Genetic lineages in the yellow fever mosquito *Aedes (Stegomyia) aegypti* (Diptera: Culicidae) from Peru

André Luis da Costa-da-Silva/+, Margareth Lara Capurro*, José Eduardo Bracco

Laboratório de Bioquímica e Biologia Molecular de Vetores, Superintendência de Controle de Endemias, Rua Paula Souza 166, 01027-000 São Paulo, SP, Brasil *Departamento de Parasitologia, Instituto Ciências Biomédicas, Universidade de São Paulo, São Paulo, SP, Brasil

The yellow fever mosquito Aedes aegypti *was introduced in Peru in 1852 and was considered to be eradicated in 1958. In 2001,* Ae. aegypti *had been recorded in 15 out of 24 Peruvian Departments. Peru has great ecological differences between the east and west sides of Andes. Because of this, we consider that* Ae. aegypti *populations of both east and west sides can have a genetically distinct population structure.*

In this study we examined genetic variability and genealogical relationships among three Ae. aegypti *Peruvian populations: Lima, Piura (west Andes), and Iquitos (east Andes) using a fragment of the ND4 gene of the mitochondrial genome. Three haplotypes were detected among 55 samples. Lima and Iquitos showed the same haplotype (Haplotype I), whereas Piura has two haplotypes (Haplotype II and III). Haplotype II is four mutational steps apart from Haplotype I, while Haplotype III is 13 mutational steps apart from Haplotype I in the network. The analysis of molecular variation showed that mostly of the detected genetic variation occurs at interpopulational level. The significant value* Φ_{st} *suggests that Piura population is structured in relation to Lima and Iquitos populations and the gene flow of the ND4 is restricted in Piura when compared to Lima and Iquitos. Genetic relationship between haplotype I and haplotype II suggests introduction of the same mtDNA lineage into those localities. However the existence of a genetically distant haplotype III also suggests introduction of at least two* Ae. aegypti *lineages in Peru.*

Key words*: Aedes aegypti* - ND4 - mitochondrial DNA - genetic variability - phylogeography

Aedes aegypti is widespread throughout tropical and subtropical areas of the world. It is vector of several arboviruses, including yellow fever, dengue (serotypes 1-4), and Chikungunya viruses, all of which can cause severe morbidity and mortality (http://www.cdc.gov/). A mosquito of African origin, it was probably imported into the American continent through African slave trade during the 18th century (Campillo 2001). In spite of a continuous eradication program carried out throughout the American continent in the 1950's and the 1960's, this species has been able to disperse and colonize several countries (Agrelo 1996).

Ae. aegypti was first reported in Peru in 1852 and it was considered eradicated in 1958. In 1984, the mosquito was reintroduced in Loreto Department and, in the same year, its presence was reported in several cities, one of which was Iquitos. A year later, the mosquito was found in the state of Piura and in 2000, larvae were collected in Lima. In 2001, 15 of the 24 Peruvian states had the *Ae. aegypti,* thus presenting a potential risk for public health in that country (Andrade et al. 2001).

Ae. aegypti shows a high adaptive capability. In fact, this vector displays variations in its morphology, physiology and behavior. At least three forms have been described (Mattingly 1957, Tabachnik & Powell 1978). Lourenço-de-Oliveira et al. (2002) reported low infection rates of the dark form of the mosquito (*Ae. aegypti formosus*) by the yellow fever virus. On the other hand, the light form (*Ae. aegypti aegypti*) presented high susceptibility rates, indicating differences in the subspecies vectorial capacities.

Mitochondrial DNA (mtDNA) polymorphism is a widely used tool in the study of genic flow in several organisms (Excoffier et al. 1992, Brower 1994, Birungi & Munstermann 2002, de Brito et al. 2002). Inferences about migration events are made based on the geographical pattern of mitochondrial genetic variability, although this pattern can be influenced by historical and/or ecological events (Templeton et al. 1995). In the analysis of *Ae. aegypti* populations of Mexico, Gorrochotegui-Escalante et al. (2000) detected, by means of random amplified polymorphic DNA-polymerase chain reaction (RAPD-PCR) markers, populations isolated by distances greater than 250 km. In this study, amplification of ND4 fragment and single-strand conformation polymorphism (SSCP) analysis allowed for the description of seven haplotypes among the 574 samples examined. The authors showed that the ND4 marker is very informative.

