





Contrasting genomic epidemiology between sympatric *Plasmodium falciparum* and *Plasmodium vivax* populations

Received: 11 April 2024

Accepted: 11 September 2024

Published online: 30 September 2024

 Check for updates

Philipp Schwabl ^{1,2,12}, Flavia Camponovo^{3,4,5,12}, Collette Clementson^{6,12}, Angela M. Early², Margaret Laws^{1,2}, David A. Forero-Peña⁷, Oscar Noya^{8,9}, María Eugenia Grillet¹⁰, Mathieu Vanhove^{1,2}, Frank Anthony⁶, Kashana James⁶, Narine Singh⁶, Horace Cox^{6,11}, Reza Niles-Robin ⁶, Caroline O. Buckee ³ & Daniel E. Neafsey ^{1,2} ✉

The malaria parasites *Plasmodium falciparum* and *Plasmodium vivax* differ in key biological processes and associated clinical effects, but consequences on population-level transmission dynamics are difficult to predict. This co-endemic malaria study from Guyana details important epidemiological contrasts between the species by coupling population genomics (1396 spatio-temporally matched parasite genomes, primarily from 2020–21) with sociodemographic analysis (nationwide patient census from 2019). We describe how *P. falciparum* forms large, interrelated subpopulations that sporadically expand but generally exhibit restrained dispersal, whereby spatial distance and patient travel statistics predict parasite identity-by-descent (IBD). Case bias towards working-age adults is also strongly pronounced. *P. vivax* exhibits 46% higher average nucleotide diversity (π) and 6.5x lower average IBD. It occupies a wider geographic range, without evidence for outbreak-like expansions, only microgeographic patterns of isolation-by-distance, and weaker case bias towards adults. Possible latency-relapse effects also manifest in various analyses. For example, 11.0% of patients diagnosed with *P. vivax* in Greater Georgetown report no recent travel to endemic zones, and *P. vivax* clones recur in 11 of 46 patients incidentally sampled twice during the study. Polyclonality rate is also 2.1x higher than in *P. falciparum*, does not trend positively with estimated incidence, and correlates uniquely to selected demographics. We discuss possible underlying mechanisms and implications for malaria control.

Malaria caused by the parasite *Plasmodium falciparum* is responsible for more than half a million deaths each year, predominantly in Sub-Saharan Africa and in children under five¹. Malaria by *Plasmodium vivax* causes less acute mortality but is more widely distributed at the global scale^{2,3} and likewise causes severe morbidity and socioeconomic impact¹. Progress towards

malaria elimination is currently stalling in many endemic regions⁴, and divergent parasite species' responses to intervention play an important role^{5–9}. *P. vivax* in particular often shows stable or increasing incidence during periods in which co-endemic *P. falciparum* is successfully being reduced⁴. It is therefore important to better understand how epidemiological

differences between the species affect the impact of control strategies and whether species-tailored approaches are warranted in co-endemic settings.

P. falciparum and *P. vivax* are not closely related (current theory suggests 30–50 million years of independent evolution^{10,11}) and differ in key developmental processes and associated clinical effects⁴. For both species, human infection begins when sporozoites enter the skin and vasculature while an infected anopheline mosquito is probing for blood. The parasite first infects the liver and later transitions to blood-stage infection where repeated erythrocyte invasion, intra-erythrocytic replication, and erythrocyte rupture are associated with febrile and paroxysmal disease. A subset of blood-stage parasites commit to onward transmission by differentiating into the sexual, mosquito-infective gametocyte form. Meiosis occurs after transmission to the mosquito via blood meal and involves outcrossing if the blood meal contains genetically distinct gametocytes (i.e., if the human infection source was polyclonal). The parasite forms oocysts under the basal lamina of the mosquito midgut and these later release sporozoites which invade the mosquito salivary glands¹².

P. vivax behaves distinctly within the human stages of this infection cycle in multiple key aspects. First, not all *P. vivax* parasites that successfully reach the liver immediately begin to produce merozoites via schizogony as is the case for all liver-stage *P. falciparum* parasites. Rather, a subset of liver-stage *P. vivax* parasites form a dormant hypnozoite reservoir which can initiate additional blood-stage infections weeks to months after initial infection¹³. Clearing this relapsing hypnozoite reservoir from the liver is hindered by incomplete drug efficacy¹⁴, toxicity¹⁵, and adherence concerns¹⁶. In the blood stage, *P. vivax* invades a smaller subset of erythrocytes¹⁷ and thus more frequently creates submicroscopic infections, especially in adults¹⁸. Furthermore, while initial blood stage onset is similar between species (12–14 days¹²), the *P. vivax* blood stage generates mature gametocytes ~11 days earlier than that of *P. falciparum*¹². This increases the likelihood of onward transmission before antimalarial treatment occurs^{19,20}. Several experimental infection studies have additionally suggested that the minimum gametocyte density required to infect susceptible mosquitoes is intrinsically lower in *P. vivax* than in *P. falciparum*^{19,21–25}.

Taken together, the above properties are thought to make *P. vivax* more resilient to clinical interventions targeting symptomatic and often pediatric infection, a main focus of *P. falciparum* control. Vector control measures are also thought to have weaker or delayed effects on *P. vivax* due to post-intervention relapse from undetected or unsuccessfully cleared hypnozoite reservoirs as well as higher probability to achieve onward transmission when mosquito abundance is low. Furthermore, long-lasting hypnozoite reservoirs and more transmissible asymptomatic blood-stage infections may make *P. vivax* more effective at long-distance dispersal and outcrossing with unrelated strains. This may also include greater potential for allochthonous transmission where malaria was previously absent or cleared^{26–28}.

The epidemiological effects of these biological differences between *P. falciparum* and *P. vivax* have been difficult to quantify because synchronized comparative study designs that avoid confounding by methodological and spatiotemporal factors are challenging to establish²⁹. Study systems representing true species sympatry are also rare, as *P. falciparum* and *P. vivax* often occupy distinct, only partially overlapping socio-geographic partitions within areas more broadly considered co-endemic^{2,3}.

This study uniquely couples a spatiotemporally matched genomic sampling scheme with epidemiological analysis of disaggregated (patient-level) malaria case records to discern co-endemic *P. falciparum* and *P. vivax* transmission dynamics. We focus on Guyana, where both species contribute substantially to the national case count (32.3% *P. falciparum*, 67.4% *P. vivax*, and <1% *P. malariae* in 2019) and where true sympatry is geographically widespread. Malaria in Guyana also

involves a diverse mix of ethnicities (e.g., strong representation from Afroguyanese, East Indian, and Amerindian groups) and dynamic population mobility patterns relating to the mining field. Next to important distinctions involving geographic and host demographic variables, we demonstrate clear species contrasts in strain ancestry, persistence, dispersal, and infection complexity within this multifaceted transmission context. We discuss possible biological drivers behind divergent species epidemiologies and identify various implications for disease control.

Results

General spatiotemporal species prevalence

We first conducted a descriptive summary of country-wide passive case detection records made available by the Guyana Ministry of Health (GMOH) for 2006 to 2019. Rolling monthly counts demonstrate well-balanced case burden between species from 2006 to 2010 followed by higher *P. falciparum* counts from 2010 to 2013 (Fig. 1a), a period during which spiking gold prices and La Niña climate conditions are thought to have fueled strong malaria peaks³⁰. *P. vivax* cases began to exceed those of *P. falciparum* in mid-2013 and remained over-represented through 2019 (end of dataset). Given the availability of individual-level case metadata curated by the GMOH for 2019, we focused on 2019 records for most subsequent epidemiological analyses. These records include 17,710 single-species cases (67.43% *P. vivax*, 32.35% *P. falciparum*, and 0.22% *P. malariae*) and 1001 mixed-species cases containing both *P. vivax* and *P. falciparum* (Supplementary Table 1). Both species show transmission throughout the year and elevated incidence between April and August (Supplementary Fig. 1).

Species intricacies of a wide co-endemic distribution

2019 case records show the extent of *P. falciparum*–*P. vivax* geographic and demographic co-endemicity in Guyana (Fig. 1b–d). Regions I, VII, and VIII account for 94.6% of *P. falciparum* cases and 89.3% of *P. vivax* cases, and regional species proportions remain stable across East Indian, Amerindian and Mixed/Other ethnicity groups. While *P. vivax* consistently predominates in these groups, case majority flips to *P. falciparum* in the Afroguyanese (Chi-squared tests, $\chi^2 = 36.13$ (I), 269.63 (VII), 125.51 (VIII), Bonferroni-adjusted (bf-adj.) $p < 0.001$ (all regions), Fig. 1b). This result is consistent with lower *P. vivax* infection risk in Duffy-negative blood types³¹. We further classified infection localities into 27 epidemiological zones (average 3,631 km² each) defined using river-road transport networks and mobility friction rasters (Fig. 1 map, Supplementary Fig. 2). Again, species sympatry is consistently maintained in all zones within Regions I, VII, and VIII. The relative proportion of *P. vivax* cases is generally higher outside of this main malaria range, especially to the south in Region IX (e.g., Upper Rupununi and Greater Lethem). Patients aged 15–54 clearly predominate in both species but are more common in *P. falciparum* (84.5%) than in *P. vivax* (79.2%—Chi-squared test, $\chi^2 = 82.89$, $p < 0.001$, Fig. 1d). The proportion of cases representing *P. falciparum* (Fig. 1e, Supplementary Fig. 3) also increases markedly with age in male patients (Kruskal-Wallis test, $H = 23.64$, $p < 0.001$ —see post-hoc significance in Fig. 1e legend). Working-age (≥ 15 years) to child (< 15 years) case ratios are less markedly elevated in Amerindians than in non-Amerindian ethnicities (3.8 to 1 vs. 13.5 to 1, respectively, for *P. falciparum* (Chi-squared test, $\chi^2 = 273.84$, $p < 0.001$) and 2.6 to 1 vs. 7.4 to 1, respectively, for *P. vivax* (Chi-squared test, $\chi^2 = 515.98$, $p < 0.001$), Supplementary Fig. 4). Elevated case bias towards working-age males, more strongly so in non-Amerindian than in Amerindian ethnicities, is consistent with male-dominated hinterland mining activity driving malaria burden in Guyana^{30,32} (see mineral resource map in Supplementary Fig. 5), especially in non-resident (e.g., ‘coastlander’³³) mining groups. Higher working-age to

child case ratio for *P. falciparum* than for *P. vivax* is also consistent with the idea that transmission in non-mining communities is less common in *P. falciparum* than in *P. vivax*.

