MOSQUITO GENETICS

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

Jacob A. Tennessen[1,2]*, Raphael Brosula[2], Estelle Chabanol[3], Sara Bickersmith[4], Angela M. Early[2], Margaret Laws[1,2], Katrina A. Kelley[1,2], Maria Eugenia Grillet[5], Dionicia Gamboa[6], Eric R. Lucas[7], Jean-Bernard Duchemin[3,8], Martha L. Quiñones[9], Maria Anice Mureb Sallum[10], Eduardo S. Bergo[11], Jorge E. Moreno[12], Sanjay Nagi[7], Nicholas J. Arisco[1], Mohini Sooklall[13], Reza Niles-Robin[13], Marcia C. Castro[1], Horace Cox[13], Mathilde Gendrin[3], Jan E. Conn[4,14], Daniel E. Neafsey[1,2]

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 *An. darlingi* (median depth 18x) from six South American countries. We observe deep geographic population structure, high genetic diversity including 13 putative segregating inversions, and no evidence for sympatric cryptic taxa despite high interpopulation divergence. Strong signals of selection are plausibly driven by insecticides, especially on cytochrome P450 genes. 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 confirmed symptomatic malaria infections per year in the Americas, causing hundreds of deaths (*1*). 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 the emergence of drug resistance 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 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 *P. falciparum* and strongly influencing their highly variable transmission rates (*6*, *7*).

[1]Harvard T.H. Chan School of Public Health, Boston, MA, USA. [2]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [3]Institut Pasteur de la Guyane, Cayenne, French Guiana. [4]New York State Department of Health, Wadsworth Center, Albany, NY, USA. [5]Instituto de Zoología y Ecología Tropical, Facultad de Ciencias, Universidad Central de Venezuela, Caracas, Venezuela. [6]Laboratorio de Malaria: Parásitos y Vectores, Laboratorios de Investigación y Desarrollo, Facultad de Ciencias e Ingeniería, Universidad Peruana Cayetano Heredia, Lima, Peru. [7]Liverpool School of Tropical Medicine, Liverpool, UK. [8]Arboviruses and Insect Vectors, Institut Pasteur, Paris, France. [9]Universidad Nacional de Colombia, Bogotá, Colombia. [10]Faculdade de Saúde Pública, Universidade de São Paulo, São Paulo, Brazil. [11]Instituto Pasteur de São Paulo, São Paulo, Brazil. [12]Instituto de Altos Estudios Dr. Arnoldo Gabaldón, Centro de Investigaciones de Campo Francesco Vitanza, Bolivar, Venezuela. [13]Vector Control Services, Ministry of Health, Georgetown, Guyana. [14]Department 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**

There is deep evolutionary divergence between malaria vectors in the neotropics relative to the rest of the world. The *Nyssorhynchus* subgenus, which contains *An. darlingi*, is confined to the neotropics (*8*) and is arguably a distinct genus (*9*). It diverged from the *Cellia* subgenus, including the principal vector of malaria in Africa (*An. gambiae*), ~90 million years ago (*10*). Despite extensive population genomic and ecological analysis of malaria vectors in Africa and Asia (*11*–*14*), comparable work on neotropical anophelines has been restricted to relatively small population genetic datasets (*15*–*17*). Other anophelines show extreme genetic diversity tied to anthropogenic population expansions over thousands of years (*11*), but genome-wide patterns of polymorphism and resultant demographic inferences are unknown for *An. darlingi*. It has also remained unclear whether *An. darlingi* is an assemblage of cryptic, morphologically identical species, as inferred in African anophelines from partial sterility in laboratory crosses and subsequently explored in genomic studies of the *An. gambiae* complex (*18*) and the *An. funestus* complex (*19*). Population genetic subdivisions occur in *An. darlingi* on either side of the Andes (*15*) and between the Amazon basin and southern Brazil (*16*, *17*), raising the hypothesis that *An. darlingi* is a cryptic species complex (*6*, *17*); however, no genomic investigation has been conducted to determine the existence of sympatric cryptic species (morphologically similar organisms in the same habitat showing partial or complete reproductive isolation).

Cryptic species have complicated efforts to accurately measure insecticide resistance phenotypes and to identify genotypic markers of insecticide resistance in African vectors since the advent of large-scale insecticidal control almost a century ago (*20*, *21*). Whereas other anophelines repeatedly evolve insecticide resistance through target site polymorphisms as well as metabolically acting factors such as cytochrome P450s (*22*–*24*), in *An. darlingi*, insecticide resistance alleles have not been detected by targeted genotyping (*25*), and observations of phenotypic resistance have been rare and restricted to certain geographic regions and classes of insecticide (*26*–*29*) (table S1).

In this study, we present a population genomic analysis of *An. 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 anophelines that have been particularly vexing to vector control goals, including strikingly high diversity characterized by large segregating inversions, partially interbreeding cryptic taxa, and insecticide resistance through target-site and metabolic mechanisms.

High genetic diversity partitioned into large haplotype blocks

We generated whole-genome sequences for 1094 adult females of *An. darlingi* from 16 locations in six South American countries (Fig. 1A and tables S2 and S3). Our sampling represents most of the species' range with some exceptions (e.g., Central America and much of southern Brazil). Excluding eight individuals with elevated heterozygosity indicative of contamination (fig. S1) and 62 individuals with median coverage ≤1x left 1024 mosquitoes from which allele frequencies were calculated [median (mean) ± standard deviation (min to max) coverage = 20x (23x) ± 15x (2 to 80x)] with 99.5% of the genome accessible for genotyping (fig. S2 and table S4). Of these, 658 individuals had ≥14x coverage, and excluding 11 individuals displaying excessive kinship relationships (>0.1 corresponding to half-sibling or closer) left 647 "exemplar" individuals used in most analyses. Collection localities were diverse in land use (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 specimens positive for human DNA ($P < 0.01$; table S3 and fig. S3). Rates of disease transmission will depend on how often mosquitoes bite humans relative to other vertebrates, which is in turn influenced by ecological factors and human behavior.

