Original Article

Hybridization boosters diversification in a Neotropical orchid group

Cecilia F. Fiorini [1,2,*], Eric de Camargo Smidt [3], L. Lacey Knowles [4], Eduardo Leite Borba [2]

[1] Departamento de Genética, Ecologia e Evolução, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, 31270-901, Brazil

[2] Departamento de Botânica, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, 31270-901, Brazil

[3] Departamento de Botânica, Universidade Federal do Paraná, Curitiba, Paraná, 81531-990, Brazil

[4] Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan, 48109-1085, USA

Running title: Hybridization boosters Neotropical orchid diversification

E-mail address of the corresponding author: ceciliafiorini@gmail.com

Abstract

Background and Aims Genetic data shows that cryptic hybrids are more common than previously thought and that hybridization and introgression are widespread processes. Regardless, studies on hybridization are scarce for the highly speciose *Bulbophyllum*. The genus presents more than 2,200 species and many examples of recent radiations, in which hybridization is expected to be frequent. Currently, only four natural *Bulbophyllum* hybrids are recognized, all of them recently described based on morphological evidence. Here we test whether genomic evidence supports the hybrid status of two Neotropical *Bulbophyllum* species, while also evaluating the impact of this phenomenon on the genomes of the putative parental species. We also assess if there is evidence of hybridization among *B. involutum* and *B. exaltatum*, sister species that diverged recently.

Methods We leverage the power of next-generation sequence data, associated with model-based analysis for three systems putatively constituted by two parental species and one hybrid. All taxa belong to the Neotropical *B.* sect. *Didactyle* clade.

Key Results We found evidence of hybridization in all studied systems. Despite the occurrence of hybridization, there are no signs of backcrossing.

Conclusions Because of the high propensity of hybridization across many taxa, the common occurrence of hybridization during the evolutionary history of *B.* sect. *Didactyle* means it is time to account for and examine its evolutionary role in these orchids.

Keywords: Hybridization, *B.* sect. *Didactyle*, Neotropics, orchids, diversity

Introduction

Hybridization is defined as the outcrossing and gene flow between populations that differ in multiple heritable characters that affect fitness (Gompert and Buerkle 2016). It was already considered an evolutionary dead end and a destructive force with little evolutionary consequences (Sætre 2013; Seehausen 2013). However, given renewed evidence, hybridization is now seen as a creative force in the evolution of plants and animals (Mallet 2007; Abbott *et al.* 2013; Seehausen 2013). 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 *et al.* 2010). Thus, both hybridization and introgressive hybridization (introgression, i.e., 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 and Larson 2014).

Hybridization can introduce alleles that had already been "tested" and act as a powerful source of adaptative variation (Arnold and Martin 2009; Whitney *et al.* 2010; Suarez-Gonzalez *et al.* 2018; Burgarella *et al.* 2019). 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 in earlier works (e.g., Key 1968; Bazykin 1969). While hybridization can slow or reverse differentiation, it may also lead to speciation by adaptative introgression (homoploid hybrid speciation) or cause fast speciation via allopolyploidization (Abbott *et al.* 2013). Hybrid speciation is defined as "a speciation event in which hybridization has played a crucial role in the evolution of reproductive barriers between a hybrid lineage and its parent lineages" and many examples of natural homoploid hybrid speciation and allopolyploidization have been described (Taylor and Larson 2019).

One of the main predictors of the chance of hybridization between two taxa is their divergence age (Paun *et al.* 2011; Abbott *et al.* 2013). Low divergence is unlikely to bring major novelties; however, as lineages diverge Dobzhansky-Muller incompatibilities (negative epistatic interactions between alleles from independent evolutionary backgrounds) increase, possibly preventing the success of hybrids individuals (Scopece *et al.* 2007; Levin 2012). As incompatibilities are subject to 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 (Seehausen 2004; Gourbière and Mallet 2010), complicating phylogenetic inference (Payseur and Rieseberg 2016; Gates *et al.* 2021, preprint). 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). In the Neotropical region hybridization studies are scarce, but suggest a possible role for hybridization in the rapid diversification of its flora (Schley *et al.* 2022).

Despite being common, hybridization is not universal with evidence of a 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 orchid hybrids are known (Yam and Arditti 2009). 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).

Regardless of the evidence suggesting hybridization might be common in orchid groups, it has not been considered one of the main drivers of diversification in *Bulbophyllum*, one of the largest genera in the family, including ~ 2,200 species (Pridgeon *et al.* 2014). Despite its late Paleogene origin (~ 25 million years ago), *Bulbophyllum* presents many examples of recent radiations (Gamisch and Comes 2019). However, only four natural *Bulbophyllum* hybrids are currently recognized – *B.* ×*chikukwa* (Africa), *B.* ×*cipoense* (South America), *B.* ×*guartelae* (South America), and *B.* ×*omerumbellatum* (Asia) – which were all described based on morphological evidence (Borba and Semir 1998a; Fibeck and Mavi 2000; Mancinelli and Smidt 2012; Lin 2022).

Among the *Bulbophyllum* hybrids described for South America, both *B.* ×*cipoense* and *B.* ×*guartelae* are putatively hybrids between species of the *B.* sect. *Didactyle*. It has been suggested that only *B. weddellii* is a pollen receptor in the formation of *B.* ×*cipoense*, since *B. weddellii*'s pollinarium size is not compatible with *B. involutum*'s stigmatic cavity (Borba and Semir 1998a; b, 1999). However, morphology indicates that introgression occurs in the opposite direction, with *B. involutum* as a pollen receptor, since there is a range of intermediate *B. involutum* forms in multiple populations (Azevedo *et al.* 2006). The hybrid origin of *B.* ×*cipoense* was tested with allozymes but there was no conclusive support for this hypothesis, probably due to marker resolution (Azevedo *et al.* 2006). Only one individual of *B.* ×*guartelae* was found in the wild, however, its existence suggests gene flow or

introgression between the parental species *B. perii* and *B. tripetalum* (Mancinelli and Smidt 2012). So far, no genetic test was performed to test the hybrid origin of *B.* ×*guartelae*.

