Temporal and spatial dynamics of *Plasmodium falciparum* clonal lineages in Guyana

Mathieu Vanhove[1,2*], Philipp Schwabl[1,2], Colette Clementson[3], Angela M. Early[1,2], Margaret Laws[1,2], Frank Anthony[3], Célia Florimond[4], Luana Mathieu[4], Kashana James[3], Cheyenne Knox[1,2], Narine Singh[3], Caroline O. Buckee[1], Lise Musset[4], Horace Cox[3,5], Reza Niles-Robin[3], Daniel E. Neafsey[1,2]

[1] Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, MA, USA
[2] Infectious Disease and Microbiome Program, Broad Institute of MIT and Harvard, Cambridge, MA, USA
[3] National Malaria Program, Ministry of Health, Georgetown, Guyana
[4] Laboratoire de parasitologie, World Health Organization Collaborating Center for Surveillance of Antimalarial Drug Resistance, Center Nationale de Référence du Paludisme, Institut Pasteur de la Guyane, Cayenne, French Guiana
[5] Caribbean Public Health Agency, Trinidad and Tobago

Corresponding author: mathieu.vanhove@gmail.com

Keywords: artemisinin resistance; *pfk13*; *Plasmodium falciparum*; malaria; South America; Guyana;

Abstract

Plasmodium parasites, the causal agents of malaria, are eukaryotic organisms that obligately undergo sexual recombination within mosquitoes. However, in low transmission settings where most mosquitoes become infected with only a single parasite clone, parasites recombine with themselves, and the clonal lineage is propagated rather than broken up by outcrossing. We investigated whether stochastic/neutral factors drive the persistence and abundance of *Plasmodium falciparum* clonal lineages in Guyana, a country with relatively low malaria transmission, but the only setting in the Americas in which an important artemisinin resistance mutation (*pfk13* C580Y) has been observed. To investigate whether this clonality was potentially associated with the persistence and spatial spread of the mutation, we performed whole genome sequencing on 1,727 *Plasmodium falciparum* samples collected from infected patients across a five-year period (2016-2021). We characterized the relatedness between each pair of monoclonal infections (n=1,409) through estimation of identity by descent (IBD) and also typed each sample for known or candidate drug resistance mutations. A total of 160 clones (mean IBD ≥ 0.90) were circulating in Guyana during the study period, comprising 13 highly related clusters (mean IBD ≥ 0.40). In the five-year study period, we observed a decrease in frequency of a mutation associated with artemisinin partner drug (piperaquine) resistance (*pfcrt* C350R) and limited co-occurence of *pfcrt* C350R with duplications of *plasmepsin 2/3*, an epistatic interaction associated with piperaquine resistance. We additionally report polymorphisms exhibiting evidence of selection for drug resistance or other phenotypes and reported a novel *pfk13* mutation (*G718S*) as well as 61 nonsynonymous substitutions that increased markedly in frequency. However, *P. falciparum* clonal dynamics in Guyana appear to be largely driven by stochastic factors, in contrast to other geographic regions.

The use of multiple artemisinin combination therapies in Guyana may have contributed to the disappearance of the *pfk13* C580Y mutation.

Author Summary

Malaria is caused by eukaryotic *Plasmodium* parasites, which undergo sexual recombination within mosquitoes. In settings with low transmission, such as Guyana, these parasites often recombine with themselves, leading to the propagation of identical clones. We explored the population genomics of *Plasmodium falciparum* malaria parasites in Guyana over five years to characterize clonal transmission dynamics and understand whether they were influenced by local drug resistance mutations under strong selection, including *pfk13* C580Y, which confers resistance to artemisinin, and *pfcrt* C350R, which confers resistance to piperaquine. Using whole genome sequencing on 1,463 samples, we identified 160 clones, in which all parasites share at least 90% of their genomes through recent common ancestry. We observed a decrease in frequency of the *pfcrt* C350R mutation, as well as the disappearance of *pfk13* C580Y. Our findings contrast with the deterministic rise of drug resistance mutations observed in other geographic regions, sometimes associated with clonality. The simultaneous use of at least two different artemisinin combination therapies may have prevented the spread of an artemisinin-resistant clone in Guyana, suggesting a strategy for resistance management in other geographic regions.

Introduction

Genomic data from pathogens, vectors, and/or human hosts can complement traditional epidemiological data on disease incidence and prevalence to inform decisions regarding control. In the case of malaria, several distinct use cases for genomic epidemiology have been previously identified [1], including the identification of imported cases and transmission hotspots [2,3], as well as informing strategies for local disease elimination by documenting connectivity among parasite populations mediated by human movement [4]. Most importantly, genomic data from malaria parasites can play an important role in surveillance of emerging drug resistance markers [5]. Resistance has arisen to every widely deployed antimalarial [6], and molecular surveillance has been endorsed by the WHO as a core intervention for maintaining the efficacy of current malaria drug treatment regimens [7].

Genetic surveillance of drug or insecticide resistance is typically conducted using genotyping data from specific polymorphisms associated with resistance [4,8]. However, whole genome sequencing (WGS) data and genome-wide genotyping assays can inform understanding of the context for the origin and spread of mutations, especially in cases where compensatory or epistatic mutations are required to generate a high-fitness resistance genotype capable of spreading quickly [9]. While measurable phenotypic resistance may be conferred by individual mutations, other genomic changes are often required for those mutations to be evolutionarily successful, with examples in *Plasmodium* malaria parasites [10–12], bacteria [13] and other pathogens [14].

Resistance has been arising in a small number of specific geographic locations to artemisinin, which is administered with one or more partner drugs as artemisinin combination therapy (ACT) as the first line treatment for malaria caused by *Plasmodium falciparum* in most of the world. Delayed parasite clearance following ACT treatment was first observed in the Greater Mekong

Subregion (GMS) of Southern Asia in early 2000s [15,16]. More recently, mutations associated with reduced susceptibility to artemisinin have also been detected in East Africa [17–21] and Papua New Guinea [22]. The most important artemisinin resistance mutation, a C to Y substitution at codon 580 (C580Y) in the propeller domain of a kelch-domain-containing protein on chromosome 13 (*pfk13*) was first observed in the Americas in samples collected in Guyana in 2010, where five out of 94 symptomatic cases were found to carry the *pfk13* C580Y mutation [23]. In 2014, a therapeutic efficacy study (TES) from Guyana failed to detect clinical artemisinin resistance [24], but sample size was likely too low to recruit subjects with low-frequency resistance mutations. *The pfk13* C580Y mutation was observed in 14 out of 854 clinical samples in a resistance surveillance study conducted in Guyana from 2016-2017, and through whole genome sequencing we determined that all of these samples represented a single clonal parasite lineage, despite being observed in disparate regions of the country [25].

This observation of a single clonal background for the *pfk13* C580Y mutation in Guyana was unexpected because *P. falciparum* is a eukaryotic parasite that undergoes sexual recombination in mosquitoes as an obligatory component of its life cycle. However, when a mosquito bites a human host with a monoclonal infection (caused by a single parasite genomic lineage), parasites do not have an opportunity to undergo sexual outcrossing in the mosquito, and instead perform selfing, resulting in the perpetuation of the genomic lineage present in the previous human host. Malaria transmission levels are low in Guyana relative to many settings in sub-Saharan Africa, and therefore most infections are monoclonal, resulting in frequent clonal transmission. Therefore, a null hypothesis to explain the observation of *pfk13* C580Y on a single clonal background could simply invoke low transmission in Guyana as a causal mechanism.

However, a plausible alternative hypothesis is that the *pfk13* C580Y mutation was observed on a single clonal background because that genomic lineage contained important compensatory or epistatic mutations, related to the phenotype of artemisinin resistance directly or resistance to one or more partner drugs commonly administered in ACTs. Historically, resistance to antimalarials has originated *de novo* in low-transmission settings like Southeast Asia or the Americas and has only later spread to sub-Saharan Africa where malaria is much more common [26], leading to the hypothesis that low sexual outcrossing rates in such settings could facilitate the emergence of high-fitness resistance genotypes by preserving key combinations of alleles (in addition to factors such as lower immunity and higher drug pressure). Clonality has been associated with the emergence of *pfk13* C580Y *P. falciparum* in Cambodia and its subsequent spread throughout the GMS [16,26], perhaps facilitated by an additional mutation in this lineage (*plasmepsin 2* and/or *plasmepsin 3* gene amplification) that confers resistance to an important artemisinin partner drug. In East Africa, studies from Uganda [19] and Eritrea [27] reported evidence of emergence of resistance through clonal propagation with an increase in prevalence of *pfk13* mutations.

The official first-line for malaria in Guyana is the ACT artemether-lumefantrine (AL), and no lumefantrine (LMF) resistance mutations are known to be segregating in Guyana *P. falciparum* populations. However, an important context for malaria transmission in Guyana is among gold-miners working in forested regions who are known to frequently self-medicate with the ACT dihydroartemisinin (DHA) -piperaquine (PPQ) -trimethoprim (TMP; DHA + PPQ + TMP; Artecom) tablets [28,29]. At least two mutations that are segregating in Guyana confer resistance to piperaquine: a point C350R mutation in the chloroquine resistance transporter (*pfcrt*) gene that is endemic to the Guiana Shield region and has been increasing in frequency over the last 20 years [28,30,31], and copy number amplification of the *plasmepsin 2* (*Pfpm2* - PF3D7_1408000) and/or

111 *plasmepsin 3* (*Pfpm3* - PF3D7_1408100) genes [30]. The *pfcrt* C350R mutation and *plasmepsin*
112 2/3 amplifications interact epistatically to yield piperaquine resistance [32], adding credibility to
113 the hypothesis that clonal transmission may be adaptive under DHA-piperaquine pressure.
114 In the present study we generated whole genome sequencing data from *P. falciparum* clinical
115 samples collected in Guyana between 2016-2021 to profile the temporal and spatial dynamics of
116 clonal parasite lineages. We identify circulating clonal components (referred to as clones), defined
117 as groups of genomically indistinguishable parasites identified under a graph-based framework
118 [33], and we explore whether limited sexual outcrossing may have been conducive to the *de novo*
119 origin of the *pfk13* C580Y mutation in Guyana. We specifically explore the representation of *pfk13*
120 C580Y, *pfcrt* C350R, and *Pfpm2/3* gene amplifications in clonal and unique parasite genomic
121 backgrounds, and their co-occurence in frequency vs. rare clonal lineages. Further, we profile new
122 signatures of selection in the local parasite population using this deep population genomic dataset
123 to determine whether previously uncharacterized mutations may also drive clonal dynamics, or
124 whether the persistence and prevalence of clonal lineages in Guyana are driven by stochastic
125 factors.
126

127 Results

128 **Temporal and Spatial Clonal Dynamics in Guyana**

129 We performed selective whole-genome amplification (sWGA) on 1,727 samples collected
130 from Guyana between 2016 and 2021 across three time periods (Fig. 1). A total of 264 genomes
131 (15.3%) did not meet the quality criteria of at least 30% of the genome covered at ≥ 5-fold
132 coverage, resulting in 1,463 samples suitable for analysis. Of this set, 54 samples were classified

as multiclonal infections ($F_{ws} < 0.7$) and were excluded from subsequent analyses. The final dataset for relatedness analysis contained 1,409 monoclonal genomes (Table S9) with an average pairwise IBD across the entire dataset of 0.283 (SD = 1.114 – Fig. S1). The final dataset obtained was composed of 736 genomes from 2016/2017; 130 genomes from 2018/2019 which were collected as part of a therapeutic efficacy study (TES); and finally 523 genomes from patient samples collected in 2019/2021. To explore patterns of relatedness due to shared recent common ancestry, pairwise identity-by-descent (IBD) values were computed between haploid genotypes (Fig. 2).