Deep genomic profiling of each species across 2020 and 2021

To search for epidemiological features that are not observable from traditional symptomatic case analysis, we sequenced 708 *P. falciparum* and 762 *P. vivax* genomes from patient blood spots collected passively by the GMOH. Collection followed informed consent using a study protocol approved by ethical committees of Harvard University and the government of Guyana. Sequencing was successful for 666 *P. falciparum* and 705 *P. vivax* genomes (Supplementary Data 1) continuously spanning January 2020 to June 2021 (Fig. 2). Of these 1371 samples, 1216 (89%) are coupled with patient travel history records used to infer geographic infection source (representing Regions I, VII, and VIII in 1156 (95.1%) samples). We additionally sequenced 13 *P. falciparum* (7 successful) and 22 *P. vivax* samples (16 successful) from 2019 representing infection sites in Venezuela and analyzed two publicly available Venezuelan *P. falciparum* genomes from 2015/16 (European Nucleotide Archive (<https://www.ebi.ac.uk>) accessions ERR1818176 and ERR2496572). Mean read-depth (MRD) variation among successfully sequenced *P. falciparum* (median MRD = 58.7) and

P. vivax samples (median MRD = 28.7) is shown in Supplementary Fig. 6.

Discrepant genomic diversity and relatedness between species

Genomic analyses focused on 46,067 and 266,141 sites with segregating single-nucleotide polymorphisms (SNPs) identified in the *P. falciparum* and *P. vivax* sample sets (respectively) following quality filtration steps. This contrast in the number of segregating sites and in average pairwise nucleotide diversity observed within each sample set (1.99×10^{-4} differences per bp in *P. falciparum* and 2.91×10^{-4} differences per bp in *P. vivax*, Supplementary Fig. 7) aligns with previous evidence that genetic diversity is higher in *P. vivax* at the global scale³⁴.

To further characterize diversity within each sample set, we applied the hidden state model hmIBD³⁵ to estimate parasite relatedness due to recent common ancestry (identity-by-descent; IBD) for 598 *P. falciparum* and 548 *P. vivax* samples identified as single-strain (monoclonal) infections using a Bayesian method called THERIALMcCOIL³⁶. The relatedness distributions generated for infections from Guyana (Fig. 3a) clearly differ between the two species (Kolmogorov-Smirnov test, $D = 0.99$, $p < 0.001$). Pairwise relatedness values in *P. vivax* exhibit a mean of 0.048, with very little standard deviation (sd = 0.032) and minimal skew or outlier occurrence. These

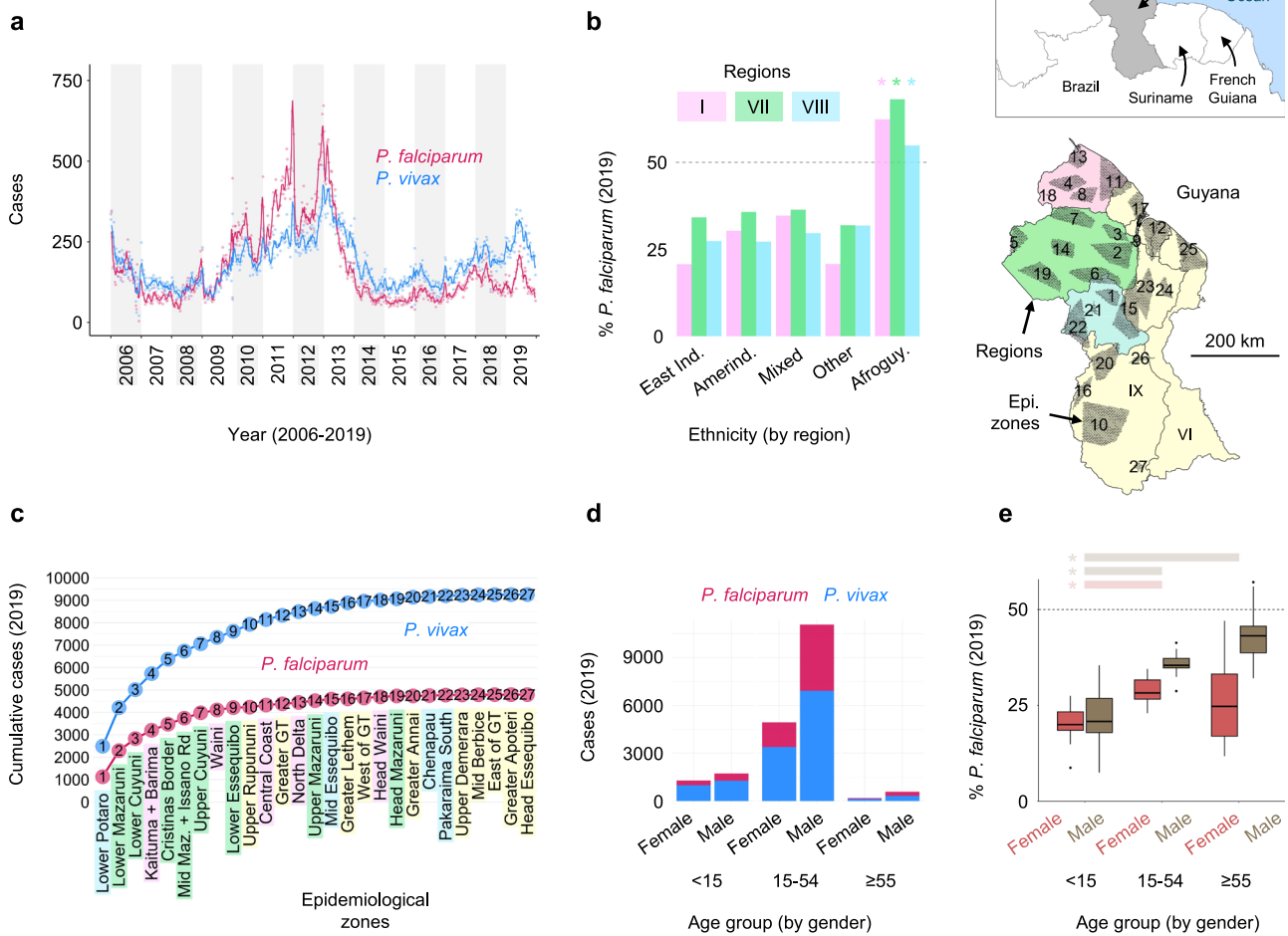


Fig. 1 | Spatiotemporal and demographic patterns of *P. falciparum* and *P. vivax* in Guyana. **a** Weekly reported cases (points) and 30-day rolling average (line) between 2006 and 2019. **b** Percent of cases representing *P. falciparum* (% Pf) by region and ethnicity in 2019. Asterisks indicate significant differences in % Pf between Afroguyanese and non-Afroguyanese cases via Chi-squared test (Bonferroni-adjusted (bf-adj.) $p = 5.5 \times 10^{-9}$ in Region I, bf-adj. $p = 4.1 \times 10^{-60}$ in Region VII, and bf-adj. $p = 1.2 \times 10^{-28}$ in Region VIII). **c** Cumulative case count across

epidemiological zones in 2019. Zones on the x-axis are ordered by total case count (descending from left to right). **d** Case counts by gender and age in 2019. **e** Boxplots summarize monthly variation (median and quartiles) in % Pf by gender and age in 2019. Asterisks indicate significant differences for age bins <15 vs. 15–54 (Benjamini-Hochberg-adjusted (bh-adj.) $p = 0.004$ in both males and females) and <15 vs. ≥55 (bh-adj. $p = 4.6 \times 10^{-6}$ in males) via two-sided Dunn test.

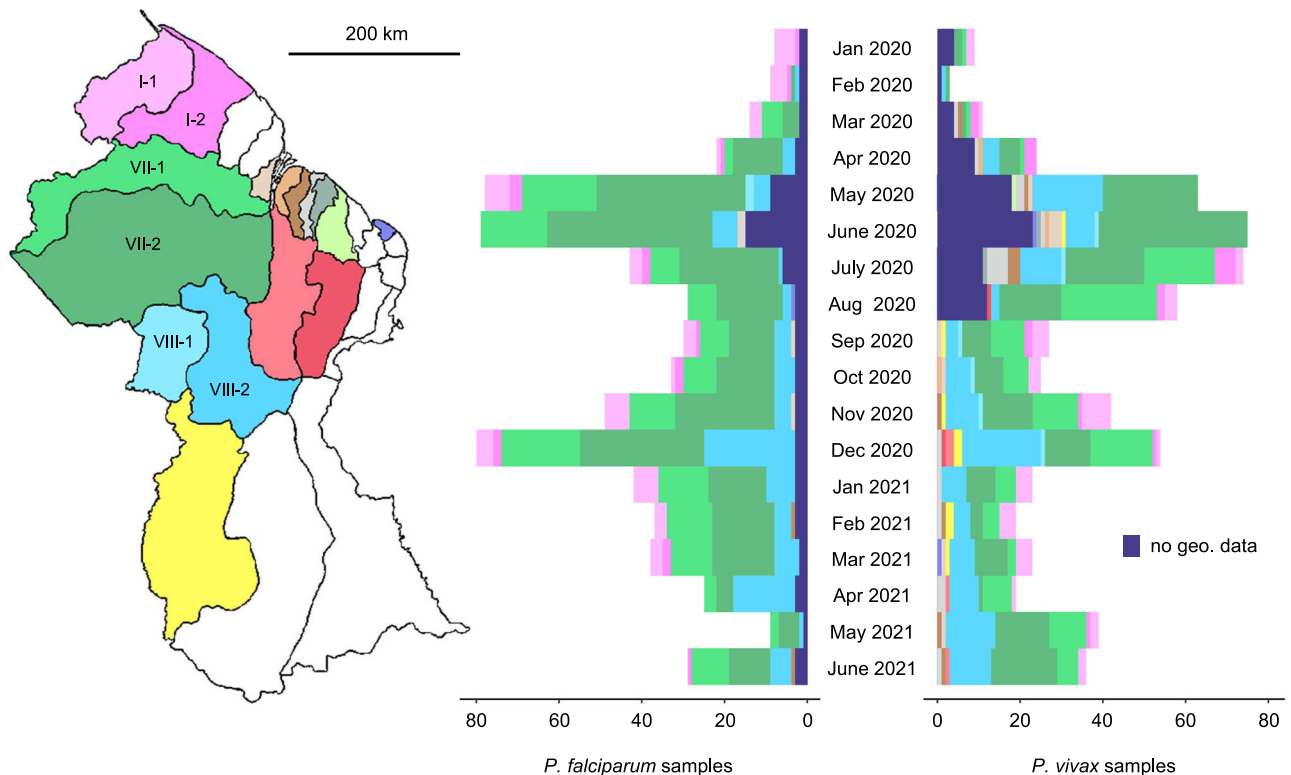


Fig. 2 | Spatiotemporally matched *P. falciparum* and *P. vivax* sampling in Guyana. Colors in the map and histograms represent the neighborhood councils (NDCs) to which infections from 2020–21 were attributed based on patient

response about location of stay 14 days prior to diagnosis. NDCs within Regions I, VII, and VIII are labeled in the map.

