

This Provisional PDF corresponds to the article as it appeared upon acceptance. Fully formatted PDF and full text (HTML) versions will be made available soon.

Rock outcrop orchids reveal the genetic connectivity and diversity of inselbergs of northeastern Brazil

BMC Evolutionary Biology 2014, **14**:49 doi:10.1186/1471-2148-14-49

Fábio Pinheiro (biopinheiro@yahoo.com.br)
Salvatore Cozzolino (cozzolin@unina.it)
David Draper (ddmunt@utpl.edu.ec)
Fábio de Barros (fdebarros@terra.com.br)
Leonardo P Félix (lpfelix@cca.ufpb.br)
Michael F Fay (M.Fay@kew.org)
Clarisse Palma-Silva (clarissepalma@yahoo.com.br)

ISSN 1471-2148

Article type Research article

Submission date 7 January 2014

Acceptance date 7 March 2014

Publication date 15 March 2014

Article URL <http://www.biomedcentral.com/1471-2148/14/49>

Like all articles in BMC journals, this peer-reviewed article can be downloaded, printed and distributed freely for any purposes (see copyright notice below).

Articles in BMC journals are listed in PubMed and archived at PubMed Central.

For information about publishing your research in BMC journals or any BioMed Central journal, go to

<http://www.biomedcentral.com/info/authors/>

© 2014 Pinheiro *et al.*

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

Rock outcrop orchids reveal the genetic connectivity and diversity of inselbergs of northeastern Brazil

Fábio Pinheiro^{1*}

* Corresponding author

Email: biopinheiro@yahoo.com.br

Salvatore Cozzolino^{2,3}

Email: cozzolin@unina.it

David Draper⁴

Email: ddmunt@utpl.edu.ec

Fábio de Barros¹

Email: fdebarros@terra.com.br

Leonardo P Félix⁵

Email: lpfelix@cca.ufpb.br

Michael F Fay⁶

Email: M.Fay@kew.org

Clarisse Palma-Silva⁷

Email: clarissepalma@yahoo.com.br

¹ Instituto de Botânica, Núcleo de Pesquisa do Orquidário do Estado, Avenida Miguel Estéfano 3687, 04301-012 São Paulo, SP, Brazil

² Dipartimento di Biologia, Complesso Universitario di Monte S. Angelo, Università degli Studi di Napoli Federico II, 80100 Napoli, Italy

³ Institute for Plant Protection, Consiglio Nazionale delle Ricerche, Via Madonna del Piano 10, I-50019 Sesto Fiorentino (FI), Italy

⁴ Departamento de Ciencias Naturales, Universidad Técnica Particular de Loja, San Cayetano Alto s/n, CP 11 01 608 Loja, Ecuador

⁵ Departamento de Fitotecnia, Laboratório de Citogenética Vegetal, Universidade Federal da Paraíba, Areia, PB, Brazil

⁶ Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK

⁷ Laboratório de Ecologia Molecular, Departamento de Ecologia, Universidade Estadual Paulista, 13506-900 Rio Claro, SP, Brazil

Abstract

Background

Because of their fragmented nature, inselberg species are interesting biological models for studying the genetic consequences of disjoint populations. Inselbergs are commonly compared with oceanic islands, as most of them display a marked ecological isolation from the surrounding area. The isolation of these rock outcrops is reflected in the high number of recorded endemic species and the strong floristic differences between individual inselbergs and adjacent habitats. We examined the genetic connectivity of orchids *Epidendrum cinnabarinum* and *E. secundum* adapted to Neotropical inselbergs of northeastern Brazil. Our goals were to identify major genetic divergences or disjunctions across the range of the species and to investigate potential demographic and evolutionary mechanisms leading to lineage divergence in Neotropical mountain ecosystems.

Results

Based on plastid markers, high genetic differentiation was found for *E. cinnabarinum* ($F_{ST} = 0.644$) and *E. secundum* ($F_{ST} = 0.636$). Haplotypes were not geographically structured in either taxon, suggesting that restricted gene flow and genetic drift may be significant factors influencing the diversification of these inselberg populations. Moreover, strong differentiation was found between populations over short spatial scales, indicating substantial periods of isolation among populations. For *E. secundum*, nuclear markers indicated higher gene flow by pollen than by seeds.

Conclusions

The comparative approach adopted in this study contributed to the elucidation of patterns in both species. Our results confirm the ancient and highly isolated nature of inselberg populations. Both species showed similar patterns of genetic diversity and structure, highlighting the importance of seed-restricted gene flow and genetic drift as drivers of plant diversification in terrestrial islands such as inselbergs.

Keywords

Brazilian Atlantic Forest, Caatinga, Chloroplast microsatellites, *Epidendrum cinnabarinum*, *Epidendrum secundum*, Genetic drift, Nuclear microsatellites, Orchidaceae

Background

Inselbergs are isolated rock outcrops typically harboring rupicolous vegetation and embedded within a landscape composed of contrasting plant communities. Because of their disconnected geographic nature, inselbergs are frequently compared with oceanic islands, since most of them display a marked ecological isolation from the surrounding area [1]. The isolation and ancient age of these rock outcrops is reflected in the high number of recorded endemic species (reviewed by [2]) and the strong floristic differences observed between individual inselbergs and their surrounding habitats [3,4]. Some authors have also suggested that inselbergs may have acted as refugia for xerophilic or cold-adapted species during

glacial/interglacial cycles [4,5]. In addition, most inselbergs maintain their typical attributes irrespective of geographic location, enabling broad-scale comparisons even between different and contrasting biomes, such as mesic forests and seasonally dry plant communities [2]. As biological islands, inselbergs are promising ecosystems for biogeographic and evolutionary studies, comparable to their oceanic counterparts [6].

Phylogeographic studies of inselberg-adapted species have provided insights into evolutionary processes leading to diversification of lineages and species [7-11]. Such studies have uncovered strong phylogeographic structure, high population differentiation and extensive genetic diversity levels, supporting the view of inselbergs as centers of species diversity and endemism. Species occurring in small isolated patches are expected to experience reduced gene flow, significant genetic drift and high levels of population divergence [6]. Indeed, genetic studies of inselberg species have provided strong support for these expectations [7,12-15]. In isolated populations, gene flow is constrained, and genetic drift is expected to be the predominant force governing allele frequencies [8,16].

Porembsky et al. [17] have noted that plant community organization on rock outcrops is driven by stochastic colonization events. Thus, phylogeographic studies of multiple co-occurring species with similar biological traits (e.g. pollination and seed dispersion) may be particularly useful for understanding the role of stochasticity (i.e. genetic drift) in the evolution of these naturally fragmented populations. By comparing the phylogeographic structures of different species, one can infer whether the current plant community has been historically stable as evidenced by geographically similar genetic structure. Alternatively, if species distributions are ephemeral over evolutionary time and more influenced by intrinsic species preferences, a mixture of phylogeographic structures is expected [18,19].

Phylogeographic studies focused on organisms associated with mesic forest communities have indicated a strong association between the occurrence of glacial cycles and the fragmentation of forest-dwelling species, supporting the classical model of tropical refugia [20-22]. On the other hand, species associated with open biomes, such as grassland, savannas and dry forests, show variable demographic responses to past glacial cycles, suggesting a more complex scenario [21,23]. In this regard, cross validation of environmental envelope models (EEMs according to terminology suggested by [24]) and molecular genetic data have confirmed different past demographic scenarios for a wide array of species. Fragmentation [25], expansion [26] and long-term persistence [27] have been detected in organisms associated with open vegetation communities. By combining EEMs and molecular genetic tools, different phylogeographic hypotheses depicting the role of past climatic fluctuations in lineage diversification can be tested.

The Caatinga is the largest, most diverse dry seasonal tropical forest biome in the Neotropical region [28]. Despite the extensive geographic area of this biome, relatively few phylogeographic studies have been conducted using species from the Caatinga [27,29-32]. This paucity of studies precludes broad conclusions regarding the impact of past climate oscillation events on the genetic structure of organisms. The high number of endemic species [28] coupled with old lineage-divergence times [27,33,34] implies an ancient origin for Caatinga communities. In addition, paleoclimate models [35] indicate that dry forest distribution within the Caatinga was stable during the Last Glacial Maximum (LGM). Indeed, high species diversity and floristic differences have been found among Caatinga rock outcrops [3,36], supporting the hypothesis of inselbergs as ancient, stable refuges of diversity [4,5].

In the present study, we used nuclear and plastid microsatellite markers to analyze phylogeographic structure and genetic diversity of *Epidendrum cinnabarinum* and *E. secundum* (Orchidaceae), two widespread, co-occurring species commonly found on inselbergs in northeastern Brazil. *Epidendrum cinnabarinum*, a polyploid species with $2n = 240$ chromosomes [37], occurs in Caatinga and Brazilian Atlantic Forest (BAF) inselbergs and on sand dune vegetation along the seashore in northeastern Brazil. *Epidendrum secundum* exhibits a diploid chromosome count of $2n = 56$ in most populations [37]; it has a much broader geographic distribution, occurring in Central America, Guiana and Andean ranges and the Brazilian Central Shield [38]. This species occurs preferentially on rock outcrops, with abundant populations observed on Caatinga and BAF inselbergs. In this study, we also applied EEMs to explore the demography of both species during important climatic oscillation events: the LGM (21,000 ka) and Last Interglacial (LIG; 120,000 ka). The following specific questions were addressed: (1) What is the current extent of genetic structure and the degree of isolation among disjoint populations on inselbergs? (2) Were the current range distributions of both species stable during Late Quaternary climatic oscillations? and (3) Are there phylogeographic breaks separating populations from different biomes (Caatinga and BAF) and ecoregions (Chapada Diamantina and Planalto da Borborema)? We also considered the phylogeographical and genetic structure of *E. cinnabarinum* and *E. secundum* in light of results based on climate modeling, paleovegetation reconstructions and the island-like nature of populations distributed on inselbergs.

