



## ECOSYSTEMS

# Phylogeography of *Drosophila buzzatii* (Diptera, Drosophilidae): responses of the species to Quaternary climates in tropical and subtropical South America

MATEUS HENRIQUE SANTOS, DORA YOVANA BARRIOS-LEAL & MAURA HELENA MANFRIN

**Abstract:** *Drosophila buzzatii* (Diptera: Drosophilidae) is a fly that breeds exclusively on decaying tissues of cacti species widely distributed in tropical and subtropical areas of South America. This distribution includes biomes in distinct climatic regimes (e.g., seasonal rain forest, semi-arid scrubs, savannas, and grasslands), which at first glance could give the false impression that the species is not sensitive to either climate or vegetation physiognomies. However, detection of historical demographic events within *D. buzzatii* reveal the interplay between climate and the population structure of the species as the Late Quaternary climate changes occurred. To understand this process, we performed a phylogeographic analysis based on sequences of the mitochondrial gene COI for 128 individuals from 43 localities. Our analyses combined coalescent methods, population genetics, and paleodistributions estimation methods. Our study reveals that the COI haplotype diversity is geographically structured, with a decreasing cline from north to south. The results suggest an ancient range expansion, dated from 610k to 550k years before present, in the northernmost region of the species distribution, the Caatinga vegetation. More recently, an intense gene flow and a population expansion were detected in the central and south portions of its distribution. The demographic events detected date back to the glacial periods of the Quaternary.

**Key words:** Evolutionary history, migration routes, population genetics, species distribution modeling.

## INTRODUCTION

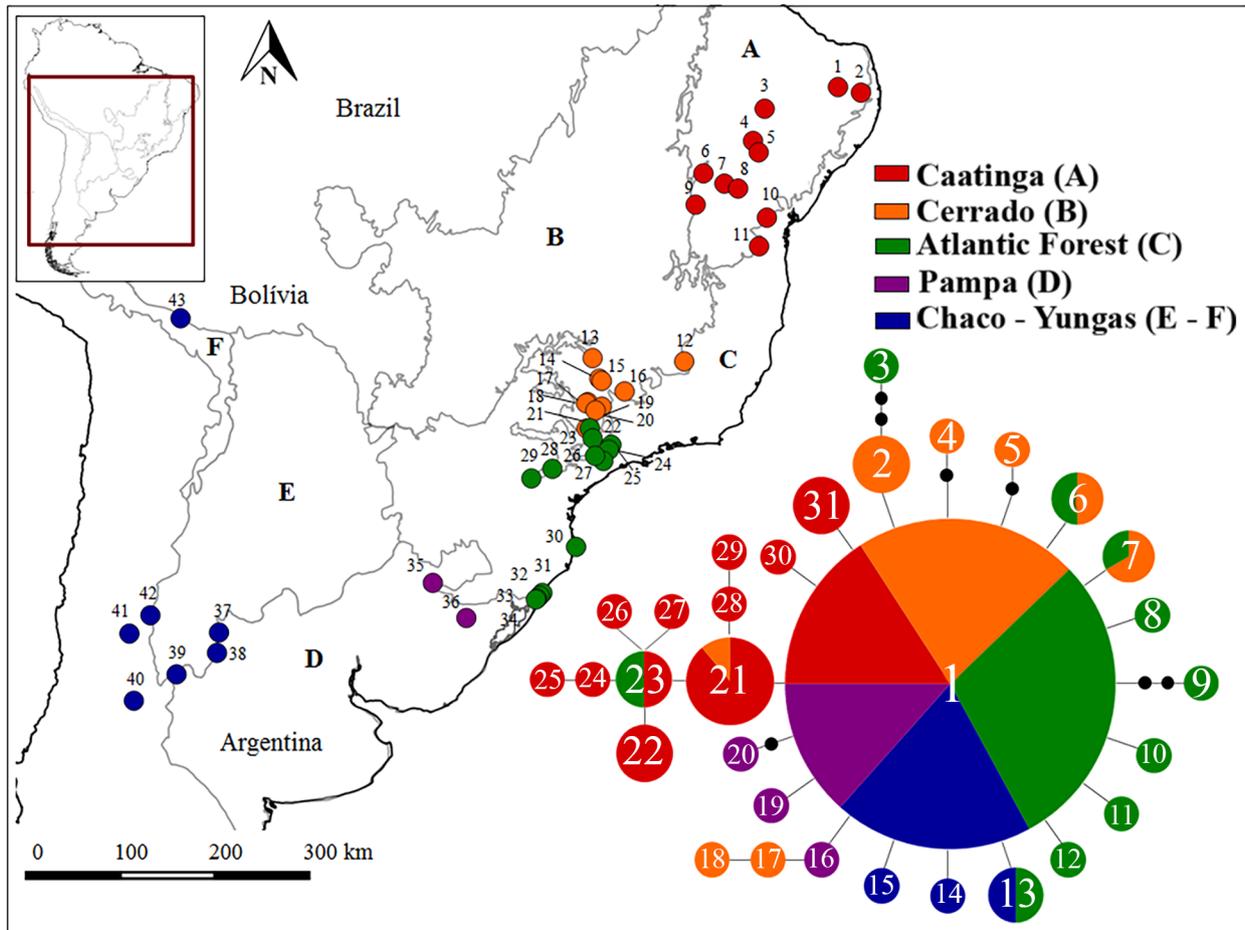
The existence of sister taxa occurring in distinct neighboring biomes is usually seen as evidence of former historical connections and disconnections between these biomes. However, under a more refined demographic approach, a species with wide distribution can generate a better understanding of events that occurred in distinct biomes, once one can detect interruptions in gene flow or migration in response to biome changes (Hewitt 2004, Bragg et al. 2015).

The *Drosophila buzzatii* Patterson and Wheeler, 1942 species cluster comprises a monophyletic groups composed of seven species that use rotting cacti tissues for larval development (Manfrin et al. 2001, Manfrin & Sene 2006). Phylogeographic analysis performed for these species revealed that the main events that explain their current genetic variation distribution are range expansion followed by isolation by distance (De Brito et al. 2002a, Franco & Manfrin 2013, Moraes et al. 2009). However, phylogeographic analyses were only

performed for species restricted to one or two biomes (Carnaval et al. 2009).

*Drosophila buzzatii* is the sister species to the remaining species included in the cluster and it has a wide geographical distribution, including distinct tropical and subtropical landscapes (Manfrin & Sene 2006). This distribution makes it an interesting biological model since an interruption in the gene flow or new migratory events in the species could be a response to changes in the dynamics of the biome. The species is found in the Caatinga (a semi-arid biome, consisting of dry forest and shrubby vegetation, located in northeast Brazil),

Cerrado (a mesic, albeit seasonal, savanna, occurring mainly in central Brazil), Chaco (a dry forest ranging from southern Bolivia and northern Paraguay to northern Argentina), Atlantic Forest (a mostly coastal forest ranging from northeast to south Brazil) and Pampa (which occurs in Uruguay, northeast Argentina and in the southernmost territories of Brazil, consisting of grassy prairies with moderate, year-round precipitation) (Ab’Saber 1977, Fiaschi & Pirani 2009, Morrone 2006, Oliveira-Filho & Ratter 1995, Oliveira-Filho et al. 2006, Overbeck et al. 2006, Zanella et al. 2010) (Fig. 1). The reproductive biology of this fly requires oviposition in



**Figure 1.** Haplotype network for COI sequences and a map of South America showing sampling points for *Drosophila buzzatii* and the biomes comprised in this study. Each line of the network represents one mutational step. The size of each circle indicates the number of individuals having that haplotype, and colors represent distinct biomes. Black circles represent putative unsampled haplotypes. Numbers shown on the map are related to populations, and the network numbers indicate haplotypes.

rotting cactus and, therefore, populations are evolutionarily associated with these plants. Cacti, on the other hand, depend on semi-arid or dry climates, or shallow soils incapable of retaining water. These conditions are observed today in all the biomes mentioned above, either as climatic phenomena (extended dry season, low average precipitations) or due to the presence of geomorphologic and geologic features which favor the maintenance of shallow soils, such as rock outcrops (found in all of the above-mentioned biomes). The obligatory association of the fly with many cactus species creates an interesting model to investigate how gene flow is currently occurring, and propose scenarios in which interruptions to this flow could help create genetic structure.

