Plant Ecology & Diversity

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tped20

Integrating morphology, niche modelling, and molecular data to disentangle taxonomic challenges in a species complex of *Calibrachoa* (Solanaceae)

Alice Backes, Pedro H. Pezzi, Leonardo T. Gonçalves, Julián A. Greppi & Loreta B. Freitas

To cite this article: Alice Backes, Pedro H. Pezzi, Leonardo T. Gonçalves, Julián A. Greppi & Loreta B. Freitas (16 Nov 2023): Integrating morphology, niche modelling, and molecular data to disentangle taxonomic challenges in a species complex of *Calibrachoa* (Solanaceae), Plant Ecology & Diversity, DOI: 10.1080/17550874.2023.2274842

To link to this article: https://doi.org/10.1080/17550874.2023.2274842

Published online: 16 Nov 2023.

PLANT ECOLOGY & DIVERSITY
https://doi.org/10.1080/17550874.2023.2274842

Taylor & Francis
Taylor & Francis Group

Integrating morphology, niche modelling, and molecular data to disentangle taxonomic challenges in a species complex of *Calibrachoa* (Solanaceae)

Alice Backes[a], Pedro H. Pezzi[a], Leonardo T. Gonçalves[a], Julián A. Greppi[b] and Loreta B. Freitas[a]

[a]Department of Genetics, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil; [b]Instituto de Floricultura, Instituto Nacional de Tecnología Agropecuaria, Buenos Aires, Argentina

ABSTRACT

Background: Plant species delimitation can be difficult when clearly visible diagnostic characteristics are lacking. Historically, phenotypic similarity has been used as a criterion to group individuals into species. However, these traits can fail to discriminate between morphologically similar species or improperly subdivide species through inaccurate interpretation of natural phenotypic diversity.
Aims: We aimed to clarify the taxonomic status of two *Calibrachoa* taxa for which some polymorphism sharing is evident.
Methods: We evaluated the diversity between the taxa in terms of molecular diversity, niche divergence, and floral and foliar differences.
Results: Overall, a high level of polymorphism sharing between the taxa was indicated, possibly reflecting their rapid divergence. However, leaf morphology was a good predictor of taxa associated with phylogenetically informative sequences. Diversity among taxa was due to differentiation more than local adaptation. Ecological niche modelling indicated no niche divergence between the two taxa.
Conclusions: Our results suggest that the recent synonymisation of *Calibrachoa thymifolia* and *C. linearis* should be reviewed, with both taxa possibly considered as valid species. More molecular markers should be used to test this possibility, ideally examining divergence across their genomes.

ARTICLE HISTORY
Received 19 April 2023
Accepted 20 October 2023

KEYWORDS
Calibrachoa; evolutionary lineages; speciation; subtropical lowland grasslands; Pampa

Introduction

There is much debate about the concept and limits of species, with numerous theories and methods proposed. Delimiting species is fundamental to understanding many evolutionary mechanisms and processes (Sites and Marshall 2003), quantifying biodiversity, and promoting effective conservation plans (Li and Wiens 2022). The species boundary defines the limits within or across which evolutionary processes operate (Barton and Gale 1993). For example, over- or under-resolving species boundaries could confound studies aimed at understanding population-level processes, as species are usually considered the fundamental unit in biogeographical, ecological, and evolutionary investigations (Sites and Crandall 1997).

The lack of clearly visible diagnostic characters makes plant species delimitation challenging. Furthermore, many evolutionary processes can interfere with establishing species limits, including recent phylogenetic divergence, introgression, high phenotypic plasticity, ongoing differentiation, and barriers to gene flow (Duminil et al. 2012). Morphological markers have traditionally been used as the main characters for plant species delimitation, as phenotypic similarity has been the criterion used historically by taxonomists to group individuals into species. However, these markers may fail to differentiate between morphologically similar species or improperly subdivide species through inaccurate interpretation of natural phenotypic diversity throughout their distributions (Duminil and Di Michele 2009). Alternatively, molecular markers provide powerful tools for species delimitation based on phylogenetic methods, where monophyly and branch support represent the diagnostic criteria (Fazekas et al. 2008).

Calibrachoa Cerv. (Solanaceae) is a young group of South American species, mainly distributed in the subtropical grasslands between 18° and 37° S. The genus comprises approximately 26 recognised species, most of which are perennial herbs and small shrubs that inhabit open areas or, rarely, forest borders. Several are narrowly distributed, although some are found in overlapping regions (Greppi et al. 2013). All species in the genus are self-incompatible, except for *C. parviflora* (Juss.) D'Arcy, and diploid. Molecular

CONTACT Loreta B. Freitas loreta.freitas@ufrgs.br

phylogenetic analyses have split the genus into two subgenera (Fregonezi et al. 2012; Mäder and Freitas 2019), with most species belonging to *C.* subg. *Stimomphis* (Raf.) Stehmann, Fregonezi & Freitas. Morphological and genetic variability suggest that adaptive radiation has occurred within the genus since its origin < 4 Mya (Särkinen et al. 2013), with the main clades, which occupy diverse biogeographic areas, exhibiting phylogenetic niche conservatism (Mäder and Freitas 2019).

Within *Calibrachoa*, two recognised species, *C. thymifolia* (A. St.-Hil.) Stehmann & Semir (Figure 1B) and *C. heterophylla* (Sendtn.) Wijsman (Figure 1C), display purple corollas with a yellow fauce surrounded by a dark-purple ring. These species are found on sandy soils in open fields from the Pampa ecosystem. The Pampa region (Figure 1A) is an open and flat grassy formation in the southernmost region of Brazil, the entire Uruguayan territory, and the Pampa province in Argentina [corresponding to the Pampa Province as described in Morrone (2006)]. The region includes coastal fields along the South Atlantic coastal plain (SACP).

Calibrachoa thymifolia occurs exclusively at the margins of the Uruguay River basin, whereas *C. heterophylla* is found within this basin, predominantly in the SACP region of Rio Grande do Sul and Santa Catarina Brazilian states. The most recent and inclusive molecular phylogenetic analysis based on nuclear and plastid markers, revealed that the two species belong to different clades (Mäder

