Zoological Studies **60:**46 (2021)
doi:10.6620/ZS.2021.60-46

Zoological Studies

Open Access

Moderate Population Structure in *Drosophila sturtevanti* from the South American Atlantic Forest Biome

Bruna Memari Trava[1], Rogério Pincela Mateus[2], Luciana Paes de Barros Machado[2], and Lilian Madi-Ravazzi[1,*]

[1]*Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP), Cristóvão Colombo Street, 2265 – Jardim Nazareth, São José do Rio Preto, SP 15054-000, Brazil.*
**Correspondence: E-mail: lilian.madi@unesp.br (Madi-Ravazzi). Tel: +55 17 3221-2389. Fax: +55 17 3221-2390.*
E-mail: brunamemaritrava@gmail.com (Trava)
[2]*Universidade Estadual do Centro-Oeste (UNICENTRO), Campus CEDETEG, Élio Antonio Dalla Vecchia Alley, 838, Vila Carli, Guarapuava, PR 85040-167, Brazil. E-mail: rogeriopmateus@gmail.com (Mateus); lpbmachado@gmail.com (Machado)*

Received 30 May 2020 / Accepted 25 May 2021 / Published 17 August 2021
Communicated by Chung-Pin Lin

Drosophila sturtevanti is a widely distributed Neotropical species. In South America, it is abundant and adapted to different phytophysiognomies of the Atlantic Forest biome. Reproductive, chromosomal and enzymatic studies have indicated the existence of a differentiation among *D. sturtevanti* populations. In this work, the level of genetic diversity and the population genetic structure were analyzed using four population groupings. One hundred and twenty-six *D. sturtevanti* males collected from nine forest fragments were analyzed for 11 species-specific microsatellite loci. A total of 109 alleles, ranging from 2 to 16 alleles per locus, were detected. The highest mean observed heterozygosity - H_O was estimated in samples from the largest collection areas, and the lowest H_O was from a population where fire events are common. A low molecular variation, around 3% among populations and negative among groups, an absence of genetic and geographic correlations and a moderate genetic differentiation - F_{ST} = 0.0663 - indicated that *D. sturtevanti* is not strongly structured. Besides no overall genetic and geographic distance correlation, the pair of closest geographically populations Matão and Nova Granada showed the lower differentiation through F_{ST}, DC and a Neighbor Joining tree. Ribeirão da Ilha - RDI, an isolated insular population, was the most differentiated according to F_{ST}, DC and a cluster-based Bayesian analysis. The isolation of RDI that resulted in significant divergence could be ancient, because of sea level regressions/transgressions, or more recently via founder effect/genetic drift by anthropic action carrying *D. sturtevanti* hosts from continent to island. This work is important for understanding the genetic variability distribution of a Neotropical forest-dwelling *Drosophila* species using for the first time, a wide population distribution approach.

Key words: *Saltans* group, Forest fragmentation, Genetic diversity, Simple sequence repeats, Population differentiation.

BACKGROUND

The Brazilian Atlantic Forest biome is formed by a complex set of vegetation types, or phytophysiognomies, which are distributed over approximately 1.3 million km^2, over 17 states, and along the coast and interior of southern Brazil (Galindo-Leal and Câmara 2003). This biome is one of the most important in terms of endemism and biodiversity in the world; however, it is also the most threatened (Mittermeier et al. 2004), and

Citation: Trava BM, Mateus RP, de Barros Machado LP, Madi-Ravazzi L. 2021. Moderate population structure in *Drosophila sturtevanti* from the South American Atlantic Forest biome. Zool Stud **60:**46. doi:10.6620/ZS.2021.60-46.

© 2021 Academia Sinica, Taiwan

only between 7 and 11% of its original area currently remains (Salgueiro et al. 2004; Tabarelli et al. 2005). The different phytophysiognomies have very particular climates—for example, the dense ombrophilous forest is hot and humid; the seasonal forest (semideciduous and deciduous) has hot and humid summers, with cold and dry winters; and the mixed ombrophilous forest has a humid and cold climate (Marques 2016).

It is known that South American biomes, especially the Atlantic Forest, contain much of the global biodiversity. However, the anthropogenic environmental impacts in the last 100 years are probably the most severe in human history. Deforestation is the dominant land-use trend in Latin America (Ramankutty and Foley 1999; Achard et al. 2002), and several biomes, including the Brazilian Cerrado and the Amazon and Atlantic rain forests, have been and continue to be severely affected by conversion to agriculture and pastures (Galindo-Leal and Câmara 2003; Klink and Machado 2005). This deforestation process is mainly due to traditional and export-oriented industrial agriculture (soybean production, for example) and cattle ranching (Hecht 1993; Grau et al. 2008), which indirectly have favored other forms of degradation, such as logging and fire (Nepstad et al. 1999). Thus, it becomes crucial to perform studies on the diversity of organisms that remain living in such threatened areas to gather information about the impact of these events on the conservation status of South American biodiversity. In agreement with this concept, the present study investigated the genetic diversity and structure of *D. sturtevanti* populations that were collected in fragments of the remaining Atlantic Forest in South America.

Drosophila sturtevanti (*saltans* group, *sturtevanti* subgroup) is a Neotropical species that occurs in the Atlantic Forest biome. It is a forest-dwelling species found in fragments of this biome and has a broad geographic distribution across almost the entire range of the *saltans* group, from Mexico to southern Brazil, including the Caribbean islands. This species is a generalist and has a high abundance in all seasons (Magalhães 1962; Tidon-Sklorz and Sene 1992; Torres and Madi-Ravazzi 2006; Valadão et al. 2019), and some of its elements have been studied previously. For example, its populations present relatively high esterase polymorphism and polytene chromosome inversion frequencies among the species of the *saltans* group (Hosaki-Kobayashi and Bicudo 1997; Bernardo and Bicudo 2009). Studies on the reproductive isolation of the populations of this species are inconsistent. For example, incipient sexual isolation among the *D. sturtevanti* populations was detected (Dobzhansky 1944; Carareto 1994), but experiments using laboratory-raised and freshly collected strains from Mexico, Central America and Brazil did not find the same results (Hosaki-Kobayashi and Bicudo 1994). Nonetheless, morphological data analysis (aedeagus and wing morphometry) indicated that there are three different genetic sets of *D. sturtevanti* Brazilian populations: those from the Northeast, Southeast and South Regions (Segala 2019). This genetic differentiation, along with the wide distribution of *D. sturtevanti*, makes this species an important model for studies of the association between habitat fragmentation and the distribution of genetic variability.

Specimens in fragmented habitats, such as the *D. sturtevanti* in the Atlantic Forest, are frequently isolated and/or experience reduction in population size, triggering bottlenecks and genetic drift effects (*e.g.*, Goodman et al. 2001; Pascual et al. 2001; Broeck et al. 2017). The population isolation of *Drosophila* species, which have low dispersion (Wallace 1966; Spencer and Heed 1975; Markow and Castrezana 2000), could result in limited gene flow, increased interpopulation genetic differentiation, and reduced intrapopulation variability and effective population sizes. These events could, over time, make the population less resilient to environmental changes (Franknam 1997; Templeton et al. 2001). Population genetics studies are very helpful for understanding the history of a species, and most investigations of genetic variation have concentrated mainly on the (nearly) neutral genetic markers as gauges of the contemporary and historical processes affecting the maintenance and distribution of genetic variation.

Therefore, in this work, for the first time, a population study of *D. sturtevanti* was performed using 11 species-specific microsatellite markers. The objectives were to evaluate the populations' levels of genetic diversity, to verify the existence of population structure and to investigate whether genetic differentiation is associated with the fragmentation of the Atlantic Forest. Thus, three different groupings were analyzed: 1 - populations of the three geographic regions of Brazil (northeast, southeast and south), as suggested by previous morphological analyses; 2 - populations of two phytophysiognomies (dense ombrophilous forest and semideciduous seasonal forest); and 3 - a putative grouping from a Bayesian analysis.