Divided by the Andes mountain range from North to South and with different climates in the coastal, mountain and jungle regions, Peru displays geographical and ecological conditions that sustain the existence of genetically differentiated and isolated *Ae. aegypti* populations.

The characterization of ND4 haplotypes and the genetic variability among different *Ae. aegypti* populations in Peru might be an important comparative data for suitable proposals to control this vector.

Financial support: Sucen, Fapesp (00/1238-7), CNPq (501529/2003-9)
+Corresponding author. E-mail: alcosta@icb.usp.br
Received 14 March 2005
Accepted 10 August 2005

This study also shows that single nucleotide polymorphism (SNP) was found in sequences of ND4 fragment of *Ae. aegypti* mtDNA from three cities in Peru. This allowed for inferences about gene flow, biogeographical patterns, and phylogeographical relations among the haplotypes detected in this study.

MATERIALS AND METHODS

Table I lists the analyzed Peruvian *Ae. aegypti* populations, locations, dates of collection, and geographic coordinates. Fig. 1 shows the geographical locations of the studied populations. The specimens (adults preserved in 100% ethanol) were obtained from the Alexander von Humboldt Tropical Medicine Institute of the Peruvian Cayetano Heredia University. The specimens were identified as *Ae. aegypti* in compliance with the description in Belkin et al. (1970), and were kept in 100% ethanol at –70°C. Before extracting the DNA, females had their abdomens removed to avoid contaminations with exogenous DNA derived from blood meal and/or semen stored in the female spermateca.

DNA extraction was performed using a phenol-chloroform extraction protocol described in Sambrook et al. (1989). After precipitation, the DNA was resuspended in 100 ml of TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0).

TABLE I

Location, dates, and coordinates of the Peruvian samples of *Aedes aegypti*

Sampling locality	Dates	Geographical coordinates
Piura	06/10/02	04º49' S 80º38' W
Lima	08/12/02	11º81' S 77º07' W
Iquitos	24/02/03	03º82' S 72º30' W

A fragment of 361 bp relative to the mitochondrial ND4 gene was amplified from each DNA sample by PCR in a final reaction volume of 50 µl with 2 mM Mg^{++}, 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 0.5 mM of each primer, 0.2 mM dNTP mix, 2 U of Taq DNA polymerase (Invitrogen), and 1 µl of template DNA. The amplified regions corresponded to nucleotides 8,521-8,882 in *Ae. albopictus* (GenBank #AY072044).

The primers used in the PCR were 5'-ATTGCCTAAGG CTCATGTAG-3' and 5'- TCGGCTTCCTAGTCGT-TCAT- 3'. For each experiment a negative control was performed. The thermocycler program was configured for an initial denaturation step at 94°C for 2 min, followed by 35 cycles at 94°C for 1 min, 56°C for 30 s, and 72°C for 1 min, and a final elongation step at 72°C for 7 min.

Fig. 1: the geographical localities sampled in the present study. Arrows indicate the sample cities.

The PCR products were precipitated by adding 1 volume of PEG solution (20% in 2.5 M NaCl), and then the tubes were gently mixed and incubated for 15 min at 37°C. The mixture was quickly agitated and incubated at 37°C for 15 min, centrifuged and the precipitate was washed twice with ice cold 80% ethanol.

Sequencing of the amplified fragments was carried out using the ABI Prism BigDye™ Terminator Cycle Sequencing Ready Reaction Kit and Ampli-Taq® DNA Polymerase. For each reaction we used 2 ml of BigDye™, 2 µl of Save Money buffer (200 mM Tris-HCl, pH 9.0, 5 mM $MgCl_2$), 3 µl of primer (10 mM), 30 to 90 ng of DNA, and 1.5 µl of autoclaved Milli-Q water. The thermocycler program was set for an initial denaturation step at 96°C for 10 s, 50°C for 20 s, and 60°C for 4 min and a final process of 4°C with a 1°C per second ramp.

Sequence reactions were purified by precipitation with isopropanol and analyzed in the ABI Prism® 377 (Applied Biosystems, US) automatic sequencer at 51°C for 7 h, with 1200 scans per hour and 1680 V of electrophoresis voltage.

