MOLECULAR ECOLOGY

Molecular Ecology (2016) **25**, 5377–5395 doi: 10.1111/mec.13866

Global genetic diversity of *Aedes aegypti*

ANDREA GLORIA-SORIA,* DIEGO AYALA,†‡ AMBICADUTT BHEECARRY,§ OLGER CALDERON-ARGUEDAS,¶ DAVE D. CHADEE,** MARINA CHIAPPERO,†† MAUREEN COETZEE,‡‡ KHOUAILDI BIN ELAHEE,§ ILDEFONSO FERNANDEZ-SALAS,§§ HANY A. KAMAL,¶¶ BASILE KAMGANG,*** EMAD I. M. KHATER,††† LAURA D. KRAMER,‡‡‡ VICKI KRAMER,§§§ ALMA LOPEZ-SOLIS,§§ JOEL LUTOMIAH,¶¶¶ ADEMIR MARTINS JR,**** MARIA VICTORIA MICIELI,†††† CHRISTOPHE PAUPY,†‡ ALONGKOT PONLAWAT,‡‡‡‡ NIL RAHOLA,†‡ SYED BASIT RASHEED,§§§§ JOSHUA B. RICHARDSON,* AMAG A. SALEH,††† ROSA MARIA SANCHEZ-CASAS,¶¶¶¶ GONÇALO SEIXAS,***** CARLA A. SOUSA,***** WALTER J. TABACHNICK,†††††† ADRIANA TROYO¶ and JEFFREY R. POWELL*

**Yale University, 21 Sachem Street, New Haven, CT 06520-8105, USA, †Laboratory MIVEGEC, Institut de Recherche pour le Développement, Montpellier 34394, France, ‡Centre International de Recherches Médicales de Franceville, Franceville, Gabon, §Vector Biology and Control Division, Ministry of Health and Quality of Life, Mauritius, Mauritius, ¶Facultad de Microbiología, Centro de Investigación en Enfermedades Tropicales, Universidad de Costa Rica, San José, Costa Rica, **Department of Life Sciences, University of the West Indies, St. Augustine, Trinidad, WI, ††Instituto de Diversidad y Ecología Animal, Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) y Universidad Nacional de Córdoba, Av. Vélez Sarsfield 299, X5000JJC, Córdoba, Argentina, ‡‡School of Pathology, Wits Research Institute for Malaria, University of Witwatersrand, Johannesburg, South Africa, §§Centro Regional de Investigación en Salud Pública INSP, Tapachula, Chiapas, Mexico, ¶¶Dallah Establishment, Pest Control Projects, Jeddah, Kingdom of Saudi Arabia, ***Research Unit Liverpool School of Tropical Medicine, Oganisation de Coordination pour la lute contre les Endemies en Afrique Centrale, Yaounde, Cameroon, †††Department of Plant Protection, College of Food and Agriculture Sciences, King Saud University, Riyadh, Kingdom of Saudi Arabia, ‡‡‡Wadsworth Center, New York State Department of Health, School of Public Health, State University of New York at Albany, Albany, NY, USA, §§§Vector Borne Disease Section, California Department of Public Health, Sacramento, CA, USA, ¶¶¶Arbovirus/Viral Hemorrhagic Fever Laboratory, Center for Virus Research, Kenya Medical Research Institute (KEMRI), P. O. Box 54628-00200, Nairobi, Kenya, ****Laboratório de Fisiologia e Controle de Artrópodes Vetores, IOC-FIOCRUZ, Rio de Janeiro, Brazil, ††††Centro de Estudios Parasitológicos y de Vectores, CONICET, La Plata, Buenos Aires, Argentina, ‡‡‡‡Department of Entomology, USAMD-AFRIMS, Bangkok, Thailand, §§§§Department of Zoology, University of Peshawar, Peshawar 25120, Pakistan, ¶¶¶¶School of Veterinary Medicine, Escobedo, Centro de Investigación y Desarrollo en Ciencias de la Salud, Monterrey, Nuevo León, Mexico, *****Global Health and Tropical Medicine, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Rua da Junqueira 100, 1349-008, Lisbon, Portugal, †††††Florida Medical Entomology Laboratory, Department of Entomology and Nematology, University of Florida, IFAS, Vero Beach, FL, USA*

Abstract

Mosquitoes, especially *Aedes aegypti*, are becoming important models for studying invasion biology. We characterized genetic variation at 12 microsatellite loci in 79 populations of *Ae. aegypti* from 30 countries in six continents, and used them to infer historical and modern patterns of invasion. Our results support the two subspecies *Ae. aegypti formosus* and *Ae. aegypti aegypti* as genetically distinct units. *Ae. aegypti aegypti* populations outside Africa are derived from ancestral African populations and are monophyletic. The two subspecies co-occur in both East Africa (Kenya) and West Africa (Senegal). In rural/forest settings (Rabai District of Kenya), the two subspecies remain genetically distinct, whereas in urban settings, they introgress freely. Populations outside Africa are highly genetically structured likely due to a combination of recent founder effects, discrete discontinuous habitats and low migration rates. Ancestral populations in sub-Saharan Africa are less genetically structured, as are the

Correspondence: Andrea Gloria-Soria, Fax: 203 432 7394;
E-mail: andrea.gloria-soria@yale.edu

populations in Asia. Introduction of *Ae. aegypti* to the New World coinciding with trans-Atlantic shipping in the 16th to 18th centuries was followed by its introduction to Asia in the late 19th century from the New World or from now extinct populations in the Mediterranean Basin. *Aedes mascarensis* is a genetically distinct sister species to *Ae. aegypti* s.l. This study provides a reference database of genetic diversity that can be used to determine the likely origin of new introductions that occur regularly for this invasive species. The genetic uniqueness of many populations and regions has important implications for attempts to control *Ae. aegypti*, especially for the methods using genetic modification of populations.

Keywords: *Aedes aegypti*, *Aedes mascarensis*, history, invasion, microsatellites

Received 16 May 2016; revision received 2 September 2016; accepted 6 September 2016

Introduction

Non-native invasive species are of great concern for ecology, conservation biology, agriculture and epidemiology. Invasive species can alter ecosystems by competing for resources with endemic counterparts in the same feeding guild, preying on native species, or transmitting pathogens and disease. Invasive species can also destabilize an ecosystem by displacing native populations while not taking over their ecological services (Mack *et al.* 2000). In the case of arthropod vectors of disease agents, an invasion can lead to the introduction of a novel disease to a naïve population or a more efficient spread of local pathogens. Mosquitoes are insect invaders of major relevance because of their connection to human disease (Lounibos 2002). Most medically important mosquitoes live in close proximity to humans and arrive to new territories with human migrations. Adaptation to human habitats almost invariably leads to the evolution of preference for the most available blood source: humans. Pathogens originally from nonhuman hosts (usually other primates) then enter the human–mosquito cycle of transmission and can have major public health consequences.

Aedes aegypti is the invasive mosquito that has caused the most human casualties worldwide, initially as the vector of devastating yellow fever epidemics, hence its common name 'the yellow fever mosquito'. Today, *Ae. aegypti* continues to plague humans as the primary vector of viruses that cause dengue fever, chikungunya and Zika. Because of its ease of rearing in the laboratory as well as its major epidemiological role, *Ae. aegypti* is the best known mosquito from all aspects of its biology (Christophers 1960; Clemens 1992, 1999) and has become a model organism over the last 15 years, thanks to an increase in molecular studies focused on disease vectors. The present report adds significantly to understanding the population genetics of this species, elucidates global invasion pathways and informs emerging management options.

Three subspecies of *Ae. aegypti* have been described. The type subspecies is *Ae. aegypti aegypti* (abbreviated Aaa) that has spread throughout the tropical and subtropical world by humans (Powell & Tabachnick 2013), is highly anthropophilic (prefers human bloodmeals; McBride *et al.* 2014) and is adapted to breed in human habitats ('domestic'). The ancestral form of the species in sub-Saharan Africa, *Ae. aegypti formosus* (Aaf), breeds in nonhuman-disturbed habitats such as forests and vegetated ecotones (Lounibos 1981) and prefers nonhuman bloodmeals (McBride *et al.* 2014). These subspecies were originally described based largely on their geographic distribution, colour and scaling patterns, the forest form Aaf being blacker with less white scaling than the brownish domestic form Aaa. However, populations are highly variable for scaling pattern (McClelland 1974; Jupp *et al.* 1991), so morphology does not always reflect the major ecological distinction between the two subspecies (Powell & Tabachnick 2013; and see below). A third subspecies, *Ae. aegypti queenslandensis*, was named by Theobald (1901) and described as a variety of *Ae. aegypti* with golden brown scales in the thorax. Mattingly (1967) suggested that populations in the Mediterranean Basin and parts of Australia and Eastern Africa conform to this subspecies or variety. It is unclear whether members of this subspecies can still be found.

Given the ambiguity of morphology matching ecological and behavioural traits, especially in sub-Saharan Africa, the precision and usefulness of subspecific designations have been questioned (Powell & Tabachnick 2013). We use the terms 'Aaa' and 'Aaf' here largely for historical continuity recognizing the limitations and sometimes ambiguity of trying to apply this dichotomy. As McClelland (1967) correctly pointed out, *Ae. aegypti* s.l. is more accurately viewed as a highly polymorphic rather than a polytypic species.

Here, we present the latest results of our ongoing studies on the genetic diversity of *Ae. aegypti*. While several previous studies have addressed the large-scale genetic patterns of *Ae. aegypti* population structure and historical movements (e.g. Brown *et al.* 2011a, 2014;

Bennett *et al.* 2016), we have genotyped a larger array of samples encompassing much of the *Aedes aegypti* geographic range, including six continents (Fig. 1). This extended sampling provides higher resolution of the genetic structure at the regional scale and allows more accurate tests of hypotheses regarding the historical movement of *Ae. aegypti* out of Africa, because each region is better represented in the data set. What we present here in no way contradicts the earlier work, but strengthens the conclusions of those studies and adds more detail. We emphasize five aspects of these comprehensive data: (i) the distinct genetic differentiation patterns observed between populations of *Ae. aegypti* in the ancestral range of Africa and populations outside Africa; (ii) what the data imply about the historical spread of *Ae. aegypti*; (iii) how these data can serve as a reference panel for determining the origin of new introductions; (iv) the validity of *Ae. mascarensis* as a genetically distinct sister species to *Ae. aegypti* s.l.; and (v) the importance of recognizing the extensive genetic variation in controlling diseases vectored by this mosquito.

Materials and methods

Mosquito collections

Aedes aegypti: Adults, larvae or eggs were received from 79 geographic locations worldwide (Table 1 and Fig. 1). Allele frequencies from some of the earlier collections have been previously reported in Brown *et al.* (2011a), Gloria-Soria *et al.* (2014) and Monteiro *et al.* (2014); because all these studies were performed in the same laboratory at Yale University, designation of alleles is consistent across this and all previous work reported from this laboratory. Mosquitoes arrived as either eggs from oviposition traps, larvae or adults in 70–100% ethanol or on silica gel. Eggs were hatched at the Yale School of Epidemiology and Public Health insectary, reared to adults and preserved in 100% ethanol at −20 °C until DNA extraction. Most mosquitoes included in this study came directly from the field, when possible; in a few cases, they were passed through one or two generations in the laboratory prior to genetic analyses: Tahiti (French Polynesia) and Pijijiapan (Mexico) – one generation; Bolivar (Venezuela), Zulia (Venezuela), Rayong (Thailand) and Prachuabkirikhan (Thailand) – two generations. These laboratory colonies were established in large cages by several hundred to over 1000 mosquitoes from the field in an attempt to be as representative as possible of field populations (Brown *et al.* 2011a). Specimens from Rabai, Kenya, were sampled as larvae from water stored in large clay vessels inside huts ('Rabai-in') and from the forests ('Rabai-out') a few hundred metres distant to the village. Previous studies have shown behavioural, morphological and genetic differentiation among these populations (Trpis & Hausermann 1975; Tabachnick *et al.* 1979; Brown *et al.* 2011a, 2014; McBride *et al.* 2014) and thus we kept them separate for the purposes of this work.

