UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
Departamento de Botânica
Programa de Pós-Graduação em Biologia Vegetal

PPGBV
U F M G

UFMG

CECILIA FONSECA FIORINI

EVOLUTIONARY PROCESSES AND THE ORIGIN OF PLANT BIODIVERSITY IN OLD NEOTROPICAL SKY-ISLANDS:
A GENOMIC AND MODEL-BASED APPROACH

Tese apresentada ao Programa de Pós-Graduação em Biologia Vegetal do Departamento de Botânica do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Doutora em Biologia Vegetal.

Área de Concentração: Morfologia, sistemática e diversidade vegetal

BELO HORIZONTE – MG

2020

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
Departamento de Botânica
Programa de Pós-Graduação em Biologia Vegetal

PPGBV
U F M G

UFMG

CECILIA FONSECA FIORINI

EVOLUTIONARY PROCESSES AND THE ORIGIN OF PLANT BIODIVERSITY IN OLD NEOTROPICAL SKY-ISLANDS:
A GENOMIC AND MODEL-BASED APPROACH

Versão final

Tese apresentada ao Programa de Pós-Graduação em Biologia Vegetal do Departamento de Botânica do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Doutora em Biologia Vegetal.

Área de Concentração: Morfologia, sistemática e diversidade vegetal

Orientador: Prof. Dr. Eduardo Leite Borba
Universidade Federal de Minas Gerais
Coorientadores: Profa. Dra. L. Lacey Knowles
University of Michigan
Prof. Dr. Eric de Camargo Smidt
Universidade Federal do Paraná

BELO HORIZONTE – MG

2020

043 Fiorini, Cecilia Fonseca.
Evolutionary processes and the origin of plant biodiversity in old neotropical sky-islands: a genomic and model-based approach [manuscrito] / Cecilia Fonseca Fiorini. – 2020.
122 f. : il. ; 29,5 cm.

Orientador: Prof. Dr. Eduardo Leite Borba. Coorientadores: Profa. Dra. L. Lacey Knowles; Prof. Dr. Eric de Camargo Smidt.
Tese (doutorado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Biologia Vegetal.

1. Morfologia vegetal. 2. Filogeografia. 3. Bulbophyllum - Teses. 4. Velloziaceae. I. Borba, Eduardo Leite. II. Knowles, L. Lacey. III. Smidt, Eric de Camargo. IV. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. V. Título.

CDU: 581

Ficha catalográfica elaborada por Fabiane C. M. Reis – CRB 6/ 2680

UFMG

Programa de Pós-Graduação em Biologia Vegetal
Universidade Federal de Minas Gerais
ICB - Departamento de Botânica

Tese defendida por Cecilia Fonseca Fiorini em 19 de fevereiro de 2020 e aprovada pela Banca Examinadora constituída pelos professores:

Dr. Eduardo Leite Borba - UFMG

Dr. Alexandre Antonelli - Royal Botanic Gardens, Kew

Dr. Cassio van den Berg - Universidade Estadual de Feira de Santana – BA

Dra. Maria Bernadete Lovato UFMG

Dr. João Aguiar Nogueira Batista UFMG

Denise Marcia Figueiredo Monteiro
SECRETÁRIA EXECUTIVA
PROGRAMA DE PÓS-GRADUAÇÃO
EM BIOLOGIA VEGETAL
ICB - UFMG

Aos povos indígenas e
aos professores de Educação Básica.

AGRADECIMENTOS

Agradeço ao Kevin por ser minha melhor companhia para todas as horas e meu maior incentivador. Eu só tenho a agradecer por cada momento que passamos juntos e por tudo que aprendi com você! Obrigada!

Agradeço aos meus pais, Bernadete e Gianfranco, e sogros, Liu e Ricardo, pelo imensurável apoio. Sem a ajuda de vocês nada disto seria possível!

Agradeço às minhas irmãs, Helena e Laura, aos meus cunhados e concunhados, Thiago, Nilo, Ian, Giu, Erica e André, e aos meus sobrinhos, Martin e Imani, pela presença e carinho durante todos esses anos. Vocês fazem a vida ser mais leve!! Agradeço também a minhas avós, tios, tias e primos por também estarem sempre presentes trazendo alegria aos meus dias.

Agradeço ao Prof. Eduardo pelo suporte. Obrigada pela confiança e por ajudar a aprimorar meu trabalho!

Deixo um agradecimento muito especial à Profa. Lacey, que mais do que uma orientadora foi uma amiga. Muito obrigada também à Rebecca, ao Emmet, à Sadie, à Mirtle e ao Asa. Kevin e eu passamos momentos incríveis com vocês!

Agradeço ao Prof. Eric por seu empenho em relação ao meu trabalho. Suas sugestões e suporte foram decisivos!

Agradeço aos demais coautores dos trabalhos que são apresentados nesta tese: Viviane, Ariane, Marina, Ubirajara, Luciana Kamino, Luciana Resende, Nara e Pedro. Trabalhar com pessoas competentes como vocês foi uma alegria e um privilégio.

Agradeço muito às minhas amigas de lab e doutorado, Gaby, Mariana e Jacqueline, por toda ajuda, companheirismo, bons momentos, apoio, aprendizagens! Foi muito bom estar com vocês, que são pessoas tão especiais. Obrigada também por compartilhar a guarda dos "Bulbofilhos", juntamente com o Fábio e a Maitê. Um agradecimento também aos demais colegas de PPG, em especial ao Junior, à Andréa e à Daniela.

Agradeço muito aos amigos da Universidade de Michigan. Luciana, Arnaud, Andréa, Jeronymo, Richie, Silvia, Leandro, Rennan, Laís, Joyce, Mariah, Tao, Rachel, Sasha, Renata, Giorgia, Raquel e Diego me ensinaram muito e me alegraram mesmo nas terríveis temperaturas abaixo de 0°C que passamos em Ann Arbor!

Agradeço a todos que colaboraram com minhas saídas de campo: Kevin (Joaquim Felício, Conceição do Mato Dentro, Catas Altas, Serra Nova, Grão Mogol e Cristália); Gian (São Roque de Minas); Prof. Eduardo (Atibaia, Tiradentes, Carrancas, São Tomé das Letras, Santa Rita de Caldas, Serra Nova, Grão Mogol, Cristália, Joaquim Felício, Diamantina, Santana

do Riacho, Caeté); Gabriela (Atibaia, Tiradentes, Carrancas, São Tomé das Letras, Santa Rita de Caldas, São Roque de Minas, Santana do Riacho, Caeté); Mariana (Tiradentes, Carrancas, São Tomé das Letras, Serra Nova, Grão Mogol, Cristália, Joaquim Felício, Caeté); Jacqueline (Alto Paraíso de Goiás); Lucas (Bahia); Marcos Alexandre dos Santos (Conceição do Mato Dentro); Rubens, Aline e Padre Lauro (Catas Altas); Francisca Izabel Rosa (São Thomé das Letras); Ismael e sua família (Santa Rita de Caldas, Ibituruna, Águas da Patra); Márcio Lucca e sua família (Diamantina); Marcelo, Gleidson e sua família, Lima-Verde e Otoch (Ceará), Erton (Pernambuco e Alagoas); Vinícius e equipe do AOBAL (Alagoas); Elaine (Caraguatatuba); e Werner (Tibagi e Bombinhas). A companhia de vocês foi fundamental para as coletas e estar com cada um de vocês foi um prazer para mim.

Meu muito obrigada também a todos os professores que colaboraram com a execução e aprimoramento deste trabalho: Profa. Ana Maria Giulietti, Dr. José Eustáquio Santos, Prof. Aron King, Prof. Alexandre Antonelli, Prof. Cássio van den Berg, Profa. Bernadete Lovato, Prof. João Batista Aguiar, Prof. Almir Pepato, Prof. Evandro Moraes, Profa. Isabel Bonatellii, Prof. Luis Eduardo Del Bem e Profa. Denise Trombert. Um agradecimento muito especial ao Prof. Fernando Silveira (Lelê) e à Profa. Luzia Modolo, que foram grandes incentivadores do meu trabalho. Agradeço também à Denise Monteiro e à Isabella Silva por toda a ajuda na secretaria do PPG e do departamento.

Obrigada aos meus amigos do peito, que não desistiram de mim mesmo nos momentos em que estive ausente. Xuxuzada e Cia., Elder, Lorena, Tabatha, Saundra, Maria Helena, Giovanna e Bárbara, muito obrigada pelo carinho e alegria!

Concluindo, um agradecimento especial aos órgãos e as agências financiadoras que viabilizaram a execução deste trabalho: IEF, SEMA, IAP, SISBIO, CAPES e CNPq.

A eles e a todos que colaboraram direta
e indiretamente com este trabalho,
muito obrigada!

"Quando despersonalizamos o rio, a montanha, quando tiramos deles os seus sentidos, considerando que isso é um atributo exclusivo dos humanos, nós liberamos esses lugares para que se tornem resíduos da atividade industrial e extrativista. Do nosso divórcio das integrações e interações com a nossa mãe, a Terra, resulta que ela está nos deixando órfãos, não só aos que em diferente graduação são chamados de índios, indígenas ou povos indígenas, mas a todos."

Ailton Krenk, *Ideias para adiar o fim do mundo,* 2019.

RESUMO

Os Neotrópicos concentram uma grande parte da biodiversidade vegetal mundial e foi demonstrado que montanhas desempenham um papel importante neste padrão. A cadeia de montanhas compreendendo Espinhaço e a Chapada-Diamantina abrigam os campos rupestres, um mosaico de vegetação herbácea-arbustiva altamente diversa, em *sky-islands* (acima de 900m) de solos rochosos ou arenosos pobres em nutrientes. Neste estudo, usamos dois grupos que ocorrem nos campos rupestres e apresentam características ecológicas contrastantes (*Bulbophyllum* sect. *Didactyle*, Orchidaceae, e *Vellozia auriculata*, Velloziaceae) para compreender o processo evolutivo que leva às elevadas diversidades de espécies e de espécies endêmicas observadas nessa vegetação. Observamos a ocorrência de estruturação geográfica, sendo mais forte para *V. auriculata*, espécie que apresenta menor vagilidade. Para *B. involutum*, o fluxo gênico entre as localidades de Espinhaço foi elevado, corroborando a noção de que as sementes pequenas e leves de orquídeas são capazes de dispersão a longa distância. No entanto, nossos dados também sustentam que características ambientais ou eventos demográficos passados podem ser fatores importantes para a diferenciação populacional. De um modo geral, a distribuição da variabilidade genética de *B.* sect. *Didactyle* reflete a geografia das populações, contudo algumas espécies da seção não foram recuperadas como monofiléticas. Além da importância das disjunções e da variabilidade ambiental para a diferenciação entre populações, demonstramos também que a hibridação pode ser um mecanismo importante para a origem e manutenção da biodiversidade de campos rupestres.

Palavras chave: Campos rupestres, *Bulbophyllum*, complexo de espécie, *Vellozia auriculata*, filogeografia

ABSTRACT

A large proportion of the world's plant biodiversity is concentrated on the Neotropics and mountains have been shown to play an important role on this pattern. The Espinhaço/Chapada-Diamantina chain harbors the highly diverse campos rupestres, a sky-island herbaceous-shrubby vegetation mosaic, that occurs above 900 m on nutrient poor, rocky and sandy soils. In this study we used two groups occurring in the campos rupestres and presenting contrasting ecological characteristics (*Bulbophyllum* sect. *Didactyle*, Orchidaceae, and *Vellozia auriculata*, Velloziaceae) to understand the evolutionary process leading to the high species and endemic diversity observed on this vegetation. We observed geographical structure, and this structure was stronger for the less vagile *V. auriculata*. For *Bulbophyllum involutum* the gene flow between Espinhaço localities was high, corroborating the notion that the small and light seeds of orchids are able of long-range dispersion. However, our data also support that environment or past demographic events might be important factors driving population differentiation. Overall, the distribution of the genetic variability of *B.* sect. *Didactyle* reflects the population geography, but some species have not been recovered as monophyletic. Beyond the importance of disjunction and environmental variability for population differentiation, here we demonstrate that hybridization might be and critical engine for the origin and maintenance of the campos rupestres biodiversity.

Keywords: Campos rupestres, *Bulbophyllum*, species complex, *Vellozia auriculata*, phylogeography.

SUMÁRIO

INTRODUCTION

The Neotropics concentrates a high proportion of the plant biodiversity of the planet (28%; Antonelli et al., 2015), and mountains have been shown to play an important role on the diversification and maintenance of its biodiversity (Flantua, O'Dea, Onstein, Giraldo, & Hooghiemstra, 2019; Rahbek et al., 2019). Besides to the well-known Andes, South America harbors another cordillera: the Espinhaço/Chapada-Diamantina chain, with altitudes up to 2,072 m and > 1,000 km of latitudinal extension in eastern Brazil. However, differently from the recent Andean cordillera, the Espinhaço/Chapada-Diamantina chain arose from Proterozoic orogenic events (Uhlein, Paim, Tassinari, & Pedreira, 2015) and was partially modified since the Pliocene by compressional tectonic event (Saadi, 1995).

The Espinhaço/Chapada-Diamantina chain and other nearby highland outcrops harbor the campos rupestres vegetation, a sky-island herbaceous-shrubby vegetation mosaic, that occurs above 900 m on nutrient poor, rocky and sandy soils. It is located in Cerrado, Caatinga and Atlantic Rain Forest provinces and is highly diverse. They maintain 15% of Brazilian plant diversity on an area representing less than 1% of the country's land area (Silveira et al., 2016). The campos rupestres are an example of old, infertile and climate-buffered landscapes (OCBILs; Silveira et al., 2016). It has been suggested that in this kind of environment the association of a buffered climate with an infertile soil is the pillar for the origin and maintenance high diversity and endemism, an idea that is supported by new evidence on montane biodiversity (Rahbek et al., 2019). An important prevision of the Ocbil theory is that high specialization and limited seed dispersal would lead to prolongedly isolated population systems, resulting in independent evolution by drift or selection (Hopper, 2009). Indeed, outcrops can maintain stable microhabitats sustaining in situ microrefugia during climatic changes (Schut et al., 2014).

The campos rupestres are especially rich in endemics, as 40% of its plants are restricted to small patches of the environment's distribution. It has been recurrently suggested that due to the disjunct aspect of the campos rupestres geographic isolation might drive independent evolution of populations, differentiation and speciation, a processes called topography-driven-isolation (Steinbauer et al., 2016). However, as happens to other Neotropical environments (Leal, Palma-da-Silva, & Pinheiro, 2016), populational studies using genetic markers do not show a consensual pattern for the campos rupestres, reflecting the intricate and complex processes of Neotropical diversification (Rull, 2013). Generally, plants presenting low dispersion capability show high spatial genetic structure (Barbosa, Fiorini, Silva-Pereira,

Mello-Silva, & Borba, 2012; Collevatti, de Castro, Lima, & Telles, 2012; Fiorini et al., 2019), while vagile species present a more homogeneously distributed genetic diversity (Barres, Batalha-Filho, Schnadelbach, & Roque, 2019).

Along with topography-driven-isolation, hybridization is another process driving population differentiation and speciation. It was considered an evolutionary dead end in the past, but its potential in creating and maintaining biodiversity is now clear (Abbott et al., 2013; Mallet, 2007; Seehausen, 2013). It is estimated that around 25% of plant and 10% of animals participate of hybridization processes (Mallet, 2007) and it is hypothesized that plant engage in hybridization more often due to their more open and plastic patterns of morphogenesis (Gottlieb, 1984). However, while hybridization is frequent in some plants, in other groups it is rare or virtually absent (Whitney, Ahern, Campbell, Albert, & King, 2010). Also, the fact that ancestral polymorphism, mutations and disruptive selection can interfere with the hybrid phenotype hampers the identification of hybrids (Rieseberg, 1995).

In this study we used two groups occurring in the campos rupestres presenting contrasting ecological characteristics in relation to the extension of their geographical distribution and mode of dispersion to understand the evolutionary process leading to the high species and endemic diversity observed on the campos rupestres. They are *Bulbophyllum* sect. *Didactyle* (Lindl.) Cogn (Orchidaceae) and *Vellozia auriculata* Mello-Silva & N.L.Menezes (Velloziaceae).

Bulbophyllum sect. *Didactyle* is composed of seven species: *Bulbophyllum perii* Schltr., *B. popayanense* F. Lehm. & Kraenzl., *B. tripetalum* Lindl. and *B. weddellii* (Lindl.) Rchb. f., and the *B. exaltatum* species complex sensu Ribeiro et al. (2008), that includes *B. exaltatum* Lindl., *B. involutum* Borba, Semir & F. Barros, *B. meridense* Rchb.f. (Smidt, 2007). The group includes also two natural hybrids: *B.* ×*cipoense* Borba & Semir and *B.* ×*guartelae* Mancinelli & E.C.Smidt (Smidt, 2007). The *B. exaltatum* species complex is distributed over the entire campos rupestres and included several other taxa previously described and currently synonymized (Smidt, 2007). Brazil is the center of diversity of the section *Didactyle*, which has wide distribution in Espinhaço and occurs less frequently also in outcrops in the states Goiás, Roraima, and other tropical countries in South America. Only one of its species occurs exclusively outside Brazil, *B. popayanense*.

Like other *Bulbophyllum* species, *B.* sect. *Didactyle* species are pollinated by flies (myiophily), their small and light seeds are anemochoric, and individuals are able to propagate vegetatively, due to its reptant rhizomes. The populations of the *B. exaltatum* complex have high habitat specificity. Although they present anemochoric dispersion, it is very frequent that

only a small portion of adjacent outcrops with virtually identical habitats are colonized. It is also important to note that those outcrops that are colonized often have a high demographic density. These facts indicate the existence of some subtle characteristic that determines the occurrence of the group.

Vellozia auriculata is part of the Velloziaceae family, whose centers of diversity and endemism are the campos rupestres in the Espinhaço Range. *V. auriculata* is one of the tallest species in the dracenoid group of *Vellozia*, reaching up to 4 m in height. It is a micro-endemic species with disjunct distribution in the Diamantina Plateau (Southern Espinhaço range) and is the only species in the dracenoid group that is flexible in terms of habitat specificity as it can be epilitic in both quartzite and canga, or psamophilic. Floral characteristics suggest that *V. auriculata* is pollinated by bees and fruit and seed characteristics suggest barocoric or slightly anemocoric dispersion, preventing the seeds from migrating over long distances. As the species of *B.* sect. *Didactyle*, *V. auriculata* is capable of vegetative propagation.

With the studies of populations of *V. auriculata* and of *B.* sect. *Didactyle* species, we intend to answer questions about the environmental regionalization (e.g.; does lineages composition mirrors the geographical structure of eastern Neotropical rock sky-islands?, Chapter 1), and the role of disjunction (e.g., does geographic disjunction among sky-island lead to genetic disruptions among orchid populations?, Chapters 3 and 4) and hybridization (e.g., are there hybrids on *B.* sect. *Didactyle*?) in promoting the diversification on the campos rupestres of the Espinhaço/Chapada-Diamantina chain and other highland outcrops on eastern Brazil. For this purpose, we scanned *B.* sect. *Didactyle* using the ddRAD methodology; additionally, to access *V. auriculata* variability we sequenced fragments from the plastid DNA and phenotyped genomic Inter Simple Sequence Repeats. The data was analyzed under a model-based approach, using coalescent-based and simulation methods, along with other multivariate techniques.

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., … Zinner, D. (2013). Hybridization and speciation. Journal of Evolutionary Biology, 26(2), 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Antonelli, A., Zizka, A., Silvestro, D., Scharn, R., Cascales-Miñana, B., & Bacon, C. D. (2015). An engine for global plant diversity: highest evolutionary turnover and emigration in

the American tropics. Frontiers in Genetics, 6, 1–14. https://doi.org/10.3389/fgene.2015.00130

Barbosa, A. R., Fiorini, C. F., Silva-Pereira, V., Mello-Silva, R., & Borba, E. L. (2012). Geographical genetic structuring and phenotypic variation in the Vellozia hirsuta (Velloziaceae) ochlospecies complex. American Journal of Botany, 99(9), 1477–1488. https://doi.org/10.3732/ajb.1200070

Barres, L., Batalha-Filho, H., Schnadelbach, A. S., & Roque, N. (2019). Pleistocene climatic changes drove dispersal and isolation of Richterago discoidea (Asteraceae), an endemic plant of campos rupestres in the central and eastern Brazilian sky-islands. Botanical Journal of the Linnean Society, 189(2), 132–152. https://doi.org/10.1093/botlinnean/boy080

Collevatti, R. G., de Castro, T. G., Lima, J. de S., & Telles, M. P. de C. (2012). Phylogeography of Tibouchina papyrus (Pohl) Toledo (Melastomataceae), an endangered tree species from rocky savannas, suggests bidirectional expansion due to climate cooling in the Pleistocene. Ecology and Evolution, 2(5), 1024–1035. https://doi.org/10.1002/ece3.236

Fiorini, C. F., Miranda, M. D., Silva-Pereira, V., Barbosa, A. R., Oliveira, U. De, Kamino, L. H. Y., … Borba, E. L. (2019). The phylogeography of Vellozia auriculata (Velloziaceae) supports low zygotic gene flow and local population persistence in the campo rupestre, a Neotropical OCBIL. Botanical Journal of the Linnean Society, 191(3), 381–398. https://doi.org/10.1093/botlinnean/boz051

Flantua, S. G. A., O'Dea, A., Onstein, R. E., Giraldo, C., & Hooghiemstra, H. (2019). The flickering connectivity system of the north Andean páramos. Journal of Biogeography, 46(8), 1808–1825. https://doi.org/10.1111/jbi.13607

Gottlieb, L. D. (1984). Genetics and morphological evolution in plants. The American Naturalist, 123(5), 681–709. Retrieved from http://www.jstor.org/stable/2461245 .

Hopper, S. D. (2009). OCBIL theory: towards an integrated understanding of the evolution, ecology and conservation of biodiversity on old, climatically buffered, infertile landscapes. Plant and Soil, 322(1), 49–86. https://doi.org/10.1007/s11104-009-0068-0

Leal, B. S. S., Palma-da-Silva, C., & Pinheiro, F. (2016). Phylogeographic studies depict the role of space and time scales of plant speciation in a highly diverse Neotropical region. Critical Reviews in Plant Sciences, 35(4), 215–230. https://doi.org/10.1080/07352689.2016.1254494

Mallet, J. (2007). Hybrid speciation. Nature, 446(7133), 279–283. https://doi.org/10.1038/nature05706

Rahbek, C., Borregaard, M. K., Antonelli, A., Colwell, R. K., Holt, B. G., Nogues-bravo, D., … Fjeldså, J. (2019). Building mountain biodiversity: geological and evolutionary processes. Science, 365(6458), 1114–1119.

Ribeiro, P. L., Borba, E. L., Smidt, E. de C., Lambert, S. M., Schnadelbach, A. S., & van den Berg, C. (2008). Genetic and morphological variation in the *Bulbophyllum exaltatum* (Orchidaceae) complex occurring in the Brazilian "campos rupestres": implications for taxonomy and biogeography. Plant Systematics and Evolution, 270(1–2), 109–137. https://doi.org/10.1007/s00606-007-0603-5

Rieseberg, L. H. (1995). The role of hybrization in evolution: old wine in new skins. American Journal of Botany, 82(7), 944–953. https://doi.org/10.2307/2445981

Rull, V. (2013). Some problems in the study of the origin of Neotropical biodiversity using palaeoecological and molecular phylogenetic evidence. Systematics and Biodiversity, 11(4), 415–423. https://doi.org/10.1080/14772000.2013.865682

Saadi, A. (1995). A geomorfologia da Serra do Espinhaço em Minas Gerais e de suas margens. Geonomos, 3(1), 41–63.

Schut, A. G. T., Wardell-Johnson, G. W., Yates, C. J., Keppel, G., Baran, I., Franklin, S. E., … Byrne, M. (2014). Rapid characterisation of vegetation structure to predict refugia and climate change impacts across a global biodiversity hotspot. PLoS ONE, 9(1), e82778. https://doi.org/10.1371/journal.pone.0082778

Seehausen, O. (2013). Conditions when hybridization might predispose populations for adaptive radiation. Journal of Evolutionary Biology, 26(2), 279–281. https://doi.org/10.1111/jeb.12026

Silveira, F. A. O., Negreiros, D., Barbosa, N. P. U., Buisson, E., Carmo, F. F., Carstensen, D. W., … Lambers, H. (2016). Ecology and evolution of plant diversity in the endangered campo rupestre: a neglected conservation priority. Plant and Soil, 403(1–2), 129–152. https://doi.org/10.1007/s11104-015-2637-8

Smidt, E. de C. (2007). Filogenia e revisão taxonômica de *Bulbophyllum* Thouars (Orchidaceae) ocorrentes no Neotrópico. PhD Dissertation. Universidade Estadual de Feira de Santana.

Steinbauer, M. J., Field, R., Grytnes, J. A., Trigas, P., Ah-Peng, C., Attorre, F., … Beierkuhnlein, C. (2016). Topography-driven isolation, speciation and a global increase of endemism with elevation. Global Ecology and Biogeography, 25(9), 1097–1107. https://doi.org/10.1111/geb.12469

Uhlein, A., Paim, P. S. G., Tassinari, C. C. G., & Pedreira, A. J. (2015). Análise estratigráfica de bacias rifte Paleo-Mesoproterozoicas dos crátons Amazônico e São Francisco, Brasil. Geonomos, 23(2), 1–13. https://doi.org/10.18285/geonomos.v23i2.707

Whitney, K. D., Ahern, J. R., Campbell, L. G., Albert, L. P., & King, M. S. (2010). Patterns of hybridization in plants. Perspectives in Plant Ecology, Evolution and Systematics, 12(3), 175–182. https://doi.org/10.1016/j.ppees.2010.02.002

CHAPTER 1 – THE GEOGRAPHY OF DIVERGENCE ON BRAZIL'S CAMPOS RUPESTRES "SKY-ISLANDS": A POPULATIONAL BIOGEOGRAPHIC ANALYSIS ACROSS MULTIPLE SPECIES (ORCHIDACEAE: *BULBOPHYLLUM*)

Abstract

The basis for the amazing diversity of Orchidaceae is still under debate. It has been shown that the high speciation rates of this family are associated with key innovations related to pollination mechanism and epiphytism, as well as with the colonization of tropical cordilleras. One of the most species rich genus in the family is *Bulbophyllum* Thouars. Within the genus, *Bulbophyllum* sect. *Didactyle* presents some of the greatest challenges, but also opportunities, for studying the orchid divergence processes, as some its species present phenotypic variability preventing the recognition of morphological distinct taxa. *B.* sect. *Didactyle* is distributed mainly across South America's disjunct outcrops, including the species rich campo rupestres vegetation. Here we use the power of genomic data and a broad geographic sampling to resolve phylogenetic relationships and reconstruct the biogeography of diversification of *B.* sect. *Didactyle*. The study of the SNP genomic variation of *B.* sect. *Didactyle* shows that lineages composition mirrors the geographical structure of eastern Neotropical rock sky-islands, in especial the classical campos rupestres' disjunctions, with some new biogeographical connections being described and breaks previously observed for other organisms being reinforced. Yet, our study sheds light on the evolutionary relationships of *B.* sect. *Didactyle*, suggesting that the group evolved trough a complex pattern of isolation, gene flow, and hybridization, and that some of the previously recognized species are polyphyletic, what shall be analyzed in deep on further research. Despite the recent diversification of *B.* sect. *Didactyle*, our study supports the hypothesis that the geographic disjunction among sky-island lead to genetic disruptions among orchid populations, a leading hypothesis for plants in general on the campos rupestres.

Introduction

With more than 28,000 species, Orchidaceae is one of the two most diverse families of angiosperms, but the basis for this astonishing species richness is still under debate (Christenhusz & Byng, 2016). The highest diversity is in the Neotropics, where it has been accumulating since it was colonized during the Cretaceous from Australia via Antarctica before

the continents drifted apart, with some subsequent trans-oceanic dispersal events (Givnish et al., 2015, 2016). The pinnacle of this diversity occurs in the hyper-diverse subfamily Epidendroideae, with over 21,000 species. This high diversity is hypothesized to reflect high speciation rates associated with key innovations related to pollination mechanism and epiphytism, as well as with the colonization of tropical cordilleras (Givnish et al., 2015, 2016; Pérez-Escobar et al., 2017).

Associated with these leading hypotheses for the extreme diversity of orchids, is a central focus on the role of genetic drift. Multiple factors might contribute to drift induced changes, as possible skews in mating success, small and disjunct populations (Phillips et al. 2012), and restrictions on pollen flow due to the pollinator behavior, especially in myophilous genus (pollinated by flies) like *Bulbophyllum* (M. T. A. Azevedo, Borba, Semir, & Solferini, 2007; Borba, Semir, & Shepherd, 2001). On the other hand, others have argued for high vagility because of the numerous, small, balloon-like, wind-dispersed seeds of orchids, which has been deemed the "paradox of orchid speciation" (Arditti & Ghani, 2000; Givnish et al., 2016). However, there is a lack of consensus regarding the effectiveness of orchid seed dispersal in maintaining cohesion between disjunct localities. Not only is there conflicting evidence on the limits of orchid seed dispersal (e.g., Hedrén and Lorenz, 2019; Phillips et al., 2012; Taylor et al., 2019 versus Helsen et al., 2016; Pinheiro et al., 2014; Trapnell et al., 2013), but recent evidence also suggests wind dispersed seeds may not enhance gene flow relative to other dispersal syndromes (Arjona, Nogales, Heleno, & Vargas, 2018; Fajardo et al., 2019), not to mention that factors unrelated to a seeds inherent potential for movement also impact plant dispersal (Taylor et al., 2019).

In addition to these general hypotheses about family-level Neotropical orchid diversity, intriguing questions about diversification lies in understanding the processes contributing to the diversity of the most species rich genus in the family – the epidendroid *Bulbophyllum* Thouars. The genus *Bulbophyllum*, that is an outlier of sorts compared to other Neotropical Orchidaceae. Although there is some notable uncertainty regarding the timing and route of its colonization of the Neotropics, *Bulbophyllum* has an undoubtedly much-truncated history in this region compared to other Epidendroideae clades whose histories generally trace back to more than 20 Mya (Gamisch & Comes, 2019; Givnish et al., 2015, 2016; Smidt, Borba, Gravendeel, Fischer, & van den Berg, 2011). As such, the number of Neotropical *Bulbophyllum* species is small compared with its diversity in Madagascar and the Asian-Pacific region where it has an extended history during which species accumulated (Gamisch & Comes, 2019). In contrast, with approximately 60 Neotropical species currently circumscribed in six sections (Smidt et al.,

2011), the processes generating this diversity over a relatively short period of time remain largely unexplored.

Within the genus, and as a group of recently diversified taxa (Gamisch & Comes, 2019), *Bulbophyllum* sect. *Didactyle*, which circumscribes seven currently recognized species and two natural hybrids (Smidt et al., 2011), presents some of the greatest challenges, but also intriguing opportunities, for studying divergence processes. For example, the group is morphologically a mosaic of clearly delimited taxa (e.g., *B. weddellii*, *B. perii*, *B. tripetalum*) and those with obscure species boundaries (i.e., *B. exaltatum* species complex sensu Ribeiro et al., 2008; Fig. 1) in which phenotypic variability makes it difficult to discern morphological distinct taxa (Ribeiro et al., 2008; Smidt, 2007). Moreover, past molecular analysis has raised questions about the reliability of proposed taxonomic designations because genetic structure corresponded to geographical barriers across the group's distribution, rather than putative species boundaries per se (Ribeiro et al., 2008). It is also possible, however, that the allozymes used in past molecular studies where simply insufficient for resolving species boundaries (see Massatti et al., 2016), and either hybridization and/or ancestral allele sharing might contribute to the lack of a strict correspondence of genetic structure with putative taxonomic boundaries (Knowles & Carstens, 2007).

The distribution of the focal group *B.* sect. *Didactyle* across South America's rocky outcrops (i.e., the campos rupestres; Fig. 2) also complicates our understanding of its diversification history (while also possibly contributing to speciation). The geographic disjunctions of the campos rupestres creates "sky-islands" of a herbaceous-shrubby vegetation mosaic above 900 m on nutrient poor rocky and sandy soils that span > 1,000 km (latitudinal) in eastern Brazil (i.e., along the Espinhaço Range, the Chapada Diamantina, and other nearby highland outcrops; Fig. 2). The campos rupestres are incredibly species rich. For example, it contains 15% of Brazil's plant species diversity, despite representing less than 1% of Brazil's land area (Silveira et al., 2016). This diversity, much of which is endemic (40% are endemic to the campos rupestres, and many are rare plants restricted to small areas; Silveira et al., 2016), suggests geographic isolation promoted by disjunctions of the campos rupestres likely contributes to the diversification process. High genetic structure documented in the few phylogeographical studies available for plants corroborate the effects of the fragmented landscape on inhabitants of the campos rupestres (e.g., Barbosa et al., 2012; Bonatelli et al., 2014). However, given hypothesized dispersal associated with the small seeds of orchids (e.g., Helsen et al., 2016; Pinheiro et al., 2014; Trapnell et al., 2013), and with all taxa in the *B.* sect. *Didactyle* occurring across multiple isolated sky-islands of the campos rupestres, including

some geographic widespread species (see Fig. 2), factors other than geographic isolation may affect the divergence process. For example, evidence of hybridization among some, but not all sky-island populations, is suggested for some taxa based on phenotypic intermediates (Fig. 1), not to mention two natural hybrids are currently recognized (i.e., *B.* ×*cipoense*, and *B.* ×*guartelae*; Borba and Semir, 1998; Mancinelli and Smidt, 2012) .

Here we undertake a biogeographic study of divergence across taxa in the focal group – *B.* sect. *Didactyle*. Specifically, we leverage the power of genomic data (Lanier, Huang, & Knowles, 2014) with broad geographic sampling of putative taxa to resolve phylogenetic relationships, and reconstruct the biogeography of diversification. We couple the population-level sampling across the geographic range of putative taxa, with consideration of phenotypical variation, and evaluate whether there is a correspondence between phenotypic and genetic variation, especially in the *B. exaltatum* species complex, which exhibits complicated morphological patterns across its distribution. As such, we present a robust analysis of the evolutionary relationships between individuals and populations of *B.* sect. *Didactyle*, and discuss what these findings suggests about putative species boundaries and biogeographic history (see Ribeiro et al. 2008). By extension, our work also speaks to the history and regionalization of the campos rupestres, which is of broad interest as a biodiversity hotspot, as well as questions surrounding diversification processes in this topographically complex and historically dynamic region (Chaves, Freitas, Vasconcelos, & Santos, 2015; Colli-Silva, Vasconcelos, & Pirani, 2019).

Methods

Sampling

We sampled 164 individuals from 47 populations across six putative species of *Bulbophyllum* sect. *Didactyle* (*sensu* Ribeiro et al., 2008; Fig. 2). Sampling spanned 35 geographic localities, including areas of sympatry (see Table 1 for vouchers codes). We did not sample *B. popayanense*, which is restricted to Northwestern South America and outside the focal geographic region of our work.

We collected individuals growing on different rocks separated by a minimum of 10 m to prevent sampling vegetative clones or closely related individuals (Hedrén & Lorenz, 2019). Individuals from field collected cuttings were propagated and maintained in the living collection of MHNJB-UFMG (Supplementary table 1). We collected all samples under issued

permits to CFF and ELB (SISBIO 52995-1, SISBIO 57116-1, SISBIO 60165-1, IEF 062/2016, SEMA 30/2017, IAP 51.16, Fundação Serra do Japi 011/2017).

Genomic library preparation and processing

We extracted genomic DNA from fresh leaves (Doyle & Doyle, 1987) and we prepared ddRAD libraries following a modified Peterson et al. (2012) protocol (Parchman et al., 2012). We size-selected fragments between 400–500 bp using Pippin Prep (Sage Science, Beverly, MA), and PCR-amplified these fragments using a high-fidelity DNA polymerase (iProof, Bio-Rad, Hercules, CA), with 8 or 12 cycles. We sequenced individuals in four lanes of an Illumina HiSeq 2500 on Rapid Run Mode (in combination with samples from other projects) at The Centre for Applied Genomics, Hospital for Sick Children, Toronto, Canada, to generate 150 bp single end reads.

Using the Stacks 2.3e pipeline (Rochette & Catchen, 2017), we processed the genomic data. We demultiplexed, filtered, and assembled sequences *de novo* with ustacks, build a catalogue of consensus loci in cstacks, identified individual genotypes with sstacks, organized data by locus with tsv2bam, and aligned the reads and called SNPs with gstacks. The assembly parameters included a minimum depth of coverage, $m = 3$, mismatches allowed between two alleles of a sample, $M = 5$, and mismatches allowed between any two alleles of the catalog, $n = 6$ (i.e., the optimal parameters based on the *r80 loci* plateau, Rochette and Catchen, 2017; see Supplementary Fig. 1), and an upper bound for $\varepsilon = 0.1$, a minimum minor allele frequency = 0.02, and a maximum observed heterozygosity = 0.5.

To maximize the number of loci, we grouped individuals from each species according to their geographic sampling localities, and retained biallelic loci from a minimum of two populations (Huang & Knowles, 2016). To guard against sequencing and assembly errors, we used a custom R script (Thomaz, Malabarba, & Knowles, 2017) to exclude SNPs with θ values within the upper 95% quantile of variability (see Supplementary Fig. 2).

Across the 164 sequenced individuals, 313,094,443 reads were generated (average of 1,909,112 $\pm$ 359,736 reads per individual; Table S1.2). We used the software plink 1.9 (Purcell et al., 2007) to identify SNPs. Because the robustness of analyses to missing data differ, we created two datasets with a maximum of 25% or 40% of missing data. The dataset with 25% of missing data, after processing and filtering the genomic data, contained 2,155 variable loci (i.e., contained at least one biallelic SNP) with a total of 3,987 SNPs; the mean coverage depth per locus was 25 $\pm$ 5 $\times$. The dataset with 40% of missing data contained 6,244 variable loci with a total of 13,273 SNPs; the mean coverage depth per locus was 22 $\pm$ 5 $\times$.

Analyses of geographic structure and phylogenetic relationships

We estimated evolutionary relationships using SVDquartets, which is a coalescent-based method that makes full use the data directly (i.e., it does not rely on summary statistics or use MCMC; Chifman and Kubatko, 2014). Given the robustness of the method to missing data (Reaz, Bayzid, & Rahman, 2014), we used the genomic dataset with 40% of missing data and analyzed it with exhaustive sampling of quartets; we performed a nonparametric bootstrap based on 100 replicates. We estimated two phylogenetic trees: one of the relationships among individuals, and another of the relationships among population lineages (i.e., multiple individuals within a sampled population are represented by a single population lineage). In addition to the coalescent based analyses, we also performed a maximum-likelihood phylogenetic analysis of individuals using concatenated unlinked SNPs with 40% of missing data in RAxML-HPC 8.2.10, on CIPRES (Miller, Pfeiffer, & Schwartz, 2010; Stamatakis, 2014), with the best-scoring ML tree retained using the GTRCAT model, and bootstrap support. For SVDquartets and RAxML *B. weddellii* was set as the outgroup, following results from Smidt et al. (2011).

To infer recent or current population structure based on nearest neighbor haplotype (co-ancestry), we used RADpainter combined with fineRADstructure (version 0.3.2; Malinsky et al. 2018). As it uses whole haplotypes, this is a powerful pipeline to infer the co-ancestry matrix from RADseq data (RADpainter) and clusterize it with a MCMC algorithm (fineRADstructure). We used the proper output from stacks as input file for RADpainter and then assigned individuals to populations using fineRADstructure. We used 100,000 burn in iterations and 100,000 sample iterations for MCMC method, sampling each 1,000 iterations. To estimate the tree, we used 100,000 sample iterations. To plot the results, we used the R script fineRADstructurePlot.R, included within the package. All SNPs from a matrix with 25% of missing data were used.

In addition to the estimation of phylogenetic relationships, we used a principal component analysis (PCA) to explore the geography of the distribution of genomic variation, using adegenet 2.1.1 (Jombart & Ahmed, 2011), in R 3.5.0 (R Core Team, 2014). Due to its potential sensitivity to missing data, we analyzed the genomic dataset with a threshold of 25% of missing data; we replaced missing data values by the per locus mean allele frequency for a given population. The analysis of all individuals was complemented by separate PCAs of different subsets of individuals, to accommodate hierarchical structuring of genetic variation (e.g., structuring of genetic variation at regional versus more local geographic areas).

Population structure within the *B. exaltatum* species complex (*B. meridense*, *B. exaltatum*, and *B. involutum*) was also inferred using fastStructure 1.0 (Raj et al. 2014). To create the bed, bim and fam files required by fastStructure, we convert ped and map files from stacks 2.43 using plink 1.9. We estimated ancestry proportions for each individual for different genetic clusters, *K*, where *K* ranged from 0 to 15, using the structure.py script (included within the package). Ten replicates for each *K* were conducted. The number of genetic clusters that best explain the data structure was inferred using the chooseK.py script (also included within the package) and results visualized with Clumpak (available at http://clumpak.tau.ac.il; Kopelman et al 2015). One random SNP from each locus from the dataset with 25% of missing data were used in these analyses.

Results

Structuring of genetic variation within and among putative taxa

Estimates of phylogenetic relationships were generally consistent between the coalescent based (i.e., estimates based on SVDquartets; Fig. 3D and Supplementary Fig. 3A), the co-ancestry based (fineRADstructure; Fig. 4), and the concatenated data (RAxML; Supplementary Fig. 3B) analyses. Specifically, geographic groups were inferred that showed varying degrees of correspondence with the species boundaries of putative taxa (Fig. 3D). In particular, samples from different locations corresponded to monophyletic taxa in *B. weddellii*, *B. perii*, *B. tripetalum* and, generally, *B. involutum*.