The consensus sequences were obtained through the analysis of the sense and antisense sequences of each PCR fragment employing the BioEdit Sequence Alignment Editor program (Hall 1999).

Analysis of genetic variation in populations must consider, besides the genetic drift, other forces that may be involved in alterations of allelic/haplotypic frequencies such as mutations, natural selection and the combination of both since the polymorphism pattern in populations is affected by selection (Tajima 1989, Fernandes-Matioli 2001). The Tajima's D (Tajima 1989) and F* & D* neutrality tests (Fu & Li 1993) as well the estimated mean number of nucleotide differences per site (π), the proportion of segregating sites (*S*) and the number of haplotypes per sample (*k*) were performed using the DnaSP program version 4.0 (Rozas et al. 2003).

The relationships between *Ae. aegypti* haplotypes were inferred in Network Software to create an unrooted haplotype network, using star contraction (Röhl 1999 *apud* Forster et al. 2001). The analysis of molecular variation (AMOVA) based on methodology described in Excoffier et al. (1992) is a technique which determines the amount of variation due to population substructure given an a priori set of population hierarchies using Φst (de Brito et al. 2002). The analysis was conducted in Arlequin 2.0 (Schneider et al. 2000). The first model tested considered that there was no hierarchy in the localities. A second model considering geographical/ecological patterns was tested. The AMOVA provided an overall estimate of population differentiation similar to F*st* following Slatkin (1991) across the three samples locations.

RESULTS

Sequencing, analysis, and sequence alignment - The fragment of 361 base pairs of the gene ND4 was sequenced in 19 individuals from Piura, 17 from Iquitos, and 19 from Lima (Table II). The alignment of consensus sequences was performed visually.

Analysis of polymorphism indexes - Table II summarizes the results of polymorphism indexes found among 55 sequences. We found three haplotypes. Haplotype I was present in Iquitos and Lima populations and haplotypes II and III were determined in Piura population. Fourteen polymorphic sites were identified. All mutations were synonymous. Nucleotide diversity was estimated for the three Peruvian populations.

Neutrality tests - Neutrality tests were applied in grouped Peruvian populations. A significant neutrality deviation was detected in the neutral selectivity test D* (Table III). Due to homogeneity of the Iquitos and Lima sequences, the tests were repeated only in the Piura population, in an attempt to show that the neutrality deviation of Peruvian sequences as a whole is being caused by the heterogeneity of the Piura sequences. As it can be seen in Table III, the Piura population showed significant neutrality deviation in test D*.

TABLE II

Polymorphism indexes of the three peruvian populations of *Aedes aegypti*

Population	N	K	Haplotype	π
Iquitos	17	1	I	0.0
Lima	19	1	I	0.0
Piura	19	2	II ; III	
0.00855				
All	55	3	I ; II ; III	0.00790

N: sample size; K: haplotypes per population; π: nucleotide diversity

TABLE III

Neutrality tests of three Peruvian populations and Piura population

Populations	Tajima's D	Fu & Li's F*	Fu & Li's D*
Lima, Iquitos, and Piura	– 0.20236[a]	1.12406[a]	1.54547[b]
Piura	– 0.06802[a]	1.16591[a]	1.43728[b]

a: P > 0.10 not significant; *b*: P < 0.05 significant

Analysis of haplotype network - The analysis resulted in an unrooted cladogram where each mutational step is represented (Fig. 2). Three different groupings were defined —haplotype I, haplotype II, haplotype III. Haplotype I, which encompasses all individuals from Iquitos and Lima, is distant from haplotype II, common to 16 individuals in the Piura population, in four mutational steps (Fig. 2). The other three individuals of the Piura population show haplotype III, which is connected to the other two haplotypes in the cladogram by 10 mutations (Fig. 2). As showed in Fig. 2, haplotype III differs from haplotype II by 11 nucleotides and it differs from haplotype I by 13 nucleotides. The same groups were found in Mexican populations (Gorrochotegui-Escalante 2000, data not showed here).