A strain presumed to be *Ae. ae. queenslandensis* was obtained from Professor David Severson (University of Notre Dame) in ethanol. This strain originated from Surabaya, Indonesia, and had undergone several

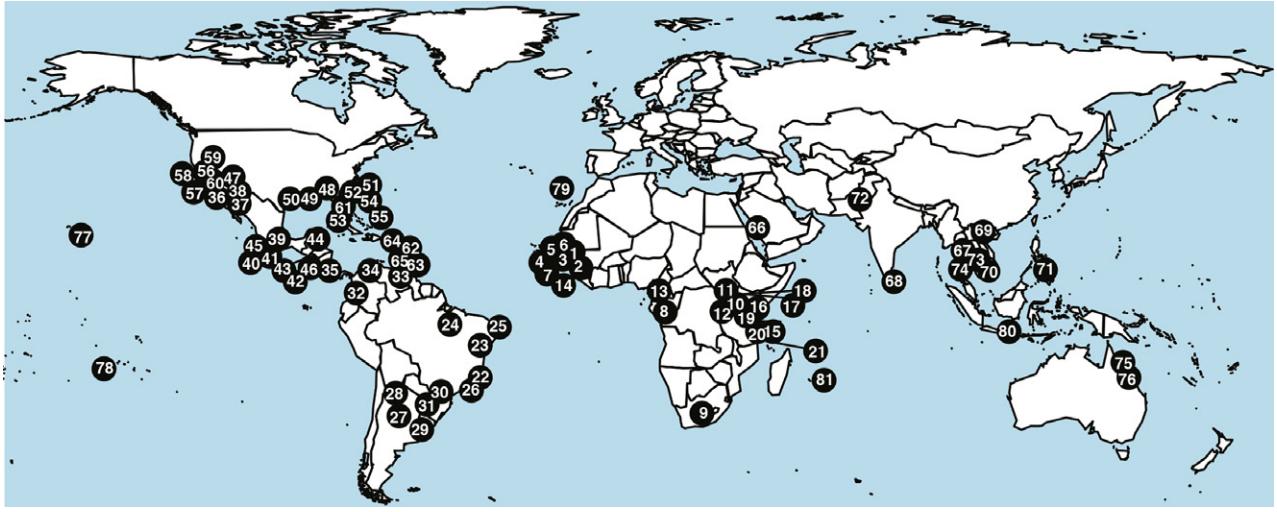

Fig. 1 Sampling locations of *Aedes aegypti* and *Aedes mascarensis* collections used in this study. Population codes are as labelled in Table 1. Putative *Ae. aegypti queenslandensis* is indicated as population 80 and *Ae. mascarensis* as population 81. Approximate locations are displayed in order to accommodate all labels. [Colour figure can be viewed at wileyonlinelibrary.com].

Table 1 *Aedes aegypti* collections included in this study

Code*	Population†	Region	Year‡	N§
1	Ngari, SN	Africa	2012	16
2	Goudiry, SN	Africa	2012	54
3	Sedhiou, SN	Africa	2012	54
4	N'goye, SN	Africa	2007	45
5	Dakar, SN	Africa	2005	44
6	PK-10, Kedougou, SN	Africa	2006	48
7	Koungheul, SN	Africa	2006	46
8	Francesville, GA	Africa	2014	54
9	Johannesburg, ZA	Africa	2015	18
10	Lunyo, UG	Africa	2013	52
11	Bundibugyo, UG	Africa	2009	47
12	Kichwamba, UG	Africa	2009	48
13	Yaounde, CM	Africa	2014	54
14	Bijagos, GW	Africa	2009	30
15	Mombasa, KE	Africa	2012	85
16	Garissa, KE	Africa	2012	19
17	Kakamega, KE	Africa	2012	8
18	Kisumu, KE	Africa	2012	34
19	Nairobi, KE	Africa	2012	54
20	Rabai-out, KE	Africa	2006, 2009	53
21	Rabai-in, KE	Africa	2006, 2009	55
22	Cachoeiro, BR	South America	2012	47
23	Jacobina, BR	South America	2013	94
24	Maraba, BR	South America	2010	48
25	Natal, BR	South America	2010	47
26	Rio de Janeiro, BR	South America	2014	39
27	Cordoba, AR	South America	2003	26
28	Salta, AR	South America	2006, 2014	27
29	La Plata, AR	South America	2014	10
30	Iguazu, AR	South America	2014	10
31	Posadas, AR	South America	2014	10
32	Cali, CO	South America	2013	80
33	Bolivar, VEN	South America	2004	48
34	Zulia, VEN	South America	2004	47
35	Siquirres, CR	Central America	2014	51
36	Tijuana, BC, MEX	North America	2013	20
37	Hermosillo, SON, MEX	North America	2013	50
38	Nogales, SON, MEX	North America	2013	51
39	Las Palomas, GRO, MEX	North America	2012	54
40	Lomas de Zapatero, GRO, MEX	North America	2012	51
41	Amacuzac, MOR, MEX	North America	2014	53
42	Mazatan, CHP MEX	North America	2012	45
43	Pijijiapan, CHP, MEX	North America	2008	47
44	Chetumal, QRO, MEX	North America	2013	54
45	Iguala, GRO, MEX	North America	2012	54
46	Tapachula Norte, CHP, MEX	North America	2012	54
47	Maricopa County, AZ	North America	2013	53
48	Musco, GA, USA	North America	2012	55
49	New Orleans, LA, USA	North America	2011	63
50	Houston, US	North America	2011	19
51	Miami, FL, USA	North America	2011	47
52	Rio, FL, USA	North America	2014	51
53	North Key West, FL, USA	North America	2013	52
54	Vaca Key, FL, USA	North America	2009	42
55	Conch Key, FL, USA	North America	2006	42
56	Madera, CA, USA	North America	2013	50
57	Clovis, CA, USA	North America	2013	60
58	San Mateo, CA, USA	North America	2013	21
59	Exeter, CA, USA	North America	2014	23
60	Los Angeles, CA, USA	North America	2014	6
61	Palm Beach, FL, USA	Caribbean	2006	42
62	Dominica, DM	Caribbean	2009	95
63	Trinidad, TT	Caribbean	2014	51
64	Patillas, PR	Caribbean	2014	54
65	Carriacou, GD	Caribbean	2015	29
66	Jeddah, SA	Asia	2012	84
67	Bangkok, TH	Asia	2013	54
68	Sri Lanka, SL	Asia	2014	7
69	Hanoi, VT	Asia	2013	54
70	Ho Chih Minh, VT	Asia	2013	54
71	Cebu City, PH	Asia	2013	108
72	Pakistan, PK	Asia	2010	49
73	Rayong, TH	Asia	2006	48
74	Prachuabkirikhan, TH	Asia	2009	47
75	Cairns, AU	Asia	2013	51
76	Townsville, AU	Asia	2009	47
77	Hawaii, USA	Pacific	2009	25
78	Tahiti, FP	Pacific	2010	48
79	Madeira, PO	Europe	2012	66
Total				3632

*Population identifier.
†Population includes the ISO 3166 two-letter geographic country code.
‡Year of collection.
§The number of individuals genotyped.

generations of artificial selection for the *queenslandensis*-scale pattern (Tsuda *et al.* 2003). *Aedes mascarensis* specimens were collected in the region known as Le Dauguet, in the suburb of Port-Louis, Capital of the Republic of Mauritius, in 2014.

DNA extraction and microsatellite genotyping

Total nucleic acids were extracted from 3682 individual *Ae. aegypti* and 26 *Ae. mascarensis* mosquitoes using the DNeasy Blood and Tissue kit (Qiagen) according to the manufacturer instructions, with an additional RNase A (Qiagen) step. Samples were stored at −20 °C until further analysis. Individual mosquitoes were genotyped as described in Brown *et al.* (2011a). The microsatellite loci analysed were as follows: A1, B2, B3, A9 (trinucleotide repeats) and AC2, CT2, AG2, AC4, AC1, AC5, AG1 and AG4 (dinucleotide repeats) (Slotman *et al.* 2007; Brown *et al.* 2011a). Polymerase chain reactions were conducted in 10-μL reaction mixture containing the Type-it

Microsatellite PCR Master Mix (Qiagen), 25 nM of each forward primer, 250 nM of each reverse primer and 500 nM of a fluorescently labelled M13 primer to allow multiplexing (Oetting *et al.* 1995; Brown *et al.* 2011a). Thermocycler conditions were as follows: 94oC × 10′, 35 × (94oC × 30″, 54oC × 30″, 72oC × 30″) and 72oC × 5′. Microsatellite primer sequences, multiplex pairings and fluorescent primers are as described in Brown *et al.* (2011a). PCR products were run for fragment analysis on an Applied Biosystems 3730xl DNA Genetic Analyser with a GS 500 Rox internal size standard (Applied Biosystems) at the DNA Analysis Facility at Science Hill at Yale University. Microsatellite alleles were scored using GENEMAPPER version 4.0 (Applied Biosystems). Raw allele frequencies are available at VectorBase.org, Population Biology Project ID: VBP0000138.

Genetic diversity

All microsatellite loci were analysed for within-population deviations from Hardy–Weinberg equilibrium (HWE) using the exact test (Weir & Cockerham 1984) with complete enumeration as implemented in GENEPOP version 4.1.0 (Raymond & Rousset 1995; Rousset 2008). Linkage disequilibrium (LD) among pairs of loci was estimated using the same software. Both tests were run with 10 000 dememorizations, 1000 batches and 10 000 iterations per batch. Bonferroni corrections were applied to the resulting matrices of both HWE and LD. Allele numbers, allelic frequencies and average observed (H_o) and expected (H_e) heterozygosities were estimated using GENALEX (Peakall & Smouse 2012). Allelic richness (AR) and private allelic richness (PAR) were calculated in HPRARE (Kalinowski 2005), which uses rarefaction to correct for unequal sample sizes ($N = 200$ for regional richness and $N = 100$ for individual population richness estimation). Pairwise genetic distances (F_{ST}) were calculated in GENODIVE 2.0b.27 (Meirmans & Van Tienderen 2004).

Population structure

Geographic and temporal population structure was evaluated via the Bayesian clustering method implemented by the software STRUCTURE version 2.3 (Pritchard *et al.* 2000), which identifies genetic clusters and assigns individuals to these clusters with no *a priori* information of sample location. The most likely number of clusters (K) was determined by conducting 20 independent runs from each $K = 1$ to 5 at the subspecies/species and global scale and from $K = 1$ to 10 for the subsequent analysis of each of the population subgroups. Each run assumed an admixture model and correlated allele frequencies using a burn-in value of 100 000 iterations followed by 500 000 repetitions. The optimal number of K clusters was determined both following the guidelines of Pritchard *et al.* (2000) and the delta K method from Evanno *et al.* (2005) with the online version of STRUCTURE HARVESTER version 0.6.94 (Earl & vonHoldt 2012). Plots of the most biologically informative number of clusters were generated with the program DISTRUCT version1.1 (Rosenberg 2004). In most cases, these plots correspond to the optimal K value as identified by the Evanno method (Evanno *et al.* 2005), but exceptions are noted in the figures. Discriminant analysis of principal components (DAPC; Jombart *et al.* 2010) was performed on allele frequencies of the same groups evaluated by STRUCTURE and plotted with the ADEGENET package (Jombart 2008) in R version 3.2.2 (R Core Team 2013).

Scale of geographic genetic differentiation

Correlation between the geographic distance (Euclidean distance in km) and genetic distance (F_{ST}) matrices (often referred to as 'isolation by distance') was made on each of the following population groups: Africa (*Ae. formosus* only), Asia, continental North America, South America and the Caribbean islands (including the islands from the Florida Keys). Central America was excluded from the analysis, because only one sample from Costa Rica was sampled from this region. Distance matrices were generated in GENEPOP version 4.1.0 (Raymond & Rousset 1995; Rousset 2008), and geographic distances were created from geographic coordinates using the GEOGRAPHIC DISTANCE MATRIX GENERATOR version 1.2.3 (Ersts 2016). Mantel tests were conducted with the *ade4* package (R Core Team 2013) using 9999 permutations and plotted with the same software.

