

## RESEARCH ARTICLE

# Population genomic evidence of structured and connected *Plasmodium vivax* populations under host selection in Latin America

Johanna Helena Kattenberg<sup>1</sup>  | Pieter Monsieurs<sup>1</sup> | Julie De Meyer<sup>1</sup> |  
Katlijn De Meulenaere<sup>1</sup> | Erin Sauve<sup>1</sup> | Thaís C. de Oliveira<sup>2</sup> | Marcelo U. Ferreira<sup>2,3</sup> |  
Dionicia Gamboa<sup>4,5</sup> | Anna Rosanas-Urgell<sup>1</sup>

<sup>1</sup>Malaria Unit, Institute of Tropical Medicine Antwerp, Antwerp, Belgium

<sup>2</sup>Department of Parasitology, Institute of Biomedical Sciences, University of São Paulo, São Paulo, Brazil

<sup>3</sup>Global Health and Tropical Medicine, Institute of Hygiene and Tropical Medicine, Nova University of Lisbon, Lisbon, Portugal

<sup>4</sup>Instituto de Medicina Tropical "Alexander von Humboldt", Universidad Peruana Cayetano Heredia, Lima, Peru

<sup>5</sup>Laboratorio de Malaria: Parásitos y Vectores, Laboratorios de Investigación y Desarrollo, Departamento de Ciencias Celulares y Moleculares, Facultad de Ciencias e Ingeniería, Universidad Peruana Cayetano Heredia, Lima, Peru

## Correspondence

Johanna Helena Kattenberg and Anna Rosanas-Urgell, Malaria Unit, Institute of Tropical Medicine Antwerp, Antwerp, Belgium.

Email: [ekattenberg@itg.be](mailto:ekattenberg@itg.be) and [arosanas@itg.be](mailto:arosanas@itg.be)

## Present address

Julie De Meyer, Integrated Molecular Plant physiology Research (IMPRES) and Plants and Ecosystems (PLECO), Department of Biology, University of Antwerp, Antwerp, Belgium

## Funding information

Conselho Nacional de Desenvolvimento Científico e Tecnológico; Department of Economy, Science and Innovation in Flanders; Fundação de Amparo à Pesquisa do Estado de São Paulo, Grant/Award Number: 2016/187409; Fonds Wetenschappelijk Onderzoek, Grant/Award Number: 1S48419N and 1SC5522N; the Belgium Development Cooperation (DGD) under the Framework Agreement Program, Grant/Award Number: FA3III 2014 2016, FA4 2017 2021 and FA5 2022 2026; Research Foundation Flanders (FWO) and the Flemish Government; Foundation for the

## Abstract

Pathogen genomic epidemiology has the potential to provide a deep understanding of population dynamics, facilitating strategic planning of interventions, monitoring their impact, and enabling timely responses, and thereby supporting control and elimination efforts of parasitic tropical diseases. *Plasmodium vivax*, responsible for most malaria cases outside Africa, shows high genetic diversity at the population level, driven by factors like sub-patent infections, a hidden reservoir of hypnozoites, and early transmission to mosquitoes. While Latin America has made significant progress in controlling *Plasmodium falciparum*, it faces challenges with residual *P. vivax*. To characterize genetic diversity and population structure and dynamics, we have analyzed the largest collection of *P. vivax* genomes to date, including 1474 high-quality genomes from 31 countries across Asia, Africa, Oceania, and America. While *P. vivax* shows high genetic diversity globally, Latin American isolates form a distinctive population, which is further divided into sub-populations and occasional clonal pockets. Genetic diversity within the continent was associated with the intensity of transmission. Population differentiation exists between Central America and the North Coast of South America, vs. the Amazon Basin, with significant gene flow within the Amazon Basin, but limited connectivity between the Northwest Coast and the Amazon Basin. Shared genomic regions in these parasite populations indicate adaptive evolution, particularly in genes related to DNA replication, RNA processing, invasion, and motility – crucial for the

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

National Institutes of Health, Grant/Award Number: U19 AI089681; Fundação para a Ciência e a Tecnologia, Portugal, Grant/Award Number: UID/04413/2020

parasite's survival in diverse environments. Understanding these population-level adaptations is crucial for effective control efforts, offering insights into potential mechanisms behind drug resistance, immune evasion, and transmission dynamics.

#### KEYWORDS

genomic epidemiology, natural selection, parasitology, phylogeography, *Plasmodium vivax*, population dynamics

#### TAXONOMY CLASSIFICATION

Biogeography, Disease ecology, Microbial ecology, Parasitology, Population genetics

## 1 | INTRODUCTION

In an era characterized by rapid environmental changes, urbanization, and increasing human-animal interactions, the dynamics of infectious diseases are evolving at an unprecedented pace. Large-scale programs are dedicated to controlling or eliminating infectious diseases with the greatest global health impact, with many of these efforts focused on neglected tropical diseases (NTDs). While NTDs encompass fungal, viral, and bacterial infections, the majority are caused by parasites, particularly protozoa and helminths. Vector-borne parasitic diseases such as malaria, trypanosomiasis, leishmaniasis, and filariasis cause the greatest incidence and mortality globally (Cholewiński et al., 2015; GBD 2019 Child and Adolescent Communicable Disease Collaborators, 2023; Pearce & Tarleton, 2014).

Effective control of NTDs relies on the ability to monitor changes in pathogen populations, ensuring that interventions stay on track toward elimination goals and enabling targeted resource allocation. However, conventional monitoring techniques face challenges in many disease-endemic countries, where diagnostic tools are often limited. This task becomes increasingly difficult as disease prevalence decreases. Genomic epidemiology, however, can provide a deep understanding of parasite population dynamics, enabling strategic planning of control interventions, monitoring their effects, and raising alerts if necessary (Kwiatkowski, 2015) and hence, support disease eradication efforts by providing actionable knowledge (Cotton et al., 2018; Gardy & Loman, 2018; Grad & Lipsitch, 2014; World Health Organization, 2022a).

While genetic data are most extensively used for diseases caused by prokaryotes and viruses (Gardy & Loman, 2018), phylogenetic tools used in viral and bacterial genomics capture both epidemiological changes and evolutionary history, due to the high mutation rates in these pathogens and measurable genetic changes within the time frame of an outbreak or epidemic (Drummond et al., 2003; Duchêne et al., 2016; Grenfell et al., 2004). However, in pathogens with a lower mutation rate and frequent recombination, such as eukaryotic parasites, inferring transmission events is more challenging (Archie et al., 2009; Prugnolle & De Meeûs, 2008). The application of genomic epidemiology for these parasitic diseases has lagged, hindered by the complexity of the parasite's life cycle and the greater size of its

genome. Genetic diversity is influenced by various factors such as its life history, population dynamics, and recent changes in population size. It is crucial to have a comprehensive understanding of pathogen populations and an accurate assessment of their population structure over time to accurately evaluate the effectiveness of control interventions (Cotton et al., 2018). This information allows for a better understanding of inbreeding patterns and gene flow that can inform the development of improved strategies for controlling current populations.

While population genetics of several parasite species have been analyzed using microsatellite regions, the rapid innovation and decreasing cost of whole-genome sequencing make it the ideal tool, since genome-wide data have more resolution and are more comparable between populations and pathogens, eliminating the need for validated and standardized marker panels. For many key parasitic diseases, essential genomic resources like annotated reference genomes are already available. Genome-wide data can provide insights into the sudden emergence and spread of new pathogen genotypes, reveal recent strong selection on certain genome regions, and population evolution in response to treatment and control interventions when signs of a significant bottleneck are detected. An example is the identification of emerging drug resistance in the malaria parasite *Plasmodium falciparum* (Miotto et al., 2015).

Malaria, caused by *Plasmodium* parasites, contributes to a very high disease burden with an estimated 247 million malaria cases in 84 malaria-endemic countries (World Health Organization, 2022b). However, in several countries across the world where control efforts have reduced overall malaria cases, there has been an increase in the proportion of *Plasmodium vivax* (Price et al., 2020). Moreover, substantial reductions in *P. vivax* prevalence over 5–10 years in several locations have not consistently resulted in changes in population structure (Feachem et al., 2010; Kattenberg et al., 2020; Neafsey et al., 2012; Waltmann et al., 2018). *P. vivax* accounts for 18.0% to 71.5% of malaria cases outside Africa, with the highest proportion in the Americas, and this region contributes approximately 0.2% of global malaria cases (World Health Organization, 2022b). Venezuela, Colombia, Brazil, and Peru are the top four countries contributing the highest number of cases (79%) in the region (World Health Organization, 2022b). In contrast to other co-endemic regions of the world, *P. falciparum* is less common except in specific regions like Colombia's Pacific coast

(Rodríguez et al., 2011). Additionally, *Plasmodium malariae* infections are under-detected in the region, despite evidence of their presence, and zoonotic transmission of *Plasmodium brasilianum* and *Plasmodium simium* between non-human primates and humans is a concern (Recht et al., 2017). Many countries in Latin America have made strong progress in malaria control, reducing the malaria burden from 1.5 to 0.6 million cases between 2000 and 2021 (World Health Organization, 2022b). However, high transmission areas remain predominantly concentrated in the Amazon rainforest regions, disproportionately affecting indigenous and remote communities. In 2021, Venezuela, Colombia, Brazil, and Peru were in the top four countries contributing the most *P. vivax* cases (79%) in the region (World Health Organization, 2022b).

Genomic diversity in malaria parasites is generated through a combination of de novo mutations during asexual replication and sexual recombination within the mosquito vector. *Plasmodium* parasites have a high recombination rate, and frequent infections with multiple genetically distinct clones, especially in the case of *P. vivax* (Nkhoma et al., 2020; Siegel & Rayner, 2020). In addition, parasite genomes are polymorphic, with a diversity of phenotypic characteristics that impact disease severity (Neafsey et al., 2021). *P. vivax* often displays a higher genetic diversity than *P. falciparum*, due to key biological factors including frequent subpatent (i.e., detectable by molecular methods but not by field diagnostics) and asymptomatic infections, along with a hidden reservoir of hypnozoites leading to a larger number of complex infections (Olliaro et al., 2016; Sattabongkot et al., 2004). The asymptomatic infections and hypnozoites contribute to this parasite's resilience and facilitate its spread and gene flow across large regions, jeopardizing the effectiveness of local and targeted elimination strategies (Angrisano & Robinson, 2022; Auburn et al., 2021; Ferreira et al., 2022). Other factors contributing to the high genetic diversity of *P. vivax* are its longer history of association with humans, larger effective population size, and fewer population bottlenecks (Cornejo & Escalante, 2006; Hupalo et al., 2016; Neafsey et al., 2012; Noviyanti et al., 2015; Rougeron et al., 2022). Finally, sexual stages of *P. vivax* parasites appear early in the infection, facilitating effective transmission to mosquitoes before treatment, even at low-level parasitemia, making the disease more difficult to eliminate (Bousema & Drakeley, 2011; Sattabongkot et al., 2004).

In Latin America, the analysis of mitochondrial genomes has previously shown that the combined effects of geographical population structure and the relatively low incidence of *P. vivax* malaria has resulted in patterns of low local but high regional genetic diversity (Taylor et al., 2013). In this study, we take a population genomic approach to investigate the spatial temporal dynamics of *P. vivax* in this region, using genome wide data identified through literature and supplemented with data from our own studies ( $n=163$ ). Using high-resolution genome wide SNP variants (1,477,945 SNPs in the core genome) of these *P. vivax* isolates, we first compare the Latin American *P. vivax* genomes ( $n=399$ ) to *P. vivax* genomes from around the world ( $n=1075$ ). Next, we investigate the population structure, admixture, relatedness and geneflow, and signatures of positive selection to study local adaptations of the parasites. With this study, we investigate if and

how the declining and more heterogenous transmission is impacting *P. vivax* population structure in this relatively recently expanded population and discuss the factors driving diversity and population structure in this ecologically diverse region. Not only is this informative for malaria control and elimination strategies, but it can also identify targets and key pathways important for *P. vivax* survival.

## 2 | MATERIALS AND METHODS

### 2.1 | Sequencing data

Based on an exhaustive literature search on PubMed, publications describing new *P. vivax* genomes were identified until October 2022, and the corresponding sequencing data was downloaded from the Sequencing Read Archive (SRA) of NCBI (Adam et al., 2022; Auburn et al., 2016, 2019; Benavente et al., 2021; Brashear et al., 2020; Buyon et al., 2020; Chan et al., 2012; Chen et al., 2017; Cowell et al., 2018; Daron et al., 2021; De Meulenaere et al., 2023; de Oliveira et al., 2017, 2020; Delgado-Ratto et al., 2016; Dharia et al., 2010; Flannery et al., 2015; Hester et al., 2013; Hupalo et al., 2016; Kattenberg et al., 2022; Neafsey et al., 2012; Pearson et al., 2016; Popovici et al., 2018). Additionally, a set of new *P. vivax* genome sequencing data produced in the context of this study, were added to the list of genomes ( $n=163$ , originating from Peru, Brazil, Vietnam, Eritrea, Ethiopia, Burundi, Mauritania, Somalia, and Sudan), which have been described in more detail in (De Meulenaere et al., 2022, 2023; Kattenberg et al., 2022). Briefly, DNA was extracted from leukocyte-depleted red blood cells, whole blood or dried blood using the QIAmp DNA Blood Mini Kit (Qiagen, Germany) following the manufacturer's protocol as previously reported (De Meulenaere et al., 2022, 2023; Kattenberg et al., 2022). Parasite species was identified and quantified by a qPCR targeting Pv mtCOX1 (Gruenberg et al., 2018) using a standard curve of light microscopy quantified control isolates. Sequencing libraries were generated using Nextera XT DNA Sample Prep Kit (Illumina), or using commercial sequencing services as previously described (De Meulenaere et al., 2022, 2023; Kattenberg et al., 2022). Details of all *P. vivax* genomes used in this study can be found in Table S1, including metadata and accession numbers. The downloaded genomes contained monkey-adapted *P. vivax* strains that were removed from population genetic analyses.

