



Landscape Genomic Conservation Assessment of a Narrow-Endemic and a Widespread Morning Glory From Amazonian Savannas

Éder C. Lanes¹, Nathaniel S. Pope², Ronnie Alves¹, Nelson M. Carvalho Filho¹, Tereza C. Giannini¹, Ana M. Giulietti¹, Vera L. Imperatriz-Fonseca¹, Waléria Monteiro¹, Guilherme Oliveira¹, Amanda R. Silva^{1,3}, José O. Siqueira¹, Pedro W. Souza-Filho^{1,4}, Santelmo Vasconcelos¹ and Rodolfo Jaffé^{1,5,6*}

OPEN ACCESS

Edited by:

Stéphane Joost,
École Polytechnique Fédérale de
Lausanne, Switzerland

Reviewed by:

S. Hoban,
Morton Arboretum, United States
Brenna R. Forester,
Colorado State University,
United States
Kevin Leempoel,
Stanford University, United States

*Correspondence:

Rodolfo Jaffé
rodolfo.jaffe@itv.org

Specialty section:

This article was submitted to
Evolutionary and Population Genetics,
a section of the journal
Frontiers in Plant Science

Received: 20 December 2017

Accepted: 05 April 2018

Published: 07 May 2018

Citation:

Lanes EC, Pope NS, Alves R,
Carvalho Filho NM, Giannini TC,
Giulietti AM, Imperatriz-Fonseca VL,
Monteiro W, Oliveira G, Silva AR,
Siqueira JO, Souza-Filho PW,
Vasconcelos S and Jaffé R (2018)
Landscape Genomic Conservation
Assessment of a Narrow-Endemic
and a Widespread Morning Glory
From Amazonian Savannas.
Front. Plant Sci. 9:532.
doi: 10.3389/fpls.2018.00532

¹ Instituto Tecnológico Vale, Belém, Brazil, ² Biological Laboratories, Department of Integrative Biology, University of Texas, Austin, TX, United States, ³ Ciências Biológicas-Botânica Tropical, Universidade Federal Rural da Amazônia/Museu Paraense Emílio Goeldi, Belém, Brazil, ⁴ Instituto de Geociências, Universidade Federal do Pará, Belém, Brazil, ⁵ Departamento de Ecologia, Universidade de São Paulo, São Paulo, Brazil, ⁶ Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém, Brazil

Although genetic diversity ultimately determines the ability of organisms to adapt to environmental changes, conservation assessments like the widely used International Union for Conservation of Nature (IUCN) Red List Criteria do not explicitly consider genetic information. Including a genetic dimension into the IUCN Red List Criteria would greatly enhance conservation efforts, because the demographic parameters traditionally considered are poor predictors of the evolutionary resilience of natural populations to global change. Here we perform the first genomic assessment of genetic diversity, gene flow, and patterns of local adaptation in tropical plant species belonging to different IUCN Red List Categories. Employing RAD-sequencing we identified tens of thousands of single-nucleotide polymorphisms in an endangered narrow-endemic and a least concern widespread morning glory (Convolvulaceae) from Amazonian savannas, a highly threatened and under-protected tropical ecosystem. Our results reveal greater genetic diversity and less spatial genetic structure in the endangered species. Whereas terrain roughness affected gene flow in both species, forested and mining areas were found to hinder gene flow in the endangered plant. Finally we implemented environmental association tests and genome scans for selection, and identified a higher proportion of candidate adaptive loci in the widespread species. These mainly contained genes related to pathogen resistance and physiological adaptations to life in nutrient-limited environments. Our study emphasizes that IUCN Red List Criteria do not always prioritize species with low genetic diversity or whose genetic variation is being affected by habitat loss and fragmentation, and calls for the inclusion of genetic information into conservation assessments. More generally, our study exemplifies how landscape genomic tools can be employed to assess the status, threats and adaptive responses of imperiled biodiversity.

Keywords: biodiversity conservation, environmental association test, *Ipomoea*, isolation by resistance, IUCN red list, landscape genetics, RAD-sequencing, SNP

INTRODUCTION

Species conservation assessments have largely neglected measures of genetic diversity, even though genetic diversity ultimately determines the ability to adapt to environmental changes and therefore underpins species long-term persistence (Reed and Frankham, 2003; Allendorf et al., 2012; Jamieson and Allendorf, 2012). For instance, the Red List criteria from the International Union for Conservation of Nature (IUCN), a widely used system to determine conservation priorities across the globe (Rodrigues et al., 2006; IUCN, 2017), does not explicitly consider genetic diversity, emphasizing instead population size, population declines and geographic range (Mace et al., 2008; IUCN, 2012). While the correlation between occurrence range and genetic diversity has been well documented in plants (Hamrick et al., 1991; Gitzendanner and Soltis, 2000; Cole, 2003; Leimu et al., 2006; Gibson et al., 2008), census size and population declines are not consistent predictors of genetic diversity. For instance, reductions in population size or geographic range may not result in diminished genetic diversity (Frankham, 1995; Honnay and Jacquemyn, 2007; Aguilar et al., 2008; Vranckx et al., 2012).

Two recent studies evaluated genetic diversity patterns across IUCN Red List Categories. Rivers et al. (2014) calculated the loss of genetic diversity under simulated range loss for two plant species in Madagascar. They found a strong relationship between loss of genetic diversity (AFLP allelic richness) and range, and a correspondence between levels of genetic diversity and thresholds for “non-threatened” vs. “threatened” IUCN Red List Categories. However, genetic diversity distinction between threatened categories was less evident, as the amount of genetic diversity at a specific threat category was not consistent across criteria (Rivers et al., 2014). Similarly, Willoughby et al. (2015) gathered 1941 microsatellite datasets from wild populations of five vertebrate classes. They then compared empirical estimates of heterozygosity and allelic richness between threatened and non-threatened species, and found that genetic diversity is generally reduced in threatened species. However, when they considered which of the IUCN Red List Criteria (including declining population size, species range extent, and the number of mature individuals) are most effective at identifying genetically depauperate species, they found that the existing criteria failed to systematically identify populations with low genetic diversity (Willoughby et al., 2015). Thus, threatened species generally show reduced genetic diversity compared to non-threatened ones (Spielman et al., 2004), but the IUCN Red List Criteria approach does not systematically prioritize the conservation of species with low genetic diversity (Rivers et al., 2014; Willoughby et al., 2015). Additionally, the genetic consequences of habitat degradation and fragmentation in relation to the IUCN Red List Categories remain largely unexplored.

As DNA sequencing and bioinformatic processing have become more affordable, genetic data have rapidly accumulated across a wide range of taxa (Vranckx et al., 2012; Frankham, 2015; Garner et al., 2017), many of which have never been assessed using the IUCN Red List Criteria. It is thus essential to integrate such data with the IUCN Red List Criteria in order to gain

deeper insights into the susceptibility and survival perspectives of imperiled biodiversity and make optimal conservation decisions (Laikre, 2010; Shafer et al., 2015; Pierson et al., 2016). However, no conservation assessment has yet compared genetic diversity and gene flow patterns between species from different IUCN Red List Categories (Rivers et al., 2014; Willoughby et al., 2015). Moreover, genomic tools have never been employed to validate this categorization.

Here we address these important knowledge gaps by performing the first genomic assessment, to our knowledge, of genetic diversity, gene flow, and patterns of local adaptation in tropical plants belonging to different IUCN Red List Categories. We chose two congeneric and sympatric morning glories (Convolvulaceae) from Amazonian savannas, one of the least known, highly threatened and under-protected tropical ecosystems (de Carvalho and Mustin, 2017). Tropical grassland mosaics associated with rocky outcrops are known to be worldwide hotspots of endemism, and the ancient inselbergs immersed in the Amazon forest are among the most iconic representatives of this ecoregion in South America. The banded iron formations known as Cangas, notorious for their extremely nutrient-poor, metal rich, flooding-prone soils, have also attracted substantial attention from mining companies given that they harbor one of the world’s largest deposits of high-grade iron ore (Skiryicz et al., 2014; Silveira et al., 2015; Viana et al., 2016).

Our model species were *Ipomoea cavalcantei*, a narrowly-distributed plant classified as *endangered*, and *I. maurandioides*, a species of *least concern* extending from Argentina, Paraguay and Bolivia throughout Brazil (Martinelli and Moraes, 2013; Simão-Bianchini et al., 2016; Wood and Scotland, 2017) (until recently this species was considered *I. carajasensis*, see Data Sheet 1). Although the pollination biology of our study species is yet to be systematically assessed, floral morphology (Simão-Bianchini et al., 2016) suggests that the small lilac flowers of *I. maurandioides* are pollinated by insects while hummingbirds pollinate the red flowers with exerted stamens of *I. cavalcantei* (Galletto and Bernardello, 2004). Fruit morphology, on the other hand, suggests wind-mediated seed dispersal, as in another congeneric species occurring in a neighboring region (Griz and Machado, 2001; Simão-Bianchini et al., 2016). The narrow-endemic *I. cavalcantei* is a flagship species for tropical conservation since its unique habitat is jeopardized by some of the world’s biggest mining projects (Souza-Filho et al., 2016; Viana et al., 2016), which involve deforestation of montane savannas, soil removal, and other types of environmental disturbances. The close phylogenetic proximity and contrasting distribution ranges of these two species offer an ideal opportunity to assess genetic diversity and the consequences of habitat loss and fragmentation on plants sharing a similar evolutionary history but belonging to different IUCN Red List Categories. Moreover, the use of thousands of genetic markers broadly distributed across the entire genome of our study species, provides high resolution measures of genetic diversity and robust estimates of gene flow and signatures of adaptation.

We hypothesized that extent of occurrence (IUCN Red List Criterion B, see Data Sheet 1) would determine genetic

diversity, gene flow, and susceptibility to habitat loss and fragmentation, thereby influencing patterns of local adaptation. We thus formulated the following predictions: (i) With a narrow distribution range *I. cavalcantei* is likely to show lower genetic diversity and higher spatial genetic structure across different montane savanna highlands than the analyzed populations of the broadly distributed *I. maurandioides*; (ii) Gene flow is probably affected by the structure of natural landscapes in both species, but landscape modifications by mining are more likely to hinder gene flow in the narrowly-distributed *I. cavalcantei*. Additionally, landscape effects on gene flow are expected to show time-lags determined by the generation time of each species (~20 years for *I. cavalcantei* and ~3 years for *I. maurandioides*); (iii) Although similar adaptations to the extreme montane savanna environments are expected in both species, genetic drift is likely stronger in the narrowly-distributed species, thus slowing adaptive responses. We therefore expect to find weaker adaptive signals (a lower proportion of candidate loci) in *I. cavalcantei*.