Results

Plastid and nuclear genetic diversity of sampled populations

Analysis of seven plastid loci recovered a total of 10 haplotypes for *E. cinnabarinum* and 12 haplotypes for *E. secundum* (Figure 1C, 1D, Table 1, Additional file 1: Table S1), with no haplotype sharing between species. One to five haplotypes were found within *E. cinnabarinum* populations, while one to six haplotypes were detected within *E. secundum* populations. Similar diversity levels were detected for both species concerning haplotype richness (*E. cinnabarinum* 0.000–3.066; *E. secundum* 0.000–4.873) and haplotype diversity (*E. cinnabarinum* 0.000–0.813; *E. secundum* 0.000–0.800) (Table 1). In two sympatric populations, TO and BZ, genetic diversity parameters were identical for both species (Table 1).

Figure 1 Geographic distribution of sampled populations. Maps showing sampled populations of *Epidendrum cinnabarinum* (A) and *E. secundum* (B), and respective plastid DNA networks for each species (C and D). Pie charts reflect the frequency of occurrence of each haplotype in each population. Haplotype colours correspond to those shown in networks. Pie charts with solid and broken outlines indicate Caatinga and Brazilian Atlantic Forest populations, respectively. Nuclear genetic groups are indicated for *E. secundum* (Nuclear cluster 1 – black; Nuclear cluster 2 – grey). In the statistical parsimony networks (C and D), the haplotype frequencies are proportional to circle sizes. The number of mutations required to explain transitions among haplotypes is indicated along the lines connecting the haplotypes by cross hatches.

Table 1 Genetic characterization of populations of *Epidendrum cinnabarinum* and *E. secundum* based on plastid markers

Population	ID	Plastid loci		
		NH	HR	HD
<i>Epidendrum cinnabarinum</i>				
Mamanguape	MA	1	0.000	0.000
Serraria	RR	2	0.990	0.505
Esperança	ES	2	0.966	0.456
Queimadas	QE	2	0.583	0.167
Jaqueira	JA	2	0.904	0.363
Cabo de Santo Agostinho	CB	1	0.000	0.000
Pirambu	PI	5	3.066	0.813
Abaeté	AB	1	0.000	0.000
<i>Epidendrum secundum</i>				
Maranguape	CE	2	1.000	0.337
Brejo da Madre de Deus	DE	2	0.800	0.100
Serra de Itabaiana	TB	2	0.800	0.100
Morro do Chapéu	MC	5	4.000	0.767
Palmeiras	PL	6	4.873	0.800
Serra da Jibóia	JI	2	0.996	0.268
Sympatric				
Pedra de Santo Antonio (<i>E. cinn.</i>)	TO	1	0.000	0.000
Pedra de Santo Antonio (<i>E. sec.</i>)		1	0.000	0.000
Bezerros (<i>E. cinn.</i>)	BZ	1	0.000	0.000
Bezerros (<i>E. sec.</i>)		1	0.000	0.000
Camocim de São Félix (<i>E. cinn.</i>)	FE	2	0.978	0.479
Camocim de São Félix (<i>E. sec.</i>)		1	0.000	0.000

Identification code of sampled populations (ID), number of haplotypes (NH), haplotype richness (HR) and haplotype diversity (HD) for seven loci.

With respect to six nuclear markers genotyped in *E. secundum*, moderate levels of genetic diversity were observed in most populations (Table 2). The number of alleles per population ranged from 19 to 23, and allelic richness per population varied from 2.28 to 2.77. Expected heterozygosity per population ranged from 0.442 to 0.608. One to nine private alleles were observed within populations, with values of private allelic richness ranging from 0.11 to 0.72. Inbreeding coefficients, which ranged from -0.032 to 0.156 , were not significantly different from zero in any population. The presence of genotyping errors due to stuttering or null alleles (frequencies ranging from 0 to 0.09, estimated using Brookfield Equation 1) was ruled out using MICRO-CHECKER tests [39].

Table 2 Genetic characterization of populations of *Epidendrum secundum* based on nuclear markers

Populations/Code	A	PA	AR	PAR	H_E	f^1	M -ratio ^{2,3}
Pedra de Santo Antonio/TO	19	1	2.28	0.11	0.462	0.156	0.872
Brejo da Madre de Deus/DE	22	0	2.35	0.19	0.466	0.073	0.895
Bezerros/BZ	20	2	2.29	0.22	0.442	-0.065	0.785
Camocim de São Félix/FE	23	9	2.39	0.35	0.459	0.154	0.794
Morro do Chapéu/MC	22	11	2.77	0.72	0.608	-0.032	0.875

¹Departures of within-population inbreeding coefficients (f) from HWE were not detected; ²A population is considered to have undergone a bottleneck if its M -ratio value falls below the lower threshold of critical M -ratio ($M_c = 0.625$) showed on Table 4; ³No bottlenecks detected.

Populations sampled, number of alleles (A), number of private alleles (PA), allelic richness (AR), private allelic richness (PAR), expected heterozygosity (H_E), the within population inbreeding coefficient f and the M -ratio values for six nuclear loci.

Genetic structure of *E. cinnabarinum*

A haplotype network based on plastid markers showed loops (ambiguities), and no apparent geographic structure was observed (Figure 1C). The most frequent haplotype in both BAF and Caatinga biomes (Hc1) was found in 46.8% of individuals and in 5 out of 11 sampled populations. Hc5, the second most frequent haplotype, was also distributed in both biomes and occupied a central position in the haplotype network (Figure 1C). Only 3 (Hc1, Hc5 and Hc6) out of 10 haplotypes were found to occur in more than one population.

High genetic differentiation was observed across populations, with $F_{ST} = 0.644$ and $G_{ST} = 0.672$. According to an analysis of molecular variance (AMOVA) (Table 3), a high proportion of the genetic variability in the haplotype data resided among populations (64.41%, $P < 0.0001$), with only 35.59% accounted for within populations. Hierarchical AMOVA did not support a division between Caatinga and BAF populations (Table 3; $P = 0.704$). Pairwise F_{ST} comparisons among populations (Additional file 2: Table S2) ranged from 1.000 to 0.000, with most values found to be significant ($P < 0.005$). Haplotype differentiation between populations separated over short spatial scales was observed between populations QE and TO (12 km apart) and between BZ and FE (18 km apart) (Additional file 2: Table S2, Figure 1A). Other than the strong genetic differentiation found among populations, no among-population phylogeographic structure or isolation by distance was detected ($P = 0.114$ and $P = 0.084$, respectively).

Table 3 Analysis of molecular variance (AMOVA) for plastid microsatellite data for *Epidendrum cinnabarinum* populations, using two different models

Source of variation	d.f.	Variance components	Variation (%)	P-value
1) Among populations	10	0.23999	64.41	$P < 0.001$
Within populations	166	0.13262	35.59	
2) Between Biomes (Caatinga and Brazilian Atlantic Forest)	1	-0.03251	-9.10	$P = 0.704$
Among populations within biomes	9	0.25725	71.99	$P < 0.001$
Within populations	166	0.13262	37.11	$P < 0.001$

One model includes all populations pooled, and the other includes populations from different biomes (Caatinga and Brazilian Atlantic Forest) separated.

Genetic structure of *E. secundum*

Analysis of plastid markers yielded a network lacking clear geographic structuring (Figure 1D). Haplotype Hs5 was shared between populations from Caatinga and BAF biomes, and also between Chapada Diamantina and Planalto da Borborema ecoregions (Figure 1B, 1D). Hs5 was also the most frequent haplotype, found in 54.5% of individuals and in six out of nine sampled populations. Seven haplotypes were restricted to Chapada Diamantina populations (Figure 1B).

The only hypothesis of population differentiation significantly supported by AMOVA was that between populations from Chapada Diamantina and the remaining localities (Table 4). The partitioning of plastid genetic diversity and structure among *E. secundum* populations was also very similar to the patterns observed for *E. cinnabarinum*. AMOVA results indicated that a high proportion of uncovered genetic variability was found among populations (63.64%, $P < 0.0001$), with only 36.36% attributed to within-population variability (Table 4). Genetic differentiation was high across all populations, with $F_{ST} = 0.636$ and $G_{ST} = 0.632$. As observed for *E. cinnabarinum*, contrasting haplotypes were detected between populations BZ and FE. Pairwise F_{ST} comparisons among populations ranged from 1.000 to 0.000, with most values significant ($P < 0.005$; Additional file 3: Table S3). No phylogeographic structure was detected among populations, as R_{ST} was not significantly larger than F_{ST} ($P = 0.503$). Isolation by distance among populations was also not detected ($P = 0.822$).

Table 4 Analysis of molecular variance (AMOVA) for nuclear and plastid microsatellite data for *Epidendrum secundum* populations, using four different models

Source of variation	d.f.	Variance components	Variation (%)	<i>P</i> -value
<i>Nuclear microsatellite</i>				
1) Among populations	4	0.97604	4.58	<i>P</i> = 0.087
Within populations	187	20.33329	95.42	
2) Between Chapada Diamantina and Planalto da Borborema ecoregions	1	1.75359	7.81	<i>P</i> = 0.205
Among populations within biomes	3	0.36609	1.63	<i>P</i> = 0.331
Within populations	187	20.33329	90.56	<i>P</i> = 0.098
<i>Plastid microsatellite</i>				
1) Among populations	8	0.22274	63.64	<i>P</i> < 0.001
Within populations	165	0.12727	36.36	
2) Between Biomes (Caatinga and Brazilian Atlantic Forest)	1	0.00302	0.86	<i>P</i> = 0.386
Among populations within biomes	7	0.22154	62.97	<i>P</i> < 0.001
Within populations	165	0.12727	36.17	<i>P</i> < 0.001
3) Between Chapada Diamantina and Planalto da Borborema ecoregions	1	0.05130	11.61	<i>P</i> = 0.201
Among populations within biomes	4	0.25806	58.42	<i>P</i> < 0.001
Within populations	108	0.13241	29.97	<i>P</i> < 0.001
4) Between Chapada Diamantina and remaining populations	1	0.85932	61.21	<i>P</i> < 0.05
Among populations within groups	7	0.22273	15.87	<i>P</i> < 0.001
Within populations	165	0.32182	22.92	<i>P</i> < 0.001

First model with all populations pooled, second model with populations from different biomes (Caatinga and Brazilian Atlantic Forest) separated, third model with populations from Chapada Diamantina and Planalto da Borborema ecoregions separated, and fourth model with populations from Chapada Diamantina separated from the remaining localities.

