Population structure and ancestry prediction of *Aedes aegypti* (Diptera: Culicidae) supports a single African origin of Colombian populations

Yoman Monsalve[1], Omar Triana-Chávez[1], Andrés Gómez-Palacio[2]/+

[1]Universidade de Antioquia, Grupo de Biología y Control de Enfermedades Infecciosas, Medellin, Colombia
[2]Universidad Pedagógica y Tecnológica de Colombia, Laboratorio de Investigación en Genética Evolutiva, Boyacá, Colombia

BACKGROUND A previous phylogeographic study revealed two *Aedes aegypti* African-related mitochondrial lineages distributed in Colombian's cities with different eco-epidemiologic characteristics with regard to dengue virus (DENV). It has been proposed these lineages might indicate independent invasion sources.

OBJECTIVES Assessing to Colombian population structure and to support evidence of its probable source origin.

METHODS We analysed a total of 267 individuals from cities of Bello, Riohacha and Villavicencio, which 241 were related to the West and East African mitochondrial lineages (termed here as WAL and EAL, respectively). Eight polymorphic microsatellite loci were analysed aiming population structure.

FINDINGS Results indicate substantial gene flow among distant and low-connected cities composing a panmictic population with incipient local differentiation of *Ae. aegypti* is placed in Colombia. Likewise, genetic evidence indicates no significant differences among individuals related to WAL and EAL is placed.

MAIN CONCLUSIONS Minimal genetic differentiation in low-connected *Ae. aegypti* populations of Colombia, and lack concordance between mitochondrial and nuclear genealogies suggest that Colombian *Ae. aegypti* shared a common demographic history. Under this scenario, we suggest current *Ae.* aegypti population structure reflects a single origin instead of contemporary migration, which founding populations have a single source from a mitochondrial polymorphic African ancient.

Key words: Colombia - genetics - *Aedes aegypti* - population structure - microsatellites

Dengue is the arthropod-borne viral disease with the fastest propagation worldwide, having a dramatic increase around 30 fold its global incidence over the past 50 years, estimated in 390 million cases by year.[1] The primary vector of dengue virus is *Aedes aegypti*, which also transmits other viruses such as yellow fever virus, chikungunya and Zika. Similar to other vector-borne diseases, arboviruses infection must be reduced mostly by controlling vector populations.[2] In this sense, studies about population genetic structure as well as gene flow between populations of *Ae. aegypti* are relevant to perform vector control and entomologic surveillance against these infections.[3] For instance, if several regions in a country share a panmictic population of the same genetic background and similar/different phenotype, uniform approach for mosquito control is warranted. Contrary, if the regions happened to have different mosquito populations that are differ with respect to their genetic background and phenotypes, the uniform approaches should be tested before it would work effectively in these regions.

Aedes aegypti population structure has been well studied around the world, and its genetic heterogeneity and historical migration routes have been unveiled at multiple several geographic scales including global scale,[4] regional such as in the Black-Sea region, and West Africa, and at local scale such as in California, Florida, Brazil, and Peru[5,6,7,8,9,10] among others. Some recent examples of *Ae. aegypti* microgeographic population structure in the Americas has disclosed diverse population dynamics (i.e., migration and diversity) is placed in eclectic ecogeographical regions such as boat traffic driving migration in Peruvian Amazon,[5] the substantial gene flow between geographically distant cities in Florida connected by an Interstate Highway,[7,8] and the recently multiple-source invasion of diverse populations into California,[9,10] and in several regions of Brazil.[11,12,13]

In Colombia, a first study based on random amplified polymorphic DNA-polymerase chain reaction products analysis (RAPD) reported local genetic population substructure in city of Cali, which was suggested involved in differences of vector competence and insecticide resistance ratios.[14] Likewise, genetic analyses based on partial nucleotide sequences of mitochondrial NADH dehydrogenase 4 (ND4) gene, disclosed local microgeographic differences among individuals of municipalities

of Sincelejo, Guaranda, Corozal and Sampués (within a maximum range of ~100 km) from Sucre department,[15,16] as well as from Medellin city.[17] These studies would suggest a complex heterogeneous population substructure is placed throughout Colombian populations mainly as consequence of local differences on selective pressures from the chemical control.

On the other hand, a more comprehensive work of *Ae. aegypti* Colombian populations reported the presence of the two mitochondrial lineages related to East and West African ancient populations, circulating in distinct frequencies across cities of Bello (BE), Riohacha (RI) and Villavicencio (VI), which having different eco-epidemiologic characteristics (i.e., dengue incidence, ecogeographical location and vector control strategies). Cities of BE, RI and VI are roughly equidistant each other (~900 km by road), and they are separated by orographic barriers such as Central and East Colombia cordilleras, and not often huge human displacement is placed among them.[18] The suggested microgeographic structure of Colombian populations and the distinct frequency of the mitochondrial lineages observed in these cities was thought to be a consequence of possible dual or multiple invasion sources.[18] Previous research hypothesised independent colonisation routes for *Ae. aegypti* mitochondrial lineages from different Africa sources to the New World and Latin American countries.[19,20,21,22,23] However, most recent literature about *Ae. aegypti* populations demography, using nuclear genetic information such as polymorphic microsatellite loci and genome-wide single nucleotide polymorphisms (SNPs) approach, unveiled lack genetic concordance between differentiation patterns observed at nuclear genetic level and East, and West African mitochondrial related populations.[4,7,10,24,25] In this sense, considering that SNPs and polymorphic microsatellite loci markers have higher genetic resolution (i.e., faster mutation rate) than mitochondrial markers, the observed mitochondrial lineage polymorphisms picture is thought occurred previous to outside Africa dispersal, and New World populations were originated by a single source from a mitochondrial polymorphic African ancient around the 16th-century, as consequence of slave trade between Africa and Americas.[4,24]

