ORIGINAL ARTICLE

Patterns of genetic variability in central and peripheral populations of *Dipteryx alata* (Fabaceae) in the Brazilian Cerrado

Thannya N. Soares · José A. F. Diniz-Filho ·
João C. Nabout · Mariana P. de Campos Telles ·
Levi C. Terribile · Lázaro J. Chaves

Abstract This study tested whether genetic parameters in *Dipteryx alata* populations, estimated from genomic and chloroplastidial microsatellite markers, were distributed according to a central-peripheral model, inferring which factors drive this spatial distribution of genetic variability within populations. For each of the 23 populations sampled throughout the species' range, the mean number of alleles per locus, expected heterozygosity and intrapopulation fixation indices were calculated using a rarefaction approach based on 54 alleles from 8 nuclear microsatellites. Explanatory variables were grouped into three subsets: the ecological suitability estimated by combining different techniques of ecological niche modeling, variables expressing human occupation, and a historical variable represented by the first eigenvector from the pairwise F_{ST} matrix based on cpDNA microsatellites. Each response variable was modeled using first (linear) and second (quadratic) order trend surface analysis (TSA). Multiple regressions were then used to evaluate the relative effects of the explanatory variables, based on AIC multi-model selection. In general, the genetic parameters did not follow a classical central-periphery model. Ecological suitability had a significance influence in all genetic parameters, so more suitable regions have higher genetic diversity and low endogamy. There was also a relationship between fixation indices and human impacts. The high genetic diversity in the southwestern region of Cerrado suggested that recent range expansion (after the Last Glacial Maximum) may also influenced the observed intrapopulation genetic patterns. Thus, complex combinations of both historical and ecological drivers, as well as contemporary human occupation, seem to drive current genetic composition within *D. alata* populations throughout its geographic range.

Keywords Conservation · Microsatellite · Niche model · Tropical tree

Handling editor: Sylvain Razafimandimbison.

T. N. Soares (✉) · M. P. de Campos Telles
Departamento de Genetica, Instituto de Ciências Biológicas, Universidade Federal de Goiás (UFG), Campus Samambaia, Caixa Postal 131, Goiânia, GO 74001-970, Brazil
e-mail: tnsoares@gmail.com

J. A. F. Diniz-Filho
Departamento de Ecologia, Instituto de Ciências Biológicas, UFG, Goiânia, GO, Brazil

J. C. Nabout
Unidade de Ciências Exatas e Tecnológicas, Universidade Estadual de Goiás (UEG), Anápolis, GO, Brazil

L. C. Terribile
Departamento de Biologia, UFG, Campus Jataí, Jataí, GO, Brazil

L. J. Chaves
Escola de Agronomia, UFG, Goiânia, GO, Brazil

Introduction

The genetic structure of populations is determined by several processes, including a complex balance among gene flow, natural selection, and genetic drift, which becomes more pronounced as population size decreases (Ginzburg and Golenberg 1985). Small population size will occur in regions with low environmental suitability with less adapted individuals, or in regions of recent range expansion (Kirkpatrick and Barton 1997). The relative magnitude of such processes is also strongly affected by ecological and life-history traits, including population

density, spatial distribution of individuals within and among local populations, dispersal mechanisms and reproductive system (Epperson 2003).

The ecological and genetic population patterns can vary throughout species' geographical range (Gaston 2003; Sagarin et al. 2006; Pfeifer et al. 2009). The simplest model describing such variation is the central-peripheral model, which predicts that central populations display larger effective sizes and gene flow rates than those in the periphery (Soulé 1973; Maurer 1994; Sagarin and Gaines 2002; Eckert et al. 2008). Under this model, peripheral populations are thought to lose genetic variability due to relatively lower population densities caused by more intense genetic drift and endogamy reducing individual fitness values and population adaptability, which may consequently lead to population extinction and help to define range limits (Kirkpatrick and Barton 1997; Eckert et al. 2008).

Studies testing the central-peripheral model started in the context of synthetic theory of evolution in the 1950s, but this model had recently regained attention in the scientific community (see Eckert et al. 2008; Lira-Noriega and Manthey 2014). However, despite the growing interest regarding investigation of central-peripheral models, these studies are operationally difficult, mainly because they require detailed local population parameters at broad geographic scales and covering most of species ranges, especially related to local abundances or densities that corroborate loss of genetic diversity due to stochastic processes (Gaston and Blackburn 2000). Consequently, comparisons between central and peripheral populations are frequently based on relatively small data sets and do not take into account the whole geographic distribution of the species (Eckert et al. 2008). Therefore, in this context, niche modeling techniques may be important tools for studying potential geographic distributions of species by providing a description of environmental suitability that may be related to population density (e.g., Torres et al. 2012), so this may help testing the relationship between genetic variability and demographic patterns (i.e., population abundance and growth rates) across species' ranges (Diniz-Filho et al. 2009a; Lira-Noriega and Manthey 2014). Also, using ENMs as predictors in central-peripheral models allows incorporating the niche dimension into expected patterns, and this may be important because genetic dynamics can be driven by environmental space (i.e., driving genetic variation from areas with more to less suitable environments) that may be not entirely coincident with geographical space (Diniz-Filho et al. 2009a).

Another important issue is that most studies testing the central-peripheral model are carried out on temperate environment species from the Northern hemisphere, resulting in a strong geographic bias (Eckert et al. 2008;

but see Diniz-Filho et al. 2009a). Although there is certainly a lack of studies in the tropical regions, it would be interesting to emphasize that processes operating in temperate and tropical regions may be different due to several aspects. For instance, there may be differences in current patterns of gene flow under more heterogeneous environments and landscapes in the tropics and, moreover, temperate regions must be more strongly affected by recent historical processes such as glaciation dynamics throughout the Pleistocene. In all cases, range dynamics after the Last Glacial Maximum (LGM) is usually an important explanation for current patterns of diversity, in northern hemisphere due to wide changes in temperature and in the southern hemisphere, particularly in Central Brazil, due to changes in precipitation (see Terribile et al. 2012).

