

Population genomics of *Anopheles darlingi*, the principal South American malaria vector mosquito

Jacob A. Tennessen^{1,2*}, Raphael Brosula², Estelle Chabanol³, Sara Bickersmith⁴, Angela M. Early², Margaret Laws^{1,2}, Katrina A. Kelley^{1,2}, Maria Eugenia Grillet⁵, Dionicia Gamboa⁶, Eric R. Lucas⁷, Jean-Bernard Duchemin³, Martha L. Quiñones⁸, Maria Anice Mureb Sallum⁹, Eduardo S. Bergo¹⁰, Jorge E. Moreno¹¹, Sanjay Nagi⁷, Nicholas J. Arisco¹, Mohini Sooklall¹², Reza Niles-Robin¹², Marcia C. Castro¹, Horace Cox¹², Mathilde Gendrin³, Jan E. Conn^{4,13}, and Daniel E. Neafsey^{1,2}

Affiliations:

¹Harvard T.H. Chan School of Public Health; Boston, MA USA.

²Broad Institute; Cambridge, MA USA.

³Institut Pasteur de la Guyane; Cayenne, French Guiana.

⁴New York State Department of Health, Wadsworth Center; Albany, NY USA.

⁵Instituto de Zoología y Ecología Tropical, Facultad de Ciencias, Universidad Central de Venezuela; Caracas, Venezuela.

⁶Laboratorio de Malaria: Parásitos y Vectores, Laboratorios de Investigación y Desarrollo, Facultad de Ciencias e Ingeniería, Universidad Peruana Cayetano Heredia; Lima, Peru.

⁷Liverpool School of Tropical Medicine; Liverpool, UK.

⁸Universidad Nacional de Colombia; Bogotá, Colombia.

⁹Faculdade de Saúde Pública, Universidade de São Paulo; São Paulo, Brazil.

¹⁰Instituto Pasteur de São Paulo; São Paulo, Brazil.

¹¹Instituto de Altos Estudios Dr. Arnoldo Gabaldón, Centro de Investigaciones de Campo Francesco Vitanza; Bolivar, Venezuela.

¹²Vector Control Services, Ministry of Health; Georgetown, Guyana.

¹³Department of Biomedical Sciences, College of Integrated Health Sciences, State University of New York at Albany; Albany, NY USA.

*Corresponding author. Email: jtennessen@hsph.harvard.edu

ABSTRACT

Malaria in South America remains a serious public health problem. *Anopheles (Nyssorhynchus) darlingi* is the most important malaria vector across tropical Latin America. Vector-targeted disease control efforts require a thorough understanding of mosquito demographic and evolutionary patterns. We present and analyze whole genomes of 1094 *A. darlingi* (median depth 18x) from six South American countries. We observe deep geographic population structure, high genetic diversity including thirteen putative segregating inversions, and no evidence for cryptic sympatric taxa despite high interpopulation divergence. Strong signals of selection are plausibly driven by insecticides, especially on cytochrome P450 genes, one of which we validated experimentally. Our results will facilitate effective mosquito surveillance and control, while highlighting ongoing challenges that a diverse vector poses for malaria elimination in the western hemisphere.

There are approximately half a million malaria cases per year in the Americas (1), and the disease remains a stubbornly persistent health crisis in this region. Some locations have recently endured rapid upswings in case numbers (1, 2), such that case counts for the Americas overall have declined little over the past decade (1), dousing optimism for imminent malaria elimination in some countries. Furthermore, South America has been a hotspot for drug resistance emergence in the *Plasmodium* spp. parasites that cause malaria (3–5), threatening control efforts around the world. Successful elimination of neotropical malaria will require a greater understanding of its transmission determinants, including the population biology and adaptations of its most widespread and diverse mosquito vector, *Anopheles (Nyssorhynchus) darlingi*. This species is the primary vector throughout the Guiana Shield and Amazon Basin and an important vector across much of the rest of tropical Latin America, efficiently transmitting both *Plasmodium vivax* and *Plasmodium falciparum* (6, 7).

There is deep evolutionary divergence between malaria vectors in the neotropics relative to the rest of the world. The *Nyssorhynchus* subgenus, which contains *A. darlingi*, is confined to the neotropics (8), and is arguably a distinct genus (9). It diverged from the *Cellia* subgenus, which contains the principal vector of malaria in Africa (*Anopheles gambiae*), approximately 90 million years ago (10). Despite extensive population genomic analysis of malaria vectors in Africa and Asia (11–14), comparable work on neotropical anophelines has been restricted to relatively small population genetic datasets. In particular, it has remained unclear whether *A. darlingi* is an assemblage of cryptic, morphologically-identical species, as genomic datasets have recently elucidated for the *A. gambiae* complex (15), *A. minimus* complex (13), and *A. funestus* complex (16). Population genetic subdivisions have been observed in *A. darlingi* on either side of the Andes (17) and between the Amazon basin and southern Brazil (18, 19), however no genomic investigation has been conducted to determine whether sympatric cryptic species exist.

Sympatric cryptic species, first detected in African anophelines through partial sterility in laboratory crossing experiments, have complicated efforts to accurately measure insecticide resistance phenotypes and identify genotypic markers of insecticide resistance in African vectors since the advent of large-scale insecticidal control almost a century ago (20, 21). While other *Anopheles* species repeatedly evolve insecticide resistance through target site polymorphisms as well as metabolically-acting factors like cytochrome P450s (22–24), in *A. darlingi* insecticide resistance alleles have not been detected (25) and observations of phenotypic resistance have been rare and restricted to certain geographic regions and classes of insecticide (26–28).

Here we present a population genomic analysis of *A. darlingi*, an animal with one of the greatest impacts on public health in the Americas. In addition to describing overall patterns of genetic diversity, a major goal of this study has been to assess whether this species exhibits the characteristic evolutionary patterns of African and Asian *Anopheles* that have been particularly vexing to vector control goals, including partially interbreeding cryptic taxa, strikingly high nucleotide diversity partitioned into large segregating haplotype blocks, and insecticide resistance via target-site and metabolic mechanisms.

Pronounced geographical population structure without sympatric cryptic taxa

We generated whole-genome sequences for 1094 adult female samples of *A. darlingi* from 16 locations in six South American countries (Fig. 1A; tables S1, and S2). Our extensive

sampling represents most of the species' range with some exceptions (e.g. Central America and much of southern Brazil). Excluding eight samples with elevated heterozygosity indicative of contamination (fig. S1), and 62 samples with median coverage $\leq 1x$, there were 1024 usable samples (median coverage 20x). Collection localities were ecologically diverse (Fig. 1B). We detected DNA from multiple vertebrate species including humans and domesticated animals, presumably deriving from recent blood meals, as well as DNA from *P. falciparum* and *P. vivax*, which was significantly more common within samples positive for human DNA ($P < 0.01$; table S2). Rates of disease transmission will depend on how often mosquitoes bite humans relative to other vertebrates, in turn influenced by ecological factors and human behavior.

We observe genetic divergence across populations, as seen in other widespread South American insects (29–32), confirming mitochondrial observations (33). Ancestry analysis with ADMIXTURE predicted eight ancestral populations, closely corresponding to contemporary collection sites, with little evidence of admixture in most locales (Fig. 2A). Ancestry models with fewer than eight populations are suboptimal but indicate that the most fundamental splits ($K = 2$ or 3) are among an Amazon/Atlantic clade (French Guiana, Amapá, Guyana, Rio de Janeiro, Pará, Amazonas, Peru), an Orinoco clade (Sifontes, Atures, Guainía), and a Chocó clade west of the Andes (fig. S2). Principal component analysis also clusters samples in a manner consistent with geography, with the largest split separating Colombian and Venezuela mosquitoes from all others (Fig. 2B). Similar results are obtained for genomic partitions such as chromosome arms (fig. S3) and for geographical subsets of samples (fig. S4). A population phylogeny with TreeMix is consistent with these results: first the Chocó clade branches off, followed by Guainía/Venezuela (Orinoco clade), leaving an Amazon/Atlantic clade within which Amazonas/Peru splits from the remaining eastern populations (Fig. 2C). The Andes represent an intuitive and frequently documented barrier to gene flow in many regional species (31), consistent with the isolation of Chocó *A. darlingi* (17). More surprising is the distinction between the two main eastern clades based around drainage basins that are not separated by obvious physical barriers preventing gene flow: Orinoco versus Amazon/Atlantic. These clades split Sifontes, Venezuela from Mahdia, Guyana, separated by only 270 km of low-elevation suitable habitat.

