Understanding *Plasmodium vivax* recurrent infections using an amplicon deep sequencing assay, PvAmpSeq, identity-by-descent and model-based classification

Authors and Affiliations

Jason Rosado[1,2]*, Jiru Han[3,4], Thomas Obadia[1,5], Jacob Munro[3,4], Zeinabou Traore[1], Kael Schoffer[4], Jessica Brewster[4], Caitlin Bourke[4], Joseph M. Vinetz[6,7], Michael White[1], Melanie Bahlo[3,4], Dionicia Gamboa[7,8,9], Ivo Mueller[1,4,10ɣ], Shazia Ruybal-Pesántez[4,10,11*ɣ^]

1. Institut Pasteur, Université Paris Cité, Infectious Diseases Epidemiology and Analytics G5 Unit, Institut Pasteur, Paris, France
2. Université Paris Cité, Institut de Recherche pour le Développement, MERIT, Paris, France
3. Genetics and Gene Regulation Division, Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia
4. Infection and Global Health Division, Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia
5. Institut Pasteur, Université Paris Cité, Bioinformatics and Biostatistics Hub, Institut Pasteur, Paris, France
6. Section of Infectious Diseases, Department of Internal Medicine, Yale School of Medicine, New Haven, CT, USA
7. Instituto de Medicina Tropical Alexander von Humboldt, Universidad Peruana Cayetano Heredia, Lima, Peru
8. Laboratorio de Malaria, Laboratorios de Investigación y Desarrollo, Facultad de Ciencias e Ingeniería, Universidad Peruana Cayetano Heredia, Lima, Peru
9. Departamento de Ciencias Celulares y Moleculares, Facultad de Ciencias e Ingeniería, Universidad Peruana Cayetano Heredia, Lima, Peru
10. Department of Medical Biology, School of Medicine and Health Sciences, University of Melbourne, Melbourne, Australia
11. Vector-borne Diseases and Tropical Public Health, Burnet Institute, Melbourne, Australia

***Correspondence**

E-mail: jjrosados@gmail.com; s.ruybal@imperial.ac.uk

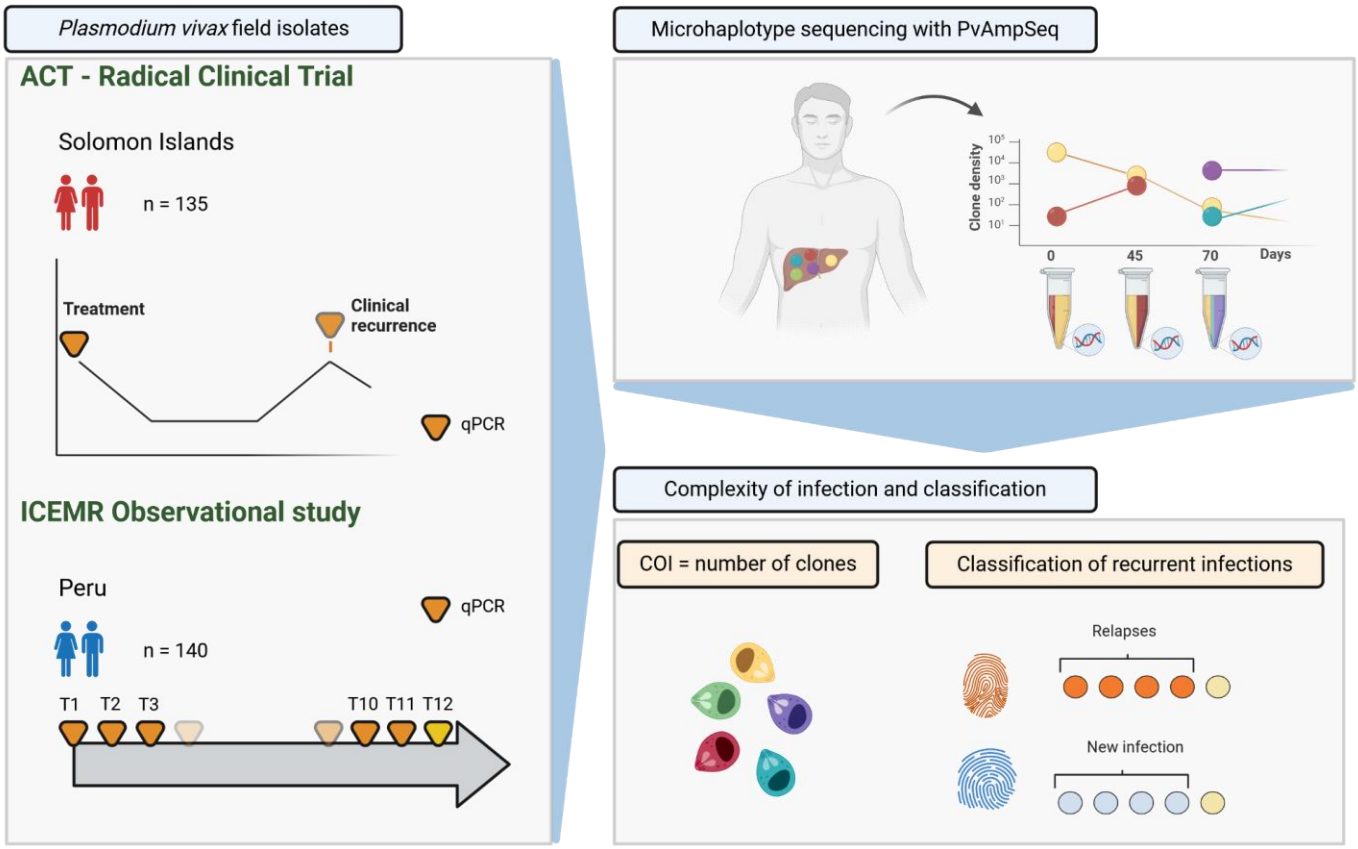
ɣ: Co-senior authors contributed equally to this work.

^ Current affiliations: MRC Centre for Global Infectious Disease Analysis, School of Public Health, Imperial College London, London, UK, and Institute of Microbiology, Universidad San Francisco de Quito, Quito, Ecuador

Summary

Plasmodium vivax infections are characterised by recurrent bouts of blood-stage parasitemia. Understanding the genetic relatedness of recurrences can distinguish whether these are caused by relapse, reinfection, or recrudescence, which is critical to understand treatment efficacy and transmission dynamics. We developed PvAmpseq, an amplicon sequencing assay targeting 11 SNP-rich regions of the *P. vivax* genome. PvAmpSeq was validated on field isolates from a clinical trial in the Solomon Islands and a longitudinal observational cohort in Peru, and statistical models were applied for genetic classification of infection pairs. In the Solomon Islands trial, where participants received antimalarials at baseline, half of the recurrent infections were caused by parasites with >50% relatedness to the baseline infection, with statistical models classifying 25% and 25% as probable relapses and recrudescences, respectively. In the Peruvian cohort, 26% of recurrences were likely relapses. PvAmpSeq provides high-resolution genotyping to characterise *P. vivax* recurrences, offering insights into transmission and treatment outcomes.

Keywords: Plasmodium vivax, recurrences, relapses, microhaplotypes, genetic diversity



Graphical abstract. Created in https://BioRender.com

Introduction

Plasmodium vivax causes an estimated 9.2 million cases per year and is the most widespread *Plasmodium* parasite infecting humans, and is endemic in 41 countries.[1] The ability of *P. vivax* to remain in the liver as dormant hypnozoites and to cause subsequent (and often multiple) blood-stage infections contributes to onward transmission and impedes efforts to eliminate malaria.[2] In areas endemic for *P. vivax,* recurrent *P. vivax* infections can result from relapses but also recrudescence (due to antimalarial treatment failure if the individual received treatment) or reinfections (from a new mosquito bite). Distinguishing between relapses, recrudescences, and reinfections is crucial for the assessment of the effectiveness of antimalarial-based control strategies and understanding *P. vivax* biology and epidemiology, but remains a major challenge.

The WHO Global Malaria Programme has recently highlighted the importance of using more sensitive, easily implemented, and reproducible tools that allow for estimating the efficacy of clinical drug trials.[3] In the case of *P. vivax*, drug efficacy trials conducted in endemic areas face the challenge of establishing whether recurrent parasitemia after treatment is due to treatment failure (recrudescence), relapse, or reinfection. Comparing parasite genotypes of baseline infections with the ones from recurrent infections allows us to better estimate drug failure rates through 'molecular correction' broadly based on the identification of the same, closely related, or distinct genotypes. Unlike *P. falciparum*, where genotyping must distinguish between recrudescences from reinfections, genotype data from *P. vivax* recurrent infections must be able to discriminate recrudescences from both reinfections and relapses for evaluation of blood-stage treatment efficacy. In addition, to evaluate the efficacy of anti-hypnozoite treatments (e.g., primaquine), it is critical to be able to resolve relapses (anti-hypnozoite treatment failure) from reinfections. Classically, genotyping infections for 'molecular correction' in drug-efficacy clinical trials has been done with microsatellite markers,[4] but more recently with amplicon sequencing approaches.[3,5,6] A limitation of microsatellite genotyping can be the insufficient resolution to detect minority clones in infections. [7] This can result in an overestimation of treatment efficacy in clinical trials as the undetected clones could be resistant parasites, recrudescent clones, or relapsing hypnozoites not detected at baseline.

The sequencing of Single Nucleotide Polymorphism (SNPs) or SNP-rich amplicons by next-generation sequencing approaches (Amplicon deep sequencing or AmpSeq) has been shown to overcome the drawbacks of microsatellites.[8] For example, i) detection of multiple SNPs in the same read allows direct detection of haplotypes, ii) sufficient sensitivity to detect minor clones and track clone dynamics over time in *Plasmodium* infections,[7,9–11] and iii) high

reproducibility. For *P. falciparum,* AmpSeq genotyping has been shown to improve the classification of recrudescence and reinfections in clinical trials.[5,12]

In addition to their use for evaluation of treatment efficacy, AmpSeq markers are also being used to identify imported malaria cases,[13] estimate transmission levels in population studies,[14] detect drug-resistant parasites,[14–19] and characterise *Plasmodium* population genetics[18,20–25]. In addition, *P. vivax* genome-wide panels of Ampseq markers and Molecular Inversion Probes (MIPs) have recently been developed for population genetics, geographic origin assignment, and detection of SNPs associated with resistance.[6,18,26–30] In a recent study by Kleinecke and colleagues, a panel of 93 genome-wide AmpSeq microhaplotypes was applied in a clinical trial to classify 108 primary and recurrent infection pairs using identity-by-descent (IBD) estimates.[6] The study showed a higher frequency of suspected relapses or recrudescence (high IBD) (84%) in patients treated with primaquine compared to those without primaquine (60%).

To advance our understanding of *P. vivax* recurrent infections, we developed a *P. vivax* AmpSeq (PvAmpSeq) assay that targets 11 highly polymorphic SNP-rich regions or microhaplotypes (across 11 chromosomes) and applied several recently developed approaches for studying infections at the clone level using IBD and probabilistic models. Here, we show the validation of PvAmpSeq for two applications: on samples from a clinical trial where genetic classification of recurrent infections is necessary for molecular correction of drug failure rates and from a community cohort where genetic classification is useful to understand *P. vivax* epidemiology and transmission patterns. We used samples from the ACT-Radical randomised-controlled clinical trial conducted in the Solomon Islands between 2018 and 2019, involving patients receiving three different drug combinations, with the overall aim to test for potential antagonistic effects of the drug combinations (James, Obadia *et al*. *under review*). Two of the trial arms included drug combinations involving the anti-hypnozoite antimalarial drug primaquine (PQ), where patients were treated to clear both blood-stage parasites and hypnozoites at baseline. We also used samples from an observational longitudinal cohort conducted in the Loreto region in the Peruvian Amazon, where individuals were followed up between December 2014 and December 2015.[31] We processed PvAmpSeq sequencing data using the AmpSeqR R package to demultiplex and analyse AmpSeq data[32] and used various approaches for genetic classification of recurrent infections. Overall, our findings point to the potential for PvAmpSeq to provide important insights about recurrent infection dynamics in both clinical trials and epidemiological studies.

Methods

Ethics

Ethical approval for the Solomon Islands ACT-Radical clinical trial was obtained from the Solomon Islands Health Research and Ethics Review Board (HRE 041-16) and the Walter and Eliza Hall Institute of Medical Research Human Research Ethics Committee (WEHI 16-08). The Peruvian cohort was approved by the Institutional Ethics Committee from the Universidad Peruana Cayetano Heredia (UPCH) (SIDISI 57395/2013) and from the University of California San Diego Human Subjects Protection Program (Project # 100765). UPCH also approved the use of the DNA samples of *Plasmodium vivax* infections at the Institut Pasteur (SIDISI 100873/2017).