The phylogenetic position of *B. weddellii* was consistently inferred across analyses with strong support, and was evolutionarily more distant from other taxa (i.e., in the unrooted tree it was separated by a relatively long branch – see Supplementary Fig. 3B, in the PCA it was highly distinct from other taxa– Fig. 3A, and it showed low co-ancestry with other taxa – see Fig. 4). Irrespective of whether relationships were estimated based on sampled populations or individuals, this relatively geographically widespread taxa showed regional population structure, with three strongly supported clades (Fig. 3D, 4, and Supplementary Fig. 3). Also, individuals and sampled population of *B. perii* and *B. tripetalum* formed single, well-supported clades in each species (Fig. 3D, 4, and Supplementary Fig. 3). Unlike *B. weddellii*, the phylogenetic relationships of these two taxa relative to the other species varied across analyses (Fig. 3D, 4, and Supplementary Fig. 3). All three of these taxa where putative taxonomic boundaries were robustly supported in phylogenetic analyses, each maintained their genetic distinctiveness despite geographic proximity with other *Bulbophyllum* taxa, including sympatry

is some cases (see Fig. 2). Their respective monophyly based on analyses of the genomic data corresponds to general phenotypic distinctiveness in *B. tripetalum*; however, *B. perii* shows morphological variation (Fig. 1).

Genomic data strongly supported regional geographic structure within *B. involutum* (Fig. 3C), with genetically distinct clusters corresponding to northern and southern groups, although *B. involutum* formed monophyletic or non-monophyletic groups, based on different analyses (i.e., Fig. 3D versus Fig. 4 and Supplementary Fig. 3). Specifically, individuals sequenced from the most northern populations (populations I01 and I02) are genetically less related to other individuals of *B. involutum*, that occasionally are inferred to share a more recent common ancestor with individuals of *B. exaltatum* (Fig. 4 and Supplementary Fig. 3), which itself is consistently inferred to be highly non-monophyletic (Fig. 3D, 4, and Supplementary Fig. 3). Morphologically, *B. involutum* exhibits marked phenotypic variation, even within each of the phylogenetically distinct groups (Fig. 1). On fastStructure (best K=5), northern and southern *B. involutum* form separated groups, but there are signals of admixture between them (Fig. 5).

B. meridense was split into clades or genetic clusters in all analyses, which corresponded to the main geographic disjunction of sampled populations and to morphological differentiation (Fig. 1, 3, 4, 5, and Supplementary Fig. 3). Specifically, individuals sampled in the northern part of *B. meridense*'s distribution formed a clade that was consistently recovered across all phylogenetic analyses and is genetically quite distinct from other taxa from the *B. exaltatum* species complex (Fig. 4, and Supplementary Fig. 3B), including individuals of *B. meridense* sampled in southern populations (i.e., M05 and M06), which were more closely related to geographically proximate populations of southern *B. exaltatum* distributed in distinct sky-islands (regions R07, R08, and R10; i.e., they are not sympatric).

Genetic clades currently recognized as *B. exaltatum* presents a genetically heterogenous assemblage, paralleling the pronounced phenotypic variation in floral morphology. The putative taxon *B. exaltatum* stands out relative to the species in that it is consistently inferred to be highly polyphyletic across analyses. There is nonetheless significant structuring of genetic variation geographically, with regional differentiation with multiple northern and southern clades and genetic groups (Fig.3, 4, 5, and Supplementary Fig. 3); note that the phylogenetic placement of northern population E06 is uncertain and shifts across analyses (Fig.3D, 4, and Supplementary Fig. 3), although it consistently appears closely related to *B. tripetalum*.

Discussion

The study of the SNP genomic variation of *Bulbophyllum* sect. *Didactyle* shows that lineages composition mirrors the geographical structure of eastern Neotropical rock sky-islands, in especial the classical campos rupestres' disjunctions (Colli-Silva et al., 2019), with some new biogeographical connections being described and breaks previously observed for other organisms being reinforced. Yet, our study sheds light on the evolutionary relationships of a thorough geographic sampling of *B.* sect. *Didactyle*, suggesting that the group evolved trough a complex pattern of isolation, gene flow and hybridization.

Geographical orchid lineages on a naturally disjunct environment

Genetic disruptions among lineages on *B.* sect. *Didactyle* are mainly coherent with campos rupestre's geographic disjunctions in different scales, suggesting that orchid gene flow among patches is low enough to result in population differentiation. At the same time, the lack of concordance on the relationships among putative species on different analysis suggest a past of differentiation with gene flow, hybridization or introgression. Indeed, the existence of hybrids is observed in this group (Borba and Semir, 1998; Mancinelli and Smidt, 2012; Chapter 2), it had been shown that species are able of interspecific outcrossing (Borba, Shepherd, & Semir, 1999), and our analyses support admixture in some populations of *B. exaltatum* and *B. involutum*.

Our study supports that the disjunction between Espinhaço Range and Chapada Diamantina is an important driver for lineage differentiation, as northern (Chapada Diamantina) and southern (Espinhaço Range) groups of *B. weddellii*, *B. meridense*, *B. involutum*, and *B. exaltatum* were recovered. Despite this major disjunction being a classical barrier for the campo rupestre environment (e.g., Chaves et al., 2015; Colli-Silva et al., 2019; Rapini et al., 2008; Ribeiro et al., 2014), the precise position of the geographical break is currently under debate. Plant occurrence data support that the limits for the Chapada Diamantina and Southern Espinhaço provinces are located between latitudes 15°-16° S (Colli-Silva et al., 2019, see Fig. 3), while bird distributions support that the barrier is located between latitudes 14°-15° S (Chaves et al., 2015, see figure 4). The southernmost population of *B. involutum* in the Bahia state (I03, at 14.69° S) groups with its conspecific southern populations (Minas Gerais state), agreeing with the bird data (Fig. 2). The phylogenetic relationships of individuals from that population suggests that the barrier is located between Caetité plateau and the lowland areas of the Contas river valley, as proposed by de Vasconcelos et al. (2012) based on the first

observation of the hummingbird *Augastes scutatus* on a nearby area. Population I03 pulls the putative position of the barrier further north, as it is located on an even lower latitude than the *A. scutatus* record.

Chapada Diamantina is usually divided in four main biogeographical sub-regions: Rio de Contas range, Sincorá range, Jacobina region and Morro-do-Chapéu region (Chaves et al., 2015; Ribeiro et al., 2014). Rio de Contas and Sincorá ranges are located on the main area of the Chapada Diamantina (region R03, Fig. 2), while Jacobina and Morro-do-Chapéu regions are located on the patchier and isolated north Chapada Diamantina (region R02, Fig. 2). The differentiation between regions R02 and R03 is supported by phylogenetic analyses and PCAs of taxa from *B.* sect. *Didactyle*, as these regions are occupied by distinct lineages. The population of *B. exaltatum* occupying an intermediate position between these regions (E05) may be viewed as another divergent lineage of hybrid or introgressed origin, as its evolutionary affinities with populations from regions R02 and R03 are not clear and it was recovered as admixed on fastStructure analysis. Also, our results partially support the existence of a biogeographical break between Rio de Contas and Sincorá ranges, as previously suggested for orchids in general (C. O. Azevedo & van den Berg, 2007) and other plant groups (e.g., Ribeiro et al., 2014). Further investigation using approaches that allow a deeper understanding of ongoing microevolutionary process is required to shed light on biogeographical and evolutionary relationships among populations of Chapada Diamantina.

Northern populations of *B. meridense* present the highest geographical disjunction of our sample in the *B. exaltatum* species complex, both within and between species. The disjunction is reflected in its high genetic differentiation in relation to the other species and among its conspecific populations, as shown on the PCA (Fig. 3A) and by the long branches on RAxML tree (Supplementary Fig. 3B). The fact that they are geographically isolated suggests that such differentiation is due to genetic drift. However, selection might also be an important driver of this pattern (Orsini, Vanoverbeke, Swillen, & Mergeay, 2013; Sexton, Hangartner, & Hoffmann, 2014; Wang & Bradburd, 2014) and a study to test the relative importance of these processes is encouraged.

Along the Serra do Espinhaço, the clustering of populations from *B. weddellii*, *B. exaltatum* and *B. involutum* localized on region R07 (Diamantina Plateau district) with their respective co-specific populations from region R08 (Iron Quadrangle district) on all phylogenetic analyses and PCAs discords with the province subdivisions proposed by Colli-Silva et al. (2019). However, it is important to highlight that the campos rupestres occur on a geomorphological mosaic (CPRM, 2010; Prístino, 2000) and species from *B.* sect. *Didactyle*

occur almost exclusively on quarzitic outcrops on this vegetation type. This kind of outcrop is abundant on the Diamantina Plateau district, but less frequent on the Iron Quadrangle district, where the singular and diverse *canga* flora grows on iron formations. Thus, our result does not defy the districts proposed by Colli-Silva et al. (2019) and other studies that identify the Iron Quadrangle as a distinct bioregion of the campos rupestres (Jacobi, Carmo, Vincent, & Stehmann, 2007; Neves et al., 2018), but complement these findings, showing that in a geomorphologically and biologically diverse environment as the campos rupestres different areas inside a same district can present alternative evolutionary-ecological histories.

The evolutionary relationship among isolated western populations (E07, E08 and E11) brings evidence to a previously unknown connection on campo rupestres vegetation. Even though region R04 (Chapada dos Veadeiros; E07) and region R09 (Serra da Canastra; E11) were not recognized as bioregions of the campos rupestres (Colli-Silva et al., 2019), our work supports that *B. exaltatum* populations from these areas are related to another isolated population from region R06 (Serra do Cabral; E08), what is supported by the overall morphological similarity of individuals from these populations. However, despite a widespread distribution across isolated outcrops of the campo rupestres, there is pronounced geographic structure, suggesting that they have remained relatively isolated. The fact that all of them showed admixture with southern *B. involutum* on fastStructure analysis (Fig. 5) suggests a hybrid origin for these populations. In spite of region R06 being included on the Diamantina Plateau district (Colli-Silva et al., 2019), it is disjunct from the main Espinhaço Range, and genetic differentiation between plant populations related to this disjunction is common (Barbosa et al., 2012; Borba, Felix, Solferini, & Semir, 2001; F. F. Jesus, Solferini, Semir, & Prado, 2001; Flavia F. Jesus, Abreu, Semir, & Solferini, 2009). As occurs to northern *B. meridense*, the long branches on RAxML tree and the high dissimilarity in relation to conspecific populations on PCA suggests that these three populations are highly differentiated due to genetic drift, which should be further investigated.

The observed genetic breaks can be driven by dispersal, environmental or interaction filters (Taylor et al., 2019). To help to disentangle the relevance of these process for population differentiation microevolutionary and ecological studies are necessary. Besides isolation by dispersal limitation, isolation by adaptation and monopolization (i.e., persistence of founder effect due to gene flow reduction associated to local adaptation of initial colonizing genotypes) had been show as two important drivers for diversification (Orsini et al., 2013).

The campos rupestres are an Old Climatically Buffered Infertile Landscape (OCBIL; Silveira et al., 2016). The OCBIL theory proposes that buffered climate and soil infertility

would lead to specific selective regimes and biotic patterns, including reduced dispersability of seeds, high diversity and high endemism (Silveira et al., 2016). Indeed, seed dispersal syndromes favoring short-distance dispersion are common on campos rupestres (Conceição et al., 2016). Population genetic studies also point to high differentiation between demes across the disjunct sky-islands on the Espinhaço Range and the Chapada Diamantina (e.g. Barbosa et al., 2012; Borba et al., 2007, 2001a; Jesus et al., 2001; Lousada et al., 2013; Pereira et al., 2007; Ribeiro et al., 2018, 2008). As our results suggest that recurrent gene flow among patches of campos rupestres is restricted even for orchids species, which present putative more vagile seeds, gene flow for plant species with low seed vagility must be even more restricted. Our results emphasize the potential for allopatric/peripatric speciation on the campos rupestres but also point to the occurrence of sporadic gene flow among the forming lineages in the past.

Despite being an important first step to understand the evolutionary history of the *B.* sect. *Didactyle*, the complexity of the speciation process on this group deserves a deeper investigation (Pennisi, 2016). The use of a more flexible method, as the model based phylogeographic approach, is required to test hypothesis about the mode of speciation and demographic relationships among populations of this group of orchids presenting an imtrincate evolution pattern (Knowles, Huang, Sukumaran, & Smith, 2018; Papadopoulou & Knowles, 2017).

Considerations about Bulbophyllum sect. Didactyle *lineages*

Our analyses show a strong geographical component of *B. weddellii* genetic variability, showing that it presented at least three geographically structured clades. Despite its morphological uniformity (Smidt, 2007), the existence of even more lineages of *B. weddellii* is expected, as the distribution of this species is wider than we sampled, reaching as far as Bolivia and Peru. *B. involutum* was also supported as a good species in our analyses. However, *B. involutum* shows signs of admixture with other species and this fact deserves further investigation to disentangle the processes related to the origin and actual dynamics of this species.

However, our analyses did not recover *B. exaltatum* and *B. meridense* as monophyletic. Hybridization events could blur the observed pattern and accelerate the differentiation among northern and southern populations of each of the currently circumscribed species (Smidt, 2007). *B. exaltatum* populations are mainly divided mainly into two geographical lineages (northern and southern), but the phylogenetic relationships within these lineages is not clear as sub-northern and -southern lineages did not emerge as monophyletic. Indeed, Ribeiro et al. (2008)

also observed that genetic identities between non-conspecific pairs from the same geographic region were greater than the similarities between conspecific pairs from different regions.

Although there are evidence that *B. exaltatum* is a polyphyletic and polymorphic group, it is not clearly polytypical and, because of that, it is difficult to accommodate this group in a formal taxonomic classification, in which the morphological variation would be congruent with the genetic structure, particularly when a clear morphological structure does not actually occur (Ribeiro et al. 2008). Such cases are not uncommon for plant species, and eventually a formal taxonomic treatment is not even applicable, as in the case of ochlospecies (Cronk, 1998), as observed in other plants species from the campos rupestres (Barbosa et al., 2012).

Furthermore, due to the extensive occurrence of processes such as reticulation and anastomosis of lineages (e.g., by hybridization and introgression), which may even occur politopically, and of temporal coexistence of parent-derivative pairs, we need to face the fact that species of plants commonly are not monophyletic. The indication of maintaining *B. exaltatum* as a taxonomic species may be more appropriate and the most practical alternative, provided that the variability and particularities of these lineages are clearly evidenced, both for the purposes of biological studies as for conservation.

Conclusion

Despite the recent diversification of *Bulbophyllum* sect. *Didactyle*, our study supports the hypothesis that the geographic disjunction among sky-island lead to genetic disruptions among orchid populations, a leading hypothesis for plants in general on the campos rupestres. Further phylogeographical studies are required to evaluate if this differentiation is driven by genetic drift or selection. Measuring gene flow rates between populations and looking for outlier loci will increase our understanding of the evolution of these orchids. Hybridization might be also an important factor on the diversification of *B.* sect. *Didactyle* and future studies shall investigate its impacts on the evolution of this group.

References

Arditti, J., & Ghani, A. K. A. (2000). Numerical and physical properties of orchid seeds and their biological implications. *New Phytologist*, *145*(3), 367–421. https://doi.org/10.1046/j.1469-8137.2000.00587.x

Arjona, Y., Nogales, M., Heleno, R., & Vargas, P. (2018). Long-distance dispersal syndromes matter: diaspore–trait effect on shaping plant distribution across the Canary Islands. *Ecography*, *41*(5), 805–814. https://doi.org/10.1111/ecog.02624

Azevedo, C. O., & van den Berg, C. (2007). Análise comparativa de áreas de campo rupestre da Cadeia do Espinhaço (Bahia e Minas Gerais, Brasil) baseada em espécies de Orchidaceae. *Sitientibus Série Ciencias Biológicas*, *7*(3), 199–210. Retrieved from http://www2.uefs.br/revistabiologia/pg7_n3.html

Azevedo, M. T. A., Borba, E. L., Semir, J., & Solferini, V. N. (2007). High genetic variability in Neotropical myophilous orchids. *Botanical Journal of the Linnean Society*, *153*(1), 33–40. https://doi.org/10.1111/j.1095-8339.2007.00589.x

Barbosa, A. R., Fiorini, C. F., Silva-Pereira, V., Mello-Silva, R., & Borba, E. L. (2012). Geographical genetic structuring and phenotypic variation in the *Vellozia hirsuta* (Velloziaceae) ochlospecies complex. *American Journal of Botany*, *99*(9), 1477–1488. https://doi.org/10.3732/ajb.1200070

Bonatelli, I. A. S., Perez, M. F., Peterson, A. T., Taylor, N. P., Zappi, D. C., Machado, M. C., … Moraes, E. M. (2014). Interglacial microrefugia and diversification of a cactus species complex: Phylogeography and palaeodistributional reconstructions for *Pilosocereus aurisetus* and allies. *Molecular Ecology*, *23*(12), 3044–3063. https://doi.org/10.1111/mec.12780

Borba, E. L., Felix, J. M., Solferini, V. N., & Semir, J. (2001). Fly-Pollinated Pleurothallis (Orchidaceae) species have high genetic variability: evidence from isozyme markers. *American Journal of Botany*, *88*(3), 419–428. https://doi.org/10.2307/2657106

Borba, E. L., Funch, R. R., Ribeiro, P. L., Smidt, E. de C., & Silva-Pereira, V. (2007). Demography, and genetic and morphological variability of the endangered *Sophronitis sincorana* (Orchidaceae) in the Chapada Diamantina, Brazil. *Plant Systematics and Evolution*, *267*(1–4), 129–146. https://doi.org/10.1007/s00606-007-0555-9

Borba, E. L., & Semir, J. (1998). *Bulbophyllum* ×*cipoense* (Orchidaceae), a new natural hybrid from the Brazilian" campos rupestres": description and biology. *Lindleyana*, Vol. 13, pp. 113–120.

Borba, E. L., Semir, J., & Shepherd, G. J. (2001). Self-incompatibility, inbreeding depression and crossing potential in five brazilian *Pleurothallis* (Orchidaceae) species. *Annals of Botany*, *88*(1), 89–99. https://doi.org/10.1006/anbo.2001.1435

Borba, E. L., Shepherd, G. J., & Semir, J. (1999). Reproductive systems and crossing potential in three species of *Bulbophyllum* (Orchidaceae) occurring in Brazilian “campo rupestre”

vegetation. *Plant Systematics and Evolution*, *217*(3–4), 205–214. https://doi.org/10.1007/BF00984366

Chaves, A. V., Freitas, G. H. S., Vasconcelos, M. F., & Santos, F. R. dos. (2015). Biogeographic patterns, origin and speciation of the endemic birds from eastern Brazilian mountaintops: a review. *Systematics and Biodiversity*, *13*(1), 1–16. https://doi.org/10.1080/14772000.2014.972477

Chifman, J., & Kubatko, L. (2014). Quartet inference from SNP data under the coalescent model. *Bioinformatics*, *30*(23), 3317–3324. https://doi.org/10.1093/bioinformatics/btu530

Christenhusz, M. J. M., & Byng, J. W. (2016). The number of known plants species in the world and its annual increase. *Phytotaxa*, *261*(3), 201–217. https://doi.org/10.11646/phytotaxa.261.3.1

Colli-Silva, M., Vasconcelos, T. N. C., & Pirani, J. R. (2019). Outstanding plant endemism levels strongly support the recognition of campo rupestre provinces in mountaintops of eastern South America. *Journal of Biogeography*, *46*(8), 1723–1733. https://doi.org/10.1111/jbi.13585

CPRM, S. G. do B. (2010). Mapa Geodiversidade do Estado de Minas Gerais. Retrieved April 9, 2015, from http://rigeo.cprm.gov.br/jspui/handle/doc/14704

Cronk, Q. C. B. (1998). The ochlospecies concept. In C. R. Huxley, J. M. Lock, & D. . Cutler (Eds.), *Chorology, taxonomy and ecology of the Floras of Africa and Madagascar* (1st ed., pp. 155–170). Richmond: Kew.

Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, *19*(1), 11–15.

Fajardo, J., Mateo, R. G., Vargas, P., Fernández-Alonso, J. L., Gómez-Rubio, V., Felicísimo, Á. M., & Muñoz, J. (2019). The role of abiotic mechanisms of long-distance dispersal in the American origin of the Galápagos flora. *Global Ecology and Biogeography*, *28*(11), 1610–1620. https://doi.org/10.1111/geb.12977

Gamisch, A., & Comes, H. P. (2019). Clade-age-dependent diversification under high species turnover shapes species richness disparities among tropical rainforest lineages of *Bulbophyllum* (Orchidaceae). *BMC Evolutionary Biology*, *19*(1), 93. https://doi.org/10.1186/s12862-019-1416-1

Givnish, T. J., Spalink, D., Ames, M., Lyon, S. P., Hunter, S. J., Zuluaga, A., … Cameron, K. M. (2015). Orchid phylogenomics and multiple drivers of their extraordinary

diversification. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1814), 20151553. https://doi.org/10.1098/rspb.2015.1553

Givnish, T. J., Spalink, D., Ames, M., Lyon, S. P., Hunter, S. J., Zuluaga, A., … Cameron, K. M. (2016). Orchid historical biogeography, diversification, Antarctica and the paradox of orchid dispersal. *Journal of Biogeography*, *43*(10), 1905–1916. https://doi.org/10.1111/jbi.12854

Hedrén, M., & Lorenz, R. (2019). Seed dispersal and fine-scale genetic structuring in the asexual *Nigritella miniata* (Orchidaceae) in the Alps. *Botanical Journal of the Linnean Society*, *190*(1), 83–100. https://doi.org/10.1093/botlinnean/boz005

Helsen, K., Meekers, T., Vranckx, G., Roldán-Ruiz, I., Vandepitte, K., & Honnay, O. (2016). A direct assessment of realized seed and pollen flow within and between two isolated populations of the food-deceptive orchid *Orchis mascula*. *Plant Biology*, *18*(1), 139–146. https://doi.org/10.1111/plb.12342

Huang, H., & Knowles, L. L. (2016). Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. *Systematic Biology*, *65*(3), 357–365. https://doi.org/10.1093/sysbio/syu046

Jacobi, C. M., Carmo, F. F. do, Vincent, R. C., & Stehmann, J. R. (2007). Plant communities on ironstone outcrops: a diverse and endangered Brazilian ecosystem. *Biodiversity and Conservation*, *16*(7), 2185–2200. https://doi.org/10.1007/s10531-007-9156-8

Jesus, F. F., Solferini, V. N., Semir, J., & Prado, P. I. (2001). Local genetic differentiation in *Proteopsis argentea* (Asteraceae), a perennial herb endemic in Brazil. *Plant Systematics and Evolution*, *226*(1–2), 59–68. https://doi.org/10.1007/s006060170073

Jesus, Flavia F., Abreu, A. G., Semir, J., & Solferini, V. N. (2009). Low genetic diversity but local genetic differentiation in endemic *Minasia* (Asteraceae) species from Brazil. *Plant Systematics and Evolution*, *277*(3–4), 187–196. https://doi.org/10.1007/s00606-008-0128-6

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*(21), 3070–3071. https://doi.org/10.1093/bioinformatics/btr521

Knowles, L. L., & Carstens, B. C. (2007). Delimiting species without monophyletic gene trees. *Systematic Biology*, *56*(6), 887–895. https://doi.org/10.1080/10635150701701091

Knowles, L. L., Huang, H., Sukumaran, J., & Smith, S. A. (2018). A matter of phylogenetic scale: distinguishing incomplete lineage sorting from lateral gene transfer as the cause

of gene tree discord in recent versus deep diversification histories. *American Journal of Botany*, *105*(3), 376–384. https://doi.org/10.1002/ajb2.1064

Lanier, H. C., Huang, H., & Knowles, L. L. (2014). How low can you go? The effects of mutation rate on the accuracy of species-tree estimation. *Molecular Phylogenetics and Evolution*, *70*(1), 112–119. https://doi.org/10.1016/j.ympev.2013.09.006

Lousada, J. M., Lovato, M. B., & Borba, E. L. (2013). High genetic divergence and low genetic variability in disjunct populations of the endemic *Vellozia compacta* (Velloziaceae) occurring in two edaphic environments of Brazilian campos rupestres. *Brazilian Journal of Botany*, *36*(1), 45–53. https://doi.org/10.1007/s40415-013-0001-x

Mancinelli, W. S., & Smidt, E. de C. (2012). *O gênero* Bulbophyllum *(Orchidaceae) na Região Sul do Brasil*. *63*(4), 803–815.

Massatti, R., Reznicek, A. A., & Knowles, L. L. (2016). Utilizing RADseq data for phylogenetic analysis of challenging taxonomic groups: A case study in *Carex* sect. *Racemosae*. *American Journal of Botany*, *103*(2), 337–347. https://doi.org/10.3732/ajb.1500315

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *2010 Gateway Computing Environments Workshop (GCE)*, 1–8. https://doi.org/10.1109/GCE.2010.5676129

Neves, D. M., Dexter, K. G., Pennington, R. T., Bueno, M. L., de Miranda, P. L. S., & Oliveira-Filho, A. T. (2018). Lack of floristic identity in campos rupestres—A hyperdiverse mosaic of rocky montane savannas in South America. *Flora*, *238*, 24–31. https://doi.org/10.1016/j.flora.2017.03.011

Orsini, L., Vanoverbeke, J., Swillen, I. N. E., & Mergeay, J. (2013). Drivers of population genetic differentiation in the wild: isolation by dispersal limitation , isolation by adaptation and isolation by colonization. 5983–5999. https://doi.org/10.1111/mec.12561

Papadopoulou, A., & Knowles, L. L. (2017). Linking micro- and macroevolutionary perspectives to evaluate the role of Quaternary sea-level oscillations in island diversification. *Evolution*, *71*(12), 2901–2917. https://doi.org/10.1111/evo.13384

Parchman, T. L., Gompert, Z., Mudge, J., Schilkey, F. D., Benkman, C. W., & Buerkle, C. A. (2012). Genome-wide association genetics of an adaptive trait in lodgepole pine. *Molecular Ecology*, *21*(12), 2991–3005. https://doi.org/10.1111/j.1365-294X.2012.05513.x

Pennisi, E. (2016). Shaking up the tree of life. *Science*, *354*(6314), 817–821. https://doi.org/10.1126/science.354.6314.817

Pereira, A. C. S., Borba, E. L., & Giulietti, A. M. (2007). Genetic and morphological variability of the endangered *Syngonanthus mucugensis* Giul. (Eriocaulaceae) from the Chapada Diamantina, Brazil: implications for conservation and taxonomy. *Botanical Journal of the Linnean Society*, *153*(4), 401–416. https://doi.org/10.1111/j.1095-8339.2007.00624.x

Pérez-Escobar, O. A., Chomicki, G., Condamine, F. L., Karremans, A. P., Bogarín, D., Matzke, N. J., … Antonelli, A. (2017). Recent origin and rapid speciation of Neotropical orchids in the world's richest plant biodiversity hotspot. *New Phytologist*, *215*(2), 891–905. https://doi.org/10.1111/nph.14629

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS One*, *7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Phillips, R. D., Dixon, K. W., & Peakall, R. (2012). Low population genetic differentiation in the Orchidaceae: implications for the diversification of the family. *Molecular Ecology*, *21*(21), 5208–5220. https://doi.org/10.1111/mec.12036

Pinheiro, F., Cozzolino, S., Draper, D., De Barros, F., Félix, L. P., Fay, M. F., & Palma-Silva, C. (2014). Rock outcrop orchids reveal the genetic connectivity and diversity of inselbergs of northeastern Brazil. *BMC Evolutionary Biology*, *14*(1), 49. https://doi.org/10.1186/1471-2148-14-49

Prístino, I. (2000). Atlas digital geoambiental: Reserva da biosfera da Serra do Espinhaço. Retrieved October 10, 2019, from https://institutopristino.org.br/atlas/

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., … Sham, P. C. (2007). *PLINK: a tool set for whole-genome association and population-based linkage analyses*. *81*(3), 559–575.

R Core Team. (2014). *R: A Language and Environment for Statistical Computing*. Retrieved from http://www.r-project.org/

Rapini, A., Ribeiro, P. L., Lambert, S., & Pirani, J. R. (2008). A flora dos campos rupestres da Cadeia do Espinhaço. *Megadiversidade*, *4*(1–2), 16–24. Retrieved from http://www.inot.org.br/artigo/Rapini_et_al_2008_A_flora_dos_campos_rupestres_da_Cadeia_do_Espinhaco.pdf

Reaz, R., Bayzid, M. S., & Rahman, M. S. (2014). Accurate phylogenetic tree reconstruction from quartets: A heuristic approach. *PLoS ONE*, *9*(8). https://doi.org/10.1371/journal.pone.0104008

Ribeiro, P. L., Borba, E. L., Smidt, E. de C., Lambert, S. M., Schnadelbach, A. S., & van den Berg, C. (2008). Genetic and morphological variation in the *Bulbophyllum exaltatum* (Orchidaceae) complex occurring in the Brazilian "campos rupestres": implications for taxonomy and biogeography. *Plant Systematics and Evolution*, *270*(1–2), 109–137. https://doi.org/10.1007/s00606-007-0603-5

Ribeiro, P. L., Pereira, A. C. S., Borba, E. L., & Giulietti, A. M. (2018). Genetic and morphological diversity and evidence of hybridization in the "sempre-vivas" (Comanthera , Eriocaulaceae) endemic to the Chapada Diamantina, Bahia, Brazil. *Flora*, *238*(November 2016), 60–71. https://doi.org/10.1016/j.flora.2017.09.013

Ribeiro, P. L., Rapini, A., Damascena, L. S., & Van Den Berg, C. (2014). Plant diversification in the Espinhaço Range: Insights from the biogeography of *Minaria* (Apocynaceae). *Taxon*, *63*(6), 1253–1264. https://doi.org/10.12705/636.16

Rochette, N. C., & Catchen, J. M. (2017). Deriving genotypes from RAD-seq short-read data using Stacks. *Nature Publishing Group*, *12*(12), 2640–2659. https://doi.org/10.1038/nprot.2017.123

Sexton, J. P., Hangartner, S. B., & Hoffmann, A. A. (2014). Genetic isolation by environment or distance: which pattern of gene flow is most common? *Evolution*, *68*(1), 1–15. https://doi.org/10.1111/evo.12258

Silveira, F. A. O., Negreiros, D., Barbosa, N. P. U., Buisson, E., Carmo, F. F., Carstensen, D. W., … Lambers, H. (2016). Ecology and evolution of plant diversity in the endangered campo rupestre: a neglected conservation priority. *Plant and Soil*, *403*(1–2), 129–152. https://doi.org/10.1007/s11104-015-2637-8

Smidt, E. de C. (2007). Filogenia e revisão taxonômica de *Bulbophyllum* Thouars (Orchidaceae) ocorrentes no Neotrópico. Universidade Estadual de Feira de Santana.

Smidt, E. de C., Borba, E. L., Gravendeel, B., Fischer, G. A., & van den Berg, C. (2011). Molecular phylogeny of the Neotropical sections of *Bulbophyllum* (Orchidaceae) using nuclear and plastid spacers. *Taxon*, *60*(4), 1050–1064.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313.

Taylor, A., Weigelt, P., König, C., Zotz, G., & Kreft, H. (2019). Island disharmony revisited using orchids as a model group. *New Phytologist*, *223*(2), 597–606. https://doi.org/10.1111/nph.15776

Thomaz, A. T., Malabarba, L. R., & Knowles, L. L. (2017). Genomic signatures of paleodrainages in a freshwater fish along the southeastern coast of Brazil: genetic structure reflects past riverine properties. *Heredity*, *119*(4), 287–294. https://doi.org/10.1038/hdy.2017.46

Trapnell, D. W., Hamrick, J. L., Ishibashi, C. D., & Kartzinel, T. R. (2013). Redefining the role of admixture and genomics in species conservation. *Molecular Ecology*, *22*(14), 3680–3692. https://doi.org/10.1111/mec.12338

Vasconcelos, M. F. de, Chaves, A. V., & Santos, F. R. dos. (2012). First record of Augastes scutatus for Bahia refines the location of a purported barrier promoting speciation in the Espinhaço Range, Brazil. *Revista Brasileira de Ornitologia*, *20*(4), 443–446.

Wang, I. J., & Bradburd, G. S. (2014). Isolation by environment. *Molecular Ecology*, *23*(23), 5649–5662. https://doi.org/10.1111/mec.12938

Figure 1. Floral morphological variation in *Bulbophyllum* sect. *Didactyle* across populations (for label localities, see Fig. 2 and Table 1). Image order corresponds to the order of tips in tree presented in Fig. 3D and the species color code is consistently used throughout the paper.

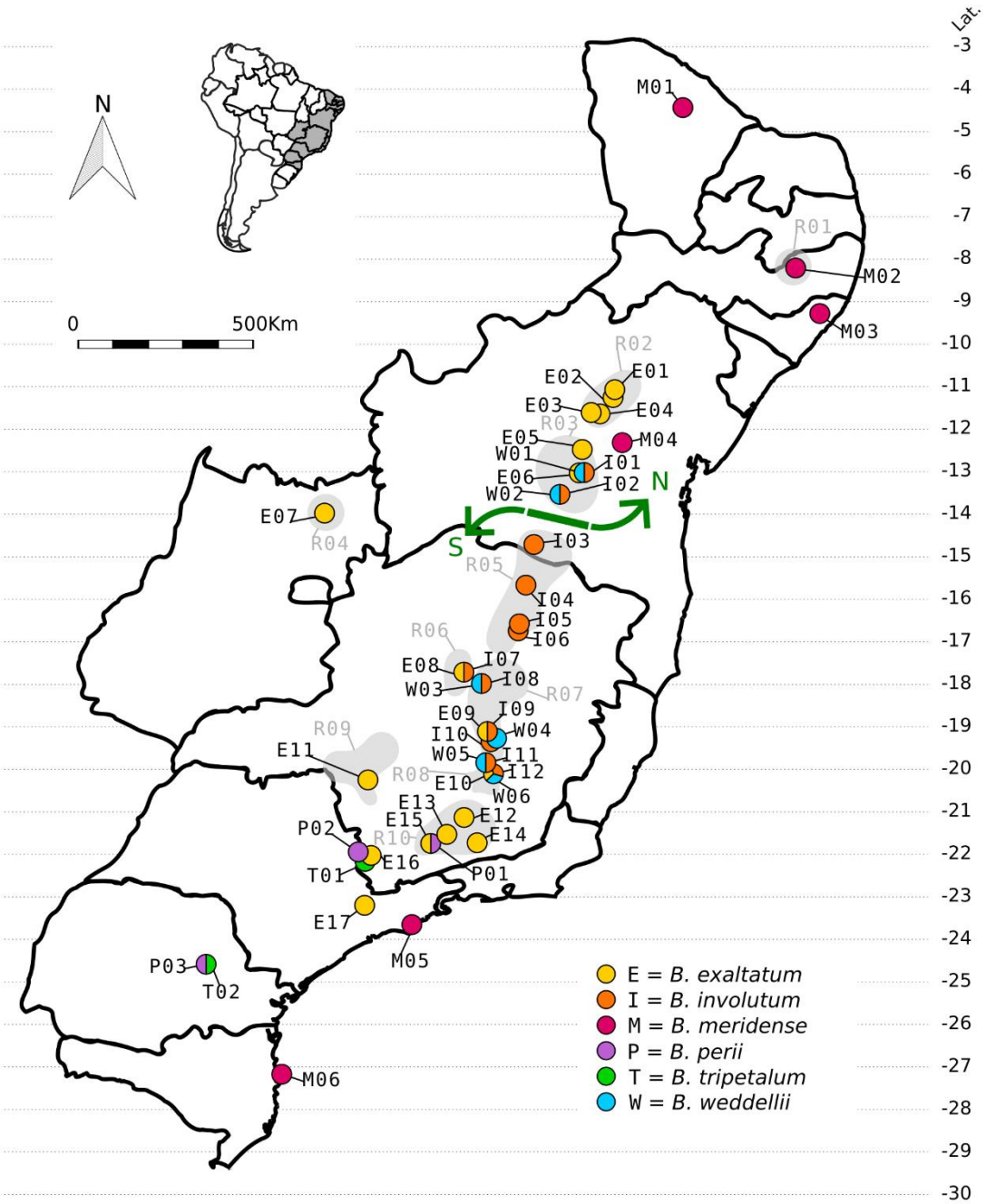


Figure 2. Sampled localities of *Bulbophyllum* sect. *Didactyle*. The circles mark sampling localities denoted by colors for each species, while spliced circles correspond to sympatric species. Campos rupestres' regions are shown in light grey and numbered for reference in the text (i.e., R01-R10). The putative position of a Northern-Southern barrier is also shown (dark green).

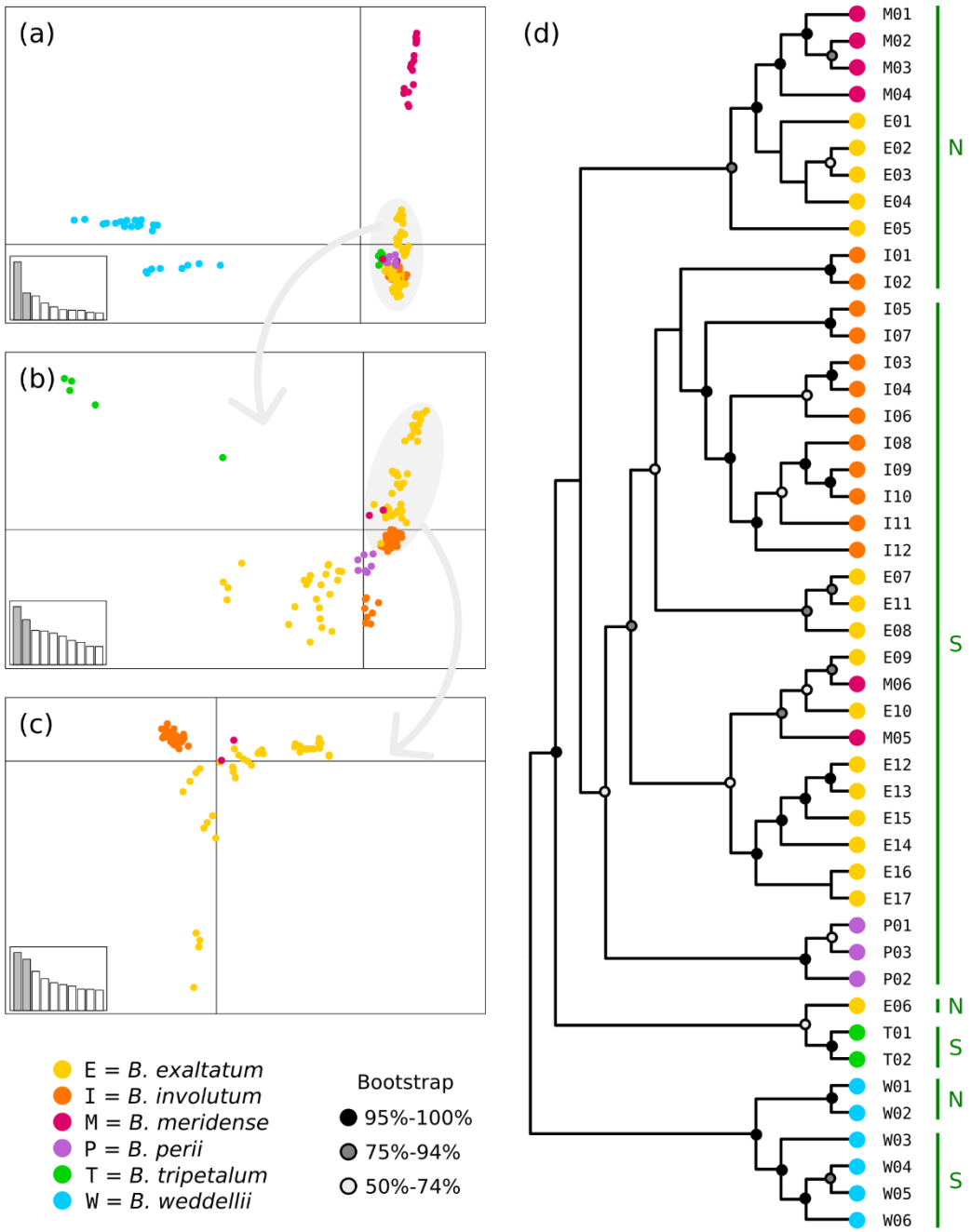


Figure 3. PCA and phylogenetic analysis for *Bulbophyllum* sect. *Didactyle*. PCAs of (a) all populations, (b) with *B. weddellii* and northern *B. meridense* excluded, and with (c) only the southern populations of *B. exaltatum*, *B. involutum*, and *B. meridense*. (d) Phylogenetic estimates of populations relationships based on SVDquartets analyses; the geographic distribution of populations is given in Fig. 2 and the northern versus southern distribution is shown.