Thus, although some studies have corroborated the central-peripheral model (see Maurer 1994 and Eckert et al. 2008 for a review), other historical or contemporary factors may alter this simple pattern (Vucetich and Waite 2003). For instance, environmental fluctuations related to historical effects (e.g., LGMs, as previously pointed out) may allow range shifts, that can be coupled with habitat loss and fragmentation due to recent anthropogenic impacts (Holderegger and Wagner 2006; Diniz-Filho et al. 2009a). In both cases, fragmentation and isolation of populations will tend to decrease population size, increasing stochastic fluctuations of allele frequencies (genetic drift) and, in the long run, allowing adaptations to local environmental conditions, creating multiple abundance peaks (Petit et al. 1997). If the patterns of range shift and fragmentation due to historical climate changes or anthropogenic impact follow a directional pattern (which is common), this will disrupt the expected stochastic pattern of differentiation under central-peripheral models (see Fig. 1 in Diniz-Filho et al. 2009a).

The Cerrado region in Central Brazil is considered one of the world's conservation hotspots, due to a high level of endemism for plants (which posses several particular traits related to adaptation to the long dry season) and high levels of threats due to human extensive activities and habitat loss (related to agricultural expansions that started in the early 1950s—see Klink and Moreira 2002; Klink and Machado 2005). The Cerrado species include several trees whose fruits have economic importance for local economies, and now many of them may be considered endangered due to habitat loss and climate changes (Nabout et al. 2011). One of these economically important Cerrado species is *Dipteryx alata* Vogel (Fabaceae), popularly known as "baru" tree, a widely distributed large tree species endemic to the biome, living in seasonal savanna habitats and growing in eutrophic and drained soils ("cerradão") (Oliveira and Sigrist 2008). The species is hermaphroditic and pollination is mainly performed by large- and medium-sized bees.

Fig. 1 Modeled geographic distribution of *Dipteryx alata*, shown as the average frequency of occurrence from 14 distinct modeling techniques. *Darker regions* indicate more overlap of models and higher ecological suitability

Suitability

0 - 0.1
0.1 - 0.2
0.2 - 0.3
0.3 - 0.4
0.4 - 0.5
0.5 - 0.6
0.6 - 0.7
0.7 - 0.8
0.8 - 0.9
0.9 - 1
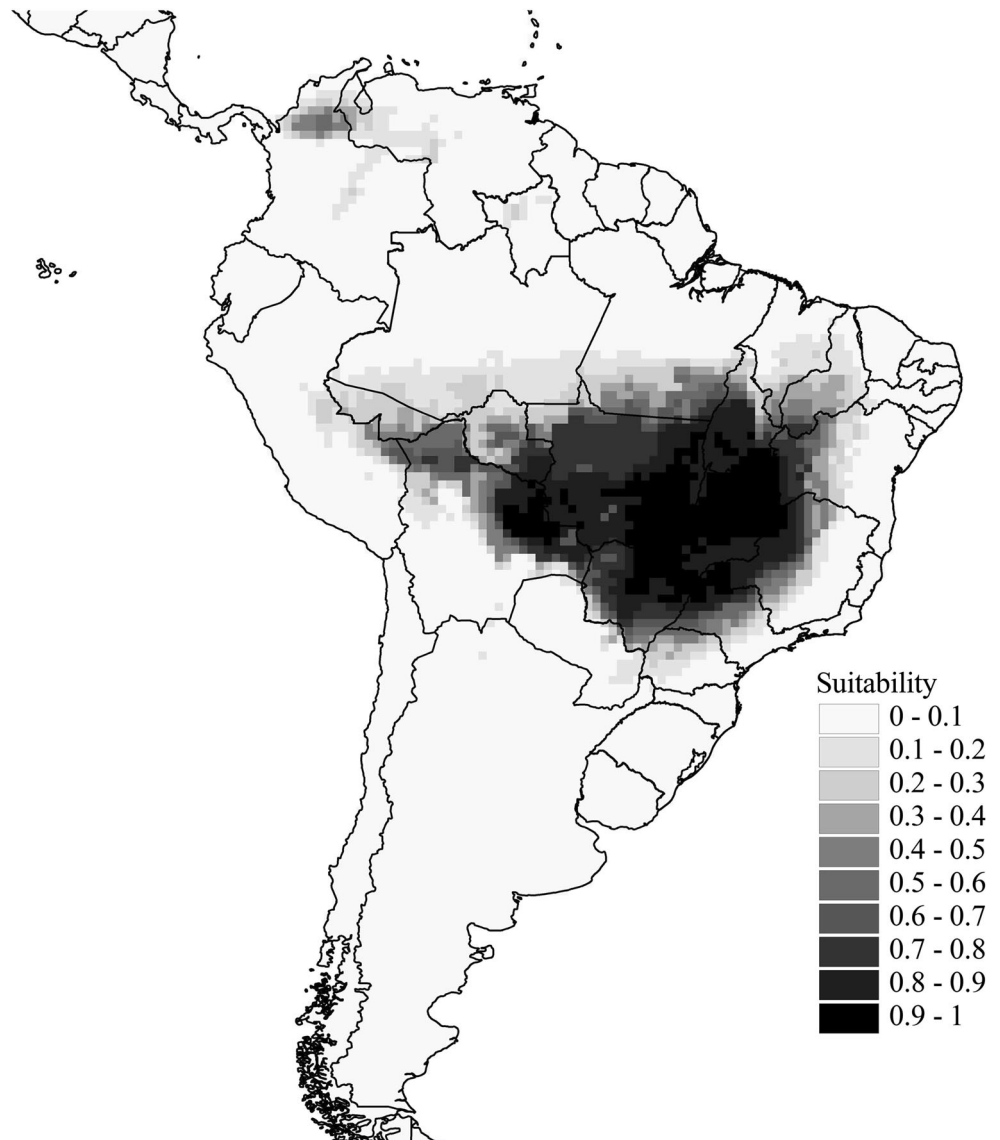
Seeds have a very woody endocarp with an edible nut that are eaten and dispersed by mammals, such as bats and monkeys, and are source of raw material for small- and middle-sized food industries, playing an important role in the local economy of Central Brazil (Sano et al. 2004). Previous analyses showed strong spatial genetic structure for this species, both among and within local populations (Soares et al. 2008; Collevatti et al. 2010, 2013a; Diniz-Filho et al. 2012).

