some of the highly differentiated variants were found in genes associated in drug resistance, rather than directly implicated. Among these, we found K753Q in the *amino acid transporter* PF3D7_0515500, which has previously been implicated in artemether response;[29] G248R in gene PF3D7_1450800, associated to artemisinin resistance in a large-scale study;[21] two amino acid changes in an ABC transporter (PF3D7_0319700) shown experimentally to be a drug target;[30] and the G233R mutation in *plasmepsin III* (*pmIII/hap*, PF3D7_1408100) which is involved in resistance to piperaquine.[31] IBD networks

restricted to flanking regions of specific drug resistance genes such as *atg18* and *pnp* confirmed that Wewak mutants clearly cluster with the Timika population (Figure 4C and 4D).

## Discussion

Over several decades, the spread of drug resistance from the GMS has rendered multiple drugs ineffective in Africa, at the cost of hundreds of thousands of lives. Therefore, the current rise of artemisinin resistance in southeast Asia is an urgent concern, since fast-acting artemisinin combination therapies are the preferred frontline treatments in nearly all endemic countries. The GMS is not the only region with the conditions for the development of drug resistance: historically, resistance to multiple drugs, such as chloroquine[32] and sulfadoxine,[16] has emerged independently on the island of New Guinea, and therefore new resistant strains could develop there.

In this study, we identified three infections from parasites carrying the *kelch13* C580Y mutation, the most widespread artemisinin resistant allele. C580Y has rapidly overtaken other *kelch13* variants, becoming dominant in large parts of the GMS,[12] and was confirmed by a transfection study to confer resistance *in vitro*.[33] The discovery of three C580Y mutants in Wewak, almost identical at genomic level, raises multiple important questions. First, are these parasites actually resistant to artemisinin? Second, were these mutants imported into PNG as a result of the recent aggressive spread of C580Y in the GMS, or are they parasites native to New Guinea that have acquired the mutation either through independent emergence, or through recombination with Asian mutants? Third, do these mutants represent a confined local phenomenon, or are they representatives of a spreading population across the island, perhaps across national borders? Fourth, if the C580Y mutations were not imported, did they result from an isolated “chance event”, or from a gradual evolutionary process that selected a genetic background to boost fitness? The results presented here provide convincing evidence to answer all these questions.

Although clinical parasite clearance rates were not measured in this study, *in vitro* tests clearly showed that the Wewak mutants are resistant to artemisinin in their ring stage, consistent with C580Y mutants elsewhere. Since no failures occurred in patients infected with these parasites after artemether-lumefantrine treatment, there is no reason to suspect that this ACT is no longer efficacious, and the clinical significance of the C580Y allele in New Guinea needs to be clarified by *in vivo* efficacy studies. However, parasite populations stand to benefit from decreased sensitivity even without treatment failures: slower clearance increases drug exposure, boosting the parasites’ chances to develop further resistance. Recent research in the GMS showed that continued drug exposure leads to accumulation of mutations that confer higher levels of resistance, facilitating the spread of multidrug resistant lineages.[12]

We found no evidence that the C580Y mutations found in Wewak have been imported from the GMS. At whole-genome level, the mutants appear more similar to New Guinean parasites than to Cambodian ones, and analyses of flanking haplotypes and IBD showed no evidence that the C580Y mutations were acquired from imported parasites. Overall, there is a strong case for an independent

emergence of the C580Y mutation on the island of New Guinea, although our data is insufficient to pinpoint a precise geographic origin.

At genome-wide level, the Wewak C580Y mutants are substantially different from the majority of our PNG samples. The majority of their genetic background is "PNG-like", in that it carries alleles commonly found in Wewak, with relatively short IDB fragments suggesting long-term interbreeding. However, a substantial proportion of C580Y mutants' genome consists of large haplotypes shared with many of the Timika parasites, suggesting these segments were acquired from a common source population, strongly connected to the Papua Indonesia parasites. Furthermore, the length of the shared haplotypes suggests that they contain regions under selection.[34] This hypothesis is bolstered by the large number of alleles associated with resistance to artemisinin and related drugs, several of which are located in IBD fragments shared with the Indonesian Papua parasites, and absent from other PNG samples. In an IBD shared segment we found the *ferredoxin* D193Y mutation, which strongly correlates with artemisinin resistance variants in the GMS,[21] and is otherwise absent from PNG. The drug resistant allele *mdr1* Y184F, known to be under selection in the GMS,[27] is also present in a large IBD segment, with another less common *mdr1* mutation, N1042D. *mdr1* is a transporter contributing to the efflux of toxic substances from the parasite's food vacuole, and its polymorphisms have been associated with resistance to multiple drugs. Another IBD segment on chromosome 10 harbours the *atg18* T38I variant, associated with decreased sensitivities to dihydroartemisinin and artemether (the artemisinin derivative used in PNG) in the China-Myanmar border area,[24] and two non-synonymous changes in the nearby *nif4* gene; other mutations in *nif4* were previously associated with artemisinin sensitivity in two independent genome-wide association studies in Southeast Asia.[21,24] We also observed a mutation in a nucleoside phosphorylase (*pnp* Q225H), recently shown to affect response to quinine and mefloquine,[28] two drugs related to lumefantrine, the partner drug used in PNG.

Such a high number of mutations implicated in drug response, both highly differentiated geographically and located in regions with high IBD counts, strongly suggests these variants have been acquired and selected for their contribution to a resistant phenotype. Therefore, we propose that one or more New Guinean lineages have accumulated a complex genetic background through recombination events and selection under artemisinin drug pressure, and that the *kelch13* C580Y mutation is a component of a complex constellation of genetic changes, rather than a standalone mutation generated by a chance event. This evolutionary process appears to be analogous to that which has produced artemisinin resistance in the GMS, where *kelch13* mutations emerged in parasites that had acquired a complex set of genetic changes.[21] In both cases, it appears that several of these changes were present before ACT treatment failures occurred, suggesting they provided improved fitness of the parasite population without major effects on clinical outcomes. In scenarios where drug resistance requires a gradual build-up of genetic changes, clinical efficacy monitoring is only likely to detect resistance at an advanced stage of the evolutionary process, when failures begin to occur. In the future, it is possible that novel analyses of evolutionary patterns, using genomic surveillance data, could allow earlier detection of drug resistance, and interventions ahead of clinical failure.

The findings presented are of great concern for public health, both locally and globally. They provide evidence of artemisinin-resistant parasites within New Guinea that have acquired a complex genetic background providing a survival advantage under artemisinin drug pressure, which may have intensified recently, due to decreasing *P. falciparum* transmission as a result of control interventions.[35]

Such a complex genetic makeup is likely to have occurred through a gradual process, rather than in a single chance event, and it is therefore reasonable to suppose that the individuals sampled in Wewak are part of a larger evolving population. Given that no evidence of *kelch13* mutants have been found by other studies on the island, it is likely that this population has emerged recently. Genetic data from the island of New Guinea is currently insufficient to demarcate the extent of the spread of this population, which may be present on either side of the Indonesia-PNG border, or both. These parasites are a potential danger to the efficacy of ACTs in New Guinea and could constitute a threat if they are not contained. Public health authorities in PNG and collaborating research institutions are currently ramping up genetic monitoring of malaria parasites at sentinel sites, and we hope this will provide much needed detailed data to map artemisinin resistant parasites and help develop containment strategies.

## Materials and methods

### Clinical blood sample collection

This study was performed as part of an ex-vivo antimalarial drugs susceptibility study, conducted in January–February 2016 and in January–February 2017 at two clinics (Town and Wirui Urban) located 2 km apart in Wewak, East Sepik, PNG.[36] All patients > 1 year of age showing malaria-suspected symptoms (axillary temperature ≥37.5C°, fever during the previous 24 h, headache, abdominal pain, nausea and/or diarrhea) were screened by Rapid Diagnosis Test (RDT) (CareStart™ Malaria HRP2/pLDH COMBO Test kit, Access Bio, USA). When a *P. falciparum*-positive result was obtained, patients were enrolled after obtaining informed consent from the patients or guardians.

Blood samples were obtained by finger prick (< 2 years, 100–500 µL) or peripheral venipuncture (≥ 2 years, 1 mL) and collected into EDTA-containing tubes and immediately transferred to the central laboratories at Wewak General Hospital. To estimate parasitemia, thick and thin blood smears were prepared and stained with 2% Giemsa for 30 min. For molecular analysis, blood samples were transferred onto chromatography filter paper (ET31CHR; Whatman Limited, Kent, UK) and separated in a plastic bag after drying at a normal temperature and stored at -20°C. *Plasmodium falciparum* DNA was extracted from a quarter of a blood spot (25 µL) using the QIAamp DNA blood Mini Kit (QIAGEN, Hilden, Germany).

In total, 257 patients were enrolled for this study: 129 at Wirui Urban clinic and 128 at Town clinic (Table 1). Nearly all background characteristics of enrolled patients were the same between the clinics: the only difference was found in the frequency of pre-treated patients, which was significantly higher at Town clinic (11.7%) than that at Wirui clinic (2.3%) ($P$ = 0.0032, Fisher-exact test). In total, 18

enrolled patients had reported ingesting antimalarials in the previous two weeks and were excluded from further analysis; artemether alone (n = 7) and chloroquine (n = 7) were the two most used forms of self-medication. No severe case was enrolled in the study.

Ethical approvals were obtained from the Medical Research Ethical Committee of Juntendo University, Tokyo, Japan (No. 13-016) and the Medical Research Advisory Committee of PNG National Department of Health (No. 14.22 & 16.41).

## Genotyping and Whole-Genome Sequencing

Genomic DNA was extracted from a quarter of blood spot (approx. 25 µL) using the QIAamp DNA blood Mini Kit (QIAGEN, Hilden, Germany) with a modified procedure.[37] The *kelch13* gene propeller domain sequence was determined by capillary sequencing of PCR products, after removal of primers and dNTPs with the ExoSAP-IT reagent (Affymetrix, CA, USA) as previously described.[38] Additional targeted genotyping was performed as described in the appendix.

Extracted DNA underwent selective whole genome amplification[39] prior to whole-genome sequencing, to enrich parasite DNA and reduce human DNA contamination. Sequence data were generated at the Wellcome Sanger Institute with Illumina short-read technology, and deposited in the European Nucleotide Archive (ENA, https://www.ebi.ac.uk/ena; Supplementary Table 7). Read counts at 1,043,334 biallelic SNPs in the nuclear genome were called with the V6.0 standardized analysis pipeline of the Plasmodium falciparum Community Project.[40,41] Genotypes were called only with a coverage of five or more reads, and alleles were disregarded if supported by fewer than 2 reads, or 5% of reads when coverage was above 50. To minimize errors and biases, we excluded samples with insufficient coverage at more than 25% of the SNPs (with the exception of one Wewak sample that had been identified as a C580Y mutant), and removed all SNPs that either were invariant after those samples were removed, or had insufficient coverage in more than 25% of the remaining samples, leaving 35,664 SNPs to be used in our analysis.

For comparisons, we used samples included in the MalariaGEN *P. falciparum* Community Project 6.2 data release (https://www.malariagen.net/projects/p-falciparum-community-project) sampled from Madang (contributed by IM), Maprik and Alotau (contributed by AB) in PNG; Indonesian Papua samples from Timika (contributed by RN); and Cambodian samples from Pailin and Pursat provinces, contributed to the Pf3K reference dataset by the Tracking Artemisinin Resistance Collaboration (TRAC)[3] (https://www.malariagen.net/projects/Pf3k). These additional samples were processed by the same methods as the Wewak samples.

To minimize errors in haplotype reconstruction due to mixed infections, we filtered samples by their $F_{WS}$ index, estimated from as previously described.[41] By removing samples with $F_{WS}$<0.95, we obtained a final set of 389 essentially monoclonal samples. To estimate from genotypes the $F_{WS}$ indexes, the allele frequencies in each population, and the differentiation measure $F_{ST}$, we applied previously published methods.[21]

## Ex-vivo ring-stage survival assay (RSA)

Ex-vivo RSA was performed as previously described[20,38] on samples with parasitemia of 0.1% or above. Initial parasite densities above 1%, parasitemia were adjusted to 1% by adding uninfected type O erythrocytes. Parasite culture mixture (100 µL per well) was dispensed with 700 nmol/L dihydroartemisinin (Tokyo Chemical Industry Co. Ltd., Tokyo, Japan). After 6 h exposure, pellets were washed to remove the drug, and incubated for 66 h. 0.1% dimethyl sulfoxide was used as control. Densities of viable parasites were determined by two investigators (MS and TM) by counting 10,000 erythrocytes, and survival rates were calculated as the ratios of parasites in exposed and non-exposed cultures.[20] Laboratory adapted artemisinin-susceptible (3D7) and resistant clones (MRA-1236 and MRA-1240, contributed by Didier Ménard of Institut Pasteur du Cambodge), were provided by the Malaria Research and Reference Reagent Resource Center (MR4) and obtained through BEI Resources, NIAID, NIH, to be used as a comparators for ex-vivo RSA.[20]

## Determinations of *kelch13* flanking haplotypes

**Using microsatellites and selected SNPs.** We genotyped *kelch13* flanking microsatellite loci and single-nucleotide polymorphisms (SNPs) located at -33.9 kb, -8.1 kb, -5.02 kb, -1.91 kb, 4.1 kb, 9.33 kb, 15.82 kb, and 43.85 kb from the *kelch13* gene (negative and positive values denote upstream and downstream distance from *kelch13*, respectively) by direct sequence as previously described[14,23]. We also genotyped four KEL1/PLA1 marker SNPs located at -137.5 kb, -14.06 kb, 6.97 kb, and 24.91 kb[13]. At positions -137.5 kb and 24.91 kb, SNPs were determined by direct sequencing of PCR products. At position 6.97 kb, SNP-specific amplification was performed with HiDi DNA polymerase (myPOLS Biotec, Konstanz, Germany). At position -14.06 kb, SNP was detected by combination of Derived Cleaved Amplified Polymorphic Sequences (dCAPS) assay and PstI digestion. These analyses produced genotypes for 18 polymorphic sites, since multiple polymorphisms were identified at positions -8.1 kb (TAT repeat and one SNP site), -5.02 kb (TAAAA, TAAA and TA repeats and one SNP site), -1.91 kb (TAA and TA repeats), 9.33 kb (TAAA and TA repeats). Primers used for these analyses were shown in Supplementary Table 8.

**Using whole-genome sequencing data.** We constructed flanking haplotypes by concatenating WGS genotypes at SNPs with minor allele frequency (MAF) >=0.01 in a region of 300kbp centered at *kelch13*.

Comparison and ranking of *kelch13* flanking haplotypes.

To assess flanking haplotypes, we assigned for each sample a score based on the extent of haplotype sequence identity with the consensus haplotype of the Wewak *kelch13* mutants. To obtain a score for the upstream (left-hand) flank, we started from the position nearest to the core locus (the kelch13 gene), and counted the number of consecutive positions (ignoring missing/heterozygous positions) in the upstream direction that carried an allele identical to that in the C580Y mutants, until a mismatch was found. The score for the downstream (right-hand) flank is determined in the same way, by traversing in the downstream direction from the end of *kelch13*. The final score for the sample was obtained by adding the scores for the two flanks.

## Genotyping of background mutations associated with artemisinin resistance

Genotyping of background alleles associated with artemisinin resistance (D193Y in *fd*, T484I in *mdr2*, V127M in *arps10*, I356T in *crt*, V1157L in *nif4* and C1484F in *pibp*)[21] were performed by multiplex PCR and probe qPCR as previously described.[42] Copy number variants of plasmepsin 2 and multidrug resistance gene 1 genes were determined using published procedures.[43] Four artemisinin-resistant clones (MRA-1236, MRA-1238, MRA-1240 and MRA-1241) and three isolates obtained in Cambodia[16] were used as comparators.

## Analysis of ancestry and relatedness

Analyses of population structure were performed using a combination of freely available tools, and custom software programs written in Java and R. Identity by descent (IBD) analysis was performed on the genotypes obtained from whole-genome sequencing using the program hmmIBD[44] using default parameters. Further estimates of IBD, producing IBD networks, were generated by the program isoRelate.[45] We constructed an $NxN$ pairwise genetic distance matrix, where $N$ is the number of samples, estimating genome-wide genetic distance using a previously published procedure.[8] In addition, we constructed an $NxN$ pairwise IBD distance matrix, estimating the IBD distance as $d_{IBD} = 1 - p_{IBD}$, where $p_{IBD}$ is the fraction of the genome that is predicted to be identical by descent in a pair of samples. PCoA analyses were carried out using the `stats` package in the R language (https://www.r-project.org/), while neighbour-joining trees were produced by the R `ape` package.[46] Population structure and admixture analyses we performed using the fastSTRUCTURE software in the Bioconda environment.[47,48] We ran fastSTRUCTURE with K=2 to 9 putative ancestral populations, and ran the `chooseK` tool to determine the value of K that provided the best explanation of the population structure (K=6).

## Role of the funding source

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

## Acknowledgments

We thank all patients who contributed samples, and their guardians; Steven Tiwara, Alpha Ao, Charlie Amai, Alphonse Coll, Douglas Tambi, Gethrude Kitipa, Julianne Gumbat, Raphael Kariwa, and Shoki Yatsushiro for their kind cooperation in the field; Takahiro Tsukahara for advice of sampling correction in the 2016 survey; the Laboratory of Molecular and Biochemical Research, Research Support Center, Juntendo University Graduate School of Medicine, for technical assistance; the staff of the PNG Institute of Medical Research and Abby Harrison for their contributions to sample collection. Genome sequencing was performed by the Wellcome Sanger Institute (WSI), and sequencing data was processed by the MalariaGEN Resource Centre; this study used data from the MalariaGEN Pf3k Project and Plasmodium falciparum Community Project. We thank the staff of the WSI Sample Logistics, Sequencing, and Informatics facilities for their contribution; Mihir Kekre and Katie Love for their support in the sample processing pipeline; Victoria Cornelius and Kim Johnson for coordinating the MalariaGEN Resource Centre; the TRAC investigators who provided the Cambodian isolates used in this work, including Elizabeth Ashley, Rick M Fairhurst, Chanaki Amaratunga, Arjen Dondorp and Nicholas J White. This study was supported by Grants-in-aid for scientific research from the Japan Society for the Promotion of Science and Foundation of Strategic Research Projects in Private Universities (26460515, 26305015, 17H04074, 18KK0231); the Ministry of Education, Culture, Sports, Science, and Technology of Japan (S0991013); Japan Agency for Medical Research and Development (AMED) (JP15km0908001); the Global Health Innovative Technology Fund (G2015-210); the Global Fund to Fight AIDS, Tuberculosis and Malaria; the National Health and Medical Research Council (NHMRC) of Australia (GNT1027108, APP 1131932), also through the Australian Centre for Research Excellence on Malaria Elimination (ACREME, APP 1134989) and Independent Research Institute Infrastructure Support Scheme; the Australian Government Department of Foreign Affairs and Trade (DFAT) through the Tropical Disease Research Regional Collaboration Initiative (TDRRCI); the Victorian State Government Operational Infrastructure Support. Sequencing, analysis, informatics and management of genomic data were supported by the Wellcome Trust (206194, 090770/Z/09/Z), Bill & Melinda Gates Foundation (OPP11188166), the Medical Research Council of the UK (G0600718) and the Department for International Development (DFID) (M006212). The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. PR is a staff member of the World Health Organization. PR alone is responsible for the views expressed in this publication and they do not necessarily represent the decisions, policy or views of the World Health Organization.

## Author Contributions

MS, SIT, MY, RN, JM, SA, RP, IM, MI, TMo, MHi, LT, MHe, ML, AB, PR, FH, TMi carried out clinical studies to obtain *P. falciparum* samples. MS, SIT, MY, SG, SA, MI, TMo, MHi, AB, JO conducted laboratory analyses. OM, MS, SIT, MY, RDP, RA, SM, AB, TMi performed data analyses. OM, DPK, TMi designed and coordinated the project. OM, TMi drafted the manuscript which was reviewed by all authors.

The authors declare they have no financial and non-financial competing interests.