There is no evidence of sympatric cryptic taxa: samples from the same site are always genetically homogeneous with closely shared ancestry (Fig. 2A,B) and do not depart systematically from Hardy-Weinberg equilibrium (fig. S5). However, further study will be needed to determine whether geographically distinct clades show any reproductive isolation.

Pairwise F_{ST} is high, especially between the major clades (median 0.18), exceeding F_{ST} values observed in *A. gambiae* and *A. coluzzii* (11, 12) (Fig. 2D, fig. S6), in which relatively low divergence ($F_{ST} < 0.01$) is typical between continental populations separated by vast physical distance up to 4500 km. F_{ST} between the *A. darlingi* major clades even surpasses F_{ST} between African species (< 0.15). While *A. gambiae* relies on high-altitude wind dispersal to facilitate gene flow (34), it is unknown whether *A. darlingi* can utilize a similar mechanism. Plausibly, long-range migration has been more adaptive in Africa than in South America due to increased aridity and seasonality of wetlands, leading to the distinct population genetic patterns on each continent.

The deep population structure we observe across South America has both positive and negative consequences for mosquito-focused interventions. Mitochondrial DNA has suggested the Amazon River as a barrier to gene flow in *A. darlingi* (18), but we observe

relatively high genetic affinity between Amapá and Pará on opposite sides of the Amazon Delta (Figs. 1 and 2), similar to the pattern seen in *A. aquasalis* across the same range (35), and mirroring the close kinship between *A. darlingi* populations separated by the Orinoco. Instead of rivers as barriers, we observe the opposite: boundaries between watersheds demarcate the major genetic divisions. Gene flow may be facilitated by waterways due to the dependence of *A. darlingi* on aquatic breeding sites. Interpopulation connectivity may also be enhanced by anthropogenic activity along rivers including deforestation leading to favorable disturbed habitat (32, 36) or transport in vehicles or boats. If a population were to be extirpated as reported in Suriname (37), there could be a long delay before the habitat is recolonized via migration from neighboring countries, unless enabled by human activity. Limited gene flow also indicates that a new drug resistance allele will not spread quickly beyond its population of origin. While we are unaware of plans to modify *A. darlingi* with gene drives as is being attempted on other species (38), our results indicate that a future gene drive implementation in these mosquitoes would likely require repeated launches in several geographic locations, potentially requiring regional coordination.

High genetic diversity partitioned into large haplotype blocks

Genetic diversity is extremely high, with heterozygosity >2% in most samples (Fig. 3A). Synonymous-site nucleotide diversity (π_s) is 3.2% on autosomes, comparable to the famously high nucleotide diversity of the *A. gambiae* complex (11), and 2.3% on the X (Fig. 3B), 2-fold higher than *A. gambiae*. X chromosome π is approximately 75% of autosomal π , reflecting the expected X/autosome effective population size ratio and suggesting that X diversity in *A. darlingi* is not impacted by the factors that reduce it further in *A. gambiae*, such as sex ratio skew.

Tajima's D in *A. darlingi* is positive and relatively close to zero (median per population 0.26, fig. S7), contrasting with the very negatively-skewed site frequency spectra of *A. gambiae* (Tajima's D < -1.5) attributable to explosive population growth facilitated by human-modified habitats (11). All *A. darlingi* clades show comparable genetic diversity and site frequency spectra, suggesting relative stability in long-term population size and persistence, with potentially slight decreases in abundance, especially in relatively low diversity populations like Chocó, Sifontes, and Pará (Fig. 3A). Thus, deforestation, urbanization, and other anthropogenic changes across Latin America have not resulted in an exponential increase in *A. darlingi* as similar changes have for African mosquitoes, likely because *A. darlingi* is less anthropophilic (39, 40). However, very recent expansions in response to anthropogenic ecological and demographic changes (36, 39, 41, 42) including mosquito-friendly habitats via open-pit mining in the Guyana shield (2) may not yet be detectable in population genetic data.

We detected thirteen large (3-19 Mb) segregating haplotype blocks — likely chromosomal inversions (43) — showing perfect linkage disequilibrium among physically distant polymorphisms (fig. S8, fig. S9; table S3). The two haplotype lineages (presumed opposite orientations) of an inversion typically show elevated sequence divergence from each other as well as substantial diversity within themselves (fig. S10). There are few examples of alternate inversion haplotypes being fixed across populations, except in lower-diversity populations like Pará in which genetic drift presumably plays a larger role (fig. S9). Heterozygosity varies considerably both among samples and across the genome (Fig. 3C), including copy number variants overlapping 6425 genes (table S4), which also vary by population (fig. S11).

Positive selection signals may reflect emerging insecticide resistance

Numerous signals of positive selection are apparent across these populations (tables S5, S6, and S7). Genome-wide F_{ST} scans between adjacent populations reveal the same 150 kb window (2RL_8.425-8.575Mb) within Inversion B to be an outlier, sometimes the top outlier, in multiple independent comparisons (Fig. 4A). G1 and G123 statistics for several populations are also outliers in this window (Fig. 4B), which overlaps 20 genes including a cluster of six cytochrome P450 genes (Fig. 4C). Even a single incidence of this selection signal would be striking, as in Peru (Fig. 4B) or Guainía (fig. S12A), yet this window is among the top outliers in approximately half our populations by at least one metric (Fig. 4D). In some populations, the selection signal is clearly concentrated on the P450 genes, especially *CYP6A1* (*ADAR2_008159*), tied to pyrethroid resistance in *A. funestus* (44) and *A. gambiae* (45). The top F_{ST} value between Peru and Amazonas is in *CYP6A1*, and missense variant *CYP6A1* T283K is one of the two highest F_{ST} values in this window between Cacao and other French Guiana populations. In Guainía, similarly high F_{ST} values occur across a wider window (2RL_8-9Mb) and many individuals exhibit near-zero heterozygosity across several megabases surrounding this genomic region (Fig. 3C; fig. S12A), although G1 peaks within the narrower window at 2RL_8.465Mb. Various derived alleles within Inversion B have been selected, rather than the entire inversion haplotype, and these alleles are not closely related across populations (fig. S12B), suggesting soft sweeps on distinct genetic backgrounds rather than a single species-wide hard sweep of a novel adaptive allele.

Given the well-established link between cytochrome P450 genes and toxin resistance (23, 44), we assessed whether the observed selection signals could be tied to land use practices near our collection sites that might expose mosquitoes to insecticides. Runoff of agricultural pesticides into Latin American aquifers (46) could be the main source of exposure, as hypothesized for resistance in African *Anopheles* (47) and possibly other non-targets of pesticides (48). We ranked the 2RL_8.425-8.575Mb window against all autosomal 150 kb windows based on F_{ST} and G1 to generate empirical P values (Fig. 4D). The lowest empirical P values (< 0.002) occurred for F_{ST} in Peru, Cacao, and Guainía, and two of these (Peru and Guainía) also had among the lowest empirical P values for G1 (< 0.01). Notably, all three of these locations, unlike most of our collection sites, have substantial land use dedicated to “mosaic of farming uses” (Fig. 1B), although there is no significant correlation between proportion of farming land use and selection signals ($P > 0.1$).