Due no comprehensive evidence for population structure of Colombian *Ae. aegypti* in this study we analysed genetic differentiation of Colombian populations, as well as addressed for whether mitochondrial West Africa lineage (WAL) and East Africa lineage (EAL) represent structured populations. Our results indicate substantial gene flow between geographically distant cities composing a panmictic population with substantial recent gene-flow of *Ae. aegypti* in Colombia, and support lack concordance between mitochondrial and nuclear genetic information. Under this scenario we propose that Colombian *Ae. aegypti* compose a unique population with incipient local genetic differentiation originated from a single source from a mitochondrial polymorphic African ancient.

MATERIALS AND METHODS

Sample collection and sites of the study - *Ae. aegypti* adult mosquitoes were collected in cities of BE, RI and VI during 2012 and 2013 using entomological networks as described elsewhere.[18] Three samplings were conducted in two neighborhoods separated by more than 1 km for each city. In order to avoid inbreeding bias in the sample less than three specimens per house were included in the genetic analyses. Entomological gathering made on private lands or in private residences were made with dweller permission and presence. A total of 267 mosquitoes, harboring 199 individuals belonging WAL and 42 to EAL were included in microsatellite analyses (Table I). Mitochondrial lineages of individuals were determined previously using partial nucleotide sequences of cytochrome oxidase I (COI) and ND4 genes and defined according to its phylogeographic clustering.[18] The same specimens were used for mitochondrial DNA and genetic analysis performed in this work.

TABLE I

Geographic origin and number of *Aedes aegypti* mosquitoes analysed in this study by microsatellites, and the corresponding number of individuals related to West African lineage (WAL) and East African lineage (EAL) as previously reported[18]

Department	City (Coordinates)	Neighborhood	*n*	Individuals related to mitochondrial lineages analysed by microsatellites	
				WAL	EAL
Antioquia	Bello (6° 20' N 75° 35' W)	Cumbre	41	20	19
		Granjas	44	17	20
La Guajira	Riohacha (11° 31' N 72° 55' W)	Aeropuerto	52	45	1
		Unión	47	44	2
Meta	Villavicencio (4° 04' N 73° 40' W)	Popular	38	34	0
		Porfía	45	39	0
Total			267	199	42

n: number of *Aedes aegypti* mosquitoes analysed.

DNA extraction and microsatellite genotyping - Genomic DNA was extracted from each adult mosquito following an insect DNA extraction protocol.[26] Microsatellite loci 1132CT1, 176TG1, 145TAAA1, 462GA1, 109CT1, 88AT1,[27] AG5 and AC5[28] were amplified in four multiplex polymerase chain reaction (PCR) (termed M1 - M4) according to its product size and using different dye set in forward primers [**Supplementary data** (Table I)]. PCR reactions for multiplex M1, M2 and M3 were conducted in a final volume of 25 µL using 3 µL of DNA template (~10 ng/µL), 1X PCR buffer (0.1 M Tris-HCl, 0.5 M KCl, and 0.015 M $MgCl_2$, pH 8.3), 1.5 mM $MgCl_2$, 0.2 mM dNTP, 0.2 µM of each primer, and 1 U of *Taq* DNA polymerase (Thermo Scientific, Waltham, MA, USA). For M4 multiplex, PCR reaction was conducted in a final volume of 20 µL using 1.5 µL of DNA template (~10 ng/µL), 1X PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTP, 0.025, and 0.5 µM of forward and reverse primers respectively, and 0.75 U of *Taq* DNA polymerase (Thermo Scientific, Waltham, MA, USA). The fragments were amplified following thermal cycling conditions previously reported.[28] Amplified fragments were visualised on 2% gel agarose stained with ethidium bromide and purified products were sent for DNA fragment detection using 400HD as internal standard size marker at Macrogen Inc., Korea. Fragment size determination with one bp resolution and allele classification were performed with Geneious v.8 software.[29]

Genetic diversity of Ae. aegypti populations - We used GenoDive v.3.0[30] to test deviations from Hardy-Weinberg equilibrium (HWE) genotypic expectations and FSTAT v.2.9.4 [31] to test significant linkage disequilibrium (LD) between locus pairs. Local false discovery rate test (fdr) was performed for correcting statistical significance for multiple comparisons, fdr was calculated using 1,000 iterations in fdrtool package for R.[32] We screened presence of null alleles and genotyping errors using Micro-Checker v.2.2.3,[33] and statistical significance was assessed with 1,000 iterations.

We used GenoDive v.3.0[30] for estimating gene diversity index such as average of number of alleles (N_a), effective number of alleles (N_e), observed (H_o), and expected (H_e) heterozygosity, and inbreeding coefficient (G_{is}). Statistical significance was assessed with 1,000 permutations as implemented in GenoDive v.3.0.[30]

Genetic structure of Ae. aegypti populations - We test population genetic subdivision using the F_{ST} analogue G_{ST},[34] which is calculated by relating the genetic diversity within populations to the overall genetic diversity using GenoDive v.3.0.[30] Significance was calculated by bootstrapping with 1,000 permutations. In order to check the effect for possible null alleles in the genetic population subdivision calculated, estimates for null allele frequencies and unbiased F_{ST} calculation was restricted to visible allele sizes (so-called as ENA method) in FreeNA software.[35] Significance was tested by *t* test using 1,000 iterations. A Mantel test was also performed in GenAlEx v.6.5 on geographic (at neighborhood) and genetic distance (pairwise *phi*PT) using 1,000 permutations.[36]