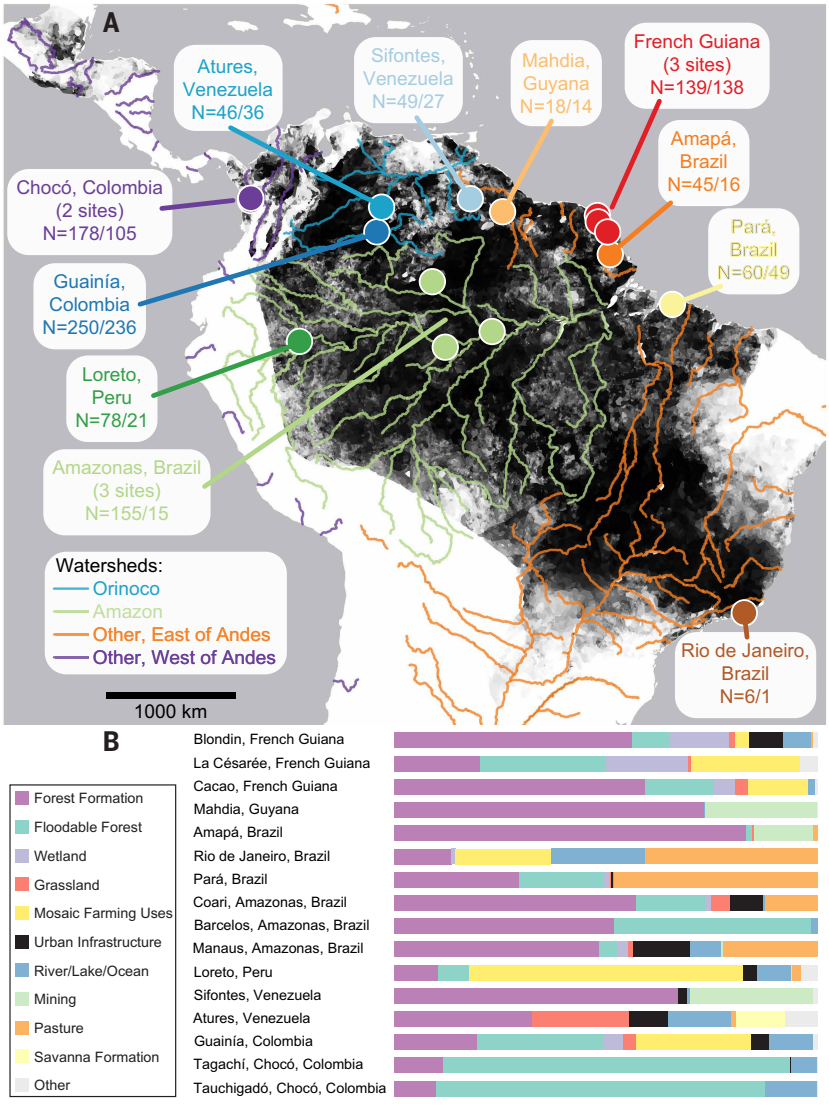

Fig. 1. Collection sites. (**A**) Collection localities with counts of usable specimens (coverage ≥2x/coverage ≥14x) in six countries. Gray scale shading shows predicted distribution of *An. darlingi* (probability of presence from white = 0 to black = 1). Major rivers are colored by watershed. (**B**) Proportion of land cover and land use type in the surrounding 5-km area for all collection sites.

Nucleotide diversity (π, the proportion of pairwise sequence differences) is extremely high, with heterozygosity (per-individual π) >2% in most individuals (Fig. 2A). Synonymous-site nucleotide diversity (π_S) is 3.2% on autosomes, which is comparable to the famously high nucleotide diversity of the *An. gambiae* complex (*11*), and 2.3% on the X (Fig. 2B), which is twofold higher than *An. gambiae*. X chromosome π is ~75% of autosomal π, reflecting the expected X/autosome effective population size ratio and suggesting that X diversity in *An. darlingi* is not affected by whatever factors may cause its further reduction in *An. gambiae* (*11*), such as sex ratio skew (an excess of breeding males reducing the X effective population size) or genetic draft (X-specific selective sweeps).

Site-frequency spectra (distributions of common and rare alleles) show largely stable population sizes with a few historical bottlenecks. Tajima's *D*, a statistic that summarizes these patterns, is close to zero (median per population −0.06; fig. S4), suggesting relative stability in long-term population size and persistence. This contrasts with the very negatively skewed Tajima's *D* value for *An. gambiae* (<−1.5), which is attributable to explosive population growth facilitated by human-modified habitats (*11*). Fitted demographic models (Fig. 2C) show that over the past 2000 years, six populations have changed little in size, while three (Cacao, La Césarée, and Loreto) experienced general declines in abundance, and three (Chocó, Sifontes, and Pará) have recovered from mild bottlenecks and retain lower-than-average diversity (Fig. 2A). Approximately 4 to 20 thousand years ago, most populations experienced bottlenecks of about one order of magnitude, which is consistent with restricted mosquito habitat in forest refugia after the Last Glacial Maximum (*6*). Linkage disequilibrium inferences of effective population sizes indicate large population fluctuations in recent decades (fig. S5), perhaps in response to anthropogenic ecological and demographic changes (*30*–*33*), including deforestation, urbanization, expanding mosquito-friendly habitats through ongoing open-pit mining in the Guiana Shield (*2*), and mosquito control efforts followed by subsequent rebounds. However, human activity across Latin America has not resulted in an exponential increase in *An. darlingi* as similar changes have for African mosquitoes, likely because *An. darlingi* is less anthropophilic (*30*, *34*).

