Comparative genomics of two invasive mosquitoes (*Aedes aegypti* and *Aedes albopictus*) from the Indo-Pacific

Thomas L Schmidt[1], Jessica Chung[1,2], Ann-Christin Honnen[3,4], Andrew R Weeks[1,5], Ary A Hoffmann[1]

[1]*School of BioSciences, Bio21 Institute, University of Melbourne, Parkville, Victoria, Australia.*

[2]*Melbourne Bioinformatics, University of Melbourne, Parkville, Victoria, Australia.*

[3]*Swiss Tropical and Public Health Institute, Basel, Switzerland*

[4] *University of Basel, Basel, Switzerland*

[4]*cesar Pty Ltd, Parkville, Victoria, Australia.*

Short Title: Genetic structure of *Aedes aegypti* and *Aedes albopictus*

Abstract

The arbovirus vectors *Aedes aegypti* (yellow fever mosquito) and *Ae. albopictus* (Asian tiger mosquito) are both common throughout the Indo-Pacific region, where 70% of global dengue transmission occurs. For *Ae. aegypti* all Indo-Pacific populations are invasive, having spread from an initial native range of Africa, while for *Ae. albopictus* the Indo-Pacific includes invasive populations and those from the native range: putatively, India to Japan to Southeast Asia. This study analyses the population genomics of 480 of these mosquitoes sampled from 27 locations in the Indo-Pacific. We compared patterns of genetic differentiation to investigate pathways of invasion and ongoing gene flow in both species, and to infer potential native-range boundaries and cryptic subspecies in *Ae. albopictus*. We also tested landscape genomic hypotheses that genetic differentiation would increase with geographical distance and be lower between locations with high connectivity to human transportation routes, the primary means of dispersal at these scales. We found that genetic distances were generally higher in *Ae. aegypti*, with Pacific populations the most highly differentiated. The most differentiated *Ae. albopictus* populations were in Vanuatu, Indonesia and Sri Lanka; the latter two may represent ancestral native-range populations or cryptic subspecies, respectively. Genetic distances in *Ae. aegypti* increased with geographical distance, while in *Ae. albopictus* they decreased with higher connectivity to human transportation routes. Contrary to the situation in *Ae. aegypti*, we found evidence of long-distance *Ae. albopictus* colonisation events, including colonisation of Mauritius from East Asia and of Fiji from Southeast Asia. Our direct genomic comparisons indicate likely differences in dispersal ecology in these species, despite their phylogenetic proximity, broadly sympatric distributions and similar use of human transport to disperse. Our findings will assist biosecurity operations to trace the source of invasive material and for biocontrol operations that benefit from matching genetic backgrounds of released and local populations.

Introduction

The Indo-Pacific region, here defined as encompassing the Indian Ocean, the western and central Pacific Ocean, and the coastal territories therein, is the site of 70% of global dengue transmission [1]. Infections are vectored by two mosquito species, *Aedes aegypti* (yellow fever mosquito) and *Ae. albopictus* (Asian tiger mosquito) [2,3], both of which have established widespread invasions in the region. *Aedes aegypti*, the primary regional vector, has a putative native range of West Africa [4] and is thought to have invaded the Indian Ocean via the Mediterranean [5], while its invasion history in the western Pacific remains unclear. *Aedes albopictus* has a hypothesised native range that includes Japan, China, northern India, and parts of Southeast Asia [6]. The southern extent of this range remains unclear and may cover as far as Indonesia and New Guinea [6–8]. Regional expansion of *Ae. albopictus* is thought to have occurred more recently than *Ae. aegypti*, with colonisation likely occurring in successive waves [7,9].

Like many invasive taxa, *Ae. aegypti* and *Ae. albopictus* (hereafter *Aedes* spp.) exhibit stratified dispersal. This consists of two distinct processes: short-range, active dispersal and long-range, passive dispersal [10]. Active dispersal of *Aedes aegypti* is generally highly localised, as measured by both traditional mark-release-recapture methods (8 – 199 m/gen [11]) or current methods using genomic estimates of relatedness (33 – 131 m/gen [12]). Similar observations have been made of *Ae. albopictus* using mark-release-recapture [13]. At the metropolitan scale (< 50 km), passive dispersal in *Aedes* spp. has been observed directly [14] or inferred from the spatial distribution of close kin [15,16]. Dispersal at broader scales such as between cities is almost exclusively by passive transportation on ships, aircraft and land vehicles [17–20].

The unrestricted spread of *Aedes* spp. has many adverse consequences. Both species are presently continuing to colonise new tropical, subtropical and, for *Ae. albopictus*, temperate regions [21], and future expansions are expected in the face of increased urbanisation and climate change [22]. In addition to range expansions, dispersal between *Aedes* spp. populations can spread advantageous alleles that reduce the efficacy of local control programs (Endersby et al. 2020: in press). Incursions of these species are frequently found on vessels such as ships and planes [18,19] (Schmidt et al. 2020: submitted), and when source locations can be identified they are often far from the point of collection [19] (Schmidt et al. 2020: submitted).

Understanding the forces that influence dispersal, gene flow, and population structure in *Aedes* spp. can help efforts to prevent, control, and prepare for future threats from these species. As well as being a critical focus for dengue control efforts, the Indo-Pacific presents a useful location to investigate these processes due to the broad sympatry of *Aedes* spp. within the region. Investigating the genetics of sympatric taxa can help link patterns of genetic differentiation to the forces that shape population structure [23,24]. For instance, when taxa share a common area but have contrasting patterns of differentiation, this may indicate intrinsic differences in how they disperse [25–27], while similar dispersal behaviour can produce similar patterns of differentiation [28].

Here we investigate these processes in *Aedes* spp. from the Indo-Pacific region at spatial scales from ~20 km (inter-city) to ~16,000 km (trans-oceanic). Dispersal at these scales largely proceeds along the network of shipping routes that link each inhabited coastal location

with every other location (Fig 1). As *Aedes* spp. are thought to make use of the same vessel types for transportation, their dispersal may take place along similar pathways, which may produce similar patterns of genetic structure [28,29]. However, in invasive populations this expectation can be complicated by heterogeneous histories of colonisation. During colonisation, stochastic processes can produce rapid change in allele frequencies [30]. This structure can be resistant to 'erosion' via gene flow when local population densities are high [31] as they generally are for *Aedes* spp. Likewise, *Ae. albopictus* may be able to survive transportation over longer distances than *Ae. aegypti* by undergoing diapause [32–34] which confers resistance to cold and desiccation [35]. *Aedes aegypti* does not undergo diapause, but like *Ae. albopictus* it can make use of egg quiescence to survive adverse conditions [34].

This study uses high-resolution markers to compare patterns of broad-scale genetic structure in *Aedes* spp. Few such comparative studies have been conducted [37,38], and none have used high-resolution molecular markers, which have demonstrated greater power to distinguish between *Aedes* populations [39]. We also test landscape genomic hypotheses of isolation by distance [40] and isolation by environment [41]; specifically, whether genetic differentiation is lower between locations with high connectivity to human transportation routes. These analyses provide a means of disentangling the putative influences of human transportation routes which may direct gene flow, high local population density following invasion which may limit the effects of gene flow, and diapause in *Ae. albopictus* which may permit gene flow between more distant populations than for *Ae. aegypti*. Recently, single-species studies using high-resolution molecular markers have reported strong divergence among *Ae. aegypti* populations [5,39] and weaker divergence among *Ae. albopictus* populations [42,43], though highly-differentiated *Ae. albopictus* populations have been recorded [44]. As broad-scale studies have typically included a greater number of samples from European, African and American *Aedes* spp. populations than from Indo-Pacific populations [5,42,43,45], this study also fills an important geographical gap in the population genomics literature of *Aedes* spp..

We found clear differences between the spatial genetic structure of *Ae. aegypti* and *Ae. albopictus*, and our landscape genomic analyses indicated that *Ae. aegypti* populations were generally structured by space, and *Ae. albopictus* populations by environment. These results may reflect important differences in the capacity of *Aedes* spp. to disperse and to invade new regions. The findings of this study will assist biosecurity operations that aim to trace the source of invasive material [19,46] and for biocontrol operations that benefit from matching genetic backgrounds of released and local populations [47].

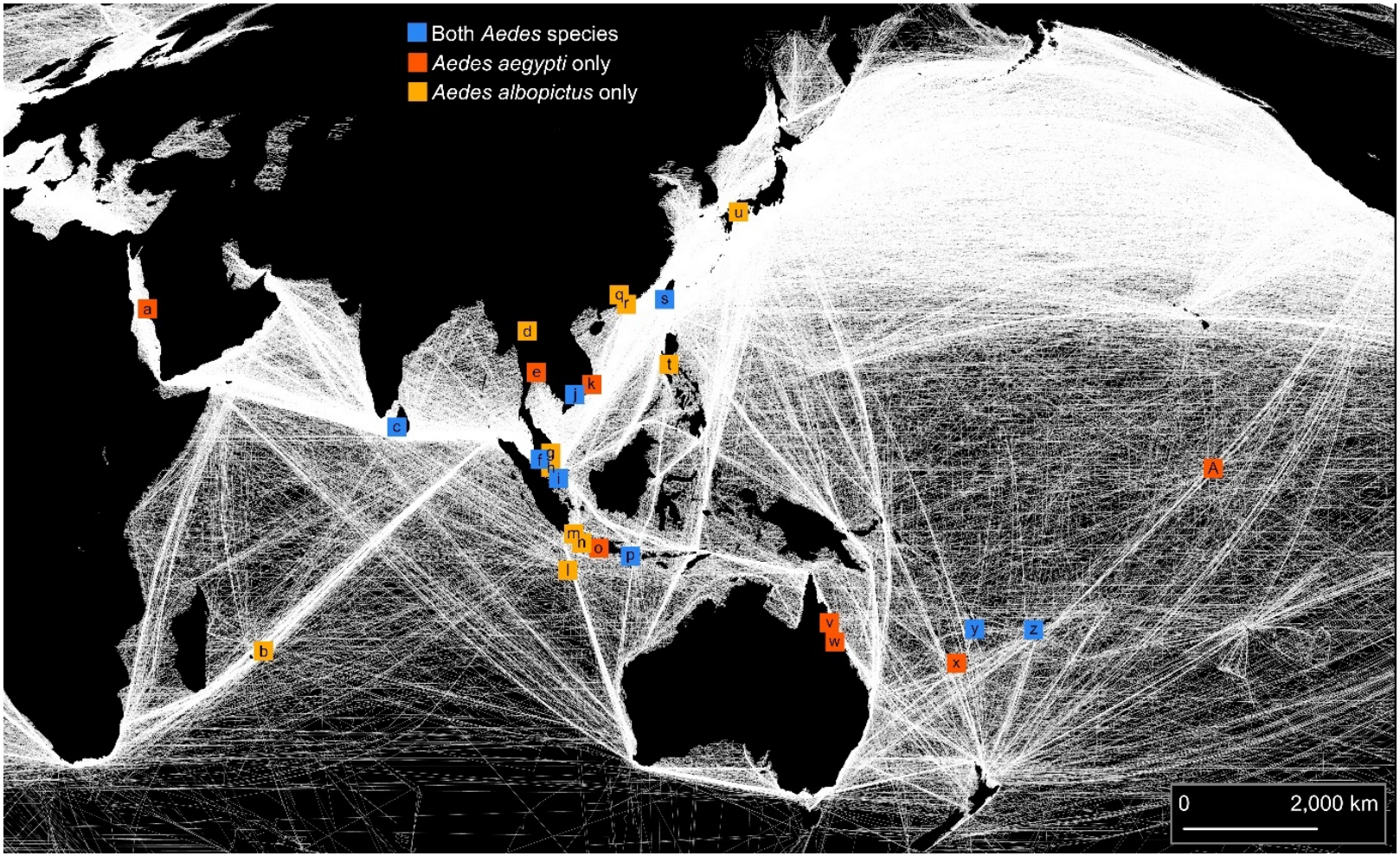

Fig 1. Sampling locations of *Aedes aegypti* and *Aedes albopictus*.