## References


1. World Health Organization. World Malaria Report 2017; 2017.
2. Dondorp AM, Nosten F, Yi P, et al. Artemisinin resistance in Plasmodium falciparum malaria. *N Engl J Med* 2009; **361**(5): 455-67.
3. Ashley EA, Dhorda M, Fairhurst RM, et al. Spread of artemisinin resistance in Plasmodium falciparum malaria. *The New England Journal of Medicine* 2014; **371**(5): 411-23.
4. Amaratunga C, Lim P, Suon S, et al. Dihydroartemisinin-piperaquine resistance in Plasmodium falciparum malaria in Cambodia: a multisite prospective cohort study. *The Lancet Infectious diseases* 2016; **16**(3): 357-65.
5. Anderson TJ, Nair S, McDew-White M, et al. Population Parameters Underlying an Ongoing Soft Sweep in Southeast Asian Malaria Parasites. *Mol Biol Evol* 2017; **34**(1): 131-44.
6. Thanh NV, Thuy-Nhien N, Tuyen NTK, et al. Rapid decline in the susceptibility of Plasmodium falciparum to dihydroartemisinin–piperaquine in the south of Vietnam. *Malaria journal* 2017; **16**(1): 27.
7. Ariey F, Witkowski B, Amaratunga C, et al. A molecular marker of artemisinin-resistant Plasmodium falciparum malaria. *Nature* 2014; **505**(7481): 50-5.
8. MalariaGEN Plasmodium falciparum Community Project. Genomic epidemiology of artemisinin resistant malaria. *Elife* 2016; **5**.
9. Mita T, Culleton R, Takahashi N, et al. Little Polymorphism at the K13 Propeller Locus in Worldwide Plasmodium falciparum Populations Prior to the Introduction of Artemisinin Combination Therapies. *Antimicrob Agents Chemother* 2016; **60**(6): 3340-7.
10. Menard D, Khim N, Beghain J, et al. A Worldwide Map of Plasmodium falciparum K13-Propeller Polymorphisms. *N Engl J Med* 2016; **374**(25): 2453-64.
11. World Health Organization. Status report on artemisinin resistance and ACT efficacy (August 2018). Geneva: World Health Organization, 2018.
12. Hamilton WL, Amato R, van der Pluijm RW, et al. Evolution and expansion of multidrug-resistant malaria in southeast Asia: a genomic epidemiology study. *Lancet Infect Dis* 2019.
13. Amato R, Pearson RD, Almagro-Garcia J, et al. Origins of the current outbreak of multidrug-resistant malaria in southeast Asia: a retrospective genetic study. *Lancet Infect Dis* 2018.
14. Chenet SM, Akinyi Okoth S, Huber CS, et al. Independent Emergence of the Plasmodium falciparum Kelch Propeller Domain Mutant Allele C580Y in Guyana. *J Infect Dis* 2016; **213**(9): 1472-5.
15. Mita T, Venkatesan M, Ohashi J, et al. Limited geographical origin and global spread of sulfadoxine-resistant dhps alleles in Plasmodium falciparum populations. *The Journal of Infectious Diseases* 2011; **204**(12): 1980-8.
16. Mita T, Venkatesan M, Ohashi J, et al. Limited geographical origin and global spread of sulfadoxine-resistant dhps alleles in Plasmodium falciparum populations. *J Infect Dis* 2011; **204**(12): 1980-8.
17. Rodriguez-Rodriguez D, Maraga S, Lorry L, et al. Repeated mosquito net distributions, improved treatment, and trends in malaria cases in sentinel health facilities in Papua New Guinea. *Malar J* 2019; **18**(1): 364.
18. Tavul L, Hetzel MW, Teliki A, et al. Efficacy of artemether-lumefantrine and dihydroartemisinin-piperaquine for the treatment of uncomplicated malaria in Papua New Guinea. *Malar J* 2018; **17**(1): 350.
19. Prosser C, Meyer W, Ellis J, Lee R. Resistance screening and trend analysis of imported falciparum malaria in NSW, Australia (2010 to 2016). *PLoS One* 2018; **13**(5): e0197369.

20. Witkowski B, Amaratunga C, Khim N, et al. Novel phenotypic assays for the detection of artemisinin-resistant Plasmodium falciparum malaria in Cambodia: in-vitro and ex-vivo drug-response studies. *The Lancet Infectious diseases* 2013; **13**(12): 1043-9.
21. Miotto O, Amato R, Ashley EA, et al. Genetic architecture of artemisinin-resistant Plasmodium falciparum. *Nat Genet* 2015; **47**(3): 226-34.
22. Wootton JC, Feng X, Ferdig MT, et al. Genetic diversity and chloroquine selective sweeps in Plasmodium falciparum. *Nature* 2002; **418**(6895): 320-3.
23. Talundzic E, Chenet SM, Goldman IF, et al. Genetic Analysis and Species Specific Amplification of the Artemisinin Resistance-Associated Kelch Propeller Domain in P. falciparum and P. vivax. *PLoS One* 2015; **10**(8): e0136099.
24. Wang Z, Cabrera M, Yang J, et al. Genome-wide association analysis identifies genetic loci associated with resistance to multiple antimalarials in Plasmodium falciparum from China-Myanmar border. *Sci Rep* 2016; **6**: 33891.
25. Sidhu AB, Valderramos SG, Fidock DA. pfmdr1 mutations contribute to quinine resistance and enhance mefloquine and artemisinin sensitivity in Plasmodium falciparum. *Mol Microbiol* 2005; **57**(4): 913-26.
26. Reed MB, Saliba KJ, Caruana SR, Kirk K, Cowman AF. Pgh1 modulates sensitivity and resistance to multiple antimalarials in Plasmodium falciparum. *Nature* 2000; **403**(6772): 906-9.
27. Vinayak S, Alam MT, Sem R, et al. Multiple genetic backgrounds of the amplified Plasmodium falciparum multidrug resistance (pfmdr1) gene and selective sweep of 184F mutation in Cambodia. *J Infect Dis* 2010; **201**(10): 1551-60.
28. Dziekan JM, Yu H, Chen D, et al. Identifying purine nucleoside phosphorylase as the target of quinine using cellular thermal shift assay. *Sci Transl Med* 2019; **11**(473).
29. Bustamante C, Folarin OA, Gbotosho GO, et al. In vitro-reduced susceptibility to artemether in P. falciparum and its association with polymorphisms on transporter genes. *J Infect Dis* 2012; **206**(3): 324-32.
30. Cowell AN, Istvan ES, Lukens AK, et al. Mapping the malaria parasite druggable genome by using in vitro evolution and chemogenomics. *Science* 2018; **359**(6372): 191-9.
31. Amato R, Lim P, Miotto O, et al. Genetic markers associated with dihydroartemisinin-piperaquine failure in Plasmodium falciparum malaria in Cambodia: a genotype-phenotype association study. *The Lancet Infectious diseases* 2016.
32. Takahashi N, Tanabe K, Tsukahara T, et al. Large-scale survey for novel genotypes of Plasmodium falciparum chloroquine-resistance gene pfcrt. *Malar J* 2012; **11**: 92.
33. Straimer J, Gnadig NF, Witkowski B, et al. Drug resistance. K13-propeller mutations confer artemisinin resistance in Plasmodium falciparum clinical isolates. *Science* 2015; **347**(6220): 428-31.
34. Sabeti PC, Reich DE, Higgins JM, et al. Detecting recent positive selection in the human genome from haplotype structure. *Nature* 2002; **419**(6909): 832-7.
35. Hetzel MW, Pulford J, Ura Y, et al. Insecticide-treated nets and malaria prevalence, Papua New Guinea, 2008-2014. *Bull World Health Organ* 2017; **95**(10): 695-705B.
36. Sekihara M, Tachibana SI, Yamauchi M, et al. Lack of significant recovery of chloroquine sensitivity in Plasmodium falciparum parasites following discontinuance of chloroquine use in Papua New Guinea. *Malar J* 2018; **17**(1): 434.
37. Sakihama N, Mitamura T, Kaneko A, Horii T, Tanabe K. Long PCR amplification of Plasmodium falciparum DNA extracted from filter paper blots. *Exp Parasitol* 2001; **97**(1): 50-4.
38. Ikeda M, Kaneko M, Tachibana SI, et al. Artemisinin-Resistant Plasmodium falciparum with High Survival Rates, Uganda, 2014-2016. *Emerg Infect Dis* 2018; **24**(4): 718-26.
39. Oyola SO, Ariani CV, Hamilton WL, et al. Whole genome sequencing of Plasmodium falciparum from dried blood spots using selective whole genome amplification. *Malar J* 2016; **15**(1): 597.

40. Pearson RD, Amato R, Kwiatkowski DP, MalariaGEN *Plasmodium falciparum* Community Project. An open dataset of Plasmodium falciparum genome variation in 7,000 worldwide samples. *bioRxiv* 2019; **824730**.
41. Manske M, Miotto O, Campino S, et al. Analysis of Plasmodium falciparum diversity in natural infections by deep sequencing. *Nature* 2012; **487**(7407): 375-9.
42. Balikagala B, Mita T, Ikeda M, et al. Absence of in vivo selection for K13 mutations after artemether-lumefantrine treatment in Uganda. *Malar J* 2017; **16**(1): 23.
43. Witkowski B, Duru V, Khim N, et al. A surrogate marker of piperaquine-resistant Plasmodium falciparum malaria: a phenotype-genotype association study. *Lancet Infect Dis* 2017; **17**(2): 174-83.
44. Schaffner SF, Taylor AR, Wong W, Wirth DF, Neafsey DE. hmmIBD: software to infer pairwise identity by descent between haploid genotypes. *Malar J* 2018; **17**(1): 196.
45. Henden L, Lee S, Mueller I, Barry A, Bahlo M. Identity-by-descent analyses for measuring population dynamics and selection in recombining pathogens. *PLoS Genet* 2018; **14**(5): e1007279.
46. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 2019; **35**(3): 526-8.
47. Raj A, Stephens M, Pritchard JK. fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* 2014; **197**(2): 573-89.
48. Gruning B, Dale R, Sjodin A, et al. Bioconda: sustainable and comprehensive software distribution for the life sciences. *Nat Methods* 2018; **15**(7): 475-6.

## Tables

**Table 1. Characteristics of enrolled patients.**

| Characteristics | 2016 | 2017 |
|---|---|---|
| Patients (n) | 123 | 134 |
| Gender (n)[1] | | |
| Male | 53 | 57 |
| Female | 68 | 76 |
| Age (years) | | |
| Median (IQR)[2] | 20 (13, 33.5) | 18 (11, 24) |
| Pretreatment | | |
| Artemether | 3 (2.5%) | 4 (2.8%) |
| Artemether/Lumefantrine | 0 | 2 (1.3%) |
| Artemether/Lumefantrine+Primaquine | 0 | 1 (0.7%) |
| Chloroquine | 2 (1.7%) | 5 (3.4%) |
| Primaquine | 1 (0.8%) | 0 |
| Parasitemia | | |
| Geometric mean (Range) | 0.11% (0.0007%-9.47%) | 0.28% (0.004%-5.35%) |
| Median (IQR)[2] | 0.14% (0.02%-0.52%) | 0.33% (0.1%-0.88%) |

[1] Unknown in three individuals

[2] IQR: Interquartile range

**Table 2. Samples used in whole-genome comparative analyses.**

| Country | Region | Location | Sample Count |
|---|---|---|---|
| PNG | East Sepik | Wewak | 73 |
| | | Maprik | 39 |
| | Madang | Madang | 20 |
| | Milne Bay | Alotau | 24 |
| Indonesia | Papua | Timika | 88 |
| Cambodia | Western Cambodia | Pailin | 70 |
| | | Pursat | 75 |
| | | Total | 389 |

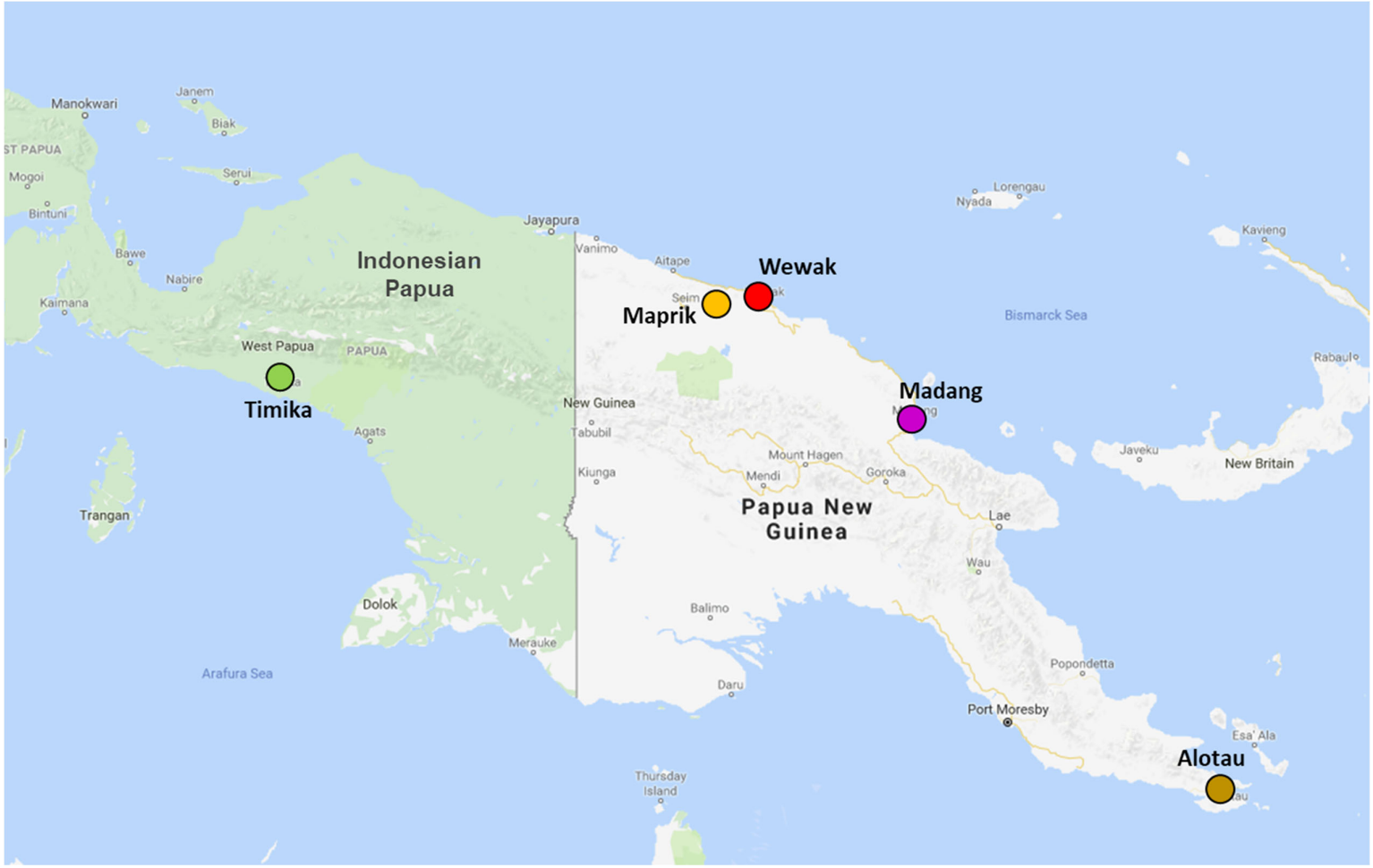
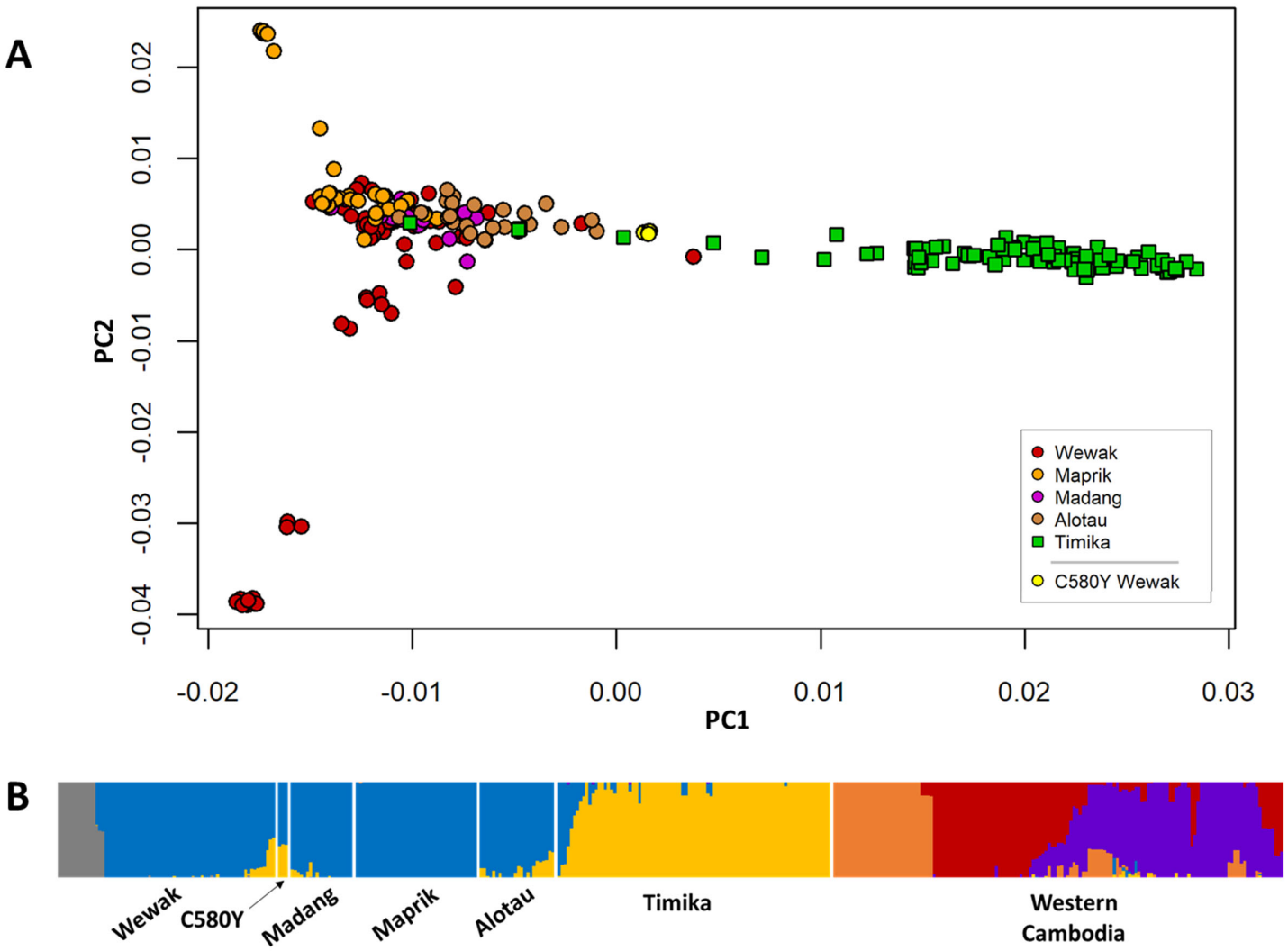



**Figure 1. Geographical location of New Guinea sites.** Sites that contributed samples in both countries that form part of the island of New Guinea (Indonesia to the west and PNG to the east) are shown by coloured circular markers.


A
PC2
0.02
0.01
0.00
-0.01
-0.02
-0.03
-0.04
-0.02
-0.01
0.00
0.01
0.02
0.03
PC1
Wewak
Maprik
Madang
Alotau
Timika
C580Y Wewak
B
Wewak
C580Y
Madang
Maprik
Alotau
Timika
Western
Cambodia

**(Previous page)**

**Figure 2. Population Structure in the New Guinea.** (A) Plot of the first two components (PC2 vs PC1) of principal coordinates analysis (PCoA) using genetic distances computed from whole-genome sequence data from all samples in New Guinea. We note a marked separation between parasites from Papua Indonesia (green squares) and PNG (circles), with a minority of samples from both regions at an intermediate position, suggesting admixture. (B) Population structure at all sites, estimated by fastSTRUCTURE, based on the hypothesis of K=6 ancestral populations; each ancestral populations is indicated by a different colour, and each vertical bar represents a single sample, coloured according to the estimated ancestry proportions. The plot suggests no recent common ancestry between New Guinea parasites and those in Cambodia. Furthermore, it shows strong differentiations between parasites in Papua Indonesia and PNG (consistent with PCoA results above), with a minority of samples from both sides (including the Wewak C580Y mutants) exhibiting significant levels of admixture. This analysis also identified a subpopulation of Wewak parasites (gray) which correspond to the cluster of samples in the lower left-hand corner of the PCoA plot.