Chromosomal inversion polymorphisms are common in other anopheline species and have been associated with adaptation (*11*, *14*, *35*). We detected 13 large (3 to 19 Mb) segregating haplotype blocks—likely chromosomal inversions (*36*)—showing perfect linkage disequilibrium among physically distant polymorphisms (fig. S6 and table S5). 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. S7). There are few examples of alternate inversion haplotypes being fixed across populations, except in lower-diversity populations such as Pará, in which genetic drift presumably plays a larger role (table S5). Heterozygosity varies considerably both among samples and across the genome, reflecting variation in selective constraints, genetic draft from selective sweeps, and wide peaks where individuals carry two divergent inversion haplotypes (Fig. 2D and fig. S8). We also identified copy number variants overlapping 6425 genes (table S6), which vary by population (fig. S9).

Pronounced geographical population structure without sympatric cryptic taxa

We observed genetic divergence across populations, as seen in other widespread South American insects (*37*–*40*) and in studies of *An. darlingi* using fewer loci (*17*, *41*). Ancestry analysis with ADMIXTURE predicted eight ancestral populations, closely corresponding to contemporary collection sites, with fewer than half of the locales displaying any substantial admixture (gene flow from multiple ancestral populations; Fig. 3A). Ancestry models with fewer than eight populations are suboptimal but indicate that the most fundamental splits [number of ancestral populations (*K*) = 2 or 3] are among an Amazon/Atlantic cluster (French Guiana, Amapá, Mahdia, Rio de Janeiro, Pará, Amazonas, and Loreto), an Orinoco cluster (Sifontes, Atures, and Guainía), and a Chocó cluster west of the Andes (figs. S10 and S11). Principal components analysis also clusters samples in a manner consistent with geography, with the largest split separating Colombian and Venezuelan mosquitoes from all others (Fig. 3B). Similar results are obtained for genomic partitions such as chromosome arms (figs. S12 and S13) and for geographical subsets of samples (fig. S14). The results of AdmixtureBayes are consistent with these results and support a model with no admixture between any populations, dividing the Amazon/

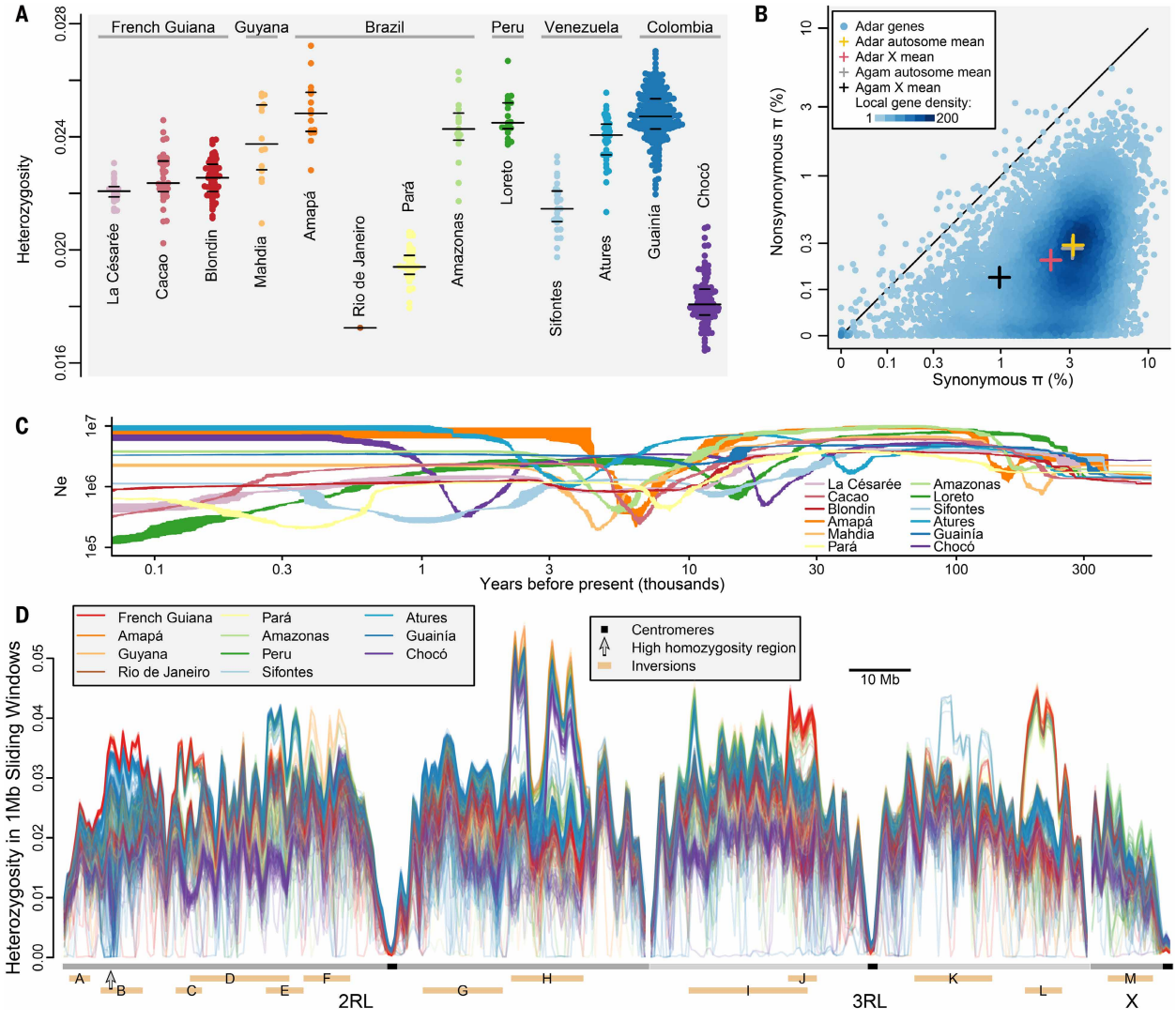

Fig. 2. Genetic diversity. (**A**) Median heterozygosity per individual, arranged by population. (**B**) Synonymous and nonsynonymous nucleotide diversity (π) per gene. Solid line = π_N/π_S (ratio of nonsynonymous to synonymous diversity) of 1. Crosses show mean values for X and autosomal genes, both for *An. darlingi* (Adar) and for the *An. gambiae* complex (Agam) as a comparison. (**C**) Stairway Plot results (range of three replicate runs per population) showing population size (Ne) changes over time. (**D**) Heterozygosity per individual in sliding 1-Mb windows. Individuals are 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 twofold higher than in homozygotes. Occasionally, otherwise diverse individuals have regions of near-zero heterozygosity for more than a megabase, especially where indicated by the gray arrow, indicating homozygosity along a long haplotype due to selection.