White lines indicate density of shipping routes. The map uses a Mollweide projection with a central meridian of 120°E. The map was produced in ARCMAP 10.5.1 using shipping route data made available from Halpern et al. [36]. Sampling locations are indicated by letters and are described in Tables 1 and 2.

Table 1. Details of *Aedes aegypti* sampled from 16 populations.

See Fig 1 for map id locations. See Materials and Methods for details regarding filtering and calculation of connectivity indices.

map id (see Fig 1)	sample	location(s)	country	year(s) collected	no. genotypes after filtering	connectivity (marine)	connectivity (aerial)
a	Jeddah	Various locations in Jeddah	Saudi Arabia	2018	18	7.92	194
c	Colombo	Kiribathgoda	Sri Lanka	2017	12	9.62	101
e	Bangkok	Seven locations in Bangkok	Thailand	2016	18	0.17	326
f	Kuala Lumpur	Bandar Sunway/UKM Specialist Centre	Malaysia	2015/2017	18	3	257
i	Singapore	Various locations in Singapore	Singapore	2015	15	2.64	408
j	Ho Chi Minh City	Various locations in Ho Chi Minh City	Vietnam	2014	18	0.43	109
k	Nha Trang	Hon Mieu Island	Vietnam	2016	11	5.26	7
o	Yogyakarta	Various locations in Yogyakarta, Java	Indonesia	2014	13	0.44	29
p	Bali	Near Denpasar Airport, Bali	Indonesia	2016/2017	18	1.24	100
s	Taiwan	Various locations in Southwest Taiwan	Republic of China	2016	8	7.25	65
v	Cairns	Six locations in urban Cairns	Australia	2015	18	1.86	56
w	Townsville	Near Townsville Airport	Australia	2014	10	2.25	16
x	New Caledonia	Various locations in Nouméa	New Caledonia	2018	13	1.24	13
y	Vanuatu	Various locations in Efate	Vanuatu	2018	14	0.71	28
z	Fiji	Various locations in Nadi	Fiji	2018	11	0.85	43
A	Kiribati	Various locations in Kiribati	Kiribati	2018	9	0.8	3

Table 2. Details of *Aedes albopictus* sampled from 20 populations.

See Fig 1 for map id locations. See Materials and Methods for details regarding filtering and calculation of connectivity indices. Colombo-1 was treated as a distinct population due to divergences (see text), and omitted (*) from landscape genomic analyses.

map id (see Fig 1)	sample name	location(s)	country	year(s) collected	no. genotypes after filtering	connectivity (marine)	connectivity (aerial)
b	Mauritius	Near Chamarel village	Mauritius	2017	7	2.68	41
c	Colombo-1	Delgoda	Sri Lanka	2017	7	*	*
c	Colombo-2	Delgoda and Kiribathgoda	Sri Lanka	2017	8	9.62	101
d	Chiang Mai	Near Wachirathan Waterfall	Thailand	2017	5	0	48
f	Selangor	Kuala Lumpur and Gombak, Selangor	Malaysia	2015/2016	11	3	257
g	Pahang	Kuantan	Malaysia	2015/2016	18	3.2	5
h	Johor	Johor Baru	Malaysia	2015/2016	18	3.28	18
i	Singapore	Near Bukit Timah Nature Reserve	Singapore	2017	14	2.07	408
j	Ho Chi Minh City	Various locations in Ho Chi Minh City	Vietnam	2017/2018	18	0.44	109
l	Christmas Island	Near Christmas Island International Airport	Australia	2017	7	1.91	2
m	Jakarta	Various locations in Jakarta, Java	Indonesia	2016	17	1.09	185
n	Bandung	Various locations in Bandung, Java	Indonesia	2017	14	0.53	21
p	Bali	Various locations in Bali	Indonesia	2016/2017	17	1.24	100
q	Guangzhou	13 locations in Guangzhou	China	2015	16	0.24	338
r	Hong Kong	Various locations in Hong Kong	China	2016	16	3.87	355
s	Taiwan	Various locations in Southwest Taiwan	Republic of China	2016	15	7.24	65
t	Manila	Manila, Quezon City and Makati City	Philippines	2017	6	1.56	193
u	Matsuyama	Near Ehime University	Japan	2017	11	1.97	12
y	Vanuatu	Various locations in Efate	Vanuatu	2018	15	0.71	28
z	Fiji	Various locations in and near to Nadi	Fiji	2018	16	0.84	43

Materials and Methods

Sample collection, genotyping, processing and subsampling

Aedes aegypti were sampled from 16 locations in the Indo-Pacific region, across a ~16,000 km range from Jeddah, Saudi Arabia to Kiribati (Fig 1, Table 1). *Aedes albopictus* were sampled from 20 locations ranging from Mauritius to Fiji to Matsuyama, Japan (Fig 1, Table 2). We considered mosquitoes collected within the same city to be from the same population, as structuring within these scales is typically weak (e.g. [16]). The smallest distance between two distinct populations was ~20 km, between Singapore and Johor, Malaysia. Among populations within the same country, the smallest distance was ~130 km, between Guangzhou and Hong Kong in China.

Mosquitoes were genotyped using a pipeline that has been described elsewhere [19] and is described here in detail in S1 Appendix. Briefly, we used the double-digest restriction site-associated (ddRAD) sequencing protocol for *Ae. aegypti* [39] to obtain sequence reads, which were processed in Stacks v2.0 [48]. We used BOWTIE V2.0 [49] to align reads to the *Ae. aegypti* [50] and *Ae. albopictus* [51] mitochondrial genome (mtDNA) assemblies. Reads that did not align to the mtDNA assemblies were aligned to the nuclear assemblies AaegL5 [52] and AaloF1 [53] to obtain nuclear genotypes for *Ae. aegypti* and *Ae. albopictus* respectively. This ensured that the nuclear genotypes would not be confounded by transpositions of mitochondrial DNA into the nuclear genome (NUMTs), a common occurrence in *Aedes* spp. [54]. As some population samples contained more close kin than others, we removed individuals in order of missing data so that no first order kin pairs remained and no population had greater than 18 genotypes, following the procedure described in S1 Appendix.

We retained 224 *Ae. aegypti* and 256 *Ae. albopictus* genotypes. We used the Stacks program REF_MAP to build catalogs containing all mosquito genotypes from each species, from which we called single nucleotide polymorphism (SNP) genotypes at RAD stacks using deault Stacks settings. Only SNPs scored in at least 75% of genotypes from each population were used in analyses.

Individual-level genetic structure

We investigated genetic structure among individual mosquito genotypes using three methods: discriminant analysis of principal components (DAPC [55]); sparse non-negative matrix factorisation (sNMF [56]); and PCA-UMAP [57], which combines principal components analysis (PCA) with uniform manifold approximation and projection (UMAP [58]). For these analyses, we filtered the 224 *Ae. aegypti* and 256 *Ae. albopictus* genotypes to retain biallelic SNPs with $< 20\%$ missing data, with minimum depth of 3, and maximum depth of 45. Genotypes were phased and missing data imputed in BEAGLE V4.1 [59] using default settings and 10 iterations.

DAPC was run in the R package *adegenet* [60]. In DAPC, each genotype is assigned to one of K genetic clusters. We ran the 'find.clusters' algorithm to determine the optimal K for $1 \leq K$

≤ N, with N equal to the number of populations (16 for *Ae. aegypti*, 20 for *Ae. albopictus*). We ran $1x10^9$ iterations and selected the value of *K* with the lowest Bayesian information criterion (BIC). We then repeated the above analysis, selecting the value of *K* with the lowest Akaike information criterion (AIC), a less conservative measure than BIC.

We ran sNMF in the R package LEA [61]. This analysis estimates individual ancestry coefficients, assuming that individual genotypes are produced from the admixture of *K* ancestral lineages, where *K* is unknown *a priori*. To determine which value of *K* was optimal for summarising the variation in each species, we set $1 \leq K \leq N$ and ran 100 iterations of the sNMF algorithm for each *K*, selecting the *K* with the lowest cross-entropy across all runs. For the chosen *K* we selected the iteration with the lowest cross-entropy for visual presentation.

To generate PCA-UMAPs, we adapted code provided in Diaz-Papkovich et al. [57]. UMAP provides a means of projecting high-dimensional data onto two-dimensional space, and has advantages over other dimensionality reduction techniques in that it better preserves the global data structure between clusters in reduced dimensions. Combining UMAP and PCA has been shown to produce optimum definition of population clusters [57]. PCA-UMAPs used the first 5 principal components of the PCA, which were projected in two dimensions via UMAP using 50 neighbours and a minimum distance of 0.5.

Population-level genetic structure

For investigations of population-level genetic structure, we were cautious to process and filter the data so that uneven population sample size (*n*) would not bias analyses, while still including as many genotypes as possible. To balance these aims, we subsampled each dataset ten times, sampling (with replacement) from each population a number of genotypes (five) equal to the minimum *n* among all populations. For each subsample we used the Stacks program POPULATIONS to calculate F_{ST}' [63] between each pair of populations, and to calculate mean observed heterozygosity (H_O) and nucleotide diversity (π) for each population. We used the results from the ten subsamples to calculate the mean and 95% confidence intervals of each estimated parameter.

