

DR. PATRÍCIA SALGUEIRO (Orcid ID : 0000-0001-8628-3651)

Article type : Original Article

## TITLE

Phylogeography and invasion history of *Aedes aegypti*, the Dengue and Zika mosquito vector in Cape Verde islands (West Africa)

Running head

Dengue and Zika vector *Aedes aegypti* in Cape Verde (West Africa)

Authors:

Patrícia Salgueiro<sup>1\*</sup>, Célia Serrano<sup>1</sup>, Bruno Gomes<sup>1, 2</sup>, Joana Alves<sup>3</sup>, Carla A. Sousa<sup>1</sup>, Ana Abecasis<sup>1</sup>, João Pinto<sup>1</sup> \*Corresponding author: patsalgueiro@hotmail.com

Affiliations:

1- Global Health and Tropical Medicine, GHTM, Instituto de Higiene e Medicina Tropical, IHMT, Universidade Nova de Lisboa, UNL, Rua da Junqueira 100, 1349-008 Lisboa, Portugal.

2- Oswaldo Cruz Institute, IOC, Fundação Oswaldo Cruz, FIOCRUZ, Av. Brasil 4365 - Pavilhão 26 sala 207, 21040-360 Rio de Janeiro, Brasil

3- Direção Geral de Saúde /Instituto Nacional de Saúde Pública, Ministério da Saúde de Cabo Verde, Praia, Cabo Verde

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/eva.12834

This article is protected by copyright. All rights reserved.

E-mail addresses:

PS: patsalgueiro@hotmail.com;

CS: cs-celia\_serrano@hotmail.com;

BG: gsilva.bruno@gmail.com;

JA: alveslima@hotmail.com;

CAS: casousa@ihmt.unl.pt;

AA: ana.abecasis@ihmt.unl.pt;

JP: jpinto@ihmt.unl.pt

## ABSTRACT

*Aedes*-borne arboviruses have spread globally with outbreaks of vast impact on human populations and health systems. The West African archipelago of Cape Verde had its first outbreak of Dengue in 2009, at the time the largest recorded in Africa, and was one of the few African countries affected by the Zika virus epidemic. *Aedes aegypti* was the mosquito vector involved in both outbreaks. We performed a phylogeographic and population genetics study of *A. aegypti* in Cape Verde in order to infer the geographic origin and evolutionary history of this mosquito. These results are discussed with respect to the implications for vector control and prevention of future outbreaks. Mosquitoes captured before and after the Dengue outbreak on the islands of Santiago, Brava and Fogo were analyzed with two mitochondrial genes COI and ND4, 14 microsatellite loci and five *kdr* mutations. Genetic variability was comparable to other African populations. Our results suggest that *A. aegypti* invaded Cape Verde at the beginning of the Holocene from West Africa. Given the historic importance of Cape Verde in the transatlantic trade of the 16<sup>th</sup>–17<sup>th</sup> centuries, a possible

This article is protected by copyright. All rights reserved.

Accepted Article  
contribution to the genetic pool of the founding populations in the New World cannot be fully discarded. However, contemporary gene flow with the Americas is likely to be infrequent. No *kdr* mutations associated with pyrethroid resistance were detected. The implications for vector control and prevention of future outbreaks are discussed.

Keywords:

*Aedes aegypti*, Africa, Cape Verde, Dengue, phylogeography, population genetics, vector-control, Zika

## INTRODUCTION

*Aedes*-borne arboviral diseases have become a major global health concern. Dengue is the most important mosquito-borne viral disease in the world, causing 10,000–20,000 deaths per year (WHO, 2012) and almost 400 million new infections are estimated to occur annually (Bhatt et al., 2013). In 2015, a new outbreak caused by Zika virus has affected 84 countries (WHO, 2017) with more than 200,000 confirmed cases just in the Americas (Mitchell, 2016). Zika virus has been associated with severe congenital neurological abnormalities in infants born to mothers infected during pregnancy, reported in 31 countries (WHO, 2017). For these arboviruses, a proportion of cases is asymptomatic (Bhatt et al., 2013; Petersen et al., 2016), which together with poor disease surveillance and lack of point-of-care diagnostic tests may be potentially misdiagnosed and underreported (Wilder-Smith & Byass, 2016). This is particularly important in Africa where the majority of febrile illnesses are treated presumptively as malaria (Amarasinghe, Kuritsky, Letson, & Margolis, 2011; Stoler & Awandare, 2016). As a consequence, little is known about the epidemiology of these arboviruses in

the African continent (Amarasinghe, Kuritsky, Letson, & Margolis, 2011; CDC, 2015; Nutt & Adams, 2017; Sang & Dunster, 2001; Were, 2012). The control of Dengue and Zika is largely dependent on efficient and sustainable vector control measures.

Arboviruses like Dengue, Zika and Yellow Fever have spread worldwide following the expansion of their main vector *Aedes aegypti*. This mosquito has originated in Africa from an ancestral sylvatic and more zoophilic form *A. aegypti formosus*, which expanded from tropical forests to urban areas giving rise to a domestic and anthropophilic form known as *A. aegypti aegypti* (Bennett et al., 2016; Brown et al., 2011; Moore et al., 2013; Powell & Tabachnick, 2013). This form was the only that succeeded in invading the rest of the world, forming a monophyletic group (Bennett et al., 2016; Brown et al., 2014, 2011; Gloria-Soria et al., 2016). *Aedes aegypti* arrived to the New World together with the first Europeans and Africans during the historical trans-Atlantic shipping traffic between 1500s-1700s, followed by the first reports of Yellow Fever and Dengue in the region (Powell & Tabachnick, 2013).

Genetic evidence has confirmed a single out of Africa colonization event for *A. aegypti*, and its historic route of expansion throughout the world (Bennett et al., 2016; Brown et al., 2014, 2011; Moore et al., 2013).

After the invasion of the Americas, a cline of reduced genetic variability suggests spreading westwards with subsequent founder events, from the American continent to Asia and Oceania (Brown et al., 2014; Gloria-Soria et al., 2016). A mitochondrial DNA study reported two ancestral clades from which *A. aegypti* populations outside Africa have arisen (Moore et al., 2013), one associated with West Africa and another with East Africa. Nevertheless, a scenario where populations outside Africa originated from a single sample with the two mtDNA lineages from ancestral Africa cannot be discarded (Gloria-Soria et al., 2016). Presently in Africa, both *formosus* and *aegypti* subspecies co-occur and interbreed. The only exception seems to be a rural/forest population from the Rabai District in Kenya, where the two subspecies remain genetically distinct

(Brown et al., 2014, 2011; Gloria-Soria et al., 2016). Most likely due to increasing urbanization, populations of *A. aegypti* in Africa can now be found in many urban settings (Kamgang et al., 2013; Paupy et al., 2008) even if they fall genetically into the *formosus* group (Gloria-Soria et al., 2016). A study (Crawford et al., 2017) using exome sequences proposed West Africa, in particular Senegal, as the source of the America invasion by *A. aegypti*. Another recent study (Kotsakiozi et al., 2018), based on SNPs from 20 African populations, suggested Angola as the most likely origin.

The archipelago of Cape Verde is located at around 500 km West of Senegal (West Africa). It comprises nine inhabited islands (Fig. 1) and a current local population of about 500,000 inhabitants. Cape Verde islands had no human occupation when they were first reached by the Portuguese in the 1450's (Lobban, 2018). During most of the 16<sup>th</sup> and 17<sup>th</sup> centuries, the islands served as port-of-call in a transatlantic commercial slave network (Russell-Wood, 1998), where most of the ships from West Africa to Santiago island would continue to the New World (Lobban, 2018; Russell-Wood, 1998).

Dengue was reported for the first time in Cape Verde at the end of September 2009, in an outbreak considered at the time the largest ever recorded in Africa, having affected all islands. Over 20,000 cases (about 5% of the country population), 174 hemorrhagic fever cases and four deaths were reported (WHO, 2009). Dengue type-3 virus (DENV-3) was confirmed both in Cape Verde and in the concomitant dengue epidemic in Senegal (Franco & Di Caro A, 2010). In 2015, DENV-2 and DENV-4 were found circulating in field mosquito samples collected in Santiago Island (Guedes et al., 2017). In October 2015, an unprecedented epidemic of Zika virus was reported in Cape Verde (WHO, 2015), with more than 7500 suspected cases. By August 2016, 18 microcephaly cases were reported in the islands of Fogo, Santiago and Maio (Monteiro, 2016). This was the first time that a Zika strain associated with neurological damage in infants was detected in Africa (WHO, 2016). Given this fact, the timing of the epidemic and the high number of travelers visiting Cape Verde from the Americas,

it was suggested that the outbreak was likely caused by the Asian genotype circulating there (Lourenço et al., 2018).