Genome-wide mean IBD estimates across samples revealed patterns of shared ancestry. Network analysis identified 160 clones (C), which were defined as groups of at least two samples with a mean pairwise IBD ≥ 90%, and 332 singletons. Some of these clones formed larger highly related clusters, defined as a group of multiple clones (n ≥ 3) which displayed a mean IBD ≥ 40% and grouped together in the hierarchically-clustered dendrogram as portrayed in Fig. 2. A total of 13 highly related clusters were present in Guyana between 2016 and 2021. Cluster 1 was composed of 7 clones and 21 singletons, including the largest component of the study (C#1), which was composed of 73 samples and disappeared in October 2018 (Fig. 3). Most parasite clones in Guyana persisted for a brief time, but others lasted multiple years. The mean duration of clones was under three months (75.0 days), but clones sampled on multiple occurrences (n=160) persisted on average for at least 8.3 months (251 days). Increasing the IBD threshold did not significantly change the number of clones (IBD ≥ 99%: 96 clones; IBD ≥ 95%: 159 clones). Four clones belonging to different highly related clusters persisted throughout the study (C#143, C#100, C#32, C#305; Fig. 3). Of the 160 clones, 138 were sampled over multiple months (≥ two months), 96 over three months, 68 over six months, and 34 over a year. Seven clones were sampled over two

155 years. A total of 69 clones were related to other clones by ≥ 0.40 mean IBD (Fig. S3) and highly
156 related cluster 3 appeared as the most related to other clusters.
157
158 **_pfk13_ C580Y was restricted to a single clonal background.**
159 The C580Y mutation in the *pfk13* gene (PF3D7_1343700) was present only on single clonal
160 background (C#268, Fig. 3) as previously reported by Mathieu et al. (2020) [25]. This clone was
161 composed of six samples and did not carry *pfcrt* C350R. The clone was part of highly related
162 cluster 10 which was composed of six clones and five singletons (Table S1). The clonal
163 background harboring *pfk13* C580Y was related to clone C#270 (n=5) and C#271 (n=2) at mean
164 pairwise IBD levels of 0.45 and 0.42, respectively. On average, clones circulated in 2.46 spatial
165 clusters and for 237.0 days. The *pfk13* C580Y-harboring clonal component C#268, last observed
166 in April 2017, was observed in six locations over 418 days (Fig 4). In terms of clonal persistence,
167 this haplotype was among the top 20% of multi-occurrence clones (n=130). We investigated
168 nonsynonymous (NSY) mutations with a similar allelic frequency (MAF = 0.007 ± 0.05) as *pfk13*
169 C580Y, screening 2,360 NSY mutations for their relative clonal size and clonal persistence (Fig.
170 5). The temporal persistence of the *pfk13* C580Y mutation above the mean clonal duration (t =
171 287.4 days, $p < 0.211$, t_{C580Y}=418.0 days) and clonal size was above average (n_{C580Y}=8.0 $p < 0.211$,
172 n_{mean}=6.0) but below the 95th percentile of polymorphisms in the same frequency class in each
173 case.
174 Two occurrences of a previously undescribed *pfk13* mutation (G718S) were also observed. The
175 two samples were collected the same week in November 2020 in Aranka River in Region 7. They
176 also carried the *pfk13* K189T mutation. These samples belonged to a clonal background (C#321)

177 composed of four samples which was first detected in April 2018, but the other members of this
178 clonal background did not have sufficient coverage at this position to permit allele identification.
179

180 **Decrease in *pfcrt* C350R frequency across the five year study period.**

181 In the *pfcrt* gene (PF3D7_0709000), which encodes a transmembrane digestive vacuole protein
182 known to modulate resistance to chloroquine and other drugs [10], Allelic positions 72, 76, 220,
183 326 (wildtype) and 356 in *pfcrt* were fixed. The frequency of *pfcrt* C350R in the dataset was
184 54.04% (n=709) and was found in 222 clones (Fig. 2). Additionally, six samples harbored a
185 previously undocumented coding polymorphism in *pfcrt:* D329N. The earliest observation of the
186 D329N mutation was obtained in September 2018 and was sampled in Georgetown as part of the
187 TES. The D329N mutation was found in three clones (C#173, C#9, C#402), which were each
188 composed of two samples. These samples exhibited the *pfcrt* C350 wildtype allele and were found
189 in different highly related clusters. Between the two study periods, a change in *pfcrt* C350R
190 frequency was observed. In 2016/17, *pfcrt* C350R was present in 73.3% of samples (n=478) while
191 in 2020/2021, the frequency of the mutation was 36.2% (n=191). This frequency reversion to the
192 wildtype allele could also be observed within a highly related cluster (Table S7). In 2016/17, highly
193 related cluster 6 displayed one predominant clone (C#134) carrying the *pfcrt* C350R mutation. The
194 clone was found primarily in Mid Essequibo but also appeared in six other spatial clusters (Fig. 3
195 & S4, Table S8). In 2020/21, samples in this highly related cluster return to the wildtype (in C#137,
196 C#135, C#136 and C#138). These samples were still widely distributed, with C#137 occuring in
197 nine spatial clusters (Fig. 6). Evidence of multiple events of *pfcrt* C350R mutation was observed.
198 The wildtype and *pfcrt* C350R were both observed in nine clones. For instance, C#45 contained

four samples with *pfcrt* C350R and seven representing the wildtype (Fig. 6). The clone C#268 harboring the *pfk13* C580Y did not carry *pfcrt* C350R.

When investigating whether *pfcrt* C350R had an impact on clonal persistence or clonal size, no significant difference was observed. The average duration of clones carrying the mutation was 268.0 days ($p < 0.810$), while the average duration of clones representing other nonsynonymous (NSY) mutations of comparable allele frequency (MAF = 0.46 ± 0.05) was 291.0 days (Fig. 5C, n = 683 NSY mutations). The size of clones harboring *pfcrt* C350R (mean = 6.9, $p < 0.526$) was also similar to the size of clones representing these comparator mutations (mean size of 6.9 samples) (Fig. 5D). Mutations significantly associated with prolonged clonal duration included one mutation in falcilysin gene (PF3D7_1360800, n= 134.5 days, $p < 0.001$) as well as two mutations in PF3D7_1133400 (AMA1 - apical membrane antigen 1, n=156.4/150.0 days, $p < 0.001$). Comparator NSY mutations associated with elevated average clone size also included variants in AMA1 (PF3D7_1133400, n=4.6 samples) and MSP1 (PF3D7_0930300, n=4.3 samples).

Co-occurrence of *plasmepsin* 2/3 duplication and *pfcrt* C350R

The combination of *pfcrt* C350R and *plasmepsin 2/3* copy number amplification has been recently demonstrated to confer piperaquine resistance in the Guiana Shield [30]. *Pfpm2/3* copy number status was recovered from 62 samples in 2016/17 (8.0%) [30]. Although the information was only available for a limited number of samples, a X^2 test revealed a significant association between *pfpm2/3* copy number and C350R ($P < 0.035$). We assessed *pfpm2/3* copy number for 401 samples in the 2020/21 dataset. A total of 96 (15.2%) samples possessed *pfcrt* C350R in combination with an increase in *pfpm2/3* copy number while 87 samples (21.7%) exhibited *pfcrt* C350R with a single copy of *pfpm2/3*. Ninety-six out of 245 (39.2%) wildtype samples presented

multiple copies of *pfpm2/3*. No significant association between *pfpm2/3* copy number and *pfcrt* C350R was observed during the 2020/21 period ($X^2 = 0.89$, $P < 0.64$). Highly related clusters appeared to carry the *pfcrt* C350R mutation heterogeneously. Highly related clusters with more than two samples displayed a frequency of *pfpm2/3* copy number variation of 37.8% (Fig. 7). Only four out of the 60 clones investigated carried the duplication homogeneously highlighting the genomic lability of this duplication.

While the frequency of *pfcrt* C350R decreased between 2016 and 2021, two highly related clusters exhibited a frequency increase and were predominantly carrying the mutation (Table S7). In cluster 4, where 61 samples were observed in 2020/21 (with only 5 samples observed across 2016/17), 43 samples (70.5%) carried *pfcrt* C350R. Among the 33 samples with *pfcrt* C350R which were tested, 14 (42.4%) displayed multiple *pfpm2/3* copies of number, whereas 19 had a single copy. In cluster 9, which was only observed in 2020/21, 37 samples (94.9%) carried *pfcrt* C350R (Fig. S8). In this cluster, 31 samples were tested for *pfpm2/3* copy of number and only 4 samples (12.9%) harbored both *pfcrt* C350R and multiple *pfpm2/3* copies.

Mutations in drug resistance genes

The rise of drug resistance polymorphisms has the potential to drive clonal dynamics. In *mdr1* (PF3D7_0523000), mutations were found at positions 1042 (n=951) and 1246 (n=1,061) with only one occurrence of the wildtype allele for each position. At position 1034, 92% (n=844) of samples possessed the double NSY mutations restoring the wildtype serine, while 64 samples carried the cysteine. Two samples displayed only the non-synonymous mutation on the second codon resulting in a threonine. In *dhfr*, positions 50, 51 and 108 (PF3D7_0417200) were monomorphic while in *dhps* (PF3D7_0810800), mutations at positions 540 (n=1,318) and 581 (n=1,319) were

near-fixed with 25 and 18 samples displaying the wildtype respectively. A new mutation in *plasmepsin 2*, G442H, was observed in 12 samples found in six clones (C#385, n=4; C#216, n=10; C#218 n=5, C#219, n=2, C#325, n=2 and C#320, n=1) and observed among different clusters.

Shift in the selection landscape in Guyana.

We searched for evidence of temporal changes in natural selection by observing the changes in allele frequencies between two time periods: 2016-2017 and 2020-2021. The frequency of highly related clusters 3, 4, 6 and 11 increased while the other clusters decreased or remained stable (Table 1 - Fig. S2 & S5). These changes were associated with the rise of NSY mutations and 61 NSY mutations spread across 41 genes, which were in the 99th percentile of change in frequencies (Table 2). These mutations included *pfk13* K189T, which increased in frequency from 34.4% (n=185) in 2016/2017 to 68.4% (n=214) in 2020/2021. The *pfkic6* Q1680K mutation in PF3D7_0609700 (*pfkic6*) encoding a Kelch13 Interacting Candidate was present in emerging clusters at 73.0% (n=) in 2020/2021 for cluster 6 and 66.7% in cluster 4 (Fig. S6A). A similar increase of K308E in PF3D7_1344000 encoding an aminomethyltransferase on chromosome 13 was observed (Fig. S6B). In highly related cluster 6, the clones present in 2020/21 harbored four mutations in PF3D7_1346400 (*pfvps13*) which were absent in C#134. The gene encodes a gametocyte-specific protein and the frequency of the H3221N mutation increased from 12.0% (n=33) to 48.0% (n=120). The I10F mutation in the FNL (falcilysin) gene on chromosome 13 (PF3D7_1360800) also increased in frequency. In PF3D7_0701900, a *Plasmodium* exported protein, six NSY mutations were observed (Table S2). Three NSY mutations in transcription factor *pfap2-g5* (PF3D7_1139300) showed a large increase in frequency: Q2468H increased from 28.3% (n=160) to 63.9% (n=205), G1901S from 48.8% (n=268) to 83.2% (n=268), and T526S from 46.5%

(n=288) to 82.8% (n=270). PF3D7_0704000 encoding for conserved *Plasmodium* membrane protein also showed a NSY mutation which increased in frequency (33.7 (n=227) to 64.7% (n=337)).