Genetic structure of the complete data set without an *a priori* assignation to a geographical origin was evaluated using the Bayesian assignment approach as implemented in STRUCTURE v.2.3.3.[37,38] Since we assume each individual has ancestry from one or more of K genetically distinct sources, we employed the Admixture ancestry model with correlated allele frequency to find the number of genetic clusters (K). We evaluated the genetic clusters creating a batch run with K = 2 (considering WAL and EAL as possible different genetic clusters), K = 3 (assuming three cities as different genetic clusters), and K = 4, 5, and 6 (assuming between and within cities (i.e., neighborhoods) as different genetic clusters). Ten replicates per each K with 500,000 iterations per run and a burn-in of 50,000 runs were performed, and a graphical representation of the results was obtained using Distruct v.1.1.[39] We determined the optimum K number using the ΔK criterion[40] as implemented in Structure Harvester v.0.6.94.[41] Individuals were considered assigned to a specific cluster if the proportion of ancestry was ≥ 80% otherwise the individual was considered unassigned.

Relationship inference and demographic analysis of Ae. aegypti populations - We estimate genetic relationship between individuals of Colombian populations as well as the relationship grade within Bello individuals belonging mitochondrial lineages, using the kinship (*ks*) coefficient[42,43] in GenoDive v.3.0.[30] Estimated *ks* values were categorised according to Iacchei et al.[44] as follows: unrelated individuals ($ks < 0.0475$); quarter-siblings ($0.0475 < ks < 0.09375$); half-siblings ($0.09375 < ks < 0.1875$); full-siblings ($0.1875 < ks < 0.375$); and nearly identical ($ks > 0.375$). We test normality of data by Anderson-Darling test, and statistical significance ($p < 0.05$) of *ks* values among geographical demes (i.e., among and within cities) as well as between lineages using Kruskal-Wallis test in *dplyr* package in R.

We used a Bayesian coalescence-based algorithm in order to address the directionality of gene flow among cities according to migration rates using Migrate-n v.4.4.4.[45] We tested five migration models assuming similar population sizes (i.e., estimated Theta) in the three populations [**Supplementary data** (Fig. 1)]: (i) full model bi-directional gene flow, (ii) one-directional gene flow out of BE to RI and VI, and bi-directional gene flow between them, (iii) one-directional gene flow from RI to VI and BE, and bi-directional gene flow between them (iv) one-directional gene flow from VI to BE and RI, and bi-directional gene flow between them (v) the null model of panmixia (i.e., the samples composes a single population with constant -not-estimated-migration rate). We used standard microsatellite parameters, with uniform prior distribution set for both Theta (min: 0 to max: 1000, and delta of 10) and M (min: 0 to max: 100, and delta of 10) similar for all populations, search strategy was performed in three independent runs and parameters were: 10,000 recorded steps [a], 10 sample increment [b], and 1 concurrent chains (replicates) [c] that summarises 1,000,000 of visited parameters [a*b*c] with a burn-in of 1,000 runs using a static

heating scheme of four chains with default temperature parameters. We selected the suitability of parameters used by checking the convergence of posterior probabilities distribution over all loci and effective sample size (ESS) > 200. We selected the best model test according to the log Bayes factor (LBF) based on the accurate marginal likelihoods of the Bezier approximation score generated for the five models.[46]

RESULTS

Microsatellite description of Ae. aegypti of Colombia - All eight loci showed less than 15% of missing data and any out 267 individuals showed above 50% of missing data, so all individuals were included in further analyses (an overall of 89 per city; Table II). After local fdr correction p-values obtained in multiple comparisons no significant HWE nor LD deviations was observed in any population-by-locus comparison [p > 0.05; **Supplementary data** (Table II)]. No genotyping errors such as allele dropouts or stuttering were detected for any locus, but presence for possible null alleles having frequencies ranking between 2% to 18% were estimated in some loci across cities [**Supplementary data** (Table II)].

All loci were polymorphic showing a mean number of alleles (N_a) per locus of 14.5, a mean effective number of alleles per locus (N_e) of 3.7, mean observed heterozygosity (H_o) of 0.53, and expected heterozygosity (H_e) of 0.67, and overall global inbreeding index (*Gis*) of 0.21 (Table II). Most of loci showed highly variable positive G_{IS} values indicating heterozygote deficit (Table II).

Genetic structure analysis of Ae. aegypti in Colombia unveils low geographical differentiation - G_{ST} statistics indicate significant, but low genetic differentiation among the three populations based on estimated overall G_{ST} (G_{ST} = 0.031; p < 0.001). Pairwise comparisons among cities indicated lower genetic differentiation between BE and VI (G_{ST} = 0.028; p < 0.001), followed for a moderate between BE and RI (G_{ST} = 0.052; p < 0.001), and RI and VI (G_{ST} = 0.056; p < 0.001). The unbiased F_{ST} per locus based on ENA correction including null alleles was not significant different from the F_{ST} estimated (*t* = 0.025; p > 0.05), indicating possible null alleles present in sample do not affect the result observed [**Supplementary data** (Table III)]. A Mantel test on the whole dataset showed significant but weak correlation (R^2 = 0.02, p < 0.001) between genetic (*phi*PT) and geographic distance (y = 11.54x + 236.85).