Compared with plastid markers, nuclear markers exhibited lower levels of genetic differentiation as estimated by Φ_{ST} (0.114), G_{ST} (0.122) and G'_{ST} (0.255). An AMOVA also generated different results than for plastid markers, with most of the genetic variation partitioned within (95.42%, $P < 0.0001$) rather than among (4.58%, $P = 0.087$) populations (Table 4). Differentiation between BAF and Caatinga populations and between Chapada Diamantina and Planalto da Borborema was not significant (Table 4). No sign of phylogeographic structure was detected among populations ($P = 0.329$). A Mantel test did not support isolation by distance among populations based on Φ_{ST} , G_{ST} or G'_{ST} (all $P > 0.05$).

As shown in Additional file 4: Figure S1, simulations performed in STRUCTURE consistently identified $K = 2$ clusters. Most analyzed individuals showed admixed ancestry, however, and genetic subdivisions among populations were unclear (Figures 1B and 2). Admixture proportions (Q) for population BZ showed strong assignment to cluster 2 (average assignment proportion = 0.92; Figure 2), whereas specimens from the southernmost population MC showed assignment proportions (0.77; Figure 2) associated with cluster 1. Intermediate assignment proportions were observed for most individuals from remaining populations (Figure 2).

Figure 2 Genetic assignment results. Summary of population structure in *Epidendrum secundum* using Bayesian assignment analysis for a $K = 2$ population model. Most individuals from population Bezerros show assignment probabilities associated to cluster 1 (black), whereas specimens from population Morro do Chapéu are mainly classified on cluster 2 (grey). Admixed proportions were found on populations Brejo da Madre de Deus, Pedra de Santo Antonio and Camocim de São Félix. The ecoregion of origin (Chapada Diamantina or Planalto da Borborema) is indicated. See Table 1 for population details.

Population size reduction and gene flow among *E. secundum* populations

Across all sites and loci, M -ratio values ranged from 0.785 to 0.895 (Table 2). The lowest observed critical value (M_C) obtained from simulations performed with different values of parameter θ , proportion of single-step mutations (pg) and average size of non-one-step mutations (Δg) was 0.589 (Table 5). According to this M_C threshold, none of the populations showed signatures of past genetic bottlenecks (Table 2).

Table 5 Demographic reduction (bottleneck) results

Parameter simulations	θ	pg	Δg	M_C
1	0.5	0.1	2.0	0.883
2	0.5	0.3	2.0	0.800
3	0.5	0.1	3.5	0.777
4	0.5	0.3	3.5	0.623
5	2.0	0.1	2.0	0.875
6	2.0	0.3	2.0	0.825
7	2.0	0.1	3.5	0.722
8	2.0	0.3	3.5	0.589
9	10.0	0.1	2.0	0.844
10	10.0	0.3	2.0	0.811
11	10.0	0.1	3.5	0.702
12	10.0	0.3	3.5	0.619

Parameters for the two-phased mutation model (TPM) used to generate critical values of M -ratio (M_C). Theta (θ), proportion of single-step mutations (pg) and average size of non one-step mutations (Δg) were used to infer M_C thresholds.

Using values of genetic differentiation among populations BZ, DE, TO, FE and MC obtained for nuclear ($G_{ST} = 0.122$) and plastid ($G_{ST} = 0.782$) markers, the ratio of pollen flow to seed flow was estimated as 25.85. The value of this ratio suggests that gene flow via pollen in *E. secundum* is > 20 times higher than that occurring via seeds.

Potential ancient distributions of *E. cinnabarinum* and *E. secundum*

Distributions under LIB, LGM and current climatic conditions were successfully estimated for *E. cinnabarinum* and *E. secundum*. As indicated by high area-under-the-curve (AUC) values (0.969 for *E. secundum* and 0.968 for *E. cinnabarinum*), the EEM analysis performed well. Climate variables contributed to both species models, with the *E. cinnabarinum* model additionally adjusted using a geological eras variable. Past distribution ranges differed between the two species, mainly at their peripheral limits. From the LIG to the present, *E. cinnabarinum* exhibited a pronounced decrease in its presence at Chapada Diamantina and

other inland mountain ranges, with a more recent expansion towards northern Planalto da Borborema and coastal regions after the LGM (Figure 3A, 3B and 3C). Populations distributed in the southern portion of the northeastern Brazilian seashore did not show marked demographic oscillations. A decrease in the inland distribution of *E. secundum* was also observed, but without any further expansion (Figure 3D, 3E and 3F) as observed for *E. cinnabarinum*. During the LIG and LGM, *E. secundum* was broadly distributed in inland mountains, being present mainly at Cadeia do Espinhaço (comprising Chapada Diamantina and southern mountain ranges) and in western portions of Planalto da Borborema. After the LGM, the species remained broadly distributed only within Chapada Diamantina and mountains to the east within the BAF biome. The model representing current climatic conditions indicated that *E. secundum* has a fragmented distribution within Planalto da Borborema, where most populations are found at the border between Caatinga and BAF biomes. Since the LIG, the distribution of this species has not been continuous between Chapada Diamantina and Planalto da Borborema.

Figure 3 Environmental envelope models results. Maps showing environmental envelope models (EEMs) for *Epidendrum cinnabarinum* (blue, AUC = 0.968) and *E. secundum* (green, AUC = 0.969) based on current (**A and D**) and two past scenarios, the Last Glacial Maximum (**B and E**) and Last inter-glacial (**C and F**) using the Maximum Entropy algorithm v. 3.3.3e. Black dots indicate localities used for EEMs analysis, and yellow dots indicates populations sampled for genetic analysis. Darker colors show areas with more suitable predicted conditions, as indicated by blue (*E. cinnabarinum*) and green scales (*E. secundum*). See Table 1 for population details.

Discussion

Genetic studies of inselberg species can provide insights into the combined effects of genetic drift and restricted gene flow on evolution and diversification of lineages restricted to disjunct populations [1,2,5]. Phylogeographic studies of inselberg species have indeed confirmed many hypotheses associated with such isolated habitats, such as low levels of gene flow and strong genetic drift [7-11]. In some extreme cases of isolation, within-inselberg interspecific gene flow between congeneric species has been found to be higher than intraspecific gene flow between different inselbergs [7,14]. To understand the effects of long-term disjoint distribution on genetic architecture and evolution of inselberg populations, we carried out a phylogeographic study of two rock-outcrop orchid species showing very similar life traits. The use of plastid markers revealed marked genetic differentiation among populations. In contrast, nuclear genetic differentiation was much lower in *E. secundum*, highlighting the important role of pollen dispersal, and consequently pollination services, in species cohesion. This result has also been observed in other plants [7,40]. Results from EEM analysis of *E. secundum* suggest long-term stability of populations from Chapada Diamantina and the southern portion of Planalto da Borborema (Figure 3D, 3E and 3F); in contrast, results for *E. cinnabarinum* imply a decrease in distribution within Chapada Diamantina, in agreement with its current rarity in this ecoregion. After the LGM, *E. cinnabarinum* expanded its distribution to northern sand dune vegetation communities. This expansion most likely occurred from inland inselberg populations, not from southern coastal ones, as extensive haplotype sharing was found among seashore and inland populations. The joint use of phylogeography and EEMs helped to clarify the origin of genetic diversity and population differentiation in these inselberg species, emphasizing the importance of such studies in terrestrial island-like environments.

Marked genetic differentiation among inselberg populations

Genetic differentiation among populations of both *E. cinnabarinum* and *E. secundum* was not significantly associated with BAF and Caatinga biomes (Tables 3 and 4). In contrast, extensive haplotype sharing was found among populations distributed in both biomes (Figure 1A and 1B). Haplotype sharing was also found between populations from inselbergs and sand dune vegetation, in agreement with floristic similarities reported in previous studies [41]. Because of their xeric nature, inselbergs can provide suitable conditions for drought tolerant species even within humid forest patches [2]. Inselbergs embedded in mesic forests often harbor xeric plant communities and can be considered refugia for species adapted to drier climatic conditions, highlighting their importance for comparative studies across different biomes [2,4]. Paleontological evidence [42] indicates that alternating dry and wet periods have influenced the BAF distribution of plant species, particularly along the northeastern coast. Inselberg populations within the BAF (populations JA, CB, TB and JI) may thus represent relicts of xeric plant communities that probably extended further towards the coast during drier periods, thereby allowing gene exchange with populations currently placed in the Caatinga biome.

The most outstanding pattern recovered by the use of plastid markers was the deep among-population differentiation within *E. cinnabarinum* and within *E. secundum* (Figure 1A and 1B, Additional files 2 and 3: Table S2 and S3). Even populations separated by a few kilometers exhibited strong differentiation based on the plastidial genome (populations TO and QE, 12 km apart, $F_{ST} = 0.985$; BZ and FE, 18 km apart, $F_{ST} = 1.000$; ES and RR, 34 km apart, $F_{ST} = 0.727$). The significant pairwise genetic differentiation among most populations, the absence of a pattern of isolation by distance, and the lack of significantly phylogeographically structured haplotypes suggest that genetic drift, as observed in other rock outcrop species [7-9,43], has been an important force shaping plastidial genetic variation in *E. cinnabarinum* and *E. secundum*. In finite populations, such as those inhabiting inselbergs and oceanic islands, genetic drift may be the predominant force governing allele frequencies of neutral loci [6,44]. Because effective population size of maternally inherited organelles is expected to be an order of magnitude lower than that of nuclear genes, genetic drift is expected to play a stronger role in shaping frequencies of organellar alleles than those of nuclear alleles [45]. In addition to the spatial dimension, genetic drift includes an intrapopulation temporal component in which a series of subpopulations diverge from one another over time. Fine-scale genetic analysis [43,46] coupled with long-term studies (e.g. [8]) may help to clarify the impact of drift on the genetic structure of inselberg populations.

