



**HAL**  
open science

## Population structure of the invasive Asian tiger mosquito, *Aedes albopictus*, in Europe

Margaret K Corley, Luciano Cosme, Peter A Armbruster, Maria Sharakhova, Alessandra Della Torre, Beniamino Caputo, Verena Pichler, Anna Bega, Sandra Urbanelli, Ademir Martins, et al.

### ► To cite this version:

Margaret K Corley, Luciano Cosme, Peter A Armbruster, Maria Sharakhova, Alessandra Della Torre, et al.. Population structure of the invasive Asian tiger mosquito, *Aedes albopictus*, in Europe. 2024. hal-04782404

**HAL Id: hal-04782404**

**<https://hal.science/hal-04782404v1>**

Preprint submitted on 14 Nov 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Population structure of the invasive Asian tiger mosquito, *Aedes albopictus*, in Europe

Margaret Corley<sup>1</sup>, Luciano Cosme<sup>2</sup>, Peter Armbruster<sup>3</sup>, Maria Sharakhova<sup>4</sup>, Alessandra della Torre<sup>5</sup>, Beniamino Caputo<sup>5</sup>, VERENA PICHLER<sup>5</sup>, Anna Bega<sup>6</sup>, Sandra Urbanelli<sup>5</sup>, Ademir Martins<sup>7</sup>, Enkelejda Velo<sup>8</sup>, Perparim Kadriaj<sup>8</sup>, Roger Eritja<sup>9</sup>, Nigel Beebe<sup>10</sup>, Chun-Hong Chen<sup>11</sup>, Jacob Crawford<sup>12</sup>, Michael Fontaine<sup>13</sup>, Sebastien Boyer<sup>14</sup>, Richard Gill<sup>15</sup>, Trang Huynh<sup>16</sup>, Kevin Maringer<sup>17</sup>, Andrew Maynard<sup>18</sup>, Shomen Mukherjee<sup>19</sup>, Leonard Munstermann<sup>20</sup>, Sinnathamby Surendran<sup>21</sup>, Isra Wahid<sup>22</sup>, Francis Schaffner<sup>23</sup>, Rubén Bueno Mari<sup>24</sup>, Antonios Michaelakis<sup>25</sup>, Georgios Balatsos<sup>25</sup>, Mustafa Akiner<sup>26</sup>, Arianna Puggioli<sup>27</sup>, Cintia Horvath<sup>28</sup>, Daniel Bravo-Barriga<sup>29</sup>, Elton Rogozi<sup>8</sup>, Francisco Collantes<sup>30</sup>, Gilles Besnard<sup>31</sup>, Mihaela Kavran<sup>32</sup>, Ognyan Mikov<sup>33</sup>, Raquel Medialdea-Carrera<sup>34</sup>, Tanya Melillo<sup>34</sup>, Maria Louise Borg<sup>34</sup>, Thomas Johnson<sup>1</sup>, Kayleigh Hackett<sup>1</sup>, Tina Wu<sup>1</sup>, Joao Pinto<sup>35</sup>, Vera Valadas<sup>35</sup>, and Adalgisa Caccone<sup>1</sup>

<sup>1</sup>Yale University

<sup>2</sup>University of California Riverside

<sup>3</sup>Georgetown University

<sup>4</sup>Virginia Polytechnic Institute and State University

<sup>5</sup>University of Rome La Sapienza

<sup>6</sup>Federal State University of Education

<sup>7</sup>IOC-FIOCRUZ

<sup>8</sup>Institute of Public Health

<sup>9</sup>Centre d'Estudis Avançats de Blanes (CEAB-CSIC)

<sup>10</sup>The University of Queensland

<sup>11</sup>National Health Research Institutes

<sup>12</sup>Verily Life Sciences LLC

<sup>13</sup>MIVEGEC

<sup>14</sup>Institut Pasteur du Cambodge

<sup>15</sup>Imperial College London

<sup>16</sup>Pasteur Institute in Ho Chi Minh City

<sup>17</sup>Pirbright Institute

<sup>18</sup>University of Queensland

<sup>19</sup>Ahmedabad University School of Arts and Sciences

<sup>20</sup>Yale University School of Public Health

<sup>21</sup>University of Jaffna

<sup>22</sup>Hasanuddin University Faculty of Medicine

<sup>23</sup>Francis Schaffner Consultancy

<sup>24</sup>Laboratorios Lokímica

<sup>25</sup>Benaki Phytopathological Institute

<sup>26</sup>Recep Tayyip Erdogan Universitesi

<sup>27</sup>Centro Agricultura Ambiente CAA "G. Nicoli"

<sup>28</sup>University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca

<sup>29</sup>University of Extremadura

<sup>30</sup>Universidad de Murcia

<sup>31</sup>Entente Interdépartementale Rhône-Alpes pour la Démoustication (EIRAD)

<sup>32</sup>University of Novi Sad Faculty of Agriculture

<sup>33</sup>National Center of Infectious and Parasitic Diseases

<sup>34</sup>Malta Health Promotion and Disease Prevention Directorate

<sup>35</sup>Universidade Nova de Lisboa Instituto de Higiene e Medicina Tropical

October 04, 2024

### Hosted file

E&E\_main\_manuscript\_text\_final.docx available at <https://authorea.com/users/833825/articles/1226654-population-structure-of-the-invasive-asian-tiger-mosquito-aedes-albopictus-in-europe>

## **Population structure of the invasive Asian tiger mosquito, *Aedes albopictus*, in Europe**

**Authors:** Margaret K. Corley, Luciano Veiga Cosme, Peter A. Armbruster, Maria Sharakhova, Alessandra della Torre, Beniamino Caputo, Verena Pichler, Anna Bega, Sandra Urbanelli, Ademir Jesus Martins, Enkelejda Velo, Perparim Kadriaj, Roger Eritja, Nigel Beebe, Chun-Hong Chen, Jacob E. Crawford, Michael Fontaine, Sebastien Boyer, Richard J. Gill, Trang Huynh, Kevin Maringer, Andrew Maynard, Shomen Mukherjee, Leonard E. Munstermann, Sinnathamby Noble Surendran, Isra Wahid, Francis Schaffner, Rubén Bueno Marí, Antonios Michaelakis, Georgios Balatsos, Muhammet Mustafa Akiner, Arianna Puggioli, Cintia Horvath, Daniel Bravo-Barriga, Elton Rogozi, Francisco Collantes, Gilles Besnard, Mihaela Kavran, Ognyan Mikov, Raquel Medialdea-Carrera, Tanya Melillo, Maria Louise Borg, Thomas Johnson, Kayleigh Hackett, Tina Wu, João Pinto, Vera Valadas, Adalgisa Caccone

### **Author Affiliations:**

Margaret K. Corley: Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA.

Luciano V. Cosme\*: Department of Entomology, University of California, Riverside, CA, USA

Peter A. Armbruster: Department of Biology, Georgetown University, Washington, DC, USA

Maria Sharakhova: Department of Entomology and the Fralin Life Science Institute, Virginia Polytechnic and State University, Blacksburg, VA, 24061, USA

Alessandra della Torre: Department of Public Health & Infectious Diseases, Sapienza University of Rome, Rome, Italy

Beniamino Caputo: Department of Public Health & Infectious Diseases, Sapienza University of Rome, Rome, Italy

Verena Pichler: Department of Public Health & Infectious Diseases, Sapienza University of Rome, Rome, Italy

Anna Bega: Federal State University of Education. Moscow region, Mytishchi, Russia & Vavilov Institute of General Genetics in the Russian Academy of Sciences, Moscow, Russia

Sandra Urbanelli: Department of Environmental Biology, Sapienza University of Rome, Rome, Italy

Ademir Jesus Martins: Laboratório de Fisiologia e Controle de Artrópodes Vetores, Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro, RJ, Brazil

Enkelejda Velo: Department of Epidemiology and Infectious Diseases Control, Institute of Public Health, Tirana, Albania

Perparim Kadriaj: Department of Epidemiology and Infectious Diseases Control, Institute of Public Health, Tirana, Albania

Roger Eritja: Centre d'Estudis Avançats de Blanes, Spanish Research Council (CEAB-CSIC), Blanes, Spain.

Nigel Beebe: School of the Environment, University of Queensland Australia, St Lucia, Australia

Chun-Hong Chen: National Health Research Institutes, National Mosquito-Borne Disease Control Research Center & National Institute of Infectious Diseases and Vaccinology, Taiwan

Jacob E. Crawford: Verily Life Sciences. South San Francisco, CA, USA

Michael Fontaine: Infectious Diseases and Vectors: Ecology, Genetics, Evolution and Control (MIVEGEC), Université de Montpellier, Montpellier, France; AND University of Groningen, Groningen Institute for Evolutionary Life Sciences, Groningen, The Netherlands

Sebastien Boyer: Medical and Veterinary Entomology Unit, Institut Pasteur du Cambodge, 5 Bd Monivong, Phnom Penh, Cambodia.

Richard J. Gill: Georgina Mace Centre for the Living Planet, Department of Life Sciences, Silwood Park, Imperial College London, UK

Trang Huynh: Department of Medical Entomology and Zoonotics, Pasteur Institute in Ho Chi Minh city, Vietnam

Kevin Maringer: The Pirbright Institute, Pirbright, England, UK

Andrew Maynard: School of the Environment, University of Queensland Australia, St Lucia, Australia

Shomen Mukherjee: Biological and Life Sciences, School of Arts and Sciences, Ahmedabad University, Gujarat, India

Leonard E. Munstermann: Yale University School of Public Health, and Yale Peabody Museum, New Haven, CT, USA

Sinnathamby Noble Surendran: Department of Zoology, University of Jaffna, Jaffna, Sri Lanka

Isra Wahid: Center for Zoonotic and Emerging Diseases HUMRC, Faculty of Medicine Hasanuddin University

Francis Schaffner: Francis Schaffner Consultancy, Riehen, Switzerland

Rubén Bueno Marí: R&D Department, Laboratorios Lokímica, Valencia, Spain; European Center of Excellence for Vector Control, Rentokil Initial; & Department of Pharmacy, Pharmaceutical Technology and Parasitology, Parasite & Health Research Group, University of Valencia, Valencia, Spain.

Antonios Michaelakis: Laboratory of Insects & Parasites of Medical Importance, Scientific Directorate of Entomology and Agricultural Zoology, Benaki Phytopathological Institute, Kifissia, Greece

Georgios Balatsos: Laboratory of Insects & Parasites of Medical Importance, Scientific Directorate of Entomology and Agricultural Zoology, Benaki Phytopathological Institute, Kifissia, Greece

Muhammet Mustafa Akiner: Department of Biology, Recep Tayyip Erdogan University, Fener, Rize, Türkiye

Arianna Puggioli: Medical and Veterinary Entomology Department, Centro Agricoltura Ambiente CAA "G. Nicoli", Crevalcore, Italy

Cintia Horvath: Department of Parasitology and Parasitic Diseases, University of Agricultural Sciences and Veterinary Medicine of Cluj-Napoca, Cluj-Napoca, Romania

Daniel Bravo-Barriga: Parasitology and Parasitic Diseases Unit, Animal Health Department, Veterinary Faculty, University of Extremadura (UEx), Cáceres, Spain

Elton Rogozi: Vectors' Control Unit, Department of Epidemiology and Control of Infectious Diseases, Institute of Public Health of Albania, Tirana, Albania

Francisco Collantes: Department of Zoology and Physical Anthropology, Universidad de Murcia, Murcia, Spain

Gilles Besnard: Entente Interdépartementale Rhône-Alpes pour la Démoustication (EIRAD), Chindrieux, France

Mihaela Kavran: Faculty of Agriculture, University of Novi Sad, Novi Sad, Serbia

Ognyan Mikov: National Center of Infectious and Parasitic Diseases, Sofia, Bulgaria

Raquel Medialdea-Carrera: Infectious Disease Prevention and Control Unit (IDCU) – Health Promotion and Disease Prevention Directorate, Malta

Tanya Melillo: Infectious Disease Prevention and Control Unit (IDCU) – Health Promotion and Disease Prevention Directorate, Malta

Maria Louise Borg: Infectious Disease Prevention and Control Unit (IDCU) – Health Promotion and Disease Prevention Directorate, Malta

Thomas Johnson: Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA

Kayleigh Hackett: Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA

Tina Wu: Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA

João Pinto: Global Health and Tropical Medicine, Instituto De Higiene E Medicina Tropical, Universidade Nova De Lisboa, Lisbon, Portugal

Vera Valadas: Global Health and Tropical Medicine, Instituto De Higiene E Medicina Tropical, Universidade Nova De Lisboa, Lisbon, Portugal.

Adalgisa Caccone: Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA

**\*Corresponding Author:** [luciano.cosme@ucr.edu](mailto:luciano.cosme@ucr.edu)

## Abstract

The Asian tiger mosquito, *Aedes albopictus*, is currently the most widespread invasive mosquito species in the world. It poses a significant threat to human health, as it is a vector for several arboviruses. We used a SNP chip to genotype 748 *Ae. albopictus* mosquitoes from 41 localities across Europe, 28 localities in the native range in Asia, and four in the Americas. Using multiple algorithms, we examined population genetic structure and differentiation within Europe and across our global dataset to gain insight into the origin of the invasive European populations. We also compared results from our SNP data to those obtained using genotypes from 11 microsatellite loci (N=637 mosquitoes from 25 European localities) to explore how sampling effort and the type of genetic marker used may influence conclusions about *Ae. albopictus* population structure. While some analyses detected more than 20 clusters worldwide, we found mosquitoes could be grouped into seven distinct genetic clusters, with most European populations originating in East Asia (Japan or China). Interestingly, some populations in Eastern Europe did not share genetic ancestry with any populations from the native range or Americas, indicating that these populations originated from areas not sampled in this study. The SNP and microsatellite datasets found similar patterns of genetic differentiation in Europe, but the microsatellite dataset could not detect the more subtle genetic structure revealed using SNPs. Overall, data from the SNP chip offered a higher resolution for detecting the genetic structure and the potential origins of invasions.

**Keywords:** population structure, population genomics, invasive species, SNP chip, microsatellites, tiger mosquito, disease vector, *Aedes albopictus*



## Introduction

Over the last 50 years, the Asian tiger mosquito, *Aedes albopictus* (Skuse 1984), has spread rapidly to become an invasive species in tropical and temperate areas worldwide, invading every continent except Antarctica (1, 2). Its range expansion has been aided by human activities and its ability to breed in artificial containers and to feed on diverse hosts (3, 4). It was first observed in Europe in the late 1970s (5), and is now widespread there, with populations extending as far north as Germany (6).

The spread of *Ae. albopictus* across Europe and other areas is concerning because this aggressive human-biting mosquito is a competent vector of many arboviruses, and a proven vector of dengue, chikungunya, and Zika viruses (7, 8). The vast extent of this species range expansion means it holds the potential to instigate major epidemics of vector-borne diseases, including in areas previously free of such tropical diseases. Locally initiated cases of arbovirus disease in Europe have been linked to invasive *Aedes* mosquitoes since 2007 (9), with additional locally acquired cases of chikungunya, dengue (10), and Zika (11) occurring thereafter. With increasing range expansion and climatic changes, the occurrence of pathogen transmission by *Ae. albopictus* in Europe is only expected to grow (12).