MATERIALS AND METHODS

Drosophila sturtevanti samples

A total of 126 individuals from nine populations of *D. sturtevanti* were collected from two phytophysiognomy types of the Atlantic Forest of three Brazilian geographic regions (Table 1, Fig. 1).

© 2021 Academia Sinica, Taiwan

The collections were performed with the permission of the Instituto Chico Mendes de Conservação da Biodiversidade – ICMBio, a regulatory agency responsible for environmental studies in Brazil (permission number 46752). The representative specimens were stored in 100% ethanol in the collection of the Museu de Zoologia da Universidade de São Paulo, Brazil. All flies were captured using closed traps containing fermented banana bait (Penariol et al. 2008). The identification of males was carried out by analyzing the aedeagus using specific identification keys (Freire-Maia and Pavan 1949; Mourão and Bicudo 1967; Vilela

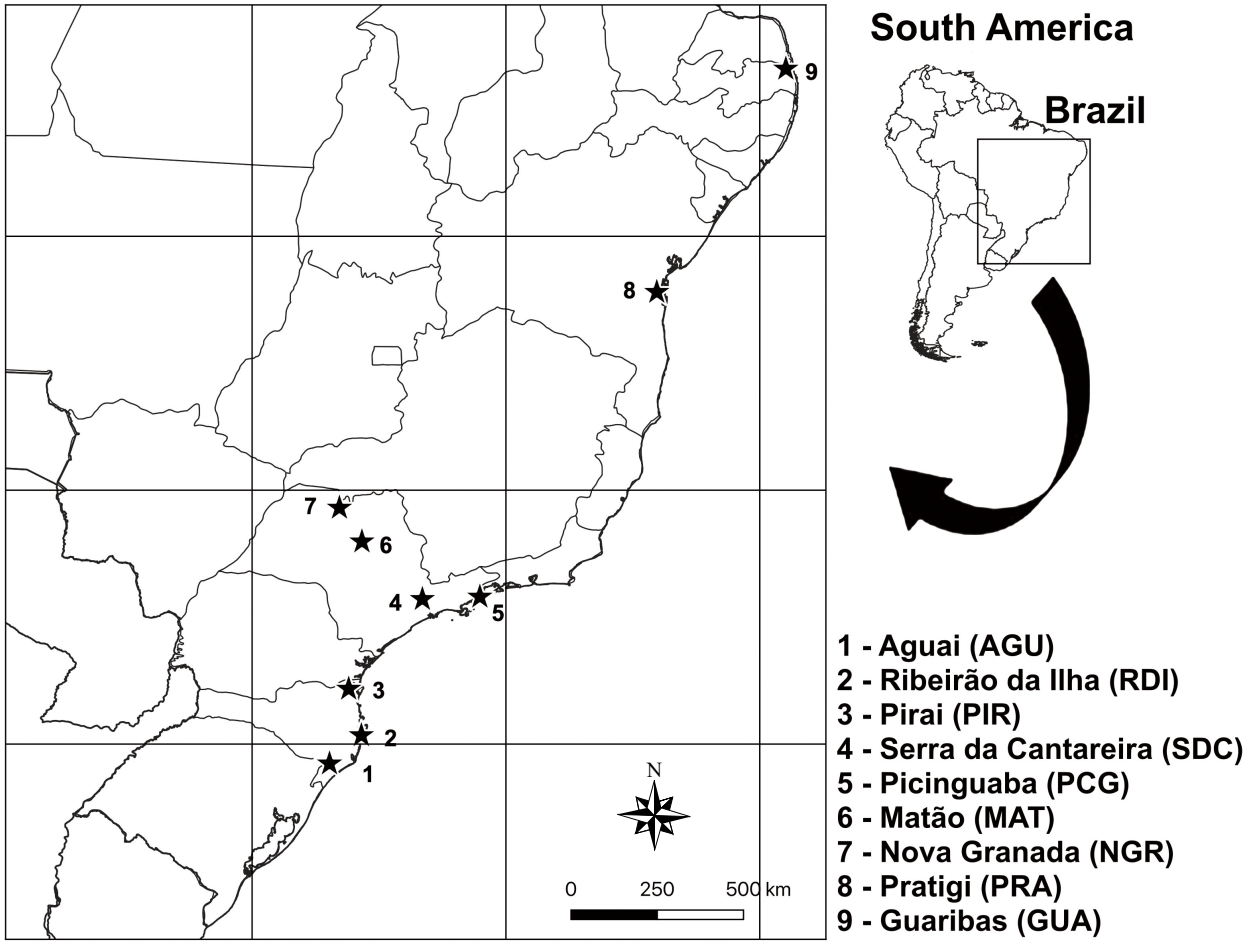

Fig. 1. Map of the Brazilian *Drosophila sturtevanti* populations from which samples were collected and analyzed.

Table 1. Geographic regions, phytophysiognomies and sizes of collection areas within the Brazilian Atlantic Forest biome of *Drosophila sturtevanti*

Geographic region	Collection site (Abbreviation)	Phytophysiognomies	Area size (ha)
South region	Aguaí (AGU)	DOF	7,652
	Ribeirão da Ilha (RDI)	DOF	10,000
	Piraí (PIR)	DOF	982.6
Southeast region	Serra da Cantareira (SDC)	DOF	64,800
	Picinguaba (PCG)	DOF	47,500
	Matão (MAT)	SSF	2,189
	Nova Granada (NGR)	SSF	1,359
Northeast region	Pratigi (PRA)	DOF	32,000
	Guaribas (GUA)	SSF	4,051

DOF = Dense Ombrophilous Forest; SSF = Semidecidual Seasonal Forest.

© 2021 Academia Sinica, Taiwan

and Bächli 1990; Souza et al. 2014). Samples were preserved in 96% ethanol and kept at 4°C until DNA extraction.

DNA extraction and microsatellite analyses

The extraction of genomic DNA was performed by individual maceration of each sampled male using the Promega kit. Eleven microsatellites (Dsturt_B, Dsturt_D, Dsturt_E, Dsturt_G, Dsturt_I, Dsturt_J, Dsturt_K, Dsturt_L, Dsturt_M, Dsturt_N and Dsturt_O) described for *D. sturtevanti* (Roman, B.E., Trava, B. M.; Madi-Ravazzi, L., submitted) were amplified through polymerase chain reaction (PCR) in a total volume of 25 µl, containing 0.5 µl of Taq DNA polymerase, 2.5 µl of 10x buffer, 0.2 mM dNTP, 0.2 pmol of each primer, 1.5 mM $MgCl_2$ and 3 ng of DNA. Touchdown PCR was performed for the Dsturt_B, Dsturt_D, Dsturt_I, Dsturt_K and Dsturt_M loci as follows: denaturation cycle at 94°C for 2 minutes; 2 repetitions of 10 cycles at 94°C for 1 minute, 65°C for 1 minute (-1°C per cycle), and 72°C for 2 minutes; ending with 18 cycles at 94°C for 1 minute, 55°C for 1 minute and 72°C for 5 minutes. For all other loci, a specific primer annealing temperature (Ta) was applied: 53°C - Dsturt_N; 55°C - Dsturt_G; 56°C - Dsturt_E and Dsturt_L; 57°C - Dsturt_O; 65°C - Dsturt_J. The PCR conditions for these loci were as follows: 94°C for 2 minutes; 30 cycles of 94°C for 1 minute, specific Ta for 1 minute and 72°C for 2 minutes; and ending with 72°C for 5 minutes. PCR amplification products were visualized in 6% polyacrylamide gel and stained with 15% silver nitrate (Sanguinetti et al. 1994).