*Aedes aegypti* is the only mosquito vector of these arboviruses so far detected in Cape Verde. This species was reported for the first time in S. Vicente Island and later on in all islands since 1964 (Ribeiro, Ramos, Capela, & Pires, 1980). Morphological identification of the subspecies was attempted with specimens from a single island (Santiago) and only *A. a. formosus* was detected (Vazeille et al., 2013). Low vector competence has been associated with African strains of *A. aegypti*, particularly *A. a. formosus*, with a lower susceptibility to DENV-2 (Diallo et al., 2013; Sylla, Bosio, Urdaneta-Marquez, Ndiaye, & Black, 2009). However, *A. aegypti* from Santiago Island collected in 2010 showed a moderate ability to transmit the epidemic DENV-3, and high susceptibility to chikungunya and yellow fever viruses (Vazeille et al., 2013). A more recent study based on collections carried out in 2012 found that mosquitoes displayed higher vector competence for DENV-2 and DENV-3 when compared to DENV-1 and DENV-4 (Moura et al., 2015).

In Cape Verde, integrated vector control strategies have been directed to both malaria and dengue vectors, *Anopheles arabiensis* and *A. aegypti*, respectively, through source reduction, diesel use, biological control with fish (*Gambusia sp.*), and chemical control with the insecticides temephos and deltamethrin (de Pina, 2013; Ministry of Health Cape Verde, WHO, & University of California, 2012). The efficacy of insecticide-based vector control has been threatened by the evolution of insecticide resistance worldwide (Smith, Kasai, & Scott, 2016). Insecticide susceptibility tests performed in *A. aegypti* from Santiago island during the dengue outbreak in 2009 revealed resistance to DDT but susceptibility to pyrethroids (Dia et al., 2012). Later on, *A. aegypti* mosquitoes collected in the same island in 2012 already exhibited resistance to pyrethroids (deltamethrin, cypermethrin) and to the organophosphate temephos (Rocha et al., 2015). A major resistance mechanism affects a gene in the insect's voltage-gated sodium channel. In this gene, several mutations, known as knockdown resistance (*kdr*) mutations, have been described and associated to DDT and pyrethroids resistance

worldwide (Du, Nomura, Zhorov, & Dong, 2016). These mutations were not detected in *A. aegypti* collected in 2012 (Rocha et al., 2015).

Since Cape Verde is in a strategic route linking Africa, Europe and the Americas, by sea or by air, thus having a high risk of introduction of new strains of arboviruses or new vectors, it is of great importance to study the local population structure of *A. aegypti*. Therefore, we have developed a comprehensive phylogenetic and population genetics study of *A. aegypti* from three islands of Cape Verde with samples collected before and after the dengue outbreak of 2009. We have measured levels of genetic variability and determine the population structure, inferred the evolutionary history and detect possible origins for this insular population and evaluated the allelic composition in three *kdr* sites.

## MATERIALS AND METHODS

### *Mosquito sampling*

*Aedes aegypti* mosquitoes were collected between November and December 2007 in three islands of the leeward group of the Cape Verde archipelago: Santiago, Fogo, and Brava, and in April 2010 from Santiago island (Fig. 1, Table 1). The samples consisted of larvae or pupae collected from 12 breeding sites. These were reared to adult stage and then morphologically identified to species (Ribeiro & Ramos, 1995; Ribeiro et al., 1980). The detailed procedure of the entomological survey has been described elsewhere (Alves et al., 2010). Adult specimens were individually preserved in tubes filled with silica gel and kept at room temperature until DNA extraction. Genomic DNA was extracted using a phenol: chloroform protocol (Donnelly, Cuamba, Charlwood, Collins, & Townson, 1999).

*Mitochondrial DNA sequencing and phylogenetic analysis*

We have analyzed two mitochondrial genes: cytochrome oxidase subunit I gene (COI) and NADH dehydrogenase subunit 4 gene (ND4). The gene COI was amplified with primers published by (Paupy et al., 2012). The gene ND4 was amplified using the primers described by (Paduan & Ribolla, 2008).

PCR conditions for both genes were optimized by (Seixas et al., 2014).

PCR products were purified and sequenced directly with the PCR primers (forward and reverse sequences obtained for each individual). Only sequences with no double peaks, suggestive of nuclear mtDNA segments (or NUMTs) were subsequently analyzed. Sequences were aligned and numbered in reference to the *A. aegypti* complete COI sequence (bases 1-1537, Genbank accession number AF390098 (Morlais & Severson, 2002)) and ND4 (bases 1-1344, Genbank accession number DQ440274 (da Costa-Ribeiro, Lourenço-De-Oliveira, & Failloux, 2007)).

Summary statistics, including the number of haplotypes (Hap), Haplotype diversity (Hd), nucleotide diversity ( $\pi$ ), standard neutrality tests: Tajima's D (Tajima, 1989), Fu's  $F_s$  (Fu, 1997) and mismatch distribution (Schneider & Excoffier, 1999) were calculated using ARLEQUIN v.3.5 (Excoffier & Lischer, 2010). When the neutrality test values were negative and significant and the mismatch distribution curve was unimodal, the population was considered to fit the sudden expansion model. Then, the expansion time was estimated (Rogers & Harpending, 1992) assuming a substitution rate of 0.023 (COI) and 0.0125 (ND4) (Brower, 1994; Yu et al., 1999) and generation time of 15 generations per year (Crawford et al., 2017). DNAsp 6 (Rozas et al., 2017) was used to identify unique haplotypes.

The haplotypes obtained from the mosquito population of Cape Verde were analyzed together with other sequences of the same genes available in GenBank, including sequences of *Aedes albopictus* used as outgroup. The 368 ND4 and 57 COI sequences are detailed in Appendixes 1 and 3, respectively. In order to increase the representativeness of African samples, other COI sequences (Bennett et al., 2016) were included in the study, resulting in a new alignment of 94 COI sequences but with only 493 base pairs (bp) (Appendix 3).

This article is protected by copyright. All rights reserved.

We used ModelFinder (Kalyaanamoorthy, Minh, Wong, Haeseler, & Jermiin, 2017) implemented in IQ-TREE web-server (Trifinopoulos, Nguyen, von Haeseler, & Minh, 2016) to determine the best fitting substitution model.

Bayesian Markov chain Monte Carlo inference, as implemented in BEAST 1.8.4 (Drummond, Suchard, Xie, & Rambaut, 2012), was used to reconstruct the phylogenies that best describe the evolutionary history of the data set, given the sequences and a priori information of the geographical traits. The substitution model HKY (Hasegawa, Kishino, & Yano, 1985) with gamma and invariant sites and three partitions into codon positions were used. Furthermore, a Bayesian skyline population growth model was used (Drummond, Rambaut, Shapiro, & Pybus, 2005) which allows for flexibility in demographic reconstruction and a strict molecular clock was assumed with a substitution rate of 0.023 for COI and 0.0125 for ND4, as previously described (Brower, 1994; Yu et al., 1999). The analyses were run in 3-5 separate independent runs in BEAST with 100,000,000 generations, sampled every 50,000 runs (ND4), 100,000,000 sampled every 10,000 runs (COI). To analyze convergence and stability, we used TRACER v1.6 software (Rambaut, Drummond, Xie, Abebe, & Suchard, 2017).

TREEANNOTATOR (Drummond et al., 2012), was used to estimate the final Maximum Clade Credibility Tree with a burn-in of 10%. Trees were visualized and edited with FIGTREE 1.4.3. (Rambaut, 2016).

In order to more accurately estimate the time to the most recent ancestor (TMRCA), the clade with Cape Verdean haplotypes obtained from the previous analysis, was reanalyzed in BEAST.

Because evolutionary relationships among populations within species can be reticulate rather than bifurcating, haplotypes were connected on median-joining networks followed by maximum parsimony to eliminate unnecessary median vectors and links (Bandelt, Forster, & Röhl, 1999) with the software NETWORK 5 available at website fluxus-engineering.com. This program was also used

to connect ND4 haplotypes from Cape Verde and Africa (54 sequences). In this case, we have applied a star contraction pre-procedure.

### *Microsatellites genotyping and data analysis*

The following 14 microsatellite loci were amplified by PCR using fluorescent-labelled primers: A1, AC2, AC4, AC7, AG1, AG2, AG4, AG5, B2, B3, CT2 (Brown et al., 2011; Slotman et al., 2007), 12ACG1, 88AAT1, 201AAT1 (Lovin et al., 2009), and the corresponding amplification conditions. Amplified fragments were separated by capillary electrophoresis and sizes scored using the software GENE-MARKER (SoftGenetics).

Allele richness (Rs) and private alleles richness (P) were calculated in HP-RARE (Steven T. Kalinowski, 2004). Estimates of expected heterozygosity (He), Hardy-Weinberg equilibrium (HWE) tests were performed with the software ARLEQUIN v.3.5 (Excoffier & Lischer, 2010). Linkage disequilibrium (LD) between pairs of loci was tested using the log likelihood ratio statistic available in GENEPOP v.4.2 (Raymond & Rousset, 1995; Rousset, 2008). To detect the presence of null alleles, we used the software MICRO-CHECKER (Van Oosterhout, Hutchinson, Wills, & Shipley, 2004).

As larval sampling methods can lead to the collection of closely related individuals, we have used the maximum-likelihood method in ML-RELATE (Kalinowski, Wagner, & Taper, 2006) to calculate proportions of related individuals within each year, in order to confirm that the assumption of independent genotypes was not violated. For each pair of individuals, log-likelihood was calculated for four relatedness categories: unrelated, parent-offspring, full-siblings, and half-siblings.

Estimates of current effective population size (Ne) were calculated by the bias-corrected LD method (Waples & Do, 2010), using NeEstimator v.2.01 (Do et al., 2014). Because rare alleles may bias LD Ne estimates, alleles with frequency below 0.05 were not considered.