Study sites and samples

Samples from the Solomon Islands were collected as part of the ACT-Radical clinical trial (registered under Australia and New Zealand Clinical Trials Registry ANZCTR 12617000329369) conducted between September 2017 and February 2019. Briefly, 374 individuals were enrolled in the study, 82% (307/374) of participants had a primary symptomatic *P. vivax* infection confirmed by qPCR and were treated with artemether-lumefantrine (AL), AL plus primaquine (AL+PQ), or dihydroartemisinin-piperaquine plus primaquine (DP+PQ). Of these, 307 individuals were actively followed for up to 168 days or until a recurrent *P. vivax* infection was confirmed either by light microscopy or qPCR. 191 individuals had *P. vivax* recurrent infections confirmed by light microscopy or qPCR. Of the positive blood samples, 218 (from 91 participants) were available for PvAmpseq evaluation. Blood samples were collected in EDTA tubes for the baseline infections (n = 91), whereas recurrent infection samples were collected as dried blood spots (DBS) (n = 137).

Details on the Peruvian cohort have been previously reported.[31,33] Briefly, a three-year-long observational cohort study was conducted in Peru from December 2012 to December 2015 in two Amazonian villages in the Loreto Region: San José de Lupuna and Cahuide.[33] Using home-to-home and community-based screening, volunteers ≥ 3 years old were invited to participate in this cohort. Between December 2014 and December 2015, 1083 out of 7612 blood samples (14.2%) were positive for *P. vivax* by qPCR. Of these positive samples, 449 DNA *P. vivax* positive samples (from 176 participants) were available at Institut Pasteur for genomic studies. As these samples were originally diagnosed by SYBR Green-based qPCR methods and most of them had low parasite density infections (1.55 parasites/µL (IQR: 0.74–8.38)),[31] we performed a diagnostic Taqman qPCR to downselect amplifiable samples. This selection resulted in 274 DNA samples from 152 participants with at least 1 qPCR positive result for *P. vivax* in the last 12 months of follow-up (1 *P. vivax* infection = 65 participants, 2 *P. vivax* infections = 56 participants, 3 *P. vivax* infections = 24 participants, 4 *P. vivax* infections = 6 participants).

All cohort participants provided written informed consent for participation in both studies. Parental written consent and assent were obtained in the case of participants <18 years in the Peruvian cohort.

DNA extraction and qPCR

In the Solomon Islands ACT-Radical clinical trial, whole blood samples from participants at baseline were collected in EDTA tubes and conserved at -20°C until DNA extraction. Genomic DNA was extracted from 200 µL of whole blood using the Favorprep 96-well genomic DNA kit (Favorgen, cat # FADWE 96004, Taiwan) and following the manufacturer's recommendations. PBS buffer was added to the samples with insufficient volume (< 200 µL). Dried blood spots (DBS) from participants with recurrent infections were collected onto filter paper and let dry at room temperature. DBS samples were conserved at -20°C until DNA extraction. DBS samples were cut into 6mm diameter bloodspot discs using a hole punch. Five bloodspot discs per sample were utilised for DNA extraction using Favorprep 96-well genomic DNA kit (Favorgen, cat # FADWE 96004, Taiwan). Genomic DNA from whole blood and DBS was eluted in 50 µL of elution buffer and stored in 96-well plates at -30°C until their use.

In the Peruvian cohort, blood samples were collected by finger prick onto filter paper and left to dry at room temperature. DNA was isolated from DBS using the E.Z.N.A. Blood DNA Mini Kit (Omega Bio-tek, Inc., Norcross, GA, US), and molecular diagnosis was performed using the Mangold method.[34,35]

For parasite density quantification, we performed a duplex TaqMan qPCR assay combining specific primers and probes targeting the *18s rRNA* gene region for *P. falciparum* and *P. vivax* in a duplex reaction, as reported by Rosanas-Urgell A *et al*., with slight modifications.[36] The reaction was prepared in a total of 13 µL containing 1X GoTaq® Probe qPCR Master Mix (Promega, USA), 769 nM of forward and reverse primers, 384 nM of probe, 1.5 µL of Nuclease-free water, and 2 µL of DNA sample. The PCR conditions consisted of an initial denaturation at 95 °C for 2 min, followed by amplification for 45 cycles of 15 s at 95 °C and 1 min at 58 °C. The assay was run in a QuantStudio TM 5 Real-Time PCR system (Applied Biosystems, USA). The number of copies of 18s rRNA DNA/µL of DNA was determined by using a standard curve from a sevenfold serial dilution to 1:10 of a plasmid at concentrations of 1×10^5 copies/µL down to 1 copy/µL in nuclease-free water. Samples with late amplification (Cq value >40 & <44; <1 copy/µL) were confirmed by an extra run. *P. vivax* and *P. falciparum* primers and probes detect 3 copies of 18s rRNA DNA per genome. This implies that a concentration of 5 copies/µL in a sample is equivalent to approximately 1-2 parasites/µL.

Marker selection and primer design

A panel of highly informative amplicons was designed based on whole genome sequencing data from the MalariaGEN *Plasmodium vivax* Genome Variation Project accessed in June 2018 (PvGV).[37] Briefly, FASTQ files were downloaded from the SRA, and SNP/indel variant calling was performed according to GATK V4 best practices against the PvP01 reference genome. Of the 354 samples processed, 154 were excluded due to low coverage (< 5x median coverage), high SNP missingness (>10%), multi-clonality checks (Fws < 0.85) or being from a country with too few samples remaining (< 15), leaving 200 samples over 6 countries (Cambodia, Colombia, Mexico, Papua New Guinea, Peru and Thailand). Genomic regions were then excluded based on the coverage distribution across remaining samples as follows: the genome was divided into 1000 bp bins, and coverage was assessed with samtools bedcov[38] for both high mapping quality (HMQ, mapping quality >= 30) and low mapping quality (LMQ, mapQ <= 30) reads. Coverage was then normalised within samples by dividing by the sample median HMQ coverage. Genomic regions were then excluded if the median across samples was greater than 1.5 or less than 0.5 or if the median proportion of HMQ coverage was less than 0.5. This method resulted in the exclusion of the majority of the hypervariable sub-telomeric regions of PvP01. The remaining genome was then searched for all regions that contained at least 4 SNPs within 140 bp, with primer design attempted using Primer3[39] with an optimal length of 22 bp and optimal melting temperature of 60 °C, avoiding SNPs and indels in the primer region.

Pv AmpSeq assay

The PvAmpSeq assay amplifies 11 SNP-rich regions, or microhaplotype markers, located across 11 chromosomes. Detailed protocols are described in Text S1. Briefly, AmpSeq libraries were prepared after amplification using a nested PCR strategy. Due to the low parasitemia of *P. vivax* infections frequently found in field samples, parasite genomic regions encompassing marker-specific sites were enriched by multiplex primary PCRs (pPCR) with 25 cycles of amplification. Individual secondary nested PCRs (nPCR) were performed using marker-specific primers with an overhang sequence in the 5' end and 25 cycles, enabling multiplexing of amplicons per sample. nPCR products were purified, quantified, and normalised at 15 ng/µL using QIAGEN MinElute 96 UF PCR Purification Kit (QIAGEN, Germany) and Quant-iT PicoGreen® dsDNA kit (Thermo Fisher Scientific, USA), respectively. Normalised nPCR products were pooled and performed index PCR (iPCR) using long fusion primers with P5/P7 adapters, index and overhang sequences in the backbone (Text S1), limited to 10 PCR cycles.

iPCR products were quantified and normalised at 20 ng/µL, then combined into pools of equal molarity. Long fusion primers (<200 nt) and long non-specific PCR products (>400 nt) were removed from the final library by double size

exclusion using 0.55 and 0.25 volumes of NucleoMag® NGS beads, respectively (Macherey-Nagel, Germany). Each pool was normalised to 10 nM and combined into a final sequencing library. Correct amplicon sizes in library pools were confirmed by Agilent D5000 ScreenTape System (Agilent Technologies, USA). The final library was sequenced on an Illumina MiSeq platform in paired-end mode using the MiSeq reagent kit v3 (600 cycles; 2 × 300 bp) with 15% spike-in of Enterobacteria phage *PhiX* control v3 (Illumina) at the Walter and Eliza Hall Institute Genomics Core (Melbourne, Australia).

Analysis of assay sensitivity: Serial dilution of samples and mock mixed infections

The dynamic range and limit of detection of AmpSeq markers were estimated by two approaches: i) serial dilution of one clinical *P. vivax* sample at 802, 80.2, 40.1, 8.02, 4.01, 0.8 parasites/µL, and ii) evaluating the detection of minority clones in mock mixed infections. We considered a marker successfully sequenced when it had >100 reads per sample. We determined the last sample dilution detected by the assay as the last sample with at least 7 of 11 successfully sequenced markers (40% of missingness). Samples with fewer than 7 of 11 successfully sequenced markers were filtered out due to low-quality amplification.

To determine the limit of detection of minority clones of the PvAmpSeq assay, we performed AmpSeq on mock mixed *P. vivax* infections. Two monoclonal *P. vivax* DNA isolates confirmed by whole genome sequencing from the Solomon Islands trial were mixed in different proportions: 1:1, 1:10, 1:50, 1:100, 1:500, 1:1000, 10:1, 50:1, 100:1, 500:1, 1000:1 and sequenced at all 11 markers. We also evaluated the detection of minority clones by artificially creating mixed infections from sequencing data using the Biostrings[40] and ShortRead[41] R packages. We generated artificial amplicon datasets from a sub-selection of the raw FASTQ sequences generated in this study from 14 monoclonal *P. vivax* samples sequenced at the 11 markers. We created artificial sequence data for each marker with known MOI and clone mixture proportions by randomly selecting the sequence from two samples with 10 different mixture proportions (0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 10%, 20%, and 50%) and 7 different read counts (100, 200, 500, 1000, 2000, 5000, and 10000). For read counts of 100, 200, and 500, the mixture proportion started at 1% because we could only extract the integer sequence from two samples. For each marker, we only artificially mixed two samples of distinct microhaplotypes and non-missing reads. Additionally, to generate more realistic amplicon datasets, we also created artificial sequence data that reflected random sampling error by utilising binomial sampling with 10 different probabilities (mixture proportions, as above) and 7 different read counts (as above) and repeated this 50 times. We then randomly selected 20 combinations of two distinct samples from each marker.

Sequencing read analysis and haplotype calling

Sequencing reads were analysed using the R package AmpSeqR version 0.1 (https://github.com/bahlolab/AmpSeqR)[32]. Sequencing reads were demultiplexed by sample and by amplicon. The overlapping sequences of paired reads were merged, and samples with a read coverage of < 1,000 reads per sample were excluded from the analysis. Index, overhang sequences, and primer sequences were removed by trimming from the forward and reverse reads. In this study, a microhaplotype was defined as an amplicon sequence variant at a given locus. The minor allele frequency (MAF) was calculated for each single nucleotide in the given dataset, and then SNPs below 0.1% in frequency were removed to exclude PCR and sequencing errors. In microhaplotypes resulting from insertion and deletion (indels), such as homopolymer regions (>3 repeated bases based on the reference sequence, e.g., "AAAA"), the sequence was adjusted to have the reference number of repeats, as the high rate of indels and errors in homopolymer regions. Microhaplotypes with low sequence identity to the *P. vivax* P01 reference genome (<75%) or with a frequency < 1%, and chimeric and singleton reads were excluded from further analysis. Microhaplotype calling required a minimum of 5 reads coverage per locus, a within-host haplotype frequency of ≥ 1%, and an occurrence of this microhaplotype in ≥ 2 samples over the entire dataset. All amplicon sequencing data are available under accession no. SAMN43387238 to SAMN43387533 at the Sequence Read Archive (SRA), and the associated BioProject ID is PRJNA1153071. Microhaplotype sequences were also deposited in the open-access Github repository: https://github.com/jrosados/PvAmpSeq.

Population-level and within-host diversity estimates

The expected heterozygosity (H_e) was calculated for each locus in the given dataset as described elsewhere.[42] The within-host microhaplotype frequency was calculated as the number of reads per microhaplotype per locus over the sum of all reads per locus in a sample. Multiplicity of infection (MOI) was calculated as the highest number of distinct microhaplotypes per locus across all loci in a given sample.

Reproducibility of AmpSeq data and comparison with available microsatellite data

As part of the optimisation of PvAmpSeq, we leveraged the availability of two sample types from the Solomon Islands ACT-RAD clinical trial: red blood cell pellets and dried blood spots from the same individual collected at the baseline visit. We compared the microhaplotype marker coverage and MOI estimates for 17 available paired sample types from the baseline.

