PROCEEDINGS OF THE ROYAL SOCIETY B

Proc. R. Soc. B (2011) **278**, 2446–2454
doi:10.1098/rspb.2010.2469
Published online 12 January 2011

Worldwide patterns of genetic differentiation imply multiple 'domestications' of *Aedes aegypti*, a major vector of human diseases

Julia E. Brown[1,*], Carolyn S. McBride[2], Petrina Johnson[3], Scott Ritchie[3], Christophe Paupy[4], Hervé Bossin[5], Joel Lutomiah[6], Ildefonso Fernandez-Salas[7], Alongkot Ponlawat[8], Anthony J. Cornel[9], William C. Black IV[10], Norma Gorrochotegui-Escalante[10], Ludmel Urdaneta-Marquez[10], Massamba Sylla[10], Michel Slotman[11], Kristy O. Murray[12], Christopher Walker[12] and Jeffrey R. Powell[1]

[1]*Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520, USA*
[2]*Laboratory of Neurogenetics and Behavior, The Rockefeller University, New York, NY 10065, USA*
[3]*School of Public Health, Tropical Medicine and Rehabilitation Sciences, James Cook University, Cairns, Queensland 4870, Australia*
[4]*Caracterisation et Controle des Populations de Vecteurs, Institut de Recherche pour le Développement (IRD), UR016, Montpelier 34394, France*
[5]*Medical Entomology Laboratory, Institut Louis Malardé, BP 30, 98713 Papeete, French Polynesia*
[6]*Arbovirology/Hemorrhagic Fevers Laboratory, Centre for Virus Research, Kenya Medical Research Institute (KEMRI), PO Box 54628, Nairobi, Kenya*
[7]*Laboratorio de Entomología Médica, Facultad de Ciencias Biológicas, Universidad Autónoma de Nuevo León, Postal 391, San Nicolás de los Garza, Nuevo León 66450, Mexico*
[8]*Vector Biology and Control Section, Department of Entomology, Armed Forces Research Institute of Medical Sciences (AFRIMS), Bangkok 10400, Thailand*
[9]*Mosquito Control Research Laboratory, Department of Entomology, University of California at Davis, Parlier, CA 93648, USA*
[10]*Department of Microbiology, Immunology and Pathology, Colorado State University, Fort Collins, CO 80523, USA*
[11]*Department of Entomology, Texas A&M University, College Station, TX 77843, USA*
[12]*Center for Infectious Diseases, University of Texas Health Science Center, School of Public Health, Houston, TX 77030, USA*

Understanding the processes by which species colonize and adapt to human habitats is particularly important in the case of disease-vectoring arthropods. The mosquito species *Aedes aegypti*, a major vector of dengue and yellow fever viruses, probably originated as a wild, zoophilic species in sub-Saharan Africa, where some populations still breed in tree holes in forested habitats. Many populations of the species, however, have evolved to thrive in human habitats and to bite humans. This includes some populations within Africa as well as almost all those outside Africa. It is not clear whether all domestic populations are genetically related and represent a single 'domestication' event, or whether association with human habitats has developed multiple times independently within the species. To test the hypotheses above, we screened 24 worldwide population samples of *Ae. aegypti* at 12 polymorphic microsatellite loci. We identified two distinct genetic clusters: one included all domestic populations outside of Africa and the other included both domestic and forest populations within Africa. This suggests that human association in Africa occurred independently from that in domestic populations across the rest of the world. Additionally, measures of genetic diversity support *Ae. aegypti* in Africa as the ancestral form of the species. Individuals from domestic populations outside Africa can reliably be assigned back to their population of origin, which will help determine the origins of new introductions of *Ae. aegypti*.

Keywords: *Aedes aegypti aegypti*; *Aedes aegypti formosus*; human habitats; microsatellites; evolution; mosquito genetics

1. INTRODUCTION

Humans have had an enormous influence on the global environment and are known to have shaped genetic variation within other species in the recent past [1–4]. Much work in conservation biology has focused on evolutionary

* Author for correspondence (julia.brown@yale.edu).

Electronic supplementary material is available at http://dx.doi.org/10.1098/rspb.2010.2469 or via http://rspb.royalsocietypublishing.org.

Received 11 November 2010
Accepted 13 December 2010

This journal is © 2011 The Royal Society

genetic changes associated with human effects on natural populations of a variety of species [5–7]. However, relatively few studies have investigated the adaptation of other species to human environments, or human commensalism, particularly at the population genetic level [1,8–10]. The processes by which certain populations of a species colonize and adapt to urban or other domestic habitats remain poorly understood.

Colonization of human habitats is particularly interesting in arthropod disease vector systems where humans provide the arthropod with the blood it needs to synthesize eggs and reproduce. Such extreme ecological overlap between human and vector populations has important implications for public health and disease transmission. Humans have a stable community structure that provides constant access to food (blood) and plentiful breeding habitats. Thus, once a founding vector population has entered a human environment, selection for specialization can occur [11]. For example, several mosquito populations that breed specifically in human domestic habitats and urban areas are specialists that have adapted to feed on humans, and prefer them over other species [12–14]. Species that breed in human-generated containers can easily spread around the world through human movement of goods [15].

The mosquito species *Aedes aegypti* is a prime study system for research on the evolution of human association with public health consequences. The species is a vector of several human diseases, including yellow fever and dengue fever, and has invaded much of the tropical and subtropical world over the past few centuries [16]. Though an effective vaccine exists for yellow fever, there is currently no vaccine for dengue. Indeed, dengue is on the rise across the world, including in sub-Saharan Africa, where major outbreaks were rare in the past [17]. The virus affects an estimated 50 million people per year, with 2.5 billion people, or 40 per cent of the global population, currently at risk of infection [18].

Aedes aegypti is often treated as a homogeneous species in its role as a disease vector, but the species is actually rich in genetic, morphological and ecological variation [19–26]. Mattingly [26] described two subspecies based on overall body colour and the extent of white scaling on the first abdominal tergite. The darker and presumably ancestral *Aedes aegypti formosus* was reportedly confined to Africa, where it tended to breed in forested habitats and was predominantly zoophilic (preferring to bite non-human animals). The lighter *Aedes aegypti aegypti*, on the other hand, was distributed throughout the tropics outside Africa, where it bred in human-generated containers and was strongly anthropophilic (preferring to bite humans). Mattingly himself, however, acknowledged that the true situation was more complex, particularly within Africa. Mattingly [26,27] and McClelland [25] documented continuous variation in one of the morphological characters (scaling patterns) that was supposed to be diagnostic of the two subspecies. Nevertheless, allozyme studies from the 1970s and 1980s identified a clear genetic structure and supported the idea that *Ae. aegypti* harbours genetically distinct forms with discrete geographical distribution and mean differences in scaling pattern and ecological preference [20–23,28]. Interestingly, two forms identified as subspecies *aegypti* and *formosus* were shown to coexist in a few places along the coast of East Africa, including the Rabai district of Kenya [23,24,28,29], where the former bred inside homes in villages and the latter bred in surrounding forests.

