


## Research Article

# Worldwide long-distance dispersal favored by epizoochorous traits in the biogeographic history of Omphalodeae (Boraginaceae)

Ana Otero<sup>1,2\*</sup> , Pedro Jiménez-Mejías<sup>3,4</sup>, Virginia Valcárcel<sup>3,4</sup>, and Pablo Vargas<sup>1</sup>

<sup>1</sup>Departamento de Biodiversidad, Real Jardín Botánico, CSIC. Pza. de Murillo, 2, Madrid 28014, Spain

<sup>2</sup>Escuela Internacional de Doctorado, Universidad Rey Juan Carlos, C/Tulipán s/n, Móstoles 28933, Spain

<sup>3</sup>Centro de Investigación en Biodiversidad y Cambio Global (CIBC-UAM), Universidad Autónoma de Madrid, Madrid 28049, Spain

<sup>4</sup>Departamento de Biología (Botánica), Universidad Autónoma de Madrid, C/Darwin, 2, Madrid 28049, Spain

\*Author for correspondence. E-mail: anaotero@ gmail.com

Received 1 October 2018; Accepted 12 April 2019; Article first published online 30 April 2019

**Abstract** Biogeographic dispersal is supported by numerous phylogenetic results. In particular, transoceanic dispersal, rather than vicariance, is suggested for some plant lineages despite current long distances between America and Europe. However, few studies on the biogeographic history of plants have also studied the role of diaspore syndromes in long-distance dispersal (LDD). Species of the tribe Omphalodeae (Boraginaceae) offer a suitable study system because the species have a wide variety of diaspore traits related to LDD and different lineages conform to patched worldwide distributions on three distant continents (Europe, America and New Zealand). Our aim is to reconstruct the biogeographical history of the Omphalodeae and to investigate the role of diaspore traits favoring LDD and current geographic distributions. To this end, a time-calibrated phylogeny with 29 of 32 species described for Omphalodeae was reconstructed using biogeographical analyses (BioGeoBEARS, Lagrange) and models (DEC and DIVA) under different scenarios of land connectivity. Character-state reconstruction (SIMMAP) and diversification rate estimations of the main lineages were also performed. The main result is that epizoochorous traits have been the ancestral state of LDD syndromes in most clades. An early diversification age of the tribe is inferred in the Western Mediterranean during late Oligocene. Colonization of the New World by Omphalodeae, followed by fast lineage differentiation, took place sometime in the Oligocene-Miocene boundary, as already inferred for other angiosperm genera. In contrast, colonization of remote islands (New Zealand, Juan Fernández) occurred considerably later in the Miocene-Pliocene boundary.

**Key words:** colonization, diversification, intercontinental disjunction, molecular dating, syndromes, Zealandia.

## 1 Introduction

Long-distance dispersal (LDD) and vicariant events have traditionally been proposed as two alternative hypotheses to account for the major disjunctions in historical biogeography. In particular, transoceanic disjunct distribution at high taxonomic ranks (e.g., subfamilies, tribes) has been traditionally explained by historical events of vicariance attributed to plate tectonics movements. For example, the disappearance of the North Atlantic Land Bridge has been linked to the current Old World-New World disjunction of temperate Tertiary relict flora of Eurasia and North America (Tiffney & Manchester, 2001), which implies that such territories were occupied by plants before the re-opening of the Bering Strait (10 000 years; Elias et al., 1996). However, recent development of biogeographic methods (from processes-based to recent event-based models; Yu et al., 2015) and divergence time estimations have revealed a geographic and temporal context that does not fit the vicariance hypothesis for

numerous transoceanic lineages (Sanmartín & Ronquist, 2004). Indeed, dispersal is revealed as fundamental to understand the organization of floristic assembly (Harris et al., 2018). In particular, recent phylogenetic studies have set most of the intercontinental disjunctions in the Northern Hemisphere between Old-New world during the Miocene, indicating LDD as the biogeographic process that accounts for disjunct distributions (Wen & Ickert-Bond, 2009). Intercontinental disjunctions in the Southern Hemisphere have also been largely discussed in the last few decades (Thornhill et al., 2015), including the Gondwana breakup around 60 million years ago (mya) in the Southern Hemisphere (Sanmartín & Ronquist, 2004). Implementation of recent biogeographic methods based on time-calibrated phylogenies allows linking the Gondwana breakup with the origin of Southern Hemisphere disjunctions in some groups of plants at high taxonomic levels (e.g., Poales, Bremer, 2002; Campanulids, Beaulieu et al., 2013). In contrast, most of the trans-Pacific disjunctions at shallow phylogenetic levels

resulted from recent LDD events (Sanmartín & Ronquist, 2004). Indeed, recent dispersal between long-distant regions, such as the Eastern Pacific (South America, SA) and the Western Pacific (Australia, Aus; New Zealand, NZ), have been inferred in both directions during the Mid Miocene (e.g., Renner et al., 2000; Crisp et al., 2009; Chen et al., 2014). In sum, time-calibrated phylogenies support LDD when transoceanic disjunctions are the result of the recent split of lineages that postdated land separation.

The occurrence of LDD in plants has traditionally been linked to specialized traits of seeds and fruits (Van der Pijl, 1982). Specializations associated with dispersal by wind (anemochory), seawater (thalassochory) and internal/external animal dispersal (endo-/epizoochory, respectively) are identified as LDD diaspora syndromes (Vargas et al., 2012). However, recent studies suggest that unspecialized diaspores play a more important role in LDD than previously believed in island colonization from continents (Heleno & Vargas, 2015). LDD syndromes are found to be more advantageous in inter-island colonization, where considerable distances have to be overcome by plant diaspores (Arjona et al., 2018). Irrespective of whether dispersal specialization is present or not and the actual means of dispersal, LDD is currently recognized as a key process for colonizing remote territories and shaping current lineage distributions, especially in the colonization of oceanic islands (Gillespie et al., 2012). Lineages with discontinuous distributions provide ideal systems to investigate biogeographical processes. In this sense, the tribe Omphalodeae (Boraginaceae: subfamily, Cynoglossoideae) offers us an ideal case of study to infer complex biogeographic histories and to understand the role of different biogeographic processes. The tribe Omphalodeae is a monophyletic group that comprises six genera (ca. 32 species) distributed in seven distant areas of endemism: (1) Eastern Mediterranean (*Omphalodes*, ca. 10 spp.); (2) Western Mediterranean (*Iberodes*, five spp.; *Gyrocarum*, monospecific); (3) Sierra Madre Oriental and Southern Texas (*Mimophytum*, ca. 11 spp.); (4) NW South America (*Selkirkia trianaeum*); (5) Chile (*S. pauciflorum* and *S. limense*); (6) Juan Fernández archipelago (*S. berteroi*); and (7) Chatham Islands (*Myosotidium*, monospecific). Most recent phylogenetic studies (Holstein et al., 2016a, 2016b) reveal at least two major intercontinental disjunctions in both hemispheres since its ancient origin around Early Miocene (Otero et al., 2019). On the one hand, Old-New world disjunctions are observed by Western Palearctic *Omphalodes* and the American *Mimophytum* and *Selkirkia* (Holstein et al., 2016a, 2016b). On the other hand, a trans-Pacific disjunction between South America (Chile and Ecuador) and New Zealand is inferred for the split of the *Selkirkia* and *Myosotidium* lineages (Holstein et al., 2016a). A considerable presence of diaspora specializations associated with LDD is observed in Cynoglossoideae (Cohen, 2013; Chacón et al., 2016; Schenk & Saunders, 2017). In particular, a variety of open and wide wings (associated with anemochorous traits) and different types of anchoring structures like uncinated hairs, dentate margins and glochids (associated with epizoochorous traits) are found in the genera of the tribe Omphalodeae (Otero et al., 2019). However, the question remains as to whether vicariant events, LDD events, or both have been

predominant in the spatiotemporal context of the biogeographic history of this tribe.

Given the diversity of the epizoochorous structures of the tribe Omphalodeae, LDD is hypothesized to be responsible for current transoceanic disjunctions. To test this hypothesis, we performed divergence time estimates for 29 of the 32 species of Omphalodeae using four DNA regions of the nuclear and plastid genomes. We also tested different biogeographic models and software packages in order to infer the most significant biogeographic events occurred in this tribe. Diversification rate analyses and ancestral-state reconstruction of LDD syndromes were additionally performed. In particular, the specific aims of this study were to (1) infer the geographic origin of the most recent common ancestor of Omphalodeae; (2) investigate the timing and direction of migration patterns responsible for the current disjunct distributions; (3) reconstruct biogeographic events that could have shaped the current diversification patterns in Omphalodeae; and (4) evaluate the role of LDD syndromes in the biogeographic events of Omphalodeae.

## 2 Material and Methods

### 2.1 Taxon sampling

All six genera and 29 of the 32 species currently recognized for the tribe Omphalodeae were sampled (Table S1). Two individuals per species were sampled, except for 10 species (*Mimophytum alienoides*, *M. australe*, *M. chiangii*, *M. mexicanum*, *M. omphalodes*, *Omphalodes lojkae*, *O. runnemarii*, *O. ripleyana*, *Selkirkia limense*, and *S. berteroi*) from which we only obtained one individual each because of the scarcity of these species in the herbaria. For the sake of simplicity, and according to our own results (see Appendix 1), we already treat *Omphalodes erecta* as *Mimophytum erectum* in the text. A total of 48 ingroup samples were sequenced, plus six taken from Genbank (Table S1). In addition, the outgroup was formed by 29 species from each of the other tribes and subtribes of Cynoglossoideae, subfamilies of Boraginaceae and families of Boraginales that were taken from Genbank (Table S2). As a result, the total sampling consisted of 35 genera and 60 species of Boraginales. Plant material was obtained from our own field collections and herbarium specimens from 11 herbaria (Table S1).

### 2.2 DNA extraction and sequence matrices

DNA was extracted from leaf tissue using a modified version of CTAB protocol from Doyle & Doyle (1987) and Shepherd & McLay (2011), including a precipitation step in isopropanol and Ammonium acetate 7.5 mol/L at -20°C overnight. Three plastid regions were sequenced: (1) the spacer *trnS-trnG*, (2) the spacer *trnL-trnF* (the *trnL* (UAA) intron), and (3) the *rps16* intron. In addition, the nuclear ribosomal DNA ITS (Internal Transcribed Spacer) was chosen based on considerable phylogenetic signal found in previous studies of Boraginaceae (Weigend et al., 2013; Chacón et al., 2017; Chacón et al., 2016). Primers for the ITS and *trnL-trnF* regions followed those of Otero et al. (2014). For the *rps16* region, we used the primers described in Otero et al. (2019). For *trnS-trnG* sequences, we used *trnS* (GCU) and *trnG* (UCC) of Hamilton (1999) as external primers. In addition, we designed two new internal primers for *trnS-trnG* to improve

the success of some amplifications: 'trnS-trnG BOR 158 F' (5' AAA GRG TCG AAA GAA TCA AC 3') and 'trnS-trnG BOR 889 R' (5' GTC TTT CCG TAA AGT GAT TC 3'). The PCR conditions for plastid and nuclear regions followed those for the *trnL-trnF* and ITS (Otero et al., 2014), respectively. PCR products were sequenced using the Macrogen Europe sequencing service (Amsterdam, The Netherlands). All DNA regions successfully sequenced were compiled in a matrix of 77 accessions that were automatically aligned using a supermatrix approach through Fast Fourier transform (MAFFT, Katoh et al., 2002) and manually reviewed using Geneious version R10 (Kearse et al., 2012). The DNA regions amplified for each taxon are specified in Tables S1, S2. The nucleotide substitution model that best fitted each partition (*trnS-trnG*, *trnL-trnF*, *rps16*, and ITS) was inferred using JModelTest v. 0.1.1. (Posada, 2008).

