Published in final edited form as:
Evolution. 2014 February ; 68(2): 514–525. doi:10.1111/evo.12281.

Human impacts have shaped historical and recent evolution in *Aedes aegypti*, the dengue and yellow fever mosquito

Julia E. Brown[1], **Benjamin R. Evans**[1], **Wei Zheng**[2], **Vanessa Obas**[1], **Laura Barrera-Martinez**[1], **Andrea Egizi**[3], **Hongyu Zhao**[4,5], **Adalgisa Caccone**[1], and **Jeffrey R. Powell**[1]

[1]Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA

[2]Keck Biostatistics Resource, Yale School of Medicine, New Haven, CT, USA

[3]Department of Entomology, Rutgers University, New Brunswick, NJ, USA

[4]Department of Biostatistics, Yale School of Public Health, New Haven, CT, USA

[5]Department of Genetics, Yale School of Medicine, New Haven, CT, USA

Abstract

Though anthropogenic impacts are often considered harmful to species, human modifications to the landscape can actually create novel niches to which other species can adapt. These "domestication" processes are especially important in the context of arthropod disease vectors, where ecological overlap of vector and human populations may lead to epidemics. Here, we present results of a global genetic study of one such species, the dengue and yellow fever mosquito, *Aedes aegypti*, whose evolutionary history and current distribution have been profoundly shaped by humans. We used DNA sequences of four nuclear genes and 1504 SNP markers developed with RAD-tag sequencing to test the hypothesis that *Ae. aegypti* originated in Africa, where a domestic form arose and spread throughout the tropical and subtropical world with human trade and movement. Results confirmed African ancestry of the species, and supported a single subspeciation event leading to the pantropical domestic form. Additionally, genetic data strongly supported the hypothesis that human trade routes first moved domestic *Ae. aegypti* out of Africa into the New World, followed by a later invasion from the New World into Southeast Asia and the Pacific. These patterns of domestication and invasion are relevant to many species worldwide, as anthropogenic forces increasingly impact evolutionary processes.

Keywords

Aedes aegypti; population genetics; human association; RAD; nuclear markers

INTRODUCTION

Over the past several thousand years, *Homo sapiens* has modified the global landscape profoundly. Though many anthropogenic impacts are known to destroy natural habitats and harm species (Chapin et al. 2000; Tilman and Lehman 2001; Pimm et al. 2006), they can also create novel niches for species living in close proximity to humans. Over time, certain species, often hardy generalists (Yeh and Price 2004; Kark et al. 2007), can evolve to live in

Corresponding author: Julia E. Brown, julia.brown@aya.yale.edu, P.O. Box 208106, 165 Prospect Street, New Haven, CT 06520-8106.

The authors have no conflicts of interest to declare.

Sequence data have been archived to GenBank with accession numbers KF360534-KF360809.

these disturbed environments and may even adapt to become human-habitat specialists (Schofield et al. 1999; Das et al. 2004; Partecke et al. 2006; Keller 2007; Evans et al. 2009). This “domestication” process, or human commensalism, has impacted the evolutionary paths of species since the beginnings of human civilization, and only continues to grow in importance over time as human-dominated landscapes represent an increasing proportion of available niches worldwide.

Some of the closest ecological relationships between humans and other species are with arthropod vectors of disease, which often depend on human blood and artificial breeding sites for their survival (Schofield et al. 1999; Harrington et al. 2001; Ayala and Coluzzi 2005; Lyimo and Ferguson 2009). These human-adapted vector species can be spread by human movement and trade, with major public health consequences (Lounibos 2002). Humans have been particularly effective at spreading anthropophilic mosquitoes over intercontinental distances, as evidenced by the invasion of the malaria vector *Anopheles arabiensis* into Brazil from West Africa in the 1930s (Parmakelis et al. 2008), and the cosmopolitan spread of the disease vectors *Aedes albopictus* and *Culex quinquefasciatus*, among many other examples (Lounibos 2002). Our study focused on *Aedes aegypti*, the dengue and yellow fever mosquito, whose evolutionary history has been profoundly impacted by its relationship with humans (Tabachnick 1991). Though an effective vaccine exists for yellow fever, there is no vaccine currently approved for use against dengue viruses. Therefore, understanding and controlling this mosquito vector is essential to preventing disease spread; dengue affects an estimated 50 million people per year, with a full 40% of the world’s population currently at risk for infection (WHO 2009).

Aedes aegypti has long been assumed to have originated in Africa, where the ancestral form was likely a generalist, zoophilic treehole breeder (Mattingly 1957; Tabachnick 1991). Ecologically similar populations of *Ae. aegypti* still exist today on the African continent as the subspecies *Aedes aegypti formosus*. The better-known specialized domestic form, *Aedes aegypti aegypti*, is found in close association with human habitats throughout the tropical and subtropical world outside of Africa. Pure *Aedes ae. aegypti* mosquitoes are not found on the African continent except for unique, isolated populations in coastal East Africa, which are described below. The two subspecies are genetically distinct with discrete geographic distributions and mean differences in scaling pattern and ecological preferences (Mattingly 1957; McClelland 1974; Tabachnick and Powell 1979; Powell et al. 1980; Wallis et al. 1983; Brown et al. 2011).

It has been hypothesized that the domestic form, *Ae. ae. aegypti*, originated from a small population of forest-dwelling *Ae. aegypti* that became isolated in North Africa when a period of severe drying began in the Sahara approximately 4,000 years ago (Petersen 1977; Tabachnick 1991). The harsh landscape could have selected for mosquitoes exhibiting domestic behaviors such as a preference for breeding in artificial water storage containers (Petersen 1977; Tabachnick 1991). As global trade increased over the centuries, highly human-adapted *Ae. ae. aegypti* were spread across much of the tropical and subtropical world. The species was likely introduced to the New World by slave trade ships between the fifteenth and eighteenth centuries, possibly on multiple occasions (Tabachnick 1991). The spread to Asia may have occurred during the late nineteenth century based on the history of urban dengue emergence in the region (Smith 1956; Tabachnick 1991).

Today, urban domestic populations of traditionally “wild” *Ae. ae. formosus* are common throughout Africa, which can make subspecies identification of these populations more difficult (Miller et al. 1989; Nasidi et al. 1989; Vazeille-Falcoz et al. 1999; Huber et al. 2008; Paupy et al. 2008; Sylla et al. 2009; Paupy et al. 2010; Brown et al. 2011). Some of these populations may have arisen from hybridization with domestic *Ae. ae. aegypti*

mosquitoes (Miller et al. 1989; Tabachnick 1991; Brown et al. 2011), but others are likely pure *Ae. ae. formosus* mosquitoes that have opportunistically moved into human-altered habitats across the African landscape. There are also isolated populations of the subspecies *Ae. ae. aegypti* that exist sympatrically with *Ae. ae. formosus* in certain locations on the East African coast (e.g. Rabai, Kenya) with little or no hybridization between forms (Tabachnick et al. 1979; Brown et al. 2011). These mosquitoes may represent a unique genetic and ecological variant of the subspecies that breeds exclusively indoors, as compared to typically outdoor-breeding pantropical *Ae. ae. aegypti* populations.

Due to its role as an invasive species and globally important disease vector, it is critical to understand the patterns by which *Ae. aegypti* populations have associated with humans, and to assess how humans have moved populations of the species around the world through time. To gain a historical perspective of *Ae. aegypti* evolution, we sequenced four nuclear markers in mosquitoes from 21 localities worldwide and used phylogenetic methods to explore the hypothesis of African ancestry in the species. We expect that the hypothesized isolation event and accompanying selection leading to the domestic subspecies, *Ae. ae. aegypti* would have resulted in a population bottleneck, detectable with our genetic markers. Additionally, a panel of 1504 SNP markers was developed using a RAD (restriction site associated DNA) sequencing approach (Baird et al. 2008) to assess fine-scale structure among diverse, global *Ae. aegypti* populations to better understand the role of humans as a disperser of the species on a more recent time scale.

