Climate cooling promoted the expansion and radiation of a threatened group of South American orchids (Epidendroideae: Laeliinae)

ALEXANDRE ANTONELLI[1]*, CHRISTIANO F. VEROLA[2], CHRISTIAN PARISOD[3] and A. LOVISA S. GUSTAFSSON[3]

[1]*Institute of Systematic Botany, Zollikerstrasse 107, CH 8008 Zurich, Switzerland*
[2]*Universidade Federal de Mato Grosso, Instituto de Biociências, Departamento de Botânica e Ecologia, Cuiabá, Avenue. Fernando Corrêa da Costa s/n, Cuiabá, CEP 78060-900, Mato Grosso, Brazil*
[3]*National Centre for Biosystematics, National History Museum, University of Oslo, PO Box 1172 Blindern, 0318 Oslo, Norway*

The Brazilian Cerrado is the most species-rich tropical savanna in the world. Within this biome, the Campos Rupestres ('rocky savannas') constitute a poorly studied and highly threatened ecosystem. To better understand how plants characteristic of this vegetation have evolved and come to occupy the now widely-separated patches of rocky formations in eastern Brazil, we reconstruct the biogeographical history of the rare orchid genus *Hoffmannseggella*. We apply parsimony and Bayesian methods to infer the phylogenetic relationships among 40 out of the 41 described species. Absolute divergence times are calculated under penalized likelihood and compared with estimates from a Bayesian relaxed clock. Ancestral ranges are inferred for all nodes of the phylogeny using Fitch optimization and statistical dispersal vicariance analysis. In all analyses, phylogenetic uncertainty is taken into account by the independent analysis of a large tree sample. The results obtained indicate that *Hoffmannseggella* underwent rapid radiation around the Middle/Late Miocene (approximately 11–14 Mya). The region corresponding today to southern Minas Gerais acted as a main source area for several independent range expansions north- and eastwards via episodic corridors. These results provide independent evidence that climate cooling following the Middle Miocene Climatic Optimum (approximately 15 Mya) led to important vegetational shifts in eastern Brazil, causing an increase in the dominance of open versus closed habitats. Polyploidy following secondary contact of previously isolated populations may have been responsible for the formation of many species, as demonstrated by the high ploidy levels reported in the genus. © 2010 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2010, **100**, 597–607.

ADDITIONAL KEYWORDS: biogeography – Cerrado – climate changes – molecular dating – Neotropics – Orchidaceae – secondary contact model.

INTRODUCTION

Orchidaceae is the largest family of flowering plants, comprising some 25 000 species, and occupying virtually all terrestrial ecosystems (Dressler, 1993). Until recently, one of the main obstacles in estimating the age of orchid groups has been the complete absence of reliable fossils in the family, which are necessary for calibrating phylogenetic trees. This situation was recently improved by the discovery of three remarkable macrofossils: an extinct bee associated with pollinia from the tribe Goodyerinae, preserved in Dominican amber approximately 15–20 Mya (Ramirez *et al.*, 2007); and fossil leaves from New Zealand assigned to genera *Earina* and *Dendrobium*

*Corresponding author.
Current address: Gothenburg Botanical Garden, Carl Skottsbergs gata 22A, 41319, Göteborg, Sweden.
E-mail: alexandre.antonelli@vgregion.se

(20–23 Mya; Conran, Bannister & Lee, 2009). The Dominican fossils have been used to perform the first fossil-calibrated molecular dating analysis focused on Orchidaceae (Ramirez *et al.*, 2007), which was further improved with the inclusion of the New Zealand fossils and the use of a Bayesian relaxed clock (Gustafsson *et al.*, 2010).

The present study attempts, for the first time, to combine molecular dating techniques with ancestral range reconstruction in a group of Neotropical orchids (genus *Hoffmannseggella*), for which we have previously sampled and sequenced 40 out of the 41 described species. By assessing when and where the group diversified, we aim to provide a plausible hypothesis for its geographical expansion and radiation.

The study group, Epidendroideae, is the largest of the five orchid subfamilies currently recognized (Chase, 2005). Within the subfamily, subtribe Laeliinae Benth. comprises 43 genera and approximately 1470 species of epiphytic or rupicolous herbs, distributed along tropical America (Dressler, 1993). On the basis of results obtained from phylogenetic studies (van den Berg *et al.*, 2000), the species of *Laelia* Lindl. were all segregated into two genera: the 'true' *Laelia*, encompassing only Central American and Mexican species, and *Sophronitis* Lindl. comprising all Brazilian species (van Den Berg & Chase, 2000). Subsequently, the Brazilian species were revised once again and split into four genera, based on differences in morphology, habitat and ecological requirements: *Hadrolaelia* Chiron & V.P. Castro, *Hoffmannseggella*, *Microlaelia* Chiron & V.P. Castro and *Dungsia* Chiron & Castro Neto, 2002). The present study focuses on one of the most ornamental groups of Laeliinae, the genus *Hoffmannseggella* (Fig. 1).

Hoffmannseggella comprises 41 species (Verola & Semir, 2007) and is endemic to Southeastern Brazil. By contrast to the closely related *Dungsia* (which comprises five epiphytic species restricted to the Brazilian state of Espírito Santo), all *Hoffmannseggella* species are strictly rupicolous (growing on rocks; Fig. 1E). They are typical elements of a vegetation type denominated 'Campos Rupestres' (Verola *et al.*, 2007). This is an open and rocky savanna-like habitat within the Cerrado biome (Fig. 1A) shown to contain a high diversity level, especially among epiphytes (Alves, Kolbek & Becker, 2008). The state of Minas Gerais is the centre of diversity of plants characteristic for this vegetation type (Harley, 1995). At the same time, this is a fast-developing region where anthropogenic activity poses serious threats to the endemic biota. A few species of the genus occur in mid-elevation remnants of the Atlantic rain forest, which is one of the world's hotspots of biodiversity (Myers, 2000). Subsequent to the colonization of

Brazil by Europeans in the 16th Century, this ecosystem has undergone massive destruction. New estimates indicate that no more than 11–16% of the original forest cover remains, of which 32–40% is constituted by secondary forests and small fragments (Ribeiro *et al.*, 2009).