MATERIALS AND METHODS

Sampling, DNA Extraction and Genome Size Estimation

We sampled individual plants distributed across our study region, which comprised the full natural occurrence range of *I. cavalcantei*, and all major montane savanna highlands of the Carajás mineral province where *I. maurandioides* is found (Simão-Bianchini et al., 2016; **Figure 1**). We note that our study region is located at the northern edge of *I. maurandioides* large distribution range (see Data Sheet 1). Leaf tissue samples of 122 individuals of *I. cavalcantei* and 254 individuals of *I. maurandioides* were collected along trails in 2016 (SISBIO collection permit N. 48272-4), observing a minimum distance of 10 m between samples to avoid re-sampling the same individuals (sample coordinates are provided in Data Sheet 2). Samples were preserved in 10 mL of a NaCl-saturated solution of 2% CTAB (Rogstad, 1992), and later frozen at -80°C . Total DNA was extracted using the DNeasy Plant Mini Kit (Qiagen, EUA), DNA concentration quantified using the Qubit High Sensitivity Assay kit (Invitrogen), and DNA integrity assessed through 1.2% agarose gel electrophoresis. All DNA samples were adjusted to a final concentration of 5 ng/ μL in a final volume of 30 μL . We used flow cytometry to estimate genome size in *I. maurandioides*, as it was already available for *I. cavalcantei* (Babiychuk et al., 2017). Nuclei were obtained from fresh leaf tissues chopped in general purpose buffer with 1% Triton X-100 and 1% PVP-30 (Loureiro et al., 2007). All the steps were conducted on ice until events acquisition on a PI fluorescence mean under a 585/42 bandpass filter. Triplicates of 1000 PI stained nuclei were finally analyzed on a BD FACS Aria II flow cytometer, using a 488 nm laser and *Petroselinum crispum* as an internal reference standard ($2C = 4.50$ pg) (Obermayer et al., 2002).

RAD Sequencing and SNP Discovery

DNA samples were shipped to SNPSaurus (<http://snpsaurus.com/>) for sequencing and bioinformatic analyses. Briefly, nextRAD genotyping-by-sequencing libraries were prepared

(Russello et al., 2015), considering genome sizes (both species showed a similar genome size, Table S1 and Figure S1). Genomic DNA was first fragmented with Nextera reagent (Illumina, Inc), which also ligates short adapter sequences to the ends of the fragments. The Nextera reaction was scaled for fragmenting 10 ng of genomic DNA, although 15 ng of genomic DNA was used for input to compensate for the amount of degraded DNA in the samples and to increase fragment sizes. Fragmented DNA was then amplified, with one of the primers matching the adapter and extending 9 nucleotides into the genomic DNA with the selective sequence GTGTAGAGC (only fragments starting with a sequence that could be hybridized by the selective sequence of the primer were efficiently amplified). The nextRAD libraries were then sequenced on an Illumina HiSeq 4000 (University of Oregon). Each species was sequenced on three lanes with 150 bp reads.

Reads were trimmed using *BBMap tools* (<http://sourceforge.net/projects/bbmap/>; `ash bbmap/bbduk.sh in=$file out=$outfile ktrim=r k=17 hdist=1 mink=8 ref=bbmap/resources/nextera.fa.gz minlen=100 ow=t qtrim=r trimq=10`) and a *de novo* reference was created by collecting 10 million reads evenly from the samples, excluding reads that had counts fewer than 10 or more than 400. The remaining loci were then aligned to each other to identify alleles. All reads were mapped to the reference with an alignment identity threshold of 95% using *BBMap tools* (`bbmap`). Genotype calling was done using *Samtools* and *bcftools* (<https://github.com/samtools/samtools>; `samtools mpileup -gu -Q 10 -t DP,DPR -f ref.fasta -b samples.txt | bcftools call -cv - >genotypes.vcf`), and the resulting set of genotypes was filtered to remove loci with a minor allele frequency of $<2.5\%$. Loci with more than 30% missing genotypes, or that were heterozygous in all samples or contained more than 2 alleles in a sample (suggesting collapsed paralogs) were removed. The absence of artifacts was checked by counting SNPs at each read nucleotide position and determining that SNP number did not increase with reduced base quality at the end of the read. This pipeline was carried out separately for each species, with a *de novo* reference created separately for each (genotype and sequence data are available in FigShare: <https://doi.org/10.6084/m9.figshare.6100004.v1>).

Genetic Diversity and Population Structure

To address our first prediction we estimated genetic diversity and assessed spatial genetic structure. The R package *r2vcftools* (<https://github.com/nspope/r2vcftools>) - a wrapper for *VCFTools* (Danecek et al., 2011) - was used to perform final quality control on the genotype data (R version 3.4.3). We filtered loci for quality (Phred score >50), read depth (20 - 200), linkage disequilibrium (LD, $r^2 < 0.5$), and strong deviations from the Hardy Weinberg Equilibrium (HWE, $p < 0.0001$). Additionally, we removed any potential loci under selection detected through environmental association and F_{ST} outlier tests (see details below). The resulting sets of neutral and independent loci were then used to calculate genetic relatedness (Yang et al., 2010), expected heterozygosity (H_E), inbreeding coefficients (F), mean per-site nucleotide diversity (π) and Tajima's D (D). Effective population size (N_e) was also estimated employing

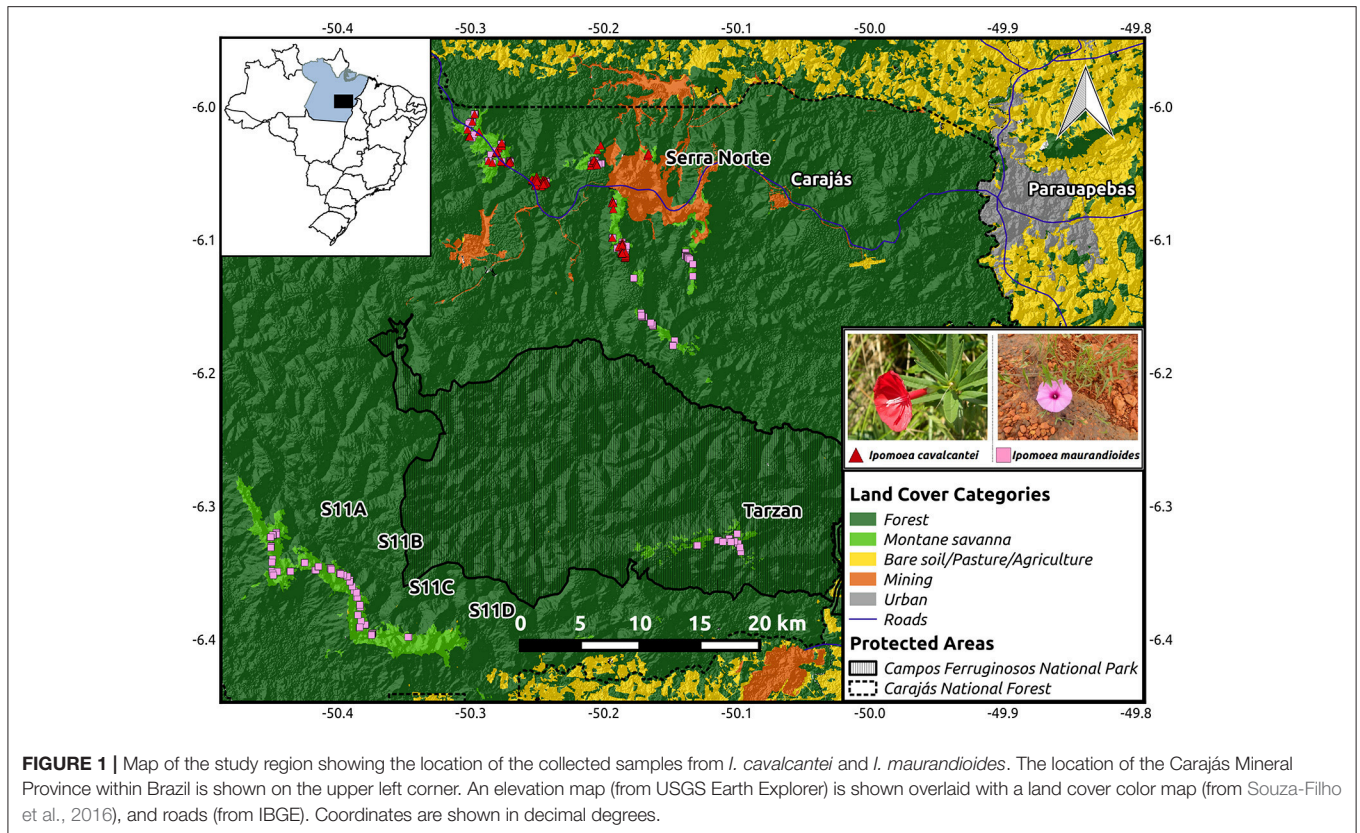


FIGURE 1 | Map of the study region showing the location of the collected samples from *I. cavalcantei* and *I. maurandioides*. The location of the Carajás Mineral Province within Brazil is shown on the upper left corner. An elevation map (from USGS Earth Explorer) is shown overlaid with a land cover color map (from Souza-Filho et al., 2016), and roads (from IBGE). Coordinates are shown in decimal degrees.

the linkage disequilibrium approach of NeEstimator 2.0.1 and a threshold lowest allele frequency value of 0.05 (Do et al., 2014). Two complementary genetic clustering software were used to assess population structure: The *snmf* function of the LEA (v2.0) package (Frichot et al., 2014; Frichot and François, 2015) and Admixture (Alexander et al., 2009). The number ancestral populations (k) was allowed to vary between 1 and 10, with 10 replicate runs for each k -value, and the best k was chosen based on cross-entropy and cross-validation errors (Frichot et al., 2014). Population structure of *I. maurandioides* was assessed a second time, considering only the samples originating from Serra Norte (where the species is sympatric with *I. cavalcantei*).