MATERIALS AND METHODS

Collections

Mosquitoes for this study were collected from 21 localities representing 13 countries worldwide (Fig. 1, Table 1). All collections were performed according to previously published methodology (Brown et al. 2011). Whole genomic DNA was extracted individually from *Ae. aegypti* mosquitoes (Table 1) and outgroup species, *Ae. mascarensis* and *Ae. bromeliae*.

Sequenced nuclear loci

Primers were developed to amplify and sequence variable sections of four nuclear genes: *apoLp-2* paralogue (Apolipophorin 2), *SDR* (Short-chain dehydrogenase/reductase), *CYP9J2* (Cytochrome P450), and *DVRF1* (dengue virus restriction factor) (Table S1, see Online Supporting Information). In total, 210 *Ae. aegypti* mosquitoes, representing 10 individuals from each of 21 *Ae. aegypti* populations (Fig. 1, Table 1) were amplified and sequenced at these four loci. Alleles were phased with TA-cloning and resequencing. The same methods were used with minor modifications (see Online Supporting Information) on four individuals each from the outgroup species *Ae. mascarensis* and *Ae. bromeliae*. The locus *DVRF1* failed to amplify in outgroup individuals with the exception of two *Ae. mascarensis* specimens where sequences matched identically to *Ae. ae. aegypti* sequences, likely a result of interspecific introgression.

Restriction-site associated DNA (RAD)

Additionally, DNA from eight mosquitoes each from a subset of 16 *Ae. aegypti* populations worldwide was submitted to Floragenex (Eugene, OR, USA) for sequenced restriction site associated DNA (RAD) library preparation, along with 8 individuals of *Ae. mascarensis* (Table 1). Individually-barcoded RAD libraries were created following published methods (Baird et al. 2008; Emerson et al. 2010; Hohenlohe et al. 2010; Chutimanitsakun et al. 2011). Illumina sequencing adaptors and individual barcodes were ligated to Sbf1-digested total genomic DNA. The libraries were sequenced on the Illumina Genome Analyzer IIx (1

lane, 40 individuals) or HiSeq 2000 (2 lanes, 48 individuals each) platform with 1 × 75 bp chemistry at the Yale Center for Genome Analysis, New Haven, CT, USA. The sequencing produced 65.4 million and 218 million short reads from one lane GAIIx and 2 lanes HiSeq 2000, respectively.

Selection of SNPs from RAD data

Sequence data were filtered and demultiplexed using the process_radtags utility included in Stacks v 0.994 (Catchen et al. 2011), which also discarded low quality reads and rescued reads where the barcode or cut site were one bp different from what was expected. Sequences for each individual were aligned to the *Ae. aegypti* reference genome (aaegypti.SUPERCONTIGS-Liverpool.AaegL1) on Vectorbase using bwa 0.5.9 (Li and Durbin 2009), allowing for no more than 5 differences across the whole ~70 bp read and no more than 3 differences in 20 base windows of each other while being aligned against the reference (options aln -n 5 -k 3 -l 20). These individual-level alignments were then pooled into population level alignments for use as pseudo-parents for input into Stacks. The reference-aligned version of the pipeline was run using the ref_map.pl script with the expectation of up to 5 differences between individuals and reference. Stacks then compared reads grouped into stacks across populations and assigned SNP calls to stacks of reads that differed across individuals (Catchen et al. 2011). This information was deposited into a database for downstream curation and filtering.

The stacks pipeline identified 184,178 unique tags, 142,448 of which were identical to the reference, i.e. were not polymorphic. Of the 41,730 polymorphic tags, 10,503 had exactly 1 SNP called. From those, we filtered for tags that were found in at least half of the individuals in 16 or 17 populations and only allowed for biallelic SNPs. Seven individuals with poor sequencing success were removed from the dataset, as reflected in Table 1. The final working dataset consisted of 1504 tags, from which we used the 1504 associated SNPs in all further analyses of the 129 included individuals (Table 1). Across the 1504 SNPs, total levels of missing data were quantified for each mosquito population (Excoffier and Lischer 2010). Levels of missing data per population varied from 8.7% to 27.7% within *Ae. aegypti* (Table 1). A full 59.7% of SNPs had no calls across individuals in the *Ae. mascarensis* population.

Population structure and evolutionary history

In order to make rooted historical inferences regarding the origins of global *Ae. aegypti* populations, individual gene trees were created for *apoLp-2* paralogue, *CYP9J2*, *SDR*, and *DVRF1*. Sequences were aligned using a slow DNA alignment with default parameters in ClustalW2 (Larkin et al. 2007; Goujon et al. 2010), and were checked by eye in MacClade 4 (Maddison and Maddison 2005). For gene trees of each locus, identical sequences were first collapsed to unique haplotypes using the program ALTER (Glez-Pena et al. 2010). These haplotype sequences were archived to GenBank with accession numbers KF360534-KF360809. The best model of nucleotide evolution was chosen using the Akaike information criterion (AIC) as implemented in the program jModelTest 0.1.1 (Posada 2008). The models of nucleotide evolution selected for the four sequenced loci were TVM+I+G (*apoLp2* paralogue), HKY+I+G (*DVRF1* and *SDR*), and SYM+I+G (*CYP9J2*). Phylogenetic trees were created using the maximum-likelihood approach implemented in the program Garli 2.0 (http://garli.googlecode.com). For each locus, the best tree was selected from 10 search replicates, and node support was assessed using 1000 bootstrap replicates.

The program *BEAST v. 1.5.3 (Heled and Drummond 2010) was used to create a species tree for the 21 *Ae. aegypti* populations using phased sequence data from each of three loci (*apoLp-2* paralogue, *CYP9J2*, and *SDR*). MCMC analyses were run for a total of 200 million

generations (sampling once every 1000 steps and excluding the first 20% as burn-in), and the analysis was repeated three times. *Aedes mascarensis* individuals were ultimately not included in the species tree due to problems with the long branch leading to the outgroup. However, the root was inferred from the individual gene trees. *DVRF1* sequences were not included in the species tree primarily because there were no outgroup sequences for this locus, and we did not feel comfortable placing a root with *DVRF1* included. Additionally, since *DVRF1* is a gene known to affect dengue virus replication in *Ae. aegypti* (Souza-Neto et al. 2009, see Online Supporting Information), the locus may be subject to selection pressures that we have not yet explored in depth, and which could bias the species tree. For each of the four sequenced nuclear loci, we used the program DnaSP v5 (Librado and Rozas 2009) to calculate nucleotide diversity (π) and haplotype diversity (H_d) of *Ae. aegypti* populations by geographical region.

Detailed relationships between 16 *Ae. aegypti* populations and one population of *Ae. mascarensis* were also assessed using the 1504 RAD-associated SNPs. Values of F_{ST} were calculated between each pair of populations in GENEPOP 4.1 (Rousset 2008). Cavalli-Sforza and Edwards's chord distance was calculated between populations in the GENDIST module of PHYLIP 3.69 and used to create a population-level neighbor-joining network in MEGA5 (Tamura et al. 2011). Node confidence was inferred with 1000 bootstrap replicates in PHYLIP 3.69 (modules SEQBOOT, GENDIST, NEIGHBOR, and CONSENSE). *Aedes mascarensis* individuals were not included in this bootstrapped neighbor-joining network due to the high levels of missing data in the species, which would have prevented a large portion of the SNPs from being included in the analysis in PHYLIP 3.69. Individual level clustering was assessed using a neighbor-joining tree and a principal components analysis (PCA) using R v12.13.1 (R Development Core Team 2008) and the adegenet package v1.3-1 (Jombart 2008; Jombart and Ahmed 2011). All missing data in the individual level analyses were replaced by population means.