The fact that 47% of all *Hoffmannseggella* species are only known from a single population, in many cases in the immediate vicinity of expanding industrial complexes and outside legally protected areas (Verola *et al.*, 2007), probably makes this genus one of the most endangered plant taxa in Brazil. The extinction threats faced by the species urge the need for studies that highlight the importance of these plants with respect to understanding the evolution of the exceedingly rich Brazilian flora.

MATERIAL AND METHODS

CHOICE OF TAXA AND SEQUENCE REGIONS

In revent years, we have been able to sample and sequence virtually all species in genus *Hoffmannseggella* (40 out of 41). To correctly root the genus in a phylogeny and include suitable calibration points for the molecular dating analysis, we included representative species of 11 additional genera (*Bifrenaria*, *Cattleya*, *Cymbidium*, *Dungsia*, *Gongora*, *Lycaste*, *Masdevallia*, *Maxillaria*, *Oncidium*, *Stanhopea*, and *Zygopetalum*). We did not opt to include a larger taxon sampling outside the *Hoffmannseggella* clade because: (1) the higher level phylogeny of subfamily Epidendroideae has already been thoroughly investigated by van den Berg *et al*. (2000, 2005), and major clades were identified with high phylogenetic support; (2) we primarily aimed to investigate fine-scale biogeographical patterns within *Hoffmannseggella*, rather than attempting to infer patterns on a continental scale; and (3) we have carried out extensive field work with *Hoffmannseggella* and are confident on the geographical distribution and morphological delimitation of all species in the genus.

In a phylogenetic study of subfamily Epidendroideae, van Den Berg *et al*. (2005) found that the internal transcribed spacer (ITS) (nuclear ribosomal DNA) contained the highest proportion of potentially parsimony informative sites (64%), compared to the *trn*L-F region (28%), the *mat*K gene (28%) and spacers (28%), and the *rbc*L gene (11%). We have therefore chosen to build our molecular matrix and perform our analyses on ITS 1 and ITS 2, intercalated by the 5.8S coding region. The addition of more rapidly evolving sequence regions, such as the external transcribed spacer, would be desirable, but would require the development of new primers specific to this group (A. L. S. Gustafsson, unpubl. data). Although several intrinsic
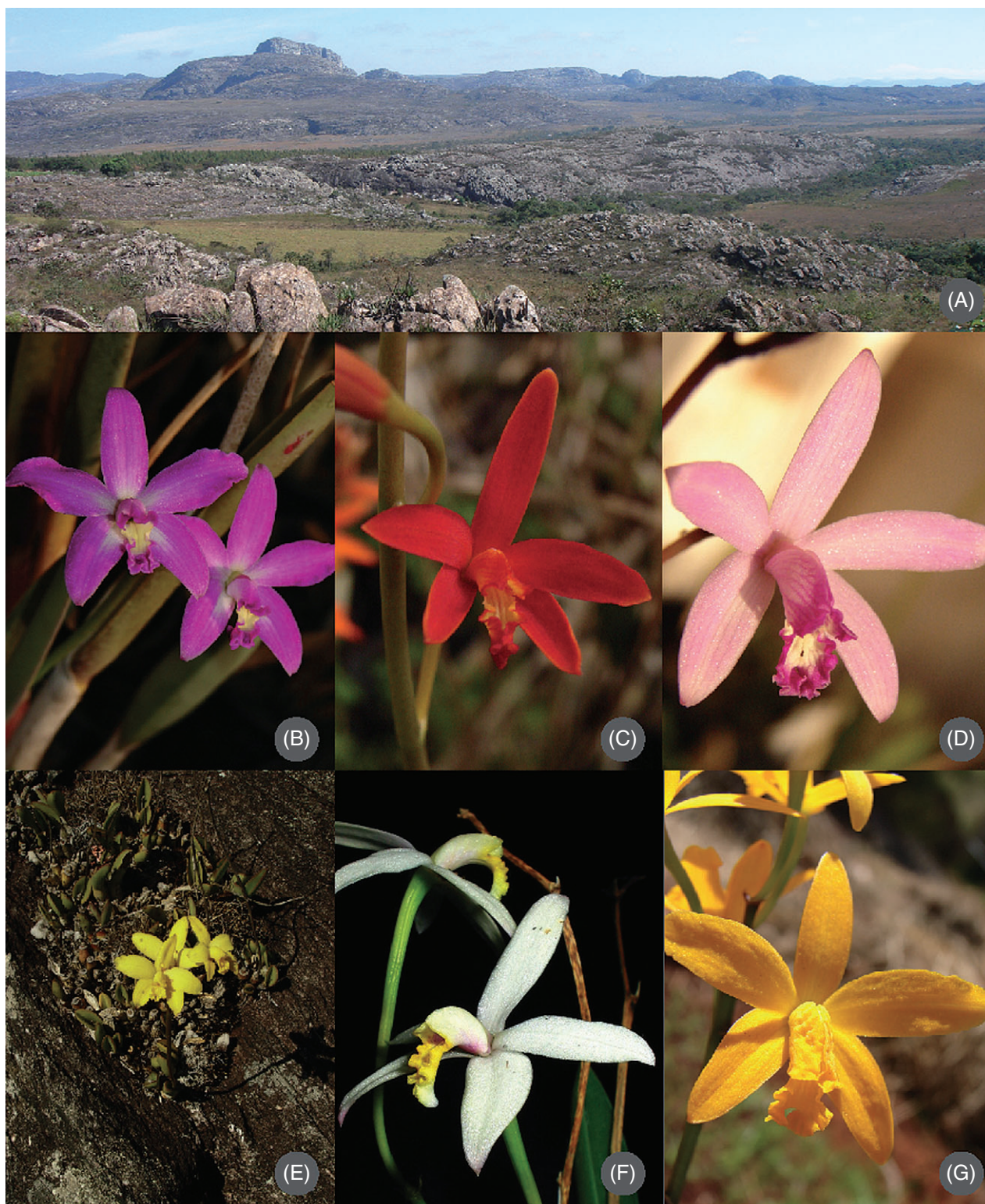
Figure 1. Floral variation among Brazilian *Hoffmannseggella* and their natural habitat. A, typical view of the Campos Rupestres vegetation near Diamantina, Minas Gerais. This is a rocky savanna-like habitat characterized by high levels of species richness and endemism, today under severe anthropogenic threat. B, *Hoffmannseggella rupestris*. C, *Hoffmannseggella angereri*. D, *Hoffmannseggella caulescens*. E, *Hoffmannseggella bradei*. F, *Hoffmannseggella fournieri*. G, *Hoffmannseggella bahiensis*. Credits: (A) A.L.S.G.; (B, C, D, E, F, G) C.F.V.

