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], Mehra Somya[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], Moses Laman[9], Alyssa Barry[6,11], Pascal Ringwald[12], Jun Ohashi[13], Francis Hombhanje[14], 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, Port Moresby, Papua New Guinea
[10] Swiss Tropical and Public Health Institute, Basel, Switzerland
[11] University of Melbourne, Melbourne, Australia
[12] World Health Organization, Geneva, Switzerland
[13] Department of Biological Sciences, Graduate School of Science, University of Tokyo, Tokyo, Japan
[14] 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 spreading on the island of New Guinea has gradually acquired a complex ensemble of variants, including *kelch13* C580Y, which may affect 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,[2,5,6] 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.[4,7]

In the GMS, slow artemisinin clearance rates have been associated with genetic mutations in two domains of the parasite's *kelch13* gene.[8] Numerous independently arising *kelch13* mutations have been found in GMS resistant parasites,[9-11] and several of these mutations have been validated in relation with clinical and *in vitro* artemisinin resistance.[12] Increasingly, the genotyping of *kelch13* is being used for the surveillance of artemisinin resistance.[5,9-11] 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.[13] Although the spread of C580Y may have been underpinned by the concurrence of a piperaquine-resistant haplotype,[14] this allele has also been found in western Thailand[6] and in Guyana,[15] suggesting it may be highly beneficial to resistant parasites.

While the GMS has been at the epicenter of resistance to antimalarials multiple times over the last sixty years, resistance has also repeatedly arisen on the island of New Guinea, producing alternative

resistant genetic variants.[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 (this portion will be referred to as "Indonesian Papua" here). Since 2010, PNG has been using artemether-lumefantrine as the first-line treatment for malaria, and delayed parasite clearance has not been observed to date. We previously reported the absence of *kelch13* mutations in PNG samples obtained in 2002-2003 in Wewak district, East Sepik;[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.[17] 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,[18] and used next-generation sequencing data to reconstruct the epidemiology of *kelch13* mutants found.

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) located 2 km apart in Wewak district, East Sepik, PNG (Figure 1).[19] 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).

All individuals showing malaria-suspected symptoms (fever ≥37.5C°, headache, abdominal pain, nausea and/or diarrhea) were screened by RDT (CareStart™ Malaria HRP2/pLDH COMBO Test kit, Access Bio, USA). Individuals older than 1 year with *P. falciparum*-positive result were enrolled, after obtaining informed consent from the patient or guardians. Blood samples were taken by finger prick (age < 2 years, 100–500 µL) or peripheral venipuncture (age ≥ 2 years, 1 mL) and collected into EDTA-containing tubes. To estimate parasitemia, thick and thin blood smears were made on site and stained 2% Giemsa for 30 minutes.

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.[20] 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.[21] Additional targeted genotyping was performed as described in the appendix.

Extracted DNA underwent selective whole genome amplification[22] 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 read counts at 1,043,334 biallelic SNPs in the nuclear genome were called with a standardized analysis pipeline[23]

(Pf6.0 release). 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 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) (Table 2). 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.[23] 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.[24]

Ex-vivo ring-stage survival assay (RSA)

Ex-vivo RSA was performed as previously described[18,21] 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% DMSO 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.[18] 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.[18]

Analysis of ancestry and relatedness

Analyses of population structure were performed using a combination of custom software programs written in Java and R. We constructed an NxN pairwise distance matrix, where N is the number of samples, using a previously published procedure.[9] A neighbour-joining tree was then produced using the `nj` implementation in the R `ape` package. Identity by descent (IBD) analysis was performed on the genotypes obtained by whole-genome sequencing using the programs hmmIBD[25] and isoRelate[26], using default parameters. IBD networks were generated by isoRelate.

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.

Results

Identification and characterization of *kelch13* C580Y mutants

This study was performed in 2016 and 2017 at two clinics in Wewak district, East Sepik in PNG (Figure 1), 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[18] 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.[18] The PNG-C580Y-1 *kelch13* mutant produced the highest survival rate (6.8%), comparable to rates measured in Cambodian C580Y parasites,[8] 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[24] (see Supplementary Notes). 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 (Supplementary Figure 3, Methods). All three *kelch13* C580Y mutants were included in this analysis, and no other parasite from Indonesia or PNG carried *kelch13* mutations. A neighbor-joining tree (NJT), derived from pairwise genome-wide genetic distances, showed a clear separation between Cambodian and New Guinean parasites, and a marked but less pronounced separation between populations in PNG and Indonesia (Figure 2). The genomes of Wewak C580Y mutants were found to be almost identical, and clustered unambiguously with New Guinea parasites. The Wewak mutants were highly differentiated from Cambodian C580Y parasites, indicating they were not recently imported from the GMS. Furthermore, they also did not cluster with most other Wewak parasites, but occupied a branch midway between the PNG and Indonesian groups, suggesting they may be genetically intermediate between the two populations.

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*.[27] 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*[15,28] in the 2017 parasites from Wewak, and eight other isolates for comparison (Supplementary Table 5, Supplementary Methods), 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 4, Supplementary Methods). 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 pronounced similarity between the C580Y mutants and parasites from Timika 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 they 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 be identical because of common ancestry rather than by chance. We estimated the proportion of IBD between the PNG-C580Y-1 genome and those in the comparison populations using hmmIBD, and plotted its distribution (Supplementary Figure 5A). No IDB with Cambodian parasites was detected, further supporting a local origin of the PNG C580Y. This was confirmed by a genome-wide IDB network using isoRelate, which showed large clusters of haplotype sharing in Cambodia and Indonesia, but no relationship to the Wewak mutants (Supplementary Figure 6). The mutants shared lower IBD with parasites from PNG sites (<2%) than is typically observed among local parasites, suggesting they did not originate at the PNG sites analyzed (Supplementary Figure 5B). However, they shared considerably higher IBD proportions with Indonesian population (~3.5-5.5%), suggesting that substantial portions of their genome shares its origin with parasites circulating in Timika. To determine how these portions are distributed across the PNG-C580Y-1 genome, we mapped the IDB predictions for each comparator population (Supplementary Figure 6). Some genome regions were IBD with populations in both PNG and Indonesia, indicating island-wide shared haplotypes. We also identified several loci, notably in chromosomes 5, 10 and 13, where long haplotypes were identical to those in Indonesia, and not found elsewhere in PNG (Supplementary Table 6). The size of these haplotypes suggests they were recently acquired from a donor population- putatively, one circulating in Indonesian Papua- and are possibly under evolutionary selection, raising the question of a possible relation to the C580Y mutation.

Genetic variants association with *kelch13* C580Y

If IBD shared segments were acquired from Indonesian parasites because they contain beneficial alleles, we would expect them to contain variants that circulate in Indonesian Papua but are rare in PNG. We estimated the F_{ST} differentiation measure to identify non-synonymous SNPs where the allele carried by Wewak C580Y mutants is common in the Indonesian samples (F_{ST} <0.2 between Timika and the mutants), but not in PNG (mean F_{ST} ≥ 0.5 between Timika and all PNG sites). These stringent criteria identified 34 SNPs, of which 25 are located in four of the nine IBD regions shared with Indonesia (Table 2).

The selected SNPs are highly enriched for variants implicated in drug response, or located in genes associated with drug resistance. IBD region #3 on chromosome 5 contains mutations Y184F and

N1042D in the *mdr1* gene, which has been implicated in resistance to multiple antimalarials[29,30] including artemisinin; the Y184F variant is common in the GMS, where it has been under selection.[31] In IBD region #9 on chromosome 13, we found *ferredoxin* D193Y, a component of the genetic background associated with artemisinin resistance in the GMS.[24] In IBD region #6 on chromosome 10, we found three noteworthy hits: the T38I mutation in the autophagy-related protein 18 gene (*atg18*), associated with artemisinin response and widespread in SE Asia; and two SNPs (N280S and 659S) in the NLI interacting factor-like phosphatase (*nif4*), implicated in conjunction with *atg18*.[32] We also observed a number of drug-related hits outside the predicted IBD regions, such as Q225H in the purine nucleoside phosphorylase gene (*pnp*) on chromosome 5, which circulates in the GMS and has been associated with drug resistance.[33] Also on chromosome 5, the K753Q mutation in the amino acid transporter PF3D7_0515500 has been implicated in artemether response.[34] Finally, we found G248R in gene PF3D7_1450800, associated to artemisinin resistance in a large-scale analysis.[24] IDB networks around specific loci confirmed that, although Wewak mutants possess haplotypes around *kelch13* that are distinct from those in both Cambodia and Indonesia, they clearly cluster with the Timika population at other genes related to drug resistance, such as *atg18*, *mdr1* and *pnp* (Supplementary Figures 8, 9).