We also evaluated patterns of population structure using the Bayesian clustering method implemented in the software program STRUCTURE v2.3.3 (Pritchard et al. 2000; Falush et al. 2003). STRUCTURE was run on all *Ae. aegypti* and *Ae. mascarensis* individuals, and no a priori information regarding sampling locations was used. The most likely number of clusters was determined following the guidelines of Pritchard et al. (Pritchard et al. 2000) and by calculating ΔK, which is based on the second order rate of change of the likelihood distribution between values of K (Evanno et al. 2005). STRUCTURE results were visualized using the program DISTRUCT (Rosenberg 2004). To determine the most likely number of clusters (K), we conducted five independent runs for each K = 1-18. For all runs, we assumed an admixture model and correlated allele frequencies and used a burn-in value of 20,000 iterations followed by 50,000 replications.

For several of our analyses, comparisons were made between major genetic groups detected here, and in previous microsatellite analyses of the species (Brown et al. 2011). The two major genetic groups, African and pantropical, correspond to the subspecies *Ae. ae. formosus* and *Ae. ae. aegypti*, respectively, and follow the same geographic boundaries. As such, in this manuscript, the pantropical form refers to all *Ae. aegypti* populations outside of Africa, as well as isolated populations on the East Coast of Africa (i.e. certain mosquitoes from Rabai, Kenya). All other *Ae. aegypti* populations from Africa included in this study represent the African genetic form.

RESULTS

Genetic diversity

Measures of genetic diversity for all sequenced nuclear loci were higher across African *Ae. ae. formosus* populations than across pantropical domestic mosquitoes (Table 2). East African populations were the most diverse group by both measures for three of the four loci, with West/Central African populations harboring an intermediate level of diversity, on average (Table 2). Within pantropical *Ae. ae. aegypti* populations, 15 of 16 values of genetic diversity were lower for Asia-Pacific populations than New World populations (Table 2).

Population genetic structure

Both the *BEAST species tree based on nuclear genes and multiple cluster analyses of 1504 SNPs clearly separated the two subspecies of *Ae. aegypti*: *Ae. ae. formosus* and *Ae. ae. aegypti* (Fig. 2, Fig. 3). One genetic cluster (African) corresponds to *Ae. ae. formosus* and is only found in Africa. The other corresponds to *Ae. ae. aegypti* (pantropical) and represents all populations from outside Africa, as well as an isolated group of mosquitoes from Rabai, Kenya. As seen in microsatellite analyses (Brown et al. 2011), the *Ae. ae. aegypti* population from Rabai, Kenya (coastal East Africa) formed a somewhat distinct genetic cluster, as did the *Ae. ae. formosus* population from the same location (Fig. 2a). Additionally, the broad-scale individual-based PCA distinguished *Ae. mascarensis* outgroup individuals from all *Ae. aegypti* individuals (Fig. 2a).

Average F_{ST} values between groups reflected similar patterns to the PCA results (Fig. 2b). Pairwise F_{ST} values between *Ae. aegypti* populations were high overall, ranging from 0.0873 to 0.3599. The average pairwise F_{ST} between pantropical *Ae. ae. aegypti* populations was 0.1523, and that among African *Ae. ae. formosus* populations was a similar value of 0.1549 (Fig. 2b). However, the average population-pairwise F_{ST} between the two groups (subspecies) was much higher at 0.3090. The outgroup species, *Ae. mascarensis*, was clearly divergent from *Ae. aegypti* populations, falling slightly closer to the African group than to pantropical populations (Fig. 2b). The Bayesian clustering algorithms implemented in STRUCTURE (Pritchard et al. 2000; Falush et al. 2003) revealed similar broad patterns as detected using PCA and F_{ST} values (Fig. 2c).

Phylogeography and Phylogenetics

All three rooted nuclear gene trees (*apoLp2* paralogue, *CYP9J2*, *SDR*) supported African ancestry in *Ae. aegypti* (Fig. 4, Tables S2-S4). Although structure was observed within Africa in all four loci, there was no clear geographic pattern of alleles within Africa in any of the gene trees (Fig. 4). Nearly all of the pantropical alleles formed a monophyletic group in the *apoLp-2* paralogue and *SDR* gene trees, (Fig.4). In both the *apoLp2* paralogue and *SDR* trees, common pantropical haplotypes were found in some African populations, mainly Ngoye and Dakar, Senegal, with a handful of alleles from Yaounde, Cameroon, and Rabai, Kenya (Fig. 4, Tables S2, S3).

Despite relatively high levels of variation in *CYP9J2* across *Ae. aegypti* populations, the subspecies relationships remained poorly resolved in this gene tree (Fig. 4). However, interestingly, all *CYP9J2* alleles from *Ae. ae. aegypti* individuals from Rabai, Kenya represented two unique haplotypes, numbers 49 and 52 (Table S4), that were not found in any other *Ae. aegypti* populations, with an exception of a single allele matching the *Ae. ae. formosus* population in the same location (Rabai, Kenya). Though the subspecies split was also largely unsupported using *DVRF1*, it is of note that at least 1 major clade in the tree was essentially missing pantropical alleles (Fig. 4, Table S5). Overall, many nodes within all gene trees remained poorly resolved.

Though *Ae. mascarensis* and *Ae. bromeliae* served as useful outgroups for *apoLp-2* paralogue, *CYP9J2*, and *SDR*, a few *SDR* and *DVRF1* alleles from *Ae. mascarensis* samples matched identically to alleles found in pantropical *Ae. aegypti* individuals (Fig. 4, Tables S3, S5), indicating some level of interspecific hybridization between *Ae. mascarensis* and *Ae. ae. aegypti*. *DVRF1* could not be amplified from most outgroup samples. Each of the four gene trees clearly depicted the large amount of variation across *Ae. ae. formosus* in Africa, as compared to the smaller number of haplotypes found in pantropical (*Ae. ae. aegypti*) populations worldwide (Fig. 4).

As in the gene trees, no clear evolutionary genetic patterns were detectable among African populations in the species tree that was created with information from three nuclear genes: *apoLp2* paralogue, *CYP9J2*, and *SDR* (Fig. 3). However, patterns and structure were consistently detected among pantropical *Ae. ae. aegypti* populations. The species tree and a neighbor-joining analysis of the RAD SNPs detected distinct New World and Asia-Pacific clusters, with the well-supported Asia-Pacific group appearing the most distant (derived) from ancestral African populations (Fig. 3). Furthermore, an individual-based neighbor-joining network provided a near complete discrimination of pantropical populations from one another, with 96.3% of pantropical individuals clustering uniquely with other individuals from their population of origin (Figure S1).