The genetic structure observed among *E. secundum* populations was higher for plastid than for nuclear markers (Figure 2, Table 4). Most haplotype diversity (63.64%) was found to reside among populations; this result is in contrast to nuclear markers, in which diversity among populations was not significant (4.58%). This difference in genetic structure between genomic compartments was translated into a pollen/seed flow ratio of 25.85, indicating that pollen-mediated gene dispersal is 25-fold more efficient than dispersal through seeds. The inference of extensive pollen flow between *E. secundum* populations is also supported by observations of low inbreeding coefficients (Table 2) and a high proportion of genetic diversity within populations (Table 4). In addition, Bayesian assignment based on nuclear markers indicated intermediate admixture proportions for many individuals and specimens assigned to different nuclear genetic clusters in the same population (Figure 1B); these results also suggest higher levels of gene flow by pollen. Efficient gene flow by pollen is normally expected for food-deceptive orchids [40,47], as pollinators avoid plants in the same patch,

thereby promoting pollen flow over long distances and reducing the chances of geitonogamous pollination. In general, low levels of nuclear genetic structure are observed in orchids (reviewed by [48]). Pollen flow has proven to be crucial to the maintenance of species cohesion in other fragmented inselberg species, highlighting the importance of pollinators in promoting gene exchange among rock outcrops [7,8,43].

Despite the strong genetic differentiation observed between populations from Chapada Diamantina and the remaining localities (Table 4), phylogeographic structure was not significant in *E. secundum*. This surprising situation is probably due to ancestral haplotype sharing (Hs5) among most populations, indicating ancient dispersal events. Another explanation for the non-significance of phylogeographic structure is the fact that multiple, phylogenetically unrelated haplotypes are found in the same populations, especially in the Chapada Diamantina ecoregion. Chapada Diamantina is an old and very stable region; demographic changes in peripheral populations, such as those inhabiting the BAF and from mountains south of Chapada Diamantina, may have contributed to the exchange of unrelated haplotypes—a pattern observed in other studies [32,34,49]. Indeed, species distribution modeling shows a broader past geographic distribution for *E. secundum*, with a former range that includes mountains to the south and north of Chapada Diamantina.

Demographic changes in *E. secundum* and *E. cinnabarinum*

For both species, results obtained using EEMs have shed light on whether haplotype sharing among populations is due to regular gene flow, past long-distance dispersal events or past continuous distribution. For *E. cinnabarinum*, a general pattern of expansion from the inland towards the seashore since the LIG is suggested (Figure 3A, 3B and 3C). This species appears to have been more commonly distributed in Chapada Diamantina during the LIG, showing a further decrease in distribution until the present. *Epidendrum cinnabarinum* is in fact currently very rare in Chapada Diamantina, where its populations are composed of only a few individuals. Colonization of the northern portion of the seashore occurred during the last 20,000 years, probably from Planalto da Borborema populations. The haplotype sharing detected among populations and the low diversity of Mamanguape and Cabo de Santo Agostinho populations are consistent with this hypothesis. Because only ancestral haplotypes are shared among populations, recurrent gene flow is unlikely; in this case, long-distance dispersal may have been responsible for the northern seashore colonization. Results from Behling [42] suggest that xeric vegetation replaced mesic plant communities along the northern portion of the northeastern Brazilian seashore in the last 10,000 years, a trend likely favoring the expansion of *E. cinnabarinum* from inland inselbergs towards dry coastal vegetation zones. Haplotype sharing between inland and coastal populations and the inferred expansion of inland populations towards the seashore have also been observed in other phylogeographic studies along the Brazilian coast, confirming the connection between the BAF and other adjacent biomes such as the Caatinga [32,49], Cerrado [50,51] and Pampas [40,52].

Analysis of EEMs recovered different demographic patterns for *E. secundum* (Figure 3D, 3E, 3F). The long-term persistence of Chapada Diamantina populations is in agreement with the high levels of genetic diversity observed in Morro do Chapéu and Palmeiras populations (Tables 1 and 2) and the absence of bottlenecks (Table 5). The genetic differentiation of Chapada Diamantina and Planalto da Borborema populations was also confirmed by the EEM analysis, which demonstrated a clear discontinuity in *E. secundum* distribution along these two ecoregions in the last 120,000 years. High levels of plant endemism and biodiversity

have been recorded for Chapada Diamantina [53,54], and floristic results indicate different species compositions between rock outcrops from Chapada Diamantina and Planalto da Borborema [3]. In addition, phylogeographic studies have indicated the occurrence of older lineages [32] and genetic differentiation among populations [27,34] occurring within Chapada Diamantina. The patterns of genetic diversity found within Chapada Diamantina populations are in agreement with refuge theory, which predicts higher levels of genetic diversity and older lineages to be present in historically climate-stable regions [5,22].

The relictual nature of inselberg populations is also evident in the structure of the haplotype networks (Figure 1C and 1D). Several mutational steps were observed in the networks calculated for *E. cinnabarinum* and *E. secundum*, suggesting substantial periods of isolation among populations of both species. Deep divisions among haplotypes may be reflective of populations with long divergence histories [16], as demonstrated by phylogeographic studies of Caatingan vertebrate animals having fragmented distributions [27,29,31,32]. Divergence times estimated for Caatingan organisms indicate major crown clade splits to have occurred in the Miocene (between 11,000 ka and 5,000 ka), followed by divergence within these major groups during the Pliocene–Pleistocene [20,27,33,34]. These results are also in agreement with the long-term persistence of populations inhabiting inselbergs [1]. Accordingly, lower levels of genetic differentiation have been observed in other *Epidendrum* species inhabiting well-connected landscapes (e.g. sand dunes and swamp vegetation) [40,55].

Conclusions

The population genetic structure recovered for both *E. cinnabarinum* and *E. secundum* reflects the isolated nature of their inselberg habitats. Similar to the results of studies of other inselberg species [7,9,11,14,15], significant differentiation at plastid loci was observed between populations separated by only a few kilometers (Additional files 2 and 3: Table S2 and S3). Genetic differentiation among populations was not correlated with geographic distance or ecoregions, suggesting that genetic drift may be a significant factor influencing the diversification of inselberg species. According to Porembski et al. [17], inselberg species composition is difficult to predict even over short distances, probably because of stochastic colonization events. Long-term persistence of populations was also supported in our study, suggesting that both species have been restricted to the specific inselberg habitat over long time frames, and have not been reduced to rock outcrops because of climatic instability. The results of EEM analysis indicate that fragmentation and reduction of *E. secundum* distribution has occurred in mountain ranges south of Chapada Diamantina. Future studies should consequently include a broad sampling of this species over adjacent mountain ranges to explore how such distribution oscillations have affected current levels of genetic diversity in these populations. In northeastern Brazil, inselbergs have different mineral origins and are distributed across different biomes and climatic clines. Thus, selection for divergent habitats may also be of great importance during lineage diversification within species under restricted gene flow. Given the contrasting patterns of genetic structure observed between nuclear and plastid markers, cytonuclear incompatibilities may contribute to the first stages of reproductive isolation among divergent lineages. Future efforts should take advantage of next-generation sequencing approaches [56] coupled with experimental studies [57,58] to clarify the role of drift, selection and cytonuclear interactions in the diversification of inselberg species.

Methods

Study species

Epidendrum L. is the largest genus of Orchidaceae in the Neotropics, with about 1,500 species distributed from the southeastern United States to northern Argentina [59]. The genus contains many species with wide distributions and high morphological diversity [38]. *Epidendrum cinnabarinum* and *E. secundum* are perennial, terrestrial species commonly found on inselbergs whose locations include sea-level to mountain elevations. Both species belong to subgenus *Amphyglottium* [60,61], a group of food-deceptive orchids whose flowers are visited by several butterfly species despite the lack of nectar reward (reviewed by [38]). Although the species are self-compatible, pollinators are necessary for seed set (Pinheiro unpub. res., [62]).

Sampling design

Samples of *E. cinnabarinum* and *E. secundum* were obtained from 11 and 9 locations, respectively (Table 6, Figure 1A and B). Sample sites covered most of the geographic distribution of these species in the Caatinga, with additional samples collected in the BAF, an adjacent biome with higher levels of humidity and species diversity (perennial mesic forest). To estimate the degree of isolation of inselberg populations, samples were collected from sites ranging from 12 to 1,000 km apart. Most populations were sampled on rock outcrops; *E. cinnabarinum* was also collected in sand dune vegetation (populations MA, PI and AB, Table 6). Both species occurred in sympatry in populations TO, BZ and FE. For molecular analysis, leaf samples were sliced into small pieces and transferred to silica gel for drying. Total genomic DNA was extracted as described by Pinheiro et al. [63].

Table 6 Geographic location and habitat description of sampled populations

Population	ID	Latitude S	Longitude W	Altitude (m)	Habitat / Biome ¹	Sample Size	
						Nuclear	Plastid
<i>Epidendrum cinnabarinum</i>							
Mamanguape	MA	6.789	34.942	50	Sand dune scrub vegetation / BAF	-	24
Serraria	RR	6.827	35.638	553	Rock outcrop within mesic forest enclave / CAA	-	20
Esperança	ES	7.009	35.899	681	Rock outcrop surrounded by SDTF/ CAA	-	19
Queimadas	QE	7.351	35.900	497	Rock outcrop surrounded by SDTF/ CAA	-	12
Jaqueira	JA	8.740	35.792	720	Rock outcrop surrounded by forest / BAF	-	14
Cabo de Santo Agostinho	CB	8.217	35.003	116	Rock outcrop surrounded by forest / BAF	-	7
Pirambu	PI	10.61	36.867	97	Shrubby Tabuleiro vegetation / BAF	-	14
Abaeté	AB	12.946	38.357	38	Sand dune scrub vegetation / BAF	-	11
<i>Epidendrum secundum</i>							
Maranguape	CE	3.894	38.722	984	Rock outcrop within mesic forest enclave / CAA	-	20
Brejo da Madre de Deus	DE	8.200	36.403	1094	Rock outcrop within mesic forest enclave / CAA	20	20
Serra de Itabaiana	TB	10.739	37.364	572	Rock outcrop surrounded by forest / BAF	-	20
Morro do Chapéu	MC	11.551	41.155	913	Rock outcrop surrounded by SDTF/ CAA	16	16
Palmeiras	PL	12.476	41.452	1082	Rock outcrop surrounded by SDTF/ CAA	-	20
Serra da Jibóia	JI	12.854	39.476	824	Rock outcrop surrounded by forest / BAF	-	20
Sympatric populations							
Pedra de Santo Antonio (C)	TO	7.344	35.798	709	Rock outcrop surrounded by SDTF/ CAA	-	15
Pedra de Santo Antonio (S)						20	20
Bezerros (C)	BZ	8.151	35.756	837	Rock outcrop within mesic forest enclave / CAA	-	21
Bezerros (S)						20	20
Camocim de São Félix (C)	FE	8.328	35.754	718	Rock outcrop surrounded by SDTF/ CAA	-	20
Camocim de São Félix (S)						20	18
Total						96	351

¹SDTF – Seasonal Dry Tropical Forest;

Populations sampled with their identification code (ID), elevation above sea level, habitat description and sample size, analysed for nuclear and plastid markers in *Epidendrum cinnabarinum* (C) and *E. secundum* (S) from Northeast Brazil, distributed in the Caatinga (CAA) and Brazilian Atlantic Forest (BAF) biomes. Populations are indicated as shown on the maps in Figure 1A and B.

