Strong positive selection in *Aedes aegypti* and the rapid evolution of insecticide resistance

R. Rebecca Love[1], Josh R. Sikder[1], Rafael J. Vivero[2], Daniel R. Matute[3], Daniel R. Schrider[1]

1 Department of Genetics, School of Medicine, University of North Carolina Chapel Hill, USA
2 Programa de Estudio y Control de Enfermedades Tropicales, PECET, Universidad de Antioquia, Colombia
3 Department of Biology, College of Arts and Sciences, University of North Carolina Chapel Hill, USA
Corresponding author: DRS, drs@unc.edu

ABSTRACT

Aedes aegypti vectors the pathogens that cause dengue, yellow fever, Zika virus, and chikungunya, and is a serious threat to public health in tropical regions. Decades of work has illuminated many aspects of *Ae. aegypti*'s biology and global population structure, and has identified insecticide resistance genes; however, the size and repetitive nature of the *Ae. aegypti* genome have limited our ability to detect positive selection in this mosquito. Combining new whole genome sequences from Colombia with publicly available data from Africa and the Americas, we identify multiple strong candidate selective sweeps in *Ae. aegypti*, many of which overlap genes linked to or implicated in insecticide resistance. We examine the voltage-gated sodium channel gene in three American cohorts, and find evidence for successive selective sweeps in Colombia. The most recent sweep encompasses an intermediate-frequency haplotype containing four candidate insecticide resistance mutations that are in near-perfect linkage disequilibrium with one another in the Colombian sample. We hypothesize that this haplotype may continue to rapidly increase in frequency and perhaps spread geographically in the coming years. These results extend our knowledge of how insecticide resistance has evolved in this species, and add to a growing body of evidence suggesting *Ae. aegypti* has an extensive genomic capacity to rapidly adapt to insecticide-based vector control.

Key words: Insecticide resistance, *Ae. aegypti*, positive selection, Colombia.

INTRODUCTION

The mosquito *Aedes (Stegomyia) aegypti* (Linnaeus, 1762) vectors the arboviruses causing dengue, Zika, yellow fever, and chikungunya (Viglietta et al. 2021), the former two of which have already caused pandemics in the first two decades of the 21st century. No vaccine is publicly available for chikungunya or Zika, and treatment for all four diseases is generally supportive or symptomatic only. While considerable uncertainty exists about the actual number of annual dengue cases worldwide (Wilder-Smith and Byass 2016, Bhatt et al. 2013), one common estimate suggests approximately 400 million cases a year as of 2010 (Bhatt et al. 2013).

Because of its importance to public health, as well as its facility as a model organism, *Ae. aegypti* has been extensively studied. Two subspecies or forms have been identified, initially distinguished on the basis of morphology with corresponding differences in behavior (Mattingly 1957), though the morphological characters in question have long been recognized as subject to variation (Mattingly 1957, Brown et al. 2011). *Ae. aegypti formosus*, recognizable by its dark abdomen, generally prefers forest and other outdoor habitats, and is limited to Africa; in contrast, *Ae. aegypti aegypti* has pale scales on the first abdominal tergite, is found widely across tropical regions of the world outside of Africa, is highly anthropophilic, and is found in close association in humans, often resting indoors and ovipositing in anthropogenic habitats (Mattingly 1957). The two subspecies are genetically distinct as well (Tabachnick and Powell 1979, Gloria-Soria, Ayala, et al. 2016, Crawford et al. 2017, Kotsakiozi, Evans, et al. 2018), though both morphological and genetic distinctions break down in some parts of West Africa (Gloria-Soria, Ayala, et al. 2016, Crawford et al. 2017, Kotsakiozi, Evans, et al. 2018). In addition, the two forms may be differentially susceptible to dengue (Sylla et al. 2009), Zika virus (Aubry et al. 2020), and yellow fever (Tabachnick et al. 1985), though the latter phenotype appears to be highly variable depending on mosquito and viral strain (Tabachnick et al. 1985, Dickson et al. 2014). A third form, *queenslandensis*, was initially recognized on the basis of morphology (Mattingly 1957) but appears to represent morphological variation of *Ae. aegypti aegypti* (Rašić et al. 2016).

Ae. aegypti tends to have low unaided dispersal (Harrington et al. 2005), though it may fly further if suitable oviposition sites or opportunities to blood-feed are lacking (Reiter et al. 1995, Codeço et al. 2007). The widespread tropical distribution of *Ae. aegypti aegypti* is

therefore attributed to its close association with humans (Christophers 1960). Originating in Africa (Tabachnick 1991, Brown et al. 2011, Brown et al. 2014), *Ae. aegypti* spread to the Americas (Brown et al. 2014, Gloria-Soria, Ayala, et al. 2016), likely due to mass transatlantic trafficking of enslaved people. It arrived in Asia more recently (Tabachnick 1991, Brown et al. 2014), possibly facilitated by the opening of the Suez Canal (Tabachnick 1991, Powell et al. 2018). The non-African domestic populations appear to result from one colonization event (Crawford et al. 2017, Rose et al. 2020), descending from a domesticated subpopulation in western Africa (Crawford et al. 2017) or possibly central Africa (Kotsakiozi, Evans, et al. 2018); human preference is linked to expression and sequence variation in the *Or4* gene (McBride et al. 2014, Rose et al. 2020) as well as dry season intensity (Rose et al. 2020).

Population structure tends to be less pronounced within Africa than outside Africa (Brown et al. 2011, Brown et al. 2014, Gloria-Soria, Ayala, et al. 2016); outside Africa, specimens show strong geographic clustering (Rašić et al. 2014, Evans et al. 2015, Schmidt et al. 2020) but extensive connectivity is seen between some geographically distant populations (Gonçalves da Silva et al. 2012, Monteiro et al. 2014, Gloria-Soria, Ayala, et al. 2016). Populations are often dynamic, showing sometimes extensive changes in genetic composition within a few years (Gloria-Soria, Kellner, et al. 2016), and reinvading areas where they had been eradicated (Kelly et al. 2021) or out-competed (Brennan et al. 2021).

Overall, the global genetic structure of *Ae. aegypti* has been well-studied; however, because of the size and repetitive nature of its genome, nearly all previous work has used either sequences from a small number of genes (Bennett et al. 2016), microsatellites (Brown et al. 2011, Gloria-Soria, Ayala, et al. 2016), mtDNA (Gonçalves da Silva et al. 2012, Bennett et al. 2016), genotyping arrays (Evans et al. 2015, Kotsakiozi, Evans, et al. 2018), or reduced representation sequencing (Brown et al. 2014, Rašić et al. 2014). Exome-based (Crawford et al. 2017, Dickson et al. 2017) and whole genome-based population resequencing have only been performed recently (Lee et al. 2019, Rose et al. 2020, Kelly et al. 2021), and the improvement of the *Ae. aegypti* reference genome to a chromosomal-level assembly (Matthews et al. 2018) will greatly facilitate such efforts moving forward. Studies using a small number of loci are extremely limited in their ability to detect recent positive selection, which skews patterns of diversity across larger genomic regions (Maynard Smith and Haigh 1974, Kaplan et al. 1989, Kim and Nielsen 2004). As a consequence, we know relatively little about the extent of recent

positive selection or the broad-scale genomic underpinnings of adaptation in *Ae. aegypti*, compared to our more extensive knowledge of the species' spread and structure. Positive selection has been observed on genomic regions linked to human preference (Rose et al. 2020), and on genes that enable females to retain eggs during periods of drought (Venkataraman et al. 2022); however, with the exception of a putative selective sweep on an insecticide resistance gene (the *Vgsc* gene, discussed below) identified in a sample from southern Mexico (Saavedra-Rodriguez et al. 2019), we know little about how its genome has responded to these selective pressures at a scale beyond that of individual loci.

In some parts of the Americas, *Ae. aegypti* has been subject to insecticide-based vector control for nearly a century. In the early part of the 20th century, some of the techniques used to target *Ae. aegypti* in Latin America included fumigation of dwellings and "oiling" of water sources to kill larvae (Severo 1955). DDT was used beginning in 1945 (Severo 1955) and extensively thereafter, enabling the campaign for continent-wide eradication in the Americas which began officially in 1947 (Soper 1965). The campaign was initially successful in eradicating *Ae. aegypti* from most of South and Central America (Soper 1965); however, DDT resistance had appeared in the Americas by 1964 (Soper 1965), and for this and other reasons, the campaign faltered. *Ae. aegypti* soon spread across South and Central America again (Jansen and Beebe 2010).

Resistance to DDT in *Ae. aegypti* is conferred by both increased metabolism as well as non-metabolic factors (Rathor and Wood 1981, Hemingway et al. 1989, Brengues et al. 2003). The best-characterized non-metabolic factor is *k*nock*d*own *r*esistance (*kdr*), first identified in houseflies, and conferring resistance to pyrethroids as well (Farnham 1977); *kdr*, caused by point mutations in the voltage-gated sodium channel gene (*Vgsc*, (Williamson et al. 1996)) has also been found in a wide variety of other insects, including *Drosophila* (Pittendrigh et al. 1997), anopheline mosquitoes (Martinez-Torres et al. 1998, Clarkson et al. 2021), tobacco budworms (Taylor et al. 1993), cockroaches (Dong and Scott 1994), and many others (Davies et al. 2007).

Unsurprisingly considering the breadth of distribution of these mutations in other insects, an extensive array of nonsynonymous point mutations associated with pyrethroid resistance have been identified in *Ae. aegypti*. (Because this resistance factor was originally identified in *Musca domestica*, the *Musca domestica* amino acid numberings are often used; to avoid ambiguity, we will follow this convention.) Some, such as F1534C (Kawada et al. 2009,

Harris et al. 2010), are globally distributed. Others appear limited to specific continents or regions; V1016I has so far only been recorded from the Americas (Fan et al. 2020) and Africa (Sombié et al. 2019), while V1016G has only been recorded from Asia (Fan et al. 2020). Similarly, S989P appears limited to Asia (Fan et al. 2020), while S723T has only been identified in the Americas (Fan et al. 2020), and V410L has been found in the Americas (Fan et al. 2020) and Africa (Ayres et al. 2020). A subset of the mutations associated with pyrethroid resistance, including F1534C (Hu et al. 2011, Du et al. 2013, Hirata et al. 2014, Haddi et al. 2017), V1016G (Du et al. 2013, Hirata et al. 2014), I1011M (Du et al. 2013) and V410L (Haddi et al. 2017) have been shown experimentally to directly cause pyrethroid resistance.

These mutations are often found in combination with each other; for example, V410L has been observed in conjunction with F1534C (Haddi et al. 2017, Saavedra-Rodriguez et al. 2018) and V1016I (Saavedra-Rodriguez et al. 2018), and S989P is very frequently found on a V1016G background (Kawada et al. 2014, Al Nazawi et al. 2017). Recurrent evolution (Fan et al. 2020) and recombination (Saavedra-Rodriguez et al. 2018) have been suggested as contributing factors to the combinations and distribution of these various mutations. A similar situation of recurrent mutation at the same site, and the presence of haplotypes comprising multiple resistance alleles, has been observed in *An. gambiae* and *An. coluzzii* (Clarkson et al. 2021).

Natural selection associated with insecticide resistance has been widely observed in other insects, including other disease vectors such as *Anopheles gambiae* and *An. coluzzii* (Miles et al. 2017, Xue et al. 2021) and *Anopheles funestus* (Weedall et al. 2020); agricultural pests such as *Leptinotarsa decemlineata* (Pélissié et al. 2022) and *Helicoverpa armigera* (Walsh et al. 2018); and the model organism *Drosophila melanogaster* (Turner et al. 2008, Kolaczkowski et al. 2011), Garud et al. 2015, Schrider et al. 2016). Thus, it seems probable that positive selection for insecticide resistance would be widespread in *Aedes* as well. In addition to selection at specific insecticide resistance loci, vector control programs may cause population contractions (Urdaneta-Marquez and Failloux 2011), which can affect genome-wide change in patterns of diversity. In Colombia, vector control relies on pyrethroids and organophosphates (temephos and malathion) (Aguirre-Obando et al. 2015).

Here, we examine patterns of genetic variation in a set of 30 *Aedes aegypti* genomes collected from Colombia. We also combine these data with previously published *Aedes* genomic

data to examine global patterns of genetic variation. Finally, we turn our attention toward detecting signatures of recent positive selection in *Aedes* samples from Africa and the Americas. In five of the six countries examined, we find strong signals of selection in regions containing genes that have been found to be associated with insecticide resistance in *Aedes* or other insects. We discuss the implications of these findings for adaptation to ongoing and future mosquito control efforts.

RESULTS

Quality control, alignment, and variant genotyping

We used a stringent filtering approach to control the quality of the sequenced samples (Methods). We removed six specimens from the Colombian sample that had mean coverage below 15x or fewer than 70% of reads mapping to the reference genome, leaving a sample of 10 specimens from Cali and 24 from Río Claro (see Table S1 for complete specimen list).

In addition, we examined publicly available sequence data from the United States of America (*n*=27) from Lee et al. 2019 (Lee et al. 2019), and from Santarém, Brazil (*n*=18), Franceville, Gabon (*n*=13), Kaya Bomu, Kenya (*n*=19), and Ngoye, Senegal (*n*=20) from Rose et al. (2020) (Rose et al. 2020), for a total of six cohorts. The Ngoye cohort comes from a population with a strong preference for human over non-human odor as well as a large "domestic" ancestry component; the Franceville cohort originates from a population with a strong preference for non-human odor, and no domestic ancestry; the Kaya Bomu population is intermediate in both traits (Rose et al. 2020).

After removing the six specimens in our focal Colombian sample that failed our quality filters, mean read depth in each population ranged from 10.16 (USA) to 21.10 (Colombia), with between 94.80% (Gabon) and 98.19% (USA) of reads mapping on average (Table S2).

Variant calling produced an initial data set of 247,108,978 variants in 131 specimens. After filtering, we retained 43,213,550 high-quality single nucleotide polymorphisms (SNPs) with a known location on one of the three autosomes, and an additional 226,687 SNPs in regions of the genome that could not be confidently assigned to chromosomes. Our analysis focused on the SNPs on the three autosomes. The number of SNPs found in each cohort, before and after applying filters based on relatedness as discussed below, is shown in Table S3.

In keeping with previous work, substantially more SNPs were called per specimen in the African sample than in the American sample (Table S3). Within the African sample, the Senegal sample had markedly fewer SNPs called per specimen than the Gabon or Kenya samples, perhaps reflecting that this is a human specialist population (Rose et al. 2020) and therefore may have experienced a bottleneck during the transition towards close proximity with humans (Crawford et al. 2017).

We also filtered genomes based on kinship coefficient after variant calling. From the American sample, ten pairs of specimens, involving 19 specimens (*i.e.*, one specimen appeared in two pairs) had a kinship coefficient greater than 0.2 (see Figure S1 for kinship coefficients between all pairs of specimens). From the African sample, four pairs involving eight specimens had a kinship coefficient greater than 0.05. We removed 13 specimens (four from Colombia, three from Brazil, two from the USA, three from Kenya, and one from Senegal), leaving a total of 118 remaining specimens.

With one exception, all pairs of specimens with a kinship coefficient greater than 0.05 were from the same country. However, one male specimen from Colombia had a kinship coefficient of 0.07 with a Brazilian specimen. While KING-robust may produce biased kinship coefficients in certain cases of admixture (Conomos et al. 2016), this finding is in keeping with previous observations of shared ancestry between South American countries (Monteiro et al. 2014, Kotsakiozi, Gloria-Soria, et al. 2018).

Patterns of genetic diversity within and between populations

To investigate the extent and patterns of population structure of our data set, we used principal component analysis. The first principal component for each chromosome explained 12-13% of the variance, and generally separated samples from Africa vs. samples from North and South America (Figure 1, Figure S2). The second principal component, explaining 4-5% of the variance, separated samples from North and South America along a cline stretching from the USA sample to the Colombian sample. The two Colombian cohorts were clearly separated, with the Brazilian sample falling in between.