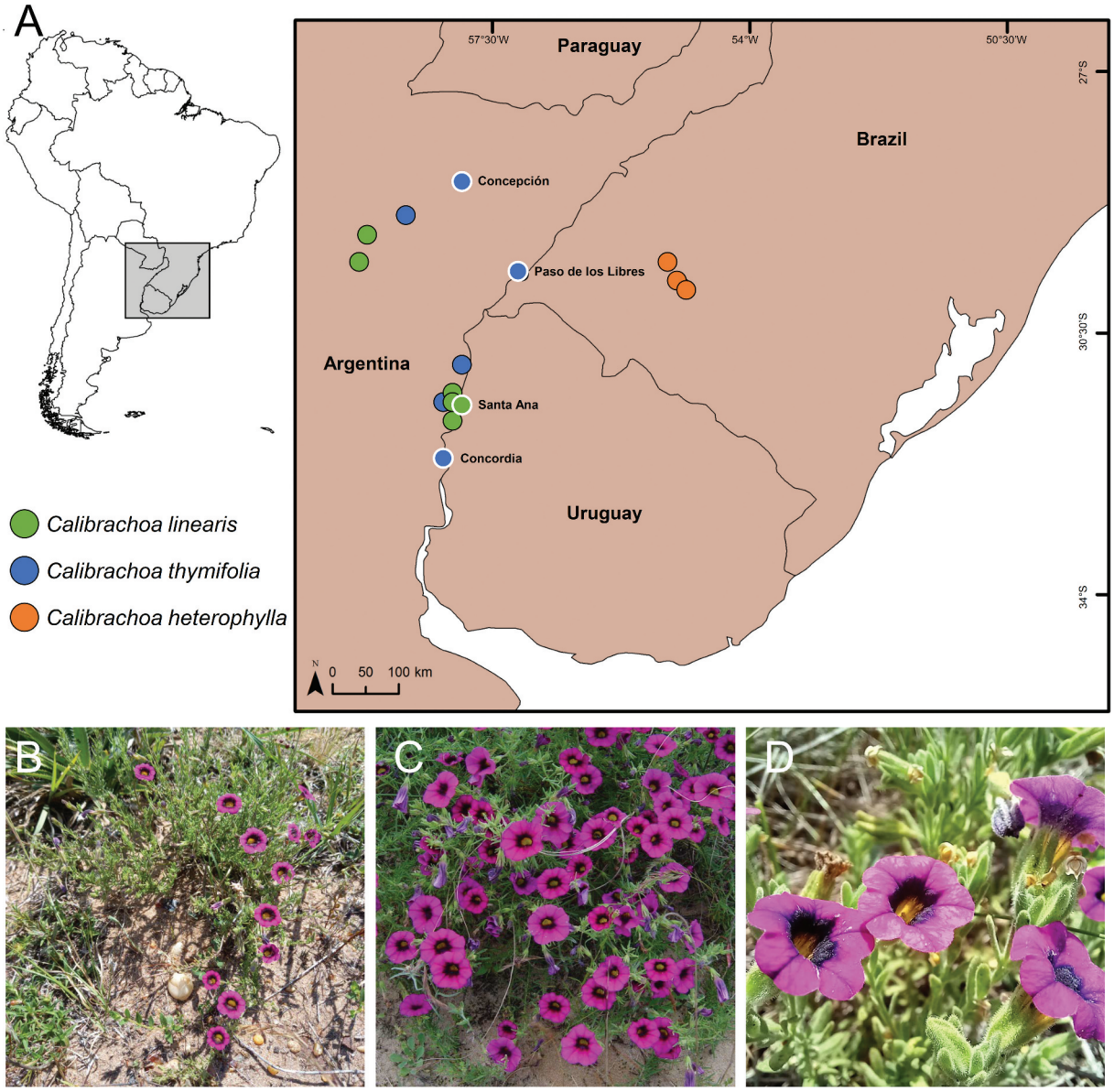

Figure 1. Map of sites from where samples of *Calibrachoa heterophylla* (orange), *C. thymifolia* (blue), and *C. linearis* (green) were collected (A). Samples from sites highlighted with a white outline were used in morphological analyses. Samples from the Concepción site were not used in molecular analyses. Flowers of *C. thymifolia* (B), *C. heterophylla* (C), and *C. linearis* (D). Photographic credit: authors' collection.

and Freitas 2019). They are distinguished by their stigma morphology with stigmas of *C. heterophylla* being very shortly bilobed and small, whereas those of *C. thymifolia* are mainly capitate. However, within *C. thymifolia* a second morphotype with truncated or under-capitated stigmas is present. This morphotype was previously recognised as an independent taxon, *C. linearis* (Hook.) Wijsman (Figure 1D) but was later synonymised as *C. thymifolia* (Greppi et al. 2013). It has been included previously in plastid-based phylogenetic analyses (Ando et al. 2005; Fregonezi et al. 2012, 2013), which showed it did not group with either *C. thymifolia* or *C. heterophylla*. Because of this, and for convenience, we revert to the former taxonomy and refer to this second morphotype of *C. thymifolia* as *C. linearis* hereafter.

To disentangle further the relationships between *C. thymifolia*, *C. linearis*, and *C. heterophylla,* we undertook a phylogeographic analysis of plastid DNA variation among several individuals of each taxon. We also screened plastid and nuclear DNA sequences for molecular diagnostic differences between *C. thymifolia* and *C. linearis*, and additionally examined morphological variation between the same two taxa. Finally, we modelled niche conservatism and niche overlap of all three taxa.

Materials and methods

Plant material and DNA extraction

We sampled multiple individuals of *C. linearis*, *C. thymifolia* and *C. heterophylla* species (Figure 1A; Table 1) from the Uruguay River basin. To optimise the sampling strategy, we took into account the size and number of populations of each taxon, ensuring that the sampling effort was proportional to their representation. Additionally, we considered the low intrapopulation diversity observed for the plastid markers in the target species (e.g. Backes et al. 2019; Barros et al. 2020). We recorded geographical coordinates using GPS (Global Positioning System) and one voucher per collection site (each representing a different population). DNA was extracted from silica-gel dried leaves using a CTAB (cetyl-trimethyl-ammonium Bromide; Sigma-Aldrich Co, St. Louis, USA)-based protocol (Roy et al. 1992), with DNA quality estimated using a Nanodrop DN-1000 Spectrophotometer (Thermo Fisher Scientific Co., Waltham, USA) and DNA quantity using a Qubit Fluorometer (Thermo Fisher).

Phylogeographic analysis based on plastid DNA sequence variation

Plastid DNA sequence variation among samples was examined by amplifying the plastid intergenic spacers of the trnH-psbA (Hamilton et al. 1999), trnS-trnG (Sang et al. 1997), rps12-rpl20 (Shaw et al. 2005), and part of the rpl32-trnL (Shaw et al. 2007) regions, using universal primers. For phylogenetic analysis, we also included some previously published sequences that employed the same markers as those utilised in this study (Mäder et al. 2013; Table S1). We followed previously published amplification protocols adapted for *Calibrachoa* species (Backes et al. 2019; Mäder and Freitas 2019) and verified PCR products by horizontal electrophoresis in 2.5% agarose gel dyed with GelRed™ (Biotium, Inc., Hayward, CA, USA). Products were purified using polyethylene glycol 20% (Dunn and Blattner 1987) before sequencing by capillary

Table 1. Sampling information for *Calibrachoa heterophylla*, *C. thymifolia* and *C. linearis* individuals included in molecular analyses from material collected in Brazil, Uruguay and Argentina.

Taxon	Collection site ID	Geographical coordinates	Voucher	Individuals (N)
C. linearis	Pop1	31°23′59.99″S, 58°4′59.98″W	BHCB 143,912	2
	Pop2	31°25′59.98″S, 58°3′59.97″W	BHCB 143,925	1
	Pop3	32°10′59.98″S, 58°9′59.97″W	BHCB 143,927	1
	Pop4	31°38′59.99″S, 58°1′59.98″W	BHCB 143,926	1
	Pop5	29°12′46.18″S, 59°13′12.86″W	NHN 1,737,224	1
	Pop6	29°34′33.77″S, 59°19′49.69″W	NA	7
C. thymifolia	Pop1	30°53′59.49″S, 57°55′52.03″W	BHCB 143,921	3
	Pop2	29°42′52.52″S, 57°50′50.53″W	NA	5
	Pop3	32°11′20.39″S, 58°10′26.97″W	BHCB 127,295	2
	Pop4	28°54′22.96″S, 58°39′41.97″W	NA	2
	Pop5	31°24′60.58″S, 58°60′50.97″W	NA	1
C. heterophylla	Pop1	29°34′59.05″S, 55°60′20.73″W	BHCB 102,097	2
	Pop2	29°53′40.88″S, 54°51′13.49″W	BHCB 117,016	3
	Pop3	29°51′17.20″S, 54°54′30.63″W	BHCB 117,021	2

BHCB, Universidade Federal de Minas Gerais Herbarium, Belo Horizonte, MG, Brazil; NHN, National Herbarium of the Netherlands, Leiden, The Netherlands; NA, not available.

electrophoresis on an ABI 3730XL genetic analyser (Thermo Fisher).

We assembled both forward and reverse strands for each plastid marker using the CHROMAS v.2.0 software (Technelysium, Helensvale, Australia), and sequences were deposited at GenBank (available at http//www.ncbi.nlm.nih.gov/genbank/-Table S1). Haplotypes were identified via DNASP v.5.10.01 (Rozas et al. 2003). DNA sequences were aligned using MEGA × (Kumar et al. 2018) with the CLUSTALW algorithm and manually edited when necessary. We coded contiguous insertion/deletion (indels) events involving more than one base pair (bp) as one mutational event (Simmons and Ochoterena 2000). We eliminated all indels that involved poly A/T because their homologies cannot be adequately accessed (Aldrich et al. 1988). We concatenated the four plastid intergenic regions and treated them as a single sequence in all analyses. ARLEQUIN v.3.5.2.2 (Excoffier and Lischer 2010) was used to estimate basic descriptive molecular diversity statistics, such as haplotype (h) and nucleotide (π) diversities per taxon, and NETWORK v.4.1.0.9 (Bandelt et al. 1999) was employed to estimate evolutionary relationships among haplotypes across all individuals and taxa.