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[35] and sulfadoxine,[36] 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,[13] and was confirmed by a transfection study to confer resistance *in vitro*.[37] 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.[13]

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, with relatively low IBD, suggesting they originated outside the sampled sites. The relatively high level of IBD shared with the Timika population raises the possibility that the C580Y mutants could have crossed the border from Indonesia, or that they may be the product of interbreeding between the two countries. In all cases, the marked differences distinguishing these mutants from previous East Sepik parasites suggests they belong to a spreading population, whose bounds within the island are presently unknown.

The sharing of long haplotypes between the mutants and the Timika parasites are strongly suggestive of a process of selection. If the Indonesian population's genetic contribution were governed by long-term interbreeding, IBD shared fragments would be short, due to breakdown from multiple recombination events. Instead, we observed long shared segments in specific parts of the genome, consistent with recently acquisition and rapid propagation, characteristic of recent selection.[38] This hypothesis is bolstered by the large number of alleles associated with resistance to artemisinin and related drugs shared by the C580Y mutants and the Indonesian Papua parasites, several of which are located in the IBD fragments, 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,[24] and is otherwise absent from PNG. The drug resistant allele *mdr1* Y184F, known to be under selection in the GMS,[31] is also present in a shared 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,[32] 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.[24,32] Additional variants related to artemisinin response, present in Indonesian Papua but rare in PNG, were found in the Wewak C580Y mutants, outside the predicted IBD segments. One of these mutations (PF3D7_0515500 K753Q) is in a transporter that has been associated with response to artemether in African parasites.[34] Another mutation in a nucleoside phosphorylase (*pnp* Q225H), recently shown to affect response to quinine

and mefloquine,[33] two drugs related to lumefantrine, the partner drug used in PNG. Such a high number of mutations implicated in drug response is unlikely to have resulted by chance; it is more plausible that these variants have been acquired and selected for their contribution to a resistant phenotype. Therefore, we propose that the Wewak C580Y lineage has accumulated a complex genetic background through multiple 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.

The findings presented are of great concern for public health, both locally and globally. They provide evidence of a population of ART-R parasites spreading within New Guinea has acquired a complex genetic background, complete with the kelch13 C580Y allele, which provides a survival advantage under artemisinin drug pressure. Such pressure may have intensified due to decreasing *P. falciparum* transmission across the island as a result of control interventions.[39] These parasites are a potential danger to the efficacy of ACTs in New Guinea and could constitute a threat if they are not contained. Genetic data from New Guinea is currently sparse and insufficient to adequately demarcate the spread of this population, which may be present on both sides of the Indonesia-PNG border. 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.

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 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.

References

1. WHO. 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. Ashley EA, Dhorda M, Fairhurst RM, et al. Spread of artemisinin resistance in Plasmodium falciparum malaria. *N Engl J Med* 2014; **371**(5): 411-23.
6. 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.
7. 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.
8. Ariey F, Witkowski B, Amaratunga C, et al. A molecular marker of artemisinin-resistant Plasmodium falciparum malaria. *Nature* 2014; **505**(7481): 50-5.
9. MalariaGEN Plasmodium falciparum Community Project. Genomic epidemiology of artemisinin resistant malaria. *Elife* 2016; **5**.
10. 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 Ch* 2016; **60**(6): 3340-7.
11. 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.
12. WHO. Artemisinin and artemisinin-based combination therapy resistance. Status report. 2016.
13. 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 Infectious Diseases* 2019; **In press**.
14. 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. *The Lancet Infectious diseases* 2018.
15. 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.
16. Mita T, Tanabe K, Takahashi N, et al. Independent evolution of pyrimethamine resistance in Plasmodium falciparum isolates in Melanesia. *Antimicrob Agents Ch* 2007; **51**(3): 1071-7.
17. 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.
18. 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.
19. 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.
20. Sakihama N, Mitamura T, Kaneko A, Horii T, Tanabe K. Long PCR amplification of Plasmodium falciparum DNA extracted from filter paper blots. *Experimental parasitology* 2001; **97**(1): 50-4.
21. 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.

22. 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.
23. 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.
24. Miotto O, Amato R, Ashley EA, et al. Genetic architecture of artemisinin-resistant Plasmodium falciparum. *Nature genetics* 2015; **47**(3): 226-34.
25. 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.
26. 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.
27. Wootton JC, Feng X, Ferdig MT, et al. Genetic diversity and chloroquine selective sweeps in Plasmodium falciparum. *Nature* 2002; **418**(6895): 320-3.
28. 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.
29. Sidhu AB, Valderramos SG, Fidock DA. pfmdr1 mutations contribute to quinine resistance and enhance mefloquine and artemisinin sensitivity in Plasmodium falciparum. *Molecular microbiology* 2005; **57**(4): 913-26.
30. 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.
31. 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.
32. 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.
33. 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).
34. 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.
35. 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.
36. 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.
37. 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.
38. 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.
39. 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.

Tables

Table 1. Characteristics of enrolled patients.

Characteristics	2016	2017
Patients (n)	123	134
Gender (n)[1]		
Male	53	57
Female	68	76
Age (year)		
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*)	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

(Next Page)

Table 3. Non-synonymous mutations shared by Wewak C580Y mutants and the Indonesian Papua population. For each position, we show (left to right): the chromosome and nucleotide position; the frequency in the Wewak C580Y mutants, and at each of the sites surveyed; the id and name of the gene; the amino acid substitution; and, if the mutation is found in a shared IBD region (orange background), the id of the region (see Supplementary Table 6). Mutations associated with antimalarial drug resistance in published studies (see main text) are listed in **bold** type.

Chr	Pos	Non-reference allele frequencies							Gene	Description	Mutation	Region
		Wewak C580Y	Wewak	Madang	Maprik	Alotau	Timika	Cambodia				
2	158568	0.83	0.06	0.20	0.05	0.29	0.86	0.13	PF3D7_0203200	conserved Plasmodium protein, unknown function	D365N	1
5	274721	1.00	0.14	0.15	0.05	0.50	0.92	0.15	PF3D7_0506500	conserved Plasmodium protein, unknown function	E2773D	
	569245	1.00	0.42	0.10	0.08	0.38	0.95	0.69	**PF3D7_0513300**	**purine nucleoside phosphorylase (PNP)**	**Q225H**	
	641983	1.00	0.15	0.15	0.08	0.08	0.94	0.12	**PF3D7_0515500**	**amino acid transporter, putative**	**K753Q**	
	719719	1.00	0.05	0.05	0.00	0.08	0.76	0.01	PF3D7_0517100	conserved Plasmodium protein, unknown function	A46V	
	756214	1.00	0.31	0.00	0.00	0.17	0.88	0.22	PF3D7_0518100	RAP protein, putative	R973C	
	827126	1.00	0.14	0.00	0.03	0.04	0.83	0.00	PF3D7_0519900	conserved Plasmodium protein, unknown function	N96S	
	920718	1.00	0.24	0.30	0.10	0.21	0.92	0.96	PF3D7_0522400	conserved Plasmodium protein, unknown function	Y4166F	3
	924826	1.00	0.38	0.35	0.03	0.17	0.93	0.94	PF3D7_0522400	conserved Plasmodium protein, unknown function	H5535Q	
	958440	1.00	0.12	0.05	0.03	0.00	0.91	0.90	**PF3D7_0523000**	**multidrug resistance protein (MDR1)**	**Y184F**	
	961013	1.00	0.09	0.00	0.03	0.00	0.91	0.00	**PF3D7_0523000**	**multidrug resistance protein (MDR1)**	**N1042D**	
8	475591	1.00	0.32	0.35	0.15	0.00	0.91	0.06	PF3D7_0809400	conserved Plasmodium protein, unknown function	F7I	
10	395382	0.67	0.11	0.10	0.05	0.04	0.83	0.00	PF3D7_1009800	conserved Pl.. membrane protein, unknown function	G871A	
	458574	1.00	0.25	0.05	0.13	0.33	0.93	0.00	PF3D7_1011800	QF122 antigen	N1059K	
	461962	1.00	0.10	0.00	0.00	0.25	0.92	0.00	PF3D7_1011900	heme oxygenase	T11I	
	471382	1.00	0.26	0.10	0.36	0.21	0.92	0.00	PF3D7_1012200	rhoptry associated adhesin	K135I	
	471727	1.00	0.12	0.15	0.10	0.13	0.91	0.00	PF3D7_1012200	rhoptry associated adhesin	A200S	
	488090	1.00	0.05	0.00	0.00	0.04	0.86	0.00	**PF3D7_1012700**	**NLI interacting factor-like phosphatase, putative**	**N280S**	6
	489227	1.00	0.05	0.00	0.00	0.04	0.89	0.00	**PF3D7_1012700**	**NLI interacting factor-like phosphatase, putative**	**N659S**	
	494817	1.00	0.32	0.10	0.33	0.08	0.93	0.04	PF3D7_1012800	conserved Plasmodium protein, unknown function	K756N	
	497461	1.00	0.07	0.00	0.00	0.04	0.92	0.94	**PF3D7_1012900**	**autophagy-related protein 18**	**T38I**	
	502210	1.00	0.34	0.15	0.11	0.04	0.88	0.02	PF3D7_1013000	zinc finger protein, putative	F518L	
	517067	1.00	0.07	0.05	0.10	0.13	0.88	0.00	PF3D7_1013300	zinc finger protein, putative	S181C	
	522059	1.00	0.03	0.00	0.00	0.00	0.85	0.00	PF3D7_1013400	conserved Plasmodium protein, unknown function	V65A	
13	666567	1.00	0.11	0.00	0.03	0.08	0.82	0.16	PF3D7_1315900	exportin 1-like protein, putative	T943A	
	748395	1.00	0.05	0.05	0.00	0.13	0.92	0.98	**PF3D7_1318100**	**ferredoxin, putative**	**D193Y**	
	757313	0.00	0.94	0.90	0.77	0.63	0.07	0.02	PF3D7_1318300	conserved Plasmodium protein, unknown function	S1135T	9
	758176	1.00	0.04	0.05	0.08	0.17	0.86	0.66	PF3D7_1318300	conserved Plasmodium protein, unknown function	D1423Y	
	760334	1.00	0.06	0.05	0.13	0.17	0.86	0.66	PF3D7_1318400	chromosome segregation protein, putative	H1217N	
	1709827	1.00	0.10	0.00	0.13	0.13	0.85	0.04	PF3D7_1343300	CDT1-like protein, putative	A285S	
	2360453	0.00	0.69	0.70	0.90	0.50	0.01	0.00	PF3D7_1359400	rRNA associated RNA binding protein, putative	K249R	
14	1287030	1.00	0.12	0.10	0.03	0.38	0.89	0.85	PF3D7_1432600	conserved Plasmodium protein, unknown function	N742S	
	2081506	1.00	0.08	0.00	0.03	0.13	0.75	0.00	**PF3D7_1450800**	**conserved Plasmodium protein, unknown function**	**G248R**	
	2432735	1.00	0.10	0.05	0.03	0.04	0.82	0.00	PF3D7_1459300	OPA3-like protein, putative	K187.	