Atlantic cluster from Orinoco/Chocó (Fig. 3C and fig. S15). Additional structure occurs among southern populations not sampled here (*17*, *41*), suggesting that the species may comprise an even larger number of genetic clusters. The Andes represent an intuitive and frequently documented barrier to gene flow in many regional species (*39*), which is consistent with the isolation of Chocó *An. darlingi* (*15*). More notable is the distinction between the two main eastern clusters based around drainage basins that are not separated by obvious physical barriers preventing gene flow: Orinoco versus Amazon/Atlantic. These clusters split Sifontes, Venezuela, from Mahdia, Guyana, separated by only 270 km of low-elevation suitable habitat that nevertheless represents a strong barrier to gene flow within an estimated effective migration surface (fig. S16). Whether these or other geographically distinct populations show any reproductive isolation, or alternatively whether gene flow is or was historically simply restricted by the external environment, cannot be determined with our data. However, within sampling locales there is no evidence of cryptic taxa occurring in sympatry; individuals from the same site are always genetically homogeneous with closely shared ancestry (Fig. 3, A and B) and do not depart systematically from Hardy-Weinberg equilibrium (fig. S17).

We observed strong population structure through pairwise F_{ST} (a metric of allele frequency divergence), especially between the major genetic clusters (median 0.18), exceeding F_{ST} observed in *An. gambiae* and *An. coluzzii* (*11*, *12*) (Fig. 3D, fig. S18, and table S7), 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 *An. darlingi* major clusters even surpasses F_{ST} between African species (<0.15). Although adaptive alleles in the *An. gambiae* complex can quickly travel >1000 km (*11*), perhaps facilitated by high-altitude wind dispersal (*42*), it is unknown whether *An. darlingi* can utilize similar mechanisms. Plausibly, long-range migration has been more adaptive in Africa than in South America owing to increased aridity and seasonality of wetlands, in contrast to the more stable and continuous forest habitat of *An. darlingi*. This or other life history