Previous studies have shown that, despite its wide distribution, *D. buzzatii* has low levels of genetic structure and a high level of gene flow between populations (De Brito et al. 2002a, Kuhn et al. 2003), except for the occurrence of private enzymatic alleles and a high level of polymorphism in mtDNA sequences in populations of the Caatinga biome (Barker et al. 1985, De Brito et al. 2002a).

Monophyly, divergence times, and evolution of host plant use were inferred from a revised phylogeny of the *Drosophila repleta* Wollaston, 1858 species group (Manfrin et al. 2001, Manfrin & Sene 2006). All these were dated as Prequaternary events and there is a consensus that the glacial cycles have affected tropical and subtropical South America by alternating drier and humid climates affect (Clapperton 1993, Ledru et al. 2005, Sene et al. 1988).

Here, we investigate the phylogeographic history of *D. buzzatii* with the mitochondrial gene COI and evaluate: (1) the population structure of *Drosophila buzzatii*; (2) how the observed genetic patterns may be related to Quaternary climate

change in South America; (3) how many and what are the directions of migration events between populations; and (4) its predicted distribution using Species Distribution Modelling (SDM) to determine potential corridors for dispersion or gene flow interruption.

## MATERIALS AND METHODS

### Samples, DNA extraction, and sequencing

A total of 128 males from 43 populations of *Drosophila buzzatii* specimens, or DNA sequences were sampled (Table I, Fig. 1) (OR582485 - OR582604).

The DNA extraction was performed using Wizard Genomics DNA Purification Kit (Promega, Madison, WI, USA). We PCR amplified and sequenced 710 bp fragments of the COI mitochondrial gene using the primers 1406f (Simon et al. 1994) and 2191r (De Brito et al. 2002a). Amplification was performed in a 20 µl reaction with 50-75 ng of DNA, 2,0 mM MgCl<sup>2+</sup>, 0,2 nmol/µl dNTPs, 4 mM each primer, and 0,14 µl of Taq polymerase (5U/µl) in a PCR buffer (pH 8.5) with Tris-HCl, KCl (PROMEGA) The PCR conditions were one step of 94°C for 90 sec., 30 cycles of 94°C for 40 sec., 47°C for 40 sec. and 72°C for 120 sec, and a final extension of 72°C for 10 min. The PCR products were visualized in 1% agarose gels stained with ethidium bromide and purified using ExoSAP-It Kit (Pharmacia). Sequencing reactions were performed according to the BigDye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer, Foster City, CA, USA). Automatic DNA sequencing was performed in an ABI Prism™ 377 sequencer (Perkin- Elmer).

### Genetic diversity and network

49 of a total of 128 DNA sequences were retrieved from Genbank and had been previously published (De Brito et al. 2002a).

**Table I. Location and geographic coordinates of samples used to model the potential current and past distributions of *Drosophila buzzatii*, with a summary of the COI sequences descriptive statistics.**

Location	Sta-Cou	Bio	Lon	Lat	Hd ± SD	$\pi$ ± SD	s	k	N	Haplotypes
1- Junco do Seridó	PB-Bra	Caa	-36.70	-7.01	---	---	---	---	1	Hap21
2- Campina Grande	PB-Bra	Caa	-35.67	-7.26	1.00±0.27	0.00370±0.00101	3	2.00	3	Hap21, 22, 27
3- Exu	PE-Bra	Caa	-39.98	-7.98	0.83±0.22	0.00216±0.00067	2	1.17	4	Hap1, 21, 31
4- Juazeiro	BA-Bra	Caa	-40.50	-9.42	0.83±0.22	0.00312±0.00113	3	1.67	4	Hap21, 26, 28, 29
5- Morro Torre	BA-Bra	Caa	-40.25	-9.93	0.50±0.26	0.00093±0.00049	1	0.50	4	Hap1 – Hap21
6- Xique-Xique	BA-Bra	Caa	-42.72	-10.87	1.00±0.25	0.00185±0.00093	1	1.00	2	Hap1 – Hap21
7- Irecê	BA-Bra	Caa	-41.78	-11.33	0.00±0.00	0.00000±0.00000	0	0.00	2	Hap1
8- Morro do Chapéu	BA-Bra	Caa	-41.17	-11.55	0.66±0.31	0.00110±0.00052	1	0.67	3	Hap1 – Hap30
9- Ibotirama	BA-Bra	Caa	-43.07	-12.27	0.50±0.26	0.00083±0.00044	1	0.50	4	Hap1 – Hap31
10- Milagres	BA-Bra	Caa	-39.88	-12.85	1.00±0.27	0.00371±0.00130	3	2.00	3	Hap22, 24, 25
11- Manoel Vitorino	BA-Bra	Caa	-40.23	-14.13	0.50±0.26	0.00167±0.00088	2	1.00	4	Hap1 – Hap23
12- Cardeal Mota	MG-Bra	Cer	-43.58	-19.28	1.00±0.50	0.00396±0.00198	2	2.00	2	Hap1 – Hap4
13- Nova Ponte	MG-Bra	Cer	-47.68	-19.13	0.28±0.19	0.00053±0.00036	1	0.28	7	Hap1 – Hap2
14- Morro ChapéuMG	MG-Bra	Cer	-47.38	-20.05	1.00±0.27	0.00370±0.00130	3	2.00	3	Hap1, 17, 18
15- Estreito	MG-Bra	Cer	-47.27	-20.15	---	---	---	---	1	Hap1
16- Furnas	MG-Bra	Cer	-46.25	-20.62	0.40±0.23	0.00074±0.00044	1	0.40	5	Hap1 – Hap2
17- Altinópolis	SP-Bra	Cer	-47.90	-21.10	0.67±0.31	0.00123±0.00058	1	0.67	3	Hap1 – Hap7
18- Sertãozinho	SP-Bra	Cer	-47.97	-21.15	1.00±0.25	0.00370±0.00185	2	2.00	2	Hap1 – Hap5
19- Cajuru	SP-Bra	Cer	-47.27	-21.30	---	----	---	---	---	---
20- São Simão	SP-Bra	Cer	-47.55	-21.47	1.00±0.50	0.00185±0.00093	1	1.00	2	Hap1 – Hap6
21- Brotas	SP-Bra	Cer	-47.95	-22.28	---	---	---	---	1	Hap21
22- Itirapina	SP-Bra	Flo	-47.80	-22.27	1.00±0.27	0.00247±0.00082	2	1.33	3	Hap1, 6, 7
23- Piracicaba	SP-Bra	Flo	-47.68	-22.72	1.00±0.50	0.00792±0.00396	4	4.00	2	Hap1 – Hap3
24- Itatiba	SP-Bra	Flo	-46.84	-23.01	0.00±0.00	0.00000±0.00000	0	0.00	1	Hap12
25- Serra Japi	SP-Bra	Flo	-46.95	-23.23	0.00±0.00	0.00000±0.00000	0	0.00	1	Hap1
26- Ibiuna	SP-Bra	Flo	-47.20	-23.73	0.00±0.00	0.00000±0.00000	0	0.00	1	Hap1
27- Sorocaba	SP-Bra	Flo	-47.57	-23.50	1.00±0.50	0.00185±0.00093	1	1.00	2	Hap1 – Hap11
28- Sengês	PR-Bra	Flo	-49.48	-24.08	0.33±0.21	0.00062±0.00040	1	0.33	6	Hap1 – Hap10
29- Tibagi	PR-Bra	Flo	-50.42	-24.51	---	---	---	---	1	Hap1
30- Florianópolis	SC-Bra	Flo	-48.42	-27.57	---	---	---	---	1	Hap23
31- Arroio Teixeira	RS-Bra	Flo	-49.93	-29.63	0.40±0.23	0.00222±0.00132	3	1.20	5	Hap1 – Hap9
32- Capão da Canoa	RS-Bra	Flo	-50.05	-29.76	---	---	---	---	1	Hap1
33- Osório	RS-Bra	Flo	-50.17	-29.88	0.00±0.00	0.00000±0.00000	0	0.00	5	Hap1
34- Tramandaí	RS-Bra	Flo	-50.22	-29.92	0.70±0.22	0.00148±0.00056	2	0.80	5	Hap1, 8, 13