For genome-wide SNP datasets, genetic differentiation can be estimated with five genotypes [62]. To determine whether using only five genotypes in each subsample affected our F_{ST}' estimates, we took ten subsamples of 16 individuals from the *Ae. aegypti* populations in Bali, Bangkok, Jeddah and Kuala Lumpur, and calculated F_{ST}' means and 95% confidence intervals as above. We compared these results with those calculated with five genotypes.

Analysis of the data indicated a deep genetic division within the Sri Lankan *Ae. albopictus* sample (see Fig 2ii and Results: Individual-level genetic structure). Seven of the 15 genotypes formed a group highly divergent from all others, including the other Sri Lankan genotypes sampled from the same region of Colombo. We therefore treated the Sri Lankan *Ae. albopictus* sample as two separate populations, denoted as Colombo-1 (highly divergent) and Colombo-2. We investigated this divergence further by calculating the folded allele frequency spectrum for each *Ae. albopictus* population in VCFTOOLS [64], using the ten subsamples of five genotypes to calculate means and 95% confidence intervals.

Landscape genomics

We investigated the geographical structuring of genetic variation in *Aedes* spp. using distance-based redundancy analysis (dbRDA), performed in the R package *vegan* [65]. These analyses tested two hypotheses: that genetic differentiation would increase with geographical distance (i.e. isolation by distance), and that genetic differentiation would decrease with greater connectivity to shipping or aerial transportation routes (i.e. isolation by environment).

Our measure for connectivity to shipping routes was derived from previously published raster data of 799,853 commercial shipping tracks from October 2004 to October 2005 (Halpern et al. [36]; see Supporting Online Material therein for full description). We used ARCMAP 10.5.1 [66] to calculate the average density of shipping tracks within a 200 km radius from each population sampling point, so that the inland site at Chiang Mai scored 0 while Jeddah, Colombo-2, and Taiwan received the highest scores (see Fig 1). For connectivity to aerial transportation routes, we used data from the OpenFlights Routes Database (https://openflights.org/data.html#route), which lists all airline routes active at June 2014. For each sampled population, we used the total number of active flight routes at the largest international airport of that city or town. This gave values of between 2 (Christmas Island) to 408 (Singapore). Shipping and aerial connectivity scores for each population are listed in Tables 1 and 2. No correlation between shipping and aerial connectivity scores were observed for either *Ae. aegypti* ($R^2 = 0.003$, $P = 0.838$) or *Ae. albopictus* ($R^2 < 0.001$, $P = 0.938$) populations. We omitted the outlier *Ae. albopictus* subpopulation Colombo-1 from analyses (see Results: Individual-level genetic structure).

The connectivity of each population pair to the shipping and aerial networks was calculated as the average of their connectivity scores for that network type. These average scores were used to construct pairwise connectivity matrices of each transportation type for each species. As independent variables in a dbRDA must be rectangular, these matrices were transformed into principle coordinates using the function *pcnm*. We likewise calculated geographical distance matrices for each species, and transformed these into principle coordinates. We used mean F_{ST}' as a measure of genetic differentiation.

To test for isolation by distance in each species, we built dbRDAs using the scaled vectors of geographical distance as the independent variable and F_{ST}' as the dependent variable. To test for isolation by environment, we first determined scaled geographical distance vectors with which to condition the model, by building dbRDAs where each vector was treated as a separate independent variable, and selecting vectors with $P < 0.01$. We assessed shipping and aerial connectivity separately. We built dbRDAs using the scaled vectors of connectivity as the independent variable and F_{ST}' as the dependent variable, with the model conditioned using the significant vectors for geographical distance. All dbRDAs were built using the function *capscale*. Marginal significance of variables was assessed using *anova.cca* with 9,999 permutations.

Results

Individual-level genetic structure

We retained 54,296 SNPs for analysis of the 224 *Ae. aegypti* and 40,016 SNPs for analysis of the 256 *Ae. albopictus*. The three analyses of genetic structure among individual genotypes, DAPC, sNMF, and PCA-UMAP, were consistent in their findings, and indicated several broad differences in genetic structure among *Aedes* spp. in the Indo-Pacific.

DAPC using lowest BIC (Fig 2) selected $K = 3$ for both *Ae. aegypti* and *Ae. albopictus*. As BIC is a conservative measure, these represent broad genetic groupings, with more detailed variation described in sNMFs (Fig 3), PCA-UMAPs (Figs 4 and 5), and DAPC using lowest AIC (S1 Fig A). For *Ae. aegypti*, DAPC divided genotypes up into three spatial groups representing the Pacific Islands (orange), Australia (cream), and all central and Western populations (purple) (Fig 2i). DAPC of *Ae. albopictus* (Fig 2ii) showed a very different pattern, with a single cluster (purple) present across the entire sampling range. One other cluster (cream) was found in Indonesia, indicating deep differences between these populations and nearby populations like Christmas Island, despite separation distances of only ~450 km. The final *Ae. albopictus* cluster was found within the Colombo population, consisting of a subset of seven genotypes that were deeply divergent from the other eight Colombo genotypes. In all further analyses the Colombo genotypes were treated as two populations, denoted as Colombo-1 (highly divergent) and Colombo-2.

The sNMF of *Ae. aegypti* (Fig 3i) found $K = 10$ was an optimal number of ancestral lineages for summarising genetic variation among genotypes. This high number relative to the number of populations (16) largely reflected the high level of structuring between populations, with five of the ten ancestral lineages almost wholly confined to single locations: Jeddah, Colombo, Taiwan, New Caledonia and Vanuatu. Proximate population pairs such as Cairns and Townsville, Kuala Lumpur and Singapore, and Ho Chi Minh City and Nha Trang each had common ancestral lineages, as did Fiji and Kiribati, despite their distance of separation. The most putatively admixed population was Bangkok, which showed genetic similarities with the other Southeast Asian populations, but not with either Taiwan or Colombo.

These results contrasted sharply with those of *Ae. albopictus* (Fig 3ii), for which $K = 5$ was optimal. Fewer ancestral lineages indicated greater genetic similarity among *Ae. albopictus* populations. The interpretation of the spatial distribution of ancestral lineages among *Ae. albopictus* requires reference to the native and invasive ranges of this species. The northernmost native range lineage, in orange, indicated a common heritage among the East Asian populations. This lineage was also dominant in Mauritius. A second native range lineage, in purple, indicated common heritage among Southeast Asian populations from Chiang Mai to Singapore, and was also found in Christmas Island and Fiji, but not Vanuatu. The population at Ho Chi Minh City was made up of roughly even contributions from the two lineages. The Indonesian *Ae. albopictus* genotypes (Fig 3ii, red) of Jakarta, Bandung and Bali were genetically distinct from the East and Southeast Asian native range genotypes.

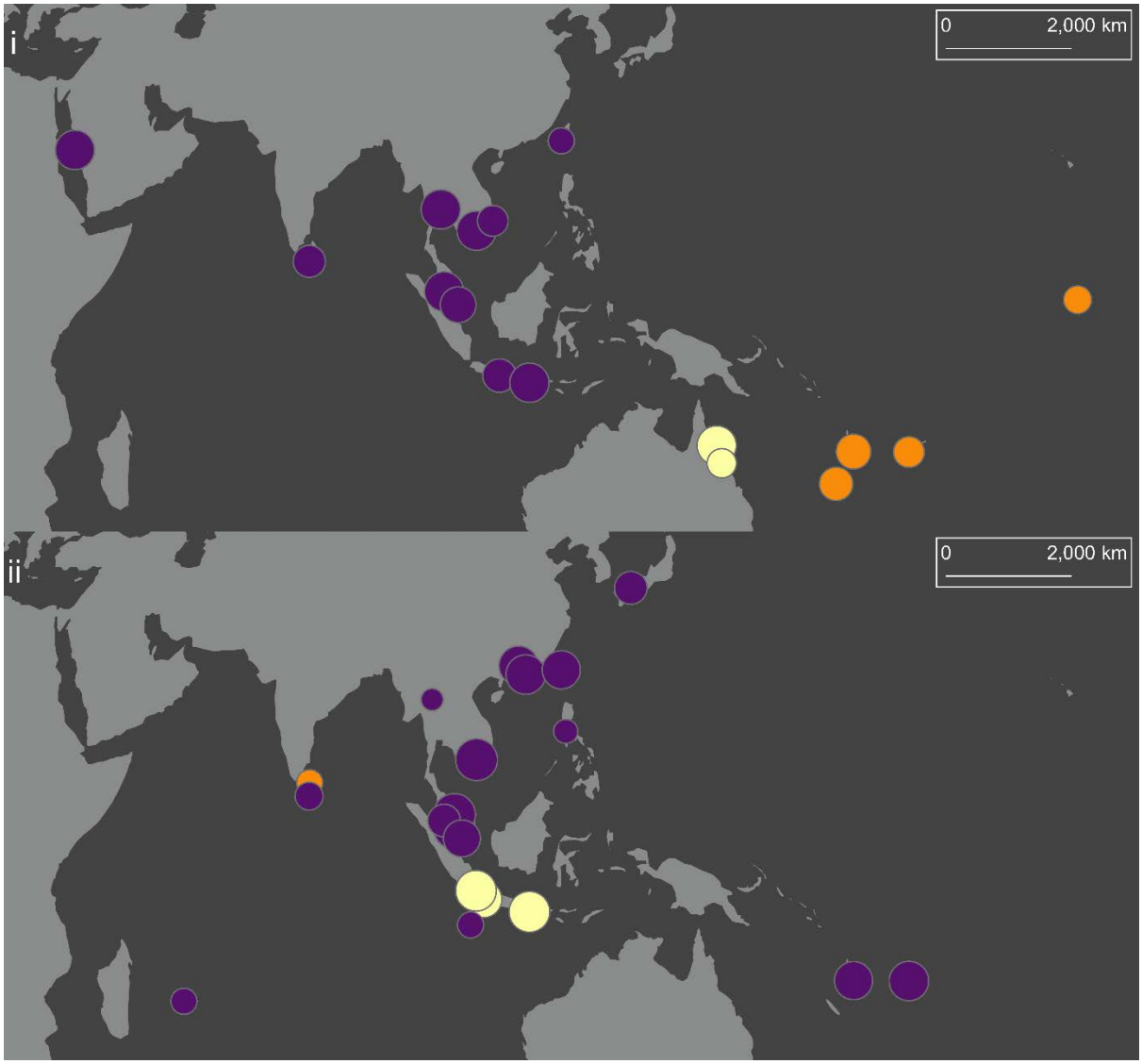

Fig 2. DAPCs of *Ae. aegypti* (i) and *Ae. albopictus* (ii) using lowest BIC (*K* = 3).