We also estimated the genetic variability among taxa based on parsimoniously informative and polymorphic sites using DNASP, and determined phylogenetic relationships of the multilocus plastid DNA dataset through Bayesian inference (BI) implemented in BEAST v.1.10 (Suchard et al. 2018). Tree support was assessed with posterior probabilities (PP) with 10^7 chains. We selected the best substitution model and gamma rate heterogeneity using JMODELTEST v.3.06 (Darriba et al. 2012) based on the Akaike information criterion (AIC) for combined intergenic plastid spacer sequences and conducted BI analysis under the Yule process and two independent runs of 10 million generations, with sampling every 1000 generations. Markov chain Monte Carlo (MCMC) convergence was assessed by examining effective sample size values (ESS >200) and likelihood plots in TRACER v.1.7 (Rambaut et al. 2018). We discarded the initial 10% of trees as burn-in, and the remaining trees were summarised to generate a maximum clade credibility tree using TREEANNOTATOR v.1.7.5 (Suchard et al. 2018) visualised with FIGTREE v.1.4.1 (http://tree.bio.ed.ac.uk/software/figtree/). *Petunia integrifolia* was used as an outgroup (Table S1).

Finally, to obtain an overview of intra- and interspecific genetic variation, we calculated pairwise p-distances among individuals using the *dist.dna()* function in APE v.5.6–2 (Paradis et al. 2019) R package. The computed distances were visualised in box plots generated using the GGPLOT v.2 3.3.6 (Wickham 2016) R package. A Wilcoxon rank sum test assessed the statistical significance between intra- and interspecific variations using the *compare_means()* function implemented in the GGPUBR v.0.4.0 R package (available at https://rpkgs.datanovia.com/ggpubr/).

Molecular diagnosis

To identify molecular diagnostic sites, we compared genetic diversity among the three dark-ringed taxa by sequencing the nuclear markers *WUS*, *WOX1*, *SOE*, and *EVG* (Segatto et al. 2016) and the plastid regions *matK* (Johnson and Soltis 1994), trnL-trnF (Taberlet et al. 1991) and psbB-psbH (Hamilton et al. 1999) for *C. linearis* following amplification and sequencing previously published protocols. We also used available sequences for *C. thymifolia* and *C. heterophylla* (Mäder and Freitas 2019). All nuclear and plastid sequences for the three taxa were aligned using the CLUSTALW algorithm in MEGA X enabling identification of molecular diagnostic sites (Filipowicz and Renner 2012). Plastid marker entries included the analysed individuals, and only conserved sites among individuals of each taxon were compared among taxa. GenBank accession numbers for all sequences are available in Table S2, including those obtained from other works.

Ecological niche modelling and niche overlapping

We retrieved occurrence data for each taxon from the SpeciesLink database (http://www.splink.net/) and Global Biodiversity Information Facility (GBIF; https://www.gbif.org/species/2928904) and filtered the results for unique entries without taxonomic uncertainties and precise geographic coordinates. This resulted in 13 entries for *C. thymifolia*, 21 for *C. linearis*, and 51 for *C. heterophylla* (Table S3). We obtained 19 climatic variables from the WorldClim 2.1 website (Fick and Hijmans 2017) at 30 arc seconds (*c.* 1 km^2) resolution and calculated the correlation between variables to minimise collinearity problems. After discarding variables with more than 0.7 of correlation, we kept four climate variables (BIO3 = isothermality; BIO8 = mean temperature of the wettest quarter; BIO13 = precipitation of wettest month; and BIO15 =

precipitation seasonality) for use in ecological niche modelling (ENM).

We ran ENM per taxon in BIOMOD2 (Thuiller et al. 2009) R package using MAXENT v.3.4.4 (available at http://biodiversityinformatics.amnh.org/open_source/maxent/) maximum entropy algorithm with 10 replicates, maximum 5,000 iterations 10,000 background points, and 70% of data used as training data. The best parameters for each model were selected in the ENMEVAL (Muscarella et al. 2014) R package. We evaluated the quality of the models based on the area under the curve (AUC; Pearce and Ferrier 2000) and calculated the environmental niche overlap among the three taxa through Schoener's *D* (Schoener 1968), the *I* statistic (Warren et al. 2008), and relative rank (RR; Warren and Seifert 2011) indices in ENMTOOLS (Warren et al. 2021) R package. Statistical significance was evaluated by comparing niche overlap results to the null hypothesis calculated by identity tests in ENMTOOLS.

Morphological diversity between *C. thymifolia* and *C. linearis*

To evaluate the morphological diversity between *C. thymifolia* and *C. linearis*, we examined 30 flowers of different individuals across both taxa (Table S4). Individuals of *C. thymifolia* came from three localities (Paso de los Libres, Santa Ana and Concepción) while those of *C. linearis* were from Concordia. We used digital callipers to measure two floral (corolla diameter and tube length) and two foliar traits (leaf length and width; Figure S1). Based on these data, we conducted a Discriminant Analysis of Principal Components (DAPC; Jombart et al. 2010) using the ADEGENET package (Jombart 2008) in R v.3.5.1 (R Core Team 2017, available at https://www.r-project.org/) and RStudio Desktop v.1.1.383 (RStudio Team 2016, available at http://www.rstud io.com/).

To assess the normality of data and homogeneity of variance, we conducted the Shapiro–Wilk test with the STATS R package and the Levene test with the CAR R package (Fox and Weisberg 2019), respectively. When the data met the assumptions of normality and homogeneity of variance, we performed a one-way analysis of variance (ANOVA; Welch 1951) to identify group differences. However, for datasets that did not meet these assumptions we used the Kruskal–Wallis test (Kruskal and Wallis 1952). Finally, we conducted a Dunn test with the DUNN.TEST in R to determine which groups differed, applying the Bonferroni correction for multiple comparisons.

Results

Genetic diversity and evolutionary relationships

The plastid intergenic spacers for 13 individuals of *C. thymifolia*, 13 individuals of *C. linearis*, and seven individuals of *C. heterophylla* resulted in a concatenated alignment 3,109 bp long (770 bp corresponding to trnS-trnG, 462 bp to trnH-psbA, 1,019 bp to rpl32-trnL, and 858 bp to rps12-rpl20). Sequences displayed 28 polymorphic sites (nine transitions, 13 transversions, and six one bp indels), 15 of which were parsimoniously informative, resulting in 12 haplotypes (see Supplementary Material, Table S1 for haplotype frequency and distribution). GC content was 31.1% in trnS-trnG, 27.5% in trnH-psbA, 27.2% in rpl32-trnL, and 32.3% in rps12-rpl20. The best substitution model obtained using JMODELTEST for the combined plastid markers was GTR+I.

Evolutionary relationships among the 12 haplotypes (Figure 2) revealed that *C. thymifolia* and *C. linearis* shared three haplotypes that were not present in *C. heterophylla* which contained seven exclusive haplotypes. The most frequent haplotype was H3 (36% of individuals), shared by individuals of *C. thymifolia* and *C. linearis*. Two to seven evolutionary steps separated haplotypes of *C. thymifolia* and *C. linearis*, whereas at least 14 mutations separated these haplotypes from those of *C. heterophylla*. Nucleotide diversity was higher in *C. linearis*, whereas haplotype diversity was similar in both *C. linearis* and *C. thymifolia* and lower than in *C. heterophylla* (Table 2).

The Bayesian tree comparing all plastid haplotypes (Figure 3) contained two main clades (PP = 1). One clade grouped all *C. heterophylla* haplotypes, whereas the other comprised the haplotypes of *C. thymifolia* and *C. linearis*. There was low support for most internal branches within each of these two clades.

