



Phylogeography of the cactophilic species *Drosophila gouveai*: demographic events and divergence timing in dry vegetation enclaves in eastern Brazil

Evandro M. Moraes^{1*}, Karla S. C. Yotoko², Maura H. Manfrin³, Vera N. Solferini⁴ and Fábio M. Sene⁵

¹Campus Sorocaba, Universidade Federal de São Carlos (UFSCar), Sorocaba, SP,

²Departamento de Biologia Geral, Universidade Federal de Viçosa (UFV), Viçosa, MG, ³Departamento de Biologia, Faculdade de Filosofia Ciências e Letras de Ribeirão Preto, Universidade de São Paulo (USP), Ribeirão Preto, SP, ⁴Departamento de Genética e Evolução, Instituto de Biologia, Universidade Estadual de Campinas (UNICAMP), Campinas, SP and ⁵Departamento de Genética, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo (USP), Ribeirão Preto, SP, Brazil

ABSTRACT

Aim The aim of this study was to assess the causal mechanisms underlying populational subdivision in *Drosophila gouveai*, a cactophilic species associated with xeric vegetation enclaves in eastern Brazil. A secondary aim was to investigate the genetic effects of Pleistocene climatic fluctuations on these environments.

Location Dry vegetation enclaves within the limits of the Cerrado domain in eastern Brazil.

Methods We determined the mitochondrial DNA haplotypes of 55 individuals (representing 12 populations) based on sequence data of a 483-bp fragment from the cytochrome *c* oxidase subunit II (COII) gene. Phylogenetic and coalescent analyses were used to test for the occurrence of demographic events and to infer the time of divergence amongst genetically independent groups.

Results Our analyses revealed the existence of two divergent subclades (G1 and G2) plus an introgressed clade restricted to the southernmost range of *D. gouveai*. Subclades G1 and G2 displayed genetic footprints of range expansion and segregated geographical distributions in south-eastern and some central highland regions, east and west of the Paraná River valley. Molecular dating indicated that the main demographic and diversification events occurred in the late to middle Pleistocene.

Main conclusions The phylogeographical and genetic patterns observed for *D. gouveai* in this study are consistent with changes in the distribution of dry vegetation in eastern Brazil. All of the estimates obtained by molecular dating indicate that range expansion and isolation pre-dated the Last Glacial Maximum, occurring during the late to middle Pleistocene, and were probably triggered by climatic changes during the Pleistocene. The current patchy geographical distribution and population subdivision in *D. gouveai* is apparently closely linked to these past events.

Keywords

Cactaceae, Cerrado, Diptera, *Drosophila*, phylogeography, Pleistocene, South America, xerophytic plants.

*Correspondence: Evandro M. Moraes, Universidade Federal de São Carlos, Campus Sorocaba, Sorocaba, SP, Brazil. E-mail: emarsola@ufscar.br

INTRODUCTION

In recent years, considerable effort has been devoted to studying the genetic effects of climatic changes and glaciations during the Pleistocene on the diversification of many temperate species in the Northern Hemisphere (e.g. Hewitt, 2004).

Most of these studies have investigated the geographical distribution of gene lineages and have shown a significant role of past climatic changes in shaping the history and population structure of terrestrial species. In contrast, much less attention has been given to the role of Pleistocene climatic changes in the diversification of species from the unglaciated, high-diversity

areas of tropical South America (Carnaval & Moritz, 2008), particularly for species restricted to dry areas in eastern Brazil (Caetano *et al.*, 2008).

This region, known as the 'dry diagonal', encompasses a large belt of dry forest that crosses Brazil in a north-eastern to south-western direction and includes three distinct domains (fig. 1 in Prado & Gibbs, 1993). Here, we use the term 'domain' to refer to a geographical area where particular morphoclimatic and phylogeographical features predominate. This dry-diagonal forest belt extends from the Caatinga domain in north-eastern Brazil to the Uruguay River valley, and occupies a position between the Amazon and Atlantic rain forests. The northern part of the dry diagonal is dominated by Caatinga vegetation, which is composed of xerophytic plants and a high density and diversity of cacti species that survive in a semiarid climate. The central and southern parts of the diagonal are covered by the Cerrado, the second largest biome in Brazil (> 2 million km²), which is dominated by a semi-deciduous savanna (Silva & Bates, 2002). The Cerrado domain harbours a mosaic of xeric vegetation patches, generally called *campos rupestres*, which occur as enclaves in areas of edaphically favourable conditions on rocky outcrops (Silva & Bates, 2002). These habitats are limited to small patches on the plateaus or mountain tops, and contain many drought-adapted plant species and a highly endemic flora (Giulietti *et al.*, 1997). The southernmost portion of the diagonal is occupied by the Chaco domain, where xeric vegetation predominates, and extends through Argentina, Bolivia and Paraguay to the eastern slopes of the Andes.

The current distribution of open/dry vegetation in eastern Brazil, including that present along the Atlantic coast, has been thought to represent the remnants of a distribution that was more extensive during the glacial periods of the Pleistocene (Ab'Saber, 1982; Bigarella & Andrade-Lima, 1982; Prado & Gibbs, 1993; Pennington *et al.*, 2000; Behling *et al.*, 2007). During the Pleistocene, the Neotropics experienced phases of climatic fluctuations, with cooler conditions and longer dry seasons than those of today in some regions (Behling & Lima da Costa, 2001; Mayle *et al.*, 2004; Ledru *et al.*, 2005). Recent palynological, palaeoecological, inventory, palaeomodelling, and phylogeographical studies have revealed evidence of fluctuations in the dry vegetation during the Pleistocene, when a drier climate favoured the expansion of drought-adapted species in eastern Brazil (Behling, 2002; Ledru *et al.*, 2006; Werneck & Colli, 2006; Behling *et al.*, 2007; Safford, 2007; Caetano *et al.*, 2008; Carnaval & Moritz, 2008).

Species historically associated with patches of xeric habitat in eastern Brazil may have experienced recurrent shifts in their distribution, with population mixing and isolation shaping their genetic structure and evolutionary history. For this reason, the enclaves of xeric vegetation within the Cerrado domain provide an unusually promising context within which to explore the evolutionary consequences of Pleistocene climatic changes in species restricted to xeric conditions. The mountainous topography in some parts of the region has created many discrete populations that allow multiple inde-

pendent observations of the effects of climatic change. Because many of these populations are surrounded by habitats unsuitable for dispersion, it is likely that they retain a signature of past distributional changes in their population genetic structure.

The *Drosophila buzzatii* cluster (*repleta* group) (Diptera, Drosophilidae) is a particularly suitable taxon with which to study the effects of Pleistocene climatic changes on population structure and species history in eastern Brazil. This taxon consists of seven related cactophilic species of flies endemic to South America [except for *D. buzzatii* (Patterson & Wheeler, 1942), which is a colonizing, subcosmopolitan species]. The species in this cluster breed exclusively in decaying cactus tissue, which therefore restricts their populations to xerophytic habitats where cacti grow. Because of this association, the evolutionary history of species in the *D. buzzatii* cluster has been used as a model with which to study the historical changes in the geographical distribution of dry vegetation caused by palaeoclimatic oscillations during the last 3 Myr in South America (Manfrin & Sene, 2006, and references therein).