Landscape Genetic Analyses

To tackle our second prediction we first tested for historical effects and then assessed fine-scale spatial genetic structure and isolation by landscape resistance (IBR) (Balkenhol et al., 2016). We tested for historical effects of size and isolation of the montane savanna highlands where our species occur, estimating area (m^2) and isolation (distance to the nearest highland in meters) using a pre-mining (1984) mapping of montane savanna highlands (Souza-Filho et al., 2016). We then ran mixed effects linear models containing individual-level genetic diversity estimates (H_E and F) as response variables, area and isolation as predictors, and highland ID as a random effect (see Figure S2). Likelihood ratio tests ($\alpha = 0.05$) were then employed to compare reduced models without each predictor variable, with full models containing them. All models were validated by plotting residual

vs. fitted values and by checking for multicollinearity and residual autocorrelation.

We assessed fine-scale spatial genetic structure by quantifying spatial autocorrelation in genetic relatedness. To do so we used local polynomial fitting (LOESS) of pairwise relatedness to pairwise geographic distance (<https://github.com/rojaff/Lplot>; Bruno et al., 2008). We then evaluated the contribution of land cover, elevation, terrain roughness, temperature, precipitation and geographic distance in explaining patterns of gene flow (Balkenhol et al., 2016). Using detailed land cover classification maps (Souza-Filho et al., 2016) we built six resistance surfaces, where we attributed different resistance values to each one of the main land cover classes (montane savanna, forest and mine). Each resistance surface thus represents a quantitative hypothesis about landscape resistance to gene flow. We first attributed low resistance values (0.1) to each main land cover class and high resistance values (0.9) to all other classes. Subsequently we did the opposite, attributing high resistance values to each main land cover class and low resistance values to all other classes. By so doing, each main land cover class was modeled as a high and a low permeability matrix. To account for possible time-lag effects (Landguth et al., 2010) we constructed similar resistance surfaces using land cover maps from three different decades (1994, 2004, and 2013, Figure S2). We chose this time span considering observations made by botanists from the Carajás Zoo Botanical Park, whom have planted both species for public exhibitions and estimate generation times of ~ 20 years for *I. cavalcantei* and ~ 3 years for *I. maurandioides*. Since elevation,

terrain roughness, mean temperature and mean precipitation have been found to influence gene flow in other plant species (Dyer, 2016), we constructed resistance surfaces employing both the raw and inverted rasters for each one of these variables (see Table S2 for spatial data specifications). Finally, we used a land cover map to create a null model for isolation by resistance (isolation by geographic distance), where all pixels were coded with identical resistance values (0.1). All rasters were set to Universal Transverse Mercator (UTM) projection and cropped to the extent of sample locations plus a buffer area of 5 km to minimize border effects (Jaffé et al., 2015). Using the program Circuitscape V4.0 (McRae, 2006) we then calculated pairwise resistance distances between all samples, employing all the resistance surfaces described above. Due to Circuitscape's computing limitations we replaced pixels containing zero values with 0.001. To assess IBR we fit mixed-effects regression models using penalized least squares and a correlation structure designed to account for the non-independence of pairwise distances (maximum-likelihood population effects or MLPE: <https://github.com/nspope/corMLPE>; Clarke et al., 2002). Yang's Relatedness between pairs of individuals (Yang et al., 2010) was used as response variable. Similar measures of relatedness have been found to be highly accurate as individual-based genetic distance metrics for landscape genetic studies (Shirk et al., 2017). Predictor variables included the different resistance distances (geographic distance, land cover, elevation, roughness, temperature and precipitation). To reduce the number of predictors we first selected which of the land cover models best explained genetic relatedness, using the sample size corrected Akaike Information Criterion (AICc) to compare univariate models from all years (1994, 2004, and 2013). Similarly, we compared univariate models containing resistance distances from raw and inverted rasters for each one of the remaining predictor variables. We then built full models for each species containing the best land cover surfaces (with lowest AICc) along with the best surfaces from the remaining predictors, and compared all models containing non-collinear predictors ($r < 0.6$) using the *dredge* function from the MuMIn (v1.4) package (https://github.com/rojaff/dredge_mc; Barton, 2015). Models with a $\Delta AICc < 2$ were included in the set of best models, and likelihood ratio tests (LRT) were performed to assess the influence of the inclusion of each predictor variable on the best model's log-likelihood. Finally, model averaging across the set of best models was used to compute parameter estimates that account for uncertainty in model selection (Burnham and Anderson, 2002; Jaffé et al., 2016).

Identification of Putative Adaptive Loci

Finally, we employed environmental association and F_{ST} outlier tests to address our third prediction. To identify putative adaptive loci we employed sets of loci filtered only by quality and depth (as described above), but not for LD or HWE. Adaptive signals were evaluated across all SNPs, even though many contigs contained more than one SNP. We thereby avoided the arbitrary exclusion of SNPs and the loss of any potentially adaptive signal. Linkage disequilibrium was nevertheless low within our 150 bp contigs (Figure S3), so most SNPs were independent markers.

We used latent factor mixed models (LFMM) to identify possible associations between SNPs and environmental variables, while accounting for the underlying population structure (Frichot et al., 2013; De Kort et al., 2014; Rellstab et al., 2015). We first ran a principal component analysis (PCA) using the 19 bioclimatic variables retrieved from WorldClim plus SRTM Elevation (Table S2), and selected the three variables showing the strongest correlation with the first three PCA axes (which explained 85 and 84% of total variance in *I. cavalcantei* and *I. maurandioides* respectively). These were minimum temperature of coldest month, precipitation of warmest quarter, and precipitation of wettest quarter for *I. cavalcantei*; and minimum temperature of coldest month, precipitation of wettest quarter, and precipitation of coldest quarter for *I. maurandioides*. Correlations between these variables were weak ($r < 0.3$), except for minimum temperature of coldest month and precipitation of wettest quarter in *I. maurandioides* ($r = 0.67$). We then used these variables to run LFMM using the LEA R package. Since incorrect assumptions about underlying demographic structure can increase both Type I and Type II errors (Storfer et al., 2018), we ran LFMM using multiple k -values (where k was the optimum number of ancestral populations detected), and then intersected candidate loci detected across all runs. Models were ran for $k \pm 2$ latent factors (*I. cavalcantei*: $k = 1, 2$, and 3; *I. maurandioides*: $k = 2, 3, 4, 5$, and 6; see results below). We used 10,000 iterations, a burn-in of 5,000 and five runs per environmental variable (Frichot and François, 2015). Finally, p -values were adjusted using the genomic inflation factor (λ) and false discovery rates were set using the Benjamini-Hochberg algorithm (Benjamini and Hochberg, 1995). We only considered as candidate loci those shared between all runs (assuming different k -values). In addition, we ran genome scans for selection using the *snmf* function from the LEA package, to identify outlier loci based on F_{ST} only (François et al., 2016). As above, p -values were adjusted using the genomic inflation factor (λ) and false discovery rates were set using the Benjamini-Hochberg algorithm. Full R scripts of LFMM and genome scans for selection can be found in the supporting information of François et al. (2016) and the LEA website (<http://membres-timc.imag.fr/Olivier.Francois/LEA/index.htm>). To evaluate if adaptive signals differed between species we used a Chi-square test to compare the number of contigs containing candidate SNPs in relation to all the tested contigs.

In order to search for the proteins coded by the genes contained in the flanking regions of our candidate SNPs, contig sequences containing candidate loci were first submitted to the EMBOSS Transeq (http://www.ebi.ac.uk/Tools/st/emboss_transeq/) to obtain its corresponding protein sequences. We used all six frames with standard code (codon table), regions (start-end), trimming (yes), and reverse(no). We then ran a functional analysis using InterPro (<https://www.ebi.ac.uk/interpro/>; `interproscan.sh -dp -appl PfamA,TIGRFAM,PRINTS,PrositesPatterns,Gene3d -goterms -pathways -f tsv -o MySequences.tsv -i MySequences.faa`), searching for go terms and pathways along the respective annotation databases (Interpro, Pfam, Tigrfam, Prints, PrositesPattern and Gene3d).

RESULTS

Genetic Diversity and Population Structure

A total of 41,881 SNPs were identified in *I. calvacantei* and 30,103 in *I. maurandioides*. After filtering these datasets (for quality, depth, LD, HWE, and candidate adaptive loci) we obtained sets of neutral markers containing 14,145 and 7,259 loci respectively. While heterozygosity and inbreeding levels were similar between *I. calvacantei* and the analyzed populations of *I. maurandioides*, nucleotide diversity, Tajima's D and effective population size were higher in *I. calvacantei* (Table 1). Two genetic clustering approaches indicated the

presence of one population in *I. calvacantei* and four populations in *I. maurandioides* (Figure 2, Figure S4, Data Sheet 3).

Landscape Genetic Analyses

We did not find evidence that pre-mining highland area and isolation were associated with current levels of heterozygosity and inbreeding (Table S3), and both species showed weak but significant genetic spatial autocorrelation for up to 5 km (Figure 3). The best models of isolation by landscape resistance (IBR) for *I. calvacantei* included precipitation, temperature, terrain roughness, and land cover, but only these last two

TABLE 1 | Genetic diversity measures for *I. calvacantei* and *I. maurandioides*.

Species	Population	N	H_E (CI)	F (CI)	π (CI)	D (CI)	N_e (CI)
<i>I. calvacantei</i>	Pop1	122	0.18 (0.18/0.18)	0.12 (0.09/0.15)	0.18 (0.18/0.19)	0.16 (0.14/0.18)	855.60 (845.90/865.60) ^a
<i>I. maurandioides</i>	Pop1	79	0.17 (0.17/0.17)	0.07 (0.02/0.11)	0.13 (0.13/0.13)	0.05 (0.01/0.08)	182.80 (180.70/184.80)
	Pop2	63	0.18 (0.18/0.18)	0.01 (-0.01/0.04)	0.15 (0.15/0.15)	0.02 (-0.004/0.05)	224.20 (221.20/227.20)
	Pop3	90	0.18 (0.18/0.18)	0.03 (-0.02/0.08)	0.14 (0.14/0.15)	-0.02 (-0.05/0.003)	129.20 (128.30/130.00)
	Pop4	22	0.23 (0.23/0.23)	0.01 (-0.05/0.07)	0.13 (0.12/0.13)	-0.01 (-0.03/0.02)	222.80 (210.70/236.40)

In the later species genetic diversity measures are grouped by genetic clusters. Sample sizes (N) are shown followed by mean expected heterozygosity (H_E), mean inbreeding coefficient (F), mean per-site nucleotide diversity (π), Tajima's D (D) and effective population size (N_e). All estimates are shown along 95% confidence intervals (CI).