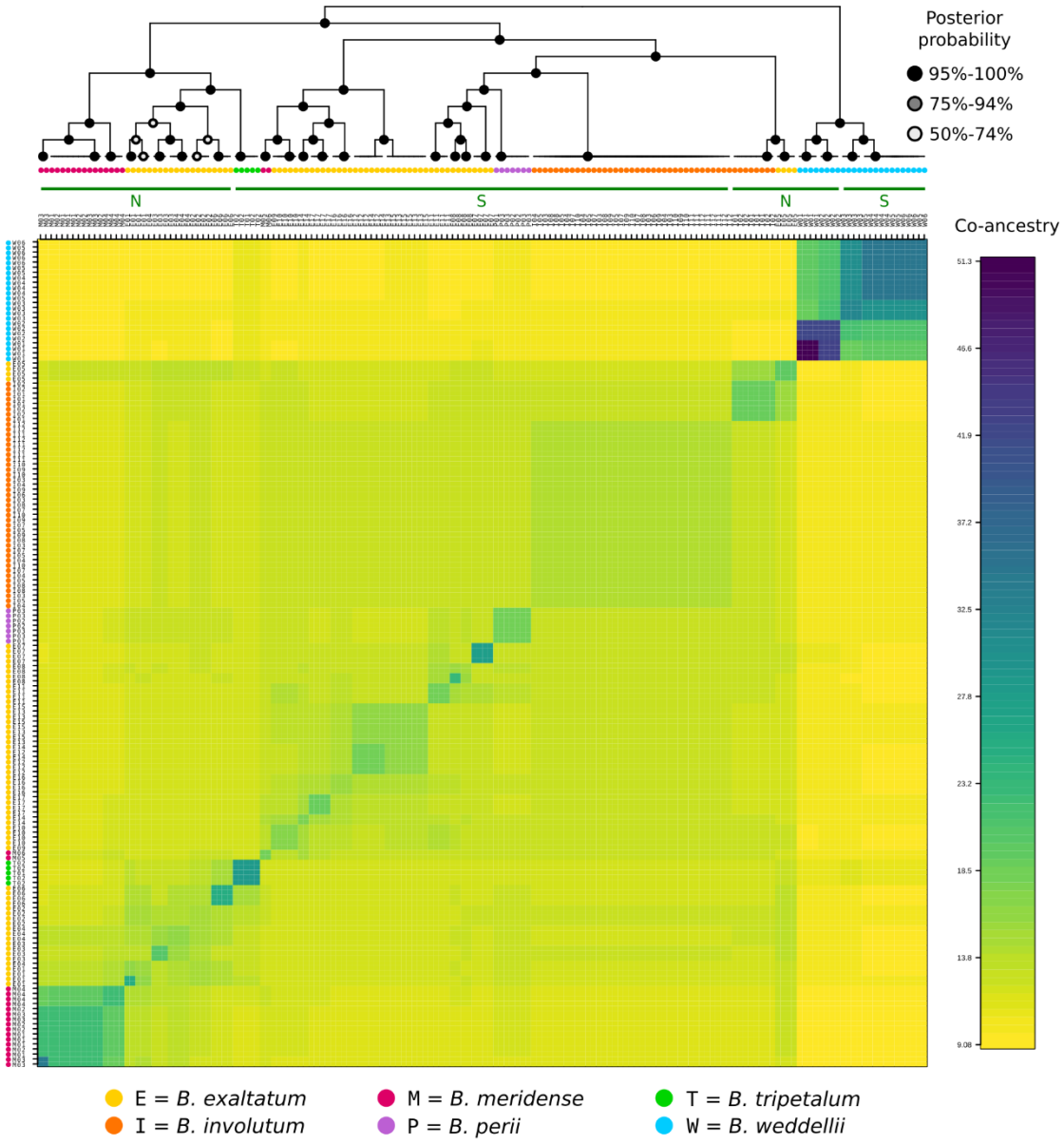


Figure 4. Phylogenetic inference and co-ancestry matrix for *Bulbophyllum* sect. *Didactyle*. The geographic distribution of populations is given in Fig. 2 and the northern versus southern distribution is shown.

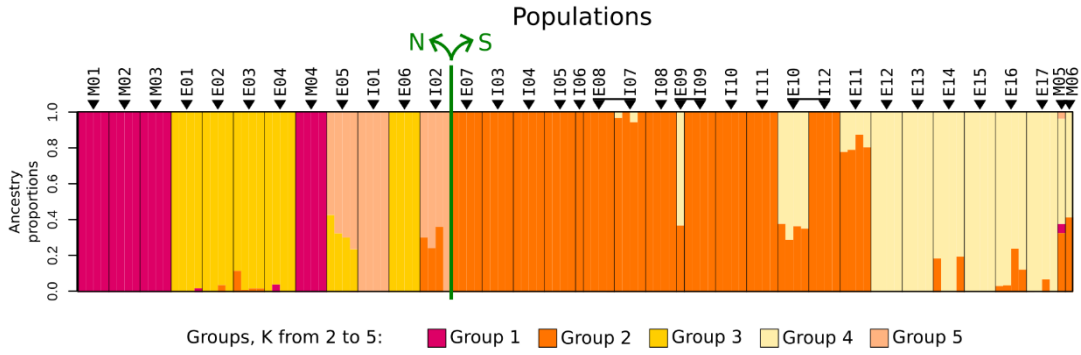


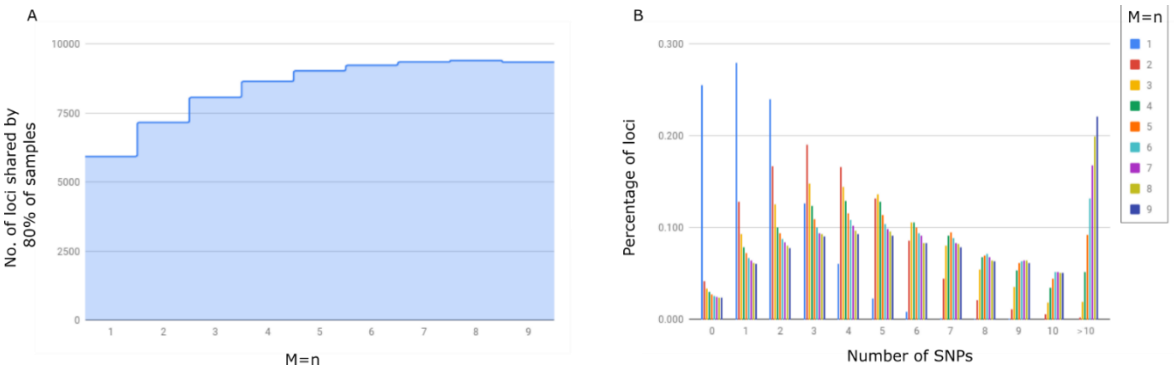
Figure 5. Admixture analysis for *Bulbophyllum exaltatum* species complex. Results of fastStructure analysis for the best K (K=5). Populations are organized by decrescent latitude, independent of putative species. Population identifications are shown. Sympatric populations are connected by a line (E08 and I07, E09 and -I09, and E10 and I12). The geographic distribution of populations is given in Fig. 2 and the northern versus southern distribution is shown in the bar.

Table 1. Geographic information of the sampled *B.* sect. *Didactyle* populations. Population names combine the first letter from of the putative species (i.e., E, *B. exaltatum*; I, *B. involutum*; M, *B. meridense*; P, *B. perii*; T, *B. tripetalum*; W, *B. weddellii*) with locality names ordered by crescent latitude. Pop: population; Lat: latitude; Lon: longitude.

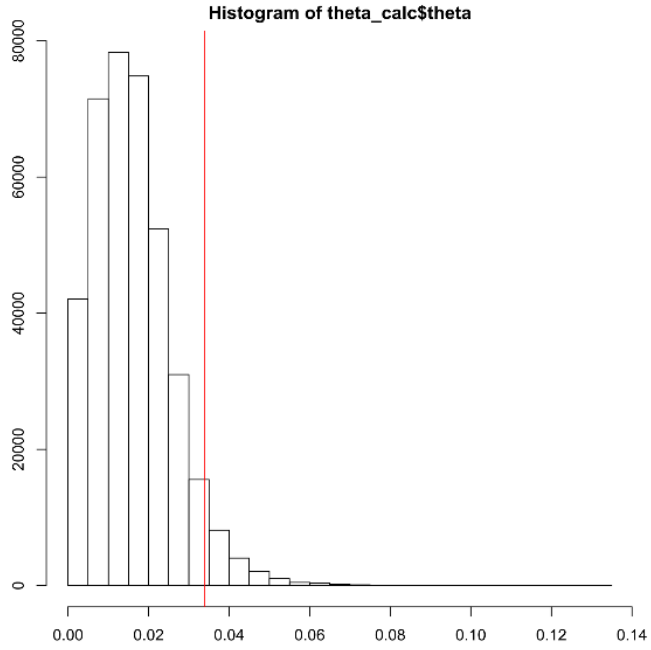
Pop	City	State	Lat	Lon	Voucher
E01	Jacobina	BA	-11.05	-40.65	HUEFS0054358
E02	Miguel Calmon	BA	-11.24	-40.70	ALCB007803
E03	Morro do Chapéu	BA	-11.59	-41.21	HUEFS0167591
E04	Ventura	BA	-11.63	-41.00	HUEFS0086104
E05	Lençóis	BA	-12.46	-41.42	ALCB007806
E06	Mucugê	BA	-13.01	-41.41	HUEFS0086117
E07	Alto Paraíso de Goiás	GO	-13.96	-47.47	UEC070739
E08	Joaquim Felício	MG	-17.69	-44.20	BHCB100401
E09	Conceição do Mato Dentro	MG	-19.09	-43.57	HUEFS0117182
E10	Catas Altas	MG	-20.08	-43.50	BHCB92776
E11	São Roque de Minas	MG	-20.23	-46.45	HUFU008211
E12	Tiradentes	MG	-21.11	-44.20	HUFSJ004023
E13	Carrancas	MG	-21.51	-44.60	UEC064706
E14	Lima Duarte	MG	-21.70	-43.89	BHCB16158
E15	São Tomé das Letras	MG	-21.72	-44.98	BHCB27981
E16	Santa Rita de Caldas	MG	-22.00	-46.38	BHCB014456
E17	Atibaia	SP	-23.17	-46.53	UEC070741
I01	Mucugê	BA	-13.00	-41.37	HUEFS0070811
I02	Rio de Contas	BA	-13.52	-41.94	HUEFS0206037
I03	Licínio de Almeida	BA	-14.69	-42.55	UFBA105815
I04	Serra Nova	MG	-15.65	-42.74	BHCB011996
I05	Grão Mogol	MG	-16.56	-42.90	IBT396396
I06	Cristália	MG	-16.72	-42.92	HUEFS0076782
I07	Joaquim Felício	MG	-17.69	-44.20	BHCB100399
I09	Conceição do Mato Dentro	MG	-19.09	-43.57	HUEFS0090623
I08	Diamantina	MG	-17.96	-43.78	NY00414802
I10	Santana do Riacho	MG	-19.33	-43.56	BHCB000352
I11	Caeté	MG	-19.82	-43.68	BHCB001030
I12	Catas Altas	MG	-20.08	-43.50	BHCB92794
M01	Aratuba	CE	-4.43	-39.05	EAC0049551
M02	Baturite	CE	-8.20	-36.41	EAC0049541
M03	Murici	AL	-9.27	-35.84	MAC0008050
M04	Ruy Barbosa	BA	-12.30	-40.48	HUEFS0117184
M05	Caraguatatuba	SP	-23.62	-45.42	HRCBNunes11
M06	Bombinhas	SC	-27.14	-48.48	JOI015001
P01	São Tomé das Letras	MG	-21.72	-44.98	HUSC11371
P02	Águas da Prata	MG	-21.92	-46.68	BHCBFiorini277
P03	Tibagi	PR	-24.56	-50.26	UPCB70034
T01	Ibituruna	MG	-22.06	-46.44	BHCBFiorini280
T02	Tibagi	PR	-24.56	-50.26	UPCB70033
W01	Mucugê	BA	-13.00	-41.37	HUEFS0085319

Pop	City	State	Lat	Lon	Voucher
W02	Rio de Contas	BA	-13.52	-41.94	HUEFS0206062
W03	Diamantina	MG	-17.96	-43.78	UEC064692
W04	Santana do Riacho	MG	-19.25	-43.51	HUEFS0162772
W05	Caeté	MG	-19.82	-43.68	BHCB56467
W06	Catas Altas	MG	-20.08	-43.50	BHCB92789

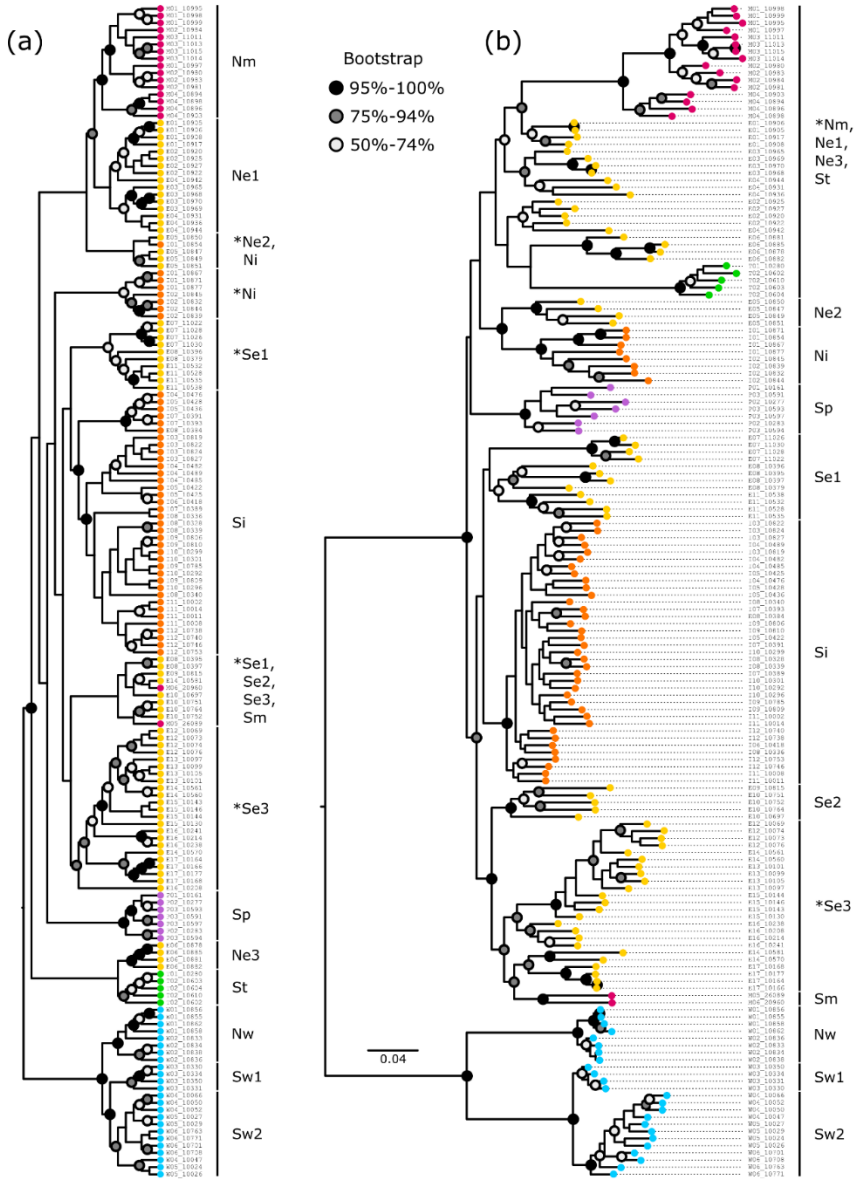
Supplementary material



Supplementary Figure 1. Selection of assembly parameters for stacks analysis. (a) Number of loci shared by 80% of samples or more and (b) the distribution of the number of SNPs per locus.



Supplementary Figure 2. Histogram of genetic variability of loci. The red line indicates the upper 95% quantile of variability.



Supplementary Figure 3. Phylogenetic analysis. (a) Individuals-partition SVDquartets tree, and (b) Individuals-partition RAxML tree. Codes of individuals are presented on b and c.

CHAPTER 2 – HERE, THERE AND EVERYWHERE: THE UBIQUITY OF HYBRIDIZATION IN AN ORCHID GROUP

Abstract

Genetic data shows that 'cryptic hybrids' are more common than previously thought and that hybridization and introgression are widespread processes in nature. Despite being common in plants, hybridization is not universal, with evidence of strong phylogenetic signal. Orchidaceae is a group with high hybridization propensity and several artificial orchids hybrids are known. Regardless of this, hybridization has not been considered one of the main drivers of diversification on this plant family. *Bulbophyllum* is one of the largest Orchidaceae genera, including 2,200 species and presents many examples of recent radiations, in which hybridization is theoretically more frequent. However, only three natural *Bulbophyllum* hybrids are currently recognized, all of them recently described based on morphological evidence. Both *B.* ×*cipoense* and *B.* ×*guartelae* are hybrids between species of the *B.* sect. *Didactyle*, and here we investigate the occurrence of hybridization in this section. We found that five from the seven species of the Neotropical *B.* sect. *Didactyle* are presently involved on hybridization in three species pairs. Despite the occurrence of hybridization, including subsequent generations of hybrids, there are no signs of backcrossing. Because hybridization presents high phylogenetic propensity, it suggests that hybridization might be a common process on the evolution of *Bulbophyllum* as a whole.

Introduction

Hybridization is defined as the outcrossing and gene flow between populations that differ at multiple heritable characters that affect fitness (Gompert & Buerkle, 2016). Hybridization was considered an "evolutionary dead end" (Seehausen, 2013), able "to lead only to deleterious effects" (Mayr, 1963), and, together with gene flow, as "mainly destructive forces with little evolutionary consequences" (Sætre, 2013). However, given renewed evidence, hybridization is now seeing as a creative force in the evolution of plants and animals (Abbott et al., 2013; Mallet, 2007). Genetic data show that 'cryptic hybrids' are found even in groups expected to show substantial barriers to gene flow, suggesting that hybridization could be a process even more common than suggested by non-molecular characters (Whitney, Ahern, Campbell, Albert, & King, 2010). Thus, both hybridization and introgressive hybridization (introgression; incorporation by hybridization and backcrossing of alleles from one species into the

gene pool of another species) are currently accepted as widespread processes in nature (Arnold, 1997; Mallet, 2005; Harrison & Larson, 2014).

Over the past century, studies on hybridization focused on two main aspects: understanding the role of hybridization in adaptation and the origin and maintenance of species reproductive barriers (Taylor & Larson, 2019). Hybridization can introduce new variation on which selection can act (Soltis & Soltis, 2009); as these new alleles had already being "tested", hybridization can act as an source of adaptative variation more powerful than mutation (Arnold & Martin, 2009; Burgarella et al., 2019; Suarez-Gonzalez, Lexer, & Cronk, 2018; Whitney et al., 2010). Loci that are not linked to reproductive isolation are more prone to introgression, and the regions promoting differentiation between lineages had been called "islands of differentiation", an idea popularized by Wu (2001), but already present on earlier works (Bazykin, 1969; Key, 1968). While hybridization can slow or reverse differentiation by means of gene flow and recombination, it may also lead to speciation by adaptative introgression (homoploid hybrid speciation) or cause instantaneous speciation via allopolypoidization (Abbott et al., 2013). Hybrid speciation is defined as "an speciation event in which hybridization has played a crucial role in the evolution of reproductive barriers between a hybrid lineage and its parental lineages" and many examples of natural homoploid hybrid speciation and allopolypoidization have been described (Taylor & Larson, 2019).

One of the main predictors of the chance of hybridization between two taxa is their divergence age (Abbott et al., 2013; Paun, Forest, Fay, & Chase, 2011). Low divergence levels are unlikely to bring major novelties; however, as lineages diverge Dobzhansky-Muller incompatibilities increase, possibly preventing the success of hybrids individuals (Levin, 2012; Scopece, Musacchio, Widmer, & Cozzolino, 2007). As incompatibilities are mainly affected by natural selection, they are not expected to evolve in clock-like steps (Mallet, 2005). Still, studies had shown that one million years are generally insufficient to generate hybrid sterility in plants, while taxa separated by more than four million years are likely to present pronounced hybrid infertility (Levin, 2012). Unsurprisingly, hybridization is exceptionally likely in rapidly diversifying adaptative radiations (Gourbière & Mallet, 2010; Seehausen, 2004), complicating phylogenetic inference (Payseur & Rieseberg, 2016; Chapter 1). The fact that hybridization is probable during early phases of divergence implies that the genetic variation of contemporary taxa could have been shaped by multiple events of hybridization in the past (Levin, 2012).

It is estimated that 25% of plant and 10% of animal species form hybrids (Mallet, 2005). The higher chance of hybridization in plants is hypothesized to be related to "the open, less integrative, and plastic patterns of plant morphogenesis", that allows larger genetic changes

(Gottlieb, 1984). Nearly 40% of the plant families and 16% of the plant genera in North America, Australia and Europe are involved in hybridization (Whitney et al., 2010). Despite being common, hybridization is not universal, with evidence of strong phylogenetic signal (λ=0.93; Whitney et al., 2010). Among the 25 larger plant families, Orchidaceae is the group with the higher hybridization propensity (weighted averages of hybridization propensities of the component genera): on average, 6% of all possible species combinations among species within genera of the family indeed form hybrids (Whitney et al., 2010). Also, a number of artificial orchids hybrids are known (Yam & Arditti, 2009). Regardless of this, hybridization has not been considered one of the main drivers of diversification on this plant family (Pace & Cameron, 2019). The absence of endosperm and the abundance of recent radiations observed in Orchidaceae has been suggested as the main hybridization boosters in this group (Johnson, 2018). Nevertheless, some orchids also present very specialized habitats and pollination systems that can act as reproductive barriers and hold hybridization (Johnson, 2018).

Bulbophyllum is one of the largest Orchidaceae genera, including 2,200 species (Pridgeon, Cribb, Chase, & Rasmussen, 2014). Despite its late Paleogene origin (~ 25 million years ago), *Bulbophyllum* presents many examples of recent radiations (Gamisch & Comes, 2019). However, only three natural *Bulbophyllum* hybrids are currently recognized (IPNI, 2020) – *B.* ×*chikukwa* (Afrotropical), *B.* ×*cipoense,* and *B.* ×*guartelae* (both Neotropical) – which were all recently described based on morphological evidence (Borba & Semir, 1998a; Fibeck & Mavi, 2000; Mancinelli & Smidt, 2012). Both *B.* ×*cipoense* and *B.* ×*guartelae* are hybrids between species of the *Bulbophyllum* sect. *Didactyle*. It has been suggested that only *B. weddellii* is a pollen receptor in the formation of *B.* × *cipoense* hybrids, since *B. weddellii*'s pollinarium size is not compatible with *B. involutum*'s stigmatic cavity, despite these species share their main pollinators (Borba & Semir, 1998a, 1998b). However, morphology indicates that introgression apparently occurs only in the opposite direction, with *B. involutum* as pollen receptor, since there is a range of intermediate *B. involutum* forms in multiple populations (Azevedo, Borba, & Van Den Berg, 2006). The hybrid origin of *B.* ×*cipoense* was tested with allozymes but there was no conclusive support for the hypothesis, probably due to marker resolution (Azevedo et al., 2006). In the case of *B.* × *guartelae*, only one individual with intermediate phenotype was found, however its existence suggests gene flow or introgression between the parental species, *B. perii* and *B. tripetalum* (Mancinelli & Smidt, 2012). No genetic test was performed to test the hybrid origin of *B.* ×*guartelae* so far. The *Didactyle* section includes also the *B. exaltatum* species complex and hybridization between the sister taxa *B. exaltatum* and *B. involutum* has been suggested, due to the continuum of morphological

variation among these species, but not formally tested or described (see Chapter 1). The polytopic origin of natural hybrids and introgression among lineages may be one of the factors responsible for the intricate morphological pattern of *Bulbophyllum* sect. *Didactyle*, especially in the *B. exaltatum* complex (Azevedo et al., 2006; Ribeiro et al., 2008; see Chapter 1).

As ancestral polymorphism, mutations and selection against intermediate characters can interfere with the hybrid phenotype, detection of hybrids is not always obvious (Leal, Chaves, Koehler, & Borba, 2016; Mallet, 2005; Pace & Cameron, 2019; Rieseberg, 1995). The advent of next-generation sequencing and genomic data sets allow more rigorous tests of hybridization (Goulet, Roda, & Hopkins, 2017; Twyford & Ennos, 2012). Due to recombination and meiosis independent assortment, unlinked loci are replicates outcomes of the hybridization process and allow precise and accurate reconstructions of the history of interbreeding (Payseur & Rieseberg, 2016). In this paper we intend to answer the following questions: (i) Does hybridization indeed occur between *B. weddellii* and *B. involutum* (*B. ×cipoense*), *B. perii* and *B. tripetalum* (*B. ×guartelae*), and *B. exaltatum* and *B. involutum*? (ii) If so, may these events relate to the complex morphological patterns observed in this group? (iii) Hybridization between the sister pair "*B. exaltatum* and *B. involutum*" is more widespread than hybridization between "*B. weddellii* and *B. involutum*", as expected due to the difference in divergence age? (iv) On sympatric localities, is it possible to find both parental and hybrid individuals?

Methodology

Sampling

To study the systems *B. weddellii*/*B. involutum*/*B. ×cipoense* (WIC), *B. tripetalum*/*B. perii*/*B. ×guartelae* (TPG), and *B. involutum*/*B. exaltatum* (IE) we sampled putative individuals of *B. weddellii* (30), *B. ×cipoense* (four, including the type specimen), *B. involutum* (77), *B. exaltatum* (80), *B. tripetalum* (10), *B. perii* (10), and *B. ×guartelae* (one, the type specimen), from 32 populations (23 localities, as some taxa are sympatric; Table 1; Fig. 2A, Fig. 3A and Fig. 4A). Sequences of part of these individuals were previously used on the biogeographical study of *Bulbophyllum* sect. *Didactyle* (see Chapter 1). We collected individuals growing on different rocks and a minimum of 10 m apart, to prevent sampling vegetative clones or closely related individuals (Hedrén & Lorenz, 2019). Vouchers were deposited in the herbaria and individuals propagated from field collected cuttings are also maintained in the living collection of MHNJB-UFMG (Supplementary table 1). All samples were collected under issued permits to CFF and ELB (SISBIO 52995-1, IEF 062/2016, and IAP 51.16).

Genomic library preparation and processing

We extracted Genomic DNA from fresh leaves (Doyle & Doyle, 1987) and prepared ddRAD libraries following a modified Peterson et al. (2012) protocol (Parchman et al., 2012). We size-selected fragments between 400–500 bp using Pippin Prep (Sage Science, Beverly, MA), and PCR-amplified these fragments using high-fidelity DNA polymerase (iProof, Bio-Rad, Hercules, CA), with 8 or 12 cycles. We sequenced individuals in four lanes of an Illumina HiSeq 2500 on Rapid Run Mode (in combination with samples from other projects) at The Centre for Applied Genomics, Hospital for Sick Children, Toronto, Canada, to generate 150 bp single end reads.

We processed genomic data using the Stacks 2.3e pipeline (Rochette & Catchen, 2017). We assembled de novo demultiplexed and filtered sequences with ustacks, build a catalogue of consensus loci in cstacks, identified individual genotypes with sstacks, organized data by locus with tsv2bam, and aligned reads and called SNPs with gstacks. The assembly parameters included a minimum depth of coverage, $m = 3$, mismatches allowed between two alleles of a sample, $M = 5$, and mismatches allowed between any two alleles of the catalog, $n = 6$ (i.e., the optimal parameters based on *r80*, see Chapter 1), and an upper bound for $\varepsilon = 0.1$, a minimum minor allele frequency = 0.02, and a maximum observed heterozygosity = 0.5.

For each of the systems, we grouped individuals from each species by populations according to their geographic sampling localities, and retained biallelic loci from a minimum of two populations, to maximize the number of loci (Huang & Knowles, 2016). To guard against sequencing and assembly errors, we used a custom R script (Thomaz, Malabarba, & Knowles, 2017) to exclude SNPs with theta values within the upper 95% quantile of variability (see Supplementary Fig. 1). We used the software plink 1.9 (Purcell et al., 2007) to identify SNPs with a maximum of 25% or 40% of missing data to create two separate datasets, because the robustness of analyses to missing data differ. The sequencing throughput for each of the systems is shown on Table 2.

Genetic differentiation and hybridization

For each of the systems, a principal component analysis (PCA) was used to visualize the distribution of genomic variation using adegenet 2.1.1 (Jombart and Ahmed, 2011), in R 3.5.0 (R Core Team, 2014). As a multivariate method, PCA summarizes the genetic similarity among populations and genotypes without requiring strong assumptions about Hardy–Weinberg equilibrium or linkage disequilibrium. However, due to its sensibility to missing data,

we used the genomic dataset with all SNPs and a threshold of 25% of missing data, with missing data values replaced by the per locus mean allele frequencies for a given population.

We used gghybrid 0.0.0.9000 (Bailey, 2018), a R package which allows hypothesis-testing on bi-allelic genomic data through Bayesian estimate of the hybrid-index (proportion of allele copies coming from one of the two parental reference sets; Buerkle, 2005). Based on morphology, we set the following populations as pure: (i) W04 (*B. weddellii*) and I10 (*B. involutum*) for system WIC; (ii) P03 (*B. perii*) and T02 (*B. tripetalum*) for system TPG; and (iii) I04 and I10 (*B. involutum*), and E12 and E17 (*B. exaltatum*) for system IE. We removed loci for which the difference in allele frequency between parental reference sets was less than 0.8 for systems WIC and TPG (resulting in a total of 190 and 167 unlinked SNPs, respectively). Given the smaller divergence time between *B. exaltatum* and *B. involutum*, we removed loci for which the difference in allele frequency between parental reference sets was less than 0.25 for system IE (resulting in a total of 213 unliked SNPs). For all systems we run a total of 10,000 MCMC iterations, including 10% of burn-in.

Also, for each of the systems the software parallelenewhybrid 1.0.1 (Wringe et al., 2017) was used to implement NewHybrids 1.1 Beta 3 in parallel (Anderson and Thompson, 2002). NewHybrids is a Bayesian model-based method capable of computing the posterior probability that each individual belongs to distinct pure or hybrid classes (F1, F2, and backcrosses), based on data from multiple markers and does not require parental species assignment, nor pure samples from the parental species. To test the existence of hybrids individuals we used 90,000 steps and a burn-in of 10,000 steps. We did not assign individuals to pure or hybrid classes a priori. For NewHybrids analysis, we used the loci set obtained after gghybrids processing, totalizing 190, 167 and 213 unlinked SNPs for systems WIC, TPG and IE, respectively.

To estimate population structure for each of the systems, we used fastStructure 1.0, an accurate variational Bayesian framework compatible with large data sets (Raj et al. 2014). To create the bed, bim and fam files required by fastStructure, we convert ped and map files from stacks 2.43 using plink 1.9. We estimate ancestry proportions for each individual for K = 2 using the structure.py script (included within the package), using 10 replicates. We visualized the results with the online application Clumpak (available at http://clumpak.tau.ac.il; Kopelman, Mayzel, Jakobsson, Rosenberg, & Mayrose, 2015). For fastStructure analysis we used a set of unlinked SNPs with a maximum of 25% of missing data.

As the IE system is expected to have diverged recently, we used HyDe 0.4.1a to infer introgression despite incomplete lineage sorting. HyDe is a Python package capable of detecting hybridization using a model that simultaneously considers coalescence and

hybridization, using phylogenetic invariants. We tested per-individual variation in the amount of hybridization using the individual_hyde_mp.py script, using a matrix with 40% of missing data and all loci. *B. weddellii* was set as the outgroup and, based on morphology, populations I04 and I10 as the pure populations for *B. involutum* and populations I12 and I17 as pure populations for *B. exaltatum*.

Results

WIC system

All analyses support the hypothesis of hybrid origin of *B. ×cipoense* individuals (Fig. 2). However, neither *B. involutum* nor *B. weddellii* showed signs of introgression, even in sympatric localities (populations I08 + W03, I11 + W05, and I12 + W06, Fig. 2). The analysis support however that *B. ×cipoense* individuals are genetically closer to *B. involutum* than to *B. weddellii* (Fig. 2B, C, and D).

The first axis of PCA clearly separates *B. involutum* and *B. weddellii*, with *B. ×cipoense* on an intermediate position. On the second axis, population W03 is segregated from other *B. weddellii* populations (Fig. 2B). FastStructure and gghybrids presented similar results, with *B. ×cipoense* showing in-between values of ancestry proportion and hybrid index, but closer to *B. involutum* (Fig. 2C and D). Both analyses support that all the other individuals belong to pure lineages, agreeing with NewHybrids results. Yet, NewHybrids indicates that *B. ×cipoense* are F2 hybrids (Fig. 2E).

TPG system

Like system WIC, all analyses support the hypothesis of hybrid origin of the *B. ×guartelae* individual (Fig. 3). Also, the genetic analysis showed that one of the individuals identified as *B. perri* based on remnants of the inflorescence was actually the second register of *B. ×guartelae*. Neither *B. perii* nor *B. tripetalum* showed signs of introgression, even in the sympatric locality (populations P03 + T02, Fig. 3). The analyses support that *B. ×guartelae* individuals are an equivalent mixture of *B. perii* and *B. tripetalum* genomes (Fig. 3B, C, and D). Both fastStructure and gghybrids presented similar results, with *B. ×guartelae* showing intermediate values of ancestry proportion and hybrid index (~0.5). Both analyses support that all the other individuals belong to pure lineages, agreeing with NewHybrids results. Yet, NewHybrids also indicates that *B. ×guartelae* are F2 hybrids (Fig. 3E).

IE system

As systems WIC and TPG, system IE shows signs of hybridization. However, on system IE individuals with hybrid genomic composition are widespread trough some *B. exaltatum* populations (E08, E10, E11, E14, E16 and, possibly, E09; Fig. 4C, D and F). *B. exaltatum* populations E13, E15 and E17 and all populations of *B. involutum* show no signs of individuals with hybrid composition.

The first axis of PCA separates *B. involutum* and *B. exaltatum*, with individuals identified as F2 by NewHybrids on intermediate position (Fig. 4B). The second axis mainly segregates *B. exaltatum* populations. As a general pattern, fastStructue and gghybrids indicate that the smaller the latitude (and closer the distance to the center of *B. involutum* distribution), the higher is the proportion of *B. involutum* genome on *B. exaltatum* individuals (Fig. 4A, C and D). Generally, HyDe results presented low significance for most individuals. Despite this, gamma values give support to the results observed in other analysis, suggesting that some *B. exaltatum* individuals are genetically closer to *B. involutum* than to other co-specific individuals (Fig. 4E). NewHybrids suggests that the individuals with hybrid ancestry are F2 hybrids, with low probability of backcrossing with *B. involutum* or *B. exaltatum* in populations E08 and E16, respectively (Fig. 4F).

Discussion

The results support our main hypothesis, confirming the existence of hybrids on systems *B. weddellii*/*B. involutum* (*B. ×cipoense*) (WIC), *B. tripetalum*/*B. perii* (*B. ×guartelae*) (TPG), and *B. involutum*/*B. exaltatum* (IE). In addition, our analyses indicate that despite the occurrence of hybridization with subsequent generations of hybrids, there are no signs of backcrossing. Therefore, five of the seven pure species currently circumscribed on *Bulbophyllum* sect. *Didactyle* are involved in the formation of hybrids. Because hybridization presents high phylogenetic propensity, it suggests that hybridization might be a common process on the evolution of *Bulbophyllum* as a whole.

Hybridization in Bulbophyllum *sect.* Didactyle

Despite species are frequently seen as discrete and fundamental units, the rise of reproductive isolation can take millions of years after initial divergence (Mallet, 2005). All the systems studied here (i.e. WIC, TPG, and IE) support this idea. The initial divergence between *Bulbophyllum* sect. *Didactyle* species occurred 2.16 million years ago (Gamisch & Comes,

2019), but at least five of the seven currently circumscribed taxa are involved in hybridization in some level. Indeed, it has been previously shown that *B. weddellii*, *B. involutum* and *B. exaltatum* are interfertile (Borba, Shepherd, & Semir, 1999). Hybrid individuals are more frequent entities in some populations of system IE, in which parentals are very closely related and floral morphology is quite similar as compared to the two other systems. Indeed, some IE populations are apparently completely formed by F2 individuals (i.e., E08, E10 and E11). In this way our results show that the level of co-ancestry between two species (see Fig 4. in Chapter 1) is related to their hybridization propensity, supporting findings of other studies (Levin, 2012). Meanwhile, *B.* ×*cipoense* (systems WIC) and *B.* ×*guartelae* (system TPG) are apparently rare (Borba & Semir, 1998a; Mancinelli & Smidt, 2012). Despite we find no backcrossing individuals, the presence of hybrids may allow some degree of gene flow and even low rates of hybridization can have impacts on all the species (Mallet, 2005).

Our study does not support the idea that the morphological variation observed in *B. involutum* is as a result of hybridization with *B. weddellii*, as suggested by Azevedo et al. (2006). *B. involutum* individuals are mainly pure, as occurs to *B. weddellii*, *B. perii*, and *B. tripetalum*. Differently, a portion of the individuals identified as *B. exaltatum* presented some degree of *B. involutum* genome. Part of the morphological obscurity in the *B. exaltatum* species complex is probably a result of the presence of these individuals of mixed ancestry. Nevertheless, given the new genomic evidence, it is now clear that individuals with mixed ancestry are morphologically distinct from pure individuals (Fig. 1), what will facilitate future field identification.

It is important to highlight the geographic factor of the distribution of populations with hybrid ancestry in *B. exaltatum*. Some authors distinguish between localized and dispersed hybridization, depending on whether individuals with mixed ancestry are found only where the two parental types are present or whether populations far from the hybrid zone are also admixed (Harrison & Larson, 2014). Our results support dispersed hybridization on the IE system, as *B. involutum* genes are present on populations outside the area of distribution of this species. However, as a matter of fact, no population from system IE could be considered a hybrid zone, as none of them presented parental species accompanied by multiple generations of hybrids. Our results suggest that individuals with mixed ancestry could form a new hybrid species (polyphyletic in IE), as no backcrossing was observed. It is not clear, however, how hybridization might had contributed to the formation of this putative new lineage (hybridization speciation versus adaptative radiation; Abbott et al., 2013). It is important to consider that "admixture could represent what remains after hybrid ancestry has been purged from critical

regions of the genome" (Taylor & Larson, 2019) and that "shared variation among populations may reflect unsorted shared ancestral polymorphism" (Payseur & Rieseberg, 2016). HyDe results support the idea of hybridization instead of incomplete lineage sorting, but the test requires a larger number of loci to give undoubtful results for all individuals (Blischak, Chifman, Wolfe, & Kubatko, 2018). Functional gene annotation and trait-based studies connecting admixture with reproductive barriers are required to confirm the existence of adaptative introgression and hybrid speciation, respectively (Abbott et al., 2013; Taylor & Larson, 2019). Both studies are highly recommended to better understand the evolutionary history and consequence of hybridization on the IE system and confirm the existence of this lineage with hybrid origin.

According to NewHybrids, the hybrids we identified are mainly F2 hybrids. However, it is important to highlight that in systems WIC and TPG hybrids individuals are rare and parental individuals are frequent, suggesting that the formation of F1 individuals must be more likely. The occurrence of incomplete lineage sorting or of insufficient sample of genetic variability (i.e. genotypes of actual individual parents of hybrids are missing) could bias our analysis, in this way we must be cautious in assuming all identified hybrids are really F2 hybrids.

It is noteworthy that *B. involutum* is considerably more abundant than *B. exaltatum* when in sympatry (pers. obs.). This fact can possibly impact hybridization outcomes, given the relevance of demographic factors to this process (Currat, Ruedi, Petit, & Excoffier, 2008; Klein, Lagache-Navarro, & Petit, 2017). The asymmetric character of hybridization in IE system (i.e. individuals morphologically assigned to *B. involutum* are pure and individuals assigned to *B. exaltatum* can be either pure or hybrids) is not uncommon in nature (Folk, Soltis, Soltis, & Guralnick, 2018) and the disjunct aspect of the campos rupestres, the herbaceous-shrubby vegetation mosaic in eastern Brazil where species from the IE system are mainly distributed (Fig. 4A), can also impact the demography of hybridization. The fact that populations are in isolated outcrops can lead to limited gene flow and rise differentiation and local adaptation (see Chapter 1).

Hybridization and the diversification of Bulbophyllum *species*

Hybridization propensity presents strong and geographically consistent phylogenetic signal (λ=0.93), suggesting that it might be an intrinsic propriety of biologic groups instead of a function of environmental conditions (Whitney et al., 2010). There are exceptions to this general pattern and environmental discontinuity and pollinator specialization may act as

hybridization hampers (Johnson, 2018). The fact that hybrids are abundant in *Bulbophyllum* sect. *Didactyle* is an indication that it might be a frequent phenomenon in *Bulbophyllum* species in general, given the abundance of sections of recent radiation (Gamisch & Comes, 2019). It has been suggested indeed that hybridization itself might be an important promoter of adaptative radiations, as it could boost the availability of genetic and phenotypic novelty (Seehausen, 2004). Also, it is expected that in herbs hybridization rates are higher than that observed for trees, due to shorter generation times (Levin, 2012). However, it is important to highlight that some *Bulbophyllum* species present slow growth, with long expected generation times (~10 years). Still, our understanding of the factors driving orchid hybridization is scarce and a better knowledge of factors driving reproductive barriers is required.

It is noteworthy to emphasize that molecular investigations are important in identifying future *Bulbophyllum* hybrids. As morphological characters are the result of the interplay of many genes and can be plastic (Rieseberg & Ellstrand, 1993), morphological intermediaries can be absent or misleading (e.g., de Hert et al., 2011; Leal et al., 2016; Pace & Cameron, 2019; Wallace, 2006).

Conservation

Despite being considered unworthy of conservation in the past (O'brien & Mayr, 1991), new policies shall be actualized considering the new evidence that natural hybrids are fundamental for the origin and maintenance of biodiversity, including hybrid speciation (Gompert & Buerkle, 2016; vonHoldt, Brzeski, Wilcove, & Rutledge, 2018). Conservation measures to protect both pure and hybrid populations.

Open questions

The study of orchid hybridization is still in its infancy and we still have much to learn about how genome evolves after hybridization (Taylor & Larson, 2019). Questions about the origin and maintenance of reproductive barriers are also still open, as how many genomic regions differentiate during speciation and how these regions are dispersed around the genome (Abbott et al., 2013). Documenting the variation of introgression rates across genome and time are also an interesting issue (Payseur & Rieseberg, 2016). As species that currently hybridize may offer exceptional insights into the genomics of hybridization, a deeper study of the hybridization process within *Bulbophyllum* sect. *Didactyle*, specially of system IE, can be a key to better understand the speciose genus *Bulbophyllum*.