results are indicative of a diverse, frequently outcrossing population without substantial sub-clustering into groups of highly related individuals. In contrast, most pairwise comparisons in *P. falciparum* exhibit greater than half-sibling-level relatedness (mean = 0.310 ± 0.124 sd). The broader, right-tailed relatedness distribution also includes a peak near 1 representing the presence of identical clones as well as values between 0.5 and 0.75 representing the presence of various inbred sibling-level relationships. A network analysis identified 79 clonal groups (>0.90 IBD) among all *P. falciparum* samples, represented by 2–39 members each (Supplementary Fig. 8). Clonal group membership detection was often spatiotemporally aggregated (e.g., >50% of group members detected within 4 months per epidemiological zone, Supplementary Fig. 9), consistent with outbreak-like patterns of prevalence³⁷. The largest clonal groups spread across multiple geographic areas (Supplementary Fig. 10a), and most appear highly interrelated (e.g., 50 groups collapse into a single network when re-clustering samples at >0.60 IBD, Supplementary Fig. 8). Four groups noted with asterisks in Supplementary Fig. 8 and mapped in Supplementary Fig. 10b appear slightly divergent (1.7 to 3.3 sd below average between-group IBD). Other clonal patterns mapped in Supplementary Fig. 10b include the clustering of groups #7, #9, and #10 around Port Kaituma and Mabaruma (Region I), groups #18, #24, and #49 along the Mazaruni and Puruni Rivers (Region VII), and groups #27, #32, and #52 around Mahdia and the lower Potaro River (Region 8). Two divergent clonal pairs (#20 and #65) also have partial membership in Venezuela (star symbols in Supplementary Fig. 10b) and likely represent transmission of imported parasite lineages (Supplementary Text 1, Supplementary Fig. 11). Group #20 additionally provides evidence for longer-term persistence of *P. falciparum* clones (detection in 2015 and 2020).

Clonal network analysis results for *P. vivax* are very different. We observe only 29 clonal groups, of which 23 contain 2 members each (i.e., are clonal pairs) and 6 contain 3 members each. Post-hoc review of

patient-specific case codes (blinded to non-GMOH authors) also clarifies that 11 of 23 clonal pairs represent repeated sampling from the same patient (median time between 1st and 2nd visit = 114 days, minimum = 40 days, maximum = 229 days). These likely represent relapses registered as new infections prior to genetic analysis. No close relatedness occurs among any clonal groups (the same number of groups created via clustering at >0.90 IBD remains when re-clustering at >0.30 IBD) or between Guyanese and Venezuelan sample sets (Supplementary Fig. 8, Supplementary Text 1). Furthermore, clonal persistence in *P. vivax* (median = 79 days, maximum = 243 days) appears more temporally limited (Wilcoxon test, $W = 55078$, $p = 0.003$) than in *P. falciparum* (median = 114 days, maximum = 500 days, excluding imported lineages (groups #20 and #65), Supplementary Fig. 12).

Regional isolation-by-distance in *P. falciparum* vs. micrographic structure in *P. vivax*

We next evaluated whether parasite genetic relatedness declines with spatial distance (Fig. 3b). The presence of ‘isolation-by-distance’ is a common indicator of population structure due to spatially limited dispersal, understanding of which is key to the success of control objectives such as limiting the spread of drug resistance^{38,39}. We applied sliding windows of spatial distance, quantifying the fraction of pairwise comparisons in each 30 km window that showed >0.50 IBD (for analysis using mean IBD and other dichotomization thresholds see Supplementary Fig. 13). Results for *P. falciparum* demonstrate a linear decrease in the frequency of >0.50 IBD as spatial distance increases between sample pairs (Pearson’s $r = 0.89$, $p < 0.001$). Exceptions occur at window starts near 100 km and 200 km where the frequency of close relatives briefly rebounds to levels observed at shorter distance classes. These upticks may partially reflect accumulation of relatedness between popular mining sites or logistics hubs (Supplementary Text 2, Supplementary Fig. 14). In *P. vivax*, sample pairs with >0.50 IBD appear very rare ($n = 87$) but also clearly overrepresented in the first spatial

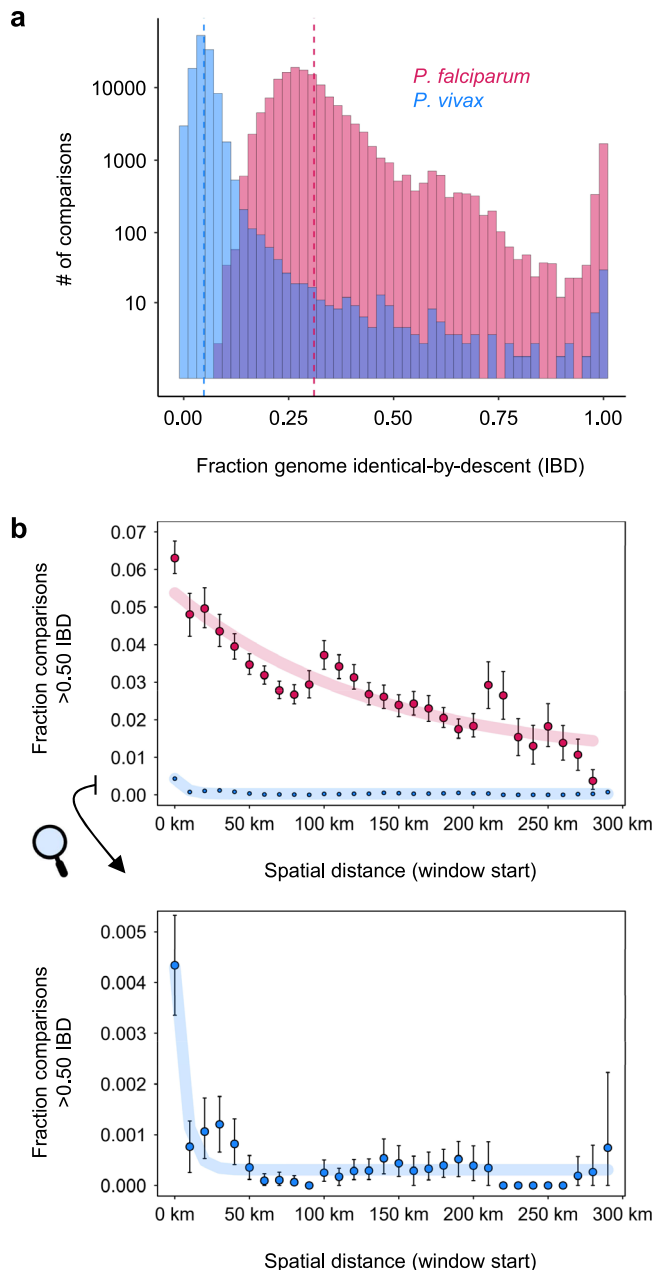


Fig. 3 | Pairwise IBD and relationships to spatial distance in *P. falciparum* and *P. vivax* in Guyana. **a** Histogram of pairwise IBD in 2020–21. The y-axis is log₁₀-scaled to help display less frequent close-relative observations. Dashed lines indicate species means. **b** Relationship of spatial distance between inferred infection sites and the frequency of >0.50 IBD in 2020–21. Points represent observations from windowed calculation using step size = 10 km starting at 0–30 km. Error bars represent 90% confidence intervals from bootstrapping 1000x. The model $y = a \cdot e^{-bx} + c$ was used to fit regression lines of linear and logarithmic shape. Distance analyses use 359 *P. falciparum* samples and 474 *P. vivax* samples (following an exclusion of *P. falciparum* samples belonging to clonal groups of ≥ 10 membership size).

distance class (0–30 km, Fig. 3). A continued decline in >0.50 IBD frequency with spatial distance is however not apparent. To help assess whether the spatial distance across which negative correlation between relatedness and spatial distance persists differs between species, we additionally quantified 99th percentile (p99) IBD, an alternative definition of close ancestry that normalizes statistical power in comparative Mantel tests (Supplementary Fig. 15, top). Correlogram results for *P. falciparum* indicated significant correlation between

genetic distance (1–p99 IBD frequency) and spatial distance for 0–30 km (Mantel $r = 0.49$, Holm-adjusted (hm-adj.) $p = 0.001$) as well as for 30–60 km distance classes (Mantel $r = 0.19$, hm-adj. $p = 0.017$). In *P. vivax*, correlation between 1–p99 IBD frequency and spatial distance loses statistical significance at 30–60 km (Mantel $r = 0.12$, hm-adj. $p = 0.089$). Shorter isolation-by-distance signal relative to *P. falciparum* also occurred when using the absolute >0.50 IBD metric to define close ancestry (Supplementary Fig. 15, bottom).

Genetic vs. non-genetic markers accord better on *P. falciparum* than on *P. vivax* dispersal

We next visualized >0.50 IBD frequency within and between epidemiological zones (Fig. 4) and examined correlations to patient travel patterns inferred from the 2019 epidemiological database. These analyses aimed to identify cases of similar, discordant, or complementary inference on parasite dispersal between data types and to understand the extent to which predictive power differs between malaria species.