A comparison of genetic distance among individuals showed that pairwise-distances were lower than 1% for the global alignment. Regarding intraspecific distances, *C. heterophylla* individuals exhibited higher genetic variation than either *C. linearis* or *C. thymifolia* (Figure 4). Additionally, the mean genetic distance between *C. linearis* individuals was significantly lower than that between *C. heterophylla*

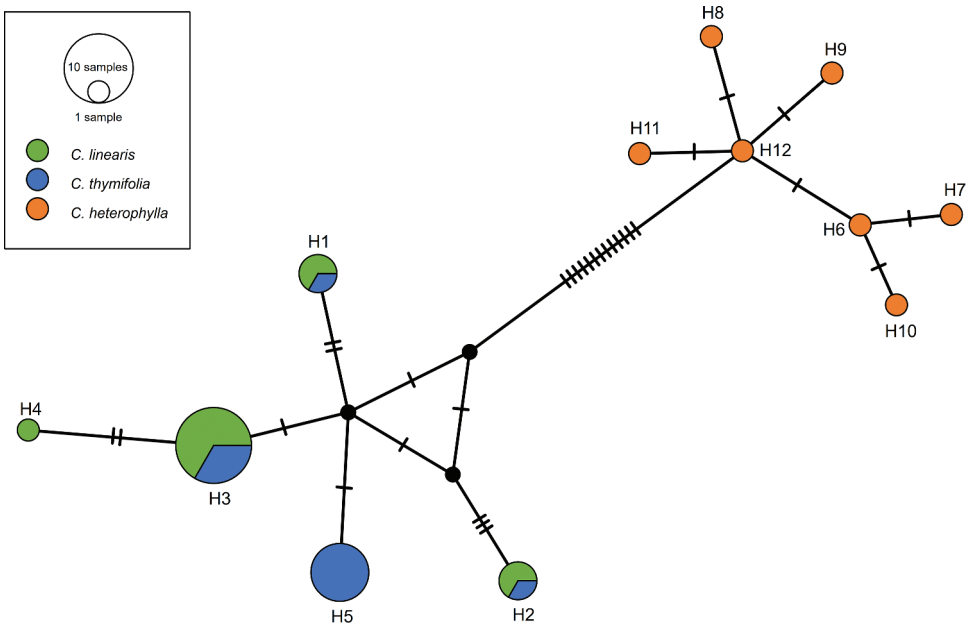

Figure 2. Evolutionary relationships of plastid haplotypes of *Calibrachoa heterophylla*, *C. thymifolia* and *C. linearis*. Haplotypes are sequentially numbered and indicated in colours according to the legend; the size of circles is proportional to haplotype frequency. Perpendicular bars indicate the evolutionary steps between haplotypes.

Table 2. Genetic diversity of *Calibrachoa heterophylla*, *C. thymifolia* and *C. linearis* based on plastid information.

Taxon	Individuals (N)	Haplotypes (N)	Nucleotide diversity π % (SD)	Haplotype diversity (SD)
C. linearis	13	4	0.08 (0.05)	0.62 (0.14)
C. thymifolia	13	4	0.06 (0.04)	0.65 (0.11)
C. heterophylla	7	7	0.06 (0.05)	1.00 (0.08)

N, number; SD, standard deviation.

and *C. linearis* ($P < 0.0001$), and similarly, *C. thymifolia* intraspecific mean distance was significantly lower than that between *C. heterophylla* and *C. thymifolia* ($P < 0.0001$). The remaining comparisons between intra- and interspecific mean distances were not statistically significant ($P > 0.05$). Thus, genetic distances were low and of a similar magnitude both within and between *C. linearis* and *C. thymifolia* (Figure 4), and only *C. heterophylla* could be clearly distinguished based on these pairwise-distance comparisons.

Molecular diagnosis

The sequence alignment for all plastid and nuclear markers in the three species was 7,117 bp long, of which ~10% of sites varied among species. *Calibrachoa linearis* and *C. thymifolia* differed for ~68% of polymorphic sites, whereas *C. heterophylla* was divergent for ~36% of polymorphic sites compared with *C. thymifolia* and ~94% compared with *C. linearis*. The most variable markers among the species were *WOX1* and rps12-rpl32, which exhibited long indel differences (6 bp and 14 bp, respectively), whereas the trnL-trnF, *SOE*, and *EVG* sequences did not vary among species. Considering only *C. thymifolia* and *C. linearis* (Supplementary Material, Table S5), we found 49 differences in the sequence between them, of which 14 were point mutations (diagnostic sites), four were multiple-base indels, and eight were single-base indels.

Ecological differentiation

The ENM analysis showed that the variable contributing most to the niche models was BIO3 (isothermality), with contributions of 51% for *C. linearis*, 78% for *C. thymifolia*, and 94% for *C. heterophylla*. The mean of AUC values was 0.91 (±0.03) for *C. linearis*, 0.82 (±0.07) for *C. thymifolia* and 0.95 (±0.03) for *C. heterophylla*, indicating high predictive power for these taxa. The ENMs exhibited a similar general pattern of suitable conditions for *C. linearis* and *C. thymifolia*, whereas *C. heterophylla* was more restricted to the ocean coast (Figure 5). Thus, a large overlap of suitable environmental area was indicated for *C. thymifolia*

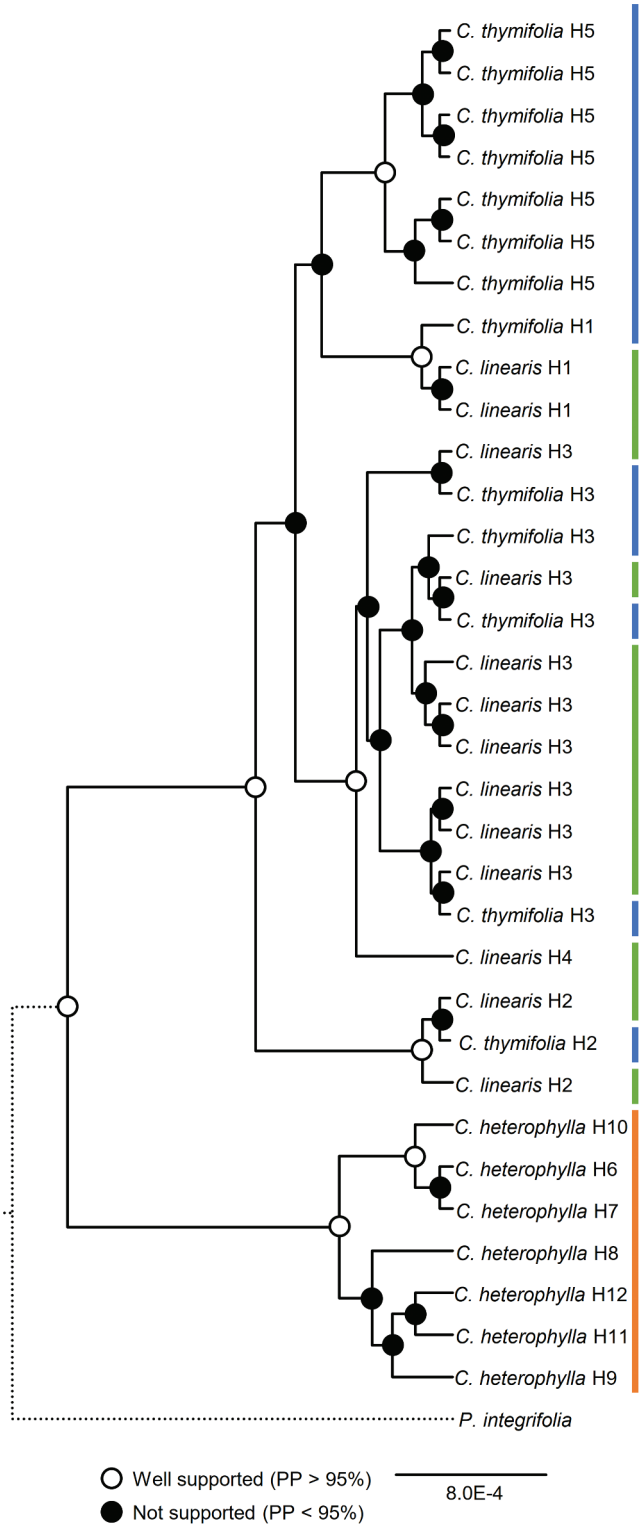

Figure 3. Phylogenetic tree obtained with a Bayesian inference for individuals based on plastid haplotypes. Vertical coloured lines indicate the taxa, *C. heterophylla* (orange), *C. thymifolia* (blue), and *C. linearis* (green). Branch supports are indicated as dark (PP > 0.95) or light (PP < 0.95) circles.