Conclusion

After the divergence of two species, parts of their genome are still likely to introgress (Mallet, 2005). Here we show that five from the seven currently circumscribed species of *Bulbophyllum* sect. *Didactyle* are presently involved on hybridization. As envisaged by the fact that species with more recent common ancestry are expected to present higher fertility rates (Levin, 2012), hybridization is much more geographically and genetically widespread on system IE than in systems WIC or TPG. We did not observe F1 or introgressed individuals in any of the studied systems, suggesting that the formation of F1 hybrids or backcrossed individuals are rare events. The geographic distribution of populations from system IE indicates yet that the formation of hybrids can be an important factor for adaptative divergence and consequent diversification (Prentis, Wilson, Dormontt, Richardson, & Lowe, 2008) of *B. exaltatum*. Future research will shed light on adaptative introgression (functional gene annotation) and connections between admixture with reproductive barriers (trait-based studies). As it has been observed that the hybridization propensity of a genus in a region is predictive of its general hybridization propensity (Whitney et al., 2010), the fact that hybridization is so abundant in *Bulbophyllum* sect. *Didactyle* may be an indication that this process is also common across other sections of the genus. If so, hybridization may play an import role on the diversification of the *Bulbophyllum*, in which recent radiations are abundant.

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., … Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, *26*(2), 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Arnold, M. L. (1997). *Natural hybridization and evolution*. New York: Oxford University Press.

Arnold, M. L., & Martin, N. H. (2009). Adaptation by introgression. *Journal of Biology*, *8*(9), 82. https://doi.org/10.1186/jbiol176

Azevedo, C. O., Borba, E. L., & Van Den Berg, C. (2006). Evidence of natural hybridization and introgression in *Bulbophyllum involutum* Borba, Semir & F. Barros and *B. weddellii* (Lindl.) Rchb. f. (Orchidaceae) in the Chapada Diamantina, Brazil, by using allozyme markers. *Revista Brasileira de Botânica*, *29*(3), 415–421. https://doi.org/10.1590/S0100-84042006000300008

Bazykin, A. D. (1969). Hypothetical mechanism of speciation. *Evolution*, *23*(4), 685–687. https://doi.org/10.1111/j.1558-5646.1969.tb03550.x

Blischak, P. D., Chifman, J., Wolfe, A. D., & Kubatko, L. S. (2018). HyDe: a python package for genome-scale hybridization detection. *Systematic Biology*, *67*(5), 821–829. https://doi.org/10.1093/sysbio/syy023

Borba, E. L., & Semir, J. (1998a). *Bulbophyllum ×cipoense* (Orchidaceae), a new natural hybrid from the Brazilian" campos rupestres": description and biology. *Lindleyana*, Vol. 13, pp. 113–120.

Borba, E. L., & Semir, J. (1998b). Wind-assisted fly pollination in three *Bulbophyllum* (Orchidaceae) species occurring in the Brazilian campos rupestres. *Lindleyana*, *13*(3), 203–218.

Borba, E. L., Shepherd, G. J., & Semir, J. (1999). Reproductive systems and crossing potential in three species of *Bulbophyllum* (Orchidaceae) occurring in Brazilian “campo rupestre” vegetation. *Plant Systematics and Evolution*, *217*(3–4), 205–214. https://doi.org/10.1007/BF00984366

Burgarella, C., Barnaud, A., Kane, N. A., Jankowski, F., Scarcelli, N., Billot, C., … Berthouly-Salazar, C. (2019). Adaptive introgression: an untapped evolutionary mechanism for crop adaptation. *Frontiers in Plant Science*, *10*(February), 1–17. https://doi.org/10.3389/fpls.2019.00004

Currat, M., Ruedi, M., Petit, R. J., & Excoffier, L. (2008). The hidden side of invasions: massive introgression by local genes. *Evolution*, *62*(8), 1908–1920. https://doi.org/10.1111/j.1558-5646.2008.00413.x

de Hert, K., Jacquemyn, H., Van Glabeke, S., Roldán-Ruiz, I., Vandepitte, K., Leus, L., & Honnay, O. (2011). Patterns of hybridization between diploid and derived allotetraploid species of *Dactylorhiza* (Orchidaceae) co-occurring in Belgium. *American Journal of Botany*, *98*(6), 946–955. https://doi.org/10.3732/ajb.1000367

Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, *19*(1), 11–15.

Ellstrand, N. C., Whitkus, R., & Rieseberg, L. H. (1996). Distribution of spontaneous plant hybrids. *Proceedings of the National Academy of Sciences*, *93*(10), 5090–5093. https://doi.org/10.1073/pnas.93.10.5090

Fibeck, W., & Mavi, S. (2000). A natural orchid hybrid from Zimbabwe. *Kirkia*, *17*(2), 147–149. Retrieved from http://www.jstor.org/stable/23502311

Folk, R. A., Soltis, P. S., Soltis, D. E., & Guralnick, R. (2018). New prospects in the detection and comparative analysis of hybridization in the tree of life. *American Journal of Botany*, *105*(3), 364–375. https://doi.org/10.1002/ajb2.1018

Gamisch, A., & Comes, H. P. (2019). Clade-age-dependent diversification under high species turnover shapes species richness disparities among tropical rainforest lineages of *Bulbophyllum* (Orchidaceae). *BMC Evolutionary Biology*, *19*(1), 93. https://doi.org/10.1186/s12862-019-1416-1

Gompert, Z., & Buerkle, C. A. (2016). What, if anything, are hybrids: enduring truths and challenges associated with population structure and gene flow. *Evolutionary Applications*, *9*(7), 909–923. https://doi.org/10.1111/eva.12380

Gottlieb, L. D. (1984). Genetics and morphological evolution in plants. *The American Naturalist*, *123*(5), 681–709. Retrieved from http://www.jstor.org/stable/2461245 .

Goulet, B. E., Roda, F., & Hopkins, R. (2017). Hybridization in plants: old ideas, new techniques. *Plant Physiology*, *173*(1), 65–78. https://doi.org/10.1104/pp.16.01340

Gourbière, S., & Mallet, J. (2010). Are species real? The shape of the species boundary with exponential failure, reinforcement, and the "missing snowball." *Evolution*, *64*(1), 1–24. https://doi.org/10.1111/j.1558-5646.2009.00844.x

Harrison, R. G., & Larson, E. L. (2014). Hybridization, introgression, and the nature of species boundaries. *Journal of Heredity*, *105*(S1), 795–809. https://doi.org/10.1093/jhered/esu033

Hedrén, M., & Lorenz, R. (2019). Seed dispersal and fine-scale genetic structuring in the asexual *Nigritella miniata* (Orchidaceae) in the Alps. *Botanical Journal of the Linnean Society*, *190*(1), 83–100. https://doi.org/10.1093/botlinnean/boz005

Huang, H., & Knowles, L. L. (2016). Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. *Systematic Biology*, *65*(3), 357–365. https://doi.org/10.1093/sysbio/syu046

IPNI. (2020). International plant names index: *Bulbophyllum*. Retrieved January 2, 2020, from https://www.ipni.org/?q=Bulbophyllum

Johnson, S. D. (2018). Natural hybridization in the orchid flora of South Africa: comparisons among genera and floristic regions. *South African Journal of Botany*, *118*, 290–298. https://doi.org/10.1016/j.sajb.2018.01.011

Key, K. H. L. (1968). The concept of stasipatric speciation. *Systematic Biology*, *17*(1), 14–22. https://doi.org/10.1093/sysbio/17.1.14

Klein, E. K., Lagache-Navarro, L., & Petit, R. J. (2017). Demographic and spatial determinants of hybridization rate. *Journal of Ecology*, *105*(1), 29–38. https://doi.org/10.1111/1365-2745.12674

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: a program for identifying clustering modes and packaging population structure inferences across *K*. *Molecular Ecology Resources*, *15*(5), 1179–1191. https://doi.org/10.1111/1755-0998.12387

Leal, B. S. S., Chaves, C. J. N., Koehler, S., & Borba, E. L. (2016). When hybrids are not hybrids: a case study of a putative hybrid zone between *Cattleya coccinea* and *C. brevipedunculata* (Orchidaceae). *Botanical Journal of the Linnean Society*, *181*(4), 621–639. https://doi.org/10.1111/boj.12437

Levin, D. A. (2012). The long wait for hybrid sterility in flowering plants. *New Phytologist*, *196*(3), 666–670. https://doi.org/10.1111/j.1469-8137.2012.04309.x

Mallet, J. (2005). Hybridization as an invasion of the genome. *Trends in Ecology and Evolution*, *20*(5), 229–237. https://doi.org/10.1016/j.tree.2005.02.010

Mallet, J. (2007). Hybrid speciation. *Nature*, *446*(7133), 279–283. https://doi.org/10.1038/nature05706

Mancinelli, W. S., & Smidt, E. de C. (2012). *O gênero* Bulbophyllum *(Orchidaceae) na Região Sul do Brasil*. *63*(4), 803–815.

Mayr, E. (1963). *Animal species and evolution*. Cambridge: Harvard University Press.

O'brien, S. J., & Mayr, E. (1991). Bureaucratic mischief: recognizing endangered species. *Source: Science, New Series*, *251*(4998), 1187–1188. Retrieved from https://pdfs.semanticscholar.org/d9c6/f4e3ddbc98ea0145484e9add7c045a260aee.pdf

Pace, M. C., & Cameron, K. M. (2019). The evolutionary and systematic significance of hybridization between taxa of *Spiranthes* (Orchidaceae) in the California Sierra Nevada and Cascade Range. *Taxon*, *68*(May 2018), 199–217. https://doi.org/10.1002/tax.12034

Parchman, T. L., Gompert, Z., Mudge, J., Schilkey, F. D., Benkman, C. W., & Buerkle, C. A. (2012). Genome-wide association genetics of an adaptive trait in lodgepole pine. *Molecular Ecology*, *21*(12), 2991–3005. https://doi.org/10.1111/j.1365-294X.2012.05513.x

Paun, O., Forest, F., Fay, M. F., & Chase, M. W. (2011). Parental divergence and hybrid speciation in angiosperms revisited. *Taxon*, *60*(5), 1241–1244. https://doi.org/10.1002/tax.605001

Payseur, B. A., & Rieseberg, L. H. (2016). A genomic perspective on hybridization and speciation. *Molecular Ecology*, *25*(11), 2337–2360. https://doi.org/10.1111/mec.13557

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS One*, *7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Prentis, P. J., Wilson, J. R. U., Dormontt, E. E., Richardson, D. M., & Lowe, A. J. (2008). Adaptive evolution in invasive species. *Trends in Plant Science*, *13*(6), 288–294. https://doi.org/10.1016/j.tplants.2008.03.004

Pridgeon, A. M., Cribb, P. J., Chase, M. W., & Rasmussen, F. N. (2014). *Genera Orchidacearum, vol. 6: Epidendroideae, part 3* (1st ed.). Oxford: Oxford University Press.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., … Sham, P. C. (2007). *PLINK: a tool set for whole-genome association and population-based linkage analyses*. *81*(3), 559–575.

Ribeiro, P. L., Borba, E. L., Smidt, E. de C., Lambert, S. M., Schnadelbach, A. S., & van den Berg, C. (2008). Genetic and morphological variation in the *Bulbophyllum exaltatum* (Orchidaceae) complex occurring in the Brazilian "campos rupestres": implications for taxonomy and biogeography. *Plant Systematics and Evolution*, *270*(1–2), 109–137. https://doi.org/10.1007/s00606-007-0603-5

Rieseberg, L. H. (1995). The role of hybrization in evolution: old wine in new skins. *American Journal of Botany*, *82*(7), 944–953. https://doi.org/10.2307/2445981

Rieseberg, L. H., & Ellstrand, N. C. (1993). What can molecular and morphological markers tell us about plant hybridization? *Critical Reviews in Plant Sciences*, *12*(3), 213–241. https://doi.org/10.1080/07352689309701902

Rochette, N. C., & Catchen, J. M. (2017). Deriving genotypes from RAD-seq short-read data using Stacks. *Nature Publishing Group*, *12*(12), 2640–2659. https://doi.org/10.1038/nprot.2017.123

Sætre, G. P. (2013). Hybridization is important in evolution, but is speciation? *Journal of Evolutionary Biology*, *26*(2), 256–258. https://doi.org/10.1111/jeb.12005

Scopece, G., Musacchio, A., Widmer, A., & Cozzolino, S. (2007). Patterns of reproductive isolation in Mediterranean deceptive orchids. *Evolution*, *61*(11), 2623–2642. https://doi.org/10.1111/j.1558-5646.2007.00231.x

Seehausen, O. (2004). Hybridization and adaptive radiation. *Trends in Ecology and Evolution*, *19*(4), 198–207. https://doi.org/10.1016/j.tree.2004.01.003

Seehausen, O. (2013). Conditions when hybridization might predispose populations for adaptive radiation. *Journal of Evolutionary Biology*, *26*(2), 279–281. https://doi.org/10.1111/jeb.12026

Soltis, P. S., & Soltis, D. E. (2009). The role of hybridization in plant speciation. *Annual Review of Plant Biology*, *60*(1), 561–588. https://doi.org/10.1146/annurev.arplant.043008.092039

Suarez-Gonzalez, A., Lexer, C., & Cronk, Q. C. B. (2018). Adaptive introgression: a plant perspective. *Biology Letters*, *14*(3). https://doi.org/10.1098/rsbl.2017.0688

Taylor, S. A., & Larson, E. L. (2019). Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nature Ecology and Evolution*, *3*(2), 170–177. https://doi.org/10.1038/s41559-018-0777-y

Thomaz, A. T., Malabarba, L. R., & Knowles, L. L. (2017). Genomic signatures of paleodrainages in a freshwater fish along the southeastern coast of Brazil: genetic structure reflects past riverine properties. *Heredity*, *119*(4), 287–294. https://doi.org/10.1038/hdy.2017.46

Twyford, A. D., & Ennos, R. A. (2012). Next-generation hybridization and introgression. *Heredity*, *108*(3), 179–189. https://doi.org/10.1038/hdy.2011.68

vonHoldt, B. M., Brzeski, K. E., Wilcove, D. S., & Rutledge, L. Y. (2018). Redefining the Role of Admixture and Genomics in Species Conservation. *Conservation Letters*, *11*(2), e12371. https://doi.org/10.1111/conl.12371

Wallace, L. E. (2006). Spatial genetic structure and frequency of interspecific hybridization in *Platanthera aquilonis* and *P. dilatata* (Orchidaceae) occurring in sympatry. *American Journal of Botany*, *93*(7), 1001–1009. https://doi.org/10.3732/ajb.93.7.1001

Whitney, K. D., Ahern, J. R., Campbell, L. G., Albert, L. P., & King, M. S. (2010). Patterns of hybridization in plants. *Perspectives in Plant Ecology, Evolution and Systematics*, *12*(3), 175–182. https://doi.org/10.1016/j.ppees.2010.02.002

Wu, C. I. (2001). The genic view of the process of speciation. *Journal of Evolutionary Biology*, *14*(6), 851–865. https://doi.org/10.1046/j.1420-9101.2001.00335.x

Yam, T. W., & Arditti, J. (2009). History of orchid propagation: a mirror of the history of biotechnology. *Plant Biotechnology Reports*, *3*(1), 1–56. https://doi.org/10.1007/s11816-008-0066-3

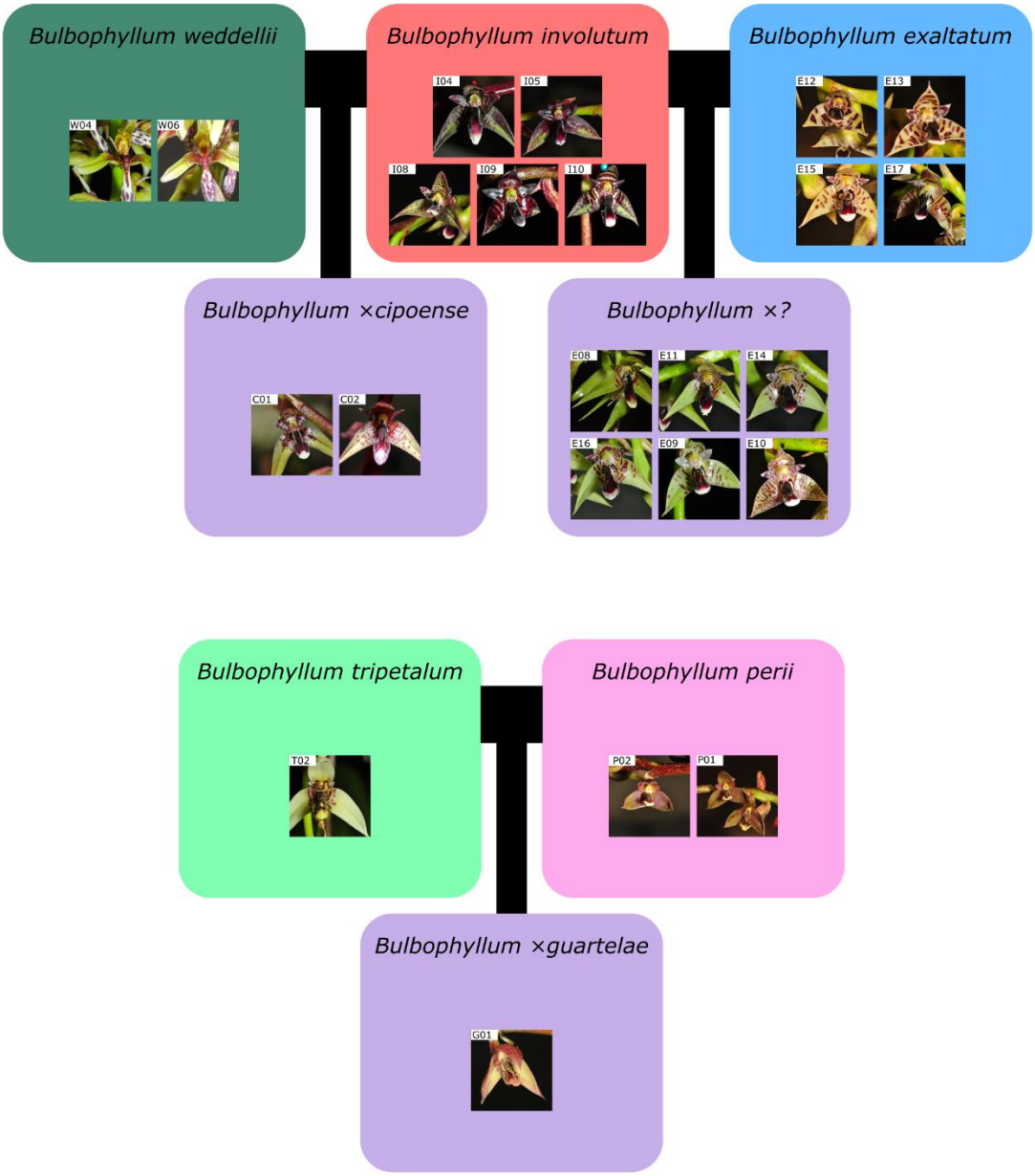


Figure 1. Morphological variability of *Bulbophyllum* sect. *Didactyle* hybrid systems. Populations names are given. When morphologically dubious, the class assignment was based on gghybrids results.

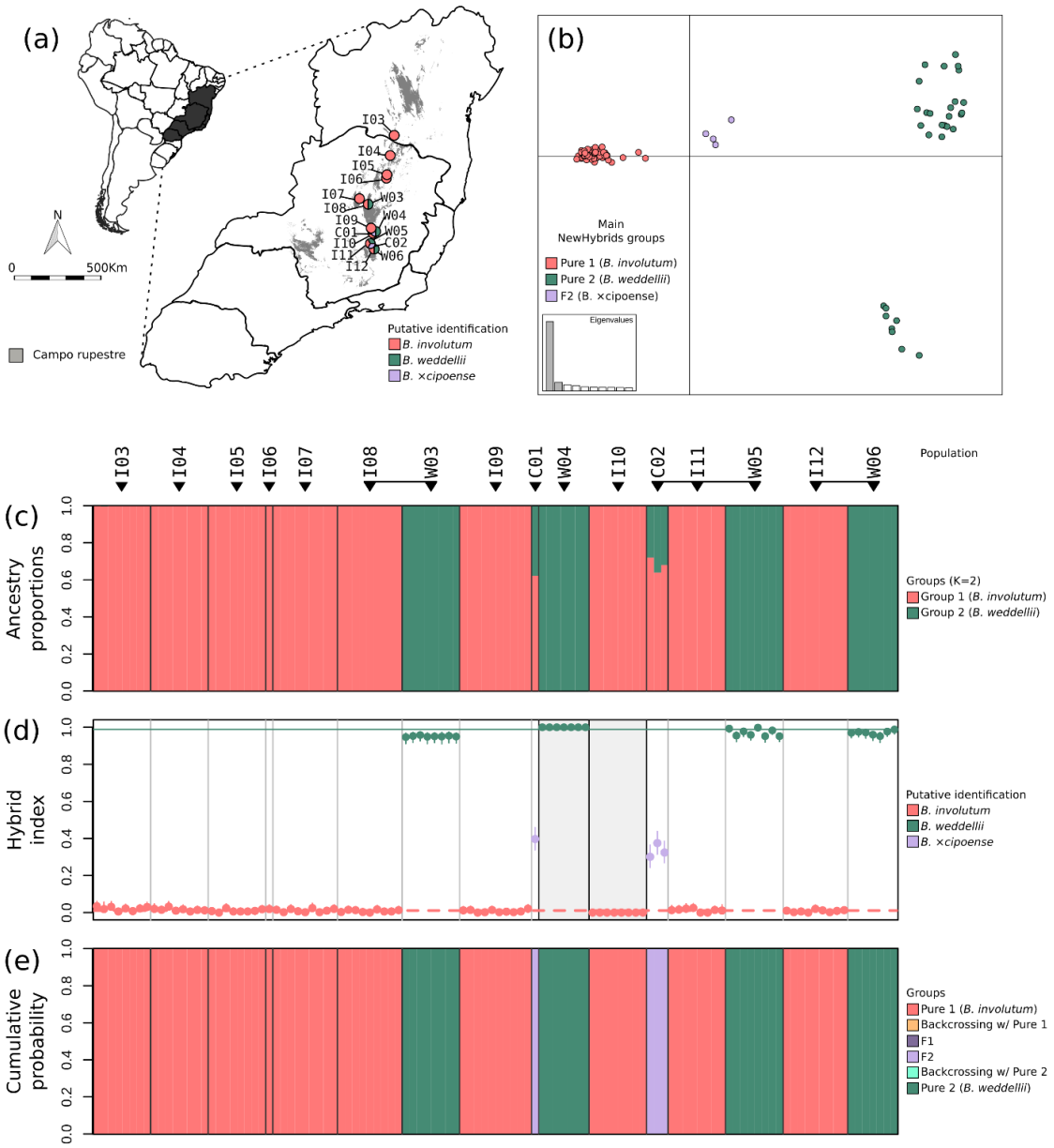


Figure 2. Hybridization in system WIC (*B. weddellii*/*B. involutum*/*B. ×cipoense*). (a) geographic distribution of populations; (b) PCA results; (c) Faststructure results for K = 2; (d) gghybrid results; (e) NewHybrids results.

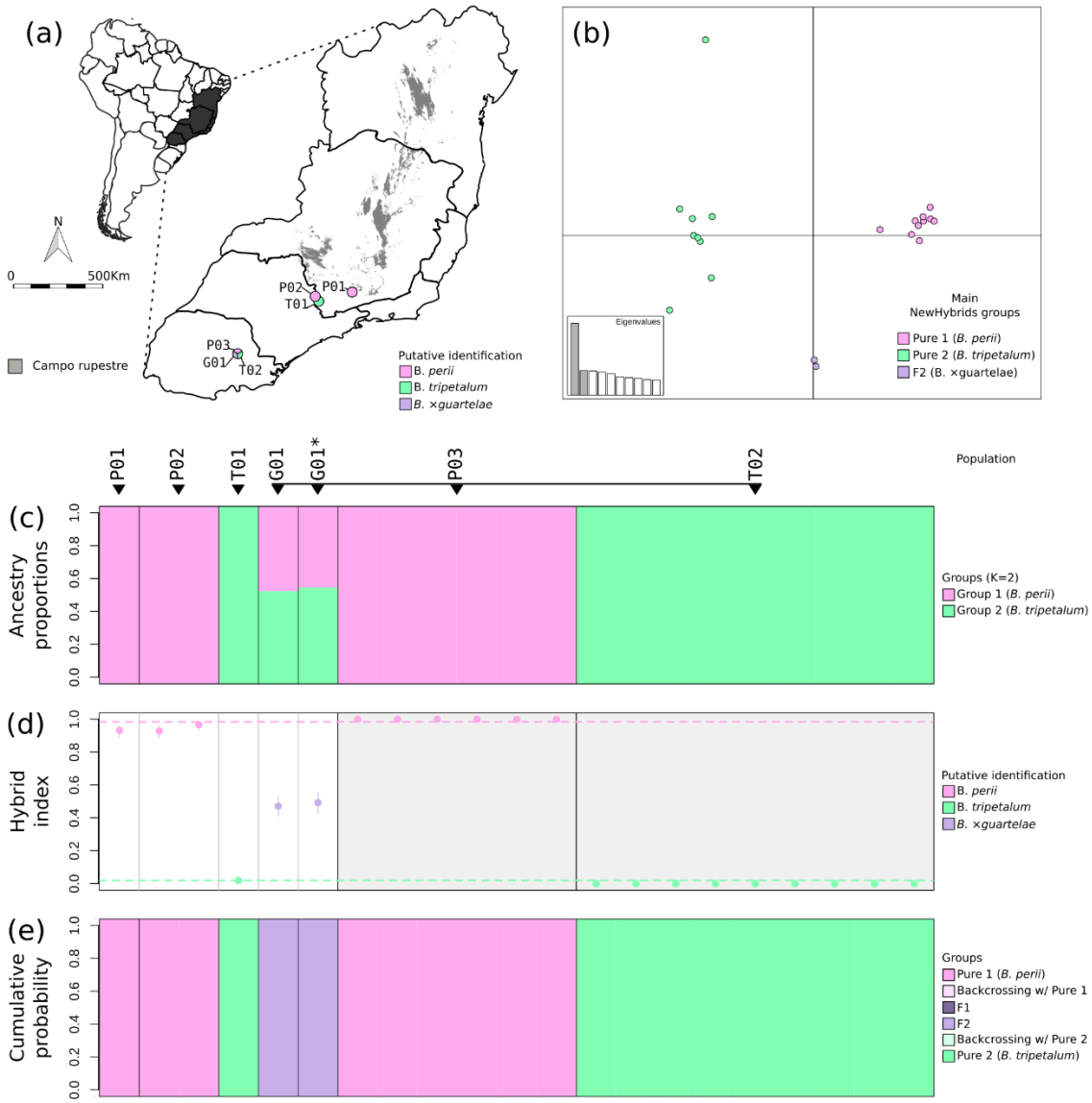


Figure 3. Hybridization in system TPG (*B. tripetalum*/*B. perii*/*B. ×guartelae*). (a) geographic distribution of populations; (b) PCA results; (c) Faststructure results for K = 2; (d) gghybrid results; (e) NewHybrids results.

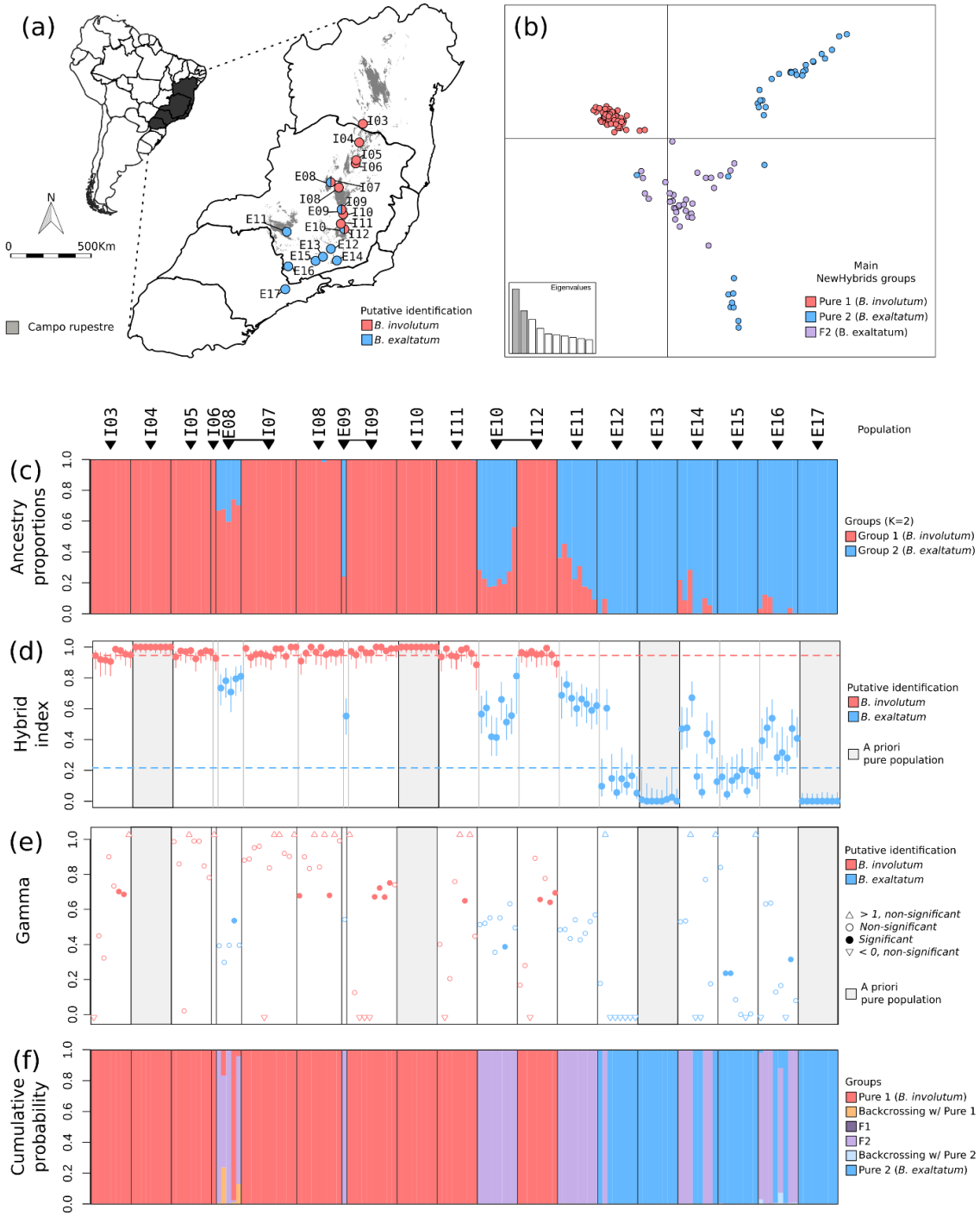


Figure 4. Hybridization in system IE (*B. involutum*/*B. exaltatum*). (a) geographic distribution of populations; (b) PCA results; (c) Faststructure results for K = 2; (d) gghybrid results; (e) HyDe results; (f) NewHybrids results.

Table 1. Information about *Bulbophyllum* sect. *Didactyle* populations analyzed in the present study. Pop: population; Lat: latitude; Lon: longitude.

System	Pop	City	State	Lat	Lon	Voucher
WIC	C01	Santana do Riacho	MG	-19.25	-43.51	UEC076050
WIC	C02	Caeté	MG	-19.82	-43.68	BHCBFiorini10
IE	E08	Joaquim Felício	MG	-17.69	-44.20	BHCB100401
IE	E09	Conceição do Mato Dentro	MG	-19.09	-43.57	HUEFS0117182
IE	E10	Catas Altas	MG	-20.08	-43.50	BHCB92776
IE	E11	São Roque de Minas	MG	-20.23	-46.45	HUFU008211
IE	E12	Tiradentes	MG	-21.11	-44.20	HUFSJ004023
IE	E13	Carrancas	MG	-21.51	-44.60	UEC064706
IE	E14	Lima Duarte	MG	-21.70	-43.89	BHCB16158
IE	E15	São Tomé das Letras	MG	-21.72	-44.98	BHCB27981
IE	E16	Santa Rita de Caldas	MG	-22.00	-46.38	BHCB014456
IE	E17	Atibaia	SP	-23.17	-46.53	UEC070741
TPG	G01	Tibagi	PR	-24.56	-50.26	UPCBMancinelli1173
WIC/IE	I03	Licínio de Almeida	BA	-14.69	-42.55	UFBA105815
WIC/IE	I04	Serra Nova	MG	-15.65	-42.74	BHCB011996
WIC/IE	I05	Grão Mogol	MG	-16.56	-42.90	IBT396396
WIC/IE	I06	Cristália	MG	-16.72	-42.92	HUEFS0076782
WIC/IE	I07	Joaquim Felício	MG	-17.69	-44.20	BHCB100399
WIC/IE	I08	Diamantina	MG	-17.96	-43.78	NY00414802
WIC/IE	I09	Conceição do Mato Dentro	MG	-19.09	-43.57	HUEFS0090623
WIC/IE	I10	Santana do Riacho	MG	-19.33	-43.56	BHCB000352
WIC/IE	I11	Caeté	MG	-19.82	-43.68	BHCB001030
WIC/IE	I12	Catas Altas	MG	-20.08	-43.50	BHCB92794
TPG	P01	São Tomé das Letras	MG	-21.72	-44.98	HUSC11371
TPG	P02	Águas da Prata	MG	-21.92	-46.68	BHCBFiorini277
TPG	P03	Tibagi	PR	-24.56	-50.26	UPCB70034
TPG	T01	Ibituruna	MG	-22.06	-46.44	BHCBFiorini280
TPG	T02	Tibagi	PR	-24.56	-50.26	UPCB70033
WIC	W03	Diamantina	MG	-17.96	-43.78	UEC064692
WIC	W04	Santana do Riacho	MG	-19.25	-43.51	HUEFS0162772
WIC	W05	Caeté	MG	-19.82	-43.68	BHCB56467
WIC	W06	Catas Altas	MG	-20.08	-43.50	BHCB92789

Table 2. Sequencing throughput for each of the *Bulbophyllum* sect. *Didactyle* systems analyzed in the present study. WIC: *B. weddellii*/*B. involutum*/*B.* ×*cipoense*; TPG: *B. tripetalum*/*B. perii*/*B.* ×*guartelae*; and IE: *B. involutum*/*B. exaltatum*.

System		WIC	TPG	IE
Individuals		112	21	149
Reads generated by sequencing		182,710,184	36,587,565	187,855,690
Average reads per individual		1,604,417 ± 443,847	1,712,451 ± 401,750	1,644,404 ± 397,696
25% missing	Variable loci	904	1,764	1,493
	SNP	1,353	2,851	2,251
	Average depth	21.3 ± 6.7 ×	24.2±6.8 ×	20.3 ± 5.2 ×
40% missing	Variable loci	2,670	3,300	4,398
	SNP	4,713	5,316	7,918
	Average depth	19.3 ± 5.6 ×	23.3±6.1 ×	18.4 ± 4.3 ×

CHAPTER 3 – DIVERSIFYING ON NEOTROPICAL SKY-ISLANDS: DISJUNCTION IS RELEVANT, BUT NOT THE ONLY FACTOR DRIVING POPULATION DIFFERENTIATION IN AN ORCHID SPECIES

Abstract

The Neotropics present one of the highest diversities on Earth and mountain ranges are believed to be important drivers of its enhanced speciation rates. On eastern South America we can find the Espinhaço/Chapada-Diamantina chain, with altitudes up to 2,072 m. This set of disjunct mountains presents a relatively old origin (Proterozoic), even though Espinhaço Range has suffered some geomorphological perturbance from the Pliocene onwards. The main vegetation on Espinhaço/Chapada-Diamantina chain are the campos rupestres vegetation. *Bulbophyllum involutum* is an orchid that is largely distributed but endemic to this environment. Although their small and light seeds are wind dispersed, very often only a small portion of adjacent outcrops with virtually identical habitats are colonized. For a more complete view of the evolutionary processes driving diversification in Orchidaceae and to better understand the diversification and connectivity patterns of geographically disjunct lineages in the campos rupestres, here we test hypotheses about genetic structure, differentiation, and demography of *B. involutum*, using ddRAD data and a model-based approach. Our work offers improved resolution to the main biogeographical breaks traditionally described to the environment, as we report gene flow estimates and differentiation measures among biogeographical units, reveling a northeastern-southwestern gradient of genetic differentiation along the Espinhaço/Chapada-Diamantina chain. The fact that Espinhaço range populations are almost panmictic, except by the slight differentiation among Northern and Southern Espinhaço Range, supports that the small and light seeds of orchids are able of long-range dispersion. However, our data also supports that environment or demography might be important factors determining population differentiation, as eastern and western populations at the Chapada Diamantina presented high differentiation. Our work does not invalidate the current main hypothesis for the origin of biodiversity on the campos rupestres (i.e., limited gene flow followed by genetic drift and local adaptation) but highlight the fact that other factors may also be import for differentiation.

Introduction

The Neotropics present one of the highest diversity on Earth, concentrating 28% of the planet's plant species (Antonelli et al., 2015). Recently the importance of mountain ranges for

the origin and maintenance of Neotropical biodiversity has being revisited using pollen records and a spatially explicit mechanistic model, showing that the Andes can act as a species pump during glacial cycles (Flantua, O'Dea, Onstein, Giraldo, & Hooghiemstra, 2019; Rangel et al., 2018). Also, mountains in general are believed to sustain high biodiversity due to enhanced speciation rates and opportunities for coexistence and persistence of lineages (Rahbek et al., 2019). In addition to the notorious Andes, we can find another mountain chain in South America: the Espinhaço/Chapada-Diamantina chain, with altitudes up to 2,072 m. In contrast to the Andean region, that presents a relatively recent origin, the Espinhaço/Chapada-Diamantina chains arose from orogenic events that occurred during the Proterozoic (Uhlein, Paim, Tassinari, & Pedreira, 2015), even though compressional tectonic event reactivated the Precambrian thrust faults and detachment planes of Espinhaço Range from the Pliocene onwards (Saadi, 1995).

Along the Espinhaço Range (mainly located in the Minas Gerais state, Brazil) and the Chapada Diamantina Plateau (Bahia state, Brazil) we can find the campos rupestres vegetation, a complex mosaic of epilithic, psammophilous and epiphytic montane flora that occurs disjunctly, mainly above 800 m (Giulietti & Pirani, 1988). Due to its high diversity and density of endemic species the campos rupestres attract attention among Brazilian vegetation types (BFG, 2015; Kew, 2017). They concentrate approximately 5,000 plant species (15% of the Brazilian plant diversity) in only 0.78% of the country's area, and about 40% of these species are endemic (BFG, 2015; Rapini, Ribeiro, Lambert, & Pirani, 2008; Silveira et al., 2016). The campos rupestres are ancient, infertile and climate-buffered landscapes (OCBILs; Silveira et al., 2016), and the buffered climate combined with the infertile soil are proposed to be the pillars for the emergence and maintenance of the high diversity and endemism observed in this environment, and idea that is reinforced by recent evidence about montane biodiversity (Rahbek et al., 2019). The high species and endemics diversities may also be related to topographic isolation among disjunct campos rupestres' areas, as interpopulation isolation may lead to local differentiation and speciation (Steinbauer et al., 2016).

Similar to what is observed for plants from other Neotropical habitats (Leal, Palma-da-Silva, & Pinheiro, 2016), phylogeographic studies of campos rupestres have shown a variety of patterns, what is coherent with the fact that different organisms can present distinct responses to montane environmental changes (Massatti & Knowles, 2014). As a potential general pattern, species presenting evidence of low vagility have shown high spatial genetic structure (Barbosa, Fiorini, Silva-Pereira, Mello-Silva, & Borba, 2012; Collevatti, de Castro, Lima, & Telles, 2012; Fiorini et al., 2019), whereas the genetic variability of vagile species are more homogeneously

distributed through space (Barres, Batalha-Filho, Schnadelbach, & Roque, 2019). However, different cacti species from the campos rupestres with similar biology present distinct levels of genetic structure (Bonatelli et al., 2014; Khan et al., 2018). In this way, it is difficult to generalize (Rull, 2013); further and higher resolution research is needed to better understand the complex modes of evolution on old Neotropical mountains.

One of the most diversified groups of plants in the Neotropics is Orchidaceae. Studies have shown that the evolutionary turnover for angiosperms is significantly higher at the American tropics and that elevational zones promoted rapid diversification in some orchids lineages (Antonelli et al., 2015; Pérez-Escobar et al., 2017). It has been proposed also that the diversity of the Orchidaceae is driven by genetic drift due to skewed mating success associated with the pollinia and the preponderance of small, disjunct populations, especially of epiphyte and lithophyte species (Phillips, Dixon, & Peakall, 2012). However, multi-loci tree-based microevolutionary studies about the diversification processes that originated the species richness of Neotropical orchids are scarce (but see Pérez-escobar et al., 2020), and the factors driving orchid speciation are still unknown.

Bulbophyllum Thouars is one of the most species rich Orchidaceae genera (~ 2,200 species; Pridgeon, Cribb, Chase, & Rasmussen, 2014). The genus presents Pantropical distribution and the ~ 60 Neotropical species form a clade (Smidt, Borba, Gravendeel, Fischer, & van den Berg, 2011). The group is pollinated by flies (myiophile) and presents a great diversity of floral forms (Smidt, 2007). The seeds are small and light (befitting anemochoric dispersion), the reptive rhizomes allow vegetative propagation, and individuals have high habitat specificity. Many *Bulbophyllum* species occur along the Espinhaço/Chapada-Diamantina chain including the *B. exaltatum* complex, a recent radiation presenting continuous morphological variation. It has been shown that disjunctions of the campos rupestres are an important factor driving the species differentiation (see Chapter 1). However, a population-based study is required for a deeper understanding of the evolutionary factors driving this process. While focusing on the history of populations of a single species, we can ask questions demanding higher resolution answers (e.g., questions about demographic parameters).