Corollary to the clear elevation in >0.50 IBD frequency at 0–30 km (Fig. 3) and significant isolation-by-distance at 0–30 km (Supplementary Fig. 15) in both *P. falciparum* and *P. vivax*, both also generally showed highest frequencies of elevated relatedness within (i.e., not between) epidemiological zones (Fig. 4). This observation from the genomic data matches the observation from the epidemiological database that the majority of symptomatic malaria cases involve non-mobile, ‘local’ infections (black pie slices, Supplementary Fig. 16), i.e., patients reporting absence of travel 14 days prior to diagnosis. The epidemiological data also suggests that the proportion of local to nonlocal infections (i.e., cases in which patients reported having stayed in a distinct epidemiological zone 14 days prior to diagnosis) is higher for *P. vivax* than for *P. falciparum* in 16 of 23 analyzed epidemiological zones (Supplementary Fig. 17). These 16 zones include non-endemic coastal areas such as Greater Georgetown, where 178 of 1,611 (11.0%) *P. vivax* patients (vs. 45 of 867 (5.2%) *P. falciparum* patients) reported absence of prior travel (Chi-squared test, $\chi^2 = 22.915$, $p < 0.001$). Elevated pairwise nucleotide diversity (+20.7% π) in *P. vivax* samples attributed to Greater Georgetown (Welch’s t -test, $t = 4.71$, $p < 0.001$, Supplementary Fig. 7) further supports the hypothesis that relapses representing various different geographic infection sources are diagnosed in this capital district. To statistically compare inference of parasite movement based on genetic vs. epidemiological data types, we correlated the relative frequency of parasite >0.50 IBD and patient travel events (‘case flow’) among all pairs of endemic epidemiological zones for which patient movement was detected in the epidemiological database (i.e., all arrowed segments in Supplementary Fig. 16). In the case of *P. falciparum*, relative frequencies of >0.50 IBD and case flow were moderately correlated (Pearson’s $r = 0.44$, $p = 0.029$) between endemic zones (Fig. 5), with most conspicuous congruence in the context of reduced connectivity between Region I and Regions VII and VIII (e.g., see short spokes of elevated >0.50 IBD frequency and intensified case flow around Kaituma and Barima in Region I, yet limited connectivity farther South in Fig. 4 and Supplementary Fig. 16). Only 0.8% of nonlocal infections sampled in Regions VII and VIII (gray pie slices, Supplementary Fig. 16) represented infection sites from Region I. Case flow to Greater Georgetown was also weaker for infections attributed to Region I (36.4%) than for infections attributed to Region VII (43.4%—Chi-squared test, $\chi^2 = 48.22$, bf-adj. $p < 0.001$) or to Region VIII (74.1%—Chi-squared test, $\chi^2 = 874.35$, bf-adj. $p < 0.001$). These results may in part reflect the absence of efficient travel routes from Region I to Georgetown (East) and to regions farther South. Rivers key to travel in Region I (e.g., the Barima and Barama) lead into a Northwest delta while those key to travel in Regions VII and VIII (e.g., the Cuyuni and Mazaruni) flow East into the Essequibo near Georgetown. A handful of semi-developed roads also parallel and connect areas of the Cuyuni, Mazaruni, and Essequibo in Regions VII and VIII.

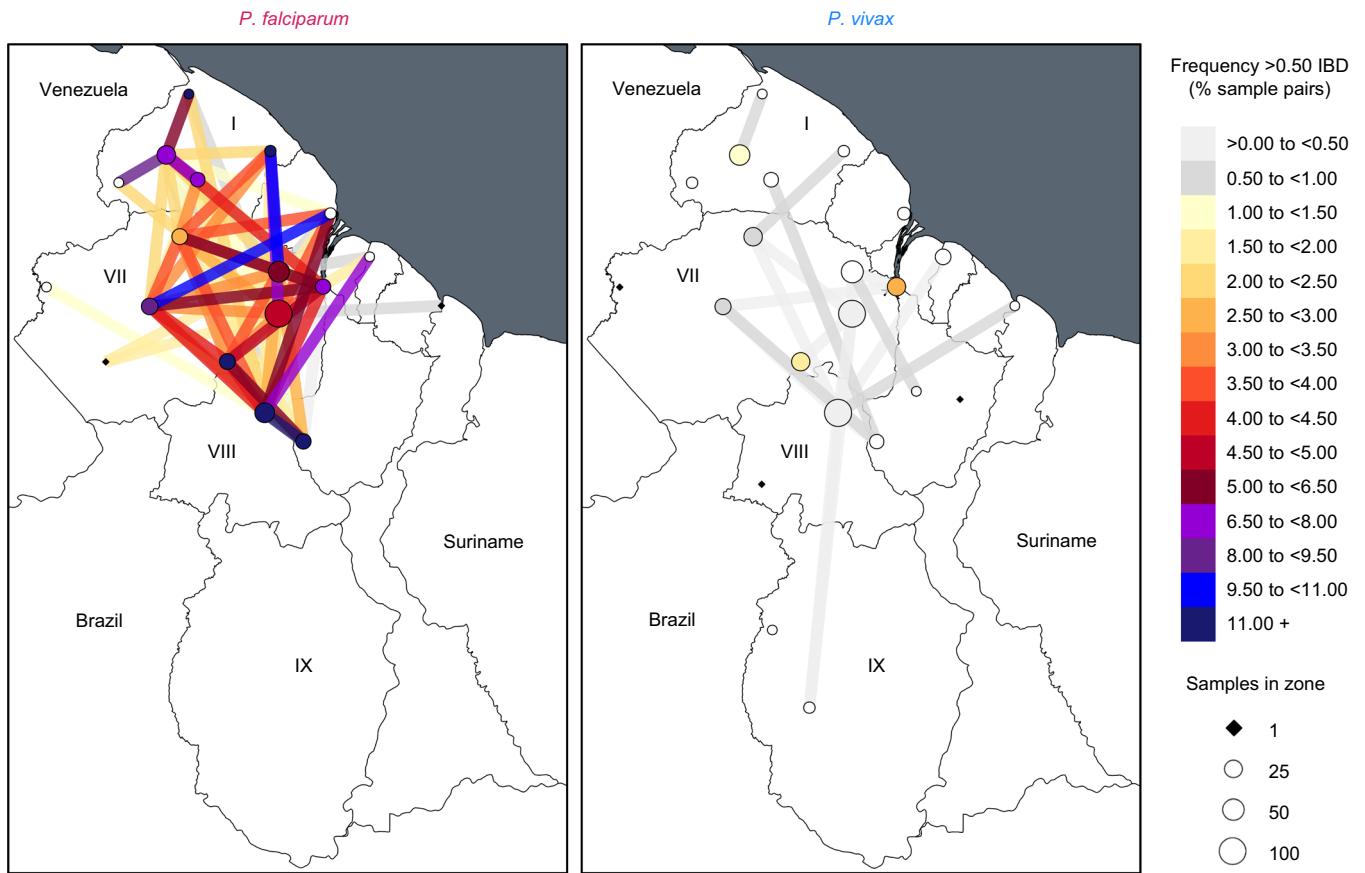


Fig. 4 | Frequency of >0.50 IBD within and between epidemiological zones in *P. falciparum* and *P. vivax* in Guyana. Nodes indicate epidemiological zones. Node and segment colors indicate frequency of >0.50 IBD within and between

epidemiological zones (respectively) in 2020–21. Node sizes are proportional to genomic sampling size. Between-zone comparisons represented by ≤ 50 comparisons are excluded.

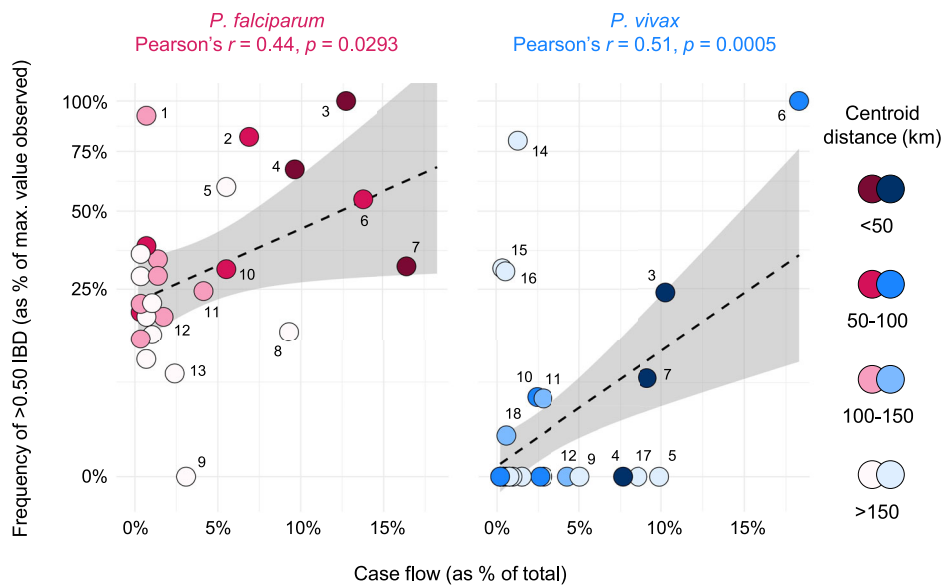


Fig. 5 | Relationship between patient case flow and the relative frequency of >0.50 IBD in *P. falciparum* and *P. vivax* in Guyana. Genetic data (y-axis) represents 2020–21 and epidemiological (case flow) data (x-axis) represents 2019. Point color indicates the spatial distance separating the zones being compared. Comparisons represented by ≤ 50 comparisons are excluded. Gray shading indicates 95% confidence intervals predicted by linear regression (dashed line). Labeled points represent observed values for (1) Central Coast (CC) vs. Lower Cuyuni (LC), (2) Head Waini (HW) vs. Kaituma and Barima (KB), (3) Lower Potaro (LP) vs. Mid

Essequibo, (4) KB vs. Waini, (5) CC vs. Lower Mazaruni (LM), (6) KB vs. North Delta (ND), (7) LC vs. LM, (8) Head Mazaruni vs. LM, (9) Cristinas Border vs. LM, (10) LM vs. Mid Mazaruni and Issano Rd, (11) LM vs. Upper Cuyuni, (12) CC vs. KB, (13) LM vs. ND, (14) East of Georgetown vs. LP, (15) LM vs. Upper Rupununi (UR), (16) LP vs. UR, (17) Greater Lethem vs. LM, and (18) LM vs. LP. Please note that significant correlation is lost from the *P. vivax* regression when outlier point #6 is omitted (Pearson's $r = 0.07$, two-sided $p = 0.675$). Significance remains when omitting this point from the *P. falciparum* regression (Pearson's $r = 0.41$, two-sided $p = 0.045$).