Genetic structure based on Bayesian clustering indicated that the optimal clustering number is K = 4 according to the highest ΔK value obtained [**Supplementary data** (Fig. 2)]. However, after assignment test according to a proportion of ancestry ≥ 80%, the highest number of individuals were grouped into main clusters (Fig. 1). Out the total individuals 27.3% were not assigned to any of the four genetic clusters. Cluster 1 was composed by BE (60%), and VI (30%); cluster 2 was mostly composed by RI (57%); cluster 3 by VI (27%); and cluster 4 showed similar percentages from individuals from the three cities [Fig. 1, **Supplementary data** (Fig. 3)]. Otherwise, individuals did not cluster based on whether they are related to the WAL or EAL, even assuming 50% of genetic mixture per individual [**Supplementary data** (Table IV)].

TABLE II
Summary of variability parameters by population and by locus

Cities	*n*	N_a	N_e	H_o	H_e	G_{IS}
Bello	85	9.01	3.78	0.49	0.65	0.24
Riohacha	99	11.38	3.96	0.56	0.69	0.20
Villavicencio	83	9.50	3.74	0.53	0.67	0.20
Overall	89	14.5 ± 2.2	3.67 ± 0.69	0.53 ± 0.07	0.67 ± 0.08	0.21 ± 0.04
Locus		N_a	N_e	H_o	H_e	G_{IS}
1132CT1		22	5.74	0.67	0.83	0.19
462GA1		20	3.23	0.42	0.69	0.39
176TG1		18	3.26	0.69	0.69	0.02
145TAAA1		8	3.09	0.49	0.68	0.29
19CT1		4	1.18	0.07	0.16	0.52
88AT1		19	5.17	0.67	0.81	0.18
AG5		12	3.98	0.62	0.76	0.18
AC5		13	3.65	0.61	0.73	0.17
Overall		14.5	3.67	0.53	0.67	0.21

n: number of individuals per population; N_a: number alleles; N_e: effective number of alleles; H_o: observed heterozygosity; H_e: expected heterozygosity; G_{IS}: inbreeding coefficient.
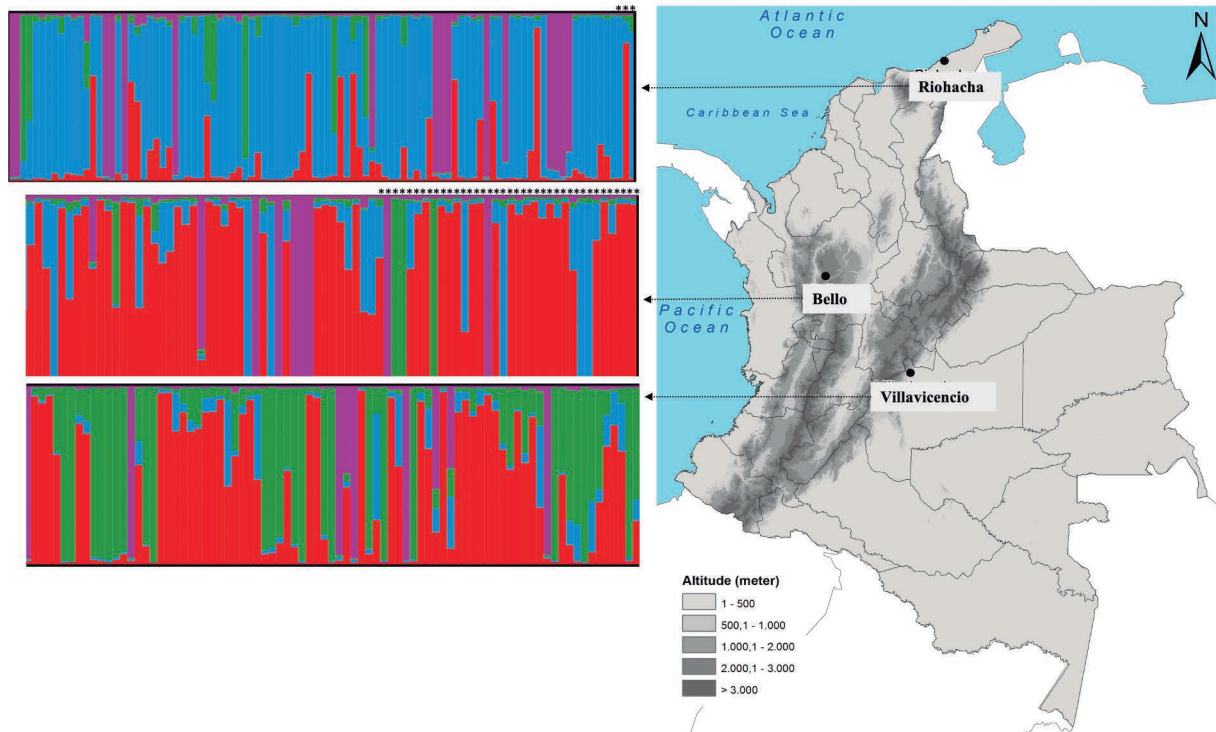
Fig. 1: map of Colombia and genetic ancestry membership of *Aedes aegypti* to genetic clusters (K = 4) found in cities of Bello, Riohacha and Villavicencio. Each bar represents an individual and the colour is the respective ancestry component to each cluster. Asterisk (*) denotes East African lineage (EAL) related individuals as previously reported,[18] see further details in Table I.