DISCUSSION

Subspeciation in *Aedes aegypti*

Our results support the hypothesis that *Ae. aegypti* originated in Africa as a generalist mosquito similar to modern *Ae. ae. formosus*, from which a specialized domestic form (*Ae. ae. aegypti*) evolved through a single subspeciation event. All three rooted nuclear gene trees (Fig. 4, Tables S2-S4) indicated African ancestry in *Ae. aegypti*, and a genetic bottleneck was apparent in the domestic pantropical form, *Ae. ae. aegypti* (Table 2, Fig. 4). Broad-scale cluster analyses cleanly separated the subspecies of *Ae. aegypti* (i.e. the African and pantropical genetic forms), and support the position of *Ae. mascarensis* as a close but distinct outgroup (Fig. 2). Together, these data are consistent with the historical hypothesis that the subspecies *Ae. ae. aegypti* originated from a population of ancestral *Ae. aegypti* that became isolated in North Africa during the start of a drying period in the region (Petersen 1977; Tabachnick 1991). However, the data are equally congruent with numerous other scenarios within Africa, and the exact origin of *Ae. ae. aegypti* on the continent remains unresolved.

An interesting finding here (Fig. 2, 3) and in previous studies (Brown et al. 2011) is that genetically pure populations of *Ae. ae. aegypti* are largely absent from Africa despite the well-supported origin of the subspecies on that continent. One possible part of the explanation lies in populations of indoor-dwelling (i.e. true domestic) *Ae. ae. aegypti* mosquitoes that used to be common throughout the arid regions of North Africa, the Middle East, and the Mediterranean until the mid-20th century (Holstein 1967). Today, it is nearly impossible to find any such populations of *Ae. ae. aegypti* due to the advent of modern plumbing and sanitation, and the resulting lack of indoor water storage (Holstein 1967). These indoor populations may have been very ecologically similar to the *Ae. ae. aegypti* ancestors in North Africa (or elsewhere) described in the scenario above. Though these populations have gone extinct throughout most of their native, albeit artificial, habitat, the *Ae. ae. aegypti* in Rabai, Kenya may represent a rare, relic population.

Uniquely, the Rabai *Ae. ae. aegypti* mosquitoes were found only indoors, despite ample breeding sites outside of homes (Brown et al. 2011). In contrast, *Ae. ae. aegypti* populations throughout the rest of the world typically exhibit peridomestic behavior. Though genetically

pantropical, as designated by previous microsatellite analyses (Brown et al. 2011), the *Ae. ae. aegypti* population from Rabai, Kenya (coastal East Africa) remained noticeably different from other groups (Fig. 2). This population also harbored private haplotypes in the *CYP9J2* nuclear gene (Fig. 4, Table S4), indicating a unique evolutionary history. Because the vast majority of sequenced haplotypes from Rabai, Kenya were identical to those from other pantropical populations (Fig. 4, Tables S2-S3, S5), we hypothesize that *Ae. ae. aegypti* were introduced to coastal East Africa, including Rabai, after the subspeciation event (Tabachnick 1991). However, without other arid, indoor-only populations to compare to, we can only make informed speculation.

Due to the exact match of Rabai, Kenya *Ae. ae. aegypti* haplotypes to those from other pantropical populations for *apoLp2*, *SDR*, and *DVRF1*, it should be noted that the apparent basal position of the *Ae. ae. aegypti* population from Rabai, Kenya to the rest of *Ae. ae. aegypti* (Fig. 3) may be attributable to low levels of introgression over time from the sympatric *Ae. ae. formosus* population (Fig. 4, Tables S2-S5), rather than to true evolutionary history. In addition, the broad-scale analyses of the SNPs (Figure 2) show that the *Ae. ae. formosus* population in Rabai also clusters distinctly from, though close to, other African populations. It is likely that, as noted above, low levels of introgression in the region may have contributed to the unique SNP-based clustering patterns of both populations.

Another part of the explanation for a lack of pure *Ae. ae. aegypti* through most of Africa may be related to the ease with which the subspecies seem to interbreed in West and Central Africa. Though the two subspecies appear distinct based on multiple analyses of many types of genetic markers, nuclear sequences revealed recent introgression between *Ae. ae. formosus* and *Ae. ae. aegypti* in a few (mostly urban) West and Central African locations (Fig. 4, Tables S2-S5). Evidence for mixing of the subspecies has also been shown using microsatellite data (Brown et al. 2011), and hybridization between subspecies is further supported by morphological and behavioral evidence in the region (Miller et al. 1989; Tabachnick 1991). Since *Ae. ae. aegypti* populations harbor less genetic diversity than *Ae. ae. formosus*, this phenomenon also provides an explanation for lower values of genetic diversity within West and Central African as compared to East Africa (Table 2). This hybridization of forms, combined with the existence of rare and/or extinct indoor populations, may help explain the apparent paradox of *Ae. ae. aegypti* moving to the New World from a region with few, if any, remaining populations of pure *Ae. ae. aegypti* mosquitoes. Historical *Aedes ae. aegypti* populations in West and Central Africa may have been absorbed into *Ae. ae. formosus* populations and or may have gone extinct.

The clear evidence of hybridization between subspecies in West Africa is very different from the situation in certain East African coastal locations like Rabai where the subspecies remain distinct, sympatric forms with low levels of introgression (Brown et al. 2011, Fig. 2, Fig. 4, Tables S2-5). Since the two subspecies, including the sympatric forms from Rabai, are fully capable of interbreeding in the laboratory (Moore 1979), the underlying reasons for this disparity remain unknown. Future molecular and ecological studies of mating behavior and niche competition in and between subspecies in East and West/Central Africa (and outside Africa) will be critical to teasing apart this complex evolutionary situation. Additional genetic and behavioral studies may also help address the evolution of the incredibly complicated range and breadth of domestication-related traits seen in populations of both subspecies of *Ae. aegypti*.

Pantropical invasions and the spread of *Aedes aegypti aegypti*

Strong among-population relationships (Fig. 3) support the historical hypothesis that humans first moved *Ae. aegypti* out of Africa into the New World with the slave trade. Within pantropical *Ae. ae. aegypti*, our results clearly support a secondary invasion of *Ae. ae.*

aegypti from the New World to the Asia-Pacific region, perhaps as recently as the 20th century (Smith 1956; Tabachnick 1991). New World populations cluster distinctly from a Southeast Asia-Pacific group, and the Southeast Asia-Pacific cluster is the most distant from ancestral African populations, suggesting a more recent invasion of the region (Smith 1956; Tabachnick 1991). This pattern appears in both the species tree and the SNP-based neighbor-joining network, though is better supported by the latter (Fig. 3). Genetic diversity measures also support a more recent founder event in the Southeast Asia-Pacific group, as nearly all values were lower for Asia-Pacific populations than New World populations (Table 2).

Population discrimination and gene flow

Though some population signal can be seen within Africa in an individual-based neighbor-joining network (Figure S1), there is very little geographic clustering and little historical information within the continent. This pattern has now been detected using SNPs, sequenced nuclear loci, and microsatellites (Brown et al. 2011). However, there is obvious geographical structuring in the rest of the world. In addition to distinct New World and Asia-Pacific clusters, the individual-based neighbor-joining network provided a near complete discrimination of pantropical populations from one another, with 96.3% of pantropical individuals clustering uniquely with other individuals from their population of origin (Figure S1).

All genetic evidence here and in our previous study (Brown et. al 2011) suggests that *Ae. ae. formosus* individuals in Africa have remained relatively panmictic over the course of the species history. Up until the past few decades, *Ae. ae. formosus* populations bred primarily in natural environments (e.g. tree holes) (Mattingly 1957; McClelland 1974), and we hypothesize that members of this subspecies would have been forced to move further to find suitable breeding locations during the dry season, as well as seek out hosts since they did not have the luxury of constant access to humans in a domestic habitat. There is evidence that a lack of local, suitable oviposition sites can lead to increased dispersal in *Ae. aegypti* (Reiter et al. 1995; Honorio et al. 2003; Maciel-de-Freitas and Lourenco-de-Oliveira 2009). This increased movement would have led to higher levels of gene flow between locations, and may explain the low population genetic signal within this subspecies.