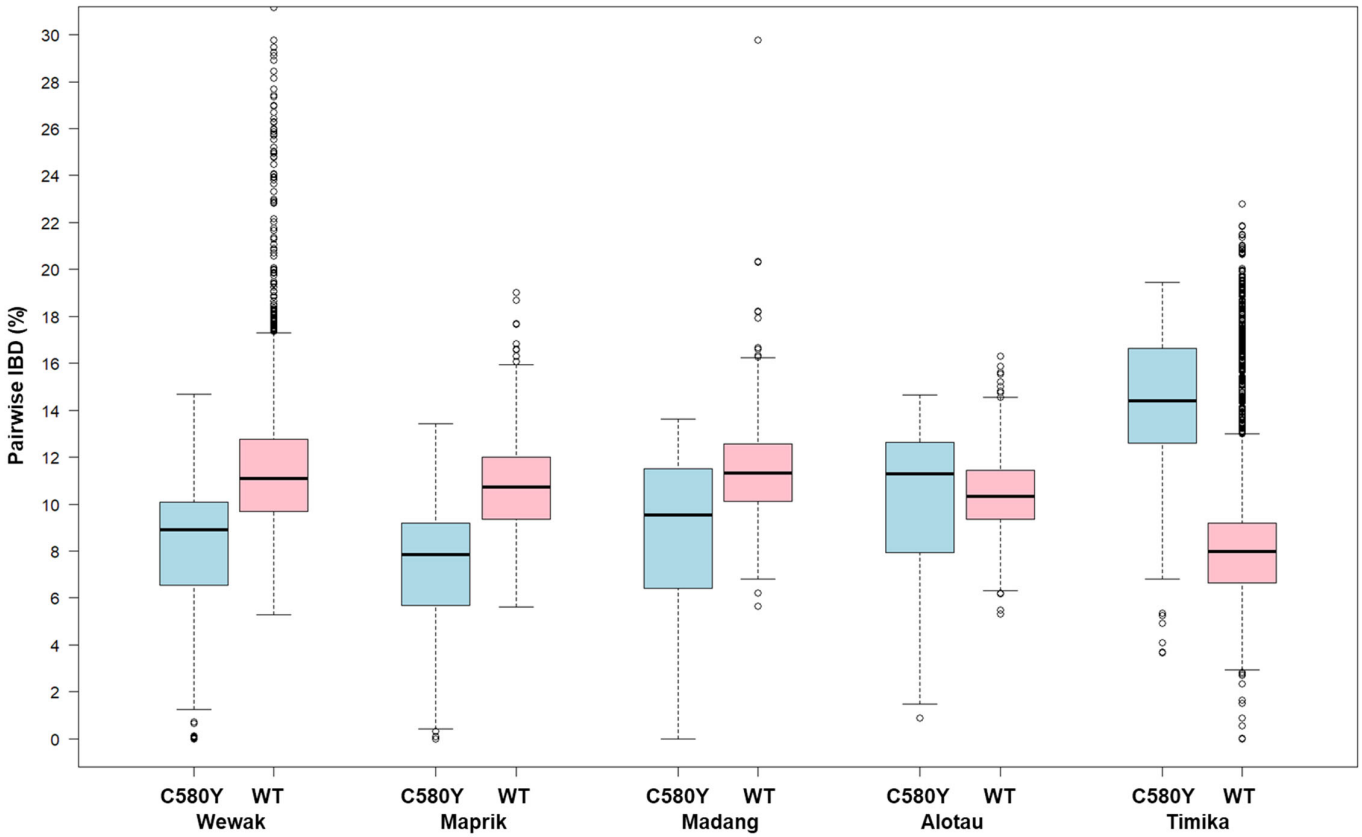
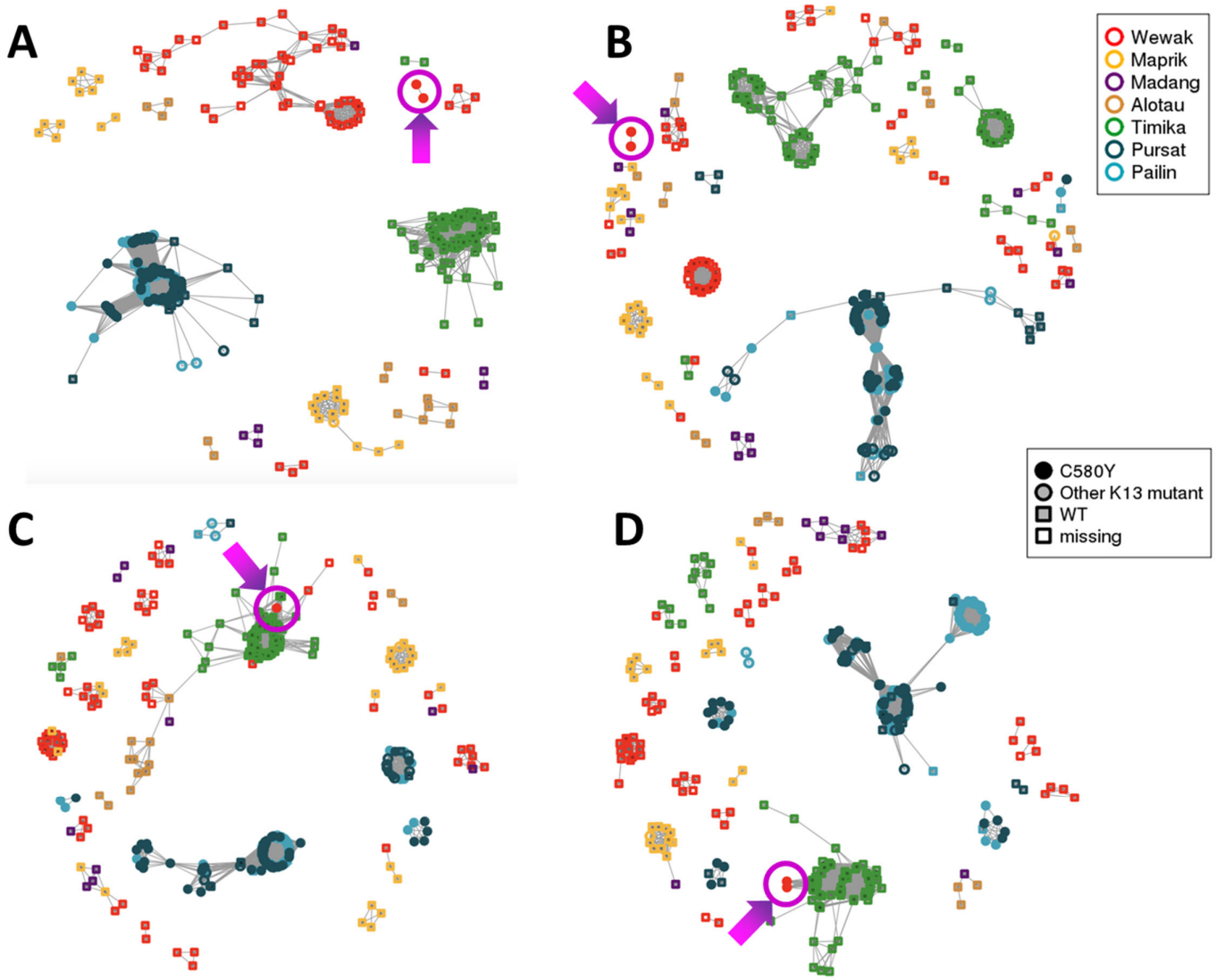



**Figure 3. IBD between parasites from Wewak and those from other New Guinea sites.** This plot shows the distributions of the level of IBD (i.e. the percentage of the genome that is predicted to be in IBD) between parasites samples in Wewak and those sampled at other sites on the island of New Guinea. For each of the five sites, two boxplots show the distribution of pairwise IBD levels for parasites from that site: either against the Wewak C580Y mutants (blue, labelled "C580Y"), or against the Wewak *kelch13* wild-type parasites (pink, "WT"). For the Wewak site (first two boxplots) we compare the two groups against Wewak WT samples only; hence, the blue boxplot shows IBD levels between Wewak C580Y mutants and Wewak WT parasites, while the pink boxplot shows IBD levels among Wewak WT parasites only.

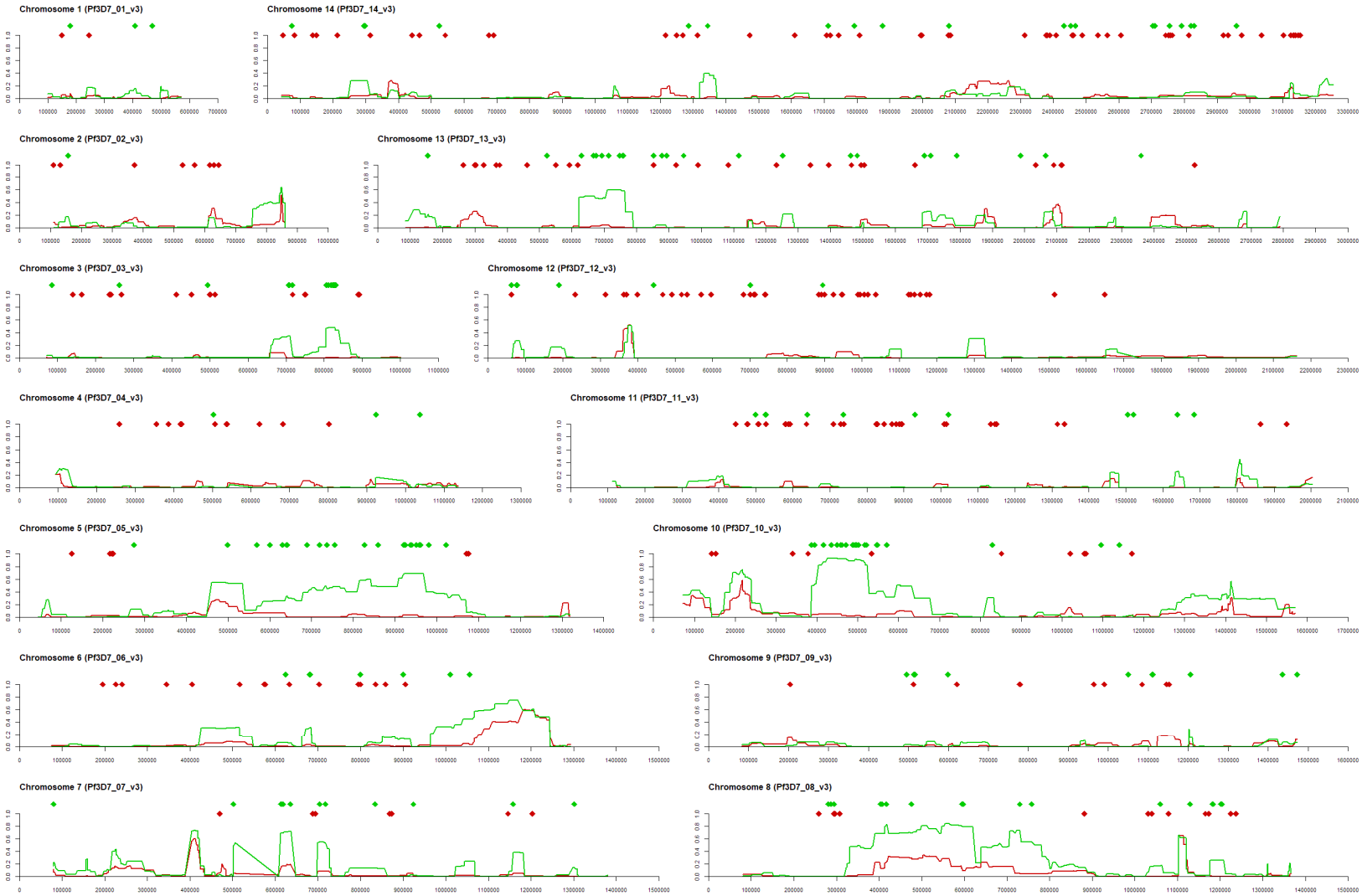

A
B
C
D
Wewak
Maprik
Madang
Alotau
Timika
Pursat
Pailin
C580Y
Other K13 mutant
WT
missing

**(Previous page)**

**Figure 4. IDB-based clustering.** These isoRelate network plots show IBD relationships between all samples analyzed for different genomic regions. The Wewak C580Y mutants are shown by a magenta circle, pointed to by an arrow. (A) Whole-genome IBD shows that, although most parasites cluster strongly by geography, the C580Y mutants form a separate group, probably reflecting their admixed structure. (B) IBD in chromosome 13 regions flanking the *kelch13* gene shows that Wewak parasites do not cluster in a single group, and the C580Y mutants form a group on their own; no New Guinea parasites cluster with the Cambodian group, confirming that *kelch13* haplotypes on the island have independent origin. (C) IBD in chromosome 10 regions flanking the *atg18* gene shows that C580Y mutants cluster with the Timika samples, as is also the case in the chromosome 5 regions flanking the *pnp* gene (D). In B-D, flanking regions are 75 kbp regions on either flank.

**(Next page)**

**Figure 5. Structure of genetic admixture in Wewak C580Y mutants.** Fourteen plots are shown, one per chromosome in the *P. falciparum* nuclear genome; the x-axis of each plot represents nucleotide positions along the chromosome. At the top of each plot, two rows of diamond markers indicate the position of SNPs that are highly differentiated between the Wewak and Timika populations ($F_{ST} \geq 0.3$); the marker colour indicates whether the allele present in the Wewak C580Y mutants matches the most common one in Timika (green diamonds) or Wewak (red diamonds). The line charts below the markers show the proportion of sample pairs in the Timika (green lines) and Wewak (red) population that are IDB with the Wewak C580Y parasites at each chromosome position.


Chromosome 1 (Pf3D7_01_v3)
Chromosome 14 (Pf3D7_14_v3)
Chromosome 2 (Pf3D7_02_v3)
Chromosome 13 (Pf3D7_13_v3)
Chromosome 3 (Pf3D7_03_v3)
Chromosome 12 (Pf3D7_12_v3)
Chromosome 4 (Pf3D7_04_v3)
Chromosome 11 (Pf3D7_11_v3)
Chromosome 5 (Pf3D7_05_v3)
Chromosome 10 (Pf3D7_10_v3)
Chromosome 6 (Pf3D7_06_v3)
Chromosome 9 (Pf3D7_09_v3)
Chromosome 7 (Pf3D7_07_v3)
Chromosome 8 (Pf3D7_08_v3)

# Emergence of artemisinin-resistant *Plasmodium falciparum* with *kelch13* C580Y mutations in Papua New Guinea

Olivo Miotto, Makoto Sekihara, Shin-Ichiro Tachibana, Masato Yamauchi, Richard D Pearson, Roberto Amato, Sonia Gonçalves, Somya Mehra, Rintis Noviyanti, Jutta Marfurt, Sarah Auburn, Ric N Price, Ivo Mueller, Mie Ikeda, Toshiyuki Mori, Makoto Hirai, Livingstone Tavul, Manuel W Hetzel, Moses Laman, Alyssa Barry, Pascal Ringwald, Jun Ohashi, Francis Hombhanje, Dominic P Kwiatkowski, Toshihiro Mita

# APPENDIX

## Contents

Supplementary Notes ..... 2
Description of sampling locations ..... 2
Supplementary Tables ..... 3
Supplementary Table 1 - Genotypes at loci of resistance for ACT component drugs. ..... 3
Supplementary Table 2 - Characteristics of the three Wewak *kelch13* C580Y mutant parasites. ..... 4
Supplementary Table 3 - Ring-stage survival assay results ..... 5
Supplementary Table 4 - Amino acid alleles at loci in the artemisinin resistance genetic background. ..... 5
Supplementary Table 5 - Haplotypes in the *kelch13* flanking region on chromosome 13, based on genotyping 18 microsatellites and SNPs. ..... 6
Supplementary Table 6 – Highly differentiated SNPs where Wewak C580Y mutants and Papua Indonesia parasites carry the same allele. ..... 8
Supplementary Table 7 – European Nucleotide Archive (ENA) identifiers for samples used in the present study. ..... 13
Supplementary Table 8 - Primers used for determination of *kelch13*-flanking SNPs. ..... 19
Supplementary Figures ..... 20
Supplementary Figure 1 - Approximate locations of residence of patients infected with *kelch13* C580Y mutant parasites. ..... 20
Supplementary Figure 2 - RSA survival rates of Wewak parasites. ..... 21
Supplementary Figure 3 – Principal coordinates analysis (PCoA) based on genome-wide pairwise genetic distances. ..... 22
Supplementary Figure 4 – Population structure in the analyzed sample set. ..... 23
Supplementary Figure 5 - Comparison of haplotypes in a ±150 kbp region on chromosome 13, flanking the kelch13 gene. ..... 24
References for Supplementary Material ..... 28

## Supplementary Notes

### Description of sampling locations

Wewak Town comprises about 25,000 inhabitants. The average temperature in the studied area is 27.3°C (min 23.8°C, max 30.9°C) with an annual rainfall of approximately 3000 mm. All four species of human malaria parasites were observed with limited seasonal variations between the wet (October to April) and dry (May to September) seasons, and were transmitted mainly by *Anopheles farauti, Anopheles punctulatus*, and *Anopheles koliensis* (Barry et al., 2013; Muller, Bockarie, Alpers, & Smith, 2003; Schultz et al., 2010).

Free distribution of long-lasting insecticidal mosquito nets (LLIN) was implemented by The Government of PNG between 2005 and 2009 and between 2009 and 2013 (M. W. Hetzel et al., 2014). Average LLIN usage was 55% in 2008 and 2009 and 32.9–67.7% in 2013–2014 (Manuel W Hetzel et al., 2012; M. W. Hetzel et al., 2017). *P. falciparum* entomological inoculation rates in Dreikikir, about 50 km from our study area, dropped from 159 infective bites/person/year in 2008 to 53 in 2011, following LLIN distribution.(Reimer et al., 2016) Current first-line regime is Artemether plus Lumefantrine, officially introduced in 2010. Free distribution of ACT and rapid diagnostic tests (RDT) through all public health facilities was initiated by the malaria control programme in late 2011 (M. W. Hetzel et al., 2014). Previously, intramuscularly artemether was sometimes used in patients who failed first-line treatment regime.

## Supplementary Tables

**Supplementary Table 1 - Genotypes at loci of resistance for ACT component drugs.**

| | 2016 | 2017 |
|---|---|---|
| ***kelch13*** | | |
| C580Y | 0 | 3 |
| Wild-type | 106 | 130 |
| ***plasmepsin 2*** | | |
| Single copy | ND | 134 |
| Multiple copies | ND | 0 |
| ***pfmdr1*** | | |
| Single copy | ND | 128 |
| Multiple copies | ND | 1 |

**Supplementary Table 2 - Characteristics of the three Wewak *kelch13* C580Y mutant parasites.**

| Sample | Collection Date | Age (Year) | Gender | Sign and symptom | Parasitemia | RSA survival rate | *plasmepsin 2/3* | *pfmdr1* |
|---|---|---|---|---|---|---|---|---|
| PNG-C580Y-1 | 25 Jan 2017 | 7 | Female | Fever, Headache, Muscle and joint pain | 0.35% | 6.8% | Single copy | Single copy |
| PNG-C580Y-2 | 23 Jan 2017 | 16 | Male | Fever, Headache, Muscle and joint pain, Nausea, Cough | 0.07% | ND* | Single copy | Single copy |
| PNG-C580Y-3 | 2 Feb 2017 | 25 | Male | Fever, Headache, Muscle and joint pain | 0.44% | >0%** | Single copy | Single copy |

* RSA was not performed because of low parasitemia (<0.1%)

** Viable parasite was detected at the drug-exposure line. Survival rate was not determined because of poor quality of blood smears at non-exposure control.

**Supplementary Table 3 - Ring-stage survival assay results**

| Outcome | 2016 | 2017 |
|---|---|---|
| Result obtained | 27 | 30 |
| Survival rate > 0% | 3* | 4** |

* Survival rates: 2.6%, 3.0% and 3.5%.
** Survival rates: 0.4%, 0.9 %, 1.6% and 6.8%.