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

Olivo Miotto, Makoto Sekihara, *et al.*

Figures

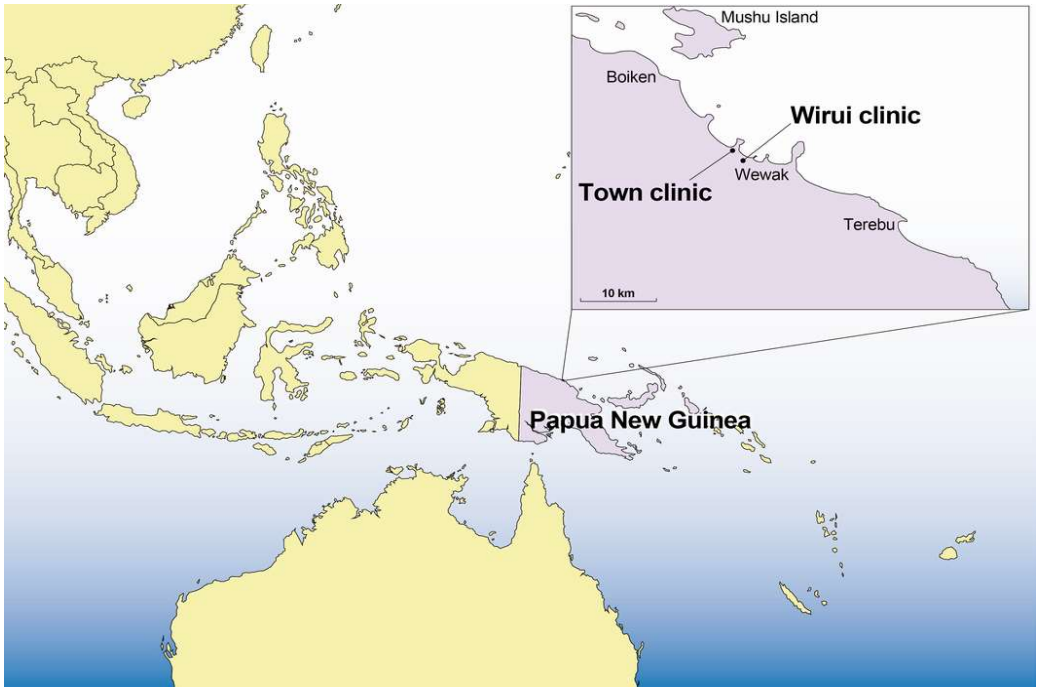


Figure 1. Geographical location of sampling sites.

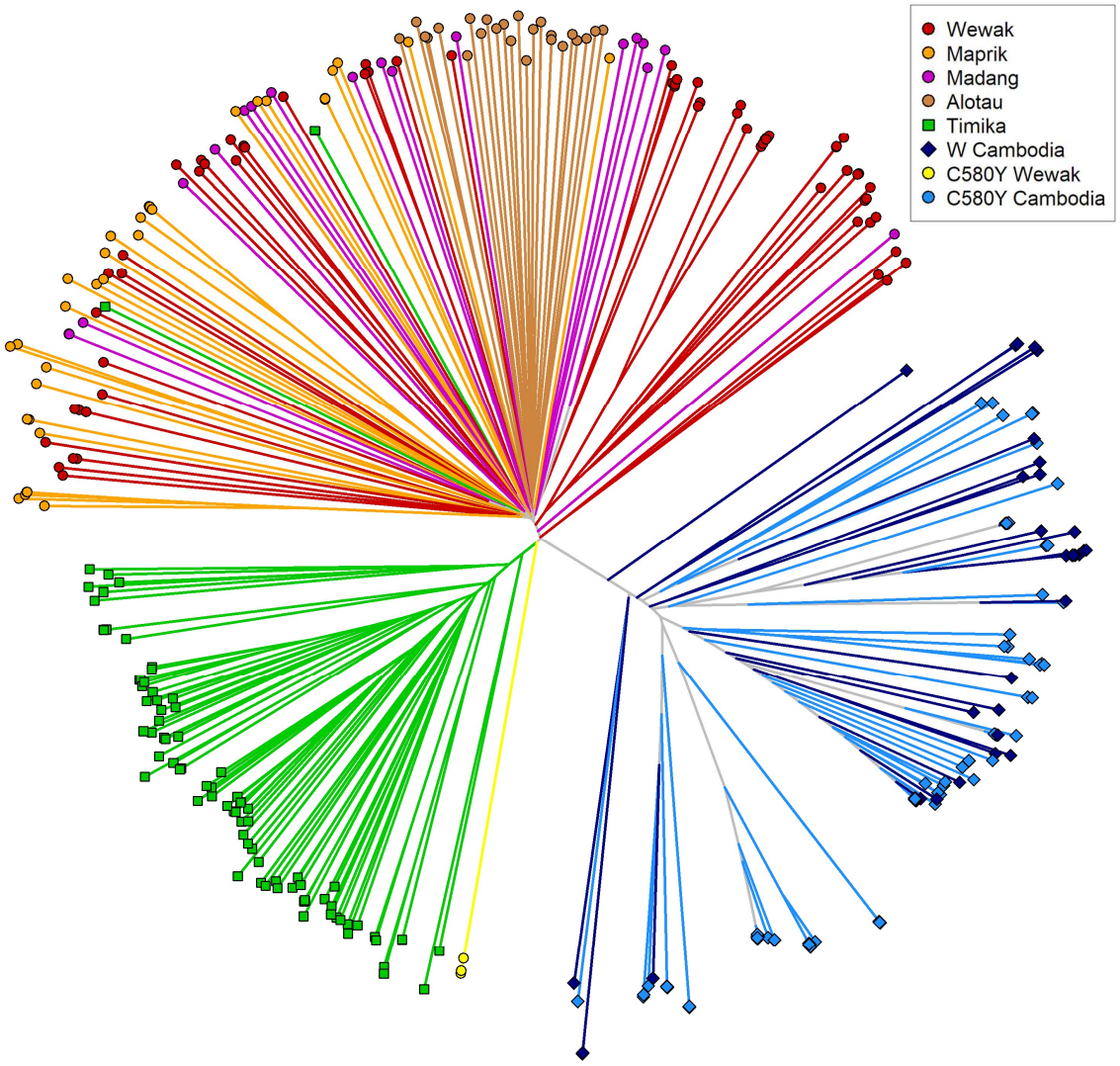


Figure 2. Neighbour-joining tree from whole-genome sequence data, showing the genetic relationship between samples from different geographical locations. Each branch represents a sample, coloured according to provenance and C580Y mutation status. Branch length represent estimated genetic distance, such that two samples separated by distant branches are more dissimilar at genome level than two samples whose branches bifurcate close to their tip.

