This is a preproof accepted article for *Parasitology*
This version may be subject to change during the production process
DOI: 10.1017/S0031182025100310

Plasmodium simium: birth and evolution of a zoonotic malaria parasite species

Nathalia Rammé M. de Albuquerque[1,2], Winni A. Ladeia[1], Ryan J. Scalsky[2], Ankit Dwivedi[2], Thomas C. Stabler[2,3,4], Priscila T. Rodrigues,[1,5] Thaís C. de Oliveira,[1,6] Joana C. Silva[2,7,8*] and Marcelo U. Ferreira[1,8*]

[1]Department of Parasitology, Institute of Biomedical Sciences, University of São Paulo, São Paulo, SP, Brazil

[2]Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD, USA

[3] Department of Medical Parasitology and Infection Biology, Swiss Tropical and Public Health Institute, Basel, Switzerland

[4]University of Basel, Basel, Switzerland

[5]Brazilian Center for Research in Energy and Materials, Campinas, SP, Brazil

[6]Department of Translational Medicine, School of Medical Sciences, University of Campinas, Campinas, SP, Brazil.

[7]Department of Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, MD, USA

[8]Global Health and Tropical Medicine (GHTM), Associate Laboratory in Translation and Innovation Towards Global Health (LA-REAL), Institute of Hygiene and Tropical Medicine, NOVA University of Lisbon, Lisbon, Portugal

*J.C.S. and M.U.F. are co-senior authors and contributed equally to this article.

Corresponding author: Nathalia R. Medeiros de Albuquerque, Joana Carneiro da Silva, or Marcelo U. Ferreira

Email: nathalia.ramme@gmail.com; jcsilva@som.umaryland.edu; muferrei@usp.br

Abstract

Plasmodium simium, a parasite of platyrrhine monkeys, is known to cause human malaria outbreaks in Southeast Brazil. It has been hypothesized that, upon the introduction of *Plasmodium vivax* into the Americas at the time of the European colonization, the human parasite adapted to Neotropical anophelines of the *Kerteszia* subgenus and to local monkeys, along the Atlantic coast of Brazil, to give rise to a sister species, *P. simium.* Here, to obtain new insights into the origins and adaptation of *P. simium* to new hosts, we analysed whole-genome sequence data from 31 *P. simium* isolates together with a global sequence dataset of 1,086 *P. vivax* isolates. Population genomic analyses revealed that *P. simium* comprises a discrete parasite lineage with greatest genetic similarity to *P. vivax* populations from Latin America – especially those from the Amazon Basin of Brazil – and to ancient European *P. vivax* isolates, consistent with Brazil as the most likely birthplace of the species. We show that *P. simium* displays half the amount of nucleotide diversity of *P. vivax* from Latin America, as expected from its recent origin. We identified pairs of sympatric *P. simium* isolates from monkeys and from humans as closely related as meiotic half-siblings, revealing ongoing zoonotic transmission of *P. simium*. Most critically, we show that *P. simium* currently causes most, and possibly all, malarial infections usually attributed to *P. vivax* along the Serra do Mar Mountain Range of Southeast Brazil.

Keywords: malaria, *Plasmodium simium*; *Plasmodium vivax;* population genomics; zoonosis; Atlantic Forest; Brazil.

Introduction

Despite the progress toward malaria elimination from Latin America over the past two decades (Ferreira and Castro, 2019), there were 548,000 clinical cases recorded across the region in 2023, with an estimated 139 million people currently at risk of infection (World Health Organization, 2024). Three fourths of malaria cases in the Americas are due to *Plasmodium vivax* and the remainder to *Plasmodium falciparum*; only 0.1% of the cases are caused by *Plasmodium malariae*. The Amazon Basin accounts for nearly 90% of the regional malaria burden, and Brazil contributed 29.8% of the laboratory-confirmed cases recorded on this continent in 2024 (World Health Organization, 2024).

Modern humans were already settled in the Americas 21,000 years ago (Pigati *et al.*, 2023), but human malaria parasites are thought to have arrived much later (Bruce-Chwatt, 1965; Carter, 2003). Current evidence suggests that human malaria was introduced in the Americas by settlers from southern Europe – mainly Portugal and Spain, where the disease was endemic at the time of the European colonization of the American continent (Bruce-Chwatt and de Zuleta, 1980) – concurrently with the forced displacement of millions of enslaved people from West and Central Africa into South, Central and North America between the mid-1500s and mid-1800s (Bruce-Chwatt, 1965; Michel *et al.*, 2024). The arrival of human malaria parasites in the Americas offers a remarkable example of relatively recent host-parasite-vector co-evolution (Rougeron *et al.*, 2022). Parasites encountered new anopheline vector species in the New World – e.g., *Anopheles darlingi* across most of South America, but mainly in the Amazonian lowlands, and *Anopheles albimanus* along the Pacific Coast of South America and in Central America, both members of the subgenus *Nyssorhynchus*. These mosquito species are evolutionarily very distant from the dominant vectors in Africa (*Anopheles gambiae* complex, subgenus *Celia*) and southern Europe

(*Anopheles atroparvus*, subgenus *Anopheles*) to which they had been previously exposed (Molina-Cruz *et al*., 2016).

Once in the Americas, *P. vivax* also adapted to anophelines of the *Kerteszia* subgenus (*Anopheles cruzii, Anopheles bellator,* and *Anopheles homunculus*) (de Azevedo *et al*., 2020) and to new vertebrate hosts along the Atlantic coast of Brazil – namely, local platyrrhine primates such as howler monkeys, woolly spider monkeys, capuchin monkeys, uakaris, and titis (de Oliveira *et al*., 2021a; Duarte *et al*., 2021). It has been proposed that one or more spillover or host-shift events gave rise to the *P. vivax* sister species *Plasmodium simium*, comprising *P. vivax*-related parasites that infect monkeys and are also associated with clusters of human malaria cases in the outskirts of major cities in Southeast Brazil, including the metropolitan areas of Rio de Janeiro and São Paulo (e.g., Brasil *et al*., 2017; Duarte *et al*., 2021). Limited genome sequence data show that *P. simium* remains nearly identical to *P. vivax* (de Oliveira *et al*., 2021b; Mourier *et al*., 2021), consistent with a sympatric speciation event at its incipience, likely facilitated by a host switch (de Oliveira *et al*., 2021a). The process(es) that enable this switch to a new vertebrate host remain unknown. Importantly, while the high genetic similarity between *P. vivax* and *P. simium* may lead to misidentification of the causative agent of malaria infections, the rate of such misclassifications is unknown.

Here, an expanded dataset of whole genome sequences of *P. simium* is explored, including isolates from human malaria cases from Southeast Brazil originally attributed to *P. vivax*, and from global *P. vivax* isolates, to further investigate the origins of *P. simium* and genetic signatures of its adaptation to new hosts.

Materials and methods

Plasmodium simium *sequence data*

Whole-genome sequence (WGS) data from 31 *P. simium* isolates (de Oliveira *et al*., 2021b; Mourier *et al*., 2021; Ibrahim *et al*., 2023) were downloaded from the Short Read Archive (SRA) of the National Center for Biotechnology Information of USA (Supplementary Table 1, Supplementary Materials). The data set comprises six isolates from nonhuman primates – four from the brown howler monkey *Allouata clamitans* (three from the Cantareira Park, São Paulo State, and one from Guapimirim, Rio de Janeiro State), one from a black-fronted titi *Callicebus nigrifrons* from the Ecological Park of Tietê, São Paulo State, and one from a black uakari *Cacajao melanocephalus* from Guapimirim, Rio de Janeiro State. There were 25 samples of human origin, from Santa Maria de Jequitibá, Espírito Santo State (n = 7), various locations in Rio de Janeiro State (n = 6), and various locations in São Paulo State (n = 12). All sampling sites map to areas that are currently, or were formerly, covered by the Atlantic Forest in southeastern Brazil (Supplementary Figure 1). The 12 human-derived isolates from São Paulo State were originally described as part of a "highly clonal, potential *P. simium* cluster" (Ibrahim *et al*., 2023); all harbored one or both single nucleotide polymorphisms (SNPs) at the positions 3535 (T>C) and 3869 (A>G) of the mitochondrial genome that allows for the differentiation of *P. simium* from *P. vivax* (de Alvarenga *et al*., 2018).

Publicly available Plasmodium vivax *sequence data*

We analysed publicly available high-quality WGS data from 1,050 *P. vivax* isolates originating from four continents. To this end, we first downloaded sequences from 779 *P. vivax* isolates from the MalariaGEN Community Pv4 dataset (Adam *et al*., 2022 and references therein; Supplementary Table 2). Next, sequence data from 271 *P. vivax* isolates from the Americas were added to the dataset: Brazil (Hupalo *et al*., 2016; de Oliveira *et al*., 2017; de Oliveira *et al*., 2020; Benavente *et al*., 2021; Mourier *et al*., 2021; Ibrahim *et al*., 2023; De Meulenaere *et al.*, 2023; Kattenberg *et al.*, 2024), Panamá (Buyon *et al.*, 2020), and Peru (Kattenberg *et al.*, 2022; De Meulenaere *et al.*, 2023; Kattenberg *et al.*, 2024) (Supplementary Table 3, Supplementary Materials). WGS data from three European *P. vivax* strains, the isolate Ebro1944 from Spain (circa 1944; van Dorp *et al.*, 2020) and the isolates STR105 and STR185 from medieval/early modern Belgium (Michel *et al.*, 2024), were also used for specific analyses.

Plasmodium vivax *whole-genome sequence data generation*

We additionally generated WGS data from 36 new clinical isolates of *P. vivax* (Supplementary Table 3). One isolate was obtained in 2004 from a patient with imported *P. vivax* malaria diagnosed in the United States after a travel to India (Rodrigues *et al.*, 2014). DNA was isolated from 200 µL of unprocessed whole blood, using the QIAamp blood DNA kit (Qiagen, Hilden, Germany). Strand displacement amplification technology, with the primer set *pvset1* (Cowell *et al.*, 2017) and phi29 DNA polymerase (New England Biolabs, Ipswich, MA), was used to enrich the sample for target parasite DNA prior to sequencing. The remaining 35 isolates were derived from patients with microscopy-confirmed *P. vivax* infection presenting between 2018 and 2019 in malaria clinics in the town of Mâncio Lima (Rodrigues *et al.*, 2024) and the periurban settlement of Vila Assis Brasil (Fontoura *et al.*, 2024), approximately 15 km apart, in the Juruá Valley region of Acre State. This region, next to the border with Peru, was the main malaria transmission hotspot of Brazil in the early 2000s (Ferreira and Castro, 2016). QIAamp DNA investigator kits (Qiagen, Hilden, Germany) were used to isolate template DNA from 50 mL of venous blood that had previously been leukocyte-depleted, as described (de Oliveira *et al.* 2017). The presence of a single-species infection was confirmed by a species-specific TaqMan assay as described (Rodrigues *et al.*, 2024). No genome amplification step prior to sequencing was applied to these samples. Illumina UDI libraries (Illumina, San Diego, CA) were prepared to generate paired-end 150 base pair-long sequence reads, on an Illumina NovaSeq 6000 platform, at the Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, United States. Raw fastq files were filtered for quality and those with mean quality scores ≤ 30 (expected base call accuracy ≤ 99.9%) were excluded. Reads of new whole-genome sequence data described in this article have been deposited into the NCBI Sequence Read Archive (accession numbers in Supplementary Table 3).