In contrast, recent but clear geographic structure can be observed across pantropical *Ae. ae. aegypti* populations using allele frequency and genotype-based analyses with both SNPs and microsatellites (Brown et al. 2011). This structure may be explained by recent founder events out of Africa and across the tropics and subtropics, as well the discontinuous nature of domestic habitats available to this highly human-adapted subspecies. Pantropical *Ae. ae. aegypti* mosquitoes tend to stay within approximately 10 to 500 m flight distance during their entire lifetimes (McDonald 1977; Trpis and Hausermann 1986; Reiter et al. 1995; Harrington et al. 2005; Maciel-de-Freitas and Lourenco-de-Oliveira 2009) when stable breeding sites and host blood are readily available, as would be likely for many domestic populations. With little dispersal or gene flow, genetic structure could be formed and maintained even on a small geographic scale.

Conclusions

We used genetic markers to show that anthropogenic forces have profoundly shaped the evolutionary history and distribution of *Ae. aegypti* across time and space. Human domestic habitats provided the ecological basis for the subspeciation of the highly invasive and epidemiologically important subspecies, *Ae. ae. aegypti*, from ancestral *Ae. ae. formosus*. Soon after, human trade and movement facilitated the invasion of *Ae. aegypti* across the tropical and subtropical world. The history of *Ae. aegypti* represents just one example of

how anthropogenic impacts can shape the evolutionary path of other species, and influence patterns of genetic diversity.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to 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, Jennifer Simpson, Carolyn S. McBride, Petrina Johnson, Scott Ritchie, Christophe Paupy, Hervé Bossin, Joel Lutomiah, Ildefonso Fernandez-Salas, Alongkot Ponlawat, Anthony J. Cornel, William C. Black, IV, Norma Gorrochotegui-Escalante, Ludmel Urdaneta-Marquez, Massamba Sylla, Michel Slotman, Kristy O. Murray, Christopher Walker, Gerard Somers, Roberto Barrera, Denise Valle, Alexandre Peixoto, and Dina Fonseca for their help in providing mosquito samples, and Beckie Symula for her help with statistical and phylogenetic analyses. Data analysis was supported by the Yale University Biomedical High Performance Computing Center. JEB was supported by NIH Pre-doctoral Training in Genetics (T32-GM07499-33) and the Yale Institute for Biospheric Studies (YIBS), Center for Field Ecology pilot grant. The collection in Brazil was supported by the Programa Nacional de Controle da Dengue (PNCD). The rest of the study was supported by NIH R01 AI046018 (JRP) and NSF Doctoral Dissertation Improvement Grant DEB-1011449 (JEB, JRP).

LITERATURE CITED

Ayala FJ, Coluzzi M. Chromosome speciation: Humans, Drosophila, and mosquitoes. Proc. Natl. Acad. Sci. USA. 2005; 102:6535–6542. [PubMed: 15851677]

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA. Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. PLoS One. 2008; 3:e3376. [PubMed: 18852878]

Brown JE, McBride CS, Johnson P, Ritchie S, Paupy C, Bossin H, Lutomiah J, Fernandez-Salas I, Ponlawat A, Cornel AJ, Black WC, Gorrochotegui-Escalante N, Urdaneta-Marquez L, Sylla M, Slotman M, Murray KO, Walker C, Powell JR. Worldwide patterns of genetic differentiation imply multiple 'domestications' of *Aedes aegypti*, a major vector of human diseases. Proc. R. Soc. Lond. B. 2011; 278:2446–2454.

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH. Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. G3: Genes, Genomes, Genetics. 2011; 1:171–182.

Chapin FS, Zavaleta ES, Eviner VT, Naylor RL, Vitousek PM, Reynolds HL, Hooper DU, Lavorel S, Sala OE, Hobbie SE, Mack MC, Diaz S. Consequences of changing biodiversity. Nature. 2000; 405:234–242. [PubMed: 10821284]

Christophers, RC. The Yellow Fever Mosquito: Its Life History, Bionomics, and Structure. Cambridge Univ. Press; Cambridge, U. K.: 1960. Aedes aegypti.

Chutimanitsakun Y, Nipper RW, Cuesta-Marcos A, Cistue L, Corey A, Filichkina T, Johnson EA, Hayes PM. Construction and application for QTL analysis of a Restriction Site Associated DNA (RAD) linkage map in barley. BMC Genomics. 2011; 12:4. [PubMed: 21205322]

Das A, Mohanty S, Stephan W. Inferring the population structure and demography of *Drosophila ananassae* from multilocus data. Genetics. 2004; 168:1975–1985. [PubMed: 15611168]

Emerson KJ, Merz CR, Catchen JM, Hohenlohe PA, Cresko WA, Bradshaw WE, Holzapfel CM. Resolving postglacial phylogeography using high-throughput sequencing. Proc. Natl. Acad. Sci. USA. 2010; 107:16196–16200. [PubMed: 20798348]

Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 2005; 14:2611–2620. [PubMed: 15969739]

Evans KL, Gaston KJ, Frantz AC, Simeoni M, Sharp SP, McGowan A, Dawson DA, Walasz K, Partecke J, Burke T, Hatchwell BJ. Independent colonization of multiple urban centres by a formerly forest specialist bird species. Proc. R. Soc. Lond. B. 2009; 276:2403–2410.

Excoffier L, Lischer HEL. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resources. 2010; 10:564–567.

Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. Genetics. 2003; 164:1567–1587. [PubMed: 12930761]

Glez-Pena D, Gomez-Blanco D, Reboiro-Jato M, Fdez-Riverola F, Posada D. ALTER: program-oriented conversion of DNA and protein alignments. Nucleic Acids Res. 2010; 38:W14–W18. [PubMed: 20439312]

Goujon M, McWilliam H, Li WZ, Valentin F, Squizzato S, Paern J, Lopez R. A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res. 2010; 38:W695–W699. [PubMed: 20439314]

Harrington LC, Edman JD, Scott TW. Why do female *Aedes aegypti* (Diptera : Culicidae) feed preferentially and frequently on human blood? J. Med. Entomol. 2001; 38:411–422. [PubMed: 11372967]

Harrington LC, Scott TW, Lerdthusnee K, Coleman RC, Costero A, Clark GG, Jones JJ, Kitthawee S, Kittayapong P, Sithiprasasna R, Edman JD. Dispersal of the dengue vector *Aedes aegypti* within and between rural communities. Am. J. Trop. Med. Hyg. 2005; 72:209–220. [PubMed: 15741559]

Heled J, Drummond AJ. Bayesian Inference of Species Trees from Multilocus Data. Mol. Biol. Evol. 2010; 27:570–580. [PubMed: 19906793]

Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA. Population Genomics of Parallel Adaptation in Threespine Stickleback using Sequenced RAD Tags. PLoS Genet. 2010; 6:e1000862. [PubMed: 20195501]

Holstein M. Dynamics of *Aedes aegypti* distribution, density and seasonal prevalence in the Mediterranean area. Bull. Wld. Hlth. Org. 1967; 36:541–543.