Here we tested the applicability of the central-peripheral model to the within-population genetic variability in *Dipteryx alata*. More specifically, our goal is to evaluate whether genetic parameters in *D. alata* populations are distributed according to a central-peripheral model and to infer which factors, including ecological habitat suitability, human occupation and historical processes, better explain such patterns.

Materials and methods

Sampling populations

The collection of biological material for this study was conducted to include the largest possible area within the natural distribution of *D. alata*. Thus, we collected samples from 23 localities ("populations" in a broad sense hereafter) distributed throughout the most continuous area of the Brazilian Cerrado. Two populations used in previous studies (Diniz-Filho et al. 2012; Collevatti et al. 2013a) were not analyzed here because data for cpDNA microsatellites were not available) A total of 583 individual trees were genotyped for 8 microsatellite loci (see Soares et al. 2012 for methodological details), with sample sizes within each population ranging from 12 to 32. The location and geographical distances between populations were

determined using the program DIVA-GIS version 5.2 (Hijmans et al. 2005), from the geographical coordinates and altitude (m) by the Global Positioning System—GPS.

Genetic analysis

Genetic data were derived from eight genomic microsatellite loci, one of which (BM164) was transferred from *Phaseolus vulgaris* to *D. alata* (Garcia et al. 2011) and seven were developed specifically for *D. alata* (Soares et al. 2012). We also analyzed two chloroplastidial microsatellite loci (ccmp02 and ccmp10) (Weising and Gardner 1999), which were used to evaluate population differentiation using multivariate analysis of F_{ST} statistics (see below).

Amplifications were performed in 15 μL reaction volumes containing 15 ng of template DNA and 1 U *Taq* DNA polymerase (Phoneutria, Belo Horizonte, Brazil, with the following conditions: 94 °C for 5 min (one cycle); 94 °C for 1 min, variable annealing temperature for 1 min, 72 °C for 1 min (30 cycles); and 72 °C for 7 min (one cycle). Polymorphisms were detected in 6 % denaturing polyacrylamide gels stained with silver nitrate (Creste et al. 2001) and sized by comparison to a 10 bp DNA ladder standard (Invitrogen, Carlsbad, California, USA).

Genetic parameters within the 23 populations were estimated based on molecular nuclear microsatellite data and include the mean number of alleles per locus (*A*), the expected heterozygosity under the Hardy–Weinberg principle (He) and the intrapopulation fixation indices (*f*) for each population. However, these estimates may be biased due to low and unequal samples sizes among populations, so we also obtained estimates of these parameters using a rarefaction approach using the function subsampling of the R package hierfstat (Goudet 2005). These parameters were re-estimated 1000 times after resampling the smallest number of individuals from each local populations (i.e., 12) and average estimates for *A*, He and *f* were then used for further analyses.

Ecological niche modeling, human occupation and historical data

The genetic variability within populations, expressed by *A*, He and *f*, was used as genetic response variables and modeled as a function of three sets of explanatory variables: environmental suitability derived from Ecological Niche Models (ENMs, used here to generate a surrogate for species abundance and environmental suitability at populations), patterns of human occupation and historical differentiation of *D. alata* populations. A detailed description of each of these subsets is given below.

First, Ecological Niche Modeling (ENM) techniques were designed to correlate species occurrence with climate and environmental data, providing a description of ecological niches that can be translated into geographic space and thus approximate the range of species (under the assumption that this range is mainly driven by broad-scale climatic patterns) (see Rangel and Loyola 2012 for a recent review). Species occurrence data, obtained from scientific literature and from the Centro de Referência em Informação Ambiental database (CRIA, http://www.cria.org.br/), were used to model the potential geographic distribution of *D. alata*, using a total of 448 occurrence records throughout the Cerrado biome, which constituted the most complete database of *D. alata* occurrences (see Nabout et al. 2010; Terribile et al. 2012, for details). Five environmental variables were used here to model species occurrence: monthly simulation for precipitation, mean, maximum and minimum temperature, subsoil pH. These climatic variables were derived from five climatic models (CCSM, GISS, MIROC e MRI) and were converted to a 0.5 degree resolution covering the New World (see details about variable estimation and selection in Terribile et al. 2012).

Fourteen different ENMs methods were employed, including methods for presence-only and presence-absence data, which are integrated in the computational platform BIOENSEMBLES (see Rangel and Loyola 2012; Terribile et al. 2012; Collevatti et al. 2013b for details of modeling and description of techniques). Methods used were BIOCLIM, Ecological Niche Factor Analysis (ENFA), Euclidian Distance (ED), Gower Distances (GD), Mahalanobis Distances (MD), Genetic Algorithm for Rule Set Production (GARP), Generalized Linear Models (GLM), Maximum Entropy (Maxent), Generalized Additive Models (GAM), Flexible Discriminant Analysis (FDA), Multivariate Adaptive Regression Splines (MARS), Generalized Boosted Models (GBM), Neural Networks (ANN) and Random Forest (RNDFOR).

For each method and climatic model, models were built using a calibration subset of 75 % of the presence cells selected at random and then evaluated with the remaining 25 %, repeating this procedure 50 times. The models were then used to generate occurrence maps based on thresholds established by the ROC curve, eliminating those models with low True Skill Statistics (TSS <0.5) (see Diniz-Filho et al. 2009b). The ENMs then generated a total of 3500 geographic distributions for *D. alata* (i.e., 14 algorithms × 50 runs × 5 climate models), which were combined so that the frequencies of models that indicate presence of the species in a grid cell can be used as a general surrogate of environmental suitability (see Terribile et al. 2012, for details) (variable "ENM" hereafter). Thus, rather than using a single method selected based on

statistics, we used an ensemble approach that combines different possible models and, by overlaying them, allow a more consensual estimation of species geographic range and suitability variation (see Araújo and New 2007).