There is also concern that as *Ae. albopictus* spreads, it may acquire new adaptations that will allow it to increase its population sizes and infect more people. Introduced species may acquire adaptations as their range increases, particularly if they colonize novel geographic areas or ecologies (13, 14). Indeed, there are indications that *Ae. albopictus* has already shifted its niche and adapted to local environments (15–17). There is mounting evidence that it can become homodynamic (i.e., breeds year-round) and thus over-winter at the adult-stage in areas of its invasive range, which could allow pathogens carried by it to become endemic in temperate regions (12, 18–20). Additionally, vector competence for arboviruses, such as chikungunya (21) and dengue (22), can rapidly evolve (23), and *Ae. albopictus* may also be under strong selective pressure to evolve insecticide resistance as management strategies to control arbovirus outbreaks in new areas intensify (24). Understanding the current population structure and past demographic history of *Ae. albopictus* has critical applications for assessing the potential risk this vector mosquito poses for human health and can offer insights into how to control these invasive populations (25–28) by enhancing efforts to determine the genetic bases of vector competence and insecticide resistance and predict invasion success in more temperate climates.

Past research has utilized a variety of genetic markers to understand the genetic diversity of *Ae. albopictus* in both its native and invasive ranges. Allozymes (31), mitochondrial DNA sequences (mtDNA) (32–40), Random Amplified Polymorphic DNA (RAPD) (41), microsatellites (33, 40, 42–44), ribosomal DNA (rDNA) (38, 39, 45–48), and more recently, double-digest RADseq (ddRADseq) (29, 49–51) have provided insights into population structure and phylogeography in Europe, to a degree. However, each method has various drawbacks and limitations, and the lack of a consistent set of markers across studies complicates comparisons, resulting in sometimes conflicting findings.

For example, some studies indicate that genetic diversity is much higher in the native (i.e. Southeast Asia) than in the invasive range, suggesting a loss of genetic diversity has occurred during the invasion process (52). However, other studies found evidence that multiple, independent introductions have resulted in relatively high levels of intra-population variability and genetic diversity in the invasive range (42). While some markers suggest a general lack of structure in Europe, others indicate genetically distinct populations: previously, Greece and Albania were found to be genetically distinct from other parts of Europe (42, 49, 51, 53), and two spatially isolated genetic lineages were identified in southern Russia (37, 54). Some of the apparent discrepancies in the observed degree of population structure, compared to gene flow and admixture, may be due to genuine differences among regions in the history of introductions and patterns of gene flow. For instance, older established invasive populations may trace their ancestry directly to populations from the native range or other early-established invasive populations. At the same time, areas that were more recently invaded may be characterized by

admixture of multiple populations that already existed in Europe. Some data support this hypothesis: populations from Italy, which is believed to have received relatively early introduction(s) from the United States and/or Japan (31, 42), were found to be genetically distinct (42, 50, 55), while ones from the Iberian Peninsula were found to have low levels of genetic structure, likely stemming from human-mediated geneflow and multiple, on-going introductions (36, 48). However, the lack of consistent markers and methods across the studies and sampling locations makes it challenging to interpret apparent differences among studies: it is often unclear whether the differences observed are genuine or merely a consequence of using different methods/marker with varying sensitivity levels.

To address outstanding questions and resolve some of the knowledge gaps from previous studies regarding the origin and spread of this invasive mosquito, we characterized the genetic structure of *Ae. albopictus* across Europe and the native range. We had three main objectives. Our first objective was to clarify the ancestry of *Ae. albopictus* populations across the invasive range in Europe and determine which population(s) in the native range are the most likely sources of invasion(s). Our second objective was to characterize the fine-scale genetic structure within Europe. Our third objective was to compare the patterns of genetic structure across Europe determined using single nucleotide polymorphisms (SNPs) to patterns determined using microsatellite markers.

To achieve these objectives, we determined the distribution of genome-wide polymorphisms in this species using the Aealbo SNP chip (56), a species specific genotyping tool that targets 175,296 SNPs and was previously demonstrated to successfully determine structure in populations across the native range (56). In contrast to previous studies, the SNP chip provides us with a more extensive set of informative genetic markers across the entire genome, allowing us to detect fine-scale structure that may not be evident when using less sensitive genetic markers. For the first time, we applied this SNP chip to characterize the genetic structure of *Ae. albopictus* in their invasive range. Specifically, we characterized the genetic structure for 688 wild-caught mosquitoes from 74 locations (41 in Europe, 28 in the native range, and four in the Americas). We used this data set to assess patterns and levels of genomic differentiation and diversity in European populations and between European and the native range. We included four American populations to capture secondary invasions into Europe that may have come from these continents (rather than directly from areas in the native range). We first evaluated global patterns of differentiation and invasion in Europe. Then we focused on selected European regions (Italy, Greece and Albania, the area around southern Russia, and the Iberian Peninsula) for which previous studies indicate somewhat conflicting conclusions regarding the history of invasion and spread and for which we have a good geographic representation. To gain further insight into the spread of *Ae. albopictus* into Europe over time, we also examined the extent of variation among temporally spaced populations from Italy, comparing mosquitoes collected in 1995 and 18-25 years later, in 2013-2020. In addition to using SNP data, we screened 637 mosquitoes from 25 locations in 15 European countries for genetic variation at 11 microsatellite loci and compared these results to the exact same 25 locations in our European SNP dataset. This provided the opportunity to explore how the type of genetic marker may influence conclusions about the population structure of *Ae. albopictus* across Europe, especially considering that all samples from the microsatellite dataset were from locations represented in the SNP dataset.

## **Materials and Methods**

### **SNP dataset**

#### *Sample Collection and Acquisition*

Mosquitoes used for SNP genotyping were collected between 1995 and 2021 and shipped to Yale University for processing. When available, we attempted to genotype 12 individuals from each sampling locality. In total, we genotyped 748 mosquitoes from 32 different countries (Figure 1): 440 individuals collected from the invasive range in Europe (41 localities in 17 countries), 260 individuals collected from the native range in Asia (28 localities in 13 countries), and 48 individuals from four locations in the Americas (two in the United States and two in Brazil) to assess whether introductions to Europe may have come from previously established invasions in North or South America rather than directly from the native range (Table S2).

#### *DNA extraction and species identification*

DNA was extracted from *Ae. albopictus* adults or larvae using the DNeasy Blood and Tissue kit (Qiagen, Valencia, CA), using modifications to the manufacturer's instructions that have been previously described for the purification of total DNA from insects (56). For larvae and other samples that could not be identified to species using morphology (N=75), we checked species assignments by sequencing a 1,537 bp fragment of the mitochondrial cytochrome oxidase subunit 1 COI gene and comparing it to the species known sequences (File S3). The protocols for PCR amplification and sequencing have been detailed previously (56). Only samples confirmed to be *Ae. albopictus* were sent for genotyping.

#### *SNP chip genotyping*

All SNP samples were genotyped using the Arealbo microarray (Affymetrix, Santa Clara, CA) in the Functional Genomics Core at the University of North Carolina at Chapel Hill. We obtained genotype calls using Axiom Analysis Suite Software v.5.1.1, running the "Best Practices Workflow" and the off-target variants (OTV) caller algorithm to identify miscalled clusters. Thresholds and other details used for quality control during data processing in Axiom Analysis Suite have been described previously (56).

We performed two distinct genotype calls: first, using the full dataset of samples from European locations and global data (from the native range and the Americas), which we refer to as the "Global" dataset, and second, exclusively on the samples from European locations, which we refer to as the "Europe" dataset. These genotype data were exported in VCF file format, and utilized for subsequent quality control filtering. The raw data used to obtain genotype calls are available at Zenodo.

#### **Microsatellite dataset**

For microsatellite analyses, 637 mosquitoes were collected between 2011 and 2020 from 25 locations in 15 European countries (Figure 1B). The number of individuals genotyped from each population ranged from 5 to 60 (mean = 26.5 mosquitoes; Table S3). Additional details on sample collection are provided in File S23.

All samples were genotyped at 11 microsatellite loci using previously published primers and protocols (42), with minor modifications. Specifically, PCR products were multiplexed in groups of two or three according to fluorescent dye (FAM, NED, HEX) and size interval and subjected to fragment analysis on an ABI 3130xl DNA analyzer (Applied Biosystems) at the Yale DNA Analysis Facility at Science Hill. At least three positive controls were used to account for inter-run variations in fragment size. Alleles were scored from electropherograms using the software GENEMARKER (SoftGenetics) (57). Additional details of the microsatellite loci and extraction and amplification methods are outlined in File S3.

#### **Statistical Analysis**

### *SNP data Quality Control*

After genotyping, we used stringent quality control measures to filter the samples and SNPs used in downstream analyses using Plink v 1.9 or v 2.0. Specifically, we removed all SNPs that did not uniquely map to autosomes or failed segregation tests based on laboratory crosses (as detailed in (56)). We also removed 1) loci with more than 10% missingness, 2) loci with a minor allele frequency smaller than 10% (or 1% for use in population structure analyses), 3) loci that failed Hardy-Weinberg tests with a threshold of 0.00001 for each population, 4) individuals with more than 20% missing loci, 5) individuals whose expected heterozygosity values deviated more than  $\pm 4$  standard deviations from the mean of all samples, which might indicate low DNA quality, contamination, or high inbreeding, and 6) related individuals using a KING kinship coefficient of  $>0.354$  to identify monozygotic twins and duplicate samples. Additional details of the quality control steps are described in [Files S6-S7](#).

### *Linkage Disequilibrium*

Linkage disequilibrium (LD), the correlation of frequencies of genes at different loci, can be increased in populations during bottleneck events and thus vary among populations depending on their history (58). We used PopLDDecay to calculate the half distance of maximum  $r^2$  (LD half-life) for the SNP dataset. Since a correlation between sample size and LD estimates has previously been observed for *Ae. albopictus* populations in the native range (56), we compared the half distance of the maximum  $r^2$  value only for populations in Europe and the native range with at least ten individuals. For populations with more than ten individuals, we randomly selected ten individuals to retain and dropped the excess individual(s) from analyses to standardize the sample size. We then compared patterns of LD decay among the sampling locations and broader geographical regions (e.g. Europe versus the native range) by plotting the LD decay for each area for each of the three chromosomes with the ggstatsplot package in R (59). Additional details of linkage disequilibrium analyses are in [File S8](#).

### *SNP Sets*

We were interested in retaining relatively rare alleles that may allow us to better distinguish between sampling locations. Therefore, we created a set of SNPs (with LD half-life pruning at  $r^2 < 0.01$ ) that only removed SNPs with a minor allele frequency (MAF) of less than 1%. This dataset contained 22,642 variants for the Global dataset and 20,968 variants for the Europe dataset ([Figure S2](#)).

For all population structure and differentiation analyses (described below), we report results for this SNP Set, which we called SNP Set 3 (filtered for LD half-life  $r^2 < 0.01$  and MAF  $> 0.01$ ). We also provide results from other sets of SNPs with different LD half-life and MAF filtering criteria for select analyses in the supplementary files and figures for comparative purposes (see [File S3](#)).

### *Population structure and differentiation*

To assess the relationship between European mosquito samples and those in the native range we included samples from 13 countries in eastern, southern, and southeastern Asia. These samples were previously analyzed by a study that examined accuracy of the Aealbo SNP chip and the population structure of *Ae. albopictus* mosquitoes in the native range (56). Since previous research found evidence that some introductions of *Ae. albopictus* to Europe may have originated in the Americas (46, 51), we included two locations from the United States and two from Brazil, which had not previously been analyzed. We performed structure analyses on the two SNP data sets: 1) the European, native range, and American samples combined (“Global” dataset, N=688), and 2) the European samples alone (“Europe” dataset, N=410).

We used three different clustering algorithms, Admixture (60), fastStructure (61), and LEA (62), on the Global and Europe datasets to cross-check consistency across methodologies and ensure our findings were robust. To facilitate comparison with the microsatellite dataset, we repeated the LEA structure analysis for a subset of 242 mosquitoes from the locations that were shared (i.e. overlapping) with the microsatellite dataset. Parameters used to run each algorithm are listed in [Table S6](#), and step-by-step procedures for these analyses are presented in [Files S9-S10](#).

To further evaluate population structure and visualize genetic distances among samples from different locations, we used principal component analysis (PCA), which was implemented using the R package LEA (62). We also used the R package adegenet to perform Discriminate Analysis of Principle Components (DAPC) (63, 64). For DAPC analyses we used the cross-validation function to choose the number of principal components (PCs) to retain. For both PCA and DAPC, we carried out analyses on both the Global and Europe datasets, as well as different subsets of European populations. We also repeated these analyses for the subset of 242 mosquitoes from the 24 locations overlapping with the microsatellite dataset. Step-by-step procedures for these analyses are presented in [Files S12-S14](#).

To quantify genetic differentiation among sampling locations, we used the R package StAMPP (65) to estimate the pairwise genetic differentiation ( $F_{ST}$ ), estimating  $F_{ST}$  values across each locus according to Weir & Cockerham (66), taking into account population size and using 100 bootstraps to estimate p-values and confidence intervals. For all  $F_{ST}$  analyses we used only locations with at least four individuals ([Files S19-S20](#)). We also examined potential patterns of isolation by distance among European populations, using a Mantel test to assess the correlation between genetic and geographic distances using the mantel.randtest function in the R package ade4 (67). We also repeated all  $F_{ST}$  and isolation by distance analyses for the subset from the 24 locations with four or more individuals that overlapped with the microsatellite dataset to facilitate comparisons between marker types.

### *Microsatellites*

We used several approaches to characterize population structure and determine the number of genetically distinct populations represented by our microsatellite data set. First, we analyzed individuals using the Bayesian clustering method of the software STRUCTURE v 2.3.4 (68–71). We carried out two versions of the analyses: the first used only genetic information to cluster our data, and the second used sampling locations as prior information to assist clustering (LOCPRIOR method). This second method is appropriate for data sets with few markers, few individuals, or where the signal of structure is relatively weak. We employed this second model to ensure that we could detect subtle population structures that may be too weak for standard structure models to detect using our 11 loci. We used an admixture model for both analyses and set parameters to 1,000,000 MCMC runs, with a burn-in of 100,000 and 10 replicates of each possible number of clusters (K) from 2 to 25. To assess the STRUCTURE results, we used the Evanno method (72) implemented in CLUMPAK (73). Then we plotted the proportions of ancestry for each individual, grouped by sampling location, in R. Additional details of STRUCTURE analyses can be found in [File S3](#).

We complemented clustering analyses with multivariate approaches, including Principal Component Analysis (PCA) and Discriminant Analysis of Principal Components (DAPC). PCA was implemented in the package LEA (62), and DAPC in the package adegenet (63, 64) in R, using the cross-validation function to choose the number of principal components (PCs) to retain. We also repeated analyses using various subsets of European populations to better visualize the clusters among samples in specific regions of Europe ([File S11](#)).