Drosophila gouveai (Tidon-Sklorz & Sene, 2001) is a member of the *D. buzzatii* cluster. Populations of *D. gouveai* are distributed from the Caatinga domain in north-eastern Brazil to the southern limit of the Cerrado domain in south-eastern Brazil (Silva & Sene, 1991), partially following the 'dry diagonal'. Populations of *D. gouveai* found outside the Caatinga domain occur in patches of xeric vegetation on summits of isolated hills or rocky outcrops associated with species of the cactus genus *Pilosocereus* and form numerous discrete populations (Silva & Sene, 1991; Tidon-Sklorz & Sene, 2001; de Brito *et al.*, 2002; Moraes & Sene, 2007). Populations from the Caatinga domain occur in more extensive areas of dry forest in which a number of cactus species belonging to the genera *Cereus*, *Opuntia*, *Pilosocereus* and *Facheiroa* are generally present (Pereira *et al.*, 1983; Moraes & Sene, 2007). All populations of *D. gouveai* are taxonomically characterized as having a morphotype B aedeagus (Silva & Sene, 1991; Tidon-Sklorz & Sene, 2001), chromosomal inversion 2e⁸ (Tosi & Sene, 1989) and karyotype I (Baimai *et al.*, 1983). As a consequence of the association with decaying cacti, differentiation amongst *D. gouveai* populations can be closely linked to the expansion and retraction of xeric vegetation in eastern Brazil during past climatic changes (Manfrin & Sene, 2006).

Microsatellite DNA and wing morphology analyses have shown that *D. gouveai* populations are highly structured (Moraes & Sene, 2007). This populational structure is attributed mainly to genetic differences amongst geographically distinct population groups, and assignment tests indicate that current gene flow amongst and within groups is very restricted. Based on these results and additional evidence, Moraes & Sene (2007) suggested that the genetic similarity amongst populations owed more to common ancestry than to ongoing genetic exchange.

The populational structure and evolutionary history of *D. gouveai* have also been investigated by analysing the mitochondrial DNA (mtDNA) cytochrome *c* oxidase subunit

I (COI) gene (Manfrin *et al.*, 2001; de Brito *et al.*, 2002). de Brito *et al.* (2002) pointed to north-eastern Brazil as the geographical origin of *D. gouveai* and found evidence for either one or two events of southward expansion in the species' range. *Drosophila gouveai* populations from the southernmost limit of the species' distribution appear to have been involved in different evolutionary events. These populations have morphological and chromosomal attributes representative of the remaining *D. gouveai* populations. However, all of the individuals analysed had COI haplotypes characteristic of *D. antonietae* (Manfrin *et al.*, 2001; de Brito *et al.*, 2002), a sibling species of the *D. buzzatii* cluster with a northern distribution limit that is *c.* 20 km from the southernmost populations of *D. gouveai*. This genetic relatedness probably resulted from secondary contact between these two species following population range expansion in this region (Manfrin *et al.*, 2001; de Brito *et al.*, 2002). Interestingly, these populations from the southernmost distribution range of *D. gouveai* appeared as outliers in microsatellite DNA and wing morphology analyses (Moraes & Sene, 2007). It is worth noting that other *D. gouveai* populations close to the southernmost introgressed populations do not exhibit mtDNA haplotypes related to *D. antonietae*, indicating a lack of gene exchange amongst neighbouring populations at the southern limit of the distribution of *D. gouveai*.

In this report, we present a new genealogy of the mtDNA cytochrome *c* oxidase subunit II (COII) gene based on specimens collected from the three main regions within the distribution of *D. gouveai* (southern border of the Caatinga, central Brazil, and southern border of the Cerrado). We used a variety of phylogenetic and coalescent analyses to test the occurrence of demographic events and to infer divergence time

amongst genetically independent groups previously identified by molecular and morphometric markers (Moraes & Sene, 2007). These data were also used to investigate the potential role of Pleistocene climatic changes as a causal mechanism underlying populational subdivision in *D. gouveai*.

MATERIALS AND METHODS

Sample and site details

Fifty-five flies of *D. gouveai* were collected from January to September of 2000 at 12 localities in eastern Brazil across the species' known range (see Fig. 1 for geographical locations and Table 1 for details of the geographical coordinates, population codes and sample sizes). The flies analysed here were a subsample of those studied by Moraes & Sene (2007). Most localities sampled were small patches (*c.* 5–10 ha) of xerophytic vegetation within the Cerrado domain, which were located at the centre, or close to the north-eastern or southern limits of this geographical area. Because *D. gouveai* occurs in these very discrete patches, geographical samples are likely to represent distinct populations. For this reason, throughout this paper, sample localities are referred to as populations. The *D. gouveai* populations at localities CCZ, CAM, FOR and MIR in the southern portion of the species' range occurred on small summits (generally < 5 ha) of isolated sandstone hills. CCZ, CAM and FOR were associated with the host cactus *Pilosocereus machrisii* (E.Y. Dawson) Backeb., and at the MIR locality, host species from the genus *Cereus* were also present. The localities FUR, CEU, CRI, FER and PIR were areas of rocky outcrops (*c.* 10 ha) and, with the exception of FER and PIR, were also associated with the host cactus *P. machrisii*. At PIR,

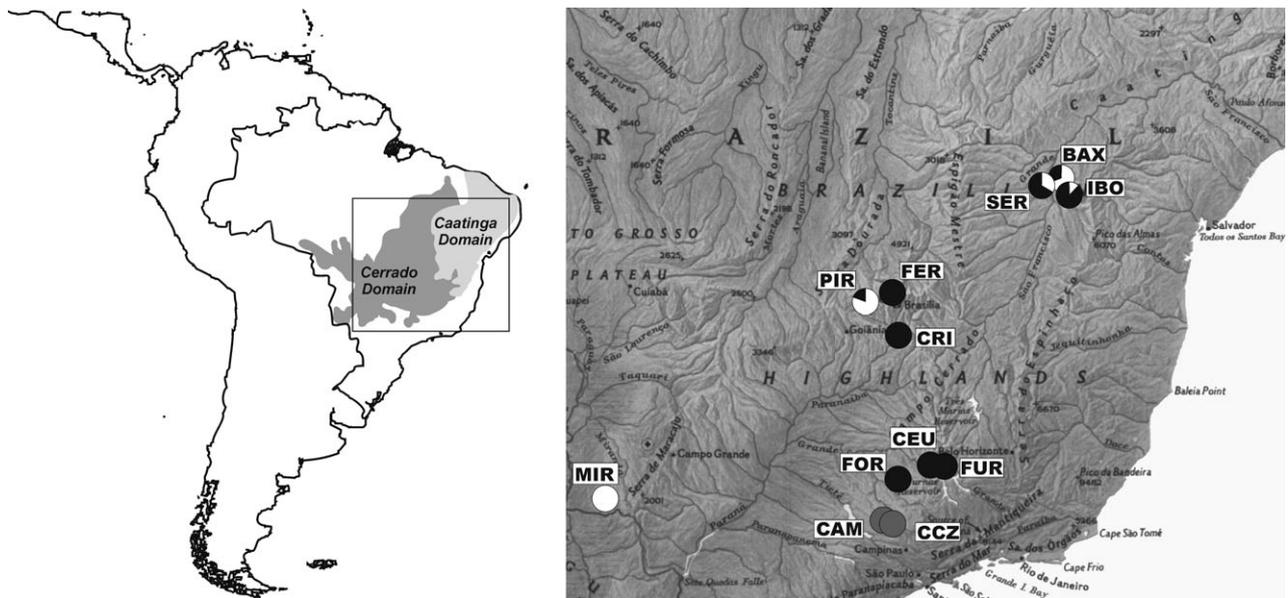


Figure 1 Map of South America showing the Caatinga and Cerrado domains and the localities sampled for *Drosophila gouveai* in this study. The colour and extent of filling of the circles indicate the proportion of haplotypes of the haplogroup Antonietae (grey) and of the subclades G1 (black) and G2 (white) of the haplogroup Gouveai in each population. Abbreviations correspond to the localities listed in Table 1.