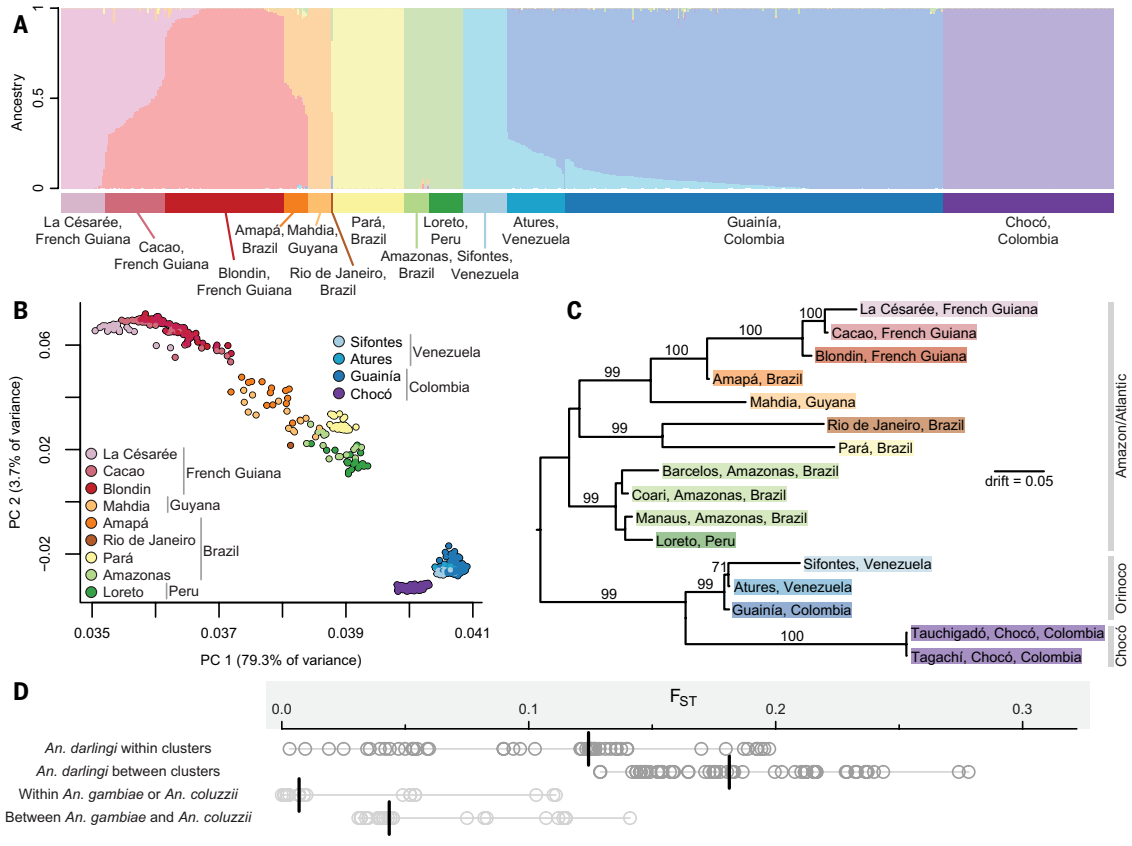

Fig. 3. Population structure. (**A**) Ancestry assignment through ADMIXTURE identifies eight ancestral populations, closely corresponding to the contemporary populations of La Césarée, Blondin, Mahdia, Pará, Amazonas/Loreto, Sifontes, Guainía, and Chocó. Cacao, Amapá, Rio de Janeiro, and Atures show intermediate ancestry. (**B**) Principal components (PC) analysis separates samples geographically. (**C**) Sampling site topology identified by AdmixtureBayes with no predicted admixture and showing three major genetic clusters: Amazon/Atlantic, Orinoco, and Chocó. Numbers on branches are posterior probabilities of nodes (not shown if <60). Branch length is in units of genetic drift, analogous to F_{ST}. (**D**) Pairwise F_{ST} among 15 *An. darlingi* collection sites (Rio de Janeiro excluded for sample size), classified as within major clusters [from (C)] (e.g., Mahdia versus Loreto) or between clusters (e.g., Mahdia versus Sifontes). Within- and between-species values for African *An. gambiae* and *An. coluzzii* are shown for comparison. For all four sets of points across both continents, median geographic distance is similarly large (1000 to 2000 km). The black vertical lines show medians.

differences could produce the distinct population genetic patterns on each continent.

The deep population structure we observed 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 *An. darlingi* (*16*), but populations in Amapá and Pará on opposite sides of the Amazon Delta, although genetically distinct, are not as divergent as some other population pairs without a riverine border (Figs. 1 and 3). Similar close relationships are seen in *An. aquasalis* across the same range (*43*) and in *An. darlingi* populations separated by the Orinoco River. Instead of rivers as barriers, we observed the opposite: Boundaries between watersheds demarcate the major genetic divisions. Gene flow may be facilitated by waterways owing to the dependence of *An. darlingi* on aquatic breeding sites, including as refugia during ancient periods of population contraction and secondary expansion (Fig. 2C). Interpopulation connectivity may also be enhanced by anthropogenic activity along rivers, including deforestation leading to favorable disturbed habitat (*33*, *40*), or transport in vehicles or boats. If a population were to be extirpated as reported in Suriname (*44*), there could be a long delay before the habitat is recolonized through migration from neighboring countries, unless enabled by human activity. Limited gene flow also indicates that a new insecticide resistance allele is less likely to spread quickly beyond its population of origin.

Positive selection signals may reflect emerging insecticide resistance

Numerous signals of positive selection are apparent across these populations (tables S8, S9, and S10). In particular, a region centered at chromosome 2RL:8.425 to 8.575 Mb within Inversion B is the top outlier in multiple populations under multiple distinct selection tests, including genome-wide F_{ST} scans between adjacent populations, G1 and G123 statistics that identify extended haplotype homozygosity, and scans for locus-specific divergent selection effects while accounting for linkage using Flink (Fig. 4A). This region shows exceptional allele frequency divergence between populations as well as overlapping windows of low heterozygosity, which is consistent with selective sweep expectations (Fig. 2D and Fig. 4B). It overlaps 20 genes, including a cluster of six cytochrome P450 genes (Fig. 4C). Even a single incidence of this selection signal would be notable, as in Loreto, Peru (Fig. 4B), or Guainía, Colombia (fig. S19A), yet this window is among the top outliers in approximately half of our populations by at least one metric (Fig. 4D). In some populations, the selection signal is clearly concentrated on the P450 genes, with a strong signal at *CYP6AA1* (*ADAR2_008159*), which is tied to pyrethroid resistance in *An. funestus* (*45*) and *An. gambiae* (*46*). 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. S19B), suggesting parallel evolution across distinct genetic backgrounds rather than a single species-wide hard sweep of a novel adaptive allele.