We evaluated the level of genetic differentiation among European sampling locations by estimating  $F_{ST}$  values using the Weir and Cockerham method (66), as implemented in the R package hierfstat (74). Using the mantel function in the R package vegan (75), we also used the Mantel test to

assess the correlation between genetic and geographic distances (76). As with the SNP datasets, we only utilized populations with at least four individuals.

## Results

### *Genotyping and Quality Control*

Of the 768 samples in the Global dataset submitted for genotyping, 712 mosquitoes passed all quality control criteria in the Axiom Analysis Suite Software's "Best Practices Workflow" and 113,823 variants were recommended after running the Off Target Variant (OTV) caller. Of the 454 samples in the Europe dataset submitted for genotyping, 423 passed all quality control criteria, and 117,981 SNPs were recommended after the OTV caller. Additional details of the quality control methods and workflow used in Axiom Analysis Suite are outlined in [Files S4-S5](#).

The SNP sets exported were subjected to additional quality control measures (see [Files S6-S7](#)). Filtering with Plink to remove the 2,047 SNPs that failed the segregation test, variants missing in more than 10% of samples, and variants with a minor allele frequency <10% resulted in a SNP set of 87,183 variants for the Global dataset and 85,306 variants for the Europe dataset ([Figure S1](#)). For both datasets, all variants passed the Hardy-Weinberg Equilibrium test. Additional details of the SNP sets are shown in [Tables S4-S5](#).

No individuals were missing >20% of variants. Four samples in the 712 genotyped individuals in the Global dataset were duplicates and were thus removed (two from KAG and two from BEN, [Table S2](#)). Of the 708 remaining samples, we removed nine with heterozygosity that deviated more than 4SD from the mean and eleven with high relatedness (> 0.354), leaving 688 individuals in the final Global dataset for subsequent analyses. Of the 423 European individuals genotyped, we removed five samples for having heterozygosity >4 SD from the mean and nine for having high relatedness (> 0.354), leaving 410 individuals in the Europe dataset. A summary of the number of individuals retained in each dataset is presented in [Table S4](#).

### *Linkage Disequilibrium*

We compared the half distance of the maximum  $r^2$  value for populations with at least ten individuals (N= 41 populations in 24 countries; [File S8](#)). The LD half-life estimates varied among locations within the European and native ranges, and the degree of LD also differed across the three chromosomes ([Figures S3-S4](#)). The LD half-distance tended to be the smallest and showed the least variability among populations for chromosome 1, while chromosome 2 had the most significant variability and the highest LD half-distance estimates ([Figure S3](#)). Longer LD half-distance estimates on chromosome 2 indicate that linkage persists over greater distances, possibly indicating a lower recombination rate or selection patterns on this chromosome.

There was not a consistent correlation between patterns of LD decay and geographic region. The mean half distance of the maximum  $r^2$  value was somewhat higher for European compared to Asian populations for chromosomes 1 and 3. In contrast, Asian populations' mean values were slightly higher than European ones for chromosome 2. However, none of these differences were statistically significant ([Figure S4](#)).

### *Genetic ancestry*

#### *Global dataset*

The three clustering algorithms indicated a K value between 20 and 23 as the best number of ancestral clusters for the Global dataset (Table S7). Figure 2A shows a representative diagram of the clustering result for K=20. There was evidence of at least some admixture in most European sampling locations and many native range and American populations. Nonetheless, there was still clear evidence of population structure (Figure 2, Figure S5).

Three clusters observed in the K=20 plot were present in both Europe and the native range. Specifically, the orange cluster, which was the most prominent in many southern and western European locations, was also observed in China (Hunan and Hainan) and, to a lesser extent, in Vietnam. The bright yellow cluster in the Iberian Peninsula, France, Armenia (and mixed in small amounts in other European locations) was also observed in the United States and in Japan (Utsonomyia). Finally, Greece showed evidence of admixture with the cyan cluster found in the south and southeast Asian locations and parts of China, particularly Yunnan.

Several European clusters were either absent or poorly represented in the native range. For example, the three Italian locations with historic samples collected in the 1990s each formed a cluster, with the teal cluster present in Desenzano (from 1995) also observed as admixture in other European regions, but not in the native range. The light green cluster that constituted most of the ancestry in another Italian location (Brescia, collected in 1995) was found admixed among individuals in both the United States and Recife, Brazil, but was completely absent from Asia. A bright green cluster observed primarily in Croatia, Albania, and Greece was also absent from the native range. Notably, the purple clusters were confined to the most eastern European locations, suggesting mosquitoes in this region are entirely distinct not only from populations sampled in the native range but also from other parts of Europe. The same was true for the blue and red clusters observed only in Spain and Aliaga, Turkey (Figure 2A). Ancestry matrices derived from other algorithms showed similar patterns (Figure S5).

The PCA of the global dataset indicates that mosquitoes from South and most of Southeast Asia cluster together, separately from other samples, except for some individuals from China. Samples from South America also form a cluster, although this cluster is less distant from the remaining samples than those formed by South and Southeast Asia. In contrast, East Asian samples spread along PC 1 substantially overlap with Southern and Western European samples. Samples from North America also overlap with this cluster, particularly with mosquitoes from Southern Europe and Japan. Mosquitoes in Eastern Europe form two distinct clusters, with some populations overlapping with East Asia and Southern Europe, and the four farthest east locations (Russia, Georgia, Armenia, and Ukraine) forming their cluster at the bottom of PC2. However, when PC3 is plotted against PC1, these four eastern locations overlap with parts of Southern Europe and East Asia (Figure 3B).

DAPC, which uses discriminant functions to accentuate the genetic differences between groups, identified six clusters in the SNP Global dataset. One central cluster consisted of most samples from southern and western Europe, locations in eastern Europe west of the Black Sea, and most of East Asia. As with the PCA, South and Southeast Asia form their own cluster, which overlaps with some East Asian samples. The four samples from Serbia also cluster somewhat apart from other populations. The three groups of mosquitoes most distinguished from the central cluster are samples from 1-the most eastern locations in Europe, 2-Albania, Croatia, Greece, and 3-Brazil (Figure S7). Plotting additional PCs (i.e., PC3 and 4) derived from the DAPC analyses shows similar patterns of clustering but also indicates that Indonesian samples are separate from the rest of the southeast Asian cluster (Figure S7.B; File S14).

PCA and DAPC analyses from all three SNP Sets for the Global dataset were very similar. Scatterplots of additional principal components and other SNP Sets are presented in Files S13-S14 and Figures S6 and S7. Additional plots highlighting PCA results for specific regions are also presented in Figure S8.

Since we were interested in investigating links between mosquitoes sampled in Europe and populations in the native range, we also plotted the ancestry matrices for K=6 and K=7 to visualize how

locations in the European invasive range may group with Asian locations. We chose to examine structures with six and seven clusters because our PCA and DAPC analyses (see [Figures S6, S7, S8](#)) indicated that samples from the most eastern locations in Europe, and possibly also samples from Croatia, Albania, and Greece, may each represent a distinct cluster not defined in the five clusters found in the native range populations in our dataset (56).

When the structure analyses for the Global dataset are plotted for K=6 and K=7, we observe that most Europe clusters are also present in the native range. However, some are represented only by small amounts of admixture ([Figure 2C, Figures S9 and S11B](#)). The three historic Italian locations share a cluster with northern Japanese locations (yellow). This yellow cluster is also observed admixed in western and southern Europe and the Americas, particularly the United States. Greece, Brazil, and Palm Beach, USA show admixture with the primary cluster observed in most south and southeast Asian locations (blue). Taiwan, Okinawa and parts of China (Hunan and Hainan) primarily comprise the orange cluster, which is also observed broadly across modern samples in eastern and southern Europe and a few locations in eastern Europe (e.g. Alushta, Ukraine). However, Hunan and Hainan, China also contain admixture with the green cluster found in samples from Croatia, Albania, and Greece. The red cluster, represented in Indonesia and some samples in Nepal, is present only in Asia and does not appear substantially in Europe or the Americas. The purple cluster is mainly represented in the most eastern European samples and is observed elsewhere only in small proportions of admixture in some eastern Asia locations ([Figures 2C and 3, Figure S9](#)). When a seventh cluster was identified (cyan), it was represented primarily in San Roque, Spain and Aliaga, Turkey ([Figures 2C, S9.B](#)), which corresponds to the dark red cluster identified in the K=20 structure plot ([Figure 2A](#)). [Table S7](#) and [Figure S9](#) present additional details of the structure analyses for different SNP Sets for comparison (also see [Figures S16-S18](#)).

#### *European datasets*

When European locations were analyzed on their own, the composition of clusters was very similar to the clusters found in the global dataset. For the European dataset, our three algorithms indicated a K value between 13 and 15 as the number of ancestral clusters for all but one set of analyses (data from SNP Set 3 run with fastStructure indicated K=18; [Table S7, Figure S10](#)). [Figure 3A](#) shows a typical clustering result for K-14. While most sampling locations showed some admixture, there was again clear evidence of structure among European populations.

One cluster (orange) had a relatively broad distribution, found throughout many locations in western and southern Europe and some parts of eastern Europe west of the Black Sea (e.g., Bulgaria, Romania). The three Italian populations sampled in the 1990s each formed unique clusters. Clusters from these historic Italian samples, especially the one in Desenzano (teal), were also found admixed in several modern populations in Italy (e.g. Sicily) and elsewhere in Europe (e.g. Portugal, Romania, Slovenia). Mosquitoes from Croatia and Albania (bright green) and Greece (dark green) formed two distinct clusters. However, in some algorithms' analyses, Greek samples were grouped with the Croatia-Albania cluster ([Figure S10](#)). Serbia formed a cluster (brown), which was not found in any other locations ([Figure 3A](#)).

The four countries east of the Black Sea, Ukraine, Russia, Georgia, and Armenia (hereafter referred to as the farthest east), were characterized by a mixture of two to three clusters (purple shades) not found in any other part of Europe ([Figure 3A, Figures S10-S11](#)). Turkish samples, including those from eastern Turkey (Hopa), had no admixture with these purple clusters, despite being very geographically close to the four farthest east European countries. Instead, the Turkish samples were primarily grouped with mosquitoes sampled in Spain. The Iberian Peninsula comprised five main clusters, though the Portuguese samples showed admixture with several other clusters (primarily those found in Desenzano). Spanish samples showed comparatively little admixture; individuals in all five Spanish sampling locations



were comprised primarily of a single cluster. The cluster in Barcelona (pink) was shared with Turkish samples and showed admixture with some modern Italian populations. The most southern Spanish location, San Roque, consisted of two distinct sets of individuals, each forming a unique cluster: blue and red. Some individuals shared the blue cluster in Magaluf, Spain, and the red cluster was shared by about half the individuals sampled in Aliaga, Turkey. The final cluster found in Spain (yellow) was detected in three locations: Catarroja, Badajoz, and Magaluf. This yellow cluster was also observed in admixture in the two French locations, some Italian locations, Slovenia, Romania, and the farthest east, particularly Armenia (Figure 3A). For comparison, additional details of the structure analyses for different SNP Sets are presented in Table S7 and Figures S10-S11 (see Figures S16-S18).

The PCA of the European dataset alone allows for visualization of finer scale structure amongst European populations. Specifically, it showed that a subset of Italian samples (i.e., those from the three locations sampled in 1995) form a cluster separate from other Italian or European populations (at the top of PC2 in Figures 3, S12). Plotting PC3 against PC1 indicates that Albania, Croatia, and Greece also cluster together, apart from other countries. Locations in Eastern Europe were spread along PC1, apart from Western and Southern Europe. As in the global dataset, mosquitoes from the four countries farthest to the east (Armenia, Georgia, Russia, and Ukraine) showed no overlap with other regions of Europe, even when additional PCs (i.e., PC2, PC3) were examined (Figures 3, S12). The DAPC showed similar clustering patterns, but also indicates that Serbian samples are somewhat distinct from the rest of Europe (Figure S13).

Overall, PCA and DAPC analyses from all three European SNP Sets were very similar. Scatterplots and pie charts of other SNP Sets are presented in Files S12-S14 and Figures S10-S12. Additional plots highlighting PCA results for specific European regions are also presented in Figure S8.

### *Genetic differentiation*

When we estimated genetic differentiation ( $F_{ST}$ ) for all localities with at least four individuals ( $N=69$  for European and native range;  $N=40$  for Europe), the overall patterns of  $F_{ST}$  were similar for all three SNP sets.  $F_{ST}$  analyses for the other SNP sets are detailed in Files S19-20, and results are shown in Figures S14-S16.

When European and native range countries were examined together, Serbia was the European country with the highest  $F_{ST}$  (mean=0.23, range 0.15-0.44), and the Maldives was the native range country with the highest  $F_{ST}$  (mean=0.25, range 0.13-0.44; Figures S14.C and S15). When European regions were examined on their own, the mean  $F_{ST}$  was similar for Southern (0.10) and Eastern Europe (0.11), and slightly lower for Western Europe (0.08). Pair-wise  $F_{ST}$  among European countries ranged from 0.01 to 0.23, with Serbia again having the highest mean  $F_{ST}$  (0.20, range 0.15-0.23, Table S8; Figure S14.B).

### *Isolation by Distance*

We created a matrix with the  $F_{ST}$  values and geographical distance in kilometers for all European locations and fit a linear regression to our  $F_{ST}$  estimates. The  $F_{ST}$  values were generally higher as the distance between the European sampling sites increased. Still the correlation was only about 4% ( $R^2=0.04$ ), indicating that only a small proportion of variation in  $F_{ST}$  could be attributed to geographic distance (Figure S17.A).

There was no strong evidence for isolation by distance (IBD) in Europe. When we fit a linear regression model for all 40 European locations, the correlation coefficient ( $R^2$ ) was 0, indicating that variation in genetic distance cannot be explained by geographic distance amongst European populations (Figure S17.B). Mantel test observations for the European SNP data were slightly negative (-0.01,  $p$ -value = 0.52), indicating that we cannot reject the null hypothesis that geographic and genetic distance are uncorrelated (Figure 4A). Patterns of IBD were similar for all three SNP sets, and analyses for SNP Sets 1

and 2 are detailed in [Files S19-S20](#). There was also no strong evidence for IDB when looking at regional subsets of European populations ([Figure 4B](#)).

### ***Microsatellites and overlapping SNP dataset in Europe***

We carried out population structure and differentiation analyses for the microsatellite dataset for 637 mosquitoes in 24 European locations. Then, to facilitate comparisons between the two sets of markers (microsatellites and SNPs), we repeated all clustering and genetic differentiation analyses for SNPs using data from the “overlapping” subset of 24 locations sampled with both types of markers ([Figure 1](#), [Tables S2](#) and [S3](#)).

#### ***Microsatellites: population structure and differentiation***

The results of the STRUCTURE analyses for the microsatellite dataset (N=24 locations) indicated that the best-supported number of clusters was K=3, and there was relatively little admixture in most locations ([Figures 5](#), [S20.A](#)). The best-supported number of clusters was also K=3 when both genetic and population information were used to analyze this same subset for the microsatellite dataset ([Figure S20.B](#)). An orange cluster mainly constituted four locations (Saint-Martin-d'Hères, France, Puglia, Italy, and the two Turkish locations), a yellow cluster accounted for most of the ancestry in Portugal, Romania, and all but one of the five Spanish locations (San Roque), and the remaining locations were mainly composed of a third (green) cluster ([Figures 5](#), [S20.A-B](#)).