Demography of Ae. aegypti indicates high gene-flow among populations and no genetic differences between individuals related to WAL and EAL in Colombia - Kinship coefficient (*ks*) for pair-wised individuals fitting to a normal distribution [Anderson-Darling normality test; A = 0.3855, p = 0.386; **supplementary data** (Fig. 4A)]. Median value for the entire dataset was *ks* = -0.01529, indicating most of individuals are unrelated (*ks* < 0.09375), as expected in a wide roughly panmictic population. Overall kinship coefficient for cities was not significant (Kruskal-Wallis X^2 = 3545.63, df = 3541, p = 0.4749; Fig. 2A), and *ks* values within-cities were comparatively higher than those calculated for individuals among-cities (Fig. 2B). No significant differences (Kruskal-Wallis X^2 = 1.1273, df = 1, p = 0.2883) in Kinship coefficient between individuals related to mitochondrial lineages were observed [**Supplementary data** (Fig. 4B-C)].

Moreover, according to the lowest LBF estimated from marginal likelihood of Bezier score (marginal likelihood = -78119.58), the best probability model (> 99.9%) test was for constant migration rate model (v), followed by similar values observed in those one-directional gene-flow based models (model (ii), marginal likelihood. = -107228.81; (iii), marginal likelihood = -108560.43, and (iv), marginal likelihood = -140668.89), and finally for the three cities for-bidirectional model (i) (marginal likelihood = -152279.34). According this result any number of migrants per generation can be accurate estimated from populations, which indicates strong evidence for possible past or recent gene-flow among the three cities.

DISCUSSION

Here we disclose genetic population structure of Colombian *Ae. aegypti* populations from three endemic cities for relevant arboviruses such as dengue, Zika and chikungunya. Our results indicate *Ae. aegypti* Colombian population composes a panmictic population with substantial recent gene-flow among distant cities with minimal evidence of isolation-by-distance process. Furthermore, result derived from Bayesian clustering analysis where four moderated mixed genetic cluster were found as well as low values of G_{ST} index between populations, and relatively low heterozygosity observed throughout in all loci analysed would indicate high inbreeding or possible Wahlund effect is placed within populations. Likewise, high within-cities Kinship values, and null model chosen as the best scenario, indicates local demographic dynamics could driving incipient genetic differentiation processes as observed in some individuals from the three cities.

The three cities analysed BE, RI and VI are distant, separated by northern Andean cordilleras, and low human displacement is placed among them. Because, the main source for *Ae. aegypti* dispersal between distant localities is passive migration associated to human movement, results raised here suggest at least two possible not mutually exclusive hypotheses can be drawn: (i) a sustained-continuous migration between intermediate populations exists without isolation by distance evidence (i.e. patchy metapopulation structure), (ii) the population structure is retained from an ancient popu-
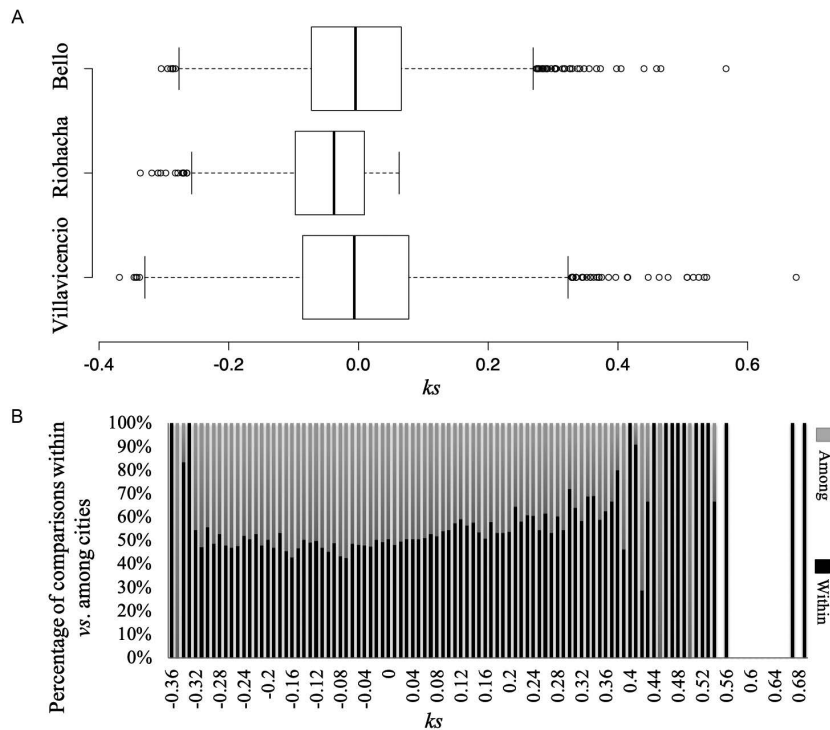
Fig. 2: Kinship coefficient (ks) of *Aedes aegypti* individuals from cities of Bello, Riohacha and Villavicencio. (a) Boxplot of ks values estimated between individuals within cities. (b) Percentage of ks values estimated between individuals among and within cities.

lation invasion occurred in the recent past, it means a large founding population that rapidly spread throughout Colombia. Thus, based on our results, we conclude that microsatellite data present here are probably mirroring the ancient event of single origin instead of contemporary migration among populations. However, although this is a plausible hypothesis, further studies involving broader populations and using more resolutive markers still need to be tested.

Microgeographic structure of Colombian populations has been slightly traced so far, but no clear picture of population demographic has been supported. Several studies at local level using dominant genetic markers such as RAPDs,[14] and nucleotide sequence analyses of mitochondrial genes, such as COI[17] and NADH dehydrogenase subunit 4 (ND4)[15,16] unveiled Colombian *Ae. aegypti* populations show moderated population differences at spatiotemporal scales, and it was suggested as a consequence of local pressures.