**Supplementary Table 4 - Amino acid alleles at loci in the artemisinin resistance genetic background.**

The table includes positions previously associated with kelch13 mutations in the GMS (see main text). Samples prefixed with "MRA" originated from the GMS, while those prefixed "PNG" were sampled in Wewak. Coloured background denoted alleles different from the reference (wild type). Light blue background reflects a mutant allele different from that previously reported in the GMS.

| Sample ID | *kelch13* | *mdr2* T484I | *pibp* C1484F | *nif4* V1157L | *fd* D193Y | *pfcrt* I356T | *arps10* V127M |
|---|---|---|---|---|---|---|---|
| MRA-1236 | **C580Y** | **I** | C | **L** | **Y** | **T** | **M** |
| MRA-1238 | **Y493H** | **I** | C | V | **Y** | **T** | **M** |
| MRA-1240 | **R539T** | **I** | C | **L** | **Y** | **T** | **M** |
| MRA-1241 | **I543T** | **I** | C | **L** | **Y** | **T** | **M** |
| PNG-C580Y-1 | **C580Y** | T | C | V | **Y** | **L** | V |
| PNG-C580Y-2 | **C580Y** | T | C | V | **Y** | **L** | V |
| PNG-C580Y-3 | **C580Y** | T | C | V | **Y** | **L** | V |
| PNG-Wild-1 | WT | T | C | V | D | **L** | V |
| PNG-Wild-2 | WT | T | C | V | D | **L** | V |
| PNG-Wild-3 | WT | T | C | V | D | **L** | V |
| PNG-Wild-4 | WT | T | C | V | D | **L** | V |
| PNG-Wild-5 | WT | T | C | V | **Y** | **L** | V |
| PNG-Wild-6 | WT | T | C | V | D | **L** | V |
| PNG-Wild-7 | WT | T | C | V | D | **L** | V |
| PNG-Wild-8 | WT | T | C | V | D | **L** | V |
| PNG-Wild-9 | WT | T | C | V | D | I | V |
| PNG-Wild-10 | WT | T | C | V | D | **L** | V |

**Supplementary Table 5 - Haplotypes in the *kelch13* flanking region on chromosome 13, based on genotyping 18 microsatellites and SNPs.**

The location of marker sites are described in terms of their distance from the *kelch13* gene (negative and positive values denote upstream and downstream distances respectively). Colour backgrounds denote an uninterrupted sequence of alleles identical to one of the Wewak C580Y haplotypes, in either direction as one traverses from the *kelch13* gene. Reference isolates prefixed with "CAM" and "MRA" originated from the GMS. The Wewak wild type haplotypes are sorted by the total length of the identical genotype sequences.

| Haplotype ID | *kelch13* allele | Sample Count | Locations of marker sites (kb) | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | -137.5 | -33.9 | -14.06 | -8.1 | | -5.02 | | | | -1.91 | | *kelch13* | 4.1 | 6.97 | 9.33 | | 15.82 | 24.91 | 43.85 |
| | | | G/C | TA | A/G | TAT | T/A | T/A | TAAAA | TAAA | TA | TA | TAA | | TA | A/T | TA | TAAA | TTA | T/C | TA |
| Reference isolates | | | | | | | | | | | | | | | | | | | | | |
| 3D7 | WT | | G | 15 | A | 13 | T | T | 2 | 3 | 15 | 25 | 3 | | 34 | A | 14 | 4 | 12 | T | 18 |
| CAM-580Y-1 | C580Y | | G | 15 | - | 13 | T | T | 2 | 3 | 9 | 7 | 5 | | 20 | A | 18 | 5 | 9 | C | 15 |
| CAM-580Y-2 | C580Y | | - | - | - | - | - | - | - | - | - | 7 | 5 | | 20 | - | 22 | 4 | 9 | - | 13 |
| CAM-580Y-3 | C580Y+S623C | | C | - | - | - | - | T | 2 | 3 | 9 | 7 | 5 | | 20 | T | - | - | 9 | C | 21 |
| MRA-1236 | C580Y | | G | 12 | G | 13 | T | T | 2 | 3 | 9 | 7 | 5 | | 20 | A | 18 | 5 | 9 | C | 16 |
| MRA-1238 | Y493H | | C | 18 | A | 10 | T | T | 2 | 3 | 9 | 9 | 5 | | 16 | A | 18 | 5 | 9 | C | 13 |
| MRA-1240 | R539T | | C | 12 | G | 13 | T | T | 2 | 3 | 9 | 7 | 5 | | 16 | A | 9 | 4 | 9 | T | 15 |
| MRA-1241 | I543T | | G | 12 | G | 13 | T | T | 2 | 3 | 9 | 7 | 5 | | 20 | T | 22 | 4 | 9 | C | 21 |
| Wewak C580Y mutants | | | | | | | | | | | | | | | | | | | | | |
| H1 | C580Y | 2 | G | 12 | G | 13 | T | T | 2 | 3 | 9 | 6 | 5 | | 20 | A | 16 | 4 | 9 | T | 13 |
| H2 | C580Y | 1 | G | 12 | G | 13 | T | T | 2 | 3 | 9 | 7 | 5 | | 20 | A | 16 | 4 | 9 | T | 13 |
| Wewak wild-type parasites | | | | | | | | | | | | | | | | | | | | | |
| H5 | WT | 5 | G | 12 | G | 13 | T | T | 2 | 3 | 9 | 7 | 5 | | 20 | A | 16 | 4 | 9 | T | 16 |
| H13 | WT | 2 | G | 12 | G | 13 | T | T | 2 | 3 | 9 | 7 | 5 | | 20 | A | 16 | 4 | 9 | T | 14 |
| H46 | WT | 1 | G | 15 | A | 13 | T | T | 2 | 3 | 9 | 7 | 5 | | 20 | A | 16 | 4 | 9 | T | 13 |
| H48 | WT | 1 | G | 18 | A | 13 | T | T | 2 | 3 | 9 | 7 | 5 | | 20 | A | 16 | 4 | 9 | T | 13 |
| H61 | WT | 1 | G | 12 | G | 13 | T | T | 2 | 3 | 9 | 7 | 5 | | 19 | A | 16 | 4 | 9 | T | 14 |
| H58 | WT | 1 | C | 12 | G | 13 | T | T | 2 | 3 | 9 | 7 | 5 | | 19 | A | 16 | 4 | 9 | T | 16 |
| H54 | WT | 1 | G | 15 | A | 12 | T | T | 2 | 3 | 9 | 7 | 5 | | 17 | A | 21 | 4 | 9 | T | 11 |
| H36 | WT | 1 | G | 15 | A | 12 | T | T | 2 | 3 | 9 | 7 | 5 | | 12 | A | 19 | 4 | 11 | T | 10 |
| H51 | WT | 1 | G | 16 | A | 12 | T | T | 2 | 3 | 9 | 7 | 5 | | 14 | A | 20 | 4 | 9 | T | 10 |
| H19 | WT | 1 | G | 16 | A | 12 | T | T | 2 | 3 | 9 | 7 | 5 | | 17 | A | 21 | 4 | 9 | T | 11 |
| H55 | WT | 1 | G | 15 | A | 11 | T | T | 2 | 3 | 10 | 7 | 5 | | 17 | A | 11 | 4 | 10 | T | 13 |
| H56 | WT | 1 | G | 15 | A | 11 | T | T | 2 | 3 | 10 | 7 | 5 | | 17 | A | 11 | 4 | 9 | T | 13 |
| H40 | WT | 1 | G | 17 | A | 13 | T | T | 2 | 4 | 16 | 7 | 5 | | 13 | A | 20 | 5 | 9 | T | 10 |
| H50 | WT | 1 | G | 15 | A | 10 | T | T | 2 | 3 | 10 | 7 | 5 | | 17 | A | 11 | 4 | 9 | T | 13 |
| H63 | WT | 1 | G | 13 | A | 13 | T | T | 2 | 4 | 10 | 7 | 5 | | 12 | A | 7 | 4 | 9 | C | 15 |
| H12 | WT | 2 | G | 15 | A | 13 | T | T | 2 | 3 | 10 | 7 | 5 | | 17 | A | 11 | 4 | 9 | T | 13 |
| H18 | WT | 1 | C | 15 | A | 13 | T | T | 2 | 3 | 13 | 7 | 5 | | 12 | A | 19 | 4 | 9 | T | 14 |
| H22 | WT | 1 | G | 15 | A | 13 | T | T | 2 | 3 | 15 | 7 | 5 | | 15 | A | 12 | 4 | 12 | T | 13 |
| H29 | WT | 1 | C | 15 | A | 11 | T | T | 2 | 3 | 13 | 7 | 5 | | 12 | A | 19 | 4 | 9 | T | 14 |
| H30 | WT | 1 | G | 20 | A | 12 | T | T | 2 | 3 | 13 | 7 | 5 | | 15 | A | 21 | 3 | 9 | T | 12 |
| H32 | WT | 1 | G | 14 | A | 13 | T | T | 2 | 3 | 10 | 7 | 5 | | 15 | A | 12 | 4 | 9 | T | 13 |
| H59 | WT | 1 | G | 20 | A | 10 | T | T | 2 | 3 | 9 | 7 | 6 | | 20 | A | 13 | 4 | 11 | T | 16 |
| H42 | WT | 1 | G | 12 | G | 13 | T | T | 2 | 3 | 10 | 15 | 5 | | 17 | A | 14 | 5 | 9 | C | 10 |
| H53 | WT | 1 | G | 15 | A | 12 | T | T | 2 | 3 | 14 | 9 | 5 | | 15 | A | 20 | 3 | 9 | T | 21 |
| H3 | WT | 9 | G | 15 | A | 10 | A | T | 3 | 3 | 9 | 9 | 4 | | 12 | A | 15 | 5 | 9 | T | 19 |
| H4 | WT | 8 | C | 15 | A | 10 | A | T | 3 | 3 | 9 | 9 | 4 | | 12 | A | 15 | 5 | 9 | T | 19 |
| H6 | WT | 3 | C | 15 | A | 10 | A | T | 3 | 3 | 9 | 9 | 4 | | 12 | A | 15 | 5 | 9 | T | 20 |

| | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| H7 | WT | 3 | C | 15 | A | 10 | A | T | 3 | 3 | 9 | 9 | 4 | 12 | A | 15 | 5 | 10 | T | 19 |
| H8 | WT | 2 | G | 15 | A | 10 | A | T | 3 | 3 | 9 | 9 | 4 | 12 | A | 15 | 5 | 9 | T | 20 |
| H9 | WT | 2 | G | 15 | A | 9 | T | A | 1 | 3 | 9 | 9 | 6 | 14 | A | 21 | 5 | 9 | T | 14 |
| H10 | WT | 2 | G | 11 | A | 9 | T | A | 1 | 3 | 9 | 9 | 6 | 14 | A | 20 | 5 | 9 | T | 13 |
| H11 | WT | 2 | G | 21 | A | 10 | A | T | 3 | 3 | 9 | 9 | 4 | 12 | A | 14 | 4 | 10 | T | 14 |
| H14 | WT | 2 | G | 15 | A | 10 | T | T | 2 | 3 | 9 | 7 | 6 | 18 | A | 18 | 4 | 10 | T | 14 |
| H15 | WT | 2 | G | 21 | A | 10 | A | T | 3 | 3 | 9 | 9 | 4 | 12 | A | 14 | 4 | 9 | T | 14 |
| H16 | WT | 1 | G | 15 | A | 10 | T | T | 2 | 3 | 14 | 9 | 6 | 14 | A | 17 | 4 | 9 | T | 14 |
| H17 | WT | 1 | G | 15 | A | 10 | T | T | 2 | 3 | 10 | 7 | 6 | 19 | A | 13 | 4 | 9 | T | 14 |
| H20 | WT | 1 | G | 15 | A | 10 | A | T | 3 | 3 | 9 | 9 | 4 | 12 | A | 15 | 5 | 10 | T | 20 |
| H21 | WT | 1 | G | 12 | A | 10 | T | A | 1 | 3 | 9 | 7 | 6 | 17 | A | 17 | 5 | 9 | T | 19 |
| H23 | WT | 1 | G | 15 | A | 10 | A | T | 3 | 3 | 0 | 9 | 4 | 12 | A | 15 | 5 | 9 | T | 19 |
| H24 | WT | 1 | C | 15 | A | 12 | T | T | 2 | 3 | 14 | 7 | 6 | 15 | A | 18 | 4 | 9 | T | 14 |
| H25 | WT | 1 | G | 14 | A | 13 | T | T | 2 | 3 | 10 | 7 | 6 | 12 | A | 20 | 3 | 9 | T | 13 |
| H26 | WT | 1 | G | 16 | A | 13 | T | T | 2 | 3 | 10 | 10 | 4 | 14 | A | 15 | 4 | 10 | C | 19 |
| H27 | WT | 1 | G | 12 | A | 10 | T | T | 2 | 3 | 15 | 7 | 6 | 17 | A | 11 | 4 | 9 | T | 13 |
| H28 | WT | 1 | G | 14 | A | 13 | T | T | 2 | 3 | 10 | 7 | 6 | 12 | A | 18 | 4 | 9 | T | 14 |
| H31 | WT | 1 | G | 15 | A | 10 | T | A | 1 | 3 | 9 | 7 | 6 | 12 | A | 9 | 4 | 9 | T | 10 |
| H33 | WT | 1 | C | 17 | A | 9 | T | A | 1 | 3 | 9 | 9 | 6 | 14 | A | 21 | 5 | 9 | T | 16 |
| H34 | WT | 1 | G | 12 | A | 10 | T | A | 1 | 3 | 9 | 7 | 6 | 17 | A | 11 | 4 | 9 | T | 13 |
| H35 | WT | 1 | G | 11 | A | 13 | T | T | 2 | 3 | 10 | 10 | 4 | 14 | A | 11 | 4 | 9 | T | 14 |
| H37 | WT | 1 | C | 11 | A | 10 | T | T | 2 | 3 | 9 | 9 | 4 | 18 | A | 26 | 4 | 10 | T | 12 |
| H38 | WT | 1 | G | 18 | A | 10 | T | T | 2 | 3 | 9 | 7 | 6 | 18 | A | 18 | 4 | 9 | T | 14 |
| H39 | WT | 1 | G | 12 | A | 13 | T | A | 1 | 3 | 9 | 7 | 6 | 17 | A | 11 | 4 | 9 | T | 13 |
| H41 | WT | 1 | G | 11 | A | 12 | T | T | 2 | 3 | 9 | 9 | 6 | 21 | A | 16 | 4 | 9 | T | 16 |
| H43 | WT | 1 | G | 15 | A | 9 | T | A | 1 | 3 | 9 | 9 | 6 | 14 | A | 20 | 5 | 9 | T | 13 |
| H44 | WT | 1 | C | 15 | A | 10 | A | T | 3 | 3 | 9 | 9 | 4 | 12 | A | 15 | 5 | 7 | T | 19 |
| H45 | WT | 1 | G | 12 | A | 10 | T | T | 2 | 3 | 14 | 7 | 6 | 15 | A | 9 | 4 | 7 | T | 12 |
| H47 | WT | 1 | C | 15 | A | 10 | A | T | 3 | 3 | 9 | 9 | 4 | 12 | A | 15 | 5 | 9 | T | 18 |
| H49 | WT | 1 | G | 16 | A | 11 | T | T | 2 | 3 | 9 | 9 | 6 | 12 | A | 14 | 4 | 9 | T | 16 |
| H52 | WT | 1 | G | 15 | A | 10 | A | T | 3 | 3 | 9 | 9 | 4 | 12 | A | 15 | 5 | 10 | T | 19 |
| H57 | WT | 1 | G | 18 | A | 10 | T | A | 1 | 3 | 9 | 7 | 6 | 14 | A | 9 | 4 | 9 | T | 10 |
| H60 | WT | 1 | G | 11 | A | 13 | T | T | 2 | 3 | 10 | 10 | 4 | 14 | A | 11 | 4 | 9 | T | 13 |
| H62 | WT | 1 | G | 15 | A | 10 | A | T | 3 | 3 | 9 | 9 | 4 | 12 | A | 17 | 4 | 9 | T | 19 |
| H64 | WT | 1 | G | 12 | A | 10 | T | A | 1 | 3 | 9 | 9 | 6 | 14 | A | 11 | 4 | 7 | T | 15 |

**Supplementary Table 6 – Highly differentiated SNPs where Wewak C580Y mutants and Papua Indonesia parasites carry the same allele.** Each row shows a non-synonymous SNP that is highly differentiated between the Wewak (PNG) and Timika (Indonesia) populations ($F_{ST} \geq 0.3$) at which the Wewak C580Y mutants carry the allele that is most prevalent in Timika. The columns show: the chromosome number and nucleotide position of the SNP; the name and ID of the gene containing the SNP (genes associated with drug resistance are shown on a coloured background); the mutation name, showing amino acid change and position; the frequency of the mutant allele, in Wewak and in Timika (background colour intensity is proportional to the value); the $F_{ST}$ between the two populations; the proportion of sample pairs that are in IBD with Wewak C580 samples at that position, in Wewak and Timika (background colour intensity is proportional to the value); and the Pubmed PMID reference of literature that implicates the gene in drug resistance, if applicable.

| Chr | Pos | Gene | | Mutation | Mutation Frequency | | $F_{ST}$ | IBD with C580Y | | Pubmed Ref |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Description | Id | | Wewak | ID | | Wewak | Timika | |
| 13 | 748395 | ferredoxin, putative | PF3D7_1318100 | D193Y | 0.01 | 0.92 | 0.82 | 1.0% | 59.8% | 25599401 |
| 13 | 757313 | conserved Plasmodium protein, unknown function | PF3D7_1318300 | S1135T | 0.97 | 0.07 | 0.82 | 1.0% | 57.6% | |
| 10 | 497461 | autophagy-related protein 18 | PF3D7_1012900 | T38I | 0.03 | 0.92 | 0.79 | 2.4% | 90.9% | 27694982 |
| 10 | 489227 | NLI interacting factor-like phosphatase, putative | PF3D7_1012700 | N659S | 0.01 | 0.89 | 0.77 | 1.4% | 88.6% | 25599401 |
| 10 | 461962 | heme oxygenase | PF3D7_1011900 | T11I | 0.06 | 0.92 | 0.74 | 2.9% | 92.0% | |
| 10 | 522059 | conserved Plasmodium protein, unknown function | PF3D7_1013400 | V65A | 0.00 | 0.85 | 0.74 | 0.0% | 81.8% | |
| 5 | 961013 | multidrug resistance protein (MDR1) | PF3D7_0523000 | N1042D | 0.05 | 0.91 | 0.74 | 1.0% | 68.6% | 10706290 |
| 11 | 528463 | Rpr2, RNAse P, putative | PF3D7_1113600 | I298V | 0.93 | 0.07 | 0.74 | 0.0% | 0.0% | |
| 10 | 488090 | NLI interacting factor-like phosphatase, putative | PF3D7_1012700 | N280S | 0.01 | 0.86 | 0.73 | 1.4% | 88.6% | 25599401 |
| 13 | 758176 | conserved Plasmodium protein, unknown function | PF3D7_1318300 | D1423Y | 0.01 | 0.86 | 0.73 | 1.0% | 57.6% | |
| 10 | 517067 | conserved Plasmodium protein, unknown function | PF3D7_1013300 | S181C | 0.03 | 0.88 | 0.72 | 1.0% | 86.4% | |
| 11 | 528525 | Rpr2, RNAse P, putative | PF3D7_1113600 | N318K | 0.92 | 0.07 | 0.72 | 0.0% | 0.0% | |
| 2 | 158568 | conserved Plasmodium protein, unknown function | PF3D7_0203200 | D365N | 0.03 | 0.86 | 0.71 | 1.0% | 17.4% | |
| 11 | 528101 | Rpr2, RNAse P, putative | PF3D7_1113600 | I177S | 0.90 | 0.07 | 0.70 | 0.0% | 0.0% | |
| 13 | 760334 | chromosome segregation protein, putative | PF3D7_1318400 | H1217N | 0.03 | 0.86 | 0.69 | 1.0% | 57.6% | |
| 10 | 471727 | rhoptry associated adhesin | PF3D7_1012200 | A200S | 0.09 | 0.91 | 0.69 | 2.9% | 92.0% | |
| 5 | 641983 | amino acid transporter, putative | PF3D7_0515500 | K753Q | 0.12 | 0.94 | 0.68 | 1.0% | 33.3% | 22615315 |
| 5 | 274721 | conserved Plasmodium protein, unknown function | PF3D7_0506500 | E2773D | 0.10 | 0.92 | 0.67 | 1.0% | 12.9% | |
| 5 | 958440 | multidrug resistance protein (MDR1) | PF3D7_0523000 | Y184F | 0.09 | 0.91 | 0.67 | 1.0% | 68.6% | 10706290 |
| 14 | 2081506 | conserved Plasmodium protein, unknown function | PF3D7_1450800 | G248R | 0.10 | 0.89 | 0.63 | 2.4% | 11.7% | 25599401 |
| 4 | 502909 | conserved Plasmodium protein, unknown function | PF3D7_0411000 | E1248K | 0.83 | 0.05 | 0.62 | 0.0% | 0.8% | |
| 13 | 1709827 | CDT1-like protein, putative | PF3D7_1343300 | A285S | 0.07 | 0.85 | 0.61 | 3.8% | 25.4% | |
| 10 | 547025 | initiation factor 2+subunit family, putative | PF3D7_1013900 | D627N | 0.21 | 0.98 | 0.60 | 4.8% | 40.5% | |