Honorio NA, Silva WD, Leite PJ, Goncalves JM, Lounibos LP, Lourenco-de-Oliveira R. Dispersal of *Aedes aegypti* and *Aedes albopictus* (Diptera : Culicidae) in an urban endemic dengue area in the State of Rio de Janeiro, Brazil. Mem. Inst. Oswaldo Cruz. 2003; 98:191–198. [PubMed: 12764433]

Huber K, Ba Y, Dia I, Mathiot C, Sall AA, Diallo M. *Aedes aegypti* in Senegal: Genetic diversity and genetic structure of domestic and sylvatic populations. Am. J. Trop. Med. Hyg. 2008; 79:218–229. [PubMed: 18689628]

Jombart T. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics. 2008; 24:1403–1405. [PubMed: 18397895]

Jombart T, Ahmed I. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics. 2011; 27:3070–3071. [PubMed: 21926124]

Kark S, Iwaniuk A, Schalimtzek A, Banker E. Living in the city: can anyone become an ‘urban exploiter’? J. Biogeogr. 2007; 34:638–651.

Keller A. *Drosophila melanogaster*’s history as a human commensal. Curr. Biol. 2007; 17:R77–R81. [PubMed: 17276902]

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. Clustal W and clustal X version 2.0. Bioinformatics. 2007; 23:2947–2948. [PubMed: 17846036]

Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25:1754–1760. [PubMed: 19451168]

Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009; 25:1451–1452. [PubMed: 19346325]

Lounibos LP. Invasions by insect vectors of human disease. Annu. Rev.Entomol. 2002; 47:233–266. [PubMed: 11729075]

Lyimo IN, Ferguson HM. Ecological and evolutionary determinants of host species choice in mosquito vectors. Trends Parasitol. 2009; 25:189–196. [PubMed: 19269900]

Maciel-de-Freitas R, Lourenco-de-Oliveira R. Presumed unconstrained dispersal of *Aedes aegypti* in the city of Rio de Janeiro, Brazil. Rev. Saude Publ. 2009; 43:8–12.

Maddison, DR.; Maddison, WP. MacClade 4: Analysis of phylogeny and character evolution. Version 4.08a2005.

Mattingly PF. Genetical aspects of the *Aedes aegypti* problem I. Taxonomy and bionomics. Ann. Trop. Med. Parasit. 1957; 51:392–408. [PubMed: 13498658]

McClelland GAH. A worldwide survey of variation in scale pattern of the abdominal tergum of *Aedes aegypti* (L.) (Diptera: Culicidae). Trans. R. Entomol. Soc. Lond. 1974; 126:239–259.

McDonald PT. Population characteristics of domestic *Aedes aegypti* (Diptera: Culicidae) in villages on the Kenya coast. II. Dispersal within and between villages. J. Med. Entomol. 1977; 14:49–53. [PubMed: 903936]

Miller BR, Monath TP, Tabachnick WJ, Ezike VI. Epidemic yellow fever caused by an incompetent mosquito vector. Trop. Med. Parasitol. 1989; 40:396–399. [PubMed: 2623418]

Moore DF. Hybridization and mating behavior in *Aedes aegypti* (Diptera: Culicidae). J. Med. Entomol. 1979; 16:223–226. [PubMed: 537005]

Nasidi A, Monath TP, Decock K, Tomori O, Cordellier R, Olaleye OD, Harry TO, Adeniyi JA, Sorungbe AO, Ajosecoker AO, Vanderlaan G, Oyediran ABO. Urban yellow fever epidemic in western Nigeria, 1987. Trans. R. Soc. Trop. Med. Hyg. 1989; 83:401–406. [PubMed: 2617590]

Parmakelis A, Russello MA, Caccone A, Marcondes CB, Costa J, Forattini OP, Sallum MAM, Wilkerson RC, Powell JR. Short report : Historical analysis of a near disaster: *Anopheles gambiae* in Brazil. Am. J. Trop. Med. Hyg. 2008; 78:176–178. [PubMed: 18187802]

Partecke J, Gwinner E, Bensch S. Is urbanisation of European blackbirds (*Turdus merula*) associated with genetic differentiation? J. Ornithol. 2006; 147:549–552.

Paupy C, Brengues C, Kamgang B, Herve JP, Fontenille D, Simard F. Gene flow between domestic and sylvan populations of *Aedes aegypti* (Diptera : Culicidae) in North Cameroon. J. Med. Entomol. 2008; 45:391–400. [PubMed: 18533431]

Paupy C, Brengues C, Ndiath O, Toty C, Herve JP, Simard F. Morphological and genetic variability within *Aedes aegypti* in Niakhar, Senegal. Infect. Genet. Evol. 2010; 10:473–480. [PubMed: 20223297]

Petersen, JL. Ph.D. thesis. University of Notre Dame; Notre Dame, IN, USA: 1977. Behavior differences in two subspecies of *Aedes aegypti* (L.) (Diptera: Culicidae) in East Africa.

Pimm S, Raven P, Peterson A, Sekercioglu CH, Ehrlich PR. Human impacts on the rates of recent, present, and future bird extinctions. Proc. Natl. Acad. Sci. USA. 2006; 103:10941–10946. [PubMed: 16829570]

Posada D. jModelTest: Phylogenetic model averaging. Mol. Biol. Evol. 2008; 25:1253–1256. [PubMed: 18397919]

Powell JR, Tabachnick WJ, Arnold J. Genetics and the origin of a vector population: *Aedes aegypti*, a case study. Science. 1980; 208:1385–1387. [PubMed: 7375945]

Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000; 155:945–959. [PubMed: 10835412]

R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; Vienna, Austria: 2008.

Reiter P, Amador MA, Anderson RA, Clark GG. Dispersal of *Aedes aegypti* in an urban area after blood feeding as demonstrated by rubidium-marked eggs. Am. J. Trop. Med. Hyg. 1995; 52:177–179. [PubMed: 7872449]

Rosenberg NA. DISTRUCT: a program for the graphical display of population structure. Mol. Ecol. Notes. 2004; 4:137–138.

Rousset F. GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol. Ecol. Resources. 2008; 8:103–106.

Schofield CJ, Diotaiuti L, Dujardin JP. The process of domestication in Triatominae. Mem. Inst. Oswaldo Cruz. 1999; 94:375–378. [PubMed: 10677759]

Smith CEG. The history of dengue in tropical Asia and its probable relationship to the mosquito *Aedes aegypti*. J. Trop. Med. Hyg. 1956; 59:243–251. [PubMed: 13368255]

Somers G, Brown JE, Barrera R, Powell JR. Genetics and morphology of *Aedes aegypti* (Diptera: Culicidae) in septic tanks in Puerto Rico. J. Med. Entomol. 2011; 48:1095–1102. [PubMed: 22238867]

Souza-Neto JA, Sim S, Dimopoulos G. An evolutionary conserved function of the JAK-STAT pathway in anti-dengue defense. Proc. Natl. Acad. Sci. USA. 2009; 106:17841–17846. [PubMed: 19805194]

Sylla M, Bosio C, Urdaneta-Marquez L, Ndiaye M, Black WC. Gene Flow, Subspecies Composition, and Dengue Virus-2 Susceptibility among *Aedes aegypti* Collections in Senegal. PLoS Negl. Trop. Dis. 2009; 3:e408. [PubMed: 19365540]

Tabachnick WJ. Evolutionary genetics and insect borne disease. The yellow fever mosquito, *Aedes aegypti*. Am. Entomol. 1991; 37:14–24.

Tabachnick WJ, Munstermann LE, Powell JR. Genetic distinctness of sympatric forms of *Aedes aegypti* in East Africa. Evolution. 1979; 33:287–295.