Colours indicate cluster membership. Circles are sized relative to population sample size. The map uses a Mollweide projection with a central meridian of 120°E. The map was produced in ARCMAP 10.5.1, using shapefiles made available by www.naturalearthdata.com.

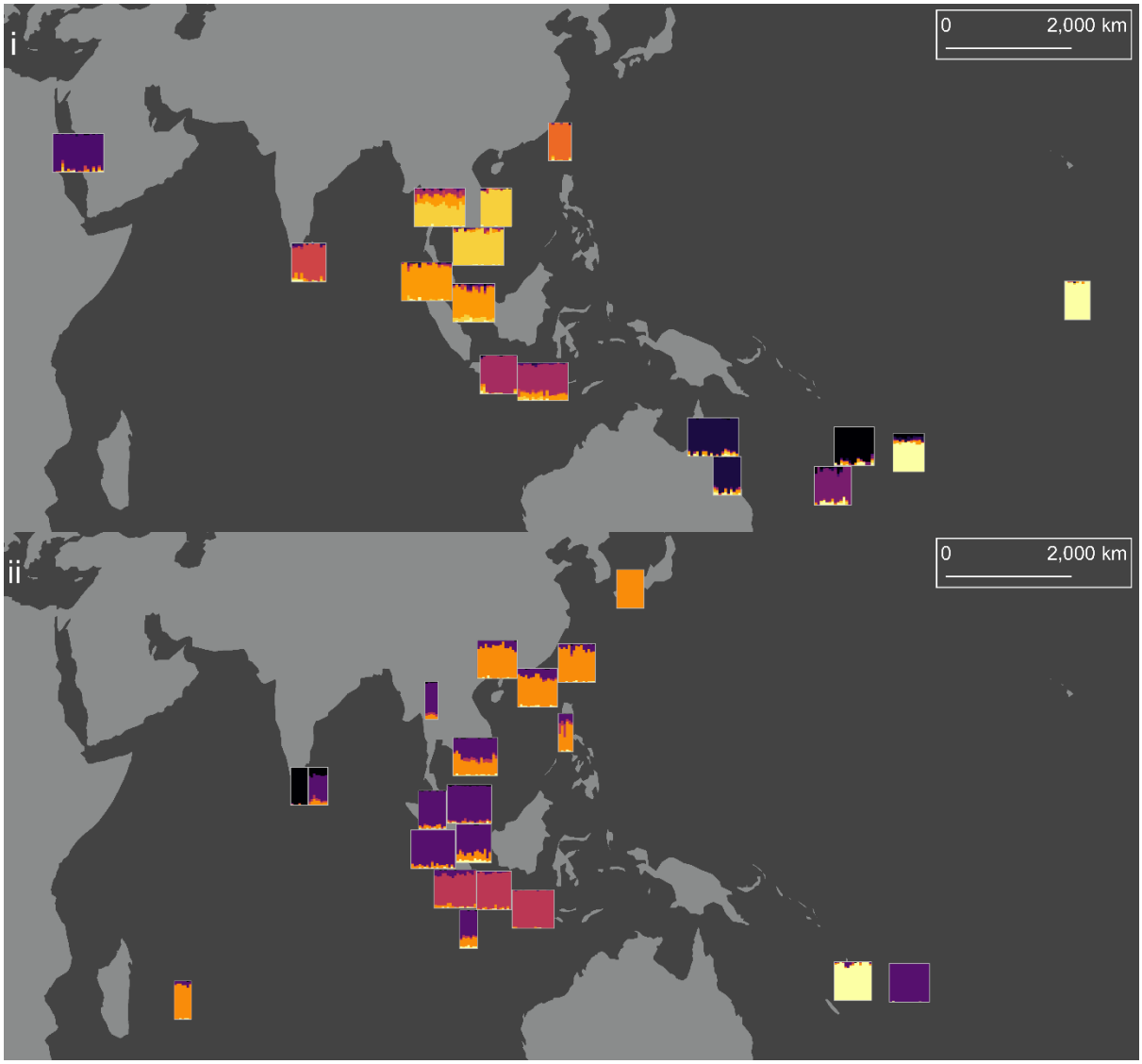

Fig 3. sNMFs of *Ae. aegypti* (i) and *Ae. albopictus* (ii).

Number of ancestral lineages (*K*) was 10 for *Ae. aegypti* and 5 for *Ae. albopictus*. Each population is a rectangle, with each genotype a vertical line made of 1-*K* colours. Colours indicate ancestral lineages. The map uses a Mollweide projection with a central meridian of 120°E. The map was produced in ARCMAP 10.5.1, using shapefiles made available by www.naturalearthdata.com.

Treating Colombo *Ae. albopictus* as two separate populations revealed that genotypes from the highly-divergent Colombo-1 population were almost entirely of one ancestral lineage (Fig 3ii, black) found only in Colombo. Colombo-2 was a composite of lineages, including the unique lineage from Colombo and the purple native range lineage.

PCA-UMAP of *Ae. aegypti* (Fig 4) showed genotypes grouping clearly by spatial location, echoing the sNMF results but providing more detailed information on differentiation between similar populations. Genotypes clustered first by population, then by region, with Pacific and Australian genotypes the most distinct. PCA-UMAP indicated similarities between genotypes

from Fiji and Kiribati, Vanuatu and New Caledonia, and Singapore and Kuala Lumpur. The westernmost samples from Jeddah and Colombo also showed evidence of similarity. Despite Taiwan's proximity to Southeast Asia, Taiwanese *Ae. aegypti* were clearly distinct from all others.

PCA-UMAP of *Ae. albopictus* (Fig 5) likewise gave results similar to the sNMF but with more detail. The Matsuyama population was revealed as distinct from others in East Asia, and there was evidence that the Fijian genotypes were most closely related to Johor, Malaysia, and those of Christmas Island were most closely related to Singapore. The Ho Chi Minh City genotypes were split into two clusters, four grouping with the East Asian genotypes and 14 with the Southeast Asian genotypes. This likely reflects Ho Chi Minh City as having ancestry from both of these groups (Fig 3ii). A corresponding split was observed in the results of DAPC using lowest AIC (S1 Fig A).

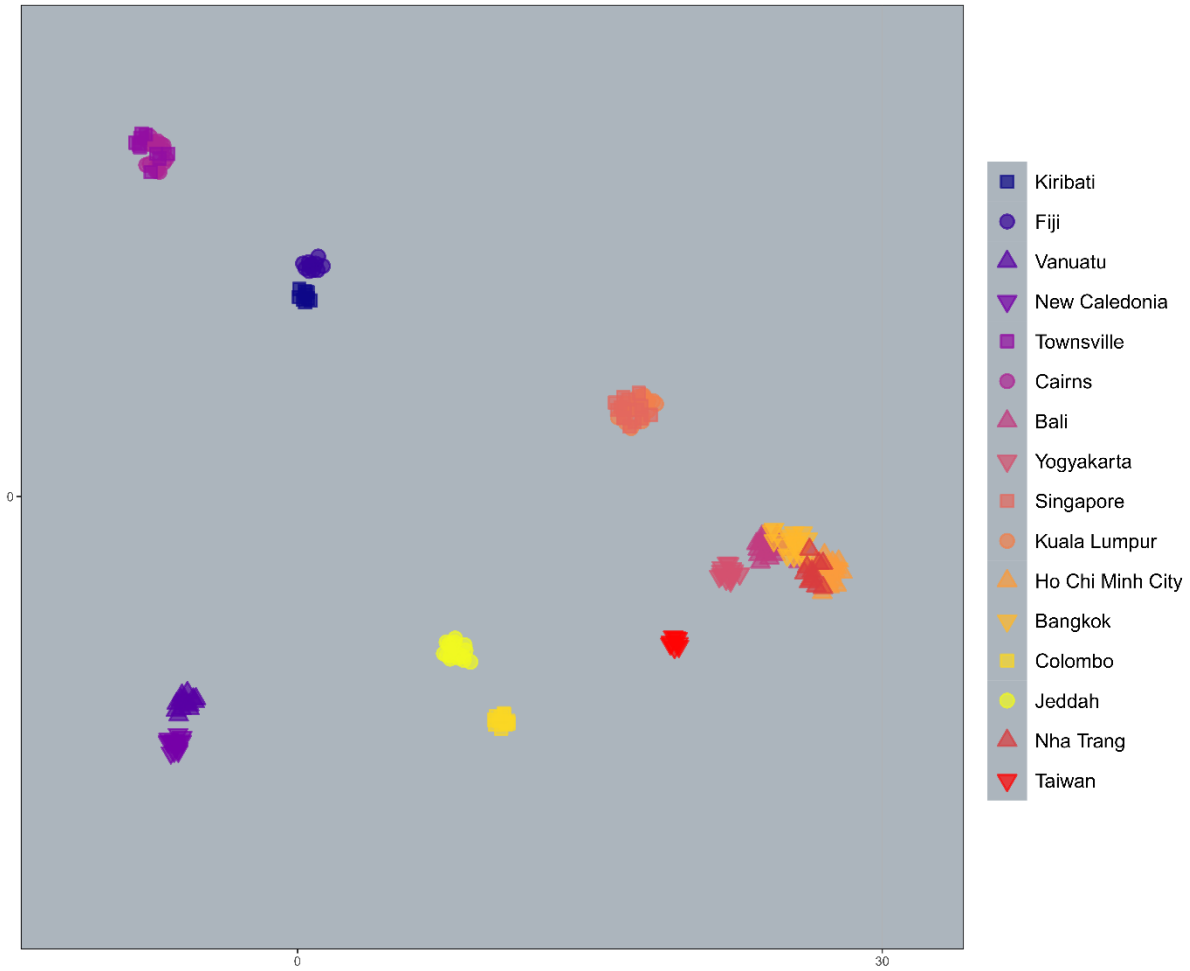

Fig 4. PCA-UMAP of 224 *Ae. aegypti* genotypes.

Genotypes are coloured by population, with similar colours given to geographically proximate populations.