| Chr | Pos | Gene | | Mutation | Mutation Frequency | | $F_{ST}$ | IBD with C580Y | | Pubmed Ref |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Description | Id | | Wewak | ID | | Wewak | Timika | |
| 5 | 719719 | conserved Plasmodium protein, unknown function | PF3D7_0517100 | A46V | 0.01 | 0.76 | 0.59 | 0.0% | 42.0% | |
| 14 | 2453031 | biotin protein ligase, putative | PF3D7_1460000 | E188K | 0.77 | 0.02 | 0.58 | 0.0% | 0.0% | |
| 5 | 950917 | zinc finger protein, putative | PF3D7_0522900 | S656Y | 0.07 | 0.83 | 0.58 | 3.3% | 68.6% | |
| 14 | 2432735 | OPA3-like protein, putative | PF3D7_1459300 | K187. | 0.07 | 0.82 | 0.57 | 0.0% | 0.0% | |
| 11 | 1639917 | conserved Plasmodium protein, unknown function | PF3D7_1140900 | P1007R | 0.74 | 0.01 | 0.56 | 0.0% | 25.8% | |
| 10 | 395382 | conserved P. membrane protein, unknown function | PF3D7_1009800 | G871A | 0.09 | 0.83 | 0.56 | 3.8% | 51.9% | |
| 13 | 629898 | ubiquitin-like protein, putative | PF3D7_1314800 | M134I | 0.11 | 0.85 | 0.55 | 1.0% | 48.1% | |
| 13 | 2360453 | rRNA associated RNA binding protein, putative | PF3D7_1359400 | K249R | 0.71 | 0.01 | 0.53 | 0.0% | 0.4% | |
| 7 | 718071 | conserved Plasmodium protein, unknown function | PF3D7_0716300 | G264D | 0.27 | 0.98 | 0.53 | 8.1% | 53.0% | |
| 3 | 806009 | endonuclease/exonuclease/phosphatase family protein | PF3D7_0319200 | T102A | 0.01 | 0.72 | 0.53 | 1.0% | 47.0% | |
| 13 | 666567 | exportin-T, putative | PF3D7_1315900 | T943A | 0.10 | 0.82 | 0.53 | 2.4% | 49.6% | |
| 3 | 827556 | ABC transporter, putative | PF3D7_0319700 | S2106C | 0.01 | 0.70 | 0.52 | 1.0% | 47.7% | 29326268 |
| 3 | 827918 | ABC transporter, putative | PF3D7_0319700 | S2227G | 0.02 | 0.70 | 0.52 | 1.0% | 47.7% | 29326268 |
| 11 | 1505488 | guanylyl cyclase (GCalpha) | PF3D7_1138400 | T2735N | 0.73 | 0.03 | 0.52 | 0.0% | 0.0% | |
| 9 | 1436778 | ring-exported protein 3+(REX3) | PF3D7_0936300 | A297V | 0.71 | 0.02 | 0.51 | 1.0% | 7.6% | |
| 5 | 920718 | conserved Plasmodium protein, unknown function | PF3D7_0522400 | Y4166F | 0.21 | 0.92 | 0.51 | 3.3% | 68.6% | |
| 5 | 827126 | conserved Plasmodium protein, unknown function | PF3D7_0519900 | N96S | 0.12 | 0.83 | 0.51 | 3.8% | 58.3% | |
| 10 | 458574 | PRE-binding protein | PF3D7_1011800 | N1059K | 0.23 | 0.93 | 0.51 | 4.8% | 92.0% | |
| 13 | 1462712 | conserved Plasmodium protein, unknown function | PF3D7_1335800 | Q178P | 0.80 | 0.09 | 0.51 | 1.0% | 0.0% | |
| 13 | 853331 | conserved Plasmodium protein, unknown function | PF3D7_1320700 | P1224T | 0.88 | 0.17 | 0.50 | 0.0% | 0.8% | |
| 1 | 409173 | phosphatidylinositol-4-phosphate 5-kinase (PIP5K) | PF3D7_0110600 | M1145I | 0.14 | 0.84 | 0.50 | 0.0% | 15.2% | 29326268 |
| 5 | 640092 | conserved Plasmodium protein, unknown function | PF3D7_0515400 | S666N | 0.71 | 0.03 | 0.49 | 1.0% | 33.3% | |
| 10 | 471382 | rhoptry associated adhesin | PF3D7_1012200 | K135I | 0.23 | 0.92 | 0.49 | 2.9% | 92.0% | |
| 11 | 1684812 | conserved Plasmodium protein, unknown function | PF3D7_1142100 | H1758R | 0.71 | 0.03 | 0.49 | 0.0% | 0.0% | |
| 11 | 639134 | heat shock protein 101 (HSP101) | PF3D7_1116800 | H57Y | 0.80 | 0.11 | 0.48 | 1.9% | 0.0% | |
| 5 | 600642 | aspartyl-tRNA synthetase, putative | PF3D7_0514300 | R199P | 0.28 | 0.94 | 0.47 | 6.7% | 24.6% | |
| 14 | 298164 | plasmepsin III,histo-aspartic protease (HAP) | PF3D7_1408100 | G233R | 0.17 | 0.85 | 0.47 | 3.8% | 28.0% | 27818095 |
| 3 | 823223 | ABC transporter, putative | PF3D7_0319700 | F662I | 0.91 | 0.24 | 0.47 | 1.0% | 47.7% | 29326268 |
| 14 | 1287030 | conserved Plasmodium protein, unknown function | PF3D7_1432600 | N742S | 0.08 | 0.75 | 0.47 | 0.0% | 0.0% | |
| 8 | 287476 | conserved Plasmodium protein, unknown function | PF3D7_0805100 | L173P | 0.75 | 0.08 | 0.46 | 0.0% | 0.0% | |
| 5 | 756356 | RAP protein, putative | PF3D7_0518100 | N1020S | 0.00 | 0.63 | 0.46 | 0.0% | 48.5% | |
| 13 | 878895 | conserved P. membrane protein, unknown function | PF3D7_1321300 | S2457N | 0.11 | 0.78 | 0.45 | 0.0% | 3.8% | |
| 13 | 1252488 | conserved Plasmodium protein, unknown function | PF3D7_1329500 | D144N | 0.03 | 0.67 | 0.45 | 0.0% | 21.2% | |
| 12 | 895415 | transcription factor with AP2 domain(s) (ApiAP2) | PF3D7_1222400 | E1612V | 0.67 | 0.03 | 0.44 | 0.0% | 0.4% | |

| Chr | Pos | Gene | | Mutation | Mutation Frequency | | $F_{ST}$ | IBD with C580Y | | Pubmed Ref |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Description | Id | | Wewak | ID | | Wewak | Timika | |
| 10 | 1140750 | peroxiredoxin (nPrx) | PF3D7_1027300 | N341D | 0.14 | 0.81 | 0.44 | 0.0% | 0.0% | |
| 3 | 818649 | elongation factor 1+(EF-1), putative | PF3D7_0319600 | G142E | 0.29 | 0.93 | 0.43 | 1.0% | 47.7% | |
| 14 | 2791430 | kelch domain-containing protein, putative | PF3D7_1468100 | Y852C | 0.87 | 0.22 | 0.43 | 4.3% | 3.4% | |
| 7 | 834765 | conserved Plasmodium protein, unknown function | PF3D7_0718800 | L50F | 0.21 | 0.86 | 0.43 | 1.4% | 13.6% | |
| 10 | 494817 | conserved Plasmodium protein, unknown function | PF3D7_1012800 | K756N | 0.30 | 0.93 | 0.42 | 2.4% | 90.9% | |
| 11 | 499091 | conserved Plasmodium protein, unknown function | PF3D7_1113000 | R2041M | 0.71 | 0.08 | 0.42 | 0.0% | 0.0% | |
| 8 | 475630 | conserved Plasmodium protein, unknown function | PF3D7_0809400 | L20I | 0.28 | 0.91 | 0.41 | 29.0% | 69.3% | |
| 13 | 1690858 | transcription factor with AP2 domain(s) (ApiAP2) | PF3D7_1342900 | E1038K | 0.09 | 0.71 | 0.41 | 3.8% | 25.0% | |
| 8 | 595033 | conserved Plasmodium protein, unknown function | PF3D7_0811800 | D127A | 0.26 | 0.89 | 0.41 | 12.4% | 79.5% | |
| 13 | 946155 | translation initiation factor EIF-2B subunit related | PF3D7_1322300 | Q943E | 0.83 | 0.19 | 0.40 | 1.0% | 0.0% | |
| 10 | 416095 | MORN repeat protein, putative | PF3D7_1010400 | G312R | 0.27 | 0.90 | 0.40 | 4.8% | 87.1% | |
| 3 | 805937 | endonuclease/exonuclease/phosphatase family protein | PF3D7_0319200 | H126N | 0.11 | 0.74 | 0.40 | 1.0% | 47.0% | |
| 3 | 717640 | kinesin-like protein, putative | PF3D7_0317500 | T1248I | 0.00 | 0.57 | 0.40 | 0.0% | 5.3% | |
| 1 | 180056 | conserved P. membrane protein, unknown function | PF3D7_0104100 | I548L | 0.28 | 0.90 | 0.39 | 8.1% | 3.0% | |
| 7 | 1158051 | cysteine desulfurase, putative (NFS) | PF3D7_0727200 | S120I | 0.20 | 0.83 | 0.39 | 4.8% | 38.6% | |
| 5 | 1022710 | ubiquitin fusion degradation protein UFD1, putative | PF3D7_0524800 | I267M | 0.13 | 0.75 | 0.39 | 0.0% | 36.4% | |
| 8 | 475591 | conserved Plasmodium protein, unknown function | PF3D7_0809400 | F7I | 0.30 | 0.91 | 0.39 | 29.0% | 69.3% | |
| 6 | 682567 | conserved Plasmodium protein, unknown function | PF3D7_0616400 | L1394Q | 0.30 | 0.91 | 0.39 | 3.8% | 31.8% | |
| 14 | 1344802 | conserved Plasmodium protein, unknown function | PF3D7_1433700 | K1398M | 0.07 | 0.67 | 0.39 | 1.9% | 39.4% | |
| 9 | 1205910 | merozoite surface protein 1+(MSP1) | PF3D7_0930300 | F1367I | 0.31 | 0.91 | 0.38 | 4.3% | 14.8% | |
| 13 | 1117095 | conserved Plasmodium protein, unknown function | PF3D7_1326600 | S3074F | 0.07 | 0.67 | 0.38 | 0.0% | 0.0% | |
| 7 | 80194 | Plasmodium exported protein, unknown function | PF3D7_0701900 | L864P | 0.24 | 0.85 | 0.38 | 11.4% | 16.7% | |
| 10 | 387292 | protein phosphatase PPM10, putative | PF3D7_1009600 | D106V | 0.16 | 0.77 | 0.38 | 3.8% | 49.6% | |
| 11 | 1639314 | conserved Plasmodium protein, unknown function | PF3D7_1140900 | R1208K | 0.26 | 0.87 | 0.38 | 0.0% | 25.8% | |
| 5 | 689481 | sporozoite surface antigen MB2 (MB2) | PF3D7_0516600 | K721I | 0.00 | 0.55 | 0.38 | 0.0% | 39.0% | |
| 8 | 1201692 | SET domain protein, putative (SET3) | PF3D7_0827800 | N945K | 0.64 | 0.06 | 0.38 | 0.0% | 26.1% | |
| 7 | 618042 | mitochondrial ribosomal protein S5 precursor, putative | PF3D7_0713600 | N37S | 0.24 | 0.85 | 0.37 | 16.7% | 69.3% | |
| 8 | 1201567 | SET domain protein, putative (SET3) | PF3D7_0827800 | Q904K | 0.16 | 0.77 | 0.37 | 0.0% | 26.1% | |
| 14 | 2820945 | aminophospholipid transporter, putative | PF3D7_1468600 | Q1562H | 0.54 | 0.00 | 0.37 | 3.8% | 9.5% | |
| 14 | 2820947 | aminophospholipid transporter, putative | PF3D7_1468600 | E1563V | 0.54 | 0.00 | 0.37 | 3.8% | 9.5% | |
| 8 | 1204112 | SET domain protein, putative (SET3) | PF3D7_0827800 | C1752S | 0.64 | 0.06 | 0.37 | 0.0% | 26.1% | |
| 5 | 630271 | phosphatidylinositol 3-kinase (PI3K) | PF3D7_0515300 | Q431K | 0.60 | 0.03 | 0.37 | 6.7% | 25.4% | 25874676 |
| 4 | 1037822 | pre-mRNA-splicing helicase BRR2, putative (BRR2) | PF3D7_0422500 | V2832I | 0.44 | 0.99 | 0.37 | 4.8% | 4.9% | |
| 13 | 1987528 | conserved Plasmodium protein, unknown function | PF3D7_1349500 | E2696V | 0.56 | 0.01 | 0.37 | 0.0% | 0.0% | |

| Chr | Pos | Gene | | Mutation | Mutation Frequency | | $F_{ST}$ | IBD with C580Y | | Pubmed Ref |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Description | Id | | Wewak | ID | | Wewak | Timika | |
| 7 | 80275 | Plasmodium exported protein, unknown function | PF3D7_0701900 | K837R | 0.25 | 0.85 | 0.37 | 11.4% | 16.7% | |
| 9 | 1110640 | zinc finger protein, putative | PF3D7_0927200 | E956D | 0.84 | 0.24 | 0.37 | 0.5% | 0.4% | |
| 8 | 403309 | tyrosyl-tRNA synthetase, putative | PF3D7_0807900 | S152T | 0.56 | 0.01 | 0.37 | 13.8% | 68.6% | |
| 6 | 900278 | conserved Plasmodium protein, unknown function | PF3D7_0622100 | P696S | 0.77 | 0.17 | 0.36 | 1.0% | 13.3% | |
| 14 | 2830173 | U2 snRNP auxiliary factor, putative | PF3D7_1468800 | P91A | 0.55 | 0.01 | 0.36 | 3.8% | 9.5% | |
| 5 | 923033 | conserved Plasmodium protein, unknown function | PF3D7_0522400 | S4938A | 0.31 | 0.90 | 0.36 | 3.3% | 68.6% | |
| 5 | 569245 | purine nucleoside phosphorylase (PNP) | PF3D7_0513300 | Q225H | 0.39 | 0.95 | 0.36 | 6.7% | 17.8% | 30602534 |
| 6 | 680890 | conserved Plasmodium protein, unknown function | PF3D7_0616400 | D835G | 0.14 | 0.74 | 0.36 | 3.8% | 31.1% | |
| 5 | 924826 | conserved Plasmodium protein, unknown function | PF3D7_0522400 | H5535Q | 0.36 | 0.93 | 0.36 | 3.3% | 68.6% | |
| 5 | 859967 | conserved Plasmodium protein, unknown function | PF3D7_0521000 | S321R | 0.07 | 0.65 | 0.36 | 1.9% | 48.5% | |
| 5 | 756214 | RAP protein, putative | PF3D7_0518100 | R973C | 0.29 | 0.88 | 0.36 | 0.0% | 48.5% | |
| 8 | 294072 | conserved Plasmodium protein, unknown function | PF3D7_0805300 | I934V | 0.62 | 0.06 | 0.36 | 0.0% | 0.0% | |
| 7 | 1301445 | transcription factor with AP2 domain(s) (ApiAP2) | PF3D7_0730300 | K1329N | 0.41 | 0.97 | 0.35 | 0.0% | 14.4% | |
| 7 | 618098 | mitochondrial ribosomal protein S5 precursor, putative | PF3D7_0713600 | V56I | 0.29 | 0.88 | 0.35 | 16.7% | 69.3% | |
| 7 | 80331 | Plasmodium exported protein, unknown function | PF3D7_0701900 | N818K | 0.26 | 0.85 | 0.35 | 11.4% | 16.7% | |
| 12 | 77177 | P. exported protein (PHISTb), unknown function | PF3D7_1201000 | L544R | 0.02 | 0.56 | 0.35 | 0.5% | 26.9% | |
| 3 | 706856 | conserved Plasmodium protein, unknown function | PF3D7_0317300 | S958C | 0.01 | 0.54 | 0.35 | 0.0% | 34.8% | |
| 3 | 706852 | conserved Plasmodium protein, unknown function | PF3D7_0317300 | F959L | 0.01 | 0.54 | 0.35 | 0.0% | 34.8% | |
| 8 | 1129486 | E3 ubiquitin-protein ligase, putative | PF3D7_0826100 | S3220C | 0.54 | 0.01 | 0.35 | 6.7% | 4.9% | |
| 7 | 635759 | conserved Plasmodium protein, unknown function | PF3D7_0713900 | T673N | 0.43 | 0.97 | 0.34 | 18.6% | 71.6% | |
| 10 | 449716 | DNA repair protein RAD23, putative | PF3D7_1011700 | G188V | 0.57 | 0.03 | 0.34 | 4.8% | 92.8% | |
| 9 | 495417 | nucleotide binding protein, putative | PF3D7_0910800 | N419Y | 0.88 | 0.31 | 0.34 | 0.0% | 3.4% | |
| 11 | 1520978 | protein phosphatase PPM2 | PF3D7_1138500 | S839C | 0.13 | 0.70 | 0.34 | 0.0% | 0.0% | |
| 8 | 1183127 | conserved Plasmodium protein, unknown function | PF3D7_0827300 | M611I | 0.93 | 0.38 | 0.34 | 0.0% | 26.1% | |
| 3 | 716603 | kinesin-like protein, putative | PF3D7_0317500 | A1594T | 0.04 | 0.58 | 0.34 | 0.0% | 5.3% | |
| 13 | 2066201 | conserved Plasmodium protein, unknown function | PF3D7_1351800 | K162I | 0.07 | 0.62 | 0.34 | 4.3% | 24.6% | |
| 11 | 1020595 | ThiF family protein, putative | PF3D7_1126100 | P634L | 0.67 | 0.10 | 0.34 | 0.0% | 0.0% | |
| 8 | 1181674 | conserved Plasmodium protein, unknown function | PF3D7_0827300 | E127G | 0.81 | 0.24 | 0.33 | 0.0% | 26.1% | |
| 7 | 502170 | sin3 associated polypeptide p18-like protein | PF3D7_0711400 | S142N | 0.19 | 0.76 | 0.33 | 5.2% | 49.2% | |
| 12 | 442815 | cGMP-specific phosphodiesterase (PDE1) | PF3D7_1209500 | E941A | 0.87 | 0.31 | 0.33 | 0.0% | 0.0% | |
| 9 | 514330 | cysteine repeat modular protein 1+(CRMP1) | PF3D7_0911300 | P945R | 0.93 | 0.39 | 0.33 | 0.0% | 2.3% | |
| 9 | 1472424 | lysophospholipase, putative | PF3D7_0937200 | K192. | 0.23 | 0.80 | 0.32 | 11.9% | 6.8% | |
| 10 | 502210 | zinc finger protein, putative | PF3D7_1013000 | F518L | 0.32 | 0.88 | 0.32 | 2.4% | 90.2% | |
| 14 | 1713410 | tRNA binding protein, putative | PF3D7_1442300 | A142D | 0.53 | 0.02 | 0.32 | 0.0% | 0.0% | |

| Chr | Pos | Gene | | Mutation | Mutation Frequency | | $F_{ST}$ | IBD with C580Y | | Pubmed Ref |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Description | Id | | Wewak | ID | | Wewak | Timika | |
| 10 | 435551 | conserved Plasmodium protein, unknown function | PF3D7_1011100 | K7M | 0.49 | 0.00 | 0.32 | 4.8% | 92.8% | |
| 13 | 523477 | translation initiation factor IF-2, putative | PF3D7_1312400 | N810Y | 0.08 | 0.61 | 0.32 | 1.9% | 0.0% | |
| 12 | 63177 | acyl-CoA synthetase (ACS7) | PF3D7_1200700 | G387A | 0.13 | 0.68 | 0.32 | 0.0% | 0.8% | |
| 3 | 263347 | conserved Plasmodium protein, unknown function | PF3D7_0305500 | P3314S | 0.03 | 0.53 | 0.32 | 0.0% | 0.0% | |
| 13 | 523164 | translation initiation factor IF-2, putative | PF3D7_1312400 | M705I | 0.08 | 0.61 | 0.31 | 1.9% | 0.0% | |
| 5 | 499537 | RNA pseudouridylate synthase, putative | PF3D7_0511500 | M4312I | 0.30 | 0.85 | 0.31 | 21.9% | 54.2% | |
| 9 | 517372 | cysteine repeat modular protein 1+(CRMP1) | PF3D7_0911300 | M1959K | 0.92 | 0.39 | 0.31 | 0.0% | 2.3% | |
| 3 | 494135 | protein kinase, putative | PF3D7_0311400 | I1725V | 0.81 | 0.25 | 0.31 | 1.4% | 0.4% | |
| 1 | 468893 | multidrug resistance-associated protein 1+(MRP1) | PF3D7_0112200 | F1390I | 0.05 | 0.55 | 0.31 | 0.0% | 0.0% | 19117944 |
| 14 | 2467244 | conserved Plasmodium protein, unknown function | PF3D7_1460500 | S385N | 0.47 | 0.00 | 0.31 | 0.0% | 0.0% | |
| 11 | 736367 | clathrin coat assembly protein, putative | PF3D7_1119500 | L100S | 0.57 | 0.06 | 0.31 | 0.0% | 0.0% | |
| 6 | 1010621 | N-acetylglucosaminylphosphatidylinositol deacetylase | PF3D7_0624700 | S33C | 0.13 | 0.67 | 0.31 | 1.9% | 32.2% | |
| 8 | 417335 | plasmepsin X | PF3D7_0808200 | R244K | 0.42 | 0.93 | 0.30 | 23.3% | 82.2% | |
| 10 | 1096446 | conserved Plasmodium protein, unknown function | PF3D7_1025900 | D278Y | 0.39 | 0.91 | 0.30 | 1.0% | 0.0% | |
| 4 | 924677 | transcription factor with AP2 domain(s) (ApiAP2) | PF3D7_0420300 | R2230C | 0.20 | 0.75 | 0.30 | 5.7% | 15.5% | |
| 10 | 829580 | conserved P. membrane protein, unknown function | PF3D7_1020600 | K585R | 0.00 | 0.46 | 0.30 | 0.0% | 30.7% | |
| 3 | 85840 | Plasmodium exported protein, unknown function | PF3D7_0301400 | V66A | 0.79 | 0.24 | 0.30 | 0.0% | 4.2% | |
| 5 | 981238 | conserved P. membrane protein, unknown function | PF3D7_0523700 | L1308P | 0.46 | 0.00 | 0.30 | 0.0% | 40.2% | |