Statistical analysis

Population genetic structure and diversities were assessed using 10–15 individuals from each of the nine populations. The mean number of alleles (N_a), effective number of alleles (N_e), expected heterozygosity (H_E), observed heterozygosity (H_O), number of private alleles (N_p), frequency of private alleles (A_p) and fixation index (F) were calculated for each population in GenAlEx software v.6.51b2 (Peakall and Smouse 2006 2012). The allelic polymorphic information content (*PIC*) was obtained using CERVUS software v.3.0.7 (Kalinowski et al. 2007). The mean allelic richness (A_r)—which is an unbiased measure of the number of alleles estimated independently of the sample size, allowing for comparison between different sample sizes (El Mousadik and Petit 1996)—was calculated in FSTAT v.2.93 (Goudet 2001). Departures from Hardy-Weinberg equilibrium (HWE) at each locus within populations were estimated in the GenAlEx software. All levels of significance were determined after a sequential Bonferroni correction for multiple tests (Holm 1979). The FreeNA software (Chapuis and Estoup 2007) was used to estimate null allele frequencies (A_n) for each locus and population following the Expectation Maximization (*EM*) algorithm of Dempster et al. (1977). A population bottleneck test, using adjusted frequencies for the presence of null alleles, was performed in BOTTLENECK software (Cornuet and Luikart 1996; Piry et al. 1999) to test the occurrence of recent demographic events. The program BOTTLENECK computed the distribution of the expected heterozygosity from the observed number of alleles when given the sample size under the assumption of mutation-drift equilibrium. The simulation of the coalescent process of n genes was performed under the two-phase model (TPM), using the Wilcoxon signed-rank test, setting the parameters as 90% single-step mutations, 10% multiple-step mutations, a variance of the geometric distribution of 12, and 1,000 iterations. These settings correspond to sensible parameter values for most microsatellites, considering that fewer than 20 loci were used (Cornuet and Luikart 1996; Piry et al. 1999).

The level of genetic differentiation among populations was verified using multiple approaches: indexes of genetic differentiation, a neighbor-joining tree, the use of a nonspatial Bayesian algorithm and analyses of molecular variance (AMOVA). The FreeNA software was used to estimate two genetic differentiation indexes based on the *ENA* method (Chapuis and Estoup 2007): D_C - Cavalli-Sforza and Edwards (1967) genetic distances, and F_{ST} (Weir 1996). For these indexes 95% confidence intervals (C.I.) were obtained using bootstrap resampling over loci. Wright (1978) qualitative classification of the genetic differentiation among populations (spatial analyses) was applied accordingly to the F_{ST} values obtained: 'low' (0–0.05), 'moderate' (0.05–0.15), 'high' (0.15–0.25) and 'very high' (> 0.25). The correlations between the genetic differentiation above (D_C and F_{ST}) and geographic distance by the Mantel tests (with 9999 random permutations) were performed in GenAlEx software. A neighbor-joining tree (Saitou and Nei 1987) was obtained through maximum likelihood analysis (Felsenstein 1981) on adjusted allele frequencies (considering the presence of null alleles) of microsatellite data using PHYLIP software (version 3.7a; Felsenstein 2009). To evaluate support for the branches, a bootstrap analysis (Felsenstein 1985) was performed 1000 times. An unrooted tree was constructed using Cavalli-Sforza and Edwards (1967) distances, using a branch and bound algorithm, the majority rule option and with random addition of populations. The cluster-based Bayesian method was performed with the software STRUCTURE v.2.3.4 (Pritchard et al. 2000),

© 2021 Academia Sinica, Taiwan

considering the presence of null alleles and the Dsturt_I locus as X or Y linked in the genotypic matrix. For this analysis, the admixture hypothesis was used assuming the existence of correlated allele frequencies, in which each sample is partially composed of the genome of each ancestral population. Together with the allele frequency model, they allow the log likelihood L (K) for the data to be obtained. This model is considered to be the most appropriate when the a priori origin and the degree of isolation of the studied populations are unknown (Pritchard et al. 2000). The prior probability, *i.e.*, the probability that an individual belongs to any reference K population, is defined as l/K. The K value was fixed from 1 to 11 using 10,000 burn-in, 500,000 Markov chain Monte Carlo (MCMC) replicates after burn-in and 25 iterations. After obtaining the results, a bar graph was generated with the CLUMPAK tool (Kopelman et al. 2015) using the best number of clusters (K) obtained by the STRUCTURE HARVESTER tool (Earl and Von Holdt 2012), according to the Evanno test (Evanno et al. 2005; Earl and Von Holdt 2012).

The AMOVA was run in the Arlequin software (version 3.5.2.2, Excoffier and Lischer 2010), using the adjusted genetic frequencies and 1,000 permutations, to test four assumptions about the distribution of genetic variability of Brazilian *D. sturtevanti* populations: 1) with no grouping; 2) grouping according to geographic regions (Northeast, Southeast and South regions); 3) grouping according to Atlantic Forest phytophysiognomies (dense ombrophilous forest and semideciduous seasonal forest); and 4) groupings according to the result of the cluster-based Bayesian method performed with STRUCTURE software. The second grouping was proposed based on a previous morphological study (Segala 2019) using *D. sturtevanti*, (Segala 2019) using *D. sturtevanti* and also to test correlation between genetic variability distribution and isolation by distance. The third grouping was tested based on phylogeography and microsatellite studies with *D. ornatifrons* (Gustani et al. 2015; Zorzato 2015), another Neotropical forest dweeling species, which suggested that there is correlation between genetic variability and Atlantic Forest phytophysionomies. In the absense of a genetic structure, slightly negative variation can be obtained only by chance, because the true value estimated is zero (Huang et al. 2021).

RESULTS

Genetic diversity analysis

The variability analysis of 11 microsatellite loci specific to *D. sturtevanti*, using 126 males from nine populations, resulted in a total of 109 alleles (Table S1). The number of alleles per locus ranged from 2 (Dsturt_B) to 16 (Dsturt_N), with an average (± standard error – S.E.) of 9.91 ± 1.30 (Table S2). Almost all loci were polymorphic, except Dsturt_B and Dsturt_D. These loci presented the lowest total numbers of alleles (2 and 3, respectively), were monomorphic in all populations except PIR (Table S1), and showed the lowest genetic diversity values (Supplementary Table S2). The Dsturt_I locus was the only one that did not show heterozygotes (H_O = 0 and F = 1.00, Table S2) despite a total of 103 males were analyzed. This indicates that this locus is most likely X- or Y-linked. The other microsatellite loci were assumed to be autosomal because all population samples showed heterozygotes.

In table S2, all the data for each locus concerning the total and mean number of alleles, polymorphism information content, mean effective number of alleles, mean observed heterozygosity, mean expected heterozygosity, fixation index and null allele frequency were reported. Among the polymorphic loci, PIC ranged from 0.815 (Dsturt_J) to 0.907 (Dsturt_N), indicating that these loci are informative for population analysis. The mean H_O (± S.E.) across loci ranged from 0.15 ± 0.04 (Dsturt_N) to 0.68 ± 0.07 (Dsturt_O), while H_E ranged from 0.73 ± 0.03 (Dsturt_E and Dsturt_J) to 0.81 ± 0.02 (Dsturt_N). Null allele frequencies ranged from 0.0774 (Dsturt_O) to 0.3682 (Dsturt_N). The Fixation index (*F*) across polymorphic loci resulted in positive values, with the Dsturt_O locus showing the lowest (0.12 ± 0.08) and the Dsturt_N locus showing the highest value (0.82 ± 0.05).