Whole-genome sequence data processing and variant calling

Newly generated sequence data (from 36 *P. vivax* isolates) and those downloaded from SRA (from 1,051 *P. vivax* and 31 *P. simium* isolates) were processed similarly. Quality-filtered fastq files were mapped onto the 20.8-megabase (Mb) core PvP01 genome (Auburn *et al.*, 2016), defined as in Table S2 of Daron *et al*. (2021), with the Burrows-Wheeler aligner (Li, 2013). Use of the PvP01 reference genome allows for comparisons with a large body of published genome data (Adam *et al.*, 2022). Alignments with an average genome sequence depth <5× were excluded. The remaining alignments were merged into Binary Alignment/Map (BAM) files and processed following the Genome Analysis Toolkit (GATK version 4.4.0) best practices (https://software.broadinstitute.org/gatk/best-practices/; McKenna *et al.*, 2010). We used the GATK tool MarkDuplicates to identify and remove duplicated reads and BaseRecalibrator and ApplyBQSR tools to detect and correct errors in base quality scores. Variant calling was carried out using the HaplotypeCaller module of GATK in the GVCF mode for joint genotyping of multiple samples. Variants were removed according to the following GATK VariantFiltration criteria: read depth (DP) < 5, variant confidence/quality by depth (QD) < 2.0, strand bias (FS) > 60.0, root mean square of the mapping quality (MQ) < 40.0, mapping quality rank sum (MQRankSum) < -12.5, read position rank sum (ReadPosRankSum) < -8.0, quality (QUAL) < 30.0. Non-biallelic SNPs and those with a minor allele frequency <0.001 (singletons) were also removed.

Complexity of infection and genetic diversity

We used the within-host diversity statistic F_{WS}, calculated with the *R* package moimix (https://github.com/bahlolab/moimix), to distinguish between single- and multiple-clone infections. Multiple-clone infections, defined by $F_{WS} < 0.95$ (Manske *et al.*, 2012), were deconvoluted with DEploid (https://github.com/DEploid-dev/DEploid). Haploid genome data from single-clone infections and phased data from dominant clones within multiple-clone infections were used in all downstream analyses. VCFtools (Danecek *et al*., 2011) was used to estimate π, defined as the average number of pairwise differences per site between pairs of DNA sequences, as a measure of genome-wide nucleotide diversity in *P. simium* and *P. vivax* populations. The ratio of π values in *P. simium* to *P. vivax* was calculated for 3,615 orthologous gene pairs to identify genes with highest or lowest diversity in *P. simium* relative to *P. vivax*.

Clustering and differentiation among **Plasmodium simium** *and* **Plasmodium vivax** *genomes*

Principal components analysis (PCA), as implemented in PLINK version 1.90 (Purcell *et al.*, 2007), was used to determine clustering patterns among *P. simium* and *P. vivax* isolates based on genome-wide genotypes. Missing genotypes were imputed using Beagle version 5.4 (Browning *et al*., 2018). Variant sites in linkage disequilibrium (LD) (r^2 value >0.2) within 50-kb windows (step size of 10 base pairs, bp) were pruned to obtain a set of unlinked SNPs. Additional analyses run with the smartPCA software (Patterson *et al*., 2006) included low-coverage genomes from European *P. vivax* isolates. We used eigenvectors computed for high-quality WGS data and projected low-coverage data from European isolates onto the axes of variation, as described elsewhere (Michel *et al*., 2024).

ADMIXTURE version 1.3.0 (Alexander and Lange, 2011) was used to investigate the genetic ancestry of *P. simium* and Latin American *P. vivax* isolates. An unsupervised ADMIXTURE analysis was run to assign isolates to *K* putative ancestral populations according to SNP frequencies. The LD-pruned data set with imputed missing genotypes was used in the analysis. We calculated cross-validation error rates to determine the most likely number of ancestral populations, with *K* between 1 and 15 (Alexander *et al*., 2009). Wright's Fixation

Index, F_{ST} (Weir and Cockerham implementation in VCFtools), was used to determine differentiation between *P. simium* and Latin American populations of *P. vivax*. To further explore the genetic relationships between *P. simium* and present-day *P. vivax* populations from Latin America, we used the *R* package *admixr* (Petr, 2024) to compute the *f4* statistic, *f4* (W, X; Y, Z) (Patterson *et al*., 2012). We compared the proportion of derived alleles shared between *P. simium* (Y) and pairs of *P. vivax* populations (W and X), using *P. cynomolgi* strain B (RefSeq assembly: GCF_000321355.1; Tachibana *et al*., 2012) as the outgroup (Z). The *P. vivax* population from Mexico (X), the genetically most distant population from *P. simium* samples in PCA and F_{ST} analysis, was used in all pairwise comparisons.

Genetic relatedness and positive selection among P. simium *isolates*

We used identity-by-descent (IBD) analysis as implemented in hmmIBD (Schaffner *et al*., 2018) to measure the genetic relatedness between pairs of *P. simium* samples from monkeys and from humans, and between *P. simium* isolates from different states (São Paulo, Rio de Janeiro, and Espírito Santo). hmmIBD implements a hidden Markov model-based approach that accounts for recombination to systematically search for genomic segments that are inferred to have descended from a common ancestor without intervening recombination and to estimate the proportion of shared ancestry between genomes. A recombination rate of 13.5 kilobases (kb) per centiMorgan was assumed (Miles *et al*., 2016) and the -n option was set to 25, to call IBD segments from common ancestors within the last 25 generations (approximately 12.5 years, assuming 2 generations per year). Genetic relatedness networks were drawn to connect sample pairs with at least 25% (equivalent to half-siblings), 50% (equivalent to meiotic siblings) or 90% of genomes IBD. We estimated the distribution of IBD coverage (i.e., the number of IBD segments overlapping each position) and identified IBD peaks (i.e., chromosome regions with IBD sharing >2 standard deviations above the mean) across the *P. simium* genome using the ibdutils command line tool (Guo *et al*., 2024; https://github.com/bguo068/ishare).

Demographic history of Plasmodium simium *and* Plasmodium vivax *populations from Brazil*

We used pairwise sequentially Markovian coalescent (PSMC) analysis (Li and Durbin, 2011; Mather *et al*., 2019) to infer the historical effective population size (N_e) of *P. simium* and *P vivax* populations from Latin America. We ran PSMC analysis (https://github.com/lh3/psmc) assuming a mutation rate (μ) of 1 x 10^{-9} and a generation time (g) of 0.18 (Daron *et al*. 2021), with the five samples with the highest sequence coverage for each species. We performed 100 bootstrap replicates by randomly resampling the variants used in the analysis.

Results

Whole genome sequence data for Plasmodium simium *and* Plasmodium vivax

We analysed whole genome sequence data from 31 *P. simium* isolates, with an average of 1.1×10^6 sequence reads per sample, for an average depth of coverage of 49.7× (range between 3.4× and 233.0× among samples). We identified 55,682 high-confidence SNPs (Supplementary Table 4, Supplementary Materials). Only three (9.7%) samples contained two or more clones; those sequence data sets were deconvoluted, with the dominant genotype retained for further analysis. Although *P. simium* sequences had been generated from unprocessed blood (i.e., leukocytes had not been removed prior to DNA extraction), most (50.7%) reads mapped to the reference *P. vivax* genome. On average, 73.9% (range: 29.4% to 93.6%) of the PvP01 core genome was covered with a read depth ≥5×.

A total of 35 new *P. vivax* whole-genome sequence data sets from Brazil were generated in this study. We obtained an average of 10.6×10^6 reads per sample, resulting in an average read depth of 32.4× (range: 5.5× to 55.6×), and a total of 60,568 high-confidence SNPs were identified after filtering (Supplementary Table 5, Supplementary Materials). Eight (22.8%)

clinical samples contained two or more clones (F_{ws} ≤0.95) and the sequence data were deconvoluted. On average, 81.4% of the reads obtained from these leukocyte-depleted samples mapped to the reference PvP01 core genome, with an average breadth of coverage of 95.9% (range, 80.0% to 98.2%) of the reference genome mapped with a read depth ≥5×.

Plasmodium simium ***originated in South America and is closest to the*** **P. vivax** ***population from Brazil***

As in previous studies (Adam *et al*., 2022; Kattenberg *et al*., 2024; Michel *et al*., 2024), PCA revealed the distinct geographic structure of *P. vivax* populations from around the globe. The first two principal components captured one-third of the overall genetic variation and defined three main clusters: one with samples from Latin America, clearly separated from samples from Africa, South and West Asia and from a third cluster with samples from East and Southeast Asia and Oceania (Figure 1). Importantly, all *P. simium* samples clustered close to Latin American *P. vivax* populations in the PCA space. Historical *P. vivax* samples from Europe that were suitable for population genetic analysis – Ebro1944 from Spain (van Dorp *et al*., 2020) and STR105 and STR185 from Belgium (Michel *et al*., 2024) – clustered with Latin American *P. vivax*, to which they are more similar than to *P. simium* (Supplementary Figure 2). These findings are consistent with the origin of *P. simium* in Latin America.

A regional PCA was done to further investigate the origin of *P. simium*. All *P. simium* isolates clustered together in the regional PCA, regardless of their state of origin in Brazil (Figure 2). The *P. simium* cluster includes 12 human-derived samples from São Paulo that were originally labelled as *P. vivax* (Ibrahim *et al*., 2023). PCA also revealed a closer affinity of *P. simium* to lineages of *P. vivax* from the Amazon Basin of Brazil (states of Acre, Amapá, Amazonas, Pará, and Rondônia), compared to *P. vivax* populations from Peru and Colombia and those from Central America (Panama, Nicaragua, and El Salvador) and Mexico. Moreover, *P. vivax* isolates from Europe clustered together with the *P. vivax* populations from Brazil and all other Latin American countries, but not with *P. simium* (Supplementary Figure 3), reflecting the extent of divergence between present-day *P. simium* and the founding European lineages of *P. vivax* introduced in Brazil.

Unsupervised ADMIXTURE analysis was used to examine shared ancestry patterns of *P. simium* and *P. vivax* at the regional and country level. Cross-validation error decreased with increasing K until $K = 15$ (Supplementary Figure 4). We chose to display results of the analyses with the smallest K value that captures most of the geographic structure in the data ($K = 4$), the K value that maximizes clustering on a country level ($K = 10$), and the K value associated with the lowest cross-validation error ($K = 15$) (Figure 3). At $K = 4$, three regional *P. vivax* clusters can be seen: Brazil (apple green); Peru, Panama, and Nicaragua (light sea green); and Colombia, El Salvador, and Mexico (purple). At $K = 10$, most *P. vivax* isolates were assigned to country-specific populations but some samples from Brazil, Panama, and Peru appear to be admixed. Indeed, samples from Brazil appear to have ancestry in one of two different ancestral populations (blue and green), or both (admixed samples), while some samples from Peru appear to share ancestry with the green population from Brazil. At $K = 15$, the *P. simium* population appeared to comprise two distinct subpopulations, one from São Paulo and the other comprising isolates from Rio de Janeiro and Espírito Santo. Importantly, at all K values, *P. simium* isolates were assigned to separate population(s), with negligible admixture with regional or country-specific *P. vivax* populations (Figure 3). Moreover, the elevated estimates of population differentiation in pairwise comparisons of *P. simium* with *P. vivax* populations from Mexico ($F_{ST} = 0.29$), Panama ($F_{ST} = 0.22$), Colombia ($F_{ST} = 0.20$), Peru ($F_{ST} = 0.20$), and Brazil ($F_{ST} = 0.16$) indicate limited historical gene flow between species (Supplementary Figure 5, Supplementary Materials).