features of ITS (compared to low or single copy genes) have been shown to potentially affect phylogenetic reconstruction and thereby influence divergence times for certain taxonomic groups (Alvarez & Wendel, 2003), ITS is among the fastest evolving sequence regions in the genome (Hughes, Eastwood & Bailey, 2006) and thus has a high potential of resolving relationships at the genus level and below. The dataset used in the present study is the same as that used in Gustafsson *et al.* (2010), where GenBank accession numbers and voucher information, sequencing and alignment procedures are provided.

PHYLOGENETIC AND DATING ANALYSIS

To test congruency of results from different methods of phylogenetic inference, we used both parsimony implemented in PAUP*, version 4.0b10 (Swofford, 1999) and Bayesian inference implemented in MrBayes, version 3.1 (Huelsenbeck & Ronquist, 2001). A heuristic search was performed in PAUP running 10 000 replicates, using tree-bisection-reconnection (TBR) branch-swapping, holding one tree at each step during stepwise addition and using the Multiple Trees option. Parsimony support values for nodes were estimated by means of bootstrapping, running 10 000 replicates, applying TBR branch swapping and saving one tree per replicate. We used the Akaike information criterion implemented in MRMODELTEST, version 2.2 (Nylander, 2004) to select which evolutionary model best fitted our data, which was then incorporated into a MrBayes block in the input file. We performed five Markov chain Monte Carlo (MCMC) analyses with 10 million generations each, sampling every 1000th generation, saving branch lengths and employing one cold and three heated chains. We saved branch lengths and used the other default settings. These analyses were performed at the Computational Biology Service Unit hosted by Cornell University (http://cbsuapps.tc.cornell.edu). The performance of the separate analyses (convergence of parameters, effective sample sizes) was assessed with the on-line software AWTY (Nylander *et al.*, 2008b) and the software TRACER, version 1.5 (Drummond & Rambaut, 2007).

Gustafsson *et al.* (2010) applied the Bayesian relaxed clock approach implemented in BEAST (Drummond & Rambaut, 2007) to estimate divergence times for a high level phylogeny of Orchidaceae, which was then used to date the *Hoffmannseggella* dataset. In the present study, we estimated divergence times by applying frequentist statistics (as opposed to Bayesian), under the penalized likelihood algorithm implemented in the software R8S, version 1.71 (Sanderson, 2002). We used the truncated Newton method and ran a cross-validation procedure to select the optimal smoothing value for our data, with the settings maxiter = 2000, num_restarts = 5, num_time_guesses = 5, penalty = add, checkgradient = yes, and increasing the log increment of the penalty function by 0.1 (Sanderson, 2002). To obtain absolute ages, we calibrated the tree using the age of the 'higher Epidendroids' estimated by Ramirez *et al.* (2007). Their youngest mean age estimate using penalized likelihood was used to constrain the split between *Oncidium* and *Cattleya* in our phylogeny at 45 Mya. To estimate the influence of phylogenetic uncertainty on age estimations, 5000 trees randomly chosen from the post burn-in Bayesian sample were individually dated and age statistics (mean, 95% confidence intervals) calculated for all nodes present in the 50% majority-rule consensus phylogram of the MrBayes analysis.

SPECIES DISTRIBUTIONS AND OPERATIONAL AREAS

To obtain distribution maps for the species analyzed, a database of species georeferences was created through a survey in Brazilian herbaria (UEC, ESA, SPF, SP, BHCB, HXBH, ESAL, MBML, R, RB, HB). Although these georeferences varied in terms of their precision depending on the source, this potential bias was not relevant for the biogeographical analyses we conducted. In addition, data points were created for field observations done by one of us (C.F.V.). We paid special care to exclude dubiously identified records.

The total distribution of the Brazilian genera surveyed in the present study was obtained by plotting all georeferences on topographical maps using the software DIVA GIS, version 5.2 (http://www.diva-gis.org/). This enabled the identification of five main areas of distribution, which are well delimited geographically and geologically (Fig. 2, inset): Chapada Diamantina in the state of Bahia (BA); northern Minas Gerais (N-MG); southern Minas Gerais (S-MG); Rio de Janeiro (RJ); and Espírito Santo (ES). The species occurring in areas BA and N-MG occur almost exclusively on soils derived from Mesoproterozoic quartzite rocks, S-MG on soils from Archeozoic rocks, and ES and RJ on Palaeoproterozoic granitic (gneiss) rocks (Fig. 2, inset; data from the US Geological Survey's EROS Data Center and Instituto Brasileiro de Geografia e Estatística). The extant distribution of all Brazilian species was coded as belonging to one or more of these five operational areas, and extra-Brazilian species were coded as occurring in Central America, the Andes, or both.

BIOGEOGRAPHICAL ANALYSIS

To reconstruct the historical biogeography of the lineages investigated, we applied two methods for

inferring the ranges of all ancestral nodes on the phylogeny: Fitch parsimony as implemented in MES-QUITE, version 2.72 (Maddison & Maddison, 2009) and statistical dispersal-vicariance analysis implemented in S-DIVA (Nylander *et al.*, 2008a; Yu, Harris & He, 2010). Although both are parsimony-based, they differ in the number of areas that an ancestral node is allowed to occupy: in Fitch parsimony, it can only be a single area (thus favouring dispersal over vicariance), whereas, in S-DIVA, a minimum of two areas need to be specified a priori (thus favouring vicariance as null biogeographical model). Because this is an arbitrary decision, two separate S-DIVA analyses were run constraining the maximum number of ancestral areas to 2 (the minimum) and 7 (the maximum). To take into account the influence of phylogenetic uncertainty on biogeographical inference, in all analyses, ancestral areas were optimized independently on the 5000 post burn-in Bayesian trees used for the dating analysis. Relative frequencies of ancestral areas for each node of a reference tree were then computed by counting all trees with a certain area uniquely assigned to the node. These results were then translated into pie charts, which included the proportion of ambiguous reconstructions and node absence across the tree sample for the MESQUITE analyses, but not for the S-DIVA analysis, because this could not be readily computed. For MESQUITE, we chose to plot the results on the 50% majority-rule consensus tree of the Bayesian analysis, whereas, for S-DIVA (which is not capable of plotting the results on a tree containing polytomies), we used the maximum a *posteriori* tree from the Bayesian analysis. We acknowledge that these differences make the results from MESQUITE and S-DIVA not directly comparable, although this appears to be irrelevant because we are more interested in evaluating the results obtained from different methods, rather than performing a formal methodological comparison.