Bulbophyllum involutum stands out among species of the *B. exaltatum* complex as the only species restrict to the Espinhaço/Chapada-Diamantina chains, offering an excellent opportunity to better understand the diversification of this mountain chain. *B. involutum* is lithophyte and, although they present anemochorous dispersion, very often only a small portion of adjacent outcrops with virtually identical habitats are colonized, while on colonized outcrops individuals are abundant (on the order of one individual/100m^2, pers. obs.). These facts indicate

the existence of some subtle characteristic that determines the occurrence of the group. For a more complete view of the evolutionary processes driving diversification in Orchidaceae and to better understand the diversification and connectivity patterns of campos rupestres geographically disjunct lineages, here we test hypothesis about genetic structure, differentiation, and demography of *B. involutum*, using ddRAD data (Peterson, Weber, Kay, Fisher, & Hoekstra, 2012) and a model-based approach. Such methodologies enable genotypes of hundreds of loci per individual to be determined, allowing more sophisticated and accurate evolutionary inferences to be made (Garrick et al., 2015). Here we seek to answer the following questions: (i) Does *B. involutum* presents genetic structure? (ii) If so, is the genetic structure associated with the main geographical regions of the Espinhaço/Chapada-Diamantina chain, namely Chapada Diamantina, Northern Espinhaço, Southern Espinhaço? (iii) What are the estimates of gene flow between populations from these regions? (iv) Are there indications that the contrasting geological history of Chapada Diamantina (prolongedly stable) and Espinhaço Range (perturbed since the Pliocene) affected *B. involutum* demography?

Methodology

Sampling

We sampled 93 individuals of *Bulbophyllum involutum* from 12 populations across the campos rupestres (Table 1). We sample also three individuals of *B. tripetalum* from two populations to be used as outgroups. We collected individuals growing on different rocks and separated by a minimum of 10 m to prevent sampling vegetative clones or closely related individuals (Hedrén & Lorenz, 2019). Individuals from field collected cuttings were propagated and maintained in the living collection of Museu de História Natural e Jardim Botânico-UFMG and vouchers were deposited in the herbaria (Table 1). All samples were collected under issued permits to CFF and ELB (SISBIO 52995-1 and IEF 062/2016).

Genomic library preparation and processing

We extracted genomic DNA from fresh leaves (Doyle & Doyle, 1987) and prepared ddRAD libraries following a modified Peterson et al. (2012) protocol (Parchman et al., 2012). Fragments between 400–500 bp were size-selected using Pippin Prep (Sage Science, Beverly, MA), and PCR-amplified using high-fidelity DNA polymerase (iProof, Bio-Rad, Hercules, CA), with 8 or 12 cycles. We sequenced individuals in four lanes of an Illumina HiSeq 2500 on Rapid Run Mode (in combination with samples from other projects) at The Centre for

Applied Genomics, Hospital for Sick Children, Toronto, Canada, to generate 150 bp single end reads.

We processed genomic data using the Stacks 2.3e pipeline (Rochette & Catchen, 2017). We assembled de novo demultiplexed and filtered sequences with *ustacks*, a catalogue of consensus loci built in *cstacks*, individual genotypes identified with *sstacks*, data organized by locus with *tsv2bam*, and reads aligned and SNPs called with *gstacks*. The assembly parameters included a minimum depth of coverage, $m = 3$, mismatches allowed between two alleles of a sample, $M = 5$, and mismatches allowed between any two alleles of the catalog, $n = 6$ (i.e., the optimal parameters based on Fiorini et al., see Chapter 1), and an upper bound for $\varepsilon = 0.1$, a minimum minor allele frequency = 0.02, and a maximum observed heterozygosity = 0.5.

To maximize the number of loci, we grouped individuals by populations according to their sampling localities, and biallelic loci from a minimum of two populations were retained (Huang & Knowles, 2016). To guard against sequencing and assembly errors, we used a custom R script (Thomaz, Malabarba, & Knowles, 2017) to exclude SNPs with theta values within the upper 95% quantile of variability (see Supplementary Fig. 1).

Across the 93 sequenced individuals, 153,537,928 reads were generated (average of 1,572,332 $\pm$ 438,400 reads per individual; Supplementary table 1). We used the software plink 1.9 (Purcell et al., 2007) to identify SNPs with a maximum of 25% or 40% of missing data, because the robustness of analyses to missing data differ. After processing and filtering the genomic data, the dataset with 25% of missing data contained 1,356 variable loci (i.e., contained at least one biallelic SNP), with a total of 1,920 SNPs and a mean coverage depth per locus of 19.3 $\pm$ 5.2 ×. The dataset with 40% of missing data contained 4,009 variable loci, with a total of 6,692 SNPs and a mean coverage depth per locus of 17.5 $\pm$ 4.4 ×.

Genetic diversity and population structure

We carried out a principal component analysis (PCA) to visualize the distribution of genomic variation using adegenet 2.1.1 (Jombart & Ahmed, 2011), in R 3.5.0 (R Core Team, 2014). Due to its sensibility to missing data, we used the genomic dataset with a threshold of 25% of missing data; missing data values were replaced by the per locus mean allele frequency for a given population to explore the geography of divergence.

To estimate population structure, we used fastStructure 1.0, an accurate variational Bayesian framework compatible with large data sets (Raj, Stephens, & Pritchard, 2014). To create the bed, bim and fam files required by fastStructure, we convert ped and map files from stacks 2.43 (Rochette & Catchen, 2017), using plink 1.9 (Purcell et al., 2007). We estimate

ancestry proportions for each individual for K from 0 to 15 using the structure.py script (included within the package), using 10 replicates. To search for the Ks that better explain the data structure, we used the chooseK.py script (also included within the package) and visualized the results for the best Ks with the online application Clumpak (available at http://clumpak.tau.ac.il; Kopelman, Mayzel, Jakobsson, Rosenberg, & Mayrose, 2015).

To infer recent or current population structure based on nearest neighbor haplotype (co-ancestry) we used RADpainter combined with fineRADstructure (version 0.3.2; Malinsky, Trucchi, Lawson, & Falush, 2018). As it uses whole haplotypes, this is a powerful pipeline to infer the co-ancestry matrix from RADseq data (RADpainter) and clusterize it with a MCMC algorithm (fineRADstructure). We used the RADpainter output from stacks as input file for RADpainter and then assigned individuals to populations using finestructure. We used 100,000 burn in iterations and 100,000 sample iterations for MCMC method, sampling each 1,000 iterations. To estimate the tree, 100,000 sample iterations were used. To plot the results, we used the R script fineRADstructurePlot.R, included within the package.

To infer historical relationships between populations we used TreeMix 1.13 (Pickrell & Pritchard, 2012), using a statistical model for inferring patterns of population splits in multiple populations, using genome-wide allele frequency data and a Gaussian approximation to genetic drift. We used *B. tripetalum* individuals as the outgrop. As we used from 1 to 10 individuals per population, we turned off sample size correction, to prevent the use of too conservative correction for small sample size. To plot the population maximum likelihood tree we used the plotting_funcs.R, included within the package.

As the rate at which genetic differentiation accumulates across space can be variable, we used LocalDiff (Duforet-Frebourg & Blum, 2014) to infer nonstationary patterns of isolation-by-distance (IBD). The software infers local genetic differentiation based on Bayesian kriging, defining the local genetic differentiation for a sampled population as the average genetic differentiation between the sampled population and fictive neighboring populations. We used a matrix composed by one SNP per locus and with 25% of missing data and the geographical coordinates of the populations as input. To visualize the results, we used the Display2D.R script, included within the package.

We used the packages adegenet 2.1.1(Jombart & Ahmed, 2011), poppr 2.8.3 (Kamvar, Brooks, & GrÃ¼nwald, 2015), and ape 5.3 (Paradis & Schliep, 2019), implemented in R 3.6.2 (R Core Team, 2019), to estimate population differentiation and φ-statistics, using the AMOVA approach (Excoffier, Smouse, & Quattro, 1992). Based on PCA, fastStructure and fineRADstructure results, we split the population set into three (regions Chapada Diamantina,

Northern Espinhaço Range, and Southern Espinhaço Range) or four groups (populations I01 and I02, and regions Northern Espinhaço Range and Southern Espinhaço Range). The significance of the results was tested by 99 permutations.

Intraspecific divergence times and migration rates

We estimate demographic parameters from the folded joint site frequency spectrum (SFS), using a composite-likelihood simulation-based approach (Excoffier, Dupanloup, Huerta-Sánchez, Sousa, & Foll, 2013), implemented in fastsimcoal2 (Excoffier & Foll, 2011). Because fastsimcoal2 does not deal with missing values, to maximize the number of loci after the removal of missing data for the calculation of the joint SFS, we subsampled pairs of the four genetic groups pointed by the PCA using a custom Python script written by Qixin He (Papadopoulou & Knowles, 2015). All possible pairs where used; since there are four groups (I01, I02, Northern Espinhaço and Southern Espinhaço; Fig. 2A and B), six pairs where evaluated. To minimize errors with allele frequency estimates, we retained only loci found in at least five individuals (10 per group. As the performance of fastsimcoal2 is improved by reducing the number of parameters estimated from the data (Excoffier et al., 2013), we calculated one effective population size (*epN*) of each dataset directly from the data, using nucleotide diversity as estimated by the module *populations* from Stacks 2.3e pipeline (Rochette & Catchen, 2017) and a mutation rate of 7.31×10^{-9} per site per generation (Krasovec, Chester, Ridout, & Filatov, 2018). The estimated parameters included the effective population size of one group (*N*), the divergence time (*T)* between groups, and the symmetrical migration rates among groups (*M*). Forty independent runs per species were performed, each run with 10^6 simulations per likelihood estimation and 40 expectation-conditional maximization (ECM) cycles, based on a stopping criterion of 0.001 relative difference between iterations. We present the global maximum likelihood solution across runs.

Results

Genetic diversity and population structure

We observed values of nucleotide diversity ranging from 0.11602 (I06) to 0.15218 (I02), values of heterozygosity ranging from 0.08276 (I02) to 0.11602 (I06), and values of FIS ranging from 0.07703 (I03) to 0.15794 (I02; Table 2; Fig. 1B, C and D).

The firs axis of the PCA clearly separates populations from Chapada Diamantina from populations of Espinhaço Range, while the second axis separates population I01 from

populations I02 (Fig. 2A). The third axis separates Espinhaço Range in two geographical regions: Northern Espinhaço Range (populations I03, I04, I05, and I06) and Southern Espinhaço Range (populations I07, I08, I09, I10, I11 and I12; Fig. 2B). However, based on the three main PCA axis, Espinhaço Range populations cannot be distinguished beyond the regional level.

FastStructure analysis support the existence of three main genetic groups: Chapada Diamantina, Northern Espinhaço Range, and Southern Espinhaço Range (Fig. 2C). Yet, this analysis suggests that individuals from populations I02 (Chapada Diamantina group) and I05 (Northern Espinhaço Range group; populations) present some admixture from the Southern Espinhaço Range genetic group (Fig. 2C). FineRADstructure also suggest the existence of three genetic groups (Fig. 3), which are consistent with fastStructure analyses. Moreover, fineRADstructure results agreed with PCA, showing that populations I01 and I02 are genetically differentiated (Fig. 2A and Fig. 3). FineRADstructure shows also that there is some subtle substructure in the Northern Espinhaço Range group, but this substructure is not coherent with the populational assignment (Fig. 3). There is no sign of genetic substructure in Southern Espinhaço Range populations, except two individuals from population I11 that present higher co-ancestry (Fig. 3). FineRADstructure cladogram mirror populations' geography (Fig. 3).

The gradient of genetic differentiation measured by localDiff and the drift parameter calculated by TreeMix suggest that there is a northeastern-southwestern slope of differentiation in *B. involutum* distribution, with a stronger sign of genetic divergence between populations I01 and I02 (Chapada Diamantina) than between Northern and Southern Espinhaço Range regions (Fig. 4). The phylogenetic relationships indicated by TreeMix are compatible with fineRADstructure phylogenetic inference (Fig. 3 and Fig. 4). Considering three genetic groups (Chapada Diamantina, Northern Espinhaço Range, and Southern Espinhaço Range) the AMOVA indicates that 18.46% of the variation is among regions, 5.86% among populations within regions, 15.27% among individuals within populations and 60.41% of the variation is within individuals (Table 2). Considering four genetic groups (I01, I02, Northern Espinhaço Range, and Southern Espinhaço Range), the AMOVA indicates that 21.54% of the variation is among regions, 3.31% among populations within regions, 15.16% among individuals within populations and 69.99% of the variation is within individuals (Table 2).

Migration rates estimated by fastsimcoal2 varied from 1.57×10^{-5} (among population I01 and Northern Espinhaço Range) and 7.48×10^{-3} (among Northern Espinhaço Range and Southern Espinhaço Range). Migration rates between populations within Chapada Diamantina are of the same order magnitude than the migration rates between these populations and

Espinhaço Range populations. Estimated population sizes often agreed with empirical populations sizes (same order of magnitude), except for the Southern Espinhaço Range group, for which estimates were one other of magnitude higher than the empirical values. Times of divergence varied from 17.02 × estimated population size (among populations I01 and I02) to 0.29 × estimated population size (among Northern Espinhaço Range and Southern Espinhaço Range; Fig. 5).

Discussion

Biogeographical work dedicated to the campos rupestres traditionally report breaks on species composition between Chapada Diamantina, Northern Espinhaço Range, and Southern Espinhaço Range (e.g., Bitencourt & Rapini, 2013; Chaves, Freitas, Vasconcelos, & Santos, 2015; Colli-Silva, Vasconcelos, & Pirani, 2019), what is also true for the diversity within some genera (e.g., Ribeiro, Rapini, Damascena, & Van Den Berg, 2014) and species complex (see Chapter 1). Here we offer high resolution estimates of gene flow and differentiation among these biogeographical regions, based on a populational assessment of the campo rupestre endemic *B. involutum*, showing a northeastern-southwestern gradient of intensity of genetic differentiation and discussing the consequences of these findings to the diversification of orchid species.

Genetic diversity and structure in a disjunct environment

The observed heterozygosity and π of *B. involutum* were high, when compared to other plant species evaluated using ddRAD markers, including other orchids (Resende-Moreira et al., 2019; Roy, Moitra, & De Sarker, 2017). This corroborate the hypothesis that genetic variability of *B. involutum* populations is secured by pollination mechanisms that promote outcrossing (Borba & Semir, 1999), despite the fact that these plants are self-compatible and pollinated by small Diptera with low capability of long range flights and behavior that favors self-pollination (Borba & Semir, 1998; Borba, Shepherd, & Semir, 1999).

Strong genetic structuring among campos rupestres populations is not rare. Studies of monocots, eudicots, invertebrates and, vertebrates using a variety of markers show that populations are usually structured across this environment (e.g., Barbosa, Fiorini, Silva-Pereira, Mello-Silva, & Borba, 2012; Bonatelli et al., 2014; A. V. Chaves, Vasconcelos, Freitas, & Santos, 2019; de Magalhães et al., 2017; Lacorte, De Sena Oliveira, & Da Fonseca, 2011; Passoni, Benozzati, & Rodrigues, 2008; Ribeiro et al., 2008). *B. involutum* also presents genetic structure mirroring the geographical distribution of the populations, with main valleys and

lowlands as putative barriers, as shown for other closely related species (see Chapter 1). Indeed, *B. involutum* genetic eastern-western structure across Chapada Diamantina is close to the genetic structure observed in other herbaceous monocots (Pereira, Borba, & Giulietti, 2007), as *B. involutum* populations I01 (Mucugê) and I02 (Rio de Contas) are distinguishable in PCA and fineRADstructure analysis. This subdivision between southeastern and southwestern Chapada Diamantina is presented also in biogeographic evaluations of other organisms (e.g. Chaves et al., 2015; Ribeiro, Rapini, Damascena, & Van Den Berg, 2014), suggesting that it is an important break for multiple species.

On the other hand, the high percentage of genetic variability within individuals detected by AMOVA and the low differentiation among populations within Northern Espinhaço Range or Southern Espinhaço Range shown by localDiff and TreMix, indicate that each of these regions are panmictic, what contrasts with observations from other plant groups. For example, on Southern Espinhaço Range genetic structure is observed among *Vellozia auriculata* (also a perennial monocot) populations as close as 4 km (Fiorini et al., 2019) and high genetic structure is also observed in other *Vellozia* from the campos rupestres (Barbosa et al., 2012; Lousada, Borba, Ribeiro, Ribeiro, & Lovato, 2011; Lousada, Lovato, & Borba, 2013). A possible explanation for the distinction in geographic structure between *Vellozia* and *B. involutum* is the fact that *Vellozia* present seeds with no apparent adaptations to long range dispersion (Fiorini et al., 2019), while *B. involutum* seeds are small and light. Indeed, *B. involutum* shows a high estimated migration rate on Espinhaço Range. However, given that all the other estimated migration rate are low, other environmental and interaction filters possibly affect its effective vagility (Taylor, Weigelt, König, Zotz, & Kreft, 2019), mainly in Chapada Diamantina.

B. involutum is not the only species showing low populational structure across Espinhaço Range. The small perennial subshrub *Richterago discoidea* (Asteraceae) presented no geographical structure based on AFLP markers (Barres et al., 2019). Also, populations of the cactus *Pilosocereus aurisetus* occurring across Southern Espinhaço Range can be clustered in a single group, even though in higher scales population genetic and geographic structure is observed in this group (Bonatelli et al., 2014).

The gradient of genetic differentiation observed for *B. involutum* can be driven by climatic or by historical and demographic factors. Chapada Diamantina is part of the São Francisco Craton, an area of high geological stability since the Precambrian (Heilbron, Cordani, & Alkmim, 2017), while the Espinhaço Range (despite its equivalent old origin) is distributed across the external border of the Craton, an area subject to more geological instability. Indeed, a compressional tectonic event reactivated the Precambrian thrust faults and detached planes of

Espinhaço Range during the Pliocene, event that may had influenced environmental and climatic characteristics on which *B. involutum* populations relay on, altering its demography and/or geographical distribution, as these geographical modifications are still in course (Saadi, 1995; A. Saadi, personal communication, Sep 27, 2017). Complementarily, it has been shown that the campos rupestres environment present strong climatic and edaphic gradients (Neves et al., 2018). Soil water deficit, mean annual temperature, temperature seasonality, and mean annual precipitation were the most important variable associated to floristic northeastern-southwestern differentiation across the campos rupestres' trees (Neves et al., 2018), but these factors are likely to influence the ecology of other flora components, including *B. involutum*.

It has been proposed that under certain combinations of taxon traits, insular regions presenting higher stability are expected to show higher population differentiation, in an opposite effect to the hypothesized "species pump" effect (Papadopoulou & Knowles, 2017). It is possible that repeated and recent colonization events had taken place on the less stable Espinhaço Range sky-islands, leading to the higher observed similarity among populations. The fact that *B. involutum* presents higher genetic diversity and inbreeding in Chapada Diamantina support this hypothesis. Indeed, Chapada Diamantina has been classified as a lineage museum, while Espinhaço Range was called a lineage cradle for an eudicot subfamily (Asclepiadoideae; Bitencourt & Rapini, 2013). The use of more sophisticated demographical models of past populational history of *B. involutum* coupled with a better understanding of the geological and geographical history of the campos rupestres would help to shed light on the process that lead to the current observed pattern (Perrigo, Hoorn, & Antonelli, 2019).

Orchid population differentiation

Here we offer the first estimate of orchid interpopulation gene flow based on ddRAD data. Currently there is a lack of consensus regarding the effectiveness of orchid seed dispersal in maintaining cohesion between disjunct localities. Not only there is conflicting evidence on the limits of orchid seed dispersal (e.g., Hedrén and Lorenz, 2019; Phillips et al., 2012; Taylor et al., 2019 versus Helsen et al., 2016; Pinheiro et al., 2014; Trapnell et al., 2013), but also recent evidence suggests the spreading capability of wind dispersed seeds may be inferior to other dispersion syndromes (Arjona, Nogales, Heleno, & Vargas, 2018; Fajardo et al., 2019). Besides, other factors unrelated to the seed's inherent potential for movement may also impact plants dispersal (Taylor et al., 2019).

Our data support that an orchid species widely distributed on the campos rupestres exhibit geographical structure of genetic markers, but that this structure is not limited

exclusively by geographical breaks, as Northern Espinhaço Range and Southern Espinhaço Range are panmictic despite the habitat discontinuity of these subregions. This indicates that the small and light seeds of orchids are able of long-range dispersion, but that other factor are also important for population differentiation (Taylor et al., 2019), as we observed a gradient of population differentiation with no evidence of a disjunction gradient (i.e. Chapada Diamantina is as apparently as fragmented as Espinhaço Range). Thus, we suggest that both habit discontinuity and environmental conditions (past and current) are responsible for differentiation of orchid populations. Future studies shall focus on the role of each of these aspects on generating orchid biodiversity.

Conclusion

Since *B. involutum* is endemic to the campos rupestres, but also present an ample geographic distribution on Espinhaço and Chapada Diamantina, it offers an unpaired opportunity to better understand the diversification of its environment. Our work offers improved resolution to the main biogeographical breaks traditionally described to the environment, as Chapada Diamantina, Northern Espinhaço Range, and Southern Espinhaço Range regions were recovered as distinct regions. We report gene flow estimates and differentiation measures among these biogeographical units, reveling a northeastern-southwestern gradient of genetic differentiation along these montane areas. It is possible that this gradient was driven by the more stable geological history of Chapada Diamantina compared to the Espinhaço range. However, other factors, such as the environmental gradient observed through the campos rupestres or climate driven past demographical events, can also play and important role into the observed pattern. The fact that Espinhaço range populations are almost panmictic, except by the slight differentiation among Northern and Southern Espinhaço Range, supports that the small and light seeds of orchids are able of long-range dispersion. However, our data also supports that environment or demography might be the most important factors determining population differentiation, as eastern and western populations in the Chapada Diamantina, presented high differentiation, despite separated by smaller distances than Espinhaço Range populations. Since many of campo rupestres species present limited vagility, it is expected that gene flow between their populations are even more limited than observed for *B. involutum*. In this way, our work does not invalidate the current main hypothesis for the origin of biodiversity on the campos rupestres (i.e., limited gene flow followed by genetic drift

and local adaptation) but highlight the fact that other factors may also be import for differentiation.

References

Antonelli, A., Zizka, A., Silvestro, D., Scharn, R., Cascales-Miñana, B., & Bacon, C. D. (2015). An engine for global plant diversity: highest evolutionary turnover and emigration in the American tropics. *Frontiers in Genetics*, *6*, 1–14. https://doi.org/10.3389/fgene.2015.00130

Arjona, Y., Nogales, M., Heleno, R., & Vargas, P. (2018). Long-distance dispersal syndromes matter: diaspore–trait effect on shaping plant distribution across the Canary Islands. *Ecography*, *41*(5), 805–814. https://doi.org/10.1111/ecog.02624

Barbosa, A. R., Fiorini, C. F., Silva-Pereira, V., Mello-Silva, R., & Borba, E. L. (2012). Geographical genetic structuring and phenotypic variation in the *Vellozia hirsuta* (Velloziaceae) ochlospecies complex. *American Journal of Botany*, *99*(9), 1477–1488. https://doi.org/10.3732/ajb.1200070

Barres, L., Batalha-Filho, H., Schnadelbach, A. S., & Roque, N. (2019). Pleistocene climatic changes drove dispersal and isolation of *Richterago discoidea* (Asteraceae), an endemic plant of campos rupestres in the central and eastern Brazilian sky-islands. *Botanical Journal of the Linnean Society*, *189*(2), 132–152. https://doi.org/10.1093/botlinnean/boy080

BFG. (2015). Growing knowledge: An overview of Seed Plant diversity in Brazil. *Rodriguesia*, *66*(4), 1085–1113. https://doi.org/10.1590/2175-7860201566411

Bitencourt, C., & Rapini, A. (2013). Centres of endemism in the Espinhaço Range: identifying cradles and museums of Asclepiadoideae (apocynaceae). *Systematics and Biodiversity*, *11*(4), 525–536. https://doi.org/10.1080/14772000.2013.865681

Bonatelli, I. A. S., Perez, M. F., Peterson, A. T., Taylor, N. P., Zappi, D. C., Machado, M. C., … Moraes, E. M. (2014). Interglacial microrefugia and diversification of a cactus species complex: Phylogeography and palaeodistributional reconstructions for *Pilosocereus aurisetus* and allies. *Molecular Ecology*, *23*(12), 3044–3063. https://doi.org/10.1111/mec.12780

Borba, E. L., & Semir, J. (1998). Wind-assisted fly pollination in three *Bulbophyllum* (Orchidaceae) species occurring in the Brazilian campos rupestres. *Lindleyana*, *13*(3), 203–218.

Borba, E. L., & Semir, J. (1999). Temporal variation in pollinarium size after its removal in species of *Bulbophyllum*: A different mechanism preventing self-pollination in Orchidaceae. *Plant Systematics and Evolution*, *217*(3–4), 197–204. https://doi.org/10.1007/BF00984365

Borba, E. L., Shepherd, G. J., & Semir, J. (1999). Reproductive systems and crossing potential in three species of *Bulbophyllum* (Orchidaceae) occurring in Brazilian "campo rupestre" vegetation. *Plant Systematics and Evolution*, *217*(3–4), 205–214. https://doi.org/10.1007/BF00984366

Chaves, A. V., Freitas, G. H. S., Vasconcelos, M. F., & Santos, F. R. dos. (2015). Biogeographic patterns, origin and speciation of the endemic birds from eastern Brazilian mountaintops: a review. *Systematics and Biodiversity*, *13*(1), 1–16. https://doi.org/10.1080/14772000.2014.972477

Collevatti, R. G., de Castro, T. G., Lima, J. de S., & Telles, M. P. de C. (2012). Phylogeography of *Tibouchina papyrus* (Pohl) Toledo (Melastomataceae), an endangered tree species from rocky savannas, suggests bidirectional expansion due to climate cooling in the Pleistocene. *Ecology and Evolution*, *2*(5), 1024–1035. https://doi.org/10.1002/ece3.236

Colli-Silva, M., Vasconcelos, T. N. C., & Pirani, J. R. (2019). Outstanding plant endemism levels strongly support the recognition of campo rupestre provinces in mountaintops of eastern South America. *Journal of Biogeography*, *46*(8), 1723–1733. https://doi.org/10.1111/jbi.13585

Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, *19*(1), 11–15.

Duforet-Frebourg, N., & Blum, M. G. B. (2014). Nonstationary patterns of isolation-by-distance: Inferring measures of local genetic differentiation with bayesian kriging. *Evolution*, *68*(4), 1110–1123. https://doi.org/10.1111/evo.12342

Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C., & Foll, M. (2013). Robust demographic inference from genomic and SNP data. *PLoS Genetics*, *9*(10), e1003905. https://doi.org/10.1371/journal.pgen.1003905

Excoffier, L., & Foll, M. (2011). fastsimcoal: a continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics*, *27*(9), 1332–1334. https://doi.org/10.1093/bioinformatics/btr124

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial

DNA restriction data. *Genetics*, *131*(2), 479–491. Retrieved from https://www.genetics.org/content/131/2/479

Fajardo, J., Mateo, R. G., Vargas, P., Fernández-Alonso, J. L., Gómez-Rubio, V., Felicísimo, Á. M., & Muñoz, J. (2019). The role of abiotic mechanisms of long-distance dispersal in the American origin of the Galápagos flora. *Global Ecology and Biogeography*, *28*(11), 1610–1620. https://doi.org/10.1111/geb.12977

Fiorini, C. F., Miranda, M. D., Silva-Pereira, V., Barbosa, A. R., Oliveira, U. De, Kamino, L. H. Y., … Borba, E. L. (2019). The phylogeography of *Vellozia auriculata* (Velloziaceae) supports low zygotic gene flow and local population persistence in the campo rupestre, a Neotropical OCBIL. *Botanical Journal of the Linnean Society*, *191*(3), 381–398. https://doi.org/10.1093/botlinnean/boz051

Flantua, S. G. A., O'Dea, A., Onstein, R. E., Giraldo, C., & Hooghiemstra, H. (2019). The flickering connectivity system of the north Andean páramos. *Journal of Biogeography*, *46*(8), 1808–1825. https://doi.org/10.1111/jbi.13607

Garrick, R. C., Bonatelli, I. a. S., Hyseni, C., Morales, A., Pelletier, T. a., Perez, M. F., … Carstens, B. C. (2015). The evolution of phylogeographic data sets. *Molecular Ecology*, *24*(6), 1164–1171. https://doi.org/10.1111/mec.13108

Giulietti, A. M., & Pirani, J. R. (1988). Patterns of geographic distribution of some plant species from the Espinhaco Range, Minas Gerais and Bahia, Brazil. *Proceedings of a Workshop on Neotropical Distribution Patterns*, (JANUARY), 39–69.

Hedrén, M., & Lorenz, R. (2019). Seed dispersal and fine-scale genetic structuring in the asexual *Nigritella miniata* (Orchidaceae) in the Alps. *Botanical Journal of the Linnean Society*, *190*(1), 83–100. https://doi.org/10.1093/botlinnean/boz005

Heilbron, M., Cordani, U. G., & Alkmim, F. F. (2017). São Francisco Craton, Eastern Brazil. In M. Heilbron, U. G. Cordani, & F. F. Alkmim (Eds.), *Regional Geology Reviews* (1st ed.). https://doi.org/10.1007/978-3-319-01715-0

Helsen, K., Meekers, T., Vranckx, G., Roldán-Ruiz, I., Vandepitte, K., & Honnay, O. (2016). A direct assessment of realized seed and pollen flow within and between two isolated populations of the food-deceptive orchid *Orchis mascula*. *Plant Biology*, *18*(1), 139–146. https://doi.org/10.1111/plb.12342

Huang, H., & Knowles, L. L. (2016). Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. *Systematic Biology*, *65*(3), 357–365. https://doi.org/10.1093/sysbio/syu046

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*(21), 3070–3071. https://doi.org/10.1093/bioinformatics/btr521

Kamvar, Z. N., Brooks, J. C., & GrÃ¼nwald, N. J. (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics*, *6*(JUN), 1–10. https://doi.org/10.3389/fgene.2015.00208

Kew, R. (2017). The state of the world's plants report (R. B. Gardens, Ed.).

Khan, G., Godoy, M. O., Franco, F. F., Perez, M. F., Taylor, N. P., Zappi, D. C., … Moraes, E. M. (2018). Extreme population subdivision or cryptic speciation in the cactus *Pilosocereus jauruensis* ? A taxonomic challenge posed by a naturally fragmented system. *Systematics and Biodiversity*, *16*(2), 188–199. https://doi.org/10.1080/14772000.2017.1359215

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: a program for identifying clustering modes and packaging population structure inferences across *K*. *Molecular Ecology Resources*, *15*(5), 1179–1191. https://doi.org/10.1111/1755-0998.12387

Krasovec, M., Chester, M., Ridout, K., & Filatov, D. A. (2018). The mutation rate and the age of the sex chromosomes in *Silene latifolia*. *Current Biology*, *28*(11), 1832-1838.e4. https://doi.org/10.1016/j.cub.2018.04.069

Leal, B. S. S., Palma-da-Silva, C., & Pinheiro, F. (2016). Phylogeographic studies depict the role of space and time scales of plant speciation in a highly diverse Neotropical region. *Critical Reviews in Plant Sciences*, *35*(4), 215–230. https://doi.org/10.1080/07352689.2016.1254494

Lousada, J. M., Borba, E. L., Ribeiro, K. T., Ribeiro, L. C., & Lovato, M. B. (2011). Genetic structure and variability of the endemic and vulnerable *Vellozia gigantea* (Velloziaceae) associated with the landscape in the Espinhaço Range, in southeastern Brazil: Implications for conservation. *Genetica*, *139*(4), 431–440. https://doi.org/10.1007/s10709-011-9561-5

Lousada, J. M., Lovato, M. B., & Borba, E. L. (2013). High genetic divergence and low genetic variability in disjunct populations of the endemic *Vellozia compacta* (Velloziaceae) occurring in two edaphic environments of Brazilian campos rupestres. *Brazilian Journal of Botany*, *36*(1), 45–53. https://doi.org/10.1007/s40415-013-0001-x

Malinsky, M., Trucchi, E., Lawson, D. J., & Falush, D. (2018). RADpainter and fineRADstructure: population inference from RADseq Data. *Molecular Biology and Evolution*, *35*(5), 1284–1290. https://doi.org/10.1093/molbev/msy023

Massatti, R., & Knowles, L. L. (2014). Microhabitat differences impact phylogeographic concordance of codistributed species: genomic evidence in montane sedges (*Carex* L.) from the Rocky Mountains. *Evolution*, *68*(10), 2833–2846. https://doi.org/10.1111/evo.12491

Neves, D. M., Dexter, K. G., Pennington, R. T., Bueno, M. L., de Miranda, P. L. S., & Oliveira-Filho, A. T. (2018). Lack of floristic identity in campos rupestres—A hyperdiverse mosaic of rocky montane savannas in South America. *Flora*, *238*, 24–31. https://doi.org/10.1016/j.flora.2017.03.011

Papadopoulou, A., & Knowles, L. L. (2015). Genomic tests of the species-pump hypothesis: Recent island connectivity cycles drive population divergence but not speciation in Caribbean crickets across the Virgin Islands. *Evolution*, *69*(6), 1501–1517. https://doi.org/10.1111/evo.12667

Papadopoulou, A., & Knowles, L. L. (2017). Linking micro- and macroevolutionary perspectives to evaluate the role of Quaternary sea-level oscillations in island diversification. *Evolution*, *71*(12), 2901–2917. https://doi.org/10.1111/evo.13384

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

Parchman, T. L., Gompert, Z., Mudge, J., Schilkey, F. D., Benkman, C. W., & Buerkle, C. A. (2012). Genome-wide association genetics of an adaptive trait in lodgepole pine. *Molecular Ecology*, *21*(12), 2991–3005. https://doi.org/10.1111/j.1365-294X.2012.05513.x

Pereira, A. C. S., Borba, E. L., & Giulietti, A. M. (2007). Genetic and morphological variability of the endangered *Syngonanthus mucugensis* Giul. (Eriocaulaceae) from the Chapada Diamantina, Brazil: implications for conservation and taxonomy. *Botanical Journal of the Linnean Society*, *153*(4), 401–416. https://doi.org/10.1111/j.1095-8339.2007.00624.x

Pérez-escobar, O. A., Bogarín, D., Schley, R., Bateman, R. M., Gerlach, G., Harpke, D., … Blattner, F. R. (2020). Molecular phylogenetics and evolution resolving relationships in an exceedingly young Neotropical orchid lineage using genotyping-by-sequencing data.

Molecular Phylogenetics and Evolution, *144*(November 2019), 106672. https://doi.org/10.1016/j.ympev.2019.106672

Pérez-Escobar, O. A., Chomicki, G., Condamine, F. L., Karremans, A. P., Bogarín, D., Matzke, N. J., … Antonelli, A. (2017). Recent origin and rapid speciation of Neotropical orchids in the world's richest plant biodiversity hotspot. *New Phytologist*, *215*(2), 891–905. https://doi.org/10.1111/nph.14629

Perrigo, A., Hoorn, C., & Antonelli, A. (2019). Why mountains matter for biodiversity. *Journal of Biogeography*, (September), 1–11. https://doi.org/10.1111/jbi.13731

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS One*, *7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Phillips, R. D., Dixon, K. W., & Peakall, R. (2012). Low population genetic differentiation in the Orchidaceae: implications for the diversification of the family. *Molecular Ecology*, *21*(21), 5208–5220. https://doi.org/10.1111/mec.12036

Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*, *8*(11), e1002967. https://doi.org/10.1371/journal.pgen.1002967

Pinheiro, F., Cozzolino, S., Draper, D., De Barros, F., Félix, L. P., Fay, M. F., & Palma-Silva, C. (2014). Rock outcrop orchids reveal the genetic connectivity and diversity of inselbergs of northeastern Brazil. *BMC Evolutionary Biology*, *14*(1), 49. https://doi.org/10.1186/1471-2148-14-49

Pridgeon, A. M., Cribb, P. J., Chase, M. W., & Rasmussen, F. N. (2014). *Genera Orchidacearum, vol. 6: Epidendroideae, part 3* (1st ed.). Oxford: Oxford University Press.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., … Sham, P. C. (2007). *PLINK: a tool set for whole-genome association and population-based linkage analyses*. *81*(3), 559–575.

R Core Team. (2014). *R: A Language and Environment for Statistical Computing*. Retrieved from http://www.r-project.org/

R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. Retrieved from https://www.r-project.org/

Rahbek, C., Borregaard, M. K., Antonelli, A., Colwell, R. K., Holt, B. G., Nogues-bravo, D., … Fjeldså, J. (2019). Building mountain biodiversity: geological and evolutionary processes. *Science*, *365*(6458), 1114–1119.

Raj, A., Stephens, M., & Pritchard, J. K. (2014). FastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*, *197*(2), 573–589. https://doi.org/10.1534/genetics.114.164350

Rangel, T. F., Edwards, N. R., Holden, P. B., Diniz-Filho, J. A. F., Gosling, W. D., Coelho, M. T. P., … Colwell, R. K. (2018). Modeling the ecology and evolution of biodiversity: biogeographical cradles, museums, and graves. *Science*, *361*(6399), eaar5452. https://doi.org/10.1126/science.aar5452

Rapini, A., Ribeiro, P. L., Lambert, S., & Pirani, J. R. (2008). A flora dos campos rupestres da Cadeia do Espinhaço. *Megadiversidade*, *4*(1–2), 16–24. Retrieved from http://www.inot.org.br/artigo/Rapini_et_al_2008_A_flora_dos_campos_rupestres_da_Cadeia_do_Espinhaco.pdf

Resende-Moreira, L. C., Knowles, L. L., Thomaz, A. T., Prado, J. R., Souto, A. P., Lemos-Filho, J. P., & Lovato, M. B. (2019). Evolving in isolation: Genetic tests reject recent connections of Amazonian savannas with the central Cerrado. *Journal of Biogeography*, *46*(1), 196–211. https://doi.org/10.1111/jbi.13468

Ribeiro, P. L., Rapini, A., Damascena, L. S., & Van Den Berg, C. (2014). Plant diversification in the Espinhaço Range: Insights from the biogeography of *Minaria* (Apocynaceae). *Taxon*, *63*(6), 1253–1264. https://doi.org/10.12705/636.16

Rochette, N. C., & Catchen, J. M. (2017). Deriving genotypes from RAD-seq short-read data using Stacks. *Nature Publishing Group*, *12*(12), 2640–2659. https://doi.org/10.1038/nprot.2017.123

Roy, S. C., Moitra, K., & De Sarker, D. (2017). Assessment of genetic diversity among four orchids based on ddRAD sequencing data for conservation purposes. *Physiology and Molecular Biology of Plants*, *23*(1), 169–183. https://doi.org/10.1007/s12298-016-0401-z

Rull, V. (2013). Some problems in the study of the origin of Neotropical biodiversity using palaeoecological and molecular phylogenetic evidence. *Systematics and Biodiversity*, *11*(4), 415–423. https://doi.org/10.1080/14772000.2013.865682

Saadi, A. (1995). A geomorfologia da Serra do Espinhaço em Minas Gerais e de suas margens. *Geonomos*, *3*(1), 41–63.

Silveira, F. A. O., Negreiros, D., Barbosa, N. P. U., Buisson, E., Carmo, F. F., Carstensen, D. W., … Lambers, H. (2016). Ecology and evolution of plant diversity in the endangered campo rupestre: a neglected conservation priority. *Plant and Soil*, *403*(1–2), 129–152. https://doi.org/10.1007/s11104-015-2637-8

Smidt, E. de C. (2007). Filogenia e revisão taxonômica de *Bulbophyllum* Thouars (Orchidaceae) ocorrentes no Neotrópico. Universidade Estadual de Feira de Santana.

Smidt, E. de C., Borba, E. L., Gravendeel, B., Fischer, G. A., & van den Berg, C. (2011). Molecular phylogeny of the Neotropical sections of *Bulbophyllum* (Orchidaceae) using nuclear and plastid spacers. *Taxon*, *60*(4), 1050–1064.