The selection landscape of these two periods 2016/2017 and 2020/2021 was also investigated using isoRelate [34] to detect genomic regions exhibiting enhanced relatedness, with a false discovery rate of 0.01. The analysis was run on clones sampled across more than three months in 2016/2017 and in 2020/2021 ($n_{2016/2017}$ = 55 clones, $n_{2020/2021}$ = 49 clones). Selection signals (relatedness peaks) were consistent across different analysis runs with different representative samples of each clonal component (Fig. S7). This allowed investigation of signals of positive selection within "successful" clones to understand whether genes present on genomic segments within these clones were particularly important in Guyana. In 2016/2017, seven segments (159 genes) contained within four strong selection signals on chromosomes 2, 4, 7 and 9 were identified among long-lasting clones (see Fig. 6 and see supplementary results for details). On chromosome 9 (chr9: 61,342-208,725 and 318,311-432,047), the selection signal found in long-lasting clones in 2016/2017 was observed in both short- and long-lasting clones in 2020/2021. Nine genes (eleven mutations) presented a large increase in mutation frequencies ($\mathbf{\Delta}_{\text{FREQUENCY}} \geq 28.4\%$, Table 3) in the past five years: PF3D7_0902400 and PF3D7_0902500 (two serine/threonine protein kinase part of the FIKK family gene), PF3D7_0903300 (unknown function), PF3D7_0904200 (PH domain-containing protein), PF3D7_0905500 (unknown function).

Table 1 – Change in highly related cluster frequencies between 2016/2017 and 2020/2021

Study Period	**2016/2017**	**2020/2021**

Highly related cluster	IBD	Number of samples	Frequency in 2016/17 (%)	IBD	Number of samples	Frequency of 2020/21 (%)
1	0.867	88	11.8	0.957	2	0.4
2	0.627	97	13.0	0.545	38	7.2
3	0.708	25	3.4	0.534	107	20.2
4	0.886	5	0.7	0.722	63	11.9
5	0.763	64	8.6	-	0	0.0
6	0.966	23	3.1	0.841	37	7.0
7	0.476	124	16.7	0.409	49	9.3
8	0.614	54	7.3	0.547	13	2.5
9	0.578	3	0.4	0.566	39	7.4
10	0.512	18	2.4	0.424	2	0.4
11	0.485	8	1.1	0.541	32	6.0
12	0.405	59	7.9	0.997	2	0.4
13	0.525	51	6.9	0.421	37	7.0
Other	0.292	125	16.8	0.299	108	20.4
Total	0.294	744	100.0	0.308	529	100.0

Table 2 – Polymorphism which increased in frequency between 2016/2017 and 2020/2021

Gene	**Chr:Position**	**Number of NSY in 99th percentile**	**Description**	**Codons**	**AA**
PF3D7_0113800	1:527107-536351	3	DBL containing protein, unknown function	3436G>A; 3445C>T; 8458G>A	Glu1146Lys; His1149Tyr; Val2820Ile
PF3D7_0216800	2:696193-699561	3	TMEM121 domain-containing protein, putative	1381T>A; 1373G>A; 938A>G	Cys461Ser; Ser458Asn; Gln313Arg
PF3D7_0418600	4:834394-840429	3	regulator of chromosome condensation, putative	2254T>G; 2255G>T; 2270C>T	Gly748Asp; Cys752Gly; Cys752Phe
PF3D7_0609000	6:370261-388494	1	nucleoporin NUP637, putative	12682G>C	Glu4228Gln
PF3D7_0609700	6:413652-419781	1	protein KIC6	5038C>A	Gln1680Lys
PF3D7_0701900	7:79890-82943	4	Plasmodium exported protein, unknown function	2654T>A; 2091T>A ;	Ile885Lys; Asp697Glu;

				2075G>A ; 2013T>A	Ser692Asn; Asn671Lys
PF3D7_0704000	7:167489-177295	1	conserved Plasmodium membrane protein, unknown function	2220A>C	Glu740Asp
PF3D7_0723800	7:993232-1000233	1	apicomplexan kinetochore protein 1, putative	2129A>T	Glu710Val
PF3D7_0828200	8:1216271-1220716	1	leucine--tRNA ligase, putative	3201A>T	Lys1067Asn
PF3D7_0831600	8:1358314-1363618	1	cytoadherence linked asexual protein 8	4099A>G	Lys1367Glu
PF3D7_0902400	9:106514-108406	1	serine/threonine protein kinase, FIKK family	58T>C	Tyr20His
PF3D7_0902500	9:109334-111329	1	serine/threonine protein kinase, FIKK family	865T>C	Cys289Arg
PF3D7_0903300	9:140977-150588	2	conserved Plasmodium membrane protein, unknown function	2518A>T; 4167A>T	Ile840Phe; Lys1389Asn
PF3D7_0904200	9:197867-198816	1	PH domain-containing protein, putative	465C>A	Asn155Lys
PF3D7_0904300	9:199051-207458	1	conserved protein, unknown function	3445T>A	Tyr1149Asn
PF3D7_0904600	9:212515-217983	1	ubiquitin specific protease, putative	2032A>G	Ile678Val
PF3D7_0905300	9:251353-269709	1	dynein heavy chain, putative	2182A>T	Asn728Tyr
PF3D7_0905500	9:278369-279088	1	conserved Plasmodium protein, unknown function	65C>G	Ala22Gly
PF3D7_1001600	10:86538-89009	1	alpha/beta hydrolase, putative	1676C>T	Ala559Val
PF3D7_1004400	10:207000-209642	2	RNA-binding protein, putative	776C>G; 575A>G	Ala259Gly; Asn192Ser
PF3D7_1005300	10:233285-234473	2	conserved Plasmodium protein, unknown function	759A>T; 11A>T	Glu253Asp; Asn4Ile
PF3D7_1005400	10:234795-235745	1	conserved Plasmodium protein, unknown function	225T>A	Asn75Lys
PF3D7_1107300	11:301283-311287	1	polyadenylate-binding protein-interacting protein 1, putative	9430A>T	Ile3144Phe
PF3D7_1129300	11:1131294-1136834	1	conserved Plasmodium protein, unknown function	1477G>A	Gly493Arg
PF3D7_1138400	11:1501011-1513691	1	guanylyl cyclase	11060A>G	Tyr3687Cys
PF3D7_1139100	11:1548186-1553561	1	RNA-binding protein, putative	4289G>A	Arg1430Lys
PF3D7_1139300	11:1556744-1565045	3	transcription factor with AP2 domain(s)	7404A>C; 5701G>A; 1577C>G	Gln2468His; Gly1901Ser; Thr526Ser
PF3D7_1342200	13:1660812-1663730	1	conserved Plasmodium membrane protein, unknown function	2902G>A	Asp968Asn
PF3D7_1343700	13:1724817-1726997	1	kelch protein K13	566A>C	Lys189Thr

PF3D7_1344000	13:1759466-1761991	1	aminomethyltransferase, putative	922A>G	Lys308Glu
PF3D7_1344100	13:1764190-1766607	1	krox-like protein, putative	2170C>G	Gln724Glu
PF3D7_1346400	13:1852898-1870864	1	VPS13 domain-containing protein, putative	9661C>A	His3221Asn
PF3D7_1346700	13:1876016-1877362	2	6-cysteine protein	1307C>T; 940T>A	Thr436Ile; Leu314Ile
PF3D7_1346800	13:1878875-1880194	2	6-cysteine protein	532A>G; 203C>T	Ile178Val; Thr68Met
PF3D7_1360800	13:2435343-2438924	1	falcilysin	28A>T	Ile10Phe
PF3D7_1417700	14:750330-751684	1	conserved Plasmodium protein, unknown function	364G>T	Asp122Tyr
PF3D7_1418900	14:783078-785819	1	ATP-dependent RNA helicase DBP4, putative	2135T>C	Val712Ala
PF3D7_1419400	14:804425-811369	3	conserved Plasmodium membrane protein, unknown function	5150G>T; 2898T>A; 2605G>A	Gly1717Val; Asn966Lys; Asp869Asn
PF3D7_1453600	14:2199371-2204708	2	RAP protein, putative	3689A>T; 3669T>G	Lys1230Ile; Asn1223Lys
PF3D7_1474200	14:3023852-3038832	1	conserved Plasmodium membrane protein, unknown function	9554C>T	Ser3185Leu
PF3D7_1478100	14:3216275-3217794	1	Plasmodium exported protein (hyp13), unknown function	721A>G	Ile241Val

Discussion

In this study, we profiled the dynamics of *P. falciparum* over a five year study period using the deepest whole genome sequencing dataset yet produced for this parasite species from a single country. In contrast to the GMS and East Africa, where clonal transmission and enhanced population relatedness were directly related to the emergence of mutations conferring resistance to ACTs [27,35], Guyana offers a different perspective on clonal dynamics in the context of drug resistance emergence in a low-transmission setting. Stochastic processes with intermittent recombination appear to be the dominant mechanism driving clonal diversity rather than a selective advantage obtained from particular polymorphisms favoring a specific clonal background.

Impact of artemisinin on clonal dynamics in Guyana

Resistant lineages can circulate at low frequencies for years before becoming dominant. In this study, a total of 160 clones aggregated into 13 highly related clusters were observed. Two highly related clusters present at the beginning of the study disappeared by 2020, while four highly related clusters increased in frequency (Table 1). Malaria transmission in the Guyana shield is largely driven by mobile populations working in gold mining or other forest-associated professions [36]. Evidence of clonal dispersal among spatial clusters was best represented by highly related cluster 6, which has two clones spreading to seven and nine spatial clusters in a limited amount of time (Fig. 4). In 2016/2017, one clonal component (C#134) dominated and was preferentially found in Mid Essequibo. By 2021, the former clonal component had disappeared and a related clone (C#137) appeared to be circulating predominantly in Potaro. A cautionary note regarding *Pf* sampling is needed as this dataset was assembled through different sampling schemes performed at different health centers. Parasite origin inferred from patient travel recollection may not be consistently precise and regions with high mining activities might be over represented (Fig. S4).

In 2004, Guyana was the first country on the continent to implement artemether-lumefantrine (COARTEM®). Therefore, constant artemisinin pressure and shifting exposure to lumefantrine, piperaquine, and perhaps other partner drugs has imposed heterogeneous selective pressure on *P. falciparum* lineages. However, only limited evidence of allelic change responding to artemisinin drug pressure has been observed. A *pfkic6* (Kelch13 Interacting Candidate; PF3D7_0609700) Q1680K polymorphism increased by 31.65% (Table S3) in the five year study period. The NSY mutation was present in clones which persisted longer than average (Δ=31.8, $p < 0.001$) and *pfkic6* is a gene which could potentially play a role in artemisinin (ART) resistance given its association

with the resistance-associated PfK13 protein [37,38]. Other polymorphisms that appeared to be favored in the Guyana landscape were associated with potential resistance to artemisinin (Table 2) and their prevalence should be closely monitored, but polymorphisms driving resistance in other regions did not show any obvious signs of selection.