RESULTS

The final matrix comprised 660 characters, of which 296 were variable and 155 were parsimony-informative. The heuristic search in PAUP found 6095 most parsimonious trees, with a tree length of 532 steps, consistency index = 0.73 and retention index = 0.80 (rescaled consistency index 0.59). The Akaike information criterion implemented in MrModelTest selected the GTR+Γ evolutionary model. All independent MCMC analyses converged and yielded effective sample sizes > 8000. Twenty-five percent of the trees in each analysis were excluded (corresponding to a burn-in at 2.5 million generations), and the remaining 37 500 trees were used to compute a 50% majority-rule consensus tree.

The Supporting information (Fig. S1) shows the 50% majority-rule consensus phylogram of the Bayesian analysis. Figure 2 shows the topology of 50% majority-rule consensus tree of the Bayesian analysis, with branch lengths proportional to mean ages as estimated in the dating analysis, as well as Bayesian posterior probabilities and parsimony bootstrap values. There was no strongly supported conflict between the results generated by the parsimony and Bayesian analyses. Pie charts at nodes indicate the relative frequencies of ancestral area optimizations across the entire tree sample, obtained from Fitch parsimony in MESQUITE. The biogeographical reconstructions carried out with S-DIVA are shown in the Supporting information (Fig. S2). The mean number of dispersals (or range expansions, two possibilities that are difficult to discern if ancestral lineages in reality occupied more than a single area) estimated in MESQUITE between the operational areas occupied by *Hoffmannseggella* are summarized in Table 1.

The overall resolution of the tree is poor, although there are several well-supported clades. Genus *Hoffmannseggella* constitutes a monophyletic group, with strong Bayesian and bootstrap support (Bayesian posterior probability = 0.99, Bootstrap support = 93). *Dungsia kautskyi* (Pabst) Chiron & V.P. Castro is sister to *Hoffmannseggella*. Confidence intervals of node ages were relatively large (Fig. 2, green bars), but most lineage splits were estimated to have taken place from the Middle Miocene onwards. The crown group of *Hoffmannseggella* was estimated at 14.2 Mya (9.69–18.6 Mya; 95% confidence intervals). Although the mean age estimates obtained here with penalized likelihood were generally older than those estimated by Gustafsson *et al*. (2010), there was considerable overlap in the 95% confidence intervals/highest posterior densities for all key nodes outlined in Figure 2 (Table 2).

DISCUSSION

RANGE EVOLUTION OF *HOFFMANNSEGGELLA*

The results obtained in the preset study (Fig. 2) appear to corroborate the hypothesis proposed by van den Berg *et al*. (2000) suggesting that the Laeliinae originated in Central America and later colonized South America. However, this could be an artefact of taxon sampling because there are several unsampled species in tribe Laelinae from other areas than southeastern Brazil (e.g. *Cattleya* species from Venezuela and Amazonia), and their inclusion could influence the optimizations at the base of the phylogenetic tree.

Both the Fitch (Fig. 2) and S-DIVA (see Supporting information, Fig. S2) analyses infer the most recent common ancestor (MRCA) of *Hoffmannseggella* (C2 in
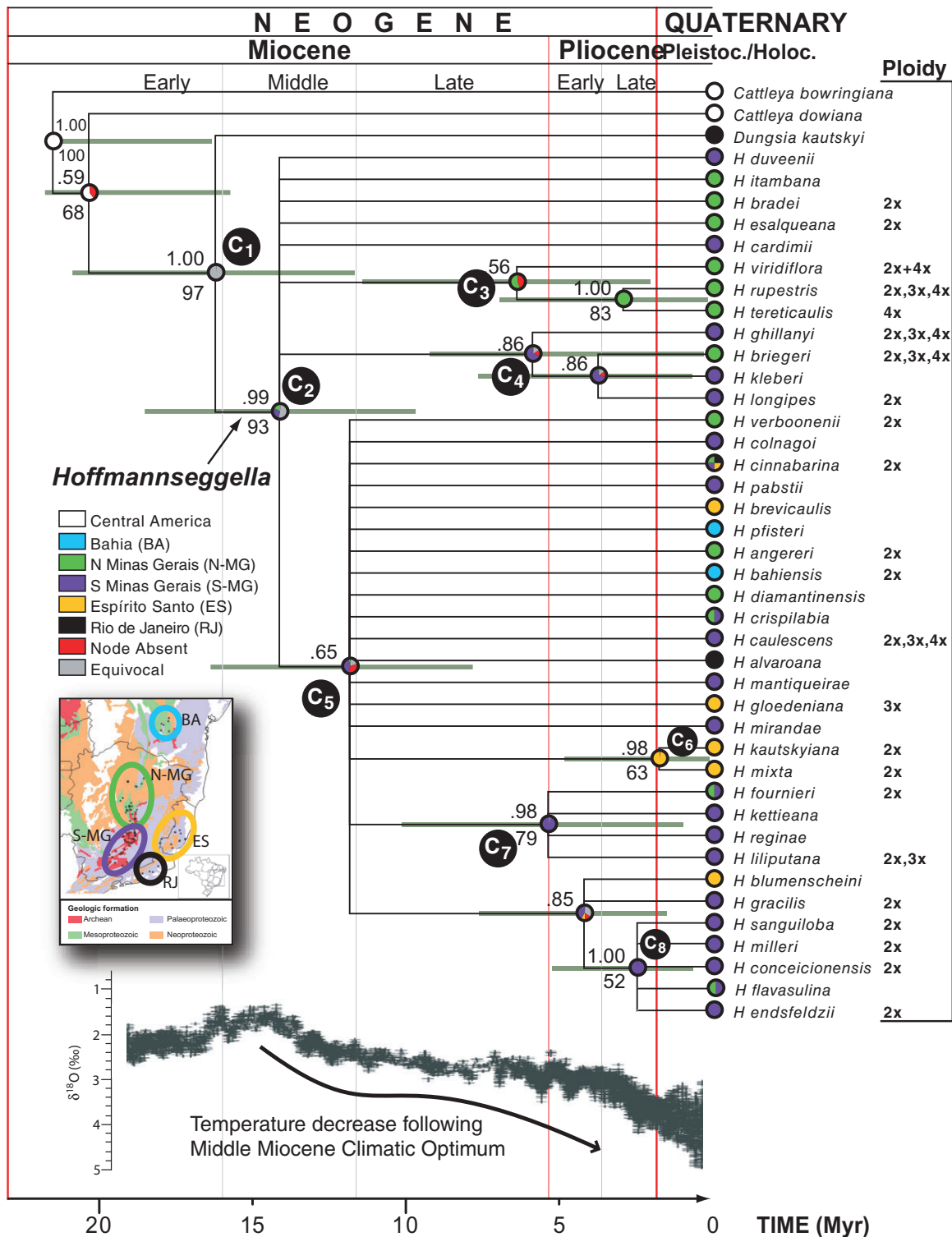
Fig. 2) to Minas Gerais. This node recurs in 4950 out of the 5000 randomly selected trees, and in approximately half of the Fitch optimizations (2651 trees) its reconstruction is ambiguous, whereas, in the other half