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

Olivo Miotto, Makoto Sekihara, *et al.*

APPENDIX

Contents

Supplementary Methods

Description of sampling locations

Wewak district comprises about 20,000 inhabitants, nearly all village-dwellers along the coast and inland. 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.[1] Average LLIN usage was 55% in 2008 and 2009 and 32.9–67.7% in 2013–2014.[2,3] *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.[4] Current first-line regime is Artemether plus Lumefantrine, officially introduced in 2010. Free distribution of ACT and rapid diagnostic tests (RDT) was initiated by the malaria control programme in 2012.[1] Previously, intramuscularly artemether was sometimes used in patients who failed first-line treatment regime.

Determinations of *kelch13* flanking haplotypes

Determinations of 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[5,6]. We also genotyped four KEL1/PLA1 marker SNPs located at -137.5 kb, -14.06 kb, 6.97 kb, and 24.91 kb[7]. 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 7.

Determinations of flanking haplotypes using whole-genome sequencing data. Comparing flanking haplotypes constructed by concatenating WGS genotypes at SNPs with minor allele frequency (MAF) >=0.01 in a region of 300kbp centered at *kelch13* (Supplementary Figure 4, Supplementary Methods).

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 positons) 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*)[8] were performed by multiplex PCR

and probe qPCR as previously described.[9] Copy number variants of plasmepsin 2 and multidrug resistance gene 1 genes were determined using published procedures.[10] Four artemisinin-resistant clones (MRA-1236, MRA-1238, MRA-1240 and MRA-1241) and three isolates obtained in Cambodia[11] were used as comparators.

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). 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. 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	*kelch13*	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 - Regions where Wewak C580Y mutants show IBD with the Papua Indonesia populations, but not with samples from PNG.

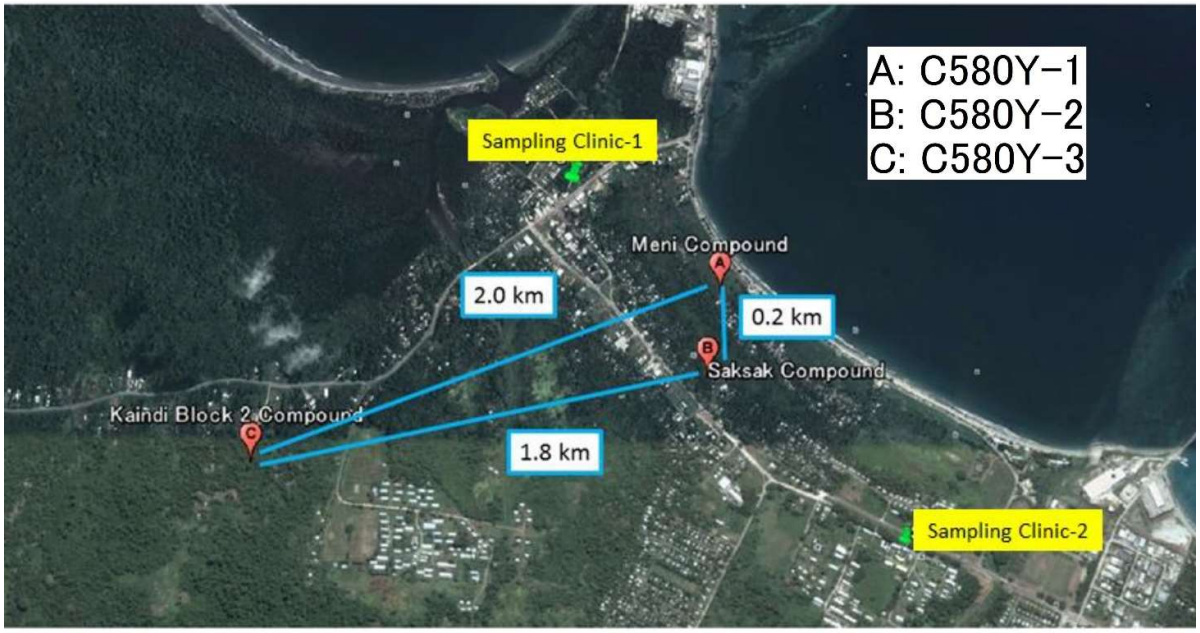
For each region, we show (left to right): an identifier; the chromosome where the region is located; start and end nucleotide positions; and the number and proportion of Indonesian parasites that are predicted to share the IBD segment.

Region	Chromosome	Start Position	End Position	IBD samples in ID	
				count	proportion
1	2	128185	169541	3	3.4%
2	3	748600	865789	9	10.2%
3	5	679000	1052779	35	39.8%
4	7	135743	177033	8	8.0%
5	8	1256966	1297695	4	4.5%
6	10	387292	631276	73	83.0%
7	10	816545	842700	5	57.0%
8	13	109859	175593	20	22.7%
9	13	658708	786950	13	14.8%

Supplementary Table 7 - Primers used for determination of K13-flanking SNPs.

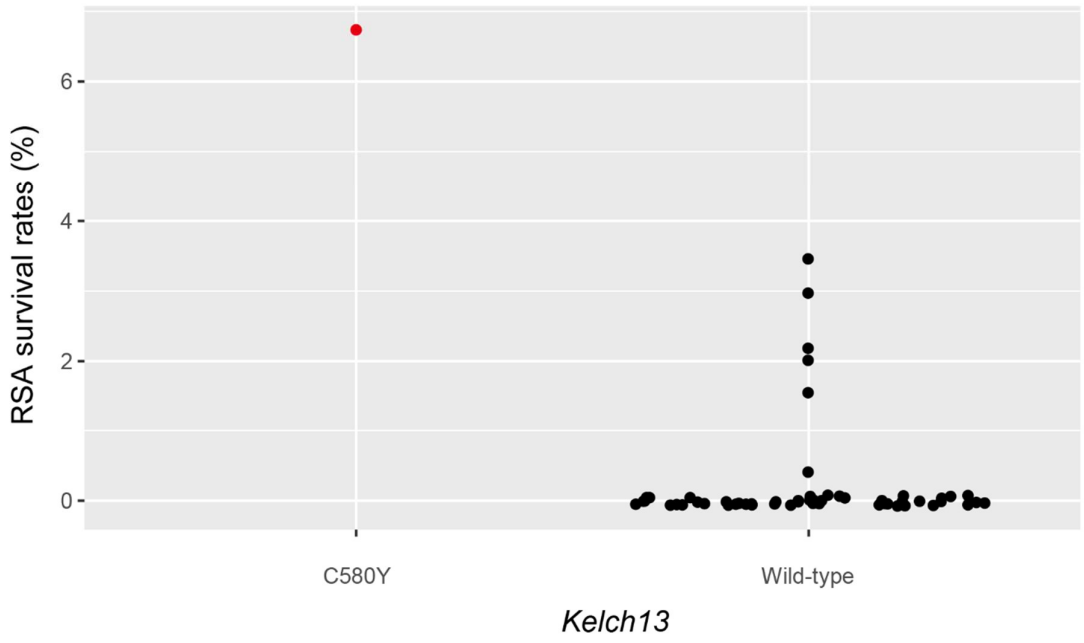
Genomic region of SNP	Primer name	Use	Sequence (5' -> 3')
Pf3D7_13_v3: 1700345	1700345_F1	1st PCR	TTCATATTTCCTGATGGTGTC
	1700345_R1	1st PCR	TATAAATAATATGATGGAAGAACC
	1700345_F2	2nd PCR & sequencing	ATTTCCTGATGGTGTCTTATC
	1700345_R2	2nd PCR	CCTTTTGGGGATGAATTTAGTAAAC
Pf3D7_13_v3: 1718288	1718288_F1	1st 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