Across populations, the total mean H_O was 0.43 ± 0.03, ranging from 0.22 ± 0.06 (GUA) to 0.39 (± 0.08 in PRA and ± 0.09 in SDC) (Table 2). Among polymorphic loci, Dsturt_O was the only locus that did not show significant deviation from Hardy-Weinberg Equilibrium (HWE) in all samples, and Dsturt_N showed deviations in all populations. The other seven polymorphic loci showed significant deviations in at least four populations, and all populations presented heterozygote deficiency. *F* values ranged from 0.33 ± 0.11 in SDC to 0.65 ± 0.08 in GUA, with a high overall mean (0.44 ± 0.04; Table S2). AGU and MAT were the populations with the highest number of loci with HWE deviations (7 - Dsturt_E, Dsturt_G, Dsturt_I, Dsturt_J, Dsturt_L, Dsturt_M and Dsturt_N), and PCG was the one with the lowest (4 – Dsturt_K, Dsturt_L, Dsturt_M and Dsturt_N). Allelic richness ranged from 3.32 in RDI to 4.10 in PRA. The frequency of null alleles ranged from 0.14 ± 0.04 (PCG) to 0.22 ± 0.04 (GUA). All populations, except PCG, presented private alleles, ranging from 1 (AGU, RDI, SDC, NGR and GUA) to 3 (MAT), and frequencies from 0.03 (SDC and NGR) to 0.20 (RDI).

© 2021 Academia Sinica, Taiwan

The results from individual population analysis using the TPM model showed that none had recent bottleneck events, defined by significant heterozygote excess.

Population Genetic Structure

The Bayesian analysis of the population structure that was conducted using STRUCTURE software indicated the existence of six genetic clusters (K = 6). The interpretation of the distribution of this genetic variability allowed the identification of two groups of populations: RDI and the other seven populations (Fig. 2). These results were used to estimate the molecular variability distribution (AMOVA), summarized in table 3. When no grouping was tested, almost 97% of the variation occurred among individuals within populations, while only 3% of total variation was detected among populations. In the other three different sets of grouped populations (geographical regions, phytophysiognomies and cluster based Bayesian method obtained above), the variations among regions were negative, among population within regions were around 4%, and the variation within populations were all the highest.

The D_C and F_{ST} indexes showed RDI to be the most genetic differentiated population (Table 4). The lower significant differentiation was between MAT and NGR (D_C = 0.3143; F_{ST} = 0.03). All pairwise D_C and F_{ST} comparisons were significantly different from zero. The overall F_{ST} result showed moderate differentiation among populations (0.0674; 95% C.I. = 0.0533–0.0808). However, no correlations between the genetic (D_C and F_{ST}) and geographic distances of the population pairs were obtained by the Mantel tests (r = 0.142, p = 0.284; r = -0.023, p = 0.569, respectively).

The neighbor-joining tree showed only one node with higher bootstrap support (> 50%), grouping MAT and NGR. GUA and PCG were the most differentiated populations (Fig. 3).

DISCUSSION

The nine populations studied showed a high number of polymorphisms and similar allelic diversity for the analyzed polymorphic microsatellite loci. These results are consistent with those from other genetic markers for this species, including chromosomal inversions (Hosaki-Kobayashi and Bicudo 1997) and esterase loci (Bernardo and Bicudo 2009), which

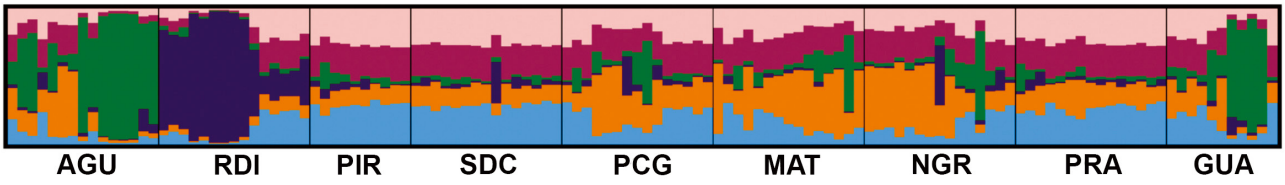

Fig. 2. Genetic structure analysis of nine Brazilian *Drosophila sturtevanti* populations using the Bayesian algorithm as implemented in the program STRUCTURE. Six clusters (K) were obtained (different colors). Population abbreviations are described in the Materials and Methods section.

Table 2. Values of genetic diversity (± standard error) across nine populations of *Drosophila sturtevanti* from the Atlantic Forest biome in Brazil using 11 species-specific microsatellite loci. N_a – mean number of alleles; N_e – mean effective number of alleles; A_r – mean allele richness; N_p – number of private alleles; A_p – mean frequency of private alleles; H_O – mean observed heterozygosity; H_E – mean expected heterozygosity; A_n – mean frequency of null alleles; F – fixation index; HWE – list of loci with departure of Hardy-Weinberg expectations

Populations	N_a	N_e	A_r	N_p	A_p	H_O	H_E	A_n	F	HWE
AGU	5.09 ± 0.62	3.54 ± 0.44	3.60	1	0.07	0.27 ± 0.08	0.61 ± 0.09	0.20 ± 0.05	0.51 ± 0.12	E, G, I, J, L, M, N
RDI	4.64 ± 0.88	3.40 ± 0.62	3.32	1	0.20	0.28 ± 0.09	0.56 ± 0.09	0.18 ± 0.04	0.54 ± 0.11	G, I, J, K, L, N
PIR	4.64 ± 0.47	3.42 ± 0.40	3.58	2	0.10	0.34 ± 0.09	0.63 ± 0.07	0.18 ± 0.05	0.37 ± 0.13	I, J, K, M, N
SDC	6.00 ± 0.80	4.15 ± 0.58	3.94	1	0.03	0.39 ± 0.09	0.65 ± 0.09	0.15 ± 0.05	0.33 ± 0.11	E, I, J, K, N
PCG	5.54 ± 0.83	3.75 ± 0.56	3.64	0	-	0.37 ± 0.08	0.61 ± 0.09	0.14 ± 0.04	0.38 ± 0.09	K, L, M, N
MAT	5.73 ± 0.62	4.04 ± 0.53	3.88	3	0.08	0.32 ± 0.08	0.65 ± 0.09	0.19 ± 0.05	0.41 ± 0.11	E, G, I, J, L, M, N
NGR	6.09 ± 0.77	4.15 ± 0.58	3.94	1	0.03	0.31 ± 0.07	0.64 ± 0.09	0.19 ± 0.05	0.46 ± 0.10	E, I, L, M, N
PRA	6.18 ± 0.71	4.37 ± 0.60	4.10	2	0.07	0.39 ± 0.08	0.66 ± 0.09	0.15 ± 0.04	0.34 ± 0.10	G, I, K, M, N
GUA	4.91 ± 0.79	3.75 ± 0.62	3.68	1	0.15	0.22 ± 0.06	0.60 ± 0.09	0.22 ± 0.04	0.65 ± 0.08	E, J, K, L, N

E = Dsturt_E; G = Dsturt_G; I = Dsturt_I; J = Dsturt_J; K = Dsturt_K; L = Dsturt_L; M = Dsturt_M; N = Dsturt_N.

© 2021 Academia Sinica, Taiwan

indicated that *D. sturtevanti* has more polymorphisms than any other species in the *saltans* group.