Molecular markers and genotyping assays

Six plastid microsatellite loci were used to genotype samples from both species (Epcp02, Epcp04, Epcp05, Epcp07, Epcp08 and Epcp09 [64]). For *E. cinnabarinum*, the intergenic *rps16-trnK* region was also sequenced to detect a 16-bp insertion. Samples of *E. secundum* from five populations (Table 6) were also analyzed at nuclear microsatellite loci Eff06, Eff26, Eff45 [63], Eff48 (forward primer 5'-TGACCGTTTGAACCTTTTGGT-3'; reverse primer 5'-ATCCAGGCATGAGCAGCA-3'), Epp96 [65] and Lspe-3 [66]. Nuclear microsatellites were not amplified for *E. cinnabarinum* samples owing to the polyploid origin of this species ($2n = 240$). All polymerase chain reaction (PCR) amplifications were performed in an Applied Biosystems 2700 thermocycler (Applied Biosystems, Foster City, CA, USA) following the protocol described by Pinheiro et al. [63]. Microsatellite alleles were resolved on an ABI 3130 Genetic Analyzer automated sequencer and were sized with LIZ (500) standard using GENEMAPPER v. 4.1 software (Applied Biosystems).

Genetic diversity of sampled populations

All sampled populations were characterized for levels of diversity based on plastid DNA markers. The number of haplotypes in each population, haplotype diversity and haplotype richness were estimated using RAREFAC v. 3.5 software [67]. Estimates of haplotype richness were corrected for differences in sample size using the rarefaction method.

Nuclear microsatellite diversity of five *E. secundum* populations was characterized according to number of alleles, number of private alleles, allelic richness, private allelic richness, expected heterozygosity and inbreeding coefficient [68], which were calculated using the programs MSA v. 4.05 [69] and HP-RARE v. 1.0 [70]. Departures from Hardy–Weinberg Equilibrium of within-population inbreeding coefficients were identified using exact tests in GENEPOP v. 4.0 [71]. The microsatellite data set was tested for genotyping errors due to stuttering, short allele dominance and null alleles by means of a Monte Carlo simulation of expected allele-size differences as implemented in MICRO-CHECKER [39].

Plastid genetic structuring

The geographical structure of genetic variation in plastid DNA was investigated for both species through several approaches. A median-joining network [72] based on plastid DNA was constructed using the program NETWORK v. 4.5.1.0 (www.fluxus-engineering.com). To assess whether observed genetic differentiation was due to drift, we tested the hypothesis that $G_{ST} = R_{ST}$ [73] (where Slatkin's R_{ST} [74] estimates the contribution of stepwise-like mutations to genetic differentiation) following Pons and Petit [75] using the program PERMUT/CpSSR. To test for the presence of phylogeographic structure among populations (i.e. R_{ST} significantly larger than G_{ST}), 10,000 permutations of R_{ST} values were performed. Partitioning of genetic diversity within and among populations and between populations from BAF (coastal) and Caatinga (inland) regions was assessed by AMOVA. For *E. secundum*, two additional models were considered, one corresponding to differentiation between populations from Chapada Diamantina and Planalto da Borborema, and the other considering populations from Chapada Diamantina separated from remaining localities. Genetic differentiation between populations was also measured by pairwise comparisons of F_{ST} using the program ARLEQUIN v. 3.5. Finally, the hypothesis that populations were differentiated because of isolation by distance was tested by assessing the correlation between pairwise geographic distances and pairwise F_{ST} values using a Mantel test in the program GENEPOP

v. 4.0. Correlation significance was estimated after performing 10,000 permutations between pairwise geographic distance and pairwise genetic differentiation matrices.

Nuclear genetic structuring

Nuclear genetic differentiation among populations of *E. secundum* was measured using the unbiased estimator of relative differentiation G_{ST} [76] and the standardized genetic differentiation measure G'_{ST} [77]. Partitioning of genetic diversity was examined at different hierarchical levels using AMOVA as implemented in ARLEQUIN v. 3.5. The hypothesis that populations were differentiated because of isolation by distance was tested by assessing the correlation between pairwise geographic distance and pairwise G'_{ST} values using a Mantel test in GENEPOP v. 4.0, similar to the approach used for plastid markers. In addition, to test for the presence of phylogeographic structure among populations, R_{ST} was calculated by permuting allele sizes among alleles for 10,000 permutations using the program SPAGEDI v. 1.3d [78], and was then compared with F_{ST} values. Phylogeographic structure was inferred when R_{ST} was significantly larger than F_{ST} .

Bayesian assignment analysis (in STRUCTURE v. 2.3.3 [79]) was used to assign individuals to genetic clusters (K) and to estimate admixture proportions (Q) for each individual. A set of models was chosen in which individuals had admixed ancestries and correlated allele frequencies. Ten replicate runs were completed for each K value for $K = 1-10$. Markov chain Monte Carlo runs consisted of 600,000 generations after an initial burn-in of 250,000 generations. The most probable number of genetic clusters (K) present in the data were defined following Evanno et al. [80] using the program Structure Harvester v. 6.0 [81].

Demographic analyses

Recent population size reductions (i.e. genetic bottlenecks) were tested in *E. secundum* with nuclear microsatellite-based M -ratios [82] using ARLEQUIN v. 3.5. Significance for each population was assessed by comparisons of M -ratios and critical values (M_C values) obtained by simulating the distribution of M -ratios under specific demographic and mutational conditions using the software CRITICAL_M.EXE (<http://swfsc.noaa.gov/textblock.aspx?Division=FED&id=3298>). The critical value M_C is set at the lower 5% tail of this distribution, and bottlenecks are detected when the M -ratio value is below the calculated M_C threshold. Different values of M_C were simulated by modifying θ (0.5, 2.0 and 10.0), pg (0.1 and 0.3) and Δg (2.0 and 3.5). Low M_C values are more conservative as a bottleneck must be of greater intensity to drop below this level [82]; for this reason, the lowest obtained M_C value was used to check bottleneck significance.

To examine the relative contribution of pollen vs. seed flow to total gene flow among *E. secundum* populations, the parameter G_{ST} for nuclear and plastid loci was compared based on populations genotyped for both plastid and nuclear markers (TO, DE, BZ, FE and MC). Pollen/seed flow ratio was estimated following Ennos [83] and Petit et al. [84] using Equation 1 presented by the latter authors.

Environmental envelope models for *E. cinnabarinum* and *E. secundum*

To test the hypothesis that inselbergs were refugia through late Quaternary climate cycles, we projected EEMs from the current situation onto late Quaternary scenarios. We used the maximum entropy algorithm of Maxent v. 3.3.3e [85,86] to obtain current and past

distributions. We used current and past bioclimatic variables derived from monthly temperature and rainfall (Bioclim scheme). Current data were obtained from WorldClim [87]. Two past scenarios were considered: LIG (~120,000–140,000 years BP), based on Otto-Bliesner et al. [88], and LGM (~21,000 years BP), based on Paleoclimate Modelling Intercomparison Project Phase II [89] considering the CCSM3 model [90]. To incorporate geological differentiation between outcrops, we also used three variables derived from Schenk et al. [91]. Geological variables were major geological types (sedimentary/igneous and metamorphic rocks), eras and geological age. Major geological types and eras were categorical and age was continuous. All variables were adjusted to a spatial resolution of 30 arcsec. Highly correlated variables were removed, and jackknifing was used to estimate variable importance. The final models were obtained considering only variables with contributions to AUC higher than 75%. Five model replicates were run for each one of the presence-only methods, with 75% of occurrences used for calibration and different subsets (25%) used for validation. Mean AUC was used to assess the performance of the models [92], where 1 was the maximum prediction and 0.5 suggested a random prediction. Past distributions were estimated by projecting the current relationship onto scenarios of past climate, assuming that current relationships between climate and distribution were maintained.

We obtained 226 *E. cinnabarinum* and 242 *E. secundum* occurrences from our field GPS records and georeferenced herbarium data extracted from the *speciesLink* project (<http://splink.cria.org.br>). All points were verified with Google Earth to ensure localities were not placed in heavily urbanized areas. The localities used for EEM analysis are available from the authors upon request to avoid illegal plant collecting.

Abbreviations

EEMs, Environmental envelope models; LGM, Last Glacial Maximum; BAF, Brazilian Atlantic Forest; LIG, Last Interglacial; AUC, Area under the curve

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

FP participated in the study design, carried out the molecular genetic studies and drafted the manuscript. SC and CPS participated in the study design, and FB coordinated the study. LPF participated in the field work and sampling strategy design. MFF contributed to chemical supply procurement and analysis. DDM conducted the EEM analysis. All authors help to draft the manuscript, and read and approved the final version.

Acknowledgements

We thank M. Alves, E. M. Pessoa, C. van den Berg, R. B. Louzada, L. W. Lima Verde, S. R. Nobrega, M. E. M. Fotunato, A. P. Prata, F. N. M. Assis and A. Alves-Araújo for help during fieldwork, Robyn S. Cowan, Dion Devey and Edith Kapinos for assistance in the laboratory, and F. Pareyn for access to the Caatinga ecoregions dataset. Funding for this study was

provided by grants from the Prance Fellowship in Neotropical Botany under the Kew Latin American Research Fellowship Programme (KLARF), the Systematics Research Fund (the Linnean Society and the Systematics Association-UK) and the International Association for Plant Taxonomy to FP, and grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq 471929/2006-9), CNPq/CNR International Cooperation grant (Brazil/Italy), and Fundação de Amparo à Pesquisa do Estado de São Paulo to FB (FAPESP 2006/54189-3), CPS (FAPESP-Biota 2009/52725-3 and 2009/17411-8), and FP (FAPESP 2009/15052-0).