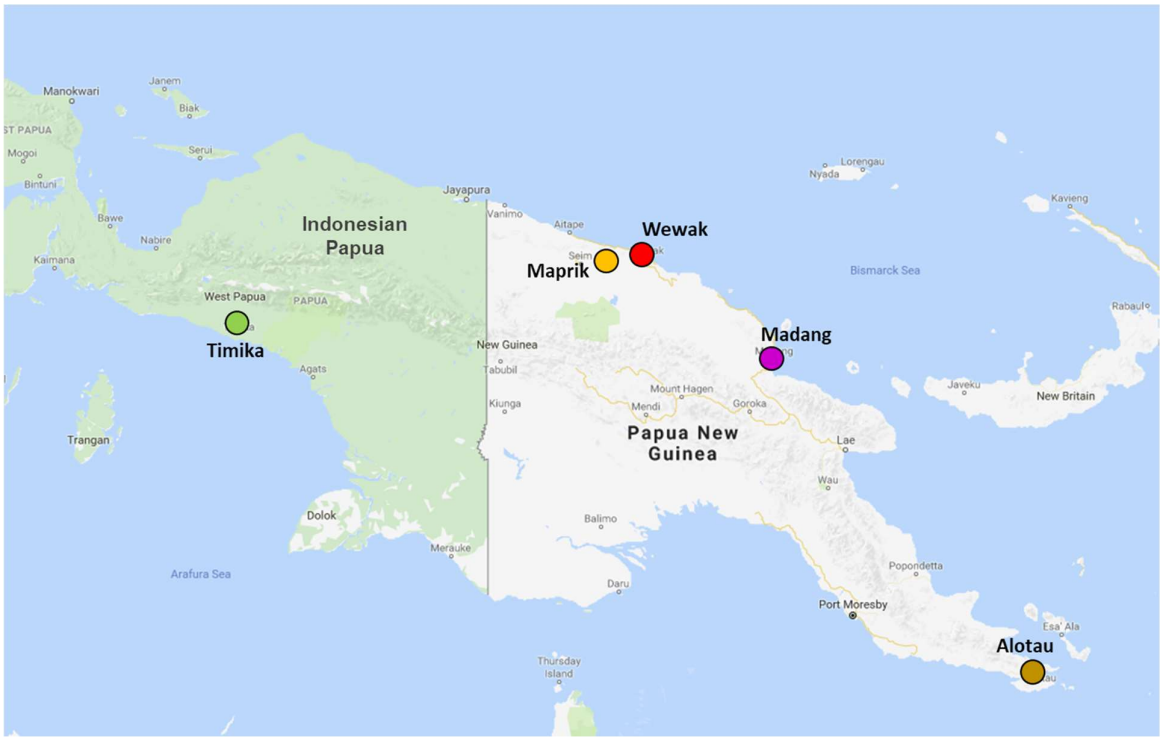


Supplementary Figure 1 - Approximate locations of residence of patients infected with *kelch13* C580Y mutant parasites.

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.

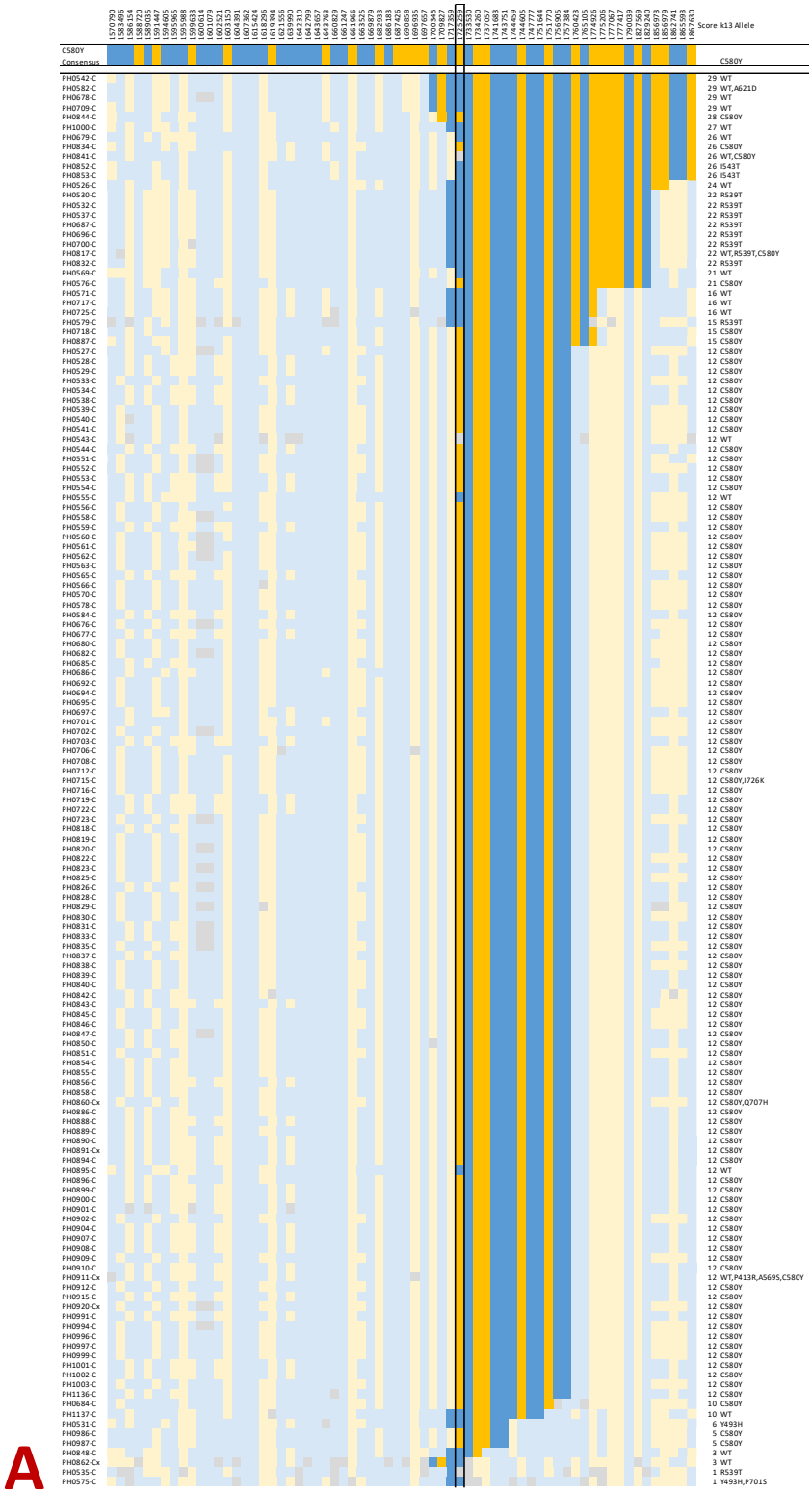


Supplementary Figure 3 – Sampling sites in New Guinea for parasites used in whole-genome analyses. Sites that contributed samples in both countries that form part of the island of New Guinea (Indonesia to the west and PNG to the east) are shown by coloured circular markers.

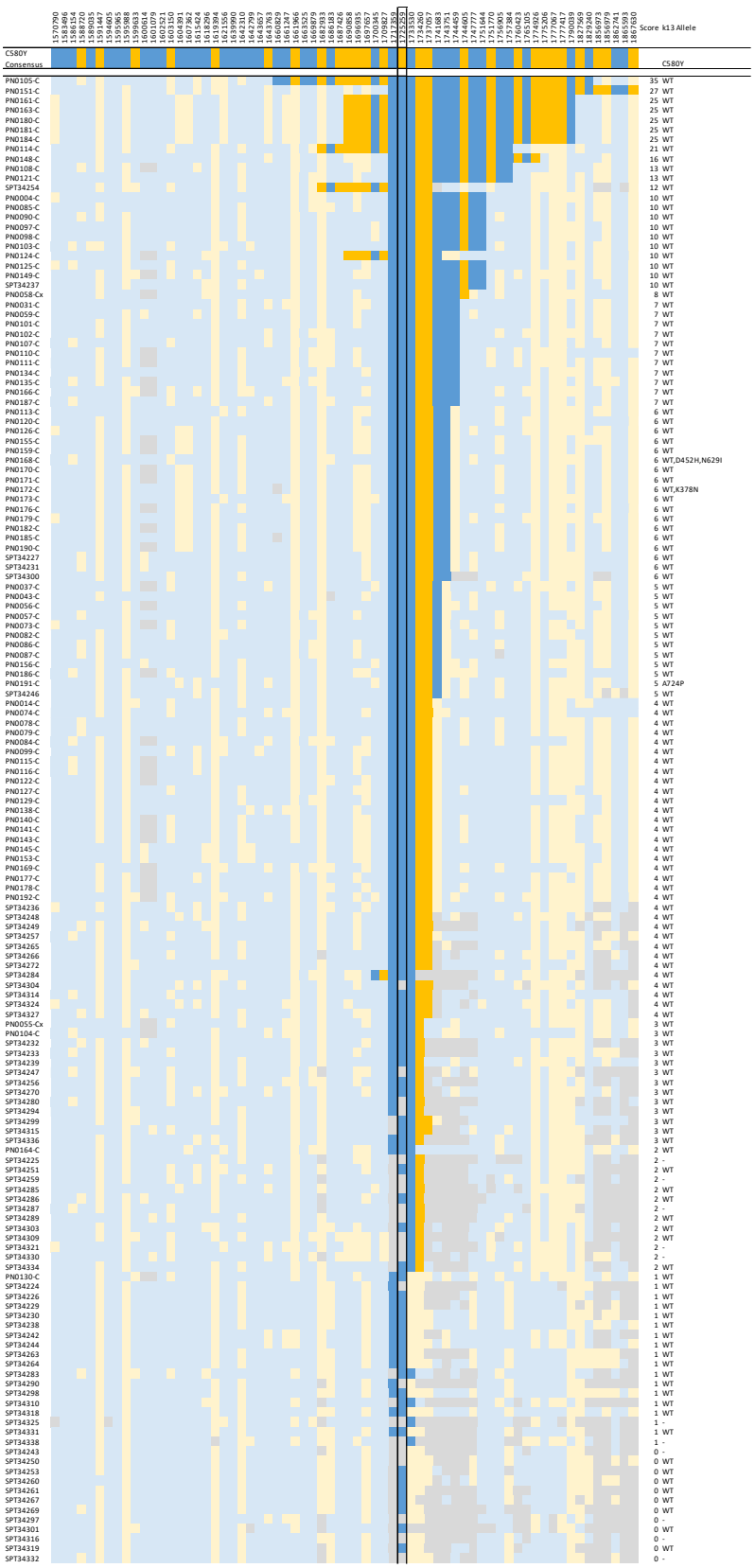
(Overleaf)

Supplementary Figure 4 - Comparison of haplotypes in a ±150 kbp region on chromosome 13, flanking the kelch13 gene.