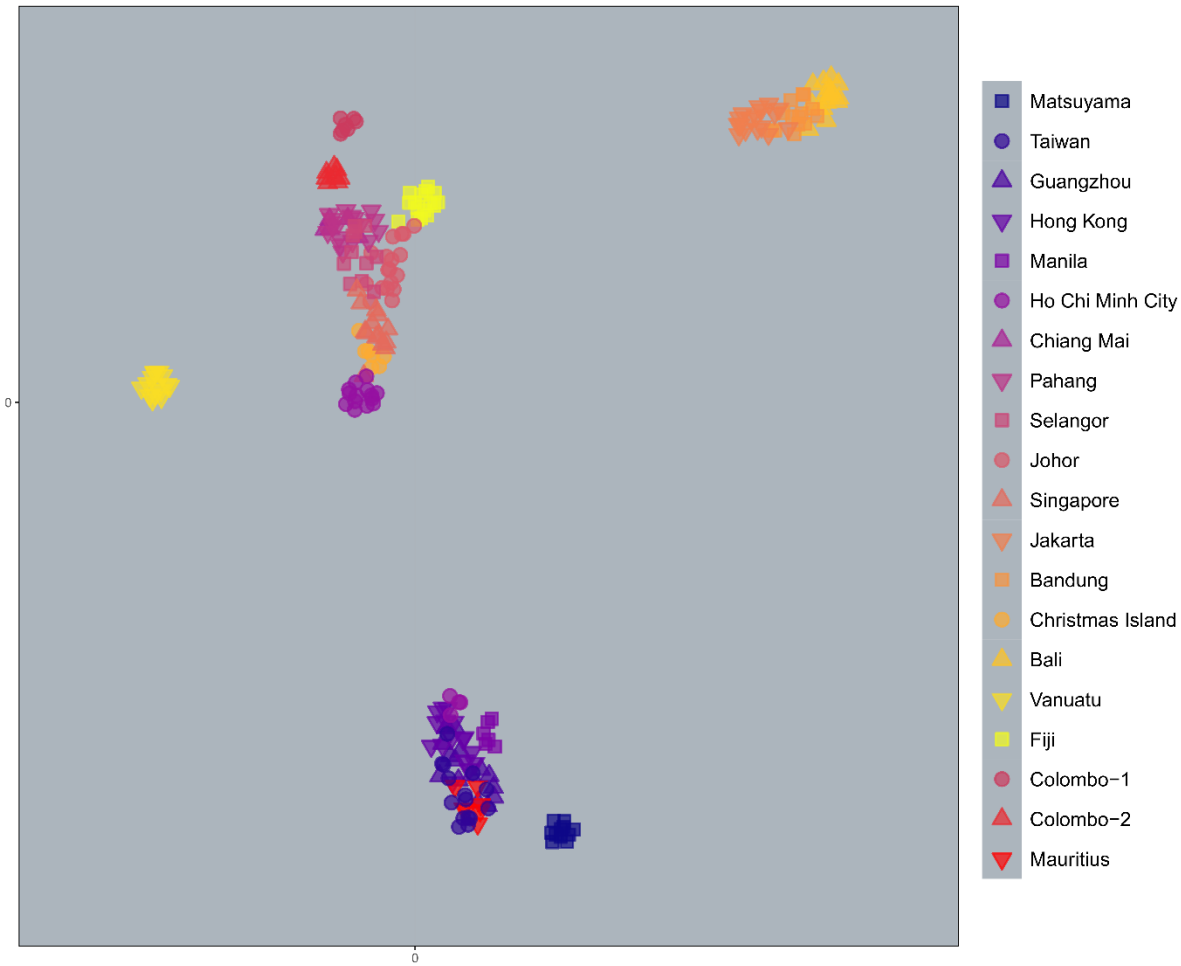

Fig 5. PCA-UMAP of 256 *Ae. albopictus* genotypes.

Genotypes are coloured by population, with similar colours given to geographically proximate populations.

Population-level genetic structure

Pairwise F_{ST}' estimates for *Ae. aegypti* and *Ae. albopictus* are listed in S1 Tables A and B respectively. Estimates are given with 95% confidence ranges, which show that pairwise F_{ST}' estimates were generally consistent across the ten subsamples. Most population pairs showed small F_{ST}' confidence intervals relative to the size of the mean (S1 Tables A - B). Mean H_O and π estimates were also mostly consistent (S1 Tables C & D). Numbers of nuclear SNPs retained for population-level analyses had ranges of 46,151 – 54,647 (*Ae. aegypti*) and 35,984 – 57,639 (*Ae. albopictus*).

Fig 6 shows all population pairs with mean pairwise F_{ST}' < 0.03, indicating high genetic similarity. Australian *Ae. aegypti* populations were genetically very similar, and there was a network of similarity among Southeast Asian populations (Fig 6i). Indonesian *Ae. albopictus* populations were genetically similar, as were a network of populations covering parts of the native range in East and Southeast Asia as well as Mauritius (Fig 6ii). Taiwan *Ae. aegypti* had the high-

est mean F_{ST}' from all others ($\bar{x}$ = 0.093; s.d. = 0.016). Colombo-1 *Ae. albopictus* had a particularly high mean F_{ST}' when compared to other populations ($\bar{x}$ = 0.153; s.d. = 0.018), including substantial differentiation from Colombo-2 ($\bar{x}$ = 0.113; s.d. = 0.008). Colombo-2 had a much lower mean F_{ST}' when compared to other populations ($\bar{x}$ = 0.070; s.d. = 0.020).

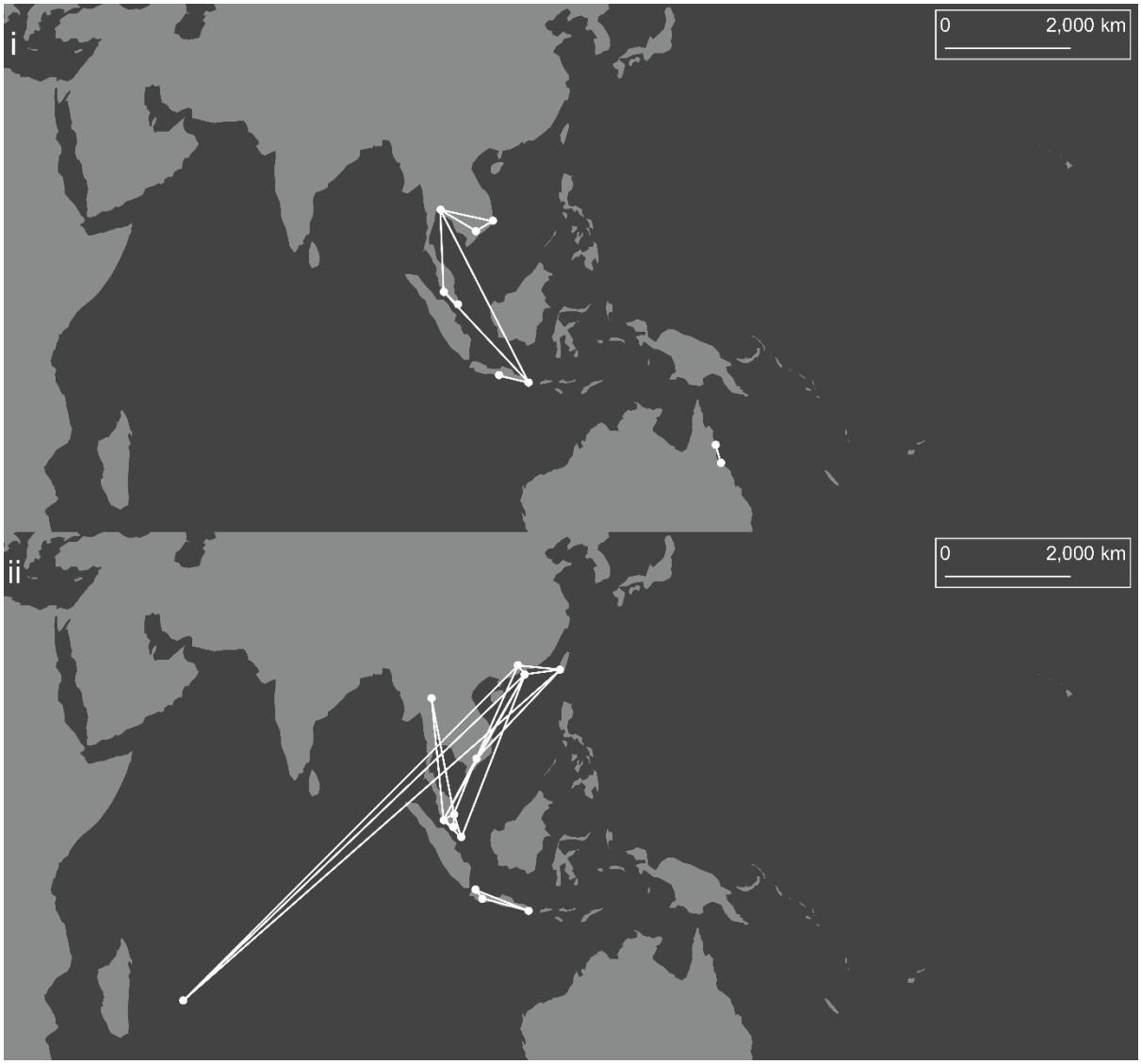

Fig 6. Population pairs of *Ae. aegypti* (i) and *Ae. albopictus* (ii) with mean pairwise F_{ST}' < 0.03.

Circles connected by lines indicate populations with mean pairwise F_{ST}' < 0.03. The map uses a Mollweide projection with a central meridian of 120°E. The map was produced in ARCMAP 10.5.1, using shapefiles made available by www.naturalearthdata.com.

When 16 genotypes were used to calculate F_{ST}', results were similar to those when 5 genotypes were used (S1 Table E). Using five genotypes tended to slightly overestimate F_{ST}', particularly for highly differentiated populations, though F_{ST}' scores were consistent relative to one another.

Analysis of the folded allele frequency spectrum in *Ae. albopictus* showed that Colombo-1 had lower frequencies of rare alleles and higher frequencies of common alleles compared with other populations (Fig 7). The proportion of loci in Colombo-1 with a minor allele frequency of 0.1 was 0.23 times higher than the mean for all populations. The proportions of loci in Colombo-1 with minor allele frequencies of 0.4 and 0.5 were 2.32 and 2.75 times higher respectively than the mean for all populations. Neither Colombo-2 nor any of the other populations had consistent differences from other populations in minor allele frequency. Colombo-1 had the largest proportion of monomorphic loci out of all populations.