**Table I. Continuation.**

35- Santiago	RS-Bra	Pam	-54.83	-29.18	0.18±0.14	0.00030±0.00024	1	0.18	11	<b>Hap1 – Hap11</b>
36- Guaritas	RS-Bra	Pam	-53.33	-30.75	1.00±0.27	0.00371±0.00130	3	2.00	3	<b>Hap1, 19, 20</b>
37- Córdoba	Argentina	Cha	-64.40	-31.40	---	---	---	---	1	<b>Hap1</b>
38- La Cruz	Argentina	Cha	-64.49	-32.31	0.70±0.22	0.00138±0.00052	2	0.80	5	<b>Hap1, 13, 14</b>
39- San Luis	Argentina	Cha	-66.30	-33.27	0.00±0.00	0.00000±0.00000	0	0.00	3	<b>Hap1</b>
40- San Raphael	Argentina	Cha	-68.21	-34.45	---	---	---	---	1	<b>Hap1</b>
41- San Juan	Argentina	Cha	-68.41	-31.45	0.67±0.31	0.00103±0.00049	1	0.67	3	<b>Hap1 – Hap15</b>
42- Vale Fertil	Argentina	Cha	-67.47	-30.63	0.00±0.00	0.00000±0.00000	0	0.00	3	<b>Hap1</b>
43- Cochabamba	Bolivia	Cha	-66.12	-17.35	0.00±0.00	0.00000±0.00000	0	0.00	2	<b>Hap1</b>
Total					0.58±0.05	0.00213±0.00028	27	0.93	128	

**Sta-Cou, State-Country; Bra, Brazil; PB, Paraíba, PE, Pernambuco, BA, Bahia; MG, Mina Gerais; SP, São Paulo; PR, Paraná; SC, Santa Catarina; RS, Rio Grande do Sul; Bio, Biome; Caa, Caatinga; Cer, Cerrado; Flo, Atlantic Florest; Pam, Pampas; Cha, Chaco; Lon, Longitude; Lat, Latitude; Hd, haplotype diversity;  $\pi$ , nucleotide diversity; s, number of variable sites; k, average number of nucleotide differences; N, number of individuals; Haplotypes, estimated in TCS.**

The COI sequences were aligned using ClustalW Multiple Alignment 1.8 (Thompson et al. 1994) and edited in BioEdit v7.0.9 (Hall 1999). In the software Mega v6.0.5 (Tamura et al. 2013), using the Akaike Information Criterion (AIC), HKY+I+G (Hasegawa et al. 1985) was chosen as the best-fit model to explain the nucleotide substitution in the sampled sequences. The haplotype diversity (Hd), nucleotide diversity ( $\pi$ ) (Nei 1987) and their standard deviations, number of polymorphic sites (s), and average number of nucleotide differences (k) were calculated in DnaSP 5.0 (Librado & Rozas 2009). A haplotype network was generated in TCS 1.21 (Clement et al. 2000).

### **Genetic structure and demographic analysis**

To evaluate population structure, we used the Bayesian clustering method in the program BAPS v6.0 (Corander et al. 2013), which was performed using the admixture analysis, with values of K ranging from 1 to 6,200 iterations, and five replicates. The group with the highest log likelihood was selected as optimal. Furthermore, we performed the analysis of global molecular variance (AMOVA) with hierarchical groups, according to biomes and the preliminary results

of the BAPS program using Arlequin 3.5 (Excoffier et al. 2005).

The Mantel test (Mantel 1967, Smouse et al. 1986) was performed using Fst values to check for significant correlations between genetic and geographic distances, estimated using the central geographical coordinate of the distributed points in each biome and the distance from this point to the other points using the equation: Distance  $i-j = 6371(A\cos(\cos(\pi(90-\text{Latitude}_i)/180)\cos((90-\text{Latitude}_j) \pi/180)) + \text{Sen}((90-\text{Latitude}_i) \pi/180) \text{Sen}((90-\text{Latitude}_j) \pi/180)\cos((\text{Longitude}_j - \text{Longitude}_i)\pi/180))$  and performed in Arlequin 3.5 (Excoffier et al. 2005).

The demographic equilibrium was tested by calculating Fu and Li's F(Fu & Li 1993) and Tajima's D (Tajima 1989) statistics and, to check the demographic expansion, we evaluated both the mismatch distribution and the use of fit between observed and expected distributions by the sum of square deviations (SSD) in DnaSP 5.0 (Librado & Rozas 2009).

Expansion times were estimated with Extended Bayesian Skyline Plot (EBSP), in Beast v1.8.4 (Drummond & Rambaut 2007), using the HKY+I+G substitution model. We opted for a

strict molecular clock to calculate the minimum and maximum substitution rates for the genus *Drosophila* Fallén, 1823 COI, 0.87% and 1.75% per million years (Barrios-Leal et al. 2018). We performed 100 million iterations with a Markov chain Monte Carlo (MCMC) algorithm, and sampled every 1,000 iterations with a 10% burn-in. The log-likelihood plot from EBSM was analyzed using the software TRACER 1.5, with a minimum ESS value of 200 to ensure the reliability of the posterior distribution of the estimated parameters (Suchard et al. 2018).

The program MIGRATE v3.2.19 (Beerli 2006) was used to calculate migration rates and effective population size between pairs of populations, using a coalescent approach (Beerli & Felsenstein 1999). To perform the analyses, we relied on a maximum likelihood estimation with 10 short and 3 long chains, discarding 10,000 trees as initial 'burn-in'. This analysis was performed five times to verify the consistency of the results using a "geofile", consisting of the distance between the geographic cores of the samples in each biome. For the first run, we used values of Theta calculated by Fst and, for subsequent runs, the input values consisted of those obtained in the previous run.

### **Species Distribution Modelling (SDM)**

Past and current potential distribution of *Drosophila buzzatii* was performed using the Species Distribution Modelling (SDM) approach. For this analysis, we used both current layers, which represent a moist climate, and layers from the Last Glaciation Maximum (21,000 years before the present), which comprised a drier climate, in the resolution of 2.5 arc-minutes. We used the following algorithms and programs to perform the models: the maximum entropy algorithm implemented in MaxEnt (Phillips et al. 2006) and the Support Vector Machine (SVM) implemented in Open Modeler v. 1.4.0 (Muñoz

et al. 2011). The sampling points used for these analyses are listed in table III. and the layers used were as follows: annual mean temperature (Bio 1); mean diurnal range (Bio 2); temperature seasonality (Bio 4); maximum temperature of warmest quarter (Bio 8); minimum temperature of coldest quarter (Bio 9); temperature annual range (Bio 7); annual precipitation (Bio 12); precipitation seasonality (Bio 15); precipitation of wettest quarter (Bio 16); and precipitation of driest quarter (Bio 17). These layers represent important environmental conditions for *D. buzzatii* and their host cacti one time extreme abiotic conditions like high humidity or high temperature variation could affect the survival of the cacti or the fly. We used the "dontextrapolate" function in MaxEnt to avoid spurious projections (Giovanelli et al. 2008, Thomé et al. 2010). The Lowest Predicted Value Threshold (LPT) was used for threshold definition, and probabilities below the threshold value were transformed to zero; a binary map was produced in ArcGis v. 10.1 (ESRI 2012). The analysis accuracy was checked using 70% of the data as training data and 30% as test data for model validation in 100 replicates. The model evaluation was made by the underlying area (AUC) (Fielding & Bell 1997, Manel et al. 2001). Models' projections (current and Last Interglacial models from Community Climate System Model CCSM, and Model for Interdisciplinary Research on Climate, MIROC) from the analysis were acquired in the Worldclim database (available at: <http://www.worldclim.org/past>).