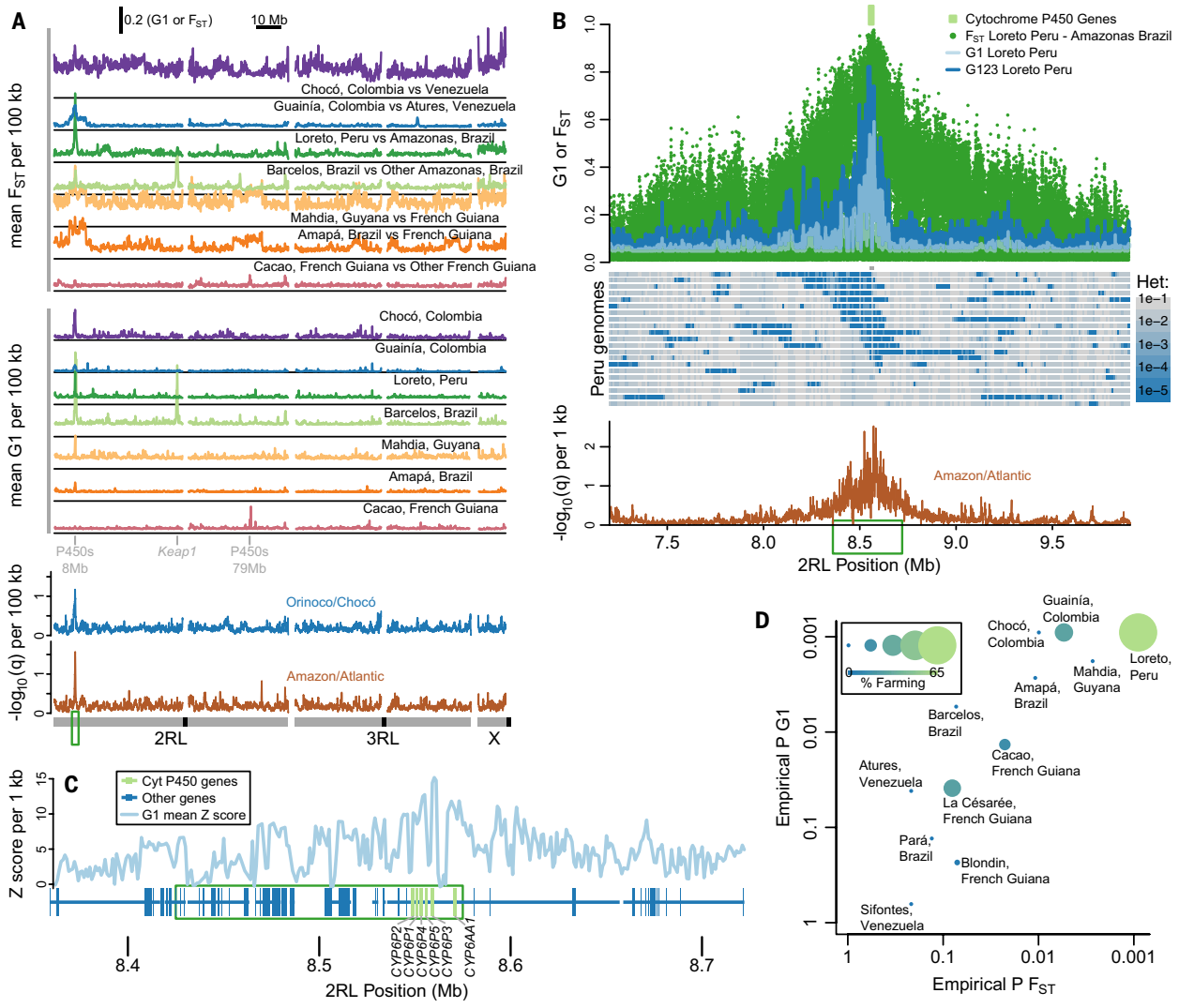

Fig. 4. Signals of selection in a P450-dense genomic region. (**A**) Genome-wide scans of F_{ST} (top), G1 (middle), and the false discovery rate for a positive locus-specific effect (q) from Flink (bottom), all in 100-kb windows. The cluster of cytochrome P450 genes at 2RL:8 Mb is a consistent outlier among metrics and among populations. Other outliers include *Keap1* and a second P450 cluster at 2RL:79 Mb. The green box indicates the region examined in (B). (**B**) Selection signals in Loreto, Peru, all peak at the P450 cluster (light green). (Top) Loreto-Amazonas F_{ST} (dark green), Loreto G1 (light blue), and Loreto G123 (dark blue). (Middle) Regions of low heterozygosity (dark blue). (Bottom) Flink q in 1-kb windows for Amazon/Atlantic group. The green box indicates the region examined in (C). (**C**) Cytochrome (Cyt) P450 cluster (genes labeled) and surrounding genes in 2RL:8 Mb selection region, with mean *Z* scores for G1 in 1-kb windows across populations. The green box indicates the region examined in (D). (**D**) For each population, the top values of G1 and F_{ST} [against genetically similar sister population(s)] in the 2RL:8.425- to 8.575-Mb window were ranked among all autosomal 150-kb windows, yielding empirical *P* values. Positive selection is indicated by low *P* (meaning that most windows have lower G1 or F_{ST}). The 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.

Given the well-established link between cytochrome P450 genes and toxin resistance (*23*, *45*), 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 (*47*) could be the main source of exposure, as hypothesized for resistance in African anophelines (*48*) and possibly other nontargets of pesticides (*49*). We ranked the 2RL:8.425-to-8.575-Mb 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 ($P < 0.001$ for both F_{ST} and G1) were observed in Loreto (Peru), which also had the highest proportion of land dedicated to "mosaic of farming uses" (Fig. 1B). Strong signals were also observed in other populations with some farming activity, such as Cacao (French Guiana) and Guainía (Colombia). Thus, there is an intriguing trend linking farming land use and selection signals, although this correlation is not significant ($P > 0.1$). A bottle assay in Cacao identified a correlation between deltamethrin resistance and the *CYP6AA1* genotype ($P < 0.05$; fig. S20), although our small ($n = 16$) and geographically restricted dataset precludes drawing general conclusions about phenotypic effects.

Other regions of the genome show selection signals beyond this window, although not as strong or as consistently across populations (Fig. 4A, figs. S21 to S24, and tables S8 to S10). F_{ST} for Cacao versus other French Guiana was highest in a different cluster of P450 genes at 2RL:79.3 Mb (fig. S24A), peaking at *CYP6M2* (*ADAR2_008623*), which is tied to pyrethroid resistance in *An. gambiae* (*50*), and also overlapping a peak in G1. 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 to 93.785 Mb, containing a set of seven adenosine triphosphate–binding cassette subfamily G genes, is an F_{ST} outlier in two independent pairwise comparisons:

between French Guiana sites and between Amazonas, Brazil sites (figs. S21 and S22 and table S8). A strong selection signal centered on 2RL:49.950 to 50.050 Mb is implicated by F_{ST}, G1, and G123 in Barcelos, Brazil (fig. S23), where pyrethroid resistance was recently reported (*29*). The top genome-wide F_{ST} for this comparison occurs in *Keap1* (*ADAR2_006983*), which also shows positive selection signals in several African anophelines (*51*, *52*), plausibly because of its role in a metabolic insecticide resistance pathway (*22*).