Estimates of demographic parameters and population history

Inferences of effective population sizes and introduction history were made using approximate Bayesian computation methods (ABC: Beaumont *et al.* 2002) as implemented by DIYABC version 2.0.4 (Cornuet *et al.* 2014). Four plausible scenarios of migration among continents were tested, and a separate analysis compared six scenarios of the origin of the indoor Aaa from Rabai, Kenya. An estimated number of 10 generations per year was chosen based on a life cycle of ~22 days from hatch to hatch of *Ae. aegypti* reared under ideal constant temperature (28 °C) and humidity conditions (85%) in our insectary, with nonoverlapping generations. This would result in ~16 continuous generations/year. If we consider that in the field temperature is not constant, neither are nutrients or humidity, that predators are

abundant and that lifespan is shortened by human intervention, the number of generations a year will be lower than under ideal conditions. Additionally, in some locations (e.g. Georgia, USA), weather conditions prevent mosquito breeding during the winter months. Other independent studies have estimated that *Ae. aegypti* can undergo 9–37 generations a year in the laboratory based on its thermal requirements (Beserra *et al.* 2006; Marinho *et al.* 2016). Thus, we consider that a mean of 10 generations a year is a good, but conservative estimate. The origin of the Asian populations was modelled by randomly subsampling 200 individuals from each of the Africa (excluding Rabai – Kenya, and Goudiry – Senegal), America, Caribbean and Asia (including the Australian populations) data sets. The origin of Rabai (Kenya) Aaa was subsequently investigated in a similar manner, by randomly drawing 32 individuals of each group to match the number of individuals available from Rabai (Kenya). Aaa from Rabai (Kenya) used for this analysis were exclusively individuals with the majority of Aaa ancestry, as determined by q values >50% in the STRUCTURE analysis (Pritchard *et al.* 2000). A mutation rate ranging from 9×10^{-6} to 1×10^{-5} was used based on the rates reported in the literature for other Diptera species (Schug *et al.* 1997; Pfeiler *et al.* 2013). Additional details of the DIYABC analysis can be found in Tables S1 and S2 (Supporting information).

Results

We have sampled and analysed 79 populations from six continents, across much of *Aedes aegypti*'s geographic range (Fig. 1). Details of these populations can be found in Tables 1 and S3 (Supporting information). Some localities have been sampled more than once, and we report the patterns of temporal stability/instability of these populations elsewhere (Gloria-Soria *et al.* 2016). All samples came directly from the field with the exception of six, which were reared no more than two generations in the laboratory (see Materials and Methods and Table S3, Supporting information). Mean and median sample sizes are 46 and 48 individuals, respectively.

Table 2 shows the allelic richness, private allelic richness and genetic distance (F_{ST}) estimates per continent. Allele richness in Africa is ~11 alleles, whereas the mean allele richness across continents outside Africa is 6.7 (North America, South America, Caribbean, Asia and Pacific). The number of private alleles in Africa is 3.4, more than an order of magnitude higher than the values for continents outside Africa. These differences are statistically significant by Wilcoxon rank sum tests at $P < 0.001$. Population genetic statistics for individual locations are reported in Tables S4 and S5 (Supporting information).

A total of 79 of 925 (8.5%) population-by-locus comparisons deviate significantly from Hardy–Weinberg expectations ($P < 0.05$, HW exact test) as indicated by F_{IS}, after sequential Bonferroni correction (Table S4, Supporting information). Such small deviations from HW are common for microsatellites most often due to rare null alleles. Of the 79 significant deviations, 67 had a positive F_{IS}, consistent with the excess of homozygotes expected in the presence of null alleles. A total of 313 of 5414 (6.0%) locus-by-locus tests for LD were significant at the 5% level after multiple test correction, with no locus pair consistently correlated across all populations. Each of the 12 microsatellites reported in this study is located on different supercontigs of the current *Ae. aegypti* genome assembly (*AaegL3* in VectorBase – Giraldo-Calderón *et al.* 2015). The low level of LD reported here (6.0% of tests significant at the 5% level) is consistent with the 12 loci being independent.

Table 2 Population genetics statistics of *Aedes aegypti* populations by continent

								Genetic distance (F_{ST})			
Region	*N*	Ho	μHe	AR (200)	PAR (200)	F_{ST} (intra)	vs. NA	vs. SA	vs. C	vs. As	vs. P
Africa	918	0.591	0.738	10.97	3.41	0.114	0.119	0.119	0.118	0.114	0.135
North America	952	0.548	0.635	6.36	0.14	0.149	—	0.053	0.017	0.065	0.132
South America	562	0.489	0.611	7.37	0.3	0.181		—	0.034	0.069	0.13
Caribbean	407	0.541	0.635	6.34	0.21	0.143			—	0.06	0.117
Asia	603	0.551	0.643	7.87	0.58	0.121				—	0.052
Pacific	73	0.557	0.597	4.75	0	0.123					—

N, sample size; Ho, observed heterozygosity; μHe, unbiased expected heterozygosity; AR and PAR, allele richness and private allelic richness estimated by rarefaction ($N = 200$ genes); F_{ST} (intra), average genetic distance between populations within the region; NA, North America; SA, South America; C, Caribbean; As, Asia; P, Pacific.

Global-scale patterns of population structure

Bayesian clustering analysis in STRUCTURE (Pritchard *et al.* 2000) and DAPC (Jombart *et al.* 2010) on the full data set of 79 *Ae. aegypti* collections from around the world (Fig. 1) support the existence of two major genetic clusters of *Ae. aegypti* s. l. (Fig. 2). The '*Ae. queenslandensis*' strain was excluded from these analyses for reasons described below. Generally, the two genetic clusters correspond to the two subspecies designations: *Aedes aegypti formosus* (Aaf) in sub-Saharan Africa and *Aedes aegypti aegypti* (Aaa) spread around the tropical and subtropical world outside Africa. Admixture between the two named subspecies was detected in four countries: Senegal (localities 2 and 4 in Fig. 2), Kenya (locality 15), Argentina (localities 27–30) and the USA (locality 59).

Differentiation within and outside Africa

Figure 3 shows the Bayesian clustering analysis of the African and out-of-Africa populations separately. Populations outside Africa are hierarchically genetically subdivided, first into three major groupings roughly representing North America, South America and Asia plus the Pacific (Fig. 3A). Caribbean populations are heterogeneous displaying a combination of affinities to North America and South America, with the exception of Dominica that clearly groups with South America (locality 62 in Fig. 3A). The Pacific region, which includes the two Australian populations, Tahiti (French Polynesia) and Hawaii (USA), clusters with Asia. Madeira, the only European sample in this data set, has both Asian and South American genetic affinities (locality 79 in Fig. 3A).

The strong population structure observed outside Africa contrasts to the relatively weak structure observed within Africa (Fig. 3B). Interestingly, while there is no indication of individual population distinctness, there is indication that Senegal populations are different from the rest of Africa (excluding the Rabai indoor collection; Fig. 3B). This difference in the degree of genetic structure within Africa compared with outside Africa is confirmed by analyses of genetic differentiation by distance (often called isolation by distance), in Fig. 4A. Populations separated by the same distance in the New World (Americas and the Caribbean) are more genetically differentiated than those in Africa. The case of Asia is discussed in the next section.

Genetic structure within continents/regions outside Africa

When each continent is analysed individually using STRUCTURE, finer geographic genetic structure is detected.

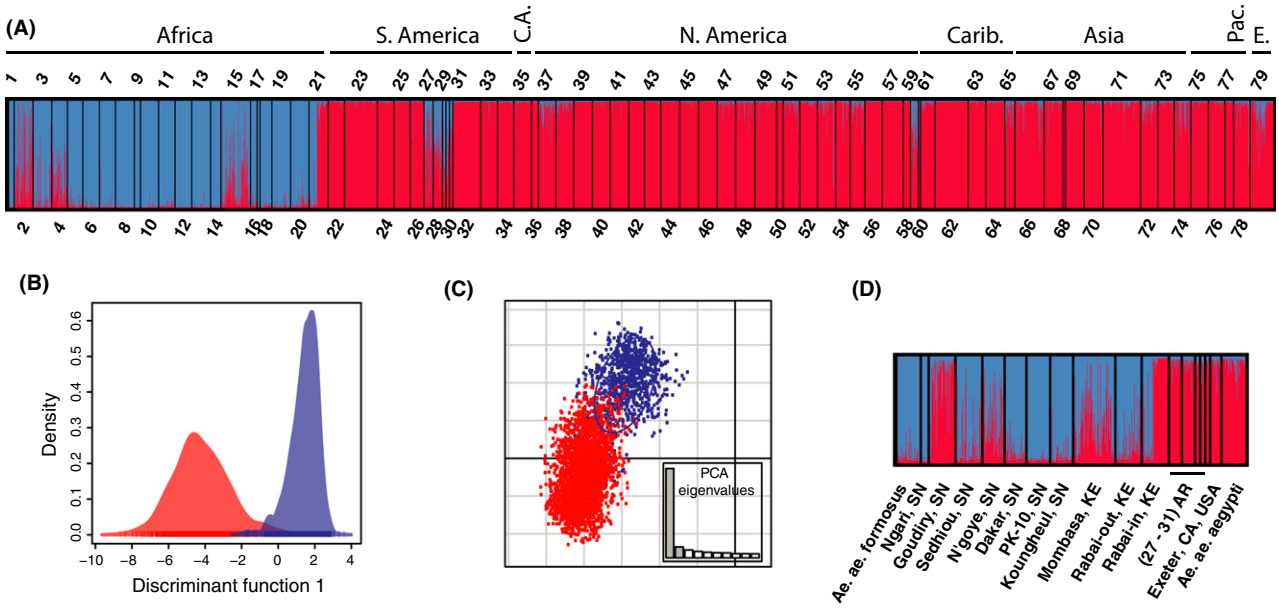

Fig. 2 Global genetic structure of *Aedes aegypti*. (A) STRUCTURE bar plot indicating genetic groupings of 79 geographic locations based on 12 microsatellite loci. Each vertical bar represents an individual. The height of each bar represents the probability of assignment to each of $K = 2$ clusters as determined using the delta K method. Each cluster is indicated by different colours: Aaa: red, and Aaf: blue. Population code numbers are in Table 1. (B) Discriminant analysis of principal components (DAPC) on microsatellite allele frequencies showing two clear genetic clusters with the minimal overlap; colours are as in A. (C) Scatter plot of the first two principal components of the same data analysed in A and B. Groups corresponding to the Aaa and Aaf genetic clusters are plotted using the same colours as in A. Most of the variation is captured by the first and second PCA, as shown by the eigenvalue graph. (D) STRUCTURE bar plot of those individual populations showing admixture in A; colours are consistent in A, B and C. [Colour figure can be viewed at wileyonlinelibrary.com].

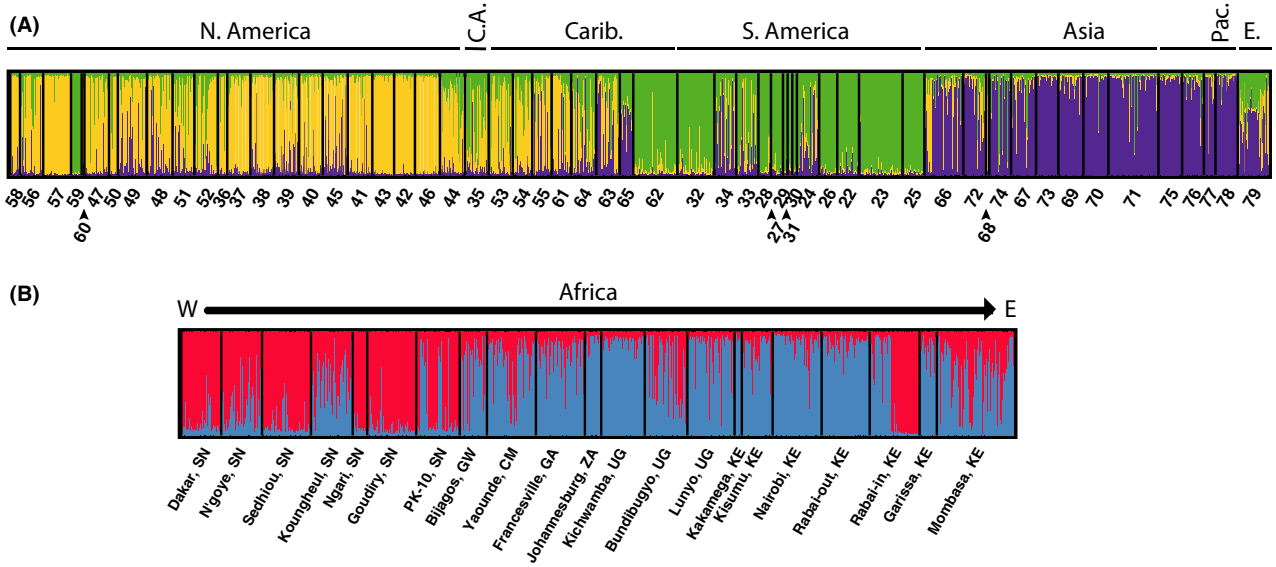

Fig. 3 Genetic structure of *Aedes aegypti* (A) out of Africa and (B) within Africa. STRUCTURE bar plots indicate relatedness among geographic locations. Population codes in A are as labelled in Table 1. Abbreviations in A top: C. A. = Central America, E. = Europe, Pac. = Australia, Tahiti and Hawaii. Populations are sorted by countries and by longitude (W: west to E: east). [Colour figure can be viewed at wileyonlinelibrary.com].