Although indoor-breeding populations from coastal East Africa provide the best-documented examples of domesticity among African populations of *Ae. aegypti*, there is ample evidence of human association in West Africa as well [25,26,30–36]. Unlike the situation in East Africa, however, domestic populations in West Africa do not appear to differ genetically from nearby forest populations [30–33]. This raises the question as to whether they represent independent incursions of *Ae. aegypti* into human habitats, or are evolutionarily related to domestic populations from other areas of the world.

We used molecular genetic markers to investigate whether all worldwide human-associated populations of *Ae. aegypti* represent a single domestication event, or whether human-associated African populations represent independent examples of domestication within the species. In this paper, we make no definitive *a priori* subspecies designations, but address how our genetic data correspond to other authors' definitions in the discussion. Using 12 polymorphic microsatellites, we examined 1152 *Ae. aegypti* individuals from 24 locations spanning 13 countries and five continents. A comprehensive dataset of this sort should allow us to assess population structure and elucidate the origin of domestic populations, particularly in West Africa.

2. METHODS

(a) *Sampling/collections*

During the period from 2004 to 2010, *Ae. aegypti* were collected from 24 locations across the world in 13 countries on five continents (table 1; see the electronic supplementary material, figure S1, for a map). Mosquitoes sampled from the field were collected as eggs or larvae, and reared to adults for identification and preservation in 70 to 100 per cent EtOH or at −80°C. To avoid sampling siblings laid by a single female in the same oviposition container, mosquitoes were collected from multiple containers across numerous sampling sites (e.g. homes). This should not be a major problem even without this precaution, however, since several studies indicate that *Ae. aegypti* females deposit eggs across a number of oviposition sites [37–41]. Mosquitoes analysed came directly from the field, except for those from Pijijiapan (Mexico) and Tahiti (French Polynesia), which passed through one generation in the laboratory before being sampled, and populations from Thailand and Venezuela, which passed through two laboratory generations (table 1). All mosquito lines that passed through the laboratory were set up to be as representative as possible of field populations. Colonies were established in large cages by several hundred to over 1000 mosquitoes from the field.

(b) *Genetic methods*

Whole genomic DNA was extracted individually from 1152 larval or adult mosquitoes with DNeasy kits (Qiagen) following the manufacturer's protocols. Individual genotypes were scored at 12 microsatellite loci (electronic supplementary material, table S1), including eight previously published markers [42] and four new loci developed for this study. The four new loci were developed by screening the genomic contigs available on VectorBase (http://aaegypti.vectorbase.org/

Table 1. Collection information for *Aedes aegypti* population samples examined at 12 microsatellite loci. A population was considered domestic if the mosquitoes were collected in or around human habitation (village, town, city). Morphological subspecies classifications, where possible, were made based on the presence or absence of white scales on the first abdominal tergite [26].

population	region	habitat type	gen. in lab	morphological subspecies classification	*n*	year collected
Vaca Key, FL, USA	North America	domestic	0	—	42	2009
Conch Key, FL, USA	North America	domestic	0	—	42	2006
Palm Beach County, FL, USA	North America	domestic	0	—	43	2006
Houston, TX, USA	North America	domestic	0	—	29	2009
Coatzacoalcos, Mexico [67][a]	North America	domestic	0	—	50	2008
Pijijiapan, Mexico [67][a]	North America	domestic	1	—	47	2008
Dominica	Caribbean	domestic	0	—	95	2009
Bolivar, Venezuela [68][a]	South America	domestic	2	—	48	2004
Zulia, Venezuela [68][a]	South America	domestic	2	—	47	2004
Rayong, Thailand	Asia	domestic	2	—	48	2009
Prachuabkhirikan, Thailand	Asia	domestic	2	—	47	2009
Tahiti, French Polynesia	South Pacific	domestic	1	—	48	2010
Cairns, Australia	Australia	domestic	0	—	48	2009
Townsville, Australia	Australia	domestic	0	—	47	2009
Rabai, Kenya	East Africa	domestic/forest	0	*aegypti* + *formosus*	71	2009
Bundibugyo, Uganda	East Africa	domestic	0	*aegypti* + *formosus*	47	2009
Kichwamba, Uganda	East Africa	domestic	0	*aegypti* + *formosus*	48	2009
Dakar, Senegal [31][a]	West Africa	domestic	0	*aegypti*	43	2005
N'goye, Senegal [31][a]	West Africa	domestic	0	*aegypti*	45	2007
Koungheul, Senegal [31][a]	West Africa	domestic	0	*aegypti* + *formosus*	46	2006
Goudiry, Senegal [31][a]	West Africa	domestic	0	*formosus*	46	2007
PK-10, Senegal [31][a]	West Africa	forest	0	*formosus*	48	2006
Yaounde, Cameroon	West Africa	domestic	0	*aegypti*	47	2009
Bijagos, Guinea-Bissau	West Africa	domestic	0	—	30	2009

[a]Samples collected for previous studies.

index.php) for trinucleotide repeats with 8–20 uninterrupted repeats that did not contain repetitive sequences in either flanking region. Primers were designed using PRIMER 3 [43], and blasted against the genome in VectorBase to ensure that the loci were single copy and located on different supercontigs to minimize the potential for genetic linkage.

We paired microsatellite loci based on non-overlapping size ranges and amplified each pair in a multiplex PCR reaction with a single fluorescent M13 primer, two forward primers with M13 tails, and two reverse primers [44]. All PCR reactions (10 μl) contained 1× Type-it Multiplex PCR Master Mix (Qiagen), 25 nM of each forward primer, 250 nM of each reverse primer and 500 nM of fluorescently labelled M13 primer. All loci were amplified according to the thermocycling conditions in Slotman *et al.* [42]. PCR products were run on an Applied Biosystems 3730*xl* DNA Genetic Analyser with a GS 500 Rox internal size standard (Applied Biosystems). Microsatellite primer sequences, multiplex pairings and fluorescent primers are given in electronic supplementary material, table S1. Microsatellite alleles were scored using the software GENEMAPPER (Applied Biosystems). We screened 30 candidate microsatellites, and from those chose 12 loci that reliably amplified in all samples and showed sufficient variability. This maximized information and allowed us to obtain a consistent dataset for a heterogeneous set of population samples. All 12 microsatellite loci reside on different genomic supercontigs (VectorBase).

(c) *Analyses*

For all analyses described below, individuals from a single geographical location were treated as a single population, with one exception. Rabai, Kenya, has previously been shown to harbour two genetically distinct forms of *Ae. aegypti* (§1). Preliminary analyses of the current dataset supported this finding, and we therefore treated Rabai individuals from the two genetic groups as independent populations.