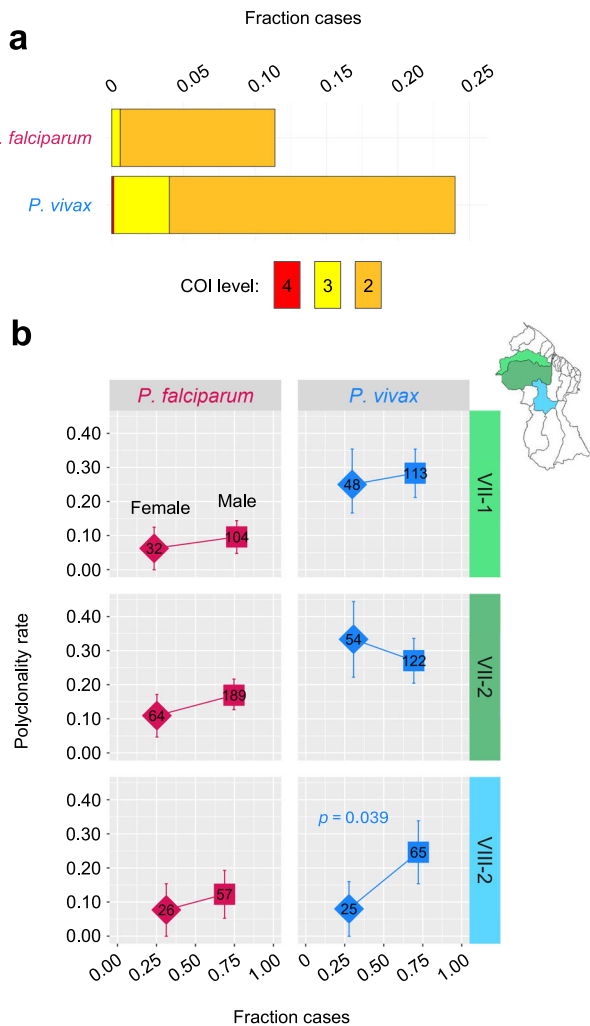


Fig. 6 | Polyclonality rates and relationships to patient gender in *P. falciparum* and *P. vivax* in Guyana. **a** Complexity of infection (COI) values in 2020–21, population-wide. **b** Polyclonality rate (2020–21) vs. patient gender fractions in neighborhood district councils (NDCs) represented by ≥ 20 genomic samples (see sample size annotations within plot). Gender fractions likewise represent 2020–21 (patient metadata from the genomic data set). Square and diamond symbols represent observed values for male and female patients, respectively, and error bars represent 90% confidence intervals from bootstrapping 1000x. Polyclonality rate comparison between male and female patients (two-proportions z-test) showed significance only for *P. vivax* in NDC VIII-2 (one-sided p -value = 0.039, without correction for multiple comparisons).

For *P. vivax*, the frequency of >0.50 IBD also correlated significantly with case flow between endemic zones (Pearson's $r = 0.51$, $p < 0.001$, Fig. 5). Unlike in *P. falciparum*, however, significance appears outlier-dependent (see Fig. 5 legend) and is lost when broadening the classification of elevated relatedness to $p99$ IBD (Supplementary Fig. 18). This observation helps clarify that the significant relationships detected when regressing >0.50 IBD frequency on spatial distance (Fig. 3b) and case flow (Fig. 5) do not mean that the *P. vivax* population is predictably structured as a whole. Results are instead consistent with a dispersive, frequently outcrossing *P. vivax* population in which pairwise relatedness is swiftly erased and local genotype associations are rarely established. In this scenario, the ephemeral presence of intact clones or sibling-level relatives creates micrographic spatio-temporal autocorrelations that are not representative of broader population dynamics.

Higher outcrossing potential in *P. vivax*—do relapse and risk carryover play a role?

We further assessed indications from relatedness analyses that outcrossing rate is elevated in *P. vivax* by examining polyclonality rate (i.e., the fraction of samples containing multiple distinct strains) and its variation with respect to geographic and demographic variables. Polyclonality rate is closely related to outcrossing rate because parasite sexual recombination in the mosquito after blood meal only results in outcrossing if multiple distinct strains are ingested simultaneously from the human host.

We observed a polyclonality rate of 24.0% in *P. vivax* vs. 11.4% in *P. falciparum* (Chi-squared test, $\chi^2 = 36.72$, $p < 0.001$, Fig. 6a). Elevated *P. vivax* outcrossing rate implied by this 2.1x polyclonality rate differential (which is likely a conservative estimate due to lower *P. vivax* read-depths—Supplementary Fig. 19) helps explain the aforementioned paucity in clonal *P. vivax* relationships (Fig. 3a) and their limited persistence over time (Supplementary Fig. 12). It is also consistent with the observation that linkage disequilibrium between SNPs declines across physical genetic distance more readily in *P. vivax* than in *P. falciparum* (Supplementary Fig. 20).

Polyclonality rate is considered a positive correlate of *P. falciparum* transmission intensity in the absence of high rates of case importation⁴⁰. We assessed this possible relationship in *P. falciparum* and *P. vivax* (in which polyclonality may additionally occur via hypozoite-based relapse) by plotting polyclonality rates per epidemiological zone against corresponding incidence estimates (Supplementary Fig. 21). Interestingly, *P. vivax* polyclonality rates remain consistently inflated, regardless of incidence estimate. Polyclonality rate and incidence estimates also trend in opposite directions for *P. falciparum* vs. *P. vivax*, though neither achieves statistical significance. Polyclonality rate was highest in *P. vivax* infections from the Lower Essequibo zone, a key area of travel convergence from the hinterlands of Region VII but one in which malaria risk is not considered high. These results may indicate that *P. vivax* samples from areas of lower endemicity are more likely to represent imported cases (possibly in the form of relapse) that involve multiple accumulated strains.

To further understand polyclonality rate relationships to transmission risk, we compared polyclonality rates in 15–54 year-old males and females within the three neighborhood councils (NDCs) for which ≥ 20 genomic samples per gender were obtained (Fig. 6b). Males represented the majority of epidemiological case records in the three NDCs, suggesting that transmission risk is generally higher in male-associated occupations. Polyclonality rate in *P. falciparum* thus followed expectations in that observed values appeared consistently higher in males than in females. Polyclonality rates in *P. vivax* were generally more variable (see confidence intervals, especially in females (Fig. 6b)), and were higher in females than in males in NDC VII-2. Interestingly, this trend was influenced by polyclonality rate enrichment in females among patients reporting Venezuelan nationality (Chi-squared test, $\chi^2 = 4.33$, $p = 0.037$, before adjusting for multiple comparisons—Supplementary Fig. 22). This enrichment may occur because infection risk is enhanced in female-biased occupations that accompany the active mining sector (e.g., in rest/supply areas where frequent contact among gametocyte-carrying male miners, mosquitoes, and female non-miners occurs). Enhanced *P. vivax* transmissibility, e.g., due to presymptomatic gametocytemia, may heighten this risk.

Discussion

In this study we coupled comparative malaria parasite population genomics with passive case record metadata analyses to compare co-endemic *P. falciparum* and *P. vivax* epidemiology in Guyana. We observed several key genomic epidemiological contrasts that indicate the need to consider species-specific approaches to malaria intervention.

Most notably, we observed dramatically different patterns of genomic relatedness between the two species. The *P. falciparum* sample set exhibited a relatively broad pairwise IBD distribution centered near values expected for half-sibling level relatedness, whereas in *P. vivax* this distribution was less variable and clearly shifted towards zero (6.5x lower average IBD). Clonal pairwise relationships (>0.90 IBD) were detected >50x more often in *P. falciparum* and most clonal *P. falciparum* groups exceeded 3 samples in membership size (max. 38). Many of these groups also displayed spatiotemporally aggregated, outbreak-like patterns of detection. In *P. vivax*, clonal groups had no more than 2–3 members and appeared to persist for shorter periods of time; many clonal *P. falciparum* relationships were detected between samples separated by >12 months whereas for *P. vivax* the maximum time between clone detection was 243 days.

This strong contrast in relatedness structure may derive mechanistically from polyclonality rate differences between the species. Polyclonality rate is a key determinant of the extent to which clones and near-relatives can persist in transmission cycles because polyclonality is a prerequisite to effective recombination between strains in the mosquito stage. In this study, *P. vivax* polyclonality was not only 2.1x higher than in *P. falciparum* at the population level but also remained consistently inflated when accounting for estimated incidence in subgeographic comparisons. This result suggests that elevated polyclonality rates in *P. vivax* are not a product of higher infection incidence alone. *P. vivax* relapse from the dormant hypnozoite liver stage likely represents an additional source of polyclonality because relapse can overlap with current infection and can involve the successive or simultaneous activation of distinct strains⁴¹.

Importantly, though humans are the common agent of dispersal throughout Guyana for both parasite species, we observed highly contrasting patterns of spatial genetic structure. The frequency of close *P. falciparum* relatives (>0.50 IBD) decayed with geographic distance, and selected clonal genotypes could be linked to specific geographic subdivisions in Guyana. In *P. vivax*, significant correlation of relatedness with spatial distance did not occur among sample pairs separated by >30 km, and meaningful regional population structure could not be resolved. Limited macrographic structure in *P. vivax* may occur largely because the dormant hypnozoite liver stage unique to this species facilitates long-distance dispersal via relapse in mobile hosts. Various study results point to the influence of relapse on *P. vivax* dispersal in the study region. At health centers in Greater Georgetown, for example, an absence of travel 14 days prior to diagnosis was reported more than twice as frequently by *P. vivax* patients (178 of 1611 (11.0%)) than by *P. falciparum* patients (45 of 867 (5.2%)). Nucleotide diversity was also significantly elevated in *P. vivax* samples representing such non-traveler cases. Given that Greater Georgetown is considered free of malaria transmission, many non-traveler *P. vivax* cases may represent reactivation of dormant parasite diversity that has accumulated in hosts over longer periods of time⁴¹. We also observed the recurrence of *P. vivax* clones in 11 patients for which blood spots were incidentally sequenced from separate malaria visits. These same-patient sample pairs represented symptomatic malaria episodes separated by a median of 114 days and the prescription of a 3 day Chloroquine + 14 day primaquine treatment course. They may thus not only represent direct evidence of relapse but also suggest efficacy and adherence challenges relating to hypnozoicidal drugs. Together, these observations suggest that differences in underlying parasite biology are contributing to divergent species responses to intervention, including flips from *P. falciparum* to *P. vivax* case majority observed over the last 15 years in various regions nearing elimination goals (e.g., Solomon islands, Myanmar, Cambodia, Lao PDR¹).