Human occupation is another important factor that may establish isolation among populations of *D. alata* and disturb demographic patterns at local scales (Manel et al. 2003; Holderegger and Wagner 2006; Diniz-Filho et al. 2009a). The Cerrado biome has been subject to intense anthropization, mainly due to the establishment of pastures and annual crops, some of which are currently being replaced by sugarcane plantations (Klink and Moreira 2002; Klink and Machado 2005; Sawyer 2008). Consequently, human occupation and habitat fragmentation indexes were paired with the geographic coordinates of each population of *D. alata* (see below) and used to evaluate their effects on observed patterns of genetic variation within local populations.

Human occupation for each population was evaluated by different variables, expressing different components of such occupation. First, explicit habitat fragmentation data were used, estimated by the proportion of remaining Cerrado natural vegetation in the area around each population (PA). Landscape data were derived from vegetation cover maps of the Brazilian biomes at the 1:250,000 spatial scales. We reclassified the maps into natural vegetation or anthropogenic land cover classes (this last classification includes agriculture, pasture lands, urban areas and mining regions), and then created 10 km-size buffers around the geographical position of each population (see details in Diniz-Filho et al. 2012; Telles et al. 2014). Further, two socio-economic variables (Human Population Density, HPD, and Human Development Index, HDI) were obtained for each municipality where populations of *D. alata* were collected (data from "Instituto Brasileiro de Geografia e Estatística—IBGE, for the year 2008).

Genetic divergence estimated from chloroplastidial DNA (cpDNA) was used to test the influence of historical relationship on the geographic pattern of intrapopulation genetic variability. Such a split between historic and current relationships could be made because chloroplastidial DNA display an uniparental heritage pattern and, in theory, suffers no crossover, so that it contains more preserved genetic variability as compared to the nuclear genome (Provan et al. 2001; Petit et al. 2005). To obtain an approximation of the relationships among the 23 populations based on cpDNA data that can be paired with the within-population genetic diversity given by A, He and f, we used the first eigenvector of the pairwise F_{ST} matrix obtained using the two cpDNA microsatellites among local populations (CP_DNA hereafter).

Modeling genetic variation within populations

The central-peripheral patterns of all genetic responses (A, He and f) across the 23 populations of *D. alata* were initially tested by Pearson correlations with geographic position variables (latitude and longitude) and distance from the center of the Cerrado biome (CENTRALITY hereafter). This may provide a fine purely geographical estimate of centrality of populations of *D. alata* analyzed here because the species is widely distributed but endemic to the biome. Next, variables were modeled using first (linear) and second (quadratic) order trend surface analysis (TSA; see Diniz-Filho et al. 2009c), which are multiple regressions in which the explanatory variables are the latitude and longitude (and their polynomial expansions) of the localities. Addition of quadratic terms was decided based on the gain in R^2 and Akaike information criterion (AIC). A better fit of a quadratic TSA model relative to the linear model indicates support of a central-peripheral pattern, better modeling a central peak of the response variable associated with decrease towards the borders.

For each genetic response we used model selection and multi-model inference approaches (Burnham and Anderson 2002) to compare the likelihood of models based on different set of predictors (composed by ecological, human occupation and historical variables). In model selection analysis, competing models were compared by evaluating the relative support of each model based on Akaike Information Criterion (AIC) (Burnham and Anderson 2002). AIC is based on the principle of parsimony (it maximizes goodness of fit and minimizes number of predictors), so there is a trade-off between prediction bias and parameter uncertainty.

Each of the genetic responses was regressed against the 63 possible combinations of the 6 explanatory variables in the three subsets (FST, ENM, PA, HPD, HDI and CENTRALITY), and the model with minimal AIC value was selected as the best. Also, multi-model inferences based on model averaging weighted by AIC relative weight (AIC w_i) were used to estimate the relative importance values of each explanatory variable (Johnson and Omland 2004).

All spatial regression analyses and multi-model inferences were performed in the SAM (Spatial Analysis in Macroecology) software, version 4.0 (Rangel et al. 2010), freely available at www.ecoevol.ufg.br/sam.

Results

Values for genetic parameters from nuclear microsatellite loci for the 23 populations of *D. alata* are shown in Table 1. According to the ENM distribution of for *D. alata*,

the largest overlap of the niche models was found mainly in the central area of the Cerrado biome (Fig. 1).

Genetic variables A and He displayed negative correlations with CENTRALITY (Table 2), indicating a trend towards lower number of alleles and lower diversity at the biome borders (Fig. 2). Negative correlations with latitude and longitude were also observed, indicating geographic trends in both north–south and east–west directions (Table 2; Fig. 2). The number of alleles A shows a better fit

for quadratic TSA, but He is better modeled by a linear fit (Table 2).

The model selection analyses indicated that the explanatory variables used were able to explain about 40 % of variation in the genetic responses (in full models or minimum models defined by AIC criterion) (Table 3). Human development index HDI was an important variable to explain the patterns in all three genetic variables, with positive relationships with all of them. As observed in simple correlation and TSA analyses, A is more related with CENTRALITY, whereas the fixation index f was negatively related to suitability from niche models. Higher values of f (i.e., more inbreeding) occurred in regions with the lower ENM (Table 3).

Table 1 Genetic parameters estimated by microsatellite loci in 23 *Dipteryx alata* populations (Pop), estimated by rarefaction approach, including number of alleles (A), expected heterozigosity (He) and fixation index (inbreeding, f)

Pop	Longitude	Latitude	A	He	f
CMT	−50.99	−14.4	2.502	0.319	0.226
ABMT	−52.04	−13.84	2.223	0.328	0.217
PGO	−49.03	−15.99	2.164	0.309	0.267
SMS	−54.7	−17.85	3.213	0.481	0.116
AMS	−53.93	−18.27	3.609	0.463	0.249
ATO	−49.12	−12.45	2.747	0.479	0.199
SMGO	−50.1	−13.23	2.409	0.324	−0.077
LGO	−48.13	−16.73	2.425	0.358	0.112
ISP	−49.22	−20.35	2.212	0.334	0.287
MAMG	−49.02	−18.98	2.582	0.445	0.118
ENGO	−49.14	−13.83	2.478	0.391	0.094
STGO	−49.63	−14.52	2.823	0.400	0.109
AMG	−46.27	−15.93	1.973	0.261	0.013
PMG	−45.17	−16.06	2.103	0.203	0.240
PMS	−52.98	−18.85	3.059	0.398	0.462
PCMS	−53.38	−19.33	2.808	0.329	0.462
CMS	−53.9	−19.53	2.861	0.408	0.004
IGO	−49.97	−17.16	2.447	0.303	0.134
RAMT	−52.19	−15.91	2.976	0.376	0.343
RAGO	−52.16	−15.95	2.865	0.384	−0.040
JGO	−50.2	−16.91	2.480	0.376	−0.066
NTO	−47.71	−11.66	2.093	0.291	−0.183
ARTO	−46.86	−12.99	2.205	0.320	−0.042