All microsatellite loci were tested for within-population deviations from Hardy–Weinberg equilibrium (HWE) using the web version of the software GENEPOP [45,46]. The same program was used to test all pairs of loci across all populations for linkage disequilibrium (LD). Markov chain parameters were set at 10 000 dememorizations, 1000 batches and 10 000 iterations per batch for both HWE and LD. GENEPOP was also used to compute allele frequencies for all loci across populations. Observed and expected heterozygosities for each population were computed using the software ARLEQUIN 3.5 [47].

We assessed overall genetic differentiation among populations by calculating F_{ST} for all population pairs in ARLEQUIN 3.5 [47]. We evaluated hierarchical patterns of population structure and tested our ability to assign individuals to their population of origin using the Bayesian clustering method implemented in the software program STRUCTURE v. 2.3 [48]. STRUCTURE identifies genetic clusters and assigns all individuals to these clusters without any *a priori* information regarding sampling locations. To determine the most likely number of clusters (K), we conducted five independent runs for each $K = 1$–26. For all runs, we assumed an admixture model and independent allele frequencies, and used a burn-in value of 100 000 iterations followed by 500 000 replications. The most likely number of clusters was determined following the guidelines of Pritchard *et al.* [48] and by calculating ΔK,

which is based on the second-order rate of change of the likelihood distribution between values of K [49]. STRUCTURE results were visualized using the program DISTRUCT [50].

Additionally, self-assignment tests were completed for a subset of populations (§3) in GENECLASS2 [51] to assess the degree to which an individual mosquito could be formally assigned to a specific population. After creating reference populations based on geography and clusters identified by STRUCTURE, this analysis allowed us to take individuals out of the analysis and then 'ask' where they would be assigned.

To assess relatedness among populations, we calculated Cavalli-Sforza & Edwards's [52] chord distance for each pair of populations in PHYLIP 3.69 (GENDIST module) and used the resulting distance matrix to create two types of plot. We first used it to construct a neighbour-joining tree in MEGA4 [53], with node confidence inferred via 1000 bootstrap replicates in PHYLIP 3.69 (modules SEQ-BOOT, GENDIST, NEIGHBOUR and CONSENSE). The neighbour-joining network should not be taken as a true phylogeny since microsatellites are not ideal markers for recovering evolutionary history [54]. Rather, the analysis should be considered as an additional method to assess genetic clustering of populations. Second, we used the distance matrix to construct a Euclidean-based non-metric multi-dimensional scaling plot in the software PAST [55].

Finally, to assess whether African *Ae. aegypti* represent the ancestral form of the species, we calculated allelic richness and private allelic richness for each population using the program HP-RARE [56,57]. HP-RARE uses rarefaction to correct for sample size. The program sampled 21 individuals (42 allele copies) at random from each population to match the smallest population sample size.

3. RESULTS

(a) *Marker validation*

Forty-two of 300 population-by-locus-specific F_{IS} values deviated significantly from Hardy–Weinberg expectations after sequential Bonferroni correction (electronic supplementary material, table S2). Of the 25 population-specific tests for each marker, anywhere from zero (AC2 or B2) to eight to nine (AC5 and A9) tests were significant. In at least one of the latter cases (A9), the significant tests resulted from an excess of homozygotes, probably owing to null alleles. The Walhund effect may also have contributed to some of the skewed values if populations were not truly panmictic since individuals were sampled across multiple homes/containers/localities with unknown substructure. Though 160 of 1650 (9.7%) locus-by-locus tests for LD remained significant after Bonferroni correction, no two loci were consistently correlated across populations, indicating that close physical genetic linkage is unlikely. This is consistent with each microsatellite residing on a different supercontig of the genome assembly. It is possible, however, that chromosomal inversions link physically distant markers in some but not all populations [58].

To assess the robustness of our results, all analyses were repeated using a 10-locus dataset that did not include the two most problematic loci (AC5 and A9). This smaller dataset produced the same pattern of population clustering and hierarchical relatedness as the full 12-locus dataset, and so we report only the latter here.

(b) *Worldwide population structure and ancestry*

The Bayesian analysis of worldwide population structure clearly indicated that the most likely number of clusters was two when evaluated using ΔK (electronic supplementary material, figure S2). The two clusters cleanly separated African and non-African populations with a single exception (figure 1*a*). The sample from Rabai, Kenya, included individuals from both clusters, a result consistent with previous work in this area [20–23,28]. The otherwise cohesive African group contained West and East African populations from urban to forest environments, and individuals that have been identified as both *Ae. ae. aegypti* and *Ae. ae. formosus* based on morphology (table 1). Very little mixing between groups was detected, other than in the population from Ngoye, Senegal (figure 1*a*). STRUCTURE was run separately on individuals from the African and non-African clusters to assess additional structuring within groups. We hereafter refer to the non-African cluster as pantropical since it encompasses populations throughout the tropics/subtropics including a single representative within Africa. Among pantropical populations, geographically relevant clustering was detected out to $K = 12$ (figure 1*b*). Individual mosquitoes were assigned correctly to their population of origin with high accuracy. African populations showed weaker (but nevertheless geographically relevant) structure at $K = 2$ and $K = 4$ (figure 1*c*), both of which appeared as peaks in a ΔK analysis. Individuals from Goudiry and PK-10, Senegal, could be assigned reliably back to their population of origin at $K = 4$ (figure 1*c*). Interestingly, the clusters detected among the African populations were not associated with ecology (urban versus forest populations) or with the two morphological subspecies.

To confirm the ability of our markers to assign unknown pantropical mosquitoes back to their population of origin, self-assignment tests were conducted in GENECLASS2 [51], which correctly assigned 90 per cent of the pantropical individuals back to their genetic cluster as determined by STRUCTURE ($K = 12$). When geographical locations were used as the reference populations, 87.7 per cent of individuals were correctly assigned back to their population of origin.

A non-metric multi-dimensional scaling analysis of all populations based on Cavalli-Sforza and Edwards's chord distances recovered the same two genetic groups as the STRUCTURE analysis. However, the pantropical form from Rabai, Kenya, is shown to be as different from the other populations in the pantropical cluster (those from outside Africa) as it is from the African populations (figure 2*a*). This result is also reflected in mean pairwise F_{ST} values between groups (figure 2*b*; raw pairwise F_{ST} data listed in electronic supplementary material, table S3). A neighbour-joining tree based on Cavalli-Sforza and Edwards's chord distances strongly supports the same separation of groups as above. African populations appear at the base of the network with no clear geographical or ecological structuring (figure 2*c*, red branches). All pantropical populations cluster together with strong support (figure 2*c*, dark blue branches).