In addition, five DNA samples from the Peruvian cohort had matched microsatellite genotyping data based on 16 markers published by Manrique *et al*.[43] We also compared the PvAmpSeq MOI data with MOI estimates based on microsatellite data.

Classification of samples into homologous and heterologous

We used the *dcifer* R package to calculate identity-by-descent (IBD) IBD r^ and the IBD r^total (i.e., inferred shared ancestry) between polyclonal infections using a statistical framework for inference that accounts for the complexity of infection (COI) and population-level allele frequencies (in our case, microhaplotype frequencies).[44] Briefly, we first estimated naive COI, and then population-level haplotype frequencies were adjusted for the estimated COI based on our data. *Dcifer* provides estimates of r^, the relatedness between two samples, and performs a likelihood-ratio test to test the null hypothesis that two samples are unrelated (H0: r^ or IBD = 0) at a significance level of $\alpha = 0.05$, adjusted for a one-sided test. For all pairs of related samples where we reject the null hypothesis, this provides statistical confidence that two pairs of samples are significantly related. In addition, in the case of samples with COI>1 we also estimate both *M'*, which is the number of related *clone* pairs between both samples, and $r^{\wedge}_{total}$, which represents the 'total' or overall relatedness between all clones. However, the method assumes there is no within-host relatedness between clones in an infection, which is violated in our case due to the possibility of genetically related relapses being present in our samples. Based on IBD estimates, infections were classified as heterologous (IBD: 0-0.25), difficult to define (IBD: 0.25-0.5), and homologous (IBD: 0.5-1).

Classification of samples into relapse, recrudescence, and reinfections

We used the recently developed *Pv3Rs* R package (https://github.com/aimeertaylor/Pv3Rs), which uses a probabilistic model-based classification framework to classify recurrent *P. vivax* infections as recrudescence, relapse, or reinfections based on genetic data.[45] The model accounts for IBD of parasite clones between recurrent episodes, COI, and population-level allele frequencies (in our case, microhaplotype frequencies) within a Bayesian framework based on informative prior probabilities for each recurrence classification state. In addition, the model estimates both marginal and joint probabilities of the different recurrence states (recrudescence, relapse, and reinfection). Marginal

probabilities provide the likelihood of each recurrence state occurring independently, while joint probabilities consider the combined likelihood across multiple episodes within an individual. Marginal probability outputs from the model may be more reliable, particularly in scenarios where model assumptions may be violated.[45]

For the Solomon Islands clinical trial samples, where participants received treatment at baseline, we applied the default prior probabilities of 0.33 for each recurrence classification state. This approach assumes that recrudescence is a possible outcome given the treatment administered at the start of the study. In contrast, for the Peru cohort, which was an observational community study without baseline treatment, we assigned prior probabilities of 0.10 for recrudescence and 0.45 each for relapse and reinfection states. These values reflect the assumption that most recurrences in this setting are likely due to relapse or reinfection rather than recrudescence. To estimate population-level microhaplotype frequencies, we used only the subset of baseline samples for the Solomon Islands clinical trial samples but all samples for the Peru cohort. In the case of the Solomon Islands, we opted for this approach to minimise the potential bias from within-host selection of recrudescent parasites, which might distort microhaplotype frequency estimates if post-treatment samples were included. In the Peru cohort, we used all samples to provide a broader representation of the circulating parasite population because most recurrences are presumed to be reinfections or relapses (both drawn from the broader mosquito population).

We performed a series of sensitivity analyses to better understand the limitations of Pv3Rs and the potential impact on our results. First, we explored the impact of inclusion/exclusion of samples for estimation of population-level microhaplotype frequencies in the case of the Solomon Islands cohort and used all samples (baseline and follow-up) to estimate population-level microhaplotype frequencies, which did not impact the classification of samples. We also ran a sensitivity analysis to estimate the false discovery rate (FDR) of relapses and recrudescences by performing a simple simulation of infection pairs using our sequence data. We randomised infection pairs (i.e., selecting entire 'infection sets' of microhaplotypes and randomly pairing them to another infection set), ensuring that pairs were not derived from the same participant and imposing a time order such that the recurrent infection always occurred after the first infection. This generated a 'null' distribution under the simple assumption that such randomly paired infections should not represent relapses or recrudescences. In the Solomon Islands dataset, we randomised baseline samples with follow-up samples, and in the Peru dataset, we randomised all samples. FDR was then calculated as the proportion of infections 'misclassified' as relapses or recrudescences out of all infection pairs under this null model. We ran 100 replicates for each cohort and calculated the mean FDR and 95% confidence interval. Finally, to assess

how particular microhaplotypes might influence the probabilistic classification, e.g., due to heterozygosity or number of alleles in the population, we iteratively re-ran the analysis, omitting one marker each time.

Statistical analysis and data accessibility

Categorical variables were compared using the two-sided Fisher's exact test or $\chi 2$ test when required. Continuous covariates were compared using two-sided Mann-Whitney, Kruskal-Wallis, or T-test when required. Relationships between parasite density and read counts were tested using the Pearson correlation coefficient, and *p-values* were adjusted by the Benjamini-Hochberg method. All statistical analyses were performed using R 4.1.0 (https://www.r-project.org/). Original data, R scripts, and algorithms developed in this study are accessible in the repository https://github.com/jrosados/PvAmpSeq. The sample and patient IDs of the original databases were not known to anyone outside the research group. De-identified datasets were generated during the current study and used to make all figures available as supplementary files or tables.

Results

PvAmpSeq microhaplotype marker selection

Of the 354 samples processed from PVGV, 200 high-quality WGS sequences from 6 countries (Cambodia, Colombia, Mexico, Papua New Guinea, Peru, and Thailand) were searched for all regions that contained at least 4 SNPs within a window of 140 bp. All successful polymorphic regions were then ranked by expected heterozygosity within each country, with the highest mean-ranked microhaplotype being selected for each of the 14 chromosomes (Figure S1). Although we excluded the majority of the hypervariable sub-telomeric regions of PvP01, microhaplotypes from Chr06 and Chr12 were selected but then excluded during the PCR optimisation as they were located in the proximity of hypervariable sub-telomeric *pir* genes. Chr04 was excluded due to the low amplification success in Solomon Islands samples. Our panel included 11 microhaplotype markers or loci across 11 chromosomes.

The final panel of PvAmpSeq microhaplotype markers comprised three loci encoding highly polymorphic surface antigens such as Merozoite Surface Protein 1 (MSP1, Chr07, PVP01_0728900), Merozoite surface protein 3 (MSP3.3, Chr10, PVP01_1031500) and Apical Membrane antigen 1 (AMA1, Chr09, PVP01_0934200); proteins involved in reticulocyte invasion, Reticulocyte binding protein 2a (RBP2a, Chr14, PVP01_1402400); enzymes such as Protein Tyrosine phosphatase putative (PTP2, Chr01, PVP01_0113700), Glyoxalase I-like protein (GILP) putative (Chr11, PVP01_1144200); pseudogenes like Lysophospholipase putative (Chr02, PVP01_0201300); and putative proteins of unknown function such as STP1 protein (Chr05, PVP01_0533300), conserved *Plasmodium* protein (Chr03, PVP01_0302600 and Chr13, PVP01_1346800) and *Plasmodium* exported protein (Chr08, PVP01_0838000). These genes contain both SNPs and indels (Text S2).

Validation on samples from two cohort studies

A total of 492 (Solomon Islands, n = 218; Peru, n = 274) samples were sequenced for the validation of the PvAmpSeq assay. 196 samples were filtered out due to a low number of reads (<1,000 reads per sample) or low identity sequences (Solomon Islands, n = 71; Peru, n = 125). Discarded samples had a median parasite density of 14.4 parasites/uL of DNA [IQR: 5.63- 39.5]. Additionally, samples with more than 40% (5 out of 11 loci) of missingness were discarded (Peru, n = 9), as well as ten samples from the Solomon Islands trial, due to sequencing failure of their paired baseline infection. Negative DNA samples and negative template controls included in sequencing runs yielded <100 reads and were filtered out.

The remaining 275 samples selected for downstream analysis corresponded to 77 participants from the Solomon Islands and 93 participants from Peru. Of the Solomon Islands participants, 41 had available samples from 1 up to 6

recurrent infections (total 58 samples), whereas only 40 Peruvian participants had follow-up samples, corresponding to 1 or 3 recurrent infections (total 47 samples). The median age of the participants was 10.1 years [IQR: 7.42-14.4] and 36.7 [18.0-51.5] for the Solomon Islands and Peru, respectively (Table 1). There was no significant difference between the parasite density of baseline and follow-up infections for both cohorts ($p > 0.05$). As expected for the clinical trial, febrile infections were more frequent at the baseline of the Solomon Islands cohort ($p < 0.05$).

Table 1. Epidemiological characteristics of participants and description of their infections

Participant characteristics	**Solomon Islands n = 77**	**Peru n = 93**
Age, median [IQR]	10.1 [7.42-14.4]	36.7 [18.0-51.5]
Sex, number female (%)	32 (41.6%)	51 (54.8%)
Treatment administered	Artemether-lumefantrine (n = 21), artemether-lumefantrine + primaquine (n = 25), dihydroartemisinin-piperaquine + primaquine (n = 31)	Chloroquine-primaquine (n=1)
Infection characteristics	**n = 135**	**n = 140**
Parasite density, median [IQR]		
Baseline	416 parasites/µL [81.7-793][a]	341 parasites/µL [40.6-8560][b]
Follow-up	151 parasites/µL [18.1-978]	337 parasites/µL [77.6-1470]
Febrile, number (%)		
Baseline	38/77 (49.4%)[c]	7/93 (7.5%)[d]
Follow-up	11/58 (19.0%)	2/47 (4.3%)

a Statistical difference between baseline and follow-up assessed by a two-sided Mann-Whitney test, $p = 0.2515$.
b Statistical difference between baseline and follow-up assessed by a two-sided Mann-Whitney test, $p = 0.7541$.
c Statistical difference between baseline and follow-up assessed by a two-sided Fisher's exact test, $p = 0.0003$.
d Statistical difference between baseline and follow-up assessed by a two-sided Fisher's exact test, $p = 0.7178$.

Data processing and read coverage

Reads were demultiplexed and filtered using AmpSeqR[16]. After discarding PCR artefacts, the rate of success per loci went from 0.70 to 1 (Table S1). The median success rate in follow-up samples was slightly lower than baseline samples but still around 0.90 for both cohorts (Solomon Islands; baseline median: 0.99[IQR: 0.97-1.00], vs follow-up median: 0.91[0.91-0.95], $p < 0.05$; Peru; baseline median: 0.98 [0.80-0.98], vs follow-up median: 0.89 [0.82-0.93], $p > 0.05$).

The median read coverage per sample was 4309 [2051-8129] and 4752 [2535.5-7524] for the Solomon Islands and the Peruvian cohort, respectively. Chr07 had the highest median read coverage per sample in both cohorts (9774, [4672.2-9964] for the Solomon Islands and 8734 [5560.5-9857.2] for Peru), whereas Chr01 had the lowest median coverage read (1763 [960-2326] in the Solomon Islands; and 2376 [1606-4180] in Peru) (Figure 1A).

No significant correlation was found between the sample parasite density (parasite/µL) and the number of reads per marker (Pearson's r range = -0.23-0.16, *adjusted p* range = 0.17-0.94) (Figure 1B). Samples with <10 parasites/µL were amplified with variable read coverage per marker (median: 4590 [2574-8499]). The high variance seen in each parasite density group suggested that the DNA quality affected the amplification performance. On the other hand, detection of a control sample (AR-246) was possible only when parasite density was >40 parasites/µL. Microhaplotypes were detected at dilutions of 40.1, 80.2, and 802 parasites/µL for the control sample (Figure 1C). Control sample dilutions with ≤ 8.02 parasites/µl were poorly sequenced for most of the loci (< 60 reads). Chr13 did not successfully amplify in the control sample and was precluded from this analysis. We detected one microhaplotype per locus in the AR-246 control sample for all dilutions, except at the highest dilution of 802 parasites/µL, where two microhaplotypes were detected in marker Chr09, one of which had a very low microhaplotype frequency (0.018) (Figure S2). This microhaplotype could be a true microhaplotype since it was also detected in other samples in this dataset and with high frequency. This indicates that, as expected, higher parasite densities can more accurately determine the complexity and diversity of sample infections.