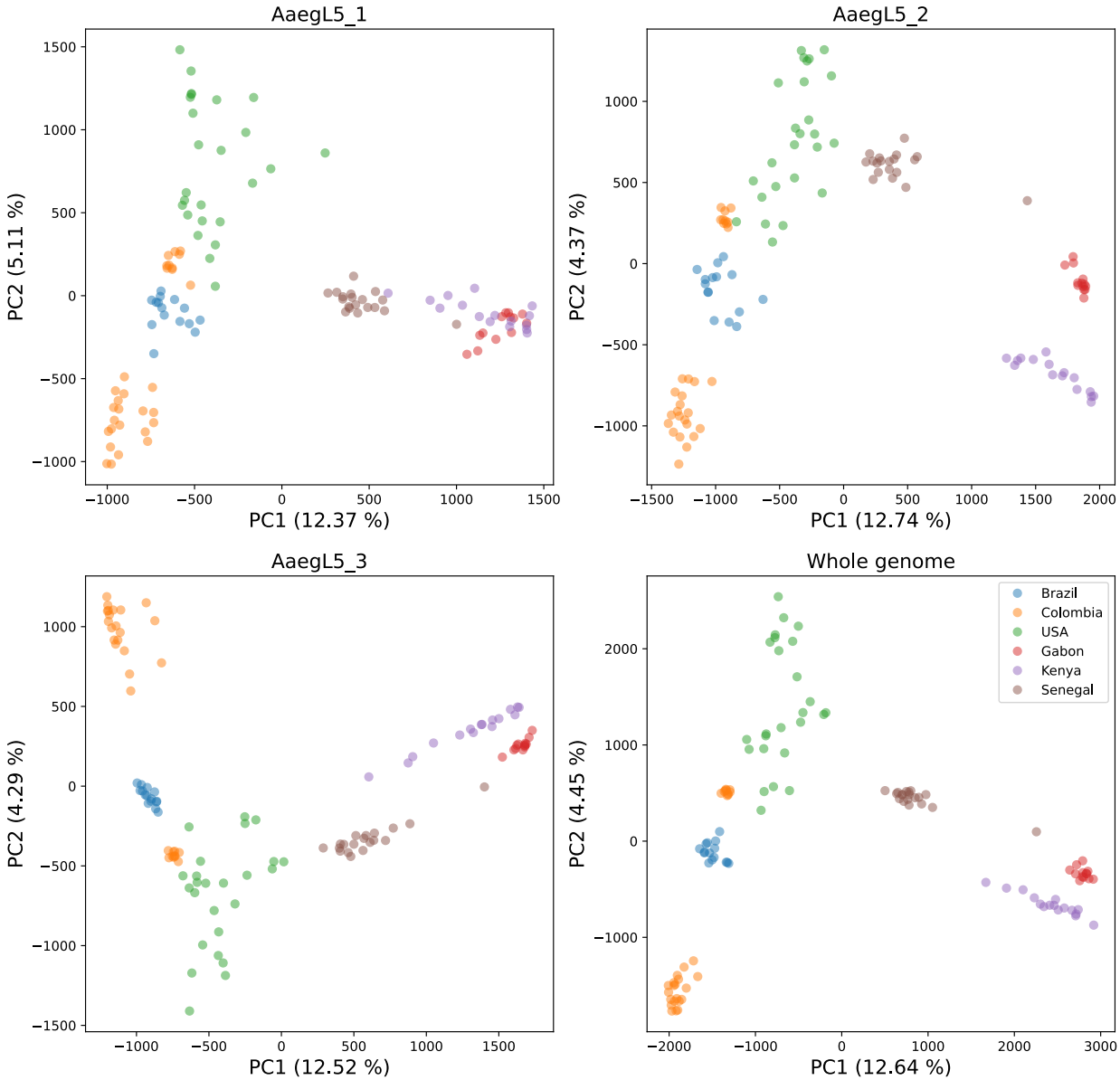

Figure 1. Principal component analysis of all filtered SNPs, showing each chromosome separately (left panels and top-right panel) and combining all three chromosomes (bottom-right panel).

On each chromosome, the Senegal sample was closer in the first principal component to the North and South American samples than were the Gabon and Kenya samples. This is in keeping with previous findings (Crawford et al. 2017, Rose et al. 2020) that have suggested a Senegalese origin of the domestic populations. On chromosome 1, the Kenya and Gabon samples are generally mixed, while they exhibit greater separation on chromosomes 2 and 3. Overall, the separation between the American and African samples is most pronounced on chromosome 1, largely due to the shifting position of the Senegal cohort on the other chromosomes: on chromosome 2, Senegal clusters particularly closely to the USA cohort, and on chromosome 3 Senegal is placed in an intermediate position relative to its position on the other two chromosomes.

Next, we quantified the extent of genetic diversity and the prevalence of rare versus common polymorphisms by measuring nucleotide diversity (π) and Tajima's *D*. Nucleotide diversity was lower in the American samples than in the African samples (0.0021 vs 0.0031 on average; Figure S3, Table S4), consistent with previous reports of greater allelic richness in African samples (Gloria-Soria, Ayala, et al. 2016). Tajima's *D* was higher in the American cohorts, consistent with the loss of rare variants caused by a population contraction. In all cohorts, nucleotide diversity tended to be lower near the telomeres, along with reductions in nucleotide diversity and Tajima's *D* observed near the centromeres (Figure S4). Tajima's *D* in the African cohorts tended to be lower towards the telomeres, a trend not observed in the American cohorts.

Next, we quantified between-population diversity using F_{ST}. As expected, mean genome-wide F_{ST} values were lowest within continents and elevated between continents (Table S5). The decreased F_{ST} between all American populations and Senegal, compared to American populations and Gabon or Kenya, are consistent with the West African origin of domestic *Ae. aegypti*.

We also measured linkage disequilibrium (LD) within 500 kb windows and found that LD in the Brazilian population was noticeably elevated on all three chromosomes, while the other five populations overlapped more closely (Figure S5, Table S6). Indeed, mean linkage disequilibrium is much higher in the Brazilian sample than in the other five populations—nearly twice as high in Colombia and Gabon, the two populations with the next-highest mean value of r^2 (Table S6).

Strong signatures of selective sweeps in insecticide resistance loci

We scanned the genome for signatures of recent selective sweeps using SweepFinder2 (DeGiorgio et al. 2016) (Methods). Strong sweep signals were observed in all populations, with the most prominent peaks in SweepFinder2's composite likelihood ratio (CLR) seen in the samples from Brazil, Colombia, and Gabon (Figure 2). CLR peaks were also observed at the beginnings or ends of chromosomes in all populations except Senegal; the size and shape of these may be a result of the paucity of recombination and the lack of polymorphisms in these regions, respectively.

We individually examined eight peaks or groups of peaks (marked with stars in Figure 2), that either contained a CLR value greater than 1,000, or had values near 1,000 in both the Americas and Africa. An example of the latter is found on chromosome 1 at approximately 209 Mb (the peak marked "*a" in Figure 2), where all cohorts except Kenya show some signal of a sweep (Fig S6a). This putative swept region overlaps 21 genes (Table S7); 19 of these code for a protein, and six are annotated with a specific predicted function, with experimental evidence available for three. The sulphate transporter gene AAEL012037 is regulated by the hormone 20-hydroxyecdysone via the ecdysone receptor (Roy et al. 2022). Another sulphate transporter, AAEL012041, is upregulated after mating (Camargo et al. 2020) . The zinc finger gene AAEL012039, part of the Hedgehog signaling pathway, is upregulated in the brains of female pupae compared to those of male pupae, and localized to the optic lobe (Tomchaney et al. 2014).

Within the American samples, the two most prominent peaks, one each on chromosomes 2 and 3, are shared to some extent by both the Brazilian and the Colombian sample. A peak on chromosome 2 ("*b"), from approximately 351 Mb to 352 Mb, overlaps 27 protein-coding genes and four non-coding RNAs (Figure 2, Fig S6b, Table S7); 18 protein-coding genes have an unspecified functional product, while the genes with specific predicted functions include an ankyrin gene (AAEL013466; discussed below) as well as a cluster of 8 glutathione transferases. Glutathione transferases have a well-known function insecticide resistance, and have been implicated in resistance to several classes of insecticides (Enayati et al. 2005).

A secondary peak is found nearby in the Brazilian sample, but not the Colombian sample, from approximately 356.3 Mb to 356.5 Mb on chromosome 2. There are one tRNA and six protein-coding genes in the vicinity of this peak (Fig S6b, Table S7), three of which have a

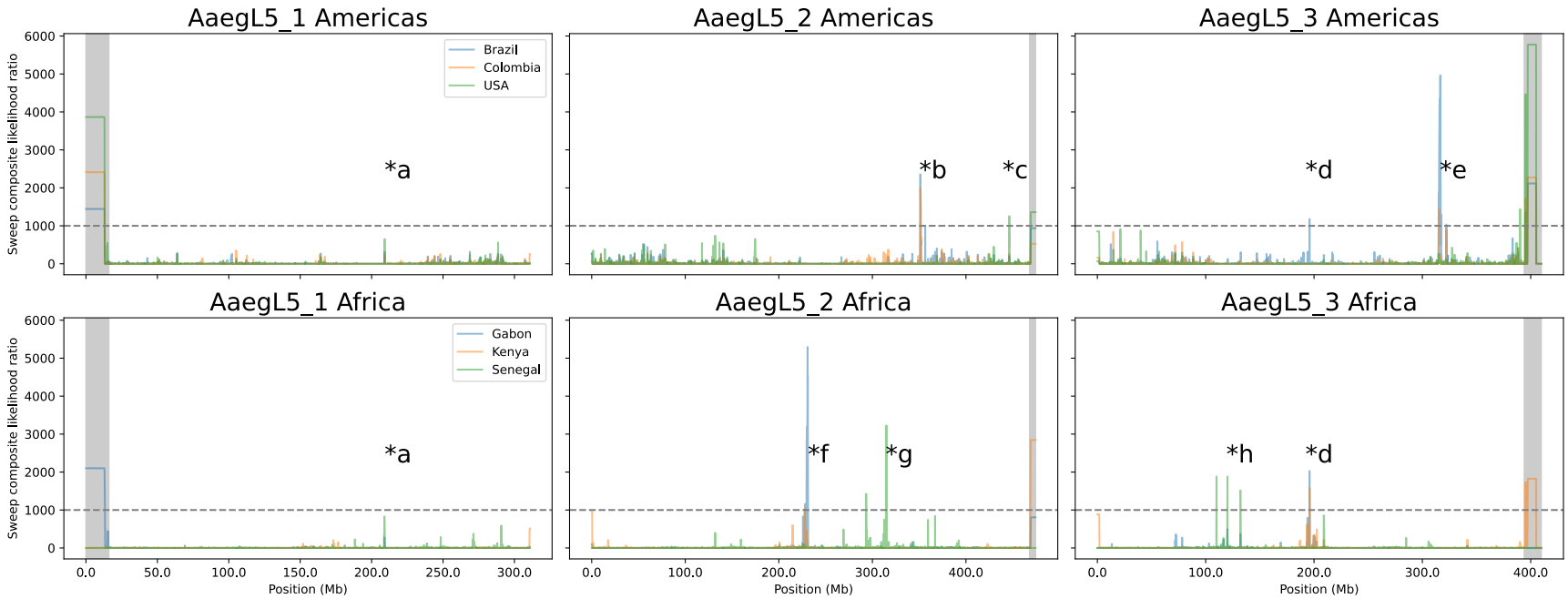

Figure 2. Genome-wide scan of candidate selective sweeps. Shaded regions at the ends of chromosomes indicate putative technical artifacts excluded from identification of outliers. Each * with a letter marks a sweep examined in more detail in the text and in the corresponding panel of Figure S6. Horizontal dashed lines indicate our threshold of CLR=1000, used in identifying sweeps examined further.

specific functional prediction. AAEL002055, a neuroendocrine protein, was previously identified as carrying a single nucleotide showing significant spatial structure within a Brazilian population (Rašić et al. 2015); *CYP4H31*/AAEL002085 is a cytochrome P450 identified as overtranscribed in larvae selected to be resistant to the neonicotinoid insecticide imidacloprid (Riaz et al. 2013). Its expression levels are altered in larvae exposed to DDT, malathion and temephos (El-Garj et al. 2016), and it has been implicated in deltamethrin resistance in *Culex* (Li et al. 2021).

Also on chromosome 2, a candidate sweep found only in the USA cohort extends from approximately 446 to 446.5 Mb, ("*c", Figure 2, Figure S6c). This region overlaps 27 genes, 18 of which are protein-coding and 8 of which have specific functional predictions (Table S7). Of these 18, three are known in the literature. The glutathione transferase *GSTX2* (AAEL010500) is upregulated in response to a variety of exposures, including to the insecticide propoxur (David et al. 2010) and to brackish water (Durant et al. 2021) in larvae, as well as to heme in a cell line and in midguts of females after a blood meal (Bottino-Rojas et al. 2015); *GSTX2* is also constitutively upregulated in a strain resistant to organophosphate and pyrethroid insecticides compared to a susceptible strain (Bottino-Rojas et al. 2018). The protein product of transcriptional regulator *ATRX* (AAEL010502) is found in eggshell preparations (Marinotti et al. 2014), and the fk506-binding protein *FKBP12* (AAEL010491) is upregulated in a population with high rates of *Wolbachia* wMEL infection (Martins et al. 2021).

On chromosome 3, Gabon, Kenya, and Brazil share a sweep at approximately 200 Mb ("*d", Figure 2). Closer examination reveals a shared peak at 196 Mb, and a peak in Kenya between 202 and 202.5 Mb, the latter of which occurs at much lower magnitude in Gabon (Figure S6d). The former peak, at 196 Mb, overlaps 11 genes, 10 of which are protein-coding and four of which have a specific predicted function. One of these, AAEL007292, is a larval lethal gene involved in the proper development of neural synapses (Hapairai et al. 2017). The latter peak overlaps one protein-coding gene and one noncoding RNA (Table S7).

The highest peak identified in any American sample occurs in the Brazilian sample, near 315 Mb on chromosome 3; this peak also occurs in the Colombian sample at a lesser magnitude ("*e", Figure 2). Closer investigation shows this peak extending from approximately 314.5 Mb to 318 Mb in the Brazilian sample (Fig S6e). This region overlaps 36 protein-coding genes and 11 non-coding RNAs (Table S7). Of the protein-coding genes, 13 have a specific functional prediction present in VectorBase; we highlight five genes that are known from the literature. The

Vanin-like protein AAEL006034 has been identified as potentially involved in metabolism of the insecticide indoxacarb (Smith et al. 2018). The calmodulin gene AAEL012326 is overexpressed in a strain susceptible to dengue compared to one refractory to dengue (Chauhan et al. 2012), and produces a protein product that is found in extracellular vesicles produced by *Ae. aegypti* cells infected by dengue, but not uninfected control cells (Gold et al. 2020). The putative 60S ribosomal protein gene AAEL011587 is upregulated in response to chikungunya infection and to chikungunya and dengue coinfection (Shrinet et al. 2017). The heme peroxidase gene *HPX1* (AAEL006014) is overexpressed in response to pyrethroid (permethrin) selection (Saavedra-Rodriguez et al. 2012).

Finally, this peak near 315 Mb on chromosome 3 also includes *Vgsc* (AAEL023266), which has been extensively associated with insecticide resistance via non-synonymous single nucleotide changes that cause resistance to pyrethroids, DDT, or both (Hu et al. 2011, Du et al. 2013, Hirata et al. 2014, Haddi et al. 2017). In the Brazilian sample, *Vgsc* is roughly centered in the putative sweep, while in our Colombian sample, the region of highest sweep likelihood is upstream of *Vgsc*. In addition to being narrower in genomic distance, the magnitude of the putative sweep is also lower in the Colombian sample than in the Brazilian one. Additionally, one of the genes identified in the prominent sweep on chromosome 2 discussed above is ankyrin gene 2,3/unc44 (AAEL013466); ankyrin proteins are associated with, and directly interact with, voltage-gated sodium channels (Bennett and Baines 2001, Bennett and Chen 2001, Jenkins and Bennett 2001). In *Ae. aegypti*, variation in ankyrin-domain-containing genes has been associated with pyrethroid resistance (Campbell et al. 2019); in addition, the ankyrin repeat region-containing gene AAEL011338 was identified as harboring a locus associated with pyrethroid resistance in a Brazilian sample (Cosme et al. 2022).

Next, we examined two peaks present in the Gabon and Kenya samples. On chromosome 2, a sweep signal present in these two samples stretches from approximately 227.5 Mb to 228.2 Mb, overlapping 9 genes ("*f", Figure 2, Figure S6f). The two genes with specific functional predictions fully encompassed by this sweep are the mitochondrial ribosomal protein L52 gene AAEL023209, and *SCRB8* (AAEL000227), a class B scavenger receptor. However, just downstream, the Gabon cohort shows another signal of a putative selective sweep, from approximately 229.5 Mb to 231.25 Mb. This peak, which is one of the strongest detected in any cohort, overlaps 39 genes, 19 of which are protein-coding and 10 of which have a specific

functional prediction (Table S7). Here, we discuss five known in the literature. *SCRB5* (AAEL011222) has been found to be upregulated in *Ae. aegypti* cells treated with human complement protein hC5a (Giraldo-Calderón et al. 2020). The gene AAEL008073, predicted to bind RNA (VectorBase release 58; (Dickson et al. 2017), is overexpressed in bloodfed *Ae. aegypti* compared to sugar-fed (Bonizzoni et al. 2011); it also displays very low heterozygosity in populations from Mexico, Thailand, and Senegal (Dickson et al. 2017). In addition, dsRNA silencing of this gene and infection with chikungunya in an *Ae. aegypti* cell line results in a significant decrease in chikungunya RNA levels (Dubey et al. 2022). When *notch* (AAEL008069), another nearby gene, is silenced, *in* vitro viral load of DENV-2 (dengue serotype 2) is significantly increased six days post-infection compared to control mosquitoes (Xiao et al. 2014). A splice variant of the gonadotropin-releasing hormone receptor *AAEL011325* is down-regulated in larvae exposed to brackish water compared to freshwater (Durant et al. 2021). Finally, the nicotinic acetylcholine receptor subunit gene *Aaea6* (AAEL008055) is upregulated in response to post-eclosion application of exogeneous juvenile hormone (Zhu et al. 2010), and has also been experimentally identified as the target of the insecticide spinosad (Lan et al. 2021); more generally, nAChR genes are known as the targets of neonicotinoid insecticides (Casida 2018).

Senegal shows a series of putative selective sweeps on chromosome 2 between 312 and 316 Mb ("*g", Figure 2, Figure S6g). Collectively, these peaks overlap 35 genes, 26 of which are protein-coding and 11 of which have a specific predicted function (Table S7). The aldo-keto reductase AAEL007275 is up-regulated after larval imidacloprid exposure (Riaz et al. 2009), and up-regulated after one generation of selection for resistance to temephos (Saavedra-Rodriguez et al. 2014). The alpha-amylases AAEL010540 and AAEL010532 contain nonsynonymous changes in putative receptors for *Bacillus thuringiensis israelensis* (Bti) Cry toxins that are nearly fixed between control strains and strains selected for resistance to individual Cry toxins (Stalinski et al. 2016). Another study showed AAEL010540 underexpression in strains resistant to certain Cry toxins compared to a control strain (Després et al. 2014). ABC transporter *ABCC13* (AAEL023524) is underexpressed in late dengue infection and overexpressed throughout yellow fever expression *in vivo* (Kumar et al. 2021). The importin alpha gene AAEL012960 is demethylated following infection with a virulent strain of *Wolbachia* (Ye et al. 2013).

In addition, five genes in this region (AAEL012960, AAEL013222, AAEL022016 [alternate gene ID: AAEL013219], AAEL010533, and AAEL023254 [previous gene ID: AAEL013215]) were identified as among the top 25 most differentiated genes between an urban population in Senegal vs. a forest population from Senegal together with a population from Uganda (Crawford et al. 2017); AAEL012960 and AAEL023254 contain exonic SNPs highly differentiated in that comparison. This sweep also overlaps a region identified as containing many variants associated with human specialist ancestry, and as highly diverged based on the Population Branch Statistic, in these specimens (Rose et al. 2020), congruent with our identification of a putative sweep.

Finally, the cohort from Senegal also shows several putative selective sweeps on chromosome 3 between 100 and 140 Mb ("*h", Figure 2, Fig S6h). The three most prominent peaks in this region overlap 25 protein coding genes, 7 of which have a specific functional prediction (Table S7). The cell division cycle 20 gene (*cdc20*, AAEL004480) is underexpressed after adult females are challenged with either *Escherichia coli* or *Micrococcus luteus* (Choi et al. 2012), and contains a SNP showing significant spatial structuring in populations in Rio de Janeiro (Rašić et al. 2015).

Selective sweep signatures are enriched for insecticide resistance genes in Colombia

In addition to examining individual sweeps, we looked for categories of genes disproportionately represented in sweeps across the genome. Here, we used the individual SweepFinder2 test sites, distributed approximately every 1 kb throughout the genome. We identified outliers in these test sites at the 99th percentile in each country, with enrichment testing performed in Gowinda with a 1 kb buffer up- and downstream of each gene (Figure S7). We found that eight GO terms were enriched in four populations, with seven (DNA binding; DNA-templated transcription, initiation; chromosome; host cell nucleus; nucleosome; nucleosome assembly; and nucleus) related to DNA or DNA binding. Further examination of the genes underlying this enrichment showed that it was partially attributable to the presence of a cluster of histone genes at approximately 340 Mb on chromosome 3, overlapped by outlier windows in the four populations (Figure S8). This highlighted the possibility of clusters of genes influencing the enrichment results, so we implemented a permutation test to test for significantly enriched GO terms after accounting for physical location (see Methods). After controlling for gene clustering this way, no GO terms

were significantly enriched at a *q* value of 0.05 when examining genes overlapping outliers windows at the 99th percentile. These results highlight the need to account for physical clustering of genes.

We also tested whether genes potentially associated with insecticide resistance genes were enriched in outlier windows. To do this, we assembled a list of 152 cytochrome P450 genes, glutathione transferase genes, and carboxylesterase genes by searching the annotated gene file available through VectorBase (release 54). We then manually added AAEL023266 (*Vgsc*), and AAEL023844 (a carboxylesterase-like gene that is part of a cluster of *CCE* genes and has been extensively associated with insecticide resistance [(Cattel, Haberkorn, et al. 2021, Cattel, Minier, et al. 2021, Sene et al. 2021)])—we note that the decision to include *Vgsc* in enrichment testing was made prior to conducting our SweepFinder2 scan. Outlier windows at the 99th percentile were significantly enriched for these genes in Colombia, but not the other five countries (however, the *p*-value for enrichment in the Brazil cohort was 0.05).

Strong but variable signatures of selection around Vgsc in South America

Given its important role in resistance to insecticides in numerous species, we conducted a more thorough analysis of *Vgsc* and the surrounding region. We began by conducting principal component analysis of SNPs inside *Vgsc* (from 315,926,360 to 316,405,639 on chromosome 3), which showed specimens clustering into an approximate triangle (Figure 3a). The point of the triangle is anchored on the right side by a cluster of Senegalese, Gabonese, and Kenyan specimens, and the *x*-axis (the first principal component, which explains 25% of the variance), approximately divides African specimens from American specimens, with, however, much intermixing. The base of the triangle, on the left, comprises American specimens only; the *y*-axis (the second principal component, explaining 15% of the variance) shows American specimens in three distinct and roughly equidistant clusters, with some spread between them. All Colombian and Brazilian specimens can be easily assigned to one of the clusters, and a subset of the specimens from the USA take intermediate positions.

We then investigated genotypes for each specimen at a number of non-synonymous SNPs previously shown to be either causally linked or associated with pyrethroid resistance (our "focal sites"): F1534C (Kawada et al. 2009, Harris et al. 2010, Hu et al. 2011, Du et al. 2013, Hirata et al. 2014), V1016I (Saavedra-Rodriguez et al. 2007, Saavedra-Rodriguez et al. 2008,

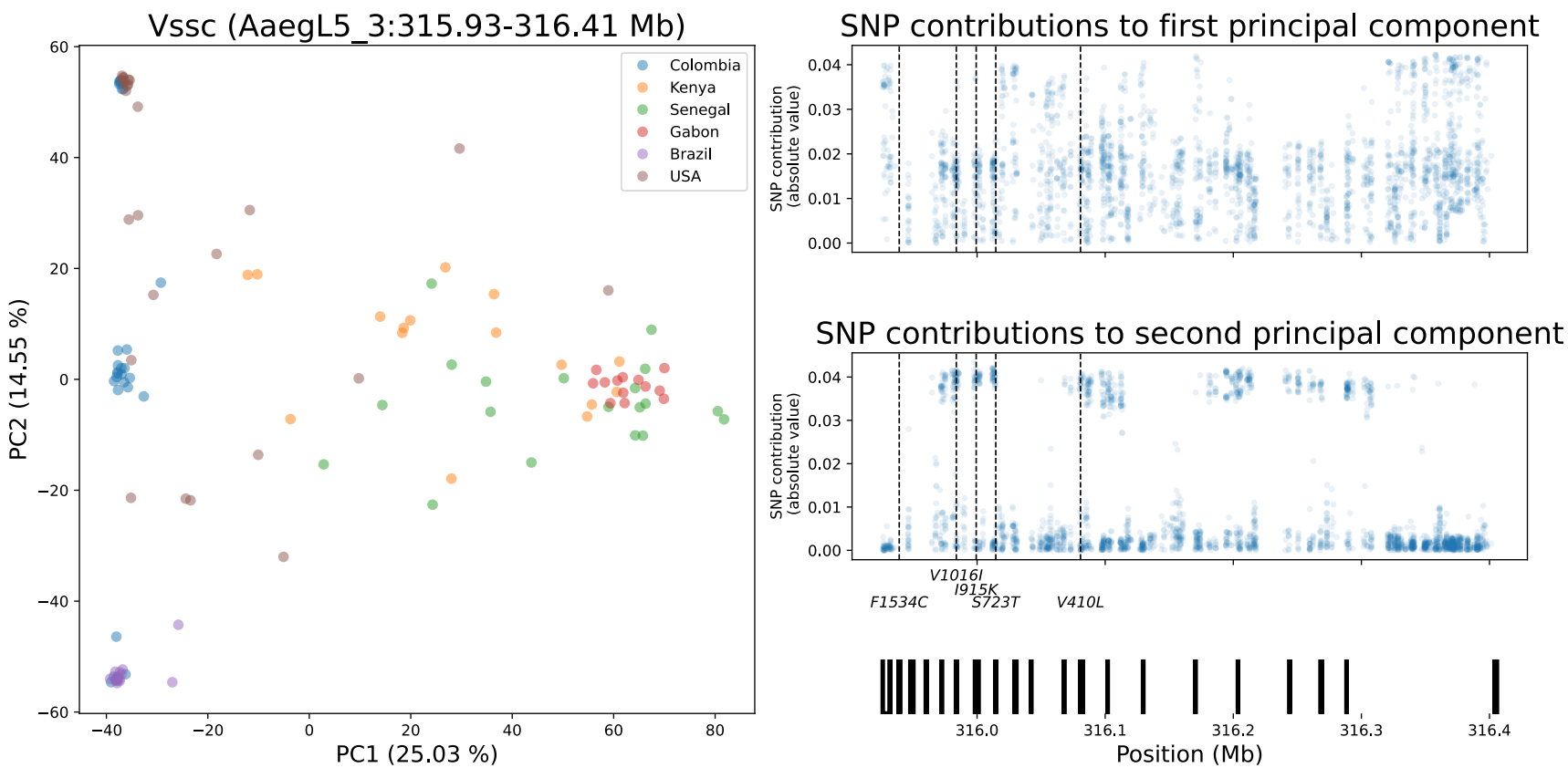

Figure 3. Principal component analysis and loadings of *Vgsc*. Vertical dashed lines on the loadings panels indicate the locations of the five focal loci listed below the plot. The vertical black bars at the bottom right show the location of exons in *Vgsc*.

Harris et al. 2010, Du et al. 2013), S723T (Saavedra-Rodriguez et al. 2019, Mack et al. 2021), V410L (Haddi et al. 2017, Saavedra-Rodriguez et al. 2019), I915K (Mack et al. 2021), and Q1853R (Kelly et al. 2021) (Table 1; we use the *Musca domestica* nomenclature). F1534C was not called in any specimen in our filtered data set, although we were later able to genotype this locus as discussed below (these genotypes were not included in the PCA but do appear in Table 1); Q1853R was confidently called as invariant in 103 specimens and either called as invariant with low confidence or uncalled in the remaining 15, and thus excluded from further analysis. The remaining four sites were confidently called and were polymorphic in our data, with high LD between them in the Colombian sample. The genotypes at these four positions also correspond to the three clusters seen on the left side of Figure 3a, with the bottom cluster corresponding to the specimens homozygous for the reference or wild-type alleles at the four sites, the top cluster corresponding to the specimens homozygous for alleles associated with resistance, and the middle group corresponding to specimens heterozygous at all four sites.

We also looked at the contribution of each SNP to the groupings in the PCA (Figure 3b), and specifically to the second principal component, along which many American specimens cluster into discrete groups. The loadings for this principal component show several clusters of SNPs with some of the highest contributions in the PCA. When we plot the location of five associated resistance SNPs, we see that, except for F1534C (which was not included in the PCA, as previously described), they correlate with clusters of SNPs that are outliers in terms of their contribution to the second principal component. In addition to the clusters associated with the known resistance loci, there are additional clusters in the proximal portion of the gene. We identified three SNPs (316268746, 316268894, and 316269032) with a PCA loading of greater than 0.03 that overlap two adjacent exons; all three SNPs are found in the untranslated region. *Vgsc* is extensively alternatively spliced (Lin et al. 2009), so these SNPs may be relevant in that regard, or may influence expression levels; however, they may also have no functional importance and simply be in high LD with the focal polymorphisms and/or another causal polymorphism not present in our filtered data set.

Examining nucleotide diversity in the region of *Vgsc*, from 310 to 320 Mb on chromosome 3, we observe a decrease in diversity in the South American populations from approximately 315 Mb to 318 Mb (Figure 4a). The specimens from the USA do not show such a drop; however, this heterogeneous cohort includes specimens with ancestry from multiple source

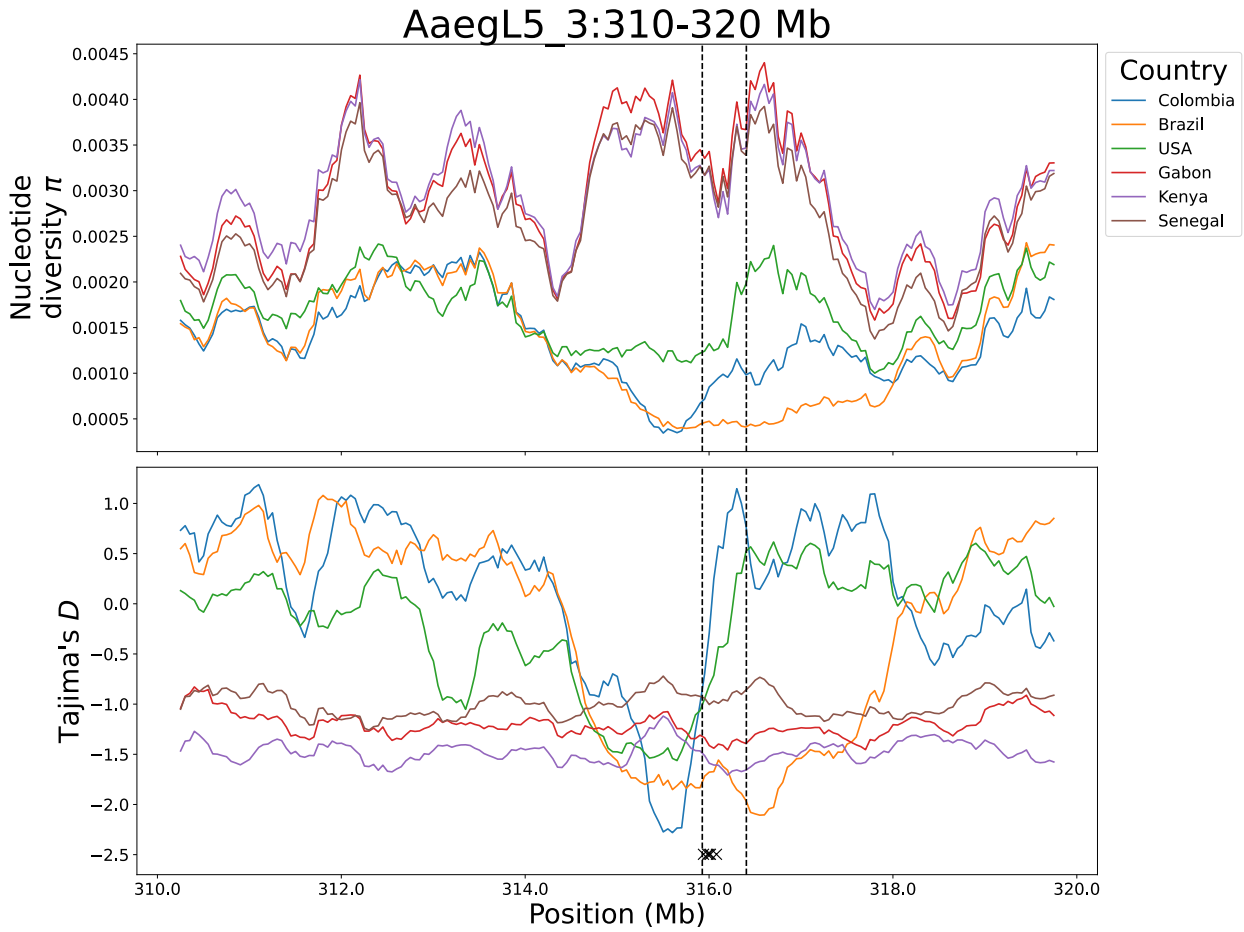

Figure 4. Nucleotide diversity and Tajima's *D* in six cohorts, on chromosome 3 from 310 to 320 Mb. The vertical dashed lines mark the boundaries of *Vgsc* (315,926,360-316,405,639). The five Xs at the bottom of the Tajima's *D* plot indicate the locations of the five focal loci.

populations (Lee et al. 2019, Kelly et al. 2021), which could give the appearance of high nucleotide diversity even if diversity is low within some ancestry groups. Relative to the Brazilian sample, the nucleotide diversity of the Colombian sample is elevated within *Vgsc* itself, and in the distal region to approximately 318 Mb, where nucleotide diversity increases in the Brazilian sample. This is in keeping with the number of clusters observed for each cohort in the PCA of the *Vgsc* region as well as the genotypes shown in Table 1: one cluster for Brazil, and three for Colombia, indicating multiple haplotypes in the latter but not the former.

Although the SweepFinder2 results suggest a strong skew in allele frequencies in this region, to more directly examine this, we calculated Tajima's *D*. In the American samples, Tajima's *D*, which is otherwise generally higher than in the African samples, drops sharply proximal to *Vgsc* (Figure 4b). This drop is especially pronounced in the Colombian sample. However, within *Vgsc* itself, Tajima's *D* is sharply elevated; this is in keeping with the presence of multiple haplotypes in this region within Colombia.

Finally, linkage disequilibrium (r^2, calculated as described in the Methods) between individual variants within *Vgsc* in the three populations from the Americas shows that regions of elevated LD are particularly prominent in the South American populations (Figure 5). One such block occurs from approximately 315.97 Mb to 316.08 Mb—we note that this block contains the V1016I, I915K, S723T, and V410L polymorphisms, which are all in high LD with one another in the Colombian sample (Table 2; compare to USA sample, Table S8). This region also shows noticeably elevated LD with the region from 316.31 Mb to the end of the gene. In addition, another block of elevated LD occurs from approximately 316.17 Mb to 316.30 Mb. The USA population, in contrast, shows a more diffuse pattern of elevated LD, spread across the gene.

These findings are particularly striking because the Brazilian sample, though wild-type for all previously identified SNPs associated with pyrethroid resistance that could be confidently genotyped in our data, still shows a sharp drop in nucleotide diversity and Tajima's *D* characteristic of a selective sweep; in addition, SweepFinder2 identified a strong candidate selective sweep in both the Brazilian and Colombian samples in this region. Indeed, SweepFinder2's CLR peak was substantially higher and broader in Brazil than Colombia. These findings, as well as the well-known role of the F1534C polymorphism in insecticide resistance, led us to re-examine this locus in our data. F1534C was also not present in the unfiltered variants, indicating this site had not been simply filtered out. In our set of unfiltered

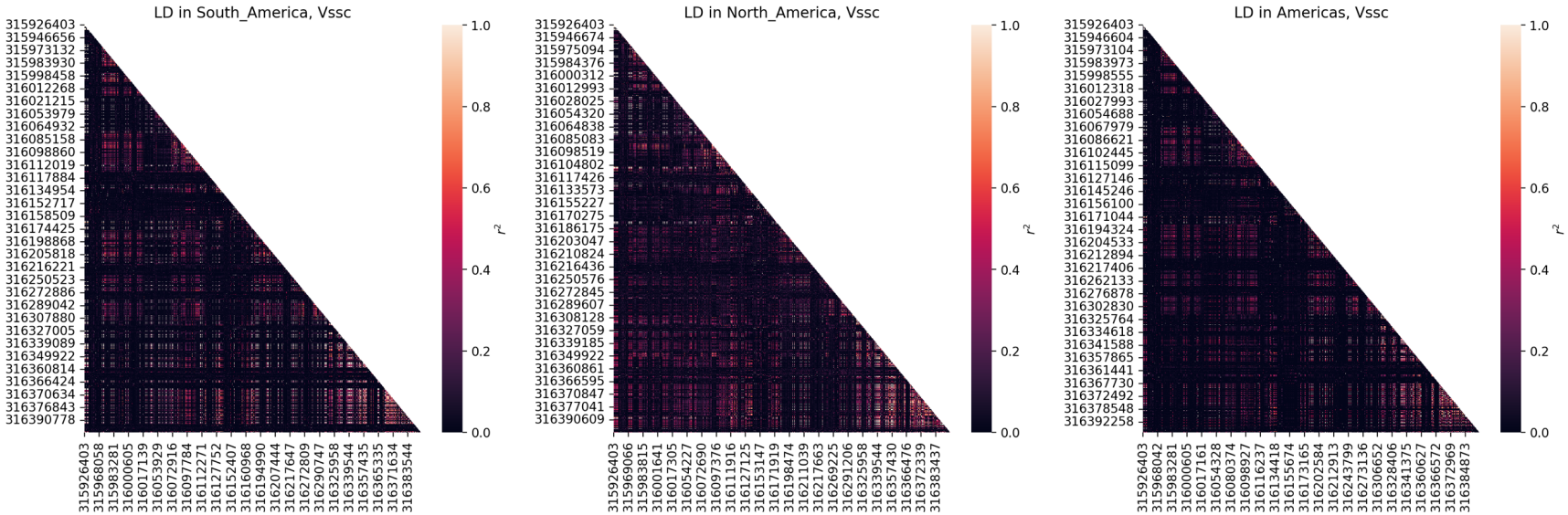

Figure 5. Linkage disequilibrium across *Vgsc* in the North American cohort (i.e. USA), the South American cohorts (Brazil and Colombia combined), and the combined American cohort (Brazil, Colombia, and USA combined).

variants, we found a total lack of coverage in all specimens in the vicinity of this SNP, spanning from approximately 315,937,705 bp to 315,945,225 bp. However, coverage in the original alignments appeared relatively normal (Figure S9a). Inspection of the alignments revealed low mapping quality in this region (Figure S9b), with a high proportion of reads aligning to this region also mapping to another region, scaffold NIGP01000811; reads that align equally well to multiple locations are given a mapping quality of 0 (Li et al. 2008) and thus are not considered by GATK's HaplotypeCaller using the default settings, which remove all reads with a mapping quality below 10 (https://gatk.broadinstitute.org/hc/en-us/articles/360036889192-MappingQualityReadFilter). The similarities between this scaffold and *Vgsc* have previously been noted (Itokawa et al. 2019). Aligning NIGP01000811 to chromosome 3 of AaegL5 revealed that most of the scaffold aligns to the vicinity of the *Vgsc* gene, but the first approximately 4.9 kb do not confidently align anywhere on chromosome 3 (Figure S10).

We therefore re-genotyped chromosome 3 from 315 to 317 Mb, as well as the entirety of NIGP01000811, using bcftools version 1.9 mpileup, which has a default minimum mapping quality of 0 (https://samtools.github.io/bcftools/bcftools.html). Using this method, we called putative genotypes at F1534C, revealing that every South American specimen is homozygous for the resistant allele (Table 1). In the Brazilian sample in particular, the genotypes are of very low quality, reflecting low confidence in the genotype calls. However, this is expected, as read mapping qualities influence the calculation of genotype likelihoods and therefore qualities (Danecek et al. 2021).

While one explanation for this apparent duplication is a technical artifact of assembly, we cannot rule out the presence of a genuine duplication, and therefore we are less confident in the genotypes at F1534C than at the other putative insecticide resistance loci; however, we present these data given the established importance of this SNP on insecticide resistance and vector control. To assess the effect of using a different variant caller on clustering observed in this region, we repeated the principal components analysis using the variants genotyped by bcftools (Figure S11); the triangular pattern observed in the calls generated by GATK is generally recapitulated, though excluding low-quality genotypes introduces a great deal of noise.

In addition, we calculated nucleotide diversity separately within two sets of South American samples: those in the top and bottom clusters of the *Vgsc* PCA (Figure S12). This reveals a sharp valley in nucleotide diversity in the bottom cluster, which contains the specimens

carrying a known resistance allele at F1534C only, and an even sharper decrease in nucleotide diversity in the top cluster, which contains specimens carrying known or putative resistance alleles at all five focal loci.

Finally, drawing firm conclusions about the North American sample published in Lee et al. (2019) (Lee et al. 2019) and analyzed here is difficult because of the multiple founding populations and small per-cluster sample sizes; however, when we decompose the portion of the sample originating in California into the clusters noted in Lee et al. (two clusters from central California, and one cluster from southern California plus one site in central California), we see decreased nucleotide diversity upstream of *Vgsc* and increased nucleotide diversity within *Vgsc* itself in two of the three clusters (Figure S13), consistent with the data observed in Colombia. In addition, Mack et al. (2021) (Mack et al. 2021) observed high frequencies of the resistant allele at the five focal loci studied here in several localities in central California; data from Kelly et al. (2021) (Kelly et al. 2021) similarly suggest an increase in the frequency of resistance alleles in Californian samples between 2014 and 2018. Finally, when examining the PCA of high-quality SNPs within *Vgsc*, a subset of the USA cohort clusters very closely with the Colombian specimens that carry all five focal insecticide resistance alleles (Figure 3). Thus, it appears likely that the resistance haplotype observed in Colombia is present in parts of North America as well.

DISCUSSION

These whole-genome data from six country-level cohorts of *Aedes aegypti* on three continents allowed us to consider the history of this species in detail. In general, our findings are consistent with the known history of this species. Tajima's *D* values are negative in all three African cohorts, while they are either positive, or very close to zero, in the three American cohorts; in addition, the Senegal population shows values of Tajima's *D* intermediate between the Gabon and Kenyan cohorts, and the American cohorts. This is in keeping with previous studies of populations from these regions (Crawford et al. 2017), and is consistent with the idea of a population bottleneck associated with domestication, perhaps followed by a second bottleneck associated with expansion out of Africa (Crawford et al. 2017). The positive Tajima's *D* in American cohorts may reflect a relatively recent contraction due to the continent-wide eradication effort, but population genetic evidence of subsequent recovery and expansion will

require larger sample sizes to detect the excess of very rare alleles produced by such a change (Tennessen et al. 2012).

Of the three American populations, Tajima's *D* is highest in the Brazilian sample; linkage disequilibrium values are also sharply elevated in this population compared to all others. Eradication of *Ae. aegypti* was claimed to be complete in Brazil (Soper 1965), but no such claim was made for Colombia. Our observations are consistent with a historical reduction in population size in Brazil, perhaps due to a founder effect during re-infestation from another country. The observation of two specimens from Brazil and Colombia, respectively, showing kinship consistent with a third-degree relationship is also consistent with previous observations of clustering between Brazilian and Colombian (Kotsakiozi et al. 2017) populations or other populations from Central and northern South America (Monteiro et al. 2014). The USA cohort has lower Tajima's *D* values, and lower linkage disequilibrium values, than the other two cohorts from the Americas; however, this sample is heterogeneous (Lee et al. 2019), so these results should be interpreted cautiously.

In addition, we present the first genome-wide scans for positive selection in *Aedes* using SNP-resolution data, identifying candidate selective sweeps in six countries. In multiple cohorts, we identify signatures of strong selective sweeps, often with CLR scores well over 1000, near loci implicated in insecticide resistance, including the voltage-gated sodium channel gene *Vgsc* (Brazil, Colombia); glutathione s-transferases (Brazil, Colombia); a cytochrome P450 (Brazil); and a nicotinic acetylcholine receptor subunit gene (Gabon).

We caution that genes associated with insecticide resistance may be disproportionately well-annotated, and are likely to be disproportionately well-investigated, because of their obvious public health implications. Some sweep signatures overlap a large number of genes, many of which whose functions are currently unannotated, making detailed conclusions about the targets of selection impossible; in addition, the peak of a sweep signature may not always correspond to the target of selection (Schrider et al. 2015). However, insecticide usage is known to cause strong selective pressure in many other insects (Garud et al. 2015, Miles et al. 2017, Walsh et al. 2018, Weedall et al. 2020, Xue et al. 2021, Pélissié et al. 2022), and the variety of genes implicated in insecticide resistance occurring near these sweeps is a potential sign of *Ae. aegypti*'s agility in responding to these insecticides.

The putative selective sweep that is perhaps the most interesting, because of the contrasting patterns shown in two different South American countries, is at *Vgsc*. Examining this genomic region in Colombia reveals two observations: one, somewhat unexpectedly, the putative selective sweep in this population appears to be localized upstream of the gene, not within *Vgsc* itself, based on results from SweepFinder2 as well as elevated Tajima's *D* within *Vgsc*. Two, the clustering of the Colombian sample into three tight groups along principal components suggests that two long-range haplotypes linking several loci associated with insecticide resistance are present in the population, with one haplotype containing the resistant alleles at at least four previously identified insecticide resistance SNPs, and the other containing the wild-type alleles. Indeed, examining the PCA loadings shows that these polymorphisms line up very well with clusters of SNPs driving the separation between PCA clusters. We note that this pattern of three equidistant clusters on a PCA is the characteristic signature of a polymorphic chromosomal inversion (Ma and Amos 2012, Love et al. 2019); however, chromosomal inversions are difficult to detect *de novo* from short read data, and long-read data are likely necessary to either confirm or rule out the possibility of an inversion in this region.

Examining this same region in the Brazilian sample reveals a very different picture. Here, the putative selective sweep is centered around *Vgsc*; however, PCA results, as well as genotypes at individual loci associated with resistance, suggest that the haplotype of resistant alleles is absent in this population. Indeed, all the Brazilian specimens carry the wild-type susceptible alleles at four of the five insecticide resistance loci examined. At the fifth loci, F1534C, a partial duplication of the *Vgsc* gene in the assembly prevents confident genotyping; however, if we assume that this duplication is technical in origin and that reads mapping to the duplicated region should instead be assigned to *Vgsc* itself, we find that the Brazilian sample is completely homozygous for the resistant allele at this locus (Table 1). This suggests that the strong selective sweep observed in this population is driven either by F1534C, by other loci that have yet to be identified, or both, but not by V410L, S723T, I915K, or V1016I, all of which are absent from our Brazilian cohort but present in the adjacent country of Colombia.

We hypothesize that the Brazilian and Colombian samples may in fact have been subject to the same selective sweep at this locus, despite the absence of the signal of such a sweep in *Vgsc* itself in the Colombian sample. Our proposed explanation to this apparent discrepancy involves the putative long-range haplotypes present in the Colombian population. The presence

of two such haplotypes, both at intermediate frequency, would necessarily skew the SFS toward intermediate-frequency alleles, thereby increasing Tajima's *D* in this region. In effect, we suggest that in Colombia, a second, and ongoing, selective sweep confined to only part of the *Vgsc* gene, has partially eroded the signal of an earlier sweep that encompassed the entire gene and some of the flanking regions. Alternatively, it could be that this haplotype is being maintained by some form of balancing selection in the Colombian sample, or its intermediate frequency has resulted from an influx of migrant haplotypes from another population, rather than ongoing positive selection. However, the high LD between V410L, S723T, I915K, or V1016I, as well as the lack of diversity among both the resistant and wild-type backgrounds (Fig S11), seem to contradict these alternative explanations; we also note that V410L and V1016I have previously been noted to be in LD with each other in other locations (Vera-Maloof et al. 2015, Saavedra-Rodriguez et al. 2018, Ayres et al. 2020, Baltzegar et al. 2021). If our hypothesis is correct, the frequency of the haplotype bearing these resistance mutations may continue to increase in the coming years under continuing selective pressure from insecticide use. Indeed, the Brazilian sample is from 2012, so allele frequencies at these loci may be very different in the present day, as the resistant alleles at V1016I and F1534C are common (Linss et al. 2014, David et al. 2018, Melo Costa et al. 2020, Cosme et al. 2022) and increasing in frequency (Linss et al. 2014) in some parts of Brazil. This situation highlights the need for an increased use of temporal data in vector genomics.

An apparent sweep in the vicinity of *Vgsc* has been observed previously in a sample from southern Mexico (Saavedra-Rodriguez et al. 2019); in addition, data from North America must be interpreted with caution for reasons previously discussed, but are consistent with reduced diversity in the vicinity of *Vgsc* and an increase in insecticide resistance alleles over time. Taken all together, these data suggest this long-range haplotype may not be limited to Colombia, and could be spreading more widely across the Americas.

An earlier origin of F1534C has been posited before (Saavedra-Rodriguez et al. 2018), and would be consistent with previous observations that F1534C is more commonly found in the absence of V410L (Haddi et al. 2017, Melo Costa et al. 2020) and V1016I (Linss et al. 2014, Vera-Maloof et al. 2015, Melo Costa et al. 2020) than vice versa; in addition, in Iquitos, Peru, F1534C was directly observed to reach near-fixation before V1016I became prevalent (Baltzegar et al. 2021). These findings also lead us to speculate about the catalyst for this putative earlier

sweep. Pyrethroids and DDT both target the voltage-gated sodium channel, and F1534C is known to confer resistance to DDT as well as to type I pyrethroids (Harris et al. 2010, Hu et al. 2011, Du et al. 2013); cross-resistance to both insecticides has been observed in *Ae. aegypti* (Chadwick et al. 1977). Populations of *Ae. aegypti* across Colombia are still highly resistant to DDT (Fonseca-González et al. 2011), though its use for vector control has been banned in that country since 1993 (Fonseca-González et al. 2011). Though the F1534C allele is present at high frequency in the USA sample and a subset of the sample shows reduced nucleotide diversity in this region, we do not see a putative selective sweep at this locus at anywhere near the magnitude of the other American populations. Although it is possible that heterogeneity within the USA sample may be somewhat masking the signature of a strong sweep, it is also possible that that the selective pressure on the sweeping allele (whether it was F1534C or otherwise), in populations in the USA, may not have been strong enough to drive one haplotype to high frequency as quickly as in South America. Unlike in South and Central America, continent-wide eradication of *Ae. aegypti* was never attempted in North America. While we can only speculate that the apparent sweep of F1534C in South America may be tied to the attempted eradication with DDT, it is worth noting that DDT, not pyrethroids, are likely to have presented the strongest anthropogenic selection pressure these populations have ever experienced.

Conclusions

Here we describe a dataset of 30 *Aedes aegypti* genomes collected from Colombia. We combined these data with whole genomes from five other countries from Africa and the Americas to examine the genomic impacts of population structure, potential past demographic events, and natural selection on this important vector species. While this analysis has revealed some clues about potential recent demographic events (e.g. domestication-associated bottlenecks, and recent migrations), future work with broader sampling should test these possibilities explicitly. Perhaps our clearest contribution is the strong signatures of positive selection that we identify around numerous insecticide resistance genes. These results paint a compelling picture of strong selective pressures, and a population that appears to be able to rapidly adapt to these pressures in part due to a fairly large number of genes that can potentially confer resistance phenotypes; these results are consistent with previous observations that a variety of loci and pathways underlie *Ae. aegypti*'s response to insecticides in general and pyrethroids in particular (Saavedra-Rodriguez et

al. 2008, Saavedra-Rodriguez et al. 2012, Paris et al. 2013, Saavedra-Rodriguez et al. 2014). We also note that because *Aedes agypti* has a large and expanding global population in terms of its census size, adaptation may not be mutation-limited even at individual sites. For example, two different resistance mutations at the same codon in *Vgsc*, V1016G and V1016I, are prevalent in Asia and the Americas, respectively, and at least one mutation in *Vgsc*, F1534C, appears to have occurred recurrently (Cosme et al. 2020). Moreover, the ability to occasionally travel long distances through the aid of human transportation may allow for the sharing of resistance alleles across populations. Thus, we argue that continued efforts to identify selected alleles in *Aedes aegypti* populations and to monitor their spread will be vital for informing vector control efforts.

METHODS

Sample collection

We sampled two urban locations in Medellín (Comuna 16 Belen, Comuna 13 San Javier) and in a rural locality near the nature reserve area of Río Claro in the department of Antioquia. Medellín has an average altitude of 1,500 meters, and a mean annual temperature of 22.5ºC. Río Claro has an average altitude of 1,750 meters, and a mean annual temperature of 22ºC. We also collected samples from Cali, capital of Valle del Cauca department (3.3817 N, -76.55237 W). Cali has an average altitude of 1,479 meters, average temperature of 24 °C, and average annual precipitation of ~1,050 mm. The Cañón del Río Claro Nature Reserve is located between the municipalities of Sonsón, Puerto Triunfo and San Luis, to the southwest of the department of Antioquia (5.90 N, -74.86 W ; (Vivero et al. 2010)).

Our collection focused on *Ae. aegypti* larval stages, and their main potential breeding grounds (pools and tanks used to store drinking water) were inspected. Larvae were then transported in Whirl-Pak plastic bags to the insectary of the Program for the Study and Control of Tropical Diseases (PECET) of the University of Antioquia, Medellín, Colombia (27 ± 2 ° C; relative humidity 80 ± 10%; photoperiod: 12 light hours), where stages III and IV were individualized and their breeding continued for entomological series, reference collection and molecular processing of adults. Larvae, pupae and adults were sorted following the protocols suggested by the Biosystematics Unit, Division of Entomology, Walter Reed 1997. To assign to species (*Ae. aegypti*), we used external morphological characters of larvae, pupae and adult females.

Ethics statement

All specimen collections were performed following the guidelines of the Colombian decree N° 1376. No specific permits were required for this study. Mosquitoes were collected on private property and permission was received from landowners before sampling.

DNA extraction and sequencing

We extracted genomic DNA from individual wild-caught adult *Aedes* in Colombia using Genetra Puregene Tissue Kits (Qiagen, Valencia, CA, USA), constructed barcoded libraries for sequencing using KAPA HyperPrep kits (Roche Sequencing, Pleasanton, CA) with a target fragment size of 500 bp. We assessed the size of the library inserts using a Tapestation 4150 (Agilent, Santa Clara, CA) and dsDNA screentape (Agilent, #5067-5365). Finally, we pooled the samples and sequenced them in a NovaSeq6000 S2 (Illumina) machine, generating 2×150bp reads. The target coverage was 15X per sample. This effort yielded 40 sequenced genomes.

To these specimens we added publicly available sequence data (Lee et al. 2019, Rose et al. 2020), as previously noted.

Quality control, alignment, and variant genotyping

We used FASTQC 0.11.9 (Andrews et al. 2010) to check initial read quality, and Trimmomatic 0.39 (Bolger et al. 2014) to remove adapters and trim regions of low quality at the beginnings and ends of reads. After trimming, we re-checked read quality using FASTQC.

For newly sequenced specimens, read group information was extracted using bash scripting. Reads were aligned to AaegL5, retrieved from VectorBase (Matthews et al. 2018, Giraldo-Calderón et al. 2022) using the mem tool of bwa-mem2 version 2.1 (Vasimuddin et al. 2019), converted to BAM format, and, for specimens split across multiple lanes, merged across lanes using samtools version 1.11 (Danecek et al. 2021).

We used MarkDuplicatesSpark in GATK version 4.1.9.0 (DePristo et al. 2011, Poplin et al. 2018) to mark optical duplicates and sort reads. BAM files were then indexed with samtools index. Quality of the resulting BAM files was assessed with qualimap version 2.2.2d (Okonechnikov et al. 2016) bamqc and with picard CollectMultipleMetrics.

Sites were genotyped with GATK HaplotypeCaller (Poplin et al. 2018) independently for each specimen, parallelized across chromosomes, using the EMIT_ALL_CONFIDENT_SITES flag. The resulting gVCF files were merged with picard GatherVcfs and indexed with GATK IndexFeatureFile. Quality metrics for each specimen were compiled from FASTQC, picard, and qualimap using multiqc version 1.9 (Ewels et al. 2016).

Genomic VCF files for each specimen were then turned into batch variant calls using GATK. Specifically, specimens were compiled using GenomicsDBImport and genotyped using GenotypeGVCFs with the include-non-variant-sites flag, parallelized across the genome as necessary for efficiency. The resulting VCFs were merged using picard GatherVcfs and filtered to SNPs only using the view command from bcftools 1.8 and 1.9 (Danecek et al. 2021). The quality control and alignment through individual variant calling steps, and the batch genotyping steps, were implemented as two nextflow v. 20.10.0 (Di Tommaso et al. 2017) workflows, available on GitHub (https://github.com/rrlove/aedes_pipeline).

The merged VCFs were converted into zarr format using scikit-allel 1.3.2 (DOI:10.5281/zenodo.4759368). Variants were filtered in python using scikit-allel. For our focal sample from Colombia, we removed specimens with a mean coverage below 15x or fewer than 70% of reads mapping to the reference genome. In the combined sample, we also removed invariant sites, and sites with indels. Based on GATK best practices for hard filtering (https://gatk.broadinstitute.org/hc/en-us/articles/360035890471-Hard-filtering-germline-short-variants), the following six hard filtering criteria were implemented, with sites required to pass five of the six filters: 1) variant quality by depth (QD) greater than or equal to 2; 2) variant strand bias as estimated by Fisher's exact test (FS) less than 40; 3) variant strand bias as estimated by the symmetric odds ratio test (SOR) less than 4; 4) mapping quality (MQ) greater than or equal to 40; 5) mapping quality rank sum test (MQRS) between -5 and 5, inclusive; 6) site position within reads rank sum test (RPRS) between -3 and 3, inclusive. In addition, sites were removed if less than 2/3 of specimens were called with a genotype quality of 20 or greater.

To identify and remove closely related specimens from our sample, we calculated relatedness using KING-robust (Manichaikul et al. 2010) with the –kinship flag for all three chromosomes. To remove any possible interference of the *m* locus and surrounding region (Fontaine et al. 2017, Matthews et al. 2018) with inferred kinship coefficients, we repeated this analysis using chromosomes 2 and 3, only. Following (Rose et al. 2020), we used a kinship

coefficient cutoff of 0.20 in American specimens and 0.05 in African specimens, to account for higher levels of inbreeding found in American specimens. If a pair of specimens had a kinship greater than this threshold, we removed the specimen with lower coverage.

Patterns of genetic diversity within and between populations

We visualized genetic structure in our sample using principal components analysis on each chromosome and on the three chromosomes together. SNPs were excluded if their minor allele frequency was 5% or less in the total sample. Genotypes were converted to the number of alternate genotypes present (0 for homozygous reference, 1 for heterozygous, 2 for homozygous alternate), and principal component analysis was done using scikit-allel. Plots were visualized using the first two principal components.

We also examined patterns of genetic diversity and divergence across the genome. Nucleotide diversity was calculated in nonoverlapping 500 kb windows for each of the six countries using the windowed_diversity function in scikit-allel. Tajima's *D* was calculated similarly, using the windowed_tajima_d function. For visualization, both statistics were also calculated in sliding 5 Mb windows with a step size of 500 kb. F_{ST} between each pair of countries was calculated using scikit-allel's moving_weir_cockerham_fst function, using nonoverlapping windows of 500,000 variants. For computational tractability, linkage disequilibrium (LD) was calculated in parallel in 50 Mb genomic regions (or one smaller segment at the end of each chromosome). In each region, linkage disequilibrium was calculated within each cohort in 500 kb nonoverlapping windows. Linkage disequilibrium was summarized by r^2 as calculated using scikit-allel's windowed_r_squared function, which uses the method of Rogers and Huff (Rogers and Huff 2009), which does not require phasing. Before calculations, genotypes with a genotype quality below 20 were masked; then, non-segregating variants were removed (these might occur because a position is invariant, or is masked due to low genotype quality, in the subset of specimens selected for analysis).

Scanning for selective sweeps

Putative selective sweeps were identified in each cohort using SweepFinder2 (DeGiorgio et al. 2016), using the empirical site frequency spectrum (SFS) generated from all SNPs segregating in the focal cohort as the background SFS, as well as a genetic map (Matthews et al. 2018). For

computational tractability, SweepFinder2 was run in chunks of 50 Mb (or, at the end of each chromosome, one segment containing the remainder), with a grid size of 1000, representing a test site approximately every kilobase. On each chromosome, the 50 Mb chunks overlapped by 5 Mb, and results were plotted to visually check for edge effects before merging. Finally, results for each test site were averaged within 100 kb and 10 kb nonoverlapping windows, for visualization at different scales.

Enrichment analysis

We used functional enrichment analysis with Gowinda version 1.12 (Kofler and Schlötterer 2012) to look for overrepresented Gene Ontology (GO) terms annotated to genes overlapping putative selective sweeps identified by SweepFinder2. First, the loci (here a "locus" refers to an individual test site, spaced approximately every 1 kb) with sweep likelihoods in the top 1% of all loci were identified. These loci were used as the candidate site file for Gowinda, with all loci queried by sweepfinder, rounded to the nearest base pair, used as the total site file. We included 1,000 bases up- and downstream from each gene to account for possible sweeps in regulatory regions. For each cutoff we ran 1,000,000 simulations, using the "gene" mode, with the following parameters: —min-significance 1 —min-genes 1 —gene-definition gene —gene-definition updownstream 1000. The gene set file was obtained by collapsing entries in the VectorBase *Ae. aegypti* version 54 gene annotation file by GO term.

Many of the genes in the gene sets with the lowest *q*-values were found in clusters, suggesting the potential for physical proximity and linkage disequilibrium to influence these results. Therefore, we implemented an additional enrichment test (https://github.com/SchriderLab/permEnrichmentTest) based on the approach of permuting the association between scores and tested loci. Specifically, this relationship was permuted by joining the three chromosomes into an artificial circular genome, then "rotating" all scores around the circle by a random interval. By preserving relationships between adjacent loci, this approach offers the ability to control for linkage. As before, we ran this enrichment analysis using the top 1% of all tested sites. For each cutoff, we ran 10,000 permutations of the SweepFinder2 scores before asking which GO terms were enriched in the sweep set relative to the permuted set. As in our analysis with Gowinda, we extended genes by 1,000 base pairs in

either direction before determining which genes contained a putative sweep in both the real and permuted SweepFinder2 scores.

Analysis of the Vgsc region

Vgsc has long been studied in *Ae. aegypti* because of its relevance to insecticide resistance; therefore, we examined this gene and its surroundings. In addition to the genome-wide approach already detailed, we used bcftools mpileup and call, with the multiallelic variant calling model, to genotype variants present between 315 and 317 Mb of AaegL5_3, and on scaffold NIGP01000811. To visualize the relationship between scaffold NIGP01000811 and chromosome 3, we aligned the two using mummer 4.0.0rc1 (Marçais et al. 2018) and visualized the alignment using Ribbon (Nattestad et al. 2021).

We performed principal component analysis as previously described on all SNPs within the bounds of *Vgsc* (AAEL023266; AaegL5_3:315,926,360-316,405,639). In addition, we visualized the "loadings" of the first two principal components, plotted against each SNP's chromosomal position. We also calculated nucleotide diversity (π) and Tajima's *D* in the region spanning 310-320 Mb on chromosome 3, in 500 kb windows slid 50 kb.

In addition, we calculated linkage disequilibrium between individual variants, again using Rogers and Huff's (2009) (Rogers and Huff 2009) estimate of r^2, within *Vgsc* and 100 kb up- and downstream, after first masking low-quality genotypes as previously described. We calculated LD within all South American specimens, all North American specimens, and then all American specimens combined.

DATA AND CODE AVAILABILITY

The nextflow pipelines used to align reads, call variants, and genotype variants can be found at https://github.com/rrlove/aedes_pipeline. Additional Python code used in data analysis can be found at <link>. Newly generated sequences from Colombia can be found at BioProject PRJNA864744.

ACKNOWLEDGEMENTS

We thank Dr. Ben Matthews for sharing an *Aedes aegypti* genetic map. RRL and DRS were supported by the National Institutes of Health under award number R35GM138286. DRM was supported by National Institutes of Health under award number R01GM125715.

REFERENCES

Aguirre-Obando OA, Bona ACD, Duque L. JE, Navarro-Silva MA. 2015. Insecticide resistance and genetic variability in natural populations of Aedes (Stegomyia) aegypti (Diptera: Culicidae) from Colombia. Zool Curitiba. 32(1):14–22. doi:10.1590/S1984-46702015000100003.

Al Nazawi AM, Aqili J, Alzahrani M, McCall PJ, Weetman D. 2017. Combined target site (kdr) mutations play a primary role in highly pyrethroid resistant phenotypes of Aedes aegypti from Saudi Arabia. Parasit Vectors. 10(1):1–10. doi:10.1186/s13071-017-2096-6.

Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Aubry F, Dabo S, Manet C, Filipović I, Rose NH, Miot EF, Martynow D, Baidaliuk A, Merkling SH, Dickson LB, et al. 2020. Enhanced Zika virus susceptibility of globally invasive Aedes aegypti populations. Science. 370(6519):991–996. doi:10.1126/science.abd3663.

Ayres CFJ, Seixas G, Borrego S, Marques C, Monteiro I, Marques CS, Gouveia B, Leal S, Troco AD, Fortes F, et al. 2020. The V410L knockdown resistance mutation occurs in island and continental populations of Aedes aegypti in West and Central Africa. PLoS Negl Trop Dis. 14(5):e0008216. doi:10.1371/journal.pntd.0008216.

Baltzegar J, Vella M, Gunning C, Vasquez G, Astete H, Stell F, Fisher M, Scott TW, Lenhart A, Lloyd AL, et al. 2021. Rapid evolution of knockdown resistance haplotypes in response to pyrethroid selection in Aedes aegypti. Evol Appl. 14(8):2098–2113. doi:10.1111/eva.13269.

Bennett KL, Shija F, Linton Y-M, Misinzo G, Kaddumukasa M, Djouaka R, Anyaele O, Harris A, Irish S, Hlaing T, et al. 2016. Historical environmental change in Africa drives divergence and admixture of Aedes aegypti mosquitoes: a precursor to successful worldwide colonization? Mol Ecol. 25(17):4337–4354. doi:10.1111/mec.13762.

Bennett V, Baines AJ. 2001. Spectrin and Ankyrin-Based Pathways: Metazoan Inventions for Integrating Cells Into Tissues. Physiol Rev. 81(3):1353–1392. doi:10.1152/physrev.2001.81.3.1353.

Bennett V, Chen L. 2001. Ankyrins and cellular targeting of diverse membrane proteins to physiological sites. Curr Opin Cell Biol. 13(1):61–67. doi:10.1016/S0955-0674(00)00175-7.

Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, Drake JM, Brownstein JS, Hoen AG, Sankoh O, et al. 2013. The global distribution and burden of dengue. Nature. 496(7446):504–507. doi:10.1038/nature12060.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 30(15):2114–2120. doi:10.1093/bioinformatics/btu170.

Bonizzoni M, Dunn WA, Campbell CL, Olson KE, Dimon MT, Marinotti O, James AA. 2011. RNA-seq analyses of blood-induced changes in gene expression in the mosquito vector species, Aedes aegypti. BMC Genomics. 12(1):1–13. doi:10.1186/1471-2164-12-82.

Bottino-Rojas V, Talyuli OAC, Carrara L, Martins AJ, James AA, Oliveira PL, Paiva-Silva GO. 2018. The redox-sensing gene Nrf2 affects intestinal homeostasis, insecticide resistance, and Zika virus susceptibility in the mosquito Aedes aegypti. J Biol Chem. 293(23):9053–9063. doi:10.1074/jbc.RA117.001589.

Bottino-Rojas V, Talyuli OAC, Jupatanakul N, Sim S, Dimopoulos G, Venancio TM, Bahia AC, Sorgine MH, Oliveira PL, Paiva-Silva GO. 2015. Heme Signaling Impacts Global Gene Expression, Immunity and Dengue Virus Infectivity in Aedes aegypti. PLOS ONE. 10(8):e0135985. doi:10.1371/journal.pone.0135985.

Brengues C, Hawkes NJ, Chandre F, McCarroll L, Duchon S, Guillet P, Manguin S, Morgan JC, Hemingway J. 2003. Pyrethroid and DDT cross-resistance in Aedes aegypti is correlated with novel mutations in the voltage-gated sodium channel gene. Med Vet Entomol. 17(1):87–94. doi:10.1046/j.1365-2915.2003.00412.x.

Brennan SA, Grob IC, Bartz CE, Baker JK, Jiang Y. 2021. Displacement of Aedes albopictus by Aedes aegypti in Gainesville, Florida. J Am Mosq Control Assoc. 37(2):93–97. doi:10.2987/20-6992.1.

Brown JE, Evans BR, Zheng W, Obas V, Barrera-Martinez L, Egizi A, Zhao H, Caccone A, Powell JR. 2014. Human Impacts Have Shaped Historical and Recent Evolution in Aedes Aegypti, the Dengue and Yellow Fever Mosquito. Evolution. 68(2):514–525. doi:10.1111/evo.12281.

Brown JE, McBride CS, Johnson P, Ritchie S, Paupy C, Bossin H, Lutomiah J, Fernandez-Salas I, Ponlawat A, Cornel AJ, et al. 2011. Worldwide patterns of genetic differentiation imply multiple 'domestications' of *Aedes aegypti* , a major vector of human diseases. Proc R Soc B Biol Sci. 278(1717):2446–2454. doi:10.1098/rspb.2010.2469.

Camargo C, Ahmed-Braimah YH, Amaro IA, Harrington LC, Wolfner MF, Avila FW. 2020. Mating and blood-feeding induce transcriptome changes in the spermathecae of the yellow fever mosquito Aedes aegypti. Sci Rep. 10(1):14899. doi:10.1038/s41598-020-71904-z.

Campbell CL, Saavedra-Rodriguez K, Kubik TD, Lenhart A, Lozano-Fuentes S, Iv WCB. 2019. Vgsc-interacting proteins are genetically associated with pyrethroid resistance in Aedes aegypti. PLOS ONE. 14(1):e0211497. doi:10.1371/journal.pone.0211497.

Casida JE. 2018. Neonicotinoids and Other Insect Nicotinic Receptor Competitive Modulators: Progress and Prospects. Annu Rev Entomol. 63(1):125–144. doi:10.1146/annurev-ento-020117-043042.

Cattel J, Haberkorn C, Laporte F, Gaude T, Cumer T, Renaud J, Sutherland IW, Hertz JC, Bonneville J, Arnaud V, et al. 2021. A genomic amplification affecting a carboxylesterase gene cluster confers organophosphate resistance in the mosquito *Aedes aegypti* : From genomic characterization to high-throughput field detection. Evol Appl. 14(4):1009–1022. doi:10.1111/eva.13177.

Cattel J, Minier M, Habchi-Hanriot N, Pol M, Faucon F, Gaude T, Gaborit P, Issaly J, Ferrero E, Chandre F, et al. 2021. Impact of selection regime and introgression on deltamethrin resistance in the arbovirus vector Aedes aegypti – a comparative study between contrasted situations in New Caledonia and French Guiana. Pest Manag Sci. 77(12):5589–5598. doi:10.1002/ps.6602.

Chadwick PR, Invest JF, Bowron MJ. 1977. An Example of Cross-resistance to Pyrethroids in DDT-resistant *Aedes aegypti*. Pestic Sci. 8(6):618–624. doi:10.1002/ps.2780080605.

Chauhan C, Behura SK, deBruyn B, Lovin DD, Harker BW, Gomez-Machorro C, Mori A, Romero-Severson J, Severson DW. 2012. Comparative Expression Profiles of Midgut Genes in Dengue Virus Refractory and Susceptible Aedes aegypti across Critical Period for Virus Infection. PLOS ONE. 7(10):e47350. doi:10.1371/journal.pone.0047350.

Choi Y-J, Fuchs JF, Mayhew GF, Yu HE, Christensen BM. 2012. Tissue-enriched expression profiles in Aedes aegypti identify hemocyte-specific transcriptome responses to infection. Insect Biochem Mol Biol. 42(10):729–738. doi:10.1016/j.ibmb.2012.06.005.

Christophers (Sir) SR. 1960. Aedes aegypti: the yellow fever mosquito. CUP Archive.

Clarkson CS, Miles A, Harding NJ, O'Reilly AO, Weetman D, Kwiatkowski D, Donnelly MJ, Consortium TA gambiae 1000 G. 2021. The genetic architecture of target-site resistance to pyrethroid insecticides in the African malaria vectors Anopheles gambiae and Anopheles coluzzii. Mol Ecol. 30(21):5303–5317. doi:10.1111/mec.15845.

Codeço CT, Lourenço-De-Oliveira R, Maciel-De-Freitas R. 2007. DAILY SURVIVAL RATES AND DISPERSAL OF AEDES AEGYPTI FEMALES IN RIO DE JANEIRO, BRAZIL. Am J Trop Med Hyg. 76(4):659–665. doi:10.4269/ajtmh.2007.76.659.

Conomos MP, Reiner AP, Weir BS, Thornton TA. 2016. Model-free Estimation of Recent Genetic Relatedness. Am J Hum Genet. 98(1):127–148. doi:10.1016/j.ajhg.2015.11.022.

Cosme LV, Gloria-Soria A, Caccone A, Powell JR, Martins AJ. 2020. Evolution of kdr haplotypes in worldwide populations of Aedes aegypti: Independent origins of the F1534C kdr mutation. PLoS Negl Trop Dis. 14(4):e0008219. doi:10.1371/journal.pntd.0008219.

Cosme LV, Lima JBP, Powell JR, Martins AJ. 2022. Genome-wide Association Study Reveals New Loci Associated With Pyrethroid Resistance in Aedes aegypti. Front Genet. 13. [accessed 2022 May 30]. https://www.frontiersin.org/article/10.3389/fgene.2022.867231.

Crawford JE, Alves JM, Palmer WJ, Day JP, Sylla M, Ramasamy R, Surendran SN, Black WC, Pain A, Jiggins FM. 2017. Population genomics reveals that an anthropophilic population of Aedes aegypti mosquitoes in West Africa recently gave rise to American and Asian populations of this major disease vector. BMC Biol. 15(1):16. doi:10.1186/s12915-017-0351-0.

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. 2021. Twelve years of SAMtools and BCFtools. GigaScience. 10(2):giab008. doi:10.1093/gigascience/giab008.

David MR, Garcia GA, Valle D, Maciel-de-Freitas R. 2018. Insecticide Resistance and Fitness: The Case of Four Aedes aegypti Populations from Different Brazilian Regions. BioMed Res Int. 2018:6257860. doi:10.1155/2018/6257860.

Davies TGE, Field LM, Usherwood PNR, Williamson MS. 2007. DDT, pyrethrins, pyrethroids and insect sodium channels. IUBMB Life. 59(3):151–162. doi:10.1080/15216540701352042.

DeGiorgio M, Huber CD, Hubisz MJ, Hellmann I, Nielsen R. 2016. SweepFinder2: increased sensitivity, robustness and flexibility. Bioinformatics. 32(12):1895–1897. doi:10.1093/bioinformatics/btw051.

DePristo M, Banks E, Poplin RE, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, et al. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Publ Group. 43(5):491–498. doi:10.1038/ng.806.

Després L, Stalinski R, Tetreau G, Paris M, Bonin A, Navratil V, Reynaud S, David J-P. 2014. Gene expression patterns and sequence polymorphisms associated with mosquito resistance to Bacillus thuringiensis israelensis toxins. BMC Genomics. 15(1):1–13. doi:10.1186/1471-2164-15-926.

Di Tommaso P, Chatzou M, Floden EW, Barja PP, Palumbo E, Notredame C. 2017. Nextflow enables reproducible computational workflows. Nat Biotechnol. 35(4):316–319. doi:10.1038/nbt.3820.

Dickson LB, Campbell CL, Juneja P, Jiggins FM, Sylla M, Black WC IV. 2017. Exon-Enriched Libraries Reveal Large Genic Differences Between Aedes aegypti from Senegal, West Africa, and Populations Outside Africa. G3 GenesGenomesGenetics. 7(2):571–582. doi:10.1534/g3.116.036053.

Dickson LB, Sanchez-Vargas I, Sylla M, Fleming K, Iv WCB. 2014. Vector Competence in West African Aedes aegypti Is Flavivirus Species and Genotype Dependent. PLoS Negl Trop Dis. 8(10):e3153. doi:10.1371/journal.pntd.0003153.

Dong K, Scott JG. 1994. Linkage of kdr-Type resistance and the para-Homologous sodium channel gene in German cockroaches (Blattella germanica). Insect Biochem Mol Biol. 24(7):647–654. doi:10.1016/0965-1748(94)90051-5.

Du Y, Nomura Y, Satar G, Hu Z, Nauen R, He SY, Zhorov BS, Dong K. 2013. Molecular evidence for dual pyrethroid-receptor sites on a mosquito sodium channel. Proc Natl Acad Sci. 110(29):11785–11790. doi:10.1073/pnas.1305118110.

Dubey SK, Mehta D, Chaudhary S, Hasan A, Sunil S. 2022. An E3 Ubiquitin Ligase Scaffolding Protein Is Proviral during Chikungunya Virus Infection in Aedes aegypti. Zhang B, editor. Microbiol Spectr. 10(3):e00595-22. doi:10.1128/spectrum.00595-22.

Durant AC, Grieco Guardian E, Kolosov D, Donini A. 2021. The transcriptome of anal papillae of Aedes aegypti reveals their importance in xenobiotic detoxification and adds significant knowledge on ion, water and ammonia transport mechanisms. J Insect Physiol. 132:104269. doi:10.1016/j.jinsphys.2021.104269.

El-Garj F, Wajidi MFF, Avicor SW. 2016. Allelic variants of cytochrome P450 monooxygenases: Constitutive and insecticide-mediated expression in a Malaysian strain of the dengue vector, Aedes aegypti (Diptera: Culicidae). Eur J Entomol. 113:507–515. doi:10.14411/eje.2016.067.

Enayati AA, Ranson H, Hemingway J. 2005. Insect glutathione transferases and insecticide resistance. Insect Mol Biol. 14(1):3–8. doi:10.1111/j.1365-2583.2004.00529.x.

Evans BR, Gloria-Soria A, Hou L, McBride C, Bonizzoni M, Zhao H, Powell JR. 2015. A Multipurpose, High-Throughput Single-Nucleotide Polymorphism Chip for the Dengue and Yellow Fever Mosquito, Aedes aegypti. G3 GenesGenomesGenetics. 5(5):711–718. doi:10.1534/g3.114.016196.

Ewels P, Magnusson M, Lundin S, Käller M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 32(19):3047–3048. doi:10.1093/bioinformatics/btw354.

Fan Y, O'Grady P, Yoshimizu M, Ponlawat A, Kaufman PE, Scott JG. 2020. Evidence for both sequential mutations and recombination in the evolution of kdr alleles in Aedes aegypti. PLoS Negl Trop Dis. 14(4):e0008154. doi:10.1371/journal.pntd.0008154.

Farnham AW. 1977. Genetics of resistance of houseflies (Musca domestica L.) to pyrethroids. I. Knock-down resistance. Pestic Sci. 8(6):631–636. doi:10.1002/ps.2780080607.

Fonseca-González I, Quiñones ML, Lenhart A, Brogdon WG. 2011. Insecticide resistance status of Aedes aegypti (L.) from Colombia. Pest Manag Sci. 67(4):430–437. doi:10.1002/ps.2081.

Fontaine A, Filipović I, Fansiri T, Hoffmann AA, Cheng C, Kirkpatrick M, Rašić G, Lambrechts L. 2017. Extensive Genetic Differentiation between Homomorphic Sex Chromosomes in the Mosquito Vector, Aedes aegypti. Genome Biol Evol. 9(9):2322–2335. doi:10.1093/gbe/evx171.

Garud NR, Messer PW, Buzbas EO, Petrov DA. 2015. Recent Selective Sweeps in North American Drosophila melanogaster Show Signatures of Soft Sweeps. PLOS Genet. 11(2):e1005004. doi:10.1371/journal.pgen.1005004.

Giraldo-Calderón GI, Calle-Tobón A, Rozo-López P, Colpitts TM, Park Y, Rua-Uribe GL, Londono-Renteria B. 2020. Transcriptome of the Aedes aegypti Mosquito in Response to Human Complement Proteins. Int J Mol Sci. 21(18):6584. doi:10.3390/ijms21186584.

Giraldo-Calderón GI, Harb OS, Kelly SA, Rund SS, Roos DS, McDowell MA. 2022. VectorBase.org updates: bioinformatic resources for invertebrate vectors of human pathogens and related organisms. Curr Opin Insect Sci. 50:100860. doi:10.1016/j.cois.2021.11.008.

Gloria-Soria A, Ayala D, Bheecarry A, Calderon-Arguedas O, Chadee DD, Chiappero M, Coetzee M, Elahee KB, Fernandez-Salas I, Kamal HA, et al. 2016. Global genetic diversity of Aedes aegypti. Mol Ecol. 25(21):5377–5395. doi:10.1111/mec.13866.

Gloria-Soria A, Kellner DA, Brown JE, Gonzalez-Acosta C, Kamgang B, Lutwama J, Powell JR. 2016. Temporal genetic stability of Stegomyia aegypti (= Aedes aegypti) populations. Med Vet Entomol. 30(2):235–240. doi:10.1111/mve.12153.

Gold AS, Feitosa-Suntheimer F, Araujo RV, Hekman RM, Asad S, Londono-Renteria B, Emili A, Colpitts TM. 2020. Dengue Virus Infection of Aedes aegypti Alters Extracellular Vesicle Protein Cargo to Enhance Virus Transmission. Int J Mol Sci. 21(18):6609. doi:10.3390/ijms21186609.

Gonçalves da Silva A, Cunha ICL, Santos WS, Luz SLB, Ribolla PEM, Abad-Franch F. 2012. Gene flow networks among American Aedes aegypti populations. Evol Appl. 5(7):664–676. doi:10.1111/j.1752-4571.2012.00244.x.

Haddi K, Tomé HVV, Du Y, Valbon WR, Nomura Y, Martins GF, Dong K, Oliveira EE. 2017. Detection of a new pyrethroid resistance mutation (V410L) in the sodium channel of Aedes aegypti: a potential challenge for mosquito control. Sci Rep. 7(1):46549. doi:10.1038/srep46549.

Hapairai LK, Mysore K, Chen Y, Harper EI, Scheel MP, Lesnik AM, Sun L, Severson DW, Wei N, Duman-Scheel M. 2017. Lure-and-Kill Yeast Interfering RNA Larvicides Targeting Neural Genes in the Human Disease Vector Mosquito Aedes aegypti. Sci Rep. 7(1):13223. doi:10.1038/s41598-017-13566-y.

Harrington LC, Scott TW, Lerdthusnee K, Coleman RC, Costero A, Clark GG, Jones JJ, Kitthawee S, Kittayapong P, Sithiprasasna R, et al. 2005. DISPERSAL OF THE DENGUE VECTOR AEDES AEGYPTI WITHIN AND BETWEEN RURAL COMMUNITIES. Am J Trop Med Hyg. 72(2):209–220. doi:10.4269/ajtmh.2005.72.209.

Harris AF, Rajatileka S, Ranson H. 2010. Pyrethroid resistance in Aedes aegypti from Grand Cayman. Am J Trop Med Hyg. 83(2):277–284. doi:10.4269/ajtmh.2010.09-0623.

Hemingway J, Boddington RG, Harris J, Dunbar SJ. 1989. Mechanisms of insecticide resistance in Aedes aegypti (L.) (Diptera: Culicidae) from Puerto Rico. Bull Entomol Res. 79(1):123–130. doi:10.1017/S0007485300018630.

Hirata K, Komagata O, Itokawa K, Yamamoto A, Tomita T, Kasai S. 2014. A Single Crossing-Over Event in Voltage-Sensitive Na+ Channel Genes May Cause Critical Failure of Dengue

Mosquito Control by Insecticides. PLoS Negl Trop Dis. 8(8):e3085. doi:10.1371/journal.pntd.0003085.

Hu Z, Du Y, Nomura Y, Dong K. 2011. A sodium channel mutation identified in Aedes aegypti selectively reduces cockroach sodium channel sensitivity to type I, but not type II pyrethroids. Insect Biochem Mol Biol. 41(1):9–13. doi:10.1016/j.ibmb.2010.09.005.

Itokawa K, Sekizuka T, Maekawa Y, Yatsu K, Komagata O, Sugiura M, Sasaki T, Tomita T, Kuroda M, Sawabe K, et al. 2019. High-throughput genotyping of a full voltage-gated sodium channel gene via genomic DNA using target capture sequencing and analytical pipeline MoNaS to discover novel insecticide resistance mutations. PLoS Negl Trop Dis. 13(11):e0007818. doi:10.1371/journal.pntd.0007818.

Jansen CC, Beebe NW. 2010. The dengue vector Aedes aegypti: what comes next. Microbes Infect. 12(4):272–279. doi:10.1016/j.micinf.2009.12.011.

Jenkins SM, Bennett V. 2001. Ankyrin-G coordinates assembly of the spectrin-based membrane skeleton, voltage-gated sodium channels, and L1 CAMs at Purkinje neuron initial segments. J Cell Biol. 155(5):739–746. doi:10.1083/jcb.200109026.

Kaplan NL, Hudson RR, Langley CH. 1989. The “hitchhiking effect” revisited. Genetics. 123(4):887–899. doi:10.1093/genetics/123.4.887.

Kawada H, Higa Y, Komagata O, Kasai S, Tomita T, Yen NT, Loan LL, Sánchez RAP, Takagi M. 2009. Widespread Distribution of a Newly Found Point Mutation in Voltage-Gated Sodium Channel in Pyrethroid-Resistant Aedes aegypti Populations in Vietnam. PLoS Negl Trop Dis. 3(10):e527. doi:10.1371/journal.pntd.0000527.

Kawada H, Oo SZM, Thaung S, Kawashima E, Maung YNM, Thu HM, Thant KZ, Minakawa N. 2014. Co-occurrence of Point Mutations in the Voltage-Gated Sodium Channel of Pyrethroid-Resistant Aedes aegypti Populations in Myanmar. PLoS Negl Trop Dis. 8(7):e3032. doi:10.1371/journal.pntd.0003032.

Kelly ET, Mack LK, Campos M, Grippin C, Chen T-Y, Romero-Weaver AL, Kosinski KJ, Brisco KK, Collier TC, Buckner EA, et al. 2021. Evidence of Local Extinction and Reintroduction of Aedes aegypti in Exeter, California. Front Trop Dis. 0. doi:10.3389/fitd.2021.703873. [accessed 2021 Nov 16]. http://www.frontiersin.org/articles/10.3389/fitd.2021.703873/full#T2.

Kim Y, Nielsen R. 2004. Linkage Disequilibrium as a Signature of Selective Sweeps. Genetics. 167(3):1513–1524. doi:10.1534/genetics.103.025387.

Kofler R, Schlötterer C. 2012. Gowinda: unbiased analysis of gene set enrichment for genome-wide association studies. Bioinformatics. 28(15):2084–2085. doi:10.1093/bioinformatics/bts315.

Kolaczkowski B, Kern AD, Holloway AK, Begun DJ. 2011. Genomic Differentiation Between Temperate and Tropical Australian Populations of Drosophila melanogaster. Genetics. 187(1):245–260. doi:10.1534/genetics.110.123059.

Kotsakiozi P, Evans BR, Gloria-Soria A, Kamgang B, Mayanja M, Lutwama J, Goff GL, Ayala D, Paupy C, Badolo A, et al. 2018. Population structure of a vector of human diseases: Aedes aegypti in its ancestral range, Africa. Ecol Evol. 8(16):7835–7848. doi:10.1002/ece3.4278.

Kotsakiozi P, Gloria-Soria A, Caccone A, Evans B, Schama R, Martins AJ, Powell JR. 2017. Tracking the return of Aedes aegypti to Brazil, the major vector of the dengue, chikungunya and Zika viruses. PLoS Negl Trop Dis. 11(7):e0005653. doi:10.1371/journal.pntd.0005653.

Kotsakiozi P, Gloria-Soria A, Schaffner F, Robert V, Powell JR. 2018. Aedes aegypti in the Black Sea: recent introduction or ancient remnant? Parasit Vectors. 11(1):396. doi:10.1186/s13071-018-2933-2.

Kumar V, Garg S, Gupta L, Gupta K, Diagne CT, Missé D, Pompon J, Kumar S, Saxena V. 2021. Delineating the Role of Aedes aegypti ABC Transporter Gene Family during Mosquito Development and Arboviral Infection via Transcriptome Analyses. Pathogens. 10(9):1127. doi:10.3390/pathogens10091127.

Lan J, Wang Z, Chen Z, Zhang L, Zhao J, Guan Q, Liao C, Liu N, Han Q. 2021. Identification of the Aedes aegypti nAChR gene family and molecular target of spinosad. Pest Manag Sci. 77(4):1633–1641. doi:10.1002/ps.6183.

Lee Y, Schmidt H, Collier TC, Conner WR, Hanemaaijer MJ, Slatkin M, Marshall JM, Chiu JC, Smartt CT, Lanzaro GC, et al. 2019. Genome-wide divergence among invasive populations of Aedes aegypti in California. BMC Genomics. 20(1):1–10. doi:10.1186/s12864-019-5586-4.

Li H, Ruan J, Durbin RM. 2008. Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res. 18(11):1851–1858. doi:10.1101/gr.078212.108.

Lin W-H, Wright DE, Muraro NI, Baines RA. 2009. Alternative Splicing in the Voltage-Gated Sodium Channel DmNav Regulates Activation, Inactivation, and Persistent Current. J Neurophysiol. 102(3):1994–2006. doi:10.1152/jn.00613.2009.

Linss JGB, Brito LP, Garcia GA, Araki AS, Bruno RV, Lima JBP, Valle D, Martins AJ. 2014. Distribution and dissemination of the Val1016Ile and Phe1534Cys Kdr mutations in Aedes aegypti Brazilian natural populations. Parasit Vectors. 7(1):1–11. doi:10.1186/1756-3305-7-25.

Love RR, Redmond SN, Pombi M, Caputo B, Petrarca V, Torre A della, Consortium TA gambiae 1000 G, Besansky NJ. 2019. In Silico Karyotyping of Chromosomally Polymorphic Malaria Mosquitoes in the Anopheles gambiae Complex. G3 Genes Genomes Genet. 9(10):3249–3262. doi:10.1534/g3.119.400445.

Ma J, Amos CI. 2012. Investigation of Inversion Polymorphisms in the Human Genome Using Principal Components Analysis. PLoS ONE. 7(7):e40224. doi:10.1371/journal.pone.0040224.

Mack LK, Kelly ET, Lee Y, Brisco KK, Shen KV, Zahid A, van Schoor T, Cornel AJ, Attardo GM. 2021. Frequency of sodium channel genotypes and association with pyrethrum knockdown time in populations of Californian Aedes aegypti. Parasit Vectors. 14(1):141. doi:10.1186/s13071-021-04627-3.

Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen W-M. 2010. Robust relationship inference in genome-wide association studies. Bioinformatics. 26(22):2867–2873. doi:10.1093/bioinformatics/btq559.

Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. 2018. MUMmer4: A fast and versatile genome alignment system. PLOS Comput Biol. 14(1):e1005944. doi:10.1371/journal.pcbi.1005944.

Marinotti O, Ngo T, Kojin BB, Chou S-P, Nguyen B, Juhn J, Carballar-Lejarazú R, Marinotti PN, Jiang X, Walter MF, et al. 2014. Integrated proteomic and transcriptomic analysis of the Aedes aegyptieggshell. BMC Dev Biol. 14(1):15. doi:10.1186/1471-213X-14-15.

Martinez-Torres D, Chandre F, Williamson MS, Darriet F, Berge JB, Devonshire AL, Guillet P, Pasteur N, Pauron D. 1998. Molecular characterization of pyrethroid knockdown resistance (kdr) in the major malaria vector Anopheles gambiae s.s. Insect Mol Biol. 7(2):179–184. doi:10.1046/j.1365-2583.1998.72062.x.

Martins M, Ramos LFC, Murillo JR, Torres A, de Carvalho SS, Domont GB, de Oliveira DMP, Mesquita RD, Nogueira FCS, Maciel-de-Freitas R, et al. 2021. Comprehensive Quantitative Proteome Analysis of Aedes aegypti Identifies Proteins and Pathways Involved in Wolbachia pipientis and Zika Virus Interference Phenomenon. Front Physiol. 12. [accessed 2022 Jun 7]. https://www.frontiersin.org/article/10.3389/fphys.2021.642237.

Matthews BJ, Dudchenko O, Kingan SB, Koren S, Antoshechkin I, Crawford JE, Glassford WJ, Herre M, Redmond SN, Rose NH, et al. 2018. Improved reference genome of Aedes aegypti informs arbovirus vector control. Nature. 563(7732):501–507. doi:10.1038/s41586-018-0692-z.

Mattingly PF. 1957. Genetical Aspects of the Aëdes Aegypti Problem. Ann Trop Med Parasitol. 51(4):392–408. doi:10.1080/00034983.1957.11685829.

Maynard Smith J, Haigh J. 1974. The hitch-hiking effect of a favourable gene. Genet Res. 23(1):23–35. doi:10.1017/S0016672300014634.

McBride CS, Baier F, Omondi AB, Spitzer SA, Lutomiah J, Sang R, Ignell R, Vosshall LB. 2014. Evolution of mosquito preference for humans linked to an odorant receptor. Nature. 515(7526):222–227. doi:10.1038/nature13964.

Melo Costa M, Campos KB, Brito LP, Roux E, Melo Rodovalho C, Bellinato DF, Lima JBP, Martins AJ. 2020. Kdr genotyping in Aedes aegypti from Brazil on a nation-wide scale from 2017 to 2018. Sci Rep. 10(1):13267. doi:10.1038/s41598-020-70029-7.

Miles A, Harding NJ, Botta G, Clarkson CS, Antao T, Kozak K, Schrider DR, Kern AD, Redmond SN, Sharakhov I, et al. 2017. Genetic diversity of the African malaria vector anopheles gambiae. Nature. 552(7683):96–100. doi:10.1038/nature24995.

Monteiro FA, Shama R, Martins AJ, Gloria-Soria A, Brown JE, Powell JR. 2014. Genetic Diversity of Brazilian Aedes aegypti: Patterns following an Eradication Program. PLoS Negl Trop Dis. 8(9):e3167. doi:10.1371/journal.pntd.0003167.

Nattestad M, Aboukhalil R, Chin C-S, Schatz MC. 2021. Ribbon: intuitive visualization for complex genomic variation. Bioinformatics. 37(3):413–415. doi:10.1093/bioinformatics/btaa680.

Okonechnikov K, Conesa A, García-Alcalde F. 2016. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics. 32(2):292–294. doi:10.1093/bioinformatics/btv566.

Paris M, Marcombe S, Coissac E, Corbel V, David J-P, Després L. 2013. Investigating the genetics of Bti resistance using mRNA tag sequencing: application on laboratory strains and natural populations of the dengue vector Aedes aegypti. Evol Appl. 6(7):1012–1027. doi:10.1111/eva.12082.

Pélissié B, Chen YH, Cohen ZP, Crossley MS, Hawthorne DJ, Izzo V, Schoville SD. 2022. Genome Resequencing Reveals Rapid, Repeated Evolution in the Colorado Potato Beetle. Mol Biol Evol. 39(2):msac016. doi:10.1093/molbev/msac016.

Pittendrigh B, Reenan R, ffrench-Constant RH, Ganetzky B. 1997. Point mutations in the Drosophila sodium channel gene para associated with resistance to DDT and pyrethroid insecticides. Mol Gen Genet MGG. 256(6):602–610. doi:10.1007/s004380050608.

Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Auwera GAV der, Kling DE, Gauthier LD, Levy-Moonshine A, Roazen D, et al. 2018. Scaling accurate genetic variant discovery to tens of thousands of samples. :201178. doi:10.1101/201178. [accessed 2022 Apr 6]. http://www.biorxiv.org/content/10.1101/201178v3.

Powell JR, Gloria-Soria A, Kotsakiozi P. 2018. Recent History of Aedes aegypti: Vector Genomics and Epidemiology Records. BioScience. 68(11):854–860. doi:10.1093/biosci/biy119.

Rašić G, Filipović I, Callahan AG, Stanford D, Chan A, Lam-Phua SG, Tan CH, Hoffmann AA. 2016. The queenslandensis and the type Form of the Dengue Fever Mosquito (Aedes aegypti L.) Are Genomically Indistinguishable. PLoS Negl Trop Dis. 10(11):e0005096. doi:10.1371/journal.pntd.0005096.

Rašić G, Filipović I, Weeks AR, Hoffmann AA. 2014. Genome-wide SNPs lead to strong signals of geographic structure and relatedness patterns in the major arbovirus vector, Aedes aegypti. BMC Genomics. 15(1):275. doi:10.1186/1471-2164-15-275.

Rašić G, Schama R, Powell R, Maciel-de Freitas R, Endersby-Harshman NM, Filipović I, Sylvestre G, Máspero RC, Hoffmann AA. 2015. Contrasting genetic structure between mitochondrial and nuclear markers in the dengue fever mosquito from Rio de Janeiro: implications for vector control. Evol Appl. 8(9):901–915. doi:10.1111/eva.12301.

Rathor HR, Wood RJ. 1981. In-vivo and in-vitro studies on DDT uptake and metabolism in susceptible and resistant strains of the mosquito *Aedes aegypti* L. Pestic Sci. 12(3):255–264. doi:10.1002/ps.2780120304.

Reiter P, Amador MA, Anderson RA, Clark GG. 1995. Short Report: Dispersal of Aedes aegypti in an Urban Area after Blood Feeding as Demonstrated by Rubidium-Marked Eggs. Am J Trop Med Hyg. 52(2):177–179. doi:10.4269/ajtmh.1995.52.177.

Riaz MA, Chandor-Proust A, Dauphin-Villemant C, Poupardin R, Jones CM, Strode C, Régent-Kloeckner M, David J-P, Reynaud S. 2013. Molecular mechanisms associated with increased tolerance to the neonicotinoid insecticide imidacloprid in the dengue vector Aedes aegypti. Aquat Toxicol. 126:326–337. doi:10.1016/j.aquatox.2012.09.010.

Riaz MA, Poupardin R, Reynaud S, Strode C, Ranson H, David J-P. 2009. Impact of glyphosate and benzo[a]pyrene on the tolerance of mosquito larvae to chemical insecticides. Role of detoxification genes in response to xenobiotics☆. Aquat Toxicol. 93(1):61–69. doi:10.1016/j.aquatox.2009.03.005.

Rogers AR, Huff C. 2009. Linkage Disequilibrium Between Loci With Unknown Phase. Genetics. 182(3):839–844. doi:10.1534/genetics.108.093153.

Rose NH, Sylla M, Badolo A, Lutomiah J, Ayala D, Aribodor OB, Ibe N, Akorli J, Otoo S, Mutebi J-P, et al. 2020. Climate and Urbanization Drive Mosquito Preference for Humans. Curr Biol. 30(18):3570-3579.e6. doi:10.1016/j.cub.2020.06.092.

Roy S, Saha TT, Ha J, Banerjee R, Aksoy E, Kulkarni A, Raikhel AS. 2022. Direct and indirect gene repression by the ecdysone cascade during mosquito reproductive cycle. Proc Natl Acad Sci. 119(11):e2116787119. doi:10.1073/pnas.2116787119.

Saavedra-Rodriguez K, Campbell CL, Lenhart A, Penilla P, Lozano-Fuentes S, Black WC. 2019. Exome-wide association of deltamethrin resistance in Aedes aegypti from Mexico. Insect Mol Biol. 28(5):591–604. doi:10.1111/imb.12575.

Saavedra-Rodriguez K, Maloof FV, Campbell CL, Garcia-Rejon J, Lenhart A, Penilla P, Rodriguez A, Sandoval AA, Flores AE, Ponce G, et al. 2018. Parallel evolution of vgsc mutations at domains IS6, IIS6 and IIIS6 in pyrethroid resistant Aedes aegypti from Mexico. Sci Rep. 8(1):6747. doi:10.1038/s41598-018-25222-0.

Saavedra-Rodriguez K, Strode C, Flores AE, Garcia-Luna S, Reyes-Solis G, Ranson H, Hemingway J, Black IV WC. 2014. Differential transcription profiles in Aedes aegypti detoxification genes after temephos selection. Insect Mol Biol. 23(2):199–215. doi:10.1111/imb.12073.

Saavedra-Rodriguez K, Strode C, Flores Suarez A, Fernandez Salas I, Ranson H, Hemingway J, Black WC. 2008. Quantitative trait loci mapping of genome regions controlling permethrin resistance in the mosquito Aedes aegypti. Genetics. 180(2):1137–1152. doi:10.1534/genetics.108.087924.

Saavedra-Rodriguez K, Suarez AF, Salas IF, Strode C, Ranson H, Hemingway J, Black IV WC. 2012. Transcription of detoxification genes after permethrin selection in the mosquito Aedes aegypti. Insect Mol Biol. 21(1):61–77. doi:10.1111/j.1365-2583.2011.01113.x.

Saavedra-Rodriguez K, Urdaneta-Marquez L, Rajatileka S, Moulton M, Flores AE, Fernandez-Salas I, Bisset J, Rodriguez M, Mccall PJ, Donnelly MJ, et al. 2007. A mutation in the voltage-gated sodium channel gene associated with pyrethroid resistance in Latin American Aedes aegypti. Insect Mol Biol. 16(6):785–798. doi:10.1111/j.1365-2583.2007.00774.x.

Schmidt TL, Chung J, Honnen A-C, Weeks AR, Hoffmann AA. 2020. Population genomics of two invasive mosquitoes (Aedes aegypti and Aedes albopictus) from the Indo-Pacific. PLoS Negl Trop Dis. 14(7):e0008463. doi:10.1371/journal.pntd.0008463.

Schrider DR, Hahn MW, Begun DJ. 2016. Parallel Evolution of Copy-Number Variation across Continents in Drosophila melanogaster. Mol Biol Evol. 33(5):1308–1316. doi:10.1093/molbev/msw014.

Schrider DR, Mendes FK, Hahn MW, Kern AD. 2015. Soft shoulders ahead: spurious signatures of soft and partial selective sweeps result from linked hard sweeps. Genetics. 200(1):267–284. doi:10.1534/genetics.115.174912.

Sene NM, Mavridis K, Ndiaye EH, Diagne CT, Gaye A, Ngom EHM, Ba Y, Diallo D, Vontas J, Dia I, et al. 2021. Insecticide resistance status and mechanisms in Aedes aegypti populations from Senegal. PLoS Negl Trop Dis. 15(5):e0009393. doi:10.1371/journal.pntd.0009393.

Severo OP. 1955. Eradication of the Aedes Aegypti Mosquito from the Americas. :23.

Shrinet J, Srivastava P, Sunil S. 2017. Transcriptome analysis of Aedes aegypti in response to mono-infections and co-infections of dengue virus-2 and chikungunya virus. Biochem Biophys Res Commun. 492(4):617–623. doi:10.1016/j.bbrc.2017.01.162.

Smith LB, Kasai S, Scott JG. 2018. Voltage-sensitive sodium channel mutations S989P + V1016G in Aedes aegypti confer variable resistance to pyrethroids, DDT and oxadiazines. Pest Manag Sci. 74(3):737–745. doi:10.1002/ps.4771.

Sombié A, Saiki E, Yaméogo F, Sakurai T, Shirozu T, Fukumoto S, Sanon A, Weetman D, McCall PJ, Kanuka H, et al. 2019. High frequencies of F1534C and V1016I kdr mutations and association with pyrethroid resistance in Aedes aegypti from Somgandé (Ouagadougou), Burkina Faso. Trop Med Health. 47(1):2. doi:10.1186/s41182-018-0134-5.

Soper FL. 1965. The 1964 status of Aedes aegypti eradication and yellow fever in the Americas. Am J Trop Med Hyg. 14(6):887–891. doi:10.4269/ajtmh.1965.14.887.

Stalinski R, Laporte F, Tetreau G, Després L. 2016. Receptors are affected by selection with each Bacillus thuringiensis israelensis Cry toxin but not with the full Bti mixture in Aedes aegypti. Infect Genet Evol. 44:218–227. doi:10.1016/j.meegid.2016.07.009.

Sylla M, Bosio C, Urdaneta-Marquez L, Ndiaye M, Iv WCB. 2009. Gene Flow, Subspecies Composition, and Dengue Virus-2 Susceptibility among Aedes aegypti Collections in Senegal. PLoS Negl Trop Dis. 3(4):e408. doi:10.1371/journal.pntd.0000408.

Tabachnick WJ. 1991. Evolutionary Genetics and Arthropod-borne Disease: The Yellow Fever Mosquito. Am Entomol. 37(1):14–26. doi:10.1093/ae/37.1.14.

Tabachnick WJ, Powell JR. 1979. A world-wide survey of genetic variation in the yellow fever mosquito, Aedes aegypti*. Genet Res. 34(3):215–229. doi:10.1017/S0016672300019467.

Tabachnick WJ, Wallis GP, Aitken TH, Miller BR, Amato GD, Lorenz L, Powell JR, Beaty BJ. 1985. Oral infection of Aedes aegypti with yellow fever virus: geographic variation and genetic considerations. Am J Trop Med Hyg. 34(6):1219–1224. doi:10.4269/ajtmh.1985.34.1219.

Taylor MFJ, Heckel DG, Brown TM, Kreitman ME, Black B. 1993. Linkage of pyrethroid insecticide resistance to a sodium channel locus in the tobacco budworm. Insect Biochem Mol Biol. 23(7):763–775. doi:10.1016/0965-1748(93)90064-Y.

Tennessen JA, Bigham AW, O'Connor TD, Fu W, Kenny EE, Gravel S, McGee S, Do R, Liu X, Jun G, et al. 2012. Evolution and Functional Impact of Rare Coding Variation from Deep Sequencing of Human Exomes. Science. 337(6090):64–69. doi:10.1126/science.1219240.

Tomchaney M, Mysore K, Sun L, Li P, Emrich SJ, Severson DW, Duman-Scheel M. 2014. Examination of the genetic basis for sexual dimorphism in the Aedes aegypti (dengue vector mosquito) pupal brain. Biol Sex Differ. 5(1):10. doi:10.1186/s13293-014-0010-x.

Turner TL, Levine MT, Eckert ML, Begun DJ. 2008. Genomic Analysis of Adaptive Differentiation in *Drosophila melanogaster*. Genetics. 179(1):455–473. doi:10.1534/genetics.107.083659.

Urdaneta-Marquez L, Failloux A-B. 2011. Population genetic structure of Aedes aegypti, the principal vector of dengue viruses. Infect Genet Evol. 11(2):253–261. doi:10.1016/j.meegid.2010.11.020.

Vasimuddin Md, Misra S, Li H, Aluru S. 2019. Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems. In: 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS). p. 314–324.

Venkataraman K, Shai N, Lakhiani P, Zylka S, Zhao J, Herre M, Zeng J, Neal LA, Molina H, Zhao L. 2022. Rapidly evolving genes underlie Aedes aegypti mosquito reproductive resilience during drought. :37.

Vera-Maloof FZ, Saavedra-Rodriguez K, Elizondo-Quiroga AE, Lozano-Fuentes S, Iv WCB. 2015. Coevolution of the Ile1,016 and Cys1,534 Mutations in the Voltage Gated Sodium Channel Gene of Aedes aegypti in Mexico. PLoS Negl Trop Dis. 9(12):e0004263. doi:10.1371/journal.pntd.0004263.

Viglietta M, Bellone R, Blisnick AA, Failloux A-B. 2021. Vector Specificity of Arbovirus Transmission. Front Microbiol. 12:773211. doi:10.3389/fmicb.2021.773211.

Vivero RJ, Muskus CE, Uribe SI, Bejarano EE, Torres C. 2010. FLEBOTOMÍNEOS (DIPTERA: PSYCHODIDAE) EN LA RESERVA NATURAL DEL CAÑÓN DEL RÍO CLARO (ANTIOQUIA), COLOMBIA. Actual Biol.:7.

Walsh TK, Joussen N, Tian K, McGaughran A, Anderson CJ, Qiu X, Ahn S-J, Bird L, Pavlidi N, Vontas J, et al. 2018. Multiple recombination events between two cytochrome P450 loci contribute to global pyrethroid resistance in Helicoverpa armigera. PLOS ONE. 13(11):e0197760. doi:10.1371/journal.pone.0197760.

Weedall GD, Riveron JM, Hearn J, Irving H, Kamdem C, Fouet C, White BJ, Wondji CS. 2020. An Africa-wide genomic evolution of insecticide resistance in the malaria vector Anopheles funestus involves selective sweeps, copy number variations, gene conversion and transposons. PLOS Genet. 16(6):e1008822. doi:10.1371/journal.pgen.1008822.

Wilder-Smith A, Byass P. 2016. The elusive global burden of dengue. Lancet Infect Dis. 16(6):629–631. doi:10.1016/S1473-3099(16)00076-1.

Williamson MS, Martinez-Torres D, Hick CA, Devonshire AL. 1996. Identification of mutations in the houseflypara-type sodium channel gene associated with knockdown resistance (kdr) to pyrethroid insecticides. Mol Gen Genet MGG. 252(1):51–60. doi:10.1007/BF02173204.

Xiao X, Liu Y, Zhang X, Wang J, Li Z, Pang X, Wang P, Cheng G. 2014. Complement-Related Proteins Control the Flavivirus Infection of Aedes aegypti by Inducing Antimicrobial Peptides. PLOS Pathog. 10(4):e1004027. doi:10.1371/journal.ppat.1004027.

Xue AT, Schrider DR, Kern AD, Ag1000g Consortium. 2021. Discovery of Ongoing Selective Sweeps within Anopheles Mosquito Populations Using Deep Learning. Mol Biol Evol. 38(3):1168–1183. doi:10.1093/molbev/msaa259.

Ye YH, Woolfit M, Huttley GA, Rancès E, Caragata EP, Popovici J, O'Neill SL, McGraw EA. 2013. Infection with a Virulent Strain of Wolbachia Disrupts Genome Wide-Patterns of Cytosine Methylation in the Mosquito Aedes aegypti. PLOS ONE. 8(6):e66482. doi:10.1371/journal.pone.0066482.

Zhu J, Busche JM, Zhang X. 2010. Identification of juvenile hormone target genes in the adult female mosquitoes. Insect Biochem Mol Biol. 40(1):23–29. doi:10.1016/j.ibmb.2009.12.004.

FIGURE LEGENDS

Figure 1. Principal component analysis of all filtered SNPs, showing each chromosome separately (left panels and top-right panel) and combining all three chromosomes (bottom-right panel).

Figure 2. Genome-wide scan of candidate selective sweeps. Shaded regions at the ends of chromosomes indicate putative technical artifacts excluded from identification of outliers. Each * with a letter marks a sweep examined in more detail in the text and in the corresponding panel of Figure S6. Horizontal dashed lines indicate our threshold of CLR=1000, used in identifying sweeps examined further.

Figure 3. Principal component analysis and loadings of *Vgsc*. Vertical dashed lines on the loadings panels indicate the locations of the five focal loci listed below the plot. The vertical black bars at the bottom right show the location of exons in *Vgsc*.

Figure 4. Nucleotide diversity and Tajima's *D* in six cohorts, on chromosome 3 from 310 to 320 Mb. The vertical dashed lines mark the boundaries of *Vgsc* (315,926,360-316,405,639). The five Xs at the bottom of the Tajima's *D* plot indicate the locations of the five focal loci.

Figure 5. Linkage disequilibrium across *Vgsc* in the North American cohort (i.e. USA), the South American cohorts (Brazil and Colombia combined), and the combined American cohort (Brazil, Colombia, and USA combined).

TABLES

	F1534C (315939224)			V1016I (315983763)			I915K (315999297)			S723T (316014588)			V410L (316080722)			total *n*
	F/F	F/C	C/C	V/V	V/I	I/I	I/I	I/K	K/K	S/S	S/T	T/T	V/V	V/L	L/L	
Brazil	0 (0)	0 (0)	15 (1)	13	0	0	13	0	0	14	0	0	14	0	0	15
Colombia	0 (0)	0 (0)	30 (19)	6	16	8	6	15	8	4	14	7	6	16	8	30
USA	2 (0)	0 (0)	23 (3)	4	4	14	2	3	14	7	3	12	4	4	13	25
Gabon	13 (2)	0 (0)	0 (0)	12	0	0	11	1	0	11	0	0	10	0	0	13
Kenya	14 (0)	0 (0)	2 (0)	16	0	0	4	10	2	14	0	0	15	0	0	16
Senegal	19 (7)	0 (0)	0 (0)	18	0	0	8	9	0	18	0	0	19	0	0	19

Table 1. Genotypes at selected loci previously identified as associated with pyrethroid resistance in *Vgsc*.[a]

	315983763 (V1016I)	315999297 (I915K)	316014588 (S723T)	316080722 (V410L)
315983763 (V1016I)				
315999297 (I915K)	1.0			
316014588 (S723T)	0.87395	0.87395		
316080722 (V410L)	1.0	1.0	0.87395	

Table 2. Linkage disequilibrium between four known or putative insecticide resistance loci in the Colombian sample.[b]

[a] Genotypes for F1534C were obtained from bcftools mpileup; the numbers in parentheses reflect the numbers remaining in each category after genotypes with qualities less than 20 were masked. Genotypes for the other four sites were derived from GATK HaplotypeCaller, and were calculated after genotypes with qualities less than 20 were masked. Total numbers within a population vary between sites because not all specimens could be confidently genotyped at all sites; the total number of specimens for each population is given in the far-right column.

[b] Linkage disequilibrium, as measured by r^2 (Methods), after masking all genotypes with a genotype quality less than 20. A fifth known insecticide resistance locus, 315939224 (F1534C), is homozygous for the alternate (resistant) allele in every specimen in this sample.