### 2.2 | Ethics

Secondary use of all samples for sequencing and analysis of *P. vivax* isolates was approved through the Institutional Review Board of the Institute of Tropical Medicine Antwerp (protocols 1417/20 and 1345/19), the ethics committee at the University Hospital of Antwerp (protocol B3002020000016 and B300201523588), and Universidad Peruana Cayetano Heredia (UPCH; Lima, Peru) (protocol 101898).

## 2.3 | Variant detection

Sequencing reads were first aligned using BWA version 0.7.17 to the human reference genome obtained from the Genome Reference Consortium Human Build 38 patch release 13 (GRCh38.p13). Reads not mapped in proper pairs to the human reference genome were extracted using samtools version 1.10 (flag-F 2), and subsequently aligned to the *P. vivax* PvP01 reference genome from PlasmoDB (version 46) using BWA. Duplicate reads were removed with Picard's MarkDuplicates (version 2.22.4). Variant detection was performed using the Genome Analysis ToolKit (GATK) version 4.1.4.1, using in a first step the HaplotypeCaller command in GVCF mode for individual chromosomes. GVCF files were merged using the GenomicsDBImport, followed by genotyping using GenotypeGVCFs, resulting in one vcf file per chromosome. The vcf files were filtered according to GATK best practices: (1) SNPs were filtered out when having a QualByDepth value lower than 2, a variant quality score (QUAL) lower than 30, a StrandOddsRatio (SOR) higher than 3, a FisherStrand (FS) value higher than 60, or Root Mean Square mapping Quality (MQ) lower than 40, (2) Indels were filtered out when having a QualByDepth value lower than 2, a variant quality score lower than 30, FisherStrand (FS) value higher than 200 and ReadPosRankSum value lower than -20. Finally, for most downstream analysis the core genome (14 chromosomes, excluding subtelomeric regions and low-complexity domains and the apicoplast and mitochondrial sequences) (Pearson et al., 2016) was selected using the BCFtools query command and samples with less than 50% of the genome covered at least 5-fold were excluded from analysis.

## 2.4 | Population structure analysis

Principal Component Analysis (PCA) was performed using PLINK software version 2.0 (Chang et al., 2015). First, only biallelic SNPs with  $MAF > 0.005$  were selected, and linkage disequilibrium (LD) pruning was performed on the vcf file encompassing all variants in the core genome using PLINK, followed by PCA analysis using the first 20 principal components. PCA results were plotted in R using the ggplot2 library. Starting from the LD pruned dataset, admixture analysis was performed with the ADMIXTURE software version 1.3.0 (Alexander et al., 2009). The optimal number of populations was determined by running ADMIXTURE for a range of  $K$ -values (i.e., number of populations) from 2 to 50. This involved a 10-fold cross-validation, and selection of the  $K$ -value for the number of populations with the lowest cross-validation error. Phylogenetic trees were constructed by first converting the vcf file to PHYLIP format using the vcf2phyip.py script (Ortiz, 2019). Phylogenetic trees were constructed using RAxML, with *P. knowlesi* defined as outgroup, using the GTR+G evolutionary model and using a bootstrapping value of 100 (Kozlov et al., 2019). The phylogenetic tree was visualized using the ggtree library in R. Nucleotide diversity was determined

by sliding across the genome in 500-bp windows over all LD-pruned SNPs of the core genome using VcfTools (Danecek et al., 2011). The multiplicity of infections was calculated by estimating Wright's inbreeding co-efficient ( $F_{WS}$ ) as a measure of the within-host parasite diversity using the *getFws* command as implemented in the *moimix* package in R (Lee & Bahlo, 2016). Infections with  $F_{WS} \geq 0.95$  were considered to contain clonal (single strain) parasites, while samples with  $F_{WS} < 0.95$ , indicating within-host diversity, were considered to contain multiple genetically distinct parasite strains.

## 2.5 | IBD relatedness and selection analysis

Shared ancestry and relatedness between isolates was estimated using Identity-by-descent (IBD). PED and MAP file formats were created using VCFtools from an LD-pruned vcf dataset of the full genome (core + (sub)telomeric and low complexity regions of the 14 chromosomes) filtered on a MAF of 0.001 based on the frequency in all included 1474 genomes. IBD-sharing between pairs of samples, using all 399 samples from LAM, was calculated using the *isoRelate* package in R, which can analyze IBD in haploid recombining microorganisms in the presence of multiclonal infections (Danecek et al., 2011; Henden et al., 2018). Genetic distance was calculated using an estimated mean map unit size from *Plasmodium chabaudi* of 13.7 kb/centimorgan (cM) (Martinelli et al., 2005; Rovira-Vallbona et al., 2021). We set the thresholds of IBD at the minimum number of SNPs ( $n=20$ ) and length of IBD segments (5000 bp) reported to reduce false-positive calls using an error of 0.001. IBD has been shown to be superior to probabilistic models such as STRUCTURE for understanding the relatedness and interconnectivity of parasite populations (Henden et al., 2018; Taylor et al., 2019; Wesolowski et al., 2018). Networks of IBD-sharing (>10% of the genome shared) between individuals were created using the *igraph* package in R, and the cumulative level of IBD-sharing between isolates in countries in the network was plotted as a connection map with Scimago graphica (Hassan-Montero et al., 2022) and used as a measure of connectivity between countries.

For the samples from Latin America, the proportion of pairs of isolates sharing IBD, as well as significance of IBD-sharing was calculated using the *isoRelate* package in R for all samples together and subdivided by population, based on country, as a measure of positive selection.

## 2.6 | Pathway enrichment analysis – GO terms

Gene Ontology (GO) categories were sourced from PlasmoDB release 46, with each gene being associated with one or more GO categories. To analyze a list of specific genes, a gene set enrichment analysis was conducted utilizing the hypergeometric distribution, which assesses the statistical significance of the overlap between a gene list and the assigned GO categories based on their respective counts.

### 3 | RESULTS

#### 3.1 | *P. vivax* genomic data summary

Based on a literature search including manuscripts published before October 2022, we identified 1311 high-quality publicly shared *P. vivax* genomes. Raw sequencing data were downloaded and all genomes were combined, including in-house sequenced *P. vivax* genomes ( $n=163$ ) samples originating from Peru, Brazil, Vietnam, and imported cases in Belgium from travelers and migrants (De Meulenaere et al., 2022, 2023; Kattenberg et al., 2022).

A total of 1474 high-quality *P. vivax* genomes (Table S1), coming from 36 countries in Asia ( $n=878$ ), Americas ( $n=399$ ), and Africa ( $n=197$ ), and collected between 2000 and 2019, were retained after removing samples with less than 50% of the genome covered at least 5-fold (Figure 1). The median sequencing coverage over the PvP01 reference genome including only retained isolates was 26-fold (range 1–763). After alignment and variant calling, a total of 2,435,842 high quality genetic variants were identified (1,983,976 SNPs and 451,866 Indels), with a total of 1,836,935 variants in the core genome region (1,477,945 SNPs and 358,990 indels).

#### 3.2 | Global population structure

*Plasmodium vivax* genomes were grouped in regional populations (following classifications from Adam et al., 2022): Africa (AFR, including isolates from all countries in sub-Saharan Africa, and returning travelers with history of travel to these countries), Eastern South East Asia (ESEA, including isolates from Cambodia, Laos, Thailand, Vietnam, and the China-Myanmar border), Latin America (LAM, which includes isolates from Mexico, Central and South America), Middle South East Asia (MSEA, including isolates from Malaysia and The Philippines), Oceania (OCE, including isolates from the island of New Guinea [i.e., Papua New Guinea and part of Indonesia]),

Western Asia (WAS, which includes Afghanistan, Bangladesh, India, Iran, Pakistan, and Sri Lanka). To investigate genetic clustering of *P. vivax* populations in these regions we used the biallelic SNPs as input for PCA and phylogenetic analysis. Both analyses (PCA + tree) reveal the presence of three major clusters consistent with their geographical origin (Figure 2a,b). Isolates from ESEA + MSEA form a differentiated cluster in the vicinity of isolates from OCE. Isolates from AFR cluster close to isolates from WAS, however, these two regions are clearly separated in the fourth principal component of the PCA (Figure S1) and form separate clades in the tree (Figure 2b). Isolates from LAM form a distinct cluster and clade in the PCA and tree, respectively. Together these results, with the nucleotide diversity (Figure S2), indicates a high genetic diversity within the global *P. vivax* population as a whole, with a structuring of populations by geographical region.

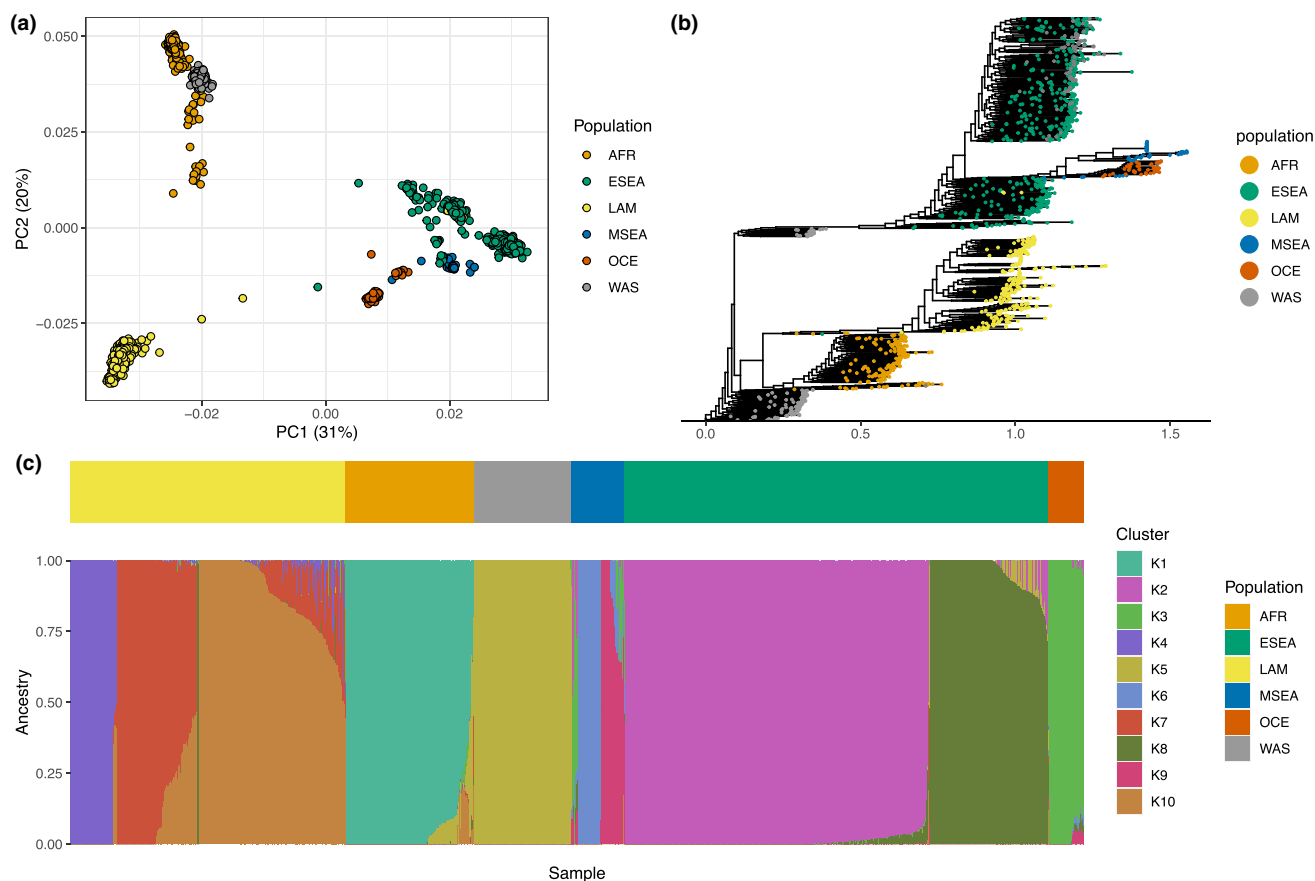
Admixture analysis estimated ten ( $K=10$ ) geographically distinct ancestral populations (Figure 2c). All genomes from AFR, WAS, and OCE were predicted to belong predominantly to a single shared ancestry within each region, while genomes from LAM, ESEA, and MSEA regions, each belong to distinct subpopulations (i.e., ancestral population within a region, Figure 2c). Admixture (predicted ancestry to more than one cluster) is mostly observed between subpopulations within a region (e.g., in LAM and ESEA), and rarely between regions, with the exception the admixture observed in AFR with WAS.