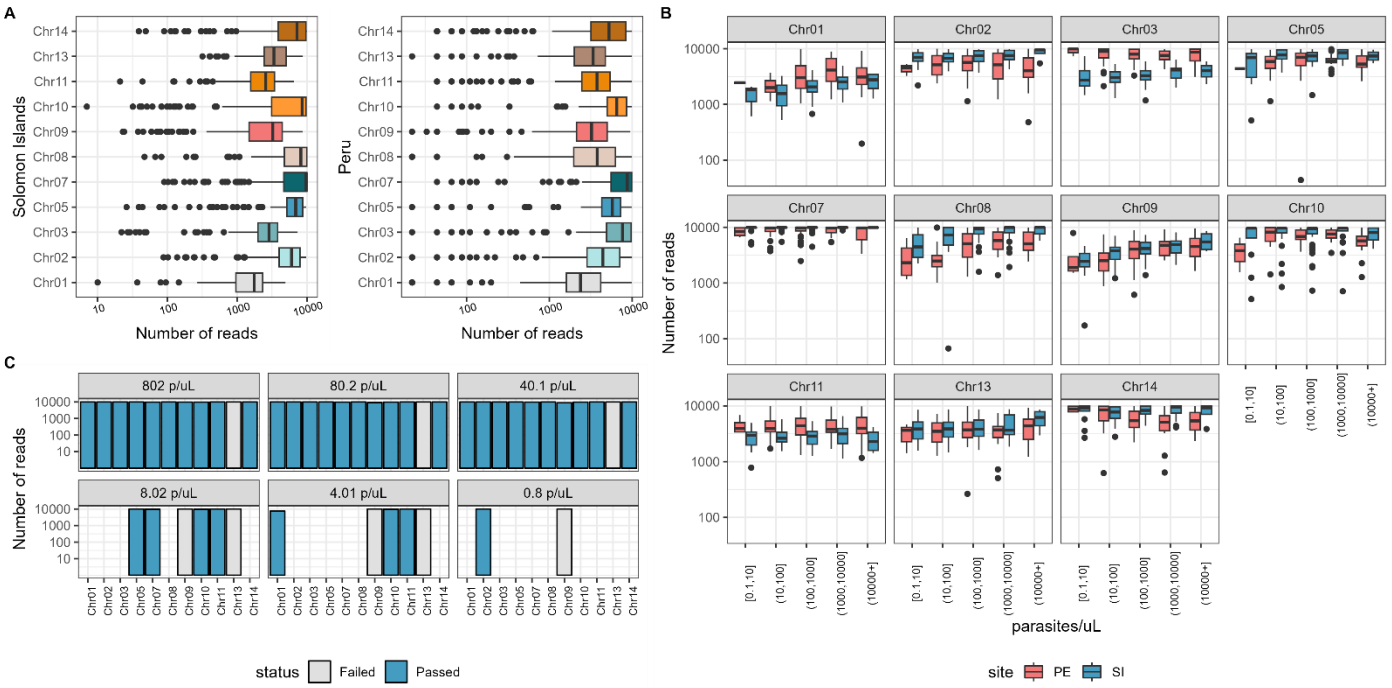


Figure 1. Marker coverage and sample parasite density. (A) Number of reads per marker on field samples. (B) Number of reads and parasite density of field samples. (C) Limit of detection for parasite density; in blue: markers that passed the quality control step; in grey: markers that failed the quality control step; empty bars: markers with no successful PCR amplification.

Limit of detection for minority clones

Two monoclonal samples from the Solomon Islands trial were selected to mimic a mixed infection (AR079 and AR093). DNA samples were normalised at 100 parasites/µL and combined in the following ratios: 1:1, 1:10, 1:50, 1:100, 1:500, 1:1000, 10:1, 50:1, 100:1, 500:1, and 1000:1. In Figure 2A, the results of the mock mixed infections are shown. Analyses were restricted to informative AmpSeq markers with distinct microhaplotypes between samples, i.e., Chr01, Chr02,

Chr03, Chr05, Chr09, and Chr14. Markers Chr07, Chr08, Chr10, and Chr11 were not included because both samples had the same microhaplotype in these loci. The correct minor microhaplotype was detected in most mixtures only at 1:1 (50%), 1:10 (10%), and 10:1 (10%) mixture proportions for Chr01, Chr02, Chr09, and Chr14, and the mixture proportion was not the same as the defined mixtures. Only microhaplotypes from the major clone were detected in most samples. We examined the minor clone sequences for each sample in the raw sequence data and found that the minor clone sequence coverage was extremely low (less than 5 reads compared to over 20,000 reads in total), which resulted in the inability to detect the minor clones. A possible reason is that the mixture experiment is based on field samples. Minority clones at 2% (1:50) were detected for Chr02. Minority clones were not detected for Chr05 and Chr03. In some samples, we detected singleton microhaplotypes with very low frequencies that were not AR-079 or AR-093 microhaplotypes, which would represent false positives (Figure S3). We compared the microhaplotype sequences of major and minor microhaplotypes in these samples and found that these minor and major microhaplotypes differed in only one position. These microhaplotypes are likely to be systematic sequencing errors, with the base-call errors generally occurring at the same genomic position from different sequence reads.[46] Systematic errors are often mistaken for heterozygous sites in individuals or SNPs in population analysis.

We also examined the limit of detection of artificially mixed infections by determining the minor clone detectability success *in silico*. At the highest read counts (5,000 and 10,000), the minor clone was robustly detected at a clone relative frequency as low as 0.5% up to 50% for all amplicon markers; however, when read counts were <1,000, the minor clone was only detected accurately when the relative frequency was >10% (Figure 2B). Regardless of the data being generated with the same mixture ratios and read counts, there were substantial differences in the detectability of the minor clone based on the microhaplotype marker. For example, at a read count of 10,000 and minor clone relative frequency of 0.1%, three markers had 100% success in detecting the minor clone (Chr03, Chr07, and Chr10), but there were very low success rates for markers Chr01, Chr08, Chr11, Chr13, and Chr14 ranging between 0% to 29%. We also created the artificially mixed infections following a binomial distribution to reflect a more realistic random sampling error and found similar results as above (Figure S4). We thus set 2% as the lower limit of detection for minority clones of PvAmpSeq and recommend at least 10,000 reads per microhaplotype amplicon for robust minor clone detection (assuming frequency may be as low as 0.1%). If the minor clone relative frequency is expected to be around 10%, we recommend at least 1,000 reads.

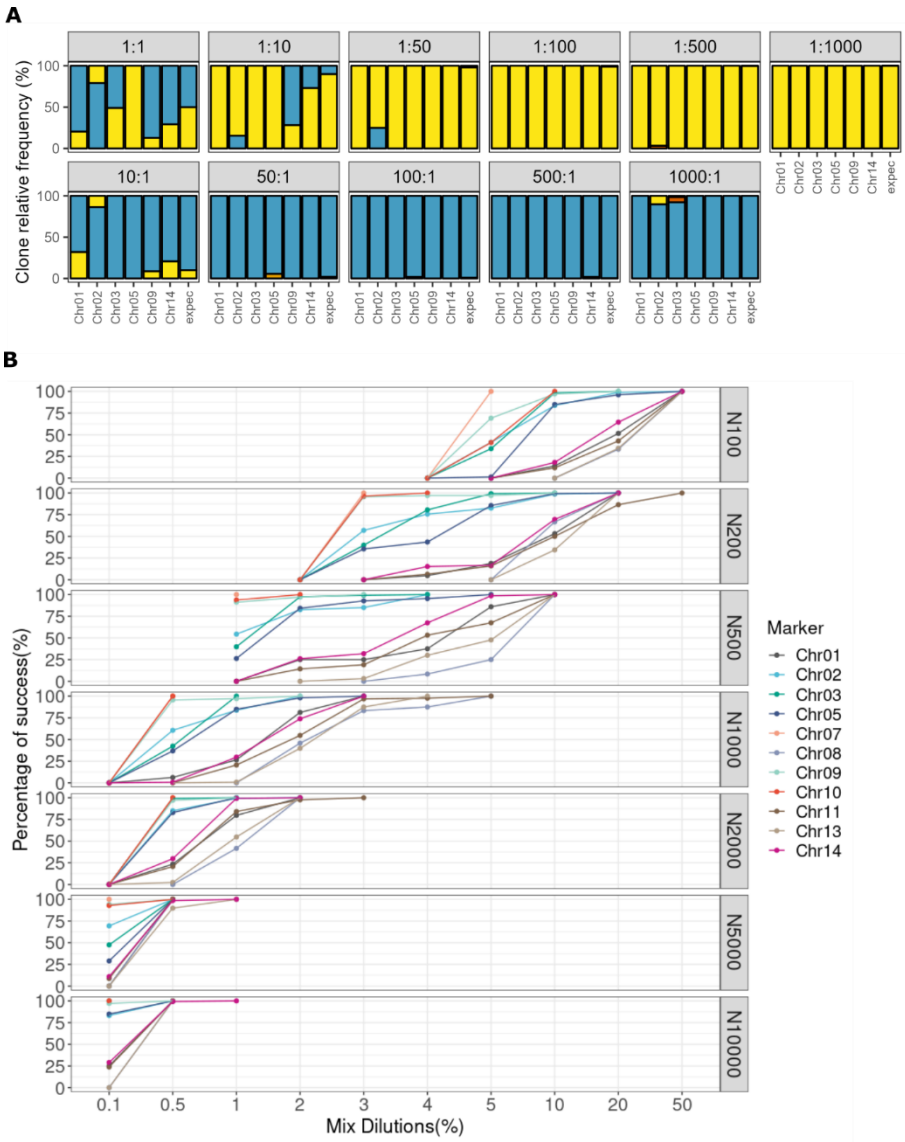


Figure 2. Detection of minority clones in mock and artificial mixed infections. (A) The Y-axis shows the sample ratios detected in mock mixed infections. The X-axis shows the markers grouped by expected sample ratios in mock infections (1:1, 1:10, 1:50, 1:100, 1:500, 1:1000, 10:1, 50:1, 100:1, 500:1 and 1000:1). In blue: clones from sample AR079; in yellow: clones from sample AR093; in red or orange: contaminants clones. expec: expected results from mocked mixed infections. Markers sharing the same microhaplotypes between samples are not displayed. (B) The detectability of the minor clone under different numbers of reads and artificial mixture ratios. The X-axis represents the mixture dilutions (%), and the Y-axis represents the success rate (%) of detecting the minor clone. Coloured by amplicon marker.

Exploring population and within-host genetic diversity metrics using PvAmpSeq

The microhaplotype markers had high diversity in both sample sets, with the mean expected heterozygosity (H_e) of the 11 markers 0.70 for the Solomon Islands samples and 0.65 for the Peruvian samples. We found approximately half of the markers had high $H_e \geq 0.70$ (7/11 markers in the Solomon Islands; 5/11 markers in Peru), but the markers with the highest heterozygosity differed in each cohort except for Chr05 and Chr07 that had high H_e in both (Figure 3A) but different microhaplotype frequencies (Figure 3B). The remaining markers had low- to moderate- heterozygosity, ranging from 0.29 to 0.69 (Figure 3A). High diversity markers ($H_e \geq 0.70$) had a median of 7 alleles (range: 4-43) and 13 microhaplotypes (range: 5-13) in the Solomon Islands and a median of 6 alleles (range: 3-34) and 7 microhaplotypes (range:5-7) in Peru (Table S2).