In order to detect recent population changes, Wilcoxon signed-ranks tests of heterozygosity were performed with BOTTLENECK v.1.2.02 (Piry, Luikart, & Cornuet, 1999). Calculations were done under the stepwise mutation model (SMM) and a two-phase model (TPM) with 30% of indels larger than one repeat, based on 1,000 replications.

Genetic differentiation between the two years was measured by pairwise  $R_{st}$  and  $F_{st}$  with ARLEQUIN v.3.5 (Excoffier & Lischer, 2010). Significance of estimates was assessed by 5000 permutation tests.

Bayesian clustering analysis implemented in STRUCTURE v.2.3.4 (Pritchard, Stephens, & Donnelly, 2000) was performed using an admixture model, without prior information on sampling locations, assuming correlated allele frequencies among populations ( $\lambda$  was set at 1, default value). Twenty independent runs with 100,000 burn-in steps and 500,000 iterations were done for each number of clusters (K). For this analysis our data was combined with a global microsatellite data set from available in VectorBase/Project ID:VBS0061862 (Gloria-Soria et al., 2016), using nine common loci (AC2, AC4, CT2, AG1, AG2, AG5, A1, B2, B3). K was tested from K=1-5 at the global scale (81 populations in total), and from K=1-10 at the African scale (23 African populations and subsequent analyses). The most likely value of K was determined with STRUCTURE HARVESTER (Earl & vonHoldt, 2012; Evanno, Regnaut, & Goudet, 2005). Results were visualized using CLUMPAK (Kopelman, Mayzel, Jakobsson, Rosenberg, & Mayrose, 2015).

In order to estimate levels of contemporary gene flow between Cape Verde and other African populations, we have also used the above mentioned microsatellite database to detect recent migrants with assignment tests performed with GENECLASS v.2 (Piry et al., 2004). The Bayesian likelihood criterion (Rannala & Mountain, 1997) and the as assignment criterion Likelihood ratio (L):  $L_{\text{home}} / L_{\text{max}}$  (Paetkau, Slade, Burden, & Estoup, 2004) was used with 1,000 simulated individuals.

Bonferroni corrections were used to adjust critical probability values for multiple tests (Rice, 1989).

### *Kdr mutations*

We have genotyped three *kdr* loci of the voltage-gated sodium channel gene in the *A. aegypti* samples from Santiago Island, in search of the following five mutations: I1011M, I1011V, V1016I, V1016G and F1534C. For the first four, we amplified and sequenced the intron spanning exons 20 and 21 (Saavedra-Rodriguez et al., 2007). For F1534C, we performed a tetra primer PCR reaction (Harris, Rajatileka, & Ranson, 2010).

## RESULTS

### *ND4 mitochondrial gene*

The alignment of 360bp of the ND4 gene sequenced in 42 Cape-Verdean *A. aegypti* samples resulted in seven distinct haplotypes (Table S1). Some of these haplotypes matched sequences already published in GenBank: ND4-1 (JQ926710-Ivory Coast, KC800689-Nigeria, KM042185-Colombia); ND4-2 (JN089748-Brazil, JN896665-Colombia); ND4-3 (JN896664-Colombia); ND4-5 (AY906847-Brazil, EF562501-Cameroon, JQ926708-Bolivia, JQ926709-Ivory Coast, JQ926714-Mexico, JQ926717-Guinea). ND4-4, -6 and -7 are unique haplotypes. This alignment includes 16 variable sites, two of which are parsimony informative, with a nucleotide diversity of  $\pi=0.002$  and a haplotype diversity of  $h=0.609$ . Most of the substitutions were transitions (15) and there was a single transversion. Both neutrality tests for Santiago 2007 showed negative results but only Fu's FS was significant (Table 2). Mismatch analysis of Santiago 2007 gave a unimodal distribution of pairwise differences among haplotypes indicative of sudden expansion. The estimator  $\tau$  (tau) was 0.762 (95%CI: 0.313-1.432) and the correspondent time since expansion (T) was 5642 years (95%CI: 2315-10605), assuming a substitution rate of 0.0125 (Yu et al., 1999).

The median-joining network based on the ND4 gene revealed closely related haplotypes in a star-like topology (Fig. S1a). The central haplotype ND4-1 was the most abundant (57%, Table S1) and the only present in the three islands sampled, while the remaining haplotypes had frequencies of less than 20%. Haplotype ND4-2 was only found in Brava Island. Comparing the temporal samples of Santiago island, in 2010 we found two haplotypes previously detected in 2007 (ND4-1 and ND4-5) and one new haplotype (ND4-7).

The alignment of these seven haplotypes with other published GenBank sequences for the same gene resulted in 368 sequences (Appendix 1a). This alignment included two sequences of *A. albopictus* (outgroup) and *A. aegypti* sequences from 26 countries, resulting in 96 haplotypes (Appendix 1b). The best fit model selected by both Akaike and Bayesian information criteria was GTR (Tavaré, 1986) with a proportion of invariable sites ( $I = 0.516$ ) and a gamma shape parameter ( $G = 0.635$ ) with 4 categories.

Given the low support of the clades in the initial analysis, possibly due to the high similarity of the sequences, a new analysis was performed with the initial alignment of the 96 haplotypes (Appendix 1b) from which 47 South American sequences were eliminated (due to their overrepresentation in GenBank, Appendix 1c). The ND4 phylogenetic tree (Fig. 2) showed two main clades supported by strong posterior probability values (i.e.  $> 0.95$ ). In the first one, haplotypes from Cape Verde occurred mainly associated with West African sequences, and some others from Asia and America, whereas the other clade contained most of East African sequences and other Asian and American sequences.

When the Cape Verdean and West African common clade (18 sequences, 16 haplotypes, Appendix 1d) was run alone in BEAST, we have obtained a height median (95%HPD) of 0.443 (0.257-0.686) corresponding to a TMRCA of 443,000 years (257,000-686,000).

From the initial set of 366 sequences (outgroups not considered), when the analysis was restricted to sequences of African origin (Fig. S2), we detected a star-like network with most haplotypes from West Africa in the center and the Cape Verdean haplotypes derived from the previous ones. In this network, many of East African haplotypes appear separated from most of West African and Cape Verdean ones by a higher number of mutations, and further apart from the center of the star.

#### *COI mitochondrial gene*

A fragment of 764bp from the COI gene was sequenced in 50 *A. aegypti* individuals. The alignment resulted in 21 polymorphic sites defining 20 distinct haplotypes (Table S2). Only the haplotype COI-1 matched one sequence already published in GenBank: JQ926693-Ivory Coast, the other 19 haplotypes seem to be unique. All observed substitutions were transitions. This alignment includes 20 variable sites, nine of which are parsimony informative, with  $\pi=0.003$  and  $h=0.943$ . Both neutrality tests for Santiago 2007 showed negative results but only Fu's FS was significant (Table 2). The mismatch analysis in Santiago 2007 revealed a unimodal distribution of pairwise differences among haplotypes consistent with a model of sudden expansion.  $\tau$ -tau was 3.031 (95% CI: 2.041-4.133) and the correspondent time since expansion  $T=5741$  years (95% CI: 3866-7827), assuming a substitution rate of 0.023 (Brower, 1994).

The median-joining network based on the COI gene revealed also related haplotypes in a star-like topology (Fig. S1b). Haplotype COI-15 in the center of the star was the only present in the three sampled islands. All except one haplotype (COI-7) were present in Santiago Island. COI-7 was only detected in Brava. In 2010, we have found four new haplotypes undetected in 2007 (COI-17, 18, 19, 20), while other three haplotypes were also present in the past sample of 2007 (COI-4, 6, 13).

The alignment of the 20 haplotypes with other published 74 GenBank sequences from 24 countries for the same gene, resulted in 94 sequences of 493bp (Appendix 3). This alignment resulted in 67

haplotypes with a nucleotide diversity of  $\pi=0.01671$ , 98 variable sites, 37 of which are parsimony informative, 61 singleton variable sites. The best fit model selected by the Bayesian information criteria was TN (Tamura & Nei, 1993) with a proportion of invariable sites ( $I=0.7948$ ) and gamma shape parameter ( $G=1.073$ ).

The resulting COI Bayesian tree (Fig. 3) showed two main clades. Haplotypes from Cape Verde were all concentrated in one of the clades grouping with all other West African sequences and some Central African ones. The other clade included all the Eastern Africa lineages. American and Asian sequences were present in both clades.

When the Cape Verdean and West African common clade (25 sequences, Appendix 3b) was run alone in BEAST, we have obtained a height median (95%HPD) of 0.126 (0.061-0.238) corresponding to a TMRCA of 126,000 years (61,000-238,000).