**Supplementary Table 7 – European Nucleotide Archive (ENA) identifiers for samples used in the present study.**

For each sample in PNG and the GMS, we detail: the MalariaGEN sample identifier; the location and country where the sample was collected; the identifier of the sample at ENA (https://www.ebi.ac.uk/ena) where the sequencing reads for the sample have been deposited; and the label used in this paper for Wewak samples that carry the C580Y mutation in *kelch13*.

| Sample ID | Location | Country | ENA | PNG C580Y Id |
|---|---|---|---|---|
| **SPT34235** | Wewak Centre | PNG | ERR2890694 | PNG-C580Y-1 |
| **SPT34323** | Wewak Wirui | PNG | ERR2890710 | PNG-C580Y-2 |
| **SPT34306** | Wewak Wirui | PNG | ERR2890779 | PNG-C580Y-3 |
| **SPT34239** | Wewak Centre | PNG | ERR2890707 | - |
| **SPT34242** | Wewak Centre | PNG | ERR2890693 | - |
| **SPT34244** | Wewak Centre | PNG | ERR2890682 | - |
| **SPT34253** | Wewak Centre | PNG | ERR2890760 | - |
| **SPT34257** | Wewak Centre | PNG | ERR2890738 | - |
| **SPT34260** | Wewak Centre | PNG | ERR2890762 | - |
| **SPT34269** | Wewak Centre | PNG | ERR2890768 | - |
| **SPT34280** | Wewak Centre | PNG | ERR2890684 | - |
| **SPT34283** | Wewak Centre | PNG | ERR2890736 | - |
| **SPT34290** | Wewak Centre | PNG | ERR2890733 | - |
| **SPT34300** | Wewak Centre | PNG | ERR2890687 | - |
| **SPT34303** | Wewak Centre | PNG | ERR2890754 | - |
| **SPT34304** | Wewak Centre | PNG | ERR2890688 | - |
| **SPT34264** | Wewak Centre | PNG | ERR2890715 | - |
| **SPT34238** | Wewak Centre | PNG | ERR2890690 | - |
| **SPT34224** | Wewak Centre | PNG | ERR2890709 | - |
| **SPT34226** | Wewak Centre | PNG | ERR2890731 | - |
| **SPT34227** | Wewak Centre | PNG | ERR2890730 | - |
| **SPT34229** | Wewak Centre | PNG | ERR2890686 | - |
| **SPT34230** | Wewak Centre | PNG | ERR2890716 | - |
| **SPT34231** | Wewak Centre | PNG | ERR2890706 | - |
| **SPT34232** | Wewak Centre | PNG | ERR2890789 | - |
| **SPT34233** | Wewak Centre | PNG | ERR2890689 | - |
| **SPT34236** | Wewak Centre | PNG | ERR2890697 | - |
| **SPT34237** | Wewak Centre | PNG | ERR2890711 | - |
| **SPT34246** | Wewak Centre | PNG | ERR2890699 | - |
| **SPT34247** | Wewak Centre | PNG | ERR2890681 | - |
| **SPT34248** | Wewak Centre | PNG | ERR2890726 | - |
| **SPT34249** | Wewak Centre | PNG | ERR2890748 | - |
| **SPT34250** | Wewak Centre | PNG | ERR2890739 | - |
| **SPT34251** | Wewak Centre | PNG | ERR2890712 | - |
| **SPT34254** | Wewak Centre | PNG | ERR2890704 | - |
| **SPT34256** | Wewak Centre | PNG | ERR2890713 | - |
| **SPT34261** | Wewak Centre | PNG | ERR2890728 | - |
| **SPT34263** | Wewak Centre | PNG | ERR2890764 | - |
| **SPT34265** | Wewak Centre | PNG | ERR2890722 | - |
| **SPT34266** | Wewak Centre | PNG | ERR2890759 | - |
| **SPT34267** | Wewak Centre | PNG | ERR2890717 | - |
| **SPT34270** | Wewak Centre | PNG | ERR2890691 | - |
| **SPT34272** | Wewak Centre | PNG | ERR2890721 | - |
| **SPT34284** | Wewak Centre | PNG | ERR2890755 | - |
| **SPT34285** | Wewak Centre | PNG | ERR2890784 | - |
| **SPT34286** | Wewak Centre | PNG | ERR2890747 | - |
| **SPT34289** | Wewak Centre | PNG | ERR2890741 | - |

| Sample ID | Location | Country | ENA | PNG C580Y Id |
|---|---|---|---|---|
| **SPT34294** | Wewak Centre | PNG | ERR2890758 | - |
| **SPT34298** | Wewak Centre | PNG | ERR2890743 | - |
| **SPT34299** | Wewak Centre | PNG | ERR2890777 | - |
| **SPT34301** | Wewak Centre | PNG | ERR2890703 | - |
| **SPT34225** | Wewak Centre | PNG | ERR2890761 | - |
| **SPT34243** | Wewak Centre | PNG | ERR2890766 | - |
| **SPT34259** | Wewak Centre | PNG | ERR2890735 | - |
| **SPT34287** | Wewak Centre | PNG | ERR2890790 | - |
| **SPT34297** | Wewak Centre | PNG | ERR2890785 | - |
| **SPT34324** | Wewak Wirui | PNG | ERR2890770 | - |
| **SPT34309** | Wewak Wirui | PNG | ERR2890729 | - |
| **SPT34310** | Wewak Wirui | PNG | ERR2890792 | - |
| **SPT34314** | Wewak Wirui | PNG | ERR2890680 | - |
| **SPT34315** | Wewak Wirui | PNG | ERR2890778 | - |
| **SPT34318** | Wewak Wirui | PNG | ERR2890787 | - |
| **SPT34319** | Wewak Wirui | PNG | ERR2890776 | - |
| **SPT34327** | Wewak Wirui | PNG | ERR2890765 | - |
| **SPT34331** | Wewak Wirui | PNG | ERR2890752 | - |
| **SPT34334** | Wewak Wirui | PNG | ERR2890718 | - |
| **SPT34336** | Wewak Wirui | PNG | ERR2890708 | - |
| **SPT34330** | Wewak Wirui | PNG | ERR2890771 | - |
| **SPT34338** | Wewak Wirui | PNG | ERR2890724 | - |
| **SPT34316** | Wewak Wirui | PNG | ERR2890775 | - |
| **SPT34321** | Wewak Wirui | PNG | ERR2890705 | - |
| **SPT34325** | Wewak Wirui | PNG | ERR2890780 | - |
| **SPT34332** | Wewak Wirui | PNG | ERR2890772 | - |
| **PN0097-C** | Alotau | PNG | ERR376140 | - |
| **PN0098-C** | Alotau | PNG | ERR376142 | - |
| **PN0099-C** | Alotau | PNG | ERR376129 | - |
| **PN0101-C** | Alotau | PNG | ERR376133 | - |
| **PN0102-C** | Alotau | PNG | ERR376144 | - |
| **PN0103-C** | Alotau | PNG | ERR376154 | - |
| **PN0104-C** | Alotau | PNG | ERR376137 | - |
| **PN0105-C** | Alotau | PNG | ERR376148 | - |
| **PN0107-C** | Alotau | PNG | ERR376149 | - |
| **PN0108-C** | Alotau | PNG | ERR376150 | - |
| **PN0110-C** | Alotau | PNG | ERR376128 | - |
| **PN0111-C** | Alotau | PNG | ERR376152 | - |
| **PN0113-C** | Alotau | PNG | ERR376153 | - |
| **PN0114-C** | Alotau | PNG | ERR376145 | - |
| **PN0115-C** | Alotau | PNG | ERR376155 | - |
| **PN0116-C** | Alotau | PNG | ERR376122 | - |
| **PN0120-C** | Alotau | PNG | ERR376161 | - |
| **PN0121-C** | Alotau | PNG | ERR376117 | - |
| **PN0122-C** | Alotau | PNG | ERR376118 | - |
| **PN0124-C** | Alotau | PNG | ERR376119 | - |
| **PN0173-C** | Alotau | PNG | ERR527436 | - |
| **PN0177-C** | Alotau | PNG | ERR527439 | - |
| **PN0178-C** | Alotau | PNG | ERR527441 | - |
| **PN0187-C** | Alotau | PNG | ERR527464 | - |
| **PN0172-C** | Maprik | PNG | ERR527452 | - |
| **PN0168-C** | Maprik | PNG | ERR527459 | - |
| **PN0125-C** | Maprik | PNG | ERR376120 | - |
| **PN0126-C** | Maprik | PNG | ERR376146 | - |
| **PN0127-C** | Maprik | PNG | ERR376135 | - |
| **PN0129-C** | Maprik | PNG | ERR376158 | - |

| Sample ID | Location | Country | ENA | PNG C580Y Id |
|---|---|---|---|---|
| **PN0130-C** | Maprik | PNG | ERR376124 | - |
| **PN0134-C** | Maprik | PNG | ERR376143 | - |
| **PN0135-C** | Maprik | PNG | ERR376130 | - |
| **PN0138-C** | Maprik | PNG | ERR376134 | - |
| **PN0140-C** | Maprik | PNG | ERR376157 | - |
| **PN0141-C** | Maprik | PNG | ERR376123 | - |
| **PN0143-C** | Maprik | PNG | ERR376127 | - |
| **PN0145-C** | Maprik | PNG | ERR376141 | - |
| **PN0148-C** | Maprik | PNG | ERR376132 | - |
| **PN0149-C** | Maprik | PNG | ERR376121 | - |
| **PN0151-C** | Maprik | PNG | ERR376136 | - |
| **PN0153-C** | Maprik | PNG | ERR376159 | - |
| **PN0155-C** | Maprik | PNG | ERR527455 | - |
| **PN0156-C** | Maprik | PNG | ERR527466 | - |
| **PN0159-C** | Maprik | PNG | ERR527440 | - |
| **PN0161-C** | Maprik | PNG | ERR527445 | - |
| **PN0163-C** | Maprik | PNG | ERR527450 | - |
| **PN0164-C** | Maprik | PNG | ERR527454 | - |
| **PN0166-C** | Maprik | PNG | ERR527462 | - |
| **PN0169-C** | Maprik | PNG | ERR527471 | - |
| **PN0170-C** | Maprik | PNG | ERR527438 | - |
| **PN0171-C** | Maprik | PNG | ERR527449 | - |
| **PN0176-C** | Maprik | PNG | ERR527442 | - |
| **PN0179-C** | Maprik | PNG | ERR527447 | - |
| **PN0180-C** | Maprik | PNG | ERR527460 | - |
| **PN0181-C** | Maprik | PNG | ERR527448 | - |
| **PN0182-C** | Maprik | PNG | ERR527451 | - |
| **PN0184-C** | Maprik | PNG | ERR527461 | - |
| **PN0185-C** | Maprik | PNG | ERR527443 | - |
| **PN0186-C** | Maprik | PNG | ERR527465 | - |
| **PN0190-C** | Maprik | PNG | ERR527470 | - |
| **PN0192-C** | Maprik | PNG | ERR527469 | - |
| **PN0191-C** | Maprik | PNG | ERR527437 | - |
| **PN0004-C** | Madang | PNG | ERR175525 | - |
| **PN0014-C** | Madang | PNG | ERR175529 | - |
| **PN0031-C** | Madang | PNG | ERR175534 | - |
| **PN0037-C** | Madang | PNG | ERR175546 | - |
| **PN0043-C** | Madang | PNG | ERR175535 | - |
| **PN0055-Cx** | Madang | PNG | ERR216461 | - |
| **PN0056-C** | Madang | PNG | ERR018911 | - |
| **PN0057-C** | Madang | PNG | ERR015321 | - |
| **PN0058-Cx** | Madang | PNG | ERR175542 | - |
| **PN0059-C** | Madang | PNG | ERR015402 | - |
| **PN0073-C** | Madang | PNG | ERR175522 | - |
| **PN0074-C** | Madang | PNG | ERR175524 | - |
| **PN0078-C** | Madang | PNG | ERR175554 | - |
| **PN0079-C** | Madang | PNG | ERR175527 | - |
| **PN0082-C** | Madang | PNG | ERR175545 | - |
| **PN0084-C** | Madang | PNG | ERR175553 | - |
| **PN0085-C** | Madang | PNG | ERR175539 | - |
| **PN0086-C** | Madang | PNG | ERR175533 | - |
| **PN0087-C** | Madang | PNG | ERR175531 | - |
| **PN0090-C** | Madang | PNG | ERR175538 | - |
| **PH0817-C** | Pailin | Cambodia | ERR175488 | - |
| **PH0911-Cx** | Pailin | Cambodia | ERR388747 | - |
| **PH0841-C** | Pailin | Cambodia | ERR180068 | - |

| Sample ID | Location | Country | ENA | PNG C580Y Id |
|---|---|---|---|---|
| **PH0848-C** | Pailin | Cambodia | ERR180090 | - |
| **PH0895-C** | Pailin | Cambodia | ERR216620 | - |
| **PH1000-C** | Pailin | Cambodia | ERR337568 | - |
| **PH1137-C** | Pailin | Cambodia | ERR404253 | - |
| **PH0832-C** | Pailin | Cambodia | ERR175516 | - |
| **PH0852-C** | Pailin | Cambodia | ERR180059 | - |
| **PH0853-C** | Pailin | Cambodia | ERR175518 | - |
| **PH0818-C** | Pailin | Cambodia | ERR175497 | - |
| **PH0819-C** | Pailin | Cambodia | ERR175504 | - |
| **PH0820-C** | Pailin | Cambodia | ERR175495 | - |
| **PH0822-C** | Pailin | Cambodia | ERR175498 | - |
| **PH0823-C** | Pailin | Cambodia | ERR175490 | - |
| **PH0825-C** | Pailin | Cambodia | ERR164732 | - |
| **PH0826-C** | Pailin | Cambodia | ERR175506 | - |
| **PH0828-C** | Pailin | Cambodia | ERR175492 | - |
| **PH0829-C** | Pailin | Cambodia | ERR175520 | - |
| **PH0830-C** | Pailin | Cambodia | ERR171561 | - |
| **PH0831-C** | Pailin | Cambodia | ERR175502 | - |
| **PH0833-C** | Pailin | Cambodia | ERR180081 | - |
| **PH0834-C** | Pailin | Cambodia | ERR175508 | - |
| **PH0835-C** | Pailin | Cambodia | ERR180066 | - |
| **PH0837-C** | Pailin | Cambodia | ERR175519 | - |
| **PH0838-C** | Pailin | Cambodia | ERR171566 | - |
| **PH0839-C** | Pailin | Cambodia | ERR180077 | - |
| **PH0840-C** | Pailin | Cambodia | ERR171567 | - |
| **PH0842-C** | Pailin | Cambodia | ERR180057 | - |
| **PH0843-C** | Pailin | Cambodia | ERR175487 | - |
| **PH0844-C** | Pailin | Cambodia | ERR180054 | - |
| **PH0845-C** | Pailin | Cambodia | ERR180086 | - |
| **PH0846-C** | Pailin | Cambodia | ERR180048 | - |
| **PH0847-C** | Pailin | Cambodia | ERR180051 | - |
| **PH0850-C** | Pailin | Cambodia | ERR171578 | - |
| **PH0851-C** | Pailin | Cambodia | ERR180055 | - |
| **PH0854-C** | Pailin | Cambodia | ERR175507 | - |
| **PH0855-C** | Pailin | Cambodia | ERR171565 | - |
| **PH0856-C** | Pailin | Cambodia | ERR180093 | - |
| **PH0858-C** | Pailin | Cambodia | ERR175511 | - |
| **PH0886-C** | Pailin | Cambodia | ERR216536 | - |
| **PH0887-C** | Pailin | Cambodia | ERR216537 | - |
| **PH0888-C** | Pailin | Cambodia | ERR216538 | - |
| **PH0889-C** | Pailin | Cambodia | ERR216551 | - |
| **PH0890-C** | Pailin | Cambodia | ERR216552 | - |
| **PH0891-Cx** | Pailin | Cambodia | ERR388746 | - |
| **PH0894-C** | Pailin | Cambodia | ERR216555 | - |
| **PH0896-C** | Pailin | Cambodia | ERR216492 | - |
| **PH0899-C** | Pailin | Cambodia | ERR216493 | - |
| **PH0900-C** | Pailin | Cambodia | ERR216557 | - |
| **PH0901-C** | Pailin | Cambodia | ERR216621 | - |
| **PH0902-C** | Pailin | Cambodia | ERR216622 | - |
| **PH0904-C** | Pailin | Cambodia | ERR216624 | - |
| **PH0907-C** | Pailin | Cambodia | ERR216558 | - |
| **PH0908-C** | Pailin | Cambodia | ERR216559 | - |
| **PH0909-C** | Pailin | Cambodia | ERR216627 | - |
| **PH0910-C** | Pailin | Cambodia | ERR216628 | - |
| **PH0912-C** | Pailin | Cambodia | ERR216629 | - |
| **PH0915-C** | Pailin | Cambodia | ERR216560 | - |