The *Didactyle* section includes also the *B. exaltatum* species complex and hybridization between the taxa *B. exaltatum* and *B. involutum* has been suggested due to the continuum of morphological variation among them. These species are interfertile, as demonstrated by experimental pollinations (Borba *et al.* 1999) and, despite some specificity in pollination systems, pollinator sharing eventually occurs (Borba and Semir 1998b). The polytopic origin of natural hybrids and introgression among lineages may be one of the factors responsible for the intricate morphological pattern of *B.* sect. *Didactyle*, especially in the *B. exaltatum* complex (Azevedo *et al.* 2006; Ribeiro *et al.* 2008).

As ancestral polymorphism, mutations and selection against intermediate characters can interfere with hybrid phenotype, and detection of hybrids is not always obvious (Rieseberg 1995; Mallet 2005; Leal *et al.* 2016; Pace and Cameron 2019). The advent of next-generation sequencing and genomic data sets allows more rigorous tests of hybridization (Twyford and Ennos 2012; Goulet *et al.* 2017). 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 and 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*, system WIC), "*B. perii* and *B. tripetalum*" (*B.* ×*guartelae*, system TPG), and "*B. exaltatum* and *B. involutum*" (system IE)? (ii) If so, may these events relate to the complex morphological patterns observed in this group? (iii) Hybridization and introgression in system IE are more widespread than in system WIC, as expected due to the difference in divergence age? (iv) On sympatric localities, is it possible to find both parental and hybrid individuals?

Materials and Methods

Sampling

To study the systems *B. weddellii*/*B. involutum*/*B.* ×*cipoense* (system WIC), *B. tripetalum*/*B. perii*/*B.* ×*guartelae* (system TPG), and *B. involutum*/*B. exaltatum* (system 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). Individuals were identified based on their morphology. We collected individuals growing on different rocks and a minimum of 10 m apart, to prevent sampling vegetative clones or closely related individuals (Hedrén and Lorenz 2019). All samples were collected under issued permits to CFF and ELB (SISBIO 52995-1, IEF 062/2016, and IAP 51.16) and voucher information can be found in Table 1.

Genomic library preparation and processing

We extracted Genomic DNA from fresh leaves (Doyle and 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, USA) and PCR-amplified these fragments using high-fidelity DNA polymerase (iProof, Bio-Rad, Hercules, CA, USA), with 8 or 12 cycles. We sequenced individuals in four lanes of an Illumina HiSeq 2500 (Illumina, San Diego, CA, USA) on Rapid Run Mode at The Centre for Applied Genomics of Hospital for Sick Children (Toronto, ON, Canada) to generate 150 bp single-end reads, in combination with samples from other projects.

We processed genomic data using the Stacks 2.3e pipeline (Rochette and Catchen 2017). We assembled de novo demultiplexed and filtered sequences with ustacks, build a catalog 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, $n = 6$ (based on the *r80 loci* plateau, Supplementary Fig. 1 [**Supplementary Information**]; Rochette and Catchen 2017), 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 and Knowles 2016). To guard against sequencing and assembly errors, we used a custom R script (Thomaz *et al.* 2017) to exclude SNPs with theta values within the upper 95% quantile of variability (Supplementary Fig. 2 [**Supplementary Information**]). For each system, we used the software plink 1.9 (Purcell *et al.* 2007) to identify SNPs with a maximum of 25% (datasets D25) or 40% missing data (datasets D40), because the robustness of analyses to missing data differs. The sequencing throughput for each of the systems is shown in Supplementary Table 1 [**Supplementary Information**]. For analyses sensitive to potential linkage disequilibrium, we built for each system a dataset with a single randomly retained SNP per locus and a maximum of 25% missing data (datasets D25U).

Genetic differentiation and hybridization

For each of the systems of putative hybrids, we generated a principal component analysis (PCA) to visualize the distribution of genomic variation using adegenet 2.1.1 (Jombart and Ahmed 2011), in R 3.5.0 (R Core Team 2019). As a multivariate method, PCA summarizes the genetic similarity among populations and genotypes without requiring strong assumptions about Hardy–Weinberg equilibrium or linkage disequilibrium. Due to its sensibility to missing data, we used the dataset D25 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) to estimate the hybrid-index (i.e., the proportion of allele copies coming from one of 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 used the dataset D25U and 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 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 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 parallelnewhybrid 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. It 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. For NewHybrids we used the same loci sets obtained by gghybrids.

To estimate population structure for each of the systems, we used fastSTRUCTURE 1.0, a variational Bayesian framework compatible with large data sets (Raj *et al.* 2014). We used the datasets D25 and 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 *et al.* 2015).

As the IE system is expected to have diverged recently, we used HyDe 0.4.1a to infer introgression despite incomplete lineage sorting (Blischak *et al.* 2018). 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 and the dataset D40. *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 the 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). However, the analysis supports 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* in 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 intermediate values of ancestry proportion and hybrid index, but slightly closer to *B. involutum* (Fig. 2C and D). Both analyses support that all the other individuals belong to pure lineages, in agreement with NewHybrids results. Yet, NewHybrids indicates that *B.* ×*cipoense* are F2 hybrids (Fig. 2E).

TPG system

Like the WIC system, 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 is actually the second record 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 showed similar results, with *B.* ×*guartelae* having intermediate values of ancestry proportion and hybrid index (~0.5). Both analyses support that all the other individuals belong to pure lineages, in agreement with the NewHybrids results. Yet, NewHybrids also indicates that *B.* ×*guartelae* are F2 hybrids (Fig. 3E).