In the case of *P. falciparum* in Guyana, malaria control strategy may benefit from further advancing decentralized case management and differentiating resource allocation based on real-time spatio-temporal surveillance. Similar to regions in Southeast Asia where these

approaches have recently reduced *P. falciparum* cases to historic lows⁴², our study indicated a *P. falciparum* metapopulation maintained by periodic expansions from high-risk transmission areas (mines in remote riparian forest) where infrastructure is limited and occupational incentives conflict with self-care. In these settings, treatment-seeking is more likely to be delayed (increasing the risk of transmission prior to treatment) and access to regional treatment centers may involve long-distance travel through endemic terrain. It may therefore be beneficial to further decentralize medical capacity from regional treatment centers to localities with strong outbreak signals (e.g., Puruni Landing) or to strategic points along travel routes that may facilitate parasite dispersal (e.g., along the Mazaruni and Cuyuni Rivers, possible conduits for clonal *P. falciparum* dispersal toward Bartica based on this study). Self-administered diagnostic and treatment approaches (e.g., Malakit^{43,44}) might additionally help cover remote areas, with monitoring to evaluate the extent to which incorrect self-treatment may negatively impact user health or contribute to resistance emergence.

In the case of *P. vivax* control, our study suggests that similar strategies of decentralized and focally intensified intervention will be less efficacious. We do not observe tractable metapopulation structure that would help prioritize intervention foci, and this lack of structure points to the risk of latent dispersal processes undermining geographically differentiated intervention. Results emphasize the importance of radical cure (treatment combining blood schizonticidal and liver hypnozoicidal drugs⁴⁵) to help reduce *P. vivax* spread, but implementing radical cure in Guyana's remote transmission contexts poses significant challenge. Primaquine, currently the only hypnozoicidal drug specified in Guyana's National Treatment Guidelines, must be taken daily for 2 weeks and requires prior glucose-6-phosphate dehydrogenase deficiency assessment of hemolysis risk. Field trials on new implementation strategies (e.g., CUREMA⁴⁶) and alternative hypnozoicidal treatments (e.g., single-dose tafenoquine) tailored to remote mining settings are therefore important. If safety and adherence concerns can be resolved, the concept of universal radical cure (extended hypnozoicidal treatment not only for *P. vivax* but also for *P. falciparum* patients) also deserves consideration. Universal radical cure has been proposed for transmission contexts in which *P. falciparum* patients are likely to carry hypnozoites from earlier *P. vivax* infection⁴⁷. This condition clearly appears relevant to Guyana given its high levels of spatiotemporal co-endemicity and clear risk profile for co-infection (adults of non-African descent involved in the mining sector) described in this study.

Fortifying preventive vector control and risk awareness programs is also important. A recent study suggests that increased risk perception is associated with increased preventive behavior against vector-borne disease in Guyana, although effect size is small (4–5%)⁴⁸. Effect size might be increased if educational programs more strongly convey risk of onward malaria transmission as opposed to risk solely towards personal health. Such altruistic risk perception may be relevant to the more rapid, pre-symptomatic spread of *P. vivax* malaria into local vectors, including upon miner return to non-mining home environments.

This study had several limitations. Importantly, the near-relative detection approach we used to describe parasite population structure focuses on outlier IBD relationships and therefore underuses available sequence content. We also did not attempt IBD estimation involving polyclonal infections⁴⁹, which constituted 17.9% of our genomic sample set. Furthermore, genomic sample collection relied on a passive sentinel method which may bias against parasite diversity found in remote, disconnected transmission cycles (e.g., Amerindian settlements) and against low-parasitemia infections. It was also challenging to address the transitory nature of human settlement in the hinterlands of Guyana, where conventional population sizes are not recorded or do not apply. This primarily complicated the estimation of

zonal incidence and its relationship to polyclonality rates. Host movement through undercharted hinterland areas also added uncertainty to geographic infection source estimates.

Despite these limitations, this study clearly exposes the discrepant profiles of co-endemic *P. falciparum* and *P. vivax* malaria epidemiology in a low transmission context and emphasizes the need for distinct approaches to surveillance and intervention. It has been very important to prioritize investment in drugs, vaccines, and monoclonal antibodies for *P. falciparum* as a more frequent cause of acute severe disease, especially in children. Weaker focus on *P. vivax* control and specialized intervention however makes it likely that malaria will persist in many co-endemic regions after *P. falciparum* is eliminated, forestalling the ultimate goal of eliminating this major disease from all regions and demographics.

Methods

Epidemiological database and definitions

The epidemiological database analyzed in this study comprised all malaria episodes detected in Guyana via passive surveillance (estimated to represent ~80% of cases¹) and reported to the National Malaria Program between 2006 and 2019. Malaria episodes were defined using both medical diagnoses and parasitological tests (blood microscopy and rapid diagnostic tests). Patient records were curated by the Vector Control Services (VCS) division of the GMOH and anonymized records of patient age, gender, nationality, and self-reported ethnicity (categorized as 'Afro-Guyanese', 'Amerindian', 'East Indian', 'Chinese', 'Portuguese', 'European', 'Mixed', or 'Other') were provided to the Harvard T.H. Chan School of Public Health (HSPH) following approval by the Harvard University Area Human Research Protection Program (protocol IRB18-1638) and by ethical committees of the GMOH. Infection localities were designated by asking 'where [the] patient stayed 2 weeks ago' (see survey form in Supplementary Fig. 23). Each voluntary travel history response was classified to one of 460 country-wide localities with known latitude/longitude coordinates. Infection localities were entered as missing (NA) if patients provided no response on prior stay (i.e., infection localities were not equated to diagnosis localities unless patients specified having stayed in the same locality 2 weeks ago). For foreign localities of prior stay (e.g., in neighboring Venezuela), geographic specifics were not retained in the database. The prior stay field simply shows the term 'imported' for these cases, which we re-code as 'foreign' in our analysis. To define epidemiological zones, we mapped malaria survey sites used by the GMOH (<https://gazetteer.gls.gov.gy/gazetteer/#7/3.650/-57.129>) onto a custom shape file containing Guyana's primary river and road coordinates and onto a motorized transport resistance raster obtained from <https://malariaatlas.org/> (Supplementary Fig. 2). Sites were clustered based on river/road connectivity in the R package RIVERDIST⁵⁰ v0.16.3, travel conductance using the R package GDISTANCE⁵¹ v1.6.4, and manual assessment of river/road and resistance layers in QGIS⁵² v2.18.4. Transmission intensity (incidence) was estimated for each epidemiological zone using population size projections for 2019 and 2021 by the LandScan Project⁵³. LandScan applies dasymetric modeling to disaggregate census counts based on remotely sensed images⁵³. To estimate incidence for an epidemiological zone, we first cropped LandScan Global raster data (<https://landscan.ornl.gov/>) to a circle of 45 km radius around the centroid of all infection localities belonging to that zone. We then divided the case count (i.e., the number of *P. falciparum* and/or *P. vivax* cases representing patients which reported prior stay in a locality belonging to the zone) by the sum of all values projected within the cropped raster.

Clinical sample material, sequencing, and variant calling

Following informed consent to analyze parasite genetic polymorphism using patient blood, samples were collected from microscopy or RDT-positive patients by spotting ~50–200 µl whole blood onto Whatman

FTA filter paper cards. Collection was carried out by VCS between January 2020 and June 2021 at medical facilities in Bartica, Georgetown, Mahdia, Lethem, and Port Kaituma, Guyana. Samples were stored at room temperature using individual desiccant packets before shipment to HSPH. The same anonymized patient metadata variables as described in the previous section (Supplementary Fig. 23) were also provided to HSPH. An additional 23 dried blood spot samples were collected in Venezuela at the Institute of Tropical Medicine, Central University of Venezuela (Caracas, bioethics permit CEC-IMT 12/2013) and Biomedical Research and Therapeutic Vaccines Institute (Ciudad Bolívar, bioethics permit CHURPCBBS-008-2019) in 2019. Metadata for these samples (collection date and reported place of stay (municipality) 2 weeks prior to diagnosis) was likewise de-identified prior to HSPH access (Supplementary Data 1).