In the phylogenetic tree, isolates from WAS form two separate clades, with the upper cluster containing isolates from India (Figure 2b). This separate subpopulation could not be confirmed in the admixture analysis that estimated one ancestral cluster in this region (Figure 2c). Therefore, while Indian isolates might be distinct from other isolates in WAS, all *P. vivax* isolates from this region share a common ancestry. The highest amount of admixture between isolates is observed between the three subpopulations in LAM (mixed ancestry proportions to K7 and K10 and to a lesser extent K4), indicating a shared ancestry or gene flow between these subpopulations (Figure 2c).



**FIGURE 1** Origin of *Plasmodium vivax* genomes per country included in the analysis. Size of the dots are proportional to the number of samples in the genome dataset, and the colors indicate the country. Dots are plotted at the centre of the country (as defined by the ggmap package in R).





**FIGURE 2** Global *Plasmodium vivax* phylogeny, admixture, and population structure. (a) Principal component analysis based on the LD-pruned biallelic SNPs using PLINK2, showing the first two principal components. The samples (dots) are colored according to the originating population (here region). (b) Phylogenetic tree based on the LD-pruned biallelic SNPs using RAXML, with *P. knowlesi* defined as outgroup. The phylogenetic tree was visualized without the outgroup to improve clarity of the *P. vivax* branches in the figure. (c) Admixture proportions for  $K=10$  populations using the ADMIXTURE software, with in the small bar on top the region of origin, (AFR=Africa, ESEA=Eastern South East Asia, LAM=Latin America, MSEA=Middle South East Asia, OCE=Oceania, WAS=Western Asia).

### 3.3 | Population structure in Central and South America

To investigate shared ancestry of *P. vivax* in Latin America at a finer geographic resolution, the population genomic analyses were repeated including only isolates from this region ( $n=399$ ). Results from both the PCA and the phylogenetic tree indicated clustering on a country level (Figure S3).

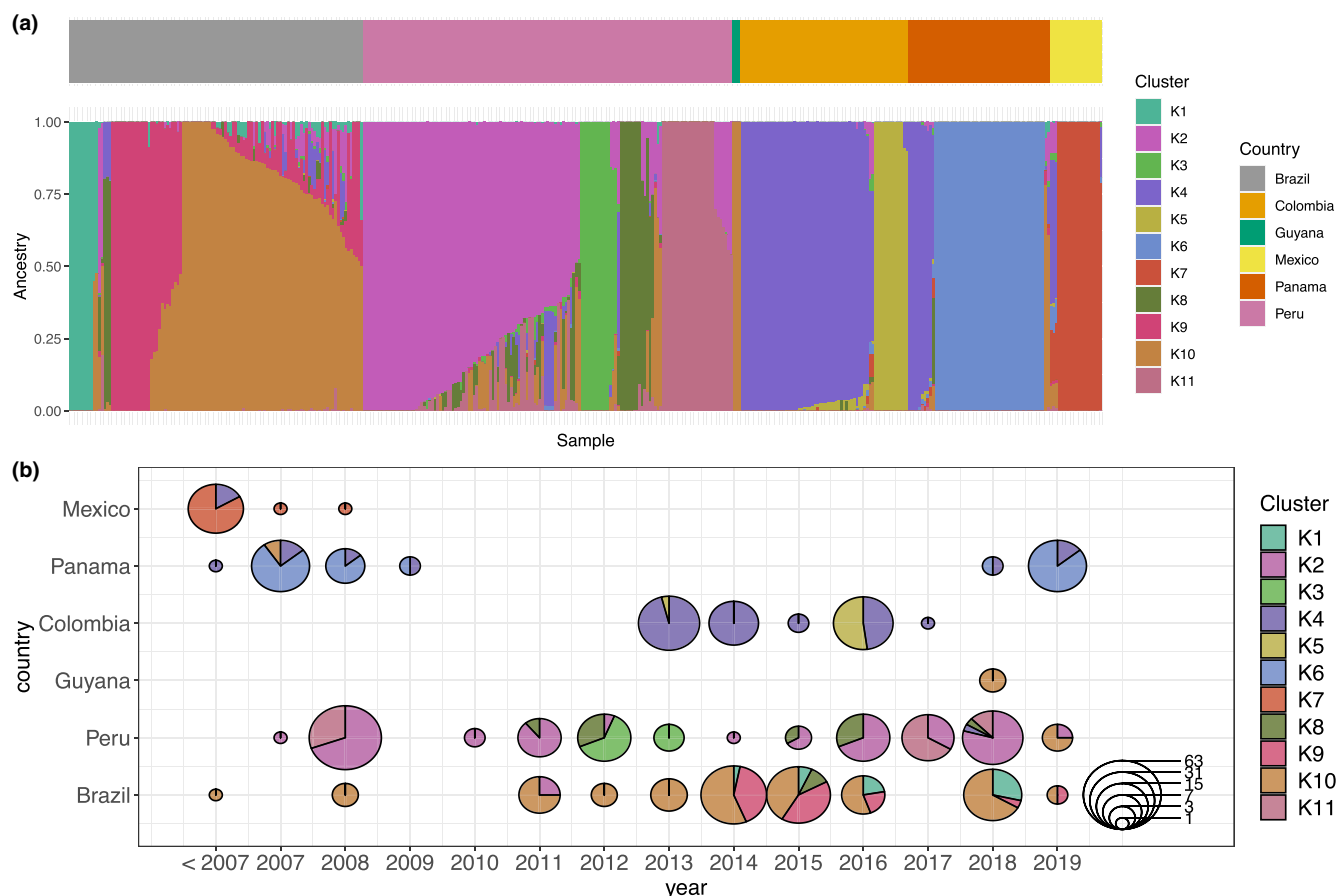
The high degree of admixture in LAM noted in the global comparison is confirmed in this analysis and constitutes, for a large part, admixed samples within Brazil and admixture between populations from Brazil and Peru (Figure 3a). Eleven ancestral clusters ( $K=11$ ) within LAM were estimated (Figure 3a), and these populations are structured geographically by country or at specific locations within a country (Figure S3). In addition, admixture is observed between isolates from Colombia, Mexico, and Panama with mixed ancestry from multiple populations across LAM. Country specific ancestral populations are seen in Mexico (K7), Panama (K6), Colombia (K5), Brazil (K1 and K9), and Peru (K3 and K11). In addition, some populations are seen in multiple countries, such as isolates from Mexico

and Panama that share ancestry with a population predominantly observed in Colombia (K4). While our dataset contains isolates sampled at different time periods, and populations are seen in multiple years (Figure 3b), we observed some distinct populations at specific locations, such as the Madre de Dios population (K3) in Peru, the K5 population in Tierralta in Colombia, and isolates from Manaus in Brazil (K9) (Figure S3).

Temporal analysis (Figure 3b) shows that the K10 sub-population that is predominant in Brazil across most years, is later also observed in other countries in the Amazon Basin (2018 in Guyana, and in 2019 in Peru in a region relatively close to the border with Brazil), and in two isolates in Panama from 2007 (Figure S3). Two additional populations are seen in Brazil that are predominant in Peru (K2 and K8).

### 3.4 | Gene flow across LAM

The connectivity between *P. vivax* populations in Latin American countries was assessed by measuring to what extent the parasite



**FIGURE 3** Spatio-temporal population dynamics in Latin America. Admixture analysis of *Plasmodium vivax* samples from LAM, using  $K=11$  populations. (a) Bar plot with admixture proportions of each sample for each ancestry cluster, with in the small bar on top the country of origin for each sample. (b) Each sample is assigned to one ancestry cluster based on the highest membership probability to that population in the admixture analysis. Pie charts represent the number of samples from each cluster in that country and year.

populations are genetically related. Pairwise IBD between all samples within and between countries was used as a measure of connectivity and parasite gene flow. From the 93,528 possible pairwise combinations of the 399 isolates from LAM, 1812 pairs of isolates (1.9%) had moderate-to-high relatedness (sharing 10–100% of their genome IBD). Among those, 638 pairs had high relatedness (more than 50% IBD, i.e., sibling or clonal pairs).

As expected, the majority of the related pairs (sharing 10–100% of their genome IBD) were observed within country (Figure 4 and Table 1), with observed relatedness between the different ancestral populations previously identified in Brazil (K1, K9, and K10) and Peru (K2, K8, K11) (Figure S4).

A small number of related pairs (sharing 10–50% of their genome IBD) was also observed between countries: between Brazil and Peru (K10–K2) and between Colombia and Panama (K4 and K5; Figure 4). This indicates connectivity and gene flow between the different ancestral clusters within and between countries in LAM.

To simplify subsequent analysis, parasites from Guyana are grouped with Brazil due to the low number of samples ( $n=3$ ) and assignment to the same population as samples from Brazil in the admixture analysis (Figure 3).

### 3.5 | Diversity

Genetic diversity of malaria parasites and frequency of multiple clone infections are often associated with transmission rates, as a result of changes in the *effective population size*. We indeed observe that the genetic diversity of *P. vivax* populations in LAM (mean nucleotide diversity ( $\pi$ ) across the genome of  $0.00071 \pm 0.00046$ ) is significantly lower ( $p < .0001$ ) than in most other regions of the world (except MSEA, which includes mostly highly related samples from Malaysia; Figure S4). Interestingly, several clonal *P. vivax* clusters (IBD-sharing  $\geq 99\%$  of the genome) are also observed within the LAM countries (Figure S5).

Genetic diversity was significantly different in all pairwise comparisons between LAM countries (pairwise *t*-tests with no assumption of equal variances,  $p < .0001$ , Figure 5b). From the LAM countries where genomes were available for this study, Brazil has the highest number of cases, which is reflected in the diversity, that was highest in Brazil, and lowest in Panama (Figure 5a,b). While currently the prevalence of malaria in Mexico is very low, observed  $\pi$  was higher than in Panama (Figure 5b). However, these samples were collected before 2007, with the majority from 2002, when reported incidence in Mexico was higher than in Panama during



**FIGURE 4** *Plasmodium vivax* IBD-based connectivity between countries in Latin America. Connectivity network of inferred IBD between *P. vivax* samples from Latin American countries. Edges connecting countries are cumulative IBD sharing between parasite pairs with at least 10% of their genomes from those countries (numbers of samples pairs are shown in Table 1). 10% IBD-sharing means that for these parasites at least 10% of their genomes descended from a common ancestor without intervening recombination, indicating distant to close relatedness. Node colors indicate the country of origin of the *P. vivax* genomes, and nodes were plotted on the map with known latitude and longitude of collection sites by district or if unknown in the respective country's capital (for example, in Guyana).

the years 2007–2009. Therefore, results conform with the pattern of lower genetic diversity observed at lower transmission levels. The majority of samples from LAM had monoclonal *P. vivax* infections (Figure 5c), but in the countries with higher numbers of cases (Colombia, Brazil, Peru), polyclonal infections ( $F_{WS} < 0.95$ ) were more frequent (15.6%, 19.0%, and 10.0% of samples, respectively).

### 3.6 | Regions of the genome under positive selection

To investigate positive selection in the *P. vivax* populations by country, the genome regions where isolates are identical (referred to as IBD-segments) was determined by scanning for IBD in 5000 bp windows. Shared regions of the genome with a high amount of IBD within these populations can be indicative of positive selection.

Some regions were very similar in many isolates (high proportion of IBD-sharing), while other regions showed more variety between isolates (low proportion of IBD sharing; Figure 6a). The IBD-segments that are shared by the greatest number of isolates are shared by a maximum of 4% of isolates (Figure 6). While this is a relatively low proportion, this is as expected from an admixed and recombining parasite population.

The highest amount of IBD-sharing was observed near and in sub-telomeric regions of the chromosomes, rather than in the core genome. Since IsoRelate can be sensitive to local differences in SNP density, and difficulties in accurate read mapping and variant calling in these regions can increase SNP density, we investigated whether the signals overlapped only in high SNP density regions (Figure S6). However, high IBD-sharing was also seen in regions where SNP density was similar this was caused by a higher SNP density (Figure S7).

High Tajima's D-values were also observed near the subtelomeric regions, confirming the selection signals (Figure S8).

The genes in regions with the highest amount of IBD-sharing between isolates (peaks in Figure 6a) in the populations are listed in Table S2. This includes, for example, two peaks on chromosome 4 (Figure 6a) that contain putative genes liver stage antigen 3 (Isa3) on the first peak and Cytosolically Exposed Rhoptry Leaflet Interacting protein 1 (cerli1) on the second peak, as well as a region with very little IBD-sharing (i.e., valley in Figure 6a, in a genomic region with high genetic variability) containing hypothetical protein PVP01\_0424500 (Figure 6a).