## **RESULTS**

### **Genetic diversity and network**

The standard diversity indexes for the 31 COI haplotypes observed from the 128 sampled sequences are shown in table I. The haplotype Hap1 was the most frequent (67.21% of all

samples) and widely distributed in all biomes. The haplotype Hap21 was the second most frequent (7.37%) and it was found only in the Caatinga and Cerrado biomes. The number of private haplotypes for each biome is as follows: 9 for Caatinga, 5 for Cerrado, 6 for Atlantic Forest, 2 for Chaco, and 3 for Pampa (Fig. 1 and Table II).

The highest nucleotide diversity and haplotype diversity were found in the Caatinga populations ( $\pi = 0.00265$ ,  $Hd = 0.79$ ), followed by Cerrado, Atlantic Forest, Pampa, and Chaco, in a decreasing north to south cline (Table II).

### Genetic structure and demographic analysis

The results of BAPS showed that the optimal genetic structure is composed of two groups: Group 2, with populations 1, 2, 3, 4, 5, 10, and 11 all belonging to the Caatinga biome, and populations 23-30 from the Atlantic Forest; Group 1, with the remaining populations (Fig. 2).

The global AMOVA indicated weak population structure ( $F_{st} = 0.31$ ,  $p < 0.05$ ). In the different scenarios tested with a hierarchical

AMOVA, we found better support for the scenario of two groups inferred by BAPS ( $F_{ct} = 0.54$ ,  $p < 0.05$ ), being those that maximized the  $F_{ct}$  and minimized the  $F_{sc}$ . The use of AMOVA to test the biomes has been proved not to be appropriate for grouping, since the difference among groups is lower than that within groups (Table III).

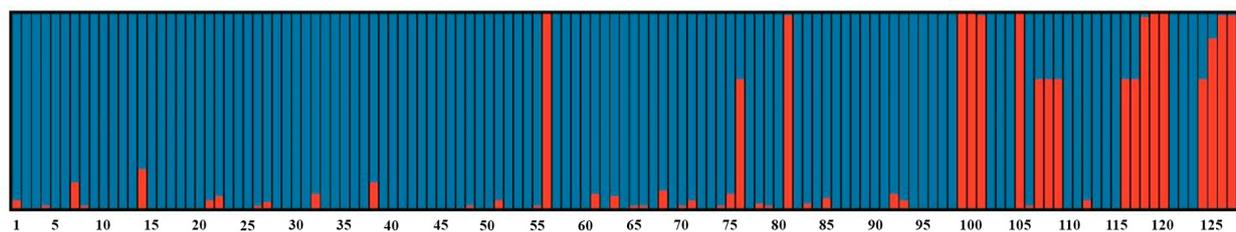
The Mantel test did not reveal a correlation between population structure ( $F_{st}$ ) and the geographic distance in *D. buzzatii* (correlation coefficient ( $r_{Y1}$ ) = 0.90;  $p > 0.05$ ).

Together, the neutrality tests and the mismatch distribution only indicated a scenario of demographic expansion for Group 1 and all *D. buzzatii* samples. However, for Group 2, the hypothesis of expansion supported by the mismatch distribution analyses cannot be rejected (Table IV; Supplementary Material - Figure S1). These results were corroborated by the EBSP analysis, for which the result (not including the value zero) was in the 95% HPD interval for the number of demographic changes (Heled & Drummond 2008), indicating an older

**Table II. Standard diversity indexes calculated for *Drosophila buzzatii*.**

Biomes	N	Hd $\pm$ SD	$\pi \pm$ SD	s	k	Hap	Pop	Fst
Caatinga	34	0.79 $\pm$ 0.05	0.00265 $\pm$ 0.00035	9	1.26	9	11	0.32*
Cerrado	28	0.59 $\pm$ 0.11	0.00208 $\pm$ 0.00055	11	1.05	5	10	0.30*
Atlantic Forrest	34	0.51 $\pm$ 0.11	0.00175 $\pm$ 0.00054	13	0.88	6	13	0.12
Pampas	14	0.39 $\pm$ 0.16	0.00106 $\pm$ 0.00051	4	0.57	2	2	0.32*
Chaco	18	0.22 $\pm$ 0.12	0.00043 $\pm$ 0.00026	2	0.22	3	7	0.23

N, number of individuals; Hd, haplotype diversity;  $\pi$ , nucleotide diversity; s, number of variable sites; k, average number of nucleotide differences; Hap, numbers of private haplotypes; Pop, population; Fst,  $p < 0.05^*$



**Figure 2. Genetic structure in *Drosophila buzzatii* inferred by BAPS (K=2). The color represents the different genetic groups, Group 1 in blue and composed of most individuals; and group 2 is in red and composed of individuals from populations 1, 2, 3, 4, 5, 10, 11, 23, 30.**

population expansion in the last 0.05–0.08 Ma (Figure S2).

Our results from MIGRATE showed migratory movements from Caatinga to Atlantic Forest; from Cerrado to Caatinga and Pampa; from Atlantic Forest to Cerrado; from Pampa to Chaco; and from Chaco to Atlantic Forest (Table V). The most intense movements occurred from Cerrado to Pampa, followed by Caatinga to Atlantic Forest, and then from Chaco to Atlantic Forest.

### Species Distribution Modelling (SDM)

The AUC values (Current time: MaxEnt: 0.9333 and SVM: 0.9111, Last Glaciation Maximun: CCSM – Maxent = 0,912 and MIROC – Maxent = 0,900) can be classified as excellent according to Swets (1988)'s method. The values generated by TSS ROC were from the current model MaxEnt: 0.7239 and SVM: 0.7177.

The Lowest Predicted Value Threshold (LPT) calculated was to the current MaxEnt model: 0.0903 and to SVM: 0.1576. Last glaciation models were plotted based on the current results in MaxEnt (Fig. 3; Figure S3). Current models showed a narrow distribution area between Caatinga, Cerrado, and Atlantic Forest and large areas potentially connecting Chaco to Pampa. Both CCSM and MIROC predictive models showed a connection between Cerrado and Chaco, although the Pantanal biome is not shown in the current models. The MIROC model showed a connection along with the current submerse coast of Brazil and Uruguay that could represent a past distribution when the ocean levels were lower than what is seen today. The overlapping map (Fig. S1) showed two main stable regions: (i) a narrow portion between Caatinga, Cerrado, and Atlantic Forest, not connected with Pampa;

**Table III. AMOVA results for *Drosophila buzzatii* populations.**

Groups	Source of variation	Sum of squares	Variance components	Percentage variation	Fixation Indices
Global	Among groups	38.42	0.18	31.39	$F_{ST} = 0.31^*$
	Within populations	33.11	0.39	68.61	
Two groups BAPS	Among groups	15.16	0.52	53.62	$F_{CT} = 0.54^*$
	Among populations within groups	23.25	0.06	6.23	$F_{SC} = 0.13^*$
	Within populations	33.11	0.39	40.15	$F_{ST} = 0.60^*$
Biome groups	Among groups	8.25	0.03	6.50	$F_{CT} = 0.06^*$
	Among populations within groups	30.05	0.15	25.69	$F_{SC} = 0.27^*$
	Within populations	33.11	0.39	67.81	$F_{ST} = 0.32^*$

\* $p \leq 0.05$ .