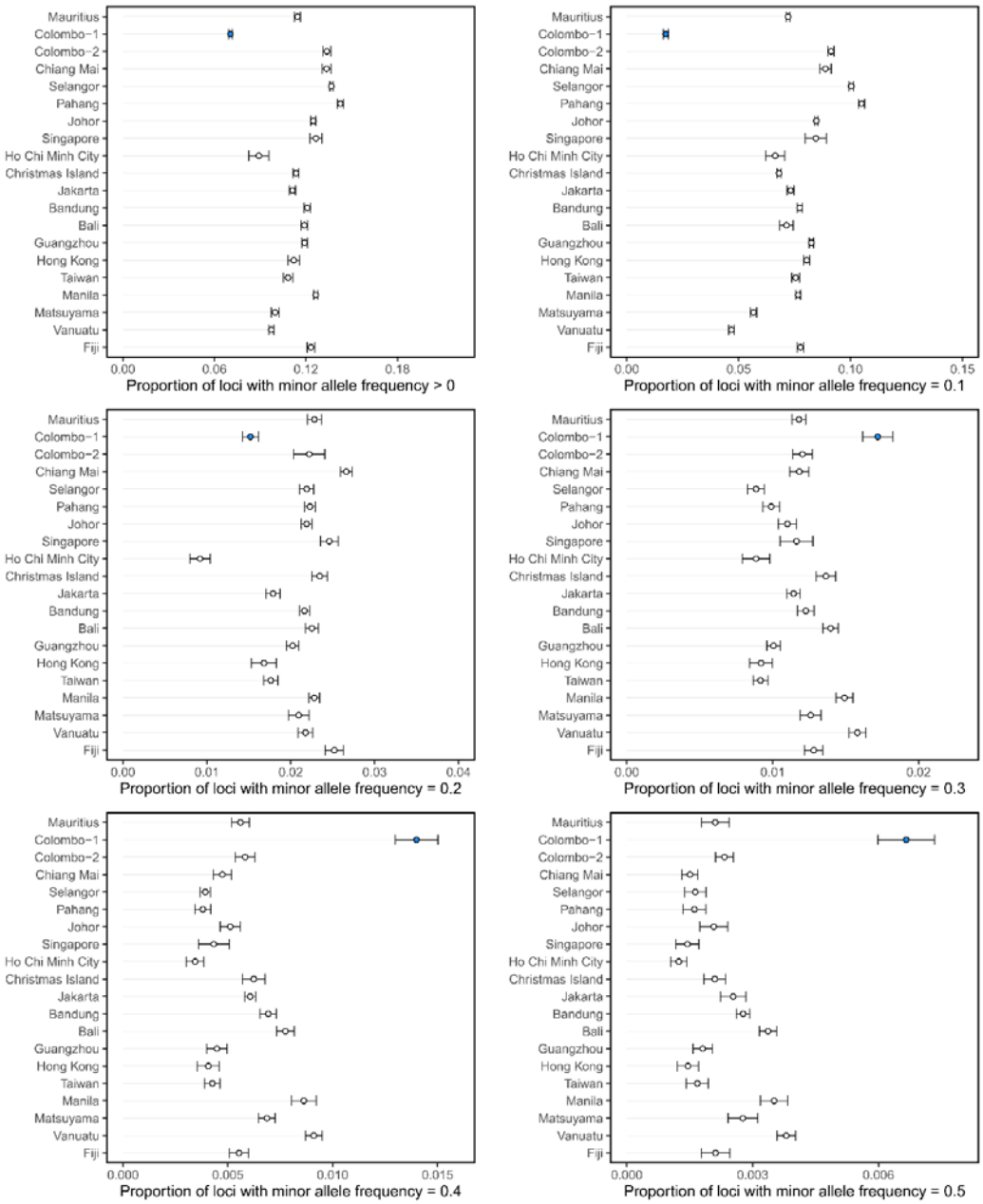

Fig 7. Minor allele frequencies of *Ae. albopictus* populations from the Indo-Pacific region.

Circles represent the mean frequency from the 10 subsamples, with 95% confidence intervals. The Colombo-1 population is indicated in blue.

Landscape genomics

The dbRDAs assessing isolation by distance indicated that geographical distance was positively associated with F_{ST}' in *Ae. aegypti* (Table 3: $F_9 = 2.136$, $P = 0.002$) but not in *Ae. albopictus* (Table 4: $F_9 = 0.965$, $P = 0.534$). When each scaled vector for geographical distance was treated as an independent variable, one was significant at $P < 0.01$ in *Ae. aegypti* and none were significant in *Ae. albopictus*.

The dbRDAs assessing connectivity to transport networks indicated that average pairwise connectivity to both shipping ($F_{15} = 2.411$, $P = 0.034$) and aerial ($F_{13} = 2.650$, $P = 0.006$) transportation routes was negatively associated with F_{ST}' in *Ae. albopictus* (Table 4). Neither was significantly associated with F_{ST}' in *Ae. aegypti* (Table 3), irrespective of whether the significant geographical distance vector was used to condition the models.

Table 3. Results of dbRDAs assessing landscape genomic hypotheses in *Ae. aegypti*.

All geographical distances and connectivity indices were transformed into principal coordinates (PCs) for analysis. The single PC with P < 0.01 when testing for isolation by distance was used to condition the isolation by environment models.

	Df	Sum of Squares	F	P
Geographical distance	9	0.026	2.136	0.002
Residual	6	0.008		
Marine connectivity	12	0.022	0.811	0.727
Geographical distance PC1	1	0.006	2.659	0.041
Residual	2	0.004		
Aerial connectivity	11	0.021	1.075	0.467
Geographical distance PC1	1	0.003	1.692	0.131
Residual	3	0.005		

Table 4. Results of dbRDAs assessing landscape genomic hypotheses in *Ae. albopictus*.

All geographical distances and connectivity indices were transformed into principal coordinates (PCs) for analysis.

	Df	Sum of Squares	F	P
Geographical distance	9	0.018	0.965	0.534
Residual	9	0.019		
Marine connectivity	15	0.034	2.411	0.034
Residual	3	0.003		
Aerial connectivity	13	0.032	2.65	0.006
Residual	5	0.005		

Discussion

Here we present the first direct comparison of genome-wide genetic structure in the globally-invasive, dengue-vectoring mosquitoes *Ae. aegypti* and *Ae. albopictus*. Our study region, the Indo-Pacific, is the site of 70% of global dengue transmission [1], and this study contributes important knowledge to the understanding of invasion histories and ongoing gene flow within this region. While both species are already widespread in the Indo-Pacific [21], increased urbanisation and climate change may allow them to invade new regions in the coming decades [22]. Furthermore, the same dispersal processes facilitating invasions into new regions can also lead to genetic invasions of advantageous alleles between established populations, as observed recently in this region (Endersby et al. 2020: in press). Our comparative approach indicated different genetic structure patterns in these species at each level of analysis: individual, population, and landscape genomic. For the most part, *Ae. aegypti* populations were genetically differentiated from each other, and differentiation increased with geographical distance. *Aedes albopictus* populations in the native range were split between two main regions of genetic similarity, one in East Asia and one in Southeast Asia, and certain populations outside the native range showed clear signs of recent invasion from these regions; specifically, Mauritius from East Asia, and Fiji and Christmas Island from Southeast Asia. *Aedes albopictus* populations in Indonesia were highly differentiated from all others, despite their proximity to other Southeast Asian populations. In Colombo, we observed two distinct *Ae. albopictus* subpopulations, one of which may represent an ancestral native-range population or cryptic subspecies. Overall these results indicate that *Aedes* spp. have established invasions in the Indo-Pacific along different pathways, and that current gene flow patterns are also different between the two species. Our findings also point to various regions of interest in the *Ae. albopictus* native range that demand further investigation.

The patterns of divergence observed in *Ae. aegypti* fit with the hypothesis that *Ae. aegypti* began its invasion of the Indian Ocean region from regions to the west [5,67]. Putatively, the Indian Ocean regions and Australia were invaded from the Mediterranean, though there is some uncertainty over the timing and direction of this invasion [68,69], and the strong genetic differentiation we observed in Australian populations confounds this somewhat. Interestingly, the most highly differentiated *Ae. aegypti* populations were in the western Pacific populations of Australia, the Pacific Islands and Taiwan, suggesting that these regions do not share an invasion history with Indian Ocean regions. Separate invasions of Australia and the Pacific Islands also seem likely, based on the strong differentiation between Australian *Ae. aegypti* and those of nearby Pacific Islands such as Vanuatu and New Caledonia.

The high differentiation of western Pacific *Ae. aegypti* populations may reflect an earlier invasion of these regions, possibly from the Americas. However, these divergences could instead be due to a lack of gene flow into these populations following invasion, with strong genetic differentiation produced by the strong effects of drift following founder events [30]. Different structural patterns in these populations may be reinforced by high local population densities, reflecting a "founder takes all" [31] system following colonisation. Considering the strong isolation by distance observed in this species (Table 3), the relative remoteness of the Pacific Islands suggests these experience limited gene flow, leading to pronounced differences among populations. By contrast, transport connectivity was an important influence on *Ae. albopictus* genetic

structure (Table 4), and there were signs of recent dispersal into Fiji from Southeast Asia (likely Malaysia); Southeast Asia is thus a likely source of the Fijian invasion of *Ae. albopictus* in 1989 [70]. These results require cautious interpretation, however, as most *Ae. albopictus* invasions of the Pacific are more recent than those of *Ae. aegypti*, and thus low differentiation among populations may reflect recent invasion rather than ongoing gene flow. Nevertheless, genetic similarities among geographically distant populations were more apparent in *Ae. albopictus* than in *Ae. aegypti*, and the isolation by distance patterns observed in *Ae. aegypti* indicate that gene flow in this species is more common between geographically proximate populations.

Our landscape genomics analyses point to different processes structuring *Aedes* spp. populations in the Indo-Pacific region. For *Ae. aegypti*, population structuring by geographical distance but not transportation routes accords with a pre-modern invasion timeline. *Aedes aegypti* is thought to have begun its invasion of the region in ~18-19th c. [67], a time when shipping routes likely differed from the 21st century routes used in our analyses. In contrast, the more recent expansion of *Ae. albopictus* occurred at a time when international shipping routes more closely resembled the 2004 routes analysed here. Although dispersal by aircraft may be less common than by shipping in this species (Schmidt et al. 2020: submitted), it is worth noting that much of the material used in this study was collected near to international airports, which may exaggerate the influence of the aerial dispersal network. Also, as the metric for aerial connectivity was derived from the activity of nearby airports, aerial connectivity scores may correlate closely with other untested variables that could influence regional genetic structure, such as economic activity.

Although the observation of isolation by distance in *Ae. aegypti* but not in *Ae. albopictus* may reflect the greater impact of long-distance transportation in modern times, the capacity of *Ae. albopictus* to undergo diapause as well as quiescence [34] could also be involved, as mosquitoes must be able to survive long-distance transportation for gene flow to occur. A comparison of incursion pathways of *Aedes* spp. species into Australia indicated that not only was each species dispersing from a different set of locations, but also that they were likely transported through different vessel types, with *Ae. albopictus* found more frequently at seaports and *Ae. aegypti* at airports (Schmidt et al. 2020: submitted). The most common source of *Ae. albopictus* incursions was southern China, where diapause has been observed in field populations [71]. Better survivorship on long ship journeys may allow *Ae. albopictus* to colonise distant locations without first colonising geographically intermediate 'stepping stones', which accords with much of the broad-scale literature for this species [42,72]. For taxa capable of long-distance dispersal, long-range colonisation becomes more probable than stepping-stone colonisation as the required number of successful colonisations is minimised [73].