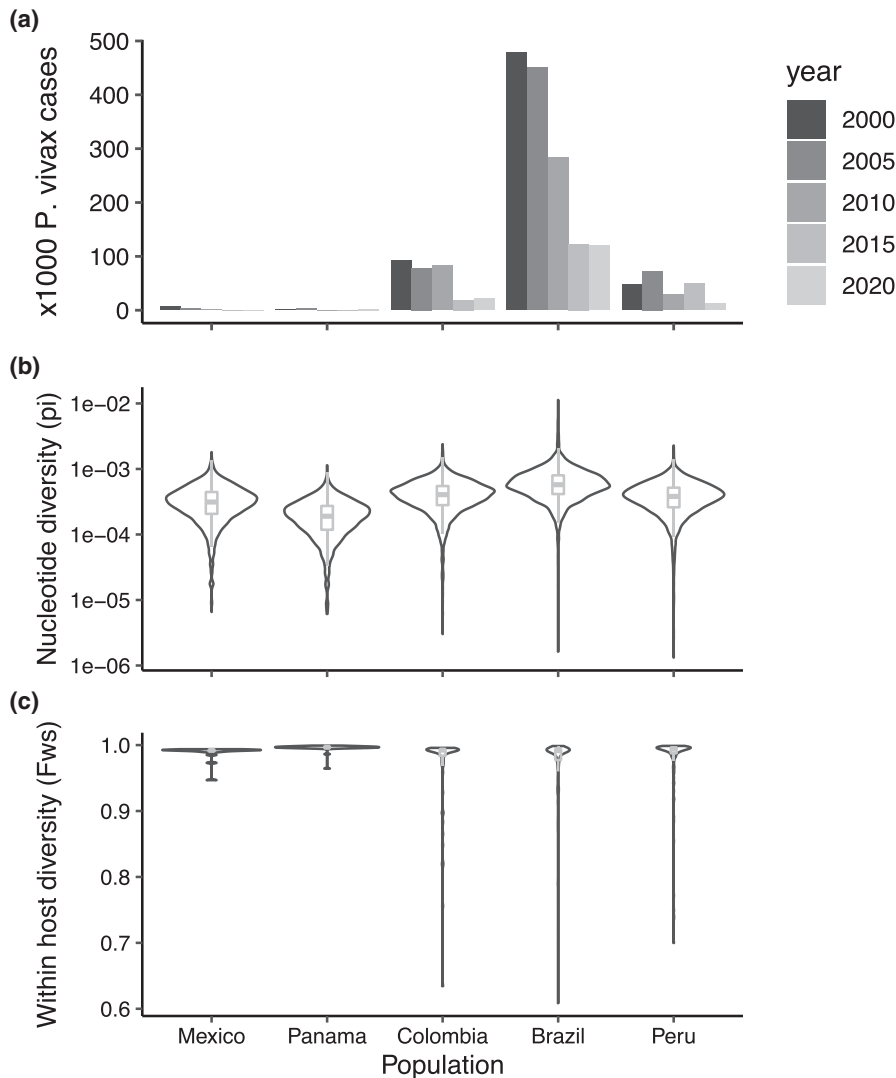
Significant IBD segments were investigated for each country resulting in some regions with significant IBD-sharing in single populations, and other regions conserved in multiple populations, such as a segment at the start of chromosome 9 (Figure 6b). Since many chromosomal regions with significant IBD segments were identified, we investigated the gene ontologies to determine which pathways were enriched in IBD segments of the different populations. Enrichment was found in pathways that are essential for (i) parasite replication, such as DNA-replication, binding and repair, protein folding, RNA transcription and processing, (ii) transport, (iii) invasion and antigenic variants, (e.g., Isa3, cerli1, merozoite surface protein 3 [msp3], tryptophan-rich proteins [trag6, trag7, trag20]), (iv) microtubule-related motility, and (v) male and female development (Table S2).

No known drug resistance associated genes or orthologues of *P. falciparum* resistance associated genes are located in the highest IBD regions, although dhfr, mdr1, and pvK13 are in areas with intermediate IBD (Figure 6a). Notably, there is an orthologue of a *P. falciparum* Kelch interacting protein (kic10) in a high IBD segment on chromosome 9, but the role of this protein and Kelch proteins in *P. vivax* drug resistance is unknown.



TABLE 1 Amount of *Plasmodium vivax* sample pairs with IBD (at  $\geq 10\%$  or  $\geq 50\%$  IBD) in pairwise analysis within and between samples from Latin American countries.

Country	Samples	Nr pairs with >50% IBD	Nr pairs with >10% IBD	Nr of possible pairs	% pairs with >50% IBD	% pairs with >10% IBD	Study references
Brazil	125	174	785	7750	2	10	Adam et al. (2022); Chan et al. (2012); De Meulenaere et al. (2023); de Oliveira et al. (2017, 2020); Hupalo et al. (2016); Kattenberg et al. (2022); Neafsey et al. (2012); Pearson et al. (2016); Benavente et al. (2021)
Colombia	65	108	149	2080	5	7	Adam et al. (2022); Hupalo et al. (2016)
Guyana	3	3	3	3	100	100	Benavente et al. (2021)
Mexico	20	41	125	190	22	66	Hupalo et al. (2016)
Panama	58	123	123	1653	7	7	Hupalo et al. (2016); Buyon et al. (2020)
Peru	162	188	591	13,041	1	5	Adam et al. (2022); Cowell et al. (2018); De Meulenaere et al. (2023); Delgado-Ratto et al. (2016); Dharia et al. (2010); Flannery et al. (2015); Hupalo et al. (2016); Kattenberg et al. (2022)
Comparison between countries							
Brazil-Peru	287	1	31	10,125	0.01	0.31	
Colombia-Panama	123	0	5	1885	0.00	0.27	



**FIGURE 5** Molecular markers for transmission intensity. From the studied countries, the number of *Plasmodium vivax* cases (a) has been highest in Brazil in all years (2010–2021), data extracted from the World Malaria report 2022 (World Health Organization, 2022b). (b) Violin plot of nucleotide diversity (pi) measured across the genome in 5000bp windows. (c) Violin plot of within-host infection complexity assessed using within sample  $F$  statistic ( $F_{WS}$ ).  $F_{WS} \geq 0.95$  was considered a proxy for a monoclonal infection.

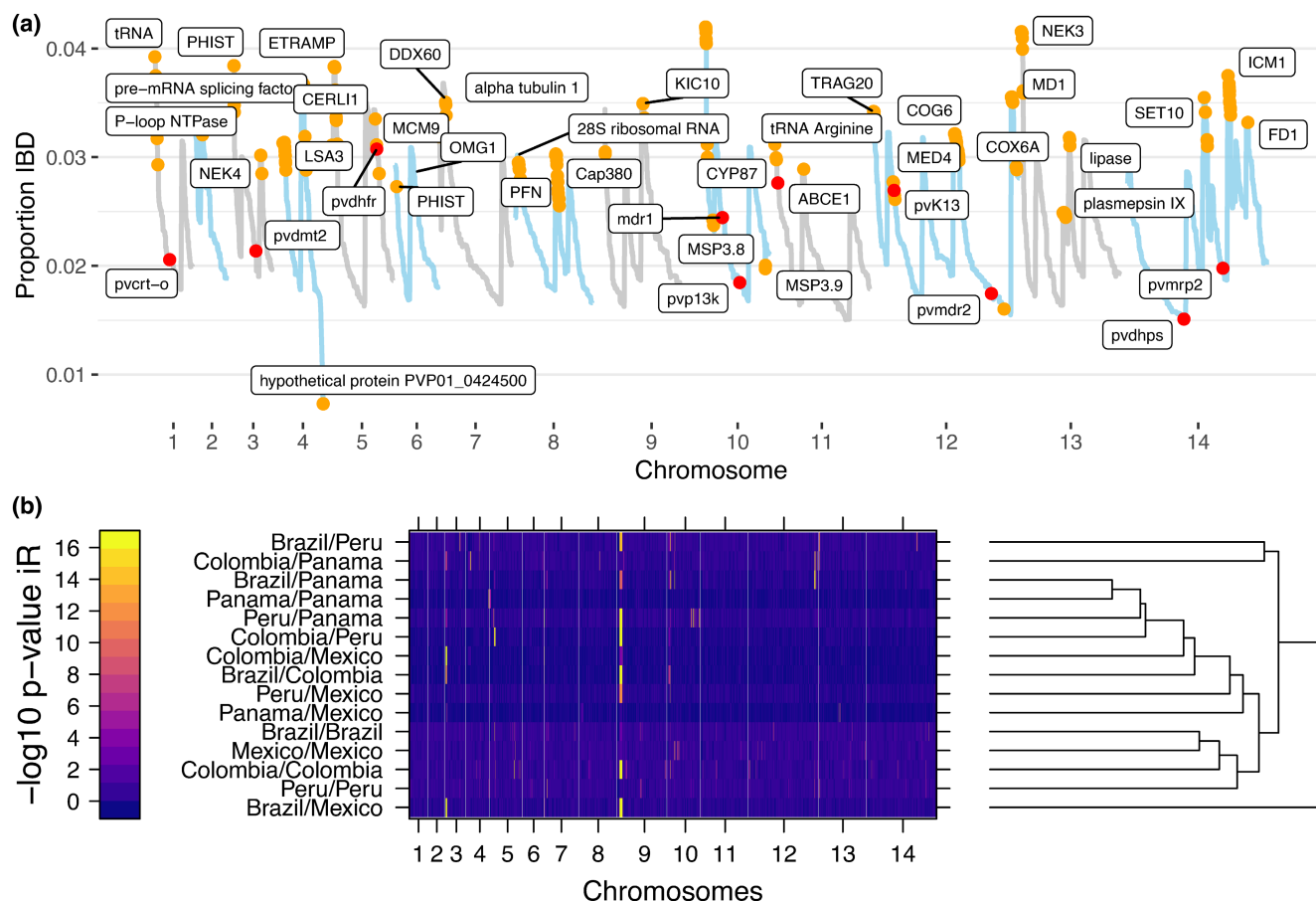
## 4 | DISCUSSION

Here, we report the genomic analysis of the largest collection ( $n=1474$ ) of high-quality *P. vivax* genomes to date, originating from 31 countries across Asia, Africa, Oceania, and Americas (Figure 1). The global *P. vivax* population displays high genetic diversity compared to high diversity *P. falciparum* populations in high transmission regions (Ocholla et al., 2014; Volkman et al., 2007), and separates into three main continental populations, i.e., Eastern Asia and Oceania, Africa and West Asia, and Latin America, similar to other reports (Adam et al., 2022; Daron et al., 2021; Ibrahim et al., 2023). The geographical clustering largely matches the previously defined zoogeographical zones with similar ecological and epidemiological characteristics and *P. vivax* relapse rates (Battle et al., 2014; Macdonald, 1957). Most of the clusters identified in the admixture analysis are relatively isolated within one region, with little admixture between regions, with the exception of Africa and West Asia, as reported previously (Benavente et al., 2021). In contrast, a high degree of admixture is observed within regions, especially in South-East Asia and in LAM. In this study we increased the number of

genomes compared to previous reports, in particular from West Africa, Vietnam, Brazil and Peru, with the largest genomic analysis of Latin American isolates ( $n=399$ ) to date. With additional samples from South-East Asia compared to earlier studies (Adam et al., 2022; Benavente et al., 2021; Neafsey et al., 2012), we do not find additional subpopulations in admixture analysis in this region. However, with the increased sample size in LAM, we do find an additional subpopulation in Latin America compared to earlier studies (Benavente et al., 2021; Cornejo et al., 2015). This increased resolution is confirmed in another very recent study with additional samples from Brazil (Ibrahim et al., 2023).

*P. vivax* isolates from Latin America form a distinct group within the global *P. vivax* population, characterized by structured subpopulations, sporadic clonal clusters at specific sites, and lower genetic diversity compared to other regions in the world. This analysis confirms the pattern of low-local but relatively high-regional genetic diversity in LAM measured previously using microsatellites (Taylor et al., 2013).