### 2.3 Phylogenetic reconstructions and topology tests

We performed Bayesian inference and Maximum Likelihood analyses using, respectively, MrBayes 3.2.6 (two runs of 50 million generations) and RAXML V8 (rapid Bootstrap analysis with automatically stop bootstrapping through majority rule criterion), through the CIPRES portal (Miller et al., 2015), in order to assess the topological robustness.

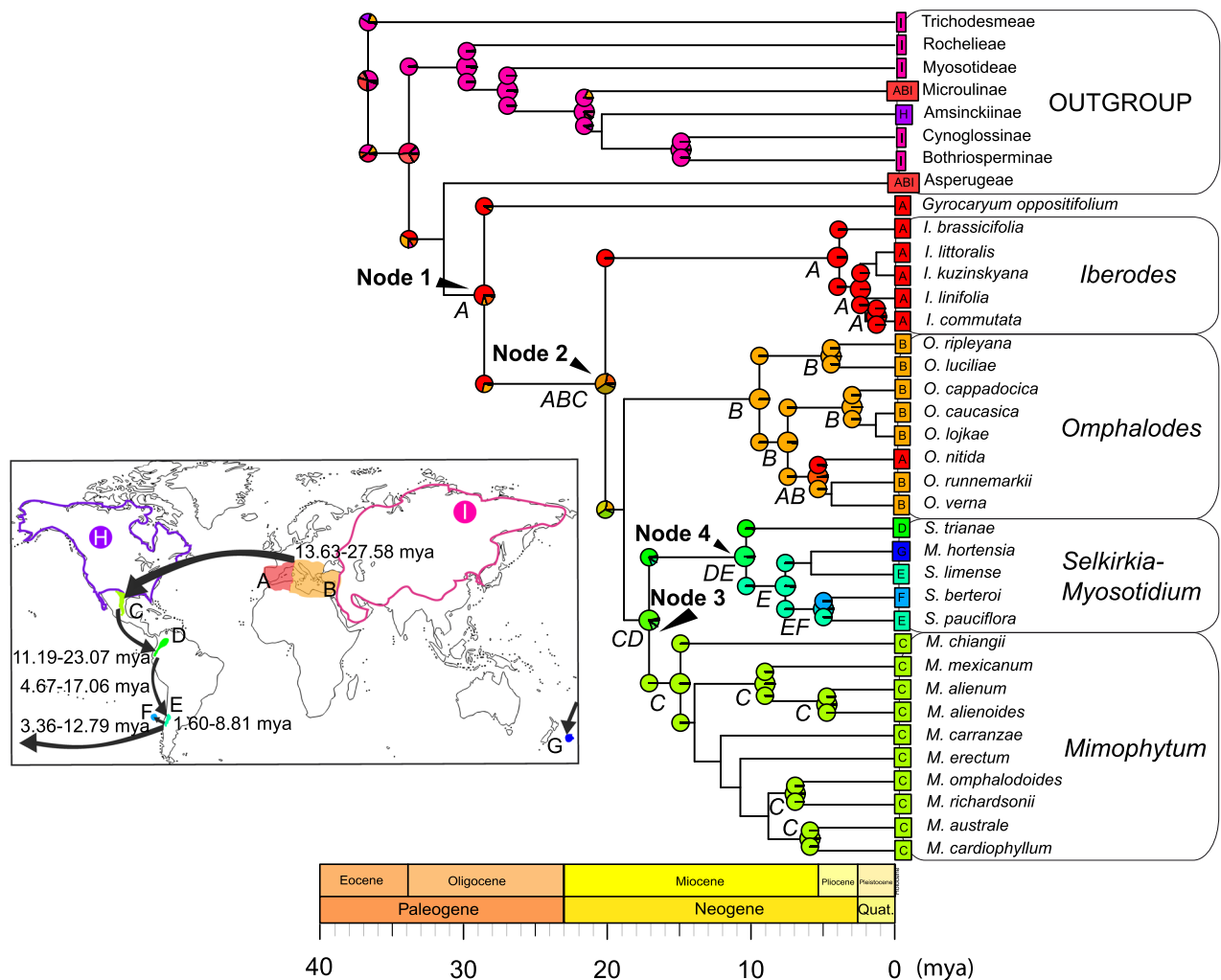
After manually reviewing the topologies that resulted from MrBayes and RAXML, hypothesis testing analyses were performed to evaluate the most likely topology for two uncertain clades (see Fig. 1, node 2: *Iberodes*, *Omphalodes* and *Myosotidium-Selkirkia-Mimophytum*; and node 4: *Myosotidium-Selkirkia*). Comparison of the topologies was performed using both the Bayes Factors (BF; Kass & Raftery, 1995; Suchard et al., 2001) and the AU test (Shimodaira, 2002). Independent constrained analyses were conducted with RAXML and MrBayes to obtain input trees for each alternative hypothesis. Topology tests were first conducted on the most basal node (node 2) using the unconstrained topology (*Iberodes*, (*Omphalodes*, *Myosotidium-Selkirkia-Mimophytum*)) and the two alternative hypotheses (H1: (*Myosotidium-Selkirkia-Mimophytum*, (*Omphalodes*, *Iberodes*)), H2: (*Omphalodes*, (*Myosotidium-Selkirkia-Mimophytum*, *Iberodes*))). Once the best-scored topology among three alternative hypotheses was chosen, the same hypothesis testing procedure was done for the node 4 using the unconstrained topology (*Myosotidium* nested within *Selkirkia*) and *Myosotidium* sister to a monophyletic *Selkirkia* as the alternative hypothesis. In all cases, to estimate the marginal likelihoods of the hypotheses for the BF we used the stepping-stone sampling method (Xie et al., 2011) through MrBayes 3.2.6. in CIPRES portal (Miller et al., 2015). The stepping-stone method provides more accurate estimates than the harmonic mean method (Xie et al., 2011; Bergsten et al., 2013). The analysis was performed for 50 steps with 980 000 generations per step in two independent parallel MCMC runs. Convergence was examined through diagnostic plots of standard deviation of the split frequencies and screening similarity in the two independent marginal likelihood estimates (Ronquist et al., 2011). Selection of the best competing hypotheses against the H0 was based on Bayes Factors evidence interpreted according to the table of Kass & Raftery (1995): Positive evidence, BF = 3–20; Strong evidence, BF = 20–150; Very Strong evidence, BF > 150. Positive values of the BF indicate preference for H0, while negative values favour H1. The AU test (Shimodaira, 2002) is very conservative when rejecting the null

hypothesis (i.e., all the trees under evaluation are equally good) and has been proven to perform better than other nonparametric bootstrapping methods (Planet, 2006). The AU test was performed in Treefinder (Jobb et al., 2004; Jobb, 2007). The likelihood of the majority-rule consensus tree obtained from the RAXML was compared using 10 000 replicates and partitioning datasets. Competing hypotheses were rejected at a significance level of 0.05.

### 2.4 Estimation of divergence times

Topology tests did not distinguish a preferred topology for one of the two uncertain clades (*Selkirkia-Myosotidium*, node 4). In order to estimate the two likely scenarios for the transoceanic disjunction of this clade, we estimated divergence times by evaluating the two different topologies obtained through Bayesian (BI) and Maximum Likelihood inferences (ML).

Firstly, divergence time reconstructions for the topology obtained through BI (node 4 unconstrained topology: *Myosotidium* nested within *Selkirkia*) were performed through an unconstrained analysis with BEAST v.1.8.6 (Drummond et al., 2012) on the CIPRES portal (Miller et al., 2015). We ran four independent runs of  $50 \times 10^6$  generations, sampling every 5000 generations in four independent Markov chain Monte Carlo. A Birth-Death tree model was set for all of the four partitions (*trnS-trnG*, *trnL-trnF*, *rps16*, and ITS). Two Uncorrelated Log-normal relaxed clocks were set separately for plastid (*trnS-trnG*, *trnL-trnF* and *rps16*; uniform distribution =  $1.0E-4$ – $1.0E-2$ ) and the nuclear region (ITS; uniform distribution =  $5.0E-4$ – $5.0E-2$ ) based on the boundaries proposed Blanco-Pastor et al. (2012) that included ranges of other angiosperm groups (Wolfe et al., 1987). Separated models of nucleotide substitution were considered for each of the four partitions. The root and crown node of Boraginaceae were calibrated based on the estimates of Luebert et al. (2017) (root node: normal prior, mean = 68.13 mya, stdev = 5.0; crown age of Boraginaceae: normal prior, mean = 53.20 mya, stdev = 4.0). In addition, four fossils were used to perform node calibration. The calibration points were set in all cases to the stem node following Otero et al. (2019). In order to account for the uncertainty of fossil calibration and to fully include the chronological information inferred from the fossil record, we fitted a lognormal probability distribution to each node. This prior is considered appropriate to summarize the paleontological information since it allows for a soft upper bound by assigning the highest point probability for the nodal age to be somewhat older than the oldest fossil (Ho & Phillips, 2009). We used four fossil calibration points whose time boundaries are in agreement with previous studies of Boraginaceae (Chacón et al., 2016; Otero et al., 2019): (1) stem node of genus *Lithospermum* (lognormal prior, logMean = 0.5, logStdev = 1.0 offset = 10.3 mya; Gabel, 1987; Thomasson, 1987); (2) stem node of clade Amsinckiinae (excluding *Andersoglossum* and *Adelinia*) (lognormal prior, logMean = 0.5, logStdev = 1.0 offset = 10.3 mya; Segal, 1966; Thomasson, 1987; Gabel et al., 1998); (3) stem node of *Lappula* and *Rochelia* (lognormal prior, logMean = 0.5, logStdev = 1.0 offset = 10.3 mya; Thomasson, 1987; Gabel et al., 1998); (4) stem node of clade Echiochileae (lognormal prior, logMean = 2.0, logStdev = 1.0 offset = 41.2 mya; Hammouda et al., 2015). Convergence and Effective Sample Size (ESS) values for all parameters were assessed using Tracer v.1.6 (Rambaut et al., 2014).



**Fig. 1.** Time-calibrated phylogeny from the unconstrained BEAST analysis and dispersal-extinction-cladogenesis (DIVA) reconstruction of Omphalodeae under M1 scheme. Names of the main taxonomic clades of Omphalodeae are framed. Biogeographic reconstruction is shown for nodes with Bayesian posterior probability (BPP) higher than 0.90. Nodes numbered 1–4 indicate the main internal nodes commented along the manuscript. Relative probability of each area combination is represented by each node pie. The most likely ancestral area including area combinations is indicated with capital letters on each node for the ingroup. Color legend and a map for each area are provided. Direction (black arrows) and time (mya) of main dispersal events between areas are shown in the map. International stratigraphic scale is included from 45 mya until present (0 mya). Geographic areas considered for biogeographic reconstructions are: A: Western Mediterranean, B: Eastern Mediterranean, C: Sierra Madre Oriental, D: NW South America, E: Chile, F: Juan Fernández Archipelago, G: Chatham islands, H: North America and I: Asia.