IE system

As with systems WIC and TPG, system IE shows signs of hybridization. However, IE individuals with hybrid genomic composition are widespread across 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 in an intermediate position (Fig. 4B). The second axis mainly segregates *B. exaltatum* populations. As a general pattern, fastSTRUCTURE and gghybrids indicate that the lower the latitude (and closer the distance to the center of *B. involutum*'s distribution), the higher the proportion of *B. involutum* genome in *B. exaltatum* individuals (Fig. 4A, C, and D). HyDe results show 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 conspecific individuals (Fig. 4E). NewHybrids suggests that the individuals with hybrid

ancestry are F2 hybrids, with a 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 in 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. Because hybridization shows high phylogenetic propensity (Whitney *et al.* 2010), it suggests that hybridization might be a common process in the evolution of *Bulbophyllum* as a whole, a hypothesis that might be better explored in the future using species from the whole *Bulbophyllum* distribution.

Hybridization in B. sect. Didactyle

The initial divergence between *B.* sect. *Didactyle* species occurred 2.16 million years ago (Gamisch and Comes 2019), but at least five of the seven currently circumscribed taxa are involved in hybridization at some level. Indeed, it has been previously shown that *B. weddellii*, *B. involutum*, and *B. exaltatum* are interfertile (Borba *et al.* 1999). Hybrid individuals are more frequent in populations of system IE, in which parentals are very closely related and floral morphology is quite similar as compared to the two other systems. In the system IE, differences in floral volatile compounds act to attract different pollinators (Silva *et al.* 1999). Although Borba and Semir (1998b) observed the occurrence of visits by pollinators of *B. exaltatum* (as *B. ipanemense*) to the flowers of *B. involutum* when they are cultivated in sympatry, the smaller size of these insects did not result in the pollination of the slightly larger flowers of the latter species. However, it seems to be clear that these barriers are not enough to maintain the integrity of the boundaries of these species when they occur in sympatry. Indeed, some IE populations are apparently completely composed of F2

individuals (i.e., E08, E10, and E11). Meanwhile, *B.* ×*cipoense* (systems WIC) and *B.* ×*guartelae* (system TPG) are apparently rare (Borba and Semir 1998a; Mancinelli and Smidt 2012). In systems WIC and TPG we find no backcrossed individuals and the formation of hybrids seems to have little effect on the fate of the parental populations, suggesting the divergence of the hybrid's flowering morphology can lead to the inefficiency of its reproductive mechanisms (Borba *et al.* 1998).

Our study does not support the idea that the morphological variation observed in *B. involutum* is a result of hybridization with *B. weddellii*, as suggested by Azevedo *et al.* (2006). *B. involutum* individuals are mainly of pure genomic makeup, as in *B. weddellii*, *B. perii*, and *B. tripetalum*. In contrast, a portion of the individuals identified as *B. exaltatum* contain some degree of *B. involutum* genome. Thus, part of the morphological obscurity in the *B. exaltatum* species complex can be viewed as a result of the presence of individuals of mixed ancestry.

It is important to highlight the geographic 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 and Larson 2014). Our results support dispersed hybridization in the IE system, as *B. involutum* genes are present in *B. exaltatum* populations outside the area sympatry. However, no population from system IE could be considered a hybrid zone, as none of them presented parental species accompanied by multiple generations of hybrids. There is evidence that individuals with mixed ancestry may form a new hybrid species, as no backcrossing was observed (Fig. 4F). It is not clear, however, how hybridization might have 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 and Larson 2019) and that "shared variation among populations may reflect unsorted shared ancestral polymorphism" (Payseur and 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 significant results for all individuals (Blischak *et al.* 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 and 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 a lineage with hybrid origin.

NewHybrids classified individuals with mixed ancestry mainly as F2 hybrids (Fig. 2E, 3E and 4F). We did not observe F1 or introgressed individuals, suggesting that the formation of F1 hybrids or backcrossed individuals are rare events. However, the occurrence of incomplete lineage sorting or of an 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 indeed F2 hybrids and that *B.* ×*cipoense* (systems WIC) and *B.* ×*guartelae* (system TPG) are reproductively isolated from their respective parental species.

It is noteworthy that *B. involutum* is considerably more abundant than *B. exaltatum* in sympatry (pers. obs.). This can possibly impact hybridization outcomes, given the relevance of demographic factors to this process (Currat *et al.* 2008; Klein *et al.* 2017). The asymmetry 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 *et al.* 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. In fact, populations that are on isolated outcrops can lead to limited gene flow and the rise of differentiation and local adaptation.

Hybridization and the diversification of Bulbophyllum species

Hybridization propensity presents strong and consistent phylogenetic signal across floras, suggesting that it might be an intrinsic propriety of biological 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 abundant hybrids in *B.* sect. *Didactyle* is an indication that it might be a frequent phenomenon in *Bulbophyllum* species in general, given the abundance of recent radiated sections (Gamisch and Comes 2019). Indeed, it has been suggested 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 exhibit slow growth, with long expected generation times. Still, our understanding of the factors driving orchid hybridization is scarce and a better knowledge of the factors driving reproductive barriers is required. It is noteworthy to emphasize that molecular investigations are important in identifying future *Bulbophyllum* hybrids and in orchids in general. As morphological characters are the result of the interplay of many genes and can be plastic (Rieseberg and Ellstrand 1993), morphological intermediaries can be absent or misleading (e.g., Wallace 2006; de Hert *et al.* 2011; Leal *et al.* 2016; Pace and Cameron 2019).