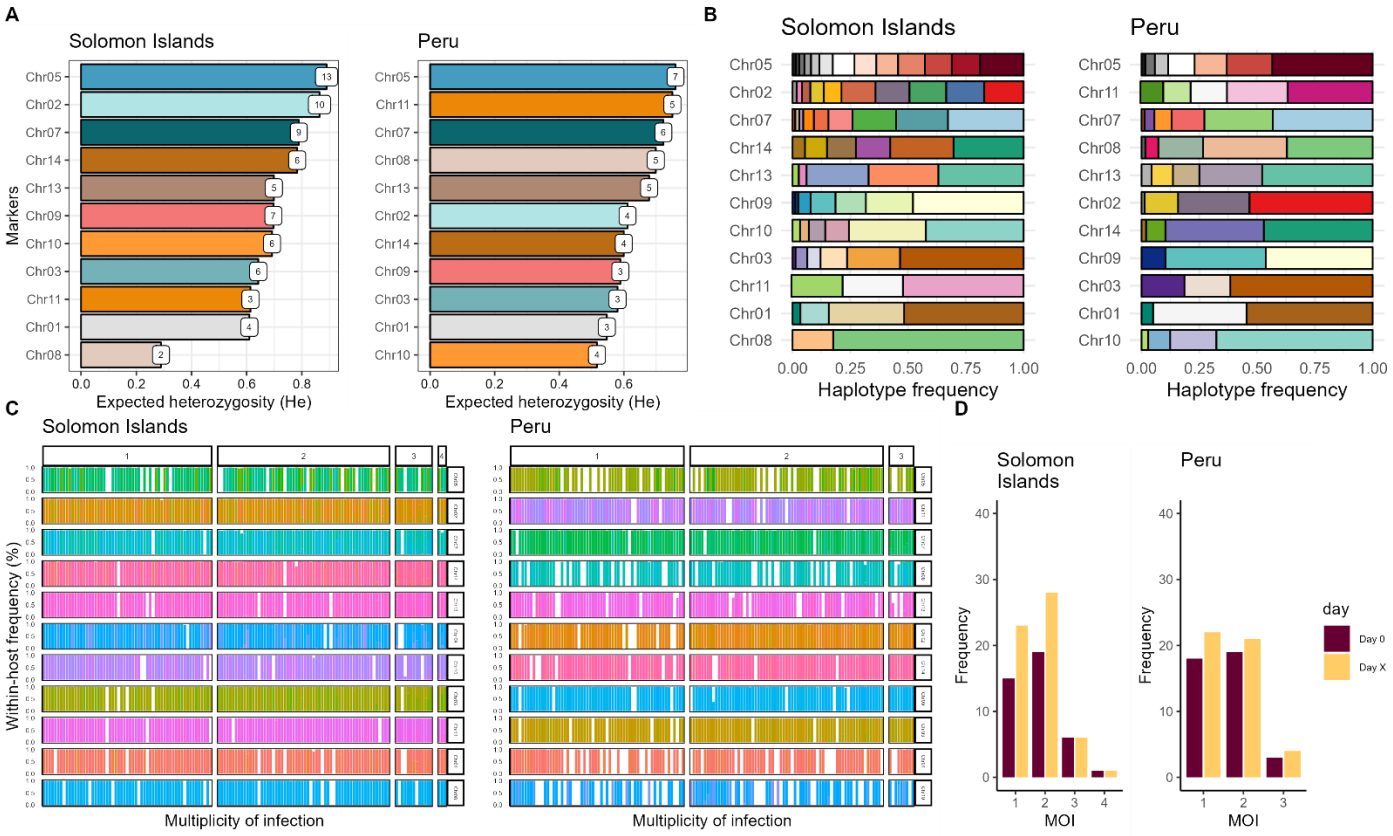


Figure 3. Genetic diversity of markers and multiplicity of infection of validation samples. (A) Markers were sorted by descending H_e. He: Expected heterozygosity calculated as the number of microhaplotypes per marker divided by microhaplotype frequency on Solomon Islands (n = 135) and Peruvian samples (n = 140). The number of microhaplotypes per marker is indicated on top of each bar. (B) Microhaplotype frequency per marker estimated on the Solomon Islands and Peruvian samples. Each colour indicates a different microhaplotype. (C) Estimated multiplicity of infection (MOI) in field samples as the highest number of distinct microhaplotypes by all markers. The Y-axis shows the within-host frequency of microhaplotypes per sample, calculated as the percentage of microhaplotype reads per sample. The X-axis shows the samples sorted by estimated MOI. Every sample is represented by a vertical bar. (D) Distribution of MOI in paired samples, where Day 0 represents the baseline infection (Solomon Islands, n = 41; Peru, n = 40) and Day X the day of recurrent infection (Solomon Islands, n = 58; Peru, n = 47).

Using the highest number of distinct microhaplotypes detected by all markers, we estimated the multiplicity of infection (MOI) in both cohorts. Out of 135 samples in the Solomon Islands cohort, 59 had a maximum MOI = 1, 60 samples had a maximum MOI = 2, 13 samples had MOI = 3, and 3 samples had MOI = 4 (Figure 3C). In the Peruvian cohort, out of 140 samples, 62 had a maximum MOI = 1, 69 samples had a maximum MOI = 2, and 9 samples had MOI = 3 (Figure 3C). No association was found between the maximum MOI obtained in the baseline samples and participant's age (Kruskal-Wallis test, Solomon Islands p = 0.986, Peru p = 0.672; Figure S5), nor with the presence of fever ($\chi 2$ test, Solomon Islands p = 0.744, Peru p = 0.823; Figure S6). There was no significant difference in MOI at baseline compared to follow-up in either cohort (Solomon Islands, baseline mean MOI = 1.83 vs follow-up mean MOI = 1.74, two-sided Mann-Whitney test, p = 0.5997; Peru, baseline mean MOI = 1.62 vs follow-up mean MOI = 1.62, two-sided Mann-Whitney test, p = 0.9244) (Figure 3D, Table S3). To test whether participants with polyclonal infections at baseline had faster times to recurrence than those with monoclonal infections at baseline, we compared the time to first recurrent infection of paired samples stratified by MOI (MOI = 1 vs MOI ≥ 2). In both cohorts, polyclonal baseline infections (MOI ≥ 2) had comparable times to first recurrent infection to monoclonal baseline infections (MOI = 1)

(two-sided Mann-Whitney test, Solomon Islands $p = 0.7555$, Peru $p = 0.0635$). Likewise, the time to the next recurrent infection was not affected by the change of MOI between recurrent infections (Kruskal-Wallis test, Solomon Islands $p = 0.3979$, Peru $p = 0.4724$; Figure S7).

As part of the assay validation, we used available microsatellite data for 5 Peruvian samples to compare MOI estimates using microsatellite genotyping and PvAmpSeq. Three out of the five samples appeared to be polyclonal by PvAmpSeq, whereas only one sample was previously reported polyclonal by microsatellites (Table S4).

Classifying recurrent infections using identity-by-descent and probabilistic estimates

Paired samples were further analysed to identify whether recurrent infections were genetically homologous or heterologous to the participants' baseline infection in the Solomon Islands trial or the first infection during the Peruvian community cohort study period. Based on IBD estimates, in the Solomon Islands ACT-Radical clinical trial, 55.2% (32/58) of recurrent infections could be classified as homologous, while 41.4% (24/58) could be classified as heterologous and 3.4% (2/58) were not easily classified as IBD was between 0.25-0.5 (Figure 5A). In participants who received primaquine (anti-hypnozoite treatment) at baseline, 47.2% (17/36) of recurrent infections were classified as homologous, while 68.2% of recurrent infections (15/22) from individuals treated without primaquine were homologous (two-sided Fisher's exact test, $p = 0.3202$, Table S5).

In the Peruvian community cohort, 23.4% (11/47) of recurrent infections could be classified as homologous, while 68.1% (32/47) could be classified as heterologous, and 8.5% (4/47) were not easily classified as IBD was between 0.25-0.5 (Figure 5B). To evaluate whether recurrent polyclonal infections contribute to the increase of IBD, we compared the frequency of heterologous infections in both cohorts per classification group. We did not find an association between MOI and the classification of recurrent infections in both cohorts (two-sided Fisher's exact test, Solomon Islands $p = 0.1279$, Peru $p = 0.3754$; Table S5). Detailed microhaplotype data and classification of recurrent samples can be seen in Text S3 and Text S4.

We found that IBD in recurrent infections changed with time for most of the recurrent episodes in both cohorts, with a maximum number of five and three recurrences experienced by the same participant in the Solomon Islands and Peru, respectively (Figure 4B). There was no clear pattern with time to recurrence; for example, we found evidence that the participants in the Solomon Islands clinical trial that had four and five recurrent infections had heterologous and homologous infections compared to the baseline infection, respectively (Figure 4C) and in Peru, most participants experienced heterologous recurrences (Figure 4D).

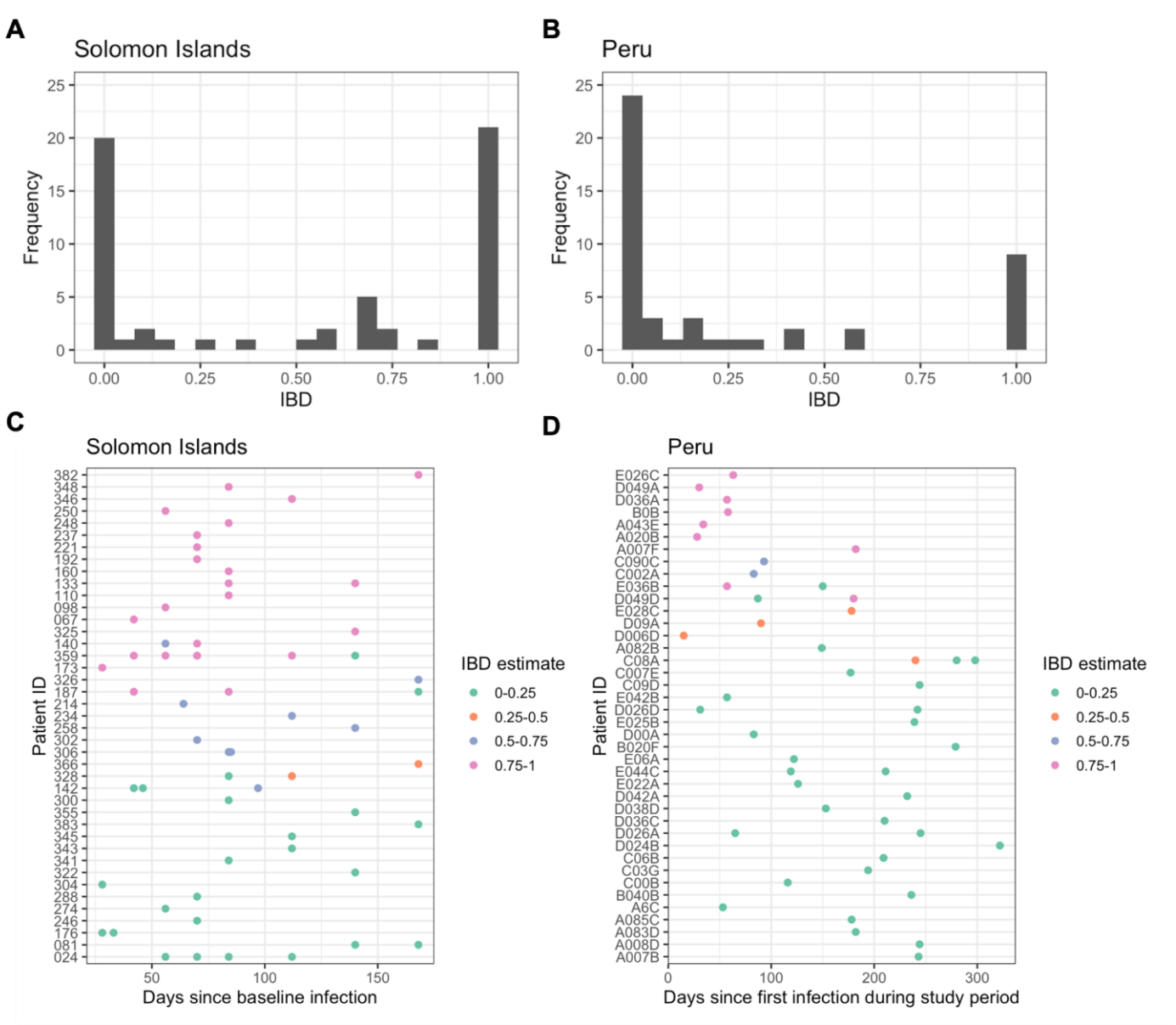


Figure 4. Classification of *Plasmodium vivax* recurrent infections with PvAmpSeq using IBD. (A-B). Distribution of IBD estimates in paired infections from the Solomon Islands and Peru, comparing baseline and recurrent infection pairs. (C-D). IBD between infection pairs in the Solomon Islands and Peru. The X-axis shows the days since baseline or first infection, the Y-axis depicts each patient's infection over time, and the colour represents the IBD estimate range for the infection pairs. Baseline and first infections (Day 0) are not displayed for easier interpretation of plots.

We also performed probabilistic classification of recurrences using the *Pv3Rs* R package and estimated both marginal and joint probability estimates of a recurrent infection being either a relapse, recrudescence, or reinfection to provide further granularity on probable recurrence classification. In Solomon Islands, we explored the most probable classification of the first recurrent infection after treatment and found that 50% of participants experienced a reinfection (mean posterior probability of estimates 0.82, range: 0.51-0.82), 25% experienced a recrudescence (mean posterior probability of estimates 0.75, range: 0.63-0.75) and 25% experienced a potential relapse (mean posterior probability of estimates 0.90, range: 0.62-0.90) (Figure 5A). In the case of participants who experienced more than one recurrence, the trends were less clear (Figure 5A), with some experiencing both reinfection and relapses (e.g., AR-024). Interestingly, for this participant, the IBD estimates for recurrences were all <0.25 compared to baseline, but the

probabilistic estimates pointed to the possibility of both reinfection and relapse over time. A sensitivity analysis using simulated reinfection Solomon Islands data showed a mean FDR (proportion of simulated reinfections being misclassified as relapses or recrudescences) of 4.1% (95% CI: 3.4-4.9%) for the Pv3Rs model.