A previous research reported heterogenous distribution of mtDNA haplotypes in those cities, which have distinct dengue incidences and vector control strategies, for instance in BE city not often chemical control is applied but community-based and biological control programs are major used strategies, unlike RI and VI where insecticide-based control is often performed.[18] In that work the authors did not support conclusive evidence for geographic or temporal genetic structuration between populations but rather the phylogeographic analysis indicated two main mitochondrial lineages related with West and East populations of Africa are differentially

distributed in these Colombian cities, where in fact, both were collected in same houses in BE city.[18] Here we support additional evidence for lack genetic concordance between local differentiation patterns observed at nuclear genetic level and West and East African mitochondrial related populations.[4,7,10,24,25].

West and East related African mitochondrial lineages have widely been reported throughout most of South American countries such as Peru,[20] Venezuela,[21] Argentina,[47] Bolivia,[23] Brazil,[22,48] and more recently in Ecuador.[49] In most of the cases, authors suggested multiple invasion sources for American populations occurred, which would explain differences in frequency for mitochondrial haplotypes observed throughout South America,[23,50,51] including Colombia.[18] However more recent research about historical demography and global genetic structure of *Ae. aegypti* using genome-wide and highly genetic resolutive markers such as polymorphic microsatellite loci and SNPs, have indicated a single source in the native range to the New World invasive range around the 16th-century, as consequence of slave trade between Africa and Americas.[4,10,24,25] Because mitochondrial genes show lower genetic resolution than SNPs and microsatellites, we conclude that the observed differentiation between mitochondrial lineages in Colombia does not evidence dual or multiple sources for Colombian populations, rather our observations indicate Colombian *Ae. aegypti* founding population is coherent with a single invasion source of a large mitochondrial polymorphic African related population that rapidly spread throughout Colombia. This hypothesis does not

exclude scenarios for multiple introductions nor colonisation routes followed by high contemporary gene flow in Colombian populations.

Despite results agreeing with previous results suggesting lacking geographic structure of Colombian populations, it is worth notice since these studies were relied on relatively small genetic markers number (i.e., partial mitochondrial COI and ND4 genes, and eight microsatellites), therefore we suggest further studies using more resolutive genetic markers must be conducted to assessing inter-population dynamics at a finer scale of *Ae. aegypti* Colombian populations. Recent advances in next-generation sequencing technology show promising results in resolving population structure. Rašić et al.[52] showed clearer separation of genome-wide SNP genotype profiles among *Ae. aegypti* populations that originated from different locations compared to microsatellite markers, and analyses for genetic relatedness of individuals from same populations was more accurately estimated when RAD-based genetic data was employed. Additionally, Lee et al.[10] utilised millions of SNPs from whole genome sequencing data and showed clearer separation of populations in close proximity (20 km apart) compared to a study based on ~25K SNPs.[9] Therefore, the use of alternative methods such as those based on high throughput data may contrast this finding and opening the possibility of pursuing more accurate results for population demography and history of Colombian populations.

AUTHORS' CONTRIBUTION

YM performed the experiments; YC and AGP designed the experiments, analysed data and wrote the manuscript; OT and AGP were performed the study design, project administration and funding acquisition. All authors read and approved the final manuscript.

REFERENCES

1. Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, et al. The global distribution and burden of dengue. Nature. 2013; 496: 504-7. https://doi.org/10.1038/nature12060.

2. Townson H, Nathan MB, Zaim M, Guillet P, Manga L, Bos R, et al. Exploiting the potential of vector control for disease prevention. Bull World Health Organ. 2005; 83(12): 942-7. Available from: https://apps.who.int/iris/handle/10665/73452.

3. Urdaneta-Marquez L, Failloux AB. Population genetic structure of *Aedes aegypti*, the principal vector of dengue viruses. Infect Genet Evol. 2011; 11(2): 253-61. https://doi.org/10.1016/j.meegid.2010.11.020.

4. Gloria-Soria A, Ayala D, Bheecarry A, Calderon-Arguedas O, Chadee DD, Chiappero M, et al. Global genetic diversity of *Aedes aegypti*. Mol Ecol. 2016; 25(21): 5377-95. https://doi.org/10.1111/mec.13866.

5. Guagliardo SAJ, Lee Y, Pierce AA, Wong J, Chu YY, Morrison AC, et al. The genetic structure of *Aedes aegypti* populations is driven by boat traffic in the Peruvian Amazon. PLoS Negl Trop Dis. 2019; 13(9): e0007552. https://doi.org/10.1371/journal.pntd.0007552.

6. Kotsakiozi P, Gloria-Soria A, Schaffner F, Robert V, Powell JR. *Aedes aegypti* in the Black Sea: recent introduction or ancient remnant? Parasit Vectors. 2018; 11(1): 396. https://doi.org/10.1186/s13071-018-2933-2.

7. Hopperstad KA, Reiskind MH, Labadie PE, Burford Reiskind MO. Patterns of genetic divergence among populations of *Aedes aegypti* L. (Diptera: Culicidae) in the southeastern USA. Parasit Vectors. 2019; 12(1): 511. https://doi.org/10.1186/s13071-019-3769-0.

8. Burford Reiskind MO, Labadie P, Bargielowski I, Lounibos LP, Reiskind MH. Rapid evolution and the genomic consequences of selection against interspecific mating. Mol Ecol. 2018; 27(18): 3641-54. https://doi.org/10.1111/mec.14821.

9. Pless E, Gloria-Soria A, Evans BR, Kramer V, Bolling BG, Tabachnick WJ, et al. Multiple introductions of the dengue vector, *Aedes aegypti*, into California. PLoS Negl Trop Dis. 2017; 11(8): e0005718. https://doi.org/10.1371/journal.pntd.0005718.