The level of heterozygote deficiency in the *D. sturtevanti* populations was high. Among the events that could generate this result, such as the null alleles, the Wahlund effect, inbreeding or demographic processes (Shoemaker and Jaenike 1997; Moraes and Sene 2002; Hurtado et al. 2004; Markow and O'Grady 2008), our data allowed us to discuss the effect of the first, null allele. Their estimated frequencies were all above 0.14, higher than those detected in a Neotropical cactophilic species of *Drosophila* (Machado et al. 2010), to which

Table 3. Analyses of Molecular Variance (AMOVA) for *Drosophila sturtevanti* populations, after correction for the presence of null alleles, without and assuming three different populational groupings (geographical regions, phytophysiognomies and after STRUCTURE analysis). DF = degrees of freedom; SS = sum of squares; VC = variance component; PV (%) = percentage of total variation

Source	DF	SS	VC	PV (%)
		Without grouping		
Among populations	8	0.092	0.00025	3.01
Within populations	111	0.900	0.00811	96.99
Total	119	0.992	0.00836	100
		Geographical regions		
Among regions	2	0.017	-0.00010	-1.17
Among populations within regions	6	0.074	0.00032	3.86
Within populations	111	0.900	0.00811	97.32
Total	119	0.992	0.00833	100
		Phytophysiognomies		
Among regions	1	0.004	-0.00016	-1.95%
Among populations within regions	7	0.088	0.00033	4.01%
Within populations	111	0.900	0.00811	97.95%
Total	119	0.992	0.00828	100%
		STRUCTURE		
Among regions	1	0.001	-0.00047	-5.88%
Among populations within regions	7	0.090	0.00037	4.59%
Within populations	111	0.900	0.00811	101.29%
Total	119	0.992	0.00801	100%

Table 4. Pairwise Cavalli-Sforza and Edwards (1967) genetic distances D_C (above diagonal) and Weir (1996) F_{ST} (below diagonal) for *Drosophila sturtevanti* populations, after correction for the presence of null alleles. Shaded numbers represent the highest (above 0.45 for D_C and 0.1 for F_{ST}), and numbers inside boxes represent the lowest values obtained

	AGU	RDI	PIR	SDC	PCG	MAT	NGR	PRA	GUA
AGU	***	0.4506	0.4310	0.4231	0.4038	0.3687	0.3522	0.3887	0.4045
RDI	0.1178	***	0.4842	0.4395	0.4458	0.4819	0.4321	0.4370	0.4315
PIR	0.0858	0.1104	***	0.4094	0.4370	0.4801	0.4546	0.4442	0.5027
SDC	0.0729	0.0870	0.0419	***	0.3665	0.3935	0.3778	0.3583	0.4295
PCG	0.0750	0.1119	0.0723	0.0523	***	0.3651	0.3379	0.3402	0.3712
MAT	0.0416	0.1219	0.0810	0.0506	0.0494	***	0.3143	0.3513	0.4539
NGR	0.0513	0.1129	0.0877	0.0582	0.0365	0.0300	***	0.3393	0.3884
PRA	0.0432	0.0929	0.0648	0.0414	0.0319	0.0448	0.0421	***	0.4276
GUA	0.0642	0.0907	0.0958	0.0608	0.0509	0.0756	0.0497	0.0427	***

© 2021 Academia Sinica, Taiwan

no effect on the HWE or population differentiation was observed. However, a study of the impact of null alleles in the Maire yew, a tree from southern China (Wu et al. 2019), that depicts null allele frequencies similar to those obtained here (0.159–0.331) showed that null alleles can have significant effects on some population genetic parameters, especially H_O, F and F_{ST}. Therefore, we considered that they could be one of the main generators of the high deficiency and should not be neglected.

Microsatellite diversity detected in *D. sturtevanti* populations (H_O = 0.43 ± 0.03) was higher than the mean diversity of the Brazilian populations of the endemic *Drosophila* species from the Atlantic Forest and open xerophytic areas (approximately 0.36 – Moraes and Sene 2007; Machado et al. 2010; Silva et al. 2015; Zorzato 2015). The highest H_O values were detected in SDC and PRA, which correspond to two largest areas of collection. The first area, SDC, is located in the Paulistano Plateau and has the largest size (64,800 ha) and highest altitude (1,215 m) of the areas. The SDC region covers São Paulo, Guarulhos, Mairiporã and Caieiras municipalities, and has great ecological importance because it is the Biosphere Reserve of the Green Belt of the city of São Paulo, which comprises 7,916.52 hectares of the Atlantic Forest. The second area, PRA, is one of the Atlantic Forest conservation regions in the Pratigi Environmental Protection Area (EPA). This EPA is located on the southern coast of Bahia state, contains 85,686 ha and is an important federal conservation unit that protects the biodiversity and hydric resources associated with the Atlantic Forest in the region. The higher genetic diversity of these two populations may reflect the ecological and climatic characteristics of these regions, with high humidity favoring the occurrence of *D. sturtevanti.*

The GUA population presented the lowest microsatellite diversity (H_O = 0.22 ± 0.06) and higher distance in neighbor-joining tree. This population was collected in the Guaribas Biological Reserve, which is located on the border of the Caatinga and is one of the largest environmental preservation areas of Atlantic Forest (IBAMA 2003). However, several episodes of fire around and inside this reservation were detected, and four occurred in GUA within the five-year period of 2007–2012, (Alencar 2014). Certainly, these incidents could have had an impact on the local fauna, including *D. sturtevanti*, and they could be the cause of the observed low diversity in GUA. However, this impact was long enough ago to not be detected by the bottleneck analysis, which only identifies a recent event, when allelic diversity was still lower than the heterozygosity (Cornuet and Luikart 1996; Piry et al. 1999).

The analysis of all population differentiation approaches resulted in some patterns that could be

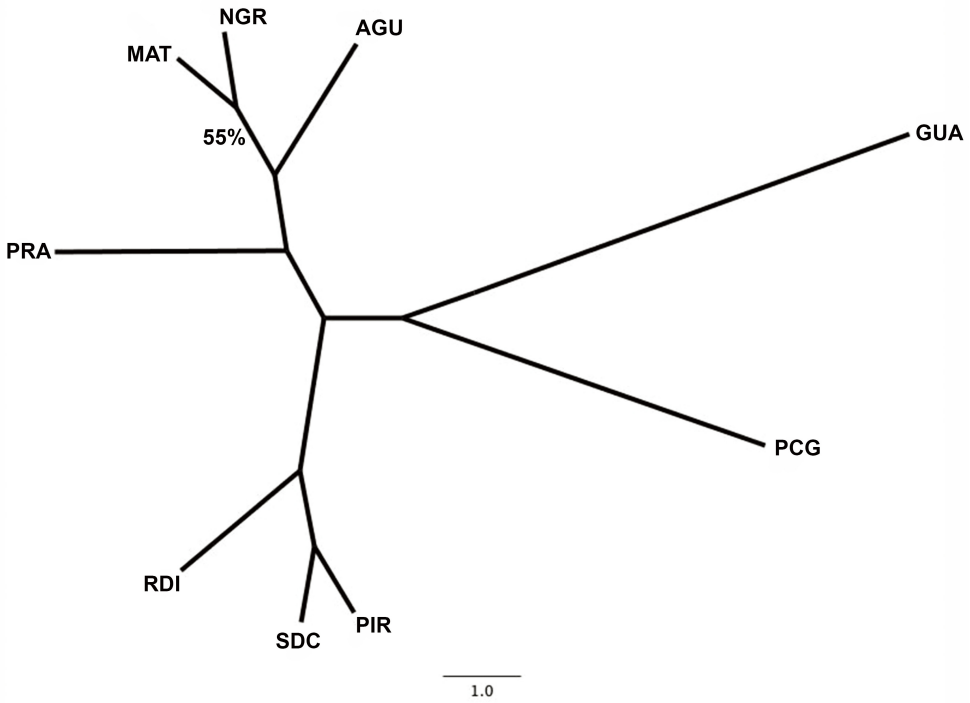

Fig. 3. Neighbor-joining tree obtained for the nine populations of *Drosophila sturtevanti* using Cavalli-Sforza and Edwards (1967) distances. Abbreviations are described in the Materials and Methods section. Node number refer to bootstrap values (only higher than 50% is shown).

© 2021 Academia Sinica, Taiwan

highlighted. The low molecular variation (3% among populations and negative among groups), the absence of genetic and geographic correlations (Mantel tests) and the moderate F_{ST} indicated that *D. sturtevanti* is not strongly structured, which is in disagreement with previous morphological data (Segala 2019) that distinguished South, Southeast and Northeast populations. This discrepancy is probably related to the nature of the markers while morphological characters should be under strong natural selection, microsatellite loci are considered nearly neutral. It could also be addressed that the microsatellite loci studied are not suffering hitchhiking effect of selection on genes that determine morphological characters analyzed.