The PCA of microsatellite data using the subset of 24 overlapping locations showed most samples cluster tightly together, except for Western Europe (France) and some individuals from Italian and Turkish locations, which were spread out along the first PC axis. Some samples from Portugal and one location from Spain also plotted somewhat separately from others ([Figure S22.B](#)). Scatterplots of the first two principal components derived from DAPC for the microsatellite dataset showed a similar pattern ([Figure S23.B](#)).

Regional estimates of average  $F_{ST}$  were slightly higher for Eastern (0.10) than Western or Southern Europe (0.10). Russia had the highest mean pairwise  $F_{ST}$  (0.12), while Slovenia and Malta had the lowest (0.07) ([Figures S24.B](#), [S25.B](#); [Table S9](#)). The correlation between  $F_{ST}$  values and geographic distance was minimal ( $R^2=0.02$ ; [Figure S26.B](#)). When we fit a linear regression model to the genetic and geographic distance, the correlation coefficient ( $R^2$ ) was 0.01 ([Figure S26.B](#); [File S20](#)). There was no evidence for isolation by distance (IBD): the Mantel test observation was slightly negative (-0.009; p-value=0.52), indicating that we cannot reject the null hypothesis that geographic and genetic distance are uncorrelated for this dataset ([Figure S27.B](#)).

#### ***Overlapping SNP dataset: population structure and differentiation***

When the subset of 24 overlapping locations with SNP data was used in clustering analyses using the R package LEA (N=242 mosquitoes), the best-supported number of clusters was K=10 ([Figures S20.C](#), [S21](#)). To facilitate comparisons between datasets, K=3 was also plotted. At K=3, Russia and Georgia formed a distinct cluster (blue), and three Spanish locations (Badajoz, Catarroja, and Magaluf) were composed almost entirely of one cluster (orange). All other locations showed substantial admixture between the orange and yellow clusters ([Figures 5](#) and [S20.D](#)).

Plots of the first two PCs showed that Eastern Europe (Russia and Georgia) and some individuals from Spain and Turkey formed discrete clusters that did not overlap with samples from other locations ([Figure S22.C](#)). Plots of the first two PCs derived from the DAPC analysis indicated that both Eastern

Europe (Russia and Georgia) and the region including Albania, Croatia, Greece, and Serbia, constituted discrete clusters, completely separated from other locations (Figure S23.C).

Regional estimates of average  $F_{ST}$  were somewhat higher for Eastern (0.11) than for Southern (0.08) or Western Europe (0.07). Serbia had the highest mean pairwise  $F_{ST}$  (0.19), followed by Russia and Georgia (0.12). Slovenia and Malta had the lowest mean pairwise  $F_{ST}$  (0.06) (Figures S24.C, S25.C; Table S9). The correlation between genetic and geographic distance was minimal: when  $F_{ST}$  and geographic distance were plotted the  $R^2$  was 0.02, and the correlation coefficient ( $R^2$ ) was 0.01 when a linear regression model fit the genetic and geographic distance (Figure S26.C). There was no evidence for isolation by distance (IBD), as the Mantel test observation was slightly negative (-0.10; p-value=0.75; Figure S27.C).

## Discussion

SNPs are widely recognized as cost-effective tools for characterizing genetic diversity and population structure, and they have recently begun to be recognized as valuable tools for assessing invasion dynamics in a variety of taxa, such as comb jellies (77), fire ants (78), and rodents (79). We used a SNP chip to characterize the genetic structure of *Ae. albopictus* across Europe and the native range to better understand the origin and spread of this invasive mosquito. This study marks the first time a SNP chip has been used to characterize the genetic structure of *Ae. albopictus* outside of its native range. Our findings support the notion of a complex history of multiple invasions from a number of different origins in both the native range (Asia) and previously invaded areas of Europe and the United States (50, 51). In addition, the Aealbo SNP chip revealed fine-scale genetic structure in *Ae. albopictus* populations across Europe, not reported in previous studies using less sensitive genetic markers (i.e. microsatellites, mitochondrial DNA).

### *Genetic structure in the native range*

To assess potential sources of invasions into Europe, we first needed to validate the patterns of population structure found in the native range. Previous analyses of the population structure in the native range that were carried out using data from the Aealbo SNP chip found support for five ancestral clusters in Asia (56). Specifically, one cluster was observed primarily in Japan, with some admixture in China. A second cluster was found in Taiwan and southern Japan (Okinawa), with admixture in parts of China. A third cluster was exclusive to Indonesia and some individuals in Nepal. A fourth cluster was found primarily in Quy Nhon City, Vietnam. The fifth cluster was the most widely distributed and constituted the primary cluster in most of south and southeast Asia, as well as parts of China (particularly Yunnan) (56).

Our analyses, which combined data from the native range with European and American populations, found more than five Asian clusters. Specifically, eight of the 20 clusters in the ancestry matrix best supported by our algorithms (shown in Figure 2A) are represented among native range populations. The primary difference from the previous analysis, which examined Asian populations independently, is that some populations previously represented primarily by a single cluster were split into two or more groups. For example, the northern Japanese cluster was divided into at least two clusters (light yellow and gold), and the single cluster that previously represented Okinawa and Taiwan was split into two clusters (orange-red and dark green). The remaining clustering pattern is very similar to previous findings, in which most of south and southeast Asia cluster together (cyan), and distinct

clusters are observed in Indonesia (along with some individuals in Nepal) and Quy Nhon City, Vietnam (56).

When we plotted the ancestry matrices for K=6 and K=7 for the global dataset, we observed four of the five clusters previously identified in the native range by the Aeab10 SNP chip (56). The main differences were that, in our results, Quy Nhon City, Vietnam clusters with the majority of south and southeast Asia (blue) rather than forming a distinct cluster, and the cluster that primarily constituted Taiwan and Okinawa in the previous analysis showed more admixture (Figures 2C, S9).

### *Origins of invasive populations and genetic structure within Europe*

In evaluating the origins and patterns of population genetic structure, we focus on European regions whose history of invasion and spread were somewhat unresolved based on previous studies and for which we have a substantial geographic representation in our sample set: Italy, Greece and Albania, the region around southern Russia, and the Iberian Peninsula. While our analyses confirmed some of the patterns suggested by earlier studies, they also revealed novel findings.

*Aedes albopictus* was first observed in northern Italy in 1990 (80, 81), and an invasive population in the United States (originating primarily from Japan) was believed to be the source of this introduction (31, 42). However, various studies have found high levels of genetic diversity in Italy, suggesting that multiple, unrelated introductions from different sources likely contributed to the establishment of Italian populations (38, 42, 50, 55). Our analyses support the latter scenario. The three locations from northern Italy sampled in 1995 (Brescia, Cesena, and Desenzano) shared a cluster with samples from northern Japan and the United States (yellow cluster in Figure 2B). While this cluster was also observed admixed in samples from western and southern Europe, the three Italian locations with historical samples were differentiated from one another (Figure 3A) and from modern Italian samples (Figures S8A, S13). In contrast to samples collected in the 1990s, modern Italian locations exhibited high levels of admixture. Still, they shared a cluster (orange), which was observed broadly across western and southern Europe, a few locations in eastern Europe (e.g. Alushta, Ukraine) (Figure 2A), China (Hunan and Hainan), and also as admixture in Taiwan, Okinawa, more northern areas of Japan, and the Americas (Figures 2B, S8A).

In summary, modern Italian samples showed intermixing of ancestry derived from historic Italian locations, other parts of Europe, and East Asia (Figure 2), suggesting that *Ae. albopictus* became established throughout Italy via a combination of spreading from the initial introductions in the north in the early 1990s and multiple new introductions from both native and other invasive regions. Furthermore, it appears that that populations from northern Italy have played an important role in establishing invasive populations in other European countries: mosquitoes throughout Europe (e.g. Portugal, Romania, Slovenia) shared ancestry with individuals sampled in Desenzano in 1995. This is consistent with indications that Italy has been a source, and even an essential stepping-stone, for *Ae. albopictus* range expansion into more northern locations, such as Switzerland (51, 82).

Mosquitoes in Greece and Albania (and some other populations in the South Balkans, including south Serbia) have repeatedly been observed to be genetically distinct from other parts of Europe (42, 49, 51, 53). *Aedes albopictus* was first identified in Albania in 1979 (5), but was not observed in Greece until 2003 (83). Previous analyses using microsatellites suggested different sources for invasions in these countries, specifically that populations in Albania originated in China and populations in Greece originated in Thailand (42). However, our SNP analyses suggested that samples from Greece and Albania are genetically similar (49, 51), and these populations may have been introduced from the United States (51).

Our data confirm that Albania and Greece are genetically distinct from most other European populations (Figures S7, S13). When we analyzed population structure in our Global dataset, Greece and

Albania were grouped into a cluster that also included Croatia (and Serbia, at  $K=7$ ). Greece also showed admixture with the largest cluster in the native range, which included most of South and Southeast Asia and parts of China (cyan cluster, [Figure 2A](#)), particularly Hunan and Hainan ([Figures 2B, S8B](#)). When we examined population structure using only European locations, Serbia constituted its own unique cluster (brown), and Greece was also differentiated from Albania and Croatia, though admixture with the Albanian-Croatian cluster was observed ([Figure 3A](#)). This suggests that Albania may have contributed, at least in part, to establishing the invasive population in Greece, but additional introductions directly from Asia or via an invasive population in the Americas have also occurred ([Figures 2B, S8B](#)).

Our analyses of populations from Eastern Europe suggest an invasion history somewhat different from that identified by previous studies. Studies based on diversity in mitochondrial genomes identified two spatially isolated genetic lineages of *Ae. albopictus* in southern Russia: one linked to Mediterranean populations and another similar to those from the USA and Japan (37). While some locations in the farthest east of Europe, namely Armenia and Ukraine, showed admixture with clusters also found in the Mediterranean area, most locations in our analyses were comprised of an admixture of two to three genetic clusters not observed anywhere else in Europe or Asia (purple clusters, [Figure 2](#)). These farthest eastern European clusters were distinct in multivariate analyses as well, showing no overlap with other samples in PCA or DAPC plots ([Figure S8C](#)), and remained differentiated from all other regions even when the number of clusters in our structure analyses was reduced to six ([Figure S9](#)). This indicates that mosquitoes in this region were introduced from somewhere in the native range not covered by our sample collection, and have not had substantial introductions from other areas of Europe. Though the four farthest east countries are separated from the rest of Europe by the Black Sea, it is unclear why most sampling sites there show so little admixture with any other localities. Samples from Turkey (Hopa) likewise showed no admixture with these purple clusters, despite being very geographically close to Georgia and southern Russia ([Figures 1, 3](#)). Another study that utilized mitochondrial genomes found that southern Russia was characterized by strong differentiation and highly restricted gene flow between local populations, and suggested either a single introduction or multiple transfers from the same source population (84). Our SNP data did not confirm this, as locations in southern Russia were genetically similar to one another and showed evidence of extensive admixture among the same three purple clusters throughout the region ([Figure S11](#)).

Across the Iberian Peninsula, previous work using mtDNA and rDNA generally identified little genetic structure, likely stemming from human-mediated gene flow and multiple, on-going introductions (36, 48). However, other analyses found that northern Spain and the Balearic Islands were somewhat differentiated from southern Spain and the rest of Europe (46, 49). The initial introduction of *Ae. albopictus* to Spain was first detected in 2004 near Barcelona (85), and the species is believed to have spread along the Spanish Mediterranean coast and inland from there (86). While previous findings suggest that Barcelona was the largest source of inter-province transfers of mosquitoes in Spain (87), our analyses do not support this scenario. We found that mosquitoes in Barcelona formed a distinct cluster (pink) that was not shared with other locations in Spain. However, this cluster was also observed in Turkey and admixed throughout other southern European countries (e.g. Italy and Slovenia; [Figure 3A](#)). On the other hand, some mosquitoes from western Spain (Badajoz), eastern Spain (Catarroja), and the Balearic islands (Magaluf) all shared ancestry (yellow in [Figures 2A, 3A](#)), indicating a common source from either Japan or the United States for mosquitoes in these geographically separated areas. While Portuguese samples showed some admixture with this cluster, they were more genetically similar to samples from France and other southern European countries. Interestingly, there was less admixture within Spanish samples compared to other European locations, though two locations contained individuals from two separate clusters: one cluster unique to Spain (blue) and the other shared with some individuals from Aliaga, Turkey (red, [Figure 3A](#)).

Overall, our findings are consistent with the hypothesis that older established invasive populations (Albania, Italy) seem to be derived from populations in the native range (e.g. Japan) or other early-established invasive populations (United States), while more recently established populations share ancestry with multiple sources, including populations already existing in Europe.

### *Factors influencing population differentiation and local adaptation*

We did not find a significant correlation between genetic and geographical distances among European samples. This lack of correlation is consistent with the results of most previous studies (42, 51, 56), and suggests that factors other than distance, such as human-facilitated transportation and ecological factors, are more important in influencing the genetic structure of *Ae. albopictus* in Europe. Studies have shown that introduced *Ae. albopictus* mosquitoes spread through a combination of natural and human-aided dispersal, and the degree of dispersal may be effected by the landscape (88). Population bottlenecks during introduction can also substantially influence the genetic variability (88). These factors could help to explain why we see different levels of genetic structuring among some European regions more than others. For example, we found that most central European populations showed high degrees of admixture, but the most eastern European populations (i.e. Armenia, Georgia, Russia, Ukraine) were largely distinct from other areas of Europe. This suggests that after *Ae. albopictus* was introduced into the area in 2011 (37) from a unique source not represented elsewhere in Europe, and the Black Sea acted as a geographic barrier to limit gene flow with other areas. However, at least one population east of the Black Sea, Hopa (Turkey), did not fit this pattern. Instead, it shares ancestry with more Western European populations (e.g., Spain), despite being geographically close to the Georgian and Armenian populations. This underscores the somewhat “chaotic” nature of dispersion identified in previous studies (42) and highlights the importance of human-mediated introductions and transport, in addition to geographic barriers, in shaping patterns of genetic structure. Our sample size per population was relatively small, with typically 12 or fewer individuals from each location genotyped. Future work should include larger numbers of individuals from each population to carry out migration analysis that looks for first or second-generation migrants. This would improve our understanding of current gene flow and how human-mediated transportation of mosquitoes may be impacting the spread and population structure of *Ae. albopictus* throughout their invasive ranges.