Discussion

To corroborate the classic central-peripheral population model (which is in turn based on the central-abundant model), it is necessary to show that as populations become more distant from the center of their geographical distribution, a progressive reduction in abundance and an increase in geographic isolation levels occur (Diniz-Filho et al. 2009a). However, no broad-scale abundance estimates exist for *D. alata* populations, so here we used niche modeling techniques to generate a surrogate for environmental suitability and, hopefully, abundance, despite discussions (see Nabout et al. 2011; Torres et al. 2012; Martinez-Meyer et al. 2012). In any case, this allowed evaluating whether peripheral populations in geographic space were indeed located in areas of lower habitat suitability, taking into account central-peripheral dynamics driven by environmental variation (Anderson and Martinez-Meyer 2004). Indeed, our analyses clearly showed the importance of suitability in explaining genetic variation within populations, at least for number of alleles A. As expected under central-periphery models, higher f values were found in regions with low environmental suitability (likely due to much reduced effective population size

Table 2 Analysis of spatial patterns in genetic parameters (A, He, and f), including the linear correlations with CENTRALITY (i.e., dist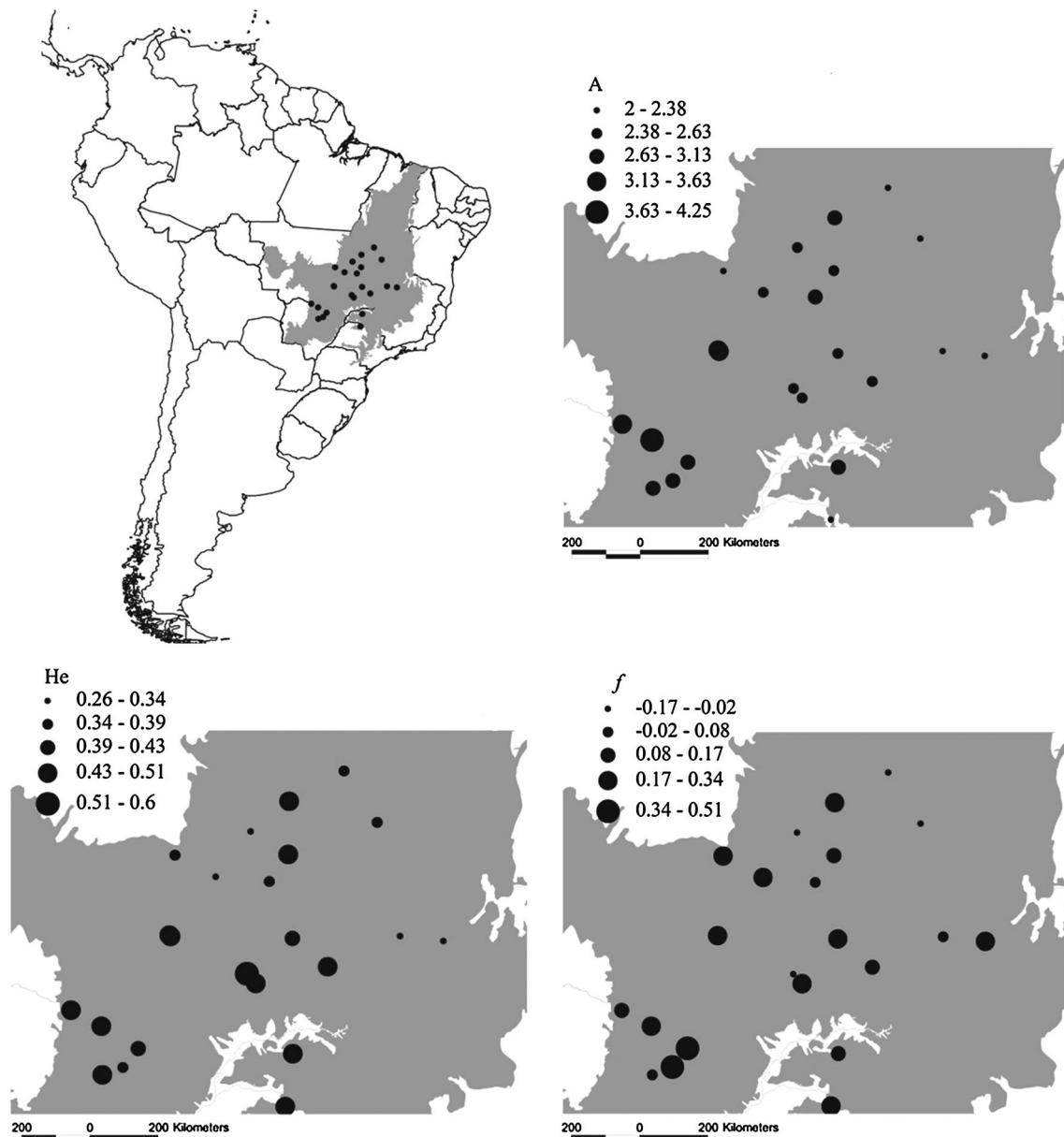ance from centre of Cerrado biome, r Biome), with latitude (r Latitude), longitude (r Longitude) and R^2, as well as Akaike Information Criterion (AIC) and ΔR^2 values from trend surface analysis (TSA) of linear and quadratic models

Genetic response	r (biome)	r (latitude)	r (longitude)	TSA				ΔR^2
				Linear		Quadratic		
				R^2	AIC	R^2	AIC	
A	**−0.656**	−0.387	**−0.808**	**0.654**	8.767	**0.697**	16.924	0.043
He	**−0.475**	−0.218	**−0.61**	**0.375**	−59.015	0.413	−49.196	0.038
f	−0.256	**−0.468**	−0.335	0.24	−13.747	0.273	−3.52	0.033

Significant values indicated in boldface ($P < 0.05$)

Fig. 2 Geographic pattern of genetic parameters estimated in 23 populations of *Dipteryx alata*. *A* average allele number, *He* expected heterozygosity *f* intrapopulation fixation indices. *Gray area* represents the Brazilian Cerrado

related to low suitability under niche models ENM, thus indirectly supporting the central-abundance model in environmental space). Coefficients for central-peripheral patterns in *f*, however, are not that high.