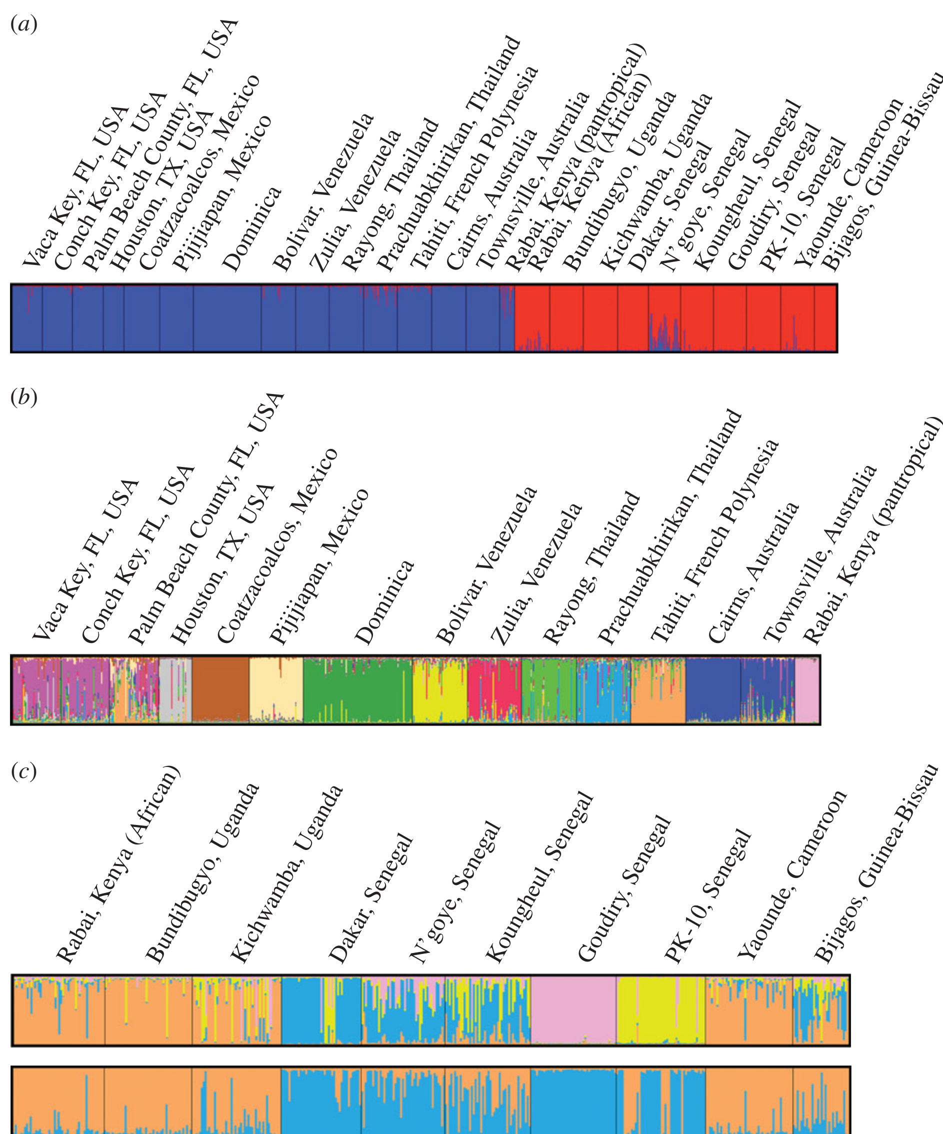

Figure 1. STRUCTURE bar plots. Each vertical bar represents a single individual. The height of each colour represents the probability of assignment to that cluster. (*a*) $K = 2$ for all sampled mosquitoes in all populations; (*b*) $K = 12$ for pantropical cluster; and (*c*) $K = 4$ (top) and $K = 2$ (bottom) for African cluster.

Genetic groupings within this pantropical group mirror geography fairly closely, but often with weak bootstrap support. The pantropical form from Rabai, Kenya, clusters with other pantropical populations, but is subtended by a long branch at the base of the group (light blue branch).

When whole regions were considered, corrected allelic richness across loci was 11.03 within the African populations and 7.46 within the pantropical populations. Unlike earlier allozyme analyses [20–22], a number of unique alleles were found in both groups. Private allelic richness was also substantially higher in the African group (5.19) than in the pantropical group (1.62). Mean observed heterozygosity was 0.613 across African populations and 0.529 in the pantropical populations. Diversity statistics for all populations are given in the electronic supplementary material, as are allele frequencies across all loci (electronic supplementary material, table S4). Microsatellite genotypes for all individuals have been deposited in the Dryad digital repository (doi:10.5061/dryad.8065).

4. DISCUSSION

(a) *Discrete genetic clusters within* Aedes aegypti

High genetic distances between pantropical populations and African populations strongly support the existence of two divergent groups within *Ae. aegypti*. These groups correspond with previous descriptions of subspecies *Ae. ae. aegypti* and *Ae. ae. formosus* based on geography and genetics [16,20–22,25,26]. However, it is now clear that the single morphological trait commonly used for subspecies identification (i.e. white scaling on the first abdominal tergite) is not a reliable method to distinguish between the two divergent genetic forms of *Ae. aegypti*, particularly in West Africa [16,25]. Indeed, we currently know of no morphological character that can reliably differentiate them, though background cuticle and scale colour, traits ignored in most previous studies, deserve further attention. In light of our results, the subspecies of *Ae. aegypti* clearly need to be revised. Without more detailed morphological and other natural history information, we do not feel comfortable trying to redefine the subspecies in this study. However, for the sake of