References

1. Porembski S, Barthlott W: *Inselbergs. Biotic diversity of isolated rock outcrops in tropical and temperate regions*. Berlin: Springer-Verlag; 2000.
2. Seine R, Porembski S, Becker U: **Phytogeography**. In *Inselbergs. Biotic diversity of isolated rock outcrops in tropical and temperate regions*. Edited by Porembski S, Barthlott W. Berlin: Springer; 2000:435–448.
3. Gomes P, Alves M: **Floristic and vegetational aspects of an inselberg in the semi-arid region of Northeast Brazil**. *Edinb J Bot* 2009, **66**:1–18.
4. Speziale KL, Ezcurra C: **The role of outcrops in the diversity of patagonian vegetation: relicts of glacial palaeofloras?** *Flora* 2012, **207**:141–149.
5. Scarano FR: **Structure, function and floristic relationships of plant communities in stressful habitats marginal to the Brazilian atlantic rainforest**. *Ann Bot* 2002, **90**:1–8.
6. MacArthur RH, Wilson EO: *The theory of island biogeography*. Princeton: Princeton University Press; 1967.
7. Palma-Silva C, Wendt T, Pinheiro F, Barbará T, Fay MF, Cozzolino S, Lexer C: **Sympatric bromeliad species (*Pitcairnia* spp.) facilitate tests of mechanisms involved in species cohesion and reproductive isolation in Neotropical inselbergs**. *Mol Ecol* 2011, **20**:3185–3201.
8. Levy F, Neal CL: **Spatial and temporal genetic structure in chloroplast and allozyme markers in *Phacelia dubia* implicate genetic drift**. *Heredity* 1999, **82**:422–431.
9. Byrne M, Hopper SD: **Granite outcrops as ancient islands in old landscapes: evidence from the phylogeography and population genetics of *Eucalyptus caesia* (Myrtaceae) in Western Australia**. *Biol J Linn Soc* 2008, **93**:177–188.
10. Collevatti RG, Rabelo SG, Vieira RF: **Phylogeography and disjunct distribution in *Lychnophora ericoides* (Asteraceae), an endangered cerrado shrub species**. *Ann Bot* 2009, **104**:655–664.
11. Barbosa AR, Fiorini CF, Silva-Pereira V, Mello-Silva R, Borba EL: **Geographical genetic structuring and phenotypic variation in the *Vellozia hirsuta* (Velloziaceae) ochlospecies complex**. *Am J Bot* 2012, **99**:1477–1488.

12. Sarthou C, Samadi S, Boisselier-Dubayle MC: **Genetic structure of the saxicole *Pitcairnia geyskesii* (Bromeliaceae) on inselbergs in French Guiana.** *Am J Bot* 2001, **88**:861–868.
13. DeChaine EG, Martin AP: **Marked genetic divergence among sky island populations of *Sedum lanceolatum* (Crassulaceae) in the Rocky Mountains.** *Am J Bot* 2005, **92**:477–486.
14. Barbará T, Martinelli G, Fay MF, Mayo SJ, Lexer C: **Population differentiation and species cohesion in two closely related plants adapted to neotropical high-altitude ‘inselbergs’, *Alcantarea imperialis* and *Alcantarea geniculata* (Bromeliaceae).** *Mol Ecol* 2007, **16**:1981–1992.
15. Barbará T, Martinelli G, Palma-Silva C, Fay MF, Mayo SJ, Lexer C: **Genetic relationships and variation in reproductive strategies in four closely related bromeliads adapted to Neotropical ‘inselbergs’: *Alcantarea glaziouana*, *A. regina*, *A. geniculata* and *A. imperialis* (Bromeliaceae).** *Ann Bot* 2009, **103**:65–77.
16. Freeland JR: *Molecular Ecology*. Chichester: John Wiley & Sons; 2005.
17. Porembski S, Martinelli G, Ohlemüller R, Barthlott W: **Diversity and ecology of saxicolous vegetation mats on inselbergs in the Brazilian Atlantic rainforest.** *Divers Distrib* 1998, **4**:107–119.
18. Bermingham E, Moritz C: **Comparative phylogeography: concepts and applications.** *Mol Ecol* 1998, **7**:367–369.
19. Arbogast BS, Kenagy GJ: **Comparative phylogeography as an integrative approach to historical biogeography.** *J Biogeogr* 2001, **28**:819–825.
20. Hughes CE, Pennington RT, Antonelli A: **Neotropical plant evolution: assembling the big picture.** *Bot J Linn Soc* 2013, **171**:1–18.
21. Turchetto-Zolet AC, Pinheiro F, Salgueiro F, Palma-Silva C: **Phylogeographic patterns shed light on evolutionary process in South America.** *Mol Ecol* 2013, **22**:1193–1213.
22. Bennett KD, Provan J: **What do we mean by ‘refugia’?** *Quat Sci Rev* 2008, **27**:2449–2455.
23. Collevatti RG, Terribile LC, Oliveira G, Lima-Ribeiro MS, Nabout JC, Rangel TF, Diniz-Filho JAF: **Drawbacks to palaeodistribution modelling: the case of South American seasonally dry forests.** *J Biogeogr* 2013, **40**:345–358.
24. Araújo MB, Peterson AT: **Uses and misuses of bioclimatic envelope modelling.** *Ecology* 2012, **93**:1527–1539.
25. Collevatti RG, Telles MPC, Nabout JC, Chaves LJ, Soares TN: **Demographic history and the low genetic diversity in *Dipteryx alata* (Fabaceae) from Brazilian Neotropical savannas.** *Heredity* 2013, **111**:97–105.

26. Jakob SS, Martinez-Meyer E, Blattner FR: **Phylogeographic analyses and paleodistribution modeling indicate Pleistocene *in situ* survival of *Hordeum* species (Poaceae) in southern Patagonia without genetic or spatial restriction.** *Mol Biol Evol* 2009, **26**:907–923.
27. Werneck FP, Gamble T, Colli GR, Rodrigues MT, Sites JW Jr: **Deep diversification and long-term persistence in the South American ‘Dry Diagonal’: integrating continent-wide phylogeography and distribution modeling of geckos.** *Evolution* 2012, **66**:3014–3034.
28. Queiroz LP: **The Brazilian Caatinga: phytogeographical patterns inferred from distribution data of the Leguminosae.** In *In Neotropical savannas and seasonally dry forests: plant diversity, biogeography and conservation*. Edited by Pennington RT, Lewis GP, Ratter JA. Boca Raton: CRC Press; 2006:121–157.
29. Carnaval AC, Bates JM: **Amphibian DNA shows marked genetic structure and tracks Pleistocene climate change in Northeastern Brazil.** *Evolution* 2007, **61**:2942–2957.
30. Caetano S, Prado D, Pennington RT, Beck S, Oliveira-Filho AT, Spichiger R, Naciri Y: **The history of seasonally dry tropical forests in eastern South America: inferences from the genetic structure of the tree *Astronium urundeuva* (Anacardiaceae).** *Mol Ecol* 2008, **17**:3147–3159.
31. Moraes EM, Yotoko KSC, Manfrin MH, Solferini VN, Sene FM: **Phylogeography of the cactophilic species *Drosophila gouveai*: demographic events and divergence timing in dry vegetation enclaves in eastern Brazil.** *J Biogeogr* 2009, **36**:2136–2147.
32. Franco FF, Manfrin MH: **Recent demographic history of cactophilic *Drosophila* species can be related to Quaternary palaeoclimatic changes in South America.** *J Biogeogr* 2012, **40**:142–154.
33. Pennington RT, Lavin M, Prado DE, Pendry CA, Pell SK, Butterworth CA: **Historical climate change and speciation: neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification.** *Philos Trans R Soc Lond B Biol Sci* 2004, **359**:515–538.
34. Novaes R, Lemos-Filho J, Ribeiro R, Lovato M: **Phylogeography of *Plathyenia reticulata* (Leguminosae) reveals patterns of recent range expansion towards northeastern Brazil and southern Cerrados in Eastern Tropical South America.** *Mol Ecol* 2010, **19**:985–998.
35. Werneck FP, Costa GC, Colli GR, Prado DE, Sites JW Jr: **Revisiting the historical distribution of seasonally dry tropical forests: new insights based on palaeodistribution modelling and palynological evidence.** *Glob Ecol Biogeogr* 2011, **20**:272–288.
36. Conceição AA, Pirani JR, Meirelles ST: **Floristics, structure and soil of insular vegetation in four quartzite-sandstone outcrops of “Chapada Diamantina”, Northeast Brazil.** *Rev Bras Bot* 2007, **30**:641–656.

37. Moraes AP, Chinaglia M, Palma-Silva C, Pinheiro F: **Interploidy hybridization in sympatric zones: the formation of *Epidendrum fulgens* × *E. puniceoluteum* hybrids (Epidendroideae, Orchidaceae).** *Ecol Evol* 2013, **3**:3824–3837.
38. Pinheiro F, Cozzolino S: ***Epidendrum* (Orchidaceae) as a model system in ecological and evolutionary studies in Neotropics.** *Taxon* 2013, **62**:77–88.
39. van Oosterhout C, Hutchinson WF, Wills DPM, Shipley PF: **Micro-checker: software for identifying and correcting genotyping errors in microsatellite data.** *Mol Ecol Notes* 2004, **4**:535–538.
40. Pinheiro F, Barros F, Palma-Silva C, Fay MF, Lexer C, Cozzolino S: **Phylogeography and genetic differentiation along the distributional range of the orchid *Epidendrum fulgens*: a Neotropical coastal species not restricted to glacial refugia.** *J Biogeogr* 2011, **38**:1923–1935.
41. Safford HD, Martinelli G: **Southeast Brazil.** In *Inselbergs. Biotic diversity of isolated rock outcrops in tropical and temperate regions.* Edited by Porembski S, Barthlott W. Berlin: Springer; 2000:339–389.
42. Behling H, Arz HW, Patzold J, Wefer G: **Late Quaternary vegetational and climate dynamics in northeastern Brazil, inferences from marine core GeoB 3104–1.** *Quat Sci Rev* 2000, **19**:981–994.
43. Mayol M, Palau C, Rosselló JA, González-Martínez SC, Molins A, Riba M: **Patterns of genetic variability and habitat occupancy in *Crepis triasii* (Asteraceae) at different spatial scales: insights on evolutionary processes leading to diversification in continental islands.** *Ann Bot* 2012, **109**:429–441.
44. Hutchison DW, Templeton AR: **Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability.** *Evolution* 1999, **53**:1898–1914.
45. Birky CW Jr: **Relaxed cellular controls and organelle heredity.** *Science* 1983, **222**:468–475.
46. Barbará T, Lexer C, Martinelli G, Mayo S, Fay MF, Heuertz M: **Within-population spatial genetic structure in four naturally fragmented species of a Neotropical inselberg radiation, *Alcantarea imperialis*, *A. geniculata*, *A. glaziouana* and *A. regina* (Bromeliaceae).** *Heredity* 2008, **101**:285–296.
47. Cozzolino S, Widmer A: **Orchid diversity: an evolutionary consequence of deception?** *Trends Ecol Evol* 2005, **20**:487–494.
48. Phillips RD, Dixon KW, Peakall R: **Low population genetic differentiation in the Orchidaceae: implications for the diversification of the family.** *Mol Ecol* 2012, **21**:5208–5220.
49. Ramos ACS, Lemos-Filho JP, Lovato MB: **Phylogeographical structure of the Neotropical forest tree *Hymenaea courbaril* (Leguminosae: Caesalpinioideae) and its**

relationship with the vicariant *Hymenaea stigonocarpa* from Cerrado. *J Heredity* 2009, **100**:206–216.