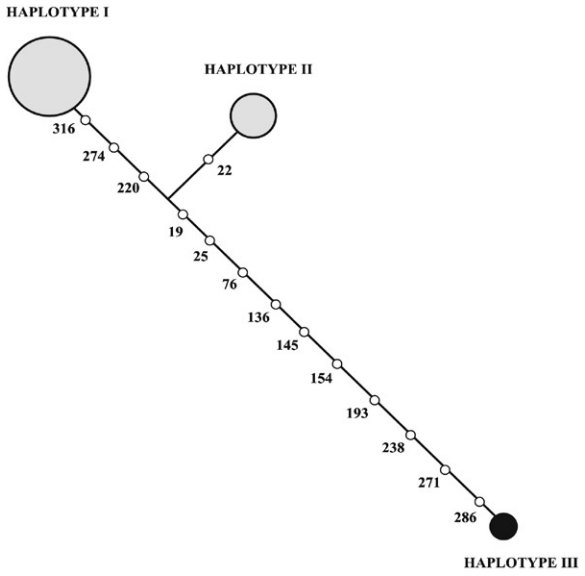

Fig. 2: haplotype network of haplotypes of *Aedes aegypti* from Peruvian populations. Arrows indicate the cities where mosquitoes were collected. Numbers are the mutational positions in a 361 bp fragment.

AMOVA analysis - We performed AMOVA to investigate population differentiation. Our first estimate considered all localities in an island model (without any specific hierarchy). This analysis produced a *Φst* value of 0.89930 ($p < 0.00001$) with 89.93% of the total variation between the populations and just 10.07% within the populations (Table IV).

In a second estimate, a geographical/ecological hierarchy was tested by separating the populations into two groups: the west of the Andes with a drier climate (Lima and Piura) and the east of the Andes with a hot and humid climate (Iquitos). Results of the analysis showed strong structuring where the highest covariation component (124.20%) is due to differences among populations within the groups ($\Phi sc = 0.77804$, $p < 0.001$, Table V). That is a result of the great distance between the Lima and Piura haplotypes, here clustered in the same group.

DISCUSSION

Several studies on population genetics have evaluated the variation and genetic structure of *Ae. aegypti* populations. Generally speaking, high genetic variability has been found in this species and the variability pattern suggests there are genetically differentiated subpopulations among the many locations analyzed (Yan et al. 1999, Gorrochotegui-Escalante et al. 2000, de Souza et al. 2000, Fraga et al. 2003).

Genetic variability in *Ae. aegypti* was found in the present study. The characterization of three Peruvian haplotypes allowed for an estimated nucleotidic diversity of $\pi = 0.00790$ (Table II). This variability is due only to the Piura samples ($\pi = 0.00855$, Table II) since they have two haplotypes, while Lima and Iquitos samples share the same haplotype. Therefore, the variability of the sampled Peruvian populations is not related to an increase in geographical distribution: the distribution of the Piura population is much smaller than that of all Peruvian populations. However, the nucleotidic diversity of Piura samples was the highest, suggesting that the isolation-by-distance mechanism is not responsible for this diversity.

The infinite-sites model seems adequate to explain the polymorphism observed in ND4 fragment, since of the 14 polymorphic sites observed, all are synonymous mutations (Table II). The only neutrality deviation detected for the Peruvian populations in test D* (Fu & Li 1989, Table III) was attributed to a great number of polymorphic sites found in Piura sequences. This fact is reinforced by the significant positive neutrality deviations in test D* carried out only for Piura sequences (Table III). As showed by de Brito et al. (2002), significantly positive results for neutrality deviation may suggest population substructuring.

The AMOVA analysis showed that: (1) there is a population structuring (Table IV) and (2) the most important component in the found variability is due to the variation among populations of the same group, i.e. given the geographical/ecological hierarchy proposed, the highest variability is a result of differences between the Lima and Piura populations (Table V). That can be explained by the genetic distances in Piura in comparison to the other two

TABLE IV

Partition (without hierarchy) of variation in the frequency of ND4 haplotypes among *Aedes aegypti* from Peru

Source of variation	d.f.	Variance components	% of variation	Φ Statistics
Among populations within groups	2	0.45120	89.93	Φst = 0.89930[a]
Within populations	15	0.05053	10.07	-

a: $P < 10^{-5}$

TABLE V

Partition of variation in the frequency of ND4 haplotypes among *Aedes aegypti* from Peru. Hierarchy: group 1 - West from Andes, hot and dry (Lima and Piura) and group 2 - East from Andes, hot and humid (Iquitos)