**Table IV. Neutrality tests results for *Drosophila buzzatii* populations.**

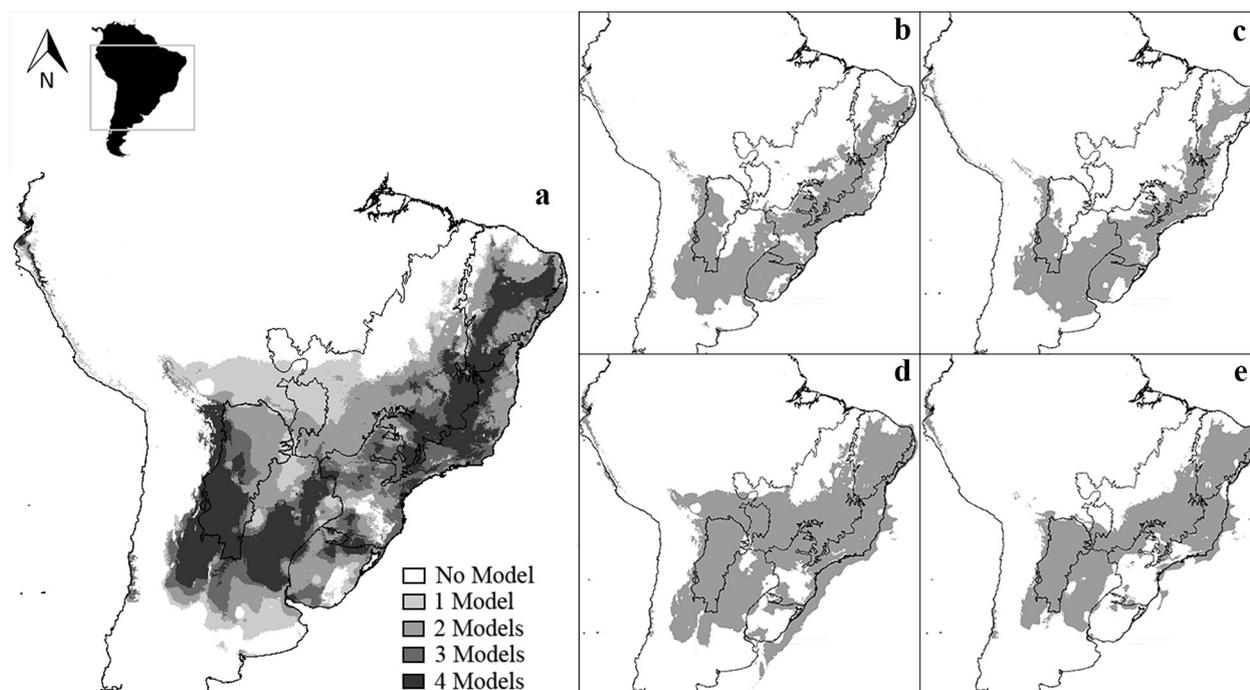
Biomes	n	Fu's Fs	Tajima's D	Mismatch Distribution
Total	128	-3.78*	-2.40**	0.0098***
Group_1	108	-4.20*	-2.40**	0.0003***
Group_2	20	-2.03	-1.51	0.0238***

Fu's Fs, Tajima's D and Mismatch Distribution.

\*  $p < 0.05$ ; \*\*  $p < 0.01$ ;  $p > 0.05$ \*\*\*.

**Table V.** Estimative of direction and number of migrants per generation calculated by the program MIGRATE. Biomes with a + indicating migrant receiving.

	Theta ( $Ne \mu$ )	Migrants per generation				
		Caatinga+	Cerrado+	Atlantic Forest+	Pampas+	Chaco+
Caatinga	0.0094	-	6.36e-007	2.62e+006	1.21e-006	1.67e-006
Cerrado	0.0077	4.91e+005	-	9.77e-006	6.45e+006	4.08e-005
Atlantic Forest	0.0049	1.40e-006	1.43e+005	-	5.37e-007	1.41e-006
Pampas	0.0067	1.08e-010	5.14e-011	3.03e-011	-	5.02e+004
Chaco	0.0024	3.02e-007	1.91e-007	1.71e+006	9.88e-008	-

**Figure 3.** Summary map of historically stable presence areas of *Drosophila buzzatii*. These projections were obtained by overlapping the predicted logistic outputs in binary maps after threshold calculation in ArcGIS 10.1. In A, stable areas according to each model; in B, the model for the current distribution using MaxEnt; in C, SVM model; in D, the model for the last glaciation calculated by MaxEnt, using MIROC prediction; and in E, the model for the last glaciation calculated by the MaxEnt, using CCSM prediction.

and (ii) Chaco with Pampa, with only a few stable areas in Pampa. Areas in Peru and Chile have also been projected, although they were not considered natural distribution areas of *D. buzzatii*.

## DISCUSSION

The *Drosophila buzzatii* cluster is composed of endemic South American species, whose larval

development occurs exclusively in the decaying tissues of different cactus species. Some species of this group show an intimate relationship with one or a few cactus species, which makes them restricted in distribution by the cactus occurrence in one or two biomes (Manfrin & Sene 2006). The current distribution of the *D. buzzatii* cluster seems to be related to glaciation events, where cooling and drought periods enabled expansion and contraction movements

of their host cacti (Moraes et al. 2009). Although *D. buzzatii*, sister to the remaining species of the clade, has a wider distribution, occurring in all biomes in the east portion of South America, in tropical and subtropical regions, it is very often found in sympatry with other species in the same group. In the present work, the demographic history of *D. buzzatii*, inferred using information on genetic variation based on COI sequences, suggests area expansion and migrant interchange between biomes in the east portion of tropical and subtropical South America.

Although this species has a widespread distribution, our results using AMOVA analyses subdivide *Drosophila buzzatii* into two main groups: (i) the Caatinga and (ii) the remaining biomes (see table V). This structure is in part corroborated by previous studies using many genetic markers in other groups, such as lizards (Werneck et al. 2012), amphibians (Carnaval & Bates 2007, Carnaval & Moritz 2008), felines (Espinosa et al. 2018), plants (Caetano et al. 2008), and closely related *Drosophila* species (De Brito et al. 2002a, Franco & Manfrin 2013, Manfrin et al. 2001, Moraes et al. 2004). Concerning *Drosophila* species, the structure found in the Caatinga is related to karyotype differences (Baimai et al. 1983), polymorphic chromosomal inversions (Ruiz et al. 2000, Tosi & Sene 1989), mitochondrial haplotypes (De Brito et al. 2002a, Franco & Manfrin 2013, Manfrin et al. 2001, Moraes et al. 2004), satDNA (Kuhn et al. 2003), and exclusive allozyme alleles (Barker et al. 1985).

Low haplotypic and nucleotide diversity indexes had been previously reported (De Brito et al. 2002a, Franco et al. 2006, Kuhn et al. 2003) although the Caatinga biome showed the highest values in our analyses. Considering the founder effect in natural species, high values of genetic diversity are often found in older populations,

while more recent populations usually show a reduction in diversity due to the foundation effect and deviation bias. A classic example is seen in the Silvereyes (*Zosterops lateralis* Latham, 1802) populations, well documented in Clegg et al. (2002). According to our data, diversity was lost in a north to south direction, suggesting that the Caatinga has the oldest remnant populations of *D. buzzatii*. Previous studies have suggested an older population of *D. buzzatii* in the Chaco, due to the high levels of chromosome inversion (Naveira et al. 1984). Nevertheless, *Drosophila buzzatii* has recently colonized areas in Spain and Australia with a significant bottleneck, and high levels of chromosomal polymorphism was found in these populations (Hasson et al. 1992). According to this, the high levels of chromosomal inversions found in the Chaco is not necessarily an indication of an ancient population, but the result of natural selection. Variation in the viability, developmental time and thorax size of *D. buzzatii* is dependent on two chromosomal inversions and host cactus species used by the fly, concluding that the cactus/karyotype relation is determinant for the adaptive value (fitness) of *D. buzzatii* (Iriarte et al. 2003, Iriarte & Hasson 2000). Similar results for *Drosophila mojavensis* have been found, showing that chromosomal inversion polymorphisms are only present in populations using the cactus *agria* [*Machaerocereus gummosus* (Engelm.) Britton & Rose] as a host in Baja California (Powell 1997).

One possible explanation for the current genetic diversity distribution is a significant and ancient movement, with the species expanding its distribution from the Caatinga biome to the southern areas, such as the Atlantic Forest biome. Similar movements were detected in other species of the cluster *D. buzzatii*. For example, in *Drosophila gouveai* Tidon-Sklorz & Sene, 2001, two groups expanded their area through mountain tops from the south to the

Caatinga biome, one in the east (Group G1) and another in the west portion of the Paran River in Central Brazil (Group G2) during glaciations (Moraes et al. 2009). Other examples are *D. serido* Vilela & Sene, 1977, *D. seriema* Tidon-Sklorz & Sene, 1995 and *D. borborema* Vilela & Sene, 1977, in the Caatinga, at the Chapada Diamantina. Specifically for *D. serido*, a migratory route from the interior of the state of Bahia to the coast of the state of Santa Catarina, in south Brazil, has been previously described (Franco & Manfrin 2013, Kokudai et al. 2011).