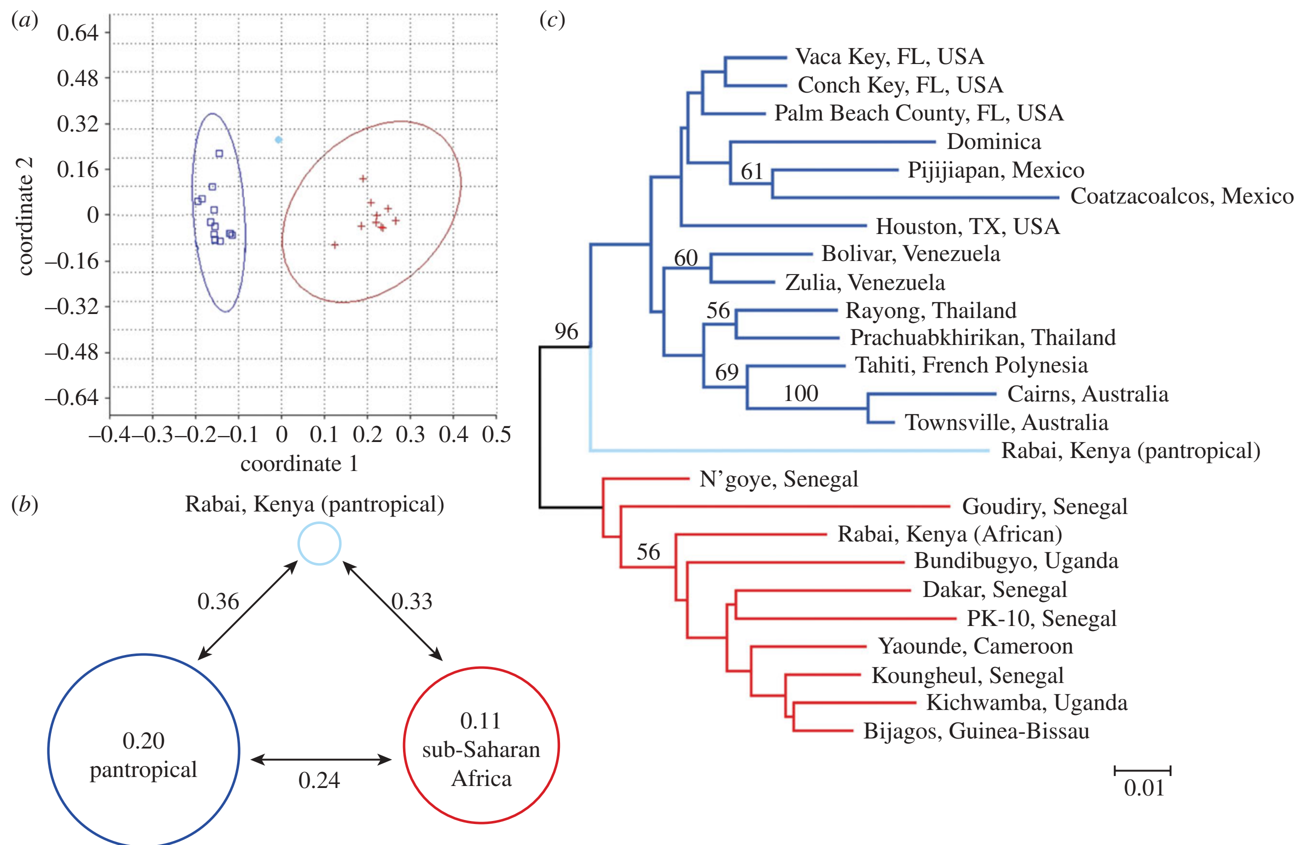

Figure 2. (*a*) Non-metric multi-dimensional scaling of *Ae. aegypti* populations based on pairwise Cavalli-Sforza and Edwards's chord distance. Pantropical populations are represented by dark blue squares with the exception of that from Rabai, Kenya, which is represented by a light blue circle. African populations are represented by red crosses. Ovals represent 95% CI. (*b*) Average pairwise F_{ST} values between major groups of *Ae. aegypti*. Circles are scaled to the number of populations within each group sampled in this study. (*c*) Neighbour-joining cluster analysis (unrooted) based on Cavalli-Sforza and Edwards's chord distance. Apparent root is for visual purposes only. Numbers indicate bootstrap values above 50%.

consistency with previous literature, throughout the rest of this discussion we will refer to all populations in the African genetic cluster as *Ae. ae. formosus* and those in the pantropical cluster, including that from Rabai, Kenya, as *Ae. ae. aegypti*.

Previous authors have inferred that Africa is the ancestral homeland of *Ae. aegypti* [27], from which a founding non-African population split allopatrically and spread across the globe [16,27,29]. Our own measures of within-population genetic diversity support this hypothesis. Values of heterozygosity, allelic richness and private allelic richness were substantially higher in *Ae. ae. formosus* than across populations of *Ae. ae. aegypti*. The distinction between the two subspecies/groups is further supported by the overall lack of mixing between them. This appeared true even in the one location, Rabai, Kenya (East Africa), where they occur sympatrically [20,21,23,28], though more extensive sampling in this region is needed. The pantropical *Ae. ae. aegypti* population in Rabai, Kenya, may represent a reintroduction of that subspecies from outside sub-Saharan Africa, which has become genetically isolated [16]. However, since the pantropical population in Rabai appears so divergent from all other *Ae. aegypti* populations, we are performing ongoing studies to determine the origin of these mosquitoes. Though pure populations of *Ae. ae. aegypti*, as defined based on our genetic results, were not found in West Africa, we cannot rule out the possibility that *Ae. ae. aegypti* has been introduced to certain locations in West Africa (e.g. N'goye, Senegal) where it has interbred with the native *Ae. ae. formosus*.

(b) ***Independent development of human association***

Mattingly [26,27] emphasized the ecological contrast between wild, zoophilic populations of *Ae. ae. formosus* in Africa and domestic, anthropophilic populations of *Ae. ae. aegypti* outside Africa. More recent studies continue to document strong domesticity/anthropophily in pantropical *Ae. ae. aegypti* populations [13,59]. However, the data presented here make it clear that the genetically cohesive *Ae. ae. formosus* is much more variable ecologically than traditional descriptions would suggest. Indeed, it encompasses both wild and domestic populations across Africa. All samples from urban domestic populations in West Africa cluster clearly with *Ae. ae. formosus*, rather than with any populations of *Ae. ae. aegypti* (figures 1 and 2). This evidence supports the view that sympatric, or near sympatric, breeding of domestic and forest forms as studied in West Africa is not a strict analogue of the situation in Rabai, Kenya (coastal East Africa), where the forms are very different from a genetic perspective [16,25,26,35]. Furthermore, even within West Africa, human-associated populations do not cluster separately from forest populations (figures 1 and 2), suggesting independent invasions of human habitats by *Ae. ae. formosus* across the African landscape. This

phenomenon of multiple independent colonizations of human habitats has been corroborated in studies of other taxa [8].

One explanation for the observed patterns is that *Ae. ae. formosus* is a generalist that invades human-associated niches where available without specifically adapting to them. Behavioural flexibility and adaptive plastic responses have been shown to affect urban invasion success in other species [60,61], and *Ae. aegypti* has been documented to be flexible in its breeding sites and behaviour in certain locations [62,63]. Alternatively, invasion of human-associated habitats by *Ae. ae. formosus* in Africa may be driven by specific adaptation to those environments, as it is in pantropical *Ae. ae. aegypti* (including populations from Rabai, Kenya) [16,23,24,28,29]. If so, the genetic dissimilarity between domestic populations within and outside of Africa suggests that this example of ecological specialization on human habitats within *Ae. ae. formosus* was independent of that associated with the split between the two subspecies. A third possibility, not exclusive of those above, is that although domestic populations of the two subspecies do not represent a single cohesive evolutionary unit, some 'domestic' alleles that segregate within *Ae. ae. formosus*, helping certain populations to exploit human habitats, are the same as those that appear to be fixed in pantropical *Ae. ae. aegypti*. They could have originated in subspecies *formosus* and risen to high frequency in the founding population of *Ae. ae. aegypti*, or they may have originated in subspecies *aegypti* and introgressed into *formosus* in modern times. In the latter case, presumed neutral markers (e.g. microsatellites) characteristic of *Ae. ae. aegypti* may not persist in African domestic populations despite introgression of the genomic regions controlling adaptive ecological traits.