Subsequently, LogCombiner v.1.8.2. (Drummond et al., 2012) was used to combine the four runs and we applied a 25% burn-in of the total states on each run. Finally, TreeAnnotator v.1.8.2. (Drummond et al., 2012) was run to obtain the maximum clade credibility (MCC) tree.

Secondly, we also ran four independent runs in BEAST v.1.8.6 with the same fossil calibrations explained above but implementing clade constraints of Myosotidium-Selkirkia clade to fit the topology of ML (node 4 constrained topology: *Mysostidium* sister to *Selkirkia*). In addition, we obtained an ultrametric tree from the ML tree through the function “chronopl” (R package “ape”, Paradis et al., 2004; R Development Core Team, 2013) to provide a starting tree for the BEAST inference and thus avoid convergence errors. As previously done

for the unconstrained BEAST analysis, the ESS values were assessed in Tracer v.1.6 (Rambaut et al., 2014), LogCombiner v.1.8.2. (Drummond et al., 2012) was used to combine the four runs with a 25% burn-in each run, and finally the MCC tree was built in TreeAnnotator v.1.8.2. (Drummond et al., 2012).

## 2.5 Biogeographic analyses

The same procedure of biogeographic analyses was repeated for the two phylogenetic dated trees obtained in BEAST (with and without clade constraints). In both cases, the ingroup was pruned in order to leave one individual per species (29 spp.) and the outgroup was reduced to subfamily Cynoglossoideae, including one tip per each of the eight subgroups defined in Otero et al. (2019). Selection of areas

for the ingroup was based on the areas of endemism of the six genera that conform to the tribe Omphalodeae. In particular, east-west division of the Mediterranean Basin (including Italy as part of the eastern Mediterranean) fits the endemism pattern of Mediterranean genera of Omphalodeae (*Gyrocarum*, *Iberodes* and *Omphalodes*) in agreement with most Mediterranean Floras and the biogeographic and paleo-botanical history of the Mediterranean (Thompson, 2005; Feliner, 2014; Iñda et al., 2014; Rundel et al., 2016). Areas were assigned to each tip according to current distributions of the included species, which represent the phylogenetic and geographic diversity of each genus of the ingroup. Thus, seven geographic areas were delimited: (1) Eastern Mediterranean Basin (E Medit), where 80% of the species of *Omphalodes* are endemic; (2) Western Mediterranean Basin (W Medit), including Italy, where *Iberodes* with five species, the monospecific *Gyrocarum* and one species of *Omphalodes* are distributed; (3) Sierra Madre Oriental (SMO), including southern Texas, where *Mimophytum* is exclusively distributed; (4) NW South America (NW SAM), including Ecuador and Peru, where *Selkirkia trianaeum* is distributed; (5) Chile (endemic area of *Selkirkia pauciflorum* and *S. limense*); (6) Juan Fernández Archipelago, where *Selkirkia berteroi* is endemic to; and (7) Chatham Islands of New Zealand, which is the endemic area of the monospecific genus *Myosotidium*. The areas assigned to the outgroup species were set by using the ancestral area obtained in Otero et al. (2019) for the eight subgroups represented by the eight species included here as the outgroup. Two more areas were considered to accommodate the distribution of outgroup species: (1) North America (endemic area for the crown node of tribe Amsinckinae; Otero et al., 2019) and (2) Asia (main centre of diversification of Cynoglossoideae lineages; Otero et al., 2019). In total, nine areas were considered for biogeographic analyses.

Vicariance is a major biogeographic event that needs to be considered in our biogeographic reconstruction, given the temporal and geographic context for the evolution of the study group. Therefore, we selected Dispersal-extinction-cladogenesis model (DEC, Ree & Smith, 2008) and parsimonious dispersal-vicariance analysis (DIVA, Ronquist, 1997) as the biogeographic models to be implemented. DEC was implemented in Lagrange 2.0.1 (Ree & Smith, 2008) and “BioGeoBEARS” package (Matzke, 2018) in R (R Development Core Team, 2013). We also implemented DIVA approach through the “BioGeoBEARS” package. In particular, the “BioGeoBEARS” package implements a likelihood interpretation of DIVA (DIVALIKE), i.e., processes are modelled in the way that DEC does but allowing only the parameters considered in each case (e.g., in DIVA, widespread vicariance is allowed but not subset sympatry; Ronquist & Sanmartín, 2011). In addition, the DIVA and DEC models of “BioGeoBEARS” were also tested adding the parameter “j” that considers the founder-event speciation in the biogeographic model. However, the results of these models were not considered for the discussion given the conceptual and statistical problems of this parameter found in recent publications (Ree & Sanmartín, 2018). In order to test the reconstruction of the ancestral area under different scenarios of connectivity and dispersal rates between areas, we proposed four schemes from a null model with no

restrictions to other more restricted: (1) M0, with no restrictions in dispersal rates and adjacency between areas; (2) M1, in which we scaled the dispersal rates between areas by assuming three levels of dispersal based on the proximity of areas involved, i.e., factor of 1 between adjacent areas (e.g., E Medit and W Medit); factor of 0.1 between non-adjacent areas connected by land (e.g., SMO and NW SAM); and factor of 0.01 for the rest of area combinations separated by oceans (e.g., W Medit and SMO); (3) M2, in which adjacency between areas is allowed only between areas that share the edge (e.g., E Medit and Asia); and (4) M3, in which both constrictions set at M1 and M2 are assumed. The four schemes were tested for both DEC and DIVA biogeographic models explained above and thus a total of eight models were performed. The number of ranges permitted in each analysis was set based on the constrictions assumed in each scheme and taking into account that only a range with a maximum of three areas were allowed to accommodate the widespread distribution of some of the subgroups that conform the outgroup. The dispersal rate matrix and adjacency matrix used for M1-M3 are shown in Tables S3 and S4. AIC calculation for the 8 models were done through the “BioGeoBEARS” package to compare likelihood estimation and choose the best fit model within each algorithm (DEC and DIVA). Given the inadequacy of likelihood comparison between different biogeographic models, the discussion of DEC or DIVA inference was done based on biological criteria (Ree & Sanmartín, 2018).

## 2.6 Analysis of diversification rates

Diversification rates of main lineages of Omphalodeae were explored using the whole-clade method of absolute diversification rates of Magallon & Sanderson (2001). Rates were calculated for both crown and stem clades, at two extremes of the relative extinction rate ( $k = 0$ , no extinction; and  $k = 0.9$ , high rate of extinction; where  $k = \text{speciation rate} / \text{extinction rate}$ ), as implemented in the R package “Geiger” (Harmon et al., 2007; R Development Core Team, 2013). We used the extreme values for the 95 per cent highest posterior density intervals inferred in the divergence time estimation obtained in BEAST.

## 2.7 Ancestral state reconstruction of LDD syndromes

The diaspore trait characterization was based on Otero et al. (2019), which includes a bibliographic review. Three dispersal syndromes were categorized based on specific sets of diaspore characters: (i) epizoochorous traits (hooked trichomes on either nutlet or calyx, nutlet glochids, toothed nutlet margins) are those that can promote dispersal externally to animals; (ii) anemochorous traits (air flotation assisted by a wide, open nutlet margin) are those that can promote dispersal by wind; and (iii) myrmecochorous traits, as a result of protrusion of the mesocarpic tissue, include particular tissues associated with dispersal by ants. Reconstruction of ancestral multi-state characters of all syndromes was performed (0 = unspecialized, 1 = epizoochorous, 2 = anemochorous, 3 = myrmecochorous). None of the species displayed multiple syndrome (polychorous) traits. Ancestral state reconstructions were done separately for both alternative topologies (with and without clade constraints) obtained in BEAST. States for the tips of the outgroup

were assigned according to the ancestral state inferred in Otero et al. (2019). Two hundred stochastic reconstructions were simulated through the stochastic mapping approach that was conducted using function “make.simmap” in the R package “PHYTOOLS” (Revell, 2012; R Development Core Team, 2013) and allowing different rates of transition between all states (model = “ARD”). As result, a summary tree of each of the 200 simulations was obtained from each character reconstruction and for each topology analyzed.

### 3 Results

#### 3.1 Phylogenetic reconstructions and topology tests

The length of the plastid matrix (*trnL-trnF*, *rps16* and *trnS-trnG*) was 3958 bp, of which 1156 informative sites were found. General time-reversible model (GTR) and gamma distribution of rates across sites was the best fitting substitution model for both plastid and nuclear regions. The Bayesian and Maximum Likelihood reconstructions obtained, respectively, with MrBayes and RAxML, were congruent with that obtained using Beast (Figs. S1, S2), except for the node of *Myosotidium* and *Selkirkia*. Although with marginal support, Bayesian inference (BEAST and MrBayes) reconstructed *Myosotidium* nested within *Selkirkia*, while in the RAxML analysis *Myosotidium* is sister to *Selkirkia* (Figs. S1, S2).

The monospecific genus *Gyrocaryum* is the earliest-diverging lineage of the tribe Omphalodeae, sister to one major clade (node 2, Fig. 1) that is formed by subclades: (1) *Iberodes*, (2) *Omphalodes* and (3) *Mimophytum-Selkirkia-Myosotidium* (node 3). Relationships between these three subclades are not totally clear because of low phylogenetic support in both Bayesian and Maximum likelihood analyses (Figs. S1, S2). The Bayes factor test to evaluate relationships between *Iberodes*, *Omphalodes* and node 3 favoured with positive evidence the unconstrained topology (H<sub>0</sub>: *Iberodes* sister to *Omphalodes* and the clade arising from node 3; Table 1). The AU test only rejected the alternative topology H<sub>2</sub> (*Iberodes* and the clade arising from node 3 sister to *Omphalodes*) (Table 1). As a result, the unconstrained topology is preferred for posterior phylogenetic-based analyses. *Iberodes* constitutes a monophyletic genus where *I. brassifolia* is sister to a lineage formed by the remaining four species: *I. commutata*, *I. linifolia*, and *I. kuzynskyanae-I. littoralis* that form a monophyletic group. *Omphalodes* is a monophyletic subclade where *O. ripleyana* is sister to *O. luciliae*, and both are sister to a lineage formed by two lineages: *O. nitida-O. runnemarkii-O. verna*, and the Caucasian endemic *O. caucasica-O. capadocica-O. lojkae*. The third major subclade (node 3) is formed by two subclades: *Mimophytum* and *Selkirkia-Myosotidium* (node 4). Internal relationships among the 10 species of *Mimophytum* are not fully resolved (Fig. S1). A close phylogenetic relationship is found for *Selkirkia* and *Myosotidium* in all the analyses. Nevertheless, *Selkirkia* was found to be paraphyletic in the unconstrained BEAST analysis (*Myosotidium* was inferred nested within it), while RAxML shows the *Selkirkia* lineage sister to *Myosotidium* although with marginal support (Fig. S2). In addition, the Bayesian inference obtained with

MrBayes differed from the tree obtained with the unconstrained BEAST, since there is no high support for *Myosotidium* nested within *Selkirkia* in the former (Figs. S1, S2). Neither the AU test nor BayesFactor could significantly discard the paraphyly of *Selkirkia* (unconstrained topology H<sub>0</sub>, Table 1).