Selection by artemisinin partner drugs

The emergence of drug resistance in the Guyana shield is of concern, considering that resistance to chloroquine and sulfadoxine-pyrimethamine emerged almost simultaneously and in an independent manner in both South America and Southeast Asia [11,39]. In Guyana, 54% of gold miners self-medicate to treat fever using Artecom (DHA+PPQ+TMP) tablets before seeking care [40]. The association of the *pfcrt* C350R allele with an amplification of plasmepsin (*xpfpm2/3*) has been shown to strengthen resistance phenotypes to piperaquine [28,30]. In the current study, we observed a reduction in frequency of *pfcrt* C350R from 73.0% to 24.2% across five years indicating a potential reduction in piperaquine pressure. We can speculate that a change in dominant ACT therapy from DHA+PPQ+TMP to artemether-lumefantrine could have occurred. Erratic use of DHA+PPQ+TMP during a period of high prevalence of the *pfcrt* C350R mutation could have contributed to the emergence of the *pfk13* C580Y mutation. Subsequent increase in the use of artemether-lumefantrine may have reduced the pressure to maintain *pfcrt* C350R, and eliminated the *pfk13* C580Y mutation primarily through success of the partner drug.

Further potential evidence for reduced DHA+PPQ+TMP self-medication in recent years is the reduced prevalence of parasite genomes containing both *pfcrt* C350R and plasmepsin duplication. In Southeast Asia, an increase in copy number of the plasmepsin 2 (*pfpm2*) and/or plasmepsin 3 (*pfpm3*) genes is associated with piperaquine resistance [35,41]. These copy number amplifications

have been observed to enhance piperaquine resistance *in vitro* through epistatic interaction with the *pfcrt* C350R mutation [30]. As observed in French Guiana [30], we found multiple mutational events for *pfcrt* C350R occurring within a short timespan. *Plasmepsin* duplication was also highly genomically labile, varying within and among conserved clonal lineages (Fig. 7). The gene amplification appeared more variable compared to the emergence of polymorphism. The frequency of these phenomena unique to this part of the world make it difficult for a clone to thrive to the extent observed in the GMS. Gene copy number may appear as a strategy for regulating expression under environmental stresses [42]. In this context, the plasticity of *pfpm2/3* might reflect a more rapid adjustment of the parasite responding to heterogeneous drug exposure. For instance, in highly related cluster 4, the dominant highly related cluster circulating in Lower Mazaruni in 2020/21, 14 of45 samples displayed both *pfcrt* C350R and *xpfpm2/3*, which might reflect localized recent selection by piperaquine.

Other candidate variants associated with clonal dynamics

Although this study primarily attributes recent spatiotemporal dynamics of parasite clones in Guyana to stochastic processes (e.g., sporadic outcrossing, periods of low-transmission bottleneck) rather than to selection towards the preservation of specific multi-locus haplotypes, we do not suggest that meaningful selective processes are entirely absent. For instance, the clonal background containing *pfk13* C580Y was observed in six spatial clusters across 418 days, where the average clone was found in 2.36 spatial clusters and lasted on average 287 days (Fig. 4). It is therefore possible that the *pfk13* C580Y mutation improved clone fitness for a period of time. We also noted the previously unobserved *pfk13* G718S mutation in C#321, further sign of autochthonous *pfk13* polymorphism in Guyana (Fig. 3). Furthermore, we observed persistent and

large clones carrying two NSY mutations in AMA1 as well as a NSY mutation (MAF = 0.46 ± 0.05) in falcilysin (PF3D7_1360800) (Table S3-S4). The latter additionally featured among the 61 NSY mutations which increased in frequency between 2016 and 2021. Falcilysin is a metalloprotease believed to be involved in hemoglobin digestion, and has been found to be a target of chloroquine, which inhibits its proteolytic activity [43]. Given that degraded products of hemoglobin activate ART [44], it is possible that this polymorphism interferes with parasite clearance.

The outcrossing rate in Guyana appears to maintain sufficient haplotypic diversity in the population to prevent the long-term dominance of specific clones. However, four clones were sampled over four years, indicating the possibility of longer-term clonal persistence in the region. The selection signal observed at *pfcrt* was conserved throughout the dataset as previously described in global *P. falciparum* populations [34] (Fig. 5). These results suggest that selection may yet be influencing clonal dynamics in Guyana, even if the impact of selection is not as stark as in the GMS [9].

Other NSY mutations which increased in frequency tended to be associated with gametocyte maturation, a process which is key to withstanding artemisinin pressure [45] because artemisinin clears only asexual parasites. Moreover, gametocyte production ultimately determines fitness because they are required for transmission. Three polymorphisms were found in transcription factor *pfap2-g5* (PF3D7_1139300). Apicomplexan-specific ApiAP2 gene family is a well-known regulator of sexual commitment and gametocyte development [46–48]. The gene appears as an important mechanism during the maturation of sexual stages through gene repression combined with other chromatin-related proteins [49]. Transcription factors (AP2 genes) involved in the gametocyte development have been previously found to display the strongest signatures of

396 selection in French Guiana [50]. Seven other genes which increased in frequency are also related
397 to gametocyte development. For instance, PF3D7_0904200 (PH domain–containing protein)
398 transcripts have been shown to be enriched in gametocytes [51] and PF3D7_1474200 was found
399 to be highly expressed in late-stage gametocytes [52].
400

401 **Relevance of *pfK13* C580Y mutation disappearance**

402 Guyana represents the first country where the *pfk13* C580Y mutation (or similar ART resistance
403 mutations) have appeared and then subsequently disappeared rather than increase in frequency.
404 The mutation was restricted to a single clonal background and was last observed in April 2017.
405 This clonal background lacked the *pfcrt* C350R mutation, making it likely susceptible to PPQ,
406 which has been subject to fluctuating use through self-medication in the country and might have
407 led to this disappearance in the presence of efficacious artemether-lumefantrine treatment.
408 Previous therapeutic efficacy studies in the region have hinted at resistance to artemether-
409 lumefantrine [53] and artesunate monotherapy [54] but evidence from TES in Guyana is lacking.
410 A modeling study exploring factors associated with the spread of *pfk13* mutations found that
411 deploying multiple first-line therapies was the best approach to postponing treatment failure [55].
412 The simultaneous use and potentially shifting balance of at least two ACTs in Guyana might have
413 therefore led to the elimination of the *pfk13* C580Y mutation and its clonal background.
414

415 Clonal turnover in Guyana appears to be different from the patterns observed in other regions
416 like South-East Asia and East Africa. In the GMS, artemisinin was initially used as monotherapy
417 facilitating rapid resistance expansion via hard selective sweep [56]. These observations indicate
418 that drug resistance emergence does not result in the same patterns of clonal dynamics in different

419 geographic locations, perhaps due to unique differences in disease epidemiology and drug pressure
420 across settings. Further molecular surveillance of clonal dynamics is warranted in settings where
421 it occurs, given the potential association of clonal transmission with both known and novel
422 mutations associated with drug resistance.
423

424 # Materials and methods

425 ## Sample collection and spatial cluster mapping

426 We evaluated 1,727 clinical samples collected from malaria-diagnosed individuals between
427 2016 and 2021 who provided informed consent for genetic analysis of their parasite samples.
428 Samples were collected as dried blood spots on Whatman FTA cards. Samples dating from 2016-
429 2017 (n=837) were collected for a resistance surveillance project [25]. Samples dating from 2018-
430 2019 (n=174) were collected in the context of a therapeutic efficacy study. Samples dating from
431 2020-2021 (n=716) were collected for a separate malaria molecular surveillance study from
432 individuals diagnosed with *P. falciparum* infection (Fig. 1). Participants provided informed
433 consent in accordance with the ethical regulations of the countries.
434 To define spatial clusters, we first matched travel history responses to a catalog of malaria survey
435 sites used by the Guyana Ministry of Health (MoH). We then mapped survey sites onto a custom
436 shape file summarizing the country's primary river and road coordinates and onto a raster map of
437 motorized transport resistance [57] available at https://malariaatlas.org/. Sites were clustered based
438 on river/road connectivity in the R package 'riverdist' [58], travel conductance using the R package
439 'gdistance' [59], and manual assessment of coordinates on river/road and resistance layers in
440 QGIS.

441 Samples were collected in specific recruitment locations and patient travel history was
442 documented. To investigate spatial patterns, 20 spatial clusters were defined following roads and
443 rivers access (Fig. 1a). Patient travel history revealed that a majority of infections were acquired
444 in Lower Mazaruni River in Region 7 (n=434, 36.1%), followed by Potaro River in Region 8
445 (n=162, 13.5%), as well as along the Cuyuni River (Table S1, S3 and supplementary results for
446 details on highly related clusters dispersal). Travel history data from the Therapeutic Efficacy
447 Study (TES) (n=174) conducted in 2018 and 2019 in Georgetown and Port Kaituma were not
448 recorded. Overall, location data were missing for 216 samples (14.9%).
449

450 **Genomic data generation**

451 DNA extraction was performed using two approaches according to year of collection. Samples
452 from 2016-2017-2018-2019 (n=1,011) were extracted from dried blood spots using the QIAamp
453 DNA mini kit according to the manufacturer's instructions (Qiagen, Hilden, Germany). For
454 samples from 2020-2021 (n=716), we performed DNA extraction on all patient samples using a
455 ThermoFisher blood and tissue kit and a ThermoFisher Kingfisher instrument. We performed
456 selective whole genome amplification (sWGA) [60] on all samples to enrich the proportion of
457 parasite DNA relative to host DNA. We performed library construction using a NEBNext kit on
458 the enriched DNA samples and sequenced them on an Illumina NovaSeq instrument using 150 bp
459 paired-end reads. We aligned reads to the *P. falciparum* 3D7 v.3 reference genome assembly and
460 called variants following the Pf3K consortium best practices
461 (https://www.malariagen.net/projects/pf3k). We used BWA-MEM [61] to align raw reads and
462 remove duplicate reads with Picard tools [62]. We called SNPs using GATK v3.5 HaplotypeCaller
463 [63]. We performed base quality score and variant quality score recalibration using a set of

Mendelian-validated SNPs, and restricted downstream population genomic analyses to SNPs observed in 'accessible' genomic regions determined to be amenable to high quality read alignment and variant calling [64]. Individual calls supported by fewer than five reads were removed and any variant within 5 nucleotides of a GATK-identified indel was also excluded. Samples exhibiting quality monoclonal genome data (>= 5x coverage for >30% of the genome) were included in relatedness analyses. The final dataset to investigate mutation comprised 74,357 SNPs.

Relatedness analysis using identity by descent

We performed analyses of relatedness by estimating pairwise identity by descent (IBD) between all monoclonal patient samples (n=1,409). We estimated IBD using the hmmIBD algorithm [65], incorporating all SNPs that were called in ≥ 90% of samples and with minor allele frequency ≥ 1%, resulting in a final set of 16,806 SNPs [65]. We used the F_{ws} metric (< 0.70) to identify and exclude samples containing multiclonal infections [66]. We conducted subsequent analyses in Python v3.8. We constructed clones using Networkx v.2.8 [67]. clones, defined as groups of statistically indistinguishable parasites identified under a graph theoretic framework [33], were obtained using a mean IBD threshold ≥ 0.90. Highly related clusters were defined as a group of clones (n ≥ 3) which clustered together in the hierarchically-clustered dendrogram (UPGMA algorithm) performed using seaborn v0.13.0 with a threshold of 3 [68] and which also displayed a mean IBD ≥ 0.40. This threshold was chosen based on this specific dataset and because it represents genomes separated by 1-2 recombination events. To identify temporal changes across the sampling period, we investigated NSY SNPs that were in the 99th percentile of change in frequencies. To investigate whether mutations in *pfcrt* were significantly associated with longer

duration or frequency of clones, we selected mutations within ± 0.05 of the minor allele frequency (MAF) of *pfcrt* C350R (MAF = 0.46). We evaluated SNP enrichment in clones with similar duration/frequency as the C350R mutation in pfcrt. Mutations within the 95th percentile were considered as significant.