North and South America generally split from each other forming two large genetic groups, with Exeter (USA) and Chetumal (Mexico) being exceptions (Fig. 5). The single Central American sample (Siquirres, Costa Rica) shows genetic affinities to both South and North America, as expected from its transitional geographic location. However, because we only had a single population sample from Central America, we could not perform more detailed analyses of the region.

South America is further divided into three genetic groups: the north including Colombian, Venezuelan and northern Brazil populations; a southern Brazil group; and Argentina (Fig. 5A). Mexico has four genetic groups that correspond geographically to the north, south, central and Caribbean regions of the country (Fig. 5C). Northern Mexican populations along the border with the United States are genetically close to Texas, Arizona and southern California (Fig. 5A). Within the United States, the southeast (Louisiana, Florida and Georgia) groups together (Fig. 5A, B) and is distinct from the southwest (Texas and Arizona) and northern California. DAPC plots for the Americas and the American regions are shown in Fig. S1 (Supporting information).

Caribbean populations are also quite distinct, although we have only sampled four locations plus the Florida Keys, USA (Fig. 6A, B). Dominica is strongly differentiated from the rest of the group and has strong affinities to South America (Figs 3A, 6A-B). Excluding Dominica, the Florida Keys (USA) break from all other islands.

Asia appears to have a similar rate of genetic differentiation by distance as Africa (Fig. 4A). However,

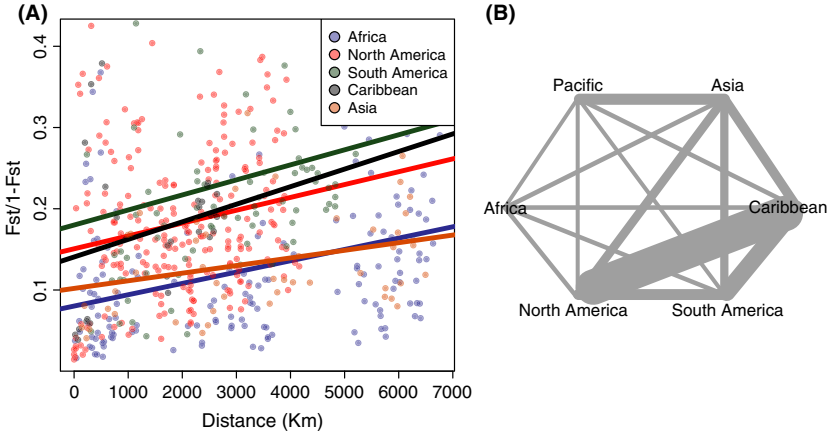

Fig. 4 Genetic differentiation of major geographic regions. (A) Scale of geographic genetic differentiation. Genetic distance is given as the linearized F_{ST} ($F_{ST}/(1 - F_{ST})$) for the analysis of 12 microsatellite loci. Statistical significance was evaluated using a Mantel test and were all significant positive slopes ($P < 0.05$) except for the Caribbean ($P = 0.18$) and South America ($P = 0.07$) populations. (B) Geneflow network between the continents or regions. The thickness of the lines is proportional to F_{ST}. [Colour figure can be viewed at wileyonlinelibrary.com].

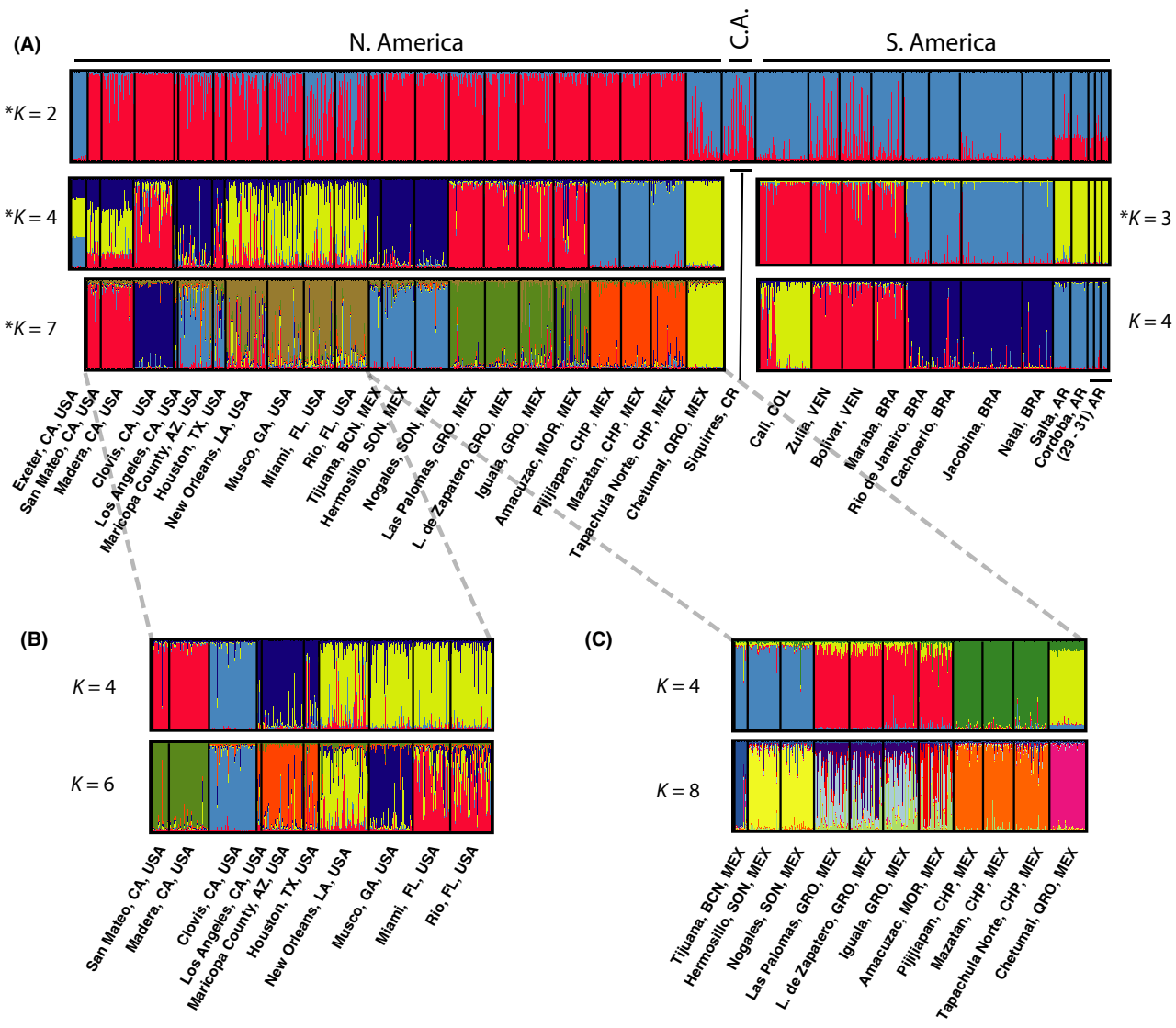

Fig. 5 Genetic structure of *Aedes aegypti* within the American continent. Panels are (A) all continental America, (B) North America (excluding Exeter, California) and (C) Mexico. STRUCTURE bar plots indicate relatedness among geographic locations. Plots representing the optimal *K* as determined by the delta *K* method are indicated by an asterisk (*). Discriminant analysis of principal component plots for these data are shown in Fig S1. [Colour figure can be viewed at wileyonlinelibrary.com].

Bayesian clustering analysis indicates a stronger genetic structure within Asia than within the African continent, similar to the genetic structure seen in the New World. When Asia and the Pacific populations are analysed together, the Pacific clusters with Pakistan and Saudi Arabia at $K = 2$, while Thailand, Vietnam and the Philippines form another group (Fig. 6C). More resolution is obtained by analysing the STRUCTURE plot at $K = 5$, with most major regions being genetically distinct (Fig. 6C). DAPC on Asian and Pacific populations is in agreement with the STRUCTURE results at $K = 5$, highlighting the distinctiveness of the Australian populations and of Hanoi, Vietnam (Fig. 6D).

Geneflow estimates

The patterns of genetic relatedness among the continents and the Caribbean are summarized in Fig. 4B. The thickness of the lines is based on F_{ST}, a measure of genetic differentiation that can also be interpreted as level of gene flow among populations ($4N_em$ for populations at equilibrium, which is unlikely to be the case for most of these relatively young populations). The data suggest that the Caribbean and North America are highly connected and that there is also a significant gene flow between the Caribbean and South America. Africa, as expected, is the most genetically isolated; that is, all lines connecting