No notable positive selection is observed at orthologs of genes conferring target-site resistance in other anophelines. Target codons Leu995 in *vgsc* (site of the *kdr* mutation), Ala296 in *rdl*, and Gly280 in *ace1* (*24*) exhibit no nonsynonymous variation (fig. S24B), which is consistent with previous results (*25*). P450s and *Keap1* underlie metabolic resistance, suggesting that this mechanism has been favored in *An. darlingi* over target-site resistance. Although the specific adaptive polymorphisms remain unidentified, there is very little evidence that gene duplications of P450 or *Keap1* are selection targets, as hypothesized for *An. gambiae* (*46*) and *An. funestus* (*53*). Copy number variants in *CYP6AA1*, *CYP6M2*, and adjacent P450 genes (table S6 and fig. S9A) 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 specimens homozygous for the selected allele at high-F_{ST} variants. We observed no *Keap1* duplicates. Thus, resistance may depend on point mutations in metabolic resistance genes, potentially including nonsynonymous variants. We saw no strong selection signals on potential *Plasmodium*-interacting genes such as *TEP1* and *APL1A*, possibly because parasite-imposed selection is weak in low-transmission settings.

Our results have several implications for vector control across Latin America. Although we are unaware of plans to modify *An. darlingi* with gene drives as is being attempted on other species (*54*), a future gene-drive implementation in these mosquitoes may require repeated launches in several geographic locations, potentially requiring regional coordination. *An. darlingi* is extremely diverse and likely harbors latent potential to adapt to novel interventions. 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 because resistance can evolve in parallel on different genetic backgrounds across habitats, as has seemingly occurred with P450s. Continued population genomic monitoring can inform disease control efforts and lead to broad insights about ecological and evolutionary diversity in neotropical vectors.

REFERENCES AND NOTES

1. World Health Organization, World Malaria Report 2024 (WHO, 2024); https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2024.
2. M. E. Grillet *et al.*, *PLOS Negl. Trop. Dis.* **15**, e0008211 (2021).
3. J. C. Wootton *et al.*, *Nature* **418**, 320–323 (2002).
4. S. M. Chenet *et al.*, *J. Infect. Dis.* **213**, 1472–1475 (2016).
5. L. C. Mathieu *et al.*, *eLife* **9**, e51015 (2020).
6. J. D. Charlwood, *Mem. Inst. Oswaldo Cruz* **91**, 391–398 (1996).
7. B. C. Carlos, L. D. P. Rona, G. K. Christophides, J. A. Souza-Neto, *Pathog. Glob. Health* **113**, 1–13 (2019).
8. R. E. Harbach, in *Anopheles Mosquitoes: New Insights into Malaria Vectors*, S. Manguin, Ed. (IntechOpen, 2013), pp. 3–55.
9. P. G. Foster *et al.*, *R. Soc. Open Sci.* **4**, 170758 (2017).
10. L. A. Freitas *et al.*, *PLOS ONE* **10**, e0134462 (2015).
11. *Anopheles gambiae* 1000 Genomes Consortium, *Nature* **552**, 96–100 (2017).
12. *Anopheles gambiae* 1000 Genomes Consortium, *Genome Res.* **30**, 1533–1546 (2020).
13. B. St. Laurent *et al.*, *Commun. Biol.* **5**, 1308 (2022).
14. M. Boddé *et al.*, *Science* **389**, eadu3596 (2025).
15. N. Naranjo-Díaz, J. E. Conn, M. M. Correa, *Infect. Genet. Evol.* **39**, 64–73 (2016).
16. P. M. Pedro, M. A. M. Sallum, *Biol. J. Linn. Soc. Lond.* **97**, 854–866 (2009).
17. K. J. Emerson, J. E. Conn, E. S. Bergo, M. A. Randel, M. A. M. Sallum, *PLOS ONE* **10**, e0130773 (2015).
18. M. C. Fontaine *et al.*, *Science* **347**, 1258524 (2015).
19. S. T. Small *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 31583–31590 (2020).
20. G. Davidson, *Nature* **178**, 863–864 (1956).
21. G. Davidson, *Bull. World Health Organ.* **31**, 625–634 (1964).
22. V. A. Ingham, P. Pignatelli, J. D. Moore, S. Wagstaff, H. Ranson, *BMC Genomics* **18**, 669 (2017).
23. J. Vontas, E. Katsavou, K. Mavridis, *Pestic. Biochem. Physiol.* **170**, 104666 (2020).
24. M. Campos *et al.*, *Sci. Rep.* **12**, 13893 (2022).
25. S. A. Bickersmith *et al.*, *Genes* **14**, 1892 (2023).
26. M. F. Suarez, M. L. Quiñones, J. D. Palacios, A. Carrillo, *J. Am. Mosq. Control Assoc.* **6**, 72–74 (1990).
27. L. I. Orjuela *et al.*, *BioMed Res. Int.* **2018**, 9163543 (2018).
28. S. Vezenegho *et al.*, *Am. J. Trop. Med. Hyg.* **108**, 424–427 (2022).
29. Q. S. Amorim *et al.*, *Malar. J.* **24**, 155 (2025).
30. L. Mirabello, J. E. Conn, *Heredity* **96**, 311–321 (2006).
31. G. B. Schoeler, C. Flores-Mendoza, R. Fernández, J. R. Davila, M. Zyzak, *J. Am. Mosq. Control Assoc.* **19**, 286–296 (2003).
32. W. Lainhart *et al.*, *Malar. J.* **14**, 375 (2015).
33. A. Y. Vittor *et al.*, *Am. J. Trop. Med. Hyg.* **81**, 5–12 (2009).
34. M. Moreno *et al.*, *PLOS Negl. Trop. Dis.* **11**, e0005337 (2017).
35. D. Ayala, A. Ullastres, J. González, *Front. Genet.* **5**, 129 (2014).
36. A. J. Cornel *et al.*, *Mem. Inst. Oswaldo Cruz* **111**, 335–346 (2016).
37. P. M. Pedro, A. Uezu, M. A. Sallum, *J. Hered.* **101**, 618–627 (2010).
38. K. M. Lewald *et al.*, *Genome Biol. Evol.* **15**, evad060 (2023).
39. A. C. Turchetto-Zolet, F. Pinheiro, F. Salgueiro, C. Palma-Silva, *Mol. Ecol.* **22**, 1193–1213 (2013).
40. C. C. Ribas *et al.*, *Nat. Reviews Biodiversity* **1**, 14–31 (2025).
41. J. E. Conn, M. G. Rosa-Freitas, S. L. Luz, H. Momen, *J. Am. Mosq. Control Assoc.* **15**, 468–474 (1999).
42. D. L. Huestis *et al.*, *Nature* **574**, 404–408 (2019).
43. T. L. Fairley, M. M. Póvoa, J. E. Conn, *J. Med. Entomol.* **39**, 861–869 (2002).
44. H. Hiwat *et al.*, *Am. J. Trop. Med. Hyg.* **86**, 649–655 (2012).
45. S. S. Ibrahim *et al.*, *Genes* **9**, 140 (2018).
46. H. Njoroge *et al.*, *Mol. Ecol.* **31**, 4307–4318 (2022).
47. S. I. Grondona, M. L. Lima, H. E. Massone, K. S. B. Miglioranza, *Sci. Total Environ.* **901**, 165992 (2023).
48. F. A. Kouadio *et al.*, *Parasit. Vectors* **16**, 270 (2023).
49. N. F. Wan *et al.*, *Nat. Commun.* **16**, 1360 (2025).
50. R. F. Djouaka *et al.*, *BMC Genomics* **9**, 538 (2008).
51. S. C. Nagi, N. J. Harding, M. K. N. Lawniczak, M. J. Donnelly, A. Miles. An atlas of positive selection in the genomes of major malaria vectors. bioRxiv 2025.07.16.664900 [Preprint] (2025). https://doi.org/10.1101/2025.07.16.664900.
52. E. R. Lucas *et al.*, *PLOS Biol.* **22**, e3002898 (2024).
53. N. M. T. Tatchou-Nebangwa *et al.*, *BMC Biol.* **22**, 286 (2024).
54. K. Kyrou *et al.*, *Nat. Biotechnol.* **36**, 1062–1066 (2018).