Table 1 Sampling information, sample sizes (n) and summary of the DNA polymorphism measures per sample locality and for each haplogroup of *Drosophila gouveai* in eastern Brazil. The number of haplotypes (H), haplotype diversity (H_d), nucleotide diversity (π_1), and mean number of pairwise differences (π_2) are shown.

Code	Locality (Country, State)	Coordinates	n	H	H_d	π_1	π_2
Species level			55	35	0.956	0.019 ± 0.002	9.49 ± 4.33
Haplogroup Antonietae			7	5	0.857	0.006 ± 0.004	2.95 ± 1.75
CCZ	Morro do Cuscuzeiro (Analândia, SP)	22°08' S 47°39' W	4	2	0.500	0.005 ± 0.004	2.50 ± 1.68
CAM	Morro do Camelo (Analândia, SP)	22°09' S 47°42' W	3	3	1.000	0.006 ± 0.005	3.33 ± 2.32
Haplogroup Gouveai			48	30	0.944	0.011 ± 0.000	5.79 ± 2.76
FOR	Morro do Forno (Altinópolis, SP)	21°02' S 47°20' W	6	2	0.333	0.000 ± 0.000	0.33 ± 0.32
FUR	Furnas (Furnas, MG)	20°37' S 46°15' W	6	6	1.000	0.008 ± 0.005	3.93 ± 2.29
CEU	Vale do Céu (Furnas, MG)	20°32' S 46°18' W	4	2	0.500	0.001 ± 0.001	0.50 ± 0.51
MIR	Pirapotanga (Aquidauana, MS)	20°21' S 55°32' W	3	2	0.667	0.001 ± 0.001	0.66 ± 0.66
CRI	Cristalina (Cristalina, GO)	16°43' S 47°40' W	4	3	0.833	0.002 ± 0.002	1.16 ± 0.92
PIR	Pirenópolis (Pirenópolis, GO)	15°48' S 48°51' W	5	5	1.000	0.012 ± 0.008	6.20 ± 3.54
FER	Fercal (Brasília, DF)	15°30' S 47°57' W	3	3	1.000	0.004 ± 0.003	2.00 ± 1.51
BAX	Baxio (Ibotirama, BA)	11°56' S 43°09' W	3	3	1.000	0.015 ± 0.012	7.33 ± 4.52
IBO	Ibotirama (Ibotirama, BA)	12°04' S 43°19' W	7	5	0.857	0.008 ± 0.005	4.19 ± 2.36
SER	Serrinha (Ibotirama, BA)	12°08' S 43°22' W	7	6	0.952	0.013 ± 0.008	6.57 ± 3.53

the host cactus species was *P. vilaboensis* (Dier and Esteves) Brawn, and at the FER locality the host species was of the genus *Cereus*. The remaining populations at BAX, IBO and SER, in north-eastern Brazil and the southern border of the Caatinga, occurred in larger areas (> 10 ha) of dry forest where a number of cactus species belonging to the genera *Cereus*, *Opuntia*, *Pilosocereus* and *Facheiroa* were generally present.

Molecular methods

The total DNA of each ethanol-preserved *D. gouveai* individual and of one individual of *D. antonietae* was isolated using the Promega DNA extraction kit (Promega Corporation, Madison, WI, USA). A 683-bp fragment of COII was amplified using the primers TL2-J-3034 and TK-N-3785, previously described by Simon *et al.* (1994). Amplification by polymerase chain reaction (PCR) was carried out in a final volume of 25 μ l using 1 U of *Taq* DNA polymerase (Invitrogen Corporation, Carlsbad, CA, USA) with primer annealing at 57°C. The PCR products were sequenced in both directions with the same primers as used in the PCR protocol. Contigs of forward and reverse sequences were formed using CONSED (Gordon *et al.*, 1998), with which the sequences were evaluated and edited.

Because of the poor quality of the tips in some sequences, only a central 483-bp fragment corresponding to positions 3173–3655 of the *D. yakuba* mtDNA sequence (Clary & Wolstenholme, 1985) was used in the analysis. The COII sequences were aligned by hand using MEGA 3.0 (Kumar *et al.*, 2004).

Genetic data analyses

Diversity parameters, including nucleotide (π_1) and haplotype (H_d) diversities, mean number of pairwise differences (π_2) and the number of distinct haplotypes (h), were estimated for each sample locality, for all individuals combined, and for each primary group of haplotypes uncovered by the phylogenetic and network analyses (see Results) using ARLEQUIN 3.0 (Excoffier *et al.*, 2005).

Phylogenetic relationships amongst all unique mtDNA haplotypes were estimated by the maximum likelihood (ML) approach using PAUP* 4.0b10 (Swofford, 2002). For ML analysis, an appropriate nucleotide substitution model that best fitted our dataset was selected using the minimum theoretical Akaike information criterion (AIC) as implemented in MODELTEST 3.06 (Posada & Crandall, 1998). The ML analyses were then implemented using a heuristic search algorithm and tree bisection–reconnection (TBR) with the

parameter estimates obtained under the best-fitting model of nucleotide substitution. Branch support on the ML tree was assessed by 100 bootstrap replications using a neighbour-joining starting tree. The haplotype sequences were also analysed with a Bayesian inference using a Markov chain Monte Carlo (MCMC) approach as implemented in MRBAYES 3.1 (Ronquist & Huelsenbeck, 2003). Because different codon positions evolve under different rates, the Bayesian phylogenetic analysis was performed under different codon-specific models. Prior models for each codon position were inferred using MRMODELTEST 2.3 (Nylander, 2008). The selected codon-specific models were Jukes–Cantor for codon position 1 and general time-reversible (GTR) for codon positions 2 and 3. The Bayesian analysis was also performed under a single substitution model, also inferred by MRMODELTEST 2.3 (Nylander, 2008), which selected a GTR model with a proportion of invariable sites of 66% and a gamma-distribution shape parameter of 0.09. Two separate runs, each containing one cold and three heated chains, were processed simultaneously for 10^6 generations, each starting from a random tree. The trees were sampled every 100 generations (10^4 samples). This procedure was used to infer the tree topology and its branch lengths. In order to avoid autocorrelation between estimated parameters and tree topologies, tree search was also performed under the single substitution model for 10^7 generations, sampling trees every 10^3 generations. In addition, in order to ensure that the heated chains sampled all the likelihood area, this analysis was performed using four distinct temperatures (0.2, 0.4, 0.6, and 1.0) for the heated chains. After running these analyses, 25% of the initial generations were burned and the remaining topologies were used to build a consensus tree for each analysis. The resulting trees were rooted using two sequences of *D. buzzatii* (GenBank accession numbers DQ202011, AF146169), the least-derived species within the *D. buzzatii* cluster species as revealed by previous phylogenetic analyses (Manfrin & Sene, 2006). In order to investigate the phylogenetic relationship between *D. gouveai* and *D. antonietae* haplotypes, one sequence of *D. antonietae* was included in the analyses. Because *D. antonietae* appears as a single clade in a previous phylogenetic analysis (Manfrin *et al.*, 2001), a single individual was enough for this purpose.

To explore the level of connectedness amongst *D. gouveai* haplotypes further we created a haplotype network using a statistical parsimony (SP) method in rcs 1.13 (Clement *et al.*, 2000), with connections constrained by a 95% limit. On the basis of this statistical parsimony network, a nested clade analysis (NCA, Templeton, 1998) was used to test for association between the genealogical relationships and geographical distributions of the haplotypes. The clades were nested by hand, and GEODIS 2.0 (Posada *et al.*, 2000) was then used to test for a significant association between the resulting nested design and the geographical distances between clades. When the analysis indicated significant genealogical–geographical associations, we used the inference key provided by Templeton (2004) to make inferences about the historical

causes and recurrent forces influencing the evolutionary history of nested clades.