of the optimizations, its range is inferred to southern (1421 trees) or northern (877 trees) Minas Gerais. Most descendant nodes outlined in Figure 2 (C3–C5, C7–C8) continue to be most frequently inferred to

Figure 2. Combined results from the phylogenetic, dating and biogeographical analysis. A, 50% majority-rule consensus from the Bayesian analysis, with branch lengths proportional to mean ages as estimated from 5000 independently-dated chronograms (outgroup excluded). Numbers above branches indicate Bayesian posterior probabilities; numbers below branches show parsimony bootstrap support. Green bars at selected nodes represent 95% confidence intervals of ages (bars extending further back in time than 23 Mya have been truncated). Pie charts at nodes represent relative frequencies of ancestral-area reconstructions obtained from 5000 optimizations using Fitch parsimony; current species distributions are shown preceding the species names (for colour coding, see legend). Ploidy levels are relative to the basic somatic chromosome number, so that one ploidy level $(x) = 2n = 20$ chromosomes (Yamagishi-Costa & Forni-Martins, 2009; C. F. Verola, unpubl. data). The oxygen–isotope curve is given as a proxy for temperature (Zachos *et al.*, 2008). The timescale is taken from Gradstein & Ogg (2004). Inset: operational areas used in the biogeographical analysis. Each dot on the map represents a population of one or more species of *Hoffmannseggella*.

Table 1. Mean number of dispersals between the biogeographical areas defined for Brazil, estimated from 5000 post burn-in Bayesian trees under Fitch parsimony as implemented in MESQUITE

		To				
From	Bahia	Northern Minas Gerais	Southern Minas Gerais	Espírito Santo	Rio de Janeiro	Totals
Bahia	–	0.32	0.14	0.35	0.01	0.82
Northern Minas Gerais	0.36	–	2.36	0.60	0.17	3.49
Southern Minas Gerais	0.73	5.60	–	2.49	1.36	10.2
Espírito Santo	0.50	0.53	0.51	–	0.01	1.55
Rio de Janeiro	0.01	0.19	0.23	0.02	–	0.45
Totals	1.60	6.64	3.24	3.46	1.55	16.5

Table 2. Comparison between divergence times estimated using penalized likelihood in r8s (present study) with ages estimated under a Bayesian relaxed clock (Gustafsson *et al.*, 2010)

Clade in Fig. 2	Penalized likelihood			BEAST		
	Mean	Lower 95% CI	Upper 95% CI	Median	Lower 95% HPD	Upper 95% HPD
C1	16.3	11.7	21.0	13.7	6.03	24.1
C2	14.2	9.69	18.6	11.2	4.53	19.9
C3	6.44	1.93	11.4	3.10	0.43	7.85
C4	5.94	1.81	10.9	3.27	0.55	7.87
C5	11.9	7.80	16.5	8.56	3.54	16.1
C6	1.83	0.00	4.78	0.43	0.00	2.51
C7	5.42	0.85	10.1	2.00	0.07	6.34
C8	2.54	0.53	5.17	2.76	0.51	6.84

Ages are given in million of years. CI, confidence interval; HPD, highest posterior density.

Minas Gerais in both optimization methods used (Fig. 2; see also Supporting information, Fig. S2).

Southern Minas Gerais was the most important source area for *Hoffmannseggella* (mean dispersals/range expansions per tree = 10.2; Table 1). Most range expansions from there reached Northern Minas Gerais (mean 5.60), but some reached Espírito Santo, Rio de Janeiro, and Bahia, in decreasing order (Table 1). Northern Minas Gerais received most lin-eages (mean 6.64), followed by Espírito Santo, Southern Minas Gerais, Bahia, and Rio de Janeiro. Virtually no biotic exchange was inferred between Bahia and Rio de Janeiro, which are two non-adjacent areas. However, geographical proximity does not appear to fully explain the inferred number of range expansions. For example, almost no dispersals were inferred between Espírito Santo and Rio de Janeiro (two adjacent areas), and more dispersals were

inferred to Bahia from Southern Minas Gerais than from Northern Minas Gerais, despite the longer distance. Accordingly, range expansions of *Hoffmannseggella* does not appear to have occurred via long-distance dispersal (e.g. seeds being carried by wind from one area to another). Rather, range expansions could have occurred gradually via biotic corridors linking those areas, as suggested for several other plant and animal taxa from high altitude areas in southeastern Brazil (Safford 2007).