The genetic differentiation (F_{ST} and D_C) and Bayesian analyses pointed out the higher isolation of RDI from the other populations. RDI is an insular population; thus, its differentiation could be explained by geographic isolation from the other continental populations and/or the consequences of the founder effect. In the first case, at the last glacial maximum (LGM) in the Pleistocene, the ocean level was between 80 and 130 meters lower than the current level (Suguio 2008), and RDI, which is today a separate island, could have been part of the continental forest at that time. The isolation may have arisen after the LGM, in the Holocene, as a consequence of sea level rise. Gustani et al. (2015) also showed that historical events involving sea level regressions/transgressions were important to the evolution of another Atlantic Forest *Drosophila* species, *D. ornatifrons*. Currently, gene flow could be restricted by continental/island isolation. The work of Tait et al. (2017), using microsatellites, on the recent invasion of *D. suzuki* in Italy is an example of higher restriction of gene flow among continental and insular populations. In the second case, considering the limited dispersion capacity of drosophilids (Spencer and Heed 1975; Markow and Castrezana 2000), some of the members of the continental *D. sturtevanti* population could have colonized the island by anthropic action, such as being carried by some hosts. The colonization of islands by *Drosophila* specimens via human activities is fairly common (Louis and David 1986; Lachaise and Silvain 2004; Jones 2005; Legrand et al. 2009 2011). For example, isolation by distance migration pattern of *D. sechellia* within the Seychelles archipelago is most likely explained by local genetic exchanges between neighboring populations due to the association of this species with the host plant *Morinda citrifolia*, which is used by humans medicinally and in fishing activities (Legrand et al. 2011).

Another point to highlight is the higher genetic proximity, accordingly to the F_{ST}, D_C and neighbor-joining, between NGR and MAT, the populations from the interior of São Paulo state that have the smallest distance between them, 140 km. Considering the hypervariability inherent to the microsatellite markers, the detection of genetic proximity between populations indicates that different combinations of alleles arose in each generation and were sorted over geographical space, according to the isolation by distance model (Sunnucks 2000). The presence of ecological corridors could be allowing genetic interchange between these populations, maintaining their genetic similarity.

Batista et al. (2018) studied six populations from the Southeast region of Brazil of *D. mediopunctata*, an almost exclusively forest-dwelling Neotropical species that belongs to the *tripunctata* group of the subgenus *Drosophila* (Vilela 1992; Yotoko et al. 2003; Hatadani et al. 2009; TaxoDros 2020) and detected low and moderate levels of population genetic structure using microsatellites and chromosome inversion polymorphisms, respectively. The authors proposed that the divergence among populations of this species might be a result of the association of climatic and geomorphological properties of the collection regions, rather than with forest fragmentation. In the case of the *D. sturtevanti* populations sampled in this work, considering the low dispersion ability of *Drosophila* (Spencer and Heed 1975; Markow and Castrezana 2000) and the fragmented and geographic isolated population distribution, the absence of a significant structure observed should be due to ancestral shared polymorphisms and/or recurrent mutations.

CONCLUSIONS

The characterization of *Drosophila sturtevanti* based on the analyses of species-specific microsatellites showed to be important as, for the first time, the genetic variability distribution of a Neotropical forest-dwelling *Drosophila* species was studied using a wide population distribution approach. This work allowed the detection of moderate genetic differentiation among its populations, without general correlation with geographic and/or Atlantic Forest phytophysiognomy distributions. The weak genetic structure among fragmented and geographic isolated populations is likely to be a consequence of ancestral shared polymorphisms and/or recurrent mutations; the similarity between MAT and NGR, geographically closest populations, might be the result of isolation by distance with gene flow provided by ecological corridors; and the higher differentiation of RDI, an insular population, could be due to its isolation after the last maximum glacial, or more recent founder effect via island colonization in association of hosts carried by anthropic action. A phylogeographical study

© 2021 Academia Sinica, Taiwan

would complement and help to clarify and distinguish the roles of the historical and ecological processes in the evolution of *D. sturtevanti*.

Acknowledgments: We are grateful to Dr. Daniela De Toni, Hermes José Schmitz and Claudia Rodhe for providing some strains of *D. sturtevanti.* We thank the members of the Laboratory of Genetics, Ecology and Evolution of *Drosophila* (IBILCE/UNESP) for collaborating in the collections and the doctoral students Samara Videira Zorzato and Cláudia Gobatto for collaborating with some analyses. We are also grateful to the following funding agencies: FAPESP – Fundação de Amparo à Pesquisa do Estado de São Paulo (processes number 2014/14059-0, 2015/17579-7, 2017/05344-0) and CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (PhD scholarship to BMT). These funding agencies were essential for financing sample collections and laboratory procedures. We finally thank American Journal Experts for English language revision (Invoice #SPBD8B5DL).

Authors' contributions: BMT and LMR collected the specimens; LMR designed the research; BMT, RPM and LPBM generated and analyzed the results and BMT, RPM, LPBM and LMR interpreted the results and wrote the paper. All authors contributed, read and approved the final manuscript.

Competing interests: The authors declare no competing interests.

Availability of data and materials: The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Consent for publication: All authors agree with the content of this article and with it publication as submitted.

Ethical approval consent to participate: All authors are in agreement.

REFERENCES

Achard F, Eva HD, Stibig HJ, Mayaux P, Gallego J, Richards T, Malingreau JP. 2002. Determination of deforestation rates in the world's humid tropical forests. Science **297:**999–1002. doi:10.1126/science.1070656.

Alencar HMQ. 2014. Análise dos incêndios combatidos em remanescentes de vegetação nativa no litoral norte da Paraíba. Undergraduate Dissertation, Universidade Federal da Paraíba.

Batista MRD, Penha RES, Sofia SH, Klaczko LB. 2018. Comparative analysis of adaptive and neutral markers of *Drosophila mediopunctata* populations dispersed among forest fragments. Ecol Evol **8:**12681–12693. doi:10.1002/ece3.4696.

Bernardo AA, Bicudo HEMC. 2009. Variability of esterase patterns in adult flies of the *saltans* species group of *Drosophila* (subgenus Sophophora). Genetics **137:**111–124. doi:10.1007/s10709-009-9357-z.

Broeck AV, Maes D, Kelager A, Wynhoff I, WallisDeVries MF, Nash DR, Oostermeijer JGB, Dyck HV, Mergeay J. 2017. Gene flow and effective population sizes of the butterfly *Maculinea alcon* in a highly fragmented, anthropogenic landscape. Biol Conserv **209:**89–97. doi:10.1016/j.biocon.2017.02.001.

Carareto CMA. 1994. Incipient ethological isolation among strains of *Drosophila sturtevanti*. Naturalia **19:**185–198.

Cavalli-Sforza LL, Edwards AWF. 1967. Phylogenetic analysis: models and estimation procedures. Evolution **21:**550–570. doi:10.1111/j.1558-5646.1967.tb03411.x.

Chapuis MP, Estoup A. 2007. Microsatellite null alleles and estimation of population differentiation. Mol Biol Evol **24:**621–631. doi:10.1093/molbev/msl191.

Cornuet JM, Luikart G. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics **144:**2001–2014. doi:10.1093/genetics/144.4.2001.

Dempster AP, Laird NM, Rubin DB. 1977. Maximum likelihood from incomplete data via the EM algorithm. J Roy Stat Soc B **39:**1–38. doi:10.1111/j.2517-6161.1977.tb01600.x.

Dobzhansky T. 1944. Experiments on sexual isolation in *Drosophila*. III. Geographic strains of *Drosophila sturtevanti.* Proc Natl Acad Sci USA **30:**335–339. doi:10.1073/pnas.30.11.335.

Earl DA, Von Holdt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour **4:**359–361. doi:10.1007/s12686-011-9548-7.