We investigated signals of selection using the genome wide test statistics ($X_{iR,s}$) in isoRelate v.0.1.0 [34] in R. $X_{iR,s}$ is a chi-squared distribution test statistic for measuring IBD. Briefly, an IBD matrix status with SNPs as rows and sample pairs as columns is created. A normalization procedure is implemented by subtracting the column mean from all rows to account for the amount of relatedness between each pair. Secondly, to adjust for differences in SNP allele frequencies, the row mean is subtracted from each row and divided by $p_i(1-p_i)$, where pi is the population allele frequency of SNP i. Then, row sums are computed and divided by the square root of the number of pairs. Summary statistics are normalized genome wide. To do this, all SNPs are binned in 100 equally sized bins partitioned on allele frequencies. Finally, the mean was subtracted and divided by the standard deviation of all values within each bin. Z-scores were squared to allow only positive values and such that the statistics followed a chi- squared distribution with 1 degree of freedom. We calculated $X_{iR,s}$ and obtained $-\log_{10}$ transformed p-values, and used a false discovery rate threshold of 0.05 to assess evidence of positive selection.

Plasmepsin 2/3 copy number estimation

DNA from selected samples was used for amplification by quantitative PCR (qPCR) to estimate the copy number of plasmepsin 2 and plasmepsin 3 (*pfpm 2/3*) using a previously published protocol that does not distinguish between the two genes [41]. *P. falciparum* tubulin primers (*Pftub*) were used as a single copy comparator locus (forward-5'-

510 TGATGTGCGCAAGTGATCC-3'; reverse-5'-TCCTTTGTGGACATTCTTCCTC-3') and
511 amplified separately from *pfpm* (forward-5'-TGGTGATGCAGAAAGTTGGAG-3'; reverse-5'-
512 TGGGACCCATAAATTAGCAGA-3'). qPCR reactions were carried out in triplicate in 20 µL
513 volumes using 384-well plates (Fisher Scientific, Hampton, NH) using 10 µL SensiFAST SYBR
514 No-ROX mix (2x) (Bioline Inc., Taunton, MA), 300 nM forward and reverse primer, 6.8 µl
515 nuclease-free H2O, and 2 uL DNA template as previously described by [69]. The reactions were
516 performed using the following conditions: initial denaturation at 95 °C for 15 minutes followed by
517 40 cycles at 95 °C for 15 seconds, 58 °C for 20 seconds, and 72 °C for 20 seconds; a melt curve
518 starting at 95 °C for 2 minutes, 68 °C for 2 minutes, followed by increments of 0.2 °C from 68 °C
519 to 85 °C for 0:05 seconds and a final step at 35 °C for 2 minutes. Copy number value was calculated
520 using the 2^-**ΔΔ**Ct method [69]. Means of *pfpm2* and *Pftub* were calculated for 3D7 (a single copy
521 control) using six replicates. Standard deviation should not be more than 25% including all
522 triplicates for the DNA samples. If the value was between 0.6 and 1.5, the copy number is
523 estimated as 1, whereas if the value was between 1.5 and 2.4, the copy number estimated was 2.
524 We use the term *xpfpm2/3* to designate the amplification of *pfpm2* or *pfpm3* and *1pfpm2*/*3* to denote
525 one copy of both genes similarly to [30].

526 Acknowledgements

527
528 We thank the participants who contributed blood samples to this study, as well as the technicians
529 who collected and processed the samples. This work was supported, in whole or in part, by the
530 Bill & Melinda Gates Foundation [INV-009416]. Under the grant conditions of the Foundation, a
531 Creative Commons Attribution 4.0 Generic License has already been assigned to the Author
532 Accepted Manuscript version that might arise from this submission. This study was also
533 supported with federal funds from the National Institute of Allergy and Infectious Diseases,

National Institutes of Health, Department of Health and Human Services, under Grant Number U19AI110818 to the Broad Institute

Data Availability Statement

Illumina-generated short-read sequence data has been deposited in the NCBI Sequence Read Archive under BioProject PRJNA809659.

References

1. Dalmat R, Naughton B, Kwan-Gett TS, Slyker J, Stuckey EM. Use cases for genetic epidemiology in malaria elimination. Malar J. 2019;18: 1–11.
2. Tessema S, Wesolowski A, Chen A, Murphy M, Wilheim J, Mupiri A-R, et al. Using parasite genetic and human mobility data to infer local and cross-border malaria connectivity in Southern Africa. Elife. 2019;8: e43510.
3. Carrasquilla M, Early AM, Taylor AR, Knudson Ospina A, Echeverry DF, Anderson TJ, et al. Resolving drug selection and migration in an inbred South American Plasmodium falciparum population with identity-by-descent analysis. PLoS Pathog. 2022;18: e1010993.
4. Wesolowski A, Taylor AR, Chang H-H, Verity R, Tessema S, Bailey JA, et al. Mapping malaria by combining parasite genomic and epidemiologic data. BMC Med. 2018;16: 1–8.
5. Neafsey DE, Taylor AR, MacInnis BL. Advances and opportunities in malaria population genomics. Nat Rev Genet. 2021;22: 502–517.
6. Haldar K, Bhattacharjee S, Safeukui I. Drug resistance in Plasmodium. Nat Rev Microbiol. 2018;16: 156–170.
7. World Health Organization. A framework for malaria elimination. World Health Organization; 2017.
8. Jacob CG, Thuy-Nhien N, Mayxay M, Maude RJ, Quang HH, Hongvanthong B, et al. Genetic surveillance in the Greater Mekong subregion and South Asia to support malaria control and elimination. Elife. 2021;10: e62997.
9. Amato R, Pearson RD, Almagro-Garcia J, Amaratunga C, Lim P, Suon S, et al. Origins of the current outbreak of multidrug-resistant malaria in southeast Asia: a retrospective genetic study. Lancet Infect Dis. 2018;18: 337–345.
10. Fidock DA, Nomura T, Talley AK, Cooper RA, Dzekunov SM, Ferdig MT, et al. Mutations in the P. falciparum digestive vacuole transmembrane protein PfCRT and evidence for their role in chloroquine resistance. Mol Cell. 2000;6: 861–871.
11. Wootton JC, Feng X, Ferdig MT, Cooper RA, Mu J, Baruch DI, et al. Genetic diversity and chloroquine selective sweeps in Plasmodium falciparum. Nature. 2002;418: 320–323.
12. Cortese JF, Caraballo A, Contreras CE, Plowe CV. Origin and dissemination of Plasmodium falciparum drug-resistance mutations in South America. J Infect Dis. 2002;186: 999–1006.

13. Corona F, Martinez JL. Phenotypic resistance to antibiotics. Antibiotics. 2013;2: 237–255.
14. Berman J, Krysan DJ. Drug resistance and tolerance in fungi. Nat Rev Microbiol. 2020;18: 319–331.
15. Noedl H, Se Y, Schaecher K, Smith BL, Socheat D, Fukuda MM. Evidence of artemisinin-resistant malaria in western Cambodia. N Engl J Med. 2008;359: 2619–2620.
16. Dondorp AM, Nosten F, Yi P, Das D, Phyo AP, Tarning J, et al. Artemisinin resistance in Plasmodium falciparum malaria. N Engl J Med. 2009;361: 455–467.
17. Stokes BH, Ward KE, Fidock DA. Evidence of artemisinin-resistant malaria in Africa. N Engl J Med. 2022;386: 1385.
18. Balikagala B, Fukuda N, Ikeda M, Katuro OT, Tachibana S-I, Yamauchi M, et al. Evidence of artemisinin-resistant malaria in Africa. N Engl J Med. 2021;385: 1163–1171.
19. Conrad MD, Asua V, Garg S, Giesbrecht D, Niaré K, Smith S, et al. Evolution of partial resistance to artemisinins in malaria parasites in Uganda. N Engl J Med. 2023;389: 722–732.
20. Mihreteab S, Anderson K, Molina-de la Fuente I, Sutherland C, Smith D, Cunningham J, et al. The spread of a validated molecular marker of artemisinin partial resistance pfkelch13 R622I and association with pfhrp2/3 deletions in Eritrea. medRxiv. 2023; 2023–10.
21. Juliano JJ, Giesbrecht DJ, Simkin A, Fola AA, Lyimo BM, Perus D, et al. Country wide surveillance reveals prevalent artemisinin partial resistance mutations with evidence for multiple origins and expansion of sulphadoxine-pyrimethamine resistance mutations in northwest Tanzania. medRxiv. 2023; 2023–11.
22. Miotto O, Sekihara M, Tachibana S-I, Yamauchi M, Pearson RD, Amato R, et al. Emergence of artemisinin-resistant Plasmodium falciparum with kelch13 C580Y mutations on the island of New Guinea. PLoS Pathog. 2020;16: e1009133.
23. Chenet SM, Akinyi Okoth S, Huber CS, Chandrabose J, Lucchi NW, Talundzic E, et al. Independent emergence of the Plasmodium falciparum kelch propeller domain mutant allele C580Y in Guyana. J Infect Dis. 2016;213: 1472–1475.
24. Rahman R, Martin MJS, Persaud S, Ceron N, Kellman D, Musset L, et al. Continued sensitivity of Plasmodium falciparum to artemisinin in Guyana, with absence of kelch propeller domain mutant alleles. Oxford University Press; 2016.
25. Mathieu LC, Cox H, Early AM, Mok S, Lazrek Y, Paquet J-C, et al. Local emergence in Amazonia of Plasmodium falciparum k13 C580Y mutants associated with in vitro artemisinin resistance. Elife. 2020;9: e51015.
26. Read A, Huijben S. Evolutionary biology and the avoidance of antimicrobial resistance. Evol Appl 2: 40–51. 2009.
27. Mihreteab S, Platon L, Berhane A, Stokes BH, Warsame M, Campagne P, et al. Increasing prevalence of artemisinin-resistant HRP2-negative malaria in Eritrea. N Engl J Med. 2023;389: 1191–1202.
28. Pelleau S, Moss EL, Dhingra SK, Volney B, Casteras J, Gabryszewski SJ, et al. Adaptive evolution of malaria parasites in French Guiana: reversal of chloroquine resistance by acquisition of a mutation in pfcrt. Proc Natl Acad Sci. 2015;112: 11672–11677.
29. Douine M, Musset L, Corlin F, Pelleau S, Pasquier J, Mutricy L, et al. Prevalence of Plasmodium spp. in illegal gold miners in French Guiana in 2015: a hidden but critical malaria reservoir. Malar J. 2016;15: 1–8.
30. Florimond C, de Laval F, Early AM, Sauthier S, Lazrek Y, Pelleau S, et al. Impact of piperaquine resistance in Plasmodium falciparum on malaria treatment effectiveness in