Steinbauer, M. J., Field, R., Grytnes, J. A., Trigas, P., Ah-Peng, C., Attorre, F., … Beierkuhnlein, C. (2016). Topography-driven isolation, speciation and a global increase of endemism with elevation. *Global Ecology and Biogeography*, *25*(9), 1097–1107. https://doi.org/10.1111/geb.12469

Taylor, A., Weigelt, P., König, C., Zotz, G., & Kreft, H. (2019). Island disharmony revisited using orchids as a model group. *New Phytologist*, *223*(2), 597–606. https://doi.org/10.1111/nph.15776

Thomaz, A. T., Malabarba, L. R., & Knowles, L. L. (2017). Genomic signatures of paleodrainages in a freshwater fish along the southeastern coast of Brazil: genetic structure reflects past riverine properties. *Heredity*, *119*(4), 287–294. https://doi.org/10.1038/hdy.2017.46

Trapnell, D. W., Hamrick, J. L., Ishibashi, C. D., & Kartzinel, T. R. (2013). Genetic inference of epiphytic orchid colonization; it may only take one. *Molecular Ecology*, *22*(14), 3680–3692. https://doi.org/10.1111/mec.12338

Uhlein, A., Paim, P. S. G., Tassinari, C. C. G., & Pedreira, A. J. (2015). Análise estratigráfica de bacias rifte Paleo-Mesoproterozoicas dos crátons Amazônico e São Francisco, Brasil. *Geonomos*, *23*(2), 1–13. https://doi.org/10.18285/geonomos.v23i2.707

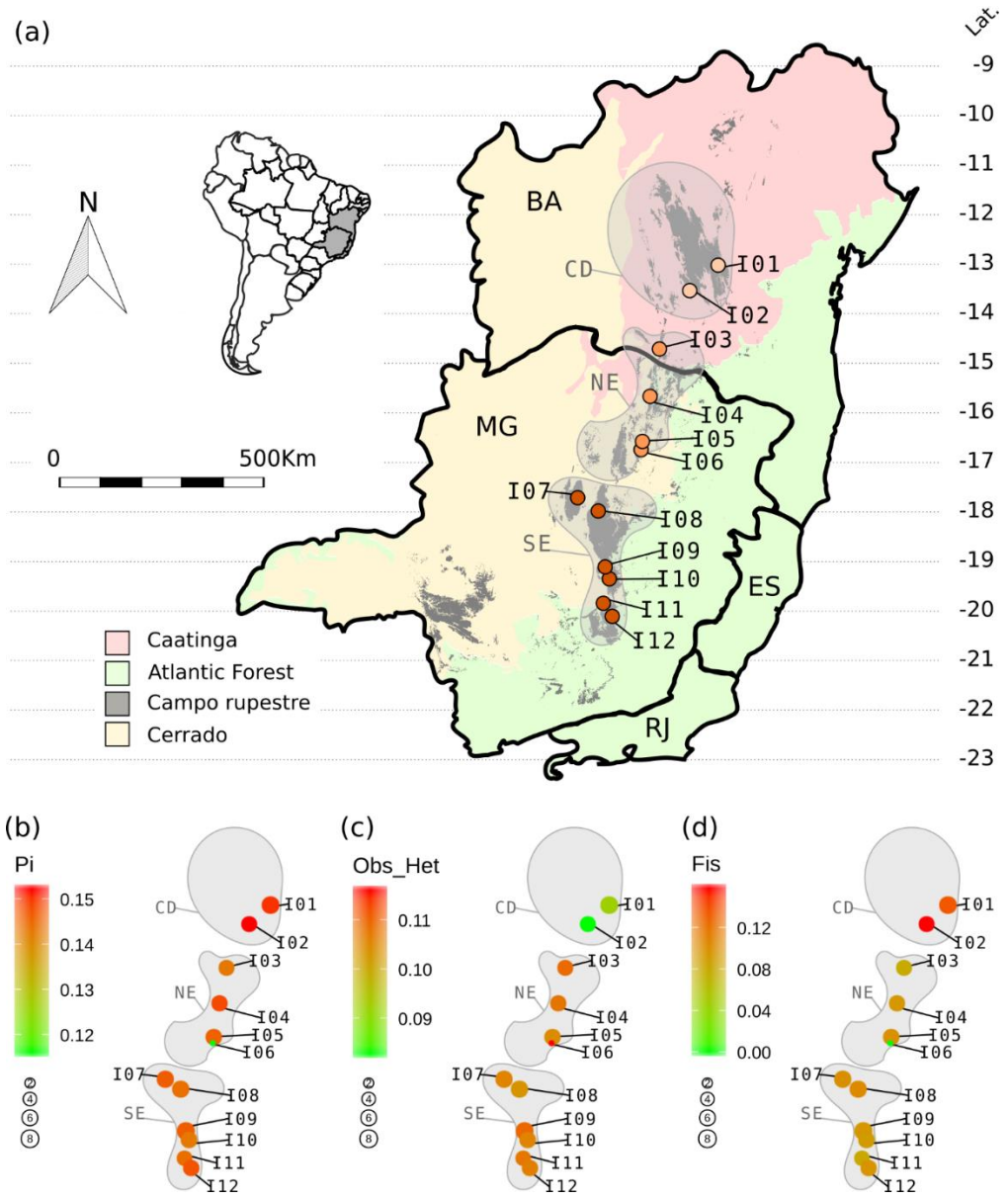


Figure 1. *Bulbophyllum involutum* sampled populations and genetic diversity and inbreeding indexes. (a) Sampled localities marked by circles (I01-I12). Groups colors are consistently used on Fig. 2, Fig. 3, and Fig. 4B. Brazilian states are shown (BA: Bahia; ES: Espírito Santo; MG: Minas Gerais; RJ: Rio de Janeiro). (b) Nucleotide diversity. (c) Observed heterozygosity. (d) F_{is}, the inbreeding coefficient. For all subfigures, campos rupestres' regions are shown in light grey (CD: Chapada Diamantina; NE: Northern Espinhaço; SE: Southern Espinhaço). For subfigures b, c, and d, the size of the circles is proportional to the number of sampled individuals and the colors of the circles indicate values on the respective scales.

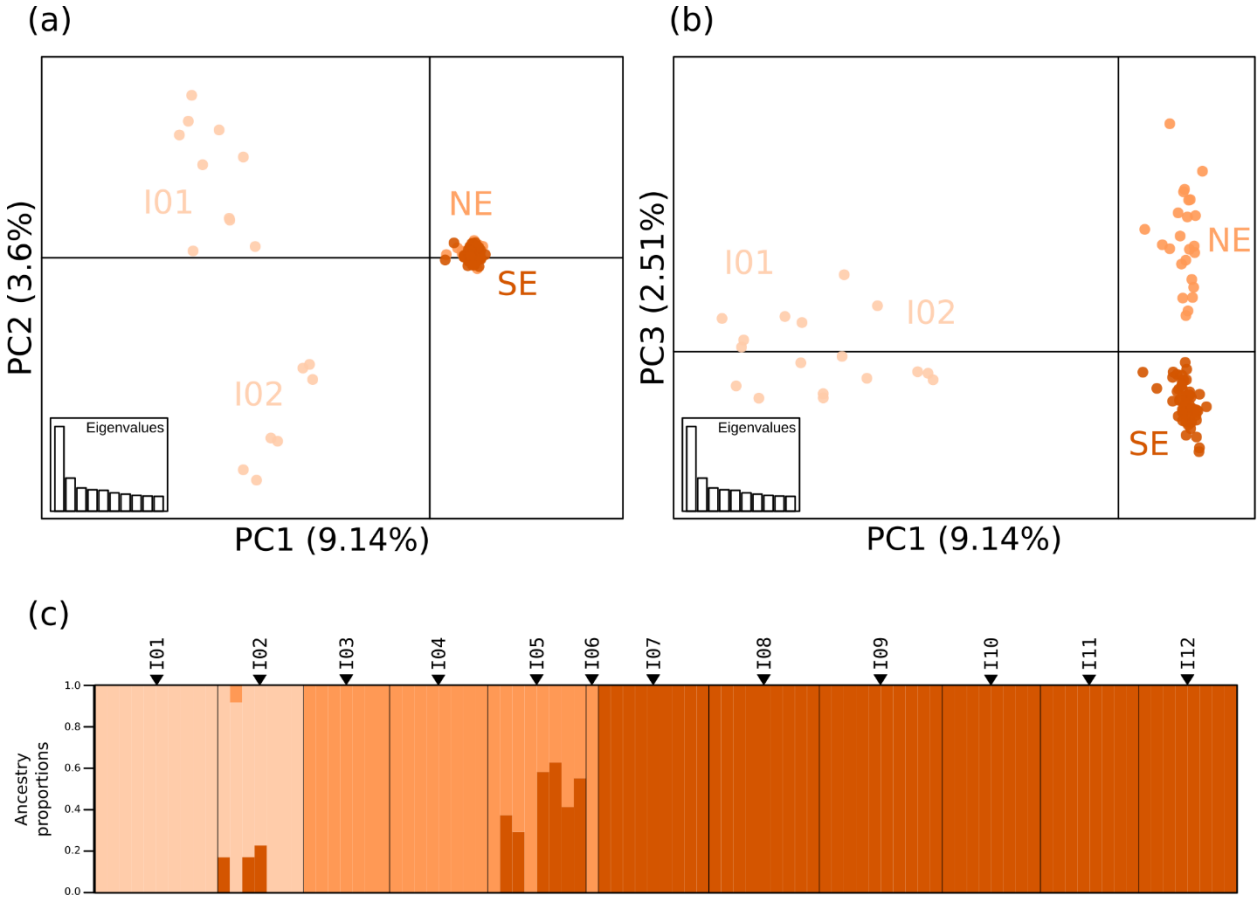


Figure 2. PCA and fastStructure analysis for *Bulbophyllum involutum*. PCA results showing axis one and two (a) and one and three (b). Axis eigenvalues are shown. Populations I01, I02 and regions NE (Northern Espinhaço) and SE (Southern Espinhaço) are indicated. (c) fastStructure results. Population names are indicated.

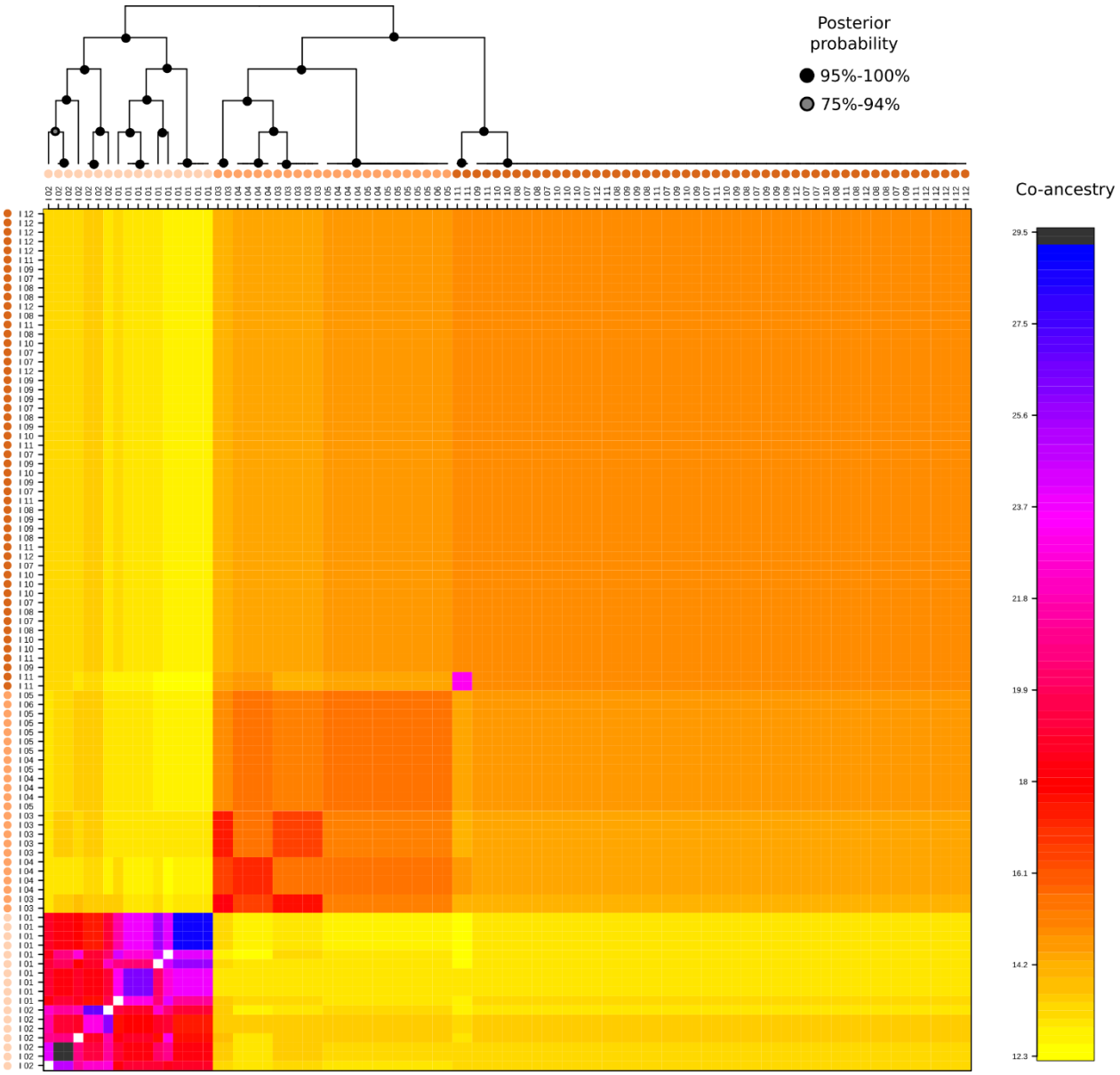


Figure 3. FineRADstructure results, including the clustered coancestry matrix and the phylogenetic inference for *Bulbophyllum involutum*. Population names for each individual are indicated.

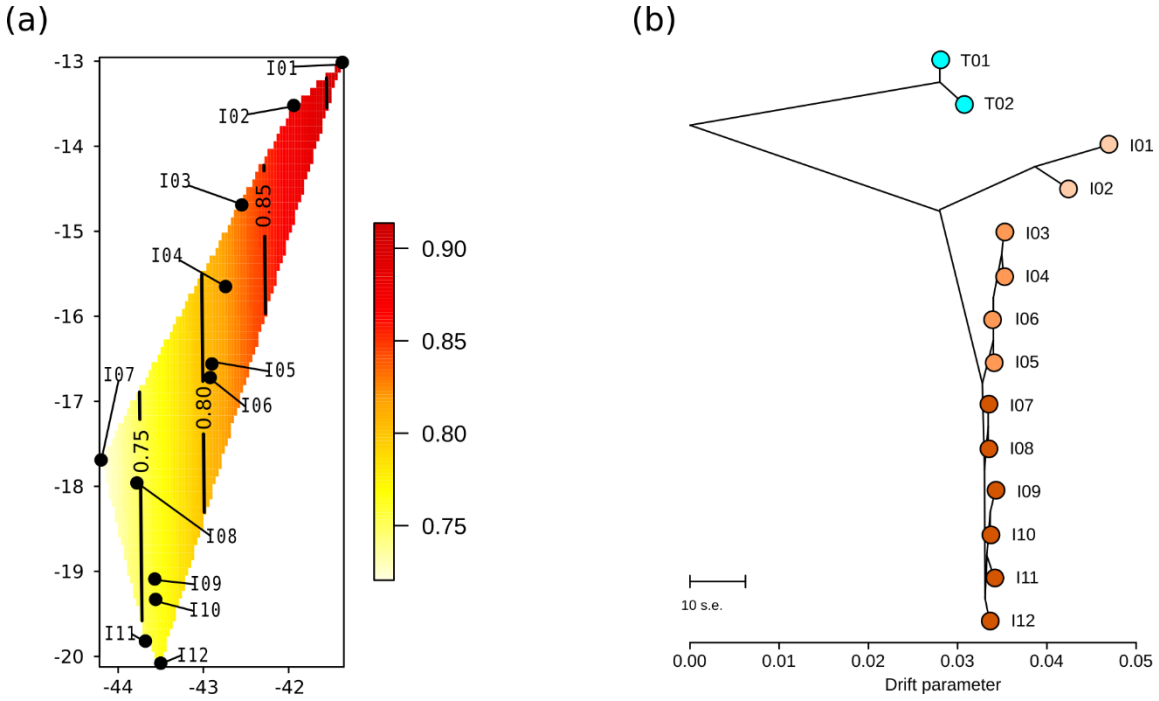


Figure 4. Populational differentiation. (a) LocalDiff heatmap showing a northeastern-southwestern slope of differentiation in *Bulbophyllum involutum* distribution. Local differentiation corresponds to one minus the correlation between sampled populations and fictive neighboring populations. Populations are indicated. (b) Treemix results. Horizontal branch lengths are proportional to the amount of genetic drift on the branch. The scale bar shows ten times the average standard error of the entries in the sample covariance matrix.

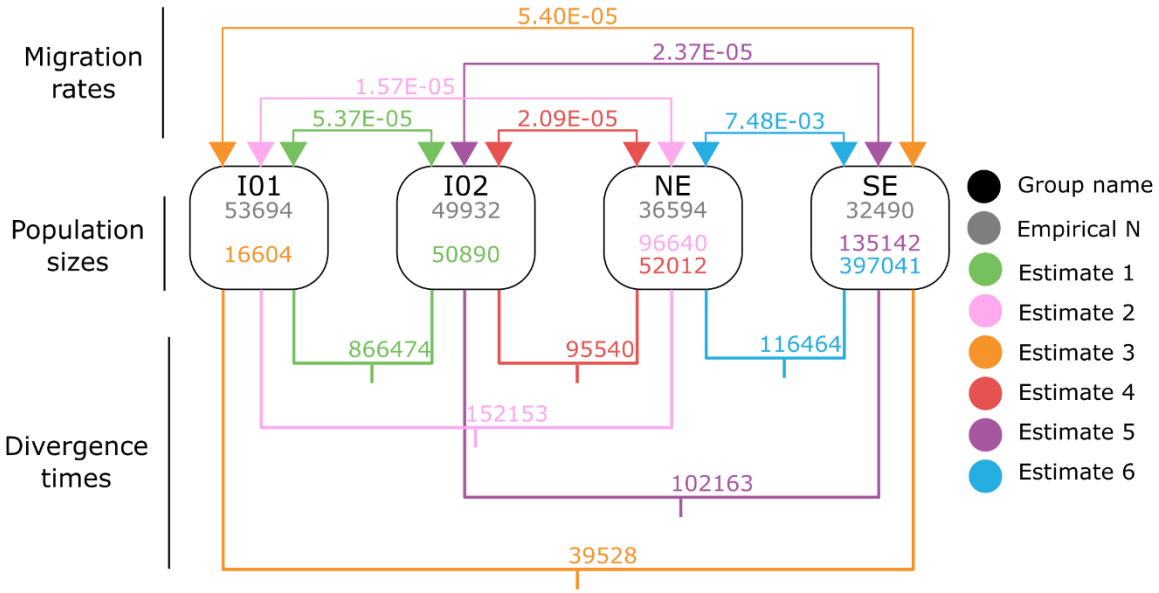


Figure 5. Fastsimcoal2 estimates among *Bulbophyllum involutum* genetic groups. For each of the six combinations of pair of genetic-geographic groups (I01, I02, Northern Espinhaço Range, and Southern Espinhaço Range) three estimates are shown: effective population size of one group (*N*), the divergence time between groups (T, in generations), and the symmetrical migration rates among groups (*M*). The empirically calculated effective population sizes are also shown.

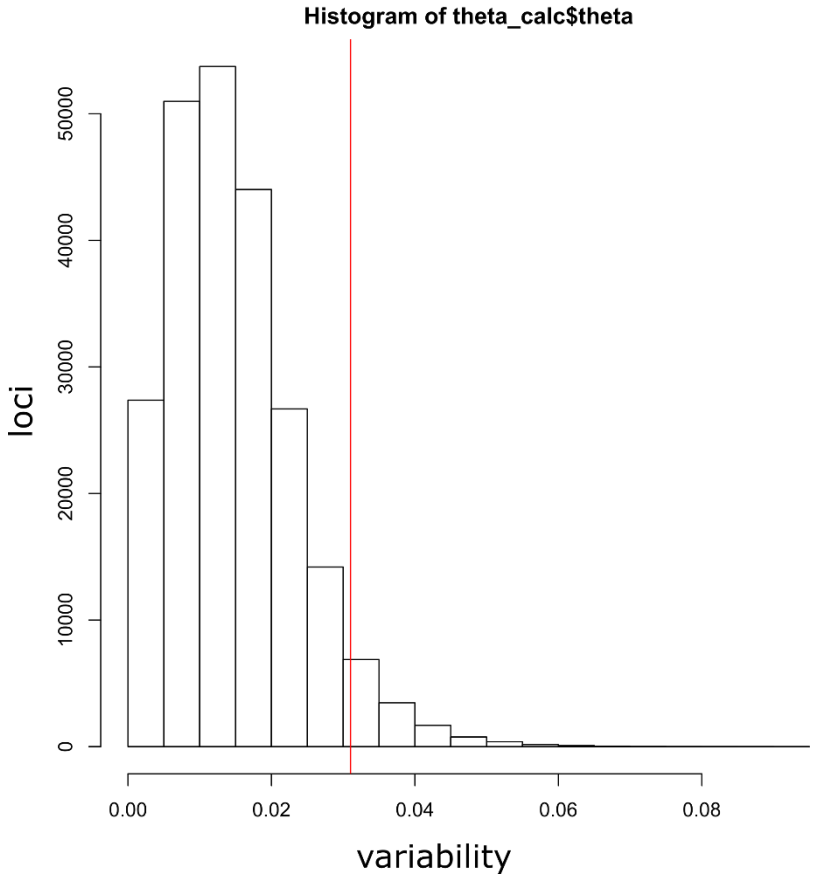
Table 1. *Bulbophyllum involutum* population information. Pop: population; N: number of samples; Lat: latitude; Lon: longitude; CD: Chapada Diamantina; NE: Northern Espinhaço; SE Southern Espinhaço.

Pop	Region	City	State	N	Lat	Lon	Voucher
I01	CD	Mucugê	BA	10	-13.00	-41.37	HUEFS0070811
I02	CD	Rio de Contas	BA	7	-13.52	-41.94	HUEFS0206037
I03	NE	Licínio de Almeida	BA	7	-14.69	-42.55	UFBA105815
I04	NE	Serra Nova	MG	8	-15.65	-42.74	BHCB011996
I05	NE	Grão Mogol	MG	8	-16.56	-42.90	IBT396396
I06	NE	Cristália	MG	1	-16.72	-42.92	HUEFS0076782
I07	SE	Joaquim Felício	MG	9	-17.69	-44.20	BHCB100399
I08	SE	Diamantina	MG	9	-17.96	-43.78	NY00414802
I09	SE	Conceição do Mato Dentro	MG	10	-19.09	-43.57	HUEFS0090623
I10	SE	Santana do Riacho	MG	8	-19.33	-43.56	BHCB000352
I11	SE	Caeté	MG	8	-19.82	-43.68	BHCB001030
I12	SE	Catas Altas	MG	8	-20.08	-43.50	BHCB92794

Table 3-2. Analysis of molecular variance (AMOVA) of *Bulbophyllum involutum* using three and four genetic groups (see text for groups explanation). DF: degrees of freedom; SUM SQ: sums of squares; Mean Sq: mean sums of squares; Sigma: variance (σ) for each hierarchical level; %: percent of the total variance; Std.Obs: variance in the randomized data; Pvalue: significance of population differentiation; ϕ: population differentiation statistics.

		DF	SUM SQ	MEAN SQ	σ	%	STD.OBS	PVALUE	φ
3 groups	Between region	2	4307.89	2153.95	32.68	18.46	4.80	0.01	$\varphi_{RT} = 0.18$
	Between pop within region	9	2874.97	319.44	10.38	5.86	19.07	0.01	$\varphi_{PR} = 0.07$
	Between individuals within pop	81	13196.54	160.93	27.02	15.27	11.49	0.01	$\varphi_{IP} = 0.20$
	Within individuals	93	10048.62	106.90	106.90	60.41	-23.90	0.01	$\varphi_{IT} = 0.40$
	Total	187	30428.02	162.72	176.97	100.00			
4 groups	Between region	3	5189.68	1729.89	38.39	21.54	4.24	0.01	$\varphi_{RT} = 0.22$
	Between pop within region	8	1993.18	249.15	5.90	3.31	13.04	0.01	$\varphi_{PR} = 0.04$
	Between individuals within pop	81	13196.54	160.93	27.02	15.16	12.49	0.01	$\varphi_{IP} = 0.20$
	Within individuals	93	10048.62	106.90	106.90	59.99	-29.31	0.01	$\varphi_{IT} = 0.40$
	Total	187	30428.02	162.72	178.20	100.00			

Supplementary material



Supplementary Fig 1. Histogram of genetic variability of loci. The red line indicates the upper 95% quantile of variability.

Supplementary Table 1. *Bulbophyllum involutum* population information. Pop: population; ID: sample identification; Sequenced: total number of sequenced reads; Retained: number reads after quality filtering; Loci_25: number of loci under a 25% of missing data filter (see text); Depth_25: average loci depth under a 25% of missing data filter (see text); Loci_40: number of loci under a 40% of missing data filter (see text); Depth_40: average loci depth under a 40% of missing data filter (see text); Live collection: live collection reference code.

Pop	ID	Sequenced	Retained	Loc_25	Dep_25	Loc_40	Dep_40	Live collection
I01	b1_10854	2131312	2116586	1615	25.97	4847	22.99	MHNJBFiorini0854
I01	b1_10869	1485466	1473633	1512	17.96	4241	17.00	MHNJBFiorini0869
I01	b1_10870	1421511	1412148	1443	16.20	3995	15.62	MHNJBFiorini0870
I01	b1_10874	1680212	1669068	1510	19.30	4353	17.58	MHNJBFiorini0874
I01	b1_10875	1347202	1338240	1440	16.04	4039	15.73	MHNJBFiorini0875
I01	b1_10876	2045596	2030161	1584	23.54	4727	20.86	MHNJBFiorini0876
I01	b2_10867	2141404	2088814	1686	28.09	5303	26.58	MHNJBFiorini0867
I01	b2_10871	2046920	1995620	1642	26.40	4996	25.12	MHNJBFiorini0871
I01	b3_10872	1817358	1757142	1582	19.34	4869	18.69	MHNJBFiorini0872
I01	b3_10877	2094313	2040760	1545	25.20	4772	24.81	MHNJBFiorini0877
I02	b2_10843	2315001	1211274	1572	16.46	4671	15.77	MHNJBFiorini0843
I02	b2_10845	2381210	2322257	1679	32.22	5346	29.61	MHNJBFiorini0845
I02	b3_10832	1391655	1355920	1459	19.09	4470	18.18	MHNJBFiorini0832
I02	b3_10839	2053947	1062281	1535	14.64	4568	14.22	MHNJBFiorini0839
I02	b3_10844	1571736	1533440	1640	20.51	5030	20.07	MHNJBFiorini0844
I02	b4_10842	1696060	1670634	1538	28.02	4591	24.42	MHNJBFiorini0842
I02	b4_10846	1272294	1253487	1566	20.33	4669	18.13	MHNJBFiorini0846
I03	b1_10820	1639162	1625625	1544	17.79	4427	15.68	MHNJBFiorini0820
I03	b1_10826	615508	609773	807	12.20	2271	11.64	MHNJBFiorini0826
I03	b1_10828	1918644	1070734	1249	12.99	3581	12.09	MHNJBFiorini0828
I03	b2_10819	1269595	1239695	1612	14.99	4917	13.75	MHNJBFiorini0819
I03	b3_10822	2007022	1947407	1595	23.43	4976	21.23	MHNJBFiorini0822
I03	b3_10824	2183688	2126654	1694	24.32	5366	21.99	MHNJBFiorini0824
I03	b3_10827	2254813	2202965	1684	24.69	5447	22.54	MHNJBFiorini0827
I04	b1_10471	1655614	1641154	1520	17.64	4413	16.15	MHNJBFiorini0471
I04	b1_10476	1940201	1923808	1584	19.88	4730	17.31	MHNJBFiorini0476
I04	b1_10488	1317076	1305893	1449	14.55	4098	13.02	MHNJBFiorini0488
I04	b2_10482	1849936	1809847	1664	20.08	5230	17.83	MHNJBFiorini0482
I04	b3_10484	1202087	1169029	1527	13.35	4703	12.65	MHNJBFiorini0484
I04	b3_10485	1329559	1298777	1577	15.82	4901	14.74	MHNJBFiorini0485
I04	b3_10489	1214017	1167178	1463	13.42	4485	12.32	MHNJBFiorini0489
I04	b4_10472	1836360	1808609	1606	26.42	5030	22.27	MHNJBFiorini0472
I05	b1_10428	2321160	2302641	1652	22.14	4924	19.61	MHNJBFiorini0428
I05	b1_10436	2190478	2174041	1632	22.09	4920	19.32	MHNJBFiorini0436
I05	b1_10438	1494492	1483950	1507	15.46	4225	14.20	MHNJBFiorini0438

Pop	ID	Sequenced	Retained	Loc_25	Dep_25	Loc_40	Dep_40	Live collection
I05	b2_10419	1031833	1006751	1567	12.25	4810	11.39	MHNJBFiorini0419
I05	b3_10422	1313286	1271068	1568	15.70	4874	14.40	MHNJBFiorini0422
I05	b3_10424	1127086	1098252	1586	14.15	4892	12.94	MHNJBFiorini0424
I05	b3_10425	1370963	1330695	1633	15.84	5036	14.34	MHNJBFiorini0425
I05	b4_10431	2136922	2111255	1616	29.75	5055	25.12	MHNJBFiorini0431
I06	b3_10418	1025853	996041	1554	12.08	4770	11.50	MHNJBFiorini0418
I07	b1_10391	2050410	2036233	1622	23.52	4917	20.07	MHNJBFiorini0391
I07	b1_10393	2387684	2367958	1676	26.41	5067	22.61	MHNJBFiorini0393
I07	b1_10401	1935005	1913801	1522	19.12	4494	16.93	MHNJBFiorini0401
I07	b2_10389	1797862	1760032	1624	21.70	5069	19.25	MHNJBFiorini0389
I07	b3_10392	958210	933662	1477	11.96	4493	11.40	MHNJBFiorini0392
I07	b4_10390	1816403	1794910	1589	24.77	4879	21.20	MHNJBFiorini0390
I07	b4_10400	1670270	1649934	1594	21.65	4949	18.38	MHNJBFiorini0400
I07	b4_10402	1753188	1729416	1602	24.71	4994	20.77	MHNJBFiorini0402
I07	b4_10405	1565007	1542744	1600	22.74	4932	19.38	MHNJBFiorini0405
I08	b2_10328	2236245	2178488	1690	27.13	5351	24.07	MHNJBFiorini0328
I08	b2_10332	1193614	1163727	1565	13.72	4768	12.61	MHNJBFiorini0332
I08	b2_10336	1428270	1392498	1564	15.72	4822	14.36	MHNJBFiorini0336
I08	b2_10337	936198	914010	1452	11.62	4401	10.84	MHNJBFiorini0337
I08	b2_10344	1185331	1153736	1544	14.75	4817	13.32	MHNJBFiorini0344
I08	b3_10338	1059266	1036205	1529	13.37	4693	12.65	MHNJBFiorini0338
I08	b3_10339	1770476	1732511	1679	21.35	5332	19.28	MHNJBFiorini0339
I08	b3_10340	1091874	1061721	1465	14.22	4572	13.25	MHNJBFiorini0340
I08	b4_10327	1743717	1722322	1641	24.24	4995	20.94	MHNJBFiorini0327
I09	b1_10786	1649323	1637723	1541	18.51	4550	16.55	MHNJBFiorini0786
I09	b1_10808	1236262	1227021	1446	14.50	4103	13.49	MHNJBFiorini0808
I09	b1_10811	1458297	1444977	1541	17.09	4444	15.05	MHNJBFiorini0811
I09	b1_10816	1845454	1831523	1593	19.13	4712	16.97	MHNJBFiorini0816
I09	b1_10818	1840808	1826378	1559	19.30	4579	17.22	MHNJBFiorini0818
I09	b2_10806	2335452	2278243	1720	29.21	5503	25.62	MHNJBFiorini0806
I09	b2_10812	1950549	1905273	1624	22.82	5201	20.20	MHNJBFiorini0812
I09	b3_10785	2322943	2257778	1713	27.01	5466	23.90	MHNJBFiorini0785
I09	b3_10809	2143208	2084792	1614	24.68	5181	22.07	MHNJBFiorini0809
I09	b3_10810	1869777	1819180	1618	21.11	5106	18.82	MHNJBFiorini0810
I10	b1_10294	1896591	1877836	1615	20.56	4834	18.02	MHNJBFiorini0294
I10	b2_10292	1860148	1807316	1711	21.05	5350	18.96	MHNJBFiorini0292
I10	b2_10296	1736655	1682633	1619	17.89	5042	16.44	MHNJBFiorini0296
I10	b2_10299	1714807	1663721	1596	18.67	5024	17.10	MHNJBFiorini0299
I10	b3_10291	1130466	1090717	1531	12.94	4641	12.01	MHNJBFiorini0291
I10	b3_10301	1714073	1660757	1627	17.88	4977	16.90	MHNJBFiorini0301
I10	b3_10302	1028471	1000556	1480	12.79	4599	12.01	MHNJBFiorini0302
I10	b4_10293	1763819	1741681	1641	24.93	4980	20.95	MHNJBFiorini0293

Pop	ID	Sequenced	Retained	Loc_25	Dep_25	Loc_40	Dep_40	Live collection
I11	b1_10006	643411	638205	1102	9.28	2900	8.85	MHNJBFiorini0006
I11	b1_10009	1477514	1465312	1524	17.60	4410	15.58	MHNJBFiorini0009
I11	b1_10016	1136891	1125323	1445	13.23	4108	12.29	MHNJBFiorini0016
I11	b2_10002	1543562	1503879	1631	19.32	5071	17.72	MHNJBFiorini0002
I11	b3_10008	1105546	1077284	1517	13.03	4604	12.28	MHNJBFiorini0008
I11	b3_10011	2074110	2022164	1709	24.56	5413	22.20	MHNJBFiorini0011
I11	b3_10014	1448664	1399420	1557	16.55	4771	15.59	MHNJBFiorini0014
I11	b3_10784	508890	489102	839	9.27	2366	9.02	MHNJBFiorini0784
I12	b1_10727	1152243	1144015	1368	14.02	3957	12.82	MHNJBFiorini0727
I12	b1_10736	1454926	1443704	1370	18.34	3935	16.37	MHNJBFiorini0736
I12	b1_10737	1545922	1532909	1490	18.13	4394	15.88	MHNJBFiorini0737
I12	b2_10733	976205	954154	1523	12.89	4653	11.69	MHNJBFiorini0733
I12	b3_10738	2069036	2018499	1693	24.45	5483	22.42	MHNJBFiorini0738
I12	b3_10740	1798907	1754941	1645	20.84	5152	19.13	MHNJBFiorini0740
I12	b3_10746	2368365	2305819	1698	28.70	5451	25.46	MHNJBFiorini0746
I12	b3_10753	2065758	2010784	1623	24.38	5149	21.62	MHNJBFiorini0753
Total		**153537928**						
Average			**1572332**		**19.3**		**17.5**	
Standard deviation			**438400**		**5.2**		**4.4**	

CHAPTER 4 – THE PHYLOGEOGRAPHY OF *VELLOZIA AURICULATA* (VELLOZIACEAE) SUPPORTS LOW ZYGOTIC GENE FLOW AND LOCAL POPULATION PERSISTENCE IN THE CAMPO RUPESTRE, A NEOTROPICAL OCBIL

Authors: Fiorini, CF; Miranda, MD; Silva-Pereira, V; Barbosa, AR; Oliveira, U; Kamino, LHY; Mota, NFO; Viana, PL; Borba, EL

Journal: Botanical Journal of the Linnean Society

Year: 2019

Volume: 191

Number: 3

Pages: 381-398

Url: https://academic.oup.com/botlinnean/article/191/3/381/5579407

DOI: 10.1093/botlinnean/boz051

ISSN: 0024-4074

Botanical Journal of the Linnean Society, 2019, **191**, 381–398. With 7 figures.

The phylogeography of *Vellozia auriculata* (Velloziaceae) supports low zygotic gene flow and local population persistence in the campo rupestre, a Neotropical OCBIL

CECILIA F. FIORINI[1,*], MARINA DUTRA MIRANDA[1], VIVIANE SILVA-PEREIRA[2], ARIANE RAQUEL BARBOSA[1], UBIRAJARA DE OLIVEIRA[3], LUCIANA HIROMI YOSHINO KAMINO[4], NARA FURTADO DE OLIVEIRA MOTA[5], PEDRO LAGE VIANA[6] and EDUARDO LEITE BORBA[1]

[1]*Departamento de Botânica, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Av. Presidente Antônio Carlos, 6627, 31270–901, Belo Horizonte, MG, Brazil*
[2]*Universidade Federal do Paraná, Setor de Ciências Biológicas, Departamento de Botânica, Av. Coronel Francisco H. Santos, 210, 81531–990, Curitiba, PR, Brazil*
[3]*Centro de Sensoriamento Remoto, Instituto de Geociências, Universidade Federal de Minas Gerais, Av. Presidente Antônio Carlos, 6627, 31270–901, Belo Horizonte, MG, Brazil*
[4]*Instituto Prístino, R. Santa Maria Goretti, 86, 30642-020, Belo Horizonte, MG, Brazil*
[5]*Instituto Tecnológico Vale de Desenvolvimento Sustentável, R. Boaventura da Silva, 955, 66055–090, Belém, PA, Brazil*
[6]*Museu Paraense Emílio Goeldi, Av. Governador Magalhães Barata, 376, 66040–170, Belém, PA, Brazil*

Received 11 December 2018; revised 14 May 2019; accepted for publication 22 July 2019

The campo rupestre is a Neotropical azonal vegetation. Its disjoint distribution and the fact that it is an old climatic buffered infertile landscape (OCBIL) have been associated with the high diversity and endemism observed in this environment. Here, we tested whether a micro-endemic species from campo rupestre shows: (1) limited zygotic gene flow; (2) lower gametic than zygotic gene flow structure; (3) substrate-driven genetic structure and (4) no evidence of Pleistocene local extinction or recolonization. By sequencing intergenic plastid regions, phenotyping inter simple sequence repeats (ISSR) and modelling present and past species suitability distributions for *Vellozia auriculata* we conclude that (1) zygotic gene flow is limited; (2) gametic gene flow is recurrent, but limited by elevation and distance; (3) there is no support for genetic structure driven by substrate and (4) Pleistocene climatic changes did not restrict the species to refugia, with local persistence. As long-term gene flow restrictions may lead to differentiation and speciation, our data helps to corroborate that the campo rupestre is both a cradle (due to low zygotic gene flow, prolonged isolation and consequent differentiation) and a lineage museum (due to local survival during climate oscillations). We highlight two distinct evolutionarily significant units (ESU), providing information for better conservation practice.

KEYWORDS: endemism – gene flow – Pleistocene – population genetics – SDM.

INTRODUCTION

The Neotropics occupy 14% of the land surface of the earth, but are home to 28% of plant diversity of the planet (Antonelli *et al.*, 2015). Localized in the Neotropical region, Brazil is the country with the highest number of plant species and the campo rupestre stands out among its several phytophysiognomies, presenting the highest density of endemic species (BFG, 2015). The Brazilian campo rupestre is a mosaic of azonal vegetations, located in the Cerrado, Caatinga and Atlantic Rain Forest provinces, characterized by disjunct areas associated with quartzite outcrops, generally above 900 m a.s.l. (Giulietti & Pirani, 1988). This phytophysiognomy occurs mainly on the Espinhaço Range, a rocky chain of > 1000 km of latitudinal extension, home to *c.* 15% of Brazilian plant species, 40% of which are endemics,

*Corresponding author. E-mail: ceciliafiorini@gmail.com

© 2019 The Linnean Society of London,

Downloaded from https://academic.oup.com/botlinnean/article-abstract/191/3/381/5579407 by guest on 28 October 2019

in only 0.78% of the area of the country (Rapini *et al.*, 2008; BFG, 2015; Silveira *et al.*, 2016).

The campo rupestre is a naturally fragmented environment and has been identified as an ancient, infertile, climatically buffered landscape (OCBIL; Silveira *et al.*, 2016). OCBIL theory proposes that buffered climate and soil infertility are important factors for maintenance of the observed diversity and endemism in those environments, and that multifactorial selective regimes would lead to specific patterns, including reduced dispersibility of seeds (Hopper, 2009). Indeed, campo rupestre species generally present seed dispersal syndromes that favour short-distance dispersion (Conceição *et al.*, 2016), and population genetic studies of campo rupestre species point to high differentiation between demes (e.g. Borba *et al.*, 2001, 2007; Jesus *et al.*, 2001; Ribeiro *et al.*, 2008; Barbosa *et al.*, 2012; Lousada, Lovato & Borba, 2013).

OCBIL theory also proposes that the combination of high specialization, limited seed dispersal and buffered climate would lead to population systems in which isolation for long periods of time and consequent independent evolution, by drift or selection, is a common process (Hopper, 2009). Due to their physical characteristics, outcrops are able to maintain stable microhabitats and would have had *in situ* microrefugia during climatic changes, as supported by theoretical studies (Main, 1997; Schut *et al.*, 2014). These factors suggest that OCBILs, as the campo rupestre, could be both lineage cradles, where new lineages are constantly generated, and lineage museums, where older lineages persist through evolutionary time (Bitencourt & Rapini, 2013).

Some groups show particularly high levels of richness, endemism and micro-endemism in the campo rupestre. Velloziaceae, which comprises *c.* 240 species and is predominantly Neotropical, exhibit its highest diversity and endemism on mountains in eastern Brazil, especially in outcrop areas of Espinhaço Range, in the Brazilian states of Minas Gerais and Bahia (Menezes, Mello-Silva & Mayo, 1993; Alcantara, Ree & Mello-Silva, 2018). Population genetic studies of species of *Vellozia* Vand. from campo rupestre support the occurrence of differentiation between demes localized in distinct rock islands, but the occurrence of cytoplasmic gene flow across short scales (distances of 10 km or less) is still an open question (Franceschinelli *et al.*, 2006; Barbosa, 2011; Lousada *et al.*, 2011, 2013; Barbosa *et al.*, 2012).

Vellozia auriculata Mello-Silva & N.L.Menezes (Velloziaceae) is one of the micro-endemic *Vellozia* spp. occurring disjunctly in outcrop landscapes along the Planalto Diamantina, between latitudes −17° and −18° of the Espinhaço Range. The distance between *V. auriculata* localities range from 2.3 km to 70.0 km (27.0 km, on average), in a total area of 1000 km^2 (Mello-Silva & Menezes, 1999). It is one of the tallest *Vellozia* spp., reaching *c.* 4 m in height, and occurs epilithically or as a psammophile, which has been shown to correlate with the pattern of genetic structure in other *Vellozia* spp. (Lousada *et al.*, 2013). There are no floral biology studies for this species, but the large bright purplish and yellow flowers of *V. auriculata* are visited, and probably pollinated, by large bees (pers. obs.), despite some species of the genus with distinct floral morphology being also pollinated by birds (Franceschinelli *et al.*, 2006). Morphological characteristics suggest that *V. auriculata* seed dispersal is barochoric, but eventually the small seeds could be secondarily dispersed by rainwater drainage or gales (pers. obs.). The individuals are capable of vegetative propagation by rooting of fallen branches.