The investigation of *Ae. albopictus* in its native range revealed a pair of native-range clades covering East and Southeast Asia, with both clades found in Ho Chi Minh City (Fig 3ii). The Indonesian populations were strongly differentiated from these and all other clades, confirming previous results from allozyme data [74]. These could represent the southern border of the native range, and whereas the East and Southeast Asian populations have become largely homogenised through gene flow the Indonesian populations have remained distinct. Alternatively, Indonesia may have been invaded earlier than other regions. A possible ancient invasion source is India, which had considerable connectivity with Java and Bali in antiquity [75].

The *Ae. albopictus* of Colombo, Sri Lanka, appeared to represent two sympatric subpopulations: Colombo-1, a highly-divergent, ancestral subpopulation of *Ae. albopictus*, possibly a

cryptic subspecies; and Colombo-2, a subpopulation seemingly produced by admixture between the ancestral Colombo-1 lineage and the Southeast Asian native-range lineage (Fig 3ii). The folded allele frequency spectrum of Colombo-1 was unlike other populations, having greater proportions of high-frequency minor alleles and fewer rare alleles (Fig 7). Several cryptic subspecies of *Ae. albopictus* have been identified in its Chinese native range [76], but the western extent of its range in India is not well-studied genetically. It is possible that Colombo is also within the native range of *Ae. albopictus*, and that Colombo-1 is an ancestral, native-range subpopulation. Alternatively, Colombo may have been invaded by an ancestral, native-range population from India or elsewhere. The putatively admixed Colombo-2 subpopulation is evidence of recent incursions into Colombo from the native range.

The different patterns of population genetic structure in the *Aedes* spp. described here has implications for biosecurity and biocontrol. Recently, genomics has been used to identify source locations of *Aedes* incursions into Australia and New Zealand [19] (Schmidt et al. 2020: submitted), which threaten not only the arrival of species themselves but also the introduction of insecticide resistance alleles (Endersby et al. 2020: in press). The genomic patterns identified here can help in identifying biosecurity threats from both species. The data are also useful for biocontrol programs which involve release of modified mosquitoes into natural populations, including those modified for dengue suppression via the introduction of *Wolbachia*. Patterns of population genetic structure can indicate areas where genetic backgrounds are likely to be similar or different to those of target populations – an important consideration for widespread species such as *Aedes* spp. that can be locally adapted to different conditions [47].

Acknowledgements

We thank Tim Hurst, Samia Elfekih, Dayanath Meegoda, Hadian Sasmita, Nazni Wa, Pattamaporn Kittayapong, Lia Faridah, Kozo Watanabe, Thaddeus Carvajal, Chi Yung Jim, Gwendolin Wong, Ashley Callahan, Jason Axford, Tim Hurst, Stephen Doggett, Elizabeth Valerie, Craig Williams, Joe Davis and other researchers for samples. We also thank Anthony van Rooyen, Nancy Endersby-Harshman and Nick Bell for sample processing. This research was facilitated by use of the Nectar Research Cloud, a collaborative Australian research platform supported by the Australian Research Data Commons (ARDC) and National Collaborative Research Infrastructure Strategy (NCRIS). We thank the Department of Agriculture and Water Resources (DAWR), Australian Government, for providing funding for this work. AAH was supported by Program and Fellowship grants from the National Health and Medical Research Council (NHMRC), no. 1037003. TLS and AAH were also supported by the Wellcome Trust UK, no. 108508.

Data Availability

Aligned .bam files for 224 *Ae. aegypti* and 256 *Ae. albopictus* will be made available at the Sequence Read Archive, NCBI, by time of publication.

References

1. Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, et al. The global distribution and burden of dengue. Nature . 2013;496(7446):504–7.

2. Gubler DJ. Dengue and dengue hemorrhagic fever. Clin Microbiol Rev . 1998;11(3):480–96.

3. Wu J-Y, Lun Z-R, James AA, Chen X-G. Dengue Fever in Mainland China. Am J Trop Med Hyg . 2010;83(3):664–71.

4. Crawford JE, Alves JM, Palmer WJ, Day JP, Sylla M, Ramasamy R, et al. Population genomics reveals that an anthropophilic population of *Aedes aegypti* mosquitoes in West Africa recently gave rise to American and Asian populations of this major disease vector. BMC Biol . 2017;15(1):16.

5. Kotsakiozi P, Gloria-Soria A, Schaffner F, Robert V, Powell JR. *Aedes aegypti* in the Black Sea: recent introduction or ancient remnant? Parasit Vectors . 2018;11(1):396.

6. Hawley WA. The biology of *Aedes albopictus*. J Am Mosq Control Assoc Suppl . 1988;1:1–39.

7. Goubert C, Minard G, Vieira C, Boulesteix M. Population genetics of the Asian tiger mosquito *Aedes albopictus*, an invasive vector of human diseases. Heredity (Edinb) . 2016;117(3):125–34.

8. Porretta D, Mastrantonio V, Bellini R, Somboon P, Urbanelli S. Glacial History of a Modern Invader: Phylogeography and Species Distribution Modelling of the Asian Tiger Mosquito *Aedes albopictus*. PLoS One . 2012;7(9):e44515.

9. Bonizzoni M, Gasperi G, Chen X, James AA. The invasive mosquito species *Aedes albopictus*: Current knowledge and future perspectives. Trends Parasitol . 2013;29(9):460–8.

10. Hengeveld R. Dynamics of biological invasions. London: Chapman & Hall; 1989.

11. Harrington LC, Scott TW, Lerdthusnee K, Coleman RC, Costero A, Clark GG, et al. Dispersal of the dengue vector *Aedes aegypti* within and between rural communities. Am J Trop Med Hyg . 2005;72(2):209–20.

12. Jasper M, Schmidt TL, Ahmad NW, Sinkins SP, Hoffmann AA. A genomic approach to inferring kinship reveals limited intergenerational dispersal in the yellow fever mosquito. Mol Ecol Resour . 2019;1–11.

13. Marini F, Caputo B, Pombi M, Tarsitani G, Della Torre A. Study of *Aedes albopictus* dispersal in Rome, Italy, using sticky traps in mark-release-recapture experiments. Med Vet Entomol . 2010;24(4):361–8.

14. Eritja R, Palmer JRB, Roiz D, Sanpera-Calbet I, Bartumeus F. Direct Evidence of Adult *Aedes albopictus* Dispersal by Car. Sci Rep . 2017;7(1):14399–413.

15. Schmidt TL, Filipović I, Hoffmann AA, Rašić G. Fine-scale landscape genomics helps explain the slow spatial spread of *Wolbachia* through the *Aedes aegypti* population in Cairns, Australia. Heredity (Edinb). 2018;120(5):386–95.

16. Schmidt TL, Rašić G, Zhang D, Zheng X, Xi Z, Hoffmann AA. Genome-wide SNPs reveal the drivers of gene flow in an urban population of the Asian Tiger Mosquito, *Aedes albopictus*. PLoS Negl Trop Dis . 2017;11(10):e0006009.

17. Tatem AJ, Hay SI, Rogers DJ. Global traffic and disease vector dispersal. Proc Natl Acad Sci U S A . 2006;103(16):6242–7.

18. Ammar SE, Mclntyre M, Swan T, Kasper J, Derraik JGB, Baker MG, et al. Intercepted Mosquitoes at New Zealand's Ports of Entry, 2001 to 2018: Current Status and Future Concerns. Trop Med Infect Dis . 2019;4(3):101–18.

19. Schmidt TL, van Rooyen AR, Chung J, Endersby-Harshman NM, Griffin PC, Sly A, et al. Tracking genetic invasions: Genome-wide single nucleotide polymorphisms reveal the source of pyrethroid-resistant *Aedes aegypti* (yellow fever mosquito) incursions at international ports. Evol Appl . 2019;12(6):1136–46.

20. Gloria-Soria A, Ayala D, Bheecarry A, Calderon-Arguedas O, Chadee DD, Chiappero M, et al. Global genetic diversity of *Aedes aegypti*. Mol Ecol . 2016;25(21):5377–95.

21. Kraemer MUG, Sinka ME, Duda KA, Mylne AQN, Shearer FM, Barker CM, et al. The global distribution of the arbovirus vectors *Aedes aegypti* and *Ae. albopictus*. Elife . 2015;4:e08347.

22. Kraemer MUG, Reiner RC, Brady OJ, Messina JP, Gilbert M, Pigott DM, et al. Past and future spread of the arbovirus vectors *Aedes aegypti* and *Aedes albopictus*. Nat Microbiol . 2019;4(5):854-863.

23. McMillen-Jackson AL, Bert TM. Disparate patterns of population genetic structure and population history in two sympatric penaeid shrimp species (*Farfantepenaeus aztecus* and *Litopenaeus setiferus*) in the eastern United States. Mol Ecol . 2003;12(11):2895–905.

24. Avise JC. The history and purview of phylogeography: A personal reflection. Mol Ecol . 1998;7(4):371–9.

25. Joseph L, Moritz C, Hugall A. Molecular support for vicariance as a source of diversity in rainforest. Proc R Soc London Ser B Biol Sci . 1995;260(1358):177–82.

26. Trewick SA, Wallis GP. Bridging the "Beech-gap": New Zealand invertebrate phylogeography implicates Pleistocene glaciation and Pliocene isolation. Evolution (N Y) . 2001;55(11):2170–80.

27. Wagner CE, McCune AR. Contrasting patterns of spatial genetic structure in sympatric rock-dwelling cichlid fishes. Evolution (N Y) . 2009;63(5):1312–26.

28. Dawson MN. Parallel phylogeographic structure in ecologically similar sympatric sister taxa. Mol Ecol . 2012;21(4):987–1004.

29. Slatkin M. Gene Flow and the Geographic Structure of Natural Populations. Source Sci

New Ser . 1987;236(4803):787–92.

30. Excoffier L, Ray N. Surfing during population expansions promotes genetic revolutions and structuration. Trends Ecol Evol . 2008;23(7):347–51.

31. Waters JM, Fraser CI, Hewitt GM. Founder takes all: density-dependent processes structure biodiversity. Trends Ecol Evol . 2013;28:78–85.