and *C. linearis*, with Schoener's $D = 0.69$ and $I = 0.78$ scores, whereas less niche overlap was indicated between these two taxa and *C. heterophylla* (Schoener's $D = 0.04$ and $I = 0.07$ for *C. heterophylla vs. C. linearis*, and Schoener's $D = 0.12$ and $I = 0.23$ for *C. heterophylla vs. C. thymifolia*). Identity tests indicated equivalence in niche suitability for *C. linearis* and *C. thymifolia*, as results were not significant for D and I statistics (both with $P = 0.75$). However, despite their similarity, these two species differed in suitability intensity of each distribution area. In contrast, identity tests revealed that *C. heterophylla* is niche differentiated compared to the other two taxa ($P < 0.01$ for all tests).

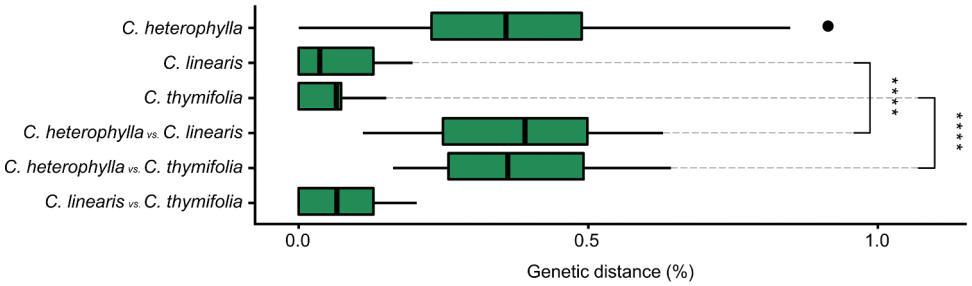

Figure 4. Overview of genetic variation among individuals, represented as boxplots of intraspecific and interspecific genetic pairwise distances for the plastid global alignment. Boxplots: vertical line, mean distance; box limits, 25th and 75th percentiles; whiskers, 1.5 times interquartile range; dots, outliers. Asterisks denote significant difference between the depicted comparisons ($P < 0.0001$).

Morphological variability between *C. thymifolia* and *C. linearis*

Considering the floral and foliar traits separately, we retained two principal components (100% of variance) and two discriminant components for the Discriminant Analysis of Principal Components (DAPC) of four *a priori* groups, with groups representing the four different localities from where plants were sampled. The DAPC analysis revealed that based on floral traits (Figure 6A), the groups exhibited no discernible differences. However, for foliar traits (Figure 6B), three groups could be distinguished, one corresponding to *C. linearis* (locality Concordia), another to the morphotype of *C. thymifolia* (localities Paso de los Libres and Santa Ana), and a third (from locality Concepción) that was not included in the molecular analyses presented here or previously.

The results showed that only corolla diameter data passed both the normality and homogeneity of variance tests. An ANOVA of these data found no significant differences among the groups ($P = 0.29$). For the other three measurements, a Kruskal–Wallis test showed no significant difference for the corolla tube length ($P = 0.13$), but a significant difference between groups for leaf length and width ($P < 0.0001$ for both). A Dunn test further revealed that groups could be distinguished by leaf length and width differences (Figure 7).

Discussion

Given the pending biodiversity crisis, increased efficacy in species delimitation is critically essential in biology (Wheeler et al. 2004). Under traditional taxonomic practices, species' discovery, delineation, and description often involve qualitative decisions on what a species should be and are thus subjected to implementing various philosophical species concepts (de Queiroz 2007). Here, we analysed three taxa that occur in sympatry and share several key morphological traits, leading to some taxonomic doubts, at least regarding two.

Our results confirmed that *Calibrachoa heterophylla* is an independent taxon exhibiting low genetic polymorphism sharing with two related species, *C. thymifolia* and *C. linearis*, as previously proposed in molecular studies (Fregonezi et al. 2012; Mäder and Freitas 2019). ENM and niche overlapping analyses reinforced the difference between *C. heterophylla* and the two other species, and, except for the dark ring in the corolla throat aperture (Figure 1), *C. heterophylla* is easily distinguished morphologically.

Regarding *C. thymifolia* and *C. linearis*, both taxa formed a group with no population structure based on multi-individual plastid sequences (Figure 3), with the sharing of some haplotypes evident (Figure 2). Furthermore, floral traits failed to distinguish the two taxa (Figures 6A and 7), while ENM (Figure 5) showed them to have overlapping suitable areas, despite some differences in extension and intensity of these areas. Although these results support the synonymisation of both taxa as previously proposed (Greppi et al. 2013), we found that they could be distinguished for leaf traits (Figures 6B and 7) and also at 49 sites across the combined sequences of several plastids and nuclear markers. Based on these differences, we propose they should be considered as independent taxa. Interestingly, our morphological analysis indicated that *C. thymifolia* might comprise an additional morphotype (from Concepción) to the standard morphotype (found at localities Paso de los Libres and Santa Ana). This additional morphotype was not subjected to molecular analysis but should be in the future.

In early stages of speciation, only a subset of genes diverges between taxa, which may confer

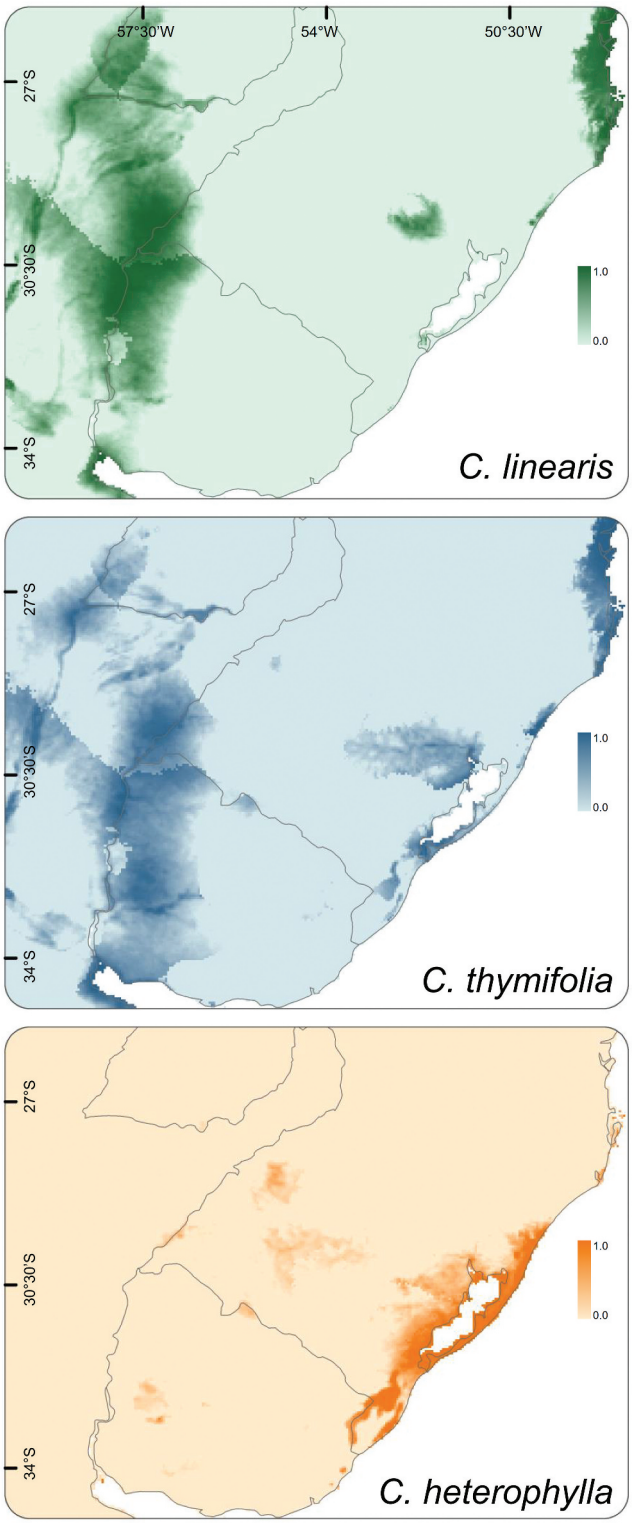

Figure 5. Environmental habitat suitability of *Calibrachoa heterophylla*, *C. thymifolia* and *C. linearis* estimated by MAXENT in biomod 2. Suitability is proportional to the heat map, where 1 indicates high suitability and 0 indicates low suitability, as shown in the colour gradient bar.