Previous analyses had suggested that *P. simium* was most closely related to present-day *P. vivax* samples from Mexico, compared to other locations in Latin America (de Oliveira *et* *al*., 2021b; Mourier *et al*., 2021). However, *f4* statistics with the expanded dataset comprising 31 *P. simium* samples from three states in southwest Brazil indicated that this species shares significantly more derived alleles with *P. vivax* samples from the Amazon Basin of Brazil, followed by Peru, Colombia, and Panama, compared to *P. vivax* samples from Mexico (Supplementary Figure 6A). These findings, consistent with our regional PCA (Figure 2) and F_{ST} results (Supplementary Figure 5), point to Brazil as the most likely birthplace of the novel parasite lineage that resulted from parasite jumps from humans to platyrrhine monkeys (de Oliveira *et al*., 2021b; Mourier *et al*., 2022).

Plasmodium simium ***has a decreasing effective population size, smaller than*** **Plasmodium vivax** ***in Latin America***

PSMC analysis revealed distinct historical trends for N_e estimates of *P. simium* and *P. vivax* populations (Figure 4). First, looking backwards in time, *P. vivax* has a much deeper genealogy than *P. simium*, consistent with the hypothesis that *P. simium* has a much more recent origin. Second, *P. vivax* shows a substantial decrease over time, followed by a recent sharp increase in N_e, suggestive of a sharp population expansion following a bottleneck. In contrast, the *P. simium* population shows a stable decrease in N_e over the entire period (Figure 4; see Supplementary Figure 7, for bootstrap replicates), consistent with a fairly small effective population size after the split from its common ancestor with *P. vivax*.

We next compared the overall nucleotide diversity of present-day *P. simium* (n = 31) with that of Latin American populations of *P. vivax* (n = 465). In line with previous studies (de Oliveira *et al*., 2021b; Mourier *et al*., 2021), we found a low diversity in the core genome of *P. simium* (mean $\pi = 2.5 \times 10^{-4}$), approximately 2 times lower than that of *P. vivax* populations from Latin America (mean $\pi = 5.0 \times 10^{-4}$, across loci) and 1.5—2.4 times lower than the diversity calculated for *P. vivax* populations from Brazil (mean $\pi = 5.0 \times 10^{-4}$), Colombia (mean $\pi = 5.9 \times 10^{-4}$), Mexico (mean $\pi = 5.8\ 0 \times 10^{-4}$), and Peru (mean π = 3.8 x

10^{-4}). The average π across loci among isolates from Panama (π = 2.4 x 10^{-4}) was similar to that for *P. simium*. Interestingly, nucleotide diversity estimates were nearly identical for *P. simium* samples from humans (mean π = 2.94 × 10^{4}; n = 25 samples) and platyrrhine monkeys (mean π = 2.90 × 10^{4}; n = 6 samples), as would be expected from repeat sampling from the same panmictic population.

Compared to a null distribution centered on 0, expected from identical nucleotide diversity per locus for both species, the density distribution of log ratios of π estimates is shifted to the left (Figure 5). This indicates less diversity in *P. simium* genes compared to their orthologs in *P. vivax*, with an average log ratio of -0.24 (standard deviation = 0.5). Interestingly, the right-hand tail of the distribution in Figure 5, which comprises genes with higher diversity in *P. simium*, includes loci that encode proteins involved in host-parasite interactions in *P. vivax* (Supplementary Table 6, Supplementary Materials). Among them are those encoding: the vacuolar protein sorting-associated protein 46 (PVP01_0704600), which may be linked to the production of extracellular vesicles in parasitized red blood cells, with a role in intercellular communication (Toda *et al*., 2020; Avalos-Padilla *et al*., 2021); the ookinete maturation gene OMG1 (PVP01_0609300), crucial for the invasive stage in the mosquito midgut in *P. berghei* (Nishi *et al*., 2022); the oocyst capsule protein Cap380, essential for oocyst development, sporozoite differentiation, and malaria transmission in *P. berghei* (Srinivasan *et al*., 2008; Nakayama *et al*., 2021); the liver merozoite formation protein (PVP01_1146600), vital for the maturation of liver merozoites in *P. berghei* (Haussig *et al*., 2011), which in *P. falciparum* appears to be essential for sporozoite formation within the oocyst (Franke-Fayard *et al*., 2022); and the AP2 domain transcription factor AP2-G5 (PVP01_0940100), crucial for gametocyte maturation in *P. falciparum* (Shang *et al*., 2021).

Contemporary **Plasmodium simium** *infects both monkeys and humans*

Plasmodium simium samples from humans from São Paulo, which were originally labelled as *P. vivax* (Ibrahim *et al.*, 2023), were found to share more derived alleles with *P. simium* from humans and from monkeys from Espírito Santo and Rio de Janeiro than with any of the *P. vivax* populations from Latin America (Supplementary Figure 6B). These findings further confirm that samples of *P. vivax*-related parasites infecting humans from São Paulo belong to the *P. simium* clade.

Parasite relatedness networks revealed some recent gene flow between *P. simium* populations across states. Although no sample pair from different states displayed ≥50% of the genome IBD (Supplementary Figure 8), examples of ≥25% IBD between pairs of isolates were found in a single cluster comprising six samples from São Paulo and three from Rio de Janeiro (Figure 6). We found only a very low proportion of IBD sharing between pairs of *P. simium* isolates originating from sites >450 km apart (Supplementary Figure 9). However, we found a pair of very closely related isolates, with ≥90% of the genome IBD, circulating within the state of São Paulo (Supplementary Figure 10), while there were no clonal lineages of *P. simium* in which isolates shared ≥99% of the genome IBD.

Notably, we found two clusters of parasites sharing ≥25% of the genome IBD that were derived from different mammalian hosts (Supplementary Figure 11). The larger cluster comprised five samples from monkeys (four from São Paulo and one from Rio de Janeiro) and four samples from humans (two from São Paulo and two from Rio de Janeiro), while the smaller cluster comprised one sample from a monkey and four from humans, all from Rio de Janeiro. These findings are consistent with recent gene flow between parasites from human and nonhuman hosts.

Positive selection in Plasmodium simium *reveals signals of adaptation to new vertebrate host and mosquito vectors*

We next searched for signatures of positive selection across the *P. simium* genome. Positive selection is expected to increase IBD sharing at the target locus and neighboring sites, generating peaks of within-population IBD sharing (Guo *et al.*, 2024). A genome-wide scan identified seven validated IBD peaks, possibly associated with selective sweeps during the parasite′s adaptation to new hosts (Figure 7). IBD peaks mapped to chromosomes 3, 9, 12 and 14 and comprise several annotated genes of potential interest (Supplementary Table 7, Supplementary Materials), including some transcription factors encoding an AP2 domain (DNA-binding) that may be involved in red blood cell invasion, gametocytogenesis, oocyst formation, and sporozoite formation (reviewed by Singhal *et al.*, 2024). For example, the chromosome 3 peak comprises genes that encode the AP2 domain transcription factor AP2-SP2 (PVP01_0303400), which is crucial for oocyst maturation in the vector (Modrzynska *et al.*, 2017), and the 6-cysteine protein P36 (PVP01_0303700), required by sporozoites for invasion and establishment of the parasitophorous vacuole within hepatocytes (Arredondo *et al.*, 2018), while one of the chromosome 9 peaks contains the gene for the red blood cell ligand apical membrane antigen 1 (PVP01_0934200), a major vaccine candidate antigen (Drew *et al.*, 2023).

Discussion

Plasmodium simium offers a compelling example of zoonotic malaria parasite originating from recent host-shift speciation events. Two main scenarios have been suggested for the origin of *P. simium*: (1) human-to-monkey shifts took place after *P. vivax* lineages were introduced in Brazil, mostly by European settlers but perhaps also by enslaved populations displaced from Africa (de Oliveira *et al.*, 2021a), and (2) *P. vivax* stocks from different source populations from Europe, Africa, and/or Asia were introduced in the New World.

While "generalist" lineages were able to infect local platyrrhine monkeys in addition to humans, giving rise to *P. simium*, more "specialist" lineages are expected to remain limited to humans (Rougeron *et al.*, 2022). Our findings support the first scenario. First, it was found low within-species nucleotide diversity among present-day lineages of *P. simium* circulating in both humans and monkeys, consistent with its recent origin from the same source population. Separate host-shift events may have occurred elsewhere in South America – e.g., in Colombia, where *P. vivax*-related parasites appear to infect wild nonhuman primates (Rondón *et al.*, 2019), but monkey-derived *P. vivax* samples outside Southeast Brazil have not been characterized at the genome level. Second, our global population structure analysis shows that all *P. simium* and Latin American *P. vivax* isolates cluster together (Figure 1) and both display genetic similarity to now-extinct European lines of *P. vivax* (Supplementary Figure 2, Supplementary Materials). The seemly common geographic origin of *P. simium* and Latin American *P. vivax* isolates argues against the hypothesis that separate introductions of parasites, from different regions or continents, originated monkey-adapted *vs.* exclusively human parasites found nowadays in Brazil. *Plasmodium simium* challenges current malaria control and elimination efforts. About 0.05% of the malaria cases recorded each year in Brazil are acquired outside the Amazon Basin (Ferreira and Castro, 2016), especially in forest fringes, where nonhuman primates serve as parasite reservoirs for spillback events (Abreu *et al.*, 2019; Duarte *et al.*, 2021) and human-vector contact is favored by environmental change (Medeiros-Sousa *et al.*, 2021). The vast majority of autochthonous human malaria cases in the Brazilian Atlantic Forest biome are reportedly caused by *P. vivax* (Garcia *et al.*, 2022), but available molecular evidence suggests that most, if not all, cases attributed to *P. vivax* in these settings are indeed caused by the zoonotic parasite *P. simium* (Brasil *et al.*, 2017; Buery *et al.*, 2017; de Oliveira *et al.*, 2021b). Here, further support to this hypothesis is provided by showing that the genome sequences of 12 parasites originally labelled as *P. vivax* circulating among humans in São Paulo (Ibrahim *et al.*, 2023) cluster together with *P. simium* lineages of human and nonhuman origin and can be clearly differentiated from *P. vivax* populations from the Amazon (Figure 2 and Supplementary Figure 6B). Importantly, zoonotic malaria transmission to humans occurs in the vicinity of Rio de Janeiro, São Paulo, and other major cities in Southeast Brazil and can undermine the country′s elimination efforts (Fornace *et al.*, 2024a; Fornace *et al.*, 2024b).

Previous analyses of substantially smaller *P. simium* sequence datasets had suggested that this parasite was most closely related to present-day *P. vivax* samples from Mexico, a relatively inbred population that has experienced a steady decline in recent years, compared to those from Brazil and other locations in Latin America (de Oliveira *et al.*, 2021b; Mourier *et al.*, 2021). However, these findings must not be over-interpreted, since *P. vivax* sequence data from Mexico originated from relatively isolated foci of residual malaria transmission in the southern part of the country (Hupalo et al., 2016). The present analysis combines additional *P. simium* sequences and a wide range of *P. vivax* sequence data from Latin America to show that *P. simium* is clearly more closely related to *P. vivax* from Brazil (Supplementary Figure 6A). However, evidence from PCA (Figure 2), ADMIXTURE (Figure 3) and F_{ST} analyses (Supplementary Figure 5) shows a clear divergence between *P. simium* and *P. vivax* from Brazil and other Latin American countries. At the species level, population differentiation follows an isolation-by-distance model, with little, if any, IBD sharing between *P. simium* samples from sites >450 km apart (Supplementary Figure 8) and across states (Figure 6). These results are consistent with focal parasite transmission among nonhuman primates, and occasionally humans, within discontinuous forest fragments intermingled with malaria-free areas in Southeast Brazil (Ferreira and Castro, 2016; Duarte *et al.*, 2021). However, the substantial IBD sharing (≥ 25%) found between parasites from humans and platyrrhine monkeys (Supplementary Figure 11, Supplementary Materials) suggests occasional cross-species transmission events (Su and Wu, 2021).