#### *Microsatellite genetic diversity*

We genotyped 70 mosquitoes from Santiago Island, 47 individuals collected in 2007 and 23 in 2010. The 14 loci were polymorphic with allele richness ranging from 4 to 8 and private allele richness from 0 to 4. The expected heterozygosity varied between 0.494 and 0.783 (Table S3). Regarding HWE tests (Table S3), we have found three significant tests in 2007 (A1, AG4, 12ACG1) and two in 2010 (AG1, 201AAT1), all due to heterozygote deficit. None of the loci deviated from HWE expectations in both samples. Null alleles were suspected in four loci (A1, 12ACG1, 88AAT1, 201AAT1). The microsatellites used in this study have been extensively validated in previous studies (Brown et al., 2011; Gloria-Soria et al., 2016; Lovin et al., 2009) where deviations from HWE were associated with null alleles, but at frequencies that did not affect the assessment of population structure. Eleven out of 182 (6%) pairwise associations of loci showed significant LD but no pair of loci was consistently

linked in both samples. The percentage of unrelated individuals was 87% in 2007 and 82% in 2010 (i.e., no alleles among pairs of individuals were identical by descent, Table S4).

The two temporal samples showed low but significant ( $P < 0.0001$ ) genetic differentiation with  $R_{st} = 0.027$  and  $F_{st} = 0.043$ . Estimates of current effective population size decreased from 55 in 2007 to 8 in 2010, with non-overlapping 95% confidence intervals, indicative of significant differences in  $N_e$  between years (Table 3). Tests performed with BOTTLENECK v.1.2.02 (Piry, Luikart, & Cornuet, 1999) revealed significant heterozygote excess in the sample of Cape Verde 2010 under the TPM (Table 3).

#### *Population structure*

Fig. 4 illustrates the results of the clustering analysis implemented by STRUCTURE v.2.3.4 (Pritchard, Stephens, & Donnelly, 2000). The best  $K$  for the global dataset with 81 world populations (Fig. 4A) was two clusters corresponding to the two subspecies *A. aegypti aegypti* and *A. aegypti formosus*, as seen previously (Gloria-Soria et al., 2016). All individuals from Cape Verde belonged to the *formosus* cluster. When only the samples from Africa were analyzed (Fig. 4B), two clusters were obtained, roughly separating most of the samples from Cape Verde, Senegal and a few from Kenya (blue cluster: populations 1-9 from West Africa, 21 and 23 from East Africa), from the rest (orange cluster). In order to better detail this clustering result, we selected individuals from the blue cluster with a probability of assignment ( $Q$ )  $> 0.90$  re-analyzed them in STRUCTURE v.2.3.4 (Pritchard, Stephens, & Donnelly, 2000) with the same previous conditions. The result is presented in Fig. 4C, including 59 out of the original 71 Cape Verdean samples. The best  $K$  was again  $K=2$ , but this time the Senegalese population of Goudiry (8 in Fig. 4) grouped with the Kenyan Rabai-outdoor collection (21 in Fig. 4), leaving the rest to another cluster, where Cape Verde was incorporated.

GENECLASS v.2 (Piry et al., 2004) assigned only one potential contemporary migrant from Goudiry, Senegal (population 8 in Fig. 4) to Cape Verde 2010, with a  $L=4.090$ . No migrants from Cape Verde to continental Africa were detected.

#### *Kdr mutations*

The analysis of the *kdr* gene from 74 individuals revealed that the *A. aegypti* population from Cape Verde was monomorphic with 100% of wild-type alleles at the three loci analyzed (i.e. 1011I, 1016V and 1534F).

#### DISCUSSION

Phylogeographic and population genetic analysis of *A. aegypti* from Cape Verde Islands revealed the main elements of the invasion history and recent demographic events of this insular mosquito population. Our results suggest an ancient West African origin and no evidence of recent founder events. We report comparable levels of genetic diversity, within the range of values for other African populations for both mitochondrial (Bennett et al., 2016; Paupy et al., 2008) and nuclear markers (Gloria-Soria et al., 2016). However, there was a low but significant inter-temporal genetic differentiation associated with a signal of population bottleneck and a reduction in effective population size in 2010. The target site mechanism of knockdown resistance was likely not present in the population of Santiago island in 2007 and 2010.

### *West African Origin*

Clustering analysis of microsatellite data indicate that the Cape-Verdean *A. aegypti* population belongs to the subspecies *formosus*. This result agrees with previous analyses based on morphological data (Vazeille et al., 2013). Microsatellite data also suggests a West African origin, possibly having *A. aegypti* from Senegal as the source population. The levels of expected heterozygosity are within the average of African populations (Gloria-Soria et al., 2016), and similar to the only African insular sample from Guinea-Bissau (sample 10-Bijagos, GW in Fig. 4). All other island populations analyzed in that study showed lower levels of genetic diversity than Cape Verde.

The matrilineages of *A. aegypti* from Cape Verde occurred in a basal clade mainly associated with West African mosquito populations for both ND4 and COI (Fig. 2-3). Some of the most abundant haplotypes had matching sequences with West African samples from Ivory Coast, Nigeria, Cameroon and Guinea. The levels of nucleotide diversity ( $\pi$ ) are similar to values reported in West Africa for both COI and ND4 (Bennett et al., 2016; Paupy et al., 2008), and much lower than the values reported for East Africa (Bennett et al., 2016), America (Eskildsen et al., 2018; Jaimes-Dueñez, Arboleda, Triana-Chávez, & Gómez-Palacio, 2015; Scarpassa, Cardoza, & Cardoso, 2008; Twerdochlib et al., 2012) and Asia (Bosio et al., 2005), possibly due to multiple introduction events in those continents (Bracco, Capurro, Lourenço-de-Oliveira, & Sallum, 2007; Powell & Tabachnick, 2013).

### *Invasion history and expansion time*

Our analysis suggests that the TMRCA of Cape-Verdean and West African mtDNA lineages is in the late Pleistocene (ND4Median= 443,000 years, COI Median= 126,000 years, 95%HPD=61,000 - 686,000).

In the three Cape-Verdean islands surveyed, we found several unique haplotypes very closely related in a star-like network in both ND4 and COI (Fig. S1). Given the highest haplotype diversity in Santiago Island and ubiquity of central haplotypes in the trees (ND4-1 and COI-15 in Fig. S1) it is likely that the colonization of the more western and smaller islands of Fogo and Brava was made from the island of Santiago (Fig. 1). However, a more detailed history of these founder events is not possible due to the reduced sampling. It would be interesting to collect mosquito samples from the other islands to unveil the complete history of *A. aegypti* in the archipelago. Cape Verde has a geological age of 25.6 Myr and is spread over 58,000 km<sup>2</sup> in two diverging chains. With older islands in the east and younger ones in the west, like Santiago Island dating to 10 mya (Duarte & Romeiras, 2009), possible different colonization scenarios can be expected.

Focusing on Santiago Island from 2007, besides the already mentioned star-like tree of very close haplotypes, we found a concordant unimodal mismatch distribution of mutations, and also a significantly negative Fu's FS neutrality test (Table 2). Together, these are genetic signatures of population expansion (Fu, 1997; Harpending, 1994). The expansion time calculated for both genes overlapped, suggesting that expansion may have started at around 5,700 years ago. This estimate predates the occupation of the island of Santiago by humans, in the 15<sup>th</sup> century AD. Therefore, genetic evidence suggests that *A. aegypti* invaded Santiago Island in the early Holocene without the aid of humans. This would correspond to a colonization of the ancestral sylvatic and zoophilic *A. aegypti formosus*, possibly blood feeding on available reptile (Heatwole & Shine, 1976) or bird (Tandon & Ray, 2000) hosts already present on the island. In Cape Verde, there is evidence for reptile presence before 6.2 mya in the late Miocene (Carranza, Arnold, Mateo, & López-Jurado, 2001) and bird species establishment during the Pleistocene (2.6-0.01 mya) (Bourne, 1951).

Cape Verde played an important role in the historical slave trade network (Lobban, 2018). Between the 16<sup>th</sup> and 17<sup>th</sup> century, Santiago island was an obligatory stop for ships going from the coast of Angola, São Tomé, Ghana and Guinea-Bissau to Brazil and the West Indies (Lobban, 2018; Russell-

Wood, 1998). This could have promoted ancient gene flow with continental Africa and the introduction of *A. aegypti* from Cape Verde in the New World by passive transportation. However, both mtDNA and microsatellites point to the presence of an *A. a. formosus* population in Cape Verde, which does not agree with the hypothesis of an early African domestication of *A. aegypti* followed by subsequent expansion to the Americas. Furthermore, Cape-Verdean mtDNA haplotypes for both ND4 and COI cluster in subclades composed almost exclusively by West African haplotypes, suggesting that recent gene flow between Cape Verde and the Americas has been infrequent. Nevertheless, knowing that the first step for the out-of-Africa dispersion of *A. aegypti* was from the African West coast (Crawford et al., 2017; Kotsakiozi et al., 2018) towards the Americas during the Atlantic slave trade, one cannot rule out the possibility that the long-established population of *A. aegypti* in Cape Verde could have contributed to the gene pool of the founding populations in the New World. It would be important to compare Cape Verde with additional populations from Senegal and Angola to better understand the relative contribution of each of these *A. a. formosus* populations to the founding of *A. a. aegypti* in the Americas.