To test if selected alleles impact insecticide resistance, we collected sixteen wild *A. darlingi* from Cacao, French Guiana and tested their resistance in a bottle assay with 12.5 $\mu\text{g/mL}$ deltamethrin (Fig. 4E). We genotyped these mosquitoes at *CYP6A1* T283K and found knockdown time of TK heterozygotes (median = 21 minutes) was significantly longer than that of either ancestral TT (median = 13.5 minutes) or derived KK (median = 13.5 minutes) homozygotes ($P < 0.05$). There was no significant phenotypic difference between homozygotes ($P > 0.1$). We note that our sample size was small with only two KK homozygotes, the overdominant genotypic effect on resistance is not straightforward to interpret, our results merely identify T283K as linked to the phenotype rather than causal, and mosquitoes of all genotypes were killed within an hour.

Other regions of the genome show selection signals beyond this window, though not as strong or as consistently across populations (figs. S13, S14, and S15; tables S5, S6, and S7). F_{ST} for Cacao versus other French Guiana was highest in a different cluster of P450 genes at

2RL_79.3 Mb (fig. S15A), peaking at *CYP6M2* (*ADAR2_008623*) tied to pyrethroid-resistance in *A. gambiae* (49). This double P450 signal strengthens our conclusion that the P450s themselves, and not unrelated linked genes, are the targets of selection. A window at 2RL_93.635-93.785 Mb, containing a set of seven ATP-binding cassette sub-family G genes, is a top F_{ST} outlier in two independent pairwise comparisons: between French Guiana sites and between Amazonas, Brazil sites (fig. S13; table S5). A strong selection signal centered on 2RL_49.950-50.050 Mb is implicated by F_{ST} , G1, and G123 in Barcelos (fig. S14). The top genome-wide F_{ST} for this comparison occurs in *Keap1* (*ADAR2_006983*), which also shows signals of positive selection in *A. gambiae* (50) and *A. arabiensis* (51), plausibly because of its role in a metabolic insecticide resistance pathway (22).

We do not observe notable positive selection at the orthologs of target-site resistance genes known in other *Anopheles*. In fact, target codons L995 in *vgsc* (site of the *kdr* mutation), A296 in *rdl*, and G280 in *ace1* (24) exhibit no nonsynonymous variation (fig. S15B), consistent with previous results (25). P450s and *Keap1* underlie metabolic resistance, suggesting that this mechanism has been favored in *A. darlingi* over target-site mechanisms of resistance. While the specific adaptive polymorphisms remain unidentified, it is notable that there is very little evidence that gene duplications of P450 or *Keap1* are selection targets, as hypothesized for *A. gambiae* (45) and *A. funestus* (52). Copy number variants in *CYP6AA1*, *CYP6M2*, and adjacent P450 genes (table S4; fig. S11A) are rare and not associated with the selected alleles. Some populations with selection signals had no or very few P450 duplications, and these were uncommon in samples homozygous for the derived allele at high- F_{ST} variants. We observed no *Keap1* duplicates.

Our results have several implications for vector control across Latin America. *A. darlingi* is extremely diverse and likely harbors latent potential to adapt to almost any novel intervention designed to kill it. Signals of positive selection suggest that this may already be occurring and indicate that broad continuing assays for emerging insecticide resistance across the continent are warranted. Although long-distance gene flow is limited, that may not matter as resistance can evolve in parallel on different genetic backgrounds across habitats, as has seemingly occurred with P450s. Continued population genomic monitoring can inform both disease control efforts and broad conclusions about ecological and evolutionary diversity in neotropical vectors.

References and Notes

1. World Health Organization (WHO), “World Malaria Report 2024” (2024); <https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2024>.
2. M. E. Grillet, J. E. Moreno, J. V. Hernández-Villena, M. F. Vincenti-González, O. Noya, A. Tami, A. Paniz-Mondolfi, M. Llewellyn, R. Lowe, A. A. Escalante, J. E. Conn, Malaria in Southern Venezuela: The hottest hotspot in Latin America. *PLoS Negl Trop Dis.* **15**, e0008211 (2021).
3. J. C. Wootton, X. Feng, M. T. Ferdig, R. A. Cooper, J. Mu, D. I. Baruch, A. J. Magill, X. Z. Su, Genetic diversity and chloroquine selective sweeps in *Plasmodium falciparum*. *Nature* **418**, 320-323 (2002).

4. S. M. Chenet, S. Akinyi Okoth, C. S. Huber, J. Chandrabose, N. W. Lucchi, E. Talundzic, K. Krishnalall, N. Ceron, L. Musset, A. Macedo de Oliveira, M. Venkatesan, R. Rahman, J. W. Barnwell, V. Udhayakumar, Independent emergence of the *Plasmodium falciparum* kelch propeller domain mutant allele C580Y in Guyana. *J. Infect. Dis.* **213**, 1472-1475 (2016).
5. L. C. Mathieu, H. Cox, A. M. Early, S. Mok, Y. Lazrek, J. C. Paquet, M. P. Ade, N. W. Lucchi, Q. Grant, V. Udhayakumar, J. S. Alexandre, M. Demar, P. Ringwald, D. E. Neafsey, D. A. Fidock, L. Musset, Local emergence in Amazonia of *Plasmodium falciparum* k13 C580Y mutants associated with in vitro artemisinin resistance. *Elife* **9**, e51015 (2020).
6. J. D. Charlwood, Biological variation in *Anopheles darlingi* Root. *Mem. Inst. Oswaldo Cruz* **91**, 391-398 (1996).
7. B. C. Carlos, L. D. P. Rona, G. K. Christophides, J. A. Souza-Neto, A comprehensive analysis of malaria transmission in Brazil. *Pathog Glob Health.* **113**, 1-13 (2019).
8. R. E. Harbach, "The phylogeny and classification of *Anopheles*" in *Anopheles Mosquitoes—New Insights into Malaria Vectors*, S. Manguin, Ed. (Rijeka, Croat.: InTech, 2013), pp. 3-55.
9. P. G. Foster, T. M. P. de Oliveira, E. S. Bergo, J. E. Conn, D. C. Sant'Ana, S. S. Nagaki, S. Nihei, C. E. Lamas, C. González, C. C. Moreira, M. A. M. Sallum, Phylogeny of Anophelinae using mitochondrial protein coding genes. *R Soc Open Sci.* **4**, 170758 (2017).
10. L. A. Freitas, C. A. Russo, C. M. Voloch, O. C. Mutaquiha, L. P. Marques, C. G. Schrago, Diversification of the genus *Anopheles* and a neotropical clade from the late Cretaceous. *PLoS One* **10**, e0134462 (2015).
11. *Anopheles gambiae* 1000 Genomes Consortium, Genetic diversity of the African malaria vector *Anopheles gambiae*. *Nature* **552**, 96-100 (2017).
12. *Anopheles gambiae* 1000 Genomes Consortium, Genome variation and population structure among 1142 mosquitoes of the African malaria vector species *Anopheles gambiae* and *Anopheles coluzzii*. *Genome Res.* **30**, 1533-1546 (2020).
13. B. St Laurent, N. Harding, N. Deason, K. Oy, C. Sok Loeun, M. Sary, R. Sunly, S. Nhep, E. Drury, K. Rockett, S. Sovannaroeth, S. Goncalves, D. Kwiatkowski, A. Miles, Population genomics reveal distinct and diverging populations of *An. minimus* in Cambodia. *Commun Biol.* **5**, 1308 (2022).
14. M. Boddé, J. Nwezeobi, P. Korlević, A. Makunin, O. Akone-Ella, S. Barasa, M. Gadji, L. Hart, E. W. Kaindoa, K. Love, E. R. Lucas, I. Lujumba, M. Máquina, S. Nagi, J. O. Otero, B. Polo, C. Sangbakembi, S. Dadzie, L. L. Koekemoer, D. Kwiatkowski, E. McAlister, E. Ochomo, F. Okumu, K. Paaijmans, D. P. Tchouassi, C. S. Wondji, D. Ayala, R. Durbin, A. Miles, M. K. N. Lawniczak, Genomic diversity of the African malaria vector *Anopheles funestus*. *bioRxiv* 2024.12.14.628470 (2024).