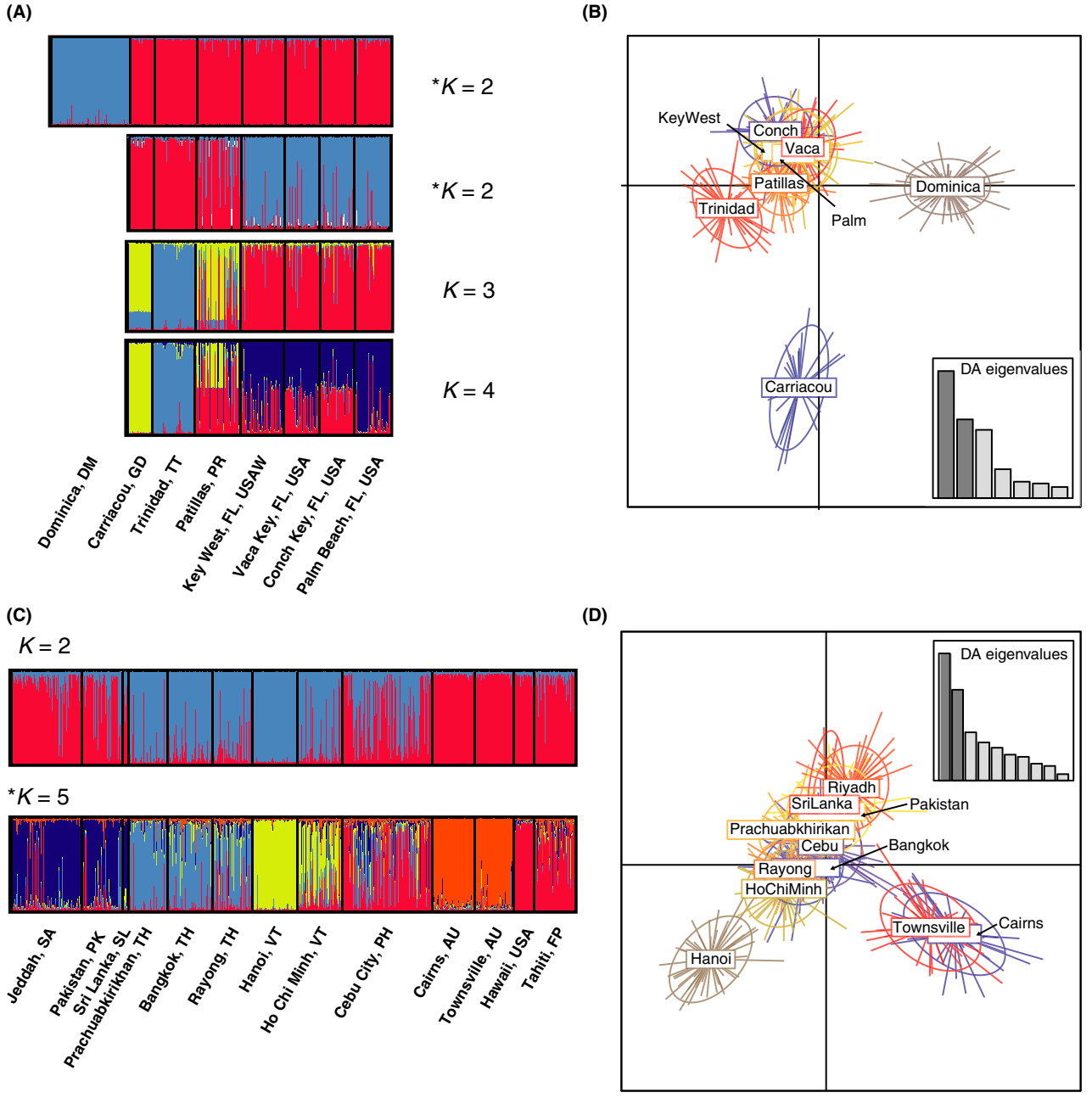

Fig. 6 Genetic structure of *Aedes aegypti* in the Caribbean and Asia/Pacific regions. (A) STRUCTURE plots of Caribbean populations (including Florida Keys) with *K* number of clusters as indicated. Plots representing the optimal *K* as determined by the delta *K* method are indicated by an asterisk (*). (B) Same populations in A in a discriminant analysis of principal components (DAPC). (C) and (D) same analyses as (A) and (B) for Asia and the Pacific (Australia, Tahiti and Hawaii). [Colour figure can be viewed at wileyonlinelibrary.com].

Africa are narrow. The Pacific (Australia plus Tahiti–French Polynesia and Hawaii) is also quite distinct, only showing some affinity with Asia.

History of Ae. aegypti

We use ABC (approximate Bayesian computation; Beaumont *et al.* 2002) to investigate the origin of *Ae. aegypti* outside Africa by testing the four plausible migration scenarios shown in Fig. 7: (i) Africa to America to Asia, (ii) Africa to Asia to America, (iii) first Africa to America/second Africa to Asia and (iv) first Africa to Asia/second Africa to America. The best-supported scenario was the one where the New World was founded from Africa and Asia was colonized from the New World (Scenario 1: $P = 0.9921$ in Fig. 7 and Table S1, Supporting

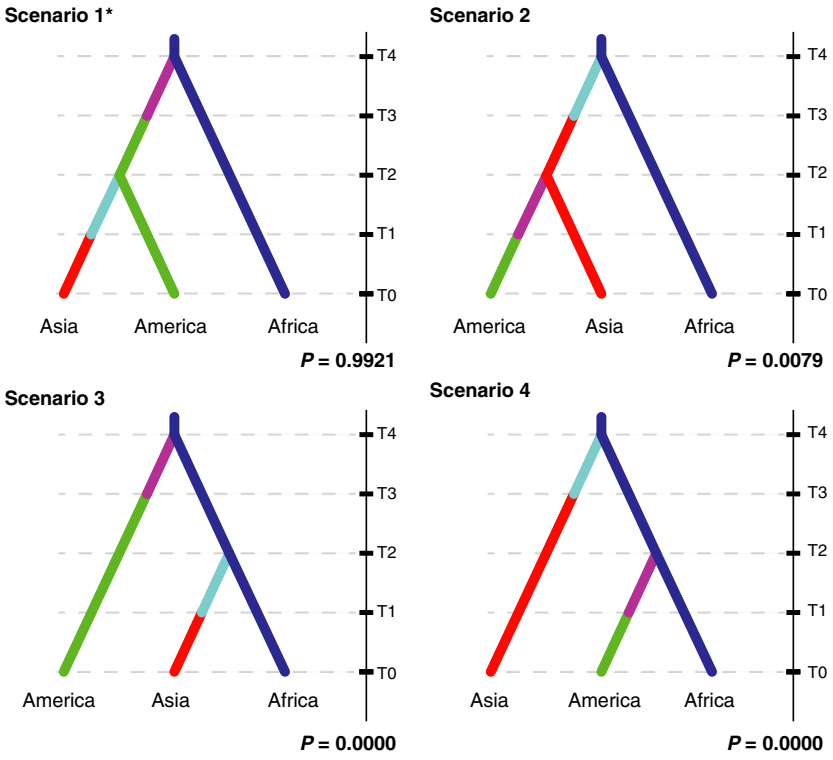

Fig. 7 Evolutionary scenarios of *Aedes aegypti* colonization of Asia, evaluated using approximate Bayesian computation inference as implemented by the DIYABC software (Cornuet *et al.* 2014). Scenarios include three populations: Africa, America and Asia, $N = 200$ for each continent. T0 represents the most recent time point and increasing values of T go back in time. Scenario 1: Africa to America to Asia; Scenario 2: Africa to Asia to America; Scenario 3: Africa to America + Africa to Asia (after America colonization); and Scenario 4: Africa to America + Africa to Asia (before America colonization). Posterior probabilities are shown for each scenario. The best-fit scenario is indicated by an asterisk (*). For more details, see Materials and methods and Table S1. [Colour figure can be viewed at wileyonlinelibrary.com].

information). Alternative scenarios were poorly supported by the analysis ($P < 0.001$; Fig. 7 and Table S1, Supporting information). The estimated mutation rate under the best-fit scenario was 9.5×10^{-6} and falls within the range of microsatellite mutation rates estimated for other Diptera (Schug *et al.* 1997; Pfeiler *et al.* 2013). The time estimated for the founding of Asian populations assuming an average of 10 generations per year is 1280–1770 generations or ~150 years ago (Table S1, Supporting information), setting priors for leaving Africa at 4500–6000 generations ago (~500 years ago) in agreement with historical records, suggesting the New World populations were founded by trade between the New World and Old World in the 15th and 16th centuries.

ABC analyses to determine the origin of the indoor Aaa from Rabai, Kenya, explored six scenarios involving a founder effect from Africa, New World or Asia. These scenarios are described in Table S2 and Fig. S2 (Supporting information). The analyses assigned moderate posterior probabilities to multiple colonization hypotheses: (1) direct colonization from the Americas after colonization of Asia (Scenario 3, $P = 0.56$), (ii) direct colonization from Asia (Scenario 1, $P = 0.24$) and (iii) colonization from the Americas prior to colonization of Asia (Scenario 4, $P = 0.20$). The scenario with highest support, Rabai being founded from the Americas after Asia (Scenario 3), predicts that the indoor Rabai Aaa populations are between 10 and 300 generations, or less than 30 years old (Table S2, Supporting information).

Other taxa

In addition to the samples of Aaf and Aaa, we had access to one laboratory strain labelled *Ae. aegypti queenslandensis* that was originally collected in Surabaya, Indonesia, in 2000. This strain underwent several rounds of selection for the scaling pattern assumed to be characteristic of the *queenslandensis* subspecies (Tsuda *et al.* 2003). The observed heterozygosity of this population was lower than any of the *Ae. aegypti* populations ($H_o = 0.276$). DAPC shows this strain as genetically distinct from all other *Ae. aegypti* (Fig. S3).

We also analysed 26 wild-caught *Ae. mascarensis*, the most closely related living relative of *Aedes aegypti* (sister taxon or species), and successfully genotyped all 12 microsatellite loci. Microsatellite diversity in this species does not overlap with populations of *Ae. aegypti* s. l. (Fig. 8).

Discussion

Global patterns

From a global perspective, *Aedes aegypti* s.l. is divided into two major genetic units (Fig. 2) that for the most

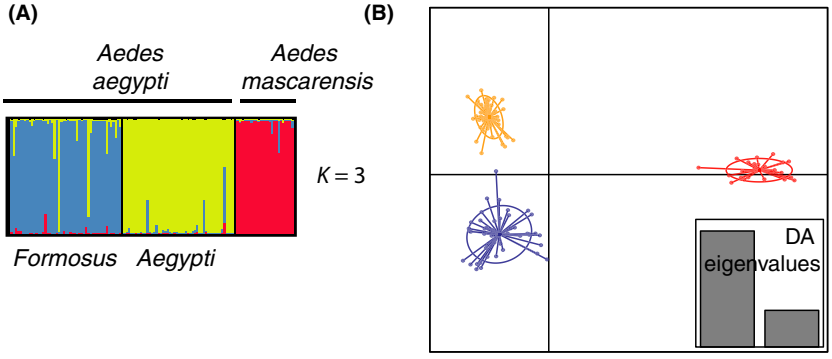

Fig. 8 Genetic structure among *Aedes aegypti* and *Aedes mascarensis* populations. (A) STRUCTURE bar plots for the 26 *Ae. mascarensis* sampled and, to avoid sample size artefacts, 50 random individuals subsampled from the large Aaa and Aaf data set, excluding those populations with large admixture levels (Fig 2D). (B) Discriminant analysis of principal components for the same samples depicted in the STRUCTURE plot. [Colour figure can be viewed at wileyonlinelibrary.com].

part, fall into the classically defined subspecies: *Ae. aegypti formosus* in Africa and *Ae. aegypti aegypti* outside Africa. These *Ae. aegypti* subspecies are distinct from each other and inhabit different geographic ranges, consistent with Mayr's definition of subspecies (Mayr 1963). The distinctness is genetic, morphological (scaling patterns and cuticle coloration), behavioural (host and oviposition choice) and ecological (larval habitats) as well as in geographic distribution (Trpis & Hausermann 1975). However, these distinctions are not absolute and may be breaking down in recent years.

Aaa and Aaf have both been reported in parts of coastal East Africa as far back as the 1950s based on morphology and habitat (Mattingly 1957) and later corroborated by genetic studies (Tabachnick *et al.* 1979; Brown *et al.* 2011a). Interestingly, in the city of Mombasa, Kenya, where both subspecies co-occur, there is free mixing among forms (locality 15 in Figs 2A, D, 3B). In contrast, in the Rabai District of Kenya, located just 20 km northwest of Mombasa, the two forms are not interbreeding even though adults can sometimes be found together indoors, or at least they remain genetically distinct (contrast locality 15 with locality 21 in Figs 2A, D, 3B). The Rabai District is rural consisting of small villages with mud wall huts surrounded by farmland and then forest. The samples we studied were larvae taken from water stored in large clay vessels inside huts ('Rabai-in') and from the forests ('Rabai-out') a few hundred metres distant (also analysed in Tabachnick *et al.* 1979; Brown *et al.* 2011a; McBride *et al.* 2014). The genetic differentiation of these populations has remained stable over nearly four decades of study in Rabai (Tabachnick *et al.* 1979). Considering the geographic closeness of Mombasa to Rabai, it is likely that ecology plays a major role on whether these subspecies freely interbreed. Mombasa is urban, whereas the Rabai District is rural with small villages adjacent to the forest. [See Lounibos (2003) for details of the ecology of mosquitoes in the Rabai District of Kenya.]

Senegal (SN) also harbours populations with affinities to domestic populations outside Africa (Goudiry and N'goye: localities 2 and 4 in Fig. 2A). This likely represents a recent introduction through shipping or other human activities (Brown *et al.* 2011a). These collections from Senegal come from cities or towns, and like in urban Mombasa in East Africa, the two forms are freely interbreeding, again contrasting to the rural Rabai District of Kenya. Huber *et al.* (2008), Sylla *et al.* (2009) and Paupy *et al.* (2010) have previously made similar observations on *Ae. aegypti* in Senegal.