We extracted total genomic DNA from dried blood spot samples (~20–35 mm² spotted area punched per sample) using KingFisher Ready DNA Ultra 2.0 Prefilled Plates on the Kingfisher Flex instrument (ThermoFisher Scientific). We subsequently applied whole-genome amplification to DNA extracts, each 50 µl sample reaction consisting of 5 µl 10x phi29 polymerase buffer (NEB B0269S), 0.125 µl (2.5 µg) recombinant albumin (NEB B9200S), 0.5 µl primer mix (10 oligos combined at 250 µM, see below), 5 µl 10 mM dNTPs (Thermo Scientific), 26.375 nuclease-free water, 3 µl (30 units) phi29 DNA polymerase (M0269L), and 10 µl DNA. Reactions were prepared on ice, with components added in the order shown. Primer mixes consisted of 'Pfl-10' for *P. falciparum*⁵⁴ and 'Pvset1' for *P. vivax*⁵⁵ (set selection based on previous microscopy or RDT-based species assignments). The 3' regions of these primers contain phosphorothioate bonds. Amplifications were generated using step down incubation (35 °C for 5 min, 34 °C for 10 min, 33 °C for 15 min, 32 °C for 20 min, 31 °C for 30 min, 30 °C for 16 h, and 65 °C for 15 min), followed by cooling to 4 °C. We then applied AMPure XP magnetic beads (Beckman Coulter A63881) at room temperature to exchange post-reaction sample buffer to 10 mM Tris-HCl + 0.1 mM EDTA. Final library construction using the NEBNext Ultra II FS DNA Library Prep Kit (NEB E6177) and 2 × 151 bp sequencing on the Illumina NovaSeq 6000 platform was completed at the Broad Institute. We aligned reads to the *P. falciparum* 3D7 and *P. vivax* P01 reference genome assemblies using BWA-MEM⁵⁶ v0.7.17-r1188 and called SNPs and INDELs using GATK⁵⁷ v3.5.0-g36282e4 'Haplotype-Caller' and 'GenotypeGVCFs' according to best practices defined by the Pf3k consortium (http://www.malariagen.net/data_package/pf3k-5/). The mapping and joint variant call process also included sample accessions ERR1818176 and ERR2496572 representing malaria infections from Venezuela (further geographic specifics unknown). All downstream genetic analyses focused on SNP sites in core regions of the genome^{58,59} and >5 bp from any INDEL call. Samples missing genotype calls for >50% SNP sites at ≥2% minor allele frequency (after application of the core region filter and INDEL masking steps) were also excluded from further analysis. Among excluded samples, coverage (% bases represented by >5 reads) averaged 12.9 in *P. falciparum* and 9.9 in *P. vivax*. Following sample exclusions, we removed sites at which >7.5% of monoclonal samples showed multi-allelic ('heterozygous') calls. These filtration steps generated the base SNP call sets for each species. Additional, analysis-specific settings are noted below.

Parasite genetic analyses

We assessed polyclonality using THEREALMcCOIL³⁶ v2, a Bayesian Markov chain Monte Carlo approach in which posterior distributions for the population allele frequency of each locus and the COI of each individual are jointly estimated to maximize the likelihood of the data using a Metropolis-Hastings sampling approach. The probability of error (i.e., incorrectly calling a homozygous locus as heterozygous and vice versa) can also be estimated alongside allele frequencies and COI by specifying `err_method = 3`. We applied the categorical method (i.e., `input_data = SNP_calls` classified binarily as homozygous or

heterozygous regardless of the signal intensity of each component allele) with `err_method = 3` to bi-allelic SNP sites with $\geq 10\%$ minor allele frequency and $\geq 90\%$ call success across samples. We further subsampled input to limit computational costs. Subsampling using the ‘–thin’ function in VCFtools⁶⁰ v0.1.15 resulted in 337 *P. falciparum* and 350 *P. vivax* input sites. We used the lower bound (quantile = 2.5%) output value as a conservative estimate of sample COI. Estimates were consistent across repeat runs (100% matching calls) and when reducing input sites to $n = 327$ for both species by further filtering out sites with ≥ 200 mean and ≥ 150 sd read-depth (99.2% matching calls). Analyses of genetic diversity and relatedness only included monoclonal samples. We estimated nucleotide diversity from the base SNP call sets (previous section) by measuring average nucleotide differences over discrete 100 kbp windows (no overlap) with VCFtools⁶⁰ v0.1.15. We measured linkage decay over bi-allelic SNP sites by recoding calls to non-reference allele counts (0, 1, or 2) and computing linkage (r^2) in sliding 10 kbp windows (200 bp steps) in PLINK⁶¹ v1.90b1g.

We estimated genetic relatedness between samples based on the concept of IBD. Unlike identity by state (IBS), IBD specifically represents genetic similarity inferred to have been inherited from a common ancestor based on factors affecting linkage likelihood such as population allele frequency, recombination rate, and chromosomal distances between variant sites. IBD generally enhances representation of recent ancestry in obligately sexual species in which divergence history is influenced more strongly by recombination than by the individual accumulation of point mutation events. We ran hmmIBD³⁵ v2.0.4 in default settings to estimate population allele frequency and infer presence or absence of IBD at all bi-allelic SNPs with $\geq 2\%$ minor allele frequency and $\geq 70\%$ call success across samples. We used the fraction of all variant sites inferred as IBD (‘`fract_sites_IBD`’) to summarize relatedness for each sample pair.

We used the R package IGRAPH⁶² v1.3.5 to generate connected graphs (clusters) in which edges represent sample pairs with relatedness values above a specified threshold. Clusters therefore represent groups of samples (nodes) in which each sample is connected by at least one direct edge to another sample (but simultaneous direct connections to additional cluster members are not required).

We also correlated relatedness values to Euclidean distances calculated from latitude and longitude (WGS 84) coordinates projected onto a common xy plane (EPSG 3786). We used Mantel tests to assess statistical significance between genetic and spatial distance matrices using the R package VEGAN⁶³ v2.6.4. We applied repeated random draws of 10 samples from each epidemiological zone (drawing until depletion without replacement) to build each distance matrix. We also mapped the frequency of pairwise relatedness between epidemiological zones using shape files from the Database of Global Administrative Boundaries v4.1⁶⁴.

Finally, we also assessed the possibility that cases of clonal IBD detection in *P. vivax* represent repeated blood spot sampling events from the same patient (e.g., a first clinical visit due to primary infection and one or more later clinical visits due to relapse or recrudescence by the same parasite genotype). HSPH authors sent a query list of 705 *P. vivax* samples to the GMOH and the GMOH flagged each according to whether it represents a patient which is also represented by another sample in the query list. The list returned by the GMOH indicated 1x (i.e., non-repeat) patient representation for 589 samples, 2x patient representation for 92 samples (necessarily representing 46 patients), and $\geq 3x$ patient representation for 22 samples (necessarily representing ≤ 7 patients). The returned list did not specify which sample sets corresponded to the same patient.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The sequence data generated by this study has been deposited in the NCBI Sequence Read Archive (<https://www.ncbi.nlm.nih.gov/sra>) under BioProjects PRJNA809659 and PRJNA1092690. The epidemiological case data (2019) is however not directly available for public use. This reflects a data use agreement with the Guyana Ministry of Health which protects against the potential for subject identification and other privacy/ethics issues. Interested parties may contact the Guyana Ministry of Health (<https://www.health.gov.gy/index.php/contact-us>) for data access through a data use agreement. All inquiries will be addressed within a reasonable timeframe (3 months or less).

Code availability

Analysis scripts and files are available at <https://doi.org/10.5281/zenodo.1335153>.

References

- World Health Organization. *WHO Malaria Report 2023* (WHO, 2023).
- Weiss, D. J. et al. Mapping the global prevalence, incidence, and mortality of *Plasmodium falciparum*, 2000–17: a spatial and temporal modelling study. *Lancet Lond. Engl.* **394**, 322–331 (2019).
- Battle, K. E. et al. Mapping the global endemicity and clinical burden of *Plasmodium vivax*, 2000–17: a spatial and temporal modelling study. *Lancet Lond. Engl.* **394**, 332–343 (2019).
- Price, R. N., Commons, R. J., Battle, K. E., Thriemer, K. & Mendis, K. *Plasmodium vivax* in the era of the shrinking *P. falciparum* map. *Trends Parasitol.* **36**, 560–570 (2020).
- Oliveira-Ferreira, J. et al. Malaria in Brazil: an overview. *Malar. J.* **9**, 115 (2010).
- Rosas-Aguirre, A. et al. Epidemiology of *Plasmodium vivax* malaria in Peru. *Am. J. Trop. Med. Hyg.* **95**, 133–144 (2016).
- Sattabongkot, J., Tsuboi, T., Zollner, G. E., Sirichaisinthop, J. & Cui, L. *Plasmodium vivax* transmission: chances for control? *Trends Parasitol.* **20**, 192–198 (2004).
- Sattabongkot, J. et al. Malaria research for tailored control and elimination strategies in the greater Mekong subregion. *Am. J. Trop. Med. Hyg.* **107**, 152–159 (2022).
- Kenangalem, E. et al. Malaria morbidity and mortality following introduction of a universal policy of artemisinin-based treatment for malaria in Papua, Indonesia: a longitudinal surveillance study. *PLoS Med.* **16**, e100281510 (2019).
- Sharp, P. M., Plenderleith, L. J. & Hahn, B. H. Ape origins of human malaria. *Annu. Rev. Microbiol.* **74**, 39–63 (2020).
- Cepeda, A. S. et al. The genome of *Plasmodium gonderi*: insights into the evolution of human malaria parasites. *Genome Biol. Evol.* **16**, evae027 (2024).
- Cowman, A. F., Healer, J., Marapana, D. & Marsh, K. Malaria: biology and disease. *Cell* **167**, 610–624 (2016).
- Krotoski, W. A. et al. Demonstration of hypnozoites in sporozoite-transmitted *Plasmodium vivax* infection. *Am. J. Trop. Med. Hyg.* **31**, 1291–1293 (1982).
- Llanos-Cuentas, A. et al. Tafenoquine versus primaquine to prevent relapse of *Plasmodium vivax* malaria. *N. Engl. J. Med.* **380**, 229–241 (2019).
- Recht, J., Ashley, E. A. & White, N. J. Use of primaquine and glucose-6-phosphate dehydrogenase deficiency testing: divergent policies and practices in malaria endemic countries. *PLoS Negl. Trop. Dis.* **12**, e0006230 (2018).
- Bruxvoort, K., Goodman, C., Kachur, S. P. & Schellenberg, D. How patients take malaria treatment: a systematic review of the literature on adherence to antimalarial drugs. *PLoS One* **9**, e84555 (2014).
- Craik, R. A note on the erythrocytes in malaria. *Lancet* **195**, 1110 (1920).