#### 3.2 Estimation of divergence times

Divergence times analyzed with BEAST gave the highest posterior density values for the split of most clades, subclades, and lineages of Omphalodeae (Figs. 1, S1; Table 2). The crown age of tribe Omphalodeae (node 1, Fig. 1) coincided with the mid-Oligocene (mean = 28.56 mya; 95% highest probability density (HPD) = 20.40-36.57 mya), while the earliest diverging clade (sister to *Gyrocaryum*) originated in the Early Miocene (node 2; mean = 20.13 mya; 95% HPD = 13.63-27.58 mya). *Iberodes* crown age is inferred during the Pliocene (mean = 3.84 mya; 95% HPD = 1.83-6.37 mya), while the crown age of *Omphalodes* was in the Late Miocene (mean crown age = 9.41 mya; 95% HPD = 4.97-14.57 mya). The third major subclade (node 3) arose during the Early Miocene (mean = 17.09 mya; 95% HPD = 11.19-23.07 mya) and *Mimophytum* short afterward (mean crown age = 14.90 mya ; 95% HPD = 9.70-20.34 mya). Interestingly, similar divergence age estimation between *Myosotidium* and the

**Table 1** Topology test with Bayes Factor and Approximately Unbiased test (AU test) for the three alternative topologies of node 2 (H<sub>0</sub>, H<sub>1</sub>, H<sub>2</sub>) and the two alternative topologies of node 4 (H<sub>0</sub>, H<sub>1</sub>). Bayes factor is represented by the quotient between marginal likelihood of each hypothesis and AU test values represent the differences between LnLikelihood of each hypothesis. Numbers in brackets are p-values obtained for AU test

Bayes Factor			
Node 2	H <sub>0</sub>	H <sub>1</sub>	H <sub>2</sub>
H <sub>0</sub> ( <i>Iberodes</i> , ( <i>Omphalodes</i> , Node 3))	-	9.31*	319.57**
H <sub>1</sub> (Node 3, ( <i>Iberodes</i> , <i>Omphalodes</i> ))	-	-	309.82**
H <sub>2</sub> ( <i>Omphalodes</i> , ( <i>Iberodes</i> , Node 3))	-	-	-
Node 4			
H <sub>0</sub> ( <i>Selkirkia</i> , <i>Myosotidium</i> )	-	-0.5	-
H <sub>1</sub> ( <i>Selkirkia trianaeum</i> , ( <i>Selkirkia</i> spp., <i>Myosotidium</i> ))	-	-	-
AU test			
Node 2			
H <sub>0</sub>	0 (0.84)	8.51 (0.22)	11.04 (0.07)
Node 4			
H <sub>0</sub>	0 (0.50)	-0.190 (0.50)	-

Asterisks indicate evidence (\* Positive, \*\* Very strong) in favor to (positive values) or against (negative values) the null hypothesis (H<sub>0</sub>) for the Bayes Factor.

**Table 2** Summary of node ages (BEAST) and ancestral area reconstruction (Biogeobears) of main clades of Omphalodeae: mean ages (mya) of each clade and highest posterior density ranges for stem and crown nodes; ancestral area and the relative probability inferred in biogeographic reconstruction with DIVA and DEC. Ancestral areas with a relative probability above 0.10 are shown. Geographic areas codes are: A: Western Mediterranean, B: Eastern Mediterranean, C: Sierra Madre Oriental, D: NW South America, E: Chile, G: Chatham islands. Hyphen (-) indicates nodes with Bayesian posterior probability (BBP) and Bootstrap value below 0.90 and 70, respectively. † indicates divergence time and ancestral area probability with the constrained BEAST analysis for node 4.

Clades	Mean Age (95% HPD) Stem node	Mean Age (95% HPD) Crown node	DIVA (relative probability)	DEC (relative probability)
Omphalodeae (Node 1)	-	28.56 (20.40-36.57)	A (0.71)	A (0.28), ABC (0.21), ABD (0.21), AB
Node 2	28.56 (20.40-36.57)	20.13 (13.63-27.58)	ABC (0.37), ABD (0.34), AB (0.20)	ACD (0.19), ABC (0.17), ABD (0.15), AC (0.11), AD (0.11)
<i>Iberodes</i>	20.13 (13.63-27.58)	3.84 (1.83-6.37)	A (0.99)	A (0.99)
<i>Omphalodes</i>	-	9.41 (4.97-14.57)	B (0.99)	AB (0.72), B (0.28)
Node 3	-	17.09 (11.19-23.07)	CD (0.77)	CD (0.60)
Node 4	17.09 (11.19-23.07)	10.33 (4.67-17.06)	DE (0.97) † DEG (0.55)	DE (0.69), CDE (0.15), DEG (0.13) † DEG (0.39)
† Node 4	† 17.78 (12.09-24.15)	† 9.25 (4.25-15.19)		
<i>Mimophytum</i>	17.09 (11.19-23.07)	14.90 (9.70-20.34)	C (0.99)	C (0.98)

South American *Selkirkia* (node 4) in the late Miocene was inferred for both the unconstrained BEAST results (mean root node age = 7.60 mya ; 95% HPD = 3.36-12.79 mya) and the constrained analysis (mean crown node age = 9.25 mya; 95% HPD = 4.25-15.19 mya) (Fig. S3). In particular, under unconstrained BEAST analysis, *Selkirkia trianae* appears to be sister to a subclade (*Myosotidium hortensia*-*S. limense* sister to *S. berteroi*-*S. pauciflora*) that is not fully resolved (mean crown age = 5.80 mya; 95% HPD = 1.84-10.44 mya).

### 3.3 Biogeographic analyses

The AIC test for the eight considered biogeographic models revealed that the best-fit model within each biogeographic model (DEC and DIVA) was the scheme M1 (scaled dispersal rates between areas based on proximity) in both DEC and DIVA, and also for the two topologies (Tables S5, S6). Both DEC and DIVA models resulted in biologically congruent reconstructions. Nevertheless, the less ambiguous reconstruction obtained with the DIVA model better resolved the biogeographic history of Omphalodeae. Probabilities of ancestral ranges obtained with DEC and DIVA for M1 are summarized for the main lineages of Omphalodeae in Table 2 and are represented in Figs. 1, S4-S6. Under unconstrained BEAST topology, the ancestral area of Omphalodeae was ambiguously inferred with DEC (Table 2; node 1, Fig. S4). In contrast, Western Mediterranean is clearly inferred as the ancestral area of Omphalodeae with a DIVA model ( $p = 0.71$ ) (Table 2; node 1, Fig. 1). Although both DEC and DIVA are congruent with a widespread ancestor for node 2 in Old and New World, uncertainty between Sierra Madre and NW South America (SAM) is inferred in DEC (Table 2; node 2, Fig. S4). Otherwise, DIVA inferred a widespread ancestor in the Mediterranean and the Sierra Madre Oriental ( $p = 0.37$ ) for node 2 (Table 2, Fig. 1). A widespread area in Sierra Madre Oriental and NW SAM is inferred for node 3 with both DEC ( $p = 0.61$ ) and DIVA ( $p = 0.77$ ). The ancestral area of *Mimophytum* is significantly inferred in Sierra Madre Oriental in both DEC ( $p = 0.99$ ) and DIVA ( $p = 0.97$ ). Finally, according

to the topology of unconstrained BEAST, the ancestral area of *Selkirkia-Myosotidium* (node 4) is inferred as a widespread range in NW SAM and Chile with both DEC ( $p = 0.69$ ) and DIVA ( $p = 0.97$ ) (Table 2; Figs. 1, S4). However, according to the topology of the constrained BEAST the ancestral area of node 4 is in a widespread area formed by NW SAM, Chile and New Zealand for both DEC ( $p = 0.39$ , Fig. S5) and DIVA ( $p = 0.55$ , Fig. S6). Two LDD from mainland to oceanic islands have been inferred for *Myosotidium* (Chatham Islands, New Zealand) and *S. berteroi* (Juan Fernández Islands, Chile). The ancestor of *Myosotidium* is widespread in NW SAM, Chile, and New Zealand according to the constrained topology and in Chile according to the unconstrained topology (Figs. 1, S4). *Selkirkia berteroi* is sister to *Selkirkia pauciflora* and Chile is inferred to be the ancestral area of their common ancestor (Figs. 1, S4-S6).

### 3.4 Analysis of diversification rates

Diversification rates inferred for main subclades of Omphalodeae at crown and stem node ages are shown in Table 3. The global rate for Omphalodeae oscillated from  $r = 0.13$  (minimum stem age, minimum extinction rate:  $\epsilon = 0.0$ ) to  $r = 0.05$  (minimum stem age, maximum extinction rate:  $\epsilon = 0.9$ ). Likewise, at the minimum crown age of Omphalodeae, diversification rates varied from  $r = 0.17$  ( $\epsilon = 0.0$ ) to  $r = 0.07$  ( $\epsilon = 0.9$ ). Diversification rates for each of the four main subclades of Omphalodeae ranged between  $r = 0.12$ - $0.21$  ( $\epsilon = 0.0$ ) and  $r = 0.02$ - $0.06$  ( $\epsilon = 0.9$ ) for the stem age, and  $r = 0.25$ - $0.88$  ( $\epsilon = 0.0$ ) and  $r = 0.07$ - $0.18$  ( $\epsilon = 0.9$ ) for the crown age. The minimum differences between stem and crown ages were found for *Mimophytum*, which presented the highest diversification rate for the stem age at minimum and maximum age ranges within HPD, and for both scenarios of extinction rates ( $\epsilon = 0.0$ ,  $\epsilon = 0.9$ ; Table 3). In contrast, *Iberodes* presented the highest difference in diversification rates between stem and crown and obtained the highest diversification rate for the crown age at minimum and

**Table 3** Diversification rates of the generic lineages of Omphalodeae and diversification rates differences between stem and crown ages of Beast analysis using different rates of extinction ( $k = 0$ , no extinction; and  $k = 0.9$ , high rate of extinction; where  $k = \text{speciation rate}/\text{extinction rate}$ ). Highest diversification rates and highest differences between stem and crown ages are marked in bold. † indicates diversification rates obtained from the constrained MCC tree.

	Stem age				Crown age				Stem age vs. Crown age			
	min		max		min		max		min		max	
	$k = 0$	$k = 0.9$	$k = 0$	$k = 0.9$	$k = 0$	$k = 0.9$	$k = 0$	$k = 0.9$	$k = 0$	$k = 0.9$	$k = 0$	$k = 0.9$
Omphalodeae	0.133	0.054	0.083	0.034	0.170	0.069	0.095	0.038	-0.037	-0.015	-0.012	-0.004
<i>Iberodes</i>	0.118	0.025	0.058	0.012	<b>0.877</b>	<b>0.183</b>	<b>0.253</b>	<b>0.053</b>	<b>-0.759</b>	<b>-0.158</b>	<b>-0.195</b>	<b>-0.041</b>
<i>Omphalodes</i>	0.169	0.047	0.083	0.023	0.463	0.129	0.158	0.044	-0.294	-0.082	-0.075	-0.021
<i>Selkirkia-Myosotidium</i>	0.144	0.030	0.070	0.015	0.345	0.072	0.070	0.020	-0.202	-0.042	0	-0.005
†	0.133	0.027	0.066	0.013	0.378	0.079	0.105	0.022	-0.245	-0.052	-0.039	-0.009
<i>Mimophytum</i>	<b>0.214</b>	<b>0.062</b>	<b>0.103</b>	<b>0.030</b>	0.247	0.071	0.118	0.034	-0.033	-0.009	-0.015	-0.004

maximum age ranges, and for both scenarios of extinction rates (Table 3).