El Mousadik A, Petit RJ. 1996. High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. Theor Appl Genet **92:**832–839. doi:10.1007/BF00221895.

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

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

Felsenstein J. 1981. Evolutionary trees from gene frequencies and quantitative characters: finding maximum likelihood estimates. Evolution **35:**1229–1242. doi:10.1111/j.1558-5646.1981.tb04991.x.

Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39:**783–791. doi:10.1111/j.1558-5646.1985.tb00420.x.

Felsenstein J. 2009. PHYLIP (Phylogeny Inference Package) version 3.7a. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle, USA.

Franknam R. 1997. Do island populations have less genetic variation than mainland populations? Heredity **78:**311–327. doi:10.1038/hdy.1997.46.

Freire-Maia N, Pavan C. 1949. Introdução ao estudo de *Drosophila*. Cultus **1:**5–71.

Galindo-Leal C, Câmara IG. 2003. The Atlantic Forest of South America: Biodiversity Status, Threats and Outlook. Island Press, Washington.

© 2021 Academia Sinica, Taiwan

Goodman SJ, Tamate HB, Wilson R, Nagata J, Tatsuzawa S, Swanson GM, Pemberton JM, McCullough D. 2001. Bottlenecks, drift and differentiation: the population structure and demographic history of sika deer (*Cervus nippon*) in the Japanese archipelago. Mol Ecol **10:**1357–1370. doi:10.1046/j.1365-294X.2001.01277.x.

Goudet J. 2001. FSTAT, A Program to Estimate and Test Gene Diversities and Fixation Indices, version 2.9.3. Available at: https://www2.unil.ch/popgen/softwares/fstat.htm. Accessed 27 May 2020.

Grau HR, Gasparri NI, Aide TM. 2008. Land use efficiency: balancing food production and nature conservation in neotropical dry forests of NW Argentina. Global Change Biol **14:**985–997. doi:10.1111/j.1365-2486.2008.01554.x.

Gustani EC, Oliveira APF, Santos MH, Machado LPB, Mateus RP. 2015. Demographic structure and evolutionary history of *Drosophila ornatifrons* (Diptera, Drosophilidae) from Atlantic Forest of Southern Brazil. Zoolog Sci **32:**141–150. doi:10.2108/zs140062.

Hatadani LM, McInerney JO, Medeiros HF, Junqueira ACM, Azeredo-Espin AML, Klaczko LB. 2009. Molecular phylogeny of the *Drosophila tripunctata* and closely related species groups (Diptera: Drosophilidae). Mol Phylogenet Evol **51:**595–600. doi:10.1016/j.ympev.2009.02.022.

Hecht SB. 1993. The logic of livestock and deforestation in Amazonia. Bioscience **43:**687–695. doi:10.2307/1312340.

Holm S. 1979. A simple sequentially rejective multiple test procedure. Scand Stat Theory App **6:**65–70.

Hosaki-Kobayashi MK, Bicudo HEMC. 1994. Productivity in mass crosses of *Drosophila sturtevanti*: a comparative study of laboratory stocks and recently collected flies. Rev Bras Genet **17:**393–399.

Hosaki-Kobayashi MK, Bicudo HEMC. 1997. Inversion polymorphism in laboratory strains and natural samples of *Drosophila sturtevanti* (*saltans* group, *sturtevanti* subgroup). Cytobios **89:**7–20.

Huang K, Wang T, Dunn DW, Zhang P, Sun H, Li B. 2021. A generalized framework for AMOVA with multiple hierarchies and ploidies. Integr Zool **16:**33–52. doi:10.1111/1749-4877.12460.

Hurtado LA, Erez T, Castrezana S, Markow TA. 2004. Contrasting population genetic patterns and evolutionary histories among sympatric Sonoran Desert cactophilic *Drosophila*. Mol Ecol **13:**1365–1375. doi:10.1111/j.1365-294X.2004.02169.x.

IBAMA. 2003. Plano de Manejo, Fase 2, da Reserva Biológica Guaribas. Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis. Available via IBAMA website. https://www.icmbio.gov.br/portal/images/stories/imgs-unidades-coservacao/rebio_guaribas.pdf. Accessed 21 May 2017.

Jones CD. 2005. The genetics of adaptation in *Drosophila sechellia*. Genetics **123:**137–145. doi:10.1007/s10709-004-2728-6.

Kalinowski ST, Taper ML, Marshall TC. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol Ecol **16:**1099–1106. doi:10.1111/j.1365-294X.2007.03089.x.

Klink CA, Machado RB. 2005. Conservation of the Brazilian Cerrado. Conserv Biol **19:**707–713. doi:10.1111/j.1523-1739.2005.00702.x.

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NN, Mayrose I. 2015. CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. Mol Ecol Resour **15:**1179–1191. doi:10.1111/1755-0998.12387.

Lachaise D, Silvain JF. 2004. How two Afrotropical endemics made two cosmopolitan human commensals: the *Drosophila melanogaster* - *D. simulans* paleogeographic riddle. Genetica **120:**17–39. doi:10.1023/B:GENE.0000017627.27537.ef.

Legrand D, Tenaillon MI, Matyot P, Gerlach J, Lachaise D, Cariou ML. 2009. Species-wide genetic variation and demographic history of *Drosophila sechellia*, a species lacking population structure. Genetics **182:**1197–1206. doi:10.1534/genetics.108.092080.

Legrand D, Vautrin D, Lachaise D, Cariou ML. 2011. Microsatellite variation suggests a recent fine-scale population structure of *Drosophila sechellia*, a species endemic of the Seychelles archipelago. Genetica **139:**909–919. doi:10.1007/s10709-011-9595-8.

Louis J, David JR. 1986. Ecological specialization in the *Drosophila melanogaster* species subgroup: a case study of *D. sechellia*. Acta Oecol **7:**215–229.

Machado LPB, Mateus RP, Sene FM, Manfrin MH. 2010. Microsatellite allele sequencing in population analyses of the South American cactophilic species *Drosophila antonietae* (Diptera: Drosophilidae). Biol J Linn Soc **100:**573–584. doi:10.1111/j.1095-8312.2010.01441.x.

Magalhães LE. 1962. Notes on the taxonomy, morphology and distribution of *saltans* group of *Drosophila*, with description of four new species. Univ Texas Publs **6205:**135–154.

Markow TA, Castrezana S. 2000. Dispersal in cactophilic *Drosophila*. Oikos **89:**378–386. doi:10.1034/j.1600-0706.2000.890219.x.

Markow TA, O'Grady P. 2008. Reproductive ecology of *Drosophila*. Funct Ecol **22:**747–759. doi:10.1111/j.1365-2435.2008.01457.x.

Marques JS. 2016. Epífitas vasculares em diferentes estágios sucessionais de um fragmento urbano de Floresta Estacional Semidecidual. Dissertation, Universidade Federal de Juiz de Fora.

Mittermeier RA, Gil PR, Hofmann M, Pilgrim J, Brooks T, Mittermeier CG, Lamoreaux J, Fonseca GAB. 2004. Hotspots Revisited: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions.: Cemex - Conservation International and Agrupacion Sierra Madre, Monterrey, Mexico.

Moraes EM, Sene FM. 2002. Breeding structure of an isolated cactophilic *Drosophila* population on a sandstone table hill. J Zool Syst EvolRes **40:**123–128. doi:10.1046/j.1439-0469.2002.00186.x.

Moraes EM, Sene FM. 2007. Microsatellite and morphometric variation in *Drosophila gouveai*: the relative importance of historical and current factors in shaping the genetic population structure. J Zoolog Syst Evol Res **45:**336–344. doi:10.1111/j.1439-0469.2007.00411.x.

Mourão CA, Bicudo HEMC. 1967. Duas novas espécies de *Drosophila* do grupo "*saltans*" (Drosophilidae, Diptera). Pap Avulsos Zool **20:**123–134.