Distinguishing between the possibilities laid out above will require phenotypic and genotypic characterization of domestic *Ae. ae. formosus* populations. The behaviours of West African populations, in particular, are severely understudied in comparison with those of pantropical *Ae. ae. aegypti*. If traditional domestic phenotypes (e.g. anthropophily, preference to breed in artificial containers) are found within *Ae. ae. formosus* populations, it would be informative to characterize their genetic basis and compare it with the genetic basis of the same behaviours in East African or worldwide *Ae. ae. aegypti* populations.

(c) *Public health implications*

Within the pantropical form, individuals could be assigned back to their population of origin with high probabilities (figure 1*b*). Owing to this strong assignment ability, our combination of markers will prove useful in tracking down the origins of new introductions of *Ae. aegypti*. The manageable number of microsatellite loci in our set will make it easier for scientists in disease endemic countries to perform these analyses locally. Overall, the pattern of genetic relatedness among pantropical populations (figure 2*c*) shows a striking resemblance to geography, despite low bootstrap support at certain nodes. Ongoing DNA sequencing of genes from these samples will potentially shed further light and add rigour to the historical inferences of the species. Since only a fraction of existing *Ae. aegypti* populations have been examined, more variants and relationships probably remain to be discovered across the world.

We have verified that most domestic populations of *Ae. aegypti* in Africa are genetically different from domestic populations found elsewhere in the world. Therefore, we cannot rely on studies of *Ae. ae. aegypti* to inform us of the threat posed by *Ae. ae. formosus* in Africa. Since *Ae. ae. formosus* is generally less competent at acquiring and transmitting dengue and yellow fever viruses than *Ae. ae. aegypti* [64–66], care must be taken with eradication campaigns against African *Ae. ae. formosus* to avoid opening up niches for potentially more virally competent worldwide *Ae. ae. aegypti* populations to invade. So far, genetically pantropical mosquitoes have been confined to restricted indoor niches in East African coastal areas, but those populations, as well as migrants brought in from elsewhere, could potentially occupy a broader range if natural *Ae. ae. formosus* populations were removed. However, despite mean differences in competence between subspecies, *Ae. ae. formosus* itself is also variable in its ability to replicate and transmit viruses [31], and more work is needed to characterize the vector competence and behaviour of diverse *Ae. ae. formosus* populations.

Whether the basis of human association in *Ae. ae. formosus* is genetic or plastic, the fact remains that African populations of *Ae. aegypti* are breeding in human environments with easy access to humans as hosts for obtaining blood. Overlap of mosquito and human populations provides opportunities for epidemic disease spread, which can even be caused by *Ae. aegypti* populations with low competence for viral replication and transmission [35]. *Aedes aegypti formosus* is apparently able to continue invading human habitats, and this phenomenon is likely to increase as the sub-Saharan African landscape becomes increasingly human-dominated and human environments represent a substantial part of available niches.

The evolution of human commensalism is a field that is likely to grow in importance as the human population continues to expand and mould environments across the world. Our findings suggest that close human association may arise multiple times within one of the most important vectors of human diseases. In this relatively new field, more research is needed across additional organisms to begin to develop a sound theoretical framework.

We thank Greg Lanzaro, Durland Fish, Charles Jeannin, Tuterarii Paoaafaite, Henri Frogier, Barry J. Beaty, Gustavo Ponce García, Adriana E. Flores, Larry Hribar, Rosemary Sang, Barry Miller and John-Paul Mutebi for their help in collecting samples, Walter Tabachnick and Leslie Vosshall for their comments on the manuscript, and Jennifer Simpson, Janelle Winters and Vanessa Obas for their efforts in the laboratory. J.E.B. was supported by NIH Predoctoral Training in Genetics (T32-GM07499-33) and the Yale Institute for Biospheric Studies (YIBS), Center for Field Ecology pilot grant. Senegal collections were supported by NIH R01AI083368 to M.Sy. and W.C.B. Collections in Rabai, Kenya, were partly supported by a grant to Richard Axel and Leslie Vosshall from the Foundation for the National Institutes of Health through the Grand Challenges in Global Health Initiative, and by an investigator award to Leslie Vosshall from the Howard Hughes Medical Institute. C.W. and collections in Texas, USA, were supported by a grant from the United States Department of Defense, Telemedicine and Advanced Technology Research Center (TexSHIELD W81XWH-07-2-0031) to K.O.M. Development of molecular markers was

supported by YIBS. The rest of the study was supported by NIH RO1 AI046018 (J.R.P.).

REFERENCES

1 Das, A., Mohanty, S. & Stephan, W. 2004 Inferring the population structure and demography of *Drosophila ananassae* from multilocus data. *Genetics* **168**, 1975–1985. (doi:10.1534/genetics.104.031567)

2 Ashley, M. V., Willson, M. F., Pergams, O. R. W., O'Dowd, D. J., Gende, S. M. & Brown, J. S. 2003 Evolutionarily enlightened management. *Biol. Conserv.* **111**, 115–123. (doi:10.1016/S0006-3207(02)00279-3)

3 Bradshaw, R. H. W. 2004 Past anthropogenic influence on European forests and some possible genetic consequences. *For. Ecol. Manage.* **197**, 203–212. (doi:10.1016/j.foreco.2004.05.025)

4 Templeton, A. R., Shaw, K., Routman, E. & Davis, S. K. 1990 The genetic consequences of habitat fragmentation. *Ann. Mo. Bot. Gard.* **77**, 13–27. (doi:10.2307/2399621)

5 Smith, T. B. & Bernatchez, L. 2008 Evolutionary change in human-altered environments. *Mol. Ecol.* **17**, 1–8. (doi:10.1111/j.1365-294X.2007.03607.x)

6 Schmidt, N. M. & Jensen, P. M. 2003 Changes in mammalian body length over 175 years—adaptations to a fragmented landscape? *Conserv. Ecol.* **7**, issue 2, article 6. See http://www.consecol.org/vol7/iss2/art6.