Source of variation	d.f.	Variance components	% of variation	Φ Statistics
Among groups	1	– 0.93545	–59.64	Φct = – 0.59636 [a]
Among populations within groups	1	1.94825	124.20	Φsc = 0.77804 [b]
Within populations	19	0.55579	35.43	Φst = 0.64568 [b]

a: not significant; *b*: $P < 0.001$

populations, since both haplotypes verified for Piura differ from haplotype I shared by the Lima and Iquitos populations. Besides, haplotype III from Piura is genetically very far from haplotype I, thus increasing the variation between populations. The highly significant F-values $\Phi_{st} = 0.64568$ and $\Phi sc = 0.77804$ indicate high structuring in the Peruvian populations (Table V). Since Lima and Iquitos share haplotype I, the structuring found reinforces the hypotheses of restricted genic flow of the ND4 gene between these two populations and the Piura population.

Concerning the low capacity of active dispersion of the *Ae. aegypti* mosquito (Trips & Hausermann 1986, Muir & Kay 1998), it was expected that the existence of a road connecting Piura and Lima (Fig. 1) could be an ultimate passive dispersion factor for the existence and maintenance of the genic flow and the genetic proximity between the populations of the two locations. Moreover, the isolation of Iquitos, caused by geographical and climatic (low temperatures) barriers represented by the Andes (Fig. 1) would be another determining factor to make the Piura and Lima samples more genetically closer between them when compared to Iquitos samples. However, this pattern could not be observed for the ND4 gene in these populations. Genetic relations between haplotype I from Lima and haplotypes II and III from Piura depicted in the network in Fig. 2 refute the hypotheses of higher genetic proximity of these two populations in comparison to Iquitos samples. This pattern of genetic distances non-related to geographical distances has already been described for *Ae. aegypti*. In a study with alozymes carried out in Argentinean populations, de Souza et al. (2000) showed that the three populations analyzed were genetically close. Nevertheless, the highest genetic distance observed was between the cities of Buenos Aires and Zárate, which were geographically closer.

As for the close relationship between haplotypes I and II (Table II, Fig. 2), this study suggests that a similar lineage was introduced in these locations and it has possibly dispersed from one region to another by human movement, considering the geographical and climatic restrictions previously mentioned.

Fonseca et al. (2001) analyzed the genetic structure of *Ae. (Finlaya) japonicus* populations in Japan and the United States using ND4 sequences. The results showed geographically differentiated Japanese populations, indicating limited gene flow. For populations of American cities, a significant genetic differentiation among samples suggested multiple introductions of the species as showed in our present study.

The existence of genetically distant haplotype III suggests the occurrence of at least two introductions of *Ae. aegypti* in Piura. Haplotype II might have derived from haplotype I (or vice versa) by means of recent mutations of this lineage. An evident division of these ND4 haplotypes into two genetically distinct groups can be observed through our data and Mexican population data (Black IV, pers. commun.). Thus, viral susceptibility, vectorial competence, resistance to insecticides and ecological adaptations can be different among the Peruvian populations of *Ae. aegypti* as showed in other populations (Lourenço-de-Oliveira et al. 2002, Fraga et al. 2003).

Therefore, knowledge of genetic variation of the vector populations related to geographical and/or historical factors can be highly informative when developing effective control strategies.

ACKNOWLEDGEMENTS

To Dr Palmira Ventosilha for supplying mosquito samples; Dr William C Black IV for kindly sending Mexican populations dataset; Kleber José Whitaker and Alexandre Santos de Moura for technical assistance.

REFERENCES

Agrelo RS 1996. *Aedes aegypti*, *Aedes albopictus* (Diptera, Culicidae) y su papel como vectores em las Américas. La situación de Uruguay. *Rev Méd Uruguay 12*: 28-36.

Andrade CS, Cáceres AG, Vaquerizo A, Bernal SI, Cachay LS 2001. Reappearance of *Aedes aegypti* (Diptera: Culicidae) in Lima, Peru. *Mem Inst Oswaldo Cruz 96*: 657-658.

Belkin JN, Heinemann SJ, Page WA 1970. Mosquitoes studies (Diptera, Culicidae). XXI. The Culicidae of Jamaica. *Contrib Am Entomol Inst (Ann Arbor) 6*: 1-458.