some specific advantages to the species, while the remaining part of the genome is freely homogenised through hybridisation (Shapiro et al. 2016; Becher et al. 2022). Strong selection on a limited number of genes, despite interspecific gene flow, can lead to rapid speciation, as appears to have occurred in *Calibrachoa* (Mäder and Freitas 2019) and the related genus, *Petunia* (Reck-Kortmann et al. 2014). Species divergence can be achieved through adaptation to microhabitats, as evidenced in the *Petunia* short corolla tube clade (Fregonezi et al. 2013), where most species exhibit highly similar flower morphology but occupy distinctive microhabitats (Segatto et al. 2017).

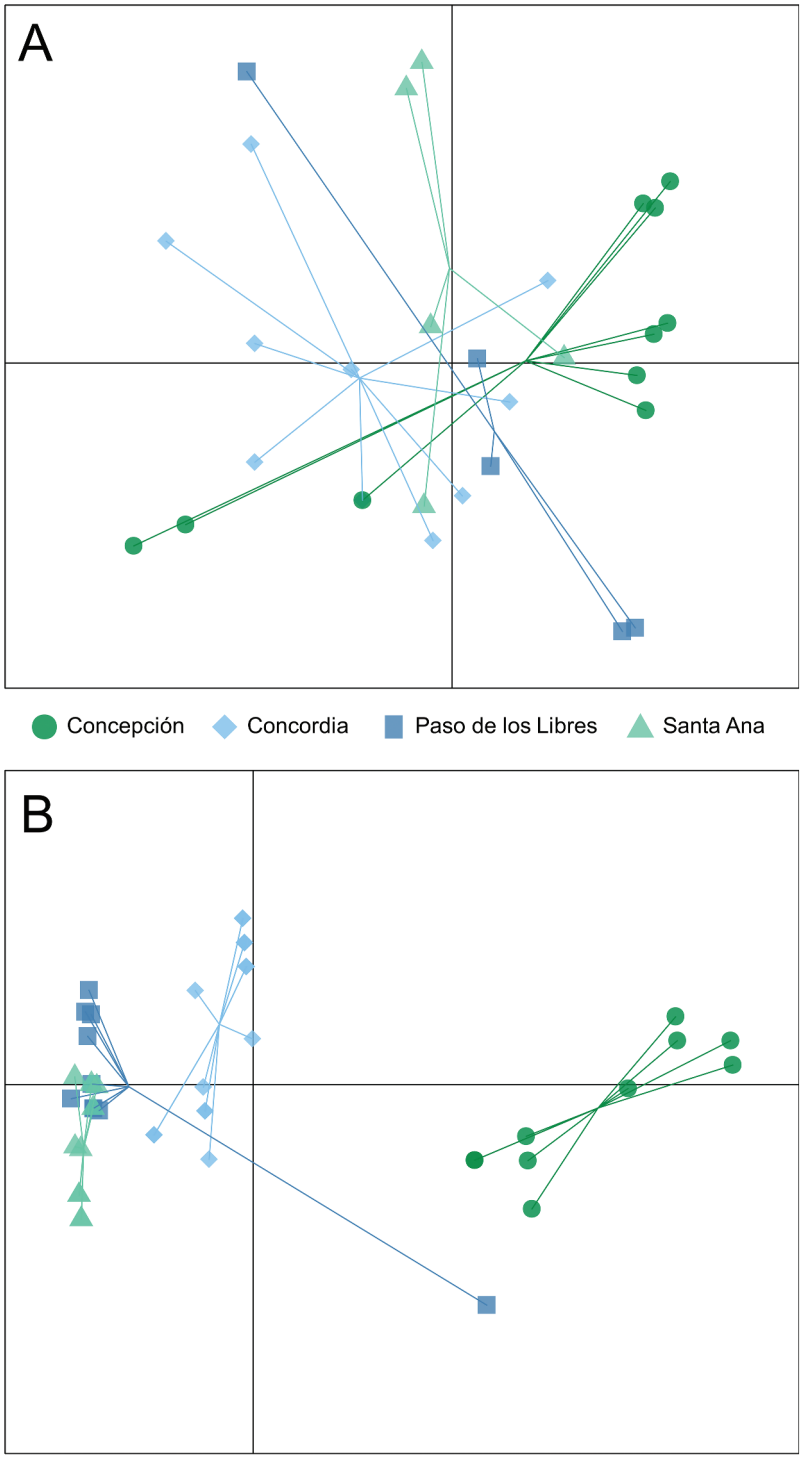

Figure 6. Cartesian plot obtained with discriminant analysis of principal components based on morphological variability in flowers (A) and leaves (B) of *Calibrachoa thymifolia* (localities Paso de los Libres and Santa Ana) and *C. linearis* (Concordia locality). Samples from the Concepción locality were assigned to the *C. thymifolia* complex according to the classification of Greppi et al. (2013) but were not included in the molecular analyses presented here or previously.

It has been shown in other Solanaceae groups in South America [e.g. *Nierembergia* Ruiz and Parv (Tate et al. 2009), *Brunfelsia* L (Filipowicz and Renner 2012), *Petunia* Juss. (Reck-Kortmann et al. 2014)] that rapid adaptive radiation has led to species diversification, though often with blurred morphological limits present between species. This also seems to be the case in *Calibrachoa* which originated < 4 Mya (Särkinen et al. 2013). It is likely that both *Petunia* and *Calibrachoa* were strongly impacted by climatic shifts during the Quaternary (Fregonezi et al. 2013), triggering speciation as in many other plant genera (Kadereit and Abbott 2021). *Calibrachoa* species generally occur in grasslands. These small shrubs and perennial herbs dispersed and diversified as did other open field-adapted species during the Pleistocene climatic cycles

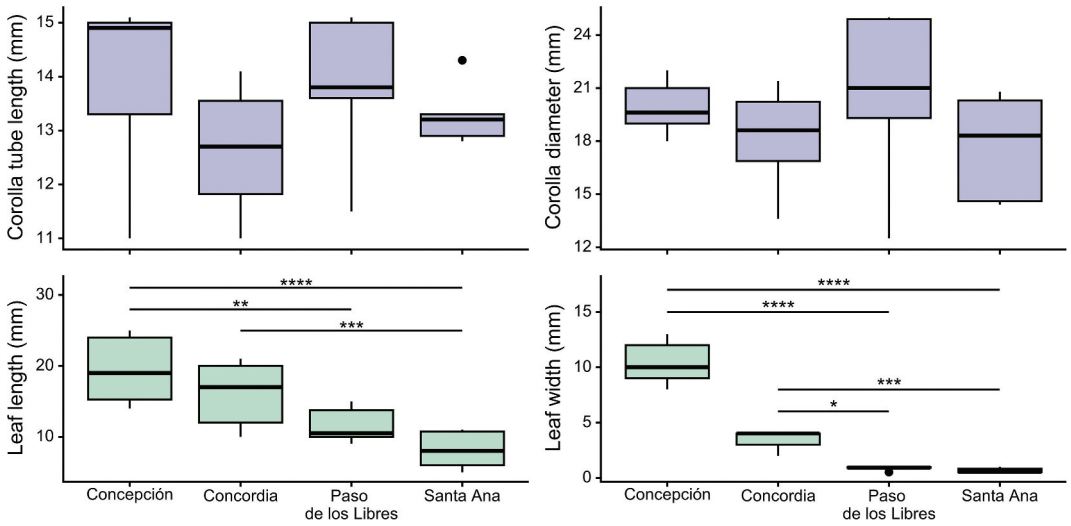

Figure 7. Comparison of floral and foliar characteristics of individuals of *C. thymifolia* (from localities Paso de los Libres and Santa Ana) and *C. linearis* (from Concordia locality). Samples from Concepción were assigned to the *C. thymifolia* complex based on their morphology in the wild and classification of Greppi et al. (2013). Boxplots: centre line, median; box limits, 25th and 75th percentiles; whiskers, 1.5 times interquartile range; dots, outliers.