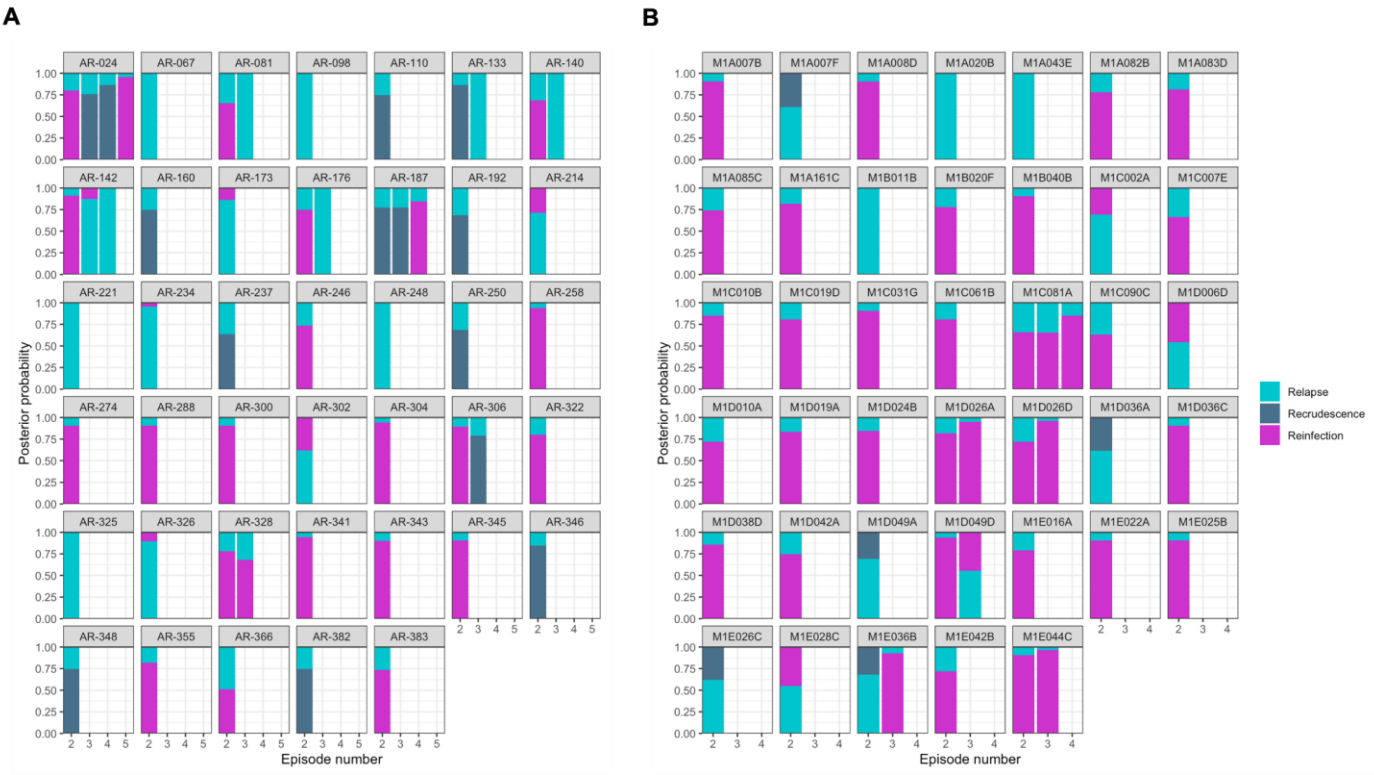


Figure 5. Model-based classification of recurrent *P. vivax* infections with PvAmpSeq. Marginal posterior probability estimates of classification outcome (relapse, recrudescence, or reinfection) of recurrent infections compared to the baseline infection faceted by each participant in (A) the Solomon Islands and (B) Peru. The X-axis shows the episode number; for example, considering that the baseline infection is the first episode, the first recurrent infection would be episode 2.

In Peru, we found that 72.5% of participants likely experienced a reinfection (mean posterior probability of estimates 0.81, range: 0.63-0.81), 27.5% experienced a potential relapse (mean posterior probability of estimates 0.73, range: 0.54-0.73), and none experienced recrudescence (Figure 5B). Most participants experiencing more than one recurrence most likely experienced reinfections (mean posterior probability estimate 0.88, range: 0.65-0.88). For the Peru dataset, our sensitivity analysis showed a mean FDR of 12.4% (95% CI: 11.3-13.4%) for the Pv3Rs model.

We explored whether the exclusion of specific markers may lead to discrepant classification and found little evidence and no clear trend for either the Solomon Islands or Peru datasets, suggesting that the marker panel is robust in different settings (Figure S8). Overall, we found moderate concordance between IBD estimate ranges and probabilistic classification results (Figure S9), with the classification of homologous infections using IBD ranges also found to be either recrudescence or relapses and similarly, heterologous infections were associated with reinfections (Figure S9A-B). This was also consistent with the microhaplotype-specific data for recurrent infections (Figure S10).

Discussion

We developed PvAmpSeq, a targeted amplicon sequencing assay that enables the characterisation of *Plasmodium vivax* infections at the clonal level by targeting 11 highly polymorphic microhaplotype markers. The PvAmpSeq assay amplifies field samples with >10 parasites/µL, yielding high-depth coverage (5510 [IQR: 3366-8789] reads) but with less coverage for samples with 1-10 parasites/µL. Both archived dried blood samples and red-blood-cell pellet samples were successfully amplified. The high coverage per marker of PvAmpSeq allows the detection of minority clones at a frequency of >10%. The PvAmpSeq microhaplotype markers comprise loci encoding surface antigens, proteins involved in reticulocyte invasion, enzymes, and proteins of unknown function. As a result of immune selection acting on these loci, their high diversity made them suitable for measuring MOI and for tracking individual clones over the course of natural infections within an individual.[47] Our results show that the PvAmpSeq markers provide high resolution to differentiate between homologous and heterologous recurrent infections. The less polymorphic PvAmpSeq markers of the panel could be suitable for studying diversity and structure between global regions; however, further exploration is needed to validate the assay with samples from various geographical origins and in comparison to other genotyping panels (e.g.,[6,26,27,29,30]).

We implemented identity-by-descent (IBD) to classify recurrent infections into homologous and heterologous infections compared to their baseline or first infection pair and modelled the probability that recurrent infections were either relapses, recrudescence, or reinfections using recently developed methods.[44,45] At the end of follow-up, homologous infections represented more than half of recurrent infections in the Solomon Islands clinical trial based on IBD estimates, which was in line with our model estimates, further classifying these recurrences into 16% probable relapses and 26% recrudescences, although the uncertainty around these classification results was wide and recovery of appropriate recrudescence and relapse classification may not be sufficiently robust in the current implementation of Pv3Rs (https://github.com/aimeertaylor/Pv3Rs/). Approximately 23% of recurrent infections in the Peruvian cohort were homologous based on IBD, which was also in line with our model estimates of 26% of recurrences being probable relapses (although, again, with relatively wide uncertainty around these estimates). Our results are consistent with low levels of malaria transmission before the ACT-Radical clinical trial was conducted in the Solomon Islands (James, Obadia *et al*. *under review*), with a corresponding less diverse parasite population. In addition, we expected more moderate transmission levels with possible parasite clonal expansion in both sampled communities, Lupuna and Cahuide, in the Peruvian Amazon, based on previous work using microsatellite genotyping.[31,33,43] Secondly, relapses occurring in patients with low hypnozoite load in their liver are usually clonal, and the relapse parasites are genetically

578 homologous to the parasite from the primary infection.[48,49] Adults in endemic areas are more likely to experience
579 relapse with heterologous parasites due to latent hypnozoites from previous inoculations.[48] In contrast, heterologous
580 recurrences were more frequent in the Peruvian cohort, suggesting they could encompass both relapses arising from
581 heterologous hypnozoites from previous inoculations and new infections.

582 IBD estimates (r^) correlated well with model-based classification for most recurrent infections in Peru and the
583 Solomon Islands, showing the potential application of the PvAmpSeq panel. In the case of discordant results, we were
584 unable to definitively ascertain whether this was due to model misspecifications/limitations or limitations of the
585 current AmpSeq panel. It is likely a combination of both since the PvAmpSeq panel includes 11 microhaplotype markers
586 and thus may not always allow appropriate classification of recurrent *P. vivax* malaria infections. In some cases, more
587 than 11 microhaplotype markers may be required, depending on the circulating parasite population's genetic diversity.
588 In addition, the *dcifer*-based IBD estimates assume no within-host relatedness between clones, which would be
589 violated in the case of *P. vivax* relapses and may impact our estimates; however, simulations by Gerlovina *et al*. showed
590 a relatively robust estimation of IBD even when this assumption is violated.[44] Our simple simulation showed robust
591 classification of Pv3Rs and resulted in a low false discovery rate of relapses and recrudescences for the Solomon Islands
592 clinical trial but above 10% for misclassifications of relapses in Peru. This is in line with the different study designs,
593 where there was a clear 'baseline' infection and rigorous follow-up period in the Solomon Islands trial. However, there
594 may be potential limitations for population-based epidemiological studies where there is no clear 'baseline' infection
595 given ongoing transmission and low treatment rates. A limitation of the current implementation of the Pv3Rs model
596 is that all recurrences are compared to the baseline infection, which is appropriate for clinical trials but less so for
597 observational studies like the Peru dataset. Another challenge for IBD-based and Pv3Rs tools is distinguishing between
598 recrudescence and intermittent patency in peripheral blood in chronic spleen-positive infections, i.e., a person might
599 be continuously infected in the spleen,[50] but parasites (all or particular clones) may only be intermittently detectable
600 in the bloodstream; in particular, in observational studies. Future work could explore refining the model-based
601 estimates by using time-to-event models to generate informative prior estimates, as was done with microsatellite
602 data,[51] and leveraging longitudinal studies with repeated sampling. Development of new models or model extensions
603 may also be needed to robustly classify recurrent infections to account for sequencing error, given the availability of
604 many *P. vivax* amplicon sequencing approaches, but this was beyond the scope of the current study. In the context of
605 newly developed marker panels (i.e., amplicon sequencing, molecular inversion probe, and microhaplotype

606 genotyping panels)[6,26,27,29,30], future work could focus on rigorous benchmarking of these analysis tools to provide
607 guidance for the community (e.g., PGEforge https://mrc-ide.github.io/PGEforge/).

608 Detecting microhaplotypes derived from persistent gametocytes or residual DNA in recurrent infections after
609 artemisinin-combination therapy (ACT) could overestimate treatment failure rates.[52,53] Unlike *P. falciparum*, *P. vivax*
610 gametocytes are commonly detected at densities of <10% of asexual parasites, and in the absence of treatment, the
611 duration of gametocytaemia is 3 days.[54] Our approach aimed for coverage of 10,000 reads per marker per sample,
612 ensuring high sensitivity for detecting minority clones at frequencies of 0.5% within-host, i.e., 50 reads. Our sensitivity
613 analysis showed we can confidently detect microhaplotypes with ≥2% within-host frequency; thus, this approach will
614 likely detect gametocyte microhaplotypes. Interestingly, in the Solomon Islands drug-efficacy clinical trial, no recurrent
615 infections were detected earlier than 28 days post-treatment, and homologous recurrent infections were more
616 frequent in patients treated without primaquine, suggestive of relapses; however, low sample sizes when stratifying
617 by treatment arm limited us from drawing definitive conclusions.

618 A limitation of our study was the number of samples included in the validation of PvAmpSeq. The original ACT-Radical
619 clinical trial involved 374 participants, of whom 307 completed follow-up (James, Obadia *et al*. *under review*), but the
620 low parasite density *P. vivax* infections during follow-up limited the number of paired baseline/follow-up samples
621 available for inclusion in PvAmpSeq validation. This could also show a potential bias of PvAmpSeq to better detect new
622 infections that usually have high parasite density. After data demultiplexing and quality control, the remaining sample
623 size per treatment arm was small; thus, we could not definitely employ molecular correction to conclude the treatment
624 efficacy assessed by PvAmpSeq. Future studies should include larger sample sizes and more paired samples; for
625 example, a recent study by Kleinecke *et al*. showed potential for amplicon sequence microhaplotype-based
626 approaches for informing on recurrences.[6] The sensitivity of detection of minority clones was affected by the lack of
627 use of genetically different *P. vivax* strains for mimicking mixed infections. However, the artificial mixed infection
628 estimated that the limit of detection of minority clones was 10% when we had 10,000 reads and 1% if we had 50,000
629 reads. With these features, the potential applications of PvAmpSeq we envision are i) the study of relapse biology in
630 longitudinal studies, ii) the contribution of hidden parasite biomass in spleen and bone-marrow to relapsing *P. vivax*
631 infections iii) parasite diversity dynamics from human-host to mosquito vector, iv) parasite population diversity within
632 mosquito collected in cross-sectional studies.