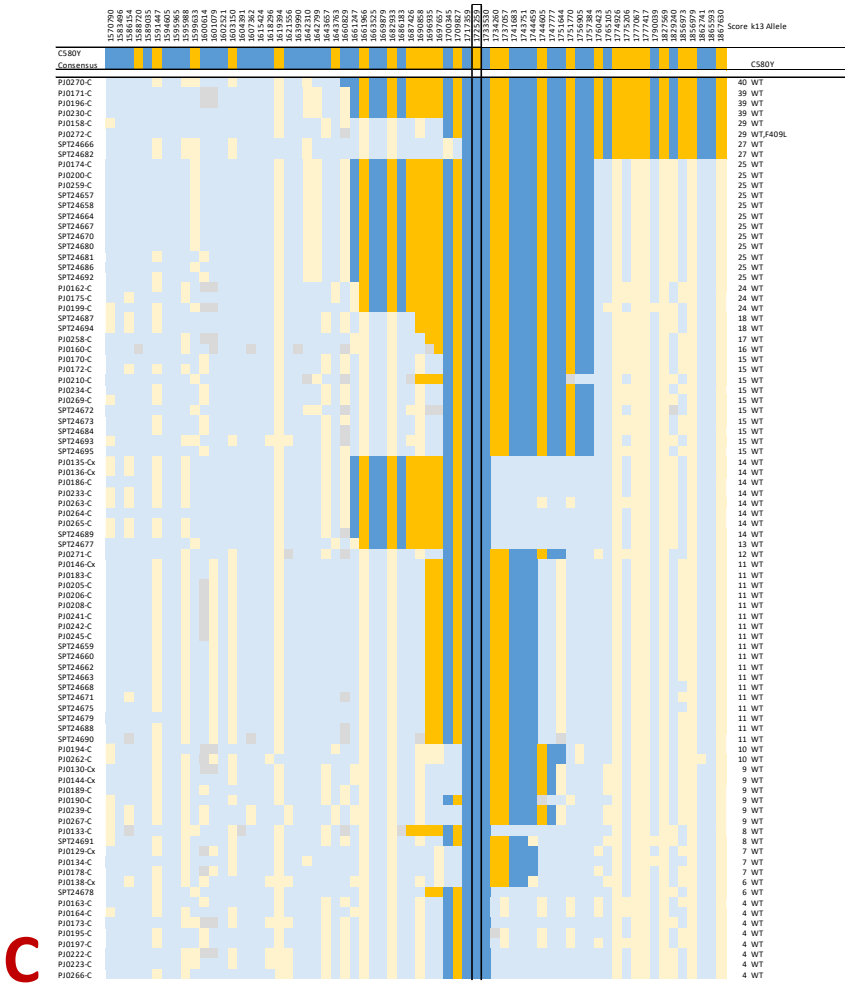
Each row represents the haplotype of a sample, the top row showing the consensus haplotype for the C580Y samples in Wewak. Each column represents a variant position; the position of the kelch13 C580Y mutations (1725259) is indicated by a box outline. Cells colours show the allele call at each position 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.