The lower genetic diversity of *P. vivax* in LAM might be explained by a founder effect of relatively recent colonization events of this



**FIGURE 6** Pairwise IBD between isolates across the five populations in LAM. (a) Line plot of median IBD shared between pairs of *Plasmodium vivax* samples along the chromosomes. IBD segments with highly significant IBD are indicated in orange and annotated genes of interest at peaks of IBD-sharing are labeled. The top genes that share significant IBD in the populations are listed in Table S2. Labeled dots in red indicate the positions and level of IBD sharing of putative drug resistant associated genes (list from Kattenberg et al., 2022). (b) Heatmap of significant pairwise IBD between populations in LAM are clustered on rows for similar patterns between populations. Low to high  $-\log_{10}$  p-values indicating significance levels of IBD sharing are color graded from blue to yellow. Significant IBD-sharing is seen at a  $-\log_{10} p > 1.3$  (i.e.,  $p < .05$ ), and a threshold of  $-\log_{10} p > 10$  was used to identify highly significant areas of IBD-sharing.

parasite species in this region, potentially through multiple waves of human migrations (Rodrigues et al., 2018; Rougeron et al., 2022; Steverding, 2020; van Dorp et al., 2020). In addition, a lower transmission intensity in LAM compared to other regions in the world can contribute to the lower diversity (Neafsey et al., 2012, 2021). In turn, Brazil, which is the country with the highest incidence of malaria – among those included in the LAM region in this study – exhibited the greatest level of genetic diversity. Furthermore, in comparison to genetic diversity of *P. falciparum* populations in high transmission African regions *P. vivax* diversity in LAM is high (Ocholla et al., 2014; Volkman et al., 2007). This is typically the effect of random genetic drift in small populations that remain relatively isolated from each other; rare alleles disappear (decreased diversity) and increased divergence between sites. Gene flow and admixture were detected across populations from Mexico, Panama, and the North Coast of Colombia, and across countries in the Amazon region, however connectivity between these two separate regions is highly limited. This marked population differentiation between these two regions might be attributed to factors such as differentiated ecological

environments, distinct vector populations, different human occupation and mobility patterns, and history of antimalarial interventions (Batista et al., 2015; Siqueira et al., 2016; Van den Eede et al., 2010). These differentiation patterns are very similar to those observed in *P. falciparum* populations in South America (Yalcindag et al., 2012). While the Amazon region and coastline regions of South and Central America are suitable habitats for various malaria vectors, there are ecological and human barriers between the Central and Northern South American coastlines and the Amazon region. These regions are separated by mountain ranges and have clear differences in climate and vegetation where the north coast region has a tropical climate with a dry season, while the Amazon rainforest has a tropical rainforest climate with year-round rainfall. The observed limited connectivity of parasite populations also suggests limited human mobility and thereby limited transmission between these regions.

Indeed, the Amazon basin is a vast territory of tropical rainforest encompassing several countries and including large regions of indigenous territories intersected by numerous rivers and scattered cities and villages that offer natural barriers contributing to the observed

subpopulation structure (Siqueira et al., 2016). On the other hand, connectivity and admixtures within this region in all likelihood is due to patterns of human mobility, which are closely related to occupational patterns like logging and illegal gold mining (Arisco et al., 2021; Castro et al., 2019; de Castro et al., 2006; Pacheco et al., 2019; Rosas-Aguirre et al., 2016; Torres et al., 2022). In addition, transmission of malaria (and gene flow) across large geographic distances is facilitated by the large human reservoir of asymptomatic and low-density *P. vivax* infections characteristic for this region (de Oliveira et al., 2020; Ferreira et al., 2022) as well as the dormant reservoir of hypnozoites characteristic for *P. vivax* in general (White et al., 2014).

Conversely, further North, in Colombia, Panama, and Mexico we observe less admixture than in the Amazon region and distinct populations by country. We identified a specific subpopulation that is predominant in Colombia, but also observed (in a lower proportion) in Panama and even into Mexico likely reflecting human migration facilitating spread of this population from the North coast of Colombia into countries further to the North. According to a recent study, the majority of *P. vivax* genomes in Panama are part of a highly clonal population that has been present in the country for at least 10 years (Buyon et al., 2020). In addition, the same study reported likely imported parasites, which we identified to belong to a Colombian population.

We identified genomic regions shared across Latin American parasite populations that exhibit signatures of positive selection. Genes in these shared regions are predominantly involved in DNA replication binding and repair, RNA transcription and processing, parasite invasion, as well as microtubule-related motility, suggesting that genes involved in these biological processes are key for *P. vivax* evolution and survival. Positive selection on DNA replication, binding, and repair genes suggests adaptive evolution to environmental challenges such as host immune responses, while positive selection on genes associated with RNA transcription and processing hints at the potential role of post-transcriptional gene regulation in parasite adaptation and survival to varying conditions (Parobek et al., 2016), such as parasite transition between hosts or life cycle stages (such as the male and female development through MD1 and FD1). Moreover, microtubules are essential components of the cytoskeleton and play a critical role in cell division, motility, and intracellular transport.

The selection signals in LAM from this study are in agreement with other studies of *P. vivax* populations from South America (Cornejo et al., 2015; Ibrahim et al., 2023), supporting the validity of our IBD-approach including polyclonal infections for selection analysis. Considering the similar approaches and findings from these studies using partly the same data, we did not further confirm selection signals with other methods. The selection signals identified in LAM are dissimilar to *P. vivax* populations in other regions, as previous investigations with global *P. vivax* isolate collections detected selective sweeps at drug resistance-associated loci (e.g., *dhfr*, *dhps*, and *mdr1*) (Hupalo et al., 2016; Pearson et al., 2016). However, these studies also found evidence of local adaptation within distinct *P. vivax* populations and differential selection on surface antigens such as *msp* genes. In contrast to other studies, we included sub-telomeric

regions in the selection analysis. We identified genes under selection putatively involved in antigenicity and host-parasite interactions, including parasite invasion, which is poorly understood in *P. vivax*. Positive selection in invasion genes has the potential to identify new candidate genes with a role in reticulocyte invasion and hence can inform vaccine development, for example, against antigens such as *Isa3* and *msp3* found under selection here. In *P. falciparum*, *Isa3* is a candidate vaccine target and has been used to elicit sterile immunity in animal models (Daubersies et al., 2008; Ghosn et al., 2018; Perlaza et al., 2008).

Sub-telomeric regions are often excluded in genome analysis due to challenges in aligning their highly variable and polymorphic short reads, particularly with the incomplete PvSall reference genome, but newer genomes like PvP01 now provide improved assembly for these regions (De Meulenaere et al., 2023; Minassian et al., 2021). Our results of positive selection in these sub-telomeric regions match well with a proposed model of *P. vivax* adaptation from a genomic study of *P. vivax* strains compared to *P. cynomolgi* and *P. knowlesi* (Cornejo et al., 2015). That study proposed that positive and negative selection might be less effective in low recombination areas (i.e., in the core genome). They suggested that in *P. vivax*, genome structure might be an adaptive mechanism to deal with changing environments, like the host's immune system, in addition to adaptations through genetic variation. In contrast, in *P. falciparum* populations, which need to survive in the host for a longer period before they can be transmitted, strong selection signals and high IBD are found in the core genome, for example, surrounding drug resistance associated genes (Amambua-Ngwa et al., 2019; Oyebola et al., 2017; Verity et al., 2020).

This study used a convenience sampling approach, limiting the ability to generalize findings. The majority of genomes included were from Brazil, Peru, and the Colombian coastal region, while genomes from other high *P. vivax* burden areas such as Venezuela, Nicaragua, and the Amazon region of Colombia are lacking. To complete the picture of genetic diversity in Latin America genomes from these regions, often remote or challenging to access, need to be generated.

This regional analysis of *P. vivax* populations in Latin America highlights the significant genetic diversity within the continent, and regional adaptation of the parasites to their hosts and different environmental challenges, which may contribute to the resilience of *P. vivax* to current malaria control strategies. Connectivity patterns between parasite populations are highly relevant for control and elimination programs in the region where human mobility is a major driver of malaria transmission. IBD-analysis confirmed connectivity between different ancestral clusters within and between countries, while showing a lack of connectivity between parasites in northern regions of Central and South America and the Amazon parasite population. While temporal differences in relatedness can serve as evidence for changes in transmission over time (Arambepola et al., 2023), the populations identified were very stable over time, indicating ongoing and stable transmission.

Genomic surveillance of parasitic pathogens, coupled with epidemiological data, offers valuable insights for control and elimination

programs. This study highlights the use of genetic information to uncover patterns to address key epidemiological questions related to *P. vivax* in Latin America, such as the observed association between transmission patterns and diversity in LAM, and lack of connectivity between populations in the North Coast and the Amazon Basin.

Finally, as malaria risk is increasing in spatial and temporal variability worldwide, the diversity of available malaria control tools continues to grow, and funding constraints on control programs increase, there is an increasing need for malaria-endemic countries to adopt intervention policies that move away from a one-size-fits-all approach to one that is specifically tailored to their subnational context. However, the high level of connectivity between regions and countries observed in this study rather shows that a wider view and regional approach remains very important in this continent and should remain incorporated in regional malaria control and elimination strategies.

## AUTHOR CONTRIBUTIONS

**Johanna Helena Kattenberg:** Conceptualization (equal); formal analysis (equal); methodology (lead); resources (equal); visualization (equal); writing – original draft (lead); writing – review and editing (lead). **Pieter Monsieurs:** Conceptualization (equal); data curation (lead); formal analysis (equal); methodology (equal); visualization (equal); writing – original draft (supporting); writing – review and editing (supporting). **Julie De Meyer:** Data curation (equal); formal analysis (supporting); writing – review and editing (supporting). **Katlijn De Meulenaere:** Data curation (equal); formal analysis (supporting); resources (equal); writing – review and editing (supporting). **Erin Sauve:** Data curation (supporting); resources (equal); writing – review and editing (supporting). **Thaís C. de Oliveira:** Resources (equal); writing – review and editing (supporting). **Marcelo U. Ferreira:** Resources (equal); writing – review and editing (supporting). **Dionicia Gamboa:** Resources (equal); writing – review and editing (supporting). **Anna Rosanas-Urgell:** Conceptualization (supporting); funding acquisition (lead); methodology (supporting); project administration (lead); supervision (lead); writing – original draft (supporting); writing – review and editing (supporting).

## ACKNOWLEDGEMENTS

We wish to thank all patients who donated their material for generation of parasite genomic data, and for all researchers that have openly shared *P. vivax* genomic data. Sample collection in Peru and molecular analysis were supported by the Belgium Development Cooperation (DGD) under the Framework Agreement Program between ITM and Peru (FA3-III 2014-2016, FA4 2017-2021, and FA5 2022-2026). This work was supported by the Department of Economy, Science and Innovation in Flanders (SOFI to ARU). Field work in Brazil and genome sequencing of Brazilian samples was supported by the Fundação de Amparo à Pesquisa do Estado de São Paulo, Brazil (2016/18740-9) and the National Institutes of Health, United States (U19 AI089681). TCO and MUF receive or received scholarships from Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil. MUF is also supported by the Fundação para a

Ciência e a Tecnologia, Portugal (GHTM project, UID/04413/2020). KDM and ES received a scholarship from the Research Foundation Flanders (1S48419N KDM and 1SC5522N ES). The computational resources and services used in this work were provided by the 1, funded by the Research Foundation Flanders (FWO) and the Flemish Government. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## OPEN RESEARCH BADGES



This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at <https://www.ncbi.nlm.nih.gov/sra>, <https://www.ebi.ac.uk/ena>, <https://www.ebi.ac.uk/eva>, and [https://github.com/pietermonsieurs/pvivax\\_genomes](https://github.com/pietermonsieurs/pvivax_genomes).

## DATA AVAILABILITY STATEMENT

Related metadata of all genomes analyzed in this study can be found in Table S1. This includes the accession numbers for all raw sequences used. Genotype data in the form of a vcf file that was used in the analysis of this study is openly available from the European Variation Archive (EVA) under project number: PRJEB65949.

Scripts used for data processing, analysis and plotting are shared on github: [https://github.com/pietermonsieurs/pvivax\\_genomes](https://github.com/pietermonsieurs/pvivax_genomes).

Benefits Generated: Benefits from this research accrue from the sharing of our scripts, data and results on public databases as described above.

## ORCID

Johanna Helena Kattenberg <https://orcid.org/0000-0002-2971-5136>

## REFERENCES

- Adam, I., Alam, M. S., Alemu, S., Amaratunga, C., Amato, R., Andrianaranjaka, V., Anstey, N. M., Aseffa, A., Ashley, E., Assefa, A., Auburn, S., Barber, B. E., Barry, A., Batista Pereira, D., Cao, J., Chau, N. H., Chotivanich, K., Chu, C., Dondorp, A. M., ... Yilma, D. (2022). An open dataset of *Plasmodium vivax* genome variation in 1,895 worldwide samples. *Wellcome Open Research*, 7, 136. <https://doi.org/10.12688/WellcomeOpenRes.17795.1>
- Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), 1655–1664. <https://doi.org/10.1101/gr.094052.109>
- Amambua-Ngwa, A., Amenga-Etego, L., Kamau, E., Amato, R., Ghansah, A., Golassa, L., Randrianarivelojosia, M., Ishengoma, D., Apinjoh, T., Maïga-Ascofaré, O., Andagalu, B., Yavo, W., Bouyou-Akotet, M., Kolapo, O., Mane, K., Worwui, A., Jeffries, D., Simpson, V., D'Alessandro, U., ... Djimde, A. A. (2019). Major subpopulations of *Plasmodium falciparum* in sub-Saharan Africa. *Science*, 365(6455), 813–816. <https://doi.org/10.1126/science.aav5427>
- Angrisano, F., & Robinson, L. J. (2022). *Plasmodium vivax* – How hidden reservoirs hinder global malaria elimination. *Parasitology International*, 87, 102526. <https://doi.org/10.1016/J.PARINT.2021.102526>
- Arambepola, R., Bérubé, S., Freedman, B., Taylor, S. M., Prudhomme O'Meara, W., Obala, A. A., & Wesolowski, A. (2023). Exploring