15. M. C. Fontaine, J. B. Pease, A. Steele, R. M. Waterhouse, D. E. Neafsey, I. V. Sharakhov, X. Jiang, A. B. Hall, F. Catteruccia, E. Kakani, S. N. Mitchell, Y. C. Wu, H. A. Smith, R. R. Love, M. K. Lawniczak, M. N. Slotman, S. J. Emrich, M. W. Hahn, N. J. Besansky NJ, Extensive introgression in a malaria vector species complex revealed by phylogenomics. *Science* **347**, 1258524 (2015).
16. S. T. Small, F. Labbé, N. F. Lobo, L. L. Koekemoer, C. H. Sikaala, D. E. Neafsey, M. W. Hahn, M. C. Fontaine, N. J. Besansky, Radiation with reticulation marks the origin of a major malaria vector. *Proc. Natl. Acad. Sci. U S A.* **117**, 31583-31590 (2020).
17. N. Naranjo-Díaz, J. E. Conn, M. M. Correa, Behavior and population structure of *Anopheles darlingi* in Colombia. *Infect. Genet. Evol.* **39**, 64-73 (2016).
18. P. M. Pedro, M. A. M. Sallum, Spatial expansion and population structure of the neotropical malaria vector, *Anopheles darlingi* (Diptera: Culicidae). *Biol. J. Linn. Soc.* **97**, 854-866 (2009).
19. K. J. Emerson, J. E. Conn, E. S. Bergo, M. A. Randel, M. A. M. Sallum, Brazilian *Anopheles darlingi* Root (Diptera: Culicidae) clusters by major biogeographical region. *PLoS One* **10**, e0130773 (2015).
20. G. Davidson, Insecticide resistance in *Anopheles gambiae* Giles: a case of simple mendelian inheritance. *Nature* **178**, 863-864 (1956).
21. G. Davidson, *Anopheles gambiae*, a complex of species. *Bull World Health Organ.* **31**, 625-634 (1964).
22. V. A. Ingham, P. Pignatelli, J. D. Moore, S. Wagstaff, H. Ranson, The transcription factor Maf-S regulates metabolic resistance to insecticides in the malaria vector *Anopheles gambiae*. *BMC Genomics* **18**, 669 (2017).
23. J. Vontas, E. Katsavou, K. Mavridis, Cytochrome P450-based metabolic insecticide resistance in *Anopheles* and *Aedes* mosquito vectors: Muddying the waters. *Pestic. Biochem. Physiol.* **170**, 104666 (2020).
24. M. Campos, J. Phelan, A. Spadar, E. Collins, A. Gonçalves, B. Pelloquin, N. M. Vaselli, A. Meiwald, E. Clark, C. Stica, J. Orsborne, M. Sylla, C. Edi, D. Camara, A. R. Mohammed, Y. A. Afrane, M. Kristan, T. Walker, L. F. Gomez, L. A. Messenger, T. G. Clark, S. Campino, High-throughput barcoding method for the genetic surveillance of insecticide resistance and species identification in *Anopheles gambiae* complex malaria vectors. *Sci. Rep.* **12**, 13893 (2022).
25. S. A. Bickersmith, J. D. Jurczynski, M. A. M. Sallum, L. S. M. Chaves, E. S. Bergo, G. A. D. Rodriguez, C. A. Morante, C. T. Rios, M. P. Saavedra, F. Alava, D. Gamboa, J. M. Vinetz, J. E. Conn, Mutations linked to insecticide resistance not detected in the *Ace-1* or *VGSC* genes in *Nyssorhynchus darlingi* from multiple localities in Amazonian Brazil and Peru. *Genes* **14**, 1892 (2023).

26. M. F. Suarez, M. L. Quiñones, J. D. Palacios, A. Carrillo, First record of DDT resistance in *Anopheles darlingi*. *J. Am. Mosq. Control Assoc.* **6**, 72-74 (1990). .
27. L. I. Orjuela, J. A. Morales, M. L. Ahumada, J. F. Rios, J. J. González, J. Yañez, A. Rosales, D. M. Cabarcas, J. Venegas, M. F. Yasnot, M. L. Quiñones, Insecticide resistance and its intensity in populations of malaria vectors in Colombia. *Biomed Res. Int.* **2018**, 9163543 (2018).
28. S. Vezenegho, R. Carinci, J. Issaly, C. Nguyen, P. Gaborit, L. Ferraro, G. Lacour, E. Mosnier, V. Pommier de Santi, Y. Epelboin, R. Girod, S. Briolant, I. Dusfour, Variation in pyrethroid resistance phenotypes in *Anopheles darlingi* in an area with residual malaria transmission: warning of suspected resistance in French Guiana. *Am. J. Trop. Med. Hyg.* **108**, 424-427 (2022).
29. P. M. Pedro, A. Uezu, M. A. Sallum, Concordant phylogeographies of 2 malaria vectors attest to common spatial and demographic histories. *J. Hered.* **101**, 618-627 (2010).
30. K. M. Lewald, C. A. Tabuloc, K. E. Godfrey, J. Arnó, C. R. Perini, J. C. Guedes, J. C. Chiu, Genome assembly and population sequencing reveal three populations and signatures of insecticide resistance of *Tuta absoluta* in Latin America. *Genome Biol. Evol.* **15**, evad060 (2023).
31. A. C. Turchetto-Zolet, F. Pinheiro, F. Salgueiro, C. Palma-Silva, Phylogeographical patterns shed light on evolutionary process in South America. *Mol. Ecol.* **22**, 1193-1213 (2013).
32. C. C. Ribas, A. O. Sawakuchi, R. P. de Almeida, F. N. Pupim, F.N., M. A. Rego, R. Batista, L. L. Knowles, The role of rivers in the origin and future of Amazonian biodiversity. *Nat. Reviews Biodiversity* **1**, 14–31 (2025).
33. J. E. Conn, M. G. Rosa-Freitas, S. L. Luz, H. Momen, Molecular population genetics of the primary neotropical malaria vector *Anopheles darlingi* using mtDNA. *J. Am. Mosq. Control Assoc.* **15**, 468-474 (1999).
34. D. L. Huestis, A. Dao, M. Diallo, Z. L. Sanogo, D. Samake, A. S. Yaro, Y. Ousman, Y. M. Linton, A. Krishna, L. Veru, B. J. Krajacich, R. Faiman, J. Florio, J. W. Chapman, D. R. Reynolds, D. Weetman, R. Mitchell, M. J. Donnelly, E. Talamas, L. Chamorro, E. Strobach, T. Lehmann, Windborne long-distance migration of malaria mosquitoes in the Sahel. *Nature* **574**, 404-408 (2019).
35. T. L. Fairley, M. M. Póvoa, J. E. Conn, Evaluation of the Amazon River delta as a barrier to gene flow for the regional malaria vector, *Anopheles aquasalis* (Diptera: Culicidae) in northeastern Brazil. *J. Med. Entomol.* **39**, 861-869 (2002).
36. A. Y. Vittor, W. Pan, R. H. Gilman, J. Tielsch, G. Glass, T. Shields, W. Sánchez-Lozano, V. V. Pinedo, E. Salas-Cobos, S. Flores, J. A. Patz, Linking deforestation to malaria in the Amazon: characterization of the breeding habitat of the principal malaria vector, *Anopheles darlingi*. *Am. J. Trop. Med. Hyg.* **81**, 5-12 (2009).