Likewise, in the Peruvian cohort, some intermediate low-density/submicroscopic infections could have been missed; thus, some of the recurrent infections might not reflect the real infection dynamics in this population. A limitation of PvAmpSeq due to the intermittent release of patent parasites to the bloodstream in spleen-positive infection DNA quality from low parasite density samples represented a challenge for PvAmpSeq as observed in samples with 1-10 parasites/µL that were discarded due to low coverage (38% of the Solomon Islands samples and 49% of the Peruvian samples). Selective Whole Genome Amplification (sWGA) has been demonstrated to improve sample coverage in very low *Plasmodium* density samples for AmpSeq or WGS methods.[16–18,27,55] However, this type of treatment to low-density samples could incur amplification biases of the most dominant clone, increase the error rate and raise costs, making it less applicable in low-resources settings.[17,18] Other factors may also influence the applicability of PvAmpSeq, like daily fluctuations in clone densities within infection, which may impede robust longitudinal tracking of clones, particularly minor clones. Although we tested 6 PvAmpSeq markers in paired dried-blood spot and red-blood-cell pellet samples, we were not able to evaluate the complete panel of markers in this sample subset due to limited DNA quantities. Nevertheless, we found that PvAmpSeq performed well on both sample types, albeit with better recovery of the number of microhaplotype variants for some of the markers. PvAmpSeq detected more polyclonal infections than microsatellite genotyping in a small subset of Peruvian samples; however, future studies should include larger sample sizes to evaluate the applicability of PvAmpSeq to quantify polyclonal infections in other settings.

AmpSeq has been proposed as a new gold standard for analysing malaria drug clinical trials.[3,12] For *P. falciparum*, five AmpSeq markers have been validated to discern between recrudescences and reinfections.[5] Regardless of the level of endemicity and the cut-off of detection for minority clones (0.1 – 2%), simulation studies showed that the use of 3 to 5 polymorphic markers was sufficient to classify *P. falciparum* recurrent infections as recrudescence or reinfections.[12] In *P. vivax* infections, genotyping highly related meiotic sibling progeny present in a relapse may not always be possible when using a limited number of markers. Whole genome sequencing analysis has demonstrated that relapses are meiotic siblings resulting from the same meiotic recombination event, contrary to what microsatellite genotyping classified as identical or clonal.[56] Recent *P. vivax* wide-genome panels of Ampseq markers could potentially address this question[6,26,27,29]; however, increasing marker coverage to 10,000 reads would represent higher costs for labs in low-resource settings. Using PvAmpSeq, we classified heterologous/homologous infections using IBD and calculated posterior probability estimates of recurrences being relapse, recrudescence or reinfections by using the information from 11 microhaplotype markers. We found that this classification using both methods was largely concordant in both

samples from the Solomon Islands clinical trial and the Peruvian community cohort. Nevertheless, simulation studies and/or studies conducted in other epidemiological contexts could inform whether the number of PvAmpSeq markers included in this study would be sufficient for classifying recurrent infections in clinical trials executed in settings with higher endemicity levels of *P. vivax* than in this study.

In conclusion, here we present a new framework to classify *P. vivax* recurrent infections using PvAmpSeq microhaplotype data for molecular correction in drug clinical trials and for studying *P. vivax* relapse biology in longitudinal cohorts, and demonstrate its potential using state-of-the-art analysis methods. We anticipate that this tool can be applied in a wide range of epidemiological studies and clinical trials to fill a pressing need in *P. vivax* genomic epidemiology to better understand relapse epidemiology and support the robust evaluation of clinical efficacy trials.

References

1. World Health Organisation (2024). World malaria report 2024.

2. Robinson, L.J., Wampfler, R., Betuela, I., Karl, S., White, M.T., Li Wai Suen, C.S., Hofmann, N.E., Kinboro, B., Waltmann, A., Brewster, J., et al. (2015). Strategies for understanding and reducing the Plasmodium vivax and Plasmodium ovale hypnozoite reservoir in Papua New Guinean children: a randomised placebo-controlled trial and mathematical model. PLoS Med. *12*, e1001891.

3. World Health Organisation (2021). Informal consultation on methodology to distinguish reinfection from recrudescence in high malaria transmission areas.

4. World Health Organisation (2008). Methods and techniques for clinical trials on antimalarial drug efficacy: genotyping to identify parasite populations.

5. Gruenberg, M., Lerch, A., Beck, H.P., and Felger, I. (2019). Amplicon deep sequencing improves Plasmodium falciparum genotyping in clinical trials of antimalarial drugs. Sci. Rep. *9*, 17790.

6. Kleinecke, M., Rumaseb, A., Sutanto, E., Trimarsanto, H., Hoon, K.S., Osborne, A., Manrique, P., Peters, T., Hawkes, D., Benavente, E.D., et al. (2024). Microhaplotype deep sequencing assays to capture*Plasmodium vivax*infection lineages. medRxiv, 2024.10.14.24315131. https://doi.org/10.1101/2024.10.14.24315131.

7. Lerch, A., Koepfli, C., Hofmann, N.E., Messerli, C., Wilcox, S., Kattenberg, J.H., Betuela, I., O'Connor, L., Mueller, I., and Felger, I. (2017). Development of amplicon deep sequencing markers and data analysis pipeline for genotyping multi-clonal malaria infections. BMC Genomics *18*, 864.

8. Ruybal-Pesántez, S., McCann, K., Vibin, J., Siegel, S., Auburn, S., and Barry, A.E. (2024). Molecular markers for malaria genetic epidemiology: progress and pitfalls. Trends Parasitol. *40*, 147–163.

9. Lerch, A., Koepfli, C., Hofmann, N.E., Kattenberg, J.H., Rosanas-Urgell, A., Betuela, I., Mueller, I., and Felger, I. (2019). Longitudinal tracking and quantification of individual Plasmodium falciparum clones in complex infections. Sci. Rep. *9*, 3333.

10. Lin, J.T., Hathaway, N.J., Saunders, D.L., Lon, C., Balasubramanian, S., Kharabora, O., Gosi, P., Sriwichai, S., Kartchner, L., Chuor, C.M., et al. (2015). Using Amplicon Deep Sequencing to Detect Genetic Signatures of Plasmodium vivax Relapse. J. Infect. Dis. *212*, 999–1008.

11. Wamae, K., Ndwiga, L., Kharabora, O., Kimenyi, K., Osoti, V., De Laurent, Z.R., Wambua, J., Musyoki, J., Ngetsa, C., Kalume, P., et al. (2022). Targeted amplicon deep sequencing of ama1 and mdr1 to track within-host P. falciparum diversity throughout treatment in a clinical drug trial [version 2; peer review: 1 approved, 1 approved with reservations]. Wellcome Open Res. *7:95*. https://doi.org/10.12688/wellcomeopenres.17736.2.

12. Jones, S., Kay, K., Hodel, E.M., Gruenberg, M., Lerch, A., Felger, I., and Hastings, I. (2021). Should Deep-Sequenced Amplicons Become the New Gold Standard for Analyzing Malaria Drug Clinical Trials? Antimicrob. Agents Chemother. *65*, e0043721.

13. He, X., Zhong, D., Zou, C., Pi, L., Zhao, L., Qin, Y., Pan, M., Wang, S., Zeng, W., Xiang, Z., et al. (2021). Unraveling the Complexity of Imported Malaria Infections by Amplicon Deep Sequencing. Front. Cell. Infect. Microbiol. *11*, 725859.

14. Wamae, K., Kimenyi, K.M., Osoti, V., de Laurent, Z.R., Ndwiga, L., Kharabora, O., Hathaway, N.J., Bailey, J.A., Juliano, J.J., Bejon, P., et al. (2022). Amplicon Sequencing as a Potential Surveillance Tool for Complexity of Infection and Drug Resistance Markers in Plasmodium falciparum Asymptomatic Infections. J. Infect. Dis. *226*, 920–927.

15. Holzschuh, A., Lerch, A., Fakih, B.S., Aliy, S.M., Ali, M.H., Ali, M.A., Bruzzese, D.J., Yukich, J., Hetzel, M.W., and Koepfli, C. (2024). Using a mobile nanopore sequencing lab for end-to-end genomic surveillance of Plasmodium falciparum: A feasibility study. PLOS Glob. Public Health *4*, e0002743.

16. LaVerriere, E., Schwabl, P., Carrasquilla, M., Taylor, A.R., Johnson, Z.M., Shieh, M., Panchal, R., Straub, T.J., Kuzma, R., Watson, S., et al. (2022). Design and implementation of multiplexed amplicon sequencing panels to serve genomic epidemiology of infectious disease: A malaria case study. Mol. Ecol. Resour. *22*, 2285–2303.

17. Kattenberg, J.H., Van Dijk, N.J., Fernandez-Minope, C.A., Guetens, P., Mutsaers, M., Gamboa, D., and Rosanas-Urgell, A. (2023). Molecular Surveillance of Malaria Using the PF AmpliSeq Custom Assay for Plasmodium falciparum Parasites from Dried Blood Spot DNA Isolates from Peru. Bio Protoc. *13*, e4621.

18. Kattenberg, J.H., Fernandez-Minope, C., van Dijk, N.J., Llacsahuanga Allcca, L., Guetens, P., Valdivia, H.O., Van Geertruyden, J.P., Rovira-Vallbona, E., Monsieurs, P., Delgado-Ratto, C., et al. (2023). Malaria Molecular Surveillance in the Peruvian Amazon with a Novel Highly Multiplexed Plasmodium falciparum AmpliSeq Assay. Microbiol. Spectr. *11*, e0096022.

19. Talundzic, E., Ravishankar, S., Kelley, J., Patel, D., Plucinski, M., Schmedes, S., Ljolje, D., Clemons, B., Madison-Antenucci, S., Arguin, P.M., et al. (2018). Next-generation sequencing and bioinformatics protocol for malaria drug resistance marker surveillance. Antimicrob. Agents Chemother. *62*. https://doi.org/10.1128/AAC.02474-17.

20. Huwe, T., Kibria, M.G., Johora, F.T., Phru, C.S., Jahan, N., Hossain, M.S., Khan, W.A., Price, R.N., Ley, B., Alam, M.S., et al. (2022). Heterogeneity in prevalence of subclinical Plasmodium falciparum and Plasmodium vivax infections but no parasite genomic clustering in the Chittagong Hill Tracts, Bangladesh. Malar. J. *21*, 218.

21. Wei, X., Malla, P., Wang, Z., Yang, Z., Cao, Y., Wang, C., and Cui, L. (2024). Genetic diversity of Plasmodium vivax population in northeast Myanmar assessed by amplicon sequencing of PvMSP1 and PvMSP3α. Acta Trop. *260*, 107461.

22. Tapaopong, P., da Silva, G., Holzschuh, A., Rungsarityotin, W., Suansomjit, C., Pumchuea, K., Manopwisedjaroen, K., Khamsiriwatchara, A., Khuntong, P., Cui, L., et al. (2025). Molecular epidemiology and genetic diversity of disappearing Plasmodium vivax in southern Thailand. Sci. Rep. *15*, 2620.

23. Reda, A.G., Huwe, T., Koepfli, C., Assefa, A., Tessema, S.K., Messele, A., Golassa, L., and Mamo, H. (2023). Amplicon deep sequencing of five highly polymorphic markers of Plasmodium falciparum reveals high parasite genetic diversity and moderate population structure in Ethiopia. Malar. J. *22*, 376.

24. Tapaopong, P., da Silva, G., Chainarin, S., Suansomjit, C., Manopwisedjaroen, K., Cui, L., Koepfli, C., Sattabongkot, J., and Nguitragool, W. (2023). Genetic diversity and molecular evolution of Plasmodium vivax Duffy Binding Protein and Merozoite Surface Protein-1 in northwestern Thailand. Infect. Genet. Evol. *113*,

105467.