### 3.5 Ancestral state reconstruction of LDD syndromes

The ancestral state reconstruction analyses indicated that the most probable LDD syndrome for the ancestor of Omphalodeae is epizoochory ( $p = 0.58$ , Fig. 2). Epizoochorous traits were maintained for all main clades and lineages except for *Omphalodes* that changed to an unspecialized ancestor ( $p = 0.85$ ) during the Miocene. Epizoochorous traits are the ancestral state for *Iberodes* ( $p = 0.76$ ), *Selkirkia-Myosotidium* ( $p = 0.80$  according to the unconstrained topology and  $p = 0.84$  according to the constrained topology) and *Mimophytum* although with low probability ( $p = 0.48$ ). Four changes from epizoochorous to anemochorous traits are observed at four terminal lineages (*I. commutata*, *Myosotidium hortensia*, *Mimophytum alienum* and *Mimophytum erectum*). Likewise, five changes from epizoochory to unspecialized syndrome were found on five lineages (*I. linifolia*, *Mimophytum chiangii*, *M. mexicanum*, *M. carranzae*, and *M. australe*). Interestingly, high probability for two reversions from unspecialized to epizoochorous traits has also been inferred within *Omphalodes* (*O. nitida* and *O. caucasica*).

## 4 Discussion

This study found a strong phylogenetic signal for colonization of distant territories via LDD (Madrean-Tethyan, amphitropical and trans-Pacific) in agreement with the general trend inferred for other groups of Boraginaceae and even for other families within Boraginales (Luebert et al., 2017). Posterior geographic isolation led to differentiation into new genera and species. Considerable differences of *in situ* diversification are observed within genera, probably due to historical conditions at a regional scale. The phylogenetic topology of Omphalodeae is congruent with previous studies (Holstein et al., 2016a, 2016b) except for the monophyly of *Selkirkia* that is not robustly supported in our study. The general patterns of colonization and speciation are congruent irrespective of phylogenetic methodologies used.

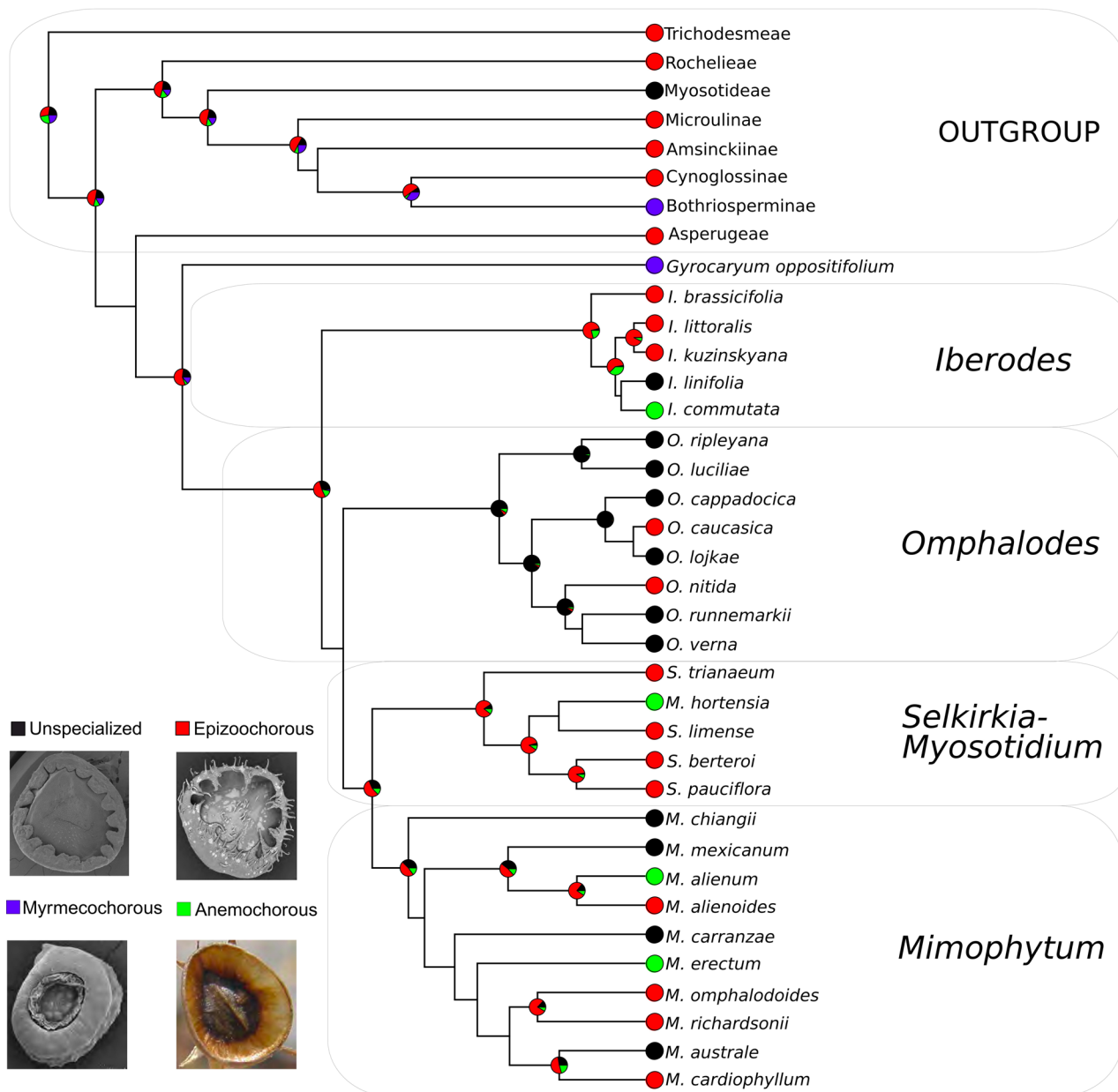
### 4.1 Ancient LDD accounts for historical colonization of North and South America

Our analyses revealed that early diversification of Omphalodeae took place in the Old World (W Mediterranean) during mid-Oligocene (node 1, Fig. 1). An expansion of Western-to-Eastern Mediterranean colonization appears to have occurred later on (Oligocene-Miocene boundary, node 2, Fig. 1). This plant migration may reflect suitable colonization conditions inasmuch as similar patterns of eastward colonization across the Mediterranean Basin have been inferred during the Miocene (*Odontites*, late Miocene, Gaudeul et al., 2016; *Narcissus*, early Miocene, Santos-Gally et al., 2012; Hyacinthaceae, early Miocene, Buerki et al., 2012; *Carex* sect. *Rhynchocystis*, late Miocene, Míguez et al., 2010).

The colonization of the New World appears to have also occurred in the Oligocene-Miocene boundary. Our DIVA analyses inferred dispersal events from The Mediterranean to Sierra Madre (North America) in the Oligocene-Miocene boundary and posterior colonization of South America later on early Miocene (nodes 2, 3; Fig. 1). Diversification of *Mimophytum* at Sierra Madre (North America) and *Selkirkia* at South America is inferred from mid Miocene to early Pliocene (node 3; Fig. 1). The colonization of remote Chatham islands (New Zealand) from South America is inferred during late Miocene (Fig. 1). The morphological differentiation of *Mimophytum*, *Selkirkia*, and *Myosotidium* coincides when gradual climate cooling at high latitudes contributed to the expansion of temperate rather than tropical floristic elements (Zachos et al., 1997; Viruel et al., 2016). The previously highlighted importance of transoceanic dispersals in Boraginaceae (e.g., Winkworth et al., 2002; Chacón et al., 2016; Luebert et al., 2017) is here described by the early colonization of Sierra Madre (Mexico) from the Mediterranean Basin, which contributes to the Madrean-Thethyan historical connectivity that is relatively well studied in other Mediterranean clades (Wen & Ickert-Bond, 2009; Kadereit & Baldwin, 2012; Vargas et al., 2014). The origin of this disjunction has been traditionally explained as an ancient migration across North Atlantic Ocean in the late Eocene when the distance between North America and Western Europe was narrow and thus, facilitated a continuous belt (Axelrod, 1975). In addition, the LDD between distant Mediterranean and Sierra Madre regions in more recent times (mainly Miocene) has been also



## Diaspore syndrome



**Fig. 2.** Ancestral character state reconstruction of diaspore syndromes using stochastic mapping (SIMMAP) with the unconstrained BEAST tree. Color legend and representative nutlet images for each state are provided.

hypothesized to explain this pattern (Raven & Axelrod, 1975). Given that the date of Madrean-Thethyan disjunction in Omphalodeae (Oligocene-Miocene boundary) is posterior to the disappearance of the North Atlantic Bridge (late Palaeocene; Tiffney, 1985; Brikiatis, 2014), the hypothesis of more recent LDD is more plausible. Indeed, there are other well-known examples of LDD from Old to New World even during Quaternary, such as the species of genus *Oligomeris* (Martín-Bravo et al., 2009).

Diversification rates in the Old and New Worlds are shown to be dissimilar. Both Old and New World generic

lineages of Omphalodeae share a common history of ancient origin and early disjunction with lineage-dependent geographic isolation, but also display very different diversification patterns (Fig. 1). On the one hand, the diversification pattern of the American lineages started close to their origin in early-mid Miocene and kept relatively constant until the Pliocene (smallest difference between its crown and stem diversification rates, Table 3). On the other hand, the Mediterranean lineages revealed large diversification gaps (largest diversification rate differences between crown and stem, Table 3), with no cladogenetic events observed

until late Miocene (E Mediterranean, *Omphalodes*; Fig. 1) or Pliocene (W Mediterranean, *Iberodes*; Fig. 1) onwards, and even without any cladogenetic event to the present (W Mediterranean, *Gyrocarium*; Fig. 1). In this sense, the fact that the monotypic *Gyrocarium* originated in the Oligocene while the more diversified *Iberodes* and *Omphalodes* originated in the Miocene is congruent with a recently uncovered pattern of an older origin for species-poor lineages than species-rich lineages in the Mediterranean (Fernández-Mazuecos et al., 2014, 2016; Vargas et al., 2018). In particular, the Pliocene diversification observed in the *Iberodes* clade after a long gap of absence of cladogenesis coincides with a common pattern of diversification detected in the Mediterranean biota in which the effect of the onset of Mediterranean climate on lineages diversification can either be explained by an adaptive radiation, Miocene stasis of diversification or mass extinction effect (Fiz-Palacios & Valcárcel, 2013). Conversely, ancient diversification in the Sierra Madre Oriental (from 20 to 10 mya) and prevalence of *Mimophytum* species up to present times is congruent with the role of the Sierra Madre Oriental, especially the highlands, as a reservoir of biodiversity and high-elevation refugia throughout interglacial periods (Mastretta-Yanes et al., 2015). Indeed, a similar diversification pattern is inferred for Californian (Mediterranean-type) clades of other tribes of Cynoglossoideae such as Amsinckiinae (unlike more recent diversification ages for South American clades, ages for North American clades are from early-to-mid Miocene, see Guilliams et al., 2017).