37. H. Hiwat, S. Mitro, A. Samjhawan, P. Sardjoe, T. Soekhoe, W. Takken, Collapse of *Anopheles darlingi* populations in Suriname after introduction of insecticide-treated nets (ITNs); malaria down to near elimination level. *Am. J. Trop. Med. Hyg.* **86**, 649-655 (2012).
38. K. Kyrou, A. M. Hammond, R. Galizi, N. Kranjc, A. Burt, A. K. Beaghton, T. Nolan, A. Crisanti, A CRISPR-Cas9 gene drive targeting doublesex causes complete population suppression in caged *Anopheles gambiae* mosquitoes. *Nat. Biotechnol.* **36**, 1062-1066 (2018).
39. L. Mirabello, J. E. Conn, Molecular population genetics of the malaria vector *Anopheles darlingi* in Central and South America. *Heredity* **96**, 311-321 (2006).
40. M. Moreno, M. P. Saavedra, S. A. Bickersmith, C. Prussing, A. Michalski, C. Tong Rios, J. M. Vinetz, J. E. Conn, Intensive trapping of blood-fed *Anopheles darlingi* in Amazonian Peru reveals unexpectedly high proportions of avian blood-meals. *PLoS Negl. Trop. Dis.* **11**, e0005337 (2017).
41. G. B. Schoeler, C. Flores-Mendoza, R. Fernández, J. R. Davila, M. Zyzak, Geographical distribution of *Anopheles darlingi* in the Amazon Basin region of Peru. *J. Am. Mosq. Control Assoc.* **19**, 286-296 (2003).
42. W. Lainhart, S. A. Bickersmith, K. J. Nadler, M. Moreno, M. P. Saavedra, V. M. Chu, P. E. Ribolla, J. M. Vinetz, J. E. Conn, Evidence for temporal population replacement and the signature of ecological adaptation in a major Neotropical malaria vector in Amazonian Peru. *Malar. J.* **14**, 375 (2015).
43. A. J. Cornel, K. K. Brisco, W. P. Tadei, N. F. Secundino, M. S. Rafael, A. K. Galardo, J. F. Medeiros, F. A. Pessoa, C. M. Ríos-Velásquez, Y. Lee, P. F. Pimenta, G. C. Lanzaro, *Anopheles darlingi* polytene chromosomes: revised maps including newly described inversions and evidence for population structure in Manaus. *Mem. Inst. Oswaldo Cruz* **111**, 335-346 (2016).
44. S. S. Ibrahim, N. Amvongo-Adjia, M. J. Wondji, H. Irving, J. M. Riveron, C. S. Wondji, Pyrethroid resistance in the major malaria vector *Anopheles funestus* is exacerbated by overexpression and overactivity of the P450 *CYP6A1* across Africa. *Genes* **9**, 140 (2018).
45. H. Njoroge, A. Van't Hof, A. Oruni, D. Pipini, S. C. Nagi, A. Lynd, E. R. Lucas, S. Tomlinson, X. Grau-Bove, D. McDermott, F. T. Wat'senga, E. Z. Manzambi, F. R. Agossa, A. Mokuba, S. Irish, B. Kabula, C. Mbogo, J. Bargul, M. J. I. Paine, D. Weetman, M. J. Donnelly, Identification of a rapidly-spreading triple mutant for high-level metabolic insecticide resistance in *Anopheles gambiae* provides a real-time molecular diagnostic for antimalarial intervention deployment. *Mol. Ecol.* **31**, 4307-4318 (2022).
46. S. I. Grondona, M. L. Lima, H. E. Massone, K. S. B. Miglioranza, Pesticides in aquifers from Latin America and the Caribbean. *Sci. Total Environ.* **901**, 165992 (2023).

47. F. A. Kouadio, N. C. Wipf, A. S. Nygble, B. K. Fodjo, C. G. Sadia, J. Vontas, K. Mavridis, P. Müller, C. S. Mouhamadou, Relationship between insecticide resistance profiles in *Anopheles gambiae* sensu lato and agricultural practices in Côte d'Ivoire. *Parasit. Vectors* **16**, 270 (2023).
48. N. F. Wan, L. Fu, M. Dainese, L. P. Kiær, Y. Q. Hu, F. Xin, D. Goulson, B. A. Woodcock, A. J. Vanbergen, D. J. Spurgeon, S. Shen, C. Scherber, Pesticides have negative effects on non-target organisms. *Nat. Commun.* **16**, 1360 (2025).
49. R. F. Djouaka, A. A. Bakare, O. N. Coulibaly, M. C. Akogbeto, H. Ranson, J. Hemingway, C. Strode, Expression of the cytochrome P450s, *CYP6P3* and *CYP6M2* are significantly elevated in multiple pyrethroid resistant populations of *Anopheles gambiae* s.s. from Southern Benin and Nigeria. *BMC Genomics* **9**, 538. 10.1186/1471-2164-9-538 (2008).
50. M. Kientega, C. S. Clarkson, N. Traoré, T. J. Hui, S. O'Loughlin, A. A. Millogo, P. S. Epopa, F. A. Yao, A. M. G. Belem, J. Brenas, A. Miles, A. Burt, A. Diabaté, Whole-genome sequencing of major malaria vectors reveals the evolution of new insecticide resistance variants in a longitudinal study in Burkina Faso. *Malar. J.* **23**, 280 (2024).
51. E. R. Lucas, S. C. Nagi, B. Kabula, B. Batengana, W. Kisinza, A. Egyir-Yawson, J. Essandoh, S. Dadzie, J. Chabi, A. E. Van't Hof, E. J. Rippon, D. Pipini, N. J. Harding, N. A. Dyer, C. S. Clarkson, A. Miles, D. Weetman, M. J. Donnelly, Copy number variants underlie major selective sweeps in insecticide resistance genes in *Anopheles arabiensis*. *PLoS Biol.* **22**, e3002898 (2024).
52. N. M. T. Tatchou-Nebangwa, L. M. J. Mugenzi, A. Muhammad, D. N. Nebangwa, M. F. M. Kouamo, C. S. D. Tagne, T. A. Tekoh, M. Tchouakui, S. M. Ghogomu, S. S. Ibrahim, C. S. Wondji, Two highly selected mutations in the tandemly duplicated *CYP6P4a* and *CYP6P4b* genes drive pyrethroid resistance in *Anopheles funestus* in West Africa. *BMC Biol.* **22**, 286 (2024).
53. E. Chabanol, O. Romoli, S. Talaga, Y. Epelboin, K. Heu, G. Prévot, M. Gendrin, Using capillary electrophoresis to identify Anopheline species in routine sampling sites. *Ecol. Evol.* **14**, e10782 (2024).
54. A. Makunin, P. Korlević, N. Park, S. Goodwin, R. M. Waterhouse, K. von Wyszetzki, C. G. Jacob, R. Davies, D. Kwiatkowski, B. St Laurent, D. Ayala, M. K. N. Lawniczak, A targeted amplicon sequencing panel to simultaneously identify mosquito species and *Plasmodium* presence across the entire *Anopheles* genus. *Mol. Ecol. Resour.* **22**, 28-44 (2022).
55. M. Vasimuddin, S. Misra, H. Li, S. Aluru, "Efficient architecture-aware acceleration of BWA-MEM for multicore systems." in 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS, 2019), pp. 314-324.
56. P. Danecek, J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies, H. Li, Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).