50. Martins FM, Templeton A, Pavan ACO, Kohlbach BC, Morgante JS: **Phylogeography of the common vampire bat (*Desmodus rotundus*): Marked population structure, Neotropical Pleistocene vicariance and incongruence between nuclear and mtDNA markers.** *BMC Evol Biol* 2009, **9**:294.

51. Pinheiro F, Cozzolino S, Barros F, Gouveia TMZM, Suzuki RM, Fay MF, Palma-Silva C: **Phylogeographic structure and outbreeding depression reveal early stages of reproductive isolation in the Neotropical orchid *Epidendrum denticulatum*.** *Evolution* 2013, **67**:2024–2039.

52. Mäder G, Fregonezi JN, Lorenz-Lemke AP, Bonatto SL, Freitas LB: **Geological and climatic changes in quaternary shaped the evolutionary history of *Calibrachoa heterophylla*, an endemic South-Atlantic species of petúnia.** *BMC Evol Biol* 2013, **13**:178.

53. Giulietti AM, Pirani JR: **Patterns of geographical distribution of some plant species from Espinhaço range, Minas Gerais and Bahia, Brazil.** In *Proceedings of a workshop on Neotropical distribution patterns*. Edited by Vanzolini PE, Heyer WR. Rio de Janeiro: Academia Brasileira de Ciências; 1988:39–69.

54. Rapini A, Ribeiro PL, Lambert S, Pirani JR: **A flora dos campos rupestres da Cadeia do Espinhaço.** *Megadiversidade* 2008, **4**:15–23.

55. Pinheiro F, Barros F, Palma-Silva C, Meyer D, Fay MF, Suzuki RM, Lexer C, Cozzolino S: **Hybridization and introgression across different ploidy levels in the Neotropical orchids *Epidendrum fulgens* and *E. puniceoluteum* (Orchidaceae).** *Mol Ecol* 2010, **19**:3981–3994.

56. Twyford AD, Ennos RA: **Next generation hybridization and introgression.** *Heredity* 2012, **108**:179–189.

57. Sambatti JBM, Ortiz-Barrientos D, Baack EJ, Rieseberg LH: **Ecological selection maintains cytonuclear incompatibilities in hybridizing sunflowers.** *Ecol Lett* 2008, **11**:1082–1091.

58. Moyle LC, Olson MS, Tiffin P: **Patterns of reproductive isolation in three angiosperm genera.** *Evolution* 2004, **58**:1195–1208.

59. Hágsater E, Soto-Arenas MA: *Epidendrum* L. In *Genera Orchidacearum Vol. 4*. Edited by Pridgeon AM, Cribb PJ, Chase MW, Rasmussen FN. Oxford: Oxford University Press; 2005:236–251.

60. Salatino ML, Salatino A, Barros F, Pinheiro F, Koehler S, Corrêa AM F: **Phylogenetic relationships and infrageneric classification of *Epidendrum* subgenus *Amphiglottium* (Laeliinae, Orchidaceae).** *Plant Systematics and Evolution* 2009, **283**:165–177.

61. Pessoa EM, Alves M, Alves-Araújo A, Palma-Silva C, Pinheiro F: **Integrating different tools to disentangle species complexes: a case study in *Epidendrum* (Orchidaceae).** *Taxon* 2012, **61**:721–734.
62. Pansarin ER, Amaral MCE: **Reproductive biology and pollination mechanisms of *Epidendrum secundum* (Orchidaceae) - floral variation: a consequence of natural hybridization?** *Plant Biol* 2008, **10**:211–219.
63. Pinheiro F, Santos MO, Barros F, Meyer D, Salatino A, Souza AP, Cozzolino S: **Isolation and characterization of microsatellite loci in the Brazilian orchid *Epidendrum fulgens*.** *Conserv Genet* 2008, **9**:1661–1663.
64. Pinheiro F, Palma-Silva C, Barros F, Félix LP, Lexer C, Cozzolino S, Fay MF: **Chloroplast microsatellite markers for the Neotropical orchid genus *Epidendrum*, and cross-amplification in other Laeliinae species (Orchidaceae).** *Conserv Genet Resour* 2009, **1**:505–511.
65. Pinheiro F, Santos MO, Palma-Silva C, Barros F, Meyer D, Salatino A, Souza AP, Cozzolino S: **Isolation and characterization of microsatellite loci in *Epidendrum puniceoluteum*, an endemic orchid from the Atlantic Rainforest.** *Mol Ecol Resour* 2008, **8**:1114–1116.
66. Cortés-Palomec AC, McCauley RA, Oyama K: **Isolation, characterization and cross-amplification of polymorphic microsatellite loci in *Laelia speciosa* (Orchidaceae).** *Mol Ecol Resour* 2008, **8**:135–138.
67. Petit RJ, El Mousadik A, Pons O: **Identifying populations for conservation on the basis of genetic markers.** *Conserv Biol* 1998, **12**:844–855.
68. Weir BS, Cockerham CC: **Estimating *F*-statistics for the analysis of population structure.** *Evolution* 1984, **38**:1358–1370.
69. Dieringer D, Schlötterer C: **Microsatellite analyzer (MSA): a platform independent analysis tool for large microsatellite data sets.** *Molecular Ecology Notes* 2003, **3**:167–169.
70. Kalinowski ST: **hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness.** *Mol Ecol Notes* 2005, **5**:187–189.
71. Raymond M, Rousset F: **GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism.** *J Hered* 1995, **86**:248–249.
72. Bandelt HJ, Forster P, Roehl A: **Median-joining networks for inferring intraspecific phylogenies.** *Mol Biol Evol* 1999, **16**:37–48.
73. Heuertz M, Freville H, Charbonnel N, Hardy OJ: **Microsatellite allele sizes: a simple test to assess their significance.** *Genetics* 2003, **163**:1467–1482.
74. Slatkin M: **A measure of population subdivision based on microsatellite allele frequencies.** *Genetics* 1995, **139**:457–462.

75. Pons O, Petit RJ: **Measuring and testing genetic differentiation with ordered vs. unordered alleles.** *Genetics* 1996, **144**:1237–1245.
76. Nei M, Chesser RK: **Estimation of fixation indices and gene diversities.** *Ann Hum Genet* 1983, **47**:253–259.
77. Hedrick P: **A standardized genetic differentiation measure.** *Evolution* 2005, **59**:1633–1638.
78. Hardy OJ, Vekemans X: **SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels.** *Mol Ecol Notes* 2002, **2**:618–620.
79. Hubisz MJ, Falush D, Stephens M, Pritchard JK: **Inferring weak population structure with the assistance of sample group information.** *Mol Ecol Resour* 2009, **9**:1322–1332.
80. Evanno G, Regnaut S, Goudet J: **Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study.** *Mol Ecol* 2005, **14**:2611–2620.
81. Earl DA, von Holdt BM: **STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method.** *Conserv Genet Resour* 2012, **4**:359–361.
82. Garza JC, Williamson EG: **Detection of reduction in population size using data from microsatellite loci.** *Mol Ecol* 2001, **10**:305–318.
83. Ennos RA: **Estimating the relative rates of pollen and seed migration among plant-populations.** *Heredity* 1994, **72**:250–259.
84. Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, Vendramin GG: **Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations.** *Mol Ecol* 2005, **14**:689–701.
85. Phillips SJ, Dudik M, Schapire RE: **A maximum entropy approach to species distribution modeling.** *ACM International Proc Ser* 2004, **69**:655.
86. Phillips SJ, Dudik M, Schapire RE: **Maximum entropy modelling of species geographic distributions.** *Ecol Model* 2006, **190**:231–259.
87. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A: **Very high resolution interpolated climate surfaces for global land areas.** *Int J Climatol* 2005, **25**:1965–1978.
88. Otto-Bliesner BL, Marshall S, Overpeck J, Miller G, Hu A, CAPE Last Interglacial Project Members: **Simulating Arctic climate warmth and icefield retreat in the last interglaciation.** *Science* 2006, **311**:1751–1753.
89. Braconnot P, Otto-Bliesner B, Harrison S, Joussaume S, Peterchmitt JY, Abe-Ouchi A, Crucifix M, Driesschaert E, Fichet T, Hewitt CD, Kageyama M, Kitoh A, Laine A, Loutre MF, Marti O, Merkel U, Ramstein G, Valdes P, Weber SL, Yu Y, Zhao Y: **Results of PMIP2 coupled simulations of the mid-Holocene and Last Glacial Maximum — Part 1: experiments and large-scale features.** *Climate of the Past* 2007, **3**:261–277.

90. Collins WD, Bitz CM, Blackmon ML, Bonan GB, Bretherton CS, Carton JA, Chang P, Doney SC, Hack JJ, Henderson TB, Kiehl JT, Large WG, McKenna DS, Santer BD, Smith RD: **The Community Climate System Model Version 3 (CCSM3)**. *J Climate* 2006, **19**:2122–2143.