Mutations, insertions, and deletions in key erythrocyte invasion ligands of malaria parasites are commonly seen in host-shift speciation events (de Oliveira *et al*., 2021a). Accordingly, all *P. simium* isolates analysed so far display a deletion of >40% of the coding sequence of the locus encoding the reticulocyte binding protein 2a, *rbp2a* (de Oliveira *et al*., 2021b; Mourier *et al*., 2021), which encodes a reticulocyte-specific parasite ligand (Malleret *et al*., 2021) and may be involved in the adaptation to binding to and/or entrance into red blood cells of platyrrhine monkeys. Our search also revealed additional putative genomic signatures of parasite-host-vector adaptation in *P. simium*: (1) an elevated nucleotide polymorphism in genes putatively involved in gametocytogenesis and ookinete and oocyst development (Supplementary Table 6) and (2) an IBD peak suggestive of a selective sweep in a genomic domain containing the gene encoding an AP2 domain transcription factor family member, AP2-SP2, which appears to modulate oocyst maturation in the vector (Figure 7). In addition, it has been suggested that adaptive changes in the gamete surface protein (P47) – the *P. vivax* ortholog of Pfs47, a surface protein that allows *P. falciparum* to evade the mosquito immune system (Molina-Cruz *et al*., 2016) – could have enhanced the compatibility between *P. vivax* and *P. simium* and New World vectors. However, no IBD peak was found around the *p47* gene and relatively few additional sequences are currently available for testing this hypothesis (de Oliveira *et al*., 2021b). Somewhat surprisingly, both *P. simium* and *P. vivax* from the Amazon appear to infect efficiently anophelines from across the globe, in addition to mosquitoes found in the Amazon and along the Atlantic Coast of South America (Collins *et al*., 2005; Shaw-Saliba *et al*., 2016).

The present study has some limitations. First, it was analysed genome sequence data from only 31 *P. simium* isolates from three states in Southeast Brazil. No WGS data were available from the southernmost range of the current distribution of *P. simium* (states of Santa Catarina and Rio Grande do Sul). Second, no WGS data was available from parasites labelled as *P. vivax* that were found to infect wild (Rondón *et al*., 2019) or captive platyrrhine monkeys (Silva *et al*., 2019) outside the currently known geographic range of *P. simium*. Whether these parasites share recent ancestry with *P. simium* lineages from Southeast Brazil remains to be investigated. Third, PCA and ADMIXTURE analyses of the genetic relatedness between *P. simium* and global or regional populations of *P. vivax* are largely exploratory and may be affected by geographic biases in parasite sampling. For example, no *P. vivax* sequence data from West and Central Africa, where *P. vivax* is rare but not absent (Baird, 2022), were available for analysis. Importantly, *P. vivax* infections acquired in Angola – the origin of nearly two thirds of the more than 4 million enslaved Africans displaced to Brazil over three centuries – have been occasionally described in migrants and travelers (e.g., Haiyambo *et al*., 2019; Martins *et al*., 2020). Consequently, whether African lineages of *P. vivax* have contributed to the ancestry of *P. vivax* and *P. simium* lineages circulating nowadays in the New World remains undetermined. In addition, isolates from the Western Amazonian state of Acre were overrepresented in our dataset of *P. vivax* sequences from Brazil (143 of 203, or 70.4% of the sequences analysed), which included data from only 23 isolates from the Eastern Amazonian states of Amapá and Pará.

Supplementary material. The Supplementary Material for this paper can be found at [DOI]

Data availability. Reads of new whole-genome sequence data described in this article have been deposited into the NCBI Sequence Read Archive and are publicly available under BioProject accession code PRJNA1242290.

Acknowledgements. We are grateful to all patients who donated blood samples for genome sequencing; to Ajucilene (Joice) G. Mota, Francisco Melo, and their team at the Health Secretary of Mâncio Lima for their overall logistic support during fieldwork in Brazil; and to

Juliana Tonini for her support in the laboratory and the field.

Author's contribution. NRMA, JCS, and MUF conceived and designed the study. WAL, PTR, and TCO gathered genome sequence data. NRMA, RJS, AD, TCS, and JCS performed bioinformatic analysis. NRMA, JCS, and MUF wrote the first draft of the article, which was reviewed and approved by all authors.

Financial support. This research was supported by grants from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, United States of America (grant U19 AI089681, subcontract to MUF; grant R01 AI141900 to JCS) and the Fundação de Amparo à Pesquisa do Estado de São Paulo, Brazil (FAPESP; grant 2016/18740-9 to MUF). We also acknowledge post-doctoral FAPESP fellowships from FAPESP to NRMdA (2022/10056-2 and 2023/12394-5), PTR (2018/03902-9), and TCdO (2021/01017-0); a doctoral FAPESP scholarship to WAL (2023/15369-1); a senior researcher scholarship from the Conselho Nacional de Desenvolvimento Científico e Tecnológico of Brazil (CNPq; 301011/2019-2) to MUF; and institutional support from the Fundação para a Ciência e Tecnologia of Portugal (FCT), through the projects UID/04413/2020 to the Global Health and Tropical Medicine research center and LA-REAL LA/P/0117/2020. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Competing interests. The authors declare there are no conflicts of interest.

Ethical standards. Clinical samples for genome sequencing were collected under a research protocol approved by the Institutional Review Board of the Institute of Biomedical Sciences, University of São Paulo and the National Committee on Ethics in Research of the Ministry of Health of Brazil (CAAE number, 6467416.6.0000.5467). Written informed consent was obtained from all study participants or their parents or legal guardians.

References

Abreu, F.V.S., Santos, E.D., Mello, A.R.L., Gomes, L.R., Alvarenga, D.A.M., Gomes, M.Q., Vargas, W.P., Bianco-Júnior, C., Pina-Costa, A., Teixeira, D.S., Romano, A.P.M., Manso, P.P.A., Pelajo-Machado, M., Brasil, P., Daniel-Ribeiro, C.T., Brito, C.F.A, Ferreira-da-Cruz, M.F. and Lourenço-de-Oliveira, R. (20190. Howler monkeys are the reservoir of malarial parasites causing zoonotic infections in the Atlantic forest of Rio de Janeiro. *PLoS Neglected Tropical Diseases* **13**, e0007906. doi: 10.1371/journal.pntd.0007906.

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., Drury, E., Echeverry, D.F., Erko, B., Espino, F., Fairhurst, R.M., Faiz, A., Fernanda Villegas, M., Gao, Q., Golassa, L., Goncalves, S., Grigg, M.J., Hamedi, Y., Hien, T.T., Htut, Y., Johnson, K.J., Karunaweera, N.D., Khan, W., Krudsood, S., Kwiatkowski, D.P., Lacerda, M., Ley, B., Lim, P., Liu, Y., Llanos-Cuentas, A., Lon, C., Lopera-Mesa, T.M., Marfurt, J., Michon, P., Miotto, O., Mohammed, R., Mueller, I., Namaik-Larp, C., Newton, P.N., Nguyen, T.N., Nosten, F., Noviyanti, R., Pava, Z., Pearson, R.D., Petros, B., Phyo, A.P., Price, R.N., Pukrittayakamee, S., Rahim, A.G., Randrianarivelojosia, M., Rayner, J.C., Rumaseb, A., Siegel, S.V., Simpson, V.J., Thriemer, K., Tobon-Castano, A., Trimarsanto, H., Urbano Ferreira, M., Vélez, I.D., Wangchuk, S., Wellems, T.E., White, N.J., William, T., Yasnot, M.F. and Yilma, D. (2022). An open dataset of *Plasmodium* vivax genome variation in 1,895 worldwide samples. *Wellcome Open Research* **7**, 136. doi: 10.12688/wellcomeopenres.17795.1.

Alexander, D.H. and Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics* **12**, 246. doi: 10.1186/1471-2105-12-246.

Alexander, D.H., Novembre, J. and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* **19**, 1655-64. doi: 10.1101/gr.094052.109.

Arredondo, S.A., Swearingen, K.E., Martinson, T., Steel, R., Dankwa, D.A., Harupa, A., Camargo, N., Betz, W., Vigdorovich, V., Oliver, B.G., Kangwanrangsan, N., Ishino, T., Sather, N., Mikolajczak, S., Vaughan, A.M., Torii, M., Moritz, R.L. and Kappe, S.H.I. (2018). The micronemal *Plasmodium* proteins P36 and P52 act in concert to establish the replication-permissive compartment within infected hepatocytes. *Frontiers in Cellular and Infection Microbiology* **8**, 413. doi: 10.3389/fcimb.2018.00413.

Auburn, S., Böhme, U., Steinbiss, S., Trimarsanto, H., Hostetler, J., Sanders, M., Gao, Q., Nosten, F., Newbold, C.I., Berriman, M., Price, R.N. and 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. doi: 10.12688/wellcomeopenres.9876.1.

Avalos-Padilla, Y., Georgiev, V.N., Lantero, E., Pujals, S., Verhoef, R., Borgheti-Cardoso, L., Albertazzi, L., Dimova, R. and Fernàndez-Busquets, X. (2021) The ESCRT-III machinery participates in the production of extracellular vesicles and protein export during *Plasmodium falciparum* infection. *PLOS Pathogens* **17**, 1009455. doi: 10.1371/journal.ppat.1009455

Baird, J.K. (2022). African *Plasmodium vivax* malaria improbably rare or benign. *Trends in Parasitology* **38**, 683-696. doi: 10.1016/j.pt.2022.05.006.

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. and Clark, T.G. (2021). Distinctive genetic structure and selection patterns in *Plasmodium vivax* from South Asia and East Africa. *Nature Communications* **12**, 3160. doi: 10.1038/s41467-021-23422-3**.**

Brasil, P., Zalis, M.G., de Pina-Costa, A., Siqueira, A.M., Júnior, C.B., Silva, S., Areas, A.L.L., Pelajo-Machado, M., de Alvarenga, D.A.M., da Silva Santelli, A.C.F., Albuquerque, H.G., Cravo, P., Santos de Abreu, F.V., Peterka, C.L., Zanini, G.M., Suárez Mutis, M.C., Pissinatti, A., Lourenço-de-Oliveira, R., de Brito, C.F.A., Ferreira-da-Cruz, M.F., Culleton, R. and Daniel-Ribeiro, C.T. (2017). Outbreak of human malaria caused by *Plasmodium simium* in the Atlantic Forest in Rio de Janeiro: a molecular epidemiological investigation. *Lancet Global Health* **5**, e1038-e1046. doi: 10.1016/S2214-109X(17)30333-9.

Browning, B.L., Zhou, Y. and Browning, S.R. (2018). A one-penny imputed genome from next-generation reference panels. *American Journal of Human Genetics* **103**, 338-348. doi: 10.1016/j.ajhg.2018.07.015.

Bruce-Chwatt, L.J. (1965). Paleogenesis and paleo-epidemiology of primate malaria. *Bulletin of the World Health Organization* **32**, 363-387.

Bruce-Chwatt, L.J. and de Zulueta, J (1980). *The Rise and Fall of Malaria in Europe: A Historico-Epidemiological Study*. Oxford, UK: Oxford University Press.