7 Conover, D. O. & Munch, S. B. 2002 Sustaining fisheries yields over evolutionary time scales. *Science* **297**, 94–96. (doi:10.1126/science.1074085)

8 Evans, K. L. *et al.* 2009 Independent colonization of multiple urban centres by a formerly forest specialist bird species. *Proc. R. Soc. B* **276**, 2403–2410. (doi:10.1098/rspb.2008.1712)

9 Keller, A. 2007 *Drosophila melanogaster*'s history as a human commensal. *Curr. Biol.* **17**, R77–R81. (doi:10.1016/j.cub.2006.12.031)

10 Partecke, J., Gwinner, E. & Bensch, S. 2006 Is urbanisation of European blackbirds (*Turdus merula*) associated with genetic differentiation? *J. Ornithol.* **147**, 549–552. (doi:10.1007/s10336-006-0078-0)

11 Schofield, C. J., Diotaiuti, L. & Dujardin, J. P. 1999 The process of domestication in Triatominae. *Mem. Inst. Oswaldo Cruz* **94**, 375–378. (doi:10.1590/S0074-02761999000700073)

12 Lyimo, I. N. & Ferguson, H. M. 2009 Ecological and evolutionary determinants of host species choice in mosquito vectors. *Trends Parasitol.* **25**, 189–196. (doi:10.1016/j.pt.2009.01.005)

13 Harrington, L. C., Edman, J. D. & Scott, T. W. 2001 Why do female *Aedes aegypti* (Diptera: Culicidae) feed preferentially and frequently on human blood? *J. Med. Entomol.* **38**, 411–422. (doi:10.1603/0022-2585-38.3.411)

14 Mukwaya, L. G. 1977 Genetic control of feeding preferences in the mosquitoes *Aedes* (*Stegomyia*) *simpsoni* and *aegypti*. *Physiol. Entomol.* **2**, 133–145. (doi:10.1111/j.1365-3032.1977.tb00091.x)

15 Lounibos, L. P. 2002 Invasions by insect vectors of human disease. *Annu. Rev. Entomol.* **47**, 233–266. (doi:10.1146/annurev.ento.47.091201.145206)

16 Tabachnick, W. J. 1991 Evolutionary genetics and insect borne disease. The yellow fever mosquito, *Aedes aegypti*. *Am. Entomol.* **37**, 14–24.

17 Gubler, D. J. 2004 The changing epidemiology of yellow fever and dengue, 1900 to 2003: full circle? *Comp. Immunol. Microbiol. Infect. Dis.* **27**, 319–330. (doi:10.1016/j.cimid.2004.03.013)

18 World Health Organization. 2009 Dengue and dengue hemorrhagic fever. WHO Fact Sheet No. 117. See http://www.who.int/mediacentre/factsheets/fs117/en/.

19 Morlais, I. & Severson, D. W. 2003 Intraspecific DNA variation in nuclear genes of the mosquito *Aedes aegypti*. *Insect Mol. Biol.* **12**, 631–639. (doi:10.1046/j.1365-2583.2003.00449.x)

20 Wallis, G. P., Tabachnick, W. J. & Powell, J. R. 1983 Macrogeographic genetic variation in a human commensal: *Aedes aegypti*, the yellow fever mosquito. *Genet. Res.* **41**, 241–258. (doi:10.1017/S0016672300021315)

21 Powell, J. R., Tabachnick, W. J. & Arnold, J. 1980 Genetics and the origin of a vector population: *Aedes aegypti*, a case study. *Science* **208**, 1385–1387. (doi:10.1126/science.7375945)

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

23 Scott, J. A. & McClelland, G. A. H. 1975 Electrophoretic differences between sympatric ecotypes. *Nature* **256**, 405–406. (doi:10.1038/256405a0)

24 Trpis, M. & Hausermann, W. 1975 Demonstration of differential domesticity of *Aedes aegypti* (L.) (Diptera, Culicidae) in Africa by mark–release–recapture. *Bull. Entomol. Res.* **65**, 199–208. (doi:10.1017/S0007485300005903)

25 McClelland, G. A. H. 1974 A worldwide survey of variation in scale pattern of the abdominal tergum of *Aedes aegypti* (L.) (Diptera: Culicidae). *Trans. R. Entomol. Soc. Lond.* **126**, 239–259. (doi:10.1111/j.1365-2311.1974.tb00853.x)

26 Mattingly, P. F. 1957 Genetical aspects of the *Aedes aegypti* problem. I. Taxonomy and bionomics. *Ann. Trop. Med. Parasitol.* **51**, 392–408.

27 Mattingly, P. F. 1967 Taxonomy of *Aedes aegypti* and related species. *Bull. World Health Organ.* **36**, 552–554.

28 Tabachnick, W. J., Munstermann, L. E. & Powell, J. R. 1979 Genetic distinctness of sympatric forms of *Aedes aegypti* in East Africa. *Evolution* **33**, 287–295. (doi:10.2307/2407619)

29 Trpis, M. & Hausermann, W. 1978 Genetics of house-entering behaviour in East African populations of *Aedes aegypti* (L.) (Diptera: Culicidae) and its relevance to speciation. *Bull. Entomol. Res.* **68**, 521–532. (doi:10.1017/S0007485300009494)

30 Paupy, C., Brengues, C., Ndiath, O., Toty, C., Herve, J. P. & Simard, F. 2010 Morphological and genetic variability within *Aedes aegypti* in Niakhar, Senegal. *Infect. Genet. Evol.* **10**, 473–480. (doi:10.1016/j.meegid.2010.03.001)

31 Sylla, M., Bosio, C., Urdaneta-Marquez, L., Ndiaye, M. & Black, W. C. 2009 Gene flow, subspecies composition, and dengue virus-2 susceptibility among *Aedes aegypti* collections in Senegal. *PLoS Negl. Trop. Dis.* **3**, e408. (doi:10.1371/journal.pntd.0000408)

32 Huber, K., Ba, Y., Dia, I., Mathiot, C., Sall, A. A. & Diallo, M. 2008 *Aedes aegypti* in Senegal: genetic diversity and genetic structure of domestic and sylvatic populations. *Am. J. Trop. Med. Hyg.* **79**, 218–229.

33 Paupy, C., Brengues, 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. *J. Med. Entomol.* **45**, 391–400. (doi:10.1603/0022-2585(2008)45[391:GFBDAS]2.0.CO;2)

34 Vazeille-Falcoz, M., Failloux, A. B., Mousson, L., Elissa, N. & Rodhain, F. 1999 Oral receptivity of *Aedes aegypti formosus* from Franceville (Gabon, Central Africa) for dengue type 2 virus. *Bull. Soc. Pathol. Exot.* **92**, 341–342.

35 Miller, B. R., Monath, T. P., Tabachnick, W. J. & Ezike, V. I. 1989 Epidemic yellow fever caused by an incompetent mosquito vector. *Trop. Med. Parasitol.* **40**, 396–399.

36 Nasidi, A. *et al.* 1989 Urban yellow fever epidemic in western Nigeria, 1987. *Trans. R. Soc. Trop. Med. Hyg.* **83**, 401–406. (doi:10.1016/0035-9203(89)90518-X)

37 Colton, Y. M., Chadee, D. D. & Severson, D. W. 2003 Natural skip oviposition of the mosquito *Aedes aegypti* indicated by codominant genetic markers. *Med. Vet. Entomol.* **17**, 195–204. (doi:10.1046/j.1365-2915.2003.00424.x)

38 Apostol, G. L., Black, W. C., Reiter, P. & Miller, B. R. 1996 Population genetics with RAPD-PCR markers: the breeding structure of *Aedes aegypti* in Puerto Rico. *Heredity* **76**, 325–334. (doi:10.1038/hdy.1996.50)

39 Apostol, G. L., Black, W. C., Reiter, P. & Miller, B. R. 1994 Use of randomly amplified polymorphic DNA amplified by polymerase chain reaction markers to estimate the number of *Aedes aegypti* families at oviposition sites in San Juan, Puerto Rico. *Am J. Trop. Med. Hyg.* **51**, 89–97.