The study of New World orchid hybridization is in development (e.g. Borba *et al.* 1999; Pinheiro *et al.* 2010; Pinheiro and Cozzolino 2016; Sujii *et al.* 2019; Leal *et al.* 2020).

Despite our knowledge of genome evolution, there is still much to discover about how the genome changes after hybridization. Questions about the origin and maintenance of reproductive barriers are also still open, such as the amount of differentiation between genomic regions during speciation and how these regions are dispersed around the genome (Abbott *et al.* 2013). As species that currently hybridize may offer exceptional insights into the genomics of hybridization, a more in depth study of the hybridization process within *B.* sect. *Didactyle*, especially of system IE, may be key to a better understanding of the speciose genus *Bulbophyllum*.

Conclusion

Here we confirm the occurrence of hybridization on systems *B. weddellii*/*B. involutum* (*B.* ×*cipoense*, WIC), *B. tripetalum*/*B. perii* (*B.* ×*guartelae*, TPG), and *B. involutum*/*B. exaltatum* (IE), species of *B.* sect. *Didactyle*. Consistent with the expectation that species with more recent common ancestry will be more interfertile (Levin 2012), hybridization is more geographically and genetically widespread in the system IE than in the systems WIC or TPG, which are more distantly related. 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 the system IE suggests that the formation of hybrids can be an important factor for adaptative divergence, and consequently divergence of *B. exaltatum*. Future research will shed light on adaptative introgression (e.g., functional gene annotation) and connections between admixture with reproductive barriers (e.g., trait-based studies) in the IE system. 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 *B.* sect. *Didactyle* may be an indication that this process is also common across

other sections of the genus, which is a hypothesis to be explored. If so, hybridization may play an important role in the diversification of the *Bulbophyllum* and its recent radiation.

Funding

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico [430515/2016-2 to E.L.B.]. We also thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior for the Doctoral, PDSE, and CAPES-PRINT fellowships [88882.184384/2018-0, 88881.188909/2018-01, 88887.683121/2022-00 to C.F.F., respectively].

Supplementary Information

Supplementary information includes the graphs used to select assembly parameters for stacks analysis, a histogram of genetic variability of loci, and the sequencing throughput information.

Acknowledgments

We thank K. L. Rodrigues, G. L. Fiorini, G. C. Lustre, M. O. Duarte, J. G. Rodrigues, L. S. B. Jordão, M. A. dos Santos; R. C. da Mota; F. I. Rosa; I. N. Vischi; M. Lucca; W. S. Mancinelli for helping in the sampling. Additionally, we would like to thank the Sagarana HPC cluster (ICB, UFMG) and the College of Literature, Science, and the Arts (UMich) for access to bioinformatics resources.

Literature cited

Abbott R, Albach D, Ansell S, *et al.* 2013. Hybridization and speciation. *Journal of Evolutionary Biology* **26**: 229–246.

Anderson EC, Thompson EA. **2002**. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* **160**: 1217–1229.

Arnold ML. **1997**. *Natural hybridization and evolution*. New York: Oxford University Press.

Arnold ML, Martin NH. **2009**. Adaptation by introgression. *Journal of Biology* **8**: 82.

Azevedo CO, Borba EL, 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**: 415–421.

Bailey RI. **2018**. gghybrid: Evolutionary Analysis of Hybrids and Hybrid Zones.

Bazykin AD. **1969**. Hypothetical mechanism of speciation. *Evolution* **23**: 685–687.

Blischak PD, Chifman J, Wolfe AD, Kubatko LS. **2018**. HyDe: a python package for genome-scale hybridization detection (D Posada, Ed.). *Systematic Biology* **67**: 821–829.

Borba EL, Semir J. **1998a**. *Bulbophyllum* ×*cipoense* (Orchidaceae), a new natural hybrid from the Brazilian "campos rupestres": Description and biology. *Lindleyana* **13**: 113–120.

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

Borba EL, 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**: 197–204.

Borba EL, Semir J, Barros F de. **1998**. *Bulbophyllum involutum* Borba, Semir & F. Barros (Orchidaceae), a new species from the Brazilian "campos rupestres." *Novon* **8**: 225–229.

Borba EL, Shepherd GJ, 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**: 205–214.

Buerkle CA. **2005**. Maximum-likelihood estimation of a hybrid index based on molecular

markers. *Molecular Ecology Notes* **5**: 684–687.

Burgarella C, Barnaud A, Kane NA, *et al.* 2019. Adaptive introgression: an untapped evolutionary mechanism for crop adaptation. *Frontiers in Plant Science* **10**: 1–17.

Currat M, Ruedi M, Petit RJ, Excoffier L. **2008**. The hidden side of invasions: massive introgression by local genes. *Evolution* **62**: 1908–1920.

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

Fibeck W, Mavi S. **2000**. A natural orchid hybrid from Zimbabwe. *Kirkia* **17**: 147–149.

Folk RA, Soltis PS, Soltis DE, Guralnick R. **2018**. New prospects in the detection and comparative analysis of hybridization in the tree of life. *American Journal of Botany* **105**: 364–375.

Gamisch A, Comes HP. **2019**. Clade-age-dependent diversification under high species turnover shapes species richness disparities among tropical rainforest lineages of *Bulbophyllum* (Orchidaceae). *BMC Evolutionary Biology* **19**: 1–16.

Gates DJ, Pilson D, Smith SD. **2021**. Inferring the history of hybridization: a case study in *Iochrominae* (Solanaceae). *EcoEvoRxiv*.

Gompert Z, Buerkle CA. **2016**. What, if anything, are hybrids: enduring truths and challenges associated with population structure and gene flow. *Evolutionary Applications* **9**: 909–923.

Gottlieb LD. **1984**. Genetics and morphological evolution in plants. *The American Naturalist* **123**: 681–709.