Genetic parameters estimated in this study have been frequently used for testing central-peripheral patterns (Garner et al. 2004; Howes and Lougheed 2008). Even though allelic richness and other estimates of genetic diversity estimates tend to be correlated, most studies use both because there were cases where they may not display the same patterns (Eckert et al. 2008). In general, allelic richness and inbreeding for *D. alata* fit the classic central-

peripheral model, with allelic richness being negatively correlated with distance from centre of the biome (lower richness at the border) and inbreeding high in regions of low suitability. These patterns result from both historical and current processes underlying them, and finding the best balance between them is not a trivial task. Our analyses, however, suggest stronger effects of current patterns related to current climate driving environmental suitability and recent human occupation.

In a historical context, some studies have demonstrated that the region of origin of a species is probably an area of higher allelic richness and genetic diversity, simply

Table 3 Standardized regression coefficients for variables explaining different estimates of genetic variability within populations of *Dipteryx alata* in Cerrado

Explanatory variables	*A*	He	*f*
CP_DNA	0.183	−0.05	0.157
HPD	−0.112	−0.14	0.014
HDI	**0.493**	**0.557**	**0.501**
CENTRALITY	**−0.380**	0.142	0.112
PA	−0.265	−0.2	−0.05
ENM	0.133	0.237	**−0.44**
R^2 full	0.452	0.434	0.249
R^2 (best model)	0.408	0.339	0.405

These variables include first eigenvector of F_{ST} based on cpDNA data (CP_DNA), Human Population Density (HPD), Human Development Index (HDI), distance from each population to the center of the biome (CENTRALITY), proportion of natural vegetation remnants (PA) and suitability from Ecological Niche Models (ENM)

Coefficients in bold expressed the variables in the best model for each response according to AIC criterion

because of more time for accumulation of mutations (Harris et al. 2002; Barnett et al. 2006). The contribution of colonization history to the genetic diversity of current populations can be evaluated in several distinct ways, including: (1) genetic diversity regression or correlation to geographic distance from the origin center (or refuge) (Garner et al. 2004) and; (2) genetic diversity regression or correlation with latitude (related to the colonization process from refuge zones) (Johansson et al. 2006; Van Rossum and Prentice 2004). For this study, it was possible to use the correlation of genetic variables with the longitude and latitude to explain the clinal patterns in He that, in turn, may be related to range expansion of *D. alata* after the LGM. Previous analyses suggest that the southwestern region is the origin center of *D. alata* (considering the genetic parameters) and that regions closer to the center of the biome have been recently (re)colonized, and therefore tend to display lower values of genetic diversity (Collevatti et al. 2013a). Thus, we expected that the north and east regions of the biome (or geographic distribution of a given species) are newer and therefore have lower values of He and *A* (and higher *f*). In fact, the results in Table 2 indicated the estimated values decline toward the north and east, at least for He. However, this hypothesis is not strongly corroborated here because of the lack of the correlation observed between "genealogic or phylogenetic" relationships among populations (first eigenvector of F_{ST} based on cpDNA data) and latitude, although this lack of support may be more likely due to problems in detecting historical signal with this approach (and not using explicit phylogeographical patterns, that are up to this time not available for *D. alata*).

In any case, current patterns of species' geographic occupation areas and genetic attributes of populations of Cerrado plants are strongly influenced by climatic variables such as temperature, humidity, fire regimes and soil drainage (Oliveira-Filho and Ratter 2002). In addition, climatic changes occurring throughout history are considered important factors that have changed the species geographic distributions (Bonaccorso et al. 2006; Nabout et al. 2011) and the genetic characteristics of populations (Hopper and Gioia 2004; Ramos et al. 2007). In species from temperate climates in the Northern hemisphere, historic geological phenomena such as glacial periods caused reductions in species' geographical ranges, creating refuges (occurrence spots and future distribution centers). Therefore, the genetic diversity of peripheral populations may be explained as resulting from recolonization of formerly glacial areas (Hewitt 1999), as suggested above. However, here we did not observe such a strong historical signal in central-periphery patterns in *D. alata*, most likely because human effects seem to be already strong and drive genetic diversity (whereas *A* and *f* seems to be related to current patterns of environmental suitability).

Moreover, the relationship found between estimated He values and human socio-economic variables (particularly HDI) does not clearly indicate that the populations studied may already be suffering the consequences of habitat fragmentation processes and changes in land use, probably caused by the recent occupation progress in the Cerrado (Klink and Moreira 2002; Klink and Machado 2005; Telles et al. 2014). The positive correlation with all genetic responses most likely reflect a coincidence in patterns of human occupation, in which higher levels of human occupation are found in the centre and southern part of the biome. Even so, the positive relationship between HDI and fixation index may indicate some effects (and also consider the partial effects of environmental suitability from niche model). An example of how even very recent human occupation may influence species genetic patterns was discussed by Telles et al. (2007). More specifically for *D. alata,* Telles et al. (2014) found correlations between proportion of natural remnants and population divergence, using a more detailed analysis of local populations. However, it is still important to consider that *D. alata* is frequently preserved in pasture areas, suggesting that these areas are important for the preservation of the species (Soares et al. 2008). Notice also that although human socio-economic variables are important, metrics for habitat fragmentation were not directly important here, perhaps due to mismatches in scale (grain) used or because these effects are not detectable at local scales but rather reflect more general patterns of human occupation, as estimated by the correlations with HDI and HPD.