10. Lee Y, Schmidt H, Collier TC, Conner WR, Hanemaaijer MJ, Slatkin M, et al. Genome-wide divergence among invasive populations of *Aedes aegypti* in California. BMC Genomics. 2019; 20(1): 204. https://doi.org/10.1186/s12864-019-5586-4.

11. Monteiro FA, Schama R, Shama R, Martins AJ, Gloria-Soria A, Brown JE, et al. Genetic diversity of Brazilian *Aedes aegypti*: patterns following an eradication program. PLoS Negl Trop Dis. 2014; 8(9): e3167. https://doi.org/10.1371/journal.pntd.0003167.

12. Louise C, Vidal PO, Suesdek L. Microevolution of *Aedes aegypti*. PLoS One. 2015; 10(9): e0137851. https://doi.org/10.1371/journal.pone.0137851.

13. Vidal PO, Suesdek L. Comparison of wing geometry data and genetic data for assessing the population structure of *Aedes aegypti*. Infect Genet Evol. 2012; 12(3): 591-6. https://doi.org/10.1016/j.meegid.2011.11.013.

14. Ocampo CB, Wesson DM. Population dynamics of *Aedes aegypti* from a dengue hyperendemic urban setting in Colombia. Am J Trop Med Hyg. 2004; 71(4): 506-13. https://doi.org/10.4269/ajtmh.2004.71.506.

15. Caldera SM, Jaramillo MC, Cochero S, Pérez-Doria A, Bejarano EE. Genetic differences between populations of *Aedes aegypti* from municipalities in northern Colombia, with high and low dengue incidence. Biomedica. 2013; 33(1): 89-98. https://doi.org/10.7705/biomedica.v33i0.1573.

16. Atencia MC, Pérez MJ, Caldera SM, Jaramillo MC, Bejarano EE. Genetic variability of *Aedes aegypti* in the department of Sucre, Colombia, by analysis of the nucleotide sequence of the mitochondrial ND4 gene. Biomedica. 2018; 38(2): 267-76. https://doi.org/10.7705/biomedica.v38i0.3728.

17. Cadavid JM, Rúa G, Campo O, Bedoya G, Rojas W. Microgeographic and temporal genetic changes of *Aedes aegypti* from Medellín, Colombia. Biomedica. 2015; 35(1): 53-61. https://doi.org/10.7705/biomedica.v35i1.2343.

18. Jaimes-Dueñez J, Arboleda S, Triana-Chávez O, Gómez-Palacio A. 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 Negl Trop Dis. 2015; 9(4): e0003553. https://doi.org/10.1371/journal.pntd.0003553.

19. Gorrochotegui-Escalante N, Gomez-Machorro C, Lozano-Fuentes S, Fernandez-Salas L, De Lourdes Munoz M, Farfan-Ale JA, et al. Breeding structure of *Aedes aegypti* populations in Mexico varies by region. Am J Trop Med Hyg. 2002; 66(2): 213-22. https://doi.org/10.4269/ajtmh.2002.66.213.

20. da Costa-da-Silva AL, Capurro ML, Bracco JE. Genetic lineages in the yellow fever mosquito *Aedes* (*Stegomyia*) *aegypti* (Diptera: Culicidae) from Peru. Mem Inst Oswaldo Cruz. 2005; 100(6): 539-44. https://doi.org/10.1590/S0074-02762005000600007.

21. Herrera F, Urdaneta L, Rivero J, Zoghbi N, Ruiz J, Carrasquel G, et al. Population genetic structure of the dengue mosquito *Aedes aegypti* in Venezuela. Mem Inst Oswaldo Cruz. 2006; 101(6): 625-33. https://doi.org/10.1590/S0074-02762006000600008.

22. Scarpassa VM, Cardoza TB, Cardoso Jr RP. Population genetics and phylogeography of *Aedes aegypti* (Diptera: Culicidae) from Brazil. Am J Trop Med Hyg. 2008; 78(6): 895-903. https://doi.org/10.4269/ajtmh.2008.78.895.

23. Paupy C, Le Goff G, Brengues C, Guerra M, Revollo J, Barja Simon Z, et al. Genetic structure and phylogeography of *Aedes aegypti*, the dengue and yellow-fever mosquito vector in Bolivia. Infect Genet Evol. 2012; 12(6): 1260-9. https://doi.org/10.1016/j.meegid.2012.04.012.

24. Kotsakiozi P, Evans BR, Gloria-Soria A, Kamgang B, Mayanja M, Lutwama J, et al. Population structure of a vector of human diseases: *Aedes aegypti* in its ancestral range, Africa. Ecol Evol. 2018; 8(16): 7835-48. https://doi.org/10.1002/ece3.4278.

25. Crawford JE, Alves JM, Palmer WJ, Day JP, Sylla M, Ramasamy R, et al. Population genomics reveals that an anthropophilic population of *Aedes aegypti* mosquitoes in West Africa recently gave rise to American and Asian populations of this major disease vector. BMC Biol. 2017; 15(1): 16. https://doi.org/10.1186/s12915-017-0351-0.

26. Collins FH, Mendez MA, Rasmussen MO, Mehaffey PC, Besansky NJ, Finnerty V. A ribosomal RNA gene probe differentiates member species of the *Anopheles gambiae* complex. Am J Trop Med Hyg. 1987; 37(1): 37-41. https://doi.org/10.4269/ajtmh.1987.37.37.