(Mäder and Freitas 2019), with range expansions occurring during glacial periods and contractions during interglacials (Behling 2002).

Closely related and young taxa, such as *C. thymifolia* and *C. linearis*, frequently show morphological similarities and low levels of genetic differentiation due to their evolutionary proximity (Segatto et al. 2017). Many morphs or ecotypes are associated with geographically distant specific habitats, where reproductive isolation results from physical separation and local adaptation (Abbott and Comes 2007). However, our ENM results and niche overlapping analyses indicated that *C. thymifolia* and *C. linearis* occupy almost the same climate niche, suggesting they did not diverge in response to climate differences between the habitats they occupy. Also, their close occurrence, at least at some locations (Figure 1), suggests that geographical distance would not be a factor preventing gene flow occurring between them, at least currently. Before ruling out ecological isolation as a cause of their divergence, however, more detailed investigation is required of possible abiotic and biotic differences in their respective habitats, which they might be differentially adapted to.

Conclusions

A frequently observed pattern among *Calibrachoa* species is the absence of reciprocal monophyly and occurrence of interspecific plastid haplotype sharing (e.g. Fregonezi et al. 2013; Backes et al. 2019; John et al. 2019; Barros et al. 2020). Our molecular and morphological results, however, suggest that the synonymisation of *C. linearis* and *C. thymifolia* (Greppi et al. 2013) needs to be revisited, with *C. linearis* considered a valid species. Additional molecular markers and other approaches should be used to solve the taxonomic issue regarding these two taxa and the possible factors important in their divergence.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico Coordenação de Aperfeiçoamento de Pessoal de Nível Superior .

ORCID

Alice Backes http://orcid.org/0000-0001-8535-5692
Pedro H. Pezzi http://orcid.org/0000-0003-4324-5078
Leonardo T. Gonçalves http://orcid.org/0000-0002-3593-8424
Loreta B. Freitas http://orcid.org/0000-0002-0728-4945

References

Abbott RJ, Comes HP. 2007. Blowin' in the wind – the transition from ecotype to species. New Phytol. 175 (2):197–200. doi:10.1111/j.1469-8137.2007.02127.x.

Aldrich J, Cherney BW, Merlin E, Christopherson L. 1988. The role of insertion/deletions in the evolution of the

intergenic region between *psbA* and *trnH* in the chloroplast genome. Curr Genet. 14(2):137–146. doi:10.1007/BF00569337.

Ando T, Kokubun H, Watanabe H, Tanaka N, Yukawa T, Hashimoto G, Marchesi E, Suárez E, Basualdo I. 2005. Phylogenetic analysis of *Petunia* sensu jussieu (Solanaceae) using chloroplast DNA RFLP. Ann Bot. 96 (2):289–297. doi:10.1093/aob/mci177.

Backes A, Mäder G, Turchetto C, Segatto ALA, Fregonezi JN, Bonatto SL, Freitas LB. 2019. How diverse can rare species be on the margins of genera distribution? AoB Plants. 11(4):lz037. doi:10.1093/aobpla/plz037.

Bandelt HJ, Forster P, Röhl A. 1999. Median joining networks for inferring intraspecific phylogenies. Mol Biol Evol. 16 (1):37–48. doi:10.1093/oxfordjournals.molbev.a026036.

Barros MJF, Silva-Arias GA, Segatto ALA, Reck-Kortmann M, Fregonezi JN, Diniz-Filho JAF, Freitas LB. 2020. Phylogenetic niche conservatism and plant diversification in South American subtropical grasslands along multiple climatic dimensions. Genet Mol Biol. 43(2):e20180291. doi:10.1590/1678-4685-gmb-2018-0291.

Barton NH, Gale K. 1993. Genetic analysis of hybrid zones. In: Harrison R, editor Hybrid zones and the evolutionary process. Oxford: Oxford University Press; pp. 13–45.

Becher H, Sampson J, Twyford AD. 2022. Measuring the invisible: the sequences causal of genome size differences in eyebrights (Euphrasia) revealed by k-mers. Front Plant Sci. 13:818410. doi: 10.3389/fpls.2022.818410.

Behling H. 2002. South and southeast Brazilian grasslands during late quaternary times: a synthesis. Palaeogeogr Palaeoclimatol Palaeoecol. 177(1–2):19–27. doi:10.1016/S0031-0182(01)00349-2.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. JModelTest 2: more models, new heuristics and parallel computing. Nat Methods. 9(8):772. doi:10.1038/nmeth.2109.

de Queiroz K. 2007. Species concepts and species delimitation. Syst Biol. 56(6):879–886. doi:10.1080/10635150701701083.

Duminil J, Di Michele M. 2009. Plant species delimitation: a comparison of morphological and molecular markers. Plant Biosyst. 143(3):528–542. doi:10.1080/11263500902722964.

Duminil J, Kenfack D, Viscosi V, Grumiau L, Hardy OJ. 2012. Testing species delimitation in sympatric species complexes: the case of an African tropical tree, *carapa* spp. (meliaceae). Mol Phylogen Evol. 62(1):275–285. doi:10.1016/j.ympev.2011.09.020.

Dunn IS, Blattner FR. 1987. Charons 36 to 40: multi–enzyme, high capacity, recombination deficient replacement vectors with polylinkers and polystuffers. Nucl Acids Res. 15(6):2677–2698. doi:10.1093/nar/15.6.2677.

Excoffier L, Lischer HEL. 2010. Arlequin suite ver. 3.5: a new series of programs to perform population genetics analyses under linux and windows. Mol Ecol Res. 10 (3):564–567. doi:10.1111/j.1755-0998.2010.02847.x.

Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, Newmaster SG, Husband BC, Percy DM, Hajibabaei M, Barrett SCH, DeSalle R. 2008. Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. PloS One. 3(7):e2802. doi:10.1371/journal.pone.0002802.

Fick SE, Hijmans RJ. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int J Climatol. 37(12):4302–4315. doi:10.1002/joc.5086.

Filipowicz N, Renner SS. 2012. *Brunfelsia* (Solanaceae): a genus evenly divided between South America and radiations on Cuba and other Antillean islands. Mol Phylogen Evol. 64(1):1–11. doi:10.1016/j.ympev.2012.02.026.

Fox J, Weisberg S. 2019. An R companion to applied regression. 3rd. ed. Thousand Oaks: Sage; p. 472.

Fregonezi JN, Freitas LB, Bonatto SL, Semir J, Stehmann JR. 2012. Infrageneric classification of *Calibrachoa* (Solanaceae) based on morphological and molecular evidence. Taxon. 61(1):120–130. doi:10.1002/tax.611009.

Fregonezi JN, Turchetto C, Bonatto SL, Freitas LB. 2013. Biogeographical history and diversification of *Petunia* and *Calibrachoa* (Solanaceae) in the neotropical pampas grassland. Bot J Linn Soc. 171(1):140–153. doi:10.1111/j.1095-8339.2012.01292.x.

Greppi JA, Hagiwara JC, Stehmann JR. 2013. Novedades en *Calibrachoa* (Solanaceae) y notas taxonómicas sobre el género para la Argentina. Darwiniana, nueva serie. 1 (1):173–187. doi:10.14522/darwiniana.2013.11.489.