While we could not explicitly test for local adaptations using our dataset, there is evidence for thermal adaptation across the wide temperature range occupied by *Ae. albopictus* in its native range (89), which may have facilitated its spread to temperate regions worldwide (90). Our findings are consistent with this idea: the most northern, temperate locations in the native range, Japan and China, were the dominant contributors to populations throughout most of Europe (Figures 2B, S11). Previous analyses of *Ae. albopictus* in southern Russia found patterns of selective sweep near genomic regions associated with neural protection, lipid conservation, and cuticle formation during diapause, which could indicate recent selection for cold adaptation within the invasive range (84). Additional systematic sampling, focusing on a range of temperatures and other climatic variables across the European continent, will be essential for determining the extent and specifics of local adaptation to cold and other ecological factors that mosquitoes have undergone as they spread to new, more temperate areas. The availability of a reliable microarray, the Aealbo SNP chip (56), can facilitate the expansion of sampling locations and allow researchers to merge existing data sets with new ones to develop a more thorough and nuanced understanding of *Ae. albopictus* range expansion and local adaptation.

### *Comparisons of microsatellite and SNP datasets*

Our SNP markers detected more fine-scale population structure than the microsatellites, even when the exact overlapping locations were examined with both sets of markers. Specifically, microsatellite data suggested that only three distinct genetic clusters were present in Europe, while SNP data indicated that the best-supported number of clusters was ten (Figure 5, Figure S20). This suggests that the microsatellite markers were unable to detect genetic structure at the same resolution as the SNPs. Additionally, the three clusters that the microsatellites identified differed from those found when we plotted  $K=3$  for the SNP data. For example, microsatellites data failed to identify the farthest eastern European locations (Russia and Georgia) as a unique cluster, which was a consistent finding from all the SNP dataset (Figures 2-3 & 5; Figures S9-S11 & S20-S21).

Multivariate analyses (PCA and DAPC) also showed less discrimination among populations when using microsatellite compared to SNP data. For the microsatellite dataset, most samples from all three European regions clustered tightly together, with only France and some individuals from Italy and Türkiye (PCA), and Portugal (DAPC) being differentiated from the rest of Europe. In contrast, in the SNP dataset, the farthest eastern Europe (Russia and Georgia) and some individuals from Spain and Turkey (PCA) as well as Albania, Croatia, Greece, and Serbia (DAPC) formed discrete clusters that did not overlap with samples from other locations (Figures S22-S23).

In both the microsatellites and SNPs datasets, the correlation between genetic and geographic distance was very small (Figure S26). Similarly, there was no substantial evidence for isolation by distance (IBD) among the overlapping locations for either the microsatellite or SNP datasets (Mantel test observations were slightly negative for both; Figure S27). Overall, these results indicate that using genome-wide markers, such as SNPs, provides not only finer resolution, but also identifies genetic structuring undescribed by microsatellite loci.

## Conclusion

Our results are consistent with previous studies that provide evidence for multiple independent invasions to Europe from both the native region and other areas of the invasive range. Our analyses also confirmed admixture among mosquitoes from different genetic ancestries in many European locations. This highlights the complexity of their spread and the need for comprehensive genetic tools to understand it fully. A more extensive set of informative genetic markers across the entire genome was found to be essential to resolve outstanding questions about population structure for invasive populations of *Ae. albopictus* in Europe that have only recently been established and may still be experiencing on-going introductions (30). Our study underscores the importance of using appropriately sensitive genetic markers when working with invasive populations, and highlights the potential for using SNP chip data to understand and manage other invasive species. The findings from this study can serve as a baseline for association studies that examine the genetic basis of key traits such as diapause, cold tolerance, arbovirus transmission, insecticide resistance, and other complex traits. Our work can also assist in predicting this species' establishment and spread into other invasive locations, with implications for understanding local adaptation, the dynamics of disease transmission, and strategies for controlling this invasive disease vector throughout Europe and the rest of its expanding range. By providing a deeper understanding of *Ae. albopictus* genetic diversity and movement patterns our research contributes to efforts to mitigate the public health impacts of this globally significant mosquito species.

## References

1. M. U. G. Kraemer, *et al.*, Past and future spread of the arbovirus vectors *Aedes aegypti* and *Aedes albopictus*. *Nature Microbiology* **4**, 854–863 (2019).
2. M. U. Kraemer, *et al.*, The global distribution of the arbovirus vectors *Aedes aegypti* and *Ae. albopictus*. *eLife* **4**, e08347 (2015).
3. P. R. D. Sprenger, The used tire trade: a mechanism for the worldwide dispersal of container breeding mosquitoes. *Journal of American Mosquito Control Association* **3**, 494 (1987).
4. T. Swan, *et al.*, A literature review of dispersal pathways of *Aedes albopictus* across different spatial scales: implications for vector surveillance. *Parasites & Vectors* **15**, 303 (2022).
5. J. Adhami, N. Murati, The presence of the mosquito *Aedes albopictus* in Albania. *Revista Mjekësore* 13–16 (1987).
6. European Centre for Disease Prevention and Control, *Aedes albopictus* - current known distribution: February 2023. (2023). Available at: <https://www.ecdc.europa.eu/en/publications-data/aedes-albopictus-current-known-distribution-february-2023> [Accessed 17 January 2024].
7. C. Paupy, H. Delatte, L. Bagny, V. Corbel, D. Fontenille, *Aedes albopictus*, an arbovirus vector: From the darkness to the light. *Microbes and Infection* **11**, 1177–1185 (2009).
8. N. G. Gratz, Critical review of the vector status of *Aedes albopictus*. *Medical and Veterinary Entomology* **18**, 215–227 (2004).
9. G. Rezza, *et al.*, Infection with chikungunya virus in Italy: an outbreak in a temperate region. *The Lancet* **370**, 1840–1846 (2007).
10. D. Tomasello, P. Schlagenhauf, Chikungunya and dengue autochthonous cases in Europe, 2007–2012. *Travel Medicine and Infectious Disease* **11**, 274–284 (2013).
11. S. Giron, *et al.*, Vector-borne transmission of Zika virus in Europe, southern France, August 2019. *Eurosurveillance* **24**, 1900655 (2019).
12. F. Schaffner, J. M. Medlock, W. V. Bortel, Public health significance of invasive mosquitoes in Europe. *Clinical Microbiology and Infection* **19**, 685–692 (2013).
13. T. Uller, R. Leimu, Founder events predict changes in genetic diversity during human-mediated range expansions. *Global Change Biology* **17**, 3478–3485 (2011).
14. A. Estoup, *et al.*, Is There a Genetic Paradox of Biological Invasion? *Annual Review of Ecology, Evolution, and Systematics* **47**, 51–72 (2016).
15. K. A. Medley, Niche shifts during the global invasion of the Asian tiger mosquito, *Aedes albopictus* Skuse (Culicidae), revealed by reciprocal distribution models. *Global Ecology and Biogeography* **19**, 122–133 (2010).



16. C. Goubert, *et al.*, High-throughput sequencing of transposable element insertions suggests adaptive evolution of the invasive Asian tiger mosquito towards temperate environments. *Molecular Ecology* **26**, 3968–3981 (2017).
17. S. Sherpa, *et al.*, Predicting the success of an invader: Niche shift versus niche conservatism. *Ecology and Evolution* **9**, 12658–12675 (2019).
18. R. Bueno-Marí, R. Jiménez-Peydró, First observations of homodynamic populations of *Aedes albopictus* (Skuse) in Southwest Europe. *Journal of vector borne diseases* **52**, 175 (2015).
19. I. Del Lesto, *et al.*, Is Asian tiger mosquito (*Aedes albopictus*) going to become homodynamic in Southern Europe in the next decades due to climate change? *Royal Society Open Science* **9**, 220967 (2022).
20. F. Collantes Alcaraz, J. A. Delgado, P. M. Alarcón Elbal, S. Delacour Estrella, J. Lucientes Curdi, First confirmed outdoor winter reproductive activity of Asian tiger mosquito (*Aedes albopictus*) in Europe. *Anales de Biología*, **36**, 2014 (2014).
21. A. Vega-Rúa, *et al.*, Vector competence of *Aedes albopictus* populations for chikungunya virus is shaped by their demographic history. *Communications Biology* **3**, 1–13 (2020).
22. O. J. Brady, *et al.*, Global temperature constraints on *Aedes aegypti* and *Ae. albopictus* persistence and competence for dengue virus transmission. *Parasites & Vectors* **7**, 1–17 (2014).
23. X. de Lamballerie, *et al.*, Chikungunya virus adapts to tiger mosquito via evolutionary convergence: a sign of things to come? *Virology Journal* **5**, 1–4 (2008).
24. V. Pichler, *et al.*, Phenotypic and genotypic pyrethroid resistance of *Aedes albopictus*, with focus on the 2017 chikungunya outbreak in Italy. *Pest Management Science* **75**, 2642–2651 (2019).
25. I. Unlu, A. Farajollahi, D. Strickman, D. M. Fonseca, Crouching Tiger, Hidden Trouble: Urban Sources of *Aedes albopictus* (Diptera: Culicidae) Refractory to Source-Reduction. *PLOS ONE* **8**, e77999 (2013).
26. F. Baldacchino, *et al.*, Control methods against invasive *Aedes* mosquitoes in Europe: a review. *Pest Management Science* **71**, 1471–1485 (2015).
27. B. Caputo, *et al.*, Assessment of the Effectiveness of a Seasonal-Long Insecticide-Based Control Strategy against *Aedes albopictus* Nuisance in an Urban Area. *PLOS Neglected Tropical Diseases* **10**, e0004463 (2016).
28. A. Farajollahi, S. P. Healy, I. Unlu, R. Gaugler, D. M. Fonseca, Effectiveness of Ultra-Low Volume Nighttime Applications of an Adulticide against Diurnal *Aedes albopictus*, a Critical Vector of Dengue and Chikungunya Viruses. *PLOS ONE* **7**, e49181 (2012).
29. S. Sherpa, D. Rioux, C. Pougnet-Lagarde, L. Després, Genetic diversity and distribution differ between long-established and recently introduced populations in the invasive mosquito *Aedes albopictus*. *Infection, Genetics and Evolution* **58**, 145–156 (2018).
30. C. Goubert, G. Minard, C. Vieira, M. Boulesteix, Population genetics of the Asian tiger mosquito *Aedes albopictus*, an invasive vector of human diseases. *Heredity* **117**, 125–134 (2016).

31. S. Urbanelli, R. Bellini, M. Carrieri, P. Sallicandro, G. Celli, Population structure of *Aedes albopictus* (Skuse): the mosquito which is colonizing Mediterranean countries. *Heredity* **84**, 331–337 (2000).
32. V. Battaglia, *et al.*, The Worldwide Spread of the Tiger Mosquito as Revealed by Mitogenome Haplogroup Diversity. *Frontiers in Genetics* **7** (2016).
33. G. Minard, *et al.*, French invasive Asian tiger mosquito populations harbor reduced bacterial microbiota and genetic diversity compared to Vietnamese autochthonous relatives. *Frontiers in Microbiology* **6** (2015).
34. S. Usmani-Brown, L. Cohnstaedt, L. E. Munstermann, Population Genetics of *Aedes albopictus* (Diptera: Culicidae) Invading Populations, Using Mitochondrial nicotinamide Adenine Dinucleotide Dehydrogenase Subunit 5 Sequences. *Annals of the Entomological Society of America* **102**, 144–150 (2009).
35. L. Mousson, *et al.*, Phylogeography of *Aedes* (*Stegomyia*) *aegypti* (L.) and *Aedes* (*Stegomyia*) *albopictus* (Skuse) (Diptera: Culicidae) based on mitochondrial DNA variations. *Genetics Research* **86**, 1–11 (2005).
36. L. Zé-Zé, *et al.*, Mitogenome diversity of *Aedes* (*Stegomyia*) *albopictus*: Detection of multiple introduction events in Portugal. *PLOS Neglected Tropical Diseases* **14**, e0008657 (2020).
37. E. Shaikevich, L. Karan, M. Fedorova, Genetic Structure of *Aedes* (*Stegomyia*) *albopictus* Populations in Russia. *Journal of Arthropod-Borne Diseases* **17**, 51–62 (2023).
38. E. Shaikevich, A. Talbalaghi, Molecular characterization of the Asian tiger mosquito *Aedes albopictus* (Skuse)(Diptera: Culicidae) in northern Italy. *International Scholarly Research Notices* **2013** (2013).
39. E. V. Shaikevich, *et al.*, Invasive mosquito species *Aedes albopictus* and *Aedes aegypti* on the Black Sea coast of the Caucasus: genetics (COI, ITS2), Wolbachia and *Dirofilaria* infections. *Vavilov Journal of Genetics and Breeding* **22**, 574–585 (2018).
40. D. Walther, D. E. Scheuch, H. Kampen, The invasive Asian tiger mosquito *Aedes albopictus* (Diptera: Culicidae) in Germany: Local reproduction and overwintering. *Acta Tropica* **166**, 186–192 (2017).
41. B. L. Apostol, W. C. Black, P. Reiter, B. R. Miller, Population genetics with RAPD-PCR markers: the breeding structure of *Aedes aegypti* in Puerto Rico. *Heredity* **76**, 325–334 (1996).
42. M. Manni, *et al.*, Genetic evidence for a worldwide chaotic dispersion pattern of the arbovirus vector, *Aedes albopictus*. *PLOS Neglected Tropical Diseases* **11**, e0005332 (2017).
43. R. Lühken, *et al.*, Microsatellite typing of *Aedes albopictus* (Diptera: Culicidae) populations from Germany suggests regular introductions. *Infection, Genetics and Evolution* **81**, 104237 (2020).
44. A. Tancredi, *et al.*, Tracing temporal and geographic distribution of resistance to pyrethroids in the arboviral vector *Aedes albopictus*. *PLOS Neglected Tropical Diseases* **14**, e0008350 (2020).

45. M. Manni, *et al.*, Molecular markers for analyses of intraspecific genetic diversity in the Asian Tiger mosquito, *Aedes albopictus*. *Parasites & Vectors* **8**, 188 (2015).
46. P. Artigas, *et al.*, *Aedes albopictus* diversity and relationships in south-western Europe and Brazil by rDNA/mtDNA and phenotypic analyses: ITS-2, a useful marker for spread studies. *Parasites & Vectors* **14**, 333 (2021).
47. N. Gojković, J. Ludoški, B. Krtinić, V. Milankov, The First Molecular and Phenotypic Characterization of the Invasive Population of *Aedes albopictus* (Diptera: Culicidae) from the Central Balkans. *Journal of Medical Entomology* **56**, 1433–1440 (2019).
48. F. Lucati, *et al.*, Multiple invasions, *Wolbachia* and human-aided transport drive the genetic variability of *Aedes albopictus* in the Iberian Peninsula. *Scientific Reports* **12**, 20682 (2022).
49. S. Sherpa, *et al.*, Unravelling the invasion history of the Asian tiger mosquito in Europe. *Molecular Ecology* **28**, 2360–2377 (2019).
50. P. Kotsakiozi, *et al.*, Population genomics of the Asian tiger mosquito, *Aedes albopictus*: insights into the recent worldwide invasion. *Ecology and Evolution* **7**, 10143–10157 (2017).
51. L. Vavassori, A.-C. Honnen, N. Saarman, A. Caccone, P. Müller, Multiple introductions and overwintering shape the progressive invasion of *Aedes albopictus* beyond the Alps. *Ecology and Evolution* **12**, e9138 (2022).
52. Z. Ruiling, L. Tongkai, M. Dezheng, Z. Zhong, Genetic characters of the globally spread tiger mosquito, *Aedes albopictus* (Diptera, Culicidae): implications from mitochondrial gene COI. *Journal of Vector Ecology* **43**, 89–97 (2018).
53. A. Mercier, *et al.*, Impact of temperature on dengue and chikungunya transmission by the mosquito *Aedes albopictus*. *Scientific Reports* **12**, 6973 (2022).
54. V. Pichler, *et al.*, Geographic distribution of the V1016G knockdown resistance mutation in *Aedes albopictus*: a warning bell for Europe. *Parasites & Vectors* **15**, 280 (2022).
55. V. Pichler, *et al.*, Complex interplay of evolutionary forces shaping population genomic structure of invasive *Aedes albopictus* in southern Europe. *PLOS Neglected Tropical Diseases* **13**, e0007554 (2019).
56. L. V. Cosme, *et al.*, A genotyping array for the globally invasive vector mosquito, *Aedes albopictus*. *Parasites Vectors* **17**, 106 (2024).
57. D. Hulce, X. Li, T. Snyder-Leiby, C. J. Liu, GeneMarker<sup>®</sup> genotyping software: Tools to increase the statistical power of DNA fragment analysis. *Journal of Biomolecular Techniques* **22**, S35 (2011).
58. J. Wang, A. Caballero, W. G. Hill, The effect of linkage disequilibrium and deviation from Hardy–Weinberg proportions on the changes in genetic variance with bottlenecking. *Heredity* **81**, 174–186 (1998).
59. I. Patil, Visualizations with statistical details: The “ggstatsplot” approach. *Journal of Open Source Software* **6**, 3167 (2021).