Vellozia auriculata is classified as 'Endangered' on the Red Lists of Threatened Species of fauna and flora of the Minas Gerais state (Drummond *et al.*, 2008), and all populations are immediately threatened by recurrent artificial fires. In addition, one of its populations occurs in a sand mining area (population b1, Fig. 1).

Due to its characteristics, *V. auriculata* is a good model for studying phylogeographic patterns of campo rupestre micro-endemic plant species. In this work, using plastid DNA regions and inter simple sequence repeat (ISSR) phenotypes, we test whether there is genetic structure across the distribution of *V. auriculata*, a pattern that would be expected under low gene flow between populations, or differential substrate-driven selection. As cytoplasmic DNA is generally maternally transmitted in monocots (Greiner, Sobanski & Bock, 2015), we expected that patterns of plastid genetic diversity will reflect the effective seed flow (seed flow followed by establishment, growth and reproduction, on a scale of many generations). Complementarily, the ISSR markers capture variation from nuclear and organelle genomes (but due to genome size disparity, it is more prone to show information about the nuclear genome), also providing information about the effective pollen flow (pollen flow followed by fertilization, seed formation, establishment, growth and reproduction, on a scale of many generations). Here, we also test whether there is evidence that during Pleistocene glacial cycles the populations survived *in situ*, showing no signs of centralized refugia or contraction followed by expressive demographic or geographical expansions. The occurrence of Pleistocene refugia has being observed in plants from other Neotropical phytophysiognomies rich in endemic species (e.g. Ramos, Lemos-Filho & Lovato, 2009; Pinheiro *et al.*, 2013; Buzatti *et al.*, 2017; Camps *et al.*, 2018; Melo *et al.*, 2018). However, the proposed climate buffering

Downloaded from https://academic.oup.com/botlinnean/article-abstract/191/3/381/5579407 by guest on 28 October 2019

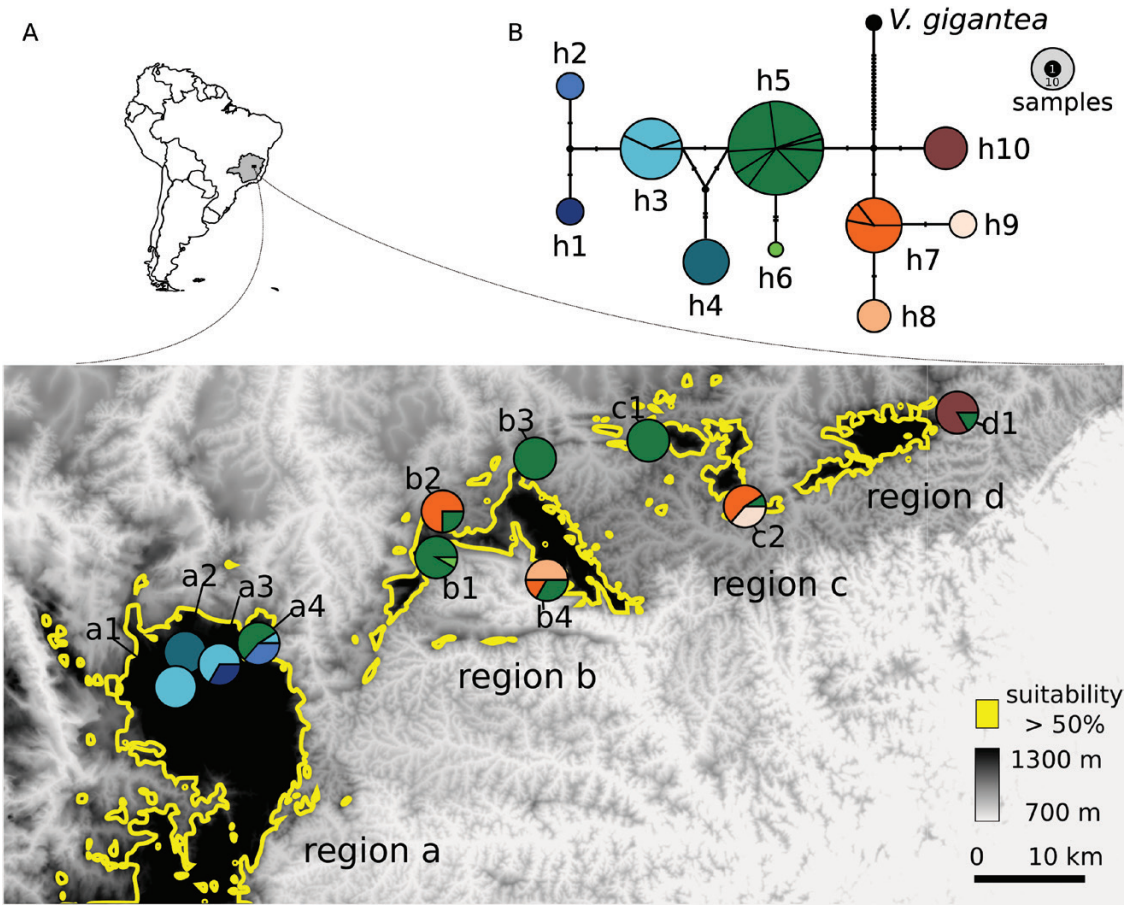


Figure 1. A, Map of the geographical distribution and B, haplotype network of the 11 populations of *Vellozia auriculata* sampled in this work. In the map, the colours indicate the occurrence of the plastid DNA haplotypes h1–h10, according to the colours presented in the network. Elevational variation is represented in greyscale. Limits of occurrence suitability > 50% are presented in yellow lines. For population names, see Table 1. In the network, the diameter of the circles is proportional to the sampling. When present in more than one population, the haplotypes circles on the network are subdivided and the angle of each slice is proportional to the frequency of the haplotypes on populations.

of the campo rupestre and the biological characteristics of its plant species (Silveira *et al.*, 2016) challenge the applicability of the refugium hypothesis to this environment.

Here, we seek to answer the following questions. (1) Are zygotic and gametic gene flow limited in a micro-endemic species from campo rupestre, despite the short distance between populations? (2) Are plastid makers more structured through space than nuclear markers, suggesting that in campo rupestre gametic gene flow may be more frequent than zygotic gene flow? (3) Is there substrate-driven genetic structure in a campo rupestre micro-endemic species? (4) Are there signs of Pleistocene refugia for a micro-endemic species from campo rupestre? We evaluated the magnitude and geographical distribution of genetic diversity, estimated the dates of divergence among lineages, investigated the patterns of cytoplasmic and nuclear genetic diversity and structure between groups of populations, and explored the occurrence of changes in demography and geographical distribution in *V. auriculata* with molecular markers and species distribution models (SDM).

MATERIAL AND METHODS

Genetic data collection

We collected and dried young leaves of individuals from 11 populations, covering the entire geographical range and edaphic occurrence of *V. auriculata* (Table 1). These localities occur in four main 'islands' of high suitability, which are isolated by elevations below 800 m (Miranda, 2012). In this work, we referred to such islands as geographical regions 'a', 'b', 'c' and 'd' (Fig. 1). We collected samples at a minimum distance

Downloaded from https://academic.oup.com/botlinnean/article-abstract/191/3/381/5579407 by guest on 28 October 2019

Table 1. Populations of *Vellozia auriculata* sampled in this study. The groups used in AMOVA, diversity indexes and geographical groups of Geneland are shown. Pop: population, X: longitude, Y: latitude, AMOVA schemes: (1) geographical with four groups, (2) geographical with three groups and (3) edaphic with two groups. *N*, number of samples; *h*, haplotype diversity, π, nucleotide diversity; σ, standard deviation; Gl, group inferred by Geneland; P, percentage of polymorphic loci; I, Shannon index and H_e, mean expected heterozygosity

Pop	Locality	X	Y	Altitude (m)	AMOVA schemes			plastid DNA					ISSR			
					1	2	3	*N*	Haplotype	*h* (σ)	π (σ)	GI	*N*	*P*	I(σ)	H_e(σ)
a1	Nascente do Córrego das Éguas	-43.35	-18.22	1537	a	a	rock	12	h3 (12)	0 (0)	0 (0)	1	18	0.73	0.36 (0.02)	0.235 (0.016)
a2	Morro do Alecrim	-43.35	-18.19	1336	a	a	rock	12	h4 (12)	0 (0)	0 (0)	2	17	0.72	0.33 (0.02)	0.219 (0.016)
a3	Pico Dois Irmãos	-43.32	-18.21	1561	a	a	rock	12	h1 (4), h3 (8)	0.4848 (0.1059)	0.000735 (0.000604)	1	16	0.76	0.35 (0.02)	0.225 (0.015)
a4	Mata do Izidoro	-43.28	-18.19	1212	a	a	rock	11	h2 (4), h3 (1), h5 (6)	0.6182 (0.1038)	0.001185 (0.000861)	3	15	0.76	0.38 (0.02)	0.251 (0.016)
b1	Serra de Pedra Menina	-43.14	-18.12	1437	b	b+c	rock	12	h5 (11), h6 (1)	0.1667 (0.1343)	0.000253 (0.000304)	4	16	0.78	0.36 (0.02)	0.235 (0.015)
b2	Serra Dois Irmãos	-43.13	-18.08	1246	b	b+c	rock	12	h5 (3), h7 (9)	0.4091 (0.1333)	0.000620 (0.000536)	5	13	0.76	0.37 (0.02)	0.244 (0.016)
b3	Penha de França	-43.06	-18.03	1013	b	b+c	sand	12	h5 (12)	0 (0)	0 (0)	4	0	-	-	-
b4	Serra do Ambrósio	-43.05	-18.13	880	b	b+c	sand	12	h5 (4), h7 (2), h8 (6)	0.6667 (0.0910)	0.001148 (0.000834)	6	17	0.73	0.35 (0.02)	0.234 (0.016)
c1	Ribeirão de Areia	-42.96	-18.02	1097	c	b+c	sand	11	h5 (11)	0 (0)	0 (0)	4	0	-	-	-
c2	Oeste da Serra Negra	-42.88	-18.07	1362	c	b+c	sand	11	h5 (1), h7 (6), h9 (4)	0.6182 (0.1038)	0.000661 (0.000565)	5	0	-	-	-
d1	Parque Estadual da Serra Negra	-42.70	-17.99	1069	d	d	sand	12	h5 (2), h10 (10)	0.3030 (0.1475)	0.000459 (0.000440)	7	15	0.73	0.35 (0.02)	0.231 (0.015)
Total								129		0.7923 (0.0255)	0.001552 (0.000971)		127	0.99	0.43 (0.02)	0.280 (0.014)

© 2019 The Linnean Society of London, *Botanical Journal of the Linnean Society*, 2019, **191**, 381–398

Downloaded from https://academic.oup.com/botlinnean/article-abstract/191/3/381/5579407 by guest on 28 October 2019

of 10 m apart, as the plants are capable of vegetative propagation. Vouchers are deposited in the herbarium BHCB of the Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais (*BHCB 106375*, *117217*, *123137*, *105142*).

For the phylogeographic study, we obtained sequences of *rpl32-trnL* and *psbD-trnT* plastid DNA intergenic regions using protocols of CTAB DNA extraction (Doyle & Doyle, 1987), PCR amplification and Sanger sequencing described in Barbosa *et al.* (2012). We built consensuses of forward and reverse sequences using Staden Package v.1.7.0 (Staden, 1996) and the alignment with Muscle implemented in MEGA 7 (Kumar, Stecher & Tamura, 2016). We visually inspected the alignment matrix to detect spurious polymorphic sites and encode indels according to the modified complex indel coding scheme (MCIC, Simmons & Ochoterena, 2000) with SeqState 1.4.1 (Müller, 2005). Sequences were deposited in GenBank (accession codes, MG953391–MG953412).

To test population geographical and edaphic structure in the genome as a whole, we used ISSR. We phenotyped with replicates a subset of individuals from three localities using 22 primers. For the seven primers that gave well-defined bands (Supplementary Table S1), we calculated the error rate as the ratio between the total number of mismatched phenotypes (band presence versus band absence) and the product of the number of replicated phenotypes and the number of replicates (Pompanon *et al.*, 2005). Including only bands ranging from 300 to 1500 bp (as the amplification of fragments out of this interval is less reliable), we observed an error rate of 5.05%, similar to that observed in another ISSR study (Casazza *et al.*, 2013). Using these seven primers, we phenotyped 127 individuals from the eight previously known populations, as three of 11 populations sampled for the plastid DNA study were discovered after we carried out the ISSR phenotyping. The addition of individuals from localities discovered after the data acquisition are not expected to change the observed patterns (as discussed below, populations from regions b and d form a group and there is evidence of isolation-by-distance, based on ISSR data; due to this, it is likely that the variability of populations not sampled was already represented). Amplification reactions were conducted in a total volume of 19 µL containing one unit Taq polymerase, 1× reaction buffer [75 mM Tris - HCl (pH 9.0), 50 mM KCl, 20 mM $(NH_4)_2SO_4$)], 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.32 µM primer and genomic DNA. The program consisted of pre-dissociation at 94 ˚C for 4 min, 37 amplification cycles with dissociation at 94 ˚C for 1 min, annealing at 47.6 ˚C or 50 ˚C for 1 min (Supplementary Table S1), extension at 72 ˚C for 2 min, and a final extension at 72 ˚C for 7 min. We included negative controls in all PCRs. We separated PCR products at 60 V for 4 h, using 1.5% agarose gel electrophoresis with 0.5× TAE buffer, stained the gel with ethidium bromide and photographed it under UV light. To construct an array of ISSR phenotypes based on presence (1) or absence (0) of bands, we used a 100 bp DNA standard (Ludwig), assuming that fragments with similar electrophoretic mobility are homologous for the same primer. To ensure that the observed patterns are not derived from association of samples from a given population or region on a single gel, we split samples from a same locality across multiple gels.

Plastid DNA analyses

We calculated haplotype (h) and nucleotide (π) diversities with Arlequin v.3.5.1.3 (Excoffier & Lischer, 2010) and inferred the plastid haplotype network with the integer neighbor-joining networks algorithm with 'old-style = 1', implemented in POPART 1.6 beta (Leigh & Bryant, 2015).

We performed analyses of molecular variance (AMOVA), φ_{PT} and φ_{RT} calculations using GenAlEx 6.5 software (Peakall & Smouse, 2006), for three distinct hierarchical schemes: (1) geographical with four groups; (2) geographical with three groups and (3) edaphic with two groups (epilithic or psammophile) (Table 1). Scheme 1 (regions a–d) follow geographical islands with > 50% suitability for the species (Miranda, 2012). As geographical regions b and c are connected by an area of higher elevation than the one which separates these regions from a and d, we set also scheme 2 (regions a, b+c and d). φ_{PT} and φ_{RT} significance values were tested by 9999 permutations.

We detected population structure using Geneland Bayesian inference of spatial clustering (Guillot, Mortier & Estoup, 2005). We tested the existence of one to 11 geographical groups in ten independent runs, composed of 100 000 iterations, with initial burnin of 20%, using the spatial model with correlated allelic frequencies. To test the occurrence of isolation-by-distance, we performed Mantel test (Mantel, 1967) from matrices of mean numbers of site-to-site genetic substitutions and geographical distances between populations, with GenAlEx 6.5 (Peakall & Smouse, 2006).

For populations (Table 1) and the species, we verified adherence of mismatch distributions to demographic and geographical expansion models with the sum of square deviations (SSD), and calculate raggedness statistic in Arlequin 3.5.1.3 (Excoffier & Lischer, 2010). The neutrality tests *D* (Tajima, 1989) and *Fs* (Fu, 1997) were also evaluated with Arlequin 3.5.1.3. We calculated the R^2 index (Ramos-Onsins & Rozas, 2002) and tested its significance by simulating 1000 replicates with DnaSP 5.10 (Librado & Rozas, 2009).

Downloaded from https://academic.oup.com/botlinnean/article-abstract/191/3/381/5579407 by guest on 28 October 2019

We specified the Bayesian phylogenetic inference with BEAUti 1.8.1 and implemented it in BEAST 1.8.1 (Drummond *et al.*, 2012). Based on the Bayesian information criterion (BIC), implemented in MEGA 7 (Darriba *et al.*, 2012; Kumar *et al.*, 2016), we selected the HKY model for partitions encoding point mutations. We included indel binary encoding as an independent data partition, using Dollo's stochastic model, based on the premise that characters lost in a deletion will not be re-acquired. We empirically calculated nucleotide frequencies, fixing it to observed proportions in the data set and we inferred trees under a constant size model. For temporal calibration, we set the divergence between *V. auriculata* and *V. gigantea* N.L.Menezes & Mello-Silva to 5.4371 Mya (σ = 1.5), following results of temporal calibration methodology presented on supplementary material. We performed 10 000 000 iterations (with initial burnin of 20%), verified convergence of results and effective sample sizes (ESS) with Tracer v.1.6 (http://tree.bio.ed.ac.uk/software/tracer/, last accessed November 2017), summarized trees with TreeAnnotator v.1.8.1 and visualized the result using FigTree v.1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/, last accessed November 2017).

We compared diversification hypotheses using approximate Bayesian computation (ABC; Beaumont, 2010), implemented in DIY-ABC 2.1.0 (Cornuet *et al.*, 2014). We evaluated five evolutionary scenarios: one scenario of fragmentation of previously continuous distribution (Fig. 2A); and four scenarios of colonization through the founder effect, one from each of the four geographical regions of high suitability for *V. auriculata* (Fig. 2B--E). Bottlenecks/founder effects were modelled by an abrupt reduction of population sizes from *Ne* to *Nb* (Fig. 2, Supplementary Table S2). We simulated 2 000 000 sequences in each of the scenarios and computed the summary statistics of simulated and observed data. We selected the most explanatory model from a logistic regression, which indicated the probability of each scenario in relation to the deviations between observed and simulated data (Fagundes *et al.*, 2007; Beaumont, 2008).

To estimate effective immigration rates (M) and effective population sizes (Θ) we used Migrate-N v.3.6.11 (Fig. 3; Beerli, 2006). Also, with Migrate-N, we compared four past demographic models, varying the number of geographical clusters and the level of connectivity between these regions. Models are represented in Fig. 3 and can be described as (1) four geographical regions (a–d) connected as stepping stones (Fig. 3B); (2) four geographical regions (a–d) interconnected (Fig. 3C); (3) three geographical regions (a, b+c and d) connected as stepping stones (Fig. 3C) and (4) three geographical regions (a, b+c and d) interconnected (Fig. 3E). Under each model, we used one long Bayesian heated chain of 2 000 000 steps (sampled every 100 generations), with a burnin of 400 000 steps, with a uniform prior ranging from 0 to 10 000 for the estimates of M and a uniform prior ranging from 0 to 0.1 for estimates of Θ. The estimated number of immigrants per generation Nm was calculated as $Nm = M_{i\text{-}j} \times (\Theta_{i} + \Theta_{j})$, as plastid markers are haploid and *V. auriculata* is monoecious. Modes of estimated M and Θ were used in the calculation.

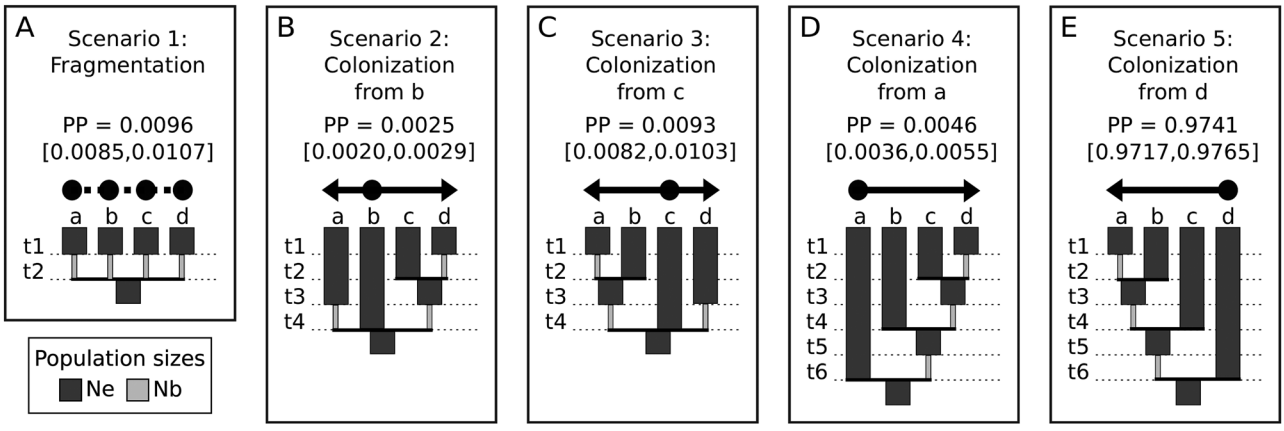


Figure 2. Scenarios tested using approximate Bayesian computation (ABC) for *Vellozia auriculata*, based on plastid DNA haplotypes. A–E, Scenario 1: fragmentation of previously connected distribution (A); scenario 2: colonization from region b (B); scenario 3: colonization from region c (C); scenario 4: colonization from west to east (D) and scenario 5: colonization from east to west (E). The positions of regions a–d are shown in Fig. 1. PP: posterior probability; t1–t6: divergence times (time is not in scale); Ne: population size, Nb: population size during a bottleneck (scenario 1) or colonization population size (founding effect, scenarios 2–5). Parameter priors for all scenarios and estimates for scenario 5 are presented on Supplementary Fig. S2.

Downloaded from https://academic.oup.com/botlinnean/article-abstract/191/3/381/5579407 by guest on 28 October 2019

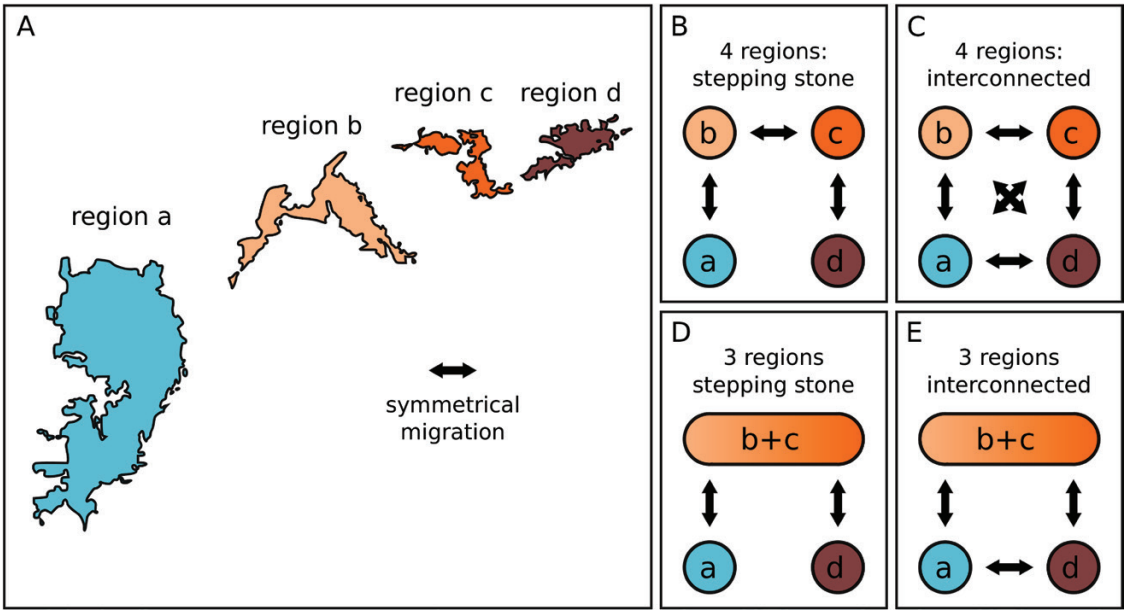


Figure 3. Scenarios tested in Migrate-N model selection for *Vellozia auriculata*, based on plastid DNA haplotypes. A–E, Scenario 1: four geographical regions (a–d) connected as stepping stones (B); scenario 2: four geographical regions (a–d) interconnected (C); scenario 3: three geographical regions (a, b+c and d) connected as stepping stones (D) and scenario 4: three geographical regions (a, b+c and d) interconnected (E). The limits of the four islands of habitat suitability are shown in A and the positions of these regions are shown in Fig. 1.

ISSR ANALYSES

We calculated diversity parameters (P, proportion of polymorphic loci; I, Shannon index of information; H_e, mean expected heterozygosity), Nei's unbiased genetic distance (Nei, 1978), AMOVA, φ_{PT} and φ_{RT} using GenAlEx v.6.5 (Peakall & Smouse, 2006). For AMOVA, we organized the group in a hierarchy according to schemes 2 and 3 (Table 1). The significance values of φ_{PT} and φ_{RT} were tested by 9999 permutations.

After conversion of the presence/absence matrix to compatible format, we constructed an unrooted neighbor-joining dendrogram from Nei's unbiased genetic distance in MEGA 7 (Kumar *et al.*, 2016) and obtained bootstrap values for this dendrogram with AFLP-SURV, using 1000 replicates (Felsenstein, 1989; Vekemans *et al.*, 2002). To test the occurrence of isolation-by-distance, we performed a Mantel test (Mantel, 1967), from matrices of Nei's unbiased genetic distance and geographical distance between populations, with GenAlEx 6.5 (Peakall & Smouse, 2006).

To test the existence of distinct genetic groups and assign individuals to populations, we carried out a Bayesian analysis using STRUCTURE 2.3.4 (Falush, Stephens & Pritchard, 2007). We varied the number of presumed gene pools (K) from one to ten, performing ten independent runs composed of 500 000 MCMC iterations for each K, with an initial burnin of 20% and models of correlated allele frequencies and admixture. To infer the number of genetic groups (populations), we calculated the mean of each probability value of K during all runs, as suggested by Pritchard, Stephens & Donnelly (2000), and the Delta K statistic according to Evanno, Regnaut & Goudet (2005) using STRUCTURE HARVESTER (Earl & Holdt, 2012).

SPECIES DISTRIBUTION MODELS

We used SDM to estimate the occurrence suitability for *V. auriculata*, testing 16 algorithms of SDM: Bioclim; boosted regression trees (BRT); classification or regression tree (CART); domain, generalized additive models (GAM); generalized boosted models (GBM); generalized linear models (GLM); maximum likelihood analysis of species occurrence probability from presence-only data for modelling species distributions (maxlike); Mahalanobis distance; Maxent; mixture discriminant analysis (MDA); multivariate adaptive regression spline (MARS); random forest (RF); recursive partitioning and regression trees (RPART); regularization paths for generalized linear models via coordinate descent (GLMnet) and support vector machine (SVM, Franklin & Miller, 2009). To test the best prediction model, we use original records obtained by field collections carried out in the distribution area of *V. auriculata* in 2011 by Miranda (2012) and, when

Downloaded from https://academic.oup.com/botlinnean/article-abstract/191/3/381/5579407 by guest on 28 October 2019

necessary, 1000 pseudoabsences in training models. Presence points were obtained using GPS of high precision (*c*. 1 m) across *V. auriculata* populations (Miranda, 2012). Pseudoabsences were selected with the upper limit of 800 m in the Espinhaço Range. The same limit was used for model predictions, because *V. auriculata* is restricted to campo rupestre. All models were run in R software package SDM (Naimi & Araujo, 2016), except Maxent, which was run in Maxent software (Steven, Miroslav & Robert, 2018). To validate models, we used presence and absences data collected in expeditions independently of training data (Miranda, 2012). We use bioclimatic variables from WorldClim and elevation (SRTM) as predictors (Hasenack *et al.*, 2010). As the elevation variable has *c*. 90 m of resolution, bioclimatic variables were resized accordingly. To reduce variable correlation, which may affect some algorithms, we tested correlation between them with a threshold of 0.7. As some variables were highly correlated, we used the five least correlated variables (elevation, annual mean temperature, mean diurnal range, isothermality and precipitation of driest month), each one representing a set of highly correlated variables. To evaluate models, we used the area under the curve (AUC), based on presences and absences independently collected (see Miranda, 2012). To build the suitability projection of past scenarios of species distribution (21 ky and 120–140 ky), we used the algorithm with the highest AUC and the climatic models CCSM and MIROC from WorldClim (http://www.worldclim.org/, last accessed February 2018). We ran new models with all points of occurrence and subtracted current prediction by each scenario prediction to show the change of suitability between past and current scenarios.

V. AURICULATA CONSERVATION

Associating results from plastid DNA and ISSR data, we described evolutionarily significant units (ESU) and independent management units (MU), according to Moritz (1994).

RESULTS

PHYLOGEOGRAPHIC ANALYSES OF PLASTID DNA

Amplifications of plastid DNA intergenic regions *rpl32-trnL* and *psbD-trnT* provided *c*. 800 and 1000 base pairs, respectively. After sequencing and trimming, the alignment of 129 individuals of *V. auriculata* and one individual of *V. gigantea* comprised 533 sites for *rpl32-trnL* and 797 sites for *psbD-trnT*. These regions were concatenated for analyses. In *V. auriculata*, we observed nine polymorphic sites (one transition and eight transversions) and one imperfect poly T microsatellite of variable length. We chose to use the poly T in the analyses because it was consistent on forward and reverse sequencing. Of the ten polymorphisms we found, eight were informative and, in combination, revealed ten haplotypes (Table 2). With addition of *V. gigantea*, 18 point mutations were included in the matrix.

Haplotype diversity for *V. auriculata* was 0.7923 ($\sigma = 0.0255$) and total nucleotide diversity was 0.001552 ($\sigma = 0.000971$). In the populations, nucleotide diversity ranged from 0 to 0.001185 ($\sigma = 0.000861$) and haplotype diversity from 0 to 0.6667 ($\sigma = 0.0910$, Table 1). The

Table 2. Description of the ten plastid DNA haplotypes of *Vellozia auriculata*, derived from the concatenation of the alignments of the *rpl32-trnL* and *psbD-trnT* plastid DNA intergenic regions. Numbers indicate positions in alignment; dots indicate that the character state is the same as that of haplotype h1 SeqState Indel codification is presented after the poly T sequence.

	rpl32-trnL														*psbD-trnT*			
	162	224	323	486	[497	498	499	500	501	502	503	504	505]	513	19	358	616	617
h1	T	G	T	C	-	-	-	-	-	-	-	-	- 1	A	C	G	T	A
h2	.	.	.	.	-	-	-	-	-	-	-	-	- 1	C	.	.	.	C
h3	.	.	G	.	-	-	-	-	-	-	-	-	- 1	C	.	.	.	.
h4	.	.	G	.	T	T	-	-	-	-	-	-	- 3	C	.	T	G	.
h5	.	.	G	.	T	-	-	-	-	-	-	-	- 2	C	.	.	.	.
h6	.	T	G	.	T	-	-	-	-	-	-	-	- 2	C	A	.	.	.
h7	.	.	G	T	T	-	-	-	-	-	-	-	- 2	.	.	.	.	.
h8	A	.	G	T	T	-	-	-	-	-	-	-	- 2	.	.	.	.	.
h9	.	.	G	T	T	-	-	-	-	-	-	-	- 2	.	.	.	.	C
h10	.	.	G	T	T	T	T	T	T	T	T	T	C 0	C	.	.	.	.

Downloaded from https://academic.oup.com/botlinnean/article-abstract/191/3/381/5579407 by guest on 28 October 2019

number of haplotypes per population ranged from one to three, with an average of 1.9, and the populations with the greatest diversity were c2, b4 and a4 (Table 1). The seven haplotypes that occur in terminal positions in the haplotype network were restricted to single populations (Fig. 1). Haplotype h5 was the most frequent (39%), occurring in the four geographical regions of high suitability for *V. auriculata*; it was present in the majority of the populations, except the three located at the extreme west of the distribution (Fig. 1).

The geographical AMOVA with four groups showed most of the variation (57%) among populations within geographical regions, with 16% of the variation between groups. In the geographical AMOVA with three groups, although most of variation was also concentrated within geographical regions (48%), we observed 28% of variation between them. In the AMOVA between populations from distinct edaphic conditions, we found only 6% of the variation between groups, with 67% of the variation being among populations within them. In all three cases, φ_{PT} values were significant ($P < 0.001$) and > 0.73; φ_{RT} values were also significant ($P < 0.001$), ranging from 0.059 to 0.277 (Table 3).

Geneland genetic-geographical analysis indicated occurrence of seven groups in all independent parallel runs (Table 1, Supplementary Fig. S1). Such groups are compatible with those defined in AMOVA with three groups, since populations of region b are grouped with populations of region c in groups 4 and 5. The Mantel test was not significant ($R = 0.279$, $P = 0.097$), suggesting no isolation-by-distance.

With Bayesian phylogenetic inference, ESS > 1300 was reached, but few branches had posterior probabilities (PP) > 95%; only the clades of haplotypes h1, h2 and h3, and h7, h8 and h9 were recovered with high PP (Fig. 4). The analysis suggests that all current haplotypes of *V. auriculata* have a common ancestor at 1.0391 Myr (95% HPD: 0.0003–2.0324 Myr) and clades formed by h1–h3 haplotypes and h7–h9 originated at 0.2727 Myr (95% HPD: 0.0.0001–0.6769 Myr) and 0.2055 Myr (95% HPD: 0.0001–0.5151 Mya), respectively.

Table 3. AMOVA of *Vellozia auriculata* populations based on plastid DNA. The populations were grouped according to Table 1 and all φ values were significant ($P < 0.001$). AMOVA schemes: (1) geographical with four groups; (2) geographical with three groups and (3) edaphic with two groups. Sum of squares and degrees of freedom are presented between parenthesis (s.s./d.f.).

Source of variation	AMOVA schemes for plastid DNA			AMOVA schemes for ISSR	
	1	2	3	2	3
Among groups	16% (40.0/3)	28% (38.9/2)	6% (13.7/1)	5% (195.8/2)	4% (120.7/1)
Among populations within groups	57% (55.6/7)	48% (56.7/8)	67% (82.3/9)	7% (255.3/5)	9% (330.5/6)
Within populations	27% (35.5/118)	25% (35.5/118)	27% (35.6/118)	87% (2569.6/119)	87% (2569.6/119)
φ_{RT}	0.161	0.277	0.059	0.052	0.042
φ_{PT}	0.735	0.753	0.731	0.127	0.127

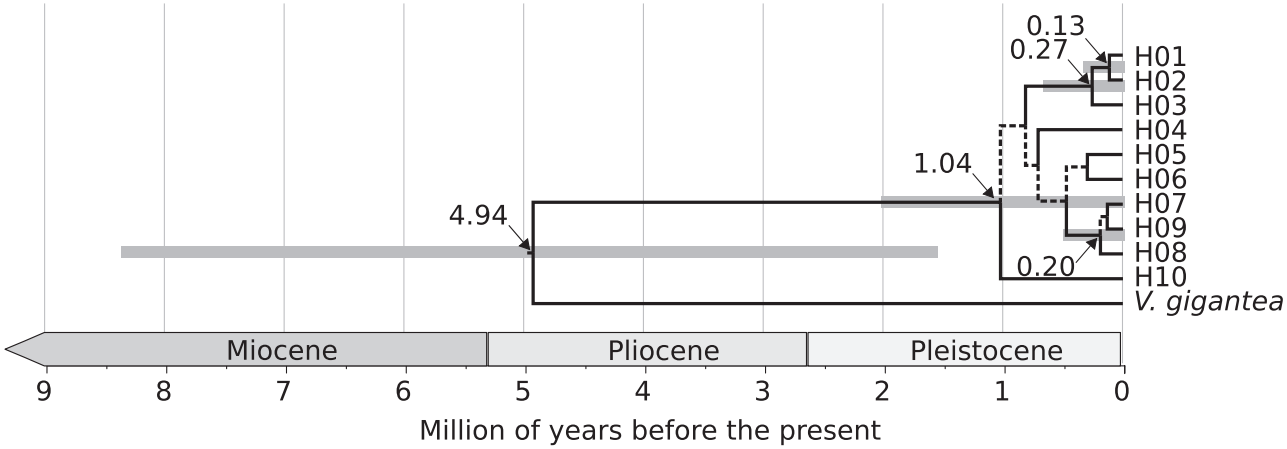


Figure 4. Bayesian phylogenetic inference of *Vellozia auriculata* plastid DNA haplotypes. Branches with posterior probability (PP) > 0.99 are shown in solid lines and branches with PP < 0.50 are shown in dashed lines. The ages of the clades with PP > 0.95 are presented by arrows and the bars indicate the 95% HPD intervals for the divergence times.

Downloaded from https://academic.oup.com/botlinnean/article-abstract/191/3/381/5579407 by guest on 28 October 2019

Table 4. Neutrality tests based on plastid DNA for *Vellozia auriculata*. *: $P < 0.001$

Population	Tajima's D	FS	R2	Raggedness index	SSD (demographic)	SSD (spatial)
a3	1.3564	2.3279	0.2424	0.7355	0.1828	0.1283
a4	1.8024	1.5454	0.2606	0.3984	0.1246	0.0871
b1	−1.4514	0.4318	0.2764	0.7500	0.0370	0.0152
b2	0.6879	1.9611	0.2045	0.6839	*0.3347	0.0914
b4	1.7228	1.5987	0.2525	0.2176	0.0598	0.0463
c2	−0.5063	0.3227	0.2096	0.1167	0.0092	0.0091
d1	−0.2481	1.3844	0.1515	0.6694	0.2787	0.0502
V. auriculata (total)	0.2793	−0.3545	0.0826	0.0628	0.0191	0.0132

The neutrality tests were not significant ($P > 0.05$). The SSDs of mismatch distributions and the raggedness statistics from populations or species (Table 4) were not significant, except the adjustment to the demographic expansion model of population b2, that had $P < 0.001$. Of the five models tested in ABC, scenario 5 presented PP significantly higher than others (0.9741, 95% HPD = 0.9717–0.9765), indicating colonization from east to west (Fig. 2E, Supplementary Fig. 2). For scenario 5, the mode of estimated population size for each island size was 8500 (95% HPD = 4690–9860) and the mode of bottleneck population size (colonization population size) was one (95% HPD = 1–10, Supplementary Table S2).

Migrate-N model selection and parameter estimate indicated that migration between regions is best explained by a model with three groups, when all these groups are interconnected (PP = 0.793, Fig. 3E, Table 5). For this scenario, estimates of M and Θ are presented in Table 6; the average number of immigrants per generation between regions a and b+c was $Nm_{a\text{-}bc} = 0.8941$, between regions a and d $Nm_{a\text{-}d} = 0.5459$ and between regions b+c and d $Nm_{bc\text{-}d} = 1.0215$.

Table 5. Migrate-N model selection. Maximum likelihood estimate (lnMLE) and model posterior probability (PP)

Model	lnMLE	PP
Four regions, stepping stone	−1853.56	0.003
Four regions, interconnected	−1849.44	0.163
Three regions, stepping stone	−1850.83	0.041
Three regions, interconnected	−1847.86	0.793

ISSR

The seven primers selected resulted in 147 loci ranging from 300 to 1500 bp, 84% of which were polymorphic (Supplementary Table S1). Eighteen percent of the individuals had missing data, ranging from 10 to 20%. Populations had mean expected heterozygosity (H_e) and Shannon index (I) averages of 0.234 and 0.356, respectively. AMOVA indicated that 87% of variation was found within populations (all schemes $\varphi_{PT} = 0.127$, $P < 0.001$; scheme 2: $\varphi_{RT} = 0.052$; scheme 3: $\varphi_{RT} = 0.042$). In the AMOVA hierarchized according to three geographical groups, we observed only 5% of variation between geographic regions and 7% between populations within regions; in the AMOVA hierarchized according to substrate types, we found only 4% of variation between edaphic groups and 9% among populations within these groups (Table 3).

The similarity dendrogram reflected the geographical distribution of populations (Figs 1, 5). The separation between region a and the others had high support (95% bootstrap), but similarity relations between populations of regions b and d had low support. The Mantel test indicated that there is a significant positive correlation between geographical and genetic distances ($R = 0.845$, $P = 0.01$). The genetic distances used in Mantel tests are presented in Supplementary Table S3.

STRUCTURE indicated the occurrence of two genetic groups, mainly represented by western (populations a1, a2, a3 and a4) and the central-eastern (populations b1, b2, b4 and d1) regions (Fig. 6). Populations b1 and b2, which occupy an intermediate position in the geographical distribution, included several individuals with mixed composition. Also, four individuals from populations a1, a2 and d1 showed mixed genetic composition.