^aA similar estimate was found when analyzing a random sample of 22 *I. calvacantei* individuals: $N_e = 901.10$ (834.80/978.70).

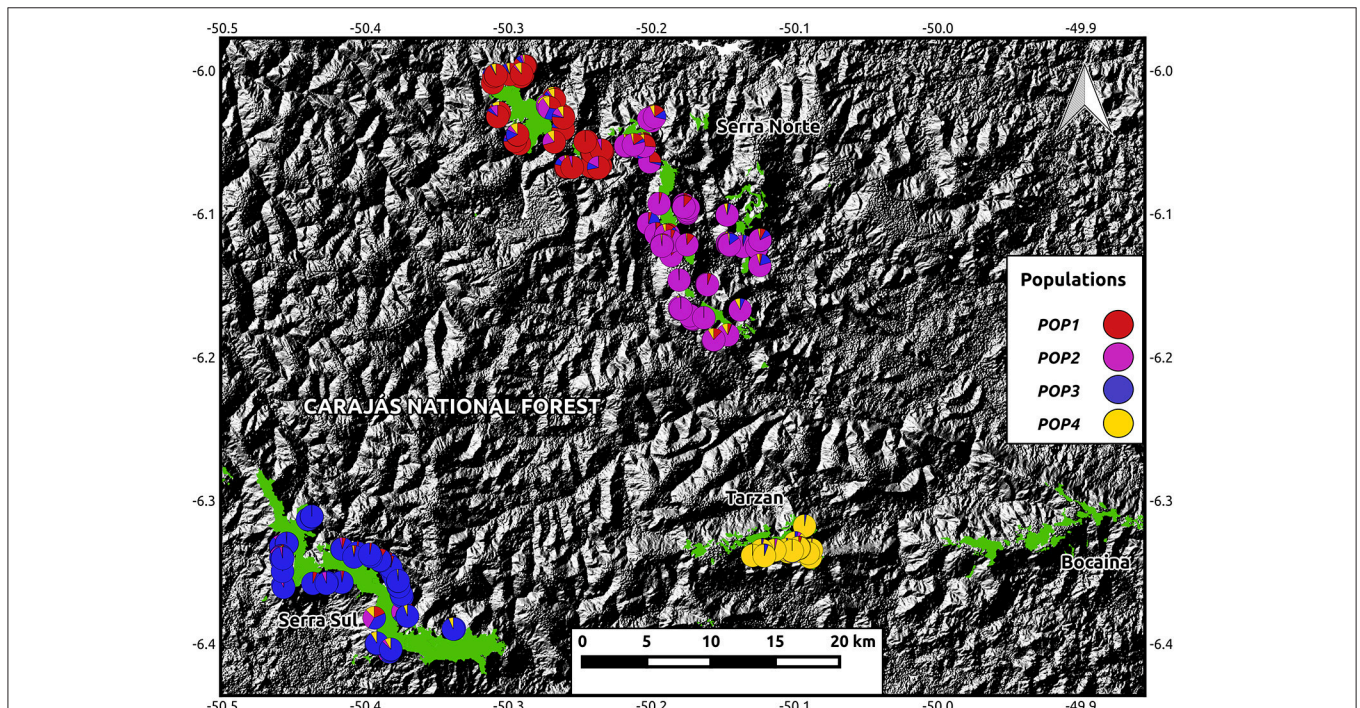
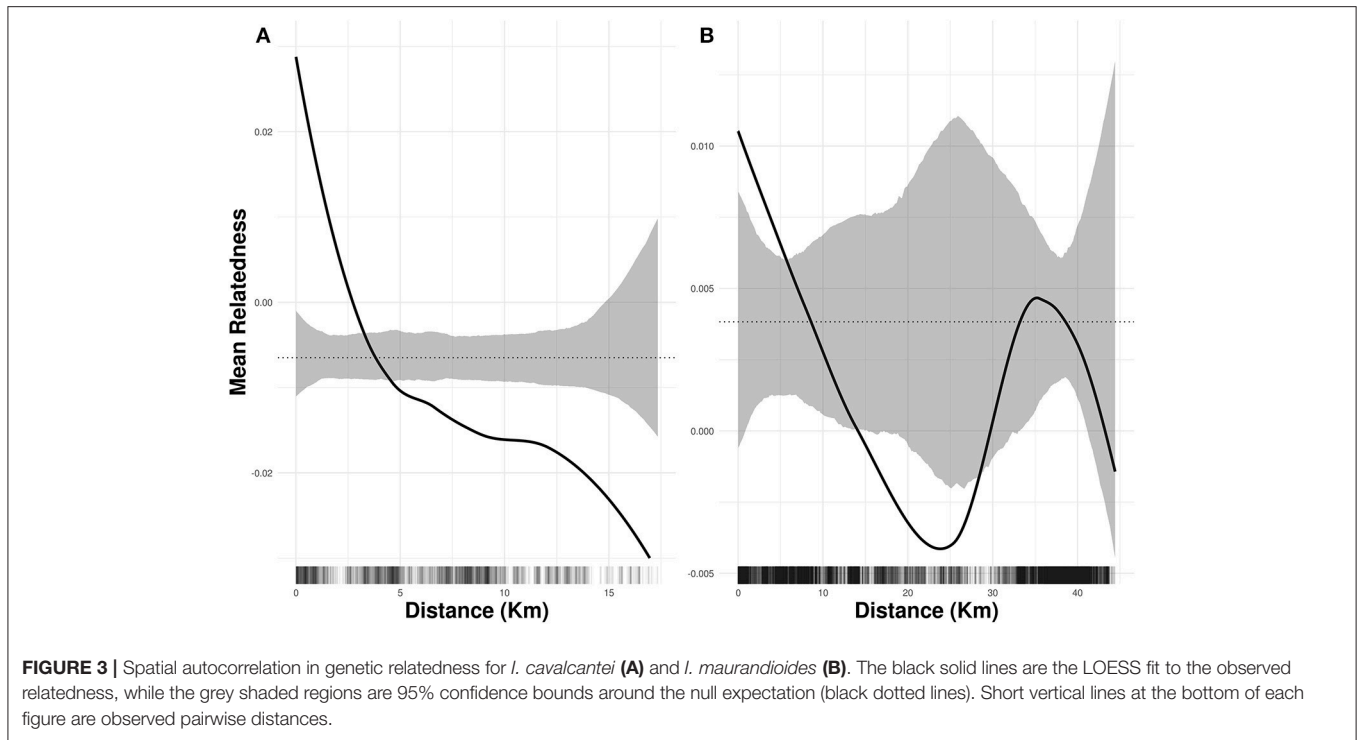


FIGURE 2 | Map showing *I. maurandioides* assignments to four genetic clusters against an elevation map (from USGS Earth Explorer). Pie charts represent ancestry coefficients determined using the LEA package. Montane savanna highlands (from Souza-Filho et al., 2016) are shown in green. Coordinates are shown in decimal degrees.



predictors significantly improved the model's log-likelihood (Table 2). Relatedness decreased with lower terrain roughness and higher land cover resistance (Tables 3, Figure S5). On the other hand, the best IBR models for *I. maurandioides* included precipitation, temperature, terrain roughness and land cover, but land cover was the only predictor that did not significantly improved the model's log-likelihood (Table 2). In this species relatedness decreased with lower precipitation, lower temperatures and higher terrain roughness (Table 3). Interestingly, univariate land cover models from all the tested years (1994, 2004, and 2013) were found to influence patterns of genetic relatedness in *I. maurandioides* but not in *I. cavalcantei*, where only land cover resistance distances from year 1994 were included in the set of best-fitting models (Table S4). The best land cover models for both species consistently revealed that forested and mining areas impose a higher resistance to gene flow than montane savanna areas (Table S5).

Identification of Putative Adaptive Loci

We found 2,733 candidate SNPs out of the 34,102 that remained after filtering for depth and quality in *I. cavalcantei*, whereas we detected 5,238 candidate SNPs out of 23,181 analyzed SNPs in *I. maurandioides* (Table 4, Figure S6). We identified a total of 1,147 contigs containing putative adaptive SNPs out of 9,625 analyzed contig sequences in *I. cavalcantei* (11.92%), and 3,065 contigs containing candidate SNPs out of 9,143 analyzed contig sequences in *I. maurandioides* (33.52%; Chi-square test: $\chi^2 = 798.19$, $p < 0.001$; Table 4, Figure S6). Within each species, some loci were identified with more than one statistical approach or were found associated with different environmental

variables (Table S6, Figure 4). Only a subset of the identified putative adaptive genes contained InterPro annotations (40 in *I. cavalcantei* and 131 in *I. maurandioides*; Data Sheets 4, 5). Interestingly, we found annotated genes shared between both species that were associated with the same environmental variables (Table S7).

DISCUSSION

Employing RAD sequencing to genotype thousands of genetic markers in two congeneric and sympatric morning glories from Amazonian savannas, we were able to compare genetic diversity, isolation by landscape resistance, and patterns of local adaptation in species belonging to different IUCN Red List Categories. Our results reveal higher genetic diversity and lower genetic differentiation in the endangered *I. cavalcantei* compared to the analyzed populations of the least concern *I. maurandioides*. Terrain roughness influenced gene flow in both species and temperature and precipitation were meaningful predictors of genetic relatedness patterns in *I. maurandioides*, whereas resistance to gene flow associated with forested and mining areas in existence two decades ago was found to influence IBR patterns in the narrowly distributed *I. cavalcantei*. Finally, we identified a set of putative adaptive genes associated to the same environmental variables in both species.