#### 4.2 Recent LDD in the colonization of Juan Fernández and New Zealand

In most recent times, plant expansion across South America followed by *in situ* differentiation took place in the New World Omphalodeae (Fig. 1, node 3). The colonization of northern South America (node 3, Fig. 1) was most likely from the Madrean region during early Miocene (ca. 17 mya) and via LDD since the closure of the Panama Isthmus was not fully completed until 8 mya (Coates et al., 2003; Bacon et al., 2015; Montes et al., 2015). The existence of island arcs between the two continents at around 15 mya (Krzywinski et al., 2001) could have facilitated this LDD. Indeed, biotic exchange for numerous plant lineages has been invoked before complete closure (e.g., Chloranthaceae, Antonelli & Sanmartin, 2011; Valerianaceae, Bell & Donoghue, 2005; *Ephedra*, Ickert-Bond et al., 2009; *Begonia*, Moonlight et al., 2011; Piperaceae, Symmank et al., 2011). Colonization of Chile via LDD from the northern Andes led to the diversification of Chilean *Selkirkia* s.l. lineage and occurred after the uplift of the eastern cordillera of the northern Andes that favored biotic connections southwards (Ayers, 1999; Weigend, 2002; Hofreiter & Rodríguez, 2006; Antonelli et al., 2009; Viruel et al., 2016). The sister relationship between clades of *Mimophytum* and *Selkirkia* can be considered as one more example of amphitropical disjunction within Boraginaceae. Indeed, Boraginaceae and, particularly the tribe Amsinckiinae, is revealed as one of the maximum exponents for American Amphitropical Disjunction (AAD; Guilliams et al., 2017) among vascular plants. Most relevant characteristics of AAD in vascular plants are shared by AAD in Omphalodeae, such as the LDD as most likely explanation, the presence of adhesion

structures (epizoochorous traits), the north-to-south event directionality, and the timing for the divergence (early-mid Miocene, 18-12 mya) (Schenk & Saunders, 2017; Simpson et al., 2017).

Colonization of remote islands by the South American clade of Omphalodeae includes two LDD events in relatively recent times. The first colonization entails a transoceanic dispersal of the ancestor of *Myosotidium* that colonized Zealandia (Chatham Islands) from South America. The divergence of *Myosotidium*, either as sister to *Selkirkia* (Figs. S5, S6) or nested within it during Miocene-Pliocene boundary (Fig. 1), predates the age of re-emergence of Chatham island (1-4 mya; Campbell, 1998). Similar patterns of divergence times before the re-emergence are inferred for other disjunct organisms in Chatham Islands (Liggins et al., 2008). In our case, LDD by stepping stone through intermediate islands, through the colonization of other non-submerged parts of Zealandia (see Heads, 2017), or even through climatically suitable areas of the Antarctic continent at that moment (Lewis et al., 2008) could be invoked to explain this pattern. Either considering the topology where *Myosotidium* is nested within *Selkirkia* or an alternative topology where the two genera are reciprocally sister lineages, the geological epoch (Miocene) inferred for this transpacific colonization would be widely in agreement with other plant lineages where relatively recent LDD challenged the traditional hypothesis of Gondwana break for this trans-Pacific disjunction pattern (Spalik et al., 2010). The second LDD entails the colonization of the Juan Fernández archipelago from Chile during early Pliocene, which is congruent with the estimated origin of most ancient islands (4.23-5.8 mya, Stuessy et al., 1984). In particular, the endemic *S. berteroi* (Juan Fernández islands) is the sister species of the Chilean *Selkirkia pauciflora*, which diverged from each other in the Pliocene (4.74 mya). Most endemic species belong to species-poor clades in the Juan Fernández archipelago. Single species are predominant within both endemic genera (*Nothomyrcia*, *Lactoris*) and widespread genera in the continent (e.g., *Erigeron*, *Berberis*, *Chenopodium*) (Landrum, 1999; Murillo-Aldana & Ruiz, 2017), which suggests anagenesis as a predominant mechanism in speciation of insular plants. Indeed, the first direct comparison of the genetic consequences of anagenetic (vs. cladogenetic) speciation in insular plants was found in *Robinsonia* (Takayama et al., 2015). Likewise, *Selkirkia berteroi* would be one more example of anagenetic event that reinforces this common speciation pattern on islands first proposed by Stuessy et al. (1990).

#### 4.3 The role of epizoochorous traits in the evolution of Omphalodeae

Epizoochory has been revealed as one of the key traits of biogeographic history of Omphalodeae. Indeed, epizoochorous traits are ancestral in both Omphalodeae and three of the four generic clades of the tribe (Fig. 2).

Dispersal within continents has been mainly related to epizoochory by large mammals, whereas colonizing remote places (especially oceanic islands) has been traditionally linked to endozoochory via birds (Ridley, 1930; Van der Pijl, 1982). Nevertheless, recent studies have found effective adhesive LDD of epizoochorous diaspores through avian

vectors as well (Aoyama et al., 2012). Indeed, epizoochorous traits are considered to be more efficient for LDD since the mechanism to transport the diaspore (externally attached to the animal) allows longer retention times, which favors migration over greater distances than by endozoochory (Sorensen, 1986). In contrast, dispersal distance by endozoochory takes fewer days, i.e., the time needed for transport through an animal's gut. Furthermore, during early divergence of Omphalodeae, epizoochorous traits were only lost for the Eastern Mediterranean lineage but maintained as ancestral specializations for those lineages that colonized other distant places via LDD, including the two transoceanic LDD dispersals. Interestingly, a shift from epizoochorous to anemochorous traits is inferred after the transoceanic colonization of the distant Chatham Islands (New Zealand) from South America (Figs. 2, S7). However, we cannot discount the possibility that the shift from epizoochory to anemochory occurred prior to long distance dispersal. The fact of losing the efficient anchoring structures (glochids) could be traditionally explained for many plant lineages (Carlquist, 1972) by the loss of dispersal ability hypothesis that predicts that islands are initially colonized by individuals with relatively good powers of dispersal, but through time, island populations evolve poorer disperser ability to avoid being swept out to sea (see Cody & Overton, 1996). Indeed, character shifts after colonization can occur in few generations (Cody & Overton, 1996). However, the interpretation of this loss of potential dispersal has been challenged by some recent studies that argue a loss of dispersal ability is unnecessarily linked to the loss of dispersal ability hypothesis (Burns, 2018). In particular, in *Myosotidium* besides the development of a soft and open nutlet margin, there has been a notably increasing of nutlet size, much larger than its relatives (Holstein et al., 2016a). Thus, the loss of dispersability could be determined by the increase in seed size instead of the loss of LDD traits, as observed in other plant examples (Fresnillo & Ehlers, 2008). Decrease of dispersal potential could be a consequence of selective pressures for increased seed size more advantageous at early stages of colonization that are unrelated to seed dispersal distances (Burns, 2018). The majority of the flora of New Zealand comes from transoceanic migrations (Winkworth et al., 2002) and character shifts in diaspore specializations from the original source is particularly common (Thorsen et al., 2009).

## Acknowledgements

We would like to thank those botanists and colleagues that help us with material sampling and particularly, Diana Íñigo, Paola Pérez, Celia Otero, David Vallecillo, Yurena Arjona, Lua López, Joaquín Ramírez, George Hinton, Javier Morente, Manuel Joao Pinto, Ruben Retuerto, Miguel Serrano, Antonio Rivas, Claude Dauge, Asli Koca, Peter Heenan, Enrique Rico, Ester Vega, Santiago Martín-Bravo, Íñigo Pulgar, and Juan Carlos Zamora. Likewise, we would also appreciate the worthy provided loans from numerous herbaria mentioned at supplementary tables. We gratefully acknowledge the work of the lab technicians of Real Jardín Botánico of Madrid, Yolanda Ruiz, Emilio Cano, Jose González, Olga Popova, and Lucia Sastre. We are grateful for the many

methodological comments of Isabel Sanmartín and Mario Fernández-Mazuecos. This study was supported by the Fundación General CSIC and Banco Santander as part of the project titled “Do all endangered species hold the same value?: origin and conservation of living fossils of flowering plants endemic to Spain.”

## References

- Antonelli A, Nylander JAA, Persson C, Sanmartín I. 2009. Tracing the impact of the Andean uplift on Neotropical plant evolution. *Proceedings of the National Academy of Sciences USA* 6: 9749–9754.
- Antonelli A, Sanmartín I. 2011. Mass extinction, gradual cooling, or rapid radiation? Reconstructing the spatiotemporal evolution of the ancient Angiosperm genus *Hedyosmum* (Chloranthaceae) using empirical and simulated approaches. *Systematic Biology* 60: 596–615.
- Aoyama Y, Kawakami K, Chiba S. 2012. Seabirds as adhesive seed dispersers of alien and native plants in the oceanic Ogasawara islands, Japan. *Biodiversity and Conservation* 21: 2787–2801.
- 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: 805–814.
- Axelrod DI. 1975. Evolution and biogeography of Madrean-Tethyan sclerophyll vegetation. *Annals of the Missouri Botanical Garden* 62: 280–334.
- Ayers T. 1999. Biogeography of *lysipomia* (Campanulaceae), a high elevation endemic: An illustration of species richness at the Huancabamba depression, Peru. *Arnaldoa* 6: 13–27.
- Bacon CD, Silvestro D, Jaramillo C, Smith BT, Chakrabarty P, Antonelli A. 2015. Biological evidence supports an early and complex emergence of the Isthmus of Panama. *Proceedings of the National Academy of Sciences USA* 112: 6110–6115.
- Beaulieu JM, Tank DC, Donoghue MJ. 2013. A southern hemisphere origin for Campanulid angiosperms, with traces of the break-up of Gondwana. *BMC Evolutionary Biology* 13: 80.
- Bell C, Donoghue MJ. 2005. Phylogeny and biogeography of Valerianaceae (Dipsacales) with special reference to the South American Valerians. *Organisms Diversity & Evolution* 5: 147–159.
- Bergsten J, Nilsson AN, Ronquist F. 2013. Bayesian tests of topology hypotheses with an example from diving beetles. *Systematic Biology* 62: 660–673.
- Blanco-Pastor JL, Vargas P, Pfeil BE. 2012. Coalescent simulations reveal hybridization and incomplete lineage sorting in Mediterranean *Linaria*. *PLoS ONE* 7: e39089.
- Bremer K. 2002. Gondwanan evolution of the grass alliance of families (Poales). *Evolution* 56: 1374–1387.
- Brikiatis L. 2014. The De Geer, Thulean and Beringia routes: Key concepts for understanding early cenozoic biogeography. *Journal of Biogeography* 41: 1036–1054.
- Buerki S, Jose S, Yadav SR, Goldblatt P, Manning JC, Forest F. 2012. Contrasting biogeographic and diversification patterns in two Mediterranean-type ecosystems. *PLoS ONE* 7: e39377.
- Burns KC. 2018. Time to abandon the loss of dispersal ability hypothesis in island plants: A comment on García-Verdugo, Mairal, Monroy, Sajeve and Caujapé-Castells (2017). *Journal of Biogeography* 45: 1219–1222.
- Campbell H. 1998. Fauna and flora of the Chatham Islands: Less than 4 my old? *Geology and Genes* 97: 15–16.