Hamilton MB, Shimoda C, Nishiumi I, Eguchi K, Yamagishi S. 1999. Isolation of microsatellite loci for paternity testing in the rufous vanga schetba rufa. Mol Ecol. 8 (3):513–514. doi:10.1046/j.1365-294x.1999.00510.x.

John ALW, Mäder G, Fregonezi JN, Freitas LB. 2019. Genetic diversity and population structure of naturally rare *Calibrachoa* species with small distribution in southern Brazil. Genet Mol Biol. 42(1):108–119. doi:10.1590/1678-4685-gmb-2017-0314.

Johnson LA, Soltis DE. 1994. *matK* DNA sequences and phylogenetic reconstruction in Saxifragaceae. s str null. 19(1):143–156. doi:10.2307/2419718.

Jombart T. 2008. Adegenet a R package for the multivariate analysis of genetic markers. Bioinformatics. 24 (11):1403–1405. doi:10.1093/bioinformatics/btn129.

Jombart T, Devillard S, Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 11(1):94. doi:10.1186/1471-2156-11-94.

Kadereit JW, Abbott RJ. 2021. Plant speciation in the quaternary. Plant Ecol Divers. 14(3–4):105–142. doi:10.1080/17550874.2021.2012849.

Kruskal WH, Wallis WA. 1952. Use of ranks in one-criterion variance analysis. J Am Stat Assoc. 47 (260):583–621. doi:10.1080/01621459.1952.10483441.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K, Battistuzzi FU. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 35(6):1547–1549. doi:10.1093/molbev/msy096.

Li X, Wiens JJEstimating global biodiversity: the role of cryptic insect speciesSyst Biol2022[accessed 2023 Apr 14692391–40310.1093/sysbio/syac069]

Mäder G, Fregonezi JN, Lorenz-Lemke AP, Bonatto SL, Freitas LB. 2013. Geological and climatic changes in quaternary shaped the evolutionary history of *Calibrachoa heterophylla*, an endemic South-Atlantic species of petunia. BMC Evol Biol. 13(1):178. doi:10.1186/1471-2148-13-178.

Mäder G, Freitas LB. 2019. Biogeographical, ecological, and phylogenetic analyses clarifying the evolutionary history of *Calibrachoa* in South American grasslands. Mol Phylogen Evol. 141:106614. doi: 10.1016/j.ympev.2019.106614.

Morrone JJ. 2006. Biogeographic areas and transition zones of Latin America and the Caribbean Islands based on panbiogeographic and cladistic analyses of the entomofauna. Annu Rev Entomol. 51(1):467–494. doi:10.1146/annurev.ento.50.071803.130447.

Muscarella R, Galante PJ, Soley-Guardia M, Boria RA, Kass JM, Uriarte M, Anderson RP, McPherson J. 2014. ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. Methods Ecol Evol. 5(11):1198–1205. doi:10.1111/2041-210X.12261.

Paradis E, Schliep K, Schwartz R. 2019. Ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics. 35(3):526–528. doi:10.1093/bioinformatics/bty633.

Pearce J, Ferrier S. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. Ecol Modell. 133(3):225–245. doi:10.1016/S0304-3800(00)00322-7.

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. Syst Biol. 67(5):901–904. doi:10.1093/sysbio/syy032.

Reck-Kortmann M, Silva-Arias GA, Segatto ALA, Mäder G, Bonatto SL, Freitas LB. 2014. Multilocus phylogeny reconstruction: new insights into the evolutionary history of the genus *Petunia*. Mol Phylogen Evol. 81:19–28. doi: 10.1016/j.ympev.2014.08.022.

Roy A, Frascaria N, MacKay J, Bousquet J. 1992. Segregating random amplified polymorphic DNAs (RAPDs) in *Betula alleghaniensis*. Theor Appl Genet. 85(2–3):173–180. doi:10.1007/BF00222856.

Rozas J, Sánchez-Del Barrio JC, Messeguer X, Rozas R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics. 19:2496–2497. doi: 10.1093/bioinformatics/btg359.

Sang T, Crawford DJ, Stuessy TF. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). Am J Bot. 84(8):1120–1136. doi:10.2307/2446155.

Särkinen T, Bohs L, Olmstead RG, Knapp S. 2013. A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): a dated 1000-tip tree. BMC Evol Biol. 13(1):214. doi:10.1186/1471-2148-13-214.

Schoener TW. 1968. The *anolis* lizards of Bimini: resource partitioning in a complex fauna. Ecology. 49(4):704–726. doi:10.2307/1935534.

Segatto ALA, Reck-Kortmann M, Turchetto C, Freitas LB. 2017. Multiple markers, niche modelling, and bioregions analyses to evaluate the genetic diversity of a plant species complex. BMC Evol Biol. 17(1):234. doi:10.1186/s12862-017-1084-y.

Segatto ALA, Thompson CE, Freitas LB. 2016. Contribution of WUSCHEL-related homeobox (*WOX*) genes to identify the phylogenetic relationships among *Petunia* species. Genet Mol Biol. 39(4):658–664. doi:10.1590/1678-4685-gmb-2016-0073.

Shapiro BJ, Leducq JB, Mallet J, Matic I. 2016. What is speciation? PLoS Genet. 12(3):e1005860. doi:10.1371/journal.pgen.1005860.

Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL. 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Am J Bot. 92(1):142–166. doi:10.3732/ajb.92.1.142.

Shaw J, Lickey EB, Shilling EE, Small RL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in Angiosperms: the tortoise and the hare III. Am J Bot. 94(3):275–288. doi:10.3732/ajb.94.3.275.

Simmons MP, Ochoterena H. 2000. Gaps as characters in sequence based phylogenetic analyses. Syst Biol. 49(2):369–381. doi:10.1093/sysbio/49.2.369.

Sites JW, Crandall KA. 1997. Testing species boundaries in biodiversity studies. Conserv Biol. 11(6):1289–1297. doi:10.1046/j.1523-1739.1997.96254.x.

Sites JW, Marshall JC. 2003. Delimiting species: A renaissance issue in systematic biology. Trends Ecol Evol. 18(9):462–470. doi:10.1016/S0169-5347(03)00184-8.

Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol. 4(16). vey. doi: 10.1093/ve/vey016.

Taberlet P, Gielly L, Pautou G, Bouvet J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol. 17(5):1105–1109. doi:10.1007/BF00037152.

Tate JA, Acosta MC, McDill J, Moscone EA, Simpson BB, Cocucci AA. 2009. Phylogeny and character evolution in *Nierembergia* (Solanaceae): molecular, morphological, and cytogenetic evidence. null. 34(1):198–206. doi:10.1600/036364409787602249.

Thuiller W, Lafourcade B, Engler R, Araújo MB. 2009. BIOMOD – a platform for ensemble forecasting of species distributions. Ecography. 32(3):369–373. doi:10.1111/j.1600-0587.2008.05742.x.

Warren DL, Glor RE, Turelli M. 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution. 62(11):2868–2883. doi:10.1111/j.1558-5646.2008.00482.x.

Warren DL, Matzke NJ, Cardillo M, Baumgartner JB, Beaumont LJ, Turelli M, Glor RE, Huron NA, Simões M, Iglesias TL, et al. 2021. Enmtools 1.0: an R package for comparative ecological biogeography. Ecography. 44(4):504–511. doi:10.1111/ecog.05485.

Warren DL, Seifert SN. 2011. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. Ecol Appl. 21(2):335–342. doi:10.1890/10-1171.1.

Welch BL. 1951. On the comparison of several mean values: an alternative approach. Biometrika. 38(3–4):330–336. doi:10.1093/biomet/38.3-4.330.

Wheeler QD, Raven PH, Wilson EO. 2004. Taxonomy: impediment or expedient? Sci. 303(5656):285. doi:10.1126/science.303.5656.285.

Wickham H. 2016. Elegant graphics for data analysis. New York (NY): Springer-Verlag. ggplot2.