Contrary to our initial prediction, we found higher genetic diversity and lower spatial genetic structure in the narrow-endemic species. For instance, we found a larger N_e in *I. cavalcantei* than in *I. maurandioides* or *I. purpurea*, another widespread morning glory (Alvarado-Serrano et al., 2017),

TABLE 2 | Model selection summary showing the best MLPE models ($\Delta AICc \leq 2$) for each species.

Species	Model	logLikelihood	AICc	$\Delta AICc$	Weight
<i>I. cavalcantei</i>	Precipitation + Inverse roughness*** + Land cover 1994 (Low montane savanna resistance)**	14847.54	-29683.10	0.00	0.53
	Temperature + Inverse roughness*** + Land cover 1994 (Low montane savanna resistance)**	14847.40	-29682.80	0.28	0.46
<i>I. maurandioides</i>	Inverse precipitation*** + Temperature*** + Roughness*** + Land cover 2013 (Low montane savanna resistance)	36747.27	-73480.50	0.00	0.86

All models contained inter-individual genetic relatedness as response variable and the different resistance distances as predictors. Likelihood Ratio Test (LRT) were performed to assess if each predictor variable significantly improved the model's log-likelihood (significance levels are highlighted with: * $p < 0.05$; ** $p < 0.01$; and *** $p < 0.001$).

TABLE 3 | Summary statistics of the best MLPE models ($\Delta AICc \leq 2$) for each species.

Species	Predictor	Estimate	df	SE	t-value	p-value
<i>I. cavalcantei</i>	Precipitation	-4.82×10^{-6}	7381	2.81×10^{-6}	1.72	0.09
	Temperature	-3.71×10^{-5}	7381	2.27×10^{-5}	1.63	0.10
	Inverse roughness	-1.83×10^{-4}	7381	3.50×10^{-5}	5.24	<0.001
	Land cover 1994 (Low montane savanna resistance)	-8.62×10^{-2}	7381	9.80×10^{-3}	8.79	<0.001
<i>I. maurandioides</i>	Inverse precipitation	-3.41×10^{-4}	32131	5.76×10^{-5}	-5.93	<0.001
	Temperature	6.91×10^{-5}	32131	1.49×10^{-5}	4.64	<0.001
	Roughness	-2.26×10^{-4}	32131	6.80×10^{-5}	-3.33	<0.001
	Land cover 2013 (Low montane savanna resistance)	1.58×10^{-3}	32131	3.41×10^{-3}	0.46	0.64

Model-averaged parameter estimates for each predictor variable are shown followed by degrees of freedom (df) standard errors (SE), t-values and p-values.

revealing that extent of occurrence (IUCN Red List Criterion B) is not a good predictor of genetic diversity in these plants (Frankham, 1996). Our large N_e estimate does not seem to be a sample size artifact (Table 1), and suggests the actual *I. cavalcantei* population is above the N_e threshold of 500 required to maintain long-term evolutionary potential (Jamieson and Allendorf, 2012). Although confidence intervals for inbreeding coefficients overlapped between species, our results reveal significant inbreeding for *I. cavalcantei* and one population of *I. maurandioides* (were 95% CI did not contain zero). While the panmictic population of *I. cavalcantei* occurs both in preserved and degraded montane savanna highlands (Figure 1), the one *I. maurandioides* population exposed to mining activities (Pop 2 in Figure 2) did not show higher inbreeding than any of the other three undisturbed populations (Table 1). Mining activities thus appear unrelated to observed inbreeding. Likewise, historical highland area and highland isolation were not found to be associated with inbreeding levels (Table S3). Observed inbreeding thus appears to be the result of some level of selfing (Leimu et al., 2006; Duncan and Rausher, 2013).

Two genetic clustering approaches identified the presence of a single panmictic population in *I. cavalcantei* (Figure S4), indicating relatively high levels of gene flow in the recent past, throughout its occurrence range. Interestingly, we found a small yet positive value of Tajima's D in *I. cavalcantei* (Table 1), which suggests the incidence of balancing selection or an ongoing population contraction (Simonsen et al., 1995), the later probably related to the expansion of mining areas (Figure S2). In contrast, we detected four spatially distinct populations in the more widely

distributed *I. maurandioides*, two of which occurred within *I. cavalcantei*'s occurrence area (Figure 2, Data Sheet 3). Since founder effects result in higher spatial genetic structure and smaller effective population sizes (Booy et al., 2000; Allendorf et al., 2012), the recent colonization of our study region by few individuals of *I. maurandioides* could explain the observed patterns of population structure and genetic diversity. The fact that our study region is located at the northern edge of *I. maurandioides* distribution range (see Data Sheet 1), indeed hints at a possible recent colonization. However, we did not find negative Tajima's D-values, indicative of recent bottlenecks or population expansions (Simonsen et al., 1995). In fact Tajima's D-values overlapped with zero or were slightly positive in one population (Table 1), suggesting most populations were at mutation-drift equilibrium. Our results therefore fail to support the incidence of founder effects.

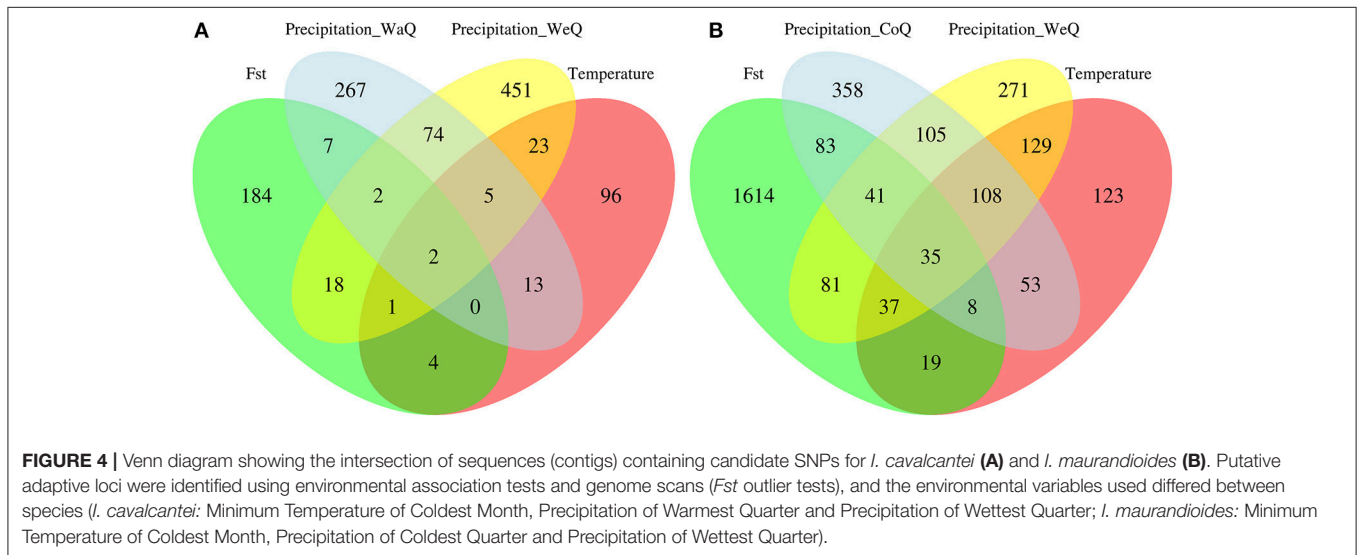
Life history differences influencing pollination and seed dispersal are known to be key predictors of gene flow and spatial genetic structure across plants (Vranckx et al., 2012; Gelmi-Candusso et al., 2017), so we believe these are likely to explain our observed patterns of genetic structure. Although our dataset does not allow disentangling the relative contribution of pollen-movement and seed dispersal to gene flow, wind-mediated seed dispersal is likely to be more restricted, as forest surrounding the montane savanna highlands constitute natural wind barriers. For instance, a recent work found strong spatial genetic structure when analyzing seven Chloroplast DNA loci in *I. cavalcantei* (Babiychuk et al., 2017), thus suggesting restricted seed dispersal in this species. Additionally, the significant (albeit

TABLE 4 | Summary of the number of adaptive signals detected using environmental association and F_{ST} outlier tests.

Species	Signal type	Total analyzed	Total under selection	Environmental association tests ^a			Fst outlier tests
				Temperature	Precipitation (WaQ/CoQ)	Precipitation(WeQ)	
<i>I. cavalcantei</i>	SNPs	34,102	2,733	327 (225)	904 (729)	1,513 (1,285)	262 (238)
	Contigs	9,625	1,147	144 (96)	370 (267)	576 (451)	218 (184)
<i>I. maurandioides</i>	SNPs	23,181	5,238	1,018 (331)	1,533 (841)	1,513 (642)	2,622 (2,331)
	Contigs	9,143	3,065	512 (123)	791 (358)	807 (271)	1,918 (1,614)

Both the number of candidate SNPs and the number of contigs (RAD tags) containing candidate SNPs are presented for each methodological approach followed by the number of independent (non-overlapping) detections in parentheses.

^aThe tested environmental variables differed between species. *I. cavalcantei*: Minimum Temperature of Coldest Month (Temperature), Precipitation of Warmest Quarter (Precipitation WaQ) and Precipitation of Wettest Quarter (Precipitation WeQ); *I. maurandioides*: Minimum Temperature of Coldest Month (Temperature), Precipitation of Coldest Quarter (Precipitation CoQ) and Precipitation of Wettest Quarter (Precipitation WeQ).



low) spatial autocorrelation of genotypes found in both species (Figure 3) indicates the existence of patches of related individuals from limited seed donors (Davies et al., 2010). This fine-space genetic structure, however, disappears at larger spatial scales, suggesting a role of pollen movement in maintaining gene flow across different montane savanna highlands (Sork and Smouse, 2006). The observed differences in spatial genetic structure between both species thus appear to be related to the dispersal capabilities of their pollinators (Sork and Smouse, 2006), and support earlier findings showing higher genetic structure in bee-pollinated plants compared to hummingbird-pollinated species (Kramer et al., 2011).