- how space, time, and sampling impact our ability to measure genetic structure across *Plasmodium falciparum* populations. *Frontiers in Epidemiology*, 3, 1058871. <https://doi.org/10.3389/fepid.2023.1058871>
- Archie, E. A., Luikart, G., & Ezenwa, V. O. (2009). Infecting epidemiology with genetics: A new frontier in disease ecology. *Trends in Ecology & Evolution*, 24(1), 21–30. <https://doi.org/10.1016/j.tree.2008.08.008>
- Arisco, N. J., Peterka, C., & Castro, M. C. (2021). Cross-border malaria in northern Brazil. *Malaria Journal*, 20(1), 135. <https://doi.org/10.1186/s12936-021-03668-4>
- Auburn, S., Böhme, U., Steinbiss, S., Trimarsanto, H., Hostetler, J., Sanders, M., Gao, Q., Nosten, F., Newbold, C. I., Berriman, M., Price, R. N., & Otto, T. D. (2016). A new *Plasmodium vivax* reference sequence with improved assembly of the subtelomeres reveals an abundance of pir genes. *Wellcome Open Research*, 1, 4. <https://doi.org/10.12688/wellcomeopenres.9876.1>
- Auburn, S., Cheng, Q., Marfurt, J., & Price, R. N. (2021). The changing epidemiology of *Plasmodium vivax*: Insights from conventional and novel surveillance tools. *PLoS Medicine*, 18(4), e1003560. <https://doi.org/10.1371/journal.pmed.1003560>
- Auburn, S., Getachew, S., Pearson, R. D., Amato, R., Miotto, O., Trimarsanto, H., Zhu, S. J., Rumaseb, A., Marfurt, J., Noviyanti, R., Grigg, M. J., Barber, B., William, T., Goncalves, S. M., Drury, E., Sriprawat, K., Anstey, N. M., Nosten, F., Petros, B., ... Price, R. N. (2019). Genomic analysis of *Plasmodium vivax* in Southern Ethiopia reveals selective pressures in multiple parasite mechanisms. *The Journal of Infectious Diseases*, 220(11), 1738–1749. <https://doi.org/10.1093/infdis/jiz016>
- Batista, C. L., Barbosa, S., Da Silva Bastos, M., Viana, S. A. S., & Ferreira, M. U. (2015). Genetic diversity of *Plasmodium vivax* over time and space: A community-based study in rural Amazonia. *Parasitology*, 142(2), 374–384. <https://doi.org/10.1017/S0031182014001176>
- Battle, K. E., Karhunen, M. S., Bhatt, S., Gething, P. W., Howes, R. E., Golding, N., Van Boeckel, T. P., Messina, J. P., Shanks, G. D., Smith, D. L., Baird, J. K., & Hay, S. I. (2014). Geographical variation in *Plasmodium vivax* relapse. *Malaria Journal*, 13(1), 144. <https://doi.org/10.1186/1475-2875-13-144>
- Benavente, E. D., Manko, E., Phelan, J., Campos, M., Nolder, D., Fernandez, D., Velez-Tobon, G., Castaño, A. T., Dombrowski, J. G., Marinho, C. R. F., Aguiar, A. C. C., Pereira, D. B., Sriprawat, K., Nosten, F., Moon, R., Sutherland, C. J., Campino, S., & Clark, T. G. (2021). Distinctive genetic structure and selection patterns in *Plasmodium vivax* from South Asia and East Africa. *Nature Communications*, 12(1), 3160. <https://doi.org/10.1038/s41467-021-23422-3>
- Bousema, T., & Drakeley, C. (2011). Epidemiology and infectivity of *Plasmodium falciparum* and *Plasmodium vivax* gametocytes in relation to malaria control and elimination. *Clinical Microbiology Reviews*, 24(2), 377–410. <https://doi.org/10.1128/CMR.00051-10>
- Brashear, A. M., Fan, Q., Hu, Y., Li, Y., Zhao, Y., Wang, Z., Cao, Y., Miao, J., Barry, A., & Cui, L. (2020). Population genomics identifies a distinct *Plasmodium vivax* population on the China-Myanmar border of Southeast Asia. *PLoS Neglected Tropical Diseases*, 14(8), e0008506. <https://doi.org/10.1371/journal.pntd.0008506>
- Buyon, L. E., Santamaria, A. M., Early, A. M., Quijada, M., Barahona, I., Lasso, J., Avila, M., Volkman, S. K., Marti, M., Neafsey, D. E., & Obaldia, N., III. (2020). Population genomics of *Plasmodium vivax* in Panama to assess the risk of case importation on malaria elimination. *PLoS Neglected Tropical Diseases*, 14(12), e0008962. <https://doi.org/10.1371/journal.pntd.0008962>
- Castro, M. C., Baeza, A., Codeço, C. T., Cucunubá, Z. M., Dal'Asta, A. P., De Leo, G. A., Dobson, A. P., Carrasco-Escobar, G., Lana, R. M., Lowe, R., Monteiro, A. M. V., Pascual, M., & Santos-Vega, M. (2019). Development, environmental degradation, and disease spread in the Brazilian Amazon. *PLoS Biology*, 17(11), e3000526. <https://doi.org/10.1371/journal.pbio.3000526>
- Chan, E. R., Menard, D., David, P. H., Ratsimbao, A., Kim, S., Chim, P., Do, C., Witkowski, B., Mercereau-Puijalon, O., Zimmerman, P. A., & Serre, D. (2012). Whole genome sequencing of field isolates provides robust characterization of genetic diversity in *Plasmodium vivax*. *PLoS Neglected Tropical Diseases*, 6(9), e1811. <https://doi.org/10.1371/journal.pntd.0001811>
- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, 4(1), 7. <https://doi.org/10.1186/s13742-015-0047-8>
- Chen, S.-B., Wang, Y., Kassegne, K., Xu, B., Shen, H.-M., & Chen, J.-H. (2017). Whole-genome sequencing of a *Plasmodium vivax* clinical isolate exhibits geographical characteristics and high genetic variation in China-Myanmar border area. *BMC Genomics*, 18(1), 131. <https://doi.org/10.1186/s12864-017-3523-y>
- Cholewiński, M., Derda, M., & Hadaś, E. (2015). Review articles parasitic diseases in humans transmitted by vectors. *Polish Parasitological Society*, 61(3), 137–157. <https://doi.org/10.17420/ap6103.01>
- Cornejo, O. E., & Escalante, A. A. (2006). The origin and age of *Plasmodium vivax*. *Trends in Parasitology*, 22(12), 558–563. <https://doi.org/10.1016/j.pt.2006.09.007>
- Cornejo, O. E., Fisher, D., & Escalante, A. A. (2015). Genome-wide patterns of genetic polymorphism and signatures of selection in *Plasmodium vivax*. *Genome Biology and Evolution*, 7(1), 106–119. <https://doi.org/10.1093/gbe/evu267>
- Cotton, J. A., Berriman, M., Dalén, L., & Barnes, I. (2018). Eradication genomics—Lessons for parasite control. *Science*, 361(6398), 130–131. [https://doi.org/10.1126/SCIENCE.AAR6609/ASSET/A7718BFD-FD13-4587-830C-40E76283F09D/ASSETS/GRAPHIC/361\\_130\\_F1.JPEG](https://doi.org/10.1126/SCIENCE.AAR6609/ASSET/A7718BFD-FD13-4587-830C-40E76283F09D/ASSETS/GRAPHIC/361_130_F1.JPEG)
- Cowell, A. N., Valdivia, H. O., Bishop, D. K., & Winzeler, E. A. (2018). Exploration of *Plasmodium vivax* transmission dynamics and recurrent infections in the Peruvian Amazon using whole genome sequencing. *Genome Medicine*, 10(1), 52. <https://doi.org/10.1186/s13073-018-0563-0>
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>
- Daron, J., Boissière, A., Boundenga, L., Ngoubangoye, B., Houze, S., Arnathau, C., Sidobre, C., Trape, J. F., Durand, P., Renaud, F., Fontaine, M. C., Prugnolle, F., & Rougeron, V. (2021). Population genomic evidence of *Plasmodium vivax* southeast Asian origin. *Science Advances*, 7(18), 3713–3741. [https://doi.org/10.1126/SCIADV.ABC3713/SUPPL\\_FILE/ABC3713\\_TABLE\\_S2.XLSX](https://doi.org/10.1126/SCIADV.ABC3713/SUPPL_FILE/ABC3713_TABLE_S2.XLSX)
- Daubersies, P., Ollomo, B., Sauzet, J.-P., Brahimi, K., Perlaza, B.-L., Eling, W., Moukane, H., Rouquet, P., de Taisne, C., & Druilhe, P. (2008). Genetic immunisation by liver stage antigen 3 protects chimpanzees against malaria despite low immune responses. *PLoS One*, 3(7), e2659. <https://doi.org/10.1371/journal.pone.0002659>
- de Castro, M. C., Monte-Mór, R. L., Sawyer, D. O., & Singer, B. H. (2006). Malaria risk on the Amazon frontier. *Proceedings of the National Academy of Sciences of the United States of America*, 103(7), 2452–2457. <https://doi.org/10.1073/pnas.0510576103>
- De Meulenaere, K., Cuypers, B., Gamboa, D., Laukens, K., & Rosanas-Urgell, A. (2023). A new *Plasmodium vivax* reference genome for south American isolates. *BMC Genomics*, 24(1), 1–14. <https://doi.org/10.1186/S12864-023-09707-5>
- De Meulenaere, K., Prajapati, S. K., Villasis, E., Cuypers, B., Kattenberg, J. H., Kasian, B., Laman, M., Robinson, L. J., Gamboa, D., Laukens, K., & Rosanas-Urgell, A. (2022). Band 3-mediated *Plasmodium vivax* invasion is associated with transcriptional variation in PvTRAG genes.