Following the Mid-Miocene climatic optimum at approximately 15 Mya, a period of global cooling began that led to drastic temperature fluctuations and the formation of Northern ice sheets from at least 11 Mya (Zachos, Dickens & Zeebe, 2008). Vegetational changes have been suggested to affect the Cerrado vegetation, by an overall range expansion during dry periods, and subsequent contraction during wetter periods (Alves & Kolbek, 1994; Pirani *et al*., 1994; Behling, 2002; Safford, 2007). The Campos Rupestres are currently confined to altitudes above 1000 m. Under this limit, shrub and tree-rich ecosystems become dominant. Assuming that, during drier periods, this limit would have decreased by 200 m (Alves & Kolbek, 1994; Pirani *et al*., 1994; Safford, 1999, 2007; Behling, 2002), we attempted to approximately determine the past distribution of the Campos Rupestres. This enabled the identification of possible migration corridors between the currently isolated 'islands' of Campos Rupestres vegetation, which correspond to the operational areas used in the biogeographical analyses. According to the dispersal events summarized in Table 1, and translated into arrows of different widths, we are then able to present a biogeographical scenario for the range expansion of *Hoffmannseggella*. This reconstruction is shown in Figure 3, where red patches indicate the present distribution of Campos Rupestres, and blue patches show areas in the range 800–1000 m, where this vegetation type presumably expanded.

CORRELATES OF SPECIATION

A direct comparison between the biogeographical results in Figure 2 and the past vegetational reconstruction shown in Figure 3 enables a general interpretation of how climatic changes may have promoted the expansion and radiation of *Hoffmannseggella*. As temperatures started to drop approximately 15 Mya (Zachos *et al*., 2008; see temperature curve at the bottom of Fig. 2) and the extend of Campos Rupestres increased in southeastern Brazil, *Hoffmannseggella* populations from mainly southern Minas Gerais were able to expand their range north- and eastwards. Although these were the two prevalent directions of range expan-

sions, the results of the present study suggest that a lower rate of dispersals have also taken place in other directions (Fig. 3, Table 1).

Assuming that range expansion following vegetational changes occurred mainly through 'diffusion' (i.e. leptokurtic dispersal resulting in a slow gain of new territories), allopatric speciation may have played a minor role in *Hoffmannseggella*. The past connectivity of areas appears to better explain the range expansions in the genus (Fig. 3) and, accordingly, hybridization and polyploidy represent likely speciation mechanisms. If populations of a common ancestral species were separated by forested areas during wetter (warmer) periods, they could have accumulated divergent chromosomal arrangements. By the time these populations came again into contact during the expansion of Campos Rupestres in drier (cooler) periods, admixture would have resulted in unfit hybrids and genome doubling (i.e. polyploidy) would have produced new lineages with restored fertility (Stebbins, 1984). Whether this secondary contact hypothesis caused the radiation of *Hoffmannseggella* in Campos Rupestres still remains speculative, although such a scenario has been documented for several polyploid lineages from regions having undergone drastic climate changes during the Pleistocene (Brochmann *et al*., 2004; Guggisberg *et al*., 2006; Suda *et al*., 2007; Parisod, Holderegger & Brochmann, 2010). Secondary contacts probably also explain the origin of polyploid lineages in the orchid genus *Dactylorhiza* (Pedersen, 2006; Hedrén *et al*., 2007) and chromosome counts in *Hoffmannseggella* support this hypothesis because almost 60% of all species are polyploid (Yamagishi-Costa & Forni-Martins, 2009; C. F. Verola, unpubl. data). Furthermore, both diploid and polyploid taxa co-occur within most clades (see ploidy counts known so far in Fig. 2). A similar pattern has been reported for the orchid genus *Epidendrum* (Ribeiro *et al*., 2009). Polyploidy is a well-known process accounting for drastic phenotypic changes over short periods of time (Soltis *et al*., 2009) and to what extent polyploidy promoted the rapid radiation of *Hoffmannseggella* remains an issue for further investigation.