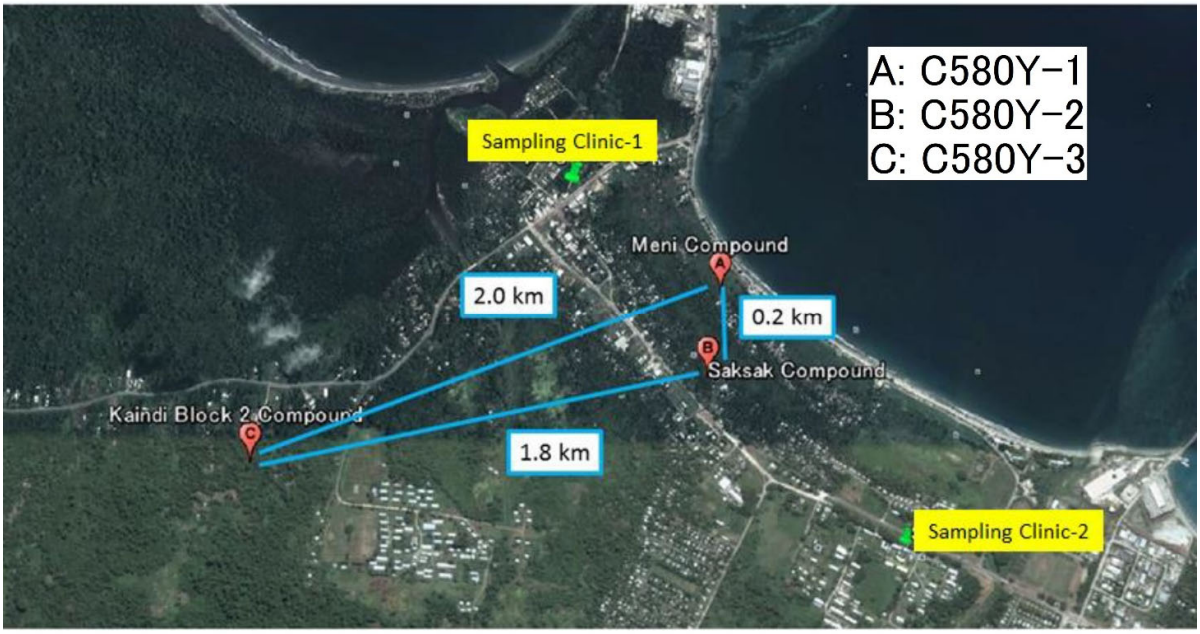
| Sample ID | Location | Country | ENA | PNG C580Y Id |
|---|---|---|---|---|
| **PH0986-C** | Pailin | Cambodia | ERR246561 | - |
| **PH0987-C** | Pailin | Cambodia | ERR246565 | - |
| **PH0991-C** | Pailin | Cambodia | ERR246568 | - |
| **PH0994-C** | Pailin | Cambodia | ERR246569 | - |
| **PH0996-C** | Pailin | Cambodia | ERR246570 | - |
| **PH0997-C** | Pailin | Cambodia | ERR246571 | - |
| **PH0999-C** | Pailin | Cambodia | ERR246572 | - |
| **PH1001-C** | Pailin | Cambodia | ERR337559 | - |
| **PH1002-C** | Pailin | Cambodia | ERR246562 | - |
| **PH1003-C** | Pailin | Cambodia | ERR337563 | - |
| **PH1136-C** | Pailin | Cambodia | ERR426026 | - |
| **PH0575-C** | Pursat | Cambodia | ERR123833 | - |
| **PH0531-C** | Pursat | Cambodia | ERR063619 | - |
| **PH0582-C** | Pursat | Cambodia | ERR123840 | - |
| **PH0526-C** | Pursat | Cambodia | ERR205942 | - |
| **PH0542-C** | Pursat | Cambodia | ERR063633 | - |
| **PH0543-C** | Pursat | Cambodia | ERR205950 | - |
| **PH0555-C** | Pursat | Cambodia | ERR123813 | - |
| **PH0569-C** | Pursat | Cambodia | ERR123827 | - |
| **PH0571-C** | Pursat | Cambodia | ERR123829 | - |
| **PH0678-C** | Pursat | Cambodia | ERR216649 | - |
| **PH0679-C** | Pursat | Cambodia | ERR175474 | - |
| **PH0709-C** | Pursat | Cambodia | ERR171651 | - |
| **PH0717-C** | Pursat | Cambodia | ERR171593 | - |
| **PH0725-C** | Pursat | Cambodia | ERR171597 | - |
| **PH0862-Cx** | Pursat | Cambodia | ERR404173 | - |
| **PH0530-C** | Pursat | Cambodia | ERR205943 | - |
| **PH0532-C** | Pursat | Cambodia | ERR205944 | - |
| **PH0535-C** | Pursat | Cambodia | ERR205941 | - |
| **PH0537-C** | Pursat | Cambodia | ERR067564 | - |
| **PH0579-C** | Pursat | Cambodia | ERR123837 | - |
| **PH0687-C** | Pursat | Cambodia | ERR337566 | - |
| **PH0696-C** | Pursat | Cambodia | ERR175481 | - |
| **PH0700-C** | Pursat | Cambodia | ERR175484 | - |
| **PH0860-Cx** | Pursat | Cambodia | ERR388741 | - |
| **PH0715-C** | Pursat | Cambodia | ERR205968 | - |
| **PH0527-C** | Pursat | Cambodia | ERR205949 | - |
| **PH0528-C** | Pursat | Cambodia | ERR063625 | - |
| **PH0529-C** | Pursat | Cambodia | ERR205948 | - |
| **PH0533-C** | Pursat | Cambodia | ERR067568 | - |
| **PH0534-C** | Pursat | Cambodia | ERR063620 | - |
| **PH0538-C** | Pursat | Cambodia | ERR067566 | - |
| **PH0539-C** | Pursat | Cambodia | ERR063618 | - |
| **PH0540-C** | Pursat | Cambodia | ERR205937 | - |
| **PH0541-C** | Pursat | Cambodia | ERR126018 | - |
| **PH0544-C** | Pursat | Cambodia | ERR067569 | - |
| **PH0551-C** | Pursat | Cambodia | ERR123809 | - |
| **PH0552-C** | Pursat | Cambodia | ERR123810 | - |
| **PH0553-C** | Pursat | Cambodia | ERR123811 | - |
| **PH0554-C** | Pursat | Cambodia | ERR123812 | - |
| **PH0556-C** | Pursat | Cambodia | ERR123814 | - |
| **PH0558-C** | Pursat | Cambodia | ERR123816 | - |
| **PH0559-C** | Pursat | Cambodia | ERR123817 | - |
| **PH0560-C** | Pursat | Cambodia | ERR123818 | - |
| **PH0561-C** | Pursat | Cambodia | ERR123819 | - |
| **PH0562-C** | Pursat | Cambodia | ERR123820 | - |

| Sample ID | Location | Country | ENA | PNG C580Y Id |
|---|---|---|---|---|
| **PH0563-C** | Pursat | Cambodia | ERR123821 | - |
| **PH0565-C** | Pursat | Cambodia | ERR123823 | - |
| **PH0566-C** | Pursat | Cambodia | ERR123824 | - |
| **PH0570-C** | Pursat | Cambodia | ERR123828 | - |
| **PH0576-C** | Pursat | Cambodia | ERR123834 | - |
| **PH0578-C** | Pursat | Cambodia | ERR123836 | - |
| **PH0584-C** | Pursat | Cambodia | ERR123842 | - |
| **PH0676-C** | Pursat | Cambodia | ERR216648 | - |
| **PH0677-C** | Pursat | Cambodia | ERR337565 | - |
| **PH0680-C** | Pursat | Cambodia | ERR175475 | - |
| **PH0682-C** | Pursat | Cambodia | ERR216651 | - |
| **PH0684-C** | Pursat | Cambodia | ERR171656 | - |
| **PH0685-C** | Pursat | Cambodia | ERR175476 | - |
| **PH0686-C** | Pursat | Cambodia | ERR175477 | - |
| **PH0692-C** | Pursat | Cambodia | ERR175479 | - |
| **PH0694-C** | Pursat | Cambodia | ERR205974 | - |
| **PH0695-C** | Pursat | Cambodia | ERR175480 | - |
| **PH0697-C** | Pursat | Cambodia | ERR175482 | - |
| **PH0701-C** | Pursat | Cambodia | ERR175485 | - |
| **PH0702-C** | Pursat | Cambodia | ERR171585 | - |
| **PH0703-C** | Pursat | Cambodia | ERR171586 | - |
| **PH0706-C** | Pursat | Cambodia | ERR171588 | - |
| **PH0708-C** | Pursat | Cambodia | ERR171590 | - |
| **PH0712-C** | Pursat | Cambodia | ERR205971 | - |
| **PH0716-C** | Pursat | Cambodia | ERR171592 | - |
| **PH0718-C** | Pursat | Cambodia | ERR171594 | - |
| **PH0719-C** | Pursat | Cambodia | ERR205963 | - |
| **PH0722-C** | Pursat | Cambodia | ERR171595 | - |
| **PH0723-C** | Pursat | Cambodia | ERR171596 | - |
| **PH0920-Cx** | Pursat | Cambodia | ERR388758 | - |

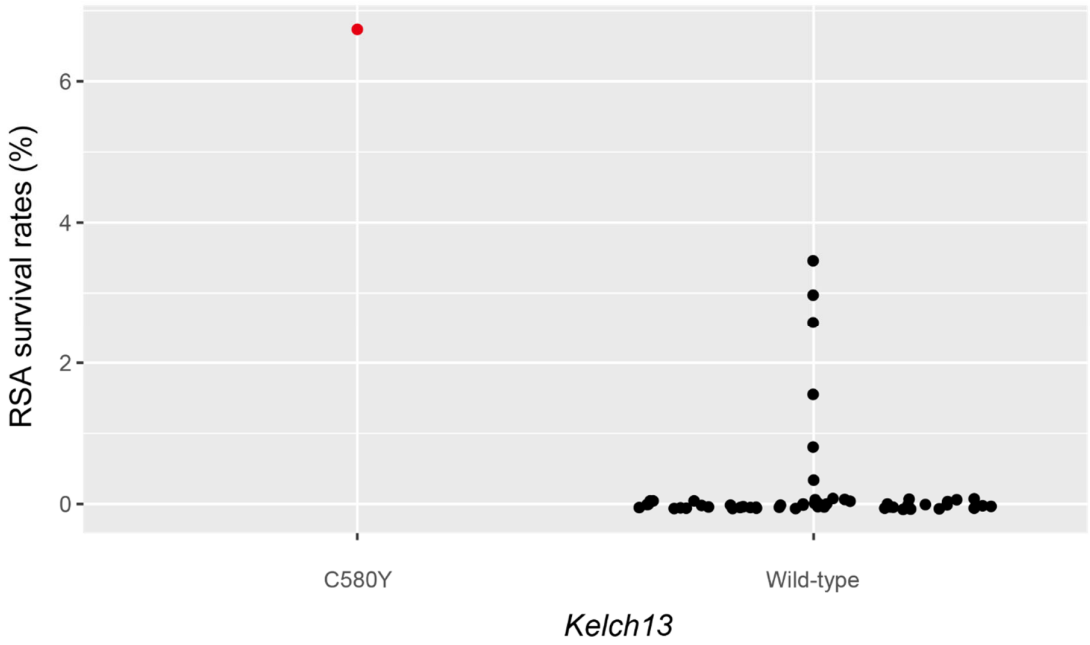
**Supplementary Table 8 - Primers used for determination of *kelch13*-flanking SNPs.**

| Genomic region of SNP | Primer name | Use | Sequence (5' -> 3') |
|---|---|---|---|
| Pf3D7_13_v3: 1700345 | 1700345_F1 | 1st PCR | TTCATATTTCCTGATGGTGTC |
| | 1700345_R1 | 1st PCR | TATAAATAATATGATGGAAGAACC |
| | 1700345_F2 | 2nd PCR & sequencing | ATTTCCTGATGGTGTCTTATC |
| | 1700345_R2 | 2nd PCR | CCTTTTGGGGATGAATTTAGTAAAC |
| Pf3D7_13_v3: 1718288 | 1718288_F1 | 1st PCR | AATAATACAGACGTGAAGAAG |
| | 1718288_R1 | 1st PCR & 2nd HiDi PCR | TAAGATTTCAAGTTCCTTTG |
| | 1718288_F2 | 2nd HiDi PCR (for wild type) | AGTAAAGTAGATCATTCCAA |
| | 1718288_F2 | 2nd HiDi PCR (for mutant) | AGTAAAGTAGATCATTCCAT |
| Pf3D7_13_v3: 1739315 | 1739315_F1 | 1st PCR | CAATAATATCCAATAAAACATCATC |
| | 1739315_R1 | 1st PCR | TTTTTCTATCCATTCAAGTCAAAG |
| | 1739315_F2 | 2nd dCAPS PCR | CTAATAATGCTTGTTCTCTGCA |
| | 1739315_R2 | 2nd dCAPS PCR | CAAGTCAAAGTAATAATTCAAC |
| Pf3D7_13_v3: 1862741 | 1862741_F1 | 1st PCR | GAAATTAGATCAGAAAGAATATAAC |
| | 1862741_R1 | 1st PCR | TCATATTTCATACCTCTC |
| | 1862741_F2 | 2nd PCR & sequencing | GATCAGAAAGAATATAACGAGTTTAG |
| | 1862741_R2 | 2nd PCR | GTAATTCACTTAAAGATGTATTGG |

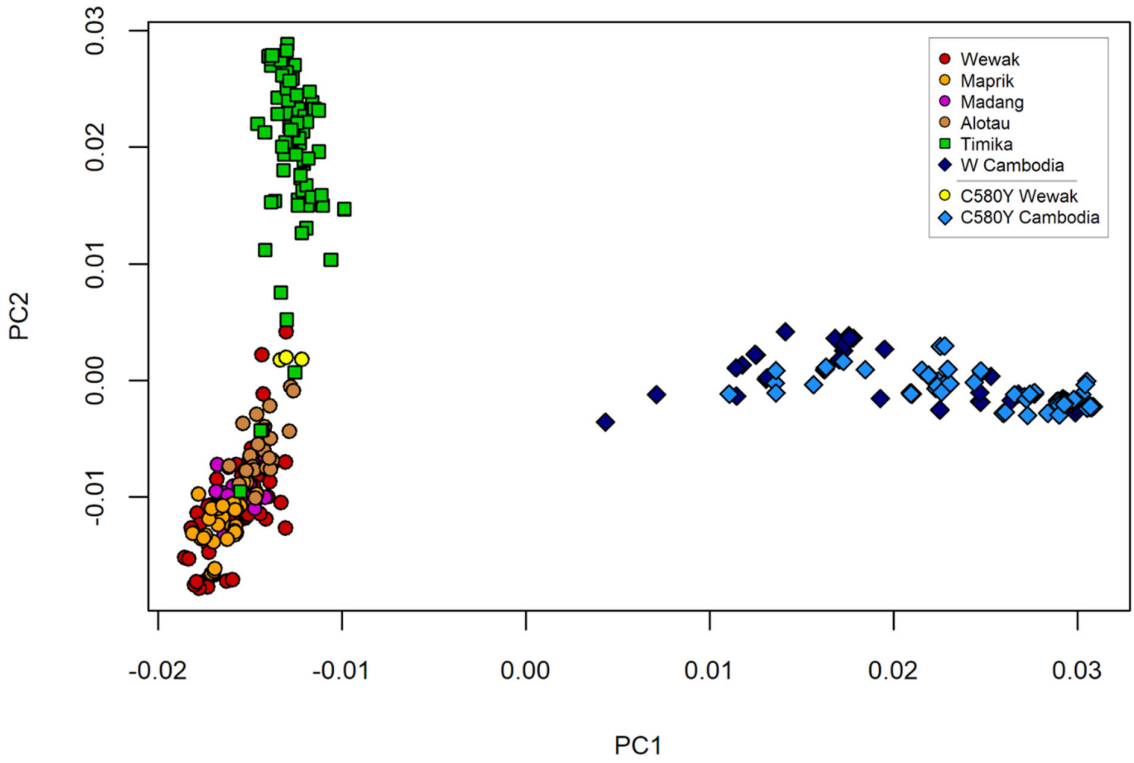
## Supplementary Figures




**Supplementary Figure 1 - Approximate locations of residence of patients infected with *kelch13* C580Y mutant parasites.**
This map of Wewak town shows the place of abode of the three patients whose parasites carried the *kelch13* C580Y allele (red markers), the distance between these locations (blue lines), and the location of the two clinics where the study was carried out (green markers)

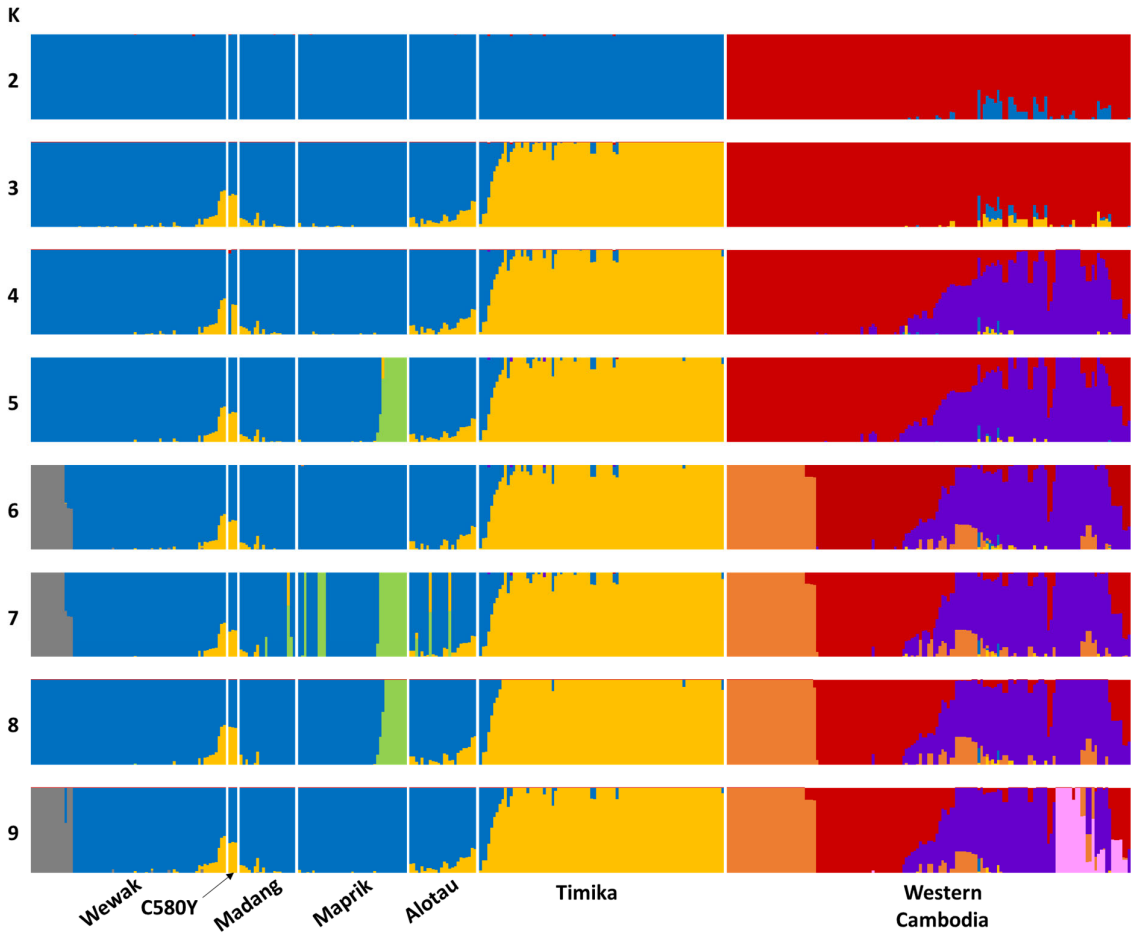



**Supplementary Figure 2 - RSA survival rates of Wewak parasites.**
The plot compares the RSA survival rate (see Methods) of one of the Wewak *kelch13* C580Y mutants (left, red marker) against those for wild-type parasites from the same area (right, black markers). RSA survival rates could not be determined for the remaining two Wewak *kelch13* C580Y mutants. Artemisinin-susceptible laboratory clone 3D7 showed no parasite at 700 nmol/L. MRA-1236 and MRA-1240 (artemisinin-resistant laboratory clones) showed survival rates 14.3% and 27.0%, respectively.




**Supplementary Figure 3 – Principal coordinates analysis (PCoA) based on genome-wide pairwise genetic distances.**

This figure shows a plot of the first two components, for all analyzed samples. The first component (PC1), which explains most of the dataset variance, separates samples from Cambodia from those in New Guinea. The second component (PC2) separates parasites from Papua Indonesia from those found in PNG. We note that there is some overlap between these two groups, and the Wewak C580Y parasites (yellow) appear at an intermediate point between PNG and Indonesian parasites.