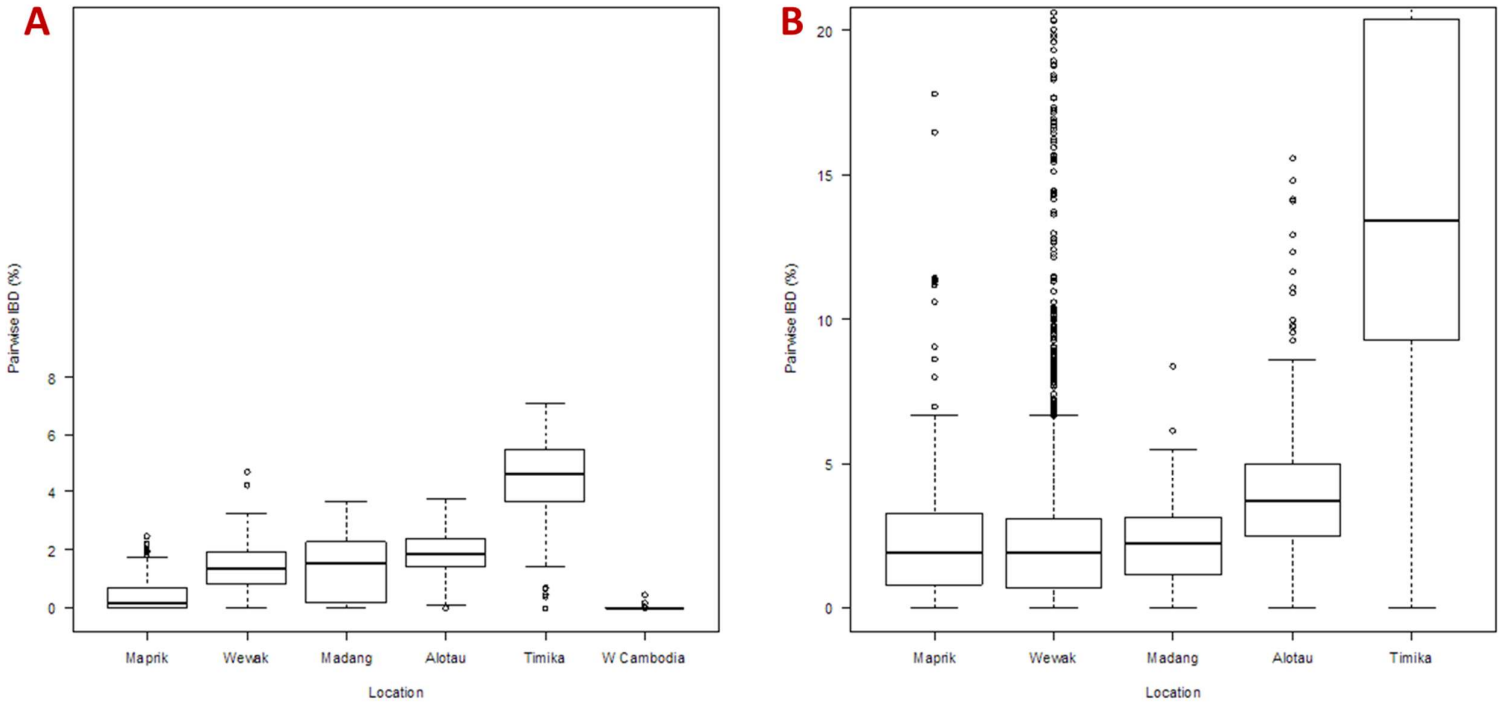
A



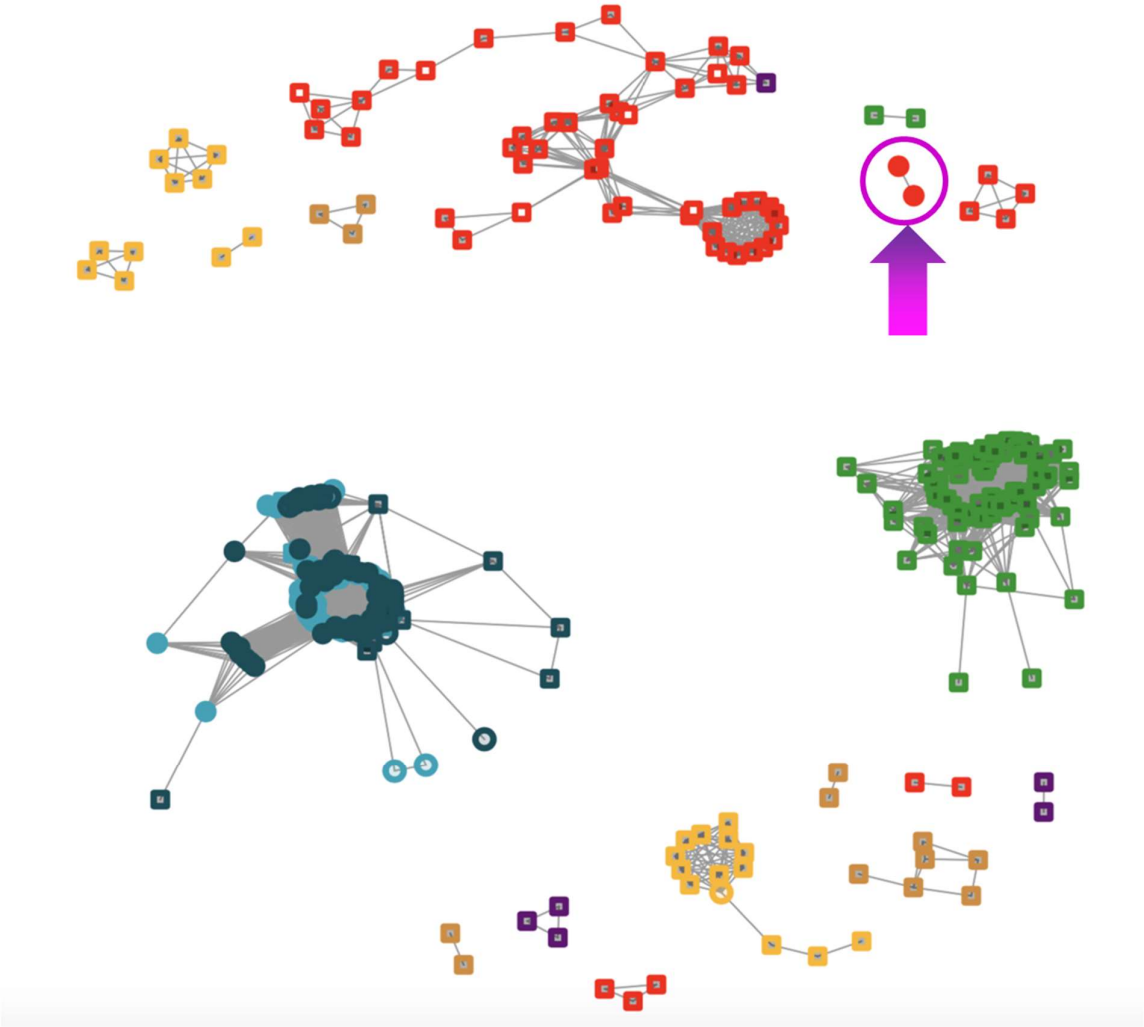
B



C



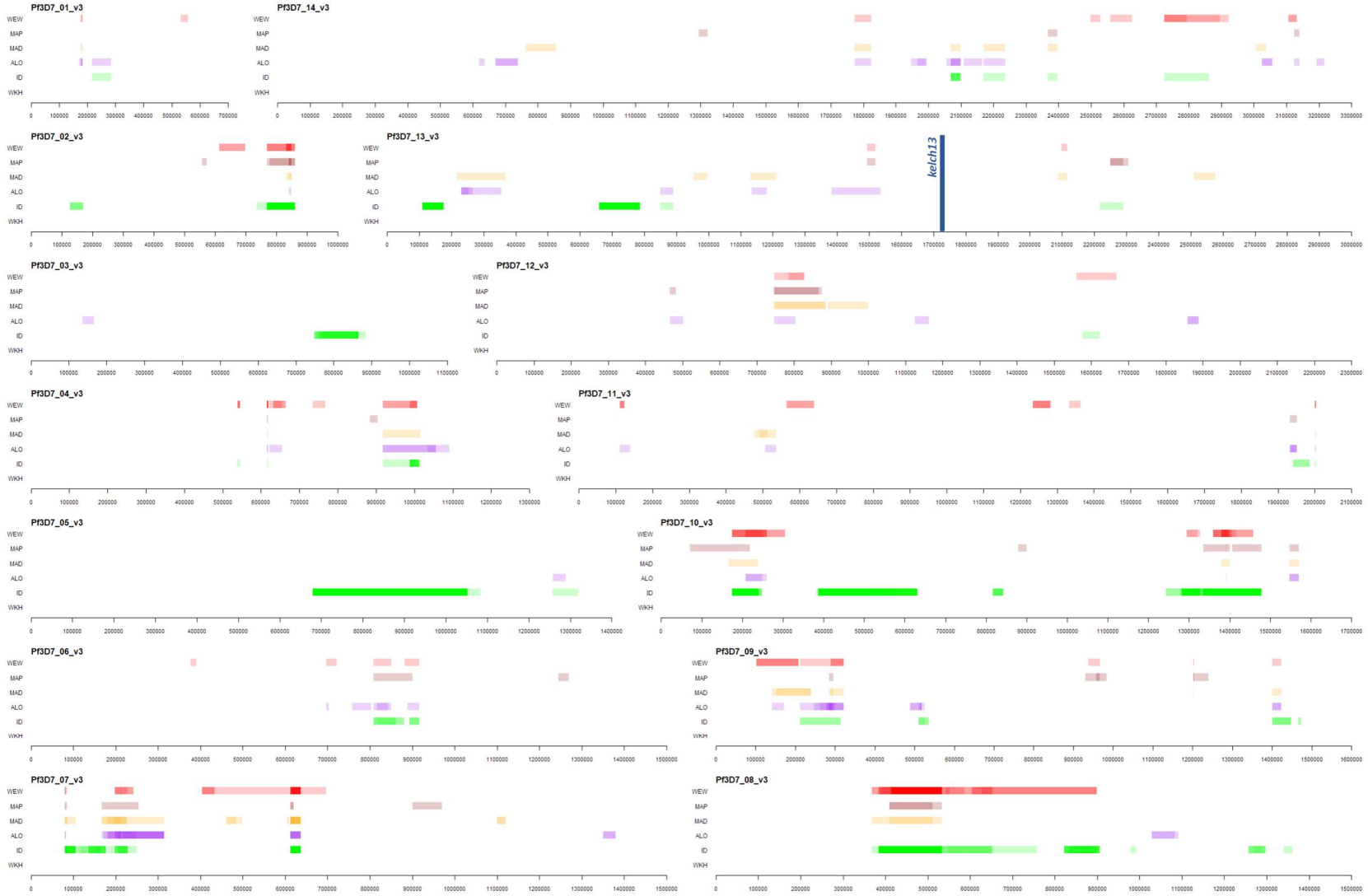
Supplementary Figure 5 - Predicted levels of IBD in the analyzed sample set.
(A) Boxplot showing the distribution of the IBD proportion between the Wewak C580Y genome (PNG-C580Y-1 sample) and those of samples in the populations included in this analysis. Practically no IBD is observed with the Cambodian samples, indicating that a GMS origin is unlikely. (B) Boxplot showing, for each of the populations in New Guinea, the distribution IBD proportions for each pair of samples from the same population.