Tabachnick WJ, Powell JR. A world-wide survey of genetic variation in the yellow fever mosquito, *Aedes aegypti*. Genet. Res. 1979; 34:215–229. [PubMed: 544311]

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol. Biol. Evol. 2011; 28:2731–2739. [PubMed: 21546353]

Tilman D, Lehman C. Human-caused environmental change: Impacts on plant diversity and evolution. PNAS. 2001; 98:5433–5440. [PubMed: 11344290]

Trpis M, Hausermann W. Dispersal and other population parameters of *Aedes aegypti* in an African village and their possible significance in epidemiology of vector-borne diseases. Am. J. Trop. Med. Hyg. 1986; 35:1263–1279. [PubMed: 3789275]

Vazeille-Falcoz M, Failloux AB, Mousson L, Elissa N, Rodhain F. Oral receptivity of *Aedes aegypti formosus* from Franceville (Gabon, Central Africa) for dengue type 2 virus. Bull. Soc. Pathol. Exot. 1999; 92:341–342. [PubMed: 10690473]

Wallis GP, Tabachnick WJ, Powell JR. Macrogeographic genetic variation in a human commensal: *Aedes aegypti*, the yellow fever mosquito. Genet. Res. 1983; 41:241–258. [PubMed: 6884770]

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

Yeh PJ, Price TD. Adaptive phenotypic plasticity and the successful colonization of a novel environment. Am. Nat. 2004; 164:531–542. [PubMed: 15459883]

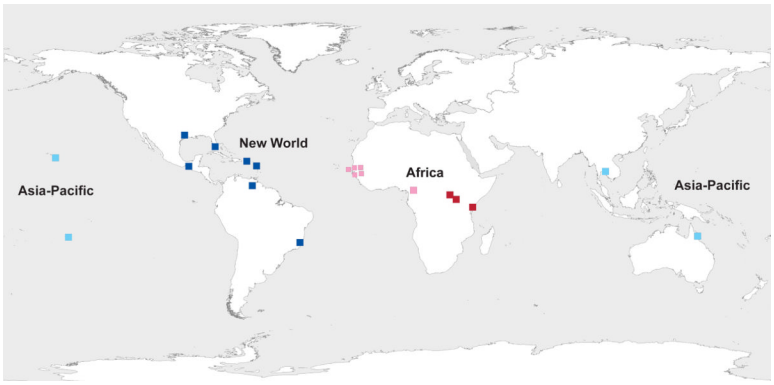

Figure 1. Map of worldwide collection sites of *Ae. aegypti* populations used in this study
Sites are color-coded by region: Asia-Pacific sites are shown in light blue, New World sites in dark blue, West/Central African sites in pink, and East African sites in red.

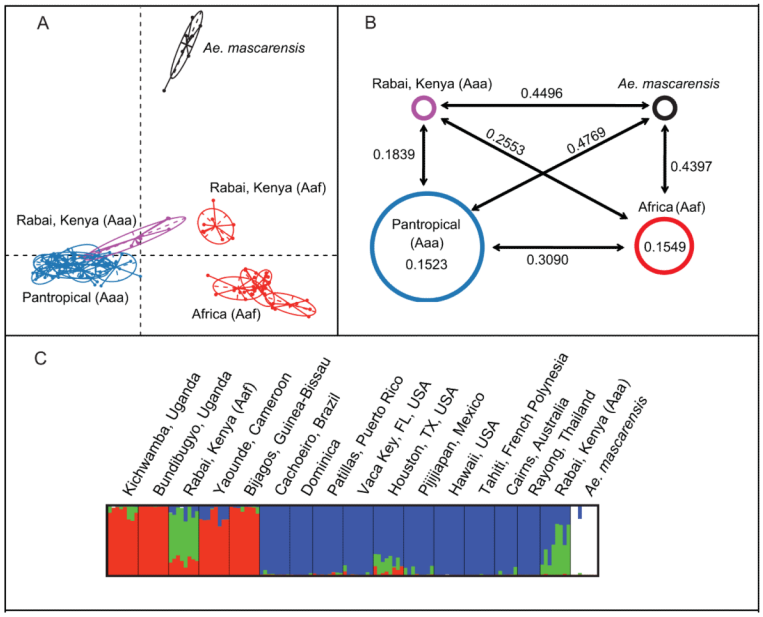

Figure 2. Broad-scale analyses of SNP data for *Ae. aegypti* and *Ae. mascarensis* populations

A) Individual-based PCA with major genetic groups denoted. African (i.e. Aaf) populations appear in red, while pantropical (i.e. Aaa) mosquitoes appear in blue. The Rabai Aaa population is shown in purple. *Ae. mascarensis* individuals are shown in black. B) Average pairwise F_{ST} values between and within major genetic groups of *Ae. aegypti* and *Ae. mascarensis*. Size of circles is proportional to number of populations sampled. C) STRUCTURE plot, K=4. Each vertical bar represents the probability of assignment of a single individual to each genetic cluster. Aaa=*Ae. ae. aegypti*, Aaf = *Ae. ae. formosus*. Of the four genetic groups identified by STRUCTURE, red corresponds roughly to African Aaf, blue to pantropical Aaa (including Rabai), and green appears to be a unique Rabai, Kenya geographic signal. The white cluster corresponds to outgroup *Ae. mascarensis* individuals.

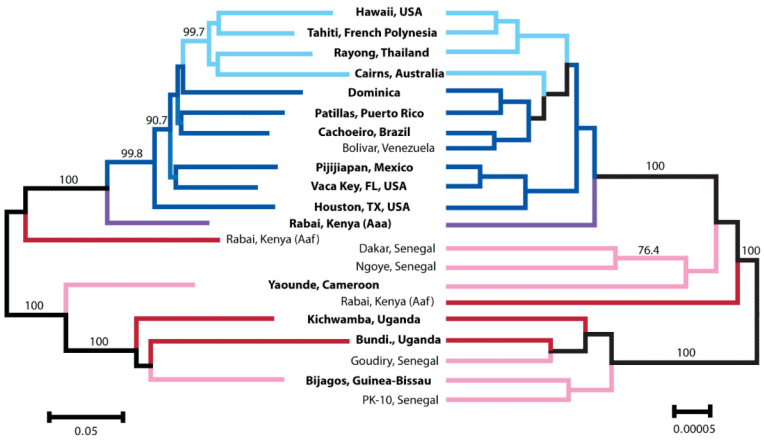

Figure 3. Evolutionary history of *Ae. aegypti* from SNPs and sequenced nuclear genes

Bootstrapped neighbor-joining network based on population pairwise chord-distances from 1504 SNPs (left). Species (i.e. population) tree based on phased *apoLp-2* paralogue, *CYP9J2*, and *SDR* sequences (right). Node support over 75% is shown on relevant branches. East African populations are shaded in red, West and Central African populations in pink, the Rabai (*Ae. ae. aegypti*) population in purple, New World populations in dark blue, and Asia-Pacific populations are shown in light blue. Rooting was inferred from nuclear gene trees (Fig. 4). Population labels in bold font refer to branches on both sides (i.e. to both the neighbor-joining network and the species tree), while labels in regular font refer only to the nearest branch.