Buery, J.C., Rodrigues, P.T., Natal, L., Salla, L.C., Loss, A.C., Vicente, C.R., Rezende, H.R., Duarte, A.M.R.C, Fux, B., Malafronte, R.S., Falqueto, A. and Cerutti, C. Jr. (2017). Mitochondrial genome of *Plasmodium vivax/simium* detected in an endemic region for malaria in the Atlantic Forest of Espírito Santo state, Brazil: do mosquitoes, simians and humans harbour the same parasite? *Malaria Journal* **16**, 437. doi: 10.1186/s12936-017-2080-9.

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 III, N. (2020). Population genomics of *Plasmodium vivax* in Panama to assess the risk of case importation on malaria elimination. *PLoS Neglected Tropical Diseases* **14**, e0008962. doi: 10.1371/journal.pntd.0008962.

Carter, R. (2003). Speculations on the origins of *Plasmodium vivax* malaria. *Trends in Parasitology* **19**, 214-9. doi: 10.1016/s1471-4922(03)00070-9.

Collins, W.E., Sullivan, J.S., Galland, G.G., Williams, A., Nace, D., Williams, T. and Barnwell, J.W. (2005). *Plasmodium simium* and *Saimiri boliviensis* as a model system for testing candidate vaccines against *Plasmodium vivax*. *American Journal of Tropical Medicine and Hygiene* **73**, 644-8.

Cowell, A.N., Loy, D.E., Sundararaman, S.A., Valdivia, H., Fisch, K., Lescano, A.G., Baldeviano, G.C., Durand, S., Gerbasi, V., Sutherland, C.J., Nolder, D., Vinetz, J.M., Hahn, B.H. and Winzeler, E.A. (2017). Selective whole-genome amplification is a robust method that enables scalable whole-genome sequencing of *Plasmodium vivax* from unprocessed clinical samples. *mBio* **8**, e02257-16. doi: 10.1128/mBio.02257-16.

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.; 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics* **27**, 2156-8. doi: 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. and

Rougeron V. (2021). Population genomic evidence of *Plasmodium vivax* Southeast Asian origin. *Science Advances* **7**, eabc3713. doi: 10.1126/sciadv.abc3713.

de Alvarenga, D.A.M., Culleton, R., de Pina-Costa, A., Rodrigues, D.F., Bianco, C. Jr, Silva, S., Nunes, A.J.D., de Souza, J.C. Jr, Hirano, Z.M.B., Moreira, S.B., Pissinatti, A., de Abreu, F.V.S., Lisboa Areas, A.L., Lourenço-de-Oliveira, R., Zalis, M.G., Ferreira-da-Cruz, M.F., Brasil, P., Daniel-Ribeiro, C.T. and de Brito, C.F.A. (2018). An assay for the identification of *Plasmodium simium* infection for diagnosis of zoonotic malaria in the Brazilian Atlantic Forest. *Scientific Reports* **8**, 86. doi: 10.1038/s41598-017-18216-x.

de Azevedo, T.S., Lorenz, C., Chiaravalloti-Neto, F. and Sallum, M.A.M. (2020). *Kerteszia cruzii* and extra-Amazonian malaria in Brazil: challenges due to climate change in the Atlantic Forest. *Infection, Genetics and Evolution* **85**, 104456. doi: 10.1016/j.meegid.2020.104456.

De Meulenaere, K., Cuypers, B., Gamboa, D., Laukens, K. and Rosanas-Urgell, A. (2023). A new *Plasmodium vivax* reference genome for South American isolates. *BMC Genomics* **24**, 606. doi: 10.1186/s12864-023-09707-5.

de Oliveira, T.C., Corder, R.M., Early, A., Rodrigues, P.T., Ladeia-Andrade, S., Alves, J.M.P., Neafsey, D.E. and 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**, e0008808. doi: 10.1371/journal.pntd.0008808.

de Oliveira, T.C., Rodrigues, P.T., Duarte, A.M.R.C., Rona, L.D.P. and Ferreira, M.U. (2021a). Ongoing host-shift speciation in *Plasmodium simium*. *Trends in Parasitology* **37**, 940-942. doi: 10.1016/j.pt.2021.08.005.

de Oliveira, T.C., Rodrigues, P.T., Early, A.M., Duarte, A.M.R.C., Buery, J.C., Bueno, M.G., Catão-Dias, J.L., Cerutti, C., Rona, L.D.P., Neafsey, D.E. and Ferreira MU. (2021b). *Plasmodium simium*: population genomics reveals the origin of a reverse zoonosis. *Journal of Infectious Diseases* **224**, 1950-1961. doi: 10.1093/infdis/jiab214.

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 Morais, G., Berná, L., Phelan, J., Robello, C., de Vasconcelos, A.T.R., Alves, J.M.P. and 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**, e0005824. doi: 10.1371/journal.pntd.0005824.

Drew, D.R., Wilson, D.W., Weiss, G.E., Yeoh, L.M.G., Henshall, I., Crabb, B.S., Dutta, S., Gilson, P.R. and Beeson, J.G. (2023). Defining species-specific and conserved interactions of apical membrane protein 1 during erythrocyte invasion in malaria to inform multi-species vaccines. *Cellular and Molecular Life Sciences* **80**, 74. doi: 10.1007/s00018-023-04712-z.

Duarte, A.M.R.C., Fernandes, L.N., Silva, F.S., Sicchi, I.L., Mucci, L.F., Curado, I., Fernandes, A., Medeiros-Sousa, A.R., Ceretti-Junior, W., Marrelli, M.T., Evangelista, E., Teixeira, R., Summa, J.L., Nardi, M.S., Garnica, M.R., Loss, A.C., Buery, J.C., Cerutti, C. Jr, Pacheco, M.A., Escalante, A.A., Sallum, M.A.M. and Laporta, G.Z. (2021). Complexity of malaria transmission dynamics in the Brazilian Atlantic Forest. *Current Research in Parasitology and Vector-Borne Diseases* **1**, 100032. doi: 10.1016/j.crpvbd.2021.100032.

Ferreira, M.U. and Castro, M.C. (2016). Challenges for malaria elimination in Brazil. *Malaria Journal* **15**, 284. doi: 10.1186/s12936-016-1335-1.

Ferreira, M.U. and Castro, M.C. (2019). Malaria situation in Latin America and the Caribbean: residual and resurgent transmission and challenges for control and elimination. *Methods in Molecular Biology* **2013**, 57-70. doi: 10.1007/978-1-4939-9550-9_4

Fontoura, P.S., Macedo, E.G., Calil, P.R., Corder, R.M., Rodrigues, P.T., Tonini, J., Esquivel, F.D., Ladeia, W.A., Fernandes, A.R.J., Johansen, I.C., Silva, M.F., Fernandes, A.O.S, Ladeia-Andrade, S., Castro, M.C. and Ferreira, M.U. (2024). Changing clinical epidemiology of *Plasmodium vivax* malaria as transmission decreases: population-based prospective panel survey in the Brazilian Amazon. *Journal of Infectious Diseases* **229**, 947-958. doi: 10.1093/infdis/jiad456.

Fornace, K.M., Drakeley, C.J., Lindblade, K.A., Jelip, J. and Ahmed, K. (2023a). Zoonotic malaria requires new policy approaches to malaria elimination. *Nature Communications* **14**, :5750. doi: 10.1038/s41467-023-41546-6.

Fornace, K.M., Laporta, G.Z., Vythilingham, I., Chua, T.H., Ahmed, K., Jeyaprakasam, N.K., Duarte, A.M.R.C, Amir, A., Phang, W.K., Drakeley, C., Sallum, M.A.M. and Lau, Y.L. (2023b). Simian malaria: a narrative review on emergence, epidemiology, and threat to global malaria elimination. *Lancet Infectious Diseases* **23**, e520-e532. doi: 10.1016/S1473-3099(23)00298-0.

Franke-Fayard, B., Marin-Mogollon, C., Geurten, F.J.A., Chevalley-Maurel, S., Ramesar, J., Kroeze, H., Baalbergen, E., Wessels, E., Baron, L., Soulard, V., Martinson, T., Aleshnick, M., Huijs, A.T.G., Subudhi, A.K., Miyazaki, Y., Othman, A.S., Kolli, S.K., Lamers, O.A.C., Roques, M., Stanway, R.R., Murphy, S.C., Foquet, L., Moita, D., Mendes, A.M., Prudêncio, M., Dechering, K.J., Heussler, V.T., Pain, A., Wilder, B.K., Roestenberg, M. and Janse, C.J. (2022). Creation and preclinical evaluation of genetically attenuated malaria parasites arresting growth late in the liver. *npj Vaccines* 7, 139. doi: 10.1038/s41541-022-00558-x

Garcia, K.K.S., Abrahão, A.A., Oliveira, A.F.M., Henriques, K.M.D., de Pina-Costa, A., Siqueira, A.M. and Ramalho, W.M. (2022). Malaria time series in the extra-Amazon region of Brazil: epidemiological scenario and a two-year prediction model. *Malaria Journal* **21**, 157. doi: 10.1186/s12936-022-04162-1.

Guo, B., Borda, V., Laboulaye, R., Spring, M.D., Wojnarski, M., Vesely, B.A., Silva, J.C., Waters N.C., O'Connor, T.D. and Takala-Harrison, S. (2024). Strong positive selection biases identity-by-descent-based inferences of recent demography and population structure in *Plasmodium falciparum*. *Nature Communications* **15**, 2499. doi: 10.1038/s41467-024-46659-0.

Haussig, J.M., Matuschewski, K. and Kooij, T.W.A. (2011). Inactivation of a *Plasmodium* apicoplast protein attenuates formation of liver merozoites. *Molecular Microbiology* **81**, 1511-1525. doi: 10.1111/j.1365-2958.2011.07787.x

Haiyambo, D.H., Uusiku, P., Mumbengegwi, D., Pernica, J.M., Bock, R., Malleret, B., Rénia, L., Greco, B. and Quaye, I.K. (2019). Molecular detection of *P. vivax* and *P. ovale* foci of infection in asymptomatic and symptomatic children in Northern Namibia. *PLoS Neglected Tropical Diseases* **13**, e0007290. doi: 10.1371/journal.pntd.0007290.

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., Baldeviano, G.C., Lescano, A.G., Laman, M., Barnadas, C., Barry, A., Mueller, I., Kazura, J.W., Eapen, A., Kanagaraj, D., Valecha, N., Ferreira, M.U., Roobsoong, W., Nguitragool, W., Sattabonkot, J., Gamboa, D., Kosek, M., Vinetz, J.M., González-Cerón, L., Birren, B.W., Neafsey, D.E. and Carlton, J.M. (2016). Population genomics studies identify signatures of global dispersal and drug resistance in *Plasmodium vivax*. *Nature Genetics* **48**, 953-8. doi: 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., Baptista, A.R.S., Machado, R.L.D., Marinho, C.R.F., Clark, T.G. and Campino, S. (2023). Population-based genomic study of *Plasmodium vivax* malaria in seven Brazilian states and across South America. *Lancet Regional Health – Americas* **18**, 100420. doi: 10.1016/j.lana.2022.100420.

Kattenberg, J.H., Monsieurs, P., De Meyer, J., De Meulenaere, K., Sauve, E., de Oliveira, T.C., Ferreira, M.U., Gamboa, D. and 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. doi: 10.1002/ece3.11103.

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., Esbroeck, M.V., Auburn, S., Nguyen, B.T.H. and 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. doi: 10.3389/fcimb.2022.953187.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM [preprint]. *arXiv*, 1303.3997. doi: 10.48550/arXiv.1303.3997.