#### *Evidence of recent demographic perturbation*

When comparing samples from 2007 and 2010 from Santiago Island, there was significant temporal genetic differentiation associated with a signal of population bottleneck and a significant reduction in effective population size in the sample of 2010. Although the reduced sample size of 2010 could have influenced estimates, these observations could also be a consequence of the increased vector control efforts implemented during the late 2009 dengue outbreak. In response to the epidemic, local health authorities have intensified treatment of breeding sites with temephos and indoor/outdoor spraying with deltamethrin (Dia et al., 2012; Rocha et al., 2015). This could have affected the effective population size of the local *A. aegypti* population. Absence of *kdr* mutations in both the 2007 and 2010 samples is in line with a field study carried out in November 2009 that

revealed susceptibility of *A. aegypti* from Cape Verde to pyrethroid insecticides (Dia et al., 2012).

However, this situation has changed in 2012 and 2014, with the first reports of resistance to these insecticides (Rocha et al., 2015). This reinforces the hypothesis of insecticide-based reduction of *N* of the mosquito population, and the increase of selection pressure for insecticide resistance during the 2009 dengue epidemic.

### *Epidemic preparedness and implications for vector control*

The role of airports and airlines in the spread of vector-borne diseases has been helpful in predicting the risks of vector-borne disease importation and establishment (Huang et al., 2012; Tatem et al., 2012). Cape Verde is in a strategic route linking Africa and South America. The major Dengue outbreak in Cape Verde was caused by a DENV3 virus that originated from Senegal (Franco & Di Caro, 2010) and the local Zika outbreak had its most likely source in Brazil (WHO, 2016). In 2015, Cape Verde received more than 7,000 travelers from Zika virus-affected countries, including direct flights from Brazil (Bogoch et al., 2016; Lourenço et al., 2018). Therefore, the risk of introduction of new virus species/strains and new vectors (e.g. *A. albopictus*) is real and surveillance should thus be prioritized. The challenges imposed by the recent Dengue and Zika epidemics highlight the need for a well-trained health care workforce capacitated for rapid response under emergency situations caused by vector-borne diseases.

Although mutations associated with *kdr* resistance have not yet been reported in Cape Verde, the first records of insecticide resistance recently detected in Santiago island raise caution for the sustainability of insecticide-based vector control. Integrated vector management is being implemented in Cape Verde given that the country is in the pre-elimination stage for malaria (Ministry of Health Cape Verde et al., 2012) and has recently suffered from Dengue and Zika outbreaks (WHO, 2009, 2015). In this context, the adoption of alternative non-insecticidal vector control strategies should be considered a priority for the Cape Verde health authorities.

This article is protected by copyright. All rights reserved.

## Acknowledgements

This work was supported by the European Union's Horizon 2020 Research and Innovation Programme (ZIKAlliance 734548), Fundação para a Ciência e a Tecnologia (FCT)/ Ministério da Ciência e Tecnologia, Portugal, through research funds from the projects: POCI/BIA-BDE/57650/2004, PEst-OE/SAU/LA0018/2011 (Project CMDT: "Past and present population genetics of the vector *Aedes aegypti*: a tool for dengue control policies"), and GHTM-UID/Multi/04413/2013. P. Salgueiro has been funded by a FCT grant (SFRH/BPD/72532/2010), part of the EDCTP2 Program supported by the European Union.

## DATA ACCESSIBILITY

Data for this study are available at GenBank (MK359818-MK359844) and VectorBase (VBP0000346).

## LITERATURE CITED

- Alves, J., Gomes, B., Rodrigues, R., Silva, J., Arez, A. P., Pinto, J., & Sousa, C. A. (2010). Mosquito fauna on the Cape Verde Islands (West Africa): an update on species distribution and a new finding. *Journal of Vector Ecology*, 35.
- Amarasinghe, A., Kuritsky, J. N., Letson, G. W., & Margolis, H. S. (2011). Dengue Virus Infection in Africa. *Emerging Infectious Diseases*, 17, 1349–1354.
- Bandelt, H. J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, 16, 37–48.

Bennett, K. L., Shija, F., Linton, Y.-M., Misinzo, G., Kaddumukasa, M., Djouaka, R., ... Walton, C.

(2016). Historical environmental change in Africa drives divergence and admixture of *Aedes aegypti* mosquitoes: a precursor to successful worldwide colonization? *Molecular Ecology*, 25, 4337–4354.

Bhatt, S., Gething, P. W., Brady, O. J., Messina, J. P., Farlow, A. W., Moyes, C. L., ... Hay, S. I. (2013).

The global distribution and burden of dengue. *Nature*, 496, 504.

Bogoch, I. I., Brady, O. J., Kraemer, M. U. G., German, M., Creatore, M. I., Brent, S., ... Khan, K. (2016).

Potential for Zika virus introduction and transmission in resource-limited countries in Africa and the Asia-Pacific region: a modelling study. *The Lancet Infectious Diseases*, 16, 1237–1245.

Bosio, C. F., Harrington, L. C., Jones, J. W., Sithiprasasna, R., Norris, D. E., & Scott, T. W. (2005).

Genetic structure of *Aedes aegypti* populations in Thailand using mitochondrial DNA. *The American Journal of Tropical Medicine and Hygiene*, 72, 434–442.

Bourne, W. R. P. (1951). The birds of the Cape Verde Islands. *Ibis*, 97, 508–556.

Bracco, J. E., Capurro, M. L., Lourenço-de-Oliveira, R., & Sallum, M. A. M. (2007). Genetic variability

of *Aedes aegypti* in the Americas using a mitochondrial gene: evidence of multiple introductions. *Memórias Do Instituto Oswaldo Cruz*, 102, 573–580.

Brower, A. V. Z. (1994). Rapid morphological radiation and convergence in the butterfly, *Heliconius*

*erato*, inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Science*, 91, 6491–6495.

Brown, J. E., Evans, B. R., Zheng, W., Obas, V., Barrera-Martinez, L., Egizi, A., ... Powell, J. R. (2014).

Human Impacts Have Shaped Historical and Recent Evolution in *Aedes Aegypti*, the Dengue and Yellow Fever Mosquito. *Evolution*, 68, 514–525.

Brown, J. E., McBride, C. S., Johnson, P., Ritchie, S., Paupy, C., Bossin, H., ... Powell, J. R. (2011).

Worldwide patterns of genetic differentiation imply multiple “domestications” of *Aedes aegypti*, a major vector of human diseases. *Proceedings of the Royal Society of London B: Biological Sciences*, 278, 2446–2454.

Carranza, S., Arnold, E. N., Mateo, J. A., & López-Jurado, L. F. (2001). Parallel gigantism and complex colonization patterns in the Cape Verde scincid lizards *Mabuya* and *Macroscincus* (Reptilia: Scincidae) revealed by mitochondrial DNA sequences. *Proceedings of the Royal Society of London B: Biological Sciences*, 268, 1595–1603.

CDC. (2015, July 28). Unveiling the Burden of Dengue in Africa | Public Health Matters Blog | Blogs | CDC. Retrieved from <https://blogs.cdc.gov/publichealthmatters/2015/07/unveiling-the-burden-of-dengue-in-africa/>

Crawford, J. E., Alves, J. M., Palmer, W. J., Day, J. P., Sylla, M., Ramasamy, R., ... Jiggins, F. M. (2017). Population genomics reveals that an anthropophilic population of *Aedes aegypti* mosquitoes in West Africa recently gave rise to American and Asian populations of this major disease vector. *BMC Biology*, 15, 16.

da Costa-Ribeiro, M. C. V., Lourenco-De-Oliveira, R., & Failloux, A. B. (2007). Low gene flow of *Aedes aegypti* between dengue-endemic and dengue-free areas in southeastern and southern Brazil. *American Journal of Tropical Medicine and Hygiene*, 77, 303–309.

de Pina, A. (2013). Estudos e experiências de controlo vectorial e doenças transmitidas. In *Bioecologia e controlo vectorial de Culicidae em Cabo Verde* (pp. 13–17). Praia: Ministério de Ensino Superior, Ciências e Inovação de Cabo Verde.

Dia, I., Diagne, C. T., Ba, Y., Diallo, D., Konate, L., & Diallo, M. (2012). Insecticide susceptibility of *Aedes aegypti* populations from Senegal and Cape Verde Archipelago. *Parasites and Vectors*, 5, 238.

This article is protected by copyright. All rights reserved.

Diallo, M., Ba, Y., Faye, O., Soumare, M. L., Dia, I., & Sall, A. A. (2013). Vector competence of *Aedes aegypti* populations from Senegal for sylvatic and epidemic dengue 2 virus isolated in West Africa. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *102*, 493–498.

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size ( $N_e$ ) from genetic data. *Molecular Ecology Resources*, *14*, 209–2014.

Donnelly, M. J., Cuamba, N., Charlwood, J. D., Collins, F. H., & Townson, H. (1999). Population structure in the malaria vector, *Anopheles arabiensis* Patton, in East Africa. *Heredity*, *83*, 408–417.

Drummond, A. J., Rambaut, A., Shapiro, B., & Pybus, O. G. (2005). Bayesian Coalescent Inference of Past Population Dynamics from Molecular Sequences. *Molecular Biology and Evolution*, *22*, 1185–1192.

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, *29*, 1969–1973.

Du, Y., Nomura, Y., Zhorov, B. S., & Dong, K. (2016). Sodium Channel Mutations and Pyrethroid Resistance in *Aedes aegypti*. *Insects*, *7*, 60.

Duarte, M. C., & Romeiras, M. M. (2009). Cape Verde Islands. In *Encyclopedia of Islands* (1st ed., pp. 143–147). Berkeley.