French Guiana: a descriptive epidemiological study. Lancet Infect Dis. 2023.
31. Mok S, Fidock DA. Determinants of piperaquine-resistant malaria in South America. Lancet Infect Dis. 2023.
32. Kane J, Li X, Kumar S, Button-Simons KA, Brenneman KMV, Dahlhoff H, et al. A Plasmodium falciparum genetic cross reveals the contributions of pfcrt and plasmepsin II/III to piperaquine drug resistance. bioRxiv. 2023.
33. Taylor AR, Echeverry DF, Anderson TJ, Neafsey DE, Buckee CO. Identity-by-descent with uncertainty characterises connectivity of Plasmodium falciparum populations on the Colombian-Pacific coast. PLoS Genet. 2020;16: e1009101.
34. 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: e1007279.
35. Amato R, Lim P, Miotto O, Amaratunga C, Dek D, Pearson RD, et al. Genetic markers associated with dihydroartemisinin–piperaquine failure in Plasmodium falciparum malaria in Cambodia: a genotype–phenotype association study. Lancet Infect Dis. 2017;17: 164–173.
36. De Santi VP, Girod R, Mura M, Dia A, Briolant S, Djossou F, et al. Epidemiological and entomological studies of a malaria outbreak among French armed forces deployed at illegal gold mining sites reveal new aspects of the disease's transmission in French Guiana. Malar J. 2016;15: 1–11.
37. Behrens HM, Schmidt S, Spielmann T. The newly discovered role of endocytosis in artemisinin resistance. Med Res Rev. 2021;41: 2998–3022.
38. Simmons CF, Gibbons J, Zhang M, Oberstaller J, Pires CV, Casandra D, et al. Protein KIC5 is a novel regulator of artemisinin stress response in the malaria parasite Plasmodium falciparum. Sci Rep. 2023;13: 399.
39. McCollum AM, Mueller K, Villegas L, Udhayakumar V, Escalante AA. Common origin and fixation of Plasmodium falciparum dhfr and dhps mutations associated with sulfadoxine-pyrimethamine resistance in a low-transmission area in South America. Antimicrob Agents Chemother. 2007;51: 2085–2091.
40. Olapeju B, Adams C, Hunter G, Wilson S, Simpson J, Mitchum L, et al. Malaria prevention and care seeking among gold miners in Guyana. Plos One. 2020;15: e0244454.
41. Witkowski B, Duru V, Khim N, Ross LS, Saintpierre B, Beghain J, et al. A surrogate marker of piperaquine-resistant Plasmodium falciparum malaria: a phenotype–genotype association study. Lancet Infect Dis. 2017;17: 174–183.
42. Zhang X, Deitsch KW, Kirkman LA. The contribution of extrachromosomal DNA to genome plasticity in malaria parasites. Mol Microbiol. 2021;115: 503–507.
43. Wirjanata G, Dziekan JM, Lin J, Sahili AE, Binte Zulkifli NE, Boentoro J, et al. Identification of an inhibitory pocket in falcilysin bound by chloroquine provides a new avenue for malaria drug development. bioRxiv. 2021; 2021–04.
44. Birnbaum J, Scharf S, Schmidt S, Jonscher E, Hoeijmakers WAM, Flemming S, et al. A Kelch13-defined endocytosis pathway mediates artemisinin resistance in malaria parasites. Science. 2020;367: 51–59.
45. Munro BA, McMorran BJ. Antimalarial drug strategies to target Plasmodium gametocytes. Parasitologia. 2022;2: 101–124.
46. Josling GA, Llinás M. Sexual development in Plasmodium parasites: knowing when it's time to commit. Nat Rev Microbiol. 2015;13: 573–587.

47. Van Biljon R, Van Wyk R, Painter HJ, Orchard L, Reader J, Niemand J, et al. Hierarchical transcriptional control regulates Plasmodium falciparum sexual differentiation. BMC Genomics. 2019;20: 1–16.
48. Poran A, Nötzel C, Aly O, Mencia-Trinchant N, Harris CT, Guzman ML, et al. Single-cell RNA sequencing reveals a signature of sexual commitment in malaria parasites. Nature. 2017;551: 95–99.
49. Singh S, Santos JM, Orchard LM, Yamada N, van Biljon R, Painter HJ, et al. The PfAP2-G2 transcription factor is a critical regulator of gametocyte maturation. Mol Microbiol. 2021;115: 1005–1024.
50. Early AM, Camponovo F, Pelleau S, Cerqueira GC, Lazrek Y, Volney B, et al. Declines in prevalence alter the optimal level of sexual investment for the malaria parasite Plasmodium falciparum. Proc Natl Acad Sci. 2022;119: e2122165119.
51. Essuman E, Grabias B, Verma N, Chorazeczewski JK, Tripathi AK, Mlambo G, et al. A novel gametocyte biomarker for superior molecular detection of the Plasmodium falciparum infectious reservoirs. J Infect Dis. 2017;216: 1264–1272.
52. Nair S, Li X, Nkhoma SC, Anderson T. Fitness costs of pfhrp2 and pfhrp3 deletions underlying diagnostic evasion in malaria parasites. J Infect Dis. 2022;226: 1637–1645.
53. Vreden SG, Jitan JK, Bansie RD, Adhin MR. Evidence of an increased incidence of day 3 parasitaemia in Suriname: an indicator of the emerging resistance of Plasmodium falciparum to artemether. Mem Inst Oswaldo Cruz. 2013;108: 968–973.
54. Vreden SG, Bansie RD, Jitan JK, Adhin MR. Assessing parasite clearance during uncomplicated Plasmodium falciparum infection treated with artesunate monotherapy in Suriname. Infect Drug Resist. 2016; 261–267.
55. Zupko RJ, Nguyen TD, Ngabonziza JCS, Kabera M, Li H, Tran TN-A, et al. Modeling policy interventions for slowing the spread of artemisinin-resistant pfkelch R561H mutations in Rwanda. Nat Med. 2023; 1–10.
56. Dondorp AM, Yeung S, White L, Nguon C, Day NP, Socheat D, et al. Artemisinin resistance: current status and scenarios for containment. Nat Rev Microbiol. 2010;8: 272–280.
57. Weiss D, Nelson A, Vargas-Ruiz C, Gligorić K, Bavadekar S, Gabrilovich E, et al. Global maps of travel time to healthcare facilities. Nat Med. 2020;26: 1835–1838.
58. Tyers M. riverdist: River network distance computation and applications. R Package Version 015 0. 2017.
59. Van Etten J. R package gdistance: distances and routes on geographical grids. 2017.
60. Oyola SO, Ariani CV, Hamilton WL, Kekre M, Amenga-Etego LN, Ghansah A, et al. Whole genome sequencing of Plasmodium falciparum from dried blood spots using selective whole genome amplification. Malar J. 2016;15: 1–12.
61. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv Prepr ArXiv13033997. 2013.
62. Picard Tools. By Broad Institute. 2020.
63. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20: 1297–1303.
64. Miles A, Iqbal Z, Vauterin P, Pearson R, Campino S, Theron M, et al. Indels, structural variation, and recombination drive genomic diversity in Plasmodium falciparum. Genome Res. 2016;26: 1288–1299.

65. 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: 1–4.
66. Manske M, Miotto O, Campino S, Auburn S, Almagro-Garcia J, Maslen G, et al. Analysis of Plasmodium falciparum diversity in natural infections by deep sequencing. Nature. 2012;487: 375–379.
67. Hagberg A, Conway D. NetworkX: Network Analysis with Python. URL Httpsnetworkx Github Io. 2020.
68. Waskom ML. Seaborn: statistical data visualization. J Open Source Softw. 2021;6: 3021.
69. Ansbro MR, Jacob CG, Amato R, Kekre M, Amaratunga C, Sreng S, et al. Development of copy number assays for detection and surveillance of piperaquine resistance associated plasmepsin 2/3 copy number variation in Plasmodium falciparum. Malar J. 2020;19: 1–10.

Figures

Figure 1 – Spatial and temporal distribution of *Plasmodium falciparum* samples in Guyana. a) Spatial clusters (n=20) delimited using an informed approach following access using roads and rivers. b) temporal distribution of samples (n=1,409) colored by patient travel history. Three sampling periods could be observed 2016/2017 (n=773), 2018/2019 where samples were collected as part of a therapeutic efficacy study (TES) (n=174) with no information on patient travel history, and 2020/2021 (n=531).

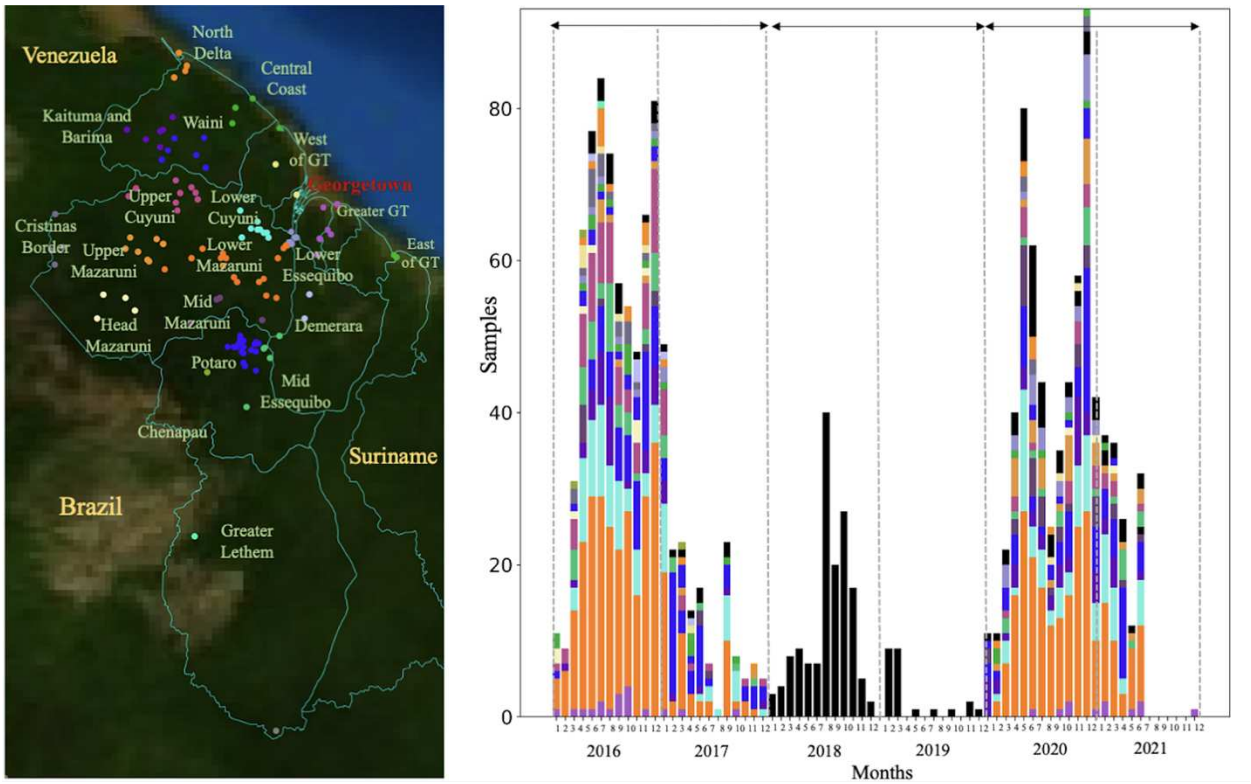


Figure 2 - The mean IBD between samples highlighting highly related clusters (IBD ≥ 0.4). Different sampling years are indicated as well as the presence/absence of *pfcrt* C350R.