60. D. H. Alexander, K. Lange, Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC bioinformatics* **12**, 1–6 (2011).
61. A. Raj, M. Stephens, J. K. Pritchard, fastSTRUCTURE: Variational Inference of Population Structure in Large SNP Data Sets. *Genetics* **197**, 573–589 (2014).
62. E. Frichot, O. François, LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution* **6**, 925–929 (2015).
63. T. Jombart, adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403–1405 (2008).
64. T. Jombart, S. Devillard, F. Balloux, Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC genetics* **11**, 94 (2010).
65. L. W. Pembleton, N. O. Cogan, J. W. Forster, StAMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources* **13**, 946–952 (2013).
66. B. S. Weir, C. C. Cockerham, Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370 (1984).
67. S. Dray, A.-B. Dufour, The ade4 Package: Implementing the Duality Diagram for Ecologists. *Journal of Statistical Software* **22**, 1–20 (2007).
68. J. K. Pritchard, M. Stephens, P. Donnelly, Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).
69. D. Falush, M. Stephens, J. K. Pritchard, Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**, 1567–1587 (2003).
70. D. Falush, M. Stephens, J. K. Pritchard, Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* **7**, 574–578 (2007).
71. M. J. Hubisz, D. Falush, M. Stephens, J. K. Pritchard, Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources* **9**, 1322–1332 (2009).
72. G. Evanno, S. Regnaut, J. Goudet, Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* **14**, 2611–2620 (2005).
73. N. M. Kopelman, J. Mayzel, M. Jakobsson, N. A. Rosenberg, I. Mayrose, Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources* **15**, 1179–1191 (2015).
74. J. Goudet, Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes* **5**, 184–186 (2005).
75. J. Oksanen, *et al.*, vegan: Community Ecology Package. (2022). Deposited 2022.
76. P. Legendre, L. Legendre, *Numerical Ecology* (Elsevier, 2012).

77. J. M. Pujolar, M. T. Limborg, M. Ehrlich, C. Jaspers, High throughput SNP chip as cost effective new monitoring tool for assessing invasion dynamics in the comb jelly *Mnemiopsis leidyi*. *Frontiers in Marine Science* **9** (2022).
78. P. Lenancker, *et al.*, Genome-wide SNPs reveal the social structure and invasion pathways of the invasive tropical fire ant (*Solenopsis geminata*). [Preprint] (2022). Available at: <https://www.biorxiv.org/content/10.1101/2022.07.20.500883v1> [Accessed 24 July 2024].
79. K. P. Oh, N. V. de Weyer, W. A. Ruscoe, S. Henry, P. R. Brown, From chip to SNP: Rapid development and evaluation of a targeted capture genotyping-by-sequencing approach to support research and management of a plaguing rodent. *PLOS ONE* **18**, e0288701 (2023).
80. G. Dalla Pozza, G. Majori, First record of *Aedes albopictus* establishment in Italy. *J Am Mosq Control Assoc* **8**, 318–320 (1992).
81. A. Sabatini, V. Raineri, G. Trovato, M. Coluzzi, *Aedes albopictus* in Italy and possible diffusion of the species into the Mediterranean area. *Parassitologia* **32**, 301–304 (1990).
82. M. N. Wymann, E. Flacio, S. Radczuweit, N. Patocchi, P. Lüthy, Asian tiger mosquito (*Aedes albopictus*) - a threat for Switzerland? *Eurosurveillance* **13**, 3–4 (2008).
83. A. Samanidou-Voyadjoglou, E. Patsoula, G. Spanakos, N. C. Vakalis, Confirmation of *Aedes albopictus* (Skuse) (Diptera: Culicidae) in Greece. *European Mosquito Bulletin* 10–12 (2005).
84. E. A. Konorov, V. Yurchenko, I. Patraman, A. Lukashev, N. Oyun, The effects of genetic drift and genomic selection on differentiation and local adaptation of the introduced populations of *Aedes albopictus* in southern Russia. *PeerJ* **9**, e11776 (2021).
85. C. Aranda, R. Eritja, D. Roiz, First record and establishment of the mosquito *Aedes albopictus* in Spain. *Medical and Veterinary Entomology* **20**, 150–152 (2006).
86. F. Collantes, *et al.*, Review of ten-years presence of *Aedes albopictus* in Spain 2004–2014: known distribution and public health concerns. *Parasites & vectors* **8**, 1–11 (2015).
87. R. Eritja, J. R. B. Palmer, D. Roiz, I. Sanpera-Calbet, F. Bartumeus, Direct Evidence of Adult *Aedes albopictus* Dispersal by Car. *Scientific Reports* **7**, 14399 (2017).
88. S. Sherpa, *et al.*, Landscape does matter: Disentangling founder effects from natural and human-aided post-introduction dispersal during an ongoing biological invasion. *Journal of Animal Ecology* **89**, 2027–2042 (2020).
89. S. Sherpa, *et al.*, Genomic Shifts, Phenotypic Clines, and Fitness Costs Associated With Cold Tolerance in the Asian Tiger Mosquito. *Molecular Biology and Evolution* **39**, msac104 (2022).
90. S. Sherpa, M. G. B. Blum, L. Després, Cold adaptation in the Asian tiger mosquito's native range precedes its invasion success in temperate regions. *Evolution* **73**, 1793–1808 (2019).

## **Acknowledgements**

We thank everyone who contributed to this research, including students at Yale University and the laboratory staff at the Yale DNA Analysis Facility at Science Hill and the Functional Genomics Core at the University of North Carolina at Chapel Hill for assistance in processing samples. We would also like to thank Lusine Paronyan for providing additional mosquito samples. This work was generously supported by a grant from the National Institutes of Health (NIH-DHHS-NIAID grant number RO1A13409), “Population genomics of a globally distributed arbovirus vector, *Aedes albopictus*”, and some authors received financial support from a regional Spanish research project (Grant Number IB16135) and as part of the Short-Term Scientific Missions (STSM) within the AIM-COST CA17108 Action to DBB.

## **Data Accessibility Statement**

This research involved international partnerships with scientists across many countries. All collaborators providing genetic samples were invited to be co-authors, and the data and results have been made available in public databases.

The codes describing the step-by-step of all analyses are available at [https://rpubs.com/margaret\\_corley](https://rpubs.com/margaret_corley), and an overview of the analyses contained in each RMarkdown file is provided in [Table S1](#). The raw data files required to reproduce the analyses are available in Zenodo (native range: <https://zenodo.org/records/10048029>, Europe & Americas: <https://doi.org/10.5281/zenodo.13760661>).

## **Ethics approval**

This research complies with all applicable laws on sampling from natural populations. All sample handling and shipping were in accordance with Yale Environmental and Health Safety protocols.

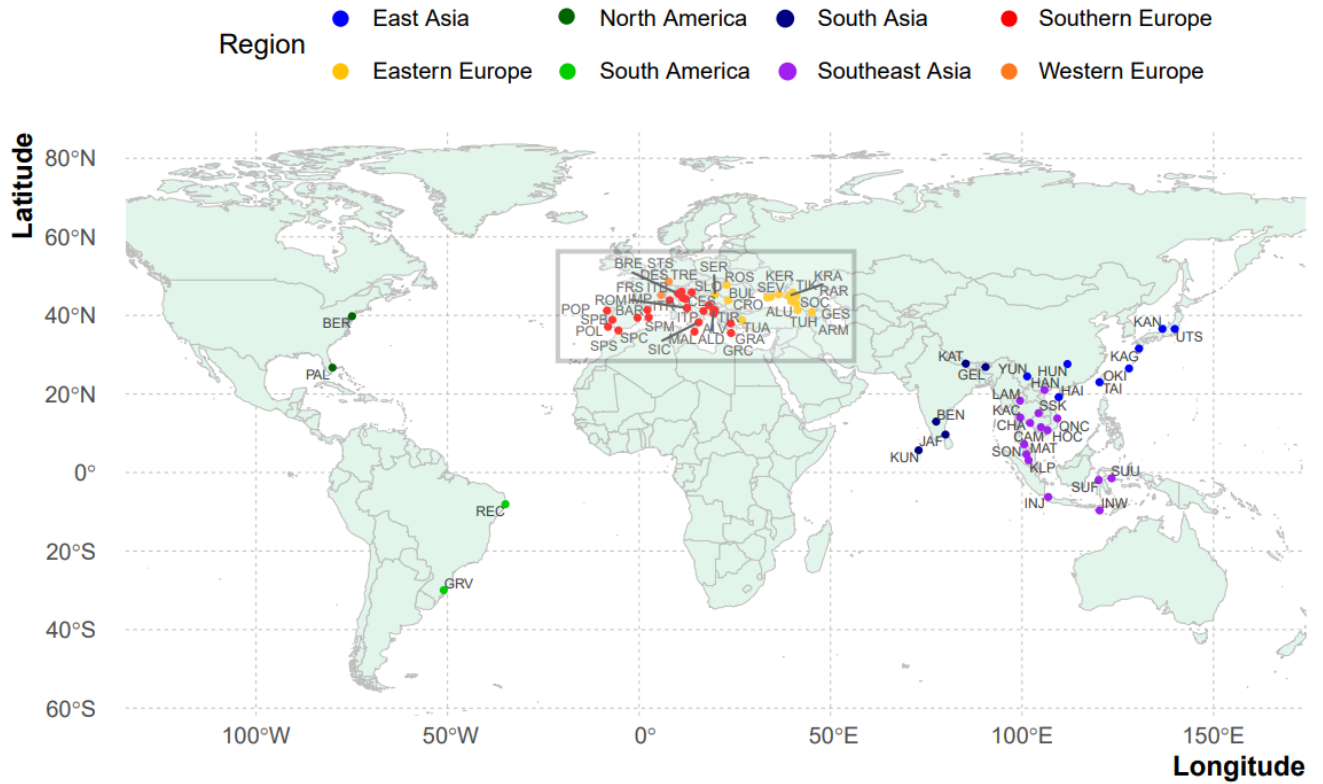
## **Author Contributions**

Data methods and analyses: MC and LVC. SNP sample preparation/investigation: MC, KH, TW, and TJ. Microsatellite sample investigation: VV. Design/Conceptualization: AC, MC, LC, JP. Project administration: AC. Wrote the paper: MC. Sample contributions: all remaining authors. All authors read and approved the manuscript prior to submission.

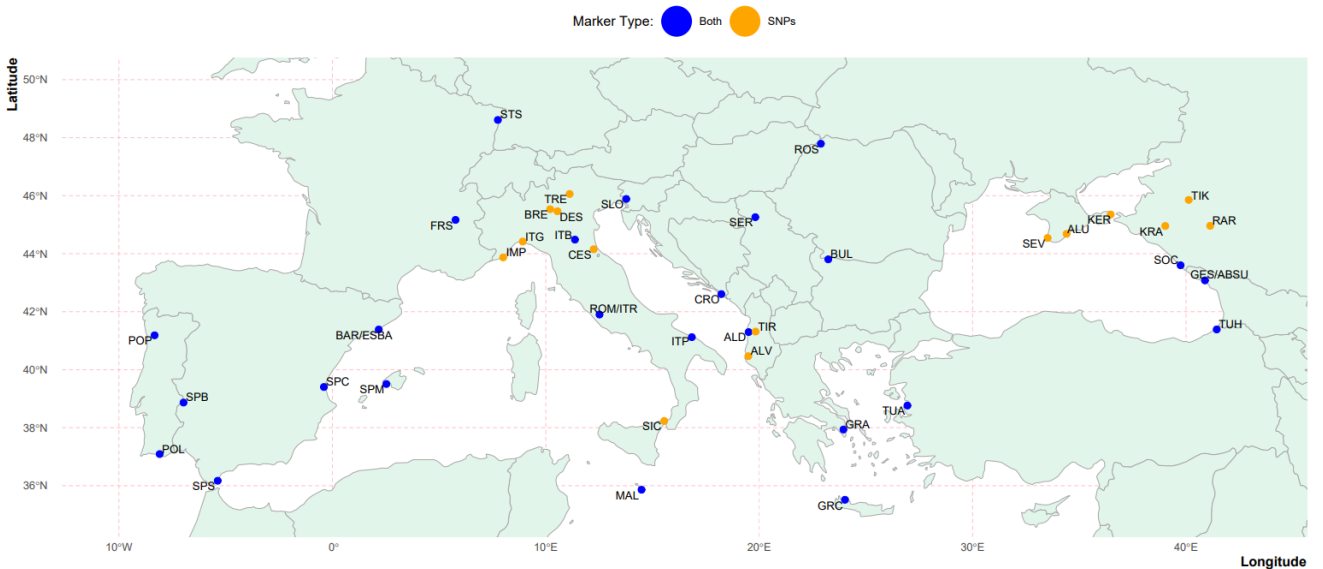
## **Conflict of Interest Statement**

The authors declare no competing interests.