Goulet BE, Roda F, Hopkins R. **2017**. Hybridization in plants: old ideas, new techniques. *Plant Physiology* **173**: 65–78.

Gourbière S, Mallet J. **2010**. Are species real? The shape of the species boundary with exponential failure, reinforcement, and the “missing snowball.” *Evolution* **64**: 1–24.

Harrison RG, Larson EL. **2014**. Hybridization, introgression, and the nature of species boundaries. *Journal of Heredity* **105**: 795–809.

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**: 83–100.

Hert K, Jacquemyn H, Van Glabeke S, *et al.* 2011. Patterns of hybridization between diploid and derived allotetraploid species of *Dactylorhiza* (Orchidaceae) co-occurring in Belgium. *American Journal of Botany* **98**: 946–955.

Huang H, Lacey Knowles L. **2016**. Unforeseen consequences of excluding missing data from next-generation sequences: Simulation study of rad sequences. *Systematic Biology* **65**: 357–365.

Johnson SD. **2018**. Natural hybridization in the orchid flora of South Africa: comparisons among genera and floristic regions. *South African Journal of Botany* **118**: 290–298.

Jombart T, Ahmed I. **2011**. adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**: 3070–3071.

Key KHL. **1968**. The concept of stasipatric speciation. *Systematic Biology* **17**: 14–22.

Klein EK, Lagache-Navarro L, Petit RJ. **2017**. Demographic and spatial determinants of hybridization rate (M Rees, Ed.). *Journal of Ecology* **105**: 29–38.

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. **2015**. CLUMPAK: A program for identifying clustering modes and packaging population structure inferences across *K*. *Molecular Ecology Resources* **15**: 1179–1191.

Leal BSS, Brandão MM, Palma-Silva C, Pinheiro F. **2020**. Differential gene expression reveals mechanisms related to habitat divergence between hybridizing orchids from the Neotropical coastal plains. *BMC Plant Biology* **20**: 1–14.

Leal BSS, Chaves CJN, Koehler S, Borba EL. **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**: 621–639.

Levin DA. **2012**. The long wait for hybrid sterility in flowering plants. *New Phytologist* **196**: 666–670.

Lin TP. **2022**. *Bulbophyllum* ×*omerumbellatum*, a natural hybrid of *B. umbellatum* and *B. omerandrum*. *Taiwania* **67**: 461–464.

Mallet J. **2005**. Hybridization as an invasion of the genome. *Trends in Ecology and Evolution* **20**: 229–237.

Mallet J. **2007**. Hybrid speciation. *Nature* **446**: 279–283.

Mancinelli WS, Smidt E de C. **2012**. O gênero *Bulbophyllum* (Orchidaceae) na Região Sul do Brasil. *Rodriguésia* **63**: 803–815.

Pace MC, Cameron KM. **2019**. The evolutionary and systematic significance of hybridization between taxa of *Spiranthes* (Orchidaceae) in the California Sierra Nevada and Cascade Range. *Taxon* **68**: 199–217.

Parchman TL, Gompert Z, Mudge J, Schilkey FD, Benkman CW, Buerkle CA. **2012**. Genome-wide association genetics of an adaptive trait in lodgepole pine. *Molecular Ecology* **21**: 2991–3005.

Paun O, Forest F, Fay MF, Chase MW. **2011**. Parental divergence and hybrid speciation in angiosperms revisited. *Taxon* **60**: 1241–1244.

Payseur BA, Rieseberg LH. **2016**. A genomic perspective on hybridization and speciation. *Molecular ecology* **25**: 2337–2360.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. **2012**. Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE* **7**: e37135.

Pinheiro F, De Barros F, Palma-Silva C, *et al.* 2010. Hybridization and introgression across

different ploidy levels in the Neotropical orchids *Epidendrum fulgens* and *E. puniceoluteum* (Orchidaceae). *Molecular Ecology* **19**: 3981–3994.

Pinheiro F, Cozzolino S. **2016**. *Epidendrum* (Orchidaceae) as a model system for ecological and evolutionary studies in the Neotropics. *Taxon* **62**: 77–88.

Pridgeon AM, Cribb PJ, Chase MW, Rasmussen FN. **2014**. *Genera Orchidacearum Volume 6: Epidendroideae (Part three)*. Oxford: Oxford University Press.

Purcell S, Neale B, Todd-Brown K, *et al.* 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics* **81**: 559–575.

R Core Team. **2019**. R: A Language and Environment for Statistical Computing.

Raj A, Stephens M, Pritchard JK. **2014**. fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* **197**: 573–589.

Ribeiro PL, Borba EL, De Camargo Smidt E, Lambert SM, Schnadelbach AS, 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**: 109–137.

Rieseberg LH. **1995**. The role of hybrization in evolution: old wine in new skins. *American Journal of Botany* **82**: 944–953.

Rieseberg LH, Ellstrand NC. **1993**. What can molecular and morphological markers tell us about plant hybridization? *Critical Reviews in Plant Sciences* **12**: 213–241.

Rochette NC, Catchen JM. **2017**. Deriving genotypes from RAD-seq short-read data using Stacks. *Nature Protocols* **12**: 2640–2659.

Sætre GP. **2013**. Hybridization is important in evolution, but is speciation? *Journal of Evolutionary Biology* **26**: 256–258.

Schley RJ, Twyford AD, Pennington RT. **2022**. Hybridization: a “double-edged sword” for

Neotropical plant diversity. *Botanical Journal of the Linnean Society* **199**: 331–356.

Scopece G, Musacchio A, Widmer A, Cozzolino S. **2007**. Patterns of reproductive isolation in Mediterranean deceptive orchids. *Evolution* **61**: 2623–2642.