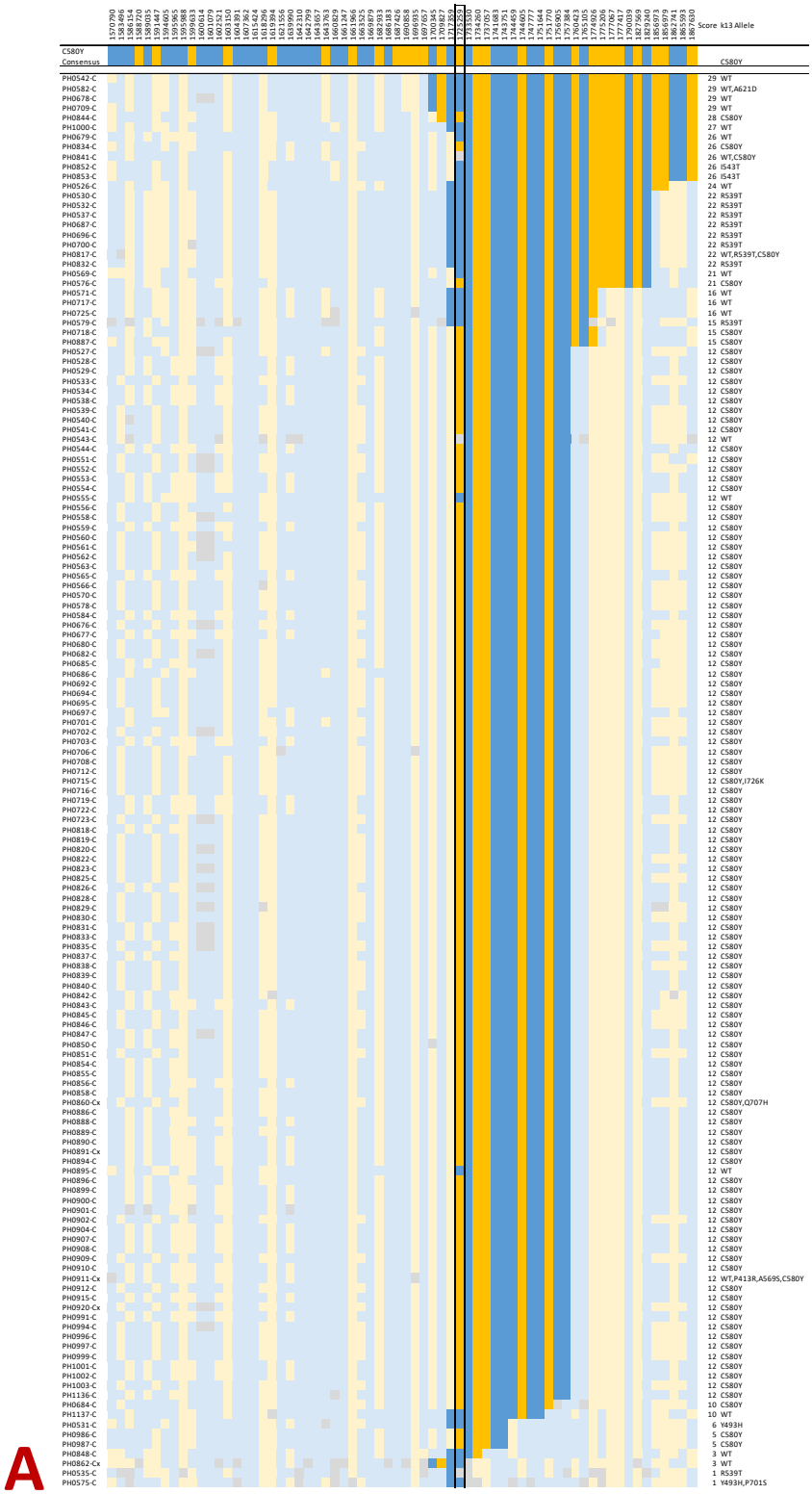



**Supplementary Figure 4 – Population structure in the analyzed sample set.**
The plots show admixture levels estimated by the fastSTRUCTURE software, based on the hypothesis of *K* ancestral populations (*K* = 2 to 9, shown on the left-hand side of the plot). Each population is represented by a different colour (*K* arbitrary colours for each plot); each sample is shown as a vertical bar, coloured according to the proportion of ancestry from each population. Samples are grouped by sampling location, as shown by labels at the bottom; Wewak C580Y mutants are shown as a separate group. From an analysis of the underlying data, the fastSTRUCTURE `chooseK` tool reported that *K*=6 is that number of populations that best explains the population structure.

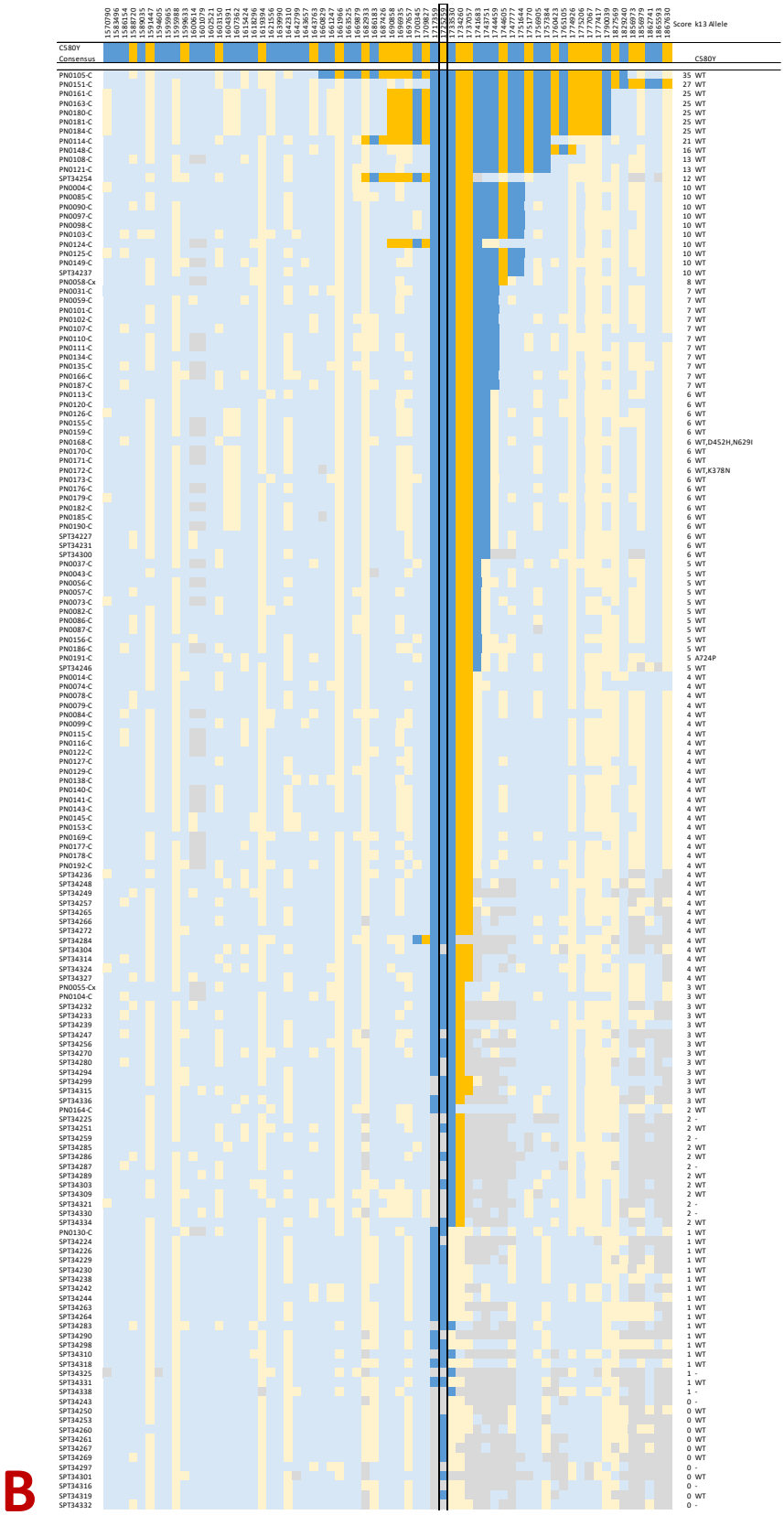
(Overleaf)

**Supplementary Figure 5 - Comparison of haplotypes in a ±150 kbp region on chromosome 13, flanking the kelch13 gene.**

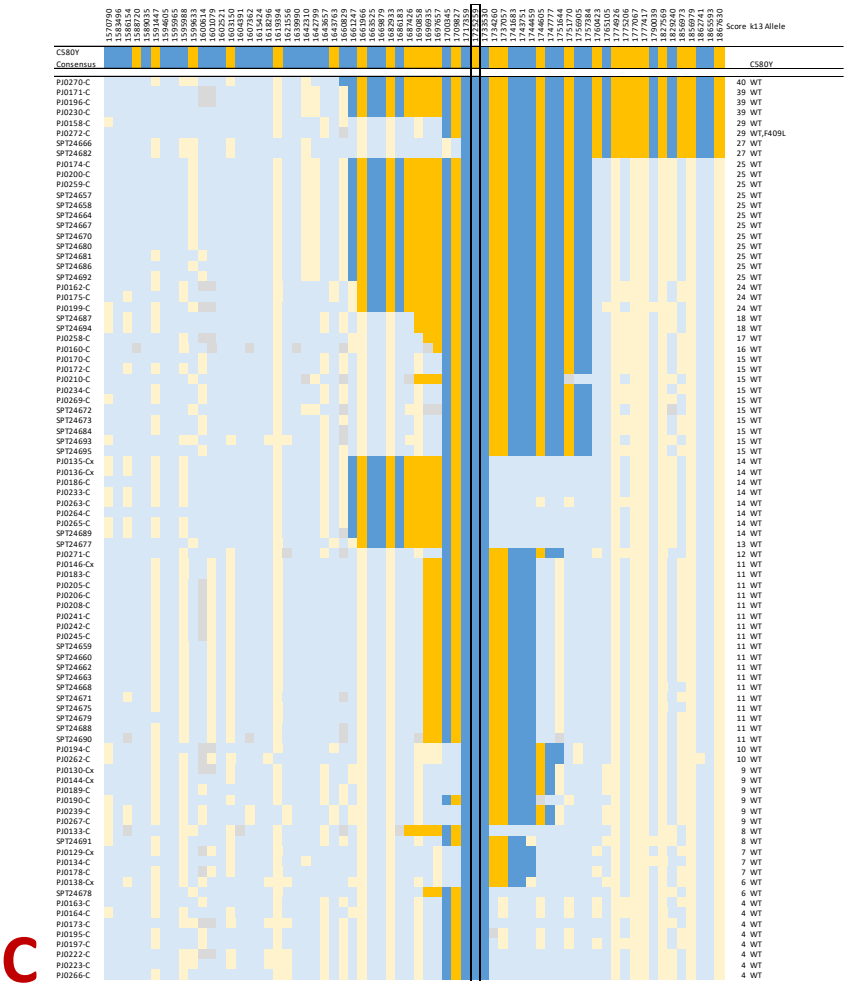
Each row represents the haplotype of a sample, the top row showing the consensus haplotype for the C580Y samples in Wewak. Each column represents a variant position; the position of the kelch13 C580Y mutations (1725259) is indicated by a box outline. Cells colours show the allele call at each position in the sample. Deep colour hues indicate a matching haplotype portion (i.e. consecutive positions within the flanking haplotypes that match the consensus haplotype), while lighter colours indicate positions after a haplotype mismatch. Blue cells symbolize the reference allele, orange the alternative allele, and gray denote a mixed allele call or insufficient coverage. Samples are grouped by country of provenance: Cambodia (panel A), Papua New Guinea (B), Indonesia (C), and sorted by decreasing matching score (the sum of length of the matching haplotype portions in the two flanks). The column on the left shows the sample identifier, while the columns on the right show the matching score and the kelch13 allele carried by the sample.


Score k13 Allele
CS80Y
Consensus
CS80Y
PH0542-C 29 WT
PH0582-C 29 WT,A621D
PH0678-C 29 WT
PH0709-C 29 WT
PH0844-C 28 CS80Y
PH1000-C 27 WT
PH0679-C 26 WT
PH0834-C 26 CS80Y
PH0841-C 26 WT,CS80Y
PH0852-C 26 I543T
PH0853-C 26 I543T
PH0526-C 24 WT
PH0530-C 22 R539T
PH0532-C 22 R539T
PH0537-C 22 R539T
PH0687-C 22 R539T
PH0696-C 22 R539T
PH0700-C 22 R539T
PH0817-C 22 WT,R539T,CS80Y
PH0832-C 22 R539T
PH0569-C 21 WT
PH0576-C 21 CS80Y
PH0571-C 16 WT
PH0717-C 16 WT
PH0725-C 16 WT
PH0579-C 15 R539T
PH0718-C 15 CS80Y
PH0887-C 15 CS80Y
PH0527-C 12 CS80Y
PH0528-C 12 CS80Y
PH0529-C 12 CS80Y
PH0533-C 12 CS80Y
PH0534-C 12 CS80Y
PH0538-C 12 CS80Y
PH0539-C 12 CS80Y
PH0540-C 12 CS80Y
PH0541-C 12 CS80Y
PH0543-C 12 WT
PH0544-C 12 CS80Y
PH0551-C 12 CS80Y
PH0552-C 12 CS80Y
PH0553-C 12 CS80Y
PH0554-C 12 CS80Y
PH0555-C 12 WT
PH0556-C 12 CS80Y
PH0558-C 12 CS80Y
PH0559-C 12 CS80Y
PH0560-C 12 CS80Y
PH0561-C 12 CS80Y
PH0562-C 12 CS80Y
PH0563-C 12 CS80Y
PH0565-C 12 CS80Y
PH0566-C 12 CS80Y
PH0570-C 12 CS80Y
PH0578-C 12 CS80Y
PH0584-C 12 CS80Y
PH0676-C 12 CS80Y
PH0677-C 12 CS80Y
PH0680-C 12 CS80Y
PH0682-C 12 CS80Y
PH0685-C 12 CS80Y
PH0686-C 12 CS80Y
PH0692-C 12 CS80Y
PH0694-C 12 CS80Y
PH0695-C 12 CS80Y
PH0697-C 12 CS80Y
PH0701-C 12 CS80Y
PH0702-C 12 CS80Y
PH0703-C 12 CS80Y
PH0706-C 12 CS80Y
PH0708-C 12 CS80Y
PH0712-C 12 CS80Y
PH0715-C 12 CS80Y,I726K
PH0716-C 12 CS80Y
PH0719-C 12 CS80Y
PH0722-C 12 CS80Y
PH0723-C 12 CS80Y
PH0818-C 12 CS80Y
PH0819-C 12 CS80Y
PH0820-C 12 CS80Y
PH0822-C 12 CS80Y
PH0823-C 12 CS80Y
PH0825-C 12 CS80Y
PH0826-C 12 CS80Y
PH0828-C 12 CS80Y
PH0829-C 12 CS80Y
PH0830-C 12 CS80Y
PH0831-C 12 CS80Y
PH0833-C 12 CS80Y
PH0835-C 12 CS80Y
PH0837-C 12 CS80Y
PH0838-C 12 CS80Y
PH0839-C 12 CS80Y
PH0840-C 12 CS80Y
PH0842-C 12 CS80Y
PH0843-C 12 CS80Y
PH0845-C 12 CS80Y
PH0846-C 12 CS80Y
PH0847-C 12 CS80Y
PH0850-C 12 CS80Y
PH0851-C 12 CS80Y
PH0854-C 12 CS80Y
PH0855-C 12 CS80Y
PH0856-C 12 CS80Y
PH0858-C 12 CS80Y
PH0860-Cx 12 CS80Y,Q707H
PH0886-C 12 CS80Y
PH0888-C 12 CS80Y
PH0889-C 12 CS80Y
PH0890-C 12 CS80Y
PH0891-Cx 12 CS80Y
PH0894-C 12 CS80Y
PH0895-C 12 WT
PH0896-C 12 CS80Y
PH0899-C 12 CS80Y
PH0900-C 12 CS80Y
PH0901-C 12 CS80Y
PH0902-C 12 CS80Y
PH0904-C 12 CS80Y
PH0907-C 12 CS80Y
PH0908-C 12 CS80Y
PH0909-C 12 CS80Y
PH0910-C 12 CS80Y
PH0911-Cx 12 WT,P413R,A569S,CS80Y
PH0912-C 12 CS80Y
PH0915-C 12 CS80Y
PH0920-Cx 12 CS80Y
PH0991-C 12 CS80Y
PH0994-C 12 CS80Y
PH0996-C 12 CS80Y
PH0997-C 12 CS80Y
PH0999-C 12 CS80Y
PH1001-C 12 CS80Y
PH1002-C 12 CS80Y
PH1003-C 12 CS80Y
PH1136-C 12 CS80Y
PH0684-C 10 CS80Y
PH1137-C 10 WT
PH0531-C 6 Y493H
PH0986-C 5 CS80Y
PH0987-C 5 CS80Y
PH0848-C 3 WT
PH0862-Cx 3 WT
PH0535-C 1 R539T
PH0575-C 1 Y493H,P701S

A




B


1570790
1584896
1585154
1588720
1580035
1594427
1594605
1595965
1596088
1599633
1600614
1601709
1602521
1603150
1604931
1607362
1613984
1621556
1628060
1642310
1642799
1643677
1643763
1660829
1661227
1661966
1663525
1668009
1682933
1686183
1687256
1690858
1696935
1697757
1700345
1700827
1701750
1722559
1733530
1734260
1734267
1741683
1743251
1744235
1744605
1747277
1749244
1751770
1759905
1759984
1760423
1763105
1769806
1775206
1777067
1777417
1790039
1822569
1824240
1856973
1859979
1862244
1865593
1867630
Score k13 Allele
CS580Y
Consensus
CS80Y
PJ0270-C 40 WT
PJ0171-C 39 WT
PJ0196-C 39 WT
PJ0230-C 39 WT
PJ0158-C 29 WT
PJ0272-C 29 WT,F409L
SPT24666 27 WT
SPT24682 27 WT
PJ0174-C 25 WT
PJ0200-C 25 WT
PJ0259-C 25 WT
SPT24657 25 WT
SPT24658 25 WT
SPT24664 25 WT
SPT24667 25 WT
SPT24670 25 WT
SPT24680 25 WT
SPT24681 25 WT
SPT24686 25 WT
SPT24692 25 WT
PJ0162-C 24 WT
PJ0175-C 24 WT
PJ0199-C 24 WT
SPT24687 18 WT
SPT24694 18 WT
PJ0258-C 17 WT
PJ0160-C 16 WT
PJ0170-C 15 WT
PJ0172-C 15 WT
PJ0210-C 15 WT
PJ0234-C 15 WT
PJ0269-C 15 WT
SPT24672 15 WT
SPT24673 15 WT
SPT24684 15 WT
SPT24693 15 WT
SPT24695 15 WT
PJ0135-Cx 14 WT
PJ0136-Cx 14 WT
PJ0186-C 14 WT
PJ0233-C 14 WT
PJ0263-C 14 WT
PJ0264-C 14 WT
PJ0265-C 14 WT
SPT24689 14 WT
SPT24677 13 WT
PJ0271-C 12 WT
PJ0146-Cx 11 WT
PJ0183-C 11 WT
PJ0205-C 11 WT
PJ0206-C 11 WT
PJ0208-C 11 WT
PJ0241-C 11 WT
PJ0242-C 11 WT
PJ0245-C 11 WT
SPT24659 11 WT
SPT24660 11 WT
SPT24662 11 WT
SPT24663 11 WT
SPT24668 11 WT
SPT24671 11 WT
SPT24675 11 WT
SPT24679 11 WT
SPT24688 11 WT
SPT24690 11 WT
PJ0194-C 10 WT
PJ0262-C 10 WT
PJ0130-Cx 9 WT
PJ0144-Cx 9 WT
PJ0189-C 9 WT
PJ0190-C 9 WT
PJ0239-C 9 WT
PJ0267-C 9 WT
PJ0133-C 8 WT
SPT24691 8 WT
PJ0129-Cx 7 WT
PJ0134-C 7 WT
PJ0178-C 7 WT
PJ0138-Cx 6 WT
SPT24678 6 WT
PJ0163-C 4 WT
PJ0164-C 4 WT
PJ0173-C 4 WT
PJ0195-C 4 WT
PJ0197-C 4 WT
PJ0222-C 4 WT
PJ0223-C 4 WT
PJ0266-C 4 WT

C

## References for Supplementary Material


Barry, A. E., Schultz, L., Senn, N., Nale, J., Kiniboro, B., Siba, P. M., . . . Reeder, J. C. (2013). High levels of genetic diversity of Plasmodium falciparum populations in Papua New Guinea despite variable infection prevalence. *Am J Trop Med Hyg, 88*(4), 718-725. doi:10.4269/ajtmh.12-0056

Hetzel, M. W., Gideon, G., Lote, N., Makita, L., Siba, P. M., & Mueller, I. (2012). Ownership and usage of mosquito nets after four years of large-scale free distribution in Papua New Guinea. *Malaria journal, 11*(1), 192.

Hetzel, M. W., Pulford, J., Maraga, S., Barnadas, C., Reimer, L. J., Tavul, L., . . . Mueller, I. (2014). Evaluation of the Global Fund-supported National Malaria Control Program in Papua New Guinea, 2009-2014. *P N G Med J, 57*(1-4), 7-29.

Hetzel, M. W., Pulford, J., Ura, Y., Jamea-Maiasa, S., Tandrapah, A., Tarongka, N., . . . Mueller, I. (2017). Insecticide-treated nets and malaria prevalence, Papua New Guinea, 2008-2014. *Bull World Health Organ, 95*(10), 695-705B. doi:10.2471/BLT.16.189902

Muller, I., Bockarie, M., Alpers, M., & Smith, T. (2003). The epidemiology of malaria in Papua New Guinea. *Trends Parasitol, 19*(6), 253-259.

Reimer, L. J., Thomsen, E. K., Koimbu, G., Keven, J. B., Mueller, I., Siba, P. M., . . . Zimmerman, P. A. (2016). Malaria transmission dynamics surrounding the first nationwide long-lasting insecticidal net distribution in Papua New Guinea. *Malar J, 15*(1), 25. doi:10.1186/s12936-015-1067-7

Schultz, L., Wapling, J., Mueller, I., Ntsuke, P. O., Senn, N., Nale, J., . . . Barry, A. E. (2010). Multilocus haplotypes reveal variable levels of diversity and population structure of Plasmodium falciparum in Papua New Guinea, a region of intense perennial transmission. *Malar J, 9*, 336. doi:10.1186/1475-2875-9-336