57. A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, M. A. DePristo, The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297-1303 (2010).
58. R. Poplin, V. Ruano-Rubio, M. A. DePristo, T. J. Fennell, M. O. Carneiro, G. A. Van der Auwera, D. E. Kling, L. D. Gauthier, A. Levy-Moonshine, D. Roazen, K. Shakir, J. Thibault, S. Chandran, C. Whelan, M. Lek, S. Gabriel, M. J. Daly, B. Neale, D. G. MacArthur, E. Banks, Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv* 201178 (2018).
59. R Core Team, R: A Language and Environment for Statistical Computing, version 4.3.3, R Foundation for Statistical Computing (2024); <https://www.R-project.org>.
60. D. H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655-1664 (2009).
61. J. K. Pickrell, J. K. Pritchard, Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics* **8**, e1002967 (2012).
62. G. I. Giraldo-Calderón, S. J. Emrich, R. M. MacCallum, G., Maslen, E. Dialynas, P. Topalis, N. Ho, S. Gesing, VectorBase Consortium, G. Madey, F. H. Collins, D. Lawson, VectorBase: an updated bioinformatics resource for invertebrate vectors and other organisms related with human diseases. *Nucleic Acids Res.* **43**, D707-13 (2015).
63. A. Stamatakis, RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313 (2014).
64. J. E. Stajich, M. W. Hahn, Disentangling the effects of demography and selection in human history. *Mol. Biol. Evol.* **22**, 63-73 (2005).
65. J. A. Tennessen, M. T. Duraisingh, Three signatures of adaptive polymorphism exemplified by malaria-associated genes. *Mol. Biol. Evol.* **38**, 1356-1371 (2021).
66. B. S. Weir, C. C. Cockerham, Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370 (1984).
67. G. Bhatia, N. Patterson, S. Sankararaman, A. L. Price, Estimating and interpreting F_{ST} : the impact of rare variants. *Genome Res.* **23**, 1514-1521 (2013).
68. A. M. Harris, N. R. Garud, and M. DeGiorgio, Detection and classification of hard and soft sweeps from unphased genotypes by multilocus genotype identity. *Genetics* **210**, 1429-1452 (2018).
69. E. R. Lucas, A. Miles, A. N. J. Harding, C. S. Clarkson, M. K. N. Lawniczak, D. P. Kwiatkowski, D. Weetman, M. J. Donnelly, *Anopheles gambiae* 1000 Genomes Consortium,

Whole-genome sequencing reveals high complexity of copy number variation at insecticide resistance loci in malaria mosquitoes. *Genome Res.* **29**, 1250-1261 (2019).

70. C. M. Souza, J. Z. Shimbo, M. R. Rosa, L. L. Parente, A. A. Alencar, B. F. T. Rudorff, H. Hasenack, M. Matsumoto, L. G. Ferreira, P. W. M. Souza-Filho, S. W. de Oliveira, W. F. Rocha, A. V. Fonseca, C. B. Marques, C. G. Diniz, D. Costa, D. Monteiro, E. R. Rosa, E. Vélez-Martin, E. J. Weber, F. E. B. Lenti, F. F. Paternost, F. G. C. Pareyn, J. V. Siqueira, J. L. Viera, L. C. Ferreira Neto, M. M. Saraiva, M. H. Sales, M. P. G. Salgado, R. Vasconcelos, S. Galano, V. V. Mesquita, T. Azevedo, Reconstructing three decades of land use and land cover changes in Brazilian biomes with Landsat archive and Earth Engine. *Remote Sens.* **12**, 2735 (2020).

Acknowledgments: We thank Nora Besansky, Romuald Carinci, Auden Cote-L’Heureux, Carlo Maria De Marco, Tatiane M. P. de Oliveira, Silvia Díaz, Elder Figueira, Pascal Gaborit, Zachary Johnson, Cheyenne Knox, Mara Lawniczak, Tovi Lehmann, Martin Llewellyn, Aleksei Makunin, Nelson Moncada, Caio C. Moreira, Monique A. Mota, Maycon Neves, Rachel Newmiller, Ruchit Panchal, Seth Rudman, Philipp Schwabl, Meg Shieh, Nithya Swaminathan, Stanislas Talaga, Ana Villarroel, and malaria control staff and local residents at various collection locations.

Funding:

National Institutes of Health grant U19AI110818 (DEN)
National Institutes of Health grant R01AI110112 (JEC, DEN)
Bill & Melinda Gates Foundation grant INV-009416 (DEN)
Agence Nationale de la Recherche grant ANR-18-CE15-0007 (MG)
National Council for Scientific and Technological Development (CNPq) grant 303382/2022-8 (MAMS)

Author contributions:

Conceptualization: JEC, DEN
Data curation: JAT, RB, AME
Formal analysis: JAT, NJA, ERL, MCC
Funding acquisition: MAMS, MG, JEC, DEN
Investigation: KAK, MAMS, JEC
Methodology: JAT, DEN
Project administration: ML, DEN
Software: SN
Resources: EC, SB, DG, ESB, JEM, MS, RN-R, MAMS, MEG, MLQ, HC, MG
Supervision: AME, DEN
Validation: J-BD

Visualization: JAT

Writing – original draft: JAT

Writing – review & editing: JAT, MEG, MAMS, JEC, DEN

Competing interests: Authors declare that they have no competing interests.

5 **Data and materials availability:** All raw sequencing data has been deposited at NCBI SRA, BioProject PRJNA1169887. Genotypes are accessible at Malaria Genome Vector Observatory. Transfer of samples was made possible by a material transfer agreement between Universidade de São Paulo and Harvard: MTA No 1017829, SISGEN Registration no. RCD3CF5, Export permit IBAMA no. 24BR0490030/DF.

10 **Supplementary Materials**

Materials and Methods

Figs. S1 to S15

Tables S1 to S7

References (53–70)

15

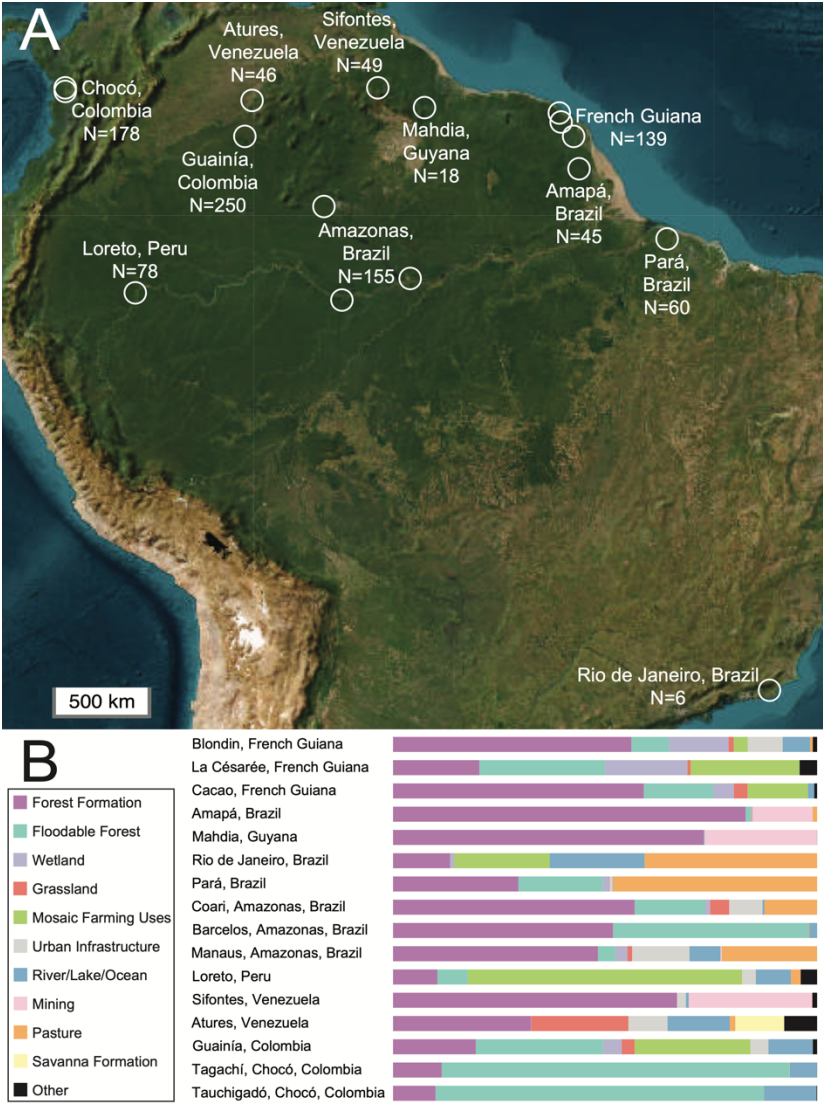


Fig. 1. Collection sites. (A) Collection localities with counts of usable samples in six countries. **(B)** Proportion of land cover and land use type in the surrounding 5 km area for all collection sites.