Li, H. and Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493-6. doi: 10.1038/nature10231

Malleret, B., El Sahili, A., Tay, M.Z., Carissimo, G., Ong, A.S.M., Novera, W., Lin, J., Suwanarusk, R., Kosaisavee, V., Chu, T.T.T., Sinha, A., Howland, S.W., Fan, Y., Gruszczyk, J., Tham, W.H., Colin, Y., Maurer-Stroh, S., Snounou, G., Ng, L.F.P., Chan, J.K.Y., Chacko, A.M., Lescar, J., Chandramohanadas, R., Nosten, F., Russell,

B. and Rénia, L. (2021). *Plasmodium vivax* binds host CD98hc (SLC3A2) to enter immature red blood cells. *Nature Microbiology* **6**, 991-999. doi: 10.1038/s41564-021-00939-3.

Manske, M., Miotto, O., Campino, S., Auburn, S., Almagro-Garcia, J., Maslen, G., O'Brien, J., Djimde, A., Doumbo, O., Zongo, I., Ouedraogo, J.B., Michon, P., Mueller, I., Siba, P., Nzila, A., Borrmann, S., Kiara, S.M., Marsh, K., Jiang, H., Su, X.Z., Amaratunga, C., Fairhurst, R., Socheat, D., Nosten, F., Imwong, M., White, N.J., Sanders, M., Anastasi, E., Alcock, D., Drury, E., Oyola, S., Quail, M.A., Turner, D.J., Ruano-Rubio, V., Jyothi, D., Amenga-Etego, L., Hubbart, C., Jeffreys, A., Rowlands, K., Sutherland, C., Roper, C., Mangano, V., Modiano, D., Tan, J.C., Ferdig, M.T., Amambua-Ngwa, A., Conway, D.J., Takala-Harrison, S., Plowe, C.V., Rayner, J.C., Rockett, K.A., Clark, T.G., Newbold, C.I., Berriman, M., MacInnis, B. and Kwiatkowski, D.P. (2012). Analysis of *Plasmodium falciparum* diversity in natural infections by deep sequencing. *Nature* **487**, 375-9. doi: 10.1038/nature11174.

Martins, J.F., Marques, C., Nieto-Andrade, B., Kelley, J., Patel, D., Nace, D., Herman, C., Barratt, J., Ponce de León, G., Talundzic, E., Rogier, E., Halsey, E.S. and Plucinski, M.M. (2020). Malaria risk and prevention in Asian migrants to Angola. *American Journal of Tropical Medicine and Hygiene* **103**, 1918-1926. doi: 10.4269/ajtmh.20-0706.

Mather, N., Traves, S.M. and Ho, S.Y.W. (2019). A practical introduction to sequentially Markovian coalescent methods for estimating demographic history from genomic data. *Ecology and Evolution* **10**, 579-589. doi: 10.1002/ece3.5888.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. and DePristo, M.A. (2010). The

Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* **20**, 1297-303. doi: 10.1101/gr.107524.110.

Medeiros-Sousa, A.R., Laporta, G.Z., Coutinho, R.M., Mucci, L.F. and Marrelli, M.T. (2021). A mathematical model for zoonotic transmission of malaria in the Atlantic Forest: exploring the effects of variations in vector abundance and acrodendrophily. *PLoS Neglected Tropical Diseases* **15**, e0008736. doi: 10.1371/journal.pntd.0008736.

Michel, M., Skourtanioti, E., Pierini, F., Guevara, E.K., Mötsch, A., Kocher, A., Barquera, R., Bianco, R.A., Carlhoff, S., Coppola Bove, L., Freilich, S., Giffin, K., Hermes, T., Hiß, A., Knolle, F., Nelson, E.A., Neumann, G.U., Papac, L., Penske, S., Rohrlach, A.B., Salem, N., Semerau, L., Villalba-Mouco, V., Abadie, I., Aldenderfer, M., Beckett, J.F., Brown, M., Campus, F.G.R., Chenghwa, T., Cruz Berrocal, M., Damašek, L., Duffett Carlson, K.S., Durand, R., Ernée, M., Fântăneanu, C., Frenzel, H., García Atiénzar, G., Guillén, S., Hsieh, E., Karwowski, M., Kelvin, D., Kelvin, N., Khokhlov, A., Kinaston, R.L., Korolev, A., Krettek, K.L., Küßner, M., Lai, L., Look, C., Majander, K., Mandl, K., Mazzarello, V., McCormick, M., de Miguel Ibáñez, P., Murphy, R., Németh, R.E., Nordqvist, K., Novotny, F., Obenaus, M., Olmo-Enciso, L., Onkamo, P., Orschiedt, J., Patrushev, V., Peltola, S., Romero, A., Rubino, S., Sajantila, A., Salazar-García, D.C., Serrano, E., Shaydullaev, S., Sias, E., Šlaus, M., Stančo, L., Swanston, T., Teschler-Nicola, M., Valentin, F., van de Vijver, K., Varney, T.L., Vigil-Escalera Guirado, A., Waters, C.K., Weiss-Krejci, E., Winter, E., Lamnidis, T.C., Prüfer, K., Nägele, K., Spyrou, M., Schiffels, S., Stockhammer, P.W., Haak, W., Posth, C., Warinner, C., Bos, K.I., Herbig, A. and Krause, J. (2024). Ancient *Plasmodium* genomes shed light on the history of human malaria. *Nature* **631**, 125-133. doi: 10.1038/s41586-024-07546-2.

Miles, A., Iqbal, Z., Vauterin, P., Pearson, R., Campino, S., Theron, M., Gould, K., Mead, D., Drury, E., O'Brien, J., Ruano Rubio, V., MacInnis, B., Mwangi, J., Samarakoon, U., Ranford-Cartwright, L., Ferdig, M., Hayton, K., Su, X.Z., Wellems, T., Rayner, J., McVean, G. and Kwiatkowski, D. (2016). Indels, structural variation, and recombination drive genomic diversity in *Plasmodium falciparum*. *Genome Research* **26**, 1288-99. doi: 10.1101/gr.203711.115.

Modrzynska, K., Pfander, C., Chappell, L., Yu, L., Suarez, C., Dundas, K., Gomes, A.R., Goulding, D., Rayner, J.C., Choudhary, J. and Billker, O. (2017). A knockout screen of *ApiAP2* genes reveals networks of interacting transcriptional regulators controlling the *Plasmodium* life cycle. *Cell Host and Microbe* **21**, 11-22. doi: 10.1016/j.chom.2016.12.003.

Molina-Cruz, A., Zilversmit, M.M., Neafsey, D.E., Hartl, D.L. and Barillas-Mury, C. (2016). Mosquito vectors and the globalization of *Plasmodium falciparum* malaria. *Annual Review of Genetics* **50**, 447-465. doi: 10.1146/annurev-genet-120215-035211.

Mourier, T., de Alvarenga, D.A.M., Kaushik, A., de Pina-Costa, A., Douvropoulou, O., Guan, Q., Guzmán-Vega, F.J., Forrester, S., de Abreu, F.V.S., Júnior, C.B., de Souza Junior, J.C., Moreira, S.B., Hirano, Z.M.B., Pissinatti, A, Ferreira-da-Cruz, M.F., de Oliveira, R.L., Arold, S.T., Jeffares, D.C., Brasil, P., de Brito, C.F.A., Culleton, R., Daniel-Ribeiro, C.T. and Pain, A. (2021). The genome of the zoonotic malaria parasite *Plasmodium simium* reveals adaptations to host switching. *BMC Biology* **19**, 219. doi: 10.1186/s12915-021-01139-5.

Nakayama, K., Kimura, Y., Kitahara, Y., Soga, A., Haraguchi, A., Hakozaki, J., Sugiyama, M., Kusakisako, K., Fukumoto, S. and Ikadai, H. (2021). Role of *Plasmodium berghei* ookinete surface and oocyst capsule protein, a novel oocyst capsule-associated protein, in ookinete motility. *Parasites and Vectors* **14**, 373. doi: 10.1186/s13071-021-04868-2

Nishi, T., Kaneko, I., Iwanaga, S. and Yuda, M. (2022). Identification of a novel AP2 transcription factor in zygotes with an essential role in *Plasmodium* ookinete development. *PLOS Pathogens* **18**, e1010510. doi: 10.1371/journal.ppat.1010510

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T. and Reich, D. (2012). Ancient admixture in human history. *Genetics* **192**, 1065-93. doi: 10.1534/genetics.112.145037.

Patterson, N., Price, A.L. and Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genetics* **2**, e190. doi: 10.1371/journal.pgen.0020190.

Petr, M. (2024). *admixr: An Interface for Running 'ADMIXTOOLS' Analyses*. R package version 0.9.1.9000, https://github.com/bodkan/admixr.

Pigati, J.S., Springer, K.B., Honke, J.S., Wahl, D., Champagne, M.R., Zimmerman, S.R.H., Gray, H.J., Santucci, V.L., Odess, D., Bustos, D. A. and Bennett, M.R. (2023). Independent age estimates resolve the controversy of ancient human footprints at White Sands. *Science* **382**, 73-75. doi: 10.1126/science.adh5007.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* **81**, 559-75. doi: 10.1086/519795.

Rodrigues, P.T., Alves, J.M., Santamaria, A.M., Calzada, J.E., Xayavong, M., Parise, M., da Silva, A.J. and Ferreira, M.U. (2014). Using mitochondrial genome sequences to track the origin of imported *Plasmodium vivax* infections diagnosed in the United States. *American Journal of Tropical Medicine and Hygiene* **90**, 1102-8. doi: 10.4269/ajtmh.13-0588.

Rodrigues, P.T., Johansen, I.C., Ladeia, W.A., Esquivel, F.D., Corder, R.M., Tonini, J., Calil, P.R., Fernandes, A.R.J., Fontoura, P.S., Cavasini, C.E., Vinetz, J.M., Castro, M.C. and Ferreira, M.U.; Mâncio Lima Cohort Study Working Group. (2024). Lower microscopy sensitivity with decreasing malaria prevalence in the urban Amazon region, Brazil, 2018-2021. *Emerging Infectious Diseases* **30**, 1884-1894. doi: 10.3201/eid3009.240378.

Rondón, S., León, C., Link, A. and González, C. (2019). Prevalence of *Plasmodium* parasites in non-human primates and mosquitoes in areas with different degrees of fragmentation in Colombia. *Malaria Journal* **18**, 276. doi: 10.1186/s12936-019-2910-z.

Rougeron, V., Daron, J., Fontaine, M.C. and Prugnolle, F. (2022). Evolutionary history of *Plasmodium vivax* and *Plasmodium simium* in the Americas. *Malaria Journal* **21**, 141. doi: 10.1186/s12936-022-04132-7.

Shaw-Saliba, K., Clarke, D., Santos, J.M., Menezes, M.J., Lim, C., Mascarenhas, A., Chery, L., Gomes, E., March, S., Bhatia, S.N., Rathod, P.K., Ferreira, M.U., Catteruccia, F., and Duraisingh, M.T. (2016). Infection of laboratory colonies of *Anopheles* mosquitoes with *Plasmodium vivax* from cryopreserved clinical isolates. *International Journal for Parasitology* **46**, 679-83. doi: 10.1016/j.ijpara.2016.06.003.

Schaffner, S.F., Taylor, A.R., Wong, W., Wirth, D.F. and Neafsey, D.E. (2018). hmmIBD: software to infer pairwise identity by descent between haploid genotypes. *Malaria Journal* **17**, 196. doi: 10.1186/s12936-018-2349-7.