A)

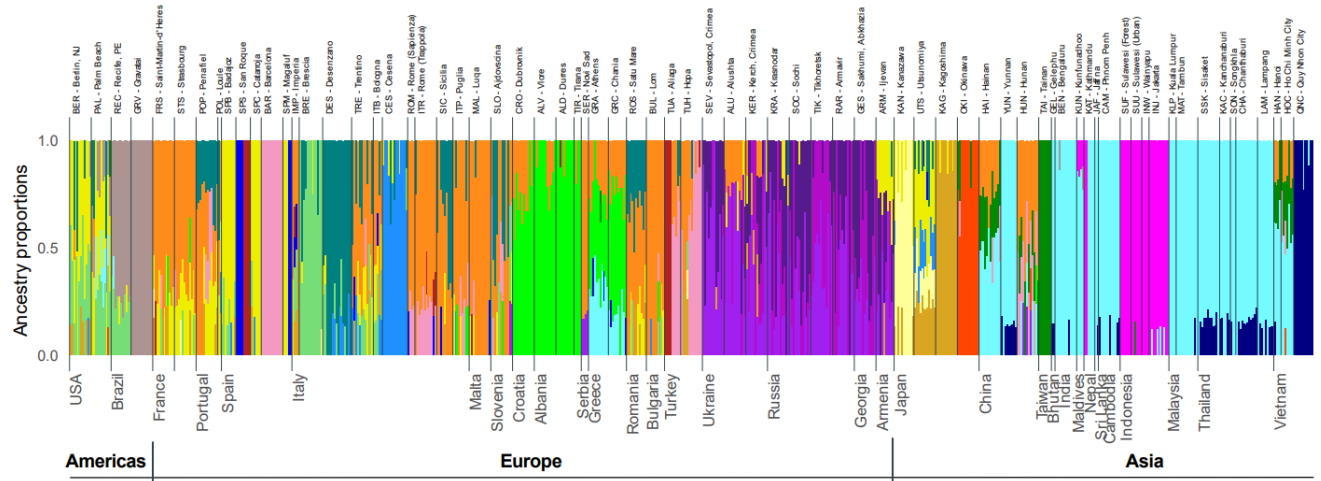


B)



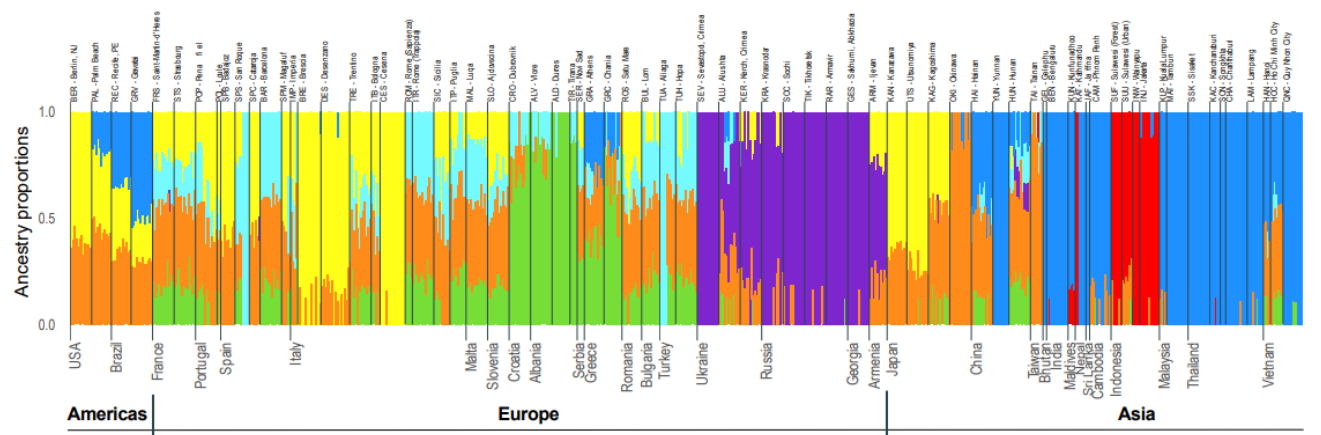
**Figure 1.** Maps of sampling locations for all genotyped *Ae. albopictus* mosquitoes. Panel A shows the 74 locations, by region, for all samples genotyped in the Global SNP Chip dataset. Panel B shows a zoomed in view of the box outlined in gray in Panel A, to highlight the locations for the 41 European sampling sites in the SNP dataset (orange) and 25 in the microsatellite dataset. Locations with samples in both the SNP and microsatellite data sets are indicated in blue. Each locality is identified by a unique three-letter abbreviation identifying the location. Details on the location, year of collection, and the number of samples per locality for both sets of samples are reported in [Tables S2](#) and [S3](#).

2A)



Each bar represents the ancestry proportions for an individual for  $k=20$ .  
LEA inference for 22,642 SNPs (SNP Set 3).

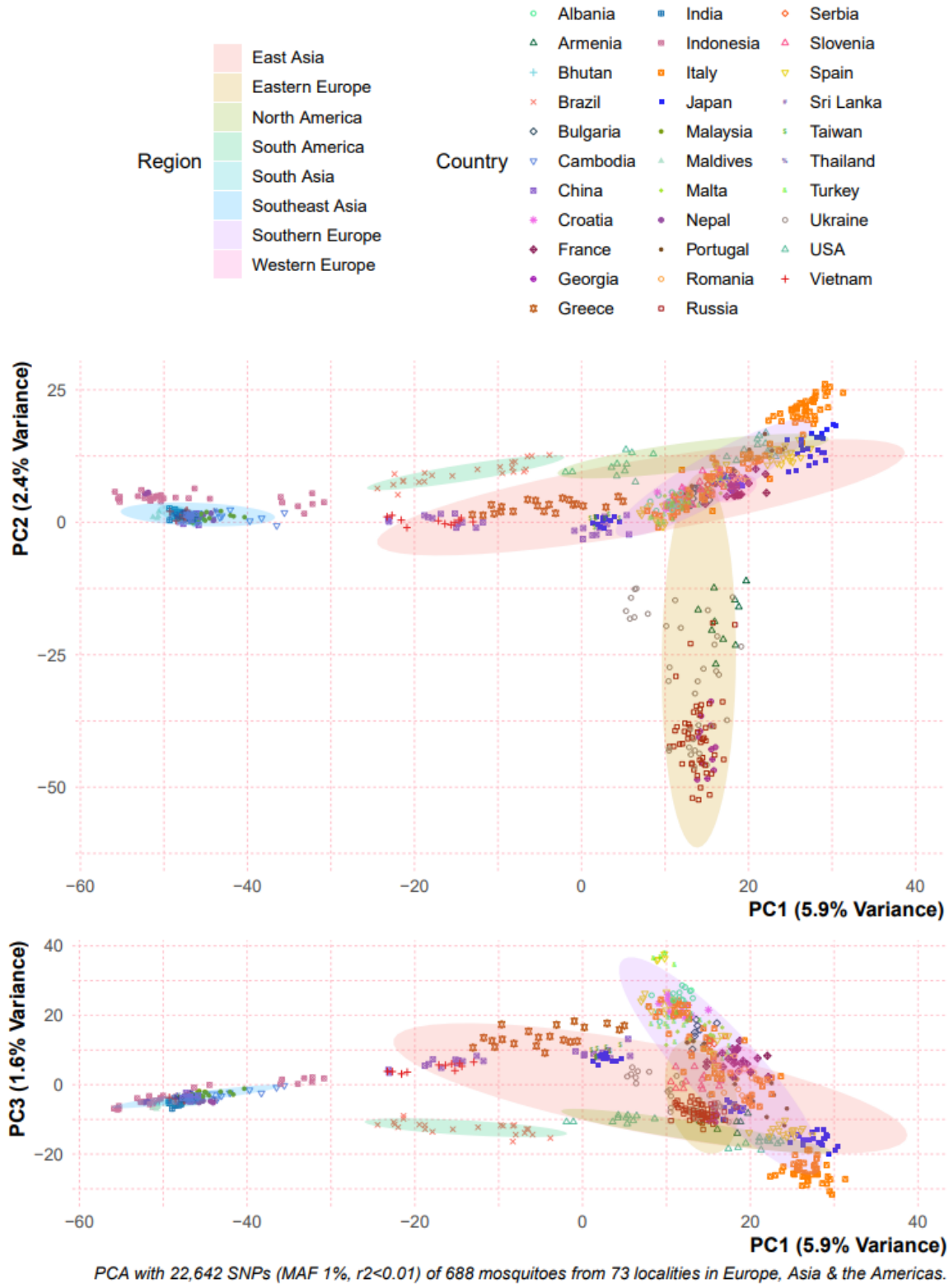
2B)



Admixture matrix  
Each bar represents the ancestry proportions for an individual for  $k=7$ .  
LEA inference for 22,642 SNPs (SNP Set 3).



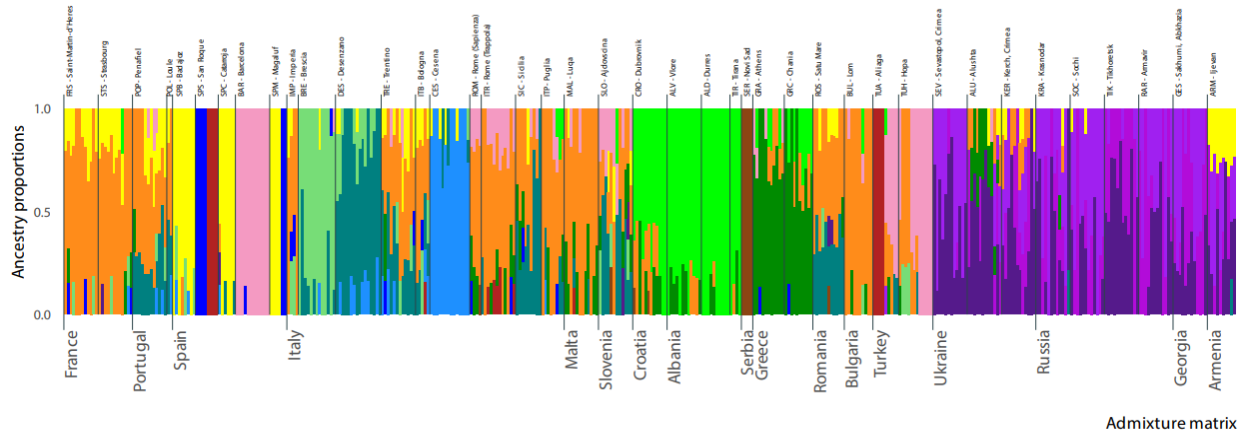
2C)



**Figure 2.** The population structure shows the best-supported number of clusters (K) identified for 688 mosquitos sampled from American, European, and native locations (K=20) (Panel A). Panel B contains the same data shown in A, but plotted for K=7. In A and B each vertical bar on the x-axis represents one

mosquito, and the y-axis shows the proportion of admixture for each individual ancestral genetic group. Panel C shows scatterplots of principal component analysis (PCA) showing the first PC on the x-axis and PCs 2 and 3 on the y-axes. Each symbol represents a mosquito, and the color and shape of the symbol indicates the country where they were sampled. Ellipses mark each region covering 80% of the samples. All plots were created in the R package LEA using data from SNP Set 3. Plots for the best K obtained from other structure algorithms (fastStructure and admixture) and using alternate SNP sets are shown in [Figures S6-S9](#), [Table S7](#), and [Supplementary Files 10](#) and [17](#).

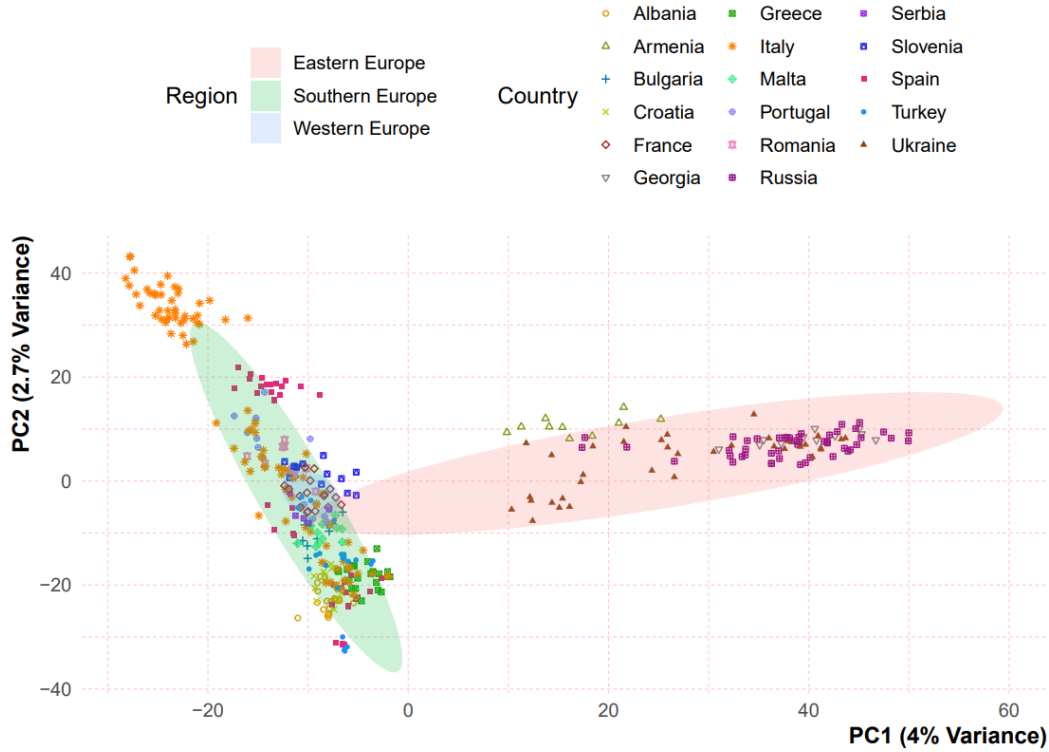
3A)

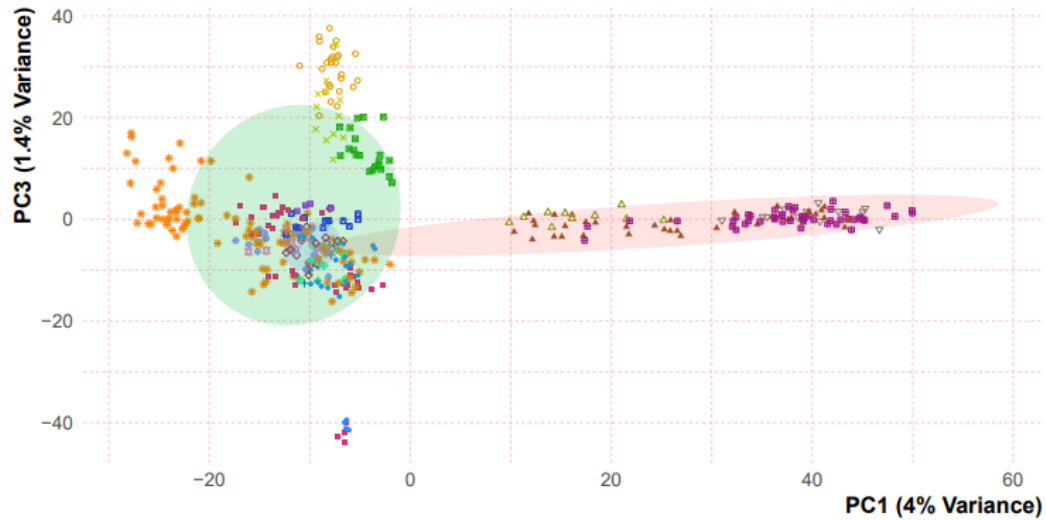


Admixture matrix

Each bar represents the ancestry proportions for an individual for k=14.  
LEA inference for 20,968 SNPs (SNP Set 3).