18. Cheng, Q., Cunningham, J. & Gatton, M. L. Systematic review of sub-microscopic *P. vivax* infections: prevalence and determining factors. *PLoS Negl. Trop. Dis.* **9**, e3413 (2015).
19. Boyd, M. F. & Kitchen, S. F. On the infectiousness of patients infected with *Plasmodium vivax* and *Plasmodium falciparum*. *Am. J. Trop. Med. Hyg.* **s1-17**, 253–262 (1937).
20. Wernsdorfer, W. H. *Malaria: Principles and Practice of Malariology*. Vol. 2092 (Edinburgh; New York, 1988).
21. Boyd, M. F., Stratman-Thomas, W. K. & Kitchen, S. F. On the relative susceptibility of *Anopheles quadrimaculatus* to *Plasmodium vivax* and *Plasmodium falciparum*. *Am. J. Trop. Med. Hyg.* **s1-15**, 485–493 (1935).
22. Boyd, M. F., Stratman-Thomas, W. K. & Muench, H. The occurrence of gametocytes of *Plasmodium vivax* during the primary attack. *Am. J. Trop. Med. Hyg.* **s1-16**, 133–138 (1936).
23. Pukrittayakamee, S. et al. Effects of different antimalarial drugs on gametocyte carriage in *P. vivax* malaria. *Am. J. Trop. Med. Hyg.* **79**, 378–384 (2008).
24. Galinski, M. R., Meyer, E. V. S. & Barnwell, J. W. *Plasmodium vivax*: modern strategies to study a persistent parasite's life cycle. In: *Advances in Parasitology* Vol. 81 (eds. Hay, S. I., Price, R. & Baird, J. K.) 1–26 (Academic Press, 2013).
25. Timinao, L. et al. Infectivity of symptomatic malaria patients to *Anopheles farauti* colony mosquitoes in Papua new Guinea. *Front. Cell. Infect. Microbiol.* **11**, 771233 (2021).
26. Blackburn, D. et al. Outbreak of locally acquired mosquito-transmitted (autochthonous) malaria — Florida and Texas, May–July 2023. *Morb. Mortal. Wkly. Rep.* **72**, 973–978 (2023).
27. Danis, K. et al. Malaria in Greece: historical and current reflections on a re-emerging vector borne disease. *Travel Med. Infect. Dis.* **11**, 8–14 (2013).
28. Bahk, Y. Y. et al. Epidemiological characteristics of re-emerging vivax malaria in the republic of Korea (1993–2017). *Korean J. Parasitol.* **56**, 531–543 (2018).
29. Lopez, L. & Koepfli, C. Systematic review of *Plasmodium falciparum* and *Plasmodium vivax* polyclonal infections: impact of prevalence, study population characteristics, and laboratory procedures. *PLoS ONE* **16**, e0249382 (2021).
30. De Salazar, P. M., Cox, H., Imhoff, H., Alexandre, J. S. F. & Buckee, C. O. The association between gold mining and malaria in Guyana: a statistical inference and time-series analysis. *Lancet Planet. Health* **5**, e731–e738 (2021).
31. Howes, R. E. et al. *Plasmodium vivax* transmission in Africa. *PLoS Negl. Trop. Dis.* **9**, e0004222 (2015).
32. Olapeju, B. et al. Malaria prevention and care seeking among gold miners in Guyana. *PLoS ONE* **15**, e0244454 (2020).
33. Hilson, G. & Laing, T. Gold mining, indigenous land claims and conflict in Guyana's hinterland. *J. Rural Stud.* **50**, 172–187 (2017).
34. Neafsey, D. E. et al. The malaria parasite *Plasmodium vivax* exhibits greater genetic diversity than *Plasmodium falciparum*. *Nat. Genet.* **44**, 1046–1050 (2012).
35. Schaffner, D. F., Taylor, A. R., Wong, W., Wirth, D. F. & Neafsey, D. E. hmmlBD: software to infer pairwise identity by descent between haploid genotypes. *Malar. J.* **17**, 196 (2018).
36. Chang, H. H. et al. THE REAL McCOIL: a method for the concurrent estimation of the complexity of infection and SNP allele frequency for malaria parasites. *PLoS Comput. Biol.* **13**, e1005348 (2017).
37. Vanhove, M. et al. Temporal and spatial dynamics of *Plasmodium falciparum* clonal lineages in Guyana. *PLoS Pathog.* **20**, e1012013 (2024).
38. Hubbard, A. et al. Implementing landscape genetics in molecular epidemiology to determine drivers of vector-borne disease: a malaria case study. *Mol. Ecol.* **32**, 1848–1859 (2023).
39. Verity, R. et al. The impact of antimalarial resistance on the genetic structure of *Plasmodium falciparum* in the DRC. *Nat. Commun.* **11**, 2107 (2020).
40. Tessema, S. et al. Using parasite genetic and human mobility data to infer local and cross-border malaria connectivity in Southern Africa. *eLife* **8**, e43510 (2019).
41. Popovici, J. et al. Genomic analyses reveal the common occurrence and complexity of *Plasmodium vivax* relapses in Cambodia. *mBio* **9**, e01888–17 (2018).
42. Manzoni, G. et al. Progress towards malaria elimination in the Greater Mekong Subregion: perspectives from the World Health Organization. *Malar. J.* **23**, 64 (2024).
43. Douine, M. et al. Malakit: an innovative pilot project to self-diagnose and self-treat malaria among illegal gold miners in the Guiana shield. *Malar. J.* **17**, 158 (2018).
44. Douine, M. et al. Self-diagnosis and self-treatment of malaria in hard-to-reach and mobile populations of the Amazon: results of Malakit, an international multicentric intervention research project. *Lancet Reg. Health Am.* **4**, 100047 (2021).
45. Thriemer, K., Ley, B. & von Seidlein, L. Towards the elimination of *Plasmodium vivax* malaria: implementing the radical cure. *PLoS Med.* **18**, e1003494 (2021).
46. Sanna, A. et al. CUREMA project: a further step towards malaria elimination among hard-to-reach and mobile populations. *Malar. J.* **23**, 271 (2024).
47. Thriemer, K. et al. Primaquine radical cure in patients with *Plasmodium falciparum* malaria in areas co-endemic for *P. falciparum* and *Plasmodium vivax* (PRIMA): a multicentre, open-label, superiority randomised controlled trial. *Lancet Lond. Engl.* **402**, 2101–2110 (2023).
48. Lopes-Rafegas, I., Cox, H., Mora, T. & Sicuri, E. The contribution of risk perception and social norms to reported preventive behaviour against selected vector-borne diseases in Guyana. *Sci. Rep.* **13**, 16866 (2023).
49. Gerlovina, I., Gerlovin, B., Rodríguez-Barraquer, I. & Greenhouse, B. Dcifer: an IBD-based method to calculate genetic distance between polyclonal infections. *Genetics* **222**, iyac126 (2022).
50. Tyers, M. riverdist: river network distance computation and applications. R package version 0.16.3 (2024).
51. van Etten, J., de Sousa, K. & Marx, A. gdistance: distances and routes on geographical grids. R package version 1.6.4 (2023).
52. QGIS Development Team. Open Source Geospatial Foundation. Software version 2.18.4 (2017).
53. Rose, A. et al. LandScan Global 2019. Oak Ridge National Laboratory, Oak Ridge, TN (2020).
54. Oyola, S. O. et al. Whole genome sequencing of *Plasmodium falciparum* from dried blood spots using selective whole genome amplification. *Malar. J.* **15**, 597 (2016).
55. Cowell, A. N. et al. Selective whole-genome amplification is a robust method that enables scalable whole-genome sequencing of *Plasmodium vivax* from unprocessed clinical samples. *mBio* **8**, e02257–16 (2017).
56. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
57. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
58. Miles, A. et al. Indels, structural variation, and recombination drive genomic diversity in *Plasmodium falciparum*. *Genome Res.* **26**, 1288–1299 (2016).
59. de Oliveira, T. C. et al. Population genomics reveals the expansion of highly inbred *Plasmodium vivax* lineages in the main malaria hotspot of Brazil. *PLoS Negl. Trop. Dis.* **14**, e0008808 (2020).
60. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).

61. Purcell, S. et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
62. Csárdi, G. et al. igraph for R: R interface of the igraph library for graph theory and network analysis. R package version 1.3.5 (2022).
63. Oksanen, J. et al. Vegan: community ecology package. R package version 2.6-4 (2022).
64. GADM. Database of Global Administrative Areas version 4.1 (2018).

Acknowledgements

We thank the participants who contributed blood samples to the study, as well as the technicians who collected and processed the samples. This study was supported by the Bill & Melinda Gates Foundation (INV-009416 to D.N. and C.B.). Under the grant conditions of the Foundation, a Creative Commons Attribution 4.0 Generic License has been assigned to the Author Accepted Manuscript version that might arise from this submission. This study was also 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. Sampling in Venezuela was additionally supported by the Scottish Funding Council Global Challenges Research Fund (GCRF) Small Grants Fund SFC/AN/12/2017 and the GCRF Vector-borne Disease Control Network (EP/T003782/1) led by the Llewellyn Lab (University of Glasgow).

Author contributions

P.S., F.C., D.N., and C.B. conceived the study. H.C., R.N., D.F., O.N., M.G., and M.L. managed sample collection and logistics. F.A., N.S., and K.J. provided oversight. C.C. curated the epidemiological database. P.S. and D.N. supervised laboratory work. P.S. and F.C. analyzed the data. A.E. and M.V. assisted in analysis and interpretation. P.S. wrote the original manuscript. F.C. and D.N. helped review and edit additional drafts. All authors reviewed and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-52545-6>.

Correspondence and requests for materials should be addressed to Daniel E. Neafsey.

Peer review information *Nature Communications* thanks Jonathan Parr, who co-reviewed with Isabela Gerdes Gyuricza, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

¹Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, MA, USA. ²Infectious Disease and Microbiome Program, Broad Institute of MIT and Harvard, Cambridge, MA, USA. ³Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA. ⁴Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute, Basel, Switzerland. ⁵University of Basel, Basel, Switzerland. ⁶National Malaria Program, Ministry of Health, Georgetown, Guyana. ⁷Biomedical Research and Therapeutic Vaccines Institute, Ciudad Bolívar, Venezuela. ⁸Institute of Tropical Medicine, Faculty of Medicine, Central University of Venezuela, Caracas, Venezuela. ⁹Center for Malaria Research, Institute of Higher Studies 'Dr. Arnoldo Gabaldón', Ministry of Popular Power for Health, Caracas, Venezuela. ¹⁰Institute of Zoology and Tropical Ecology, Central University of Venezuela, Caracas, Venezuela. ¹¹Caribbean Public Health Agency, Port of Spain, Trinidad and Tobago. ¹²These authors contributed equally: Philipp Schwabl, Flavia Camponovo, Collette Clementson.

✉ e-mail: neafsey@broadinstitute.org