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

Eskildsen, G. A., Rovira, J. R., Smith, O., Miller, M. J., Bennett, K. L., McMillan, W. O., & Loaiza, J. (2018). Maternal invasion history of *Aedes aegypti* and *Aedes albopictus* into the Isthmus of

Panama: Implications for the control of emergent viral disease agents. *PloS One*, *13*, e0194874.

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

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. *Molecular Ecology Resources*, *10*, 564–567.

Franco, L., & Di Caro A, C. F., Vapalahti O, Renaudat C, Zeller H, Tenorio A. (2010). Recent expansion of dengue virus serotype 3 in West Africa. *Euro Surveill*, *15*, 19490.

Fu, Y. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, *147*, 915–925.

Gloria-Soria, A., Ayala, D., Bheecarry, A., Calderon-Arguedas, O., Chadee, D. D., Chiappero, M., ... Powell, J. R. (2016). Global Genetic Diversity of *Aedes aegypti*. *Molecular Ecology*, *25*, 5377–5395.

Grobbelaar, A. A., Weyer, J., Moolla, N., Jansen van Vuren, P., Moises, F., & Paweska, J. T. (2016). Resurgence of Yellow Fever in Angola, 2015–2016. *Emerging Infectious Diseases*, *22*, 1854–1855.

Guedes, D. R. D., Gomes, E. T. B., Paiva, M. H. S., Melo-Santos, M. A. V. de, Alves, J., Gómez, L. F., & Ayres, C. F. J. (2017). Circulation of DENV2 and DENV4 in *Aedes aegypti* (Diptera: Culicidae) mosquitoes from Praia, Santiago Island, Cabo Verde. *Journal of Insect Science*, *17*.  
doi:10.1093/jisesa/iex057

Harpending, H. (1994). Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Hum Biol*, *66*, 591–600.

Harris, A. F., Rajatileka, S., & Ranson, H. (2010). Pyrethroid Resistance in *Aedes aegypti* from Grand Cayman. *American Journal of Tropical Medicine and Hygiene*, *83*, 277–284.

Hasegawa, M., Kishino, H., & Yano, T. A. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, *22*, 160–174.

Heatwole, H., & Shine, R. (1976). Mosquitoes feeding on ectothermic vertebrates: a review and new data. *Australian Zoologist*, *9*, 69–75.

Huang, Z., Das, A., Qiu, Y., & Tatem, A. J. (2012). Web-based GIS: the vector-borne disease airline importation risk (VBD-AIR) tool. *International Journal of Health Geographics*, *11*, 33.

Jaimes-Dueñez, J., Arboleda, S., Triana-Chávez, O., & Gómez-Palacio, A. (2015). Spatio-Temporal Distribution of *Aedes aegypti* (Diptera: Culicidae) Mitochondrial Lineages in Cities with Distinct Dengue Incidence Rates Suggests Complex Population Dynamics of the Dengue Vector in Colombia. *PLOS Neglected Tropical Diseases*, *9*, e0003553.

Kalinowski, S. T. (2004). hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, *5*, 187–189.

Kalinowski, S. T., Wagner, A. P., & Taper, M. L. (2006). ML-RELATE: a computer program for maximum likelihood estimation of relatedness and relationship. *Molecular Ecology Notes*, *6*, 576–579.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Haeseler, A. von, & Jermini, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*, 587.

Kamgang, B., Ngoagouni, C., Manirakiza, A., Nakouné, E., Paupy, C., & Kazanji, M. (2013). Temporal Patterns of Abundance of *Aedes aegypti* and *Aedes albopictus* (Diptera: Culicidae) and

Mitochondrial DNA Analysis of *Ae. albopictus* in the Central African Republic. *PLoS Neglected Tropical Diseases*, 7, e2590.

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, 15, 1179–1191.

Kotsakiozi, P., Evans, B. R., Gloria-Soria, A., Kamgang, B., Mayanja, M., Lutwama, J., ... Badolo, A. (2018). Population structure of a vector of human diseases: *Aedes aegypti* in its ancestral range, Africa. *Ecology and Evolution*. doi:10.1002/ece3.4278

Lobban, R. A. (2018). *Cape Verde: Crioulo Colony To Independent Nation*. Routledge. Retrieved from [https://www.goodreads.com/book/show/1532713.Cape\\_Verde](https://www.goodreads.com/book/show/1532713.Cape_Verde)

Lourenço, J., Monteiro, M. de L., Valdez, T., Rodrigues, J. M., Pybus, O., & Faria, N. R. (2018). Epidemiology of the Zika Virus Outbreak in the Cabo Verde Islands, West Africa. *PLOS Currents Outbreaks*. doi:10.1371/currents.outbreaks.19433b1e4d007451c691f138e1e67e8

Lovin, D., Washington, K., deBruyn, B., Hemme, R., Mori, A., Epstein, S., ... Severson, D. (2009). Genome-based polymorphic microsatellite development and validation in the mosquito *Aedes aegypti* and application to population genetics in Haiti. *BMC Genomics*, 10, 590.

Ministry of Health Cape Verde, WHO, & University of California. (2012). *Eliminating malaria: case study 2. Moving towards sustainable elimination in Cape Verde*. Geneva: WHO. Retrieved from <http://www.who.int/malaria/publications/atoz/9789241504386/en/>

Mitchell, C. (2016, September 1). PAHO WHO | Zika Cumulative Cases. Retrieved October 4, 2017, from [http://www.paho.org/hq/index.php?option=com\\_content&view=article&id=12390:zika-cumulative-cases&catid=8424:contents&Itemid=42090&lang=en](http://www.paho.org/hq/index.php?option=com_content&view=article&id=12390:zika-cumulative-cases&catid=8424:contents&Itemid=42090&lang=en)

Monteiro, M. L. (2016, August 12). Evolução da Epidemia de Zika em Cabo Verde, 2015 - 2016.

Retrieved from <http://www.minsaude.gov.cv/index.php/documentosite/381-evolucao-da-epidemia-de-zika-cabo-verde-2015-2016mariadelourdes/file>

Moore, M., Sylla, M., Goss, L., Burugu, M. W., Sang, R., Kamau, L. W., ... Black, W. C. (2013). Dual African Origins of Global *Aedes aegypti* s.l. Populations Revealed by Mitochondrial DNA. *PLoS Neglected Tropical Diseases*, *7*, e2175.

Morlais, I., & Severson, D. W. (2002). Complete mitochondrial DNA sequence and amino acid analysis of the cytochrome C oxidase subunit I (COI) from *Aedes aegypti*. *DNA Sequence: The Journal of DNA Sequencing and Mapping*, *13*, 123–127.

Moura, A. J. F. da, Santos, M. A. V. de M., Oliveira, C. M. F., Guedes, D. R. D., Carvalho-Leandro, D. de, Brito, M. L. da C., ... Ayres, C. F. J. (2015). Vector competence of the *Aedes aegypti* population from Santiago Island, Cape Verde, to different serotypes of dengue virus. *Parasites & Vectors*, *8*, 114.

Nutt, C., & Adams, P. (2017). Zika in Africa—the invisible epidemic? *The Lancet*, *389*, 1595–1596.

Paduan, K., & Ribolla, P. (2008). Mitochondrial DNA polymorphism and heteroplasmy in populations of *Aedes aegypti* in Brazil. *Journal of Medical Entomology*, *45*, 59–67.

Paetkau, D., Slade, R., Burden, M., & Estoup, A. (2004). Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology*, *13*, 55–65.

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. *Journal of Medical Entomology*, *45*, 391–400.

Paupy, C., Le Goff, G., Brengues, C., Guerra, M., Revollo, J., Barja Simon, Z., ... Fontenille, D. (2012).

Genetic structure and phylogeography of *Aedes aegypti*, the dengue and yellow-fever mosquito vector in Bolivia. *Infection, Genetics and Evolution*, *12*, 1260–1269.

Petersen, E. E., Staples, J. E., Meaney-Delman, D., Fischer, M., Ellington, S. R., Callaghan, W. M., &

Jamieson, D. J. (2016). Interim Guidelines for Pregnant Women During a Zika Virus Outbreak — United States, 2016. *MMWR. Morbidity and Mortality Weekly Report*, *65*, 30–33.

Piry, S., Alapetite, A., Cornuet, J., Paetkau, D., Baudouin, L., & Estoup, A. (2004). GENECLASS2: a

software for genetic assignment and first-generation migrant detection. *Journal of Heredity*, *95*, 536–539.

Piry, S., Luikart, G., & Cornuet, J. M. (1999). BOTTLENECK: a computer program for detecting recent

reductions in the effective population size using allele frequency data. *Journal of Heredity*, *90*, 502–503.

Powell, J. R., & Tabachnick, W. J. (2013). History of domestication and spread of *Aedes aegypti* - A

Review. *Memórias Do Instituto Oswaldo Cruz*, *108*, 11–17.

Pritchard, J., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus

genotype data. *Genetics*, *155*, 945–959.

Rambaut, A. (2016). FigTree (Version 1.4.3). Retrieved from

<http://tree.bio.ed.ac.uk/software/figtree/>

Rambaut, A., Drummond, A., Xie, D., Abebe, G., & Suchard, M. (2017). Tracer (Version 1.6). Retrieved

from <http://beast.community/tracer>

Rannala, B., & Mountain, J. L. (1997). Detecting immigration by using multilocus genotypes.