The use of niche modeling techniques, human occupation process indicators and inference over historical processes may help to outline a possible multi-effect scenario based on different factors affecting genetic parameters throughout the spatial distribution of local *D. alata* populations. Thus, the high genetic diversity values found in the southwestern region of the Cerrado biome, coupled with the spatially structured patterns of genetic variability from that region, suggest that historical factors related to range expansion after the last Glacial Maximum may be important to understand the patterns observed here, although more detailed analyses are necessary. Our analyses, on the other hand, show a much strong effect related to contemporary phenomena, including human occupation (shifting genetic diversity) and environmental patterns (driving patterns of inbreeding, potentially due to small populations sizes in regions of low suitability). These two factors are clearly important for conservation actions, by considering the current high rates of disturbance in the Cerrado region and the global climate change patterns.

References

Anderson RP, Martinez-Meyer E (2004) Modeling species' geographic distributions for preliminary conservation assessments: an implementation with the spiny pocket mice (Heteromys) of Ecuador. Biol Conservation 116:167–179

Araújo MB, New M (2007) Ensemble forecasting of species distributions. Trends Ecol Evol 22:42–47

Barnett R, Yamaguchi N, Barnes I, Cooper A (2006) The origin, current diversity and future conservation of the modern lion (*Panthera leo*). Proc Royal Soc B Biol Sci 273:2119–2125

Bonaccorso E, Koch I, Peterson AT (2006) Pleistocene fragmentation of Amazon species' ranges. Diversity Distrib 12:157–164

Burnham KP, Anderson DR (2002) Model selection and multi-model inference: a practical information-theoretic approach. Springer, Heidelberg

Collevatti RG, Lima JS, Soares TN, Telles MPC (2010) Spatial genetic structure and life history traits in Cerrado tree species: inferences for conservation. Nat Conservation 8:54–59

Collevatti RG, Telles MPC, Nabout JC, Chaves LJ, Soares TN (2013a) Demographic history and the low genetic diversity in *Dipteryx alata* (Fabaceae) from Brazilian Neotropical savannas. Heredity 111:97–105

Collevatti RG, Terribile LC, Oliveira G, Lima-Ribeiro MS, Nabout JC, Rangel TF, Diniz-Filho JAF (2013b) Drawbacks in paleo-distribution modelling: the case of South American seasonally dry forests. J Biogeogr 40:345–358

Creste C, Tulmann Neto A, Figueira A (2001) Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. Pl Molec Biol Rep 19:299–306

Diniz-Filho JAF, Nabout JC, Bini LM, Soares TN, Telles MPC, De Marco P Jr, Collevatti RG (2009a) Niche modelling and landscape genetics of *Caryocar brasiliense* ("Pequi" tree: Caryocaraceae) in Brazilian Cerrado: an integrative approach for evaluating central–peripheral population patterns. Tree Genet Genomes 5:617–627

Diniz-Filho JAF, Bini LM, Rangel TF, Loyola RD, Hof C, Nogues-Bravo D, Araújo MB (2009b) Partitioning and mapping uncertainties in ensembles of forecasts of species turnover under climate change. Ecography 32:897–906

Diniz-Filho JAF, Nabout JC, Telles MPC, Soares TN, Rangel TFLVB (2009c) A review of techniques for spatial modeling in geographical, conservation and landscape genetics. Genet Molec Biol 32:203–211

Diniz-Filho JAF, Melo DB, Oliveira G, Collevatti RG, Soares TN, Nabout JC, Lima JS, Dobrowolski R, Chaves LJ, Naves RV, Telles MPC (2012) Planning for optimal conservation of geographical genetic variability within species. Conservation Genet 13:1085–1093

Eckert CG, Samis KE, Lougheed SC (2008) Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. Molec Ecol 17:1170–1188

Epperson BK (2003) Geographical Genetics. Princeton University Press, Princeton

Garcia RAV, Rangel PN, Brondani C, Martins WS, Melo LC, Carneiro MS, Borba TCO, Brondani RPV (2011) The characterization of a new set of EST—derived simple sequence repeat (SSR) markers as a resource for the genetic analysis of *Phaseolus vulgaris*. BMC Genet 12(41):1–14

Garner TWJ, Pearman PB, Angelone S (2004) Genetic diversity across a vertebrate species' range: a test of the central-peripheral hypothesis. Molec Ecol 13:1047–1053

Gaston KJ (2003) The structure and dynamics of geographic ranges. Oxford University Press, Oxford

Gaston KJ, Blackburn TM (2000) Pattern and process in macroecology. Blackwell, London

Ginzburg LR, Golenberg EM (1985) Lectures in theoretical population biology. Prentice-Hall, Englewood Cliffs

Goudet J (2005) Hierfstat, a package for R to compute and test variance components and F-statistics. Molec Ecol Notes 5:184–186

Harris SA, Robinson JP, Juniper BE (2002) Genetic clues to the origin of the apple. Trends Genet 18:426–430

Hewitt GM (1999) Post-glacial re-colonization of European biota. Biol J Linn Soc 68:87–112

Hijmans RJ, Guarino L, Bussink C, Barrantes I, Rojas E (2005) DIVA-GIS, Version 5.2. http://www.diva-gis.org. Accessed 8 Apr 2014

Holderegger R, Wagner HH (2006) A brief guide to landscape genetics. Landscape Ecol 21:793–796

Hopper SD, Gioia P (2004) The Southwest Australian floristic region: evolution and conservation of a global hot spot of biodiversity. Annual Rev Ecol Evol Sys 35:623–650

Howes BJ, Lougheed SC (2008) Genetic diversity across the range of a temperate lizard. J Biogeogr 35:1269–1272

Johansson M, Primmer CR, Merila J (2006) History vs. current demography: explaining the genetic population structure of the common frog (*Ranatemporaria*). Molec Ecol 15:975–983

Johnson JP, Omland KS (2004) Model selection in ecology and evolution. Trends Ecol Evol 19:101–108

Kirkpatrick M, Barton NH (1997) Evolution of a species' range. Amer Nat 150:1–23

Klink CA, Machado RB (2005) Conservation of the Brazilian Cerrado. Conservation Biol 19:707–713