Birungi J, Munstermann LE 2002. Genetic structure of *Aedes albopictus* (Dipera: Culicidae) populations based on mitochondrial ND5 sequences: evidence for an independent invasion into Brazil and United States. *Genetics 95:* 125-132.

Brower AVZ 1994. Rapid morphological radiation and convergence among races of the butterfly *Helicinius erato* inferred from patterns of mitochondrial DNA evolution. *Proc Natl Acad Sci USA 91*: 6491-6495.

Campillo MC 2001. Las grandes epidemias en lá América colonial. *Arch Zootec 50*: 597-612.

de Brito RA, Manfrin MH, Sene FM 2002. Mitochondrial DNA phylogeography of Brazilian populations of *Drosophila buzzatii*. *Gen Mol Biol 25*: 161-171.

de Souza GB, Avilés G, Gardenal CN 2000. Allozymic polymorphism in *Aedes aegypt*i populations from Argentina. *J Am Mosq Control Assoc 16*: 206-209.

Excoffier L, Smouse PE, Quattro JM 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics 131*: 479-491.

Fernandes-Matioli FMC 2001. Genealogias e o processo de coalescência. In SR Matioli, *Biologia Molecular e Evolução*, Holos, Ribeirão Preto, p. 162-171.

Fonseca DM, Campbell S, Crans WJ, Mogi M, Miyagi I, Toma T, Bullians M, Andreadis TG, Berry RL, Pagac B, Sardelis MR, Wilkerson RC 2001. *Aedes (Finlaya) japonicus* (Diptera:Culicidae), a newly recognized mosquito in the United States: analyses of genetic variation in the United States and putative source populations. *J Med Entomol 38*: 135-146.

Forster P, Torroni A, Renfrew C, Röhl A 2001. Phylogenetic star contraction applied to Asian and Papuan mtDNA evolution. *Mol Biol Evol 18*: 1864-1881.

Fraga EC, Santos JMM, Maia JF 2003. Enzimatic variability in *Aedes aegypti* (Diptera: Culicidae) populations from Manaus-AM, Brazil. *Gen Mol Biol 26*: 181-187.

Fu YX, Li WH 1993. Statistical tests of neutrality mutations. *Genetics 133*: 693-709.

Gorrochotegui-Escalante N, Munoz ML, Fernandez-Salas I, Beaty BJ, Black IV WC 2000. Genetic isolation by distance among *Aedes aegypti* populations along the northeastern coast of Mexico. *Am J Trop Med Hyg 62*: 200-209.

Hall TA 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser 41*: 95-98.

Lourenço-de-Oliveira R, Vazeille M, Filippis AMB, Failloux AB 2002. Oral susceptibility to yellow fever virus of *Aedes aegypti* from Brazil. *Mem Inst Oswaldo Cruz 97*: 437-439.

Mattingly PF 1957. Genetical aspects of the *Aedes aegypti* problem. I. Taxonomy and bionomics. *Ann Trop Med Parasitol 51*: 392-408.

Muir LE, Kay BH 1998. *Aedes aegypti* survival and dispersal estimated by mark-release-recapture in northern Australia. *Am J Trop Med Hyg 58*: 277-282.

Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics 19*: 2496-2497.

Sambrook J, Fritsch EF, Maniatis T 1989. *Molecular Cloning. A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, New York.

Scheneider S, Roessli D, Excoffier L 2000. Arlequin ver. 2000: A software for population genetic data analysis, Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Slatkin M 1991. Inbreeding coefficients and coalescence times. *Genet Res 58*: 167-175.

Tabachnick WJ, Powell JR 1978. Genetic structure of the east African domestic populations of *Aedes aegypti*. *Nature 272*: 535-537.

Tajima F 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics 123*: 585-595.

Templeton AR, Routman E, Phillips A 1995. Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the Tiger Salamander, *Ambystoma tigrinum*. *Genetics 140*: 767-782.

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

Yan G, Romero-Severson J, Walton M, Chadee DD, Severson DW 1999. Population genetics of the yellow fever mosquito in Trinidad: comparisons of amplified fragment length polymorphism (AFLP) and restriction fragment length polymorphism (RFLP) markers. *Mol Ecol 8*: 951-963.