Species distribution models

Maxent gave the highest AUC in the validation (0.95; details in Supplementary Fig. S5 and S6) and thus was chosen as the best model. This model predicted high suitability zones in regions where populations of *V. auriculata* were sampled for the current scenario (0 ky, Fig. 7A). Past model projections indicated decrease in suitability in all scenarios and climatic models (Fig. 7B, D, F), but showed persistence of the current pattern of suitability discontinuity (Fig. 7C, E, G). No scenario presented evidences for connection of populations in the past and the change in suitability in all scenarios

Downloaded from https://academic.oup.com/botlinnean/article-abstract/191/3/381/5579407 by guest on 28 October 2019

Table 6. Migrate-N distribution of the estimates of parameters Θ and M, for the best model (scenario 4: three interconnected regions, Fig. 3 E). Q, quantile

Parameter	Mean	Mode	Q2.5	Q25	Q50	Q75	Q97.5
Θ_{a}	0.0011	0.0010	0.0000	0.0003	0.0014	0.0017	0.0029
Θ_{bc}	0.0008	0.0008	0.0000	0.0001	0.0012	0.0013	0.0025
Θ_{d}	0.0015	0.0008	0.0000	0.0001	0.0012	0.0013	0.0031
$M_{a\text{-}bc}$	1072.2	496.7	0.0	180.0	876.7	1000.0	2720.0
$M_{a\text{-}d}$	859.4	303.3	0.0	46.7	656.7	706.7	2366.7
$M_{bc\text{-}d}$	1413.0	663.3	0.0	266.7	1103.3	1273.3	3686.7

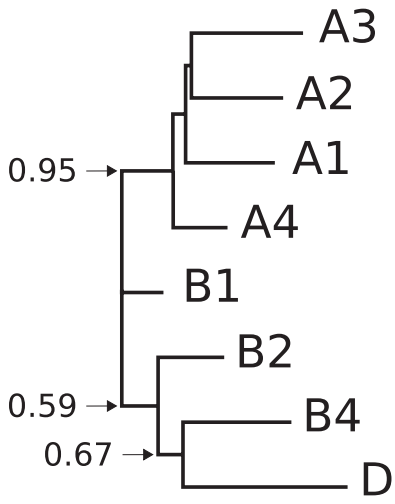


Figure 5. Unrooted neighbor-joining dendrogram of *Vellozia auriculata* populations, using Nei's unbiased genetic distance, based on 147 ISSR loci. Bootstrap values > 50% are presented. For population names see Table 1.

presented similar results (correlation above 0.95). Even if all scenarios presented highly suitability reduction in current areas of occurrence of *V. auriculata* (Fig. 7C, E, G), they showed persistence of suitability, even when subtle, in all regions corresponding to current suitability islands (regions a–d), except region d during the last interglacial (Fig. 7F).

DISCUSSION

Vellozia auriculata, a micro-endemic species from campo rupestre, shows limited zygotic migration among populations and low genetic structure of gametic flow markers (ISSR) when compared to plastid DNA markers; there is no sign of substrate-driven genetic structure and no evidence of expansion from Pleistocene refugia or habitat fragmentation due to Pleistocene climate oscillations. These results indicate that (1) seed flow between populations is low and even virtually absent between major areas of the species occurrence, (2) gene flow between populations occurs mainly via pollen and (3) populations persisted locally during Pleistocene (i.e. presented prolonged local persistence).

The pattern of restricted zygotic gene flow among *V. auriculata* populations is supported by the strong structure highlighted by Geneland and the plastid DNA AMOVA and by the low average number of immigrants per generation between regions estimated by Migrate-N, in accordance with the absence of obvious mechanism for seed dispersion in this species. The isolation driving this pattern is not expected to be recent, as SDM supports discontinuity of habitat suitability during Pleistocene, in both glacial and interglacial periods. Also, haplotypes that originated *c*. 270 000 kya are not shared between *V. auriculata* populations that are separated by distances < 4 km, suggesting that even short distances of low suitability could be gene flow barriers for this and other micro-endemic species from the campo rupestre. Other plants occurring in the campo rupestre and in outcrops across the globe show similar patterns, and this has been associated with low effective seed vagility in these environments (e.g. Byrne & Hopper, 2008; Palma-Silva *et al.*, 2011; Barbosa *et al.*, 2012; Collevatti *et al.*, 2012; Bonatelli *et al.*, 2014; Pinheiro *et al.*, 2014).

Genetic drift has been an important factor on the evolution of *V. auriculata*, for colonization of new environments (founder effect) and for the isolation-driven differentiation between populations (Steinbauer *et al.*, 2016). The model selected in ABC suggests that the current distribution of *V. auriculata* was attained by occasional dispersion events of few individuals, as the mode of posterior distribution for colonization population size is one. This is also supported by the absence of correlation between genetic and geographical distances for plastid DNA data and the low intrapopulation haplotype diversity, again suggesting low capacity of effective seed dispersion of this micro-endemic species from the campo rupestre. The scheme of three geographical groups tested in the plastid DNA AMOVA (scheme 2) retained most of the species interregional structuring and was compatible

Downloaded from https://academic.oup.com/botlinnean/article-abstract/191/3/381/5579407 by guest on 28 October 2019

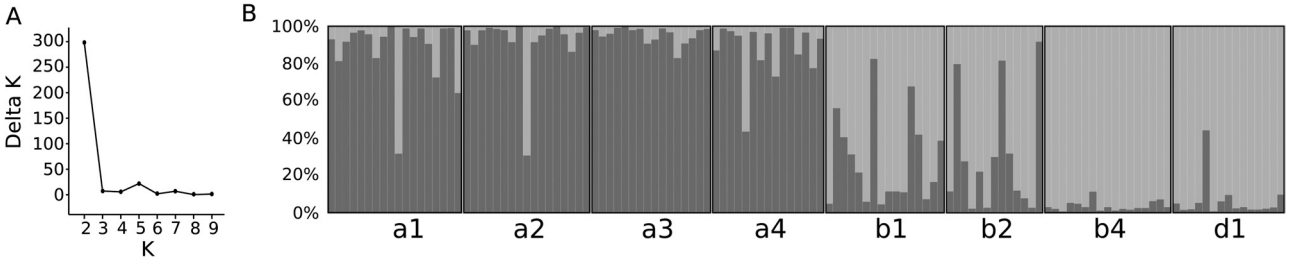


Figure 6. Bayesian clustering analysis of eight populations of *Vellozia auriculata*, based on 147 ISSR loci. The delta K graphic (A) and the summary plot of estimates of ancestry vector (B) are presented. On B, the shades of grey represent different genetic clusters and populations are separated by vertical bars. For population names see Table 1.

with the geographical division proposed by Geneland and Migrate-N model selection, suggesting that for *V. auriculata* elevations of < 700 m are more restrictive to occasional events of gene flow than those between 700 and 800 m, as corroborated by the SDM estimates of environmental unsuitability on lower lands.

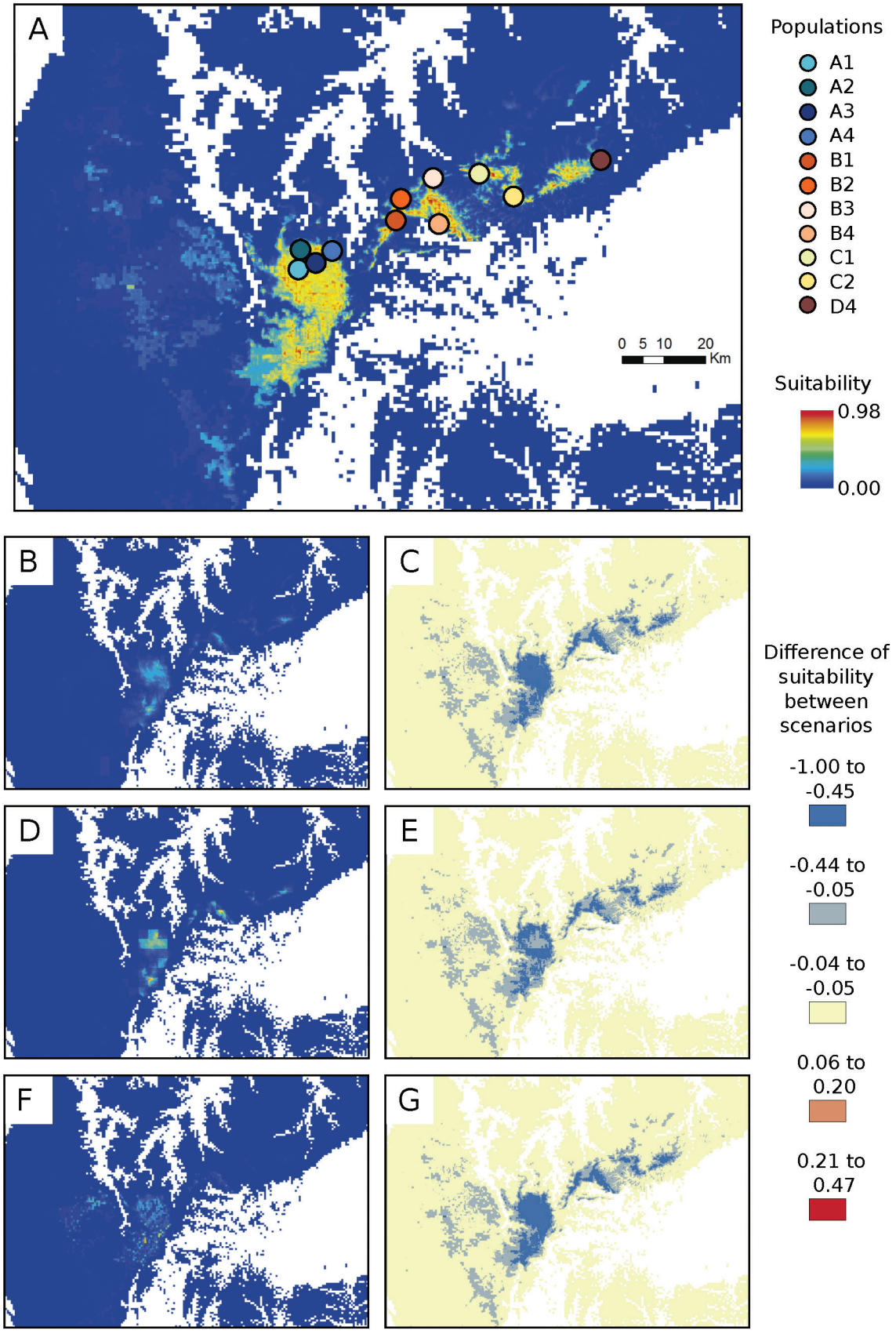
The plastid data and the ISSR phenotypes suggest that substrate has little relevance for the genetic structure, contrasting with the pattern observed in *V. compacta* Mart. (Lousada *et al.*, 2013). For *V. auriculata*, the small portion of variation observed between edaphic groups may even include geographical structuring, since psammophile populations are distributed to the east and epilithic populations to the west of the species distribution. Survival in different substrate types is expected to require specific adaptations (Poot, Hopper & van Diggelen, 2012), and it is possible that the genome of isolated populations of plants occupying different substrates will be at least partially differentiated (Wu, 2001). However, according to the ABC scenario with the highest PP, *V. auriculata* emerged as a psammophile species, subsequently reaching rocky environments. Next-generation sequencing studies could be performed to examine the existence of nuclear genes related to adaptation or edaphic plasticity in *V. auriculata* and to obtain better parameter estimates in ABC for this and other species from the campo rupestre.

The population differentiation in *V. auriculata* is the lowest reported among *Vellozia* spp. phenotyped with ISSR markers (Barbosa, 2011; Lousada *et al.*, 2011, 2013), probably due to its smaller range of occurrence and smaller distance between populations. However, STRUCTURE analysis and the similarity dendrogram indicate the existence of regional groups a and b+c+d for *V. auriculata*, suggesting that gametic gene flow is partially limited, despite the small distance between populations. The existence of groups due to limited gametic gene flow should, however, be interpreted with caution, as the occurrence of isolating by distance indicated by the Mantel test could bias the clustering (Perez *et al.*, 2018). Also, despite the large number of ISSR loci analysed and the geographical and biological coherence of the results presented here, it is important to be aware of ISSR reproducibility, homology and dominance limitations (Roux & Wieczorek, 2009). Future studies with more sophisticated techniques are desirable to deeper explore the intensity of gametic gene flow and existence of genetic clusters across the distributions of campo rupestre endemic species, such as *V. auriculata*.

Our study corroborates the hypothesis that gametic gene flow in the campo rupestre may be more frequent than zygotic gene flow, as plastid DNA haplotypes are more structured through space than ISSR phenotypes. As the mutation rate giving origin to ISSR polymorphisms is expected to be higher than the mutation rate of plastid DNA (Roux & Wieczorek, 2009), this result suggests prolonged low zygotic gene flow compared to gametic gene flow and can be interpreted with confidence. If, instead, ISSR was expected to have lower mutation rates than plastid DNA, the observed similarity between populations could represent incomplete lineage sorting and not a sign of recurrent gene flow. Other plants occurring in OCBILs exhibit the same pattern of greater structuring of plastid markers, indicating that this could be a common motive in this type of vegetation (Barbosa, 2011; Palma-Silva *et al.*, 2011; Pinheiro *et al.*, 2014; Tapper *et al.*, 2014a, b; Hmeljevski *et al.*, 2017). Even when pollen gene flow is recurrent, isolation by low seed gene flow may lead to conflicts between nuclear and plastid genes and consequent reproductive isolation and speciation (Greiner *et al.*, 2011; Greiner & Bock, 2013; Barnard-Kubow, So & Galloway, 2016). Speciation by gene conflict (Crespi & Nosil, 2013) may be common in OCBILs, and studies exploring this issue should be performed.

Geographical retractions and expansions caused by Pleistocene climatic changes have been invoked to explain the differentiation of campo rupestre populations (e.g. Barbosa *et al.*, 2012; Collevatti *et al.*, 2012; Barres *et al.*, 2019). However, in agreement with the hypothesis of prolonged persistence of OCBIL

Downloaded from https://academic.oup.com/botlinnean/article-abstract/191/3/381/5579407 by guest on 28 October 2019



Downloaded from https://academic.oup.com/botlinnean/article-abstract/191/3/381/5579407 by guest on 28 October 2019

lineages and the high environmental specificity of campo rupestre plants (Hopper, 2009; Silveira *et al.*, 2016), the phylogeographical patterns of *V. auriculata* do not corroborate that Pleistocene glacial and interglacial periods were triggers for its population differentiation. The SDM indicates lasting suitability discontinuity between regions and shows that, despite the reduction of suitability in past scenarios, SDM indicates that current areas presenting highest suitability match the suitable areas of glacial and interglacial periods. Thus, our analyses support the prolonged persistence in isolation of *V. auriculata* populations during Pleistocene climate oscillations across its current distribution. There is evidence that local persistence during Pleistocene climatic oscillations also occurred in *Minaria* T.U.P.Konno & Rapini (Ribeiro *et al.*, 2014), *Encholirium horridum* L.B.Sm. (Hmeljevski *et al.*, 2017) and in plant species occurring in Australian OCBILs (Byrne & Hopper, 2008; Tapper *et al.*, 2014a, b; Nistelberger *et al.*, 2015). It is expected that other plants from the campo rupestre had prolonged persistence on isolation. The spread of this pattern in contrast to the pattern of recolonization from refugia frequently observed in currently more continuous Neotropical environments, such as savannas and forests (Leal, Palma-Silva & Pinheiro, 2016), should be addressed in future work. Plants with widespread distribution in the campo rupestre could offer interesting opportunities to test this hypothesis. It is important to highlight that Neotropical biodiversity evolution was a complex process (Rull, 2013; Rull, 2015; Antonelli *et al.*, 2018), and thus it is not expected that campo rupestre flora presents strictly concordant responses to past climatic changes. Organisms with distinct niche and life strategies are expected to experience different demographic processes that will lead to idiosyncratic genetic patterns (Massatti & Knowles, 2014).

The low suitability of the eastern region (d) during Pleistocene glacial periods showed by SDM seems to contrast with ABC result, which suggests that *V. auriculata* originated in the eastern region of its distribution. However, the origin of *V. auriculata* pre-dates Pleistocene climatic oscillations, and thus these results are not contradictory. Instead, combined with the occurrence and abundance of the endemic haplotype h10 in population d1, these results suggest that *V. auriculata* persisted locally in eastern populations, even when suitable habitats were limited in this region, again supporting the occurrence of prolonged persistence with isolation of this micro-endemic species from the campo rupestre. Also, niche evolution could be a possible explanation for the estimated low suitability using current data, as Miranda (2012) observed a niche gradient across *V. auriculata* distribution.

The results we presented here support that the western region of occurrence of *V. auriculata* (a) comprise an ESU (Moritz, 1994) distinct from the centre-east region (b+c and d). In addition, the occurrence of exclusive plastid DNA haplotypes in populations a2, a3, a4, b1, b4, c2 and d1 makes each of them an independent MU (Moritz, 1994). It is fundamental that *in situ* and *ex situ* management and conservation plans take into account the high interregional and interpopulation differentiation of *V. auriculata*, including comprehensive protection actions for effective conservation of genetic variability of this species. The extinction of populations in the centre would be especially critical, since it could restrict gametic gene flow between the species distribution edges, potentially decreasing genetic diversity and compromising the future viability of populations. Currently, these populations are not included in conservation units and are subject to anthropogenic impacts that must be urgently interrupted.

We conclude that for *V. auriculata* (1) zygotic gene flow is limited even on short scales, (2) gametic gene flow is recurrent, but limited by distance and environmental unsuitability and (3) climatic changes of Pleistocene did not restrict the species to refugia, with local persistence of populations. Our data contribute to the understanding of the high diversity observed in the campo rupestre by providing more arguments in support of the hypothesis that campo rupestre is both a cradle (due to low zygotic gene flow between populations, which favours population isolation and differentiation) and a lineage museum (since it is able to maintain *in situ* survival of populations during climate changes, Silveira *et al.*, 2016). As genetic studies of non-model organisms are now more accessible than ever, we hope to see new

Figure 7. SDM of *Vellozia auriculata* using Maxent, which presented the best performance between the algorithms tested, for current (A) and past (B, D, F) modelled distributions. B, 21k (maximum glacial) model CCSM; D, 21k (maximum glacial) model MIROC and F, 120–140k. C, E and G show changes in suitability in relation to current conditions, where negative values indicate reduction and positive values indicate increase of suitability. C, 21k (maximum glacial) model CCSM; E, 21k (maximum glacial) model MIROC and G, 120–140k.

Downloaded from https://academic.oup.com/botlinnean/article-abstract/191/3/381/5579407 by guest on 28 October 2019

advancements on phylogeographical understanding of campo rupestre flora in the coming years.

ACKNOWLEDGEMENTS

This work was supported by grants from the Brazilian Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq proc. #567718/2008-5) and Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG proc. #CRA-5316-5.02-07). C.F.F. and M.D.M. received scholarships from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and CNPq, respectively. E.L.B is supported by productivity grant (PQ1B) from CNPq (proc. #311918/2014-0). We thank E.C. Smidt for help with analyses and discussion. We also thank A.M.G. Harley, F.A.O. Silveira, J.E. Santos J., G.M. Mori and four anonymous reviewers for their valuable suggestions. Samples were collected under SISBIO sampling permit number: 10.870–1.

REFERENCES

Alcantara S, **Ree RH**, **Mello-Silva R. 2018.** Accelerated diversification and functional trait evolution in Velloziaceae reveal new insights into the origins of the campos rupestres' exceptional floristic richness. *Annals of Botany* **122:** 165–180.

Antonelli A, **Ariza M**, **Albert J**, **Andermann T**, **Azevedo J**, **Bacon C**, **Faurby S**, **Guedes T**, **Hoorn C**, **Lohmann LG**, **Matos-Maraví P**, **Ritter CD**, **Sanmartín I**, **Silvestro D**, **Tejedor M**, **Ter Steege H**, **Tuomisto H**, **Werneck FP**, **Zizka A**, **Edwards SV. 2018.** Conceptual and empirical advances in Neotropical biodiversity research. *PeerJ* **6:** e5644.

Antonelli A, **Zizka A**, **Silvestro D**, **Scharn R**, **Cascales-Miñana B**, **Bacon CD. 2015.** An engine for global plant diversity: highest evolutionary turnover and emigration in the American tropics. *Frontiers in Genetics* **6:** 130.

Barbosa AR. 2011. *Biossistemática do complexo* Vellozia hirsuta *(Velloziaceae) baseada em análise filogeográfica e genética de populações.* Unpublished Masters Dissertation, Universidade Federal de Minas Gerais.

Barbosa AR, **Fiorini CF**, **Silva-Pereira V**, **Mello-Silva R**, **Borba EL. 2012.** Geographical genetic structuring and phenotypic variation in the *Vellozia hirsuta* (Velloziaceae) ochlospecies complex. *American Journal of Botany* **99:** 1477–1488.

Barnard-Kubow KB, **So N**, **Galloway LF. 2016.** Cytonuclear incompatibility contributes to the early stages of speciation. *Evolution* **70:** 2752–2766.

Barres L, **Batalha-Filho H**, **Schnadelbach AS**, **Roque N. 2019.** Pleistocene climatic changes drove dispersal and isolation of *Richterago discoidea* (Asteraceae), an endemic plant of campos rupestres in the central and eastern Brazilian sky islands. *Botanical Journal of the Linnean Society* **189:** 132–152.

Beaumont MA. 2008. Joint determination of topology, divergence time, and immigration in population trees. In: Matsumura S, Forster P, Renfrew C, eds. *Simulation, genetics, and human prehistory*. Cambridge: McDonald Institute for Archaeological Research, 135–154.

Beaumont MA. 2010. Approximate Bayesian computation in evolution and ecology. *Annual Review of Ecology, Evolution, and Systematics* **41:** 379–406.

Beerli P. 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* **22:** 341–345.

BFG (The Brazil Flora Group). 2015. Growing knowledge: an overview of seed plant diversity in Brazil. *Rodriguésia* **66:** 1085–1113.

Bitencourt C, Rapini A. 2013. Centres of endemism in the Espinhaço Range: identifying cradles and museums of Asclepiadoideae (Apocynaceae). *Systematics and Biodiversity* **11**: 525–536.

Bonatelli IA, **Perez MF**, **Peterson AT**, **Taylor NP**, **Zappi DC**, **Machado MC**, **Koch I**, **Pires AH**, **Moraes EM. 2014.** Interglacial microrefugia and diversification of a cactus species complex: phylogeography and palaeodistributional reconstructions for *Pilosocereus aurisetus* and allies. *Molecular Ecology* **23:** 3044–3063.

Borba EL, **Félix JM**, **Solferini VN**, **Semir J. 2001.** Fly-pollinated *Pleurothallis* (Orchidaceae) species have high genetic variability: evidence from isozyme markers. *American Journal of Botany* **88:** 419–428.

Borba EL, **Funch RR**, **Ribeiro PL**, **Smidt EC**, **Silva-Pereira V. 2007.** Demography, and genetic and morphological variability of the endangered *Sophronitis sincorana* (Orchidaceae) in the Chapada Diamantina, Brazil. *Plant Systematics and Evolution* **267:** 129–146.

Buzatti RSO, **Lemos-Filho JP**, **Bueno ML**, **Lovato MB. 2017.** Multiple Pleistocene refugia in the Brazilian cerrado: evidence from phylogeography and climatic niche modelling of two *Qualea* species (Vochysiaceae). *Botanical Journal of the Linnean Society* **185:** 307–320.

Byrne M, **Hopper SD. 2008.** Granite outcrops as ancient islands in old landscapes: evidence from the phylogeography and population genetics of *Eucalyptus caesia* (Myrtaceae) in Western Australia. *Biological Journal of the Linnean Society* **93:** 177–188.

Camps GA, **Martínez-Meyer E**, **Verga AR**, **Sérsic AN**, **Cosacov A. 2018.** Genetic and climatic approaches reveal effects of Pleistocene refugia and climatic stability in an old giant of the Neotropical dry forest. *Biological Journal of the Linnean Society* **125:** 401–420.

Casazza G, **Grassi F**, **Zecca G**, **Mariotti MG**, **Guerrina M**, **Minuto L. 2013.** Phylogeography of *Primula allionii* (Primulaceae), a narrow endemic of the Maritime Alps. *Botanical Journal of the Linnean Society* **173:** 637–653.

Collevatti RG, **de Castro TG**, **de Souza Lima J**, **de Campos Telles MP. 2012.** Phylogeography of *Tibouchina papyrus* (Pohl) Toledo (Melastomataceae), an endangered tree species from rocky savannas, suggests bidirectional expansion due to climate cooling in the Pleistocene. *Ecology and Evolution* **2:** 1024–1035.

Downloaded from https://academic.oup.com/botlinnean/article-abstract/191/3/381/5579407 by guest on 28 October 2019

Conceição AA, **Rapini A**, **Carmo FF**, **Brito JC**, **Silva GA**, **Neves SPS**, **Jacobi CM. 2016.** Rupestrian grassland vegetation, diversity, and origin. In: Fernandes GW, ed. *Ecology and conservation of mountaintop grasslands in Brazil*. Cham: Springer, 105–127.

Cornuet JM, **Pudlo P**, **Veyssier J**, **Dehne-Garcia A**, **Gautier M**, **Leblois R**, **Marin JM**, **Estoup A. 2014.** DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* **30:** 1187–1189.

Crespi B, **Nosil P. 2013.** Conflictual speciation: species formation via genomic conflict. *Trends in Ecology & Evolution* **28:** 48–57.

Darriba D, **Taboada GL**, **Doallo R**, **Posada D. 2012.** jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9:** 772.

Doyle JJ, **Doyle JL. 1987.** A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin, Botanical Society of America* **19:** 11–15.

Drummond AJ, **Suchard MA**, **Xie D**, **Rambaut A. 2012.** Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29:** 1969–1973.

Drummond G, **Machado A**, **Martins C**, **Mendonça MP**, **Stehmann J. 2008.** *Listas vermelhas das espécies da fauna e da flora ameaçadas de extinção em Minas Gerais*. Belo Horizonte: Fundação Biodiversitas.

Earl DA, **von Holdt BM. 2012.** STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4:** 359–361.

Evanno G, **Regnaut S**, **Goudet J. 2005.** Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14:** 2611–2620.

Excoffier L, **Lischer HE. 2010.** Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10:** 564–567.

Fagundes NJR, **Ray N**, **Beaumont M**, **Neuenschwander S**, **Salzano FM**, **Bonatto SL**, **Excoffier L. 2007.** Statistical evaluation of alternative models of human evolution. *Proceedings of the National Academy of Sciences U.S.A.* **104:** 17614–17619.

Falush D, **Stephens M**, **Pritchard JK. 2007.** Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* **7:** 574–578.

Felsenstein J. 1989. Phylogeny inference package (Version 3.2). *Cladistics* **5:** 164–166.

Franceschinelli EV, **Jacobi CM**, **Drummond MG**, **Resende MF. 2006.** The genetic diversity of two Brazilian *Vellozia* (Velloziaceae) with different patterns of spatial distribution and pollination biology. *Annals of Botany* **97:** 585–592.

Franklin J, **Miller JA. 2009.** *Mapping species distributions spatial inference and prediction*. Cambridge: Cambridge University Press.

Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147:** 915–925.

Giulietti A, **Pirani JR. 1988.** Patterns of geographic distribution of some plant species from the Espinhaço Range, Minas Gerais and Bahia, Brazil. In: Vanzolini PE, Heyer WR, eds. *Proceedings of a Workshop on Neotropical Distribution Patterns*. Rio de Janeiro: Academia Brasileira de Ciências, 39–69.

Greiner S, **Bock R. 2013.** Tuning a ménage à trois: co-evolution and co-adaptation of nuclear and organellar genomes in plants. *BioEssays* **35:** 354–365.

Greiner S, **Rauwolf U**, **Meurer J**, **Herrmann RG. 2011.** The role of plastids in plant speciation. *Molecular Ecology* **20:** 671–691.

Greiner S, **Sobanski J**, **Bock R. 2015.** Why are most organelle genomes transmitted maternally? *Bioessays* **37:** 80–94.

Guillot G, **Mortier F**, **Estoup A. 2005.** GENELAND: a computer package for landscape genetics. *Molecular Ecology Notes* **5:** 712–715.

Hasenack H, **Weber E**, **Fernandes SMC**, **Jonas RJ**, **Duarte EHG. 2010.** *Processamento e organização de modelos digitais de elevação contínuos para os países de língua portuguesa a partir do SRTM*. Available at: http://www.ecologia.ufrgs.br/labgeo, last accessed February 2018.

Hmeljevski KV, **Nazareno AG**, **Leandro Bueno M**, **Dos Reis MS**, **Forzza RC. 2017.** Do plant populations on distinct inselbergs talk to each other? A case study of genetic connectivity of a bromeliad species in an OCBIL landscape. *Ecology and Evolution* **7:** 4704–4716.

Hopper SD. 2009. OCBIL theory: towards an integrated understanding of the evolution, ecology and conservation of biodiversity on old, climatically buffered, infertile landscapes. *Plant and Soil* **322:** 49–86.

Jesus FF, **Solferini VN**, **Semir J**, **Prado PI. 2001.** Local genetic differentiation in *Proteopsis argentea* (Asteraceae), a perennial herb endemic in Brazil. *Plant Systematics and Evolution* **226:** 59–68.

Kumar S, **Stecher G**, **Tamura K. 2016.** MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33:** 1870–1874.

Leal BSS, **Palma-Silva C**, **Pinheiro F. 2016.** Phylogeographic studies depict the role of space and time scales of plant speciation in a highly diverse Neotropical region. *Critical Reviews in Plant Sciences* **35:** 215–30.

Leigh JW, **Bryant D. 2015.** POPART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution* **6:** 1110–1116.

Librado P, **Rozas J. 2009.** DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25:** 1451–1452.

Lousada JM, **Borba EL**, **Ribeiro KT**, **Ribeiro LC**, **Lovato MB. 2011.** Genetic structure and variability of the endemic and vulnerable *Vellozia gigantea* (Velloziaceae) associated with the landscape in the Espinhaço Range, in southeastern Brazil: implications for conservation. *Genetica* **139:** 431–440.

Lousada JM, **Lovato MB**, **Borba EL. 2013.** High genetic divergence and low genetic variability in disjunct populations

Downloaded from https://academic.oup.com/botlinnean/article-abstract/191/3/381/5579407 by guest on 28 October 2019

of the endemic *Vellozia compacta* (Velloziaceae) occurring in two edaphic environments of Brazilian campo rupestre. *Brazilian Journal of Botany* **36:** 45–53.

Main BY. 1997. Granite outcrops: a collective ecosystem. *Journal of the Royal Society of Western Australia* **80:** 113–122.

Mantel N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* **27:** 209–220.

Massatti R, **Knowles LL. 2014.** Microhabitat differences impact phylogeographic concordance of codistributed species: genomic evidence in montane sedges (*Carex* L.) from the Rocky Mountains. *Evolution* **68:** 2833–2846.

Mello-Silva R, **Menezes NL. 1999.** Two new Brazilian Velloziaceae, *Vellozia auriculata* and *Vellozia gigantea*, and a key to the related dracenoid species of *Vellozia*. *Novon* **9:** 536–541.

Melo WA, **Freitas CG**, **Bacon CD**, **Collevatti RG. 2018.** The road to evolutionary success: insights from the demographic history of an Amazonian palm. *Heredity* **121:** 183–195.

Menezes NL, **Mello-Silva R**, **Mayo SJ. 1993.** A cladistic analysis of the Velloziaceae. *Kew Bulletin* **49:** 71–92.

Miranda MD. 2012. *Modelos de distribuição de espécies de* Vellozia *(Velloziaceae) endêmicas da Cadeia do Espinhaço e o efeito amostral sobre os mapas preditivos.* Unpublished Masters Dissertation, Universidade Federal de Minas Gerais.

Moritz C. 1994. Defining 'evolutionarily significant units' for conservation. *Trends in Ecology & Evolution* **9:** 373–375.

Müller K. 2005. SeqState-primer design and sequence statistics for phylogenetic DNA data sets. *Applied Bioinformatics* **4:** 65–69.

Naimi B, **Araujo MB. 2016.** sdm: a reproducible and extensible R platform for species distribution modelling. *Ecography* **39:** 368–375.

Nei M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89:** 583–590.

Nistelberger HM, **Byrne M**, **Coates D**, **Roberts JD. 2015.** Genetic drift drives evolution in the bird-pollinated, terrestrial island endemic *Grevillea georgeana* (Proteaceae). *Botanical Journal of the Linnean Society* **178:** 155–168.

Palma-Silva C, **Wendt T**, **Pinheiro F**, **Barbará T**, **Fay MF**, **Cozzolino S**, **Lexer C. 2011.** Sympatric bromeliad species (*Pitcairnia* spp.) facilitate tests of mechanisms involved in species cohesion and reproductive isolation in Neotropical inselbergs. *Molecular Ecology* **20:** 3185–3201.

Peakall R, **Smouse PE. 2006.** GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6:** 288–295.

Perez MF, **Franco FF**, **Bombonato JR**, **Bonatelli IAS**, **Khan G**, **Romeiro-Brito M**, **Fegies AC**, **Ribeiro PM**, **Silva GAR**, **Moraes EM. 2018.** Assessing population structure in the face of isolation by distance: are we neglecting the problem? *Diversity and Distributions* **24:** 1883–1889.

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

Pinheiro F, **Cozzolino S**, **Draper D**, **de Barros F**, **Félix LP**, **Fay MF**, **Palma-Silva C. 2014.** Rock outcrop orchids reveal the genetic connectivity and diversity of inselbergs of northeastern Brazil. *BMC Evolutionary Biology* **14:** 49.

Pompanon F, **Bonin A**, **Bellemain E**, **Taberlet P. 2005.** Genotyping errors: causes, consequences and solutions. *Nature Reviews. Genetics* **6:** 847–859.

Poot P, **Hopper SD**, **van Diggelen JM. 2012.** Exploring rock fissures: does a specialized root morphology explain endemism on granite outcrops? *Annals of Botany* **110:** 291–300.

Pritchard JK, **Stephens M**, **Donnelly P. 2000.** Inference of population structure using multilocus genotype data. *Genetics* **155:** 945–959.

Ramos ACS, **Lemos-Filho JP**, **Lovato MB. 2009.** Phylogeographical structure of the Neotropical forest tree *Hymenaea courbaril* (Leguminosae: Caesalpinioideae) and its relationship with the vicariant *Hymenaea stigonocarpa* from Cerrado. *Journal of Heredity* **100:** 206–216.

Ramos-Onsins SE, **Rozas J. 2002.** Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution* **19:** 2092–2100.

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

Ribeiro PL, **Borba EL**, **Smidt EC**, **Lambert SM**, **Schnadelbach AS**, **van den Berg C. 2008.** Genetic and morphological variation in the *Bulbophyllum exaltatum* (Orchidaceae) complex occurring in the Brazilian 'campo rupestre': implications for taxonomy and biogeography. *Plant Systematics and Evolution* **270:** 109–137.

Ribeiro PL, **Rapini A**, **Damascena LS**, **van den Berg C. 2014.** Plant diversification in the Espinhaço range: insights from the biogeography of *Minaria* (Apocynaceae). *Taxon* **63:** 1253–1264.

Roux J, **Wieczorek AM. 2009.** Molecular systematics and population genetics of biological invasions: towards a better understanding of invasive species management. *Annals of Applied Biology* **154:** 1–17.

Rull V. 2013. Some problems in the study of the origin of neotropical biodiversity using palaeoecological and molecular phylogenetic evidence. *Systematics and Biodiversity* **11:** 415–423.

Rull V. 2015. Pleistocene speciation is not refuge speciation. *Journal of Biogeography* **42:** 602–604.

Schut AG, **Wardell-Johnson GW**, **Yates CJ**, **Keppel G**, **Baran I**, **Franklin SE**, **Hopper SD**, **Van Niel KP**, **Mucina L**, **Byrne M. 2014.** Rapid characterisation of vegetation structure to predict refugia and climate change impacts across a global biodiversity hotspot. *PLoS One* **9:** e82778.

Silveira FAO, **Negreiros D**, **Barbosa NPU**, **Buisson E**, **Carmo FF**, **Carstensen DW**, **Conceição AA**, **Cornelissen TG**, **Echternacht L**, **Fernandes GW**, **Garcia QS**, **Guerra TJ**, **Jacobi CM**, **Lemos-Filho JP**, **Stradic S**, **Morellato LPC**, **Neves FS**, **Oliveira RS**, **Schaefer CE**, **Viana PL**, **Lambers H. 2016.** Ecology and evolution of plant diversity in the endangered campo

Downloaded from https://academic.oup.com/botlinnean/article-abstract/191/3/381/5579407 by guest on 28 October 2019

rupestre: a neglected conservation priority. *Plant and Soil* **403:** 129–152.

Simmons MP, **Ochoterena H. 2000.** Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology* **49:** 369–381.

Staden R. 1996. The Staden sequence analysis package. *Molecular Biotechnology* **5:** 233–241.

Steinbauer MJ, **Field R**, **Grytnes JA**, **Trigas P**, **Ah-Peng C**, **Attorre F**, **Birks HJB**, **Borges PAV**, **Cardoso P**, **Chou CH**, **De Sanctis M**, **de Sequeira MM**, **Duarte MC**, **Elias RB**, **Fernández-Palacios JM**, **Gabriel R**, **Gereau RE**, **Gillespie RG**, **Greimler J**, **Harter DEV**, **Huang TJ**, **Irl SDH**, **Jeanmonod D**, **Jentsch A**, **Jump AS**, **Kueffer C**, **Nogué S**, **Otto R**, **Price J**, **Romeiras MM**, **Strasberg D**, **Stuessy T**, **Svenning JC**, **Vetaas OR**, **Beierkuhnlein C. 2016.** Topography-driven isolation, speciation and a global increase of endemism with elevation. *Global Ecology and Biogeography* **25:** 1097–1107.

Steven JP, **Miroslav D**, **Robert ES. 2018.** *Maxent software for modeling species niches and distributions (Version 3.4.1).* Available at: http://biodiversityinformatics.amnh.org/open_source/maxent/, last accessed February 2018.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123:** 585–595.

Tapper SL, **Byrne M**, **Yates CJ**, **Keppel G**, **Hopper SD**, **van Niel K**, **Schut AGT**, **Mucina L**, **Wardell-Johnson GW. 2014a.** Isolated with persistence or dynamically connected? Genetic patterns in a common granite outcrop endemic. *Diversity and Distributions* **20:** 987–1001.

Tapper SL, **Byrne M**, **Yates CJ**, **Keppel G**, **Hopper SD**, **van Niel K**, **Schut AGT**, **Mucina L**, **Wardell-Johnson GW. 2014b.** Prolonged isolation and persistence of a common endemic on granite outcrops in both mesic and semi-arid environments in south-western Australia. *Journal of Biogeography* **41:** 2032–2044.

Vekemans X, **Beauwens T**, **Lemaire M**, **Roldán-Ruiz I. 2002.** Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology* **11:** 139–151.

Wu CI. 2001. The genic view of the process of speciation. *Journal of Evolutionary Biology* **14:** 851–865.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Supplementary Figure S1. Geneland Bayesian inference of spatial clustering. A, Number of clusters along the chain after burnin. B, Spatial distribution of genetic-geographic groups.
Supplementary Figure S2. Model selection and model checking for scenario 2.
Supplementary Figure S3. Parameter priors and subsequent estimates based on the most likely scenario, colonization from east to west (scenario 5).
Supplementary Figure S4. SDM of *V. auriculata* for the study area, training of Maxent algorithm (best performance between the algorithms tested) for current and past scenarios.
Supplementary Figure S5. Validation between algorithms training for *V. auriculata* for the study area.
Supplementary Figure S6. Detail of result of validation of SDM between algorithms, for the occurrence area of *V. auriculata*.
Supplementary Table S1. Primers used in the ISSR amplifications of *V. auriculata* and their respective polymorphism rates.
Supplementary Table S2. Parameters priors for all scenarios and estimates for Scenario 5 (scenario with highest posterior probability on the ABC analysis), based on cpDNA of *Vellozia auriculata*. Ne: population size after colonization, Nb: population size during a bottleneck (scenario 1) or colonization population size (founding effect, scenarios 2–5), t1–t6, divergence times; Dist, prior distribution; uni, uniform; Min, minimum prior value; Max, maximum prior value; Con, constraints and Q, quantile.
Supplementary Table S3. Genetic distance between *V. auriculata* populations. Number of mean substitutions for cpDNA (below diagonal) and unbiased distance of Nei (1978) from ISSR phenotypes (above diagonal).

Downloaded from https://academic.oup.com/botlinnean/article-abstract/191/3/381/5579407 by guest on 28 October 2019

FINAL CONSIDERATIONS

The association between the high diversity of the campos rupestres, an old Neotropical sky-islands environment on eastern South America, and its disjunct aspect have long been stated. Here we confirm, refine and renew this idea using two groups presenting contrasting biologic characteristic. *Bulbophyllum* sect. *Didactyle* (Orchidaceae) evolved trough a complex pattern of isolation, gene flow and hybridization, and some of its previously recognized species are not recovered as monophyletic. The distribution of the genetic diversity of *Bulbophyllum* sect. *Didactyle* mirrors the geographic distributions of the populations and the main geographical breaks of the Espinhaço/Chapada-Diamantina chain, while a new link among Chapada dos Veadeiros, Serra do Cabral, and Serra da Canastra (regions R04, R06 and R09, respectively, on Chapter 1) was observed.

On a species level, both *B. involutum* and *Vellozia auriculata* (Velloziaceae) presented geographical structure, despite this structure was stronger for the less vagile *V. auriculata*. For both species the overall gene flow estimates between populations was low, despite for *B. involutum* they were high across the Espinhaço range. The high gene flow between Espinhaço range *B. involutum* populations corroborates the notion that the small and light seeds of orchids are able of long-range dispersion, while the fact that eastern and western *B. involutum* populations in the Chapada Diamantina are highly differentiated supports that the environment or the past demography of this species might be important factors driving population differentiation. Since many of campo rupestres species present limited vagility, as occurs to *V. auriculata*, it is expected that gene flow between their populations are generally limited, in accordance with the topography-driven-isolation theory, but it is likely that environmental factor also play a role in this differentiation, as probably occurs in *B. exaltatum*.

Beyond the importance of disjunction and environmental variability for population differentiation, this study shows that hybridization might be and critical engine for the origin and maintenance of the campos rupestres biodiversity. Hybridization is a widespread phenomenon on *Bulbophyllum* sect. *Didactyle*, as at least five from its seven currently circumscribed species are presently involved on interspecific crosses. As the phylogenetic propensity of hybridization is high, it is likely that hybridization is also an important event on other sections of *Bulbophyllum*, influencing the high diversity of this highly speciose genus. Further phylogeographical, niche modeling and landscape genetics work might help to improve the disentanglement of the complex patterns related to the evolution of biodiversity on the campos rupestres. Also, the genetic of hybridization might be explored in deep, using

functional gene annotation and trait-based studies, to shed light on the role adaptative introgression and connections between admixture with reproductive barriers generating the exuberant diversity of orchids and the campos rupestres.