The occurrence of historical demographic events, such as population growth or range expansion, was also tested in independent clades uncovered by ML and Bayesian phylogenetic analyses (subclades G1 and G2; see Results) using two approaches. Initially, population equilibrium was assessed by applying Tajima's *D*-test (Tajima, 1989) and Fu's *F_s*-test (Fu, 1997). The demographic history of *D. gouveai* was further inferred by testing deviation from the expected mismatch distributions under the sudden-expansion model, using the generalized least-squares method. In this method, if the distribution fits the expansion model then the time elapsed since the demographic event can be estimated by the parameter τ , which is the date of the demographic event measured in units of mutational time (Rogers & Harpending, 1992). All historical demographic analyses were performed in ARLEQUIN 3.0 (Excoffier *et al.*, 2005), which provided the 95% confidence interval by a parametric bootstrapping approach.

The divergence time between pairs of populations from distinct geographical regions was examined under the isolation-with-migration (IM) model, which explicitly incorporates parameters for time of population splitting, bidirectional gene flow, and population sizes, including the size of the ancestral population (Hey & Nielsen, 2004). Because evidence of population expansion was provided by neutrality tests and mismatch analyses, a divergence model allowing for independent changes in population sizes (Hey, 2005) was applied. The model was fitted using a Bayesian framework that provides estimates for the posterior probability density of the model parameters, given the data, as implemented in the IM program (Hey & Nielsen, 2004). All parameters in the IM model are scaled by the neutral mutation rate (see above). Wide uninformative prior distributions were assigned from multiple test runs of the IM program, ensuring that the posterior distribution fell completely within the prior distribution. To mitigate potential problems related to small sample size, populations belonging to haplogroup Gouveai (see Results) were grouped into three regional groups previously defined based on microsatellite DNA and morphometric analyses (Moraes & Sene, 2007), namely northern (BAX, IBO, SER), central (FER, PIR, CRI) and southern (CEU, FUR, FOR). Population MIR, a southern population, was not included in the southern group because it appeared as an outlier within this group (see Results).

Three independent MCMC runs per data set with the same prior distributions and different seed numbers were carried out using the Hasegawa–Kishino–Yano mutation model (HKY). Each was run for 10^8 generations and preceded by a burn-in period of 10^7 generations. Under these conditions, autocorrelations amongst parameters were low ($r < 0.03$ for all parameters), and effective sample sizes of all parameters exceeded 300 (generally up to 1000), suggesting adequate exploration of the sample space. The peaks of the posterior distributions were taken as ML estimates of the parameters, and credibility intervals (CIs) were assessed from the 90%

highest posterior density (HPD) interval (i.e. the boundaries of the shortest span that incorporates 90% of the posterior density of a parameter). Because all three runs for each dataset gave very similar parameter estimates and 90% HPDs, we report results of divergence time (T_{DIV}) from the longest run only. Estimates of time to the most recent common ancestor (T_{MRCA}) were also used to date the coalescence times of major clades within the phylogenetic tree based on the methods described above.

The scaled divergence time provided by the IM program was converted to years by assuming 10 generations a year, a generation time commonly experienced in laboratory for a large number of *Drosophila* species. For all of these dating analyses we used a mutation rate (μ) of 5.88×10^{-7} mutations per sequence per generation. This mutation rate was obtained by estimating the genetic distance (using the Kimura 2-parameter correction) of the 483-bp fragment analysed here between the Hawaiian drosophilids *D. picticornis* (GenBank accession number AY006430) and the species *D. planitibia* (AY006438), *D. differens* (AY006434), *D. heteroneura* (AY006434) and *D. silvestris* (AY006433). The time of divergence between *D. picticornis* (endemic to Kauai Island; Carson & Yoon, 1982) and the remaining species was assumed to be 5.1 Ma based on the geological dating of the formation of Kauai (Rowan & Hunt, 1991). The substitution rate estimated here was within the range of 1–2% of substitutions per million years of the rate determined for other *Drosophila* species (Russo *et al.*, 1995).

RESULTS

Drosophila gouveai mtDNA contained 51 polymorphic sites, 33 of which were parsimony-informative, and defined 35 unique haplotypes (GenBank accession numbers EU7161760–EU716210), with an overall haplotypic diversity of 0.956 and a nucleotide diversity of 0.019 ± 0.002 (Table 1). The sample size was 3–7 individuals amongst sample localities, with 2–6 distinct haplotypes within each locality. The northern populations, along with the central population PIR, had the highest nucleotide diversity and the largest mean number of pairwise differences.

Phylogenetic and network analyses

The ML and Bayesian methods generated essentially identical tree topologies. Hence, only the tree topology and branch lengths inferred by the Bayesian approach are presented here, along with the posterior probabilities of each node and the bootstrap values obtained by ML analysis (Fig. 2). We also ran a separate Bayesian analysis that included only the haplogroup Gouveai (see Results below; Fig. 2). The substitution model inferred from this analysis was very similar to that inferred for all haplotypes, but with a proportion of invariable sites of 82% and a gamma-distribution shape parameter of 0.07. All Bayesian analyses reached the same topology, with only minor differences in some posterior probability node supports, meaning that the phylogenetic relationships presented were

robust, supporting the existence of different clades and subclades in *D. gouveai* (see below).

The statistical parsimony method used in the network analysis with a 95% reconnection limit between haplotypes produced two subnetworks that corresponded to the same primary clades as revealed by the phylogenetic analysis (Fig. 3). The following observations can be drawn from both the phylogenetic and network analyses.

1. The 35 haplotypes of *D. gouveai* formed two highly supported clades with an average nucleotide divergence of 4.6%. These clades are hereafter referred to as haplogroups Gouveai and Antonietae. In agreement with the Bayesian tree, the SP network also resolved the haplotypes of *D. gouveai* as two subnetworks separated by a minimum of 19 mutational steps between haplotypes H35 (haplogroup Antonietae) and H30 or H10 (haplogroup Gouveai).
2. Haplogroup Antonietae consisted exclusively of all haplotypes found in the southernmost populations (CCZ and CAM, Fig. 1). These haplotypes were closely related to the sequence of *D. antonietae*, with haplotype H35 being identical to that species.
3. Most haplotypes in haplogroup Gouveai were clustered in a highly supported subclade (G1), and the remaining set of haplotypes appeared to form a polytomy. However, this polytomy was completely resolved in an additional Bayesian analysis that included only the haplogroup Gouveai, and revealed a second highly supported subclade (G2; insert box A in Fig. 2). These subclades had partially overlapping distributions in north-eastern and central regions (Fig. 1) and were separated by a net average nucleotide distance of 1.2%.
4. Subclade G1 had a wide geographical distribution: its haplotypes were found in all populations except for MIR. Most of the haplotypes in subclade G1 were restricted to a unique population.
5. Haplotypes from subclade G2 were restricted to northern populations, to population PIR in the central region, and to the MIR population. The latter population contained only haplotypes from subclade G2, whereas the northern populations BAX, IBO and SER, and the central population PIR, contained haplotypes from subclades G1 and G2. In the SP network, subclade G2 (clade 3-3 of NCA) appeared as a 'star-shaped network' around haplotype H10, the most frequent and widespread haplotype.