Klink CA, Moreira AG (2002) Past and current human occupation, and land use. In: Oliveira PS, Marquis RJ (eds) The Cerrados of Brazil. Columbia University Press, New York, pp 69–88

Lira-Noriega A, Manthey JD (2014) Relationship of genetic diversity and niche centrality: a survey and analysis. Evolution 68:1082–1093

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. Trends Ecol Evol 18:189–197

Martinez-Meyer E, Diáz-Porras D, Peterson AT, Yánez-Arenas (2012) Ecological niche structure and rangewide abundance patterns of species. Biol Lett 9:20120637

Maurer BA (1994) Geographical population analysis: tools for the analysis of biodiversity. Blackwell Scientific Publications, Oxford

Nabout JC, Soares TN, Diniz-Filho JAF, De Marco Júnior P, Telles MPC, Naves RV, Chaves LJ (2010) Combining multiple models to predict the geographical distribution of the Baru tree (*Dipteryx alata Vogel*) in the Brazilian Cerrado. Brazil J Biol 70(4):911–919

Nabout JC, Oliveira G, Magalhães G, Magalhães MR, Terribile LC (2011) Global climate change and the production of Pequi fruits (Caryocar brasiliense) in the Brazilian Cerrado. Nat Conservation 9:55–60

Oliveira MIB, Sigrist MR (2008) Fenologia reprodutiva, polinização e reprodução de Dipteryx alata Vog. (Leguminosae-Papilionoidae) em Mato Grosso do Sul. Revista Brasil Bot 31:195–207

Oliveira-Filho AT, Ratter JA (2002) Vegetation physiognomies and woddy flora of the cerrado Biome. In: Oliveira PS, Marquis RJ (eds) The Cerrados of Brazil. Columbia University Press, New York, pp 91–120

Petit RJ, Pineau E, Demesure B, Bacilieri R, Ducousso A, Kremer A (1997) Chloroplast DNA footprints of postglacial recolonization by oaks. Proc Natl Acad Sci USA 94:9996–10001

Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, Vendramin GG (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. Molec Ecol 14:689–701

Pfeifer M, Schatz B, Xavier PF, Passalacqua NG, Fay MF, Carey PD, Jeltsch F (2009) Phylogeography and genetic structure of the orchid *Himantoglossum hircinum* (L.) Spreng. across its European central–marginal gradient. J Biogeogr 36:2353–2365

Provan J, Powell W, Hollingsworth PM (2001) Chloroplast microsatellites: new tools for studies in plant ecology and evolution. Trends Ecol Evol 16:142–147

Ramos ACS, Lemos-Filho JP, Ribeiro RA, Santos FR, Lovato MB (2007) Phylogeography of the tree *Hymenaea stigonocarpa* (Fabaceae: Caesalpinioideae) and the influence of quaternary climate changes in the Brazilian Cerrado. Ann Bot (Oxford) 100:1219–1228

Rangel TF, Loyola RD (2012) Labeling ecological niche models. Nat Conservation 10:119–126

Rangel TF, Diniz-Filho JAF, Bini LM (2010) SAM: a comprehensive application for spatial analysis in macroecology. Ecography 33:46–50

Sagarin RD, Gaines SD (2002) The 'abundant centre' distributions: to what extent is it a biogeographical rule? Ecol Lett 5:137–147

Sagarin RD, Gaines SD, Gaylord B (2006) Moving beyond assumptions to understand abundance distributions across the range of species. Trends Ecol Evol 21:524–530

Sano SM, Ribeiro JF, Brito MA (2004) Baru: Biologia e uso. Embrapa, Planaltina

Sawyer D (2008) Climate change, biofuels and eco-social impacts in the Brazilian Amazon and Cerrado. Philos Trans Ser B 363:1747–1752

Soares TN, Chaves LJ, Telles MPC, Diniz-Filho JAF, Resende LV (2008) Landscape conservation genetics of *Dipteryx alata* ("baru" tree: Fabaceae) from Cerrado region of central Brazil. Genetica 132:9–19

Soares TN, Melo DB, Resende LV, Vianello RP, Chaves LJ, Collevatti RG, Telles MPC (2012) Development of microsatellite markers for the neotropical tree species *Dipteryx alata* (Fabaceae). Amer J Bot 99:e72–e73

Soulé M (1973) The epistasis cycle: a theory of marginal populations. Annual Rev Ecol Syst 4:165–187

Telles MPC, Diniz-Filho JAF, Bastos RP, Soares TN, Guimarães LD, Lima LP (2007) Landscape genetics of *Physalaemus cuvieri* in Brazilian Cerrado: correspondence between population structure and patterns of human occupation and habitat loss. Biol Conservation 139:37–46

Telles MPC, Dobrovolski R, Souza KS, Lima JS, Collevatti RG, Soares TN, Chaves LJ, Diniz-Filho JAF (2014) Disentangling landscape effects on population genetic structure of a Neotropical savana tree. Nat Conservation 12:65–70

Terribile LC, Lima-Ribeiro MS, Araújo M, Bizao N, Collevatti RG, Dobrovolski R, Franco A, Guilhaumon F, Lima JS, Murakami DM, Nabout JC, Oliveira G, Oliveira LK, Rabello SG, Rangel TF, Simon LM, Soares TN, Telles MPC, Diniz-Filho JAF (2012) Areas of climate stability of species ranges in the Brazilian Cerrado: disentangling uncertainties through time. Nat Conservation 10:152–159

Torres NM, De Marco Jr P, Santos T, Silveira L, Jácomo AT, Diniz-Filho JAF (2012) Can species distribution modelling provide estimates of population densities? A case study with jaguars in the Neotropics. Diver Distrib 18:615–627

Van Rossum F, Prentice HC (2004) Structure of allozyme variation in Nordic *Silene nutans* (Caryophyllaceae): population size, geographical position and immigration history. Biol J Lin Soc 81:357–371

Vucetich JA, Waite TA (2003) Spatial patterns of demography and genetic processes across the species' range: null hypotheses for landscape conservation genetics. Conservation Genet 4:639–645

Weising K, Gardner RC (1999) A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genome of dicotyledonous angiosperms. Genome 42:9–19