These migratory movements were possible due to climatic connections reflected by vegetation continuity between the biomes studied. The analysis of floristic composition in the seasonally dry tropical forest (SDTF) by Prado (2000) showed three cores - I: Chaco and related vegetation; II: Amazonia, Atlantic Forest, and Cerrado vegetation; III: tropical seasonal forests of South America, Caatinga, the calcareous outcrops in Minas Gerais/Mato Grosso do Sul and the Planalto forests of So Paulo, to the Upper Uruguay and Paran River valley vegetation, and the Subandean Piedmont Forests. Those cores were connected by what this author named "tracks" (see Prado 2000) or paths. The first nucleus (Chaco) is connected to the second (Cerrado, Amazonia, and Atlantic Forest) by two paths: the northern one links these regions by the calcareous outcrops in the Cerrado, the central portion of Brazil, and the western portion of Bahia; the southern one extends along the valleys of the So Francisco and Jequitinhonha rivers and to the south, in the Rio das Velhas to the Belo Horizonte area and the states of Rio de Janeiro, So Paulo, Paran, and Santa Catarina. According to our SMD models, an area of stability connecting these biomes is estimated (Fig. 3). Interestingly, the cactus *Cereus hildmannianus* K. Schum is found in non-flooding portions of the Pantanal biome,

known as "Capes" (personal observation). The second nucleus (Paraguay-Parana basin core) links the three nucleuses and has vegetation formations that are very similar to the Caatinga arboreal formation. The third nucleus extends from the southeastern region of Catamarca in northwestern Argentina, to Bolivia and the Brazilian state of Acre (Prado 2000).

Using a molecular clock and a coalescent-based analysis, the expansion movement was estimated at approximately 510,000 ybp, with an exponential growth shown by the BSP analysis (Fig. 2). This population expansion matches the end of the Pre-Illinoian glaciation stage (2.5–0.5 Ma), a global dry/cold period that might have opened the way to the expansion of open vegetation and its associated fauna in South America. The expansion of open vegetation in South America during dry/cold glaciation events has been suggested by many authors (Ab’Saber 1977, Ledru et al. 2005, 2006, Pennington et al. 2000, Prado 2000, Prado & Gibbs 1993). According to this hypothesis, in dry/cold environmental conditions, plant species would disperse to new areas and very likely be followed by migration of its associated fauna as well. One example is the expansion of the seasonally dry tropical forest (SDTF) through the south and north routes proposed by Prado (2000). According to Pennington et al. (2009), the SDTF is distributed in clades, which they named nucleuses, that are interpreted as ancient biomes, and in South America, one of them formed an arc of plant species adapted to dry conditions, called the Pleistocene Arc (Prado 2000, Prado & Gibbs 1993). The connections between these cores in dry/cold periods of vegetation expansion have been possibly used as a bridge to the population expansion in *D. buzzatii*.

The BSP’s results from each biome (data not shown) suggests a population growth in the Illinoian period (0.3 – 0.13 Ma) with exception

of the Caatinga biome (Pre-Illinoian, around 0.51 Ma). According to our results, the oldest population growth movement was detected in the Caatinga biome, and it followed a south route towards the Atlantic Forest, around 0.51 Ma ago. This movement was identified by the Migrate analyses with an effective migrant number of  $2.10e^3$  (table IV). Caatinga, Cerrado and Atlantic Forest are adjacent biomes and share the STDF vegetation (Pennington et al. 2009). Connection areas in the model overlapping showed similar results to the South Route of SDF dispersion as proposed by Prado (2000).

Another expansion movement detected in this work occurred from the Atlantic Forest to Cerrado, around 295,000 ybp. The Cerrado is located alongside the Caatinga and the Atlantic Forrest, sharing many species with these biomes. This geographic position could allow species exchange between those biomes during dry/cold periods by opening routes promoted by vegetation expansion. Another expansion movement was found around 240,000 ybp, from Cerrado to Caatinga and Pampa. Similar dispersion events were found in other *Drosophilidae* species in a very short period (De Brito et al. 2002b, Moraes et al. 2009). Moraes et al. (2009) proposed that *D. gouveai* expanded its area around 313,000 ybp (see table II) (Moraes et al. 2009), which is in agreement with our period of expansion, at least for the movement from Cerrado to Caatinga. Another similar pattern was proposed for *D. antonietae* Tidon-Sklorz & Sene, 2001, which has an intimate association with the cactus *Cereus hildmaniannus*, occurring in mesophilic forests along rivers of the Parana-Paraguay basin. For this species, migration pattern from the state of São Paulo, up north, reaching the states of Paraná and Rio Grande do Sul was found (De Brito et al. 2002a).

The most recent expansion movement was from Pampa to Chaco, which could not be further

analyzed, since we did not sampled specimens from areas in between these two biomes. One probable explanation for this pattern could be that an expansion of *Cereus hildmannianus* to Chaco connected these two biomes, and the movements could be similar to those found in *D. antonietae* (Mateus & Sene 2007). However, we detected an expansion movement from the Chaco to the Atlantic Forest at around 270,000 ybp, which might indicate that *D. buzzatii* was in the Atlantic Forest before reaching the Pampas. One hypothesis to explain this could be an ancient connection between these biomes through the Pantanal (Fig. 3). However, this connection does not exist today, and this could be a result possibly generated due to the fact that no samples from these regions were collected.

Our data suggest that *D. buzzatii* has a high dispersion capacity, with high gene flow and low genetic structure between biomes. The dispersion/expansion movements, in general, follow a north to south pattern crossing many biomes in cold/dry climatic periods. These patterns of dispersions help to understand the dynamics of historical migratory events by species that have been very influenced by the glaciations in South America. Conservation strategies must consider these patterns when their action are planned, and could influence the natural dynamics of fauna and flora sensitive to these changes.

### Acknowledgments

We would like to thank Reinaldo O. A. A. de Brito (Universidade Federal de São Carlos – UFSCAR) and Esteban R. Hasson (Universidad de Buenos Aires – UBA) for kindly giving us some samples; Fábio de Melo Sene (Universidade de São Paulo – USP) for sharing his knowledge and for the interesting conversations during the study; and Nilce Martinez Rossi (Faculdade de Medicina de Ribeirão Preto – FMRP-USP) and Mendelson Mazucato (FMRP-USP) for laboratory support; Diego A. Fachin (USP), Lívia R. Pinheiro (Universidade Federal

de São Carlos – UFSCar), Marco Antônio T. Marinho (Universidade Federal de Pelotas – UFPel), Rafaela L. Falaschi (UEPG) and Simeão S. Moraes (Universidade Estadual de Campinas – UNICAMP) for their suggestions on the first drafts. Special thanks to the Departamento de Genética (FMRP-USP) for giving structural and personnel support. During the preparation of this paper, MHS benefited from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES – Finance Code 001), DYBL from Fundação de Amparo à Pesquisa do Estado de São Paulo grant (FAPESP 2011//51652-2) and MHM from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq 447697/2014-5).