32. Armbruster PA. Photoperiodic Diapause and the Establishment of *Aedes albopictus* (Diptera: Culicidae) in North America. J Med Entomol . 2016;53(5):1013–23.

33. Medley KA, Westby KM, Jenkins DG. Rapid local adaptation to northern winters in the invasive Asian tiger mosquito *Aedes albopictus*: A moving target. J Appl Ecol . 2019;1365-2664.13480.

34. Diniz DFA, De Albuquerque CMR, Oliva LO, De Melo-Santos MAV, Ayres CFJ. Diapause and quiescence: Dormancy mechanisms that contribute to the geographical expansion of mosquitoes and their evolutionary success. Parasit Vectors . 2017;10(1):310–22.

35. Urbanski JM, Benoit JB, Michaud MR, Denlinger DL, Armbruster P. The molecular physiology of increased egg desiccation resistance during diapause in the invasive mosquito, *Aedes albopictus*. Proc R Soc B Biol Sci . 2010;277(1694):2683–92.

36. Halpern BS, Walbridge S, Selkoe KA, Kappel C V, Micheli F, D'Agrosa C, et al. A global map of human impact on marine ecosystems. Science . 2008;319(5865):948–52.

37. Mousson L, Dauga C, Garrigues T, Schaffner F, Vazeille M, Failloux AB. Phylogeography of *Aedes* (*Stegomyia*) *aegypti* (L.) and *Aedes* (*Stegomyia*) *albopictus* (Skuse) (Diptera: Culicidae) based on mitochondrial DNA variations. Genet Res . 2005;86(1):1–11.

38. Eskildsen GA, Rovira JR, Smith O, Miller MJ, Bennett KL, McMillan WO, et al. Maternal invasion history of *Aedes aegypti* and *Aedes albopictus* into the Isthmus of Panama: implications for the control of emergent viral disease agents. PLoS One. 2018;13(3):e0194874.

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

40. Wright S. Isolation by Distance. Genetics . 1943;28(2):114–38.

41. Wang IJ, Bradburd GS. Isolation by environment. Mol Ecol . 2014;23(23):5649–62.

42. Kotsakiozi P, Richardson JB, Pichler V, Favia G, Martins AJ, Urbanelli S, et al. Population genomics of the Asian tiger mosquito, *Aedes albopictus* : insights into the recent worldwide invasion. Ecol Evol . 2017;7(23):10143–57.

43. Pichler V, Kotsakiozi P, Caputo B, Serini P, Caccone A, Torre A Della. Complex interplay of evolutionary forces shaping population genomic structure of invasive *Aedes albopictus* in southern Europe. PLoS Negl Trop Dis. 2019;13(8).

44. Sherpa S, Rioux D, Pougnet-Lagarde C, Després L. Genetic diversity and distribution differ between long-established and recently introduced populations in the invasive mosquito *Aedes albopictus*. Infect Genet Evol . 2018 ;58:145–56.

45. Sherpa S, Blum MGB, Capblancq T, Cumer T, Rioux D, Després L. Unravelling the invasion history of the Asian tiger mosquito in Europe. Mol Ecol . 2019;28(9):2360–77.

46. Hamelin RC, Roe AD. Genomic biosurveillance of forest invasive alien enemies: a story written in code. Evol Appl . 2019;eva.12853.

47. Garcia G de A, Sylvestre G, Aguiar R, da Costa GB, Martins AJ, Lima JBP, et al. Matching the genetics of released and local *Aedes aegypti* populations is critical to assure *Wolbachia* invasion. PLoS Negl Trop Dis . 2019;13(1):e0007023.

48. Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. Stacks: an analysis tool set for population genomics. Mol Ecol . 2013;22(11):3124–40.

49. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods . 2012;9(4):357–9.

50. Behura SK, Lobo NF, Haas B, DeBruyn B, Lovin DD, Shumway MF, et al. Complete sequences of mitochondria genomes of *Aedes aegypti* and *Culex quinquefasciatus* and comparative analysis of mitochondrial DNA fragments inserted in the nuclear genomes. Insect Biochem Mol Biol . 2011;41(10):770–7.

51. Zhang H, Xing D, Wang G, Li C, Zhao T. Sequencing and analysis of the complete mitochondrial genome of *Aedes albopictus* (Diptera: Culicidae) in China. Mitochondrial DNA . 2016;27(4):2787–8.

52. Matthews BJ, Dudchenko O, Kingan SB, Koren S, Antoshechkin I, Crawford JE, et al. Improved reference genome of *Aedes aegypti* informs arbovirus vector control. Nature . 2018;1.

53. Chen X-G, Jiang X, Gu J, Xu M, Wu Y, Deng Y, et al. Genome sequence of the Asian Tiger mosquito, *Aedes albopictus*, reveals insights into its biology, genetics, and evolution. Proc Natl Acad Sci . 2015;112(44):E5907–15.

54. Hlaing T, Tun-Lin W, Somboon P, Socheat D, Setha T, Min S, et al. Mitochondrial pseudogenes in the nuclear genome of *Aedes aegypti* mosquitoes: Implications for past and future population genetic studies. BMC Genet . 2009;10(1):11.

55. Jombart T, Devillard S, Balloux F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet . 2010 Oct 15;11(1):94–108.

56. Frichot E, Mathieu F, Trouillon T, Bouchard G, François O. Fast and efficient estimation of individual ancestry coefficients. Genetics. 2014 Apr 1;196(4):973–83.

57. Diaz-Papkovich A, Anderson-Trocme L, Gravel S. Revealing multi-scale population structure in large cohorts. bioRxiv . 2019;423632. Available from: https://www.biorxiv.org/content/10.1101/423632v2.abstract

58. McInnes L, Healy J, Melville J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. bioRxiv . 2018; Available from: http://arxiv.org/abs/1802.03426

59. Browning BL, Browning SR. Genotype Imputation with Millions of Reference Samples. Am J Hum Genet . 2016 Jan 7 [cited 2019 Aug 7];98(1):116–26.

60. Jombart T. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics . 2008;24(11):1403–5.

61. Frichot E, François O. LEA : An R package for landscape and ecological association studies. O'Meara B, editor. Methods Ecol Evol . 2015;6(8):925–9.

62. Willing E-M, Dreyer C, van Oosterhout C. Estimates of Genetic Differentiation Measured by FST Do Not Necessarily Require Large Sample Sizes When Using Many SNP Markers. PLoS One . 2012;7(8):e42649.

63. Meirmans PG. Using the AMOVA framwwork to estimate a standardized genetic differentiation measure. Evolution (N Y) . 2006;60(11):2399–402.

64. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. Bioinformatics . 2011;27(15):2156–8.

65. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, et al. vegan: Community Ecology Package . 2019. p. 2395–6.

66. ESRI. ArcGIS. Redlands, CA: Environmental Systems Research Institute; 2017.

67. Powell JR, Gloria-Soria A, Kotsakiozi P. Recent History of *Aedes aegypti* : Vector Genomics and Epidemiology Records. Bioscience . 2018;68(11):854–60.

68. Kuno G. Letter to the Editor. Bioscience . 2019;69(3):161–161.

69. Powell JR, Gloria-Soria A, Kotsakiozi P. Response to Kuno. Bioscience . 2019;69(3):161–2.

70. Guillaumot L, Ofanoa R, Swillen L, Singh N, Bossin HC, Schaffner F. Distribution of *Aedes albopictus* (Diptera, Culicidae) in southwestern Pacific countries, with a first report from the Kingdom of Tonga. Parasites and Vectors. 2012;5(1):247.

71. Xia D, Guo X, Hu T, Li L, Teng PY, Yin QQ, et al. Photoperiodic diapause in a subtropical population of *Aedes albopictus* in Guangzhou, China: Optimized field-laboratory-based study and statistical models for comprehensive characterization. Infect Dis Poverty . 2018;7(1):89.

72. Manni M, Guglielmino CR, Scolari F, Vega-Rúa A, Failloux A-B, Somboon P, et al. Genetic evidence for a worldwide chaotic dispersion pattern of the arbovirus vector, *Aedes albopictus*. PLoS Negl Trop Dis . 2017;11(1):e0005332.

73. Gillespie RG, Baldwin BG, Waters JM, Fraser CI, Nikula R, Roderick GK. Long-distance dispersal: a framework for hypothesis testing. Trends Ecol Evol . 2012;27(1):47–56.

74. Urbanelli S, Bellini R, Carrieri M, Sallicandro P, Celli G. Population structure of *Aedes albopictus* (Skuse): The mosquito which is colonizing Mediterranean countries. Heredity (Edinb) . 2000;84(3):331–7.

75. Ardika IW, Bellwood P. Sembiran: the beginnings of Indian contact with Bali. Antiquity . 1991;65(247):221–32.

76. Guo Y, Song Z, Luo L, Wang Q, Zhou G, Yang D, et al. Molecular evidence for new sympatric cryptic species of *Aedes albopictus* (Diptera: Culicidae) in China: A new threat from *Aedes albopictus* subgroup? Parasit Vectors . 2018;11(1):228–41.

Supporting Information Legends

S1 Appendix

This appendix contains additional information on sample processing and filtering, from the DNA extraction stage onwards.

S1 Tables A – E

These tables show upper and lower confidence ranges of: (A) pairwise F_{ST}' for *Ae. aegypti* in the Indo-Pacific region; (B) pairwise F_{ST}' for *Ae. albopictus* in the Indo-Pacific region; (C) heterozygosity and nucleotide diversity for *Ae. aegypti* in the Indo-Pacific region; (D) heterozygosity and nucleotide diversity for *Ae. albopictus* in the Indo-Pacific region; and (E) comparison of pairwise F_{ST}' using 5 and 16 genotypes.

S1 Fig A DAPCs of *Ae. aegypti* (i; $K = 15$) and *Ae. albopictus* (ii; $K = 11$) using lowest AIC.

Colours indicate cluster membership. Circles are sized relative to population sample size. The map uses a Mollweide projection with a central meridian of 120°E. The map was produced in ARCMAP 10.5.1, using shapefiles made available by www.naturalearthdata.com.

Author contributions

Conceptualization: TLS, AAH, A-CH

Data Curation: TLS, JC

Formal Analysis: TLS, JC

Funding Acquisition: ARW, AAH

Investigation: TLS, JC, A-CH

Methodology: TLS, JC

Project Administration: TLS, ARW, AAH

Resources: TLS, A-CH, ARW, AAH

Software: TLS, JC

Supervision: AAH

Validation: TLS, JC

Visualization: TLS, JC

Writing – Original Draft Preparation: TLS

Writing – Review & Editing: TLS, JC, A-CH, ARW, AAH