- Frontiers in Cellular and Infection Microbiology*, 12, 1011692. <https://doi.org/10.3389/fcimb.2022.1011692>
- de Oliveira, T. C., Corder, R. M., Early, A., Rodrigues, P. T., Ladeia-Andrade, S., Alves, J. M. P., Neafsey, D. E., & Ferreira, M. U. (2020). Population genomics reveals the expansion of highly inbred *Plasmodium vivax* lineages in the main malaria hotspot of Brazil. *PLoS Neglected Tropical Diseases*, 14(10), e0008808. <https://doi.org/10.1371/journal.pntd.0008808>
- de Oliveira, T. C., Rodrigues, P. T., Menezes, M. J., Gonçalves-Lopes, R. M., Bastos, M. S., Lima, N. F., Barbosa, S., Gerber, A. L., Loss de Moraes, G., Berná, L., Phelan, J., Robello, C., de Vasconcelos, A. T. R., Alves, J. M. P., & Ferreira, M. U. (2017). Genome-wide diversity and differentiation in new world populations of the human malaria parasite *Plasmodium vivax*. *PLoS Neglected Tropical Diseases*, 11(7), e0005824. <https://doi.org/10.1371/journal.pntd.0005824>
- Delgado-Ratto, C., Gamboa, D., Soto-Calle, V. E., van den Eede, P., Torres, E., Sánchez-Martínez, L., Contreras-Mancilla, J., Rosanas-Urgell, A., Rodríguez Ferrucci, H., Llanos-Cuentas, A., Erhart, A., Van Geertruyden, J.-P., & D'Alessandro, U. (2016). Population genetics of *Plasmodium vivax* in the Peruvian Amazon. *PLoS Neglected Tropical Diseases*, 10(1), e0004376. <https://doi.org/10.1371/journal.pntd.0004376>
- Dharia, N. V., Bright, A. T., Westenberger, S. J., Barnes, S. W., Batalov, S., Kuhen, K., Borboa, R., Federe, G. C., McClean, C. M., Vinetz, J. M., Neyra, V., Llanos-Cuentas, A., Barnwell, J. W., Walker, J. R., & Winzeler, E. A. (2010). Whole-genome sequencing and microarray analysis of ex vivo *Plasmodium vivax* reveal selective pressure on putative drug resistance genes. *Proceedings of the National Academy of Sciences*, 107(46), 20045–20050. <https://doi.org/10.1073/pnas.1003776107>
- Drummond, A. J., Pybus, O. G., Rambaut, A., Forsberg, R., & Rodrigo, A. G. (2003). Measurably evolving populations. *Trends in Ecology & Evolution*, 18(9), 481–488. [https://doi.org/10.1016/S0169-5347\(03\)00216-7](https://doi.org/10.1016/S0169-5347(03)00216-7)
- Duchêne, S., Holt, K. E., Weill, F.-X., Le Hello, S., Hawkey, J., Edwards, D. J., Fourment, M., & Holmes, E. C. (2016). Genome-scale rates of evolutionary change in bacteria. *Microbial Genomics*, 2(11), e000094. <https://doi.org/10.1099/mgen.0.000094>
- Feachem, R. G. A., Phillips, A. A., Hwang, J., Cotter, C., Wielgosz, B., Greenwood, B. M., Sabot, O., Rodriguez, M. H., Abeyasinghe, R. R., Ghebreyesus, T. A., & Snow, R. W. (2010). Shrinking the malaria map: Progress and prospects. *The Lancet*, 376(9752), 1566–1578. [https://doi.org/10.1016/S0140-6736\(10\)61270-6](https://doi.org/10.1016/S0140-6736(10)61270-6)
- Ferreira, M. U., Corder, R. M., Johansen, I. C., Kattenberg, J. H., Moreno, M., Rosas-Aguirre, A., Ladeia-Andrade, S., Conn, J. E., Llanos-Cuentas, A., Gamboa, D., Rosanas-Urgell, A., & Vinetz, J. M. (2022). Relative contribution of low-density and asymptomatic infections to *Plasmodium vivax* transmission in the Amazon: Pooled analysis of individual participant data from population-based cross-sectional surveys. *The Lancet Regional Health – Americas*, 9, 100169. <https://doi.org/10.1016/j.lana.2021.100169>
- Flannery, E. L., Wang, T., Akbari, A., Corey, V. C., Gunawan, F., Bright, A. T., Abraham, M., Sanchez, J. F., Santolalla, M. L., Baldeviano, G. C., Edgel, K. A., Rosales, L. A., Lescano, A. G., Bafna, V., Vinetz, J. M., & Winzeler, E. A. (2015). Next-generation sequencing of *Plasmodium vivax* patient samples shows evidence of direct evolution in drug-resistance genes. *ACS Infectious Diseases*, 1(8), 367–379. <https://doi.org/10.1021/acsinfecdis.5b00049>
- Gardy, J. L., & Loman, N. J. (2018). Towards a genomics-informed, real-time, global pathogen surveillance system. *Nature Reviews Genetics*, 19(1), 9–20. <https://doi.org/10.1038/nrg.2017.88>
- GBD 2019 Child and Adolescent Communicable Disease Collaborators. (2023). The unfinished agenda of communicable diseases among children and adolescents before the COVID-19 pandemic, 1990–2019: A systematic analysis of the global burden of disease study 2019. *The Lancet*, 402(10398), 313–335. [https://doi.org/10.1016/S0140-6736\(23\)00860-7](https://doi.org/10.1016/S0140-6736(23)00860-7)
- Ghosn, S., Chamat, S., Prieur, E., Stephan, A., Druilhe, P., & Bouharoun-Tayoun, H. (2018). Evaluating human immune responses for vaccine development in a novel human spleen cell-engrafted NOD-SCID-IL2r<sup>γ</sup>Null mouse model. *Frontiers in Immunology*, 9, 601. <https://doi.org/10.3389/fimmu.2018.00601>
- Grad, Y. H., & Lipsitch, M. (2014). Epidemiologic data and pathogen genome sequences: A powerful synergy for public health. *Genome Biology*, 15(11), 538. <https://doi.org/10.1186/s13059-014-0538-4>
- Grenfell, B. T., Pybus, O. G., Gog, J. R., Wood, J. L. N., Daly, J. M., Mumford, J. A., & Holmes, E. C. (2004). Unifying the epidemiological and evolutionary dynamics of pathogens. *Science*, 303(5656), 327–332. <https://doi.org/10.1126/science.1090727>
- Gruenberg, M., Moniz, C. A., Hofmann, N. E., Wampfler, R., Koepfli, C., Mueller, I., Monteiro, W. M., Lacerda, M., de Melo, G. C., Kuehn, A., & Siqueira, A. M. (2018). *Plasmodium vivax* molecular diagnostics in community surveys: pitfalls and solutions. *Malaria Journal*, 17, 55. <https://doi.org/10.1186/s12936-018-2201-0>
- Hassan-Montero, Y., De-Moya-Anegón, F., & Guerrero-Bote, V. P. (2022). SCLmago Graphica: A new tool for exploring and visually communicating data. *Profesional de La Información*, 31(5), e310502. <https://doi.org/10.3145/epi.2022.sep.02>
- Henden, L., Lee, S., Mueller, I., Barry, A., & Bahlo, M. (2018). Identity-by-descent analyses for measuring population dynamics and selection in recombining pathogens. *PLoS Genetics*, 14(5), e1007279. <https://doi.org/10.1371/journal.pgen.1007279>
- Hester, J., Chan, E. R., Menard, D., Mercereau-Puijalon, O., Barnwell, J., Zimmerman, P. A., & Serre, D. (2013). De novo assembly of a field isolate genome reveals novel *Plasmodium vivax* erythrocyte invasion genes. *PLoS Neglected Tropical Diseases*, 7(12), e2569. <https://doi.org/10.1371/journal.pntd.0002569>
- Hupalo, D. N., Luo, Z., Melnikov, A., Sutton, P. L., Rogov, P., Escalante, A., Vallejo, A. F., Herrera, S., Arévalo-Herrera, M., Fan, Q., Wang, Y., Cui, L., Lucas, C. M., Durand, S., Sanchez, J. F., Christian Baldeviano, G., Lescano, A. G., Laman, M., Barnadas, C., ... Carlton, J. M. (2016). Population genomics studies identify signatures of global dispersal and drug resistance in *Plasmodium vivax*. *Nature Genetics*, 48(8), 953–958. <https://doi.org/10.1038/ng.3588>
- Ibrahim, A., Manko, E., Dombrowski, J. G., Campos, M., Benavente, E. D., Nolder, D., Sutherland, C. J., Nosten, F., Fernandez, D., Vélez-Tobón, G., Castaño, A. T., Aguiar, A. C. C., Pereira, D. B., da Silva Santos, S., Suarez-Mutis, M., Di Santi, S. M., de Souza, R., Baptista, A., Dantas Machado, R. L., ... Campino, S. (2023). Population-based genomic study of *Plasmodium vivax* malaria in seven Brazilian states and across South America. *The Lancet Regional Health – Americas*, 18, 100420. <https://doi.org/10.1016/j.lana.2022.100420>
- Kattenberg, J. H., Razook, Z., Keo, R., Koepfli, C., Jennison, C., Lautugumal, D., Fola, A. A., Ome-Kaius, M., Barnadas, C., Siba, P., Felger, I., Kazura, J., Mueller, I., Robinson, L. J., & Barry, A. E. (2020). Monitoring *Plasmodium falciparum* and *Plasmodium vivax* using microsatellite markers indicates limited changes in population structure after substantial transmission decline in Papua New Guinea. *Molecular Ecology*, 29(23), 4525–4541. <https://doi.org/10.1111/mec.15654>
- Kattenberg, J. H., Van Nguyen, H., Nguyen, H. L., Sauve, E., Nguyen, N. T. H., Chopop-Pizarro, A., Trimarsanto, H., Monsieurs, P., Guetens, P., Nguyen, X. X., Van Esbroeck, M., Auburn, S., Nguyen, B. T. H., & Rosanas-Urgell, A. (2022). Novel highly-multiplexed AmpliSeq targeted assay for *Plasmodium vivax* genetic surveillance use cases at multiple geographical scales. *Frontiers in Cellular and Infection Microbiology*, 12, 953187. <https://doi.org/10.3389/fcimb.2022.953187>
- Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAXML-NG: A fast, scalable and user-friendly tool for maximum

- likelihood phylogenetic inference. *Bioinformatics*, 35(21), 4453–4455. <https://doi.org/10.1093/bioinformatics/btz305>
- Kwiatkowski, D. (2015). Malaria genomics: Tracking a diverse and evolving parasite population. *International Health*, 7(2), 82–84. <https://doi.org/10.1093/INTHEALTH/IHV007>
- Lee, S., & Bahlo, M. (2016). *Moimix: An R package for assessing clonality in high-throughput sequencing data*. <https://doi.org/10.5281/ZENODO.58257>
- Macdonald, G. (1957). *The epidemiology and control of malaria*. Oxford University Press.
- Martinelli, A., Hunt, P., Fawcett, R., Cravo, P. V. L., Walliker, D., & Carter, R. (2005). An AFLP-based genetic linkage map of *Plasmodium chabaudi chabaudi*. *Malaria Journal*, 4, 11. <https://doi.org/10.1186/1475-2875-4-11>
- Minassian, A. M., Themistocleous, Y., Silk, S. E., Barrett, J. R., Kemp, A., Quinkert, D., Nielsen, C. M., Edwards, N. J., Rawlinson, T. A., Lopez, F. R., Roobsoong, W., Ellis, K. J. D., Cho, J. S., Aunin, E., Otto, T. D., Reid, A. J., Bach, F. A., Labbé, G. M. C., Poulton, I. D., ... Draper, S. J. (2021). Controlled human malaria infection with a clone of *Plasmodium vivax* with high-quality genome assembly. *JCI Insight*, 6(23), e152465. <https://doi.org/10.1172/JCI.INSIGHT.152465>
- Miotto, O., Amato, R., Ashley, E. A., Macinnis, B., Almagro-Garcia, J., Amaratunga, C., Lim, P., Mead, D., Oyola, S. O., Dhorda, M., Imwong, M., Woodrow, C., Manske, M., Stalker, J., Drury, E., Campino, S., Amenga-Etego, L., Thanh, T. N. N., Tran, H. T., ... Kwiatkowski, D. P. (2015). Genetic architecture of artemisinin-resistant *Plasmodium falciparum*. *Nature Genetics*, 47(3), 226–234. <https://doi.org/10.1038/ng.3189>
- Neafsey, D. E., Galinsky, K., Jiang, R. H., Young, L., Sykes, S. M., Saif, S., Gujja, S., Goldberg, J. M., Young, S., Zeng, Q., Chapman, S. B., Dash, A. P., Anvikar, A. R., Sutton, P. L., Birren, B. W., Escalante, A. A., Barnwell, J. W., & Carlton, J. M. (2012). The malaria parasite *Plasmodium vivax* exhibits greater genetic diversity than *Plasmodium falciparum*. *Nature Publishing Group*, 44(9), 1046–1050. <https://doi.org/10.1038/ng.2373>
- Neafsey, D. E., Taylor, A. R., & MacInnis, B. L. (2021). Advances and opportunities in malaria population genomics. *Nature Reviews Genetics*, 22(8), 502–517. <https://doi.org/10.1038/s41576-021-00349-5>
- Nkhoma, S. C., Trevino, S. G., Gorena, K. M., Nair, S., Khoswe, S., Jett, C., Garcia, R., Daniel, B., Dia, A., Terlouw, D. J., Ward, S. A., Anderson, T. J. C., & Cheeseman, I. H. (2020). Co-transmission of related malaria parasite lineages shapes within-host parasite diversity. *Cell Host & Microbe*, 27(1), 93–103. <https://doi.org/10.1016/j.chom.2019.12.001>
- Noviyanti, R., Coutrier, F., Utami, R. A. S., Trimarsanto, H., Tirta, Y. K., Trianty, L., Kusuma, A., Sutanto, I., Kosasih, A., Kusriastuti, R., Hawley, W. A., Laihad, F., Lobo, N., Marfurt, J., Clark, T. G., Price, R. N., & Auburn, S. (2015). Contrasting transmission dynamics of co-endemic *Plasmodium vivax* and *P. falciparum*: Implications for malaria control and elimination. *PLoS Neglected Tropical Diseases*, 9(5), e0003739. <https://doi.org/10.1371/journal.pntd.0003739>
- Ocholla, H., Preston, M. D., Mipando, M., Jensen, A. T. R., Campino, S., MacInnis, B., Alcock, D., Terlouw, A., Zongo, I., Oudraogo, J.-B., Djimde, A. A., Assefa, S., Doumbo, O. K., Borrmann, S., Nzila, A., Marsh, K., Fairhurst, R. M., Nosten, F., Anderson, T. J. C., ... Montgomery, J. (2014). Whole-genome scans provide evidence of adaptive evolution in Malawian *Plasmodium falciparum* isolates. *The Journal of Infectious Diseases*, 210(12), 1991–2000. <https://doi.org/10.1093/infdis/jiu349>
- Olliaro, P. L., Barnwell, J. W., Barry, A., Mendis, K., Mueller, I., Reeder, J. C., Shanks, G. D., Snounou, G., & Wongsrichanalai, C. (2016). Implications of *Plasmodium vivax* biology for control, elimination, and research. *The American Journal of Tropical Medicine and Hygiene*, 95(6 Suppl), 4–14. <https://doi.org/10.4269/AJTMH.16-0160>
- Ortiz, E. M. (2019). *vcf2phylip v2.0: Convert a VCF matrix into several matrix formats for phylogenetic analysis*. <https://doi.org/10.5281/ZENODO.2540861>
- Oyebola, K. M., Idowu, E. T., Olukosi, Y. A., Awolola, T. S., & Amambua-Ngwa, A. (2017). Pooled-DNA sequencing identifies genomic regions of selection in Nigerian isolates of *Plasmodium falciparum*. *Parasites & Vectors*, 10(1), 320. <https://doi.org/10.1186/s13071-017-2260-z>
- Pacheco, M. A., Schneider, K. A., Céspedes, N., Herrera, S., Arévalo-Herrera, M., & Escalante, A. A. (2019). Limited differentiation among *Plasmodium vivax* populations from the northwest and to the south Pacific coast of Colombia: A malaria corridor? *PLoS Neglected Tropical Diseases*, 13(3), e0007310. <https://doi.org/10.1371/journal.pntd.0007310>
- Parobek, C. M., Lin, J. T., Saunders, D. L., Barnett, E. J., Lon, C., Lanteri, C. A., Balasubramanian, S., Brazeau, N., DeConti, D. K., Garba, D. L., Meshnick, S. R., Spring, M. D., Chuor, C. M., Bailey, J. A., & Juliano, J. J. (2016). Selective sweep suggests transcriptional regulation may underlie *Plasmodium vivax* resilience to malaria control measures in Cambodia. *Proceedings of the National Academy of Sciences of the United States of America*, 113(50), E8096–E8105. <https://doi.org/10.1073/pnas.1608828113>
- Pearce, E. J., & Tarleton, R. L. (2014). Overview of the parasitic pathogens. In S. H. E. Kaufmann, A. Sher, & R. Ahmed (Eds.), *Immunology and infectious diseases* (pp. 39–52). American Society for Microbiology Press.
- Pearson, R. D., Amato, R., Auburn, S., Miotto, O., Almagro-Garcia, J., Amaratunga, C., Suon, S., Mao, S., Noviyanti, R., Trimarsanto, H., Marfurt, J., Anstey, N. M., William, T., Boni, M. F., Dolecek, C., Tran, H. T., White, N. J., Michon, P., Siba, P., ... Kwiatkowski, D. P. (2016). Genomic analysis of local variation and recent evolution in *Plasmodium vivax*. *Nature Genetics*, 48(8), 959–964. <https://doi.org/10.1038/ng.3599>
- Perlaza, B.-L., Valencia, A. Z., Zapata, C., Castellanos, A., Sauzet, J.-P., Blanc, C., Cohen, J., Arévalo-Herrera, M., Corradin, G., Herrera, S., & Druilhe, P. (2008). Protection against *Plasmodium falciparum* challenge induced in Aotus monkeys by liver-stage antigen-3-derived long synthetic peptides. *European Journal of Immunology*, 38(9), 2610–2615. <https://doi.org/10.1002/eji.200738055>
- Popovici, J., Friedrich, L. R., Kim, S., Bin, S., Run, V., Lek, D., Cannon, M. V., Menard, D., & Serre, D. (2018). Genomic analyses reveal the common occurrence and complexity of *Plasmodium vivax* relapses in Cambodia. *MBio*, 9(1), e01888-17. <https://doi.org/10.1128/mBio.01888-17>
- Price, R. N., Commons, R. J., Battle, K. E., Thriemer, K., & Mendis, K. (2020). *Plasmodium vivax* in the era of the shrinking P. Falciparum map. *Trends in Parasitology*, 36(6), 560–570. <https://doi.org/10.1016/j.pt.2020.03.009>
- Prugnolle, F., & De Meeûs, T. (2008). The impact of clonality on parasite population genetic structure. *Parasite*, 15(3), 455–457. <https://doi.org/10.1051/parasite/2008153p455>
- Recht, J., Siqueira, A. M., Monteiro, W. M., Herrera, S. M., Herrera, S., & Lacerda, M. V. G. (2017). Malaria in Brazil, Colombia, Peru and Venezuela: Current challenges in malaria control and elimination. *Malaria Journal*, 16(1), 273. <https://doi.org/10.1186/s12936-017-1925-6>
- Rodrigues, P. T., Valdivia, H. O., de Oliveira, T. C., Alves, J. M. P., Duarte, A. M. R. C., Cerutti-Junior, C., Buery, J. C., Brito, C. F. A., de Souza, J. C., Hirano, Z. M. B., Bueno, M. G., Catão-Dias, J. L., Malafronte, R. S., Ladeia-Andrade, S., Mita, T., Santamaria, A. M., Calzada, J. E., Tantular, I. S., Kawamoto, F., ... Ferreira, M. U. (2018). Human migration and the spread of malaria parasites to the New World. *Scientific Reports*, 8(1), 1993. <https://doi.org/10.1038/s41598-018-19554-0>
- Rodríguez, J. C. P., Uribe, G. Á., Araújo, R. M., Narváez, P. C., & Valencia, S. H. (2011). Epidemiology and control of malaria in Colombia. *Memórias do Instituto Oswaldo Cruz*, 106(suppl 1), 114–122. <https://doi.org/10.1590/S0074-02762011000900015>
- Rosas-Aguirre, A., Gamboa, D., Manrique, P., Conn, J. E., Moreno, M., Lescano, A. G., Sanchez, J. F., Rodriguez, H., Silva, H., Llanos-Cuentas,