Nested clade analysis

Because haplotypes within haplogroup Antonietae were restricted to two nearby localities, NCA was carried out only on the network connecting haplotypes of the haplogroup Gouveai (Fig. 3). For all one-step clades, NCA could not reject the null hypothesis of no geographical association. Within higher-level clades, the analysis revealed significant support for restricted gene flow, with isolation by distance for clades 2-1, 3-1 and 3-2. These clades had overlapping distributions, occurred in all of the geographical regions studied, and were connected to one another by a low number of mutational

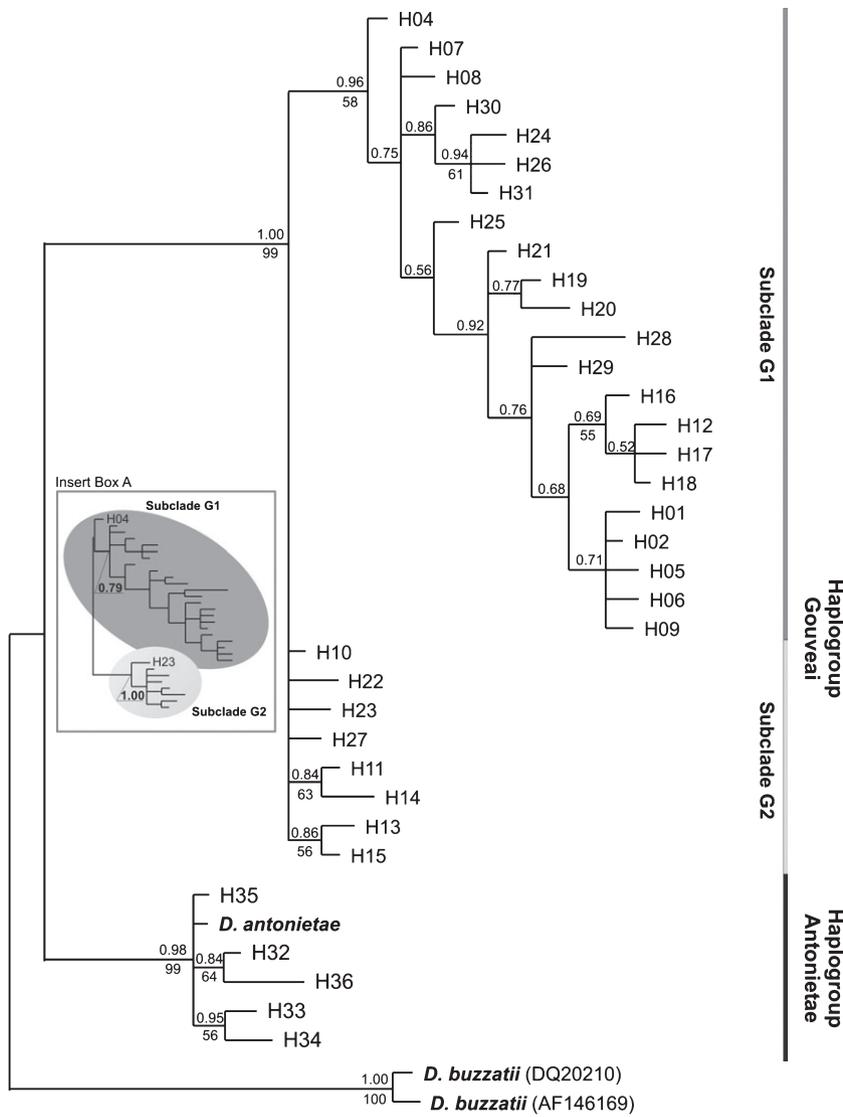


Figure 2 Bayesian phylogenetic tree for *Drosophila gouveai* based on unique haplotypes of the 483 bp of the mtDNA COII gene. Haplotypes from *D. buzzatii* and *D. antonietae* were also included. Clades and subclades are identified by vertical lines and the corresponding name. The haplotype names are shown on the terminal nodes. The nodal indices are bootstrap supports (above) and Bayesian posterior probabilities (below). The insert box shows an additional Bayesian phylogenetic tree that includes only haplotypes from the haplogroup Gouveai.

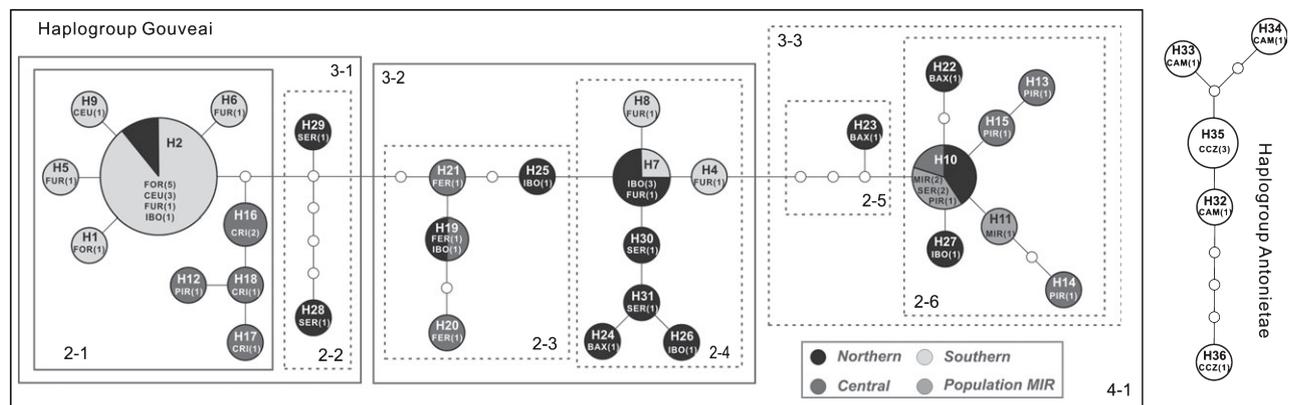


Figure 3 Statistical parsimony haplotype networks (95% limit of statistical parsimony) for *Drosophila gouveai*. The haplogroups (Gouveai and Antonietae) revealed by phylogenetic analyses are indicated. The sampled haplotypes are indicated by circles, with size proportional to the number of individuals bearing the haplotype. The colour and extent of filling of the circles indicate the partition of haplotype numbers amongst regions. The codes (see Table 1) and numbers in parentheses indicate the population where each haplotype was found and the number of individuals that were involved. Small open circles indicate missing haplotypes that were not observed, and the line between them represents a single mutational step. The figure also shows the nested clade design for 2, 3 and 4 steps in haplogroup Gouveai. Dotted lines demarcate clades for which the nested clade analysis yielded non-significant results.

Table 2 Summary of historical demographic expansion analyses for *Drosophila gouveai* in eastern Brazil.

Phylogenetic groups	Neutrality tests		Mismatch analysis		Expansion dates (ka) (95% CI)
	D-test	F_S -test	SSD (P-value)	τ (95% CI)	
Subclade G1	-0.994 ^{NS}	-13.698*	0.012 (0.91)	5.45 (2.60–8.25)	463 (221–701)
Subclade G2	-1.704 ^{NS}	-4.136*	0.036 (0.89)	2.06 (0–3.69)	175 (0–313)

D-test, Tajima's D -test (Tajima, 1989); F_S -test, Fu's F_S -test (Fu, 1997); SSD, sum of squared deviations between the observed and the expected mismatch; τ , scaled time elapsed since the demographic event; NS, nonsignificant; * P -value < 0.01; ka, thousands of years ago; CI, confidence interval.

steps. NCA also indicated that the phylogeographical structure of clade 4-1, the highest-level clade corresponding to the entire haplogroup Gouveai, could be explained by a contiguous range expansion.