*Proceedings of the National Academy of Sciences*, *94*, 9197–9201.

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

Ribeiro, H., & Ramos, H. (1995). Guia ilustrado para a identificação dos mosquitos de Angola. *Boletim Da Sociedade Portuguesa de Entomologia, Suppl. 4*, 287.

Ribeiro, H., Ramos, H., Capela, R., & Pires, C. (Eds.). (1980). *Os mosquitos de Cabo Verde (Diptera: Culicidae): sistemática, distribuição, bioecologia e importância médica* (Vol. 135). Lisboa: Junta de Investigação Científica do Ultramar.

Rice, W. (1989). Analyzing Tables of Statistical Tests. *Evolution*, *43*, 223–225.

Rocha, H., Paiva, M., Silva, N. M., Araújo, A. P. de, Camacho, D., Moura, A., ... Santos, M. A. (2015). Susceptibility profile of *Aedes aegypti* from Santiago Island, Cabo Verde, to insecticides. *Acta Tropica*, *152*, 66–73.

Rogers, A. R., & Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology Evolution*, *9*, 552–569.

Rousset, F. (2008). Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, *8*, 103–106.

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sánchez-Gracia, A. (2017). DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Molecular Biology and Evolution*, *34*, 3299–3302.

Russell-Wood, A. J. R. (1998). *The Portuguese Empire, 1415-1808: A World on the Move*. JHU Press.

Saavedra-Rodriguez, K., Urdaneta-Marquez, L., Rajatileka, S., Moulton, M., Flores, A. E., Fernandez-Salas, I., ... Black 4th, W. C. (2007). A mutation in the voltage-gated sodium channel gene associated with pyrethroid resistance in Latin American *Aedes aegypti*. *Insect Molecular Biology*, *16*, 785–798.

This article is protected by copyright. All rights reserved.

Sang, R. C., & Dunster, L. M. (2001). The growing threat of arbovirus transmission and outbreaks in Kenya: a review. *East African Medical Journal*, *78*, 655–661.

Scarpassa, V. M., Cardoza, T. B., & Cardoso, R. P., Junior. (2008). Population Genetics and Phylogeography of *Aedes aegypti* (Diptera: Culicidae) from Brazil. *American Journal of Tropical Medicine and Hygiene*, *78*, 895–903.

Schneider, S., & Excoffier, L. (1999). Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, *152*, 1079–1089.

Seixas, G., Salgueiro, P., Silva, A. C., Campos, M., Spenassatto, C., Reyes-Lugo, M., ... Sousa, C. A. (2014). *Aedes aegypti* on Madeira Island (Portugal): genetic variation of a recently introduced dengue vector. *Memórias Do Instituto Oswaldo Cruz*, *108*, 3–10.

Slotman, M. A., Kelly, N., Harrington, L., Kitthawee, S., Jones, J. W., Scott, T. W., ... Powell, J. R. (2007). Polymorphic microsatellite markers for studies of *Aedes aegypti* (Diptera: Culicidae), the vector of dengue and yellow fever. *Molecular Ecology Notes*, *7*, 168–171.

Smith, L. B., Kasai, S., & Scott, J. G. (2016). Pyrethroid resistance in *Aedes aegypti* and *Aedes albopictus*: Important mosquito vectors of human diseases. *Pesticide Biochemistry and Physiology*, *133*, 1–12.

Stoler, J., & Awandare, G. A. (2016). Febrile illness diagnostics and the malaria-industrial complex: a socio-environmental perspective. *BMC Infectious Diseases*, *16*, 683. doi:10.1186/s12879-016-2025-x

Sylla, M., Bosio, C., Urdaneta-Marquez, L., Ndiaye, M., & Black, W. C. I. V. (2009). Gene flow, subspecies composition, and Dengue virus-2 susceptibility among *Aedes aegypti* collections in Senegal. *Plos Neglected Tropical Diseases*, *3*, e408.

- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*, 585–595.
- Tamura, K., & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology Evolution*, *10*, 512–526.
- Tandon, N., & Ray, S. (2000). Host feeding pattern of *Aedes aegypti* and *Aedes albopictus* in Kolkata India. *Dengue Bulletin*, *24*, 117–120.
- Tatem, A. J., Huang, Z., Das, A., Qi, Q., Roth, J., & Qiu, Y. (2012). Air travel and vector-borne disease movement. *Parasitology*, *139*, 1816–1830.
- Tavaré, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on Mathematics in the Life Sciences*, *17*, 57–86.
- Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A., & Minh, B. Q. (2016). W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Research*, *44*, W232–W235.
- Twerdochlib, A. L., Bonna, A. C. D., Leite, S. S., Chitolina, R. F., Westphal, B., & Navarro-Silva, M. A. (2012). Genetic variability of a population of *Aedes aegypti* from Paraná, Brazil, using the mitochondrial ND4 gene. *Revista Brasileira de Entomologia*, *56*, 249–256.
- Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., & Shipley, P. F. (2004). Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, *4*, 535–538.
- Vazeille, M., Yebakima, A., Lourenço-de-Oliveira, R., Andriamahefazafy, B., Correira, A., Rodrigues, J. M., ... Failloux, A. B. (2013). Oral Receptivity of *Aedes aegypti* from Cape Verde for Yellow Fever, Dengue, and Chikungunya Viruses. *Vector-Borne and Zoonotic Diseases*, *13*, 37–40.

Waples, R. S., & Do, C. (2010). Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, 3, 244–262.

Were, F. (2012). The dengue situation in Africa. *Paediatrics and International Child Health*, 32, 18–21.

WHO. (2009). *Dengue fever in Cape Verde - update 1*. WHO. Retrieved from [http://www.who.int/csr/don/2009\\_11\\_18/en/index.html](http://www.who.int/csr/don/2009_11_18/en/index.html)

WHO. (2012). Dengue and severe dengue. Retrieved from <http://www.who.int/mediacentre/factsheets/fs117/en/>

WHO. (2015). *Zika virus infection – Cape Verde*. Retrieved from <http://www.who.int/csr/don/21-december-2015-zika-cape-verde/en/>

WHO. (2016). WHO confirms Zika virus strain imported from the Americas to Cabo Verde. Retrieved from <http://www.who.int/mediacentre/news/releases/2016/zika-cabo-verde/en/>

WHO. (2017). WHO | Zika situation report. Retrieved October 4, 2017, from <http://www.who.int/emergencies/zika-virus/situation-report/10-march-2017/en/>

Wilder-Smith, A., & Byass, P. (2016). The elusive global burden of dengue. *The Lancet Infectious Diseases*, 16, 629–631.

Yu, H., Wang, W., Fang, S., Zhang, Y.-P., Lin, F.-J., & Geng, Z.-C. (1999). Phylogeny and Evolution of the *Drosophila nasuta* Subgroup Based on Mitochondrial ND4 and ND4L Gene Sequences. *Molecular Phylogenetics and Evolution*, 13, 556–565.

[dataset] Salgueiro P., Serrano C., Gomes B., Alves J., Sousa C., Ana Abecasis A., Pinto J.; 2019; *Aedes aegypti* MtDNA haplotypes; GenBank; MK359818-MK359844.

[dataset] Salgueiro P., Serrano C., Gomes B., Alves J., Sousa C., Ana Abecasis A., Pinto J.; 2019; *Aedes aegypti* microsatellite data; VectorBase; VBP0000346.

TABLES

Table 1

**Table 1.** Number of *Aedes aegypti* mosquitoes used for different analysis per year and collection site.

Year	date	Island	Site	DNA extracted	COI	ND4	<i>kdr</i>	microsatellites
2007	November-December	Santiago	Praia	53	22	21	47	47
		Brava	N <sup>a</sup> Sr <sup>a</sup> do Monte	6	6	6	-	-
		Fogo	S. Filipe	1	1	1	-	-
2010	April	Santiago	Barragem do Poilão	27	21	14	27	27
Total				87	50	42	74	74

Table 2

**Table 2** Mitochondrial DNA genetic diversity and neutrality tests results of four populations of *Aedes aegypti* from Cape Verde

Gene	Analyses	Statistics	2007			2010		Mean	Global
			Santiago	Brava	Fogo	Santiago			
ND4	Genetic diversity	Hap	4	2	1	3	3	6	
		Hd	0.471	0.600	NA	0.703	0.592	0.609	
		Pi	0.001	0.002	NA	0.003	0.002	0.002	
								s.d.	
	Sample size	21	6	1	14	10.5	8.813		
	Neutrality tests	Tajima's D	-1.299	1.445	0	1.459	0.401	1.324	
D p-value		NS	NS	NA	NS	NS	NS		
Fu's FS		-2.387	0.795	0	0.783	0.202	1.503		
FS p-value		0.005	NS	N.A	NS	NA	NA		
							Global		
COI	Genetic diversity	Hap	15	3	1	7	6.5	20	
		Hd	0.949	0.733	NA	0.832	0.838	0.944	
		Pi	0.004	0.001	NA	0.003	0.003	0.003	
								s.d	
	Sample size	23	6	1	20	12.5	10.661		
	Neutrality tests	Tajima's D	-1.195	-0.050	0	0.413	-0.208	0.690	
D p-value		NS	NS	NS	NS	NS	NS		
Fu's FS		-9.777	-0.427	0	-0.496	-2.675	4.740		
FS p-value		0.000	NS	NA	NS	NA	NA		

Number of haplotypes (Hap); Haplotype diversity (Hd); Nucleotide diversity (Pi); Standard neutrality tests: Tajima's D (Tajima, 1989), Fu's Fs (Fu, 1997); standard deviation (s.d.) calculated for the haplotypes identified in Cape Verde; P values  $\geq 0.05$  were considered non-significant (NS); Not applicable (NA).