Seehausen O. **2004**. Hybridization and adaptive radiation. *Trends in Ecology and Evolution* **19**: 198–207.

Seehausen O. **2013**. Conditions when hybridization might predispose populations for adaptive radiation. *Journal of Evolutionary Biology* **26**: 279–281.

Silva UF, L. Borba E, Semir J, Marsaioli AJ. **1999**. A simple solid injection device for the analyses of *Bulbophyllum* (Orchidaceae) volatiles. *Phytochemistry* **50**: 31–34.

Suarez-Gonzalez A, Lexer C, Cronk QCB. **2018**. Adaptive introgression: a plant perspective. *Biology Letters* **14**: 20170688.

Sujii PS, Cozzolino S, Pinheiro F. **2019**. Hybridization and geographic distribution shapes the spatial genetic structure of two co-occurring orchid species. *Heredity* **123**: 458–469.

Taylor SA, Larson EL. **2019**. Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nature Ecology and Evolution* **3**: 170–177.

Thomaz AT, Malabarba LR, Knowles LL. **2017**. Genomic signatures of paleodrainages in a freshwater fish along the southeastern coast of Brazil: Genetic structure reflects past riverine properties. *Heredity* **119**: 287–294.

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

Wallace LE. **2006**. Spatial genetic structure and frequency of interspecific hybridization in *Platanthera aquilonis* and *P. dilatata* (Orchidaceae) occurring in sympatry. *American Journal of Botany* **93**: 1001–1009.

Whitney KD, Ahern JR, Campbell LG, Albert LP, King MS. **2010**. Patterns of

hybridization in plants. *Perspectives in Plant Ecology, Evolution and Systematics* **12**: 175–182.

Wringe BF, Stanley RRE, Jeffery NW, Anderson EC, Bradbury IR. **2017**. parallelnewhybrid: an R package for the parallelization of hybrid detection using newhybrids. *Molecular Ecology Resources* **17**: 91–95.

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

Yam TW, Arditti J. **2009**. History of orchid propagation: a mirror of the history of biotechnology. *Plant Biotechnology Reports* **3**: 1–56.

Table 1. Information about *Bulbophyllum* sect. *Didactyle* populations analysed 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

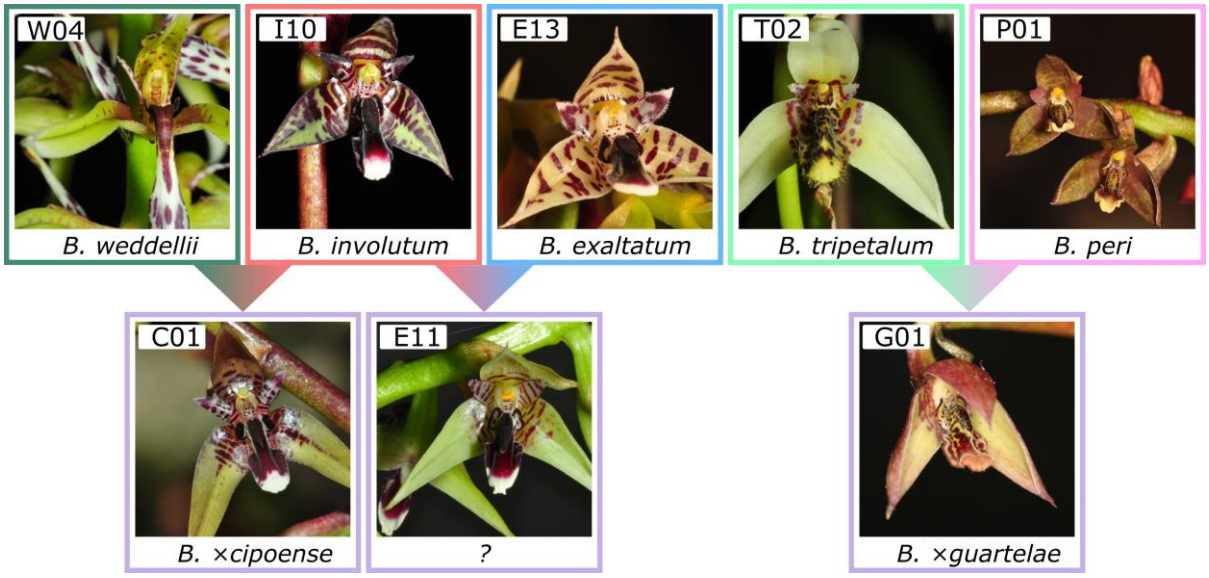

Figure 1. Morphological variability of *Bulbophyllum* sect. *Didactyle* hybrid systems. Populations names are given within pictures. Species colours are consistently used trough the paper. For populations information see Table 1.