Population expansion and divergence timing

Neutrality tests and mismatch analyses were performed separately for subclades G1 and G2 (Table 2). Tajima's D -test did not differ significantly from zero when subclades G1 and G2 were considered separately. However, the highly significant negative values of the F_S -test indicated an excess of recent mutations (Fu, 1997) and supported the occurrence of population range expansion for each of the two phylogenetic groups. In agreement with the F_S -tests, mismatch distributions showed unimodal distributions for the phylogenetic groups analysed (Fig. 4), with no significant deviation between the observed and expected distributions (Table 2). Estimates of the parameter τ were used to determine the time elapsed since the demographic events based on a mutation rate (μ) of 5.88×10^{-7} mutations per sequence per generation, and 10 generations a year (Table 2). Based on these calculations, the population expansion event in subclade G1 was estimated to have occurred during the middle Pleistocene, *c.* 450 ka, whereas in subclade G2 the expansion event was estimated to have occurred *c.* 170 ka. However, because the confidence interval includes zero (0–0.313), we are unable to discriminate between Holocene and Pleistocene events for subclade G2.

The coalescent-based analysis conducted under the isolation-with-migration model indicated that the time to most recent common ancestor (T_{MRCA}) shared by subclades G1 and G2 was 586 ka (90% HPD = 0.892–0.375), which can be interpreted as the initial timing of diversification of *D. gouveai*. Estimated divergence times (T_{DIV}) between population pairs were *c.* 140 ka (90% HPD = 0.300–0.054) between southern and central populations, *c.* 88 ka (90% HPD = 0.260–0.024) between southern and northern populations, and *c.* 173 ka (90% HPD = 0.565–0.073) between central and northern populations.

DISCUSSION

Drosophila gouveai consists of two main clades, referred to as haplogroups Gouveai and Antonietae. Haplogroup Antonietae is restricted to the southernmost populations of the distribution range of *D. gouveai* and consists exclusively of haplotypes

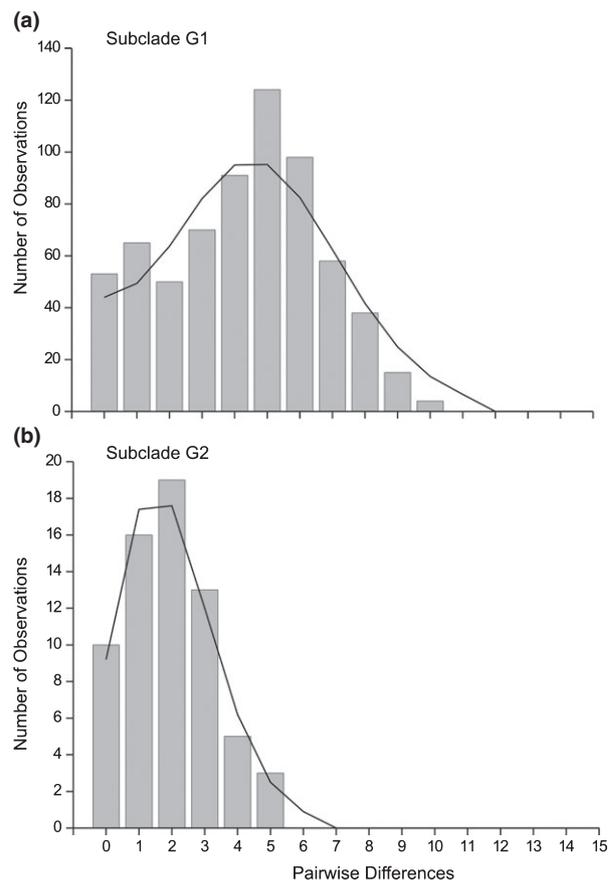


Figure 4 Pairwise mismatch distribution of the mtDNA sequences for (a) subclade G1 and (b) subclade G2 of haplogroup Gouveai of *Drosophila gouveai*. Solid histograms represent the observed differences, and the lines represent the expected distributions compatible with a sudden-expansion model.

closely related to the sequence of *D. antonietae*. These results produced an identical phylogenetic pattern to that previously described for COI sequences from these populations (Manfrin *et al.*, 2001; de Brito *et al.*, 2002). Because haplogroup Antonietae probably arose by introgression of the mtDNA genome of *D. antonietae* into these populations (Manfrin *et al.*, 2001), only haplogroup Gouveai should bear genetic footprints of demographic events that occurred in *D. gouveai*.

Haplogroup Gouveai contains subclades G1 and G2, which are partially segregated geographically. These subclades have overlapping geographical distributions in north-eastern Brazil

and at a single locality (PIR) in central Brazil; there is no overlap in the south-eastern region. The coalescence and neutrality-based tests indicate range expansion for each subclade, whereas NCA suggests range expansion for all of haplogroup Gouveai (clade 4-1 in the SP network). The widespread distributions of some internal haplotypes, such as haplotypes H2 (subclade G1) and H10 (subclade G2), are also indicative of a rapid range expansion. de Brito *et al.* (2002) also found evidence for two events of range expansion in *D. gouveai* and, based on the geographical distribution of the ancestral haplotypes, concluded that these expansions occurred from the north-eastern to the southern populations. A north-to-south expansion range is corroborated by the higher nucleotide diversity of the northern populations in our study, which suggests that the northern areas are representatives of the ancestral populations of *D. gouveai*, as previously observed by de Brito *et al.* (2002).

Although subclades G1 and G2 have been interpreted as two lineages that experienced independent demographic events, they could represent divergent lineages that were present in a single polymorphic expanding population. However, the segregated geographical distributions of subclades G1 and G2 in south-eastern and some central regions indicate independent expansions. Because the host cactus species of *D. gouveai* generally occur on summits of Brazilian highlands, the expansion of these xerophytic plants and, consequently, of *D. gouveai* populations probably occurred along the axes of the major mountain streams. Accordingly, a tentative hypothesis that matches the distribution of these subclades is that the expansion of subclade G1 occurred from north-eastern populations to southern areas along mountain tops east of the Paraná River valley, whereas range expansion in subclade G2 occurred independently from highlands in central Brazil southwards, via mountain tops west of the Paraná River valley. This hypothesis would predict that all populations occurring west of the Paraná River valley in central Brazil should have only haplotypes related to subclade G2.

There are no reciprocally monophyletic populations within haplogroup Gouveai. However, haplotypes are differentially sorted following the distribution of regional groups of populations. The haplotype network indicates that most haplotypes occurring within a geographical region are closely related. This finding is similar to that revealed by microsatellite DNA analysis, in which *D. gouveai* populations were clustered into northern, central and southern groups (Moraes & Sene, 2007). A geographical association amongst haplotypes is supported by NCA for haplotypes clustered within subclade G1 (clades 2-1, 3-1 and 3-2 of the SP network). When populations show incomplete lineage sorting, it is difficult to determine whether populations differ because of differences in the levels of isolation or because of migration. A given level of population structure may be the result of an ancient divergence followed by a more recent exchange of genes, or it may reflect a more recent divergence with little subsequent gene flow.

The haplotype network of *D. gouveai* suggests very restricted gene flow amongst populations. Accordingly, the more geo-

graphically widespread haplotypes (H2, H7 and H10) are located internally or at the centre of the gene tree, whereas haplotypes at the tips of the gene tree have a more restricted geographical distribution. The population pairwise analysis from the IM model also indicates the absence of gene flow amongst geographical regions (data not shown). All comparisons in those analyses provide estimates of migration rate per generation near to zero (m -values $< 10^{-5}$, including the upper values of credibility intervals). Therefore, persistence of ancestral polymorphism, rather than recent gene flow, is a more likely scenario for describing the association of the majority of haplotypes within haplogroup Gouveai. The absence of gene flow amongst regional groups of *D. gouveai* seen here is corroborated by microsatellite DNA analysis (Moraes & Sene, 2007) and agrees with the patchy distribution of *D. gouveai*.