27. Hickner PV, Debruyn B, Lovin DD, Mori A, Behura SK, Pinger R, et al. Genome-based microsatellite development in the *Culex pipiens* complex and comparative microsatellite frequency with *Aedes aegypti* and *Anopheles gambiae*. PLoS One. 2010; 5(9): e13062. https://doi.org/10.1371/journal.pone.0013062.

28. Slotman MA, Kelly NB, Harrington LC, Kitthawee S, Jones JW, Scott TW, et al. Polymorphic microsatellite markers for studies of *Aedes aegypti* (Diptera: Culicidae), the vector of dengue and yellow fever. Mol Ecol Notes. 2007; 7(1): 168-71. https://doi.org/10.1111/j.1471-8286.2006.01533.x.

29. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012; 28(12): 1647-9. https://doi.org/10.1093/bioinformatics/bts199.

30. Meirmans PG. GENODIVE version 3.0: Easy-to-use software for the analysis of genetic data of diploids and polyploids. Mol Ecol Resour. 2020; 20: 1126-31. https://doi.org/10.1111/1755-0998.13145.

31. Goudet J. Fstat (ver. 2.9.4), a program to estimate and test population genetics parameters. 2003. Available from: http://www2.unil.ch/popgen/softwares/fstat.htm.

32. Strimmer K. fdrtool: a versatile R package for estimating local and tail area-based false discovery rates. Bioinformatics. 2008; 24(12): 1461-2. https://doi.org/10.1093/bioinformatics/btn209.

33. Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes. 2004; 4(3): 535-8. https://doi.org/10.1111/j.1471-8286.2004.00684.x.

34. Nei M. Molecular evolutionary genetics. New York: Columbia University Press; 1987.

35. Chapuis MP, Estoup A. Microsatellite null alleles and estimation of population differentiation. Mol Biol Evol. 2007; 24(3): 621-31. https://doi.org/10.1093/molbev/msl191.

36. Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. Bioinformatics. 2012; 28(19): 2537-9. https://doi.org/10.1093/bioinformatics/bts460.

37. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics. 2003; 164(4): 1567-87.

38. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000; 155(2): 945-59.

39. Rosenberg NA. distruct: a program for the graphical display of population structure. Mol Ecol Notes. 2004; 4(1): 137-8. https://doi.org/10.1046/j.1471-8286.2003.00566.x.

40. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 2005; 14(8): 2611-20. https://doi.org/10.1111/j.1365-294X.2005.02553.x.

41. Earl D, vonHoldt B. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour. 2012; 4: 359-61. https://doi.org/10.1007/s12686-011-9548-7.

42. Loiselle BA, Sork VL, Nason J, Graham C. Spatial genetic structure of a Tropical understory shrub, *Psychotria officinalis* (Rubiaceae). Am J Bot. 1995; 82(11): 1420-5. https://doi.org/10.2307/2445869.

43. Jasper M, Schmidt TL, Ahmad NW, Sinkins SP, Hoffmann AA. A genomic approach to inferring kinship reveals limited intergenerational dispersal in the yellow fever mosquito. Mol Ecol Resour. 2019; 19(5): 1254-64. https://doi.org/10.1111/1755-0998.13043.

44. Iacchei M, Ben-Horin T, Selkoe KA, Bird CE, García-Rodríguez FJ, Toonen RJ. Combined analyses of kinship and FST suggest potential drivers of chaotic genetic patchiness in high gene-flow populations. Mol Ecol. 2013; 22(13): 3476-94. https://doi.org/10.1111/mec.12341.

45. Beerli P. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. Bioinformatics. 2006; 22(3): 341-5. https://doi.org/10.1093/bioinformatics/bti803.

46. Beerli P, Palczewski M. Unified framework to evaluate panmixia and migration direction among multiple sampling locations. Genetics. 2010; 185(1): 313-26. https://doi.org/10.1534/genetics.109.112532.

47. Dueñas JC, Llinás GA, Panzetia-Dutari GM, Gardenal CN. Two different routes of colonization of *Aedes aegypti* in Argentina from neighboring countries. J Med Entomol. 2009; 46(6): 1344-54. https://doi.org/10.1603/033.046.0613.

48. Lima RS, Scarpassa VM. Evidence of two lineages of the dengue vector *Aedes aegypti* in the Brazilian Amazon, based on mitochondrial DNA ND4 gene sequences. Genet Mol Biol. 2009; 32(2): 414-22. https://doi.org/10.1590/S1415-47572009005000036.

49. Cevallos V, Benítez D, Coloma J, Carrazco A, Wang C, Holecheck S, et al. Genetic variation of *Aedes aegypti* populations from Ecuador. bioRxiv. 2019: 2019.12.17.875591. https://doi.org/10.1101/2019.12.17.875591.

50. Bracco JE, Capurro ML, Lourenço-de-Oliveira R, Sallum MA. Genetic variability of *Aedes aegypti* in the Americas using a mitochondrial gene: evidence of multiple introductions. Mem Inst Oswaldo Cruz. 2007; 102(5): 573-80. https://doi.org/10.1590/S0074-02762007005000062.

51. Paupy C, Brengues C, Kamgang B, Hervé 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(3): 391-400. https://doi.org/10.1603/0022-2585(2008)45[391:GFBDAS]2.0.CO;2.

52. Rašić G, Filipović I, Weeks AR, Hoffmann AA. Genome-wide SNPs lead to strong signals of geographic structure and relatedness patterns in the major arbovirus vector, *Aedes aegypti*. BMC Genomics. 2014; 15: 275. https://doi.org/10.1186/1471-2164-15-275.