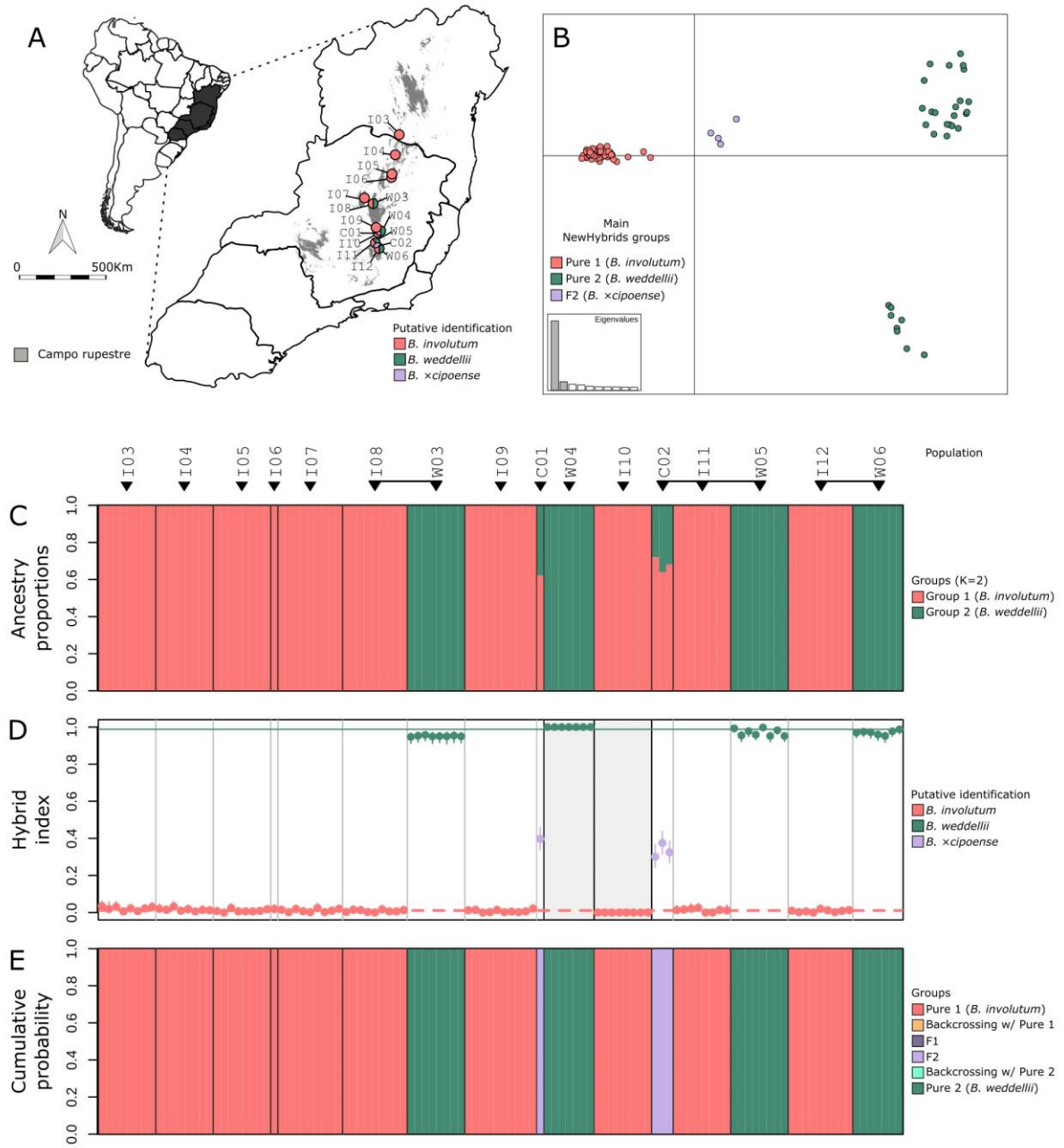

Figure 2. Hybridization in system WIC (*B. weddellii*/*B. involutum*/*B. ×cipoense*). (A) geographic distribution of populations; (B) PCA results; (C) fastSTRUCTURE results for *K* = 2; (D) gghybrid results; (E) NewHybrids results. For populations information see Table 1.