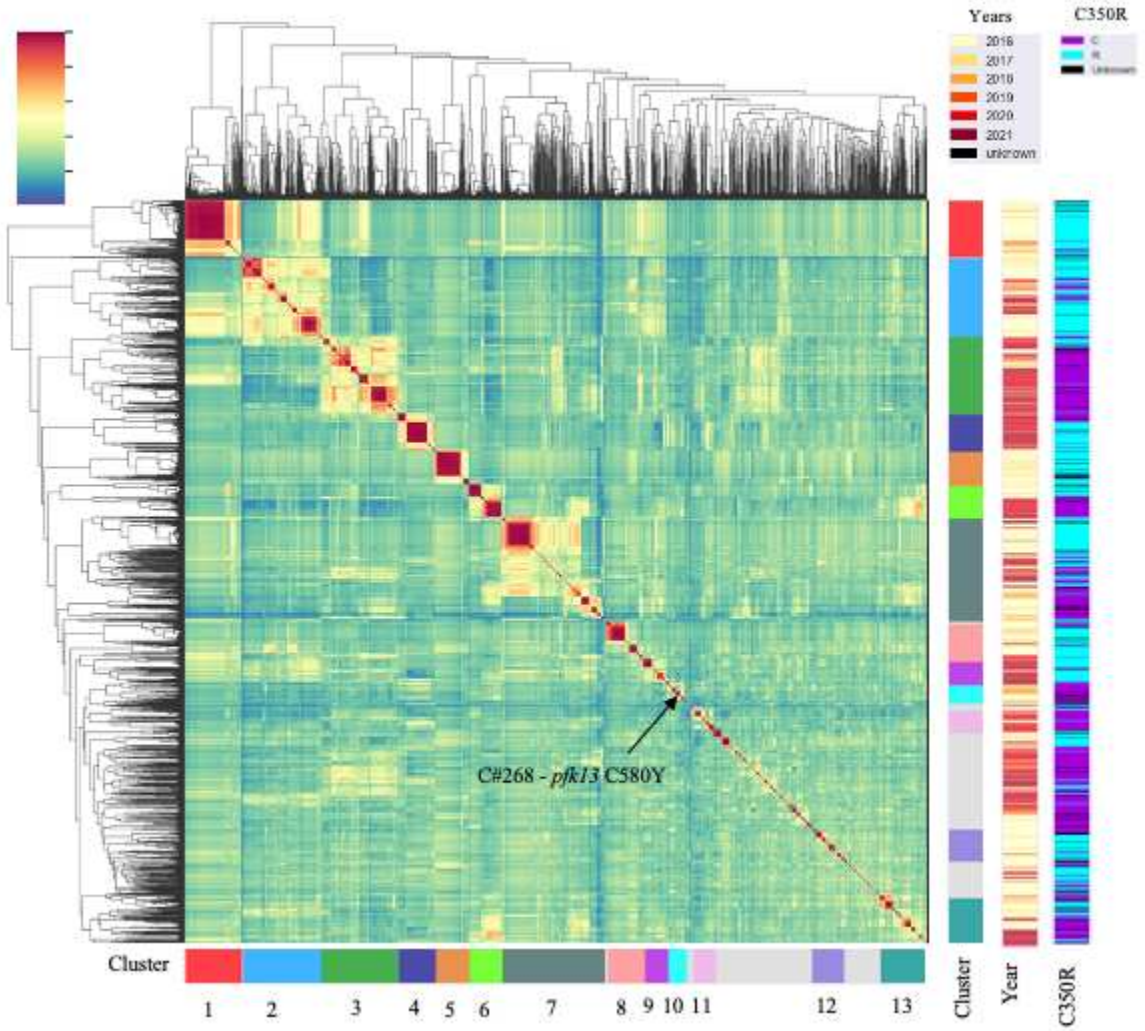


Figure 3 – Clonal temporal dynamics between 2016 and 2021. Clone (clonal component) C#1 in highly related cluster 1 was the largest clone present in the dataset (n=73 samples). C#268 is the clonal background harboring *pfk13* C580Y, while C#321 carried the *pfk13* G718S. All clones highlighted on the figure are referenced in the text.

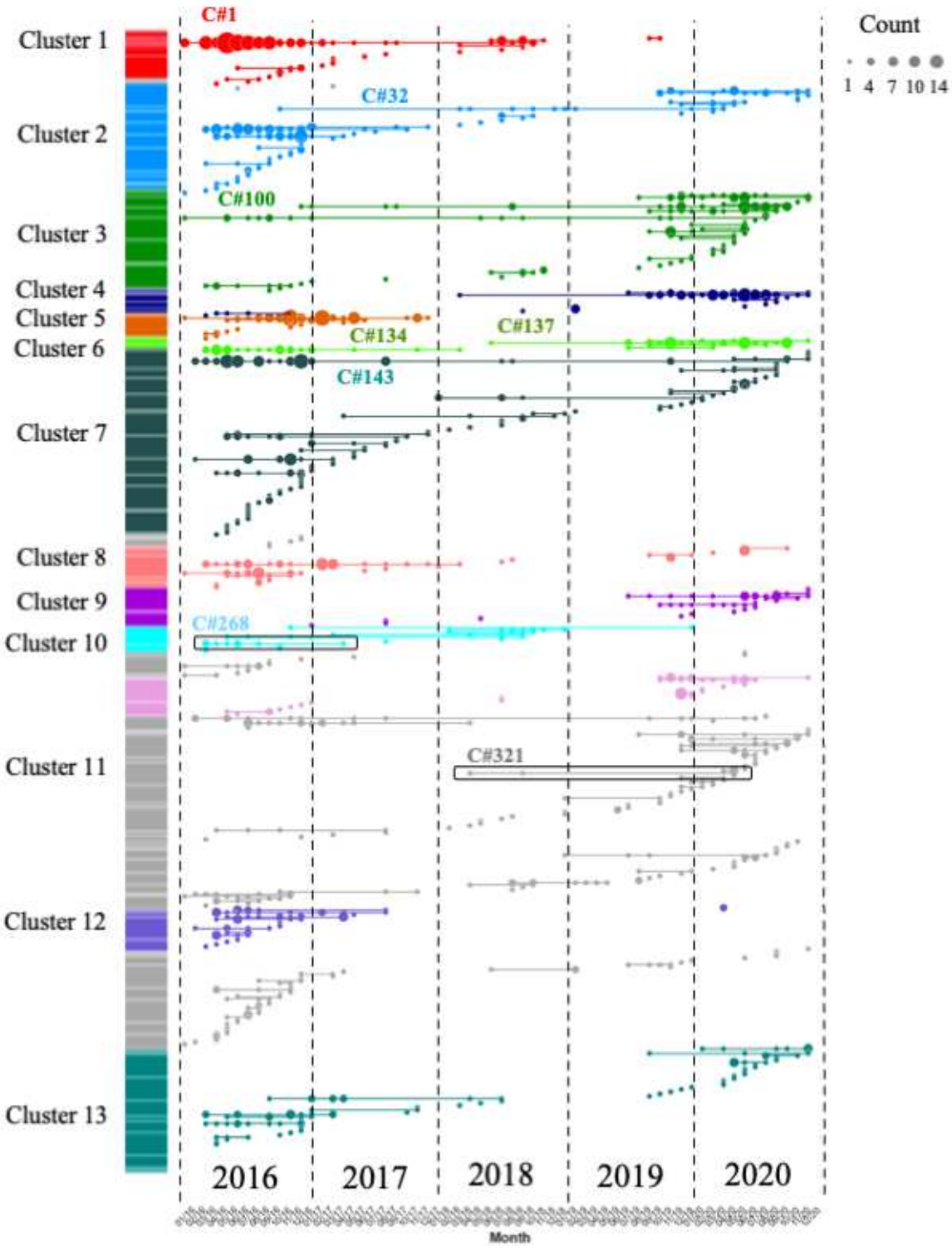


Figure 4 – Clone persistence and the number of spatial clusters reached (> 1 sample and at least one spatial location, n=130, mean =287 days). Clones are colored by highly related clusters. C#268 carrying the *pfk13* C580Y mutation is highlighted right on the persistence 80th percentile (vertical line). The horizontal and vertical dashed lines represent the 95th percentiles. All clones highlighted on the figure are ref in the text.

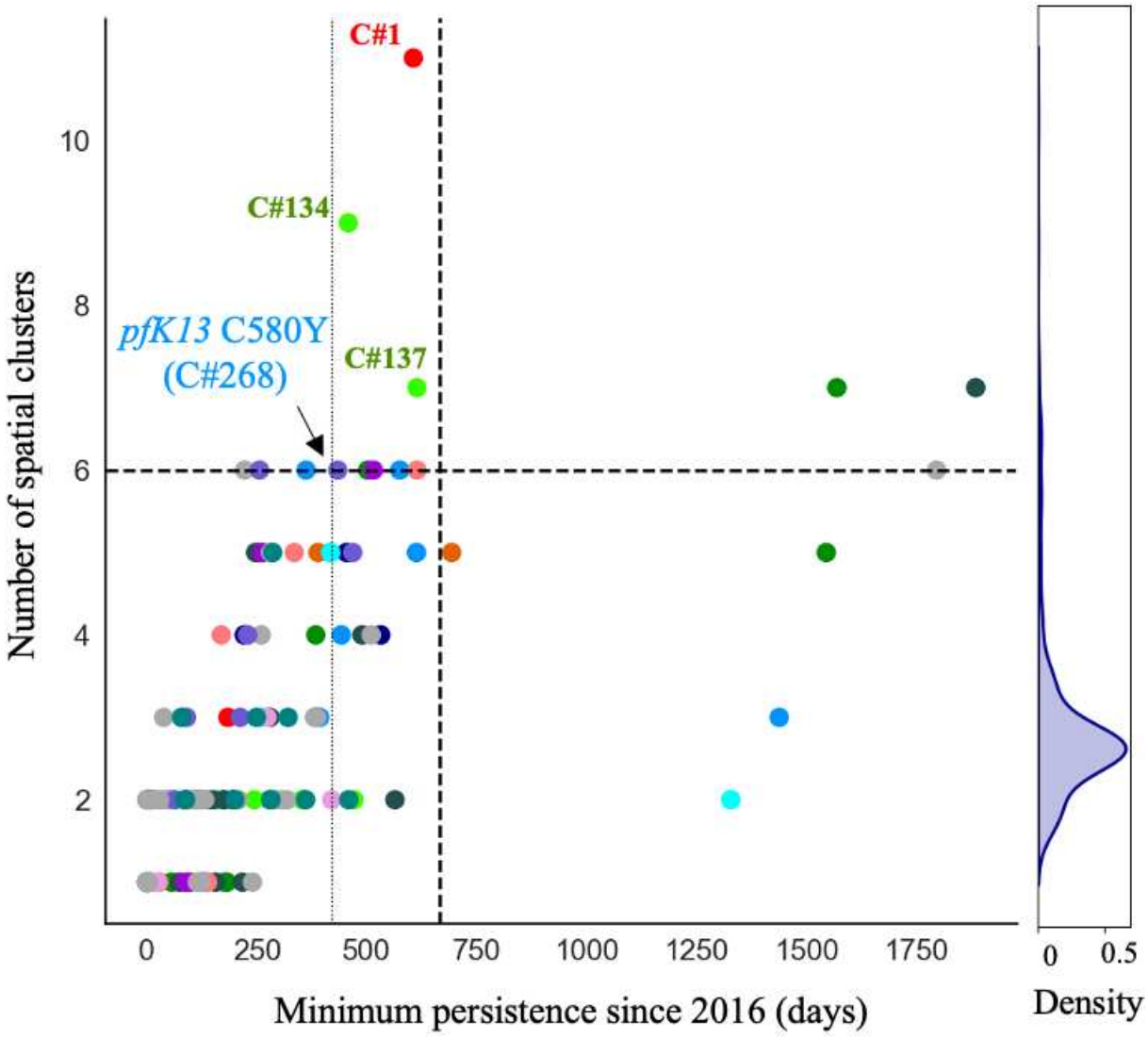


Figure 5 – Clone persistence and clone size for NSY mutation with similar (± 0.05) MAF (a-b) *pfk13* C580Y (MAF = 0.007 ± 0.05 n = 2,360), c-d) *pfcrt* C350R (MAF = 0.46 ± 0.05 n = 683) and (e-f) Eleven NSY (nine genes) on chromosome 9 which increased in frequency (in 99th percentile: $\Delta_{FREQUENCY} \geq 28.4\%$ - MAF = 0.29 ± 0.05 n = 853). Vertical black lines represent the mean of the distribution, red vertical lines are the mutations observed, and the dashed line is 95th percentile at the particular MAF.