The apparent recent Aaf ancestry in Argentinian populations is harder to explain (localities 27–31 in Fig. 2A). The presence of recent African ancestry may extend to populations of southern Bolivia and Paraguay (Rondan Duenas *et al.* 2009; Llinas & Gardenal 2012; Paupy *et al.* 2012), regions not sampled by us. This genetic result is consistent with the observations of *Ae. aegypti* breeding in tree holes in Argentina, a typical Aaf larval breeding habitat (Mangudo *et al.* 2015). In attempting to interpret this, it is important to note that *Ae. aegypti* was declared eradicated in much of South America including Argentina, Bolivia and Paraguay by 1970 (Tonn *et al.* 1982; Schatzmayr 2000) and recolonized since that time. Although the analyses of Brazilian populations were consistent with a complete eradication in that country (Monteiro *et al.* 2014), it is conceivable that given the remoteness of the region, eradication was never fully achieved. The domestication event leading to Aaa likely occurred before *Ae. aegypti* migrated from Africa. Thus, if there was a relict population in Argentina that escaped eradication efforts, it should have been Aaa. Furthermore, eradication was never achieved in many other parts of the New World including northern South America, Caribbean and the United States (Gubler 1998), and yet there is no indication of recent Aaf ancestry today in these noneradicated regions. An alternative explanation is that human trade or migration has recently introduced African Aaf to Argentina and the adjacent regions to the northwest and is now interbreeding with local Aaa populations.

History of Ae. aegypti introductions

Data reported here, as well as much previous work, support the ancestral status of the African populations (Brown *et al.* 2011a, 2014; Bennett *et al.* 2016). The elevated number of alleles and especially private alleles found in African populations compared with outside Africa is the strongest genetic evidence in favour of an ancestral African type (Table 2).

Movement of *Ae. aegypti* from the Old World likely occurred with trans-Atlantic travel beginning in the 1500s (Tabachnick 1991; Powell & Tabachnick 2013), with populations outside Africa being monophyletic (Brown *et al.* 2011a, 2014; Bennett *et al.* 2016). Our interpretation of all populations outside Africa having a single origin contrasts with the interpretation of mtDNA diversity. Moore *et al.* (2013) suggested a dual origin of Aaa populations outside Africa based on the diversity of mtDNA haplotypes. However, their data are also consistent with populations outside Africa originating from a single sample of two mtDNA lineages from ancestral Africa. If Aaa originated as a response to the expanding Sahara Desert (as hypothesized in the next paragraph), these original proto-domestic populations were likely large enough to have captured multiple mtDNA lineages from the ancestral populations south of the Sahara.

It is well established that Aaa arrived in Asia after its arrival in the New World, likely in the 1890s (Powell & Tabachnick 2013). Our data support a scenario where the colonization of Asia occurred from the New World (Scenario 1: in Fig. 7 and Table S1, Supporting information). However, an alternative colonization scenario not covered by our analysis is plausible. It has been hypothesized that the domestication of *Ae. aegypti* was the result of the Sahara Desert expansion (4000–6000 years ago). Human habitats would have been the only reliable water sources north of the desert, so *Ae. aegypti* evolved to breed in human-generated containers and to take bloodmeals from humans (Peterson 1977; Tabachnick 1991; Powell & Tabachnick 2013). Aaa persisted in the Mediterranean Basin until about 1950 (Curtin 1967; Holstein 1967). It is conceivable that these populations were the original proto-Aaa form preadapted to surviving aboard ships travelling from the Mediterranean (Spain and Portugal) to the New World. These Mediterranean populations could be the source of Asian *Ae. aegypti,* their introduction aided by the opening of the Suez Canal in 1869 (Tabachnick 1991). The time estimated by the ABC analysis for the founding of Asia is remarkably consistent with the historical record ~150 years ago in the late 19th century (Table S1, Supporting information). Urban dengue fever was first recorded in Asia in the 1890s and *Ae. aegypti* was the only Asian urban vector of dengue at that time (Smith 1956). Mediterranean populations therefore could represent a missing 'ghost taxon' (*sensu* Norell 1996) that would change the outcome of the ABC results, if included.

We also tested hypotheses concerning the origin of the indoor-collected populations of Aaa in Rabai (Kenya), most likely a re-introduction from outside Africa (Fig. S2, Table S2, Supporting information). Interestingly, the scenario with highest support, Rabai being founded from the Americas after Asia, predicts a very young age for the Rabai indoor populations, 10–300 generations or less than 30 years ago (Table S2, Supporting information). This result is inconsistent with the observation that Aaa has been in coastal Kenya for at least 65 years (Mattingly 1957). It is possible that, like for the New World and Asia, the now extinct Mediterranean Basin populations may have been the immediate source of the Rabai Aaa indoor form, especially given the proximity of the Suez Canal connecting the Kenyan coast to the Mediterranean.

New world population structure

In interpreting New World population structure for *Ae. aegypti*, it is important to consider that this species was officially eradicated from much of this area in the 1950s and 1960s, with recolonization starting in the 1970s. The genetic data available in Monteiro *et al.* (2014) were interpreted as consistent with Brazil being colonized by *Ae. aegypti* from two directions, from South American countries to the north and from the Caribbean, regions where eradication was never attained. Frequencies of mtDNA haplotypes (Braco *et al.* 2007) and alleles at the insecticide resistance gene, *kdr*, are consistent with these two major groups in Brazil (Linss *et al.* 2014). Argentina, the third genetic group, is an enigma in being the only region outside Africa with evidence of recent ancestry of mosquitoes from Africa (discussed previously).

The genetic pattern observed in Mexico is consistent with previous work using the collections taken about 20 years prior to our sampling (Gorrochotegui-Escalante *et al.* 2000), attesting to the temporal stability of Aaa population genetic patterns. Similarly, in the USA, the genetic break in *Ae. aegypti* approximately at the border between Louisiana and Texas was documented ~40 years ago using allozymes (Powell *et al.* 1980; Wallis *et al.* 1983), suggesting that this division is temporally stable, at least for that time frame. California populations, some of which are thought to be quite recently established (Gloria-Soria *et al.* 2014), are complex and are the subject of ongoing independent analyses that will be reported in the near future.

Aaf and Aaa population structure and introgression

Aaf in Africa are less genetically differentiated and structured compared with Aaa populations outside Africa (Figs 3 and 4). This pattern could be explained by the fact that populations outside Africa are young, have experienced founder effects, inhabit highly discontinuous habitats and active migration in this species is restricted (Reiter 2007). Being the ancestral region, African populations are much older and there is evidence of historical admixture events within the region during the Holocene (1000–12 000 ya), when the forest expanded from glacial refugia (Bennett *et al.* 2016) and likely led to the relative homogenization of gene frequencies.

It is also important to note that this relatively low level of genetic differentiation among African populations belies the greater habitat diversity now occupied by African populations of *Ae. aegypti*. Historically, African populations bred almost exclusively in forests. Today, populations of *Ae. aegypti* in Africa can be found in urban habitats (Paupy *et al.* 2008; Kamgang *et al.* 2013) even if they fall genetically into the Aaf group, for example localities 5 (Dakar, Senegal), 8 (Franceville, Gabon) and 13 (Yaounde, Cameroon) in Fig. 2A. There is evidence that these urban and forest African populations are genetically differentiated to certain degree (Paupy *et al.* 2008), yet remain within the large Aaf genetic group.

Interestingly, one might expect the younger Asian populations to be similarly differentiated as New World populations, but this is not the case with regard to the overall genetic differences as measured by F_{ST} (Fig. 4). The observed genetic homogeneity of Asian populations relative to the New World might be due to different time and patterns of passive migration due to human commerce. From the time of the first appearance of Aaa in Asia in the late 19th century, trade by rail, sea and overland routes in Asia was much more developed compared with commerce in the New World of the 16th to 18th centuries (e.g. Tracy 1990) when *Ae. aegypti* populations were first established there (Powell & Tabachnick 2013).

Given the genetic, morphological, ecological and behavioural differences between Aaa and Aaf, it is perhaps surprising that there is little or no evidence of reproductive isolation between these forms. Hybrids and backcrosses are fertile and viable (Moore 1979). In areas where the subspecies coexist (e.g. Senegal, Kenya and Argentina), they seem to freely interbreed, as indicated by the vast majority of loci being in H-W proportions. For example, localities 2, 3, 4, 15, 27, 28 and 29 in Fig. 2 have less than two loci deviating from H-W expectations at the 5% significance level (Table S4, Supporting information) similar to the number of deviations observed for all population samples, which we attribute to the presence of rare null alleles (see Results). Even in Rabai, Kenya, where Aaa and Aaf retain genetic distinctness in sympatry, no pre- or post-mating isolation can be detected (Moore 1979) [low fertility has recently been reported in F_1 offspring between laboratory colonies derived from Senegal Aaf and laboratory colonies of Aaa (Dickson *et al.* 2016)].

Introgression among divergent populations resulting from multiple introductions could lead to the increased invasiveness due to cumulative genetic variation. This mechanism has been postulated for other invasive taxa (reviewed in Bock *et al.* 2015), as well as for the invasive mosquito *Culex pipiens* (Fonseca *et al.* 2009). From the data presented here, it appears as if the two classically defined subspecies Aaa and Aaf are increasingly coming into contact and hybridizing (e.g. Senegal, Kenya and Argentina). Will this lead to the increased invasiveness of *Ae. aegypti* in these localities? There is presently no information on this.

Other taxa

It is unclear whether the third named subspecies, *Ae. aegypti queenslandensis*, is supported by genetic data. We had access to one laboratory strain that had been labelled *Ae. aegypti queenslandensis* and was originally collected in Surabaya, Indonesia, in 2000. This strain underwent several rounds of selection for the scaling pattern assumed to be characteristic of that subspecies (Tsuda *et al.* 2003). The strain is genetically distinct from all other *Ae. aegypti* populations (Fig. S3), but its low heterozygosity suggests that this could be due to the long laboratory rearing and/or artificial selection performed on this single line (Tsuda *et al.* 2003). Rašić *et al.* (2016) recently showed that specimens corresponding morphologically to *Ae. ae. queenslandensis* collected in Australia and Singapore were not genetically distinct from the specimens collected in the same locality that corresponded morphologically to *Ae. ae. aegypti*. With regard to morphology, Mattingly (1957) stated that *queenslandensis* '...does not differ notably from specimens from many other parts of the world'. As far as known, there are no extant populations of this dubious subspecies.

Aedes mascarensis is the closest living relative of *Aedes aegypti* (sister taxon or species). Its distribution is restricted to a single island in the Indian Ocean, Mauritius (MacGregor 1924). This species can form fertile F_1 hybrids with *Ae. aegypti*, although hybrids breakdown in further generations (Hartberg & Craig 1970). In fact, this species is so closely related to *Ae. aegypti* that Delatte *et al.* (2011) suggested it should be considered

simply a 'form' or subspecies of *Ae. aegypti*, based on mtDNA data. Our results, however, indicate strong genetic differentiation not overlapping with the genetic diversity within Aaa and Aaf (Fig. 8). Identical findings were made by Brown *et al.* (2014) using 1504 SNPs. These data therefore suggest that *Ae. mascarensis* is a genetically distinct species and not simply another 'form' within *Ae. aegypti* s.l.

Implications for disease control

These results suggest that microsatellite data can serve to identify the likely origin of new introductions of *Ae. aegypti* that place resident human populations at risk for diseases transmitted by this mosquito. Populations of *Ae. aegypti* vary considerably in efficiency of transmitting disease-causing viruses (Tabachnick *et al.* 1985; Black *et al.* 2002; Sim *et al.* 2013). A recent study on *Ae. albopictus* highlights the importance of population genetic differentiation of mosquitoes and their involvement in the transmission of pathogenic viruses. Only certain genotypes of *Ae. albopictus* can preferentially transmit the chikungunya virus variant that led to the massive epidemic in La Reunion Island (Vazeille *et al.* 2016). Likewise, populations vary in their degree of resistance to insecticides (e.g. Montella *et al.* 2007; Linss *et al.* 2014). Recognizing the source of the introduction allows assessment of the level of health risks associated with the invasion and can guide control measures.

As an example of how analysis of hierarchical genetic structure can narrow down the origin of a new invasion, we consider the Caribbean (Fig. 6A). Assume an initial analysis on the total data set indicated that the new introduction is Aaa (Fig. 2), subsequent regional analysis then points to the New World and finally to the Caribbean (Figs 3 and 6). When the Caribbean alone is considered, using $K = 2$ in STRUCTURE analyses, Dominica can be distinguished from all other samples of the region (Fig. 6A). Removing Dominica from the analysis allows, first, to separate the Florida Keys from the other islands. Then at $K = 3$, Trinidad stands out as genetically distinct from the other populations, with Puerto Rico and Carriacou starting to differentiate. Differentiation between these two latter islands, however, becomes more obvious at $K = 4$. So, while our data do not always allow the precise identification of a single population, it does allow identification to a reasonably narrow geographic range. Microsatellites were used in a similar manner to identify the origin of a 2010 temporary introduction into the Netherlands (Brown *et al.* 2011b) and a presumed recent introduction into California (Gloria-Soria *et al.* 2014), similar to studies of colonization routes of other pest species such as the hemlock woolly adelgid from Japan to North America (Havill *et al.* 2016).