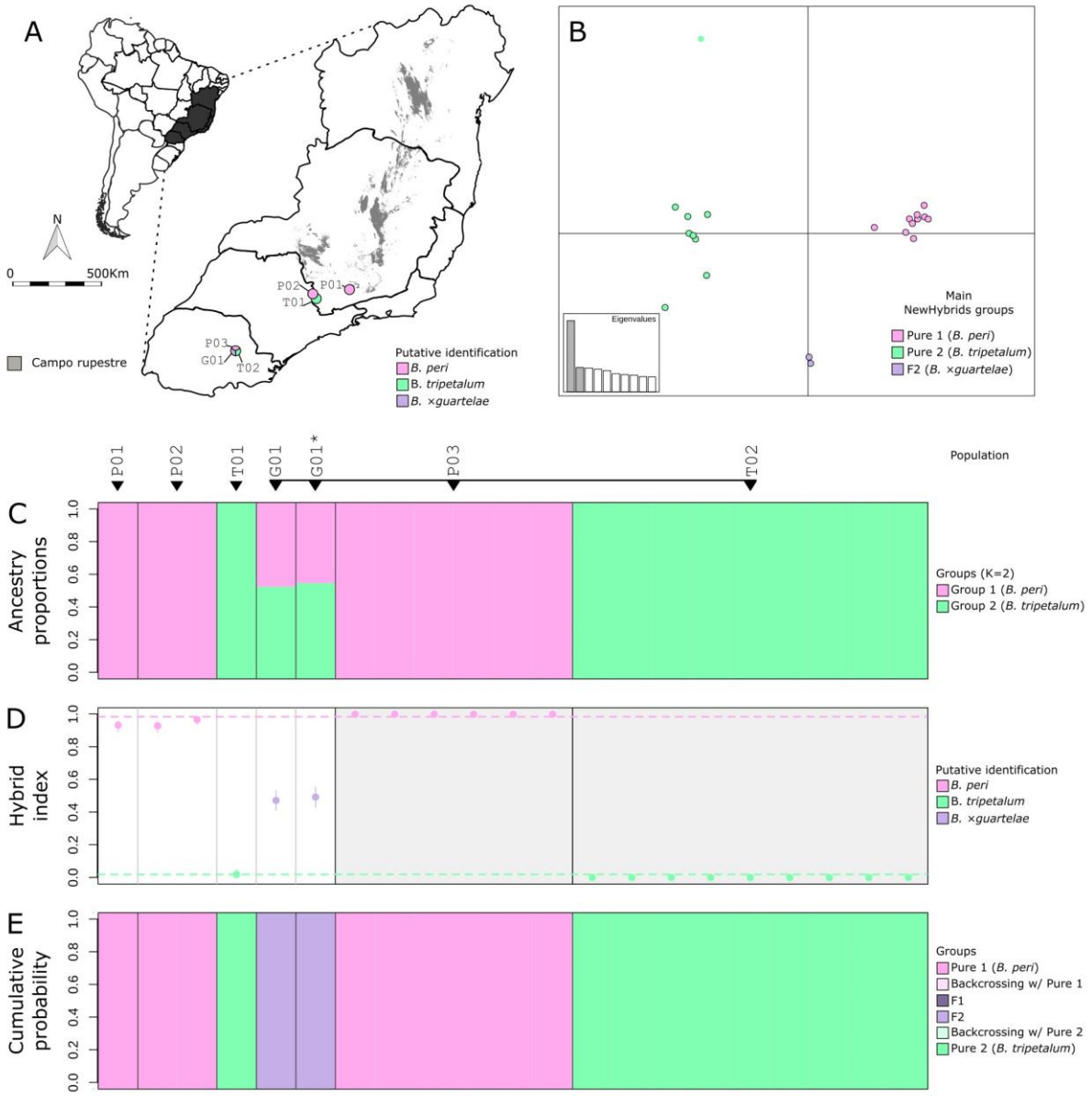

Figure 3. Hybridization in system TPG (*B. tripetalum*/*B. perii*/*B.* ×*guartelae*). (A) geographic distribution of populations; (B) PCA results; (C) fastSTRUCTURE results for *K* = 2; (D) gghybrid results; (E) NewHybrids results. For populations information see Table 1.

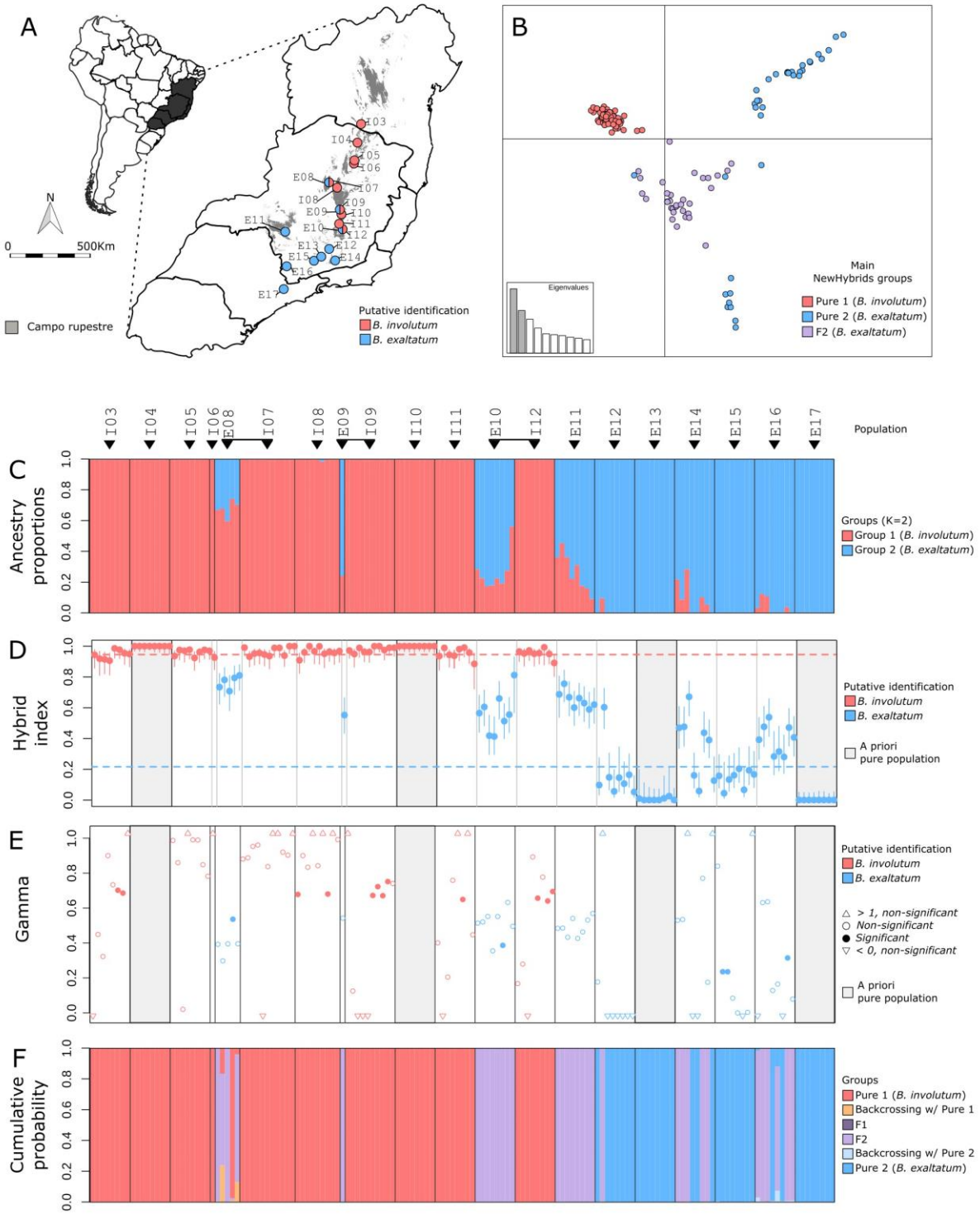

Figure 4. Hybridization in system IE (*B. involutum*/*B. exaltatum*). (A) geographic distribution of populations; (B) PCA results; (C) fastSTRUCTURE results for K = 2; (D) gghybrid results; (E) HyDe results; (F) NewHybrids results. For populations information see Table 1.