ACKNOWLEDGMENTS

We thank N. Besansky, R. Carinci, A. Cote-L'Heureux, C. Maria De Marco, T. M. P. de Oliveira, S. Díaz, E. Figueira, P. Gaborit, Z. Johnson, C. Knox, M. Lawniczak, T. Lehmann, M. Llewellyn, A. Makunin, N. Moncada, C. C. Moreira, M. A. Mota, M. Neves, R. Newmiller, R. Panchal, S. Rudman, V. Sanchez, P. Schwabl, M. Shieh, N. Swaminathan, S. Talaga, A. Villarroel, and malaria control staff and local residents at various collection locations. **Funding:** This work was supported by NIH grant U19AI110818 (D.E.N.), NIH grant R01AI110112 (J.E.C.), Bill & Melinda Gates Foundation grant INV-009416 (D.E.N.), Agence Nationale de la Recherche grant ANR-18-CE15-0007 (M.G.), and National Council for Scientific and Technological Development (CNPq) grant 303382/2022-8 (M.A.M.S.). **Author contributions:** Conceptualization: J.E.C., D.E.N.; Data curation: J.A.T., R.B., A.M.E.; Formal analysis: J.A.T., N.J.A., E.R.L., M.C.C.; Funding acquisition: M.A.M.S., M.G., J.E.C., D.E.N.; Investigation: K.A.K., M.A.M.S., J.E.C.; Methodology: J.A.T., D.E.N.; Project administration: M.L., D.E.N.; Software: S.N.; Resources: E.C., S.B., D.G., E.S.B., J.E.M., M.S., R.N.-R., M.A.M.S., M.E.G., M.L.Q., H.C., M.G.; Supervision: A.M.E., D.E.N.; Validation: J.-B.D.; Visualization: J.A.T.; Writing – original draft: J.A.T.; Writing – review & editing: J.A.T., M.E.G., M.A.M.S., J.E.C., D.E.N. **Competing interests:** The authors declare that they have no competing interests. **Data, code, and materials availability:** All raw sequencing data has been deposited at NCBI SRA, BioProject PRJNA1169887. Genotypes are accessible at Malaria Genome Vector Observatory: https://www.malariagen.net/project/anopheles-darlingi-genomic-surveillance-project/. 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. All DNA has been depleted, and thus no new materials were generated for this study.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adw9761
Materials and Methods; Figs. S1 to S24; Tables S1 to S10; References (*55–89*); Reproducibility Checklist

Submitted 24 February 2025; resubmitted 27 August 2025; accepted 15 January 2026

10.1126/science.adw9761

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

Jacob A. Tennessen, Raphael Brosula, Estelle Chabanol, Sara Bickersmith, Angela M. Early, Margaret Laws, Katrina A. Kelley, Maria Eugenia Grillet, Dionicia Gamboa, Eric R. Lucas, Jean-Bernard Duchemin, Martha L. Quiñones, Maria Anice Mureb Sallum, Eduardo S. Bergo, Jorge E. Moreno, Sanjay Nagi, Nicholas J. Arisco, Mohini Sooklall, Reza Niles-Robin, Marcia C. Castro, Horace Cox, Mathilde Gendrin, Jan E. Conn, and Daniel E. Neafsey

Science **391** (6792), . DOI: 10.1126/science.adw9761

Editor's summary

Malaria is a major burden in South America and is spread mainly by the mosquito *Anopheles darlingi*. However, genetic studies of this species have so far been limited, leaving many questions about insecticide resistance and even population structure. Tennessen *et al*. sequenced the genomes of 1094 mosquitos collected across six South American countries. They found no evidence of cryptic taxa but did identify 13 large inversions, some with signs of positive selection. Additionally, they found that insecticide resistance in *A. darlingi* may be based more in metabolic genes such as those coding for cytochrome P450 enzymes than in gene targets that are common in other mosquito species. These results have important implications for public health measures and potential future gene drive efforts. —Corinne Simonti