- Carlquist S. 1972. Island biology: We've only just begun. *Bioscience* 22: 221–225.
- Chacón J, Luebert F, Hilger HH, Ovchinnikova S, Selvi F, Cecchi L, Williams CM, Hasenstab-Lehman K, Sutóry K, Simpson MG, Weigend M. 2016. The Borage family (Boraginaceae s.str.): A revised infrafamilial classification based on new phylogenetic evidence, with emphasis on the placement of some enigmatic genera. *Taxon* 65: 523–546.
- Chacón J, Luebert F, Weigend M. 2017. Biogeographic events are not correlated with diaspore dispersal modes in Boraginaceae. *Frontiers in Ecology and Evolution* 5: 26.
- Chen LY, Zhao SY, Mao KS, Les DH, Wang QF, Moody ML. 2014. Historical biogeography of Haloragaceae: An out-of-Australia hypothesis with multiple intercontinental dispersals. *Molecular Phylogenetics and Evolution* 78: 87–95.
- Coates AG, Aubry M-P, Berggren WA, Collins LS, Kunk M. 2003. Early Neogene history of the Central American arc from Bocas del Toro, western Panama. *Geological Society of America Bulletin* 115: 271–287.
- Cody ML, Overton JM. 1996. Short-term evolution of reduced dispersal in island plant populations. *Journal of Ecology* 84: 53–61.
- Cohen JL. 2013. A phylogenetic analysis of morphological and molecular characters of Boraginaceae: evolutionary relationships, taxonomy, and patterns of character evolution. *Cladistics* 30: 1–31.
- Crisp MD, Arroyo MT, Cook LG, Gandolfo MA, Jordan GJ, McGlone MS, Weston PH, Westoby M, Wilf P, Linder HP. 2009. Phylogenetic biome conservatism on a global scale. *Nature* 458: 754–756.
- Doyle J, Doyle J. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemistry Bulletin* 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.
- Elias SA, Short SK, Nelson CH, Birks HH. 1996. Life and times of the Bering Land Bridge. *Nature* 382: 60.
- Feliner GN. 2014. Patterns and processes in plant phylogeography in the Mediterranean Basin. A review. *Perspectives in Plant Ecology, Evolution and Systematics* 16: 265–278.
- Fernández-Mazuecos M, Jiménez-Mejías P, Martín-Bravo S, Buide M, Álvarez I, Vargas P. 2016. Narrow endemics on coastal plains: Miocene divergence of the critically endangered genus *Avellara* (Compositae). *Plant Biology* 18: 729–738.
- Fernández-Mazuecos M, Jiménez-Mejías P, Rotllan-Puig X, Vargas P. 2014. Narrow endemics to Mediterranean islands: Moderate genetic diversity but narrow climatic niche of the ancient, critically endangered *Naufraga* (Apiaceae). *Perspectives in Plant Ecology, Evolution and Systematics* 16: 190–202.
- Fiz-Palacios O, Valcárcel V. 2013. From Messinian crisis to Mediterranean climate: A temporal gap of diversification recovered from multiple plant phylogenies. *Perspectives in Plant Ecology, Evolution and Systematics* 15: 130–137.
- Fresnillo B, Ehlers B. 2008. Variation in dispersability among mainland and island populations of three wind dispersed plant species. *Plant Systematics and Evolution* 270: 243–255.
- Gabel ML. 1987. A fossil *Lithospermum* (Boraginaceae) from the tertiary of South Dakota. *American Journal of Botany* 74: 1690–1693.
- Gabel ML, Backlund DC, Haffner J. 1998. The Miocene macroflora of the northern Ogallala group, northern Nebraska and southern South Dakota. *Journal of Paleontology* 72: 388–397.
- Gaudeul M, Vela E, Rouhan G. 2016. Eastward colonization of the Mediterranean Basin by two geographically structured clades: The case of *Odontites* Ludw. (Orobanchaceae). *Molecular Phylogenetics and Evolution* 96: 140–149.
- Gillespie RG, Baldwin BG, Waters JM, Fraser CI, Nikula R, Roderick GK. 2012. Long-distance dispersal: A framework for hypothesis testing. *Trends in Ecology and Evolution* 27: 47–56.
- Guilliams CM, Hasenstab-Lehman KE, Mabry ME, Simpson MG. 2017. Memoirs of a frequent flier: Phylogenomics reveals 18 long-distance dispersals between North America and South America in the popcorn flowers (Asteraceae). *American Journal of Botany* 104: 1717–1728.
- Hamilton M. 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology* 8: 521–523.
- Hammouda SA, Weigend M, Mebrouk F, Chacón J, Bensalah M, Ensikat HJ, Adaci M. 2015. Fossil nutlets of Boraginaceae from the continental Eocene of Hamada of Méridja (southwestern Algeria): The first fossil of the borage family in Africa. *American Journal of Botany* 102: 2108–2115.
- Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W. 2007. Geiger: Investigating evolutionary radiations. *Bioinformatics* 24: 129–131.
- Harris AJ, Ickert-Bond S, Rodríguez A. 2018. Long distance dispersal in the assembly of floras: A review of progress and prospects in North America. *Journal of Systematics and Evolution* 56: 430–448.
- Heads MJ. 2017. *Biogeography and evolution in New Zealand*. Boca Raton, Florida: Taylor & Francis.
- Heleno R, Vargas P. 2015. How do islands become green? *Global Ecology and Biogeography* 24: 518–526.
- Ho SY, Phillips MJ. 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology* 58: 367–380.
- Hofreiter A, Rodríguez E. 2006. The Alstroemeriaceae in Peru and neighbouring areas. *Revista peruana de biología* 13: 5–69.
- Holstein N, Chacon J, Hilger HH, Weigend M. 2016a. No longer shipwrecked—*Selkirkia* (Boraginaceae) back on the mainland with generic rearrangements in South American “*Omphalodes*” based on molecular data. *Phytotaxa* 270: 231–251.
- Holstein N, Chacon J, Otero A, Jiménez-mejías P, Weigend M. 2016b. Towards a monophyletic *Omphalodes*—or an expansion of North American *Mimophytum*. *Phytotaxa* 288: 131–144.
- Ickert-Bond SM, Rydin C, Renner SS. 2009. A fossil-calibrated relaxed clock for *Ephedra* indicates an Oligocene age for the divergence of Asian and New World clades and Miocene dispersal into South America. *Journal of Systematics and Evolution* 47: 444–456.
- Inda LA, Sanmartín I, Buerki S, Catalán P. 2014. Mediterranean origin and Miocene–Holocene Old World diversification of meadow fescues and ryegrasses (*Festuca* subgenus *Schedonorus* and *Lolium*). *Journal of Biogeography* 41: 600–614.
- Jobb G. 2007. Treefinder version of november 2007. Available from <http://www.treefinder.de/>.
- Jobb G, VonHaeseler A, Strimmer K. 2004. Retracted article: Treefinder: A powerful graphical analysis environment for molecular phylogenetics. *BMC Evolutionary Biology* 4: 18.
- Kadereit JW, Baldwin BG. 2012. Western Eurasian–western North American disjunct plant taxa: The dry-adapted ends of formerly widespread north temperate mesic lineages—and examples of long-distance dispersal. *Taxon* 61: 3–17.
- Kass RE, Raftery AE. 1995. Bayes factors. *Journal of the American Statistical Association* 90: 773–795.

- Katoh K, Misawa K, Kuma Ki, Miyata T. 2002. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* 30: 3059–3066.
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C. 2012. Geneious basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28: 1647–1649.
- Krzywinski J, Wilkerson RC, Besansky NJ. 2001. Toward understanding Anophelinae (Diptera, Culicidae) phylogeny: Insights from nuclear single-copy genes and the weight of evidence. *Systematic Biology* 50: 540–556.
- Landrum LR. 1999. Revision of *Berberis* (Berberidaceae) in Chile and adjacent southern Argentina. *Annals of the Missouri Botanical Garden* 86: 793–834.
- Lewis AR, Marchant DR, Ashworth AC, Hedenäs L, Hemming SR, Johnson JV, Leng MJ, Machlus ML, Newton J, Raine I, Willenbring JK, Williams M, Wolfe AP. 2008. Mid-Miocene cooling and the extinction of tundra in continental Antarctica. *Proceedings of the National Academy of Sciences USA* 105: 10676–10680.
- Liggins L, Chapple DG, Daugherty CH, Ritchie PA. 2008. Origin and post-colonization evolution of the Chatham Islands skink (*Oligosoma nigriplantare nigriplantare*). *Molecular Ecology* 17: 3290–3305.
- Luebert F, Couvreur TLP, Gottschling M, Hilger HH, Miller JS, Weigend M. 2017. Historical biogeography of Boraginales: West Gondwanan vicariance followed by long-distance dispersal? *Journal of Biogeography* 44: 158–169.
- Magallon S, Sanderson MJ. 2001. Absolute diversification rates in Angiosperm clades. *Evolution* 55: 1762–1780.
- Martín-Bravo S, Vargas P, Luceño M. 2009. Is *Oligomeris* (Reseda-ceae) indigenous to North America? Molecular evidence for a natural colonization from the Old World. *American Journal of Botany* 96: 507–518.
- Mastretta-Yanes A, Moreno-Letelier A, Piñero D, Jorgensen TH, Emerson BC. 2015. Biodiversity in the Mexican highlands and the interaction of geology, geography and climate within the Trans-Mexican volcanic belt. *Journal of Biogeography* 42: 1586–1600.
- Matzke NJ. 2018. BioGeoBEARS: BioGeography with Bayesian (and likelihood) Evolutionary Analysis with R Scripts. version 1.1.1, published on GitHub on November 6, 2018. doi: <http://dx.doi.org/10.5281/zenodo.1478250>.
- Míguez M, Gehrke B, Maguilla E, Jiménez-Mejías P, Martín-Bravo S. 2017. *Carex* sect. *Rhynchocystis* (Cyperaceae): A Miocene subtropical relict in the Western Palearctic showing a dispersal-derived rand flora pattern. *Journal of Biogeography* 44: 2211–2224.
- Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *Proceedings of the Gateway Computing Environments Workshop, New Orleans, Louisiana, USA*. 1–8.
- Montes C, Cardona A, Jaramillo C, Pardo A, Silva J, Valencia V, Ayala C, Pérez-Angel L, Rodríguez-Parra L, Ramirez V, Niño H. 2015. Middle Miocene closure of the Central American seaway. *Science* 348: 226–228.
- Moonlight PW, Richardson JE, Tebbitt MC, Thomas DC, Hollands R, Peng C-I, Hughes M, Silman M. 2015. Continental-scale diversification patterns in a megadiverse genus: The biogeography of neotropical *Begonia*. *Journal of Biogeography* 42: 1137–1149.
- Murillo-Aldana J, Ruiz E. 2011. Revalidación de *Nothomyrcia* (Myrtaceae), un género endémico del archipiélago de Juan Fernández. *Gayana Botánica* 68: 129–134.
- Nesom G. 2013. A third species of *Mimophytum* s. str. and three new species of *Omphalodes* (Boraginaceae) from North America. *Phytoneuron* 64: 1–23.
- Otero A, Jiménez-Mejías P, Valcárcel V, Vargas P. 2014. Molecular phylogenetics and morphology support two new genera (*Memoremea* and *Nihon*) of Boraginaceae s.s. *Phytotaxa* 173: 241–277.
- Otero A, Jiménez-Mejías P, Valcárcel V, Vargas P. 2019. Being in the right place at the right moment? Parallel diversification bursts related to persistence of ancient epizoochorous traits and hidden factors in Cynoglossoideae. *American Journal of Botany* 106: 1–15.
- Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20: 289–290.
- Posada D. 2008. Jmodeltest: Phylogenetic model averaging. *Molecular Biology and Evolution* 25: 1253–1256.
- R Development Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from <http://www.R-project.org/>.
- Rambaut A, Suchard M, Xie D, Drummond A. 2014. Tracer v1. 6. Available from <http://tree.bio.ed.ac.uk/software/tracer/>.
- Raven PH, Axelrod DI. 1975. History of the flora and fauna of Latin America: The theory of plate tectonics provides a basis for reinterpreting the origins and distribution of the biota. *American Scientist* 63: 420–429.
- Ree R, Sanmartín I. 2018. Conceptual and statistical problems with the dec+j model of founder-event speciation and its comparison with dec via model selection. *Journal of Biogeography* 45: 741–749.
- Ree R, Smith S. 2008. Lagrange: Software for likelihood analysis of geographic range evolution. *Systematic Biology* 57: 4–14.
- Renner S, Foreman D, Murray D. 2000. Timing transantarctic disjunctions in the Atherospermataceae (Laurales): Evidence from coding and noncoding chloroplast sequences. *Systematic Biology* 49: 579–591.
- Revell LJ. 2012. Phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3: 217–223.
- Ridley HN 1930. *The dispersal of plants throughout the world*. Ashford, Kent: L. Reeve & Company, Limited.
- Ronquist F. 1997. Dispersal-vicariance analysis: A new approach to the quantification of historical biogeography. *Systematic Biology* 46: 195–203.
- Ronquist F, Huelsenbeck J, Teslenko M 2011. Draft MrBayes version 3.2 manual: tutorials and model summaries. Available from [http://mrbayes.sourceforge.net/mb3.2\\_manual.pdf](http://mrbayes.sourceforge.net/mb3.2_manual.pdf)
- Ronquist F, Sanmartín I. 2011. Phylogenetic methods in biogeography. *Annual Review of Ecology, Evolution, and Systematics* 42: 441–464.
- Rundel PW, Arroyo MT, Cowling RM, Keeley JE, Lamont BB, Vargas P. 2016. Mediterranean biomes: Evolution of their vegetation, floras, and climate. *Annual Review of Ecology, Evolution, and Systematics* 47: 383–407.
- Sanmartín I, Ronquist F. 2004. Southern hemisphere biogeography inferred by event-based models: Plant versus animal patterns. *Systematic Biology* 53: 216–243.
- Santos-Gally R, Vargas P, Arroyo J. 2012. Insights into Neogene Mediterranean biogeography based on phylogenetic relationships