The accuracy of such assignments depends on two factors. One is thoroughness of sampling. This points to the importance of adding more samples to this database in the future. Second, how long a database like this is reliable in determining origin of new introductions depends on the temporal genetic stability of populations. We addressed the temporal genetic stability/instability of *Ae. aegypti* populations in a previous study (Gloria-Soria *et al.* 2016) and found that some populations remain quite stable at least over 2–7 years, whereas others change. However, almost all temporal changes are minor compared to spatial stability and do not distort the geographic patterns used to identify the origin of new introductions. In earlier sections of this study, we also discussed the evidence that genetic patterns in Rabai Kenya, the southern United States and Mexico have been stable for at least 20–40 years.

Our results should be considered in any attempts to use genetic modification of populations to control disease transmission (McGraw & O'Neill 2013). The genetic distinctness of populations, especially outside Africa, means that ease with which introduced genes can be driven into a population may vary depending on the genetic composition of the target population. Different genetic strategies may need to be designed for genetically distinct target populations. Genetic modification using a method based on standing genetic variation in populations (e.g. Powell & Tabachnick 2014) will need to be specific to the population. Finally, the low rates of migration implied by the high genetic structuring of *Ae. aegypti* mean that the released genetic material will not rapidly spread beyond the local population; that is, migration rates and distances are highly restricted. This evidence for low dispersal based on genetic patterns is also supported by direct behavioural and ecological studies (Reiter 2007).

Conclusions

We have documented the great degree of genetic diversity and genetic structure of *Aedes aegypti*. In the laboratory, no signs of reproductive isolation between the *Ae. aegypti* subspecies or between any populations have been observed. The evident-free interbreeding in African cities where the domestic form has been introduced (Senegal and Mombasa, Kenya) is evidence that there is no reproductive isolation under field conditions. Thus, *Ae. aegypti* can be considered a single genetic species.

In addition to the domestication event that led to *Ae. aegypti*'s spread out of Africa ~500 years ago, it is clear that additional independent domestication events

are occurring in Africa, probably in response to expanding urbanization. A prime example is Yaounde, Cameroon (locality 13 in Fig. 2A), where populations were collected in an urban environment, yet are genetically much more similar to forest-breeding populations typical of Aaf, than to Aaa outside Africa. These recent ecologically domestic types in Africa have retained the black cuticle typical of forest-breeding populations. Mirroring this shift of classically defined Aaf from forest to domestic breeding are classically defined Aaa domestic populations that revert to breeding in natural sites such as rock holes, bromeliads, plant axials and tree holes, best documented in the Caribbean (Chadee *et al.* 1998).

This mosquito may be the most genetically diverse species of insect ever studied. The rapid and highly successful adaptation to human habitats and its subsequent spread is proof of the adaptive flexibility this genetic variation confers. This adaptability presents a challenge to control populations of *Ae. aegypti* in efforts to decrease their impact on human health.

Acknowledgements

We thank generous colleagues who assisted in providing the samples for this study; we thank J. Pinto, R. Butlin, M. Boots, D. Severson, R. Kelly, K. McCarley, Y. Lourenco, R. Martinez, Dr E. Nchoutpouen, S. Kluh, L. Irby, M. Garcia, A. Scalzo, S. Mulligan, J. Holeman, R. Gay, N. Zahiri, B. Weber and D. J. Gubler. We thank David A. Kellner and Robert Hallberg who provided technical assistance in the laboratory. Financial support was provided by NIAID RO1 AI101112 awarded to JRP. A. Gloria-Soria was supported by NIAID 3R01AI091646-04S1. G. Seixas was funded by a PhD fellowship of Fundação para a Ciência e Tecnologia/MCTES (SFRH/BD/98873/2013). EIK and AAS were funded by the National Plan for Science, Technology and Innovation (MAARIFAH), King Abdulaziz City for Science and Technology, Kingdom of Saudi Arabia (Award Number BIO1483), to EIK.

References

Beaumont MA, Zhang W, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics*, **162**, 2025–2035.

Bennett KL, Shija F, Linton Y-M *et al.* (2016) Historical environmental change in Africa drives divergence and admixture of Aedes aegypti mosquitoes: a precursor to successful worldwide colonisation? *Molecular Ecology*, **25**, 4337–4354.

Beserra EB, de Castro Jr FP, dos Santos JW, Santos TDS, Fernandes CR (2006) Biology and thermal exigency of *Aedes aegypti* (L.) (Diptera: Culicidae) from four bioclimatic localities of Paraiba. *Neotropical Entomology*, **35**, 853–860.

Black WC, Bennett KE, Gorrochotegui-Escalante N *et al.* (2002) Flavivirus susceptibility in *Aedes aegypti*. *Archives of Medical Research*, **33**, 379–388.

Bock DG, Caseys C, Cousens RD *et al.* (2015) What we still don't know about invasion genetics. *Molecular Ecology*, **24**, 2277–2297.

Braco JE, Lourenco-de-Oliveira R, Salluym MA (2007) Genetic variability of *Aedes aegypti* in the Americas using a mitochondrial gene: evidence of multiple introductions. *Memorias do Instituto Oswaldo Cruz*, **102**, 573–580.

Brown JE, McBride CS, Johnson P *et al.* (2011a) Worldwide patterns of genetic differentiation imply multiple "domestications" of *Aedes aegypti*, a major vector of human diseases. *Proceedings of the Royal Society of London B: Biological Sciences*, **278**, 2446–2454.

Brown JE, Scholte E-J, Dik M, Hartog WD, Beeuwkes J, Powell JR (2011b) *Aedes aegypti* mosquitoes imported into the Netherlands, 2010. *Emerging Infectious Diseases*, **17**, 2335–2337.

Brown JE, Evans B, Zheng W *et al.* (2014) Human impacts have shaped historical and recent evolution in *Aedes aegypti*, the dengue and yellow fever mosquito. *Evolution*, **68**, 514–525.

Chadee DD, Ward RA, Novak RJ (1998) Natural habitats of *Aedes aegypti* in the Caribbean—a review. *Journal of the American Mosquito Control Association*, **13**, 5–11.

Christophers R (1960) *Aedes aegypti, the Yellow Fever Mosquito*. Cambridge University Press, Cambridge, UK.

Clemens AN (1992) *The Biology of Mosquitoes. Vol. 1: Development, Nutrition and Reproduction*. Chapman and Hall, London.

Clemens AN (1999) *The Biology of Mosquitoes Vol 2: Sensory Reception and Behaviour*. CABI, University Press, Cambridge.

Cornuet J-M, Pudlo P, Veyssier J *et al.* (2014) DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, **30**, 1187–1189.

Curtin TJ (1967) Status of Aedes aegypti in the eastern Mediterranean. *Journal of Medical Entomology*, **4**, 48–50.

Delatte H, Bagny L, Brengue C, Bouetard A, Pauly C, Fontenille D (2011) The invaders: Phylogeography of dengue and chikungunya viruses Aedes vectors on the southwest islands of the India Ocean. *Infection, Genetics and Evolution*, **11**, 1769–1781.

Dickson LB, Sharakhova MV, Timoshevskiy VA *et al.* (2016) Reproductive incompatibility involving Senegalese *Aedes aegypti* (L.) is associated with chromosome rearrangements. *PLOS Neglected Tropical Diseases*, **10**, 1–28.

Earl SA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

Ersts PJ (2016) Internet 2016 GEOGRAPHIC DISTANCE MATRIX GENERATOR (version 1.2.3). American Museum of Natural History, Center for Biodiversity and Conservation. Available from http://biodiversityinformatics.amnh.org/open_source/gdmg. Accessed on Feb. 2, 2016.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Fonseca DM, Smith JL, Kim H-C, Mogi M (2009) Population genetics of the mosquito *Culex pipiens pallens* reveals sex-linked asymmetric introgression by *Culex quiquefaciatus*. *Infection, Genetics and Evolution*, **9**, 1197–1203.

Giraldo-Calderón GI, Emrich SJ, MacCallum RM *et al.* (2015) VectorBase: an updated bioinformatics resource for invertebrate vectors and other organisms related with human diseases. *Nucleic Acids Research*, **43**, D707–D713.

Gloria-Soria A, Brown JE, Kramer V, Hardstone Yoshimizu M, Powell JR (2014) Origin of the dengue fever mosquito, Aedes aegypti, in California. *PLOS Neglected Tropical Diseases*, **8**, e3029.

Gloria-Soria A, Kellner D, Brown JE *et al.* (2016) Temporal genetic stability of *Stegomyia aegypti* (= *Aedes aegypti*) populations. *Medical and Veterinary Entomology*, **30**, 235–240.

Gorrochotegui-Escalante N, De Lourdes Munoz M, Fernandez-Salas I, Beaty BJ, Black WC (2000) Genetic isolation by distance among *Aedes aegypti* populations along the northeastern coast of Mexico. *American Journal of Tropical Medicine and Hygiene*, **62**, 200–209.

Gubler DJ (1998) Dengue and dengue hemorrhagic fever. *Clinical Microbiology Reviews*, **11**, 480–496.

Hartberg WK, Craig GB Jr (1970) Reproductive isolation in *Stegomyia* mosquitoes. II. Hybrid breakdown between *Aedes aegypti* and *A. mascarensis*. *Evolution*, **24**, 692–703.

Havill NP, Shiyake S, Lamb Galloway A *et al.* (2016) Ancient and modern colonization of North America by hemlock woolly adelgid, Adelges tsugae (Hemiptera: Adelgidae), an invasive insect from East Asia. *Molecular Ecology*, **25**, 2065–2080.

Holstein M (1967) Dynamics of *Aedes aegypti* distribution, density and prevalence in the Mediterranean area. *Bulletin of the World Health Organization*, **36**, 541–543.

Huber K, Ba Y, Dia I, Mathiot C, Sall AA, Diallo M (2008) *Aedes aegypti* in Senegal: genetic diversity and genetic structure of domestic and sylvatic populations. *American Journal of Tropical Medicine and Hygiene*, **79**, 218–229.

Jombart T (2008) Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403–1405.

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, **11**, 1–15.

Jupp PG, Kemp A, Frangos C (1991) The potential for dengue in South Africa: morphology and the taxonomic status of *Aedes aegypti* populations. *Mosquito Systematics*, **23**, 182–190.

Kalinowski ST (2005) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, **5**, 187–189.

Kamgang B, Ngoagouni C, Manirankiza A, Nakoune E, Paupy C, Kazanji M (2013) Temporal patterns of abundance of *Aedes aegypti* and *Aedes albopictus* (Diptera: Culicidae) and mitochondrial DNA analysis of *Ae. albopictus* in Central African Republic. *PLoS Neglected Tropical Diseases*, **7**, e2590.

Linss JG, Brito LP, Garcia GA *et al.* (2014) Distribution and dissemination of the Val1016Ile and Phe1534Cys Kdr mutation in *Aedes aegypti* Brazilian natural populations. *Parasites Vectors*, **7**, 1–11.

Llinas GA, Gardenal CN (2012) Phylogeography of *Aedes aegypti* in Argentina: long-distance colonization and rapid restoration of fragmented relicts after a continental control campaign. *Vector-Borne and Zoonotic Diseases*, **12**, 254–261.

Lounibos LP (1981) Habitat segregation among African treehole mosquitoes. *Ecological Entomology*, **6**, 129–154.

Lounibos LP (2002) Invasions by insect vectors of human disease. *Annual Review of Entomology*, **47**, 233–266.

Lounibos LP (2003) Genetic-control trials and the ecology of *Aedes aegypti* in the Kenya coast. In: *Ecological Aspects for Application of Genetically Modified Mosquitoes* (eds Takken W, Scott TW), pp. 33–43. Kluwer Academic Publishers, Dordrecht, The Netherlands.