The inferred ages for most nodes in *Hoffmannseggella* (Fig. 2, Table 2) can be compared with the finding obtained by Simon *et al*. (2009) indicating that plant clades in the Cerrado diversified within the last 10 Mya, especially within the last 4 Mya, when C4 grasses presumably increased in ecological dominance (Simon *et al*. 2009). Although the Campos Rupestres differ from the typical savanna-like habitats characteristic of the Cerrado, it is logical that an expansion of the Cerrado in southeastern Brazil would also have increased the exposure of rocky formations. The fact that several *Hoffmannseggella* clades pre-date 4 Mya, however, might indicate that open patches of Campos
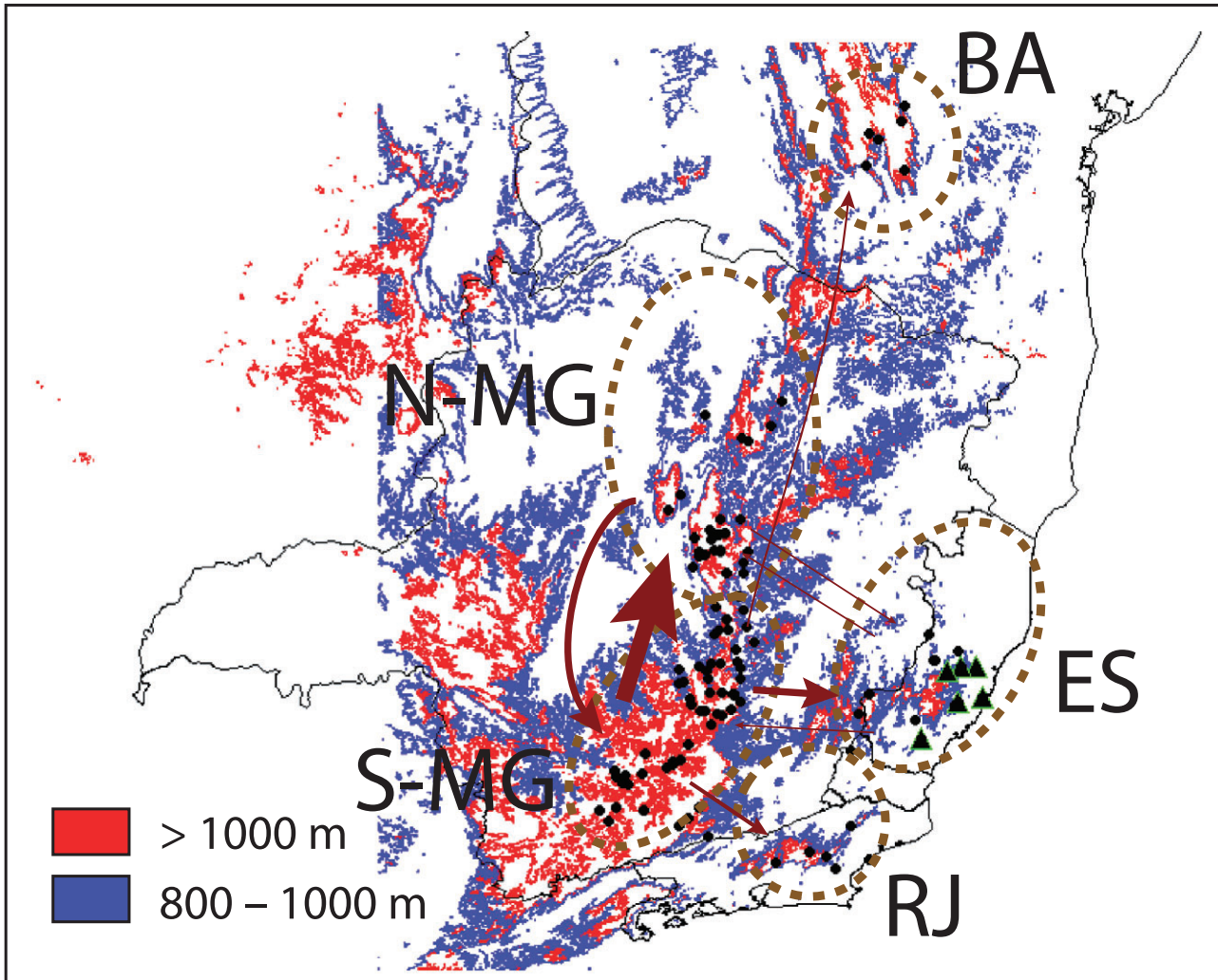
Figure 3. Range expansion of *Hoffmannseggella* in eastern Brazil. The present-day distribution of *Hoffmannseggella* (black dots) and its sister genus *Dungsia* (black triangles) are plotted. Patches in red represent areas higher than 1000 m and are equivalent to the current distribution of rocky savannas (Campos Rupestres). Patches in blue indicate areas in the range 800–1000 m, into which this vegetation type is expected to have expanded during dry (cooler) periods following the Mid-Miocene Climatic Optimum. The expansion of the Campos Rupestres appears to have created biotic corridors for the range expansion of *Hoffmannseggella*. Arrows represent the inferred directions of dispersals, with their width drawn in proportion to the relative frequency of dispersals between areas (Table 1). BA, Chapada Diamantina in the state of Bahia; N-MG, northern Minas Gerais; S-MG, southern Minas Gerais; RJ, Rio de Janeiro; ES, Espírito Santo.

Rupestres may have existed prior to the large-scale expansion of the Cerrado. Clearly, more lineages endemic to the Campos Rupestres need to be investigated to advance our understanding of the historical assembly of this species-rich ecosystem.

REFERENCES

Alvarez I, Wendel JF. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* **29:** 417–434.

Alves RJV, Kolbek J. 1994. Plant species endemism in savanna vegetation on table mountains (Campo Rupestre) in Brazil. *Vegetation* **113:** 125–139.

Alves RJV, Kolbek J, Becker J. 2008. Vascular epiphyte vegetation in rocky savannas of southeastern Brazil. *Nordic Journal of Botany* **26:** 101–117.

Behling H. 2002. South and southeast Brazilian grasslands during Late Quaternary times: a synthesis. *Palaeogeography, Palaeoclimatology, Palaeoecology* **177:** 19–27.

van den Berg C, Chase MW. 2000. Nomenclatural notes on Laeliinae-I. *Lindleyana* **15:** 115–119.

van den Berg C, Goldman DH, Freudenstein JV, Pridgeon AM, Cameron KM, Chase MW. 2005. An overview of the phylogenetic relationships within Epidendroideae inferred from multiple DNA regions and recircumscription of Epidendreae and Arethuseae (Orchidaceae). *American Journal of Botany* **92:** 613–624.

van den Berg C, Higgins WE, Dressler RL, Whitten WM, Arenas MAS, Culham A, Chase MW. 2000. A phylogenetic analysis of Laeliinae (Orchidaceae) based on sequence data from internal transcribed spaces (ITS) of nuclear ribosomal DNA. *Lindleyana* **15:** 96–114.

Brochmann C, Brysting AK, Alsos IG, Borgen L, Grundt HH, Scheen AC, Elven R. 2004. Polyploidy in arctic plants. *Biological Journal of the Linnean Society* **82:** 521–536.

Chase MW. 2005. Classification of Orchidaceae in the age of DNA data. *Curtis's Botanical Magasin* **22:** 2–7.

Chiron GR, Castro Neto VP. 2002. Révision des espèces brésiliènnes du genre Laelia Lindley. *Richardiana* **2:** 4–28.

Conran JG, Bannister JM, Lee DE. 2009. Earliest orchid macrofossils: early Miocene dendrobium and earina (Orchidaceae: Epidendroideae) from New Zealand. *American Journal of Botany* **96:** 466–474.

Dressler RL. 1993. *Phylogeny and classification of the orchid family*. Cambridge: Cambridge University Press.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7:** 8.

Gradstein FM, Ogg JG. 2004. Geologic time scale 2004 – why, how, and where next! *Lethaia* **37:** 175–181.

Guggisberg A, Mansion G, Kelso S, Conti E. 2006. Evolution of biogeographic patterns, ploidy levels, and breeding systems in a diploid–polyploid species complex of Primula. *New Phytologist* **171:** 617–632.

Gustafsson ALS, Verola CF, Antonelli A. 2010. Reassessing the temporal evolution of orchids with new fossils and a Bayesian relaxed clock, with implications for the diversification of the rare South American genus *Hoffmannseggella* (Orchidaceae: Epidendroideae). *BMC Evolutionary Biology*. doi:10.1186/1471-2148-10-177

Harley RM. 1995. *Flora of the Pico das Almas, Chapada Diamantina-Bahia, Brazil*. Kew: Royal Botanic Gardens.