- A., & Vinetz, J. M. (2016). Epidemiology of *Plasmodium vivax* malaria in Peru. *The American Journal of Tropical Medicine and Hygiene*, 95(6 Suppl), 133–144. <https://doi.org/10.4269/ajtmh.16-0268>
- Rougeron, V., Daron, J., Fontaine, M. C., & Prugnolle, F. (2022). Evolutionary history of *Plasmodium vivax* and *Plasmodium simium* in the Americas. *Malaria Journal*, 21(1), 141. <https://doi.org/10.1186/s12936-022-04132-7>
- Rovira-Vallbona, E., Van Hong, N., Kattenberg, J. H., Mah Huan, R., Thi Huong Binh, N., Thi Hong Ngo, N., Guetens, P., Luong Hieu, N., Thi, N., Hien, T., Thi Sang, V., Duc Long, N., Sauve, E., Thanh Duong, T., Xuan Xa, N., Erhart, A., & Rosanas-Urgell, A. (2021). High proportion of genome-wide homology and increased pre-treatment pvcrt levels in *Plasmodium vivax* late recurrences: A chloroquine therapeutic efficacy study. *Antimicrobial Agents and Chemotherapy*, 65(8), e00095–21. <https://doi.org/10.1128/AAC.00095-21>
- Sattabongkot, J., Tsuboi, T., Zollner, G. E., Sirichaisinthop, J., & Cui, L. (2004). *Plasmodium vivax* transmission: Chances for control? *Trends in Parasitology*, 20(4), 192–198. <https://doi.org/10.1016/j.pt.2004.02.001>
- Siegel, S. V., & Rayner, J. C. (2020). Single cell sequencing shines a light on malaria parasite relatedness in complex infections. *Trends in Parasitology*, 36(2), 83–85. <https://doi.org/10.1016/j.pt.2019.12.007>
- Siqueira, A. M., Mesones-Lapouble, O., Marchesini, P., Sampaio, V. S., Brasil, P., Tauil, P. L., Fontes, C. J., Costa, F. T. M., Daniel-Ribeiro, C. T., Lacerda, M. V. G., Damasceno, C. P., & Santelli, A. C. S. (2016). *Plasmodium vivax* landscape in Brazil: Scenario and challenges. *The American Journal of Tropical Medicine and Hygiene*, 95(6 Suppl), 87–96. <https://doi.org/10.4269/ajtmh.16-0204>
- Steverding, D. (2020). The spreading of parasites by human migratory activities. *Virulence*, 11(1), 1177–1191. <https://doi.org/10.1080/21505594.2020.1809963>
- Taylor, A. R., Jacob, P. E., Neafsey, D. E., & Buckee, C. O. (2019). Estimating relatedness between malaria parasites. *Genetics*, 212(4), 1337–1351. <https://doi.org/10.1534/genetics.119.302120>
- Taylor, J. E., Pacheco, M. A., Bacon, D. J., Beg, M. A., MacHado, R. L., Fairhurst, R. M., Herrera, S., Kim, J. Y., Menard, D., Póvoa, M. M., Villegas, L., Mulyanto, Snounou, G., Cui, L., Zeyrek, F. Y., & Escalante, A. A. (2013). The evolutionary history of *Plasmodium vivax* as inferred from mitochondrial genomes: Parasite genetic diversity in the Americas. *Molecular Biology and Evolution*, 30(9), 2050–2064. <https://doi.org/10.1093/MOLBEV/MST104>
- Torres, K., Ferreira, M. U., Castro, M. C., Escalante, A. A., Conn, J. E., Villasis, E., da Silva Araujo, M., Almeida, G., Rodrigues, P. T., Corder, R. M., Fernandes, A. R. J., Calil, P. R., Ladeia, W. A., Garcia-Castillo, S. S., Gomez, J., do Valle Antonelli, L. R., Gazzinelli, R. T., Golenbock, D. T., Llanos-Cuentas, A., ... Vinetz, J. M. (2022). Malaria resilience in South America: Epidemiology, vector biology, and immunology insights from the Amazonian International Center of Excellence in malaria research network in Peru and Brazil. *The American Journal of Tropical Medicine and Hygiene*, 107(4\_Suppl), 168–181. <https://doi.org/10.4269/ajtmh.22-0127>
- Van den Eede, P., Van der Auwera, G., Delgado, C., Huyse, T., Soto-Calle, V. E., Gamboa, D., Grande, T., Rodriguez, H., Llanos, A., Anné, J., Erhart, A., & D'Alessandro, U. (2010). Multilocus genotyping reveals high heterogeneity and strong local population structure of the *Plasmodium vivax* population in the Peruvian Amazon. *Malaria Journal*, 9(1), 151. <https://doi.org/10.1186/1475-2875-9-151>
- van Dorp, L., Gelabert, P., Rieux, A., de Manuel, M., de-Dios, T., Gopalakrishnan, S., Carøe, C., Sandoval-Velasco, M., Fregel, R., Olalde, I., Escosa, R., Aranda, C., Huijben, S., Mueller, I., Marquès-Bonet, T., Balloux, F., Gilbert, M. T. P., & Lalueza-Fox, C. (2020). *Plasmodium vivax* malaria viewed through the lens of an eradicated European strain. *Molecular Biology and Evolution*, 37(3), 773–785. <https://doi.org/10.1093/molbev/msz264>
- Verity, R., Aydemir, O., Brazeau, N. F., Watson, O. J., Hathaway, N. J., Mwandagaliwa, M. K., Marsh, P. W., Thwai, K., Fulton, T., Denton, M., Morgan, A. P., Parr, J. B., Tumwebaze, P. K., Conrad, M., Rosenthal, P. J., Ishengoma, D. S., Ngondi, J., Gutman, J., Mulenga, M., ... Juliano, J. J. (2020). The impact of antimalarial resistance on the genetic structure of *Plasmodium falciparum* in the DRC. *Nature Communications*, 11(1), 2107. <https://doi.org/10.1038/s41467-020-15779-8>
- Volkman, S. K., Sabeti, P. C., DeCaprio, D., Neafsey, D. E., Schaffner, S. F., Milner, D. A., Daily, J. P., Sarr, O., Ndiaye, D., Ndir, O., Mboup, S., Duraisingh, M. T., Lukens, A., Derr, A., Stange-Thomann, N., Waggoner, S., Onofrio, R., Ziaugra, L., Mauceli, E., ... Wirth, D. F. (2007). A genome-wide map of diversity in *Plasmodium falciparum*. *Nature Genetics*, 39(1), 113–119. <https://doi.org/10.1038/ng1930>
- Waltmann, A., Koepfli, C., Tessier, N., Karl, S., Fola, A., Darcy, A. W., Wini, L., Harrison, G. L. A., Barnadas, C., Jennison, C., Karunajeewa, H., Boyd, S., Whittaker, M., Kazura, J., Bahlo, M., Mueller, I., & Barry, A. E. (2018). Increasingly inbred and fragmented populations of *Plasmodium vivax* associated with the eastward decline in malaria transmission across the Southwest Pacific. *PLoS Neglected Tropical Diseases*, 12(1), e0006146. <https://doi.org/10.1371/journal.pntd.0006146>
- Wesolowski, A., Taylor, A. R., Chang, H.-H., Verity, R., Tessema, S., Bailey, J. A., Alex Perkins, T., Neafsey, D. E., Greenhouse, B., & Buckee, C. O. (2018). Mapping malaria by combining parasite genomic and epidemiologic data. *BMC Medicine*, 16(1), 190. <https://doi.org/10.1186/s12916-018-1181-9>
- White, M. T., Karl, S., Battle, K. E., Hay, S. I., Mueller, I., & Ghani, A. C. (2014). Modelling the contribution of the hypnozoite reservoir to *Plasmodium vivax* transmission. *eLife*, 3, e04692. <https://doi.org/10.7554/eLife.04692>
- World Health Organization. (2022a). *Global genomic surveillance strategy for pathogens with pandemic and epidemic potential 2022–2032*. World Health Organization.
- World Health Organization. (2022b). *World malaria report 2022*. World Health Organization.
- Yalcindag, E., Elguero, E., Arnathau, C., Durand, P., Akiana, J., Anderson, T. J., Aubouy, A., Balloux, F., Besnard, P., Bogreau, H., Carnevale, P., D'Alessandro, U., Fontenille, D., Gamboa, D., Jombart, T., Le Mire, J., Leroy, E., Maestre, A., Mayxay, M., ... Prugnolle, F. (2012). Multiple independent introductions of *Plasmodium falciparum* in South America. *Proceedings of the National Academy of Sciences of the United States of America*, 109(2), 511–516. [https://doi.org/10.1073/PNAS.1119058109/SUPPL\\_FILE/SAPP.PDF](https://doi.org/10.1073/PNAS.1119058109/SUPPL_FILE/SAPP.PDF)

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Kattenberg, J. H., Monsieurs, P., De Meyer, J., De Meulenaere, K., Sauve, E., de Oliveira, T. C., Ferreira, M. U., Gamboa, D., & Rosanas-Urgell, A. (2024). Population genomic evidence of structured and connected *Plasmodium vivax* populations under host selection in Latin America. *Ecology and Evolution*, 14, e11103. <https://doi.org/10.1002/ece3.11103>