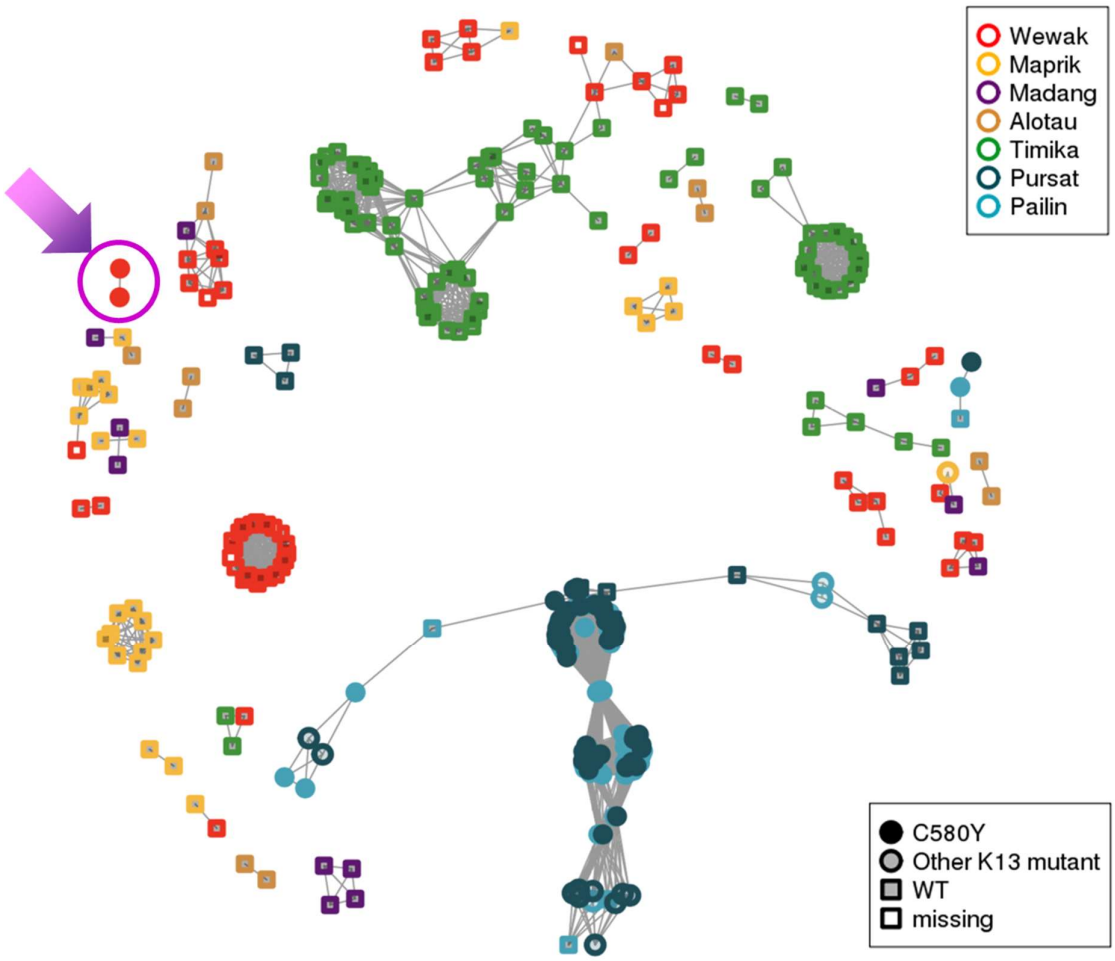


Supplementary Figure 6 - Whole-genome IBD network.
This network diagram uses genome-wide IBD proportions, estimated by IsoRelate, to connect samples into clusters; any pair of samples is connected if they share >=10% IBD across the genome. Large clusters of closely related parasites are found in Cambodia (blue), Timika (green), Wewak (red) and Maprik (yellow), consistent with inbreeding in the these populations. The Wewak C580Y mutants (red filled circles, indicated by a magenta circle and arrow) do not share this level of IBD with any other individual within the dataset, and cluster with none of the large groups.

(Overleaf)

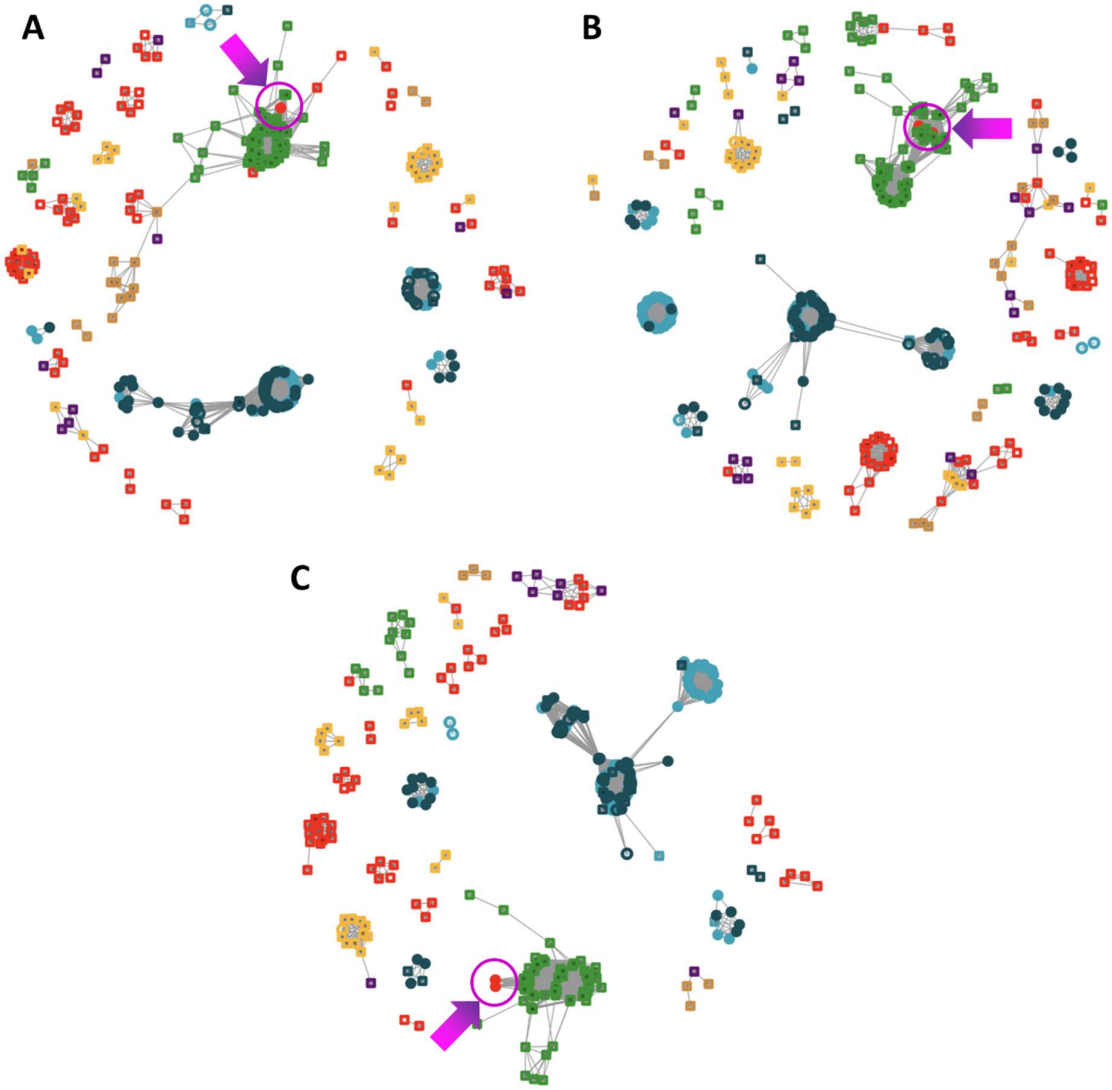
Supplementary Figure 7 – Genome map of IBD segments.
Each of the 14 plots shows the length of one of the nuclear chromosomes. Coloured rectangles are drawn in regions where we predict of the Wewak C580Y genome (PNG-C580Y-1 sample) is IBD with samples from the comparator populations (WEW = Wewak, red; MAP = Maprik, brown; MAD = Madang, orange; ALO = Alotau, purple; ID = Indonesia, green; and WKH = Western Cambodia, blue). IBD regions with more saturated colour are shared with more samples than those in light shades. A blue line indicates the position of the *kelch13* gene on chromosome 13.


Pf3D7_01_v3
Pf3D7_14_v3
Pf3D7_02_v3
Pf3D7_13_v3
kelch13
Pf3D7_03_v3
Pf3D7_12_v3
Pf3D7_04_v3
Pf3D7_11_v3
Pf3D7_05_v3
Pf3D7_10_v3
Pf3D7_06_v3
Pf3D7_09_v3
Pf3D7_07_v3
Pf3D7_08_v3
WEW
MAP
MAD
ALO
ID
WKH



Supplementary Figure 8 - IBD network using genomic segments surrounding the *kelch13* gene.

This network diagram uses IBD predicted by IsoRelate for a region of 150kbp surrounding *kelch13* to connect samples into clusters; any pair of samples is connected if they share any IBD in this region. The Wewak C580Y mutants (red filled circles, indicated by a magenta circle and arrow) do not share IDB with any other sample; crucially, they do not appear to have common ancestry with Cambodian parasites (blue cluster) nor with the Timika parasites (green cluster), unlike some wild-type Wewak samples. Consistent results are obtained if the left and right flanks are analyzed separately.



Supplementary Figure 9 - IBD networks using genomic segments surrounding genes containing genes related to drug resistance.

This network diagram uses IBD predicted by IsoRelate for a region of 150kbp surrounding genes *atg18* (A), *mdr1* (B) and *pnp* (C) to connect samples into clusters; any pair of samples is connected if they share any IBD in this region. Each of these genes contains one or more SNPs with an allele associated with drug resistance that is found in Wewak C580Y parasites, but not elsewhere in PNG. The Wewak C580Y mutants (red filled circles, indicated by a magenta circle and arrow) share IDB with the Timika samples at these loci, but not with Cambodian or other PNG parasites. Cambodian parasites form large groups at these loci, indicating they share haplotypes, but different from those in Indonesian Papua.

References

1. Hetzel MW, Pulford J, Maraga S, et al. Evaluation of the Global Fund-supported National Malaria Control Program in Papua New Guinea, 2009-2014. *Papua and New Guinea medical journal* 2014; **57**(1-4): 7-29.
2. 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.
3. Hetzel MW, Gideon G, Lote N, Makita L, Siba PM, Mueller I. Ownership and usage of mosquito nets after four years of large-scale free distribution in Papua New Guinea. *Malaria journal* 2012; **11**(1): 192.
4. Reimer LJ, Thomsen EK, Koimbu G, et al. Malaria transmission dynamics surrounding the first nationwide long-lasting insecticidal net distribution in Papua New Guinea. *Malar J* 2016; **15**(1): 25.
5. 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.
6. 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.
7. 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. *The Lancet Infectious diseases* 2018.
8. Miotto O, Amato R, Ashley EA, et al. Genetic architecture of artemisinin-resistant Plasmodium falciparum. *Nature genetics* 2015; **47**(3): 226-34.
9. 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.
10. Witkowski B, Duru V, Khim N, et al. A surrogate marker of piperaquine-resistant Plasmodium falciparum malaria: a phenotype-genotype association study. *The Lancet Infectious diseases* 2017; **17**(2): 174-83.
11. 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.