MacGregor ME (1924) *Aedes (Stegomyia) mascarensis*, MacGregor: a new mosquito from Mauritius. *Bulletin of Entomological Research*, **14**, 409–412.

Mack RN, Simberloff D, Lonsdale WM, Evans H, Clout M, Bazzaz FA (2000) Biotic invasions: causes, epidemiology, global consequences, and control. *Ecological Applications*, **10**, 689–710.

Mangudo C, Aparicio JP, Gleiser RM (2015) Tree holes as larval habitats for *Aedes aegypti* in urban, suburban and forest habitats in a dengue affected area. *Bulletin of Entomological Research*, **105**, 679–684.

Marinho RA, Beserra EB, Bezerra-Gusmão MA *et al.* (2016) Effects of temperature on the life cycle, expansion, and dispersion of *Aedes aegypti* (Diptera: Culicidae) in three cities in Paraiba, Brazil.

Mattingly PF (1957) Genetical aspects of the *Aedes aegypti* problem I. Taxonomy and bionomics. *Annals of Tropical Medicine and Parasitology*, **51**, 392–408.

Mattingly PF (1967) Taxonomy of *Aedes aegypti* and related species. *Bulletin of the World Health Organization*, **36**, 552–554.

Mayr E (1963) *Animal Species and Evolution*. Belknap Press, Harvard, Cambridge, MA.

McBride CS, Baier F, Omondi AB *et al.* (2014) Evolution of mosquito preference for humans linked to an odorant receptor. *Nature*, **515**, 222–227.

McClelland GAH (1967) Speciation and evolution in Aedes. In: *Genetics of Insect Vectors of Disease* (eds Wright JW, Pal R), pp. 277–311. Elsevier, Amsterdam, The Netherlands.

McClelland GAH (1974) A worldwide survey of variation in scale pattern of the abdominal tergum of *Aedes aegypti* (L.) (Diptera: Culicidae). *Transactions of the Entomological Society of London*, **126**, 239–259.

McGraw EA, O'Neill SL (2013) Beyond insecticides: new thinking on an ancient problem. *Nature Reviews Microbiology*, **11**, 181–193.

Meirmans PG, Van Tienderen PH (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, **4**, 792–794.

Monteiro FA, Shama R, Martins A, Gloria-Soria A, Brown JE, Powell JR (2014) Genetic diversity of Brazilian *Aedes aegypti*: patterns following and eradication program. *PLoS Neglected Tropical Diseases*, **8**, e3167.

Montella IR, Martins AJ, Viana-Medeiros PF, Pereira Lima JB, Braga IA, Valle D (2007) Insecticide resistance mechanisms of Brazilian *Aedes aegypti* populations form 2001 to 2004. *American Journal of Tropical Medicine and Hygiene*, **77**, 467–477.

Moore DF (1979) Hybridization and mating behavior of *Aedes aegypti* (Diptera: Culicidae). *Journal of Medical Entomology*, **16**, 223–226.

Moore M, Sylla M, Goss L *et al.* (2013) Dual African origins of global *Aedes aegypti* s.l. populations revealed by mitochondrial DNA. *PLoS Neglected Tropical Diseases*, **74**, e2175.

Norell MA (1996) Ghost taxa, ancestors, and assumptions: a comment on Wagner. *Paleobiology*, **22**, 453–455.

Oetting WS, Lee HK, Flanders DJ, Wiesner GL, Sellers TA, King RA (1995) Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. *Genomics*, **30**, 450–458.

Paupy C, Kamgang B, Herve J-P, Fontenille D, Simard F (2008) Gene flow between domestic and sylvan populations of *Aedes aegypti* (Diptera: Culicidae) in north Cameroon. *Journal of Medical Entomology*, **45**, 391–400.

Paupy C, Brengues C, Ndiath O, Toty C, Herve J-P, Simard F (2010) Morphology and genetic variability within *Aedes aegypti* in Niakhar, Senegal. *Infection, Genetics and Evolution*, **10**, 473–480.

Paupy C, Le Goff G, Brengues C *et al.* (2012) Genetic structure and phylogeography of *Aedes aegypti,* the dengue and yellow-fever mosquito vector in Bolivia. *Infection, Genetics and Evolution*, **12**, 1260–1269.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, **28**, 2537–2539.

Peterson JL (1977) Behavior differences in two subspecies of *Aedes aegypti* (L.) (Diptera: Culicidae) in East Africa. Ph.D. dissertation, University of Notre Dame, Notre Dame, Ind.

Pfeiler E, Flores-López CA, Mada-Vélez JG, Escalante-Verdugo J, Markow TA (2013) Genetic diversity and population genetics of mosquitoes (Diptera: Culicidae: Culex spp.) from the Sonoran Desert of North America. *Scientific World Journal*, **2013**, 724609.

Powell JR, Tabachnick WJ (2013) History of domestication and spread of *Aedes aegypti*. *Memorias do Instituto Oswaldo Cruz*, **108**, 11–17.

Powell JR, Tabachnick WJ (2014) Genetic shifting: a novel approach for controlling vector-borne diseases. *Trends in Parasitology*, **30**, 282–288.

Powell JR, Tabachnick WJ, Arnold J (1980) Genetics and the origin of a vector population: *Aedes aegypti,* a case study. *Science*, **208**, 1385–1387.

Pritchard JK, Stephens M, Donnelly P (2000) 2000 Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

R Core Team (2013) ISBN 3-900051-07-0. [http://www.R-project.org/].

Rašić G, Filipović I, Callahan AG *et al.* (2016) The queenslandensis and the type form of the dengue fever mosquito (Aedes aegypti L.) are genomically indistinguishable. *bioRxiv*, doi: 10.1101/063792.

Raymond M, Rousset F (1995) GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Reiter P (2007) Oviposition, dispersal, and survival in *Aedes aegypti:* implications for efficacy of control strategies. *Vector-Borne and Zoonotic Diseases*, **7**, 261–273.

Rondan Duenas JC, Llinas GA, Panzetta-Dutari GM, Gradenal CN (2009) Two different routes of colonization of *Aedes aegypti* in Argentina from neighboring countries. *Journal of Medical Entomology*, **46**, 1344–1354.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.

Rousset F (2008) Genepop '07: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.

Schatzmayr HG (2000) Dengue situation in Brazil by year 2000. *Memorias do Instituto Oswaldo Cruz*, **95**, 179–181.

Schug MD, Mackay TF, Aquadro CF (1997) Low mutation rates of microsatellite loci in Drosophila melanogaster. *Nature Genetics*, **15**, 99–102.

Sim S, Jupatanakul N, Rabirez JL *et al.* (2013) Transcriptomic profiling of diverse *Aedes aegypti* strains reveals increased basal-level immune activation in dengue virus-refractory populations and identifies novel virus-vector molecular interactions. *PLoS Neglected Tropical Diseases*, **7**, e2295.

Slotman MA, Kelly NB, Harrington LC *et al.* (2007) Polymorphic microsatellite markers for studies of *Aedes aegypti* (Diptera: Culicidae), the vector of dengue and yellow fever. *Molecular Ecology Notes*, **7**, 168–171.

Smith CEG (1956) The history of dengue in tropical Asia and its probable relationship to the mosquito *Aedes aegypti*. *American Journal of Tropical Medicine and Hygiene*, **59**, 3–11.

Sylla M, Bosio C, Urdaneta-Marquez L, Ndlaye M, Black WC (2009) Gene flow, subspecies composition, and dengue virus-2 susceptibility among *Aedes aegypti* collections in Senegal. *PLoS Neglected Tropical Diseases*, **3**, e408.

Tabachnick WJ (1991) Evolutionary genetics and the yellow fever mosquito. *American Entomologist*, **37**, 14–24.

Tabachnick WJ, Munstermann L, Powell JR (1979) Genetic distinctness of sympatric forms of *Aedes aegypti*. *Evolution*, **33**, 287–295.

Tabachnick WJ, Wallis GP, Aitken THG *et al.* (1985) Oral infection of *Aedes aegypti* with yellow fever virus: geographical variation and genetic considerations. *American Journal of Tropical Medicine and Hygiene*, **34**, 1219–1224.

Theobald FV (1901) Monograph of the Culicidae of the World. Vols I–V. Brit Mus (Nat Hist), London.

Tonn RJ, Figueredo R, Uribe LJ (1982) *Aedes aegypti*, the yellow fever and dengue mosquito in the Americas. *Mosquito News*, **42**, 497–505.

Tracy JD (ed.) (1990) *The Rise of Merchant Empires: Long Distance Trade in the Early Modern World*. Cambridge Univ. Press, Cambridge, UK.

Trpis M, Hausermann W (1975) Demonstration of differential domesticity of *Aedes aegypti* (L.) (Diptera, Culicidae) in Africa by mark-release-recapture. *Bulletin of Entomological Research*, **65**, 199–208.

Tsuda Y, Yotopranoto S, Bendryman SS, Rossmanida R, Dachlan YP, Takagi M (2003) Seasonal changes in variation of dorsal scale pattern of *Aedes aegypti* (L.) (Diptera: Culicidae) in Surabaya, Indonesia. *Medical Entomology and Zoology*, **54**, 73–80.

Vazeille K, Zouache A, Vega-Rua J-M *et al.* (2016) Importance of mosquito "quasispecies" in selecting an epidemic arthropod-borne virus. *Nature Scientific Reports*, **6**, 29564.

Wallis GP, Tabachnick WJ, Powell JR (1983) Macrogeographic genetic variation in a human commensal: *Aedes aegypti,* the yellow fever mosquito. *Genetical Research*, **41**, 241–258.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

A.G.S. and J.R.P. conceived and designed the experiments, analysed the data and wrote the manuscript. A.G.S. and J.B.R. performed the experiments. D.A., A.B., O.C.-A., D.D.C., M.Ch., M.C., J.b.E., I.F.-S., H.A.K., B.K., E.I.M.K., L.D.K., V.K., A.L.-S., J.L., A.M.Jr., M.V.M., C.P., A.P., N.R., S.B.R., A.A.S., R.M.S.-C., G.S., C.A.S., W.J.T. and A.T. contributed reagents/materials and performed field collections.

Data accessibility

Raw allele frequencies are available at VectorBase.org, Population Biology Project ID: VBP0000138.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Discriminant Analysis of Principal Components on microsatellite allele frequencies for A) The Americas, B) United States of America (excluding Exeter population), c) Mexico, and D) South America.

Fig. S2 Evolutionary scenarios of *Aedes aegypti aegypti* re-colonization of Rabai, Kenya, evaluated using Approximate Bayesian Computation inference as implemented by the DIYABC software (Cornuet *et al*. 2014). Scenarios include four populations: Africa (Aaf), America (Aaa), Asia (Aaa), and Rabai (Aaa). $N = 32$ for each sampled population. All samples available from Aaa Rabai were included. Subsampling was required from the larger populations (see Materials and methods). T0 represents the most recent time point and increasing values of T go back in time. Scenario 1: Africa to America to Asia to Rabai; Scenario 2: Africa to Rabai (before Asia colonization); Scenario 3: Africa to America to Rabai (after Asia colonization); Scenario 4: Africa to America to Rabai (before Asia colonization); Scenario 5: Africa to Rabai (before America colonization); and Scenario 6: Africa to Rabai (after Asia colonization). Posterior probabilities are shown for each scenario. The best-fit scenario was Scenario 3, indicated by an asterisk (*). For more details see Materials and methods and Table S5.

Fig. S3 Discriminant Analysis of Principal Components showing two main genetic clusters: 1) the putative *Ae. aegypti queenlandensis* strain and 2) the main *Ae. aegypti* cluster including all Aaa and Aaf populations.

Table S1 ABC analysis Africa to Asia.

Table S2 ABC analysis Africa to Rabai.

Table S3 Detailed collection information of *Aedes aegypti* populations including geographic region, number of individuals per sample, geographic coordinates of each location, and number of generations reared in the laboratory previous to genetic analysis.

Table S4 Population genetics statistics for the 79 individual Ae. aegypti populations analyzed in this study.

Table S5 F_{ST} population pair-wise matrix including the 79 Ae. *Aedes aegypti* populations from this study.