91. Schenk CJ, Viger RJ, Anderson CP: *Maps showing geology, oil and gas fields, and geologic provinces of South America*. Denver Colorado: 97-470D. U.S. Geological Survey; 1999. <http://pubs.usgs.gov/of/1997/ofr-97-470/OF97-470D/>.

92. Fielding AH, Bell JF: **A review of methods for the assessment of prediction errors in conservation presence/absence models**. *Environ Conserv* 1997, **24**:38–49.

Additional files

Additional_file_1 as DOC

Additional file 1: Table S1. Description of 10 chloroplast microsatellite haplotypes of *Epidendrum cinnabarinum* and 12 haplotypes of *E. secundum*, characterized at six cpSSR loci. For *E. cinnabarinum*, the presence (1) and absence (0) of a 16 bp fragment from intergenic region *rps16–trnK* is also reported. The frequency of occurrence (*n*) in total collection screened is indicated for both species.

Additional_file_2 as DOC

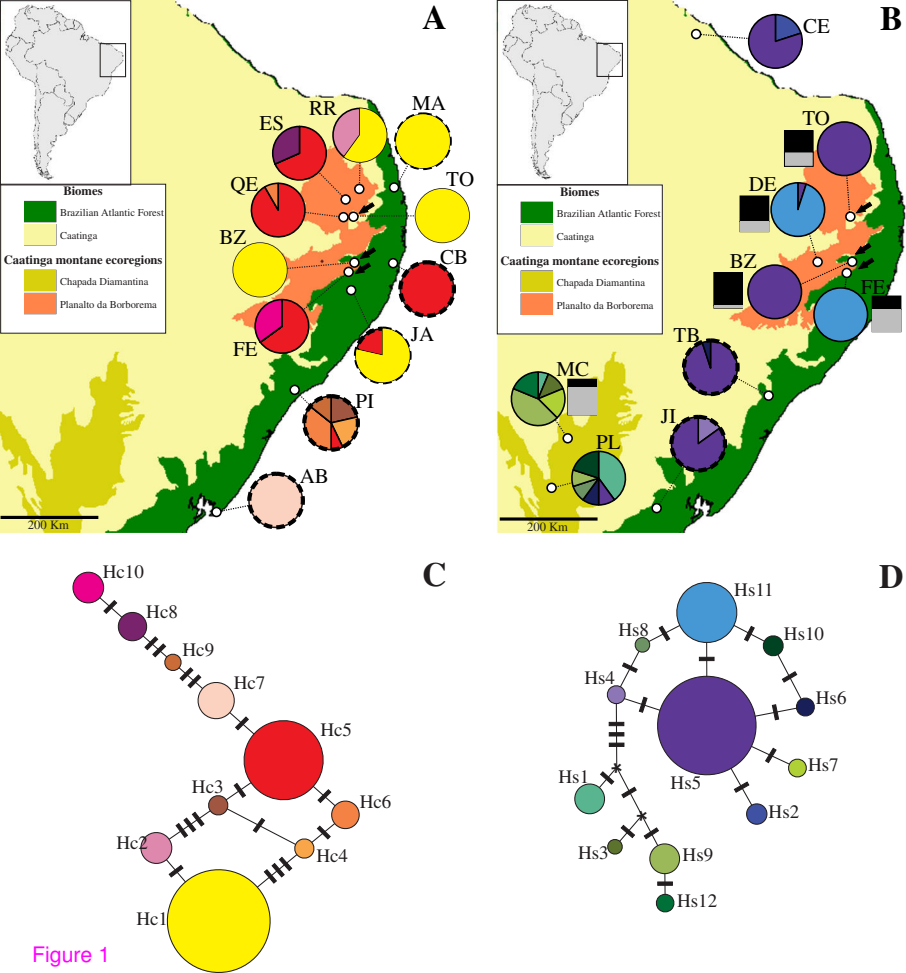
Additional file 2: Table S2. Pairwise comparisons of F_{ST} between populations of *Epidendrum cinnabarinum* based on plastid markers.

Additional_file_3 as DOC

Additional file 3: Table S3. Pairwise comparisons of F_{ST} between populations of *Epidendrum secundum* based on plastid markers.

Additional_file_4 as PDF

Additional file 4: Figure S1. Magnitude of ΔK from STRUCTURE analysis as a function of *K* (number of genetic groups, details in Methods) calculated according to the simulations described by Evanno et al. [80]. The modal value of these distributions indicates the true *K* or the uppermost level of structure—in the present case, two genetic clusters.



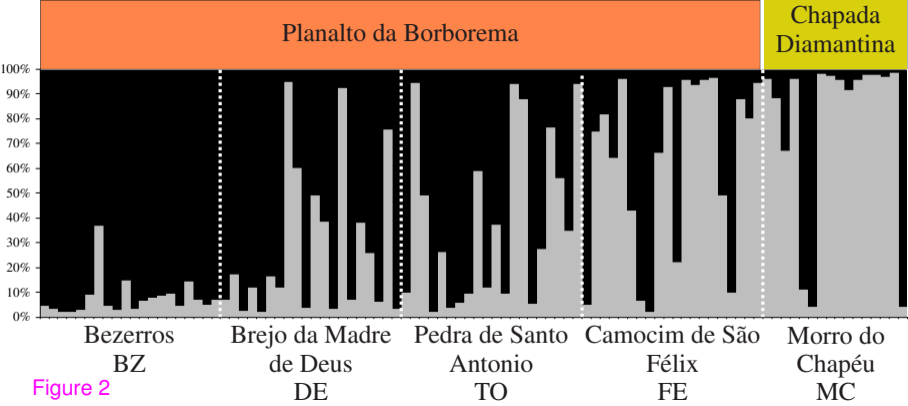
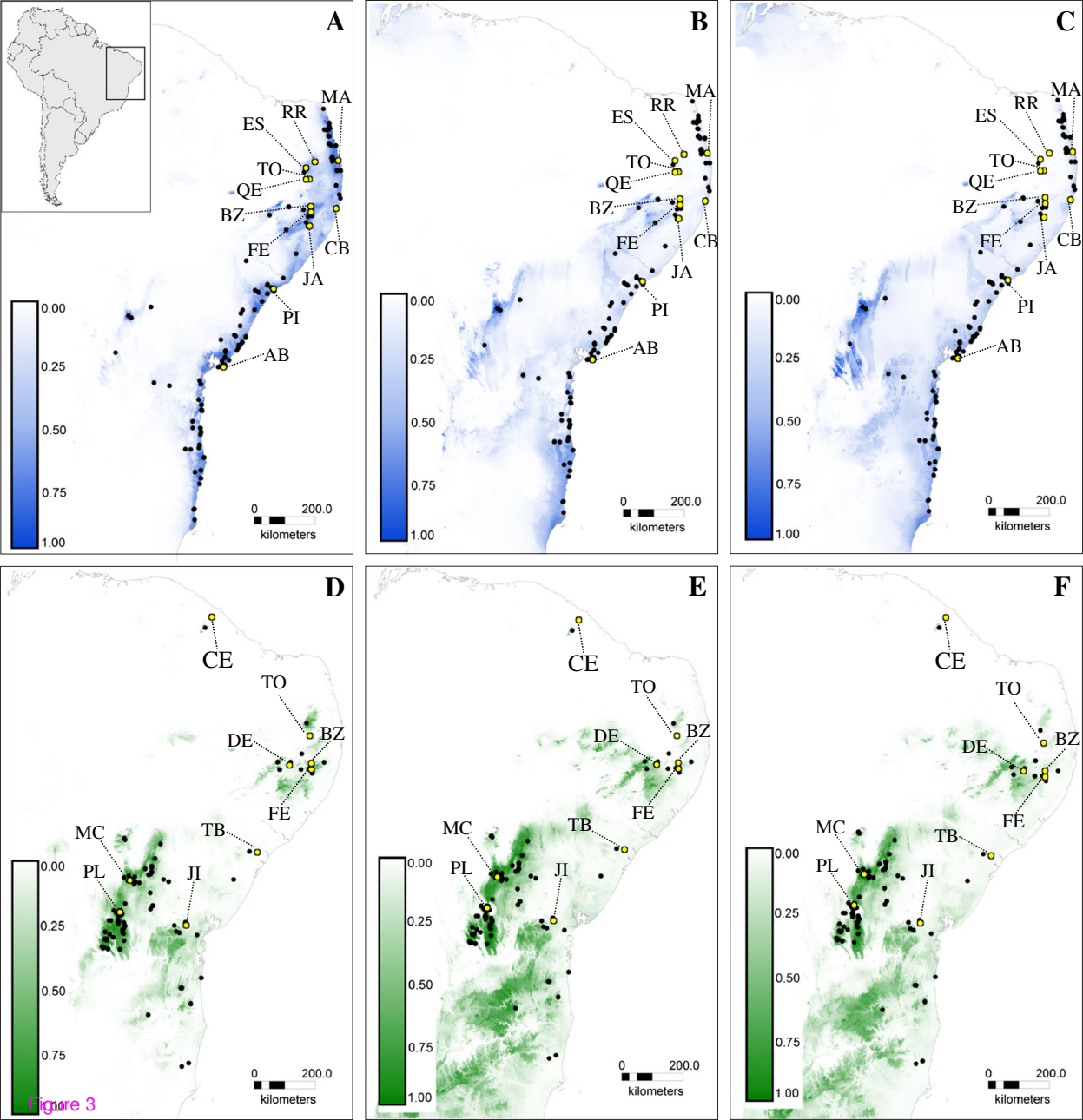


Figure 2



Additional files provided with this submission:

Additional file 1: 5053972611801867_add1.doc, 56K

<http://www.biomedcentral.com/imedia/1019607452124057/supp1.doc>

Additional file 2: 5053972611801867_add2.doc, 36K

<http://www.biomedcentral.com/imedia/1671561147124057/supp2.doc>

Additional file 3: 5053972611801867_add3.doc, 33K

<http://www.biomedcentral.com/imedia/3841292911240578/supp3.doc>

Additional file 4: 5053972611801867_add4.pdf, 438K

<http://www.biomedcentral.com/imedia/1836016764124057/supp4.pdf>