40 Apostol, G. L., Black, W. C., Miller, B. R., Reiter, P. & Beaty, B. J. 1993 Estimation of the number of full sibling families at an oviposition site using RAPD-PCR markers: applications to the mosquito *Aedes aegypti*. *Theor. Appl. Genet.* **86**, 991–1000. (doi:10.1007/BF00211052)

41 Chadee, D. D. 1992 Seasonal incidence and horizontal distribution patterns of oviposition by *Aedes aegypti* in an urban environment in Trinidad, West Indies. *J. Am. Mosq. Contr. Assoc.* **8**, 281–284.

42 Slotman, M. A., Kelly, N. B., Harrington, L. C., Kitthawee, S., Jones, J. W., Scott, T. W., Caccone, A. & Powell, J. R. 2007 Polymorphic microsatellite markers for studies of *Aedes aegypti* (Diptera: Culicidae), the vector of dengue and yellow fever. *Mol. Ecol. Not.* **7**, 168–171. (doi:10.1111/j.1471-8286.2006.01533.x)

43 Rozen, S. & Skaletsky, H. 2000 Primer 3 on the WWW for general users and for biologists programmers. In *Bioinformatics methods and protocols: methods in molecular biology* (eds S. Krawetz & S. Misener), pp. 365–386. Totowa, NJ: Humana Press.

44 Boutin-Ganache, I., Raposo, M., Raymond, M. & Deschepper, C. F. 2001 M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele-sizing methods. *Biotechniques* **31**, 24–26.

45 Rousset, F. 2008 GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Mol. Ecol. Resour.* **8**, 103–106. (doi:10.1111/j.1471-8286.2007.01931.x)

46 Raymond, M. & Rousset, F. 1995 GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.* **86**, 248–249.

47 Excoffier, L. & Lischer, H. E. L. 2010 Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**, 564–567. (doi:10.1111/j.1755-0998.2010.02847.x)

48 Pritchard, J. K., Stephens, M. & Donnelly, P. 2000 Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.

49 Evanno, G., Regnaut, S. & Goudet, J. 2005 Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**, 2611–2620. (doi:10.1111/j.1365-294X.2005.02553.x)

50 Rosenberg, N. A. 2004 DISTRUCT: a program for the graphical display of population structure. *Mol. Ecol. Not.* **4**, 137–138. (doi:10.1046/j.1471-8286.2003.00566.x)

51 Piry, S., Alapetite, A., Cornuet, J. M., Paetkau, D., Baudouin, L. & Estoup, A. 2004 GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Heredity* **95**, 536–539. (doi:10.1093/jhered/esh074)

52 Cavalli-Sforza, L. L. & Edwards, A. W. 1967 Phylogenetic analysis. Models and estimation procedures. *Am. J. Hum. Genet.* **19**, 233–257.

53 Tamura, K., Dudley, J., Nei, M. & Kumar, S. 2007 MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**, 1596–1599. (doi:10.1093/molbev/msm092)

54 Takezaki, N. & Nei, M. 1996 Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* **144**, 389–399.

55 Hammer, Ø., Harper, D. A. T. & Ryan, P. D. 2001 PAST: Paleontological statistics software package for education and data analysis. *Palaeontol. Electron.* **4**, 9.

56 Kalinowski, S. T. 2005 HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Mol. Ecol. Not.* **5**, 187–189. (doi:10.1111/j.1471-8286.2004.00845.x)

57 Kalinowski, S. T. 2004 Counting alleles with rarefaction: private alleles and hierarchical sampling designs. *Conserv. Genet.* **5**, 539–543. (doi:10.1023/B:COGE.0000041021.91777.1a)

58 Bernhardt, S. A., Blair, C., Sylla, M., Bosio, C. & Black, W. C. 2009 Evidence of multiple chromosomal inversions in *Aedes aegypti formosus* from Senegal. *Insect Mol. Biol.* **18**, 557–569. (doi:10.1111/j.1365-2583.2009.00895.x)

59 Ponlanwat, A. & Harrington, L. C. 2005 Blood feeding patterns of *Aedes aegypti* and *Aedes albopictus* in Thailand. *J. Med. Entomol.* **42**, 844–849.

60 Kark, S., Iwaniuk, A., Schalimtzek, A. & Banker, E. 2007 Living in the city: can anyone become an 'urban exploiter'? *J. Biogeogr.* **34**, 638–651. (doi:10.1111/j.1365-2699.2006.01638.x)

61 Yeh, P. J. & Price, T. D. 2004 Adaptive phenotypic plasticity and the successful colonization of a novel environment. *Am. Nat.* **164**, 531–542. (doi:10.1086/423825)

62 Chadee, D. D., Ward, R. A. & Novak, R. J. 1998 Natural habitats of *Aedes aegypti* in the Caribbean—a review. *J. Am. Mosq. Control Assoc.* **14**, 5–11.

63 Wallis, G. P. & Tabachnick, W. J. 1990 Genetic analysis of rock hole and domestic *Aedes aegypti* on the Caribbean island of Anguilla. *J. Am. Mosq. Control Assoc.* **6**, 625–630.

64 Failloux, A. B., Vazeille, M. & Rodhain, F. 2002 Geographic genetic variation in populations of the dengue virus vector *Aedes aegypti*. *J. Mol. Evol.* **55**, 653–663. (doi:10.1007/s00239-002-2360-y)

65 Bosio, C. F., Beaty, B. J. & Black, W. C. 1998 Quantitative genetics of vector competence for Dengue-2 virus in *Aedes aegypti*. *Am. J. Trop. Med. Hyg.* **59**, 965–970.

66 Tabachnick, W. J., Wallis, G. P., Aitken, T. H. G., Miller, B. R., Amato, G. D., Lorenz, L., Powell, J. R. & Beaty, B. J. 1985 Oral infections of *Aedes aegypti* with yellow fever virus: geographic variation and genetic considerations. *Am. J. Trop. Med. Hyg.* **34**, 1219–1224.

67 Garcia, G. P. *et al.* 2009 Recent rapid rise of a permethrin knock down resistance allele in *Aedes aegypti* in Mexico. *PLoS Negl. Trop. Dis.* **3**, 10. (doi:10.1371/journal.pntd.0000531)

68 Urdaneta-Marquez, L., Bosio, C., Herrera, F., Rubio-Palis, Y., Salasek, M. & Black, W. C. 2008 Genetic relationships among *Aedes aegypti* collections in Venezuela as determined by mitochondrial DNA variation and nuclear single nucleotide polymorphisms. *Am. J. Trop. Med. Hyg.* **78**, 479–491.