25. Holzschuh, A., Lerch, A., Gerlovina, I., Fakih, B.S., Al-Mafazy, A.-W.H., Reaves, E.J., Ali, A., Abbas, F., Ali, M.H., Ali, M.A., et al. (2023). Multiplexed ddPCR-amplicon sequencing reveals isolated Plasmodium falciparum populations amenable to local elimination in Zanzibar, Tanzania. Nat. Commun. *14*, 3699.

26. Siegel, S.V., Trimarsanto, H., Amato, R., Murie, K., Taylor, A.R., Sutanto, E., Kleinecke, M., Whitton, G., Watson, J.A., Imwong, M., et al. (2024). Lineage-informative microhaplotypes for recurrence classification and spatio-temporal surveillance of Plasmodium vivax malaria parasites. Nat. Commun. *15*, 6757.

27. Kattenberg, J.H., Nguyen, H.V., Nguyen, H.L., Sauve, E., Nguyen, N.T.H., Chopo-Pizarro, A., Trimarsanto, H., Monsieurs, P., Guetens, P., Nguyen, X.X., et al. (2022). Novel highly-multiplexed AmpliSeq targeted assay for Plasmodium vivax genetic surveillance use cases at multiple geographical scales. Front. Cell. Infect. Microbiol. *12*, 953187.

28. Kattenberg, J.H., Cabrera-Sosa, L., Figueroa-Ildefonso, E., Mutsaers, M., Monsieurs, P., Guetens, P., Infante, B., Delgado-Ratto, C., Gamboa, D., and Rosanas-Urgell, A. (2024). Plasmodium vivax genomic surveillance in the Peruvian Amazon with Pv AmpliSeq assay. PLoS Negl. Trop. Dis. *18*, e0011879.

29. Popkin-Hall, Z.R., Niaré, K., Crudale, R., Simkin, A., Fola, A.A., Sanchez, J.F., Pannebaker, D.L., Giesbrecht, D.J., Kim, I.E., Jr, Aydemir, Ö., et al. (2024). High-throughput genotyping of Plasmodium vivax in the Peruvian Amazon via molecular inversion probes. Nat. Commun. *15*, 10219.

30. Hubbard, A., Solares, E., Bradley, L., Jeang, B., Yewhalaw, D., Janies, D., Lo, E., Yan, G., and Hemming-Schroeder, E. (2025). PvGAP: Development of a globally-applicable, highly-multiplexed microhaplotype amplicon panel for Plasmodium vivax. medRxiv. https://doi.org/10.1101/2025.04.30.25326751.

31. Rosado, J., White, M.T., Longley, R.J., Lacerda, M., Monteiro, W., Brewster, J., Sattabongkot, J., Guzman-Guzman, M., Llanos-Cuentas, A., Vinetz, J.M., et al. (2021). Heterogeneity in response to serological exposure markers of recent Plasmodium vivax infections in contrasting epidemiological contexts. PLoS Negl. Trop. Dis. *15*, e0009165.

32. Han, J., Munro, J.E., and Bahlo, M. (2023). AmpSeqR: an R package for amplicon deep sequencing data analysis [version 1; peer review: awaiting peer review]. F1000Res. *12*. https://doi.org/10.12688/f1000research.129581.1.

33. Rosas-Aguirre, A., Guzman-Guzman, M., Chuquiyauri, R., Moreno, M., Manrique, P., Ramirez, R., Carrasco-Escobar, G., Rodriguez, H., Speybroeck, N., Conn, J.E., et al. (2021). Temporal and Microspatial Heterogeneity in Transmission Dynamics of Coendemic Plasmodium vivax and Plasmodium falciparum in Two Rural Cohort Populations in the Peruvian Amazon. J. Infect. Dis. *223*, 1466–1477.

34. Carrasco-Escobar, G., Miranda-Alban, J., Fernandez-Minope, C., Brouwer, K.C., Torres, K., Calderon, M., Gamboa, D., Llanos-Cuentas, A., and Vinetz, J.M. (2017). High prevalence of very-low Plasmodium falciparum and Plasmodium vivax parasitaemia carriers in the Peruvian Amazon: insights into local and occupational mobility-related transmission. Malar. J. *16*, 415.

35. Mangold, K.A., Manson, R.U., Koay, E.S., Stephens, L., Regner, M., Thomson, R.B., Jr, Peterson, L.R., and Kaul, K.L. (2005). Real-time PCR for detection and identification of Plasmodium spp. J. Clin. Microbiol. *43*, 2435–2440.

36. Rosanas-Urgell, A., Mueller, D., Betuela, I., Barnadas, C., Iga, J., Zimmerman, P.A., del Portillo, H.A., Siba, P., Mueller, I., and Felger, I. (2010). Comparison of diagnostic methods for the detection and quantification of the four sympatric Plasmodium species in field samples from Papua New Guinea. Malar. J. *9*, 361.

37. Pearson, R.D., Amato, R., Auburn, S., Miotto, O., Almagro-Garcia, J., Amaratunga, C., Suon, S., Mao, S., Noviyanti, R., Trimarsanto, H., et al. (2016). Genomic analysis of local variation and recent evolution in Plasmodium vivax. Nat. Genet. *48*, 959–964.

38. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, Subgroup (2009). The Sequence Alignment/Map format and SAMtools.

Bioinformatics *25*, 2078–2079.

39. Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G. (2012). Primer3--new capabilities and interfaces. Nucleic Acids Res. *40*, e115.

40. Pagès, H., Aboyoun, P., Gentleman, R., and DebRoy, S. (2024). Biostrings: Efficient manipulation of biological strings. R package version 2.72.1.

41. Morgan, M., Anders, S., Lawrence, M., Aboyoun, P., Pages, H., and Gentleman, R. (2009). ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. Bioinformatics *25*, 2607–2608.

42. Havryliuk, T., and Ferreira, M.U. (2009). A closer look at multiple-clone Plasmodium vivax infections: detection methods, prevalence and consequences. Mem. Inst. Oswaldo Cruz *104*, 67–73.

43. Manrique, P., Miranda-Alban, J., Alarcon-Baldeon, J., Ramirez, R., Carrasco-Escobar, G., Herrera, H., Guzman-Guzman, M., Rosas-Aguirre, A., Llanos-Cuentas, A., Vinetz, J.M., et al. (2019). Microsatellite analysis reveals connectivity among geographically distant transmission zones of Plasmodium vivax in the Peruvian Amazon: A critical barrier to regional malaria elimination. PLoS Negl. Trop. Dis. *13*, e0007876.

44. Gerlovina, I., Gerlovin, B., Rodriguez-Barraquer, I., and Greenhouse, B. (2022). Dcifer: an IBD-based method to calculate genetic distance between polyclonal infections. Genetics *222*. https://doi.org/10.1093/genetics/iyac126.

45. Taylor, A.R., Foo, Y.S., and White, M.T. (2022). *Plasmodium vivax*relapse, reinfection and recrudescence estimation using genetic data. bioRxiv, 2022.11.23.22282669. https://doi.org/10.1101/2022.11.23.22282669.

46. Meacham, F., Boffelli, D., Dhahbi, J., Martin, D.I., Singer, M., and Pachter, L. (2011). Identification and correction of systematic error in high-throughput sequence data. BMC Bioinformatics *12*, 451.

47. Barry, A.E., Waltmann, A., Koepfli, C., Barnadas, C., and Mueller, I. (2015). Uncovering the transmission dynamics of Plasmodium vivax using population genetics. Pathog. Glob. Health *109*, 142–152.

48. Imwong, M., Boel, M.E., Pagornrat, W., Pimanpanarak, M., McGready, R., Day, N.P., Nosten, F., and White, N.J. (2012). The first Plasmodium vivax relapses of life are usually genetically homologous. J. Infect. Dis. *205*, 680–683.

49. Chen, N., Auliff, A., Rieckmann, K., Gatton, M., and Cheng, Q. (2007). Relapses of Plasmodium vivax infection result from clonal hypnozoites activated at predetermined intervals. J. Infect. Dis. *195*, 934–941.

50. Kho, S., Qotrunnada, L., Leonardo, L., Andries, B., Wardani, P.A.I., Fricot, A., Henry, B., Hardy, D., Margyaningsih, N.I., Apriyanti, D., et al. (2021). Hidden biomass of intact malaria parasites in the human spleen. N. Engl. J. Med. *384*, 2067–2069.

51. Taylor, A.R., Watson, J.A., Chu, C.S., Puaprasert, K., Duanguppama, J., Day, N.P.J., Nosten, F., Neafsey, D.E., Buckee, C.O., Imwong, M., et al. (2019). Resolving the cause of recurrent Plasmodium vivax malaria probabilistically. Nat. Commun. *10*, 5595.

52. Hasugian, A.R., Purba, H.L., Kenangalem, E., Wuwung, R.M., Ebsworth, E.P., Maristela, R., Penttinen, P.M., Laihad, F., Anstey, N.M., Tjitra, E., et al. (2007). Dihydroartemisinin-piperaquine versus artesunate-amodiaquine: superior efficacy and posttreatment prophylaxis against multidrug-resistant Plasmodium falciparum and Plasmodium vivax malaria. Clin. Infect. Dis. *44*, 1067–1074.

53. Pukrittayakamee, S., Imwong, M., Singhasivanon, P., Stepniewska, K., Day, N.J., and White, N.J. (2008). Effects of different antimalarial drugs on gametocyte carriage in P. vivax malaria. Am. J. Trop. Med. Hyg. *79*, 378–384.

54. Bousema, T., and Drakeley, C. (2011). Epidemiology and infectivity of Plasmodium falciparum and Plasmodium vivax gametocytes in relation to malaria control and elimination. Clin. Microbiol. Rev. *24*, 377–410.

55. Oyola, S.O., Ariani, C.V., Hamilton, W.L., Kekre, M., Amenga-Etego, L.N., Ghansah, A., Rutledge, G.G., Redmond,

S., Manske, M., Jyothi, D., et al. (2016). Whole genome sequencing of Plasmodium falciparum from dried blood spots using selective whole genome amplification. Malar. J. *15*, 597.

56. Bright, A.T., Manary, M.J., Tewhey, R., Arango, E.M., Wang, T., Schork, N.J., Yanow, S.K., and Winzeler, E.A. (2014). A high resolution case study of a patient with recurrent Plasmodium vivax infections shows that relapses were caused by meiotic siblings. PLoS Negl. Trop. Dis. *8*, e2882.

Acknowledgments

We thank the participants who anonymously agreed to participate in the ACT-Radical clinical trial and the ICEMR study. We acknowledge the field teams that contributed to sample collection in Solomon Islands and Peru, and Soazic Gardais for logistical support at Institut Pasteur Paris. We thank Leanne Robinson for provision of the control DNA samples. We thank Aimee Taylor for her input and advice on running Pv3Rs. This work was also made possible through the Victorian State Government Operational Infrastructure Support and Australian Government National Health and Medical Research Council Independent Research Institute Infrastructure Support Scheme.

Author's contributions

JR, IM, and SRP designed the study. JR and SRP designed PvAmpSeq protocols. ZT, KS, JB and CB did the genotyping. JH, JM, and MB developed AmpSeqR and preprocessed sequencing data, with contributions from JR, SRP, and CB. JR and SRP analysed the data and performed statistical analyses. JMV, MTW, MB, DG, and IM provided administrative and logistical support. JR and SRP wrote the first draft of the manuscript. All authors had full access to all the data. All authors contributed to the manuscript's writing.

Declaration of interest

The authors declare no conflicts of interest.

Funding

The study was supported by a grant from the Bill and Melinda Gates Foundation (funding grant ID number: OPP1151132). IM was supported by an Australian National Health and Medical Research Council (NHMRC) Investigator Grant (#2016726). MB was supported by an Australian NHMRC Investigator Grant (GNT1195236). The Peruvian cohort was part of the Amazonian International Center of Excellence in Malaria Research that received funding through the Cooperative Agreement U19AI089681 from the US Public Health Service, National Institutes of Health/National Institute of Allergy and Infectious Diseases, USA to JMV, and from a FOGARTY global infectious disease training grant (2D43TW007120-11A1, NIH-USA to JMV). JR postdoctoral fellowship was funded by Institut de Recherche pour le Développement. SR-P acknowledges funding from the MRC Centre for Global Infectious Disease Analysis (reference MR/X020258/1), funded by the UK Medical Research Council (MRC). This UK-funded award is carried out in the frame of the Global Health EDCTP3 Joint Undertaking.

Supplemental information

Document S1. Figures S1–S10

Document S2. Table S1–S5

Text S1. Detailed standard operational protocols for PvAmpSeq.

Text S2. Excel file containing additional data on PvAmpSeq markers, too large to fit in a PDF.

Text S3. Data of IBD classification in Solomon Islands cohort.

Text S4. Data of IBD classification in the Peruvian cohort.