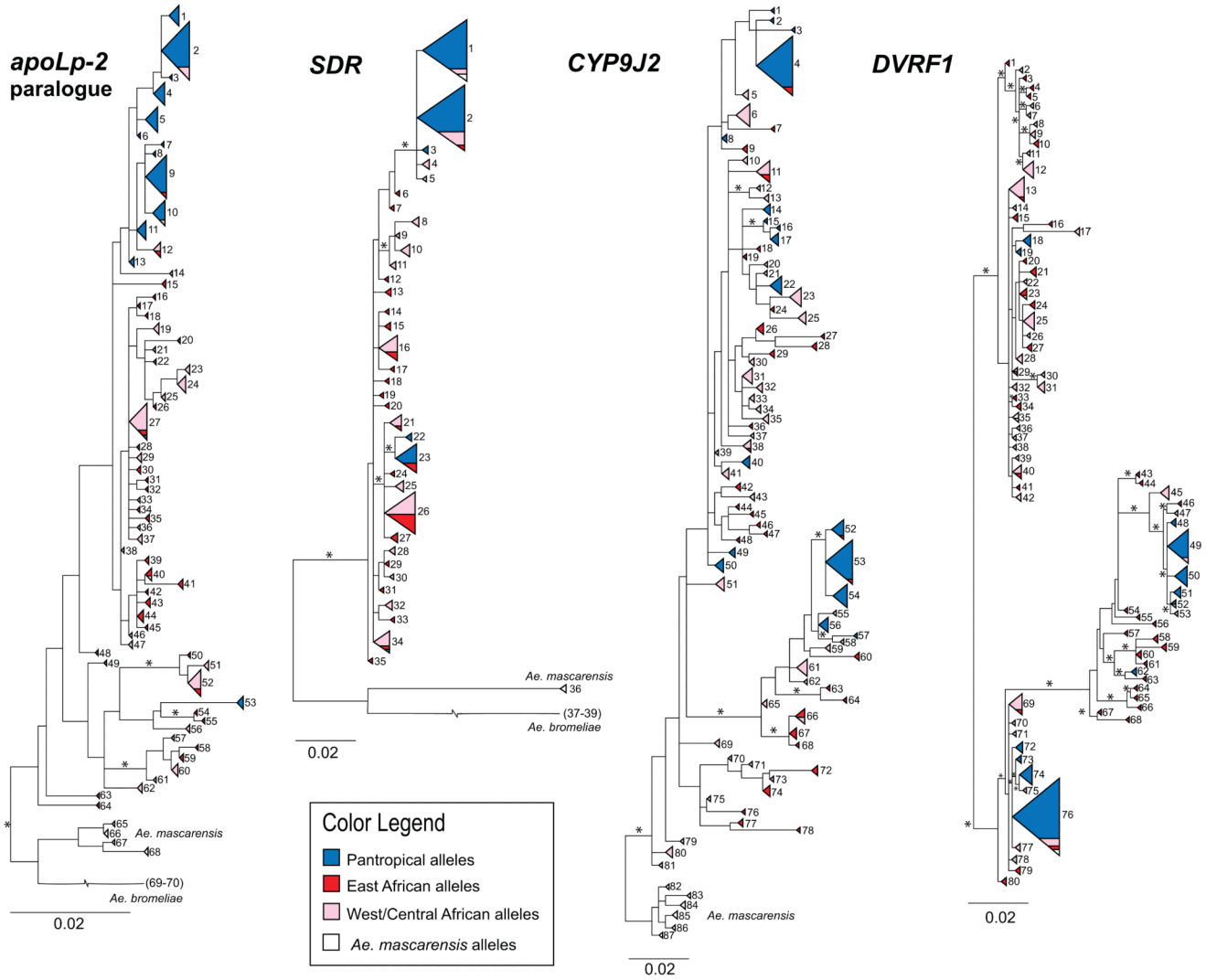

Figure 4.

Gene trees for *apoLp-2* paralogue, *SDR*, *CYP9J2*, and *DVRF1* (from left to right). Each triangle represents a single, unique haplotype, and the area of the triangle is proportional to the number of alleles with that sequence. Alleles from East Africa are colored red, West/Central African alleles are shown in pink, and pantropical alleles are colored blue. Alleles from *Ae. mascarensis* individuals are shown in white. Numbers next to each triangle/haplotype represent the haplotype number as denoted on the keys (Tables S2-S5). Nodes with bootstrap support over 60% are indicated with an asterisk. The broken line leading to *Ae. bromeliae* samples indicates long branch lengths. The distance to *Ae. bromeliae* is approximately 0.11 average substitutions per site to the tips for the *apoLp-2* paralogue and 0.12 substitutions per site for *SDR*. *Aedes bromeliae* haplotypes were pruned from the *CYP9J2* gene tree due to long branch lengths and unclear placement. The *DVRF1* gene tree is not rooted by outgroup sequences.

Table 1

Population information, including: collection site, global region, number of individual mosquitoes (N) analyzed with sequenced nuclear loci (nuDNA) and SNPs, number of generations in the lab, and year collected. % Missing data across all SNP loci for each population is shown in the final column.

Collection Site	Region	N (nuDNA/ SNPs)	Gen. in Lab	Year collected	% Missing
Kichwamba, Uganda*	E. Africa	10/8	0	2009	25.9%
Bundibugyo, Uganda*	E. Africa	10/8	0	2009	23.2%
Rabai, Kenya (*formosus*)*	E. Africa	10/8	0	2009	26.4%
Yaounde, Cameroon*	C. Africa	10/8	0	2009	23.2%
Dakar, Senegal*	W. Africa	10/-	0	2005	-
Goudiry, Senegal*	W. Africa	10/-	0	2007	-
Ngoye, Senegal*	W. Africa	10/-	0	2007	-
PK-10, Senegal*	W. Africa	10/-	0	2006	-
Bijagos, Guinea-Bissau*	W. Africa	10/8	0	2009	26.5%
Bolivar, Venezuela*	New World	10/-	2	2004	-
Cachoeiro de Itapemirim, Brazil	New World	10/8	1	2010	9.3%
Dominica*	New World	10/6	0	2009	8.7%
Patillas, Puerto Rico**	New World	10/8	0	2010	27.7%
Vaca Key, FL, USA*	New World	10/8	0	2009	10.4%
Houston, TX, USA*	New World	10/8	0	2009	12.1%
Pijijiapan, Mexico*	New World	10/8	1	2008	11.5%
Hawaii, USA	Asia/Pacific	10/8	0	2010	12.9%
Tahiti, French Polynesia*	Asia/Pacific	10/8	1	2010	10.8%
Cairns, Australia*	Asia/Pacific	10/6	0	2009	15.4%
Rayong, Thailand*	Asia/Pacific	10/6	2	2009	16.4%
Rabai, Kenya (*aegypti*)*	E. Africa	10/8	0	2009	13.5%
All Sites	-	210/129	-	-	-
Outgroup Species	**Country**	**N (nuDNA/ SNPs)**	**Gen. in Lab**	**Year collected**	**% Missing**
Ae. mascarensis	Mauritius	4/7	0	2011	59.7%
Ae. bromeliae	Kenya	4/-	0	2009	-

* Among those analyzed with microsatellite markers in Brown et al. 2011

** Analyzed with microsatellites in Somers et al. 2011

Table 2

Average values of nucleotide diversity (π) and haplotype diversity (Hd) across populations in selected regions for each of four nuclear loci.

Group/ Region	*apoLp-2* paralogue		*CYP9J2*		*SDR*		*DVRF1*	
	Hd	π	Hd	π	Hd	π	Hd	π
Africa	0.7152	0.0086	0.7908	0.0224	0.7638	0.0060	0.8871	0.0193
East	0.9493	0.0129	0.9213	0.0306	0.7263	0.0050	0.9613	0.0220
West/Central	0.5982	0.0065	0.7255	0.0183	0.7825	0.0065	0.8500	0.0179
Pantropical	0.5785	0.0031	0.4938	0.0178	0.5180	0.0032	0.4685	0.0134
New World	0.6406	0.0038	0.5150	0.0190	0.5700	0.0039	0.4949	0.0123
Asia-Pacific	0.4843	0.0021	0.4158	0.0130	0.4343	0.0019	0.3750	0.0127