The IBR patterns detected here are thus likely to reflect the influence of local landscape on pollinator-mediated pollen transfer. Terrain roughness, for instance, seems to have facilitated humming-bird mediated pollination but hindered insect-mediated pollen transfer (Table 3). Indeed topographic complexity has been found to influence gene flow in other organisms (Murphy et al., 2010; Nogueras et al., 2016), and elevation has been associated to IBR in other tropical plants (Castilla et al., 2016). On the other hand, climatic factors (temperature and precipitation) were only found to influence

gene flow in *I. maurandioides*, indicating they could have affected bee dispersal (Jaffé et al., 2016) or caused plant phenological differences (Kramer et al., 2011).

Two lines of reasoning suggest our genetic diversity estimates mainly reflect gene flow taking place in the recent past. First, our estimates were based on neutral and independent nuclear loci. Second, heterozygosity and inbreeding were unaffected by historical highland area or highland isolation (Table S3). Interestingly, we found that only the oldest land cover map (1994) was able to explain the observed genetic relatedness patterns in *I. cavalcantei*, while all maps (1994–2013) were equally informative in the case of *I. maurandioides* (Table S4). This result was expected, given that generation times are much longer in the former species, and indicate a longevity dependent time-lag between land cover changes and genetic differentiation patterns (Landguth et al., 2010). Thus, our *I. cavalcantei* data do not allow us assessing restrictions to gene flow due to land cover changes that occurred after 1994.

In spite of these temporal limitations, land cover significantly influenced gene flow in the narrow-endemic but not in the broadly distributed species, as we initially predicted (Table 3). Specifically, forested and mining areas were found to impose

a higher resistance to gene flow than montane savanna areas (Table S5), a result that confirms the importance of this unique ecosystem as the exclusive habitat of this morning glory (Skiryicz et al., 2014; Viana et al., 2016). Forested and mining areas thus appear to hinder hummingbird-mediated pollination, suggesting an effect of habitat (montane savanna) fragmentation (Hadley et al., 2014). This effect, however, did not influence overall genetic structure, indicating that a substantial level of gene flow is still maintained across forested areas. The impact of mining on population-wide genetic structure is nevertheless more difficult to assess, as such an effect would only appear after several generations exposed to mining (Landguth et al., 2010).

Our results support our prediction of weaker adaptive signals (a lower proportion of candidate loci) in the narrow-endemic species, thereby suggesting that stronger genetic drift is slowing adaptive responses in *I. cavalcantei* (Allendorf et al., 2012). Although local climatic variability was limited, due to the resolution of bioclimatic variables (see Table S8), we believe that our environmental association tests were able to identify true adaptive signals. For instance, our environmental association tests can be considered conservative, since they accounted for false discovery rates and showed genomic inflation factors below 1 (Table S6; François et al., 2016). Moreover, LFMM results are robust to deviations from the underlying demographic structure, as candidate loci were intersected across different k -values. On the other hand, F_{ST} outlier tests showed genomic inflation factors above 1 (Table S6; François et al., 2016), indicating these were more liberal tests. We therefore caution about a higher incidence of Type I errors in our F_{ST} outlier tests, and highlight the need for follow-up studies involving different analytical approaches or functional validation. Likewise, we note that many other genes occurring in the flanking regions of our candidate SNPs could be responsible for the detected adaptive signals, so fine genomic mapping is necessary to confirm these genes.

In many cases we did not find any matches for our translated proteins, or found matches with uncharacterized proteins. This explains the lower number of putative adaptive proteins identified compared to the initial number of candidate loci. One of these proteins (reverse transcriptase, RNA-dependent DNA polymerase) was found associated to all environmental variables and was also identified through F_{ST} outlier tests in both species (Table S7). This protein indicates the presence of retrotransposons, as does ribonuclease H-like domain, retroviral aspartyl protease, retrotransposon gag domain, and integrase. Retrotransposons are abundant and ancient components of plant genomes associated to stress response (Grandbastien, 1998). Indeed reverse transcriptase acts as key initiator of defense responses in flowering plants (Sarris et al., 2016), while Leucine-rich repeat domain, CAP domain proteins, knottin and defensin plant proteins are also involved in plant defenses against pathogens (Gibbs et al., 2008; Benko-Iseppon et al., 2010; Tavormina et al., 2015). On the other hand, WPP domain and CCHC zinc finger proteins are involved in plant cell division (Patel et al., 2004; Clay and Nelson, 2005), and death domain proteins in apoptotic signaling (Coll et al., 2011). Finally, ClpP/crotonases, hydrolases, esterases, helicases and pectinesterases carry out a wide range of metabolic functions

(Holden et al., 2001; Xu et al., 2004). Interestingly, in both species we found proteins identified using F_{ST} outlier tests that did not overlap with those found by environmental association tests (Table 4). Our results thus suggest that these proteins are signaling general responses, unrelated to environmental variations. In sum, most of the identified putative proteins appear to be related to pathogen resistance and physiological adaptations to life in the extreme environments of montane savannas (Skiryicz et al., 2014; Silveira et al., 2015; Viana et al., 2016). These potential adaptations are relevant to conservation and restoration initiatives seeking the protection of adaptive genetic diversity or the introduction of individuals adapted to adverse environmental conditions into degraded areas (Shryock et al., 2017).

Most conservation geneticists have used the 50/500 rule to define minimum viable populations (Jamieson and Allendorf, 2012; Franklin et al., 2014): An N_e of 50 is deemed necessary to avoid short-term inbreeding depression, while an N_e of 500 is recommended to maintain long-term evolutionary potential. More recently, Willoughby et al. (2015) presented a novel approach for identifying species of conservation need by estimating the expected loss of genetic diversity. Their method, however, requires access to a large database of heterozygosity values from related species in order to establish a cut off value. Since we do not have access to such a database we cannot implement this novel approach to re-assess IUCN Red List Categories. Instead, we base our conservation recommendations on the 50/500 rule. In the case of the narrow-endemic *I. cavalcantei*, which showed comparatively higher genetic diversity and high levels of gene flow across its entire occurrence range, conservation efforts should be oriented to preserve long-term evolutionary potential ($N_e > 500$). On the other hand, precautionary conservation measures should aim at maintaining $N_e > 50$ within each population to avoid short term inbreeding depression in *I. maurandioides*. Although land cover was found to influence gene flow in the narrow-endemic species, overall population panmixia indicates that a substantial level of genetic connectivity is maintained across the fragmented habitat of *I. cavalcantei*. Additionally, we identified possible adaptations for life in Amazonian savannas, which could inform future conservation and restoration programs.

Our study exemplifies how landscape genomic tools can be applied to assess the conservation status, susceptibility to global change, and adaptive responses of imperiled biodiversity. More specifically, our genomic assessment revealed unexpected patterns of genetic diversity, spatial genetic structure and gene flow in tropical plant species belonging to different IUCN Red List Categories. Given the level of resolution of our genetic diversity and gene flow estimates, our findings emphasize that the IUCN Red List Criteria approach does not systematically prioritize the conservation of species with low genetic diversity (Rivers et al., 2014; Willoughby et al., 2015) or highest susceptibility to habitat loss and fragmentation, and call for the inclusion of genetic information into conservation assessments (an issue currently evaluated by the IUCN Conservation Genetics Specialist Group: <http://www.cgsg.uni-freiburg.de/>). We are well aware this will be no easy endeavor, as funding limitations, country-specific conservation policies and

communication barriers between scientists and practitioners are among the issues to overcome (Hoban et al., 2013; Bowman et al., 2016; Pierson et al., 2016; Taylor et al., 2017), but would like to stress out that the demographic parameters employed to assess conservation status two decades ago have remained unchanged in the age of next-generation-sequencing and meta-genomics.

AUTHOR CONTRIBUTIONS

RJ: conceived, designed and coordinated the project; RJ and VI-F: initiated the project; RJ, ÉL, and AG: coordinated the field work and sampling; ÉL, NC, WM, and SV: performed laboratory work; RJ, ÉL, NP, RA, WM, AS, and PS-F: performed the data analysis. The first draft of the paper was written by ÉL with input from RJ. All authors contributed to discussing the results and editing the paper.

REFERENCES

- Aguilar, R., Quesada, M., Ashworth, L., Herrerias-Diego, Y., and Lobo, J. (2008). Genetic consequences of habitat fragmentation in plant populations: Susceptible signals in plant traits and methodological approaches. *Mol. Ecol.* 17, 5177–5188. doi: 10.1111/j.1365-294X.2008.03971.x
- Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. doi: 10.1101/gr.094052.109
- Allendorf, F. W., Luikart, G. H., and Aitken, S. N. (2012). *Conservation and the Genetics of Populations*. Oxford, UK: Wiley.
- Alvarado-Serrano, D. F., Van Etten, M. L., Chang, S.-M., and Baucom, R. S. (2017). Landscape connectivity of a noxious invasive weed: promises and challenges of landscape genomics for knowledge-based weed management. *bioRxiv*:13:43. doi: 10.1101/054122
- Babiychuk, E., Kushnir, S., Vasconcelos, S., Dias, M. C., Carvalho-Filho, N., Nunes, G. L., et al. (2017). Natural history of the narrow endemics *Ipomoea cavalcantei* and *I. marabaensis* from Amazon Canga savannahs. *Sci. Rep.* 7:7493. doi: 10.1038/s41598-017-07398-z
- Balkenhol, N., Cushman, S., Storfer, A., and Waits, L. (eds). (2016). *Landscape Genetics: Concepts, Methods, Applications*. Hoboken, NJ: John Wiley & Sons.
- Barton, K. (2015). *MuMIn: Multi-Model Inference*. R package version 1.15.1. Version 1:18.
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57, 289–300.
- Benko-Iseppon, A. M., Lins Galdino, S., Calsa, T. Jr., Akio Kido, E., Tossi, A., Carlos Belarmino, L., et al. (2010). Overview on Plant Antimicrobial Peptides. *Curr. Protein Pept. Sci.* 11, 181–188. doi: 10.2174/138920310791112075
- Booy, G., Hendriks, R. J. J., Smulders, M. J. M., Van Groenendaal, J. M., and Vosman, B. (2000). Genetic diversity and the survival of populations. *Plant Biol.* 2, 379–395. doi: 10.1055/s-2000-5958
- Bowman, J., Greenhorn, J. E., Marrotte, R. R., McKay, M. M., Morris, K. Y., Prentice, M. B., et al. (2016). On applications of landscape genetics. *Conser. Genet.* 17, 753–760. doi: 10.1007/s10592-016-0834-5
- Bruno, C., Macchiavelli, R., and Balzarini, M. (2008). Non-parametric smoothing of multivariate genetic distances in the analysis of spatial population structure at fine scale. *Theor. Appl. Genet.* 117, 435–447. doi: 10.1007/s00122-008-0788-1
- Burnham, K. P., and Anderson, D. R. (2002). *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach*. New York, NY: Springer.
- Castilla, A. R., Pope, N., Jaffé, R., and Jha, S. (2016). Elevation, not deforestation, promotes genetic differentiation in a pioneer tropical tree. *PLoS ONE* 11:e0156694. doi: 10.1371/journal.pone.0156694