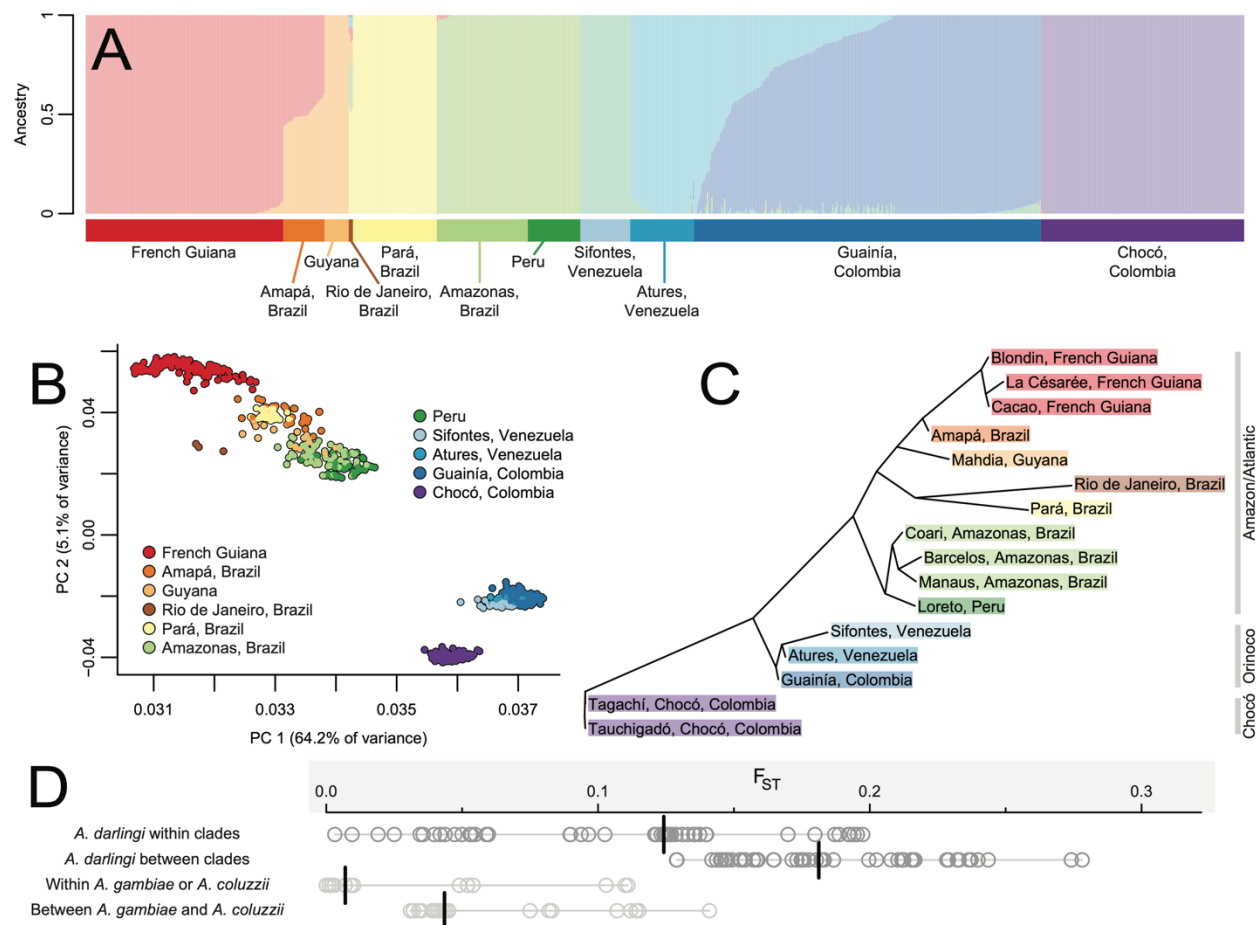


Fig. 2. Population structure. (A) Ancestry assignment via ADMIXTURE identifies eight ancestral populations, closely corresponding to the contemporary populations of French Guiana, Guyana, Pará, Amazonas/Peru, Sifontes, Atures, Guainía, and Chocó. Only Amapá, Rio de Janeiro, and Guainía show substantial admixture. (B) Principal component analysis separates samples geographically. (C) Sampling site phylogeny generated with TreeMix, showing three major clades: Amazon/Atlantic, Orinoco, and Chocó. (D) Pairwise F_{ST} among fifteen *A. darlingi* collection sites (Rio de Janeiro excluded for sample size), classified as within major clades (from part C) (e.g. Guyana vs. Peru) or between clades (e.g. Guyana vs. Sifontes). Within- and between-species values for African *A. gambiae* and *A. coluzzii* shown for comparison. For all four sets of points across both continents, median geographic distance is similarly large (1000-2000 km). Black vertical lines show medians.