3B)

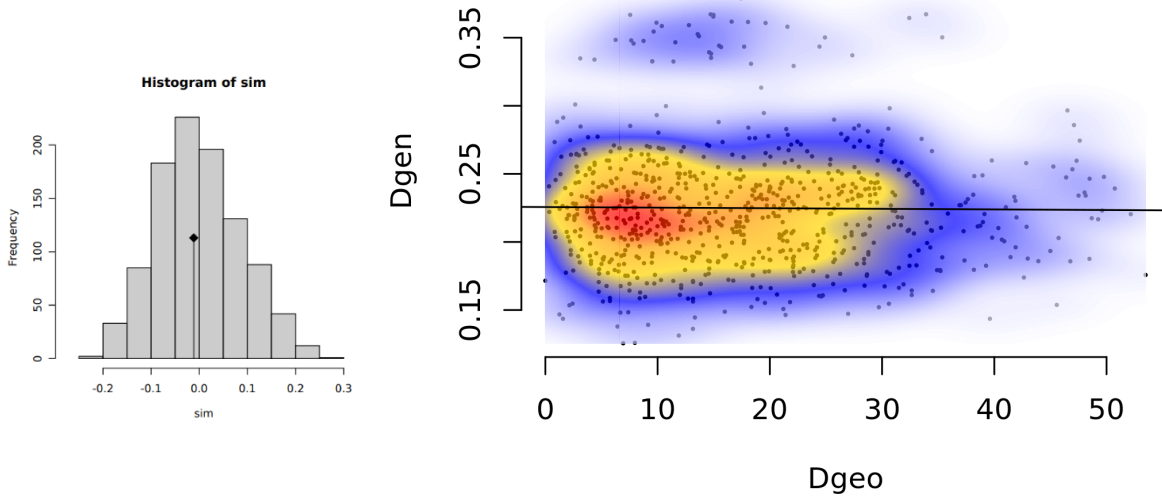




**Figure 3.** Panel A shows population structure for the best-supported number of clusters (K) identified for 410 mosquitos sampled from 41 European locations (K=14). Each vertical bar on the x-axis represents one mosquito, and the y-axis shows the proportion of admixture for each individual ancestral genetic group. Panel B shows scatterplots of principal component analysis (PCA) showing the first PC on the x-axis and PCs 2 and 3 on the y-axes. Each symbol represents a mosquito, and the color and shape of the symbol indicates the country where they were sampled. Ellipses mark each region covering 80% of the samples. All plots were created using LEA with data from SNP Set 3. Plots for the best K obtained from other structure algorithms (fastStructure and admixture) and using alternate SNP sets are shown in [Figures S10-S11](#), [Table S7](#), and [Supplementary files 9 and 16](#).

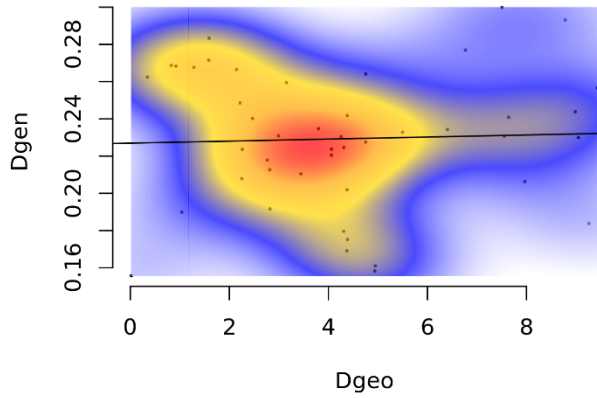
4A)

### Isolation by distance

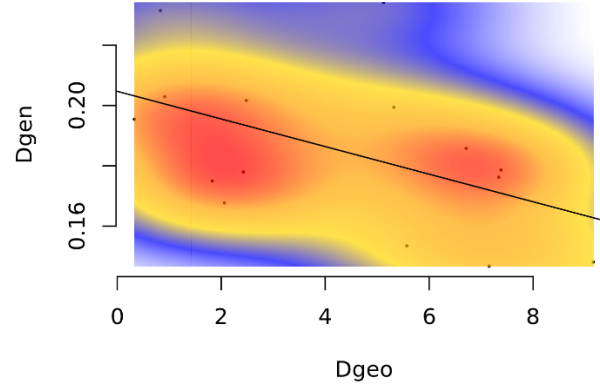


4B)

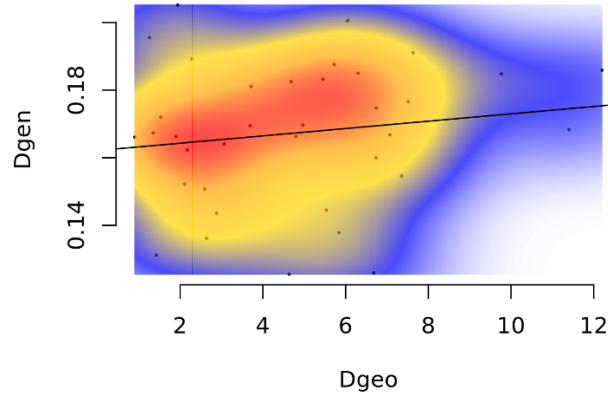
i)



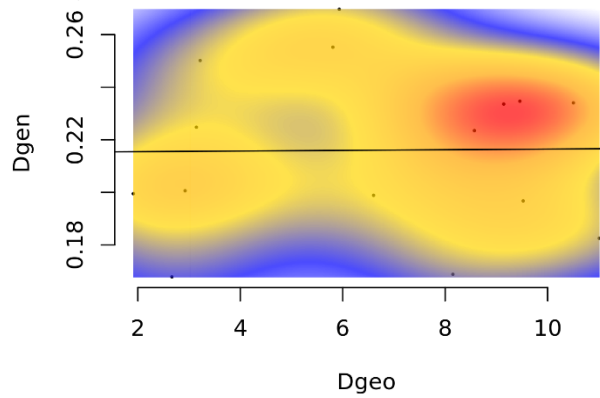
ii)



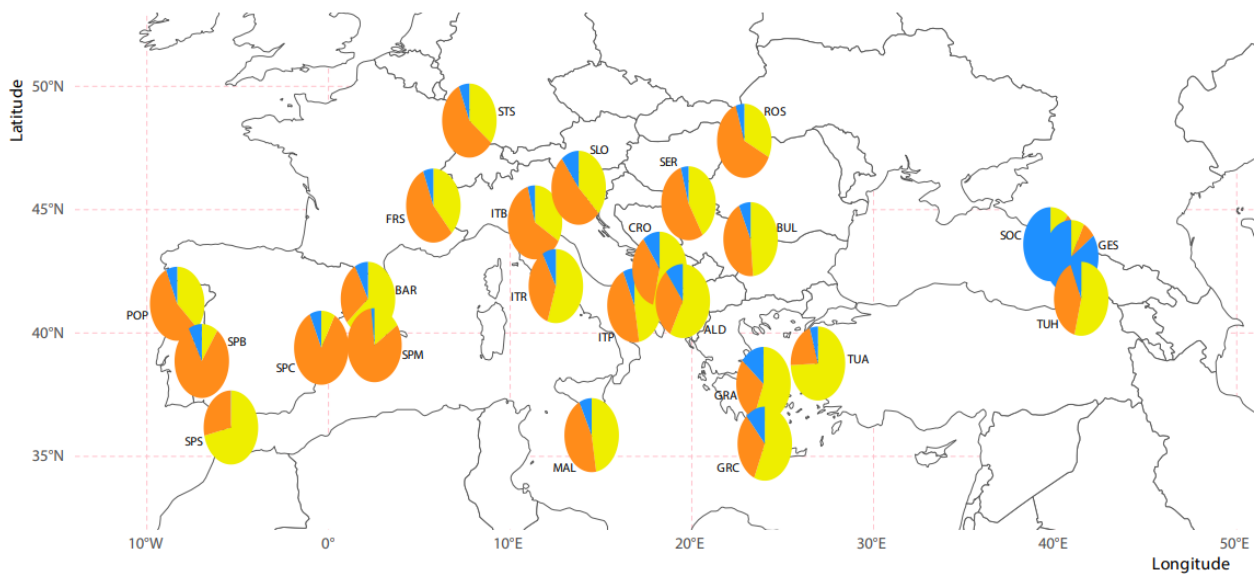
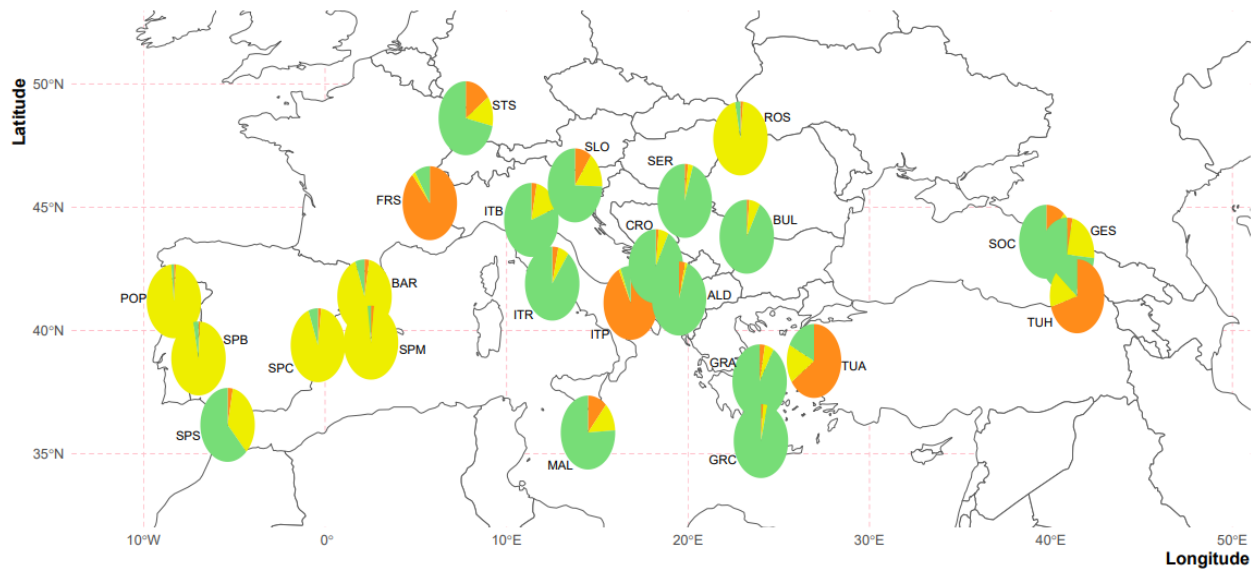
iii)

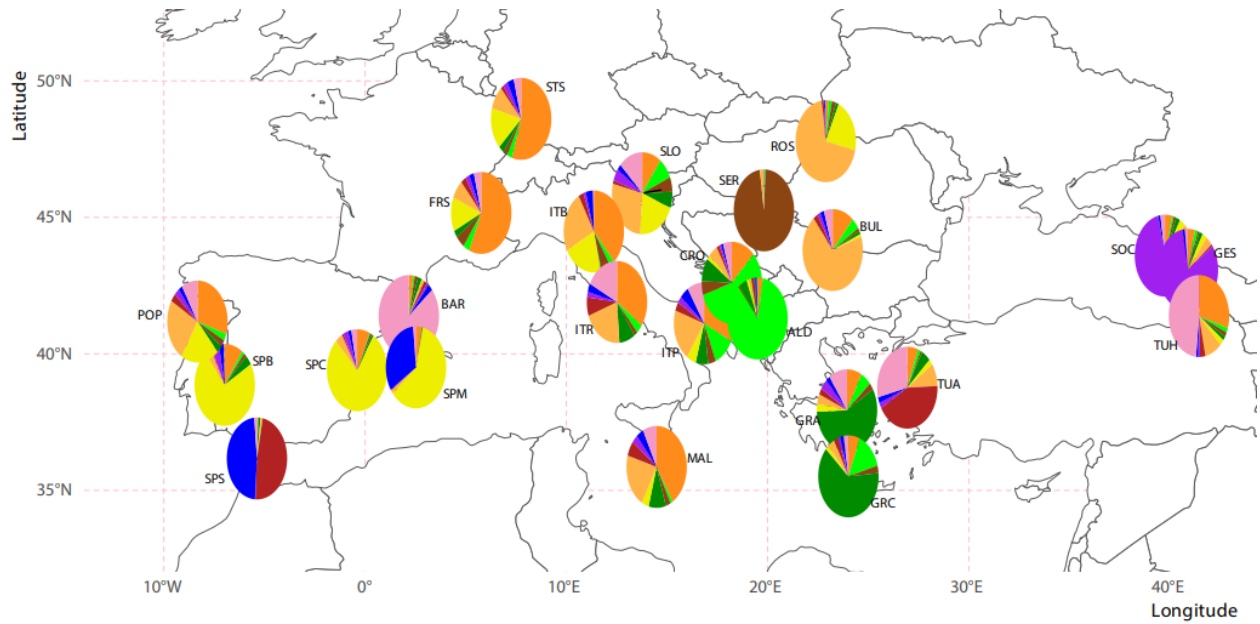


iv)



**Figure 4.** A) Plots of the Mantel test for isolation-by-distance (IBD) using the 40 European populations with at least four mosquitoes ([Table S3](#)) for SNP Set 3 (20,968 SNPs). The histogram shows the distribution of correlation coefficients between genetic ( $D_{gen}$ ) and geographic ( $D_{geo}$ ) distances obtained from 999 random permutations. The arrow indicates the observed correlation coefficient (-0.011) derived from the data, which suggests no significant deviation from random expectations ( $p < 0.519$ ). The scatterplot shows geographic distance ( $D_{geo}$ ) against genetic distance ( $D_{gen}$ ). The line from the linear regression model fit to the data had an  $R^2=0.01$ . The density of overlapping points is represented by color, with warmer shades indicating more overlap. Panel B shows scatterplots of geographic distance ( $D_{geo}$ ) against genetic distance ( $D_{gen}$ ) for subsets of European samples by region: i) Italy ( $R^2=0.0$ ), ii) Greece, Albania and Croatia ( $R^2=-0.25$ ), iii) Eastern Europe ( $R^2=0.02$ ), and iv) Iberian Peninsula ( $R^2=0.0$ ). Additional plots for subsets are shown in [Supplementary Figures S18-19](#).





**Figure 5.** Maps of the ancestry matrices obtained using the 24 locations shared between the microsatellite and SNP datasets. The top panel shows results when the clustering analysis was run for 637 mosquitos in the microsatellite dataset using only genetic data for inferencing in STRUCTURE (best supported number of clusters:  $K=3$ ). The middle panel show the clusters obtained using LEA to analyze structure for the same 24 locations with SNP Set 3 dataset. For the SNP data,  $K=3$  is shown to facilitate comparison to microsatellites.  $K=10$ , which was the best supported number of clusters for SNPs is shown in the bottom panel. For more detailed maps showing sampling locations in the microsatellite dataset see [Supplementary file 18](#).