ACKNOWLEDGMENTS

Funding was provided by Instituto Tecnológico Vale, CNPq grants 307479/2016-1 (GO) and 300714/2017-3 (ÉL), CAPES MSc scholarship 88882.161651/2017-01 (WM), and an NSF predoctoral fellowship (NP). We thank Elena Babiitouchouk, Alexandre Castilho, Cesar Neto and Xavier Prous for assistance in the field, Rafael M. Brito for GIS support, and Manoel Lopes for help in the laboratory. Tais Fernandes, Antonio R. Castilla, and three Referees substantially improved earlier versions of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2018.00532/full#supplementary-material>

- Clarke, R. T., Rothery, P., and Raybould, A. F. (2002). Confidence limits for regression relationships between distance matrices: estimating gene flow with distance. *J. Agric. Biol. Environ. Stat.* 7, 361–372. doi: 10.1198/108571102320
- Clay, N. K., and Nelson, T. (2005). The recessive epigenetic swillmap mutation affects the expression of two step II splicing factors required for the transcription of the cell proliferation gene *Struwwelpeter* and for the timing of cell cycle arrest in the Arabidopsis leaf. *Plant Cell* 17, 1994–2008. doi: 10.1105/tpc.105.032771
- Cole, C. (2003). Genetic variation in rare and common plants. *Ann. Rev. Ecol. Evol. Syst.* 34, 213–237. doi: 10.1146/annurev.ecolsys.34.030102.151717
- Coll, N. S., Epple, P., and Dangel, J. L. (2011). Programmed cell death in the plant immune system. *Cell Death Differentiat.* 18, 1247–1256. doi: 10.1038/cdd.2011.37
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi: 10.1093/bioinformatics/btr330
- Davies, S. J., Cavers, S., Finegan, B., Navarro, C., and Lowe, A. J. (2010). Genetic consequences of multigenerational and landscape colonisation bottlenecks for a neotropical forest pioneer tree, *Vochysia ferruginea*. *Trop. Plant Biol.* 3, 14–27. doi: 10.1007/s12042-010-9040-7
- de Carvalho, W. D., and Mustin, K. (2017). The highly threatened and little known Amazonian savannahs. *Nat. Ecol. Evol.* 1:100. doi: 10.1038/s41559-017-0100
- De Kort, H., Vandepitte, K., Bruun, H. H., Closset-Kopp, D., Honnay, O., and Mergeay, J. (2014). Landscape genomics and a common garden trial reveal adaptive differentiation to temperature across Europe in the tree species *Alnus glutinosa*. *Mol. Ecol.* 23, 4709–4721. doi: 10.1111/mec.12813
- Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., and Ovenden, J. R. (2014). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Mol. Ecol. Resour.* 14, 209–214. doi: 10.1111/1755-0998.12157
- Duncan, T. M., and Rausher, M. D. (2013). Evolution of the selfing syndrome in *Ipomoea*. *Front. Plant Sci.* 4:301. doi: 10.3389/fpls.2013.00301
- Dyer, R. J. (2016). “Landscapes and Plant Population Genetics,” in *Landscape Genetics: Concepts, Methods, Applications*. Hoboken, eds N. Balkenhol, S. Cushman, A. Storfer, and L. Waits (Oxford, UK: Wiley Online Library), 181–198.
- François, O., Martins, H., Caye, K., and Schoville, S. D. (2016). Controlling false discoveries in genome scans for selection. *Mol. Ecol.* 25, 454–469. doi: 10.1111/mec.13513
- Frankham, R. (1995). Effective population size/adult population size ratios in wildlife: a review. *Genet. Res.* 66, 95–107. doi: 10.1017/S0016672300034455
- Frankham, R. (1996). Relationship of genetic variation to population size in wildlife. *Conserv. Biol.* 10, 1500–1508. doi: 10.1046/j.1523-1739.1996.10061500.x

- Frankham, R. (2015). Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. *Mol. Ecol.* 24, 2610–2618. doi: 10.1111/mec.13139
- Franklin, I. R., Allendorf, F. W., and Jamieson, I. G. (2014). The 50/500 rule is still valid—Reply to Frankham et al. *Biol. Conserv.* 176, 284–285. doi: 10.1016/j.biocon.2014.05.004
- Frichot, E., and François, O. (2015). LEA: an R package for landscape and ecological association studies. *Methods Ecol. Evol.* 6, 925–929. doi: 10.1111/2041-210X.12382
- Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., and François, O. (2014). Fast and efficient estimation of individual ancestry coefficients. *Genetics* 196, 973–983. doi: 10.1534/genetics.113.160572
- Frichot, E., Schoville, S. D., Bouchard, G., and François, O. (2013). Testing for associations between loci and environmental gradients using latent factor mixed models. *Mol. Biol. Evol.* 30, 1687–1699. doi: 10.1093/molbev/mst063
- Galetto, L., and Bernardello, G. (2004). Floral nectaries, nectar production dynamics and chemical composition in six Ipomoea species (Convolvulaceae) in relation to pollinators. *Ann. Bot.* 94, 269–280. doi: 10.1093/aob/mch137
- Garner, B. A., Hand, B. K., Amish, S. J., Bernatchez, L., Foster, J. T., Miller, K. M., et al. (2017). Genomics in conservation: case studies and bridging the gap between data and application. *Trends Ecol. Evol.* 31, 81–83. doi: 10.1016/j.tree.2015.10.009
- Gelmi-Candusso, T. A., Heymann, E. W., and Heer, K. (2017). Effects of zoochory on the spatial genetic structure of plant populations. *Mol. Ecol.* 26, 5896–5910. doi: 10.1111/mec.14351
- Gibbs, G. M., Roelants, K., and O'Bryan, M. K. (2008). The CAP superfamily: Cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1 proteins - Roles in reproduction, cancer, and immune defense. *Endocr. Rev.* 29, 865–897. doi: 10.1210/er.2008-0032
- Gibson, J., Rice, S. A., and Stucke, C. M. (2008). Comparison of population genetic diversity between a rare, narrowly distributed species and a common, widespread species of *Alnus* (Betulaceae). *Am. J. Bot.* 95, 588–596. doi: 10.3732/ajb.2007316
- Gitzendanner, M. A., and Soltis, P. S. (2000). Patterns of genetic variation in rare and widespread plant congeners. *Am. J. Bot.* 87, 783–792. doi: 10.2307/2656886
- Grandbastien, M. A. (1998). Activation of plant retrotransposons under stress conditions. *Trends Plant Sci.* 3, 181–187. doi: 10.1016/S1360-1385(98)01232-1
- Griz, L. M. S., and Machado, I. C. S. (2001). Fruiting phenology and seed dispersal syndromes in caatinga, a tropical dry forest in the northeast of Brazil. *J. Trop. Ecol.* 17, 303–321. doi: 10.1017/S0266467401001201
- Hadley, A. S., Frey, S. J. K., Robinson, W. D., Kress, W. J., and Betts, M. G. (2014). Tropical forest fragmentation limits pollination of a keystone understory herb. *Ecology* 95, 2202–2212. doi: 10.1890/13-0929.1
- Hamrick, J. L., Godt, M. J. W., Murawski, D. A., and Loveless, M. D. (1991). “Correlations between species traits and allozyme diversity: implications for conservation biology,” in *Genetics and conservation of rare plants*, eds D. A. Falk, K. E. Holsinger (New York, NY: Oxford University Press), 75–86.
- Hoban, S. M., Hauffe, H. C., Pérez-Espona, S., Arntzen, J. W., Bertorelle, G., Bryja, J., et al. (2013). Bringing genetic diversity to the forefront of conservation policy and management. *Conserv. Genet. Resour.* 5, 593–598. doi: 10.1007/s12686-013-9859-y
- Holden, H. M., Benning, M. M., Haller, T., and Gerlt, J. A. (2001). The crotonase superfamily: divergently related enzymes that catalyze different reactions involving acyl coenzyme a thioesters. *Acc. Chem. Res.* 34, 145–157. doi: 10.1021/ar000053l
- Honnay, O., and Jacquemyn, H. (2007). Susceptibility of common and rare plant species to the genetic consequences of habitat fragmentation. *Conserv. Biol.* 21, 823–831. doi: 10.1111/j.1523-1739.2006.00646.x
- IUCN (2012). *IUCN Red List Categories and Criteria: Version 3.1*. Gland; Cambridge: IUCN.
- IUCN (2017). *The IUCN Red List of Threatened Species*. Available online at www.iucnredlist.org (Accessed July 21, 2017).
- Jaffé, R., Castilla, A., Pope, N., Imperatriz-Fonseca, V. L., Metzger, J. P., Arias, M. C., et al. (2015). Landscape genetics of a tropical rescue pollinator. *Conserv. Genet.* 17, 267–278. doi: 10.1007/s10592-015-0779-0
- Jaffé, R., Pope, N., Acosta, A. L., Alves, D. A., Arias, M. C., De la Rúa, P., et al. (2016). Beekeeping practices and geographic distance, not land use, drive gene flow across tropical bees. *Mol. Ecol.* 25, 5345–5358. doi: 10.1111/mec.13852
- Jamieson, I. G., and Allendorf, F. W. (2012). How does the 50/500 rule apply to MVPs? *Trends Ecol. Evol.* 27, 578–584. doi: 10.1016/j.tree.2012.07.001
- Kramer, A. T., Fant, J. B., and Ashley, M. V. (2011). Influences of landscape and pollinators on population genetic structure: examples from three Penstemon (Plantaginaceae) species in the Great Basin. *Am. J. Bot.* 98, 109–121. doi: 10.3732/ajb.1000229
- Laikre, L. (2010). Genetic diversity is overlooked in international conservation policy implementation. *Conserv. Genet.* 11, 349–354. doi: 10.1007/s10592-009-0037-4
- Landguth, E. L., Cushman, S. A., Schwartz, M. K., McKelvey, K. S., Murphy, M., and Luikart, G. (2010). Quantifying the lag time to detect barriers in landscape genetics. *Mol. Ecol.* 19, 4179–4191. doi: 10.1111/j.1365-294X.2010.04808.x
- Leimu, R., Mutikainen, P., Koricheva, J., and Fischer, M. (2006). How general are positive relationships between plant population size, fitness and genetic variation? *J. Ecol.* 94, 942–952. doi: 10.1111/j.1365-2745.2006.01150.x
- Loureiro, J., Rodriguez, E., Doležel, J., and Santos, C. (2007). Two new nuclear isolation buffers for plant DNA flow cytometry: a test with 37 species. *Ann. Bot.* 100, 875–888. doi: 10.1093/aob/mcm152
- Mace, G. M., Collar, N. J., Gaston, K. J., Hilton-Taylor, C., Akçakaya, H. R., Leader-Williams, N., et al. (2008). Quantification of extinction risk: IUCN's system for classifying threatened species. *Conserv. Biol.* 22, 1424–1442. doi: 10.1111/j.1523-1739.2008.01044.x
- Martinelli, G., and Moraes, M. A. (eds.). (2013). *Livro Vermelho da Flora do Brasil*. Rio de Janeiro: CNCFlora, Centro Nacional de Conservação da Flora.
- McRae, B. H. (2006). Isolation by resistance. *Evolution* 60, 1551–1561. doi: 10.1111/j.0014-3820.2006.tb00500.x
- Murphy, M. A., Dezzani, R., Pilliod, D. S., and Storfer, A. (2010). Landscape genetics of high mountain frog metapopulations. *Mol. Ecol.* 19, 3634–3649. doi: 10.1111/j.1365-294X.2010.04723.x
- Noguerales, V., Cordero, P. J., and Ortego, J. (2016). Hierarchical genetic structure shaped by topography in a narrow-endemic montane grasshopper. *BMC Evol. Biol.* 16:96. doi: 10.1186/s12862-016-0663-7
- Obermayer, R., Leitch, I. J., Hanson, L., and Bennett, M. D. (2002). Nuclear DNA C-values in 30 species double the familial representation in pteridophytes. *Ann. Bot.* 90, 209–217. doi: 10.1093/aob/mcf167
- Patel, S., Rose, A., Meulia, T., Dixit, R., Cyr, R. J., and Meier, I. (2004). Arabidopsis WPP-domain proteins are developmentally associated with the nuclear envelope and promote cell division. *Plant Cell* 16, 3260–3273. doi: 10.1105/tpc.104.026740
- Pierson, J. C., Coates, D. J., Oostermeijer, J. G. B., Beissinger, S. R., Bragg, J. G., Sunnucks, P., et al. (2016). Genetic factors in threatened species recovery plans on three continents. *Front. Ecol. Environ.* 14, 433–440. doi: 10.1002/fee.1323
- Reed, D. H., and Frankham, R. (2003). Correlation between fitness and genetic diversity. *Conserv. Biol.* 17, 230–237. doi: 10.1046/j.1523-1739.2003.01236.x
- Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., and Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Mol. Ecol.* 24, 4348–4370. doi: 10.1111/mec.13322
- Rivers, M. C., Brummitt, N. A., Nic Lughadha, E., and Meagher, T. R. (2014). Do species conservation assessments capture genetic diversity? *Glob. Ecol. Conserv.* 2, 81–87. doi: 10.1016/j.gecco.2014.08.005
- Rodrigues, A. S. L., Pilgrim, J. D., Lamoreux, J. F., Hoffmann, M., and Brooks, T. M. (2006). The value of the IUCN Red List for conservation. *Trends Ecol. Evol. (Amst)*. 21, 71–76. doi: 10.1016/j.tree.2005.10.010
- Rogstad, S. H. (1992). Saturated NaCl-CTAB Solution as a Means of Field Preservation of Leaves for DNA Analyses. *Taxon* 41, 701–708. doi: 10.2307/1222395
- Russello, M. A., Waterhouse, M. D., Etter, P. D., and Johnson, E. A. (2015). From promise to practice: pairing non-invasive sampling with genomics in conservation. *PeerJ* 3:e1106. doi: 10.7717/peerj.1106
- Sarris, P. F., Cevik, V., Dagdas, G., Jones, J. D. G., and Krasileva, K. V. (2016). Comparative analysis of plant immune receptor architectures uncovers host proteins likely targeted by pathogens. *BMC Biol.* 14:8. doi: 10.1186/s12915-016-0228-7
- Shafer, A. B. A., Wolf, J. B. W., Alves, P. C., Bergström, L., Bruford, M. W., Brännström, I., et al. (2015). Genomics and the challenging translation into conservation practice. *Trends Ecol. Evol.* 30, 78–87. doi: 10.1016/j.tree.2014.11.009