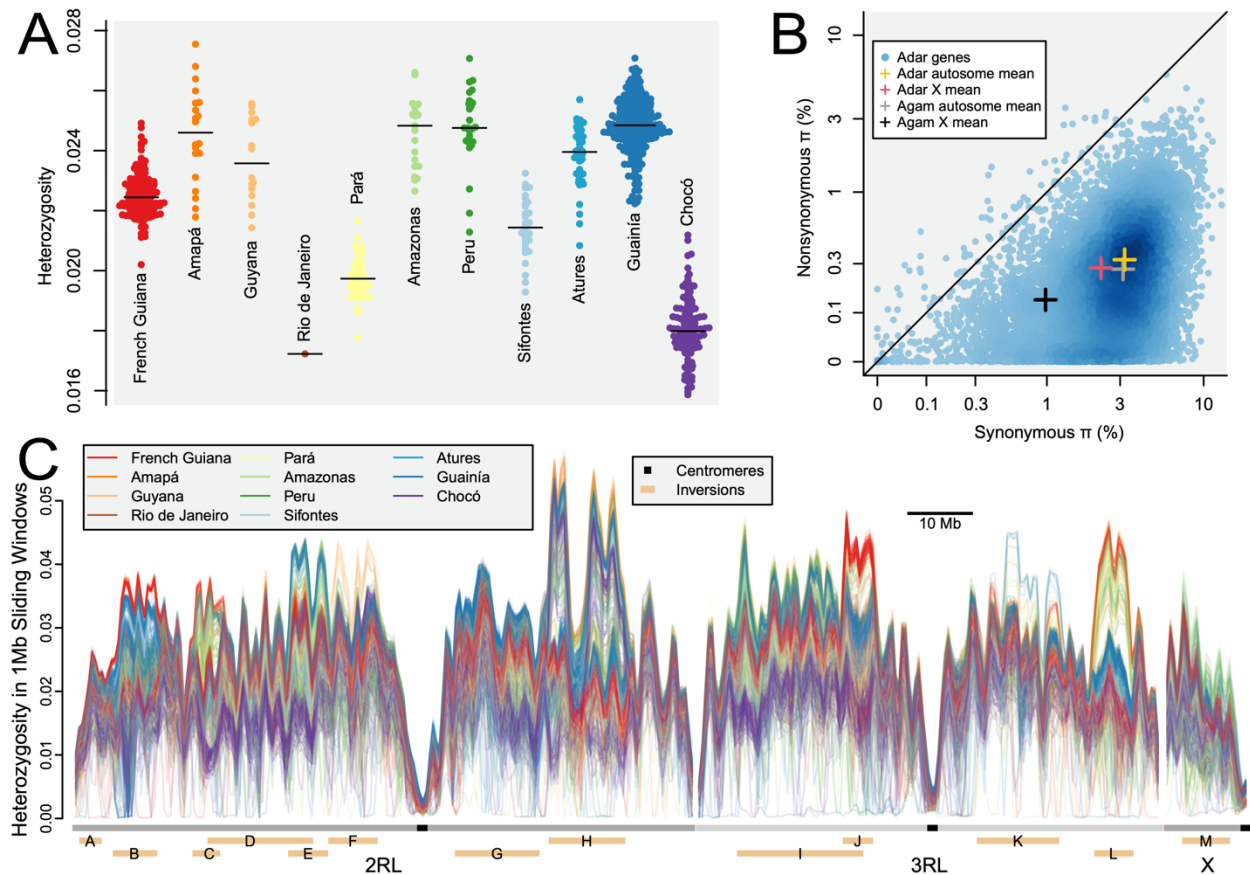


Fig. 3. Genetic diversity. (A) Median heterozygosity per sample, arranged by population. (B) Synonymous and nonsynonymous nucleotide diversity (π) per gene. Dotted lines = mean values. Solid line = π_N/π_S of 1. Crosses show mean values for X and autosomal genes, both for *A. darlingi* (Adar) and for the *A. gambiae* complex (Agam) as a comparison. (C) Heterozygosity per sample in sliding 1 Mb windows. Samples colored by population. Tan rectangles (bottom) indicate inversions, designated by capital letters. Peaks of elevated heterozygosity in some individuals typically overlap inversions and represent heterozygotes for the two haplotype lineages, with heterozygosity often up to two-fold higher than homozygotes. Occasionally, otherwise diverse individuals have regions of near-zero heterozygosity for over a megabase, indicating homozygosity along a long haplotype.

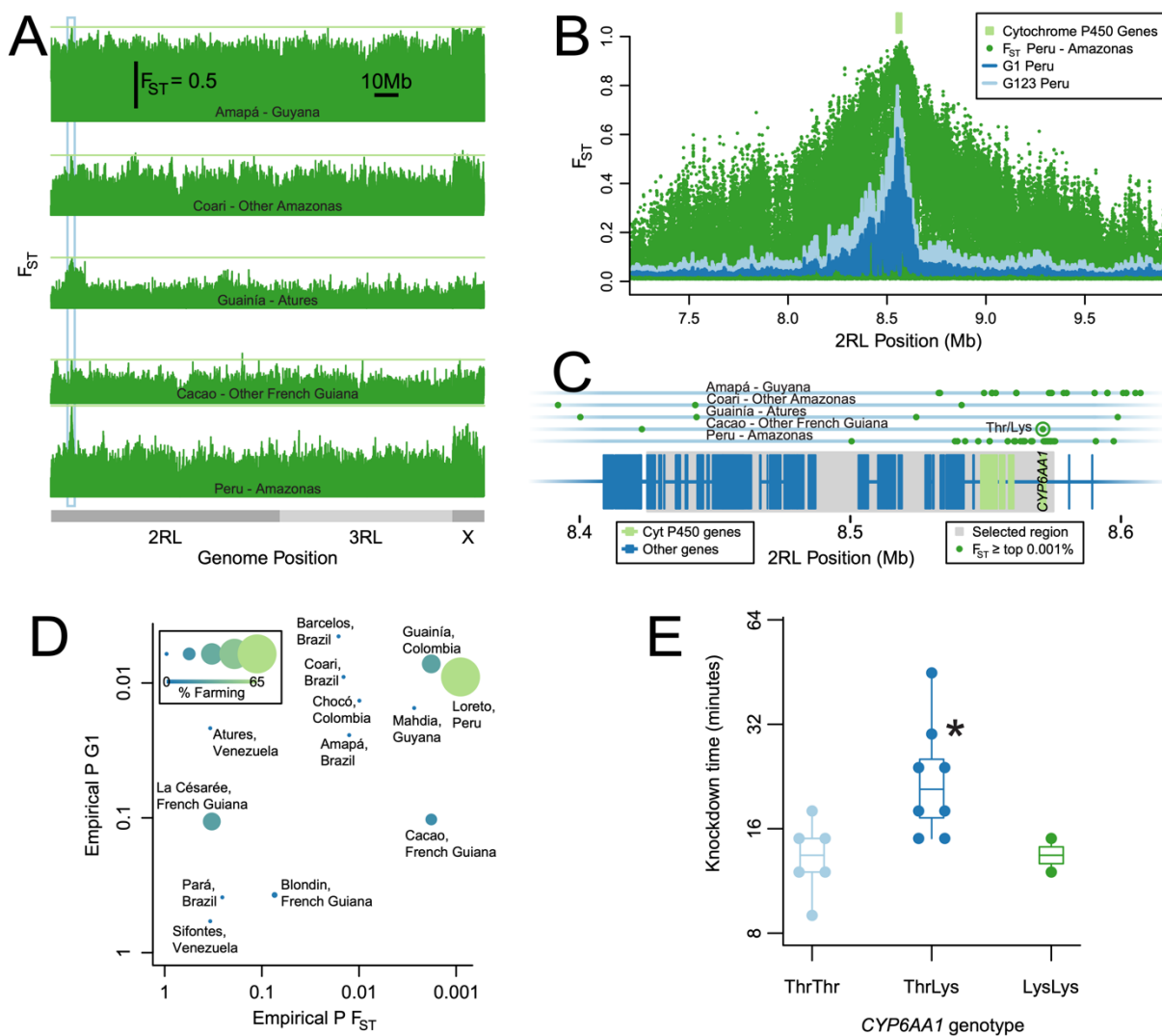


Fig. 4. Signals of selection in a P450-dense genomic region. (A) In genome-wide F_{ST} scans of five independent pairwise comparisons, the same region (blue-outlined box) has the highest or nearly-highest F_{ST} (light green horizontal lines). (B) Peru-Amazonas F_{ST} (dark green), Peru G1 (dark blue), and Peru G123 (light blue) all peak around 2RL 8.5 Mb, near a cluster of cytochrome P450 genes (light green). (C) Genes in the 150 kb region (2RL_8.425-8.575Mb, grey box) showing a positive selection signal in multiple populations. Top F_{ST} outliers (0.001% threshold) from the five comparisons (A) are plotted (dark green dots). The *CYP6A1* Thr283Lys polymorphism is circled. (D) For each population, the top values of G1 and F_{ST} (against genetically similar sister population(s)) in the 2RL_8.425-8.575 Mb window were ranked among all autosomal 150 kb windows, yielding empirical P values. Positive selection is indicated by low P (meaning most windows have lower G1 or F_{ST}). Size and color of points indicate the proportion of surrounding land devoted to “mosaic of farming uses”, which is high for several strongly-selected populations. (E) Knockdown bottle assay with 12.5 $\mu\text{g/mL}$ deltamethrin of wild-caught Cacao *A. darlingi* genotyped at *CYP6A1* Thr283Lys polymorphism. Heterozygotes survive significantly longer ($P < 0.05$).