It is important to point out one caveat with our sampling. The sample size was 3–7 individuals amongst the 12 sample localities. Amongst the 35 distinct haplotypes sampled, 29 were present in single individuals, and only four were shared by individuals from different localities. The small sample size means that singletons may not necessarily be rare, but could occur with a higher frequency in their local populations. However, the small number of shared haplotypes is unlikely to have been caused by limited sampling, as no significant linear trend between the number of unique haplotypes and sample size per locality was evident ($r = 0.39$; $P = 0.2036$). This indicates that the geographical association amongst haplotypes probably was not an artefact of sample size and that the segregated geographical distribution of subclades G1 and G2 in some regions could, in fact, be caused by idiosyncratic demographic events.

Based on the evidence of range expansion and the current patchy distribution of *D. gouveai*, the phylogeographical structure seen in haplogroup Gouveai could have resulted from isolation following a reduction in the distributional range of dry vegetation. The evidence for past ecological links and range contraction in plant and animal taxa currently restricted to enclaves of open/dry vegetation within the Cerrado domain is strengthened by inventory studies that have compared species assemblages in these environments (Werneck & Colli, 2006; Safford, 2007), as well as by palynological (Behling *et al.*, 2007), phylogenetic (Almeida *et al.*, 2007) and phylogeographical (Caetano *et al.*, 2008) approaches. Caetano *et al.* (2008) investigated the genetic variation in DNA sequences of chloroplast spacers and microsatellite loci of the tree *Astronium urundeuva*, a species restricted to seasonally dry tropical forests (SDTF) of eastern South America. In the Cerrado domain, *A. urundeuva* shows a distribution pattern similar to that of *D. gouveai*, occurring in enclaves of dry forest, but with the two species present in different localities and vegetation types. They found that the populations are structured in three regional groups (Northeast, Central and Southwest) and suggested that the fragmented distribution of *A. urundeuva* is the result of vicariance of a previously more continuous formation of SDTF in eastern South America. Furthermore,

the chloroplast variation in *A. urundeuva* provided signals of posterior southward and northward range expansions leading to secondary contact in the Central group. Taking into account the partial overlap in the distribution range, along with the occurrence in dry areas of *A. urundeuva* and *D. gouveai*, the phylogeographical structures of the two species have probably been shaped by the same palaeoclimatic events in eastern South America.

The time of divergence amongst regional populations, estimated to be 173–88 ka, fell in a period subject to cycles of range shifts in dry vegetation that began in the late Pliocene and extended throughout the Pleistocene (Prado & Gibbs, 1993; Pennington *et al.*, 2004). Such results can be interpreted as a crude date estimate of the break of ecological connectedness of the habitat of *D. gouveai*. Although these results should be taken with caution, because of the somewhat wide credibility intervals that ranged from 24 to 565 ka, all estimates are within the Pleistocene. A clear portrait emerging from these dating results is that the main demographic events experienced by *D. gouveai* pre-date the arid period that occurred during the Last Glacial Maximum (LGM, 18–30 ka). The pre-LGM period has been identified as an age of intraspecific differentiation, range expansion and speciation for other species associated with open/dry vegetation in eastern Brazil, for example the Neotropical rattlesnake *Crotalus durissus* (Wüster *et al.*, 2005) and a group of large-bodied rodent species of the genus *Calomys* (Almeida *et al.*, 2007). In both of these instances, the events mentioned above were attributed to climatic changes and shifts in the distribution of dry vegetation in eastern Brazil during the mid-Pleistocene (Wüster *et al.*, 2005; Almeida *et al.*, 2007).

In summary, the phylogeographical and genetic patterns observed for *D. gouveai* in this study are consistent with changes in the distribution of dry vegetation in eastern Brazil. The initial diversification of *D. gouveai* took place in the mid-Pleistocene, probably in north-eastern Brazil, and this was followed by rapid expansion. Our data suggest the occurrence of two independent expansions from north-eastern populations to southern areas, probably occurring along mountain tops east and west of the Paraná River valley. After the expansion of the distributional range, *D. gouveai* populations diverged into three regionally isolated groups (southern, central and northern). All of the estimates obtained by molecular dating indicate that range expansion and isolation occurred during the middle to late Pleistocene, prior to the Last Glacial Maximum, and were probably triggered by climatic changes during the middle Pleistocene. The current patchy geographical distribution and subdivision of *D. gouveai* populations appear to be closely linked to these past events. A general result that emerges from the few studies previously performed is that many currently isolated areas of dry/open vegetation in eastern Brazil were more widely distributed during the Pleistocene, probably with ecological links amongst them (Behling, 2002; Ledru *et al.*, 2006; Behling *et al.*, 2007; Safford, 2007; Caetano *et al.*, 2008). With multiple recurrences of range fluctuation, plant and animal

species restricted to these habitats were provided with opportunities for range expansion, colonization, migration, vicariance, genetic divergence, and speciation. More genetic studies on species living in these habitats are needed to provide additional independent evidence of such evolutionary events and to reveal the role of the Pleistocene climatic oscillations in the diversification of species from eastern Brazil and other Neotropical regions.

ACKNOWLEDGEMENTS

We thank Erica C.C. Silva and Fernando F. Franco for help with the nested clade analysis, Fernando N.J. Villela for assistance with the map, and Paulo R. Epifanio, Antonio Hernandez-Junior and Leonardo Teofilo for assistance with the field sampling. We also thank Aluana G. Abreu, Therese Markow and an anonymous reviewer for helpful comments on the manuscript. This work was supported by FAPESP (grant no. 00/05979-5 to E.M.M.), CNPq, CAPES, FAEP/UNICAMP and USP.

REFERENCES

- Ab'Saber, A.N. (1982) The paleoclimate and paleoecology of Brazilian Amazonia. *Biological diversifications in the tropics* (ed. by G.T. Prance), pp. 41–59. Columbia University Press, New York.
- Almeida, F.C., Bonvicino, C.R. & Cordeiro-Estrela, P. (2007) Phylogeny and temporal diversification of *Calomys* (Rodentia, Sigmodontinae): implications for the biogeography of an endemic genus of the open/dry biomes of South America. *Molecular Phylogenetics and Evolution*, **42**, 449–466.
- Baimai, V., Sene, F.M. & Pereira, M.A.Q.R. (1983) Heterochromatin and karyotypic differentiation of some Neotropical cactus-breeding species of the *Drosophila repleta* species group. *Genetica*, **60**, 81–92.
- Behling, H. (2002) Late Quaternary vegetation and climate dynamics in southeastern Amazonia inferred from Lagoa da Confusão in Tocantins State, northern Brazil. *Amazoniana*, **17**, 27–39.
- Behling, H. & Lima da Costa, M. (2001) Holocene vegetational and coastal environmental changes from the Lago Crispim record in northeastern Pará State, eastern Amazonia. *Review of Palaeobotany and Palynology*, **114**, 145–155.
- Behling, H., Dupont, L., Safford, H.D. & Wefer, G. (2007) Late Quaternary vegetation and climate dynamics in the Serra da Bocaina, southeastern Brazil. *Quaternary International*, **161**, 22–31.
- Bigarella, J.J. & Andrade-Lima, D. (1982) Paleoenvironmental changes in Brazil. *Biological diversifications in the tropics* (ed. by G.T. Prance), pp. 27–40. Columbia University Press, New York.
- de Brito, R.A., Manfrin, M.H. & Sene, F.M. (2002) Nested cladistic analysis of Brazilian populations of *Drosophila serido*. *Molecular Phylogenetics and Evolution*, **22**, 131–143.