Shang X, Shen S, Tang J, He, X., Zhao, Y., Wang, C., He, X., Guo, G., Liu, M., Wang, L., Zhu, Q., Yang, G., Jiang, C., Zhang, M., Yu, X., Han, J., Culleton, R., Jiang, L., Cao, J., Gu, L. and Zhang, Q. (2021). A cascade of transcriptional repression determines sexual commitment and development in *Plasmodium falciparum*. *Nucleic Acids Research* **16**, 9264-9279. doi: 10.1093/nar/gkab683.

Silva, T.R.M., Barros, F.N.L., Bahia, M., Sampaio Junior, F.D., Santos, S.S.F., Inoue, L.S., Gonçalves, T.S., Chiesorin Neto, L., Faria, D.C.L.O., Tochetto, C., Viana, G.M.R., Monteiro, F.O.B., Góes-Cavalcante, G. and Scofield, A. (2019). *Plasmodium vivax* and *Plasmodium falciparum* infection in Neotropical primates in the western Amazon, Brazil. *Zoonoses and Public Health* **66**, 798-804. doi: 10.1111/zph.12626.

Srinivasan, P., Fujioka, H. and Jacobs-Lorena, M. (2008). PbCap380, a novel oocyst capsule protein, is essential for malaria parasite survival in the mosquito. *Cellular Microbiology* **10,** 1304-1312. doi: 10.1111/j.1462-5822.2008.01127.x

Su, X.Z. and Wu, J. (2021). Zoonotic transmissions and host switches of malaria parasites. *Zoonoses* (Burlington) **1**, 11. doi: 10.15212/zoonoses-2021-0015.

Tachibana, S., Sullivan, S.A., Kawai, S., Nakamura, S., Kim, H.R., Goto, N., Arisue, N., Palacpac, N.M., Honma, H., Yagi, M., Tougan, T., Katakai, Y., Kaneko, O., Mita, T., Kita, K., Yasutomi, Y., Sutton, P.L., Shakhbatyan, R., Horii, T., Yasunaga, T., Barnwell, J.W., Escalante, A.A., Carlton, J.M. and Tanabe, K. (2012). *Plasmodium cynomolgi* genome sequences provide insight into *Plasmodium vivax* and the monkey malaria clade. *Nature Genetics* **44**, 1051-5. doi: 10.1038/ng.2375.

Toda, H., Diaz-Varela, M., Segui-Barber, J. Roobsoong, W., Baro, B., Garcia-Silva, S., Galiano, A., Gualdrón-López, M., Almeida, A.C.G., Brito, M.A.M., de Melo, G.C., Aparici-Herraiz, I., Castro-Cavadía, C., Monteiro, W.M., Borràs, E., Sabidó, E., Almeida, I.C., Chojnacki, J., Martinez-Picado, J., Calvo, M., Armengol, P., Carmona-Fonseca, J., Yasnot, M.F., Lauzurica, R., Marcilla, A., Peinado, H., Galinski, M.R., Lacerda, M.V.G., Sattabongkot, J., Fernandez-Becerra, C. and del Portillo, H.A. (2020). Plasma-derived extracellular vesicles from *Plasmodium vivax* patients signal spleen fibroblasts via NF-kB facilitating parasite cytoadherence. *Nature Communications* **11**, 2761 https://doi.org/10.1038/s41467-020-16337-y

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. and Lalueza-Fox, C. (2020). *Plasmodium vivax* malaria viewed through the lens of an eradicated European strain. *Molecular Biology and Evolution* **37**, 773-785. doi: 10.1093/molbev/msz264.

World Health Organization (2024) World Malaria Report 2024: addressing inequity in the global malaria response. Geneva: World Health Organization; 2024 (https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2024).

Figure 1. Global *Plasmodium vivax* and *Plasmodium simium* population structure revealed by standard principal component analysis (PCA). Data analysed corresponds to linkage disequilibrium-pruned biallelic SNPs. We show the first two principal components (PCs), which together account for 35.9% of the overall variance. Each symbol – circles for *P. simium* and triangles for *P. vivax* – represents a single isolate and was coloured according to the country of origin of the sample.

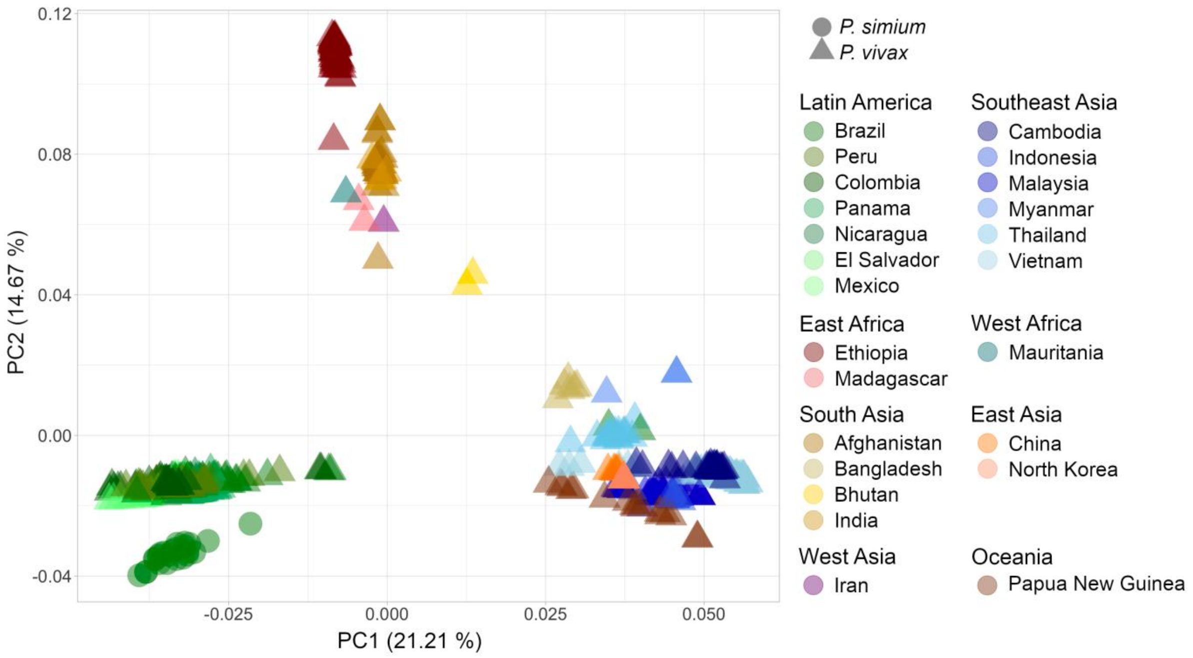


Figure 2. *P. simium* and *P. vivax* population structure in Latin America revealed by principal component analysis (PCA). Analysis included a total of 495 isolates (*P. simium*: n = 31; *P. vivax*: n = 464) We display the first three principal components (PCs), which together account for 45.9% of the overall variance. Each symbol – circles for *P. simium* and squares for *P. vivax* – represents a single isolate and was coloured according to the country or state (within Brazil) of origin of the sample. Locations of each state in Brazil are shown in Supplementary Figure 1.

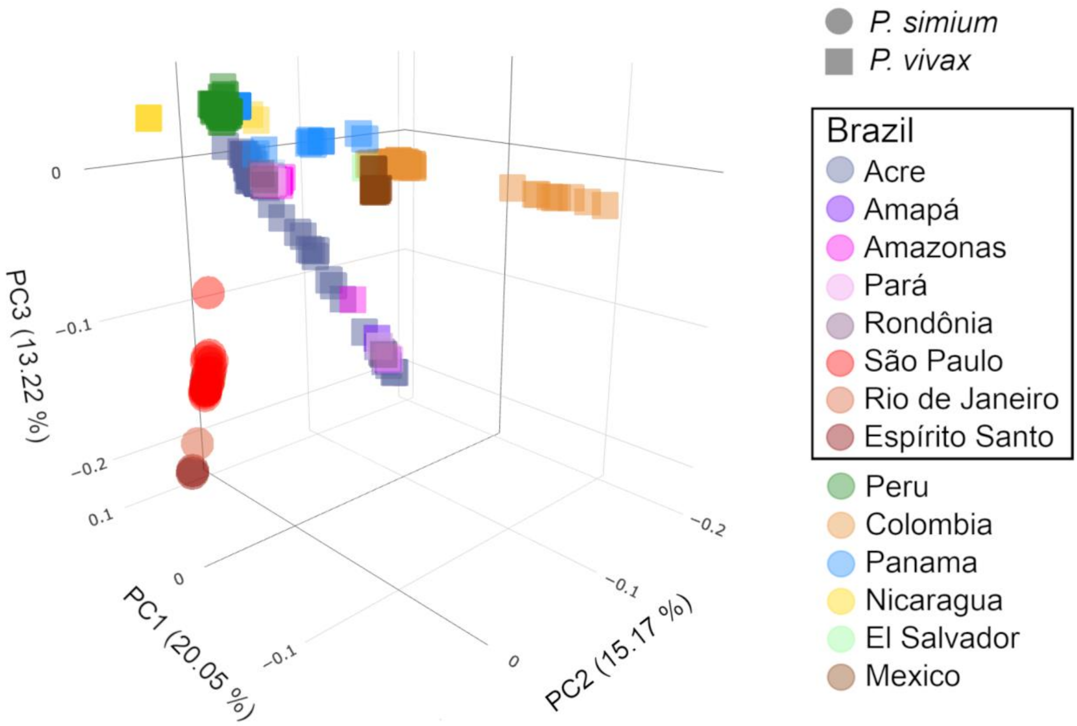


Figure 3. Unsupervised ADMIXTURE analysis of *P. simium* and *P. vivax* from Latin America. Three layers correspond to $K = 4$, $K = 10$ and $K = 15$ populations. In ADMIXTURE bar plots, each isolate is represented by a bar that is coloured to indicate the proportion of the genome (from 0 to 1) with ancestry from each of K putative ancestral source populations. Admixed samples (those with ancestry from more than one source population) are represented by bar segments of different colors.