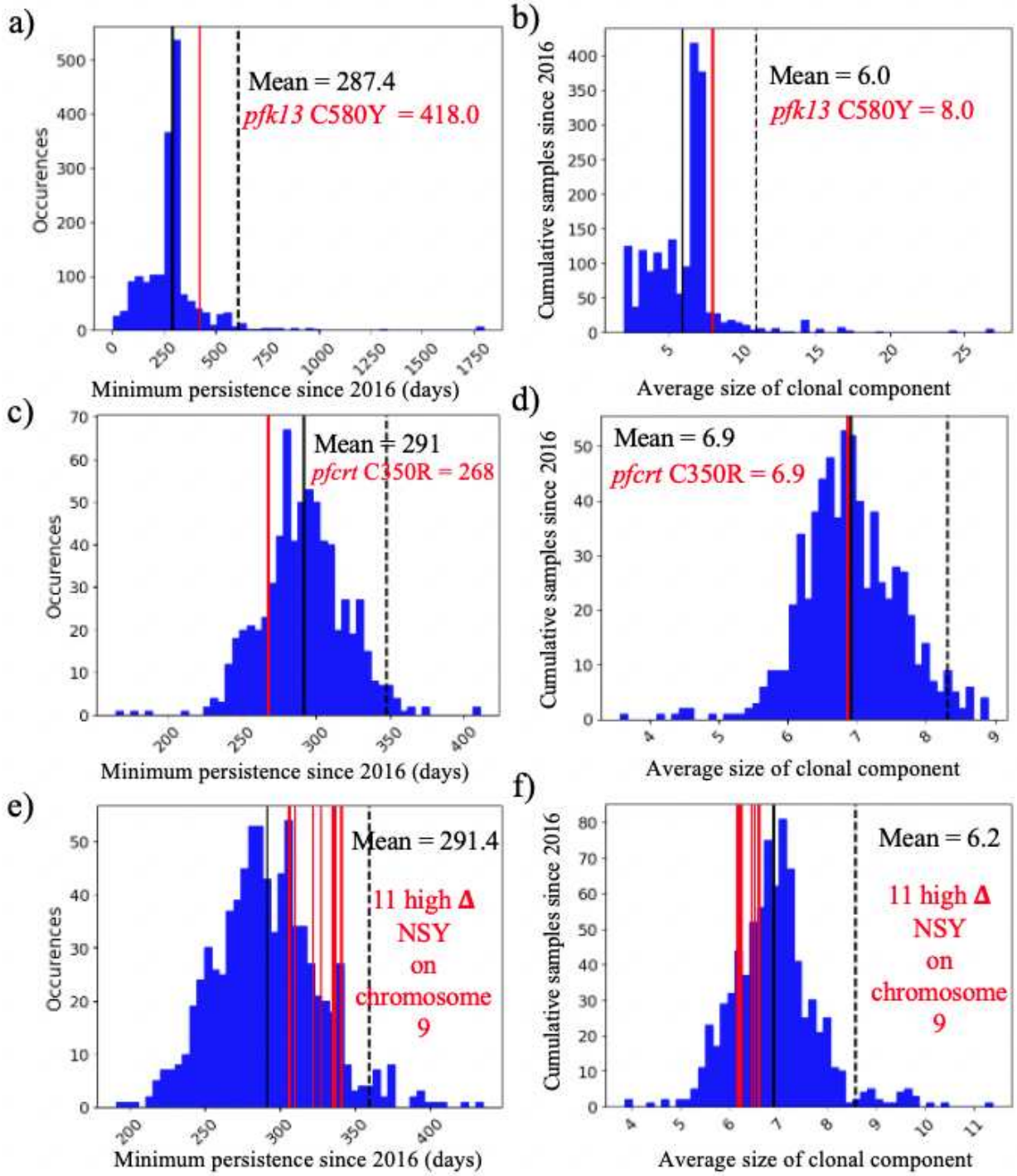


Figure 6 – Selection signals from isoRelate in long-lasting clones (sampled over three months) and short-lived clones sampled over two study periods (a-b) 2016/2017 and (c-d) 2020/2021. Dashed lines represent the threshold for the different selection signals investigated using a false discovery rate of 0.01.

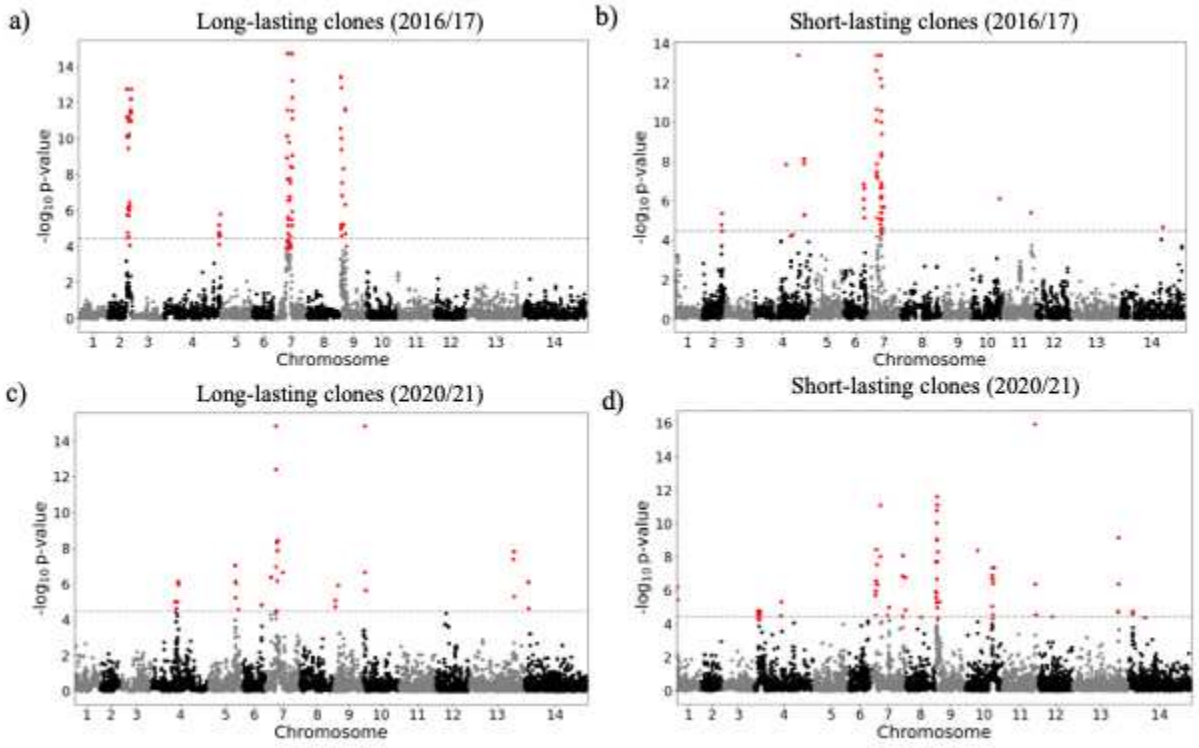
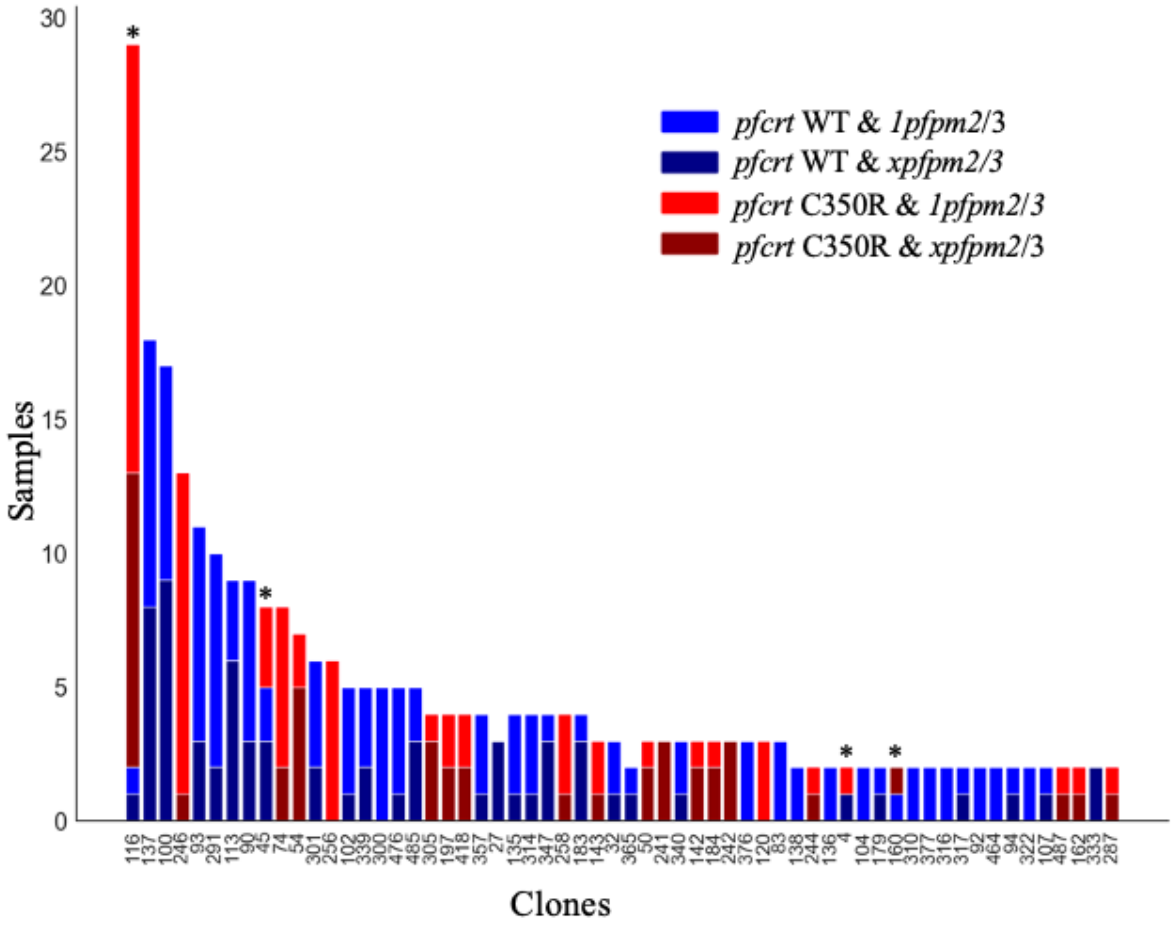


Figure 7 – *pfcrt* C350R and plasmepsin amplification (*xpfpm2/3*) in the different clones (> 1 sample) from 2020/2021. Copy number in plasmepsin appears not consistent among clones ($P < 0.0349$) and recurrent mutational events of *pfcrt* C350R were observed in four clones (C#116, C#45, C#4, C#160).



813