## REFERENCES

- AB'SABER AN. 1977. Espaços ocupados pela expansão dos climas secos na América do Sul, por ocasião dos períodos glaciais quaternários. *Paleoclimas* 3: 1-19.
- BAIMAI V, SENE FM & PEREIRA M. 1983. Heterochromatin and karyotypic differentiation of some neotropical cactus-breeding species of the *Drosophila repleta* species group. *Genética* 60: 81-92.
- BARKER JSF, SENE FM, EAST PD & PEREIRA MAQR. 1985. Allozyme and chromosomal polymorphism of *Drosophila buzzatii* in Brazil and Argentina. *Genética* 67: 161-170.
- BARRIOS-LEAL DY, FRANCO FF, SILVA ECC, SANTOS CKB, SENE FM & MANFRIN MH. 2018. Deep intraspecific divergence in *Drosophila meridionalis*, a cactophilic member of the New World *Drosophila repleta* group. *Biol J Linn Soc* 123(1): 163-178.
- BEERLI P. 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* 22: 341.
- BEERLI P & FELSENSTEIN J. 1999. Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics* 152: 763.
- BRAGG JG, POTTER, S, BI K & MORITZ C. 2015. Exon capture phylogenomics: efficacy across scales of divergence. *Mol Ecol Resour* 15(4): 795-808.
- CAETANO S, PRADO D, PENNINGTON R, BECK S, OLIVEIRA-FILHO A, SPICHIGER R & NACIRI Y. 2008. The history of seasonally dry tropical forests in eastern South America: inferences from the genetic structure of the tree *Astronium urundeuva* (Anacardiaceae). *Mol Ecol* 17: 3147-3159.
- CARNAVAL A & MORITZ C. 2008. Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *J Biogeogr* 35: 1187-1201.
- CARNAVAL AC & BATES JM. 2007. Amphibian DNA shows marked genetic structure and tracks Pleistocene climate change in northeastern Brazil. *Evolution* 61: 2942-2957.
- CARNAVAL AC, HICKERSON MJ, HADDAD CF, RODRIGUES MT & MORITZ C. 2009. Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science* 323(5915): 785-789.
- CLAPPERTON CM. 1993. Nature of environmental changes in South America at the Last Glacial Maximum. *Palaeogeogr Palaeoclimatol Palaeoecol* 101: 189-208.
- CLEGG SM, DEGNAN SM, KIKKAWA J, MORITZ C, ESTOUP A & OWENS IPF. 2002. Genetic consequences of sequential founder events by an island-colonizing bird. *Proc Natl Acad Sci USA* 99: 8127-8132.
- CLEMENT M, POSADA D & CRANDALL K. 2000. TCS: a computer program to estimate gene genealogies. *Mol Ecol* 9: 1657-1659.
- CORANDER J, MARTTINEN P, SIREN J & TANG J. 2013. Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinform* 14(1): 1-10.
- DE BRITO RA, MANFRIN MH & SENE FM. 2002a. Mitochondrial DNA phylogeography of Brazilian populations of *Drosophila buzzatii*. *Genet Mol Biol* 25: 161-171.
- DE BRITO RA, MANFRIN MH & SENE FM. 2002b. Nested cladistic analysis of Brazilian populations of *Drosophila serido*. *Mol Phylogenet Evol* 22: 131-143.
- DRUMMOND A & RAMBAUT A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7: 214.
- ESPINOSA CC, TRIGO TC, TIRELLI FP, SILVA LG, EIZIRIK E, QUEIROLO D, MAZIM FD, PETERS FB, FAVARINI MO & FREITAS TRO. 2018. Geographic distribution modeling of the margay (*Leopardus wiedii*) and jaguarundi (*Puma yagouaroundi*): a comparative assessment. *J Mammal* 99: 252-262.
- ESRI - ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE. 2012. ArcGIS Release 10.1.
- EXCOFFIER L, LAVAL G & SCHNEIDER S. 2005. Arlequin version 3.0: an integrated software package for population genetics data analysis. *Evol Bioinform Online* 1: 47-50.
- FALLÉN CF. 1823. *Geomyzides Sveciae*. *Berlingianis, Lundae* [= Lund], 8 p. [1823.05.24].
- FIASCHI P & PIRANI JR. 2009. Review of plant biogeographic studies in Brazil. *J Syst Evol* 47: 477-496.

- FIELDING A & BELL J. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ Conserv* 24: 38-49.
- FRANCO FF, KUHN GC, SENE FM & MANFRIN MH. 2006. Conservation of pBuM-2 satellite DNA sequences among geographically isolated *Drosophila gouveai* populations from Brazil. *Genética* 128: 287-295.
- FRANCO FF & MANFRIN MH. 2013. Recent demographic history of cactophilic *Drosophila* species can be related to Quaternary palaeoclimatic changes in South America. *J Biogeogr* 40: 142-154.
- FU Y & LI W. 1993. Statistical tests of neutrality of mutations. *Genetics* 133: 693.
- GIOVANELLI J, HADDAD C & ALEXANDRINO J. 2008. Predicting the potential distribution of the alien invasive American bullfrog (*Lithobates catesbeianus*) in Brazil. *Biol Invasions* 10: 585-590.
- HALL TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41: 95-98.
- HASEGAWA M, KISHINO H & YANO T-A. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol* 22: 160-174.
- HASSON E, NAVIERA H & FONTDEVILA A. 1992. The breeding sites of the argentinian species of the *Drosophila mulleri* complex (subgenus *Drosophila, repleta* group). *Rev Chil Hist Nat* 65: 319-326.
- HELED J & DRUMMOND AJ. 2008. Bayesian inference of population size history from multiple loci. *BMC Evol Biol* 8(1): 1-15.
- HEWITT GM 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philos Trans R Soc Lond B Biol Sci* 359(1442): 183-195.
- IRIARTE PF & HASSON E. 2000. The role of the use of different host plants in the maintenance of the inversion polymorphism in the cactophilic *Drosophila buzzatii*. *Evolution* 54: 1295-1302.
- IRIARTE PJF, NORRY FM & HASSON ER. 2003. Chromosomal inversions effect body size and shape in different breeding resources in *Drosophila buzzatii*. *Heredity* 91: 51-59.
- KOKUDAI C, SENE F & MANFRIN M. 2011. Sympatry and asymmetric introgression between the cactophilic species *Drosophila serido* and *Drosophila antonietae* (Diptera: Drosophilidae). *Ann Entomol Soc Am* 104: 434-442.
- KUHN GCS, FRANCO FF, SILVA JR WA & MARTINEZ-ROSSI NM. 2003. On the pBuM189 satellite DNA variability among South American populations of *Drosophila buzzatii*. *Hereditas* 139: 161-166.
- LEDRU M, ROUSSEAU D & CRUZ F. 2005. Paleoclimate changes during the last 100,000 yr from a record in the Brazilian Atlantic rainforest region and interhemispheric comparison. *Quat Res* 64: 444-450.
- LEDRU MP, CECCANTINI G, GOUVEIA SEM, LÓPEZ-SÁEZ JA, PESSEDA LCR & RIBEIRO AS. 2006. Millennial-scale climatic and vegetation changes in a northern Cerrado (Northeast, Brazil) since the Last Glacial Maximum. *Quat Sci Rev* 25: 1110-1126.
- LIBRADO P & ROZAS J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451.
- MANEL S, WILLIAMS H & ORMEROD S. 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. *J Appl Ecol* 38: 921-931.
- MANFRIN MH, DE BRITO RA & SENE FM. 2001. Systematics and evolution of the *Drosophila buzzatii* (Diptera: Drosophilidae) cluster using mtDNA. *Ann Entomol Soc Am* 94: 333-346.
- MANFRIN MH & SENE FM. 2006. Cactophilic *Drosophila* in South America: A model for evolutionary studies. *Genética* 126: 57-75.
- MANTEL N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res* 27: 209-220.
- MATEUS RP & SENE FM. 2007. Population genetic study of allozyme variation in natural populations of *Drosophila antonietae* (Insecta, Diptera). *J Zool Syst Evol Res* 45: 136-143.
- MORAES E, YOTOKO K, MANFRIN M, SOLFERINI VN & SENE FM. 2009. Phylogeography of the cactophilic species *Drosophila gouveai*: demographic events and divergence timing in dry vegetation enclaves in eastern Brazil. *J Biogeogr* 36: 2136-2147.
- MORAES EM, MANFRIN MH, LAUS AC, ROSADA RS, BOMFIN SC & SENE FM. 2004. Wing shape heritability and morphological divergence of the sibling species *Drosophila mercatorum* and *Drosophila paranaensis*. *Heredity* 92: 466-473.
- MORRONE JJ. 2006. Biogeographic areas and transition zones of Latin America and the Caribbean islands based on panbiogeographic and cladistic analyses of the entomofauna. *Annu Rev Entomol* 51: 467-494.
- MUÑOZ MES, GIOVANNI R, SIQUEIRA MF, SUTTON T, BREWER P, PEREIRA RS, CANHOS DAL & CANHOS VP. 2011. openModeller:

a generic approach to species' potential distribution modelling. *Geoinformatica* 15(1): 111-135.