Table 3

**Table 3** Estimates of effective population size and heterozygosity tests for *Aedes aegypti* from Cape Verde

Populations	Ne	CI	Bottleneck	
			P(SMM)	P(TPM)
CV2007	54.8	39.0-84.5	NS	NS
CV2010	7.7	5.5-10.6	NS	0.0017

Ne: current effective population size based on the bias-corrected LD method (Waples and Do 2010); CI: parametric 95% confidence interval; P(SMM) and P(TPM): P-value for heterozygote excess under SMM and TPM, after Bonferroni correction; NS: Non-significant

#### FIGURE LEGENDS

**Fig. 1:**

Map with the location of the Cape Verde archipelago, naming the nine inhabited islands. The *A. aegypti* samples from the present study have been collected from the islands in bold (Santiago, Brava and Fogo).

**Fig. 2:**

Bayesian phylogenetic tree (see text for details) of relationships among 49 haplotypes of *A. aegypti* based on a 360bp ND4 fragment and using *A. albopictus* as outgroup (in the Figure it is written “output” by mistake). The support values written on the branches correspond to posterior probabilities (values < 0.900 are not presented). The haplotype codes and their origins are listed in

Appendix 1c. The branch lines are colored by geographic distribution of the corresponding haplotypes. Time scale: substitutions/Myr

**Fig. 3:**

Bayesian phylogenetic tree (see text for details) of relationships among 94 sequences of *A. aegypti* based on a 493bp COI fragment and using *A. albopictus* as outgroup. The support values written on the branches correspond to posterior probabilities (values < 0.800 are not presented). The haplotype codes and their origins are listed in Appendix 3. The branch lines are colored by geographic distribution of the corresponding haplotypes. Time scale: substitutions/Myr

*A. albopictus*: 155albopictus\_JQ004524; 1: 100Mdgr\_HQ688298; 2: 190TZ\_KX446465; 3: 186TZ\_KX446459; 4: 169U4\_KX446426; 5: 183TZ\_KX446456; 6: 027Mex\_JQ926698; 7: 078Thail\_JQ926692; 8: 008BR\_JQ926703; 9: 025Mart\_JQ926697; 10: 026Mart\_JQ926697; 11: 003Bol\_JQ926678; 12: 002Bol\_JQ926677; 13: 023Guyana\_HQ688297; 14: 001Bol\_JQ926676; 15: 102TZA\_JQ926704; 16: 189TZ\_KX446464; 17: 187TZ\_KX446460; 18: 180U10\_KX446447; 19: 086Vietn\_JQ926686; 20: 087Vietn\_JQ926687; 21: 172U5\_KX446430; 22: 192TZ\_KX446469; 23: 089Eur\_HQ688296; 24: 173U5\_KX446431; 25: 191TZ\_KX446467; 26: 175U9\_KX446436; 27: 185TZ\_KX446458; 28: 168U4\_KX446425; 29: 181U11\_KX446450; 30: 184TZ\_KX446457; 31: 188TZ\_KX446463; 32: 182U12\_KX446451; 33: 178U10\_KX446442; 34: 171U4\_KX446429; 35: 165U2\_KX446420; 36: 166U4\_KX446423; 37: 174U9\_KX446434; 38: 179U10\_KX446444; 39: 177U9\_KX446438; 40: 170U4\_KX446427; 41: 176U9\_KX446437; 42: 164U1\_KX446418; 43: 167U4\_KX446424; 44: 154Formosus\_AY056597; 45: 011Col\_KM203198; 46: 012Col\_KM203248; 47: 028Mex\_JQ926699; 48: 010Col\_KM203140; 49: 052FP\_HQ688295; 50: 111Madeira\_KF909122; 51: 036USA\_JQ926684; 52: 085Vietn\_JQ926685; 53: 077Thail\_JQ926691; 54: 040Ven\_JQ926701; 55: 048Camb\_JQ926690; 56: 084Vietn\_HQ688292; 57: 046Camb\_JQ926688; 58: 047Camb\_JQ926689; 59: 045Camb\_HQ688294; 60: 145RCI\_JQ926694; 61: 163Ben1\_KX446415; 62: 118Cam\_JQ926702; 63: 144RCI\_JQ926693; 64: 161Ben1\_KX446401; 65: 142Guin\_JQ926700; 66: 157Ben2\_KX446393; 67: 141Guin\_HQ688293; 68: 160Ben1\_KX446397; 69: 162Ben1\_KX446413; 70: 156Ben2\_KX446392; 71: 159Ben1\_KX446396; 72: 146RCI\_JQ926695; 73: 121CV\_COI03; 74: 158Ben1\_KX446395; 75: 135CV\_COI17; 76: 130CV\_COI12; 77: 120CV\_COI02; 78: 124CV\_COI06; 79: 134CV\_COI16; 80: 136CV\_COI18; 81: 122CV\_COI04; 82: 129CV\_COI11; 83: 123CV\_COI05; 84: 132CV\_COI14; 85: 131CV\_COI13; 86: 138CV\_COI20; 87: 119CV\_COI01; 88: 125CV\_COI07; 89: 128CV\_COI10; 90: 127CV\_COI09; 91: 133CV\_COI15; 92: 126CV\_COI08; 93: 137CV\_COI19

**Fig. 4:**

Global genetic structure of *Aedes aegypti*. STRUCTURE bar plots indicating genetic groupings of different geographic locations based on nine microsatellite loci. Each vertical bar represents an individual. The height of each bar represents the probability of assignment (Q) to each of K = N clusters as determined using the Delta K method. Each cluster is indicated by different colors.

(A) Genetic groupings of 81 geographic locations and 3703 individuals, sorted by countries, best K = 2 clusters, *A. a. aegypti*: red and *A. a. formosus*: green. Population code numbers are: 1-Santiago 2010, CV; 2-Santiago 2007, CV; 3-Dakar, SN; 4-N'goye, SN; 5-Sedhiou, SN; 6-Koungheul, SN; 7-Ngari, SN; 8-Goudiry, SN; 9-PK-10, SN; 10-Bijagos, GW; 11-Yaounde, CM; 12-Francesville, GA; 13-Johannesburg, ZA; 14-Kichwamba, UG; 15-Bundibugyo, UG; 16-Lunyo, UG; 17-Kakamenga, KE; 18-Kisimu, KE; 19-Nairobi, KE; 20-Rabai-out, KE; 21-Rabai-in, KE; 22-Garissa, KE; 23-Mombasa, KE; 24-Cordoba, AR; 25-Salta, AR; 26-Iguazu, AR; 27-La Plata, AR; 28-Posadas, AR; 29-Cachoeiro, BR; 30-Maraba, BR; 31-Natal, BR; 32-Jacobina, BR; 33-Rio de Janeiro, BR; 34-Cali, CO; 35-Bolivar, VE; 36-Zulia, VE; 37-Siquirres, CR; 38-Dominica, DM; 39-Carriacou, GD; 40-Trinidad, TT; 41-Patillas, PR; 42-Tijuana, MX; 43-Pijijiapan, MX; 44-Las Palomas, MX; 45-Lomas de Zapatero, MX; 46-Iguala, MX; 47-Mazatan, MX; 48-Tapachula Norte, MX; 49-Hermosillo, MX; 50-Nogales, MX; 51-Chetumal, MX; 52-Amacuzac, MX; 53-Maricopa, US; 54-Madera, US; 55-Clovis, US; 56-New Orleans, US; 57-Houston, US; 58-Vaca Key, US; 59-Miami, US; 60-Musco, US; 61-San Mateo, US; 62-Rio FL, US; 63-Exeter, US; 64-Los Angeles, US; 65-Conch Key, US; 66-Palm Beach, US; 67-North Key West, US; 68-Madeira, PT; 69-Riyadh, SA; 70-Pakistan, PK; 71-Sri Lanka, LK; 72-Rayong, TH; 73-Prachuabkhirikan, TH; 74-Bangkok, TH; 75-Ho Chi Minh, VN; 76-Hanoi, VN; 77-Cebu, PH; 78-Townsville, AU; 79-Cairns, AU; 80-Tahiti, FP; 81-Hawaii, US. All details about the non-Cape Verdean samples are described elsewhere (Gloria-Soria et al., 2016). (B) Genetic groupings of 23 African populations and 989 individuals from A), sorted by countries and by longitude (W: west to E: east), best K = 2 clusters (blue and orange). (C) The same 989 individuals from 23 African populations as in

B), sorted by Q. (D) Genetic groupings of 13 African populations and 258 individuals selected from individuals with the  $Q > 0.90$  of assignment to the blue cluster in A), best  $K=2$ .