- Shirk, A. J., Landguth, E. L., and Cushman, S. A. (2017). A comparison of individual-based genetic distance metrics for landscape genetics. *Mol. Ecol. Resour.* 17, 1308–1317. doi: 10.1111/1755-0998.12684
- Shryock, D. F., Havrilla, C. A., DeFalco, L. A., Esque, T. C., Custer, N. A., and Wood, T. E. (2017). Landscape genetic approaches to guide native plant restoration in the Mojave Desert. *Ecol. Appl.* 27, 429–445. doi: 10.1002/eap.1447
- Silveira, F. A. O., Negreiros, D., Barbosa, N. P. U., Buisson, E., Carmo, F. F., Carstensen, D. W., et al. (2015). Ecology and evolution of plant diversity in the endangered campo rupestre: a neglected conservation priority. *Plant Soil* 403, 129–152. doi: 10.1007/s11104-015-2637-8
- Simão-Bianchini, R., Vasconcelos, L. V., and Pastore, M. (2016). Flora das cangas da Serra dos Carajás, Pará, Brasil: convolvulaceae. *Rodriguésia* 67, 1301–1318. doi: 10.1590/2175-7860201667530
- Simonsen, K. L., Churchill, G. A., and Aquadro, C. F. (1995). Properties of statistical tests of neutrality for DNA polymorphism data. *Genetics* 141, 413–429.
- Skrycz, A., Castilho, A., Chaparro, C., Carvalho, N., Tzotzos, G., and Siqueira, J. O. (2014). Canga biodiversity, a matter of mining. *Front. Plant Sci.* 5:653. doi: 10.3389/fpls.2014.00653
- Sork, V. L., and Smouse, P. E. (2006). Genetic analysis of landscape connectivity in tree populations. *Landsc. Ecol.* 21, 821–836. doi: 10.1007/s10980-005-5415-9
- Souza-Filho, P. W. M., de Souza, E. B., Silva Júnior, R. O., Nascimento, W. R., Versiani de Mendonça, B. R., Guimarães, J. T. F., et al. (2016). Four decades of land-cover, land-use and hydroclimatology changes in the Itacaiúnas River watershed, southeastern Amazon. *J. Environ. Manage.* 167, 175–184. doi: 10.1016/j.jenvman.2015.11.039.
- Spielman, D., Brook, B. W., and Frankham, R. (2004). Most species are not driven to extinction before genetic factors impact them. *Proc. Natl. Acad. Sci. U.S.A.* 101, 15261–15264. doi: 10.1073/pnas.0403809101
- Storfer, A., Patton, A., and Fraik, A. K. (2018). Navigating the interface between landscape genetics and landscape genomics. *Front. Genet.* 9:68. doi: 10.3389/fgene.2018.00068
- Tavormina, P., De Coninck, B., Nikonorova, N., De Smet, I., and Cammue, B. P. A. (2015). The Plant Peptidome: an Expanding Repertoire of Structural Features and Biological Functions. *Plant Cell* 27, 2095–2118. doi: 10.1105/tpc.15.00440
- Taylor, H., Dussex, N., and van Heezik, Y. (2017). Bridging the conservation genetics gap by identifying barriers to implementation for conservation practitioners. *Glob. Ecol. Conserv.* 10, 231–242. doi: 10.1016/j.gecco.2017.04.001
- Viana, P. L., Mota NF de, O., Gil A dos, SB., Salino, A., Zappi, D. C., Harley, R. M., et al. (2016). Flora of the cangas of the Serra dos Carajás, Pará, Brazil: history, study area and methodology. *Rodriguésia* 67, 1107–1124. doi: 10.1590/2175-7860201667501
- Vranckx, G., Jacquemyn, H., Muys, B., and Honnay, O. (2012). Meta-Analysis of Susceptibility of Woody Plants to Loss of Genetic Diversity through Habitat Fragmentation. *Conserv. Biol.* 26, 228–237. doi: 10.1111/j.1523-1739.2011.01778.x
- Willoughby, J. R., Sundaram, M., Wijayawardena, B. K., Kimble, S. J. A., Ji, Y., Fernandez, N. B., et al. (2015). The reduction of genetic diversity in threatened vertebrates and new recommendations regarding IUCN conservation rankings. *Biol. Conserv.* 191, 495–503. doi: 10.1016/j.biocon.2015.07.025
- Wood, J. R. I., and Scotland, R. W. (2017). Misapplied names, synonyms and new species of Ipomoea (Convolvulaceae) from South America. *Kew Bulletin* 72:9. doi: 10.1007/s12225-017-9680-y
- Xu, Z., Escamilla-Treviño, L., Zeng, L., Lalgondar, M., Bevan, D., Winkel, B., et al. (2004). Functional genomic analysis of *Arabidopsis thaliana* glycoside hydrolase family 1. *Plant Mol. Biol.* 55, 343–367. doi: 10.1007/s11103-004-0790-1
- Yang, J., Benyamin, B., McEvoy, B. P., and Gordon, S. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* 42, 565–569. doi: 10.1038/ng.608

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Lanes, Pope, Alves, Carvalho Filho, Giannini, Giulietti, Imperatriz-Fonseca, Monteiro, Oliveira, Silva, Siqueira, Souza-Filho, Vasconcelos and Jaffé. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.