NAVEIRA H, HAUSCHTECK-JUNGEN E & FONTDEVILA A. 1984. Spermiogenesis of inversion heterozygotes in backcross hybrids between *Drosophila buzzatii* and *D. serido*. *Genética* 65: 205-214.

NEI M. 1987. *Molecular Evolutionary Genetics*. New York: Columbia University Press, 512 p.

OLIVEIRA-FILHO AT, JARENKOW JA & RODAL M. 2006. Floristic relationships of seasonally dry forests of eastern South America based on tree species distribution patterns. In: PENNINGTON RT ET AL. (Eds), *Neotropical Savannas and Seasonally Dry Forests: Plant diversity, Biogeography and Conservation*, 1st ed., Boca Raton: CRC Press, p. 159-192.

OLIVEIRA-FILHO AT & RATTER JA. 1995. A study of the origin of central Brazilian forests by the analysis of plant species distribution patterns. *Edinburgh J Bot* 52: 141-194.

OVERBECK GE, MÜLLER S, PILLAR VD & PFADENHAUER J. 2006. Floristic composition, environmental variation and species distribution patterns in burned grassland in southern Brazil. *Braz J Biol* 66: 1073-1090.

PENNINGTON R, PRADO D & PENDRY C. 2000. Neotropical seasonally dry forests and Quaternary vegetation changes. *J Biogeogr* 27: 261-273.

PENNINGTON RT, LAVIN M & OLIVEIRA-FILHO A. 2009. Woody plant diversity, evolution, and ecology in the tropics: perspectives from seasonally dry tropical forests. *Annu Rev Ecol Syst* 40: 437-457.

PHILLIPS S, ANDERSON R & SCHAPIRE R. 2006. Maximum entropy modeling of species geographic distributions. *Ecol Model* 190: 231-259.

POWELL JR. 1997. *Progress and prospects in evolutionary biology: the Drosophila model*. Oxford University Press, USA, 576 p.

PRADO D. 2000. Seasonally dry forests of tropical South America: from forgotten ecosystems to a new phytogeographic unit. *Edinb J Bot* 57: 437-461.

PRADO D & GIBBS P. 1993. Patterns of species distributions in the dry seasonal forests of South America. *Ann Missouri Bot Gard* 80: 902-927.

RUIZ A, CANSIAN AM, KUHN GCS, ALVES MAR & SENE FM. 2000. The *Drosophila serido* speciation puzzle: putting new pieces together. *Genética* 108: 217-227.

SENE FM, PEREIRA MAQR & VILELA CR. 1988. Contrasting patterns of differentiation inferred from traditional genetic markers in the process of speciation. *Pac Sci* 42: 81-88.

SIMON C, FRATI F, BECKENBACH A, CRESPI B, LIU H & FLOOK P. 1994. Evolution, Weighting, and Phylogenetic Utility of Mitochondrial Gene Sequences and a Compilation of Conserved Polymerase Chain-Reaction Primers. *Ann Entomol Soc Am* 87(6): 651-701.

SMOUSE PE, LONG JC & SOKAL RR. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst Zool* 35: 627-632.

SUCHARD MA, LEMEY P, BAELE G, AYRES DL, DRUMMOND AJ & RAMBAUT A. 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol* 4(1): vey016.

TAJIMA F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123(3): 585-595.

TAMURA K, STECHER G, PETERSON D, FILIPSKI A & KUMAR S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol Biol Evol* 30: 2725-2729.

THOMÉ MTC, ZAMUDIO KR, GIOVANELLI JGR, HADDAD CFB, BALDISSERA JR FA & ALEXANDRINO J. 2010. Phylogeography of endemic toads and post-Pliocene persistence of the Brazilian Atlantic Forest. *Mol Phylogenet Evol* 55: 1018-1031.

THOMPSON J, HIGGINS D & GIBSON T. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673.

TIDON-SKLORZ R & SENE FM. 1995. *Drosophila seriema*: A new member of the *Drosophila serido* (Diptera, Drosophilidae) superspecies taxon. *Ann Entomol Soc Am* 88: 139-142.

TIDON-SKLORZ R & SENE FM. 2001. Two new species of the *Drosophila serido* sibling set (Diptera, Drosophilidae). *Iheringia Ser Zool* (90): 141-146.

TOSI D & SENE FM. 1989. Further studies on chromosomal variability in the complex taxon *Drosophila serido* (Diptera, Drosophilidae). *Rev Bras de Genét* 12: 729-745.

VILELA CR & SENE FM. 1977. Two new neotropical species of the *repleta* group of the Genus *Drosophila* (Diptera, Drosophilidae). *Rev Bras Entomol* 30: 295-299.

WERNECK FP, NOGUEIRA C, COLLI GR, SITES JR JW & COSTA GC. 2012. Climatic stability in the Brazilian Cerrado: implications for biogeographical connections of South American savannas, species richness and conservation in a biodiversity hotspot. *J Biogeogr* 39: 1695-1706.

WOLLASTON TV. 1858. Brief diagnostic characters of undescribed Madeiran insects [concl.] Ann Mag nat Hist (3)1: 18-28. [1858.??]. 113-125, 2 pls. [1858.02.01].

ZANELLA NRZ, PRUDENCIO M & CASTELLANI TT. 2010. Análise da cobertura vegetal em duna semifixa dez anos após a aplicação de técnicas de restauração no Parque Municipal das Dunas da Lagoa da Conceição, Florianópolis, Santa Catarina. Biotemas 23: 49-58.

<sup>2</sup>Universidade de São Paulo, Faculdade de Medicina de Ribeirão Preto, Departamento de Genética, Rua Monte Alegre, Av. Bandeirantes, 3900, Subsetor Oeste - 11 (N-11), 14040-900 Ribeirão Preto, SP, Brazil

<sup>3</sup>Universidade de São Paulo, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Departamento de Biologia, Av. Bandeirantes, 3900, Subsetor Oeste - 11 (N-11), 14040-900 Ribeirão Preto, SP, Brazil

## SUPPLEMENTARY MATERIAL

### Figures S1, S2, S3.

#### How to cite

SANTOS MH, BARRIOS-LEAL DY & MANFRIN MH. 2023. Phylogeography of *Drosophila buzzatii* (Diptera, Drosophilidae): responses of the species to Quaternary climates in tropical and subtropical South America. An Acad Bras Cienc 95: e20220846. DOI 10.1590/0001-3765202320220846.

Manuscript received on December 7, 2022;  
accepted for publication on March 21, 2023

#### MATEUS HENRIQUE SANTOS<sup>1</sup>

<https://orcid.org/0000-0003-3063-1455>

#### DORA YOVANA BARRIOS-LEAL<sup>2</sup>

<https://orcid.org/0000-0002-7996-2792>

#### MAURA HELENA MANFRIN<sup>2,3</sup>

<https://orcid.org/0000-0001-6925-2167>

<sup>1</sup>Pós-Graduação em Biologia Evolutiva. Universidade Estadual de Ponta Grossa, Departamento de Biologia Estrutural, Molecular e Genética, Av. General Carlos Cavalcanti, 4748, Uvaranas, 84030-900 Ponta Grossa, PR, Brazil

Correspondence to: **Mateus Henrique Santos**

E-mail: [marangonsantos@gmail.com](mailto:marangonsantos@gmail.com)

### Author contributions

Mateus Henrique Santos designed the experiments, analyzed the data and drafted the manuscript; Dora Yohana Barrios-Leal, performed the BAPS analyzes and review the manuscript; Maura Helena Manfrin, head of the laboratory, financial and intellectual support, review the manuscript.