Hedrén M, Norström S, Hovmalm HAP, Pedersen HA, Hansson S. 2007. Patterns of polyploid evolution in Greek marsh orchids (Dactylorhiza; Orchidaceae) as revealed by allozymes, AFLPS, and plastid DNA data. *American Journal of Botany* **94:** 1205–1218.

Huelsenbeck JP, Ronquist F. 2001. MrBayes: Bayesian inference of phylogenetic. *Bioinformatics* **17:** 754–755.

Hughes CE, Eastwood RJ, Bailey CD. 2006. From famine to feast? Selecting nuclear DNA sequence loci for plant species-level phylogeny reconstruction. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences* **361:** 211–225.

Maddison WP, Maddison DR. 2009. *Mesquite: a modular system for evolutionary analysis, Version 2.72.* Available at: http://mesquiteproject.org

Myers N. 2000. Biodiversity hotspots for conservation priorities. *Nature* **403:** 853–858.

Nylander JAA. 2004. *MrModeltest*, Version 2. Program distributed by the author. Uppsala University: Evolutionary Biology Centre.

Nylander JAA, Olsson U, Alström P, Sanmartín I. 2008a. Accounting for phylogenetic uncertainty in biogeography: a Bayesian approach to dispersal-vicariance analysis of the thrushes (Aves: Turdus). *Systematic Biology* **57:** 257–268.

Nylander JAA, Wilgenbusch JC, Warren DL, Swofford DL. 2008b. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* **24:** 581–583.

Parisod C, Holderegger R, Brochmann C. 2010. Evolutionary consequences of autopolyploidy. *New Phytologist* **186:** 5–17.

Pedersen HE. 2006. Systematics and evolution of the Dactylorhiza romana/sambucina polyploid complex (Orchidaceae). *Botanical Journal of the Linnean Society* **152:** 405–434.

Pirani JR, Giulietti AM, Mello-Silva R, Meguro M. 1994. Checklist and patterns of geographical distribution of the vegetation of Serra do Ambrósio, Minas Gerais, Brazil. *Revista Brasileira de Botânica* **17:** 133–147.

Ramirez SR, Gravendeel B, Singer RB, Marshall CR, Pierce NE. 2007. Dating the origin of the Orchidaceae from a fossil orchid with its pollinator. *Nature* **448:** 1042–1045.

Ribeiro MC, Metzger JP, Martensen AC, Ponzoni FvJ, Hirota MrM. 2009. The Brazilian Atlantic Forest: how much is left, and how is the remaining forest distributed? Implications for conservation. *Biological Conservation* **142:** 1141–1153.

Safford HD. 1999. Brazilian Páramos I. An introduction to the physical environment and vegetation of the campos de altitude. *Journal of Biogeography* **26:** 693–712.

Safford HD. 2007. Brazilian Páramos IV. Phytogeography of the campos de altitude. *Journal of Biogeography* **17:** 1–22.

Sanderson MJ. 2002. *Software r8s,* Version 1.70. Available at: http://loco.biosci.arizona.edu/r8s/

Simon MF, Grether R, de Queiroz LP, Skema C, Pennington RT, Hughes CE. 2009. Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. *Proceedings of the National*

Academy of Sciences of the United States of America **106:** 20359–20364.

Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson AH, Zheng C, Sankoff D, DePamphilis CW, Wall PK, Soltis PS. 2009. Polyploidy and angiosperm diversification. *American Journal of Botany* **96:** 336–348.

Stebbins GL. 1984. Polyploidy and the distribution of the arctic-alpine flora: new evidence and a new approach. *Botanica Helvetica* **94:** 1–13.

Suda J, Weiss-Schneeweiss H, Tribsch A, Schneeweiss GM, Trávnícek P, Peter Schönswetter P. 2007. Complex distribution patterns of di-, tetra-, and hexaploid cytotypes in the European high montain plant *Senecio carniolicus* (Asteraceae). *American Journal of Botany* **94:** 1391–1401.

Swofford DL. 1999. *PAUP* Phylogenetic analysis using parsimony (*and other methods)*. Sunderland, MA: Sinauer Associates.

Verola CF, Semir J. 2007. *Hoffmannseggella viridiflora* (Orchidaceae, Laeliinae) a New species from Brazilian Campos Rupestres. *Novon* **17:** 125–129.

Verola CF, Semir J, Antonelli A, Koch I. 2007. Biosystematic studies in the brazilian endemic genus *Hoffmannseggella* H. G. Jones (Orchidaceae: Laeliinae): a multiple approach applied to conservation. *Lankesteriana* **7:** 419–422.

Yamagishi-Costa J, Forni-Martins ER. 2009. Hybridization and polyploidy: cytogenetic indications for *Hoffmannseggella* (Orchidaceae) species evolution. *International Journal of Botany* **5:** 93–99.

Yu Y, Harris AJ, He X. 2010. S-DIVA (statistical dispersal-vicariance analysis). Available at: http://mnh.scu.edu.cn/S-diva/

Zachos JC, Dickens GR, Zeebe RE. 2008. An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature* **451:** 279–283.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1. 50% majority-rule consensus phylogram derived from 37 500 trees from the stationary (post burn-in) Bayesian sample. The scale bar indicates the amount of expected nucleotide changes per site.

Figure S2. Ancestral area optimization using statistical dispersal-vicariance analysis (S-DIVA; Yu *et al.*, 2010). The tree is the maximum a posteriori tree from all the Bayesian runs. Pie charts at nodes represent relative frequencies of ancestral-area reconstructions obtained from 5000 optimizations. Relative frequencies of ancestral ranges were only calculated from trees containing the node in the reference tree (meaning that some nodes may have very low phylogenetic support, although their ancestor still appear to be unambiguously reconstructed; compare with Fig. 2). Current species distributions are shown by capital letters preceding the species names (see legend at bottom). Left: S-DIVA results constraining ancestral nodes to occupy a maximum of two areas. Right: The same analysis, but with no maximum area constraint.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.