- Caetano, S., Prado, D., Pennington, R.T., 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). *Molecular Ecology*, **17**, 3147–3159.
- Carnaval, A.C. & Moritz, C. (2008) Historical climate modeling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *Journal of Biogeography*, **35**, 1187–1201.
- Carson, H.L. & Yoon, J.S. (1982) Genetics and evolution of Hawaiian *Drosophila*. *The genetics and biology of Drosophila*, Vol. 3b (ed. by M.J. Ashburner and H.L. Carson), pp. 298–344. Academic Press, New York.
- Clary, D.O. & Wolstenholme, D.R. (1985) The mitochondrial-DNA molecule of *Drosophila yakuba* – nucleotide sequence, gene organization, and genetic code. *Journal of Molecular Evolution*, **22**, 252–271.
- Clement, M., Posada, D. & Crandall, K.A. (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.
- Excoffier, L., Laval, G. & Schneider, S. (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.
- Fu, Y.-X. (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.
- Giulietti, A.M., Pirani, J.R. & Harley, R.M. (1997) Espinhaço Range Region, eastern Brazil. *Centres of plant diversity: a guide and strategy for their conservation*, Vol. 3: *The Americas*, (ed. by S.D. Davis, V.H. Heywood, O. Herrera-Macbride, J. Villa-Lobos and A. C Hamilton), pp. 397–404. IUCN Publication, Cambridge.
- Gordon, D., Abajian, C. & Green, P. (1998) A graphical tool for sequence finishing. *Genome Research*, **8**, 195–202.
- Hewitt, G.M. (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 183–195.
- Hey, J. (2005) On the number of new world founders: a population genetic portrait of the peopling of the Americas. *PLoS Biology*, **3**, e193.
- Hey, J. & Nielsen, R. (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747–760.
- Kumar, S., Tamura, K. & Nei, M. (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics*, **5**, 150–163.
- Ledru, M.P., Rousseau, D.D., Cruz, F.W., Jr, Riccomini, C., Karmann, I. & Martin, L. (2005) Paleoclimatic changes during the last 1,00,000 year from a record in the Brazilian Atlantic rainforest region and interhemispheric comparison. *Quaternary Research*, **64**, 444–450.
- Ledru, M.P., Ceccantini, G., Gouveia, S.E.M., López-Saez, J.A., Pessenda, L.C.R. & Ribeiro, A.S. (2006) Millennial-scale climatic and vegetation changes in a northern Cerrado (Northeast, Brazil) since the Last Glacial Maximum. *Quaternary Science Reviews*, **25**, 1110–1126.
- Manfrin, M.H. & Sene, F.M. (2006) Cactophilic *Drosophila* in South America: a model for evolutionary studies. *Genetica*, **126**, 57–75.
- Manfrin, M.H., de Brito, R.O.A. & Sene, F.M. (2001) Systematics and evolution of the *Drosophila buzzatii* (Diptera: Drosophilidae) cluster using mtDNA. *Annals of the Entomological Society of America*, **94**, 333–346.
- Mayle, F.E., Beerling, D.J., Gosling, W.D. & Bush, M.B. (2004) Responses of Amazonian ecosystems to climatic and atmospheric carbon dioxide changes since the last glacial maximum. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 499–514.
- Moraes, E.M. & Sene, F.M. (2007) Microsatellite and morphometric variation in *Drosophila gouveai*: the relative importance of historical and current factors in shaping the genetic population structure. *Journal of Zoological Systematics and Evolutionary Research*, **45**, 336–344.
- Nylander, J.A.A. (2008) *MrModeltest v2.3. Program distributed by the author*. Evolutionary Biology Centre, Uppsala University, Uppsala.
- Patterson, J.T. & Wheeler, M.R. (1942) Description of new species of the subgenera *Hirtodrosophila* and *Drosophila*. *University of Texas Publication*, **4213**, 67–109.
- Pennington, R.T., Prado, D.E. & Pendry, C.A. (2000) Neotropical seasonally dry forests and Quaternary vegetation changes. *Journal of Biogeography*, **27**, 261–273.
- Pennington, R.T., Lavin, M., Prado, D.E., Pendry, C.A., Pell, S.K. & Butterworth, C.A. (2004) Historical climate change and speciation: neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 515–537.
- Pereira, M.A.Q.R., Vilela, C.R. & Sene, F.M. (1983) Notes on breeding and feeding sites of some species of the replete group of the genus *Drosophila* (Diptera, Drosophilidae). *Ciência e Cultura*, **35**, 1313–1319.
- Posada, D. & Crandall, K.A. (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.
- Posada, D., Crandall, K.A. & Templeton, A.R. (2000) GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology*, **9**, 487–488.
- Prado, D.E. & Gibbs, P.E. (1993) Patterns of species distributions in the dry seasonal forests of South America. *Annals of the Missouri Botanical Garden*, **80**, 902–927.
- Rogers, A.R. & Harpending, H.C. (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.
- Ronquist, F. & Huelsenbeck, J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.
- Rowan, R.G. & Hunt, J.A. (1991) Rates of DNA change and phylogeny from the DNA-sequences of the alcohol-dehydrogenase gene for five closely related species of Hawaiian *Drosophila*. *Molecular Biology and Evolution*, **8**, 49–70.

- Russo, C.A.M., Takezaki, N. & Nei, M. (1995) Molecular phylogeny and divergence times of drosophilid species. *Molecular Biology and Evolution*, **12**, 391–404.
- Safford, H.D. (2007) Brazilian Páramos IV. Phylogeography of the campos de altitude. *Journal of Biogeography*, **34**, 1701–1722.
- Silva, J.M.C. & Bates, J.M. (2002) Biogeographic patterns and conservation in the South American Cerrado: a tropical savanna hotspot. *BioScience*, **52**, 225–233.
- Silva, A.F.G. & Sene, F.M. (1991) Morphological geographic variability in *Drosophila serido* (Diptera, Drosophilidae). *Revista Brasileira de Entomologia*, **35**, 455–468.
- 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. *Annals of the Entomological Society of America*, **87**, 651–701.
- Swofford, D.L. (2002) *PAUP*: phylogenetic analysis using parsimony (*and other methods)*. Beta version 4.0b10. Sinauer Associates, Sunderland, MA.
- Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.
- Templeton, A.R. (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology*, **7**, 381–397.
- Templeton, A.R. (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology*, **13**, 789–809.
- Tidon-Sklorz, R. & Sene, F.M. (2001) Two new species of the *Drosophila serido* sibling set (Diptera, Drosophilidae). *Iheringia*, **90**, 141–146.
- Tosi, D. & Sene, F.M. (1989) Further studies on chromosomal variability in *Drosophila serido* (Diptera, Drosophilidae). *Revista Brasileira de Genética*, **1**, 729–746.
- Werneck, F.P. & Colli, G.R. (2006) The lizard assemblage from Seasonally Dry Tropical Forest enclaves in the Cerrado biome, Brazil, and its association with the Pleistocenic Arc. *Journal of Biogeography*, **33**, 1983–1992.
- Wüster, W., Ferguson, J.E., Quijada-Mascareñas, J.A., Pook, C.E., Salomão, M.G. & Thorpe, R.S. (2005) Tracing an invasion: landbridges, refugia, and the phylogeography of the Neotropical rattlesnake (Serpentes: Viperidae: *Crotalus durissus*). *Molecular Ecology*, **14**, 1095–1108.

BIOSKETCH

Evandro M. Moraes is an evolutionary geneticist interested in the historical and current processes underlying the phylogeography and population genetic structure of animal and plant species associated with open vegetation, with a particular focus on Neotropical enclaves of xeric vegetation.

Editor: David Hafner