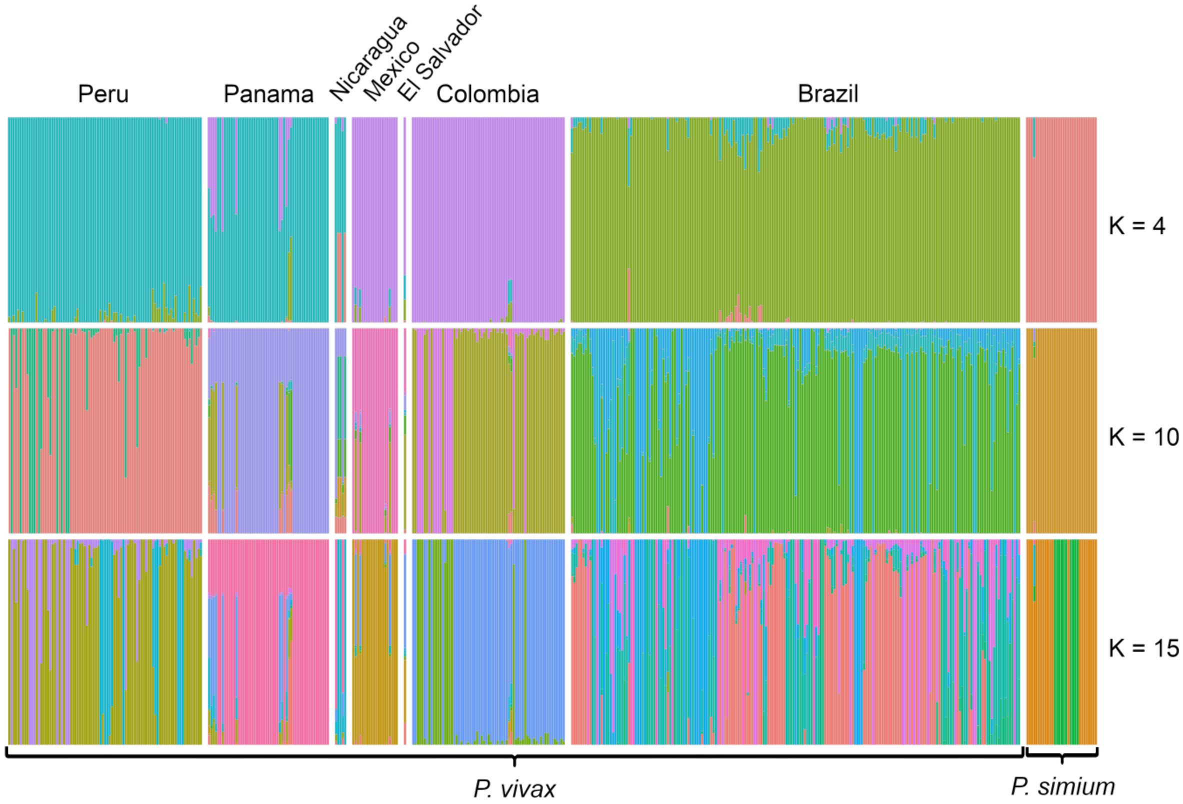


Figure 4. History of effective population size, N_e, for *P. simium* and *P. vivax*. Pairwise sequentially Markovian coalescent (PSMC) model estimation of historical population size changes in *P. simium* and *P. vivax* populations from Brazil. We used a mutation rate (μ) of 1×10^{-9} and generation time (g) of 0.18. Results are shown for the five samples with the highest sequence coverage for each species.

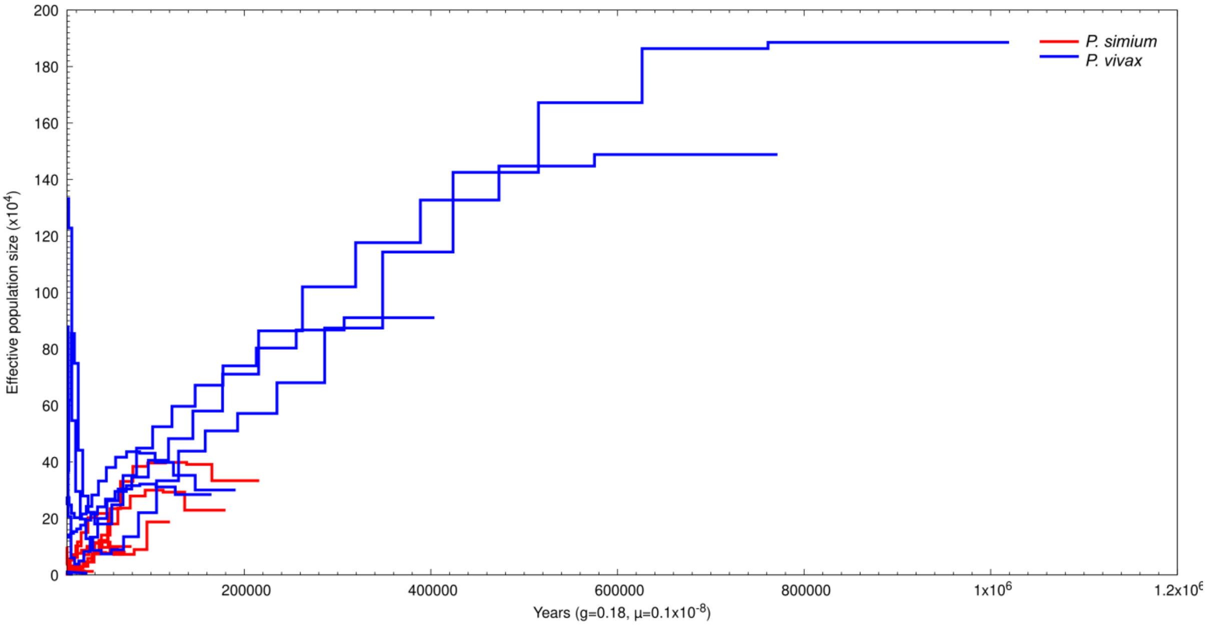


Figure 5. Nucleotide diversity in *P. simium* and *P. vivax*. Empirical density distribution of the log ratios of nucleotide diversity (π) estimates for *P. simium* and *P. vivax* ortholog genes (bars). The red line shows the null distribution centred on 0 as expected under identical nucleotide diversity between species. Data corresponds to a total of 3,615 orthologous gene pairs between *P. vivax* and *P. simium*.

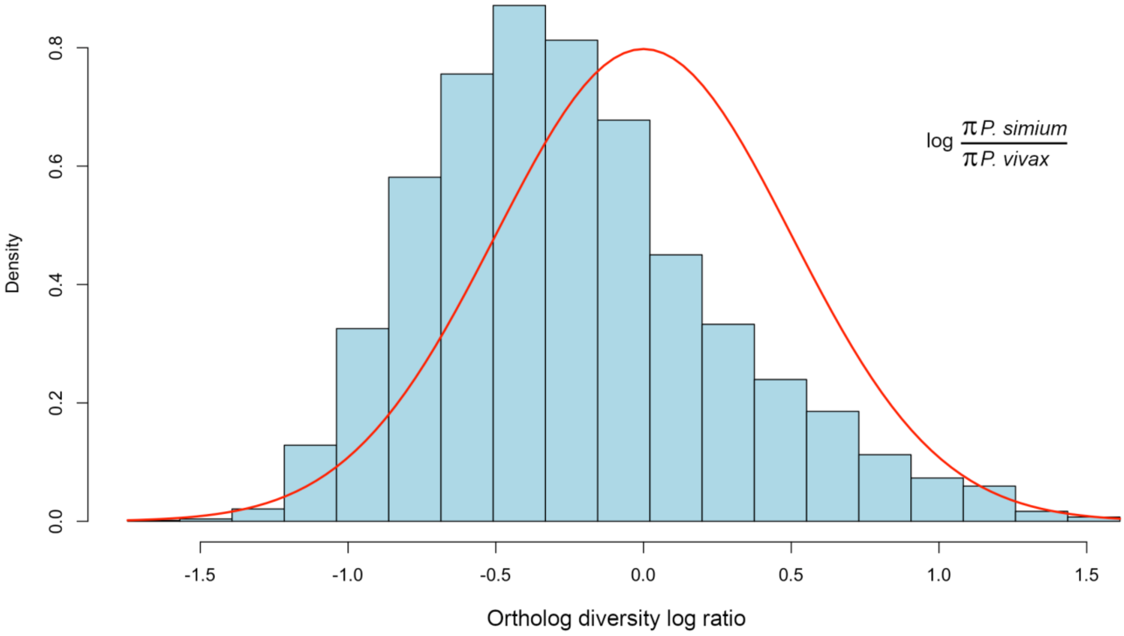


Figure 6. Relatedness network of *Plasmodium simium* samples. Samples were collected from humans and platyrrhine monkeys from the states of São Paulo, Rio de Janeiro, and Espírito Santo, southeastern Brazil (n = 31). Nodes represent individual samples that are coloured according to the state of origin; edges connect samples with mean pairwise ancestry sharing ≥ 0.25 (equivalent to half-siblings). Unconnected nodes indicate isolates that do not share at least 25% of their genome-wide ancestry with other isolates from the same or different states.

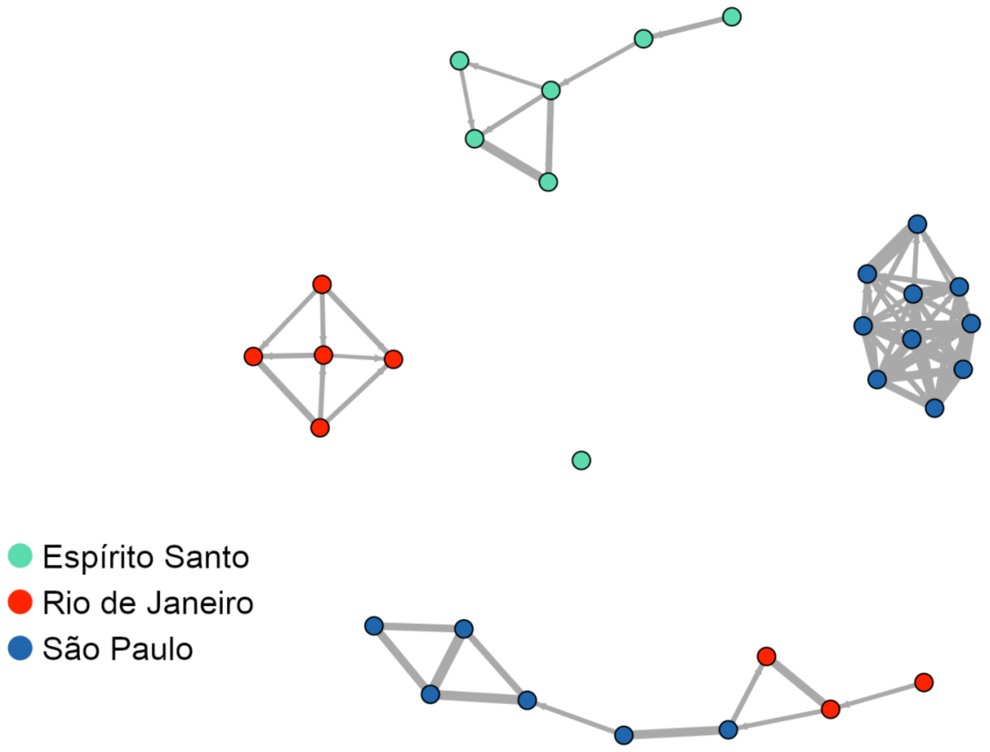
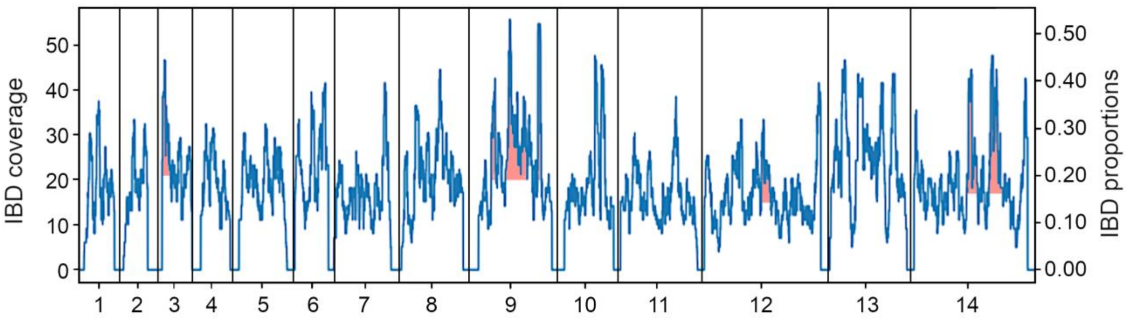


Figure 7. Genomic regions in *P. simium* under positive selection. Domains putatively under strong positive selection revealed by identity-by-descent (IBD) analysis of the *P. simium* genome. We display peaks of IBD coverage and proportion of shared ancestry along the 14 chromosomes. Red shading indicates validated peaks likely associated with selective sweeps (Guo *et al*., 2024).



Graphical Abstract: