Zoological Studies **63**:58 (2024)
doi:10.6620/ZS.2024.63-58

Zoological Studies

Open Access

Assessing Phylogeographic Patterns and Genetic Diversity in *Culex quinquefasciatus* (Diptera: Culicidae) via mtDNA Sequences from Public Databases

Gian Carlo García-Escobar[1,2], Juan José Trujillo González[1,3], and Oscar Alexander Aguirre-Obando[1,3,*]

[1]*Escuela de investigación en Biomatemática, Universidad del Quindío. Carrera 15, Calle 12 Norte, Armenia, Quindío, Colombia. E-mail: gcgarciae@uqvirtual.edu.co (García-Escobar)*
[2]*Programa de Licenciatura en Biología y Educación Ambiental, Facultad de Educación, Universidad del Quindío. Carrera 15, Calle 12 Norte, Armenia, Quindío, Colombia*
[3]*Programa de Biología, Facultad de Ciencias Básicas y Tecnologías, Universidad del Quindío. Carrera 15, Calle 12 Norte, Armenia, Quindío, Colombia. *Correspondence: E-mail: oscaraguirre@uniquindio.edu.co (Aguirre-Obando) E-mail: juanj.trujillog@uqvirtual.edu.co (González)*

Received 28 May 2024 / Accepted 27 November 2024 / Published 31 December 2024
Communicated by John Wang

To identify the worldwide genetic structure, gene flow, and diversity of *Culex quinquefasciatus*, we conducted phylogeographic and population genetics analyses utilizing publicly available mtDNA sequences. Therefore, the aim of this study was to investigate the genetic structure and diversity of natural populations of *C. quinquefasciatus* worldwide, using available genetic data reflecting its natural distribution. Our study focused on the cytochrome *c* oxidase subunit I (*COI*) gene, mirroring the species' distribution pattern. We examined *COI* gene sequences from *C. quinquefasciatus* populations across Asia (n = 1,698), America (n = 334), Africa (n = 30), Oceania (n = 21), and Europe (n = 1), identifying 69 haplotypes. Genetic links were observed between Asian populations and those from other continents. Global genetic diversity was 0.531, varying from 0.095 in Oceania to 0.648 in South America. Neutrality tests indicated demographic expansions at the continental level in the Americas, North America, and Asia, as well as in some countries within these regions. In contrast, at both global and continental levels (South America, Oceania, and Africa), and in most countries within these continents, neutral populations were observed. AMOVA revealed genetic structuring among and within countries, with no genetic isolation observed (R^2 = 0.03144; p > 0.05). Despite lower genetic diversity, Asian populations facilitated gene flow with other continents, suggesting a possible native origin of the species in Asia. The dispersal of this mosquito to new regions, coupled with its ability to transmit various arboviruses, underscores its significance as a potential public health threat.

Key words: Epidemiology, Genetic diversity, Public health, Southern house mosquito, Vector capacity

BACKGROUND

The mosquito *C. quinquefasciatus* (Say 1823), commonly known as the southern house mosquito, is a widely distributed dipteran inhabiting urban, peri-urban, and rural environments. Its larval stages thrive in stagnant water containing organic debris, often found in artificial containers, subterranean systems, septic tanks, clogged drains, and abandoned wells (Calhoun et al. 2007). Primarily active at night, particularly the females,

Citation: García-Escobar GC, González JJT, Aguirre-Obando OA. 2024. Assessing phylogeographic patterns and genetic diversity in *Culex quinquefasciatus* (Diptera: Culicidae) via mtDNA sequences from public databases. Zool Stud **63**:58. doi:10.6620/ZS.2024.63-58.

adults of this species feed on the blood of both humans and animals. During daylight hours, they seek refuge in shadowed corners, shelters, sewers, and vegetation (Hickner et al. 2011).

The taxonomic classification of *C. quinquefasciatus* has been a topic of debate. Some studies classify it as *C. pipiens* forma *pipiens* or *C. fatigans*, while others include it within a species complex alongside *C. pipiens* f. *pipiens*, *C. pipiens* f. *molestus*, *C. pallens*, *C. australicus*, and *C. globocoxitus* (Harbach 2012). In this study, *C. quinquefasciatus* is considered as part of the *C. pipiens* complex. Hybridization has been observed between populations of *C. quinquefasciatus* and *C. pipiens* f. *pipiens* and f. *molestus* in Southeast Asia, North America, Argentina, and Madagascar. Although mating between the three taxa is possible under controlled conditions, hybrids often have lower egg fertility and viability, although notable differences between populations persist in South Africa (Aardema et al. 2020). Morphologically, females and interspecies hybrids are nearly identical, necessitating molecular analysis or detailed examination of behavioral and physiological traits such as male genitalia for differentiation (Smith and Fonseca 2004; Harbach 2012).

Among the *Culex* genus, *C. quinquefasciatus* emerges as the most anthropophilic and endophagic species (Gouge et al. 2019). It serves as a vector for filarial nematodes such as *Dirofilaria immitis* and *Wuchereria bancrofti* (Thanchomnang et al. 2013), as well as protozoa like *Plasmodium relictum* (Ferreira et al. 2022). Moreover, it is implicated in the transmission of a myriad of diseases, including Ross River virus (RRV) (Harley et al. 2001), Venezuelan equine encephalitis (VEE), Eastern equine encephalitis (EEE), Japanese encephalitis (JE), St. Louis encephalitis (SLE), West Nile virus (WNV), Myxoma virus (MV), avian reticuloendotheliosis virus (REV) (Gouge et al. 2019), and Rift Valley fever virus (RVFV) (Sang et al. 2010). Given its pivotal role, comprehending the patterns of genetic structuring and flow within *C. quinquefasciatus* populations is crucial, as it influences pathogen transmission, vector capacity, and competence (Van Den Eynde et al. 2022).

Molecular markers play a pivotal role in unraveling the biology and population dynamics of disease vectors. Mitochondrial DNA (mtDNA) stands out for its small size, rapid evolutionary rate, and exclusive maternal inheritance with minimal genetic recombination. In the phylogeographic context, molecular markers have been extensively employed at the mtDNA level for *C. quinquefasciatus* in regions like the United States and Australia (Behura et al. 2011). Notably, genes such as dehydrogenase subunit 4 (*ND4*), cytochrome *b* (*cytb*), and cytochrome *c* oxidase subunit I (*COI*) have been utilized. *ND4* data is available from the United States, Thailand, and South Africa (Rasgon et al. 2006; Chaulk et al. 2016), while *cytb* data spans regions including Benin, China, the Philippines, the United States, Réunion, South Africa, and Sri Lanka (Ishtiaq et al. 2008; Atyame et al. 2011). Similarly, *COI* information is accessible from various countries across Asia, Africa, Oceania, the Americas and Europe (Hasan et al. 2009; Shaikevich and Zakharov 2010; Huang et al. 2011; Quintero and Navarro 2012; Pfeiler et al. 2013; Sharma et al. 2013; Low et al. 2014; Ashfaq et al. 2014; Wilke et al. 2014; Daravath et al. 2015; Gunay et al. 2015; Murugan et al. 2016; Shaikevich et al. 2016; Dumas et al. 2016; Talaga et al. 2017; Koosha et al. 2017; Anoopkumar et al. 2019; Cane et al. 2020; Lorenz et al. 2021; Maekawa et al. 2021; Aremu et al. 2022; Panda and Barik 2022; Thankachan et al. 2023). However, a comprehensive analysis of all available genetic data for *C. quinquefasciatus* is lacking. Thus, this study aimed to investigate the global genetic structure, migration patterns, and diversity of natural populations of *C. quinquefasciatus*, utilizing existing genetic datasets that adhere to the species' natural distribution pattern.

MATERIALS AND METHODS

Sequence retrieval, download, and analysis

We conducted an exhaustive search in the NCBI and BoldSystem databases to acquire data on nuclear and mitochondrial genomes, as well as 13 mitochondrial genes. The search query employed the species name (in parentheses) combined with the boolean operator "AND" and the phrases "complete nuclear genome" or "complete mitochondrial genome," enclosed in quotation marks, to retrieve genomic information. For the retrieval of mitochondrial genes, we used the species name in parentheses, followed by the "AND" operator, the specific gene name enclosed in quotation marks, and the "OR" operator followed by the gene's abbreviations, also enclosed in quotation marks (Table S1) (Waldbieser et al. 2003).

From the results obtained, genetic information was selected for further analysis based on a distribution pattern similar to that of *C. quinquefasciatus* as provided by GBIF (https://www.gbif.org/). Only genetic sequences demonstrating a similarity percentage above 98% following BLAST analysis (https://blast.ncbi.nlm.nih.gov/Blast.cgi) were considered (Donkor et al. 2014). Given that the species under study is part of a species complex, it was necessary to ensure that the genetic information used was exclusive to *C.*

quinquefasciatus. To achieve this, two phylogenetic clustering analyses were conducted: one employing the maximum likelihood (ML) method (Fig. 4) and the other utilizing Bayesian inference (BI) (Holder and Lewis 2003). Both trees included species from *pipiens* complex, *C. pipiens* f. *pipiens* (GenBank access number: HQ724616), *C. pipiens* f. *molestus* (GenBank access number: MN389460), *C. pallens* (GenBank access number: KT851543), *C. australicus* (GenBank access number: NC054314), and *C. globocoxitus* (GenBank access number: KU495003). *C. tarsalis* (GenBank access number: JX259917), and *Aedes aegypti* (GenBank access number: MF194022) were used as an outgroup. Exclusive information within the internal group, comprising the genetic data under analysis, was considered valid and was the information used in this study.

Based on the previous results, sequences of *C. quinquefasciatus* were classified according to their geographical location, continent, country, type of coverage, and altitude (Table S2). Sequences without geographical information were excluded from the analyses. The classification of coverage types (urban, peri-urban, and rural) and altitude (in meters) was determined utilizing the geographic coordinates sourced from occurrence data retrieved from the GBIF platform, as well as sequences from GenBank and BoldSystem. This process involved overlaying coverage and altitude layers obtained through DIVA-GIS (https://www.diva-gis.org/) (Guilherme et al. 2024). Subsequently, a map was generated using QGIS (https://www.qgis.org/es/site/) to visualize the distribution of geographical information derived from genetic and occurrence data (Fig. 1).

The alignment of sequences and construction of a haplotype network

We conducted multiple sequence alignment using the MEGA software (Katoh et al. 2018). Haplotypes (H) were identified based on their frequency, with H1 assigned to the most common haplotype and subsequent numbering for others. To identify potential nuclear mitochondrial DNA sequences (NUMTs) among the haplotypes, a stop codon search was performed in the alignment (Leite 2012) (Table S3). Subsequently, to elucidate the genetic relationships between worldwide populations of *C. quinquefasciatus*, we constructed a haplotype network using NUMT-free haplotypes (Fig. 2). For this purpose, we employed the Population Analysis with Reticulate Trees (PopART) with a parsimony approach (Clement et al. 2000). Additionally, based on the interpretation of the haplotype network, which contains the haplotypes, their connections through lines with their respective mutational steps, and the countries they belong to, we plotted the genetic flow on a world map using arrows to represent the connections suggested by the network, starting from the most frequent haplotype to all other haplotypes. These arrows connect countries that share the same haplotypes and, in turn, those with similar or different haplotypes and/or countries.

Population genetic analysis

Based on the aligned, trimmed, and country-organized sequences, we conducted a population genetic analysis that included calculation of haplotype diversity (*Hd*), nucleotide diversity (π), and neutrality tests (Tajima's *D*), at both continental and country levels within each continent (Table 1). In these country-level analyses, we considered each country and all the associated sequences as a distinct population due to geographical differences and potential barriers to dispersal that could influence the genetic structure of the mosquito under study. Additionally, an analysis of molecular variance (AMOVA) was performed at a global and continental level to assess the distribution of genetic variation among and within defined populations (countries). These analyses were carried out using the R environment (CoreTeam 2017) with the following packages: ape (Didier 2024), adegenet (Jombart 2023), pegas (Paradis et al. 2016), poppr (Kamvar et al. 2014), usethis (Maintainer 2024), devtools (Hadley et al. 2022), and mmod (Winter et al. 2013). Population genetic structuring was evaluated using the fixation index F_{ST}. Pairwise F_{ST} values between all populations were calculated as proposed by Weir and Cockerham (1984) implemented in the hierfstat package (Goudet et al. 2022) in R. Gene flow (Nm) was estimated from the F_{ST} values using the formula $Nm = 1\text{-}F_{ST}/2.F_{ST}$, assuming an island model. The F_{ST} values obtained were represented using boxplots for each country, showing the distribution of F_{ST} values in comparison with all other countries. These plots were generated using the ggplot2 package in R. In these analyses, singleton populations (countries with only one sample) were excluded due to their impact on the precise estimation of F_{ST} values and to avoid biases in interpreting genetic differentiation (Hale et al. 2012). Furthermore, we examined the potential relationship between genetic (F_{ST}) and geographic (Km) distances using the Mantel test (Fig. 5). The Mantel test was performed using the mantel.randtest function from the ade4 package (Dray and Dufour 2007), with 999 permutations to assess statistical significance. Geographic distances between populations were calculated using the central geographic coordinates of each country. Google Earth

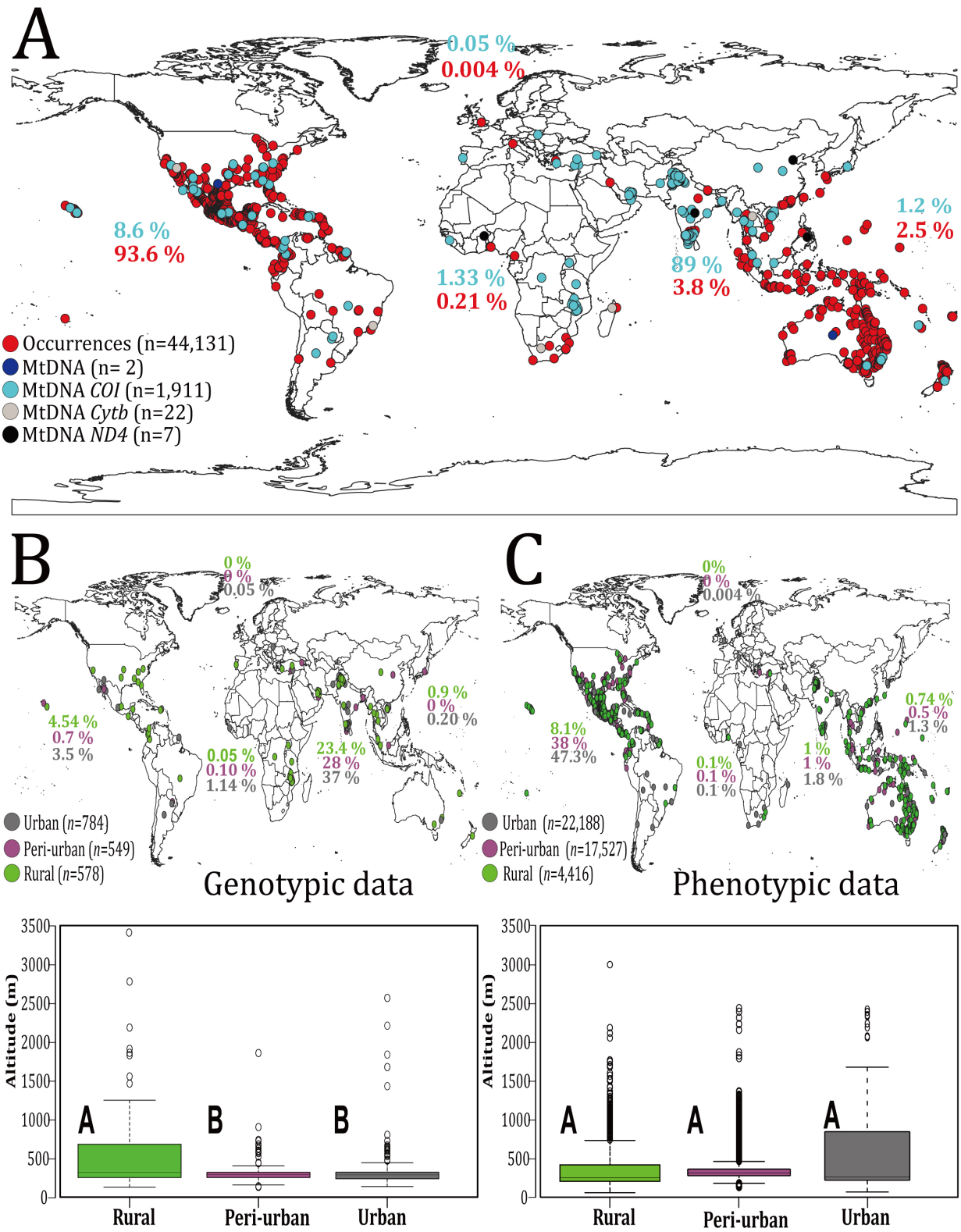

Fig. 1. World map for *Culex quinquefasciatus* presenting information for: A, Distribution of information retrieved from genetic and geographical databases. B, Genotypic distribution of information retrieved from genetic databases associated with three coverages (urban, peri-urban, and rural) and boxplot representing the results of the Kruskal-Wallis analysis for the variables coverage and altitude. C, Phenotypic distribution of information retrieved from the geographical database associated with three coverages (urban, peri-urban, and rural) and boxplot representing the results of the Kruskal-Wallis analysis for the variables coverage and altitude.

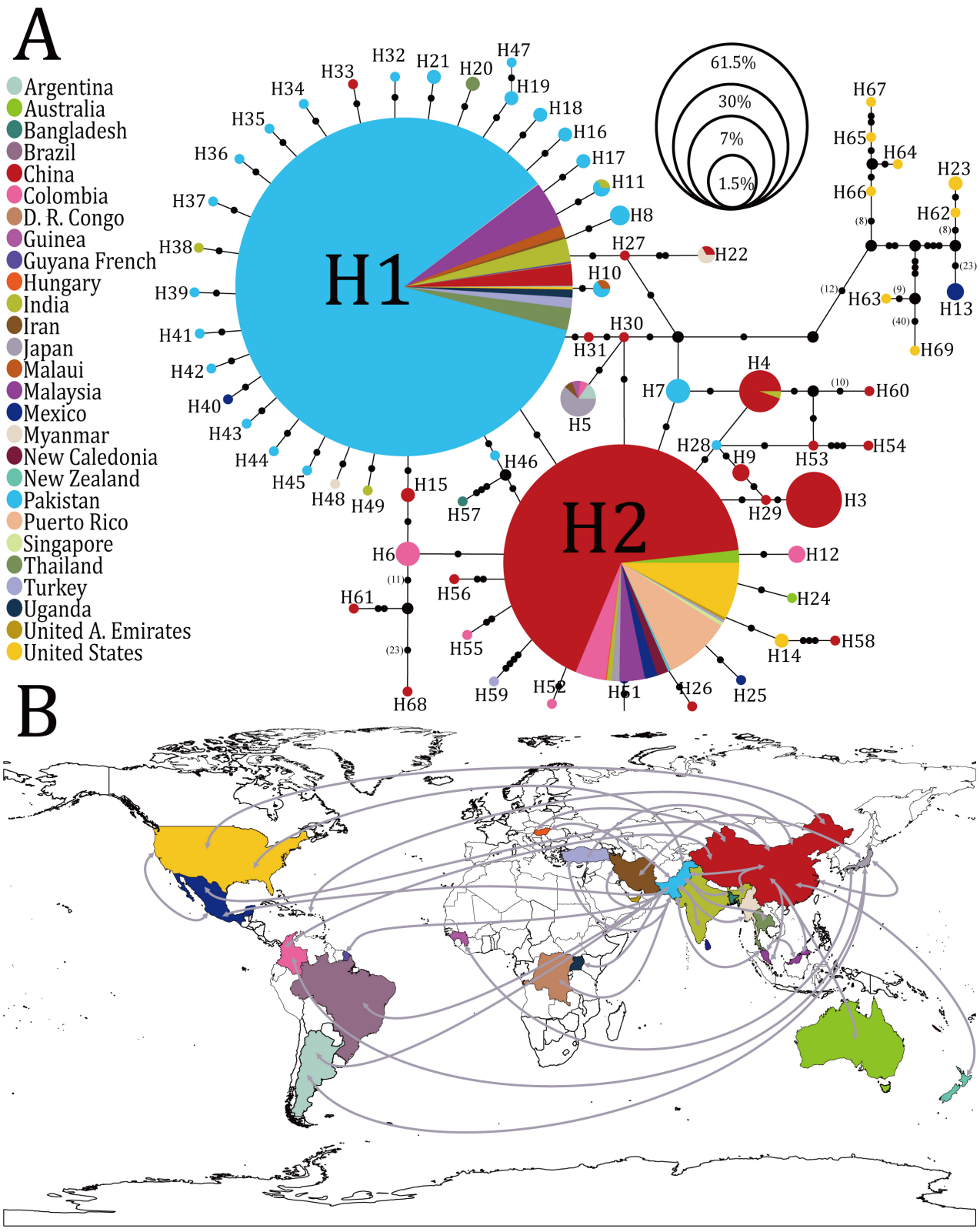

Fig. 2. Global haplotype network (A) and map representing gene flow for the *COI* gene (B) in *C. quinquefasciatus*. In the haplotype network, the size of the circles is proportional to their frequency, and each colored circle belongs to a respectively numbered haplotype; black circles and numbers in parentheses represent mutational steps. In map B, color indicates the location of the sequences used, and arrows represent the genetic relationships between natural populations of *Culex quinquefasciatus*.

was used to calculate the geographic (Km) distances used in the Mantel test.

RESULTS

Figure 1 shows the world map for *C. quinquefasciatus* and the distribution of data from genetic and geographical databases (Fig. 1A), including genetic data categorized by urban, peri-urban, and rural coverage (Fig. 1B), with a boxplot of Kruskal-Wallis analysis for coverage and altitude variables, and geographical data categorized similarly with the corresponding boxplot (Fig. 1C). Figure 1A shows the distribution of genetic (GenBank + BoldSystems) and phenotypic information (GBIF) of *C. quinquefasciatus*, including the *COI* gene data (n = 1,911) used in this study. When statistically comparing whether countries with phenotypic data

Table 1. Global genetic diversity presenting data on haplotypic diversity (*Hd*), nucleotide diversity (π), and results of the neutrality test (Tajima's *D*), calculated for natural populations of the southern house mosquito, by continents and countries within each continent

Countries by continent	Number of sequences	Genetic diversity				
		Number of haplotypes	*Hd*	π	Tajima's *D*	*P*-value
Global	**1,911**	**69**	**0.531**	**0.003**	**-2.570**	**0.10**
America	**167**	**27**	**0.383**	**0.012**	**-2.239**	**0.025***
South America	**42**	**10**	**0.648**	**0.002**	**-1.322**	**0.186**
Colombia	36	6	0.533	0.002	-1.476	0.140
Argentina	3	2	0.667	0.003	NaN	NaN
Guayana Francesa	2	1	0	0	NaN	NaN
Brazil	1	1	NaN	NaN	NaN	NaN
North America	**125**	**17**	**0.268**	**0.015**	**-2.080**	**0.038***
United Stated	56	11	0.384	0.022	-1.776	0.076
Puerto Rico	53	1	0	0	NaN	NaN
Mexico	16	5	0.600	0.034	0.164	0.870
Oceania	**21**	**4**	**0.095**	**0.0002**	**-1.164**	**0.245**
Australia	11	2	0.182	0.0005	-1.129	0.259
New Caledonia	9	1	0	0	NaN	NaN
New Zealand	1	1	NaN	NaN	NaN	NaN
Europe	**1**	**1**	**NaN**	**NaN**	**NaN**	**NaN**
Hungary	1	1	NaN	NaN	NaN	NaN
Africa	**25**	**5**	**0.157**	**0.001**	**-1.214**	**0.225**
Malawi	14	2	0.143	0.001	-1.155	0.248
Uganda	9	1	0	0	NaN	NaN
Guinea	1	1	NaN	NaN	NaN	NaN
D. R. of the Congo	1	1	NaN	NaN	NaN	NaN
Asia	**1,697**	**68**	**0.480**	**0.002**	**-2.521**	**0.012***
Pakistan	1,044	25	0.073	0.0001	-2.345	0.019*
China	477	20	0.341	0.002	-2.599	0.009*
Malasia	72	2	0.407	0.001	1.196	0.232
India	34	6	0.410	0.001	-1.576	0.115
Thailand	27	2	0.142	0.001	-0.954	0.340
Turkey	15	3	0.362	0.002	-1.451	0.147
Japan	14	2	0.527	0.001	1.434	0.152
Iran	4	2	0.500	0.003	-0.710	0.478
Myanmar	4	3	0.833	0.003	0.592	0.554
Singapore	3	1	0	0	NaN	NaN
United Arab Emirates	2	1	0	0	NaN	NaN
Bangladesh	1	1	0	NaN	NaN	NaN

NaN: Not enough sequences to obtain the results, or no segregating sites in the sample. *: indicate statistical significance.

also have genetic data, and analyzing the difference between both distributions, the Fischer test indicates a significant difference ($p \leq 0.05$), with an odds ratio of 0 (95% CI: 0.0000000–0.1859958), suggesting that the presence of genetic data is not associated with the presence of the species. However, it is observed that several countries where the species has been recorded also have genetic information, suggesting that despite the statistical difference, our results generally reflect a global distribution analysis of *C. quinquefasciatus*. Additionally, it illustrates the distribution of this information based on genetic data (Fig. 1B) and occurrences (Fig. 1C) categorized by type of coverage. Upon separate analysis of the genetic data and occurrence data using a Kruskal-Wallis test associated with three coverage categories (urban, peri-urban, and rural) concerning altitude, a statistically significant difference is observed in the genetic data ($p < 0.0001$; Fig. 1B), whereas no such difference is evident in the occurrence data ($p > 0.9999$; Fig. 1C). This suggests that the available information for genetic data obtained worldwide is not uniform regarding coverage and altitude categories, whereas it is for occurrence data.

The species exhibits a global distribution based on genetic and occurrence data, across altitudinal ranges from 1 to 3,540 m. The genetic data suggest that the widest altitudinal distribution is observed in the Americas, with records ranging from 3 to 3,540 m (North America: 4 to 1,545 m, and South America: 3 to 3,540 m). In contrast, in Africa, altitudes range from 22 to 1,209 m, while for Europe, only one record exists at 128 m. Regarding occurrence data, worldwide, the species is recorded altitudinally within a range of 1 to 2,912 m (North America: 1 to 2,638 m, and South America: 1 to 2,912 m), with the lowest altitude in Europe, with records from 168 to 155 m. Distribution by coverage reveals that the majority of records, both in genetic data and occurrence data, are found in urban areas worldwide. Table S2 further elaborates on this information by continent and country, delineating coverage type and altitude for *COI* gene genetic data and occurrence data sourced from GBIF.

Between September 2022 and January 2023, a total of 2,055 genetic sequences were acquired, primarily from the GenBank database (93%) and BoldSystem (7%). These sequences spanned various continents, with distribution as follows: Asia (88.3%), America (8.6%, comprising 6.2% in North America and 2.4% in South America), Africa (1.9%), Oceania (1.1%), and Europe (0.1%). The average sequence length was 698 bp, ranging from 300 to 1,000 bp. Subsequently, following alignment and trimming processes, 1,911 sequences with a consistent length of 388 bp each were retained, covering all aforementioned continents.

Asia (88.829%) emerged as the most represented continent, with sequences originating from countries such as Pakistan (54.6%), China (25%), Malaysia (3.8%), India (1.8%), Thailand (1.4%), Turkey (0.8%), Japan (0.7%), Iran (0.21%), Myanmar (0.16%), Singapore (0.2%), United Arab Emirates (0.10%), and Bangladesh (0.05%). North America (6.5%) was represented by sequences from the United States (2.9%), Puerto Rico (2.8%), and Mexico (0.8%). In South America (2.16%), sequences were recorded from Colombia (1.9%), Argentina (0.16%), French Guiana (0.10%), and Brazil (0.05%). Africa (1.33%) was represented by sequences from Malawi (0.73%), Uganda (0.5%), Guinea (0.05%), and the Democratic Republic of the Congo (0.05%). Oceania (1.15%) was represented by sequences from Australia (0.6%), New Caledonia (0.5%), and New Zealand (0.05%). Lastly, Europe (0.05%) was represented by a single sequence from Hungary (0.05%).

Table S3 for the *COI* gene comprises the sequence alignment for each of the 69 identified haplotypes (H), along with their amino acid translations. H1 is the most prevalent, accounting for 61.54% of the sequences, followed by H2 (30.09%), H3 (1.65%), H4 (0.94%), H5 (0.68%), H6 to H7 (each at 0.31%), H8 (0.21%), H9 to H13 (each at 0.16%), H14 to H23 (each at 0.10%), and H24 to H69 (each at 0.05%). While H1 was the most abundant and found in 15 countries, H2 exhibited the broadest geographical distribution, being present in 16 countries.

Figure 2 illustrates the geographic origin of the collected genetic material, depicting the distribution of these haplotypes across various continents and countries. Both the haplotype network (Fig. 2A) and the genetic flow map (Fig. 2B) demonstrate genetic connections between Asia and other regions, including America, Africa, Europe, and Oceania, highlighting these connections across multiple populations of *C. quinquefasciatus*. In Asia, genetic connections are observed among populations from Pakistan and various countries such as China, Myanmar, Malaysia, Thailand, Turkey, India, and Iran. Additionally, China exhibits connections with Pakistan, Japan, and Iran. Asia and America also display genetic relationships through Pakistan, which links with the United States, Mexico, Argentina, Brazil, and French Guiana. Furthermore, China is linked to America, establishing connections with the United States, Puerto Rico, Mexico, and Colombia, while Japan has genetic connections with Argentina and Colombia. Asia is connected to Africa via Pakistan, which links with Uganda and the Democratic Republic of the Congo, and Japan shows genetic connections with Guinea. Moreover, Asia is associated with Europe through Pakistan and China, which connect

with Hungary. Lastly, Asia also connects with Oceania, particularly through China, which has genetic ties with Australia, New Zealand, and New Caledonia. These intercontinental genetic connections reflect the diversity and dispersal history of *C. quinquefasciatus* on a global scale.

Additionally, in a global context, we conducted two additional analyses based on the results obtained for the haplotypes. In the first analysis, the objective was to evaluate whether there are significant differences in the altitude at which different haplotypes are found. A non-parametric Kruskal-Wallis test was applied, where the dependent variable was altitude and the independent variable was the haplotypes (mitochondrial lineages H1 to H69). This test was appropriate because altitude does not follow a normal distribution, and there are multiple groups. The results were visualized using boxplots, allowing us to observe the distribution of altitudes associated with each haplotype. In the second analysis, we aimed to determine whether the altitudinal distribution of haplotypes varies according to zones (urban, peri-urban, and rural) and if there is an interaction between the zone and different haplotypes with respect to altitude. A two-way Analysis of Variance (ANOVA) was performed, considering zone and haplotypes as factors and altitude as the dependent variable. Before conducting the ANOVA, we assessed the assumptions of normality and homoscedasticity. The results were presented through a heat map, showing variations in mean altitude among different haplotypes and zones, facilitating the identification of patterns and potential interactions. In both analyses, the altitude data were associated with the results presented in table S2, considering the distribution of haplotypes. All analyses were performed using R (CoreTeam 2017), utilizing the stats package for statistical tests, ggplot2 for data visualization, and reshape2 for data manipulation.

Figure 3 presents the results of the Kruskal-Wallis tests (Fig. 3A) and the two-way ANOVA (Fig. 3B). In both cases, the analyses were significant: Kruskal-Wallis ($d.f.$ = 68, $p \leq 0.05$) and ANOVA (Zone (Z): $d.f.$ = 2, $p \leq 0.05$; Haplotype (H): $d.f.$ = 68, $p \leq 0.05$; Zone x Haplotype Interaction (Z x H): $d.f.$ = 14, $p \leq 0.05$). Despite these results, the bias in the distribution of genetic information is highlighted again (Tables 1 and S2). Overall, figure 3A suggests that haplotypes are distributed differently across the various altitudes where they are reported, and many of them, due to the limited amount of data, do not show altitudinal variation. It is noteworthy that haplotypes 1 and 2, although the most frequent, do not exhibit the widest range of altitudinal distribution; in contrast, haplotype 6 does show a greater altitudinal range. On the other hand, figure 3B indicates that the distribution of haplotypes based on urban, rural, and peri-urban areas concerning altitude differs. In rural areas, a greater number of haplotypes is observed, which are also more widely distributed altitudinally. In contrast, peri-urban and urban areas have fewer haplotypes compared to rural areas, although the haplotypes present in these zones tend to be concentrated at lower altitudes. Interestingly, haplotype 6 is present in both rural and urban areas. The phylogenetic tree constructed using the maximum likelihood (ML) method indicates that all genetic information for the *COI* gene of *C. quinquefasciatus* is exclusive to this species (Fig. 4). However, branch supports were generally ≤ 50 in most cases, indicating limited reliability in the clusters within the tree for certain nodes. However, it is observed that the American populations are different from the Asian ones, and the latter may have been derived from the American populations. Conversely, the phylogenetic tree generated using the Bayesian inference (BI) method did not achieve stabilization of the Markov chains after 1,000 million interactions, and thus it's results are not presented.

Table 1 presents the results for haplotype diversity (*Hd*), nucleotide diversity (π), and the neutrality test for different continents and countries. The global *Hd* was 0.531, showing significant variation among continents, ranging from 0.095 in Oceania to 0.648 in South America. At the country level, *Hd* ranged from 0.0 (observed in places like French Guiana, Puerto Rico, New Caledonia, Uganda, Singapore, and the United Arab Emirates) to 0.833 in Myanmar. Regarding π, the global value was 0.003, with regional disparities ranging from 0.0 (recorded in French Guiana, Puerto Rico, New Caledonia, Uganda, Singapore, and the United Arab Emirates) to 0.034 in Mexico. Our results based on the number of haplotypes distributed by continents suggest the highest diversity in Asia, followed by America, Africa, and Oceania. Tajima's *D* test results suggest that, at both global and continental levels (South America, Oceania, and Africa), and in most countries within these continents, negative values (the majority) and positive values (only a few) were observed, though none were statistically significant. In contrast, at the continental level, the Americas, North America, and Asia presented significant negative values. Additionally, in some countries in Asia (Pakistan and China), significant negative Tajima's *D* values were also observed. These findings underscore diverse patterns of diversity and natural selection across different geographical regions for *C. quinquefasciatus*.

The results of the analysis of molecular variance (AMOVA) indicated significant genetic structuring both at the level of countries within the same continent ($d.f.$ = 22, ss = 642.23, $p < 0.05$) and within countries

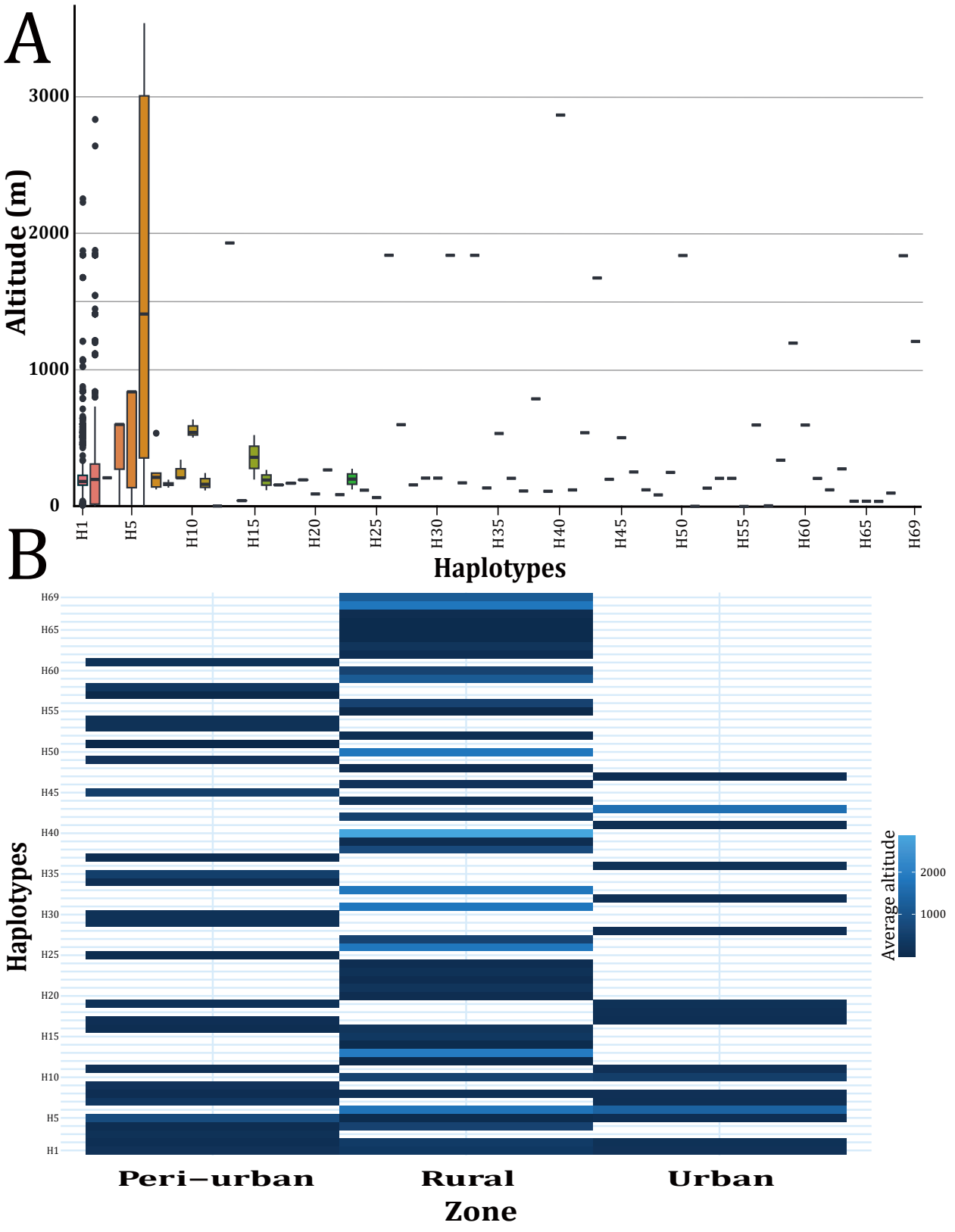

Fig. 3. Boxplot representing the altitudinal distribution of each observed haplotype (A) and heatmap visualizing the interaction between the factors zone and haplotype with respect to altitude (B). Each cell shows the average altitude of a haplotype in a specific zone, with a blue color scale indicating the value: lighter colors represent lower altitude values, while darker colors indicate higher altitude values.

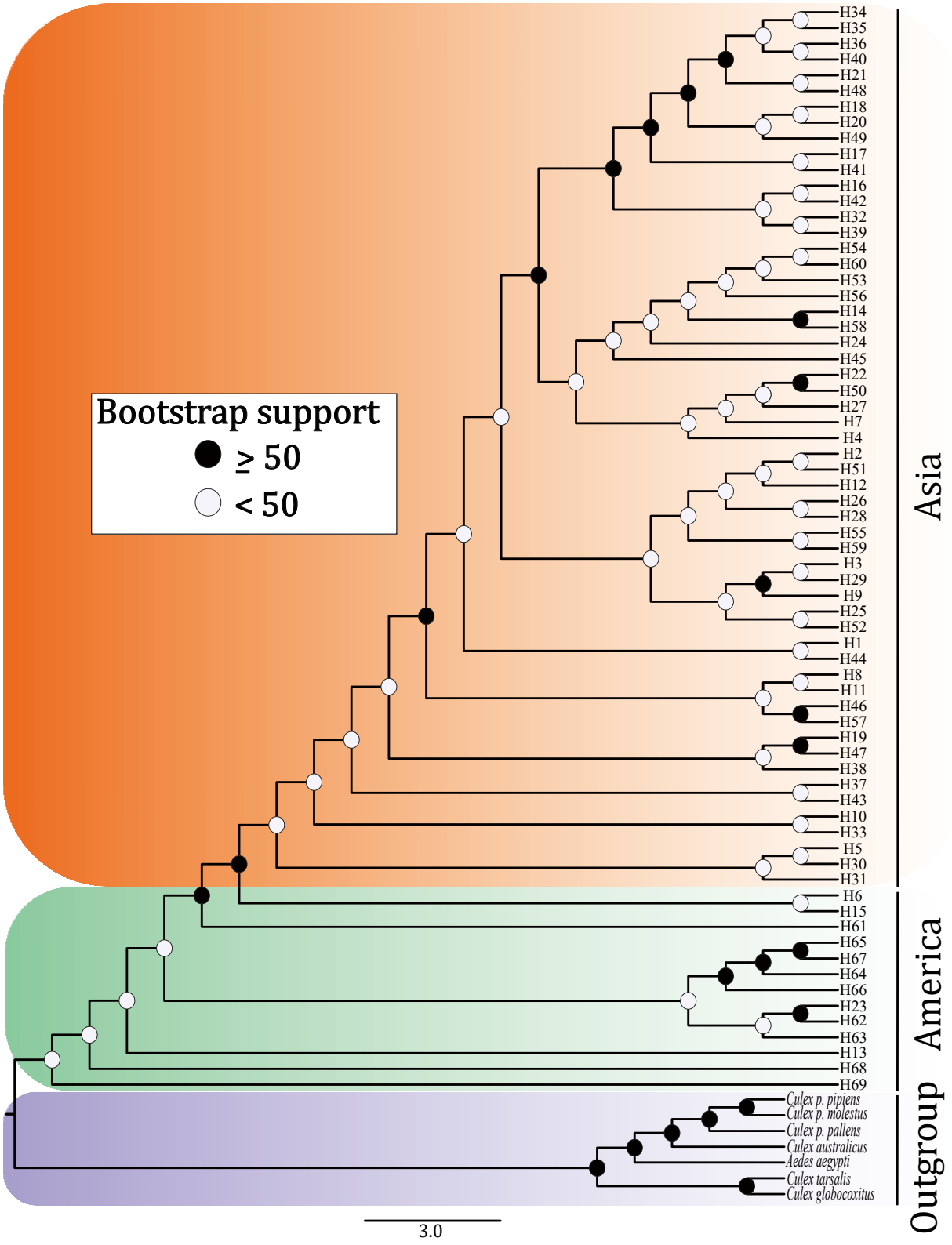

Fig. 4. Phylogenetic tree of *Culex quinquefasciatus* populations constructed from 69 *COI* gene haplotypes using Maximum Likelihood (ML). Evolutionary history was inferred using the Jukes Cantor evolutionary model, and the tree was obtained using 1,000 boostrap replicates. Branch support is indicated by Bootstrap values: black circles bootstrap values ≥ 50 and white circles bootstrap values < 50. The vertical lines indicate the continent Asia and America and the outgroup.

themselves (*d.f.* = 1.88, ss = 270.88, $p < 0.05$). However, no genetic structuring was found between continents (*d.f.* = 4; ss = 160.71, $p > 0.05$). The majority of genetic variation was observed among countries within continents (64.97%), followed by variation between continents (19.50%) and within countries (15.52%). These results indicate significant genetic differentiation among populations of *C. quinquefasciatus* at the country level within their respective continents. Conversely, the Mantel test did not reveal a significant correlation between genetic distance (F_{ST}) (Fig. 5) and geographic distance (km) ($R^2 = 0.03144$; $p > 0.05$).

DISCUSSION

The gathered information provides a comprehensive insight into the diversity and ecological adaptability of *C. quinquefasciatus*, marking the first study to cover its wide altitudinal range, from 1 m to 3,540 m. This suggests a broad adaptation to diverse habitats despite intraspecific competition. However, the concentration of occurrence records in urban areas raises the possibility of a sampling bias towards urban environments, potentially influencing its comparison with rural or peri-urban areas (Escalante 2003), which could explain its designation as a domestic mosquito in the South. This phenomenon may be linked to its predominant feeding on humans and animals, contrasting with its presence in urban areas (Rose et al. 2020).

The presence of the mosquito *C. quinquefasciatus* has been recorded in subtropical and tropical areas, leading to debates about its origin. Some authors suggest its origin in West Africa (Harbach 2012; Huang et al. 2023), while others point to Southeast Asia (Fonseca et al. 2006; Bhattacharya and Basu 2016; Negi and Verma 2018). Although the genetic information analyzed is distributed unevenly, that is, it is skewed, our results support the hypothesis of a possible Asian origin for *C. quinquefasciatus* and its global dispersion. This result, while consistent with the hypothesis of Asia as the species' origin (Fonseca et al. 2006; Bhattacharya and Basu 2016; Negi and Verma 2018), requires further study using mitochondrial or nuclear genomic information to corroborate our hypothesis. Globally, it has been observed that in other invasive mosquito species, global trade routes such as maritime or air transportation have facilitated their spread worldwide (Ibañez-Justicia 2020). Additionally, it has been documented that vehicles such as cars, trucks, and buses also contribute to its local dispersal, influencing the spread of mosquito-borne diseases (Hulme 2021). This passive, human-mediated flow could explain its presence on all continents, except polar regions, where samples have been found in container breeding sites such as metal, plastic, and tire containers (Bhattacharya and Basu 2016).

In our comprehensive review of the genetic information of *C. quinquefasciatus*, the mitochondrial gene *COI* emerged as the most utilized for intraspecific analyses due to its easy extraction, availability in public repositories, and maternal evolutionary history (Torres-Gutierrez et al. 2016). We identified 69 haplotypes, predominantly in Asia, suggesting this region as its possible native area. However, further studies incorporating nuclear or mitochondrial genetic

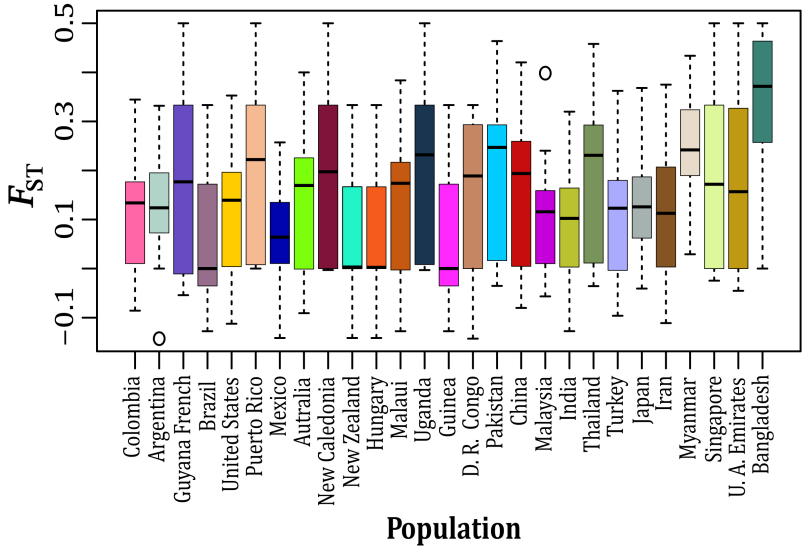

Fig. 5. Box plot of the genetic differentiation coefficient (F_{ST}) by countries indicating genetic structuring in a global context for natural populations of the southern house mosquito.

information are necessary to confirm this hypothesis (Bhattacharya and Basu 2016; Negi and Verma 2018).

In general, *COI* gene sequences were detected in a large proportion of the localities where *C. quinquefasciatus* has been recorded. However, the distribution of these sequences is notably skewed towards Asia, with 89% of the data originating from that region. Interestingly, South America, which contributes only 2.2% of the total data, exhibits greater global nucleotide and haplotype diversity than Asia, the presumed native region of *C. quinquefasciatus*. This finding is unexpected, as native regions of species typically show higher genetic diversity due to long-term evolutionary processes, while recently colonized areas tend to display reduced diversity due to recent introduction events, such as a recent founder effect (Amsellem et al. 2000; Cartaxo et al. 2011).

We propose two scenarios to explain this unusual pattern. The first relates to the uneven distribution of sequences across regions. Although Asia has a large number of sequences, if these come from genetically similar populations or if the sampling was more homogeneous—as seems to be the case, given that 1,044 of the 1,697 sequences come from Pakistan—this could result in lower haplotype diversity (*Hd*) compared to South America, which has only 42 sequences. In contrast, if the South American samples come from genetically diverse populations or different localities, this could increase *Hd* values in South America despite the smaller sample size (Nei 1987).

The second scenario may involve the occurrence of multiple independent introductions of genetically diverse individuals into South America, which could have increased genetic diversity through admixture with previously established populations in the region (Aguirre-Obando et al. 2015; Weaver et al. 2021). We propose that, in this context, historical trade routes and human migration may have facilitated these introductions, bringing mosquitoes from diverse source populations and consequently increasing the genetic diversity of South American populations (Tatem et al. 2006; Moreno-García et al. 2023).

Our results, although biased towards certain geographical localities, suggest that haplotypes are distributed differently according to altitude and the zones where they have been recorded. The distribution of haplotypes in mosquitoes in altitudinal and regional contexts is an area of study that is continually evolving. Although we did not find specific studies addressing the distribution of haplotypes by zones and/or land cover, there are investigations that explore genetic diversity in altitudinal contexts. One example is a study on *Aedes albopictus*—a vector of diseases such as dengue, Zika, and chikungunya—conducted in Ecuador, where the genetic diversity of mosquito populations was investigated. In this case, two haplotypes were identified, distributed across different regions of the country: one predominant in coastal areas and the Amazon basin, and another in the lowlands of the northeast (Carrazco-Montalvo et al. 2022). This pattern suggests a possible differential adaptation of haplotypes to various ecological and climatic conditions. Additionally, studies in the Swiss Alps have analyzed how different groups of terrestrial arthropods vary in their diversity and abundance along altitudinal gradients, relating these variations to factors such as available surface area and soil characteristics, which influence species composition and distribution (Gilgado et al. 2022). These findings highlight the lack of detailed analyses on the distribution of haplotypes in altitudinal contexts and the importance of identifying these patterns to better understand the genetic diversity of mosquito populations.

When evaluating patterns of genetic diversity, it's crucial to consider the geographic scale of studies. For instance, Pfeiler et al. (2013) conducted a study at a regional scale, focusing on data from specific states within the United States and Mexico. They found that *C. quinquefasciatus* exhibited very low haplotypic and nucleotide diversity ($Hd = 0.071$, $\pi = 0.00012$) compared to *C. tarsalis*. While this study offers a detailed perspective on genetic diversity within populations localized to a specific area, our results in North America showed considerably higher values ($Hd = 0.384$, $\pi = 0.022$). It's possible that the number of sequences used influenced our results; with a larger dataset, our sampling may be more robust and comprehensive (Escalante 2003).

Our neutrality test results suggest the presence of both significant and non-significant values at the global level. In our analysis of the genetic neutrality of *C. quinquefasciatus* at the continental level in the Americas, North America, and Asia, as well as in some countries within these continents for the *COI* gene, we observed negative Tajima's *D* values with statistical significance. These findings suggest an excess of rare polymorphisms, which is characteristic of a recent population expansion or purifying selection. In this case, the population has experienced rapid growth or the removal of unfavorable variants, leading to a deficit of intermediate alleles (Tajima 1989; Ramírez-Soriano et al. 2008; Porretta et al. 2012; Maynard et al. 2017).

Recent studies on mosquitoes have shown that negative Tajima's *D* values may be associated with vector control measures. These negative values indicate an excess of rare polymorphisms, suggesting a recent population expansion or a 'selective sweep' due to the removal of individuals susceptible to insecticides and

the proliferation of resistant genotypes. In studies on *Culex pipiens pallens* in China, negative Tajima's *D* values indicated a recent population expansion, possibly driven by selective pressure from vector control interventions, suggesting that populations respond to these measures by expanding resistant individuals (Zang et al. 2024). These observations highlight that human interventions, such as the use of insecticides, not only reduce mosquito populations but also impact the genetic structure of populations, promoting the selection of resistant variants.

Conversely, studies on *C. quinquefasciatus* have revealed populations with non-significant Tajima's *D* values, indicating evolution under neutrality. For instance, a genetic analysis of *C. pipiens* populations in Europe, including *C. quinquefasciatus*, showed genetic diversity without significant selective pressures, with Tajima's *D* values close to zero, suggesting that these populations are evolving according to the neutral model. These results suggest that *Culex* populations in some regions are not under intense selective pressure, such as that exerted by natural selection or vector control interventions, and are largely influenced by genetic drift and neutral mutations (Shaikevich and Zakharov 2010; Shaikevich et al. 2016). Understanding these patterns provides deep insight into the evolutionary processes that shape the population dynamics and genetic diversity of *C. quinquefasciatus* (Tajima 1989; Láruson and Reed 2021).

Understanding the genetic structure of *C. quinquefasciatus* is pivotal for devising effective strategies to control mosquito-borne diseases (Thiemann et al. 2012), as it provides insights into the genetic similarities and differences among the studied populations. In our research, the analysis of molecular variance (AMOVA) revealed significant genetic differentiation among populations at the country level, corroborating previous studies that suggest a complex population structure and extensive genetic diversity in *C. quinquefasciatus* (Wilke et al. 2014). However, some countries across all continents with genetic information for *C. quinquefasciatus* were represented by a single sequence, which led to the exclusion of these data from our analyses. F_{ST} is a measure of genetic differentiation between populations and depends on genetic variation within and between them. When a population is represented by a single individual, it is not possible to adequately estimate intrapopulation genetic variability. Studies have shown that small sample sizes can lead to biased estimates of F_{ST}, as genetic variability is not fully represented (Waples 2015). Hale et al. (2012) suggest that a minimum sample size of 25 to 30 individuals per population is necessary to accurately estimate allele frequencies and, consequently, measures of genetic differentiation. In the case of singleton populations, the absence of intrapopulation genetic variation can artificially inflate F_{ST} values, giving a false impression of greater genetic differentiation (Leberg 2002). Therefore, although our analyses were conducted on a broad geographic scale, some populations, despite having genetic information, did not have enough samples to be included in the analyses. We recommend increasing the number of genetic studies in these areas to improve understanding of the diversity and genetic structure of the species. Greater sampling efforts in these underrepresented regions could provide a more comprehensive view of the global genetic structure of *C. quinquefasciatus*, which is essential for developing control strategies for this vector.

CONCLUSIONS

Our study represents a pioneering effort to comprehensively assess the genetic diversity of *C. quinquefasciatus* using global data. Through our analyses, and despite the bias in genetic distribution, we propose the possibility that Asia is the origin of this species. While our findings underscore the presence of genetic variability on a global scale and across diverse regions, reflecting the impact of environmental and geographical factors on the species' genetic structure, we advocate for future research employing Bayesian methods to delve deeper into migration scenarios between continents. Given that the colonization of new regions by *C. quinquefasciatus* facilitates the spread of various mosquito-borne diseases—such as filariasis nematodes, protozoa, and arboviruses—our understanding of genetic structure, flow, and diversity holds significant implications for public health. This knowledge can inform programs aimed at assessing insecticide resistance, conducting molecular epidemiology studies, modeling and predicting disease risk, and implementing effective vector control measures.

Acknowledgments: The authors thank Universidad del Quindío.

Authors' contributions: All the authors contributed to the study conception and design. GCGE, JTS and OAAO prepared the material, collected the data, and contributed with analyses. The manuscript was written by GCGE, JJTS and OAAO. All the authors read and approved the final draft of the manuscript.

Competing interests: All authors declare that they have no conflict of interest.

Availability of data and materials: Raw data used in this article are available on the Zoological Studies server.

Consent for publication: Not applicable.

Ethics approval consent to participate: Not applicable.

REFERENCES

Aardema ML, Vonholdt BM, Fritz ML, Davis SR. 2020. Global evaluation of taxonomic relationships and admixture within the *Culex pipiens* complex of mosquitoes. Parasites and Vectors **13:**1–17. doi:10.1186/s13071-020-3879-8.

Aguirre-Obando OA, Dalla Bona AC, Duque LJE, Navarro-Silva MA. 2015. Insecticide resistance and genetic variability in natural populations of *Aedes* (*Stegomyia*) *aegypti* (Diptera: Culicidae) from Colombia. Zoologia **32:**14–22. doi:10.1590/S1984-46702015000100003.

Amsellem L, Noyer JL, Bourgeois TLE, Hossaert-Mckey M. 2000. Comparison of genetic diversity of the invasive weed *Rubus alceifolius* Poir. (Rosaceae) in its native range and in areas of introduction, using amplified fragment length polymorphism (AFLP) markers. Mol Ecol **9:**443–455. doi:10.1046/j.1365-294X.2000.00876.x.

Anoopkumar AN, Puthur S, Rebello S, Aneesh EM. 2019. Molecular characterization of *Aedes*, *Culex*, *Anopheles*, and Armigeres vector mosquitoes inferred by mitochondrial cytochrome oxidase I gene sequence analysis. Biologia (Bratisl) **74:**1125–1138. doi:10.2478/s11756-019-00231-0.

Aremu HK, Iyoha A, Ajayi E. 2022. Population Genetic Structure of *Culex quinquefasciatus* in Osogbo Metropolis, Nigeria inferred from Mitochondrial *COI* sequences. Research Square, pp. 1–18. doi:10.21203/rs.3.rs-2198391/v1.

Ashfaq M, Hebert PDN, Mirza JH, Khan AM, Zafar Y, Mirza MS. 2014. Analyzing mosquito (Diptera: Culicidae) diversity in Pakistan by DNA barcoding. PLoS ONE **9:**1–12. doi:10.1371/journal.pone.0097268.

Atyame CM, Delsuc F, Pasteur N, Weill M, Duron O. 2011. Diversification of *Wolbachia* endosymbiont in the *Culex pipiens* mosquito. Mol Biol Evol **28:**2761–2772. doi:10.1093/molbev/msr083.

Behura SK, Lobo NF, Haas B, de Bruyn B, Lovin DD, Shumway MF, Puiu D, Romero-Severson J, Nene V, Severson DW. 2011. Complete sequences of mitochondria genomes of *Aedes aegypti* and *Culex quinquefasciatus* and comparative analysis of mitochondrial dna fragments inserted in the nuclear genomes. Insect Biochem Mol Biol **28:**2761–2772. doi:10.1016/j.ibmb.2011.05.006.

Bhattacharya S, Basu P. 2016. The Southern House Mosquito, *Culex quinquefasciatus*: profile of a smart vector. J Entomol Zool Stud **4:**73–81.

Calhoun LM, Avery M, Jones LA, Gunarto K, King R, Roberts J, Burkot TR. 2007. Combined sewage overflows (CSO) are major urban breeding sites for *Culex quinquefasciatus* in Atlanta, Georgia. Am J Trop Med Hyg **77:**478–484. doi:10.4269/ajtmh.2007.77.478.

Cane RP, Li D, Turbitt E, Chambers GK. 2020. Molecular phylogenetic analysis of New Zealand mosquito species. New Zeal J Zool **47:**324–349. doi:10.1080/03014223.2020.1770305.

Carrazco-Montalvo A, Ponce P, Villota SD, Quentin E, Muñoz-Tobar S, Coloma J, Cevallos V. 2022. establishment, genetic diversity, and habitat suitability of *Aedes albopictus* populations from Ecuador. Insects **13:**305. doi:10.3390/insects13030305.

Cartaxo MFS, Ayres CFJ, Weetman D. 2011. Loss of genetic diversity in *Culex quinquefasciatus* targeted by a lymphatic filariasis vector control program in Recife, Brazil. Trans R Soc Trop Med Hyg **105:**491–499. doi:10.1016/j.trstmh.2011.05.004.

Chaulk AC, Carson KP, Whitney HG, Fonseca DM, Chapman TW. 2016. The arrival of the northern house mosquito *Culex pipiens* (Diptera: Culicidae) on Newfoundland's avalon peninsula. J Med Entomol **53:**1364–1369. doi:10.1093/jme/tjw105.

Clement M, Posada D, Crandall KA. 2000. TCS: A computer program to estimate gene genealogies. Mol Ecol **9:**1657–1659. doi:10.1046/j.1365-294X.2000.01020.x.

CoreTeam R. 2017. R: A Language and Environment for Statistical Computing. 2.

Daravath SS, Siddaiah M, ReddyaNaik B. 2015. Molecular Characterization and Phylogenetic Analysis of *Culex quinquefasciatus* by DNA Barcoding. Adv Entomol **03:**118–124. doi:10.4236/ae.2015.33014.

Didier G. 2024. Package "ape": Analyses of Phylogenetics and Evolution, pp. 1–298.

Donkor ES, Dayie Nicholas TKD, Adiku TK. 2014. Bioinformatics with basic local alignment search tool (BLAST) and fast alignment (FASTA). J Bioinforma Seq Anal **6:**1–6. doi:10.5897/ijbc2013.0086.

Dray S, Dufour AB. 2007. The ade4 package: Implementing the duality diagram for ecologists. J Stat Softw **22:**1–20. doi:10.18637/jss.v022.i04.

Dumas E, Atyame CM, Malcolm CA, Goff GL, Unal S, Makoundou, Pasteur N, Weill M, Duron N. 2016. Molecular data reveal a cryptic species within the *Culex pipiens* mosquito complex. Insect Mol Biol **25:**800–809. doi:10.1111/imb.12264.

Escalante T. 2003. ¿Cuántas especies hay? los estimadores no paramétricos de Chao. Elementos **52:**53–56.

Ferreira FC, Videvall E, Seidl CM, Wagner N, Kilpatrick AM, Fleischer RC, Fonseca DM. 2022. Transcriptional response of individual Hawaiian *Culex quinquefasciatus* mosquitoes to the avian malaria parasite *Plasmodium relictum*. Malar J **21:**1–13. doi:10.1186/s12936-022-04271-x.

Fonseca DM, Smith JL, Wilkerson RC, Fleischer RC. 2006. Pathways of expansion and multiple introductions illustrated by large genetic differentiation among worldwide populations of the southern house mosquito. Am J Trop Med Hyg **74:**284–289. doi:10.4269/ajtmh.2006.74.284.

Gilgado JD, Rusterholz HP, Braschler B, Zimmermann S, Chittaro Y, Baur B. 2022. Six groups of ground-dwelling arthropods show different diversity responses along elevational gradients in the Swiss Alps. PLoS ONE **17:**1–23. doi:10.1371/journal.pone.0271831.

Goudet J, Jombart T, Kamvar Z, Hardy O. 2022. Package "hierfstat" Estimation and Tests of Hierarchical *F*-Statistics, pp. 1–73.

Gouge DH, Li S, Walker K, Sumner C. 2019. Mosquitoes: Biology and Integrated Mosquito Management. CALS Publications Archive. The University Of Arizona. Available at: https://arizona.openrepository.com/handle/10150/625410.

Guilherme F, Vicente JR, Carretero MA, Farinha-Marques P. 2024. Mapping multigroup responses to land cover legacy for urban biodiversity conservation. Biol Conserv **291:**110508. doi:10.1016/j.biocon.2024.110508.

Gunay F, Alten B, Simsek F, Aldemir A, Linton YM. 2015. Barcoding Turkish *Culex* mosquitoes to facilitate arbovirus vector incrimination studies reveals hidden diversity and

new potential vectors. Acta Trop **143:**112–120. doi:10.1016/j.actatropica.2014.10.013.

Hadley A, Hester J, Chang W. 2022. Package "devtools" Tools to Make Developing R Packages Easier, pp. 1–33.

Hale ML, Burg TM, Steeves TE. 2012. Sampling for microsatellite-based population genetic studies: 25 to 30 individuals per population is enough to accurately estimate allele frequencies. PLoS ONE **7:**e45170. doi:10.1371/journal.pone.0045170.

Harbach RE. 2012. *Culex pipiens*: Species versus species complex - Taxonomic history and perspective. J Am Mosq Control Assoc **28:**10–23. doi:10.2987/8756-971X-28.4.10.

Harley D, Sleigh A, Ritchie S. 2001. Ross river virus transmission, infection, and disease: A cross-disciplinary review. Clin Microbiol Rev **14:**909–932. doi:10.1128/CMR.14.4.909-932.2001.

Hasan AU, Suguri S, Ahmed SMU, Fujimoto C, Harada M, Rahman SM, Zaman RU, Kakehi Y. 2009. Molecular phylogeography of *Culex quinquefasciatus* mosquitoes in central Bangladesh. Acta Trop **112:**106–114. doi:10.1016/j.actatropica.2009.07.005.

Hickner PV, Debruyn B, Lovin DD, Mori A, Saski CA, Severson DW. 2011. Enhancing genome investigations in the mosquito *Culex quinquefasciatus* via BAC library construction and characterization. BMC Res Notes **4:**358. doi:10.1186/1756-0500-4-358.

Holder M, Lewis PO. 2003. Phylogeny estimation: Traditional and Bayesian approaches. Nat Rev Genet **4:**275–284. doi:10.1038/nrg1044.

Huang S, Molaei G, Andreadis TG. 2011. Reexamination of *Culex pipiens* hybridization zone in the eastern United States by ribosomal DNA-based single nucleotide polymorphism markers. Am J Trop Med Hyg **85:**434–441. doi:10.4269/ajtmh.2011.10-0679.

Huang X, Athrey GN, Kaufman PE, Fredregill C, Slotman A. 2023. Effective population size of *Culex quinquefasciatus* under insecticide-based vector management and following Hurricane Harvey in Harris County, Texas. Front Genet **14:**1297271. doi:10.3389/fgene.2023.1297271.

Hulme PE. 2021. Unwelcome exchange: International trade as a direct and indirect driver of biological invasions worldwide. One Earth **4:**666–679 . doi:10.1016/j.oneear.2021.04.015.

Ibañez-Justicia A. 2020. Pathways for introduction and dispersal of invasive *Aedes* mosquito species in Europe: a review. J Eur Mosq Control Assoc **38:**1–10.

Ishtiaq F, Guillaumot L, Clegg SM, Phillimore AB, Black RA, Owens IPF, Mundy NI, Sheldon BC. 2008. Avian haematozoan parasites and their associations with mosquitoes across Southwest Pacific Islands. Mol Ecol **17:**4545–4555. doi:10.1111/j.1365-294X.2008.03935.x.

Jombart T. 2023. Package "adegenet" Exploratory Analysis of Genetic and Genomic Data. Bioinformatics, pp. 1403–1405.

Kamvar ZN, Tabima JF, Grünwald NJ. 2014. *Poppr*: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ **4:**1–14. doi:10.7717/peerj.281.

Katoh K, Rozewicki J, Yamada KD. 2018. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform **20:**1160–1166. doi:10.1093/bib/bbx108.

Koosha M, Oshaghi MA, Sedaghat MM, Vatandoost H, Azari-Hamidian S, Abai MR, Mohtarami F. 2017. Sequence analysis of mtDNA *COI* barcode region revealed three haplotypes within *Culex pipiens* assemblage. Exp Parasitol **181:**102–110. doi:10.1016/j.exppara.2017.08.003.

Láruson ÁJ, Reed FA. 2021. Population Genetics with R: An Introduction for Life Scientists.

Leberg PL. 2002. Estimating allelic richness: Effects of sample size and bottlenecks. Mol Ecol **11:**2445–2449. doi:10.1046/j.1365-294x.2002.01612.x.

Leite LAR. 2012. Mitochondrial pseudogenes in insect DNA barcoding: differing points of view on the same issue. Biota Neotrop **12:**301–308. doi:10.1590/S1676-06032012000300029.

Lorenz C, Alves JMP, Foster PG, Suesdek L, Sallum MAM. 2021. Phylogeny and temporal diversification of mosquitoes (Diptera: Culicidae) with an emphasis on the Neotropical fauna. Syst Entomol **46:**798–811. doi:10.1111/syen.12489.

Low VL, Lim PE, Chen CD, Lim YA, Tan TK, Norma-Rashid Y, Lee HL, Sofian-Azirun M. 2014. Mitochondrial DNA analyses reveal low genetic diversity in *Culex quinquefasciatus* from residential areas in Malaysia. Med Vet Entomol **28:**157–168. doi:10.1111/mve.12022.

Maekawa Y, Pemba D, Kumala J, Gowelo S, Higa Y, Futami K, Sawabe K, Tsuda Y. 2021. DNA barcoding of mosquitoes collected through a nationwide survey in 2011 and 2012 in Malawi, Southeast Africa. Acta Trop **213:**105742. doi:10.1016/j.actatropica.2020.105742.

Maintainer JB. 2024. Package "usethis" Automate Package and Project Setup, pp. 3–77.

Maynard AJ, Ambrose L, Cooper RD, Chow WK, Davis JB, Muzari MO, Van den Hurk AF, Hall-Medelin S, Hasty JM, Burkot TR, Bangs MJ, Reimer LJ, Butafa C, Lobo NF, Syafruddin D, Maung TN, Ahmad R, Beebe NW. 2017. Tiger on the prowl: Invasion history and spatio-temporal genetic structure of the Asian tiger mosquito *Aedes albopictus* (Skuse 1894) in the Indo-Pacific. PLoS Negl Trop Dis **11:**e0005546. doi:10.1371/journal.pntd.0005546.

Moreno-García M, González-Acosta C, Barón Olivares H, Correa-Morales F. 2023. Ecología evolutiva de la resistencia a insecticidas en el mosquito *Aedes* (*Stegomyia*) *aegypti*. Inventio **19:**1–13. doi:10.30973/inventio/2023.19.47/1.

Murugan K, Vadivalagan C, Karthika P, Panneerselvam C, Paulpandi M, Subramaniam J, Wei H, Aziz AT, Alsalhi MS, Devanesan S, Nicoletti M, Paramasivan R, Parajulee MN, Benelli G. 2016. DNA barcoding and molecular evolution of mosquito vectors of medical and veterinary importance. Parasitol Res **115:**107–121. doi:10.1007/s00436-015-4726-2.

Nei M. 1987. Molecular evolutionary genetics. Columbia University Press, NY, USA.

Negi C, Verma P. 2018. Review on *Culex quinquefasciatus*: Southern House Mosquito. Int J Life-Sciences Sci Res **4:**1563–1566.

Panda D, Barik TK. 2022. Molecular characterization and genetic divergence of seven *Culex* mosquito (Diptera: Culicidae) species using Mt *COI* gene from Odisha State, India. J Basic Appl Zool **83:**41. doi:10.1186/s41936-022-00305-7.

Pfeiler E, Flores-López CA, Mada-Vélez JG, Escalante-Verdugo J, Markow TA. 2013. Genetic diversity and population genetics of mosquitoes (Diptera: Culicidae: *Culex* spp.) from the sonoran desert of North America. Sci World J **2013:**1–11. doi:10.1155/2013/724609.

Porretta D, Mastrantonio V, Bellini R, Somboon P, Urbanelli S. 2012. Glacial history of a modern invader: phylogeography and species distribution modelling of the Asian tiger mosquito *Aedes albopictus*. PLoS ONE **7:**e44515. doi:10.1371/journal.pone.0044515.

Paradis E, Jombart T, Schliep K, Potts A, Winter D. 2016. Package 'PEGAS': Population and evolutionary genetics analysis system, pp. 1–82. Available at: http://ape-package.ird.fr/pegas.html.

Quintero L, Navarro JC. 2012. Filogenia intraespecífica y variabilidad genética de *Culex quinquefasciatus* Say (Diptera: Culicidae) con los genes mitocondriales *ND5* y *COI*. Bol Malariol y Salud Ambient **52:**46–65.

Ramírez-Soriano A, Ramos-Onsins SE, Rozas J, Calafell F, Navarro A. 2008. Statistical power analysis of neutrality tests under demographic expansions, contractions and bottlenecks with recombination. Genetics **179:**555–567. doi:10.1534/genetics.107.083006.

Rasgon JL, Cornel AJ, Scott TW. 2006. Evolutionary history of a mosquito endosymbiont revealed through mitochondrial hitchhiking. Proc R Soc B Biol Sci **273:**1603–1611. doi:10.1098/rspb.2006.3493.

Rose NH, Sylla M, Badolo A, Lutomiah J, Ayala D, Aribodor OB, Ibe N, Akorli J, Otoo S, Mutebi JP, Kriete A, Ewing EG, Sang R, Gloria-Soria A, Powell JR, Baker RE, White BJ, Crawford JE, McBride CS. 2020. Climate and Urbanization Drive Mosquito Preference for Humans. Curr Biol **30:**3570–3579.e6. doi:10.1016/j.cub.2020.06.092.

Sang R, Kioko E, Lutomiah J, Warigia M, Ochieng C, Guinn MO, Lee JS, Koka H, Godsey M, Hoel D, Hanafi H, Miller B, Schnabel D, Breiman RF, Richardson J. 2010. Rift Valley fever virus epidemic in Kenya, 2006/2007: The entomologic investigations. Am J Trop Med Hyg **83(2_Suppl):**28–37. doi:10.4269/ajtmh.2010.09-0319.

Shaikevich EV, Vinogradova EB, Bouattour A, de Almeida APG. 2016. Genetic diversity of *Culex pipiens* mosquitoes in distinct populations from Europe: Contribution of *Cx. quinquefasciatus* in Mediterranean populations. Parasites and Vectors **9:**1–16. doi:10.1186/s13071-016-1333-8.

Shaikevich E, Zakharov I. 2010. Polymorphism of mitochondrial *COI* and nuclear ribosomal ITS2 in the *Culex pipiens* complex and in *Culex torrentium* (Diptera: Culicidae). Acta Trop **4:**161–174. doi:10.3897/compcytogen.v4i2.45.

Sharma AK, Chandel K, Tyagi V, Mendki MJ, Tikar SN, Sukumaran D. 2013. Molecular phylogeney and evolutionary relationship among four mosquito (Diptera: Culicidae) species from India using PCR-RFLP. J Mosq Res **3:**58–64.

Smith JL, Fonseca DM. 2004. Rapid assays for identification of members of the *Culex* (*Culex*) *pipiens* complex, their hybrids, and other sibling species (Diptera: Culicidae). Am J Trop Med Hyg **70:**339–345. doi:10.4269/ajtmh.2004.70.339.

Tajima F. 1989. Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. Pharmatherapeutica **3:**607–612.

Talaga S, Leroy C, Guidez A, Dusfour I, Girod R, Dejean A, Murienne J. 2017. DNA reference libraries of French Guianese mosquitoes for barcoding and metabarcoding. PLoS ONE **12:**e0176993. doi:10.1371/journal.pone.0176993.

Tatem AJ, Hay SI, Rogers DJ. 2006. Global traffic and disease vector dispersal. Proc Natl Acad Sci USA **103:**6242–6247. doi:10.1073/pnas.0508391103.

Thanchomnang T, Intapan PM, Tantrawatpan C, Lulitanond V, Chungpivat S, Taweethavonsawat P, Kaewkong W, Sanpool O, Janwan P, Choochote W, Maleewong W. 2013. Rapid detection and identification of *Wuchereria bancrofti*, *Brugia malayi*, *B. pahangi*, and *Dirofilaria immitis* in mosquito vectors and blood samples by high resolution melting real-time PCR. Korean J Parasitol **51:**645–650. doi:10.3347/kjp.2013.51.6.645.

Thankachan M, Surya P, Sebastian CD. 2023. Molecular Identification and phylogenetic analysis of mosquito vectors from Mananthavady Taluk, Wayanad, Kerala, India. J Vector Borne Dis **60:**8. doi:10.4103/0972-9062.361166.

Thiemann TC, Lemenager DA, Kluh S, Carrol BD, Lothrop HD, Reisen WK. 2012. Spatial variation in host feeding patterns of *Culex tarsalis* and the *Culex pipiens* complex (Diptera: Culicidae) in California. J Med Entomol **49:**903–916. doi:10.1603/ME11272.

Torres-Gutierrez C, Bergo ES, Emerson KJ, de Oliveira TMP, Greni S, Sallum MAM. 2016. Mitochondrial *COI* gene as a tool in the taxonomy of mosquitoes *Culex* subgenus Melanoconion. Acta Trop **164:**137–149. doi:10.1016/j.actatropica.2016.09.007.

Van Den Eynde C, Sohier C, Matthijs S, De Regge N. 2022. Japanese encephalitis virus interaction with Mosquitoes: a review of vector competence, vector capacity and mosquito immunity. Pathogens **11:**317. doi:10.3390/pathogens11030317.

Waldbieser GC, Bilodeau AL, Nonneman DJ. 2003. Complete sequence and characterization of the channel catfish mitochondrial genome. DNA Seq - J DNA Seq Mapp **14:**265–277. doi:10.1080/1042517031000149057.

Waples RS. 2015. Testing for hardy-weinberg proportions: Have we lost the plot? J Hered **106:**1–19. doi:10.1093/jhered/esu062.

Weaver SC, Forrester NL, Liu J, Vasilakis N. 2021. Population bottlenecks and founder effects: implications for mosquito-borne arboviral emergence. Nat Rev Microbiol **19:**184–195. doi:10.1038/s41579-020-00482-8.

Weir B, Cockerham CC. 1984. Estimating *F*-Statistics for the Analysis of Population Structure. Evolution **38:**1358–1370. doi:10.2307/2408641.

Wilke A, Vidal P, Suesdek L, Marrelli M. 2014. Population genetics of neotropical *Culex quinquefasciatus* (Diptera: Culicidae). Parasit Vectors **7:**468. doi:10.1186/preaccept-1270950864130552.

Winter D, Green P, Kamvar Z, Gosselin T. 2013. Package "mmod" Modern Measures of Population Differentiation. R package version **1:**1–7.

Zang C, Wang X, Liu Y, Wang H, Sun Q, Cheng P, Liu H. 2024. *Wolbachia* and mosquitoes: Exploring transmission modes and coevolutionary dynamics in Shandong Province, China. PLoS Negl Trop Dis **18:**e0011944. doi:10.1371/journal.pntd.0011944.

Supplementary materials

Table S1. List of mitochondrial gene names and their abbreviations. Adapted from Waldbieser et al. (2003), for *Culex quinquefasciatus*. (download)

Table S2. Global compilation for *Culex quinquefasciatus* of altitude and coverage data derived from information retrieved from genetic and geographical databases. (download)

Table S3. Global haplotypic diversity from the *COI* gene for natural populations of *Culex quinquefasciatus*. The translated amino acid sequence is represented in uppercase blue letters, above the third nucleotide of its corresponding codon. Invariable sites are indicated with dots, whereas alternative nucleotides are represented in uppercase black letters. (download)