Nepstad DC, Verssimo A, Alencar A, Nobre C, Lima E, Lefebvre P, Schlesinger P, Potter C, Moutinho P, Mendoza E, Cochrane M, Brooks V. 1999. Large-scale impoverishment of Amazonian forests by logging and fire. Nature **398:**505–508. doi:10.1038/19066.

Pascual M, Aquadro CF, Soto V, Serra L. 2001. Microsatellite variation in colonizing and palearctic populations of *Drosophila subobscura*. Mol Biol Evol **18:**731–740. doi:10.1093/oxfordjournals.molbev.a003855.

Peakall R, Smouse PE. 2006. GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes **6:**288–295. doi:10.1111/j.1471-8286.2005.01155.x.

Peakall R, Smouse PE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. Bioinformatics **28:**2537–2539. doi:10.1093/bioinformatics/bts460.

© 2021 Academia Sinica, Taiwan

Penariol L, Bicudo HEMC, Madi-Ravazzi L. 2008. On the use of open or closed traps in the capture of drosophilids. Biota Neotrop **8:**47–51. doi:10.1590/S1676-06032008000200004.

Piry S, Luikart G, Cornuet JM. 1999. BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. J Heredity **90:**502–503. doi:10.1093/jhered/90.4.502.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics **155:**945–959. doi:10.1093/genetics/155.2.945.

Ramankutty N, Foley JA. 1999. Estimating historical changes in global land cover: croplands from 1700 to 1992. Global Biogeochem Cycles **13:**997–1027. doi:10.1029/1999GB900046.

Saitou N, Nei M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic tree. Mol Biol Evol **4:**406–425. doi:10.1093/oxfordjournals.molbev.a040454.

Salgueiro F, Felix D, Caldas JF, Margis-Pinheiro M, Margis R. 2004. Even population differentiation for maternal and biparental gene markers in *Eugenia uniflora*, a widely distributed species from the Brazilian coastal Atlantic rain forest. Divers Distrib **10:**201–210. doi:10.1111/j.1366-9516.2004.00078.x.

Sanguinetti C, Dias Neto E, Simpson AJG. 1994. Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. Biotechniques **17:**209–214.

Segala LF. 2019. Diferenciação populacional em *Drosophila sturtevanti* (subgrupo *sturtevanti*, grupo *saltans*) avaliada por morfometria geométrica da asa e do edeago. PhD dissertation, Universidade Estadual Paulista, Brazil.

Shoemaker DD, Jaenike J. 1997. Habitat continuity and the genetic structure of *Drosophila* populations. Evolution **51:**1326–1332. doi:10.1111/j.1558-5646.1997.tb03981.x.

Silva DC, Machado LPB, Mateus RP. 2015. Migration rate and genetic diversity of two *Drosophila maculifrons* (Duda, 1927) populations from Highland Araucaria Forest Fragments in Southern Brazil. Braz J Biol **75:**254–255. doi:10.1590/1519-6984.09914.

Souza TAJ, Noll FB, Bicudo HEMC, Madi-Ravazzi L. 2014. Scanning electron microscopy of male terminalia and its application to species recognition and phylogenetic reconstruction in the *Drosophila saltans* group. PLoS ONE **9:**1–12. doi:10.1371/journal.pone.0097156.

Spencer J, Heed WB. 1975. Dispersal of *Drosophila*: the effect of *baiting* on the behavior and distribution of natural populations. Am Nat **109:**207–216. doi:10.1086/282987.

Suguio K. 2008. Mudanças Ambientais da Terra. 1st Ed. Instituto Geológico, São Paulo.

Sunnucks P. 2000. Efficient genetic markers for population biology. Trends Ecol Evol **15:**199–203. doi:10.1016/S0169-5347(00)01825-5.

Tabarelli M, Pinto LP, Silva JMC, Hirota MM, Bedê LC. 2005. Desafios e oportunidades para a conservação da biodiversidade na Mata Atlântica brasileira. Megadiversidade **1:**132–138.

Tait G, Vezzulli S, Sassù F, Antonini G, Biondi A, Baser N, Sollai G, Cini A, Tonina L, Ometto L, Anfora G. 2017. Genetic variability in Italian populations of *Drosophila suzukii*. BMC Genetics **18:**87–96. doi:10.1186/s12863-017-0558-7.

TaxoDros. 2020. The database on Taxonomy of Drosophilidae v1.04 – Database 2020/1, Zürich. https://www.taxodros.uzh.ch. Accessed 24 Feb. 2020.

Templeton AR, Robertson RJ, Brisson J, Strasburg J. 2001. Disrupting evolutionary processes: the effect of habitat fragmentation on collared lizards in the Missouri Ozarks. Proc Natl Acad Sci USA **98:**5426–5432. doi:10.1073/pnas.091093098.

Tidon-Sklorz R, Sene FM. 1992. Vertical and temporal distribution of *Drosophila* (Diptera, Drosophilidae) species in a wooded area in the state of São Paulo, Brazil. Rev Bras Biol **52:**311–317.

Torres FR, Madi-Ravazzi L. 2006. Seasonal variation in natural populations of *Drosophila* spp. (Diptera) in two woodlands in the state of São Paulo, Brazil. Iheringia Ser Zool **96:**437–444. doi:10.1590/S0073-47212006000400008.

Valadão H, Proença CEB, Kuhlmann MP, Harris SA, Tidon R. 2019. Fruit-breeding drosophilids (Diptera) in the Neotropics: playing the field and specializing in generalism? Ecol Entomol **44:**721–737. doi:10.1111/een.12769.

Vilela CR. 1992. On the *Drosophila tripunctata* species group (Diptera, Drosophilidae). Rev Bras Entomol **36:**197–221.

Vilela CR, Bächli G. 1990. Taxonomics studies on Neotropical species of seven genera of Drosophilidae (Diptera). Mitt Schweiz Entomol Ges **63:**1–332.

Wallace B. 1966. On the dispersal of *Drosophila*. Am Nat **100:**551–563. doi:10.1086/282449.

Weir BS. 1996. Genetic Data Analysis II. Sinauer Associates, Sunderland.

Wright S. 1978. Evolution and the genetics of population: variability within and among natural populations. Vol. 4. University of Chicago Press, Chicago, USA.

Wu X, Wang L, Zhang D, Wen Y. 2019. Microsatellite null alleles affected population genetic analyses: a case study of Maire yew (*Taxus chinensis var. mairei*). J Forest Res **24:**230–234. doi:10.1080/13416979.2019.1634230.

Yotoko KSC, Medeiros HF, Solferini VN, Klaczko LB. 2003. A molecular study of the systematics of the *Drosophila tripunctata* group and the *tripunctata* radiation. Mol Phylogenet Evol **28:**614–619. doi:10.1016/S1055-7903(03)00218-5.

Zorzato SV. 2015. Análise da Estrutura Genética Populacional de *Drosophila ornatifrons* (Diptera; Drosophilidae). Master Dissertation. Universidade Estadual Centro-Oeste, Brazil.

Supplementary materials

Table S1. Allele frequencies for 11 microsatellite loci of nine Brazilian populations of *Drosophila sturtevanti*. Population descriptions are in the Material and Methods section. n = number of individuals genotyped. (download)

Table S2. Values of genetic diversity (± standard error) across 11 loci of nine populations of *D. sturtevanti* from the Atlantic Forest biome in Brazil. M – monomorphic; P – polymorphic; A – autosomic; S – sex chromosome linked (X or Y); N_T = Total number of alleles of each locus; *PIC* = Polymorphism Information Content; N_a – mean number of alleles; N_e – mean effective number of alleles; H_O – mean observed heterozygosity; H_E – mean expected heterozygosity; F – fixation index; A_n – null allele frequency estimate. (download)

© 2021 Academia Sinica, Taiwan