- of mountain and lowland lineages of *Narcissus* (Amaryllidaceae). *Journal of Biogeography* 39: 782–798.
- Schenk JJ, Saunders K. 2017. Inferring long-distance dispersal modes in American amphitropically disjunct species through adaptive dispersal structures. *American Journal of Botany* 104: 1756–1764.
- Segal R. 1966. Taxonomic study of the fossil species of the genus *Cryptantha* (Boraginaceae). *The Southwestern Naturalist* 11: 205–210.
- Shepherd LD, McLay TG. 2011. Two micro-scale protocols for the isolation of DNA from polysaccharide-rich plant tissue. *Journal of Plant Research* 124: 311–314.
- Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. *Systematic Biology* 51: 492–508.
- Simpson MG, Johnson LA, Villaverde T, Williams CM. 2017. American amphitropical disjuncts: Perspectives from vascular plant analyses and prospects for future research. *American Journal of Botany* 104: 1600–1650.
- Sorensen AE. 1986. Seed dispersal by adhesion. *Annual Review of Ecology, Evolution, and Systematics* 17: 443–463.
- Spalik K, Pivczyński M, Danderson CA, Kurzyńska-Młynik R, Bone TS, Downie SR. 2010. Amphitropic amphiantarctic disjunctions in Apiaceae subfamily Apioideae. *Journal of Biogeography* 37: 1977–1994.
- Stuessy TF, Crawford DJ, Marticorena C. 1990. Patterns of phylogeny in the endemic vascular flora of the Juan Fernandez islands, Chile. *Systematic Botany* 338–346.
- Stuessy TF, Foland KA, Sutter JF, Sanders RW, Mario Silva O. 1984. Botanical and geological significance of potassium-argon dates from the Juan Fernández islands. *Science* 225: 49–51.
- Suchard MA, Weiss RE, Sinsheimer JS. 2001. Bayesian selection of Continuous-Time Markov Chain evolutionary models. *Molecular Biology and Evolution* 18: 1001–1013.
- Symmank L, Samain M-S, Smith JF, Pino G, Stoll A, Goetghebeur P, Neinhuis C, Wanke S. 2011. The extraordinary journey of *Peperomia* subgenus *tildenia* (Piperaceae): Insights into diversification and colonization patterns from its cradle in Peru to the trans-mexican volcanic belt. *Journal of Biogeography* 38: 2337–2349.
- Takayama K, López-Sepúlveda P, Greimler J, Crawford DJ, Peñailillo P, Baeza M, Ruiz E, Kohl G, Tremetsberger K, Gatica A. 2015. Relationships and genetic consequences of contrasting modes of speciation among endemic species of *Robinsonia* (Asteraceae, Senecioneae) of the Juan Fernández archipelago, Chile, based on AFLPs and SSRs. *New Phytologist* 205: 415–428.
- Thomasson JR. 1987. Late Miocene plants from northeastern Nebraska. *Journal of Paleontology* 61: 1065–1079.
- Thompson JD. 2005. *Plant evolution in the Mediterranean*. New York: Oxford University Press.
- Thornhill AH, Ho SY, Kulheim C, Crisp MD. 2015. Interpreting the modern distribution of Myrtaceae using a dated molecular phylogeny. *Molecular Phylogenetics Evolution* 93: 29–43.
- Thorsen MJ, Dickinson KJM, Seddon PJ. 2009. Seed dispersal systems in the New Zealand flora. *Perspectives in Plant Ecology, Evolution and Systematics* 11: 285–309.
- Tiffney BH. 1985. The Eocene North Atlantic land bridge: its importance in Tertiary and modern phytogeography of the Northern Hemisphere. *Journal of the Arnold Arboretum* 66: 243–273.
- Tiffney BH, Manchester SR. 2001. The use of geological and paleontological evidence in evaluating plant phylogeographic hypotheses in the Northern Hemisphere Tertiary. *International Journal of Plant Sciences* 162: S3–S17.
- Van der Pijl L. 1982. *Principles of dispersal*. Berlin: Springer.
- Vargas P, Fernandez-Mazuecos M, Heleno R. 2018. Phylogenetic evidence for a Miocene origin of Mediterranean lineages: Species diversity, reproductive traits and geographical isolation. *Plant Biology* 20(Suppl. 1): 157–165.
- Vargas P, Heleno R, Traveset A, Nogales M. 2012. Colonization of the Galápagos islands by plants with no specific syndromes for long-distance dispersal: A new perspective. *Ecography* 35: 33–43.
- Vargas P, Valente LM, Blanco-Pastor JL, Liberal I, Guzmán B, Cano E, Forrest A, Fernández-Mazuecos M. 2014. Testing the biogeographical congruence of palaeofloras using molecular phylogenetics: Snapdragons and the Madrean-Tethyan flora. *Journal of Biogeography* 41: 932–943.
- Viruel J, Segarra-Moragues JG, Raz L, Forest F, Wilkin P, Sanmartín I, Catalán P. 2016. Late Cretaceous-early Eocene origin of yams (*Dioscorea*, Dioscoreaceae) in the Laurasian Palaeartic and their subsequent Oligocene-Miocene diversification. *Journal of Biogeography* 43: 750–762.
- Weigend M. 2002. Observations on the biogeography of the Amotape-Huancabamba zone in Northern Peru. *The Botanical Review* 68: 38–54.
- Weigend M, Luebert F, Selvi F, Brokamp G, Hilger HH. 2013. Multiple origins for hound's tongues (*Cynoglossum* L.) and navel seeds (*Omphalodes* Mill.) – the phylogeny of the borage family (Boraginaceae s.str.). *Molecular Phylogenetics and Evolution* 68: 604–618.
- Wen J, Ickert-Bond SM. 2009. Evolution of the Madrean-Tethyan disjunctions and the North and South American amphitropical disjunctions in plants. *Journal of Systematics and Evolution* 47: 331–348.
- Winkworth RC, Wagstaff SJ, Glenny D, Lockhart PJ. 2002. Plant dispersal news from New Zealand. *Trends in Ecology and Evolution* 17: 514–520.
- Wolfe KH, Li W-H, Sharp PM. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences USA* 84: 9054–9058.
- Xie W, Lewis PO, Fan Y, Kuo L, Chen MH. 2011. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Systematic Biology* 60: 150–160.
- Yu Y, Harris AJ, Blair C, He X. 2015. Rasp (reconstruct ancestral state in phylogenies): A tool for historical biogeography. *Molecular Phylogenetics Evolution* 87: 46–49.
- Zachos JC, Flower BP, Paul H. 1997. Orbitally paced climate oscillations across the Oligocene/Miocene boundary. *Nature* 388: 567–570.

#### Appendix I. Taxonomic rearrangement derived from the phylogenetic findings.

There are some taxonomic consequences needed in the light of a more natural classification of Omphalodeae.

The reconstruction of the Southern Mexican endemic *M. australe* nested within the more widespread *M. cardiophyllum* (Nesom, 2013) points to overemphasizing the geographic distribution gap in this group. The minor morphological differences between the two species (Nesom, 2013) have probably been overvalued under a putative morphogeographic compartmentalization scenario (e.g., Míguez et al., 2017), leading to the misconception

tion of *M. australe* as distinct enough to be considered a different species from *M. cardiophyllum*. In any case, before simply lump one species under the other, further studies are needed to confirm the fine taxonomic structure revealed by our phylogeny, given the limited considered sampling.

In addition, *Omphalodes erecta* has been found to be nested within *Mimophytum*, which supports its transfer to this North American endemic genus. Most North American species of *Mimophytum* were previously circumscribed in *Omphalodes* until the rearrangement proposed by Holstein et al. (2016b), who combined the accepted names under *Mimophytum*. However, the deviant morphological characters of *O. erecta* in comparison with other *Mimophytum* (primarily the cuneate leaves and the erect growth of *O. erecta* vs. the basally constricted leaves and the mostly ascending growth of *Mimophytum*) prevented them from addressing the formal nomenclatural changes. Remarkably, they pointed to several morphological affinities that put *O. erecta* closer to *Mimophytum* than to the other Omphalodeae (i.e., spreading pubescence, nutlets with a flat toothed margin), but considered that such evidence should be taken with caution until further molecular data was available. After our phylogenetic results, the transfer of *O. erecta* to *Mimophytum* is fully supported, and we make the following formal combination:

*Mimophytum erectum* (I.M.Johnst.) A.Otero, Jim.Mejías, Valcárcel & P. Vargas, **comb. nov.**

≡ *Omphalodes erecta* I.M.Johnst., J. Arnold Arbor. 16: 204. 1935 [Basionym].

Holotype: México, Nuevo León, Alamar to Taray, ca. 15 mi SW of Galeana, C. H. & M. T. Mueller 992 (GH digital image!; isotype: F digital image!, MICH digital image!, TEX digital image!).

## Supplementary Material

The following supplementary material is available online for this article at <http://onlinelibrary.wiley.com/doi/10.1111/jse.12504/supinfo>:

**Table S1.** List of voucher specimen and GenBank numbers of species of the ingroup. Asterisks mark accessions newly sequenced.

**Table S2.** List of GenBank accessions for the outgroup.

**Table S3.** Dispersal rate scale used for biogeographic schemes M1 and M3.

**Table S4.** Adjacency matrix used for biogeographic schemes M2 and M3.

**Table S5.** Log-likelihood and AIC values obtained for the biogeographic models analyzed under different connectivity schemes (M0-M3) with the MCC from the unconstrained BEAST analysis (H0: *Myosotidium* is nested within *Selkirkia*).

**Table S6.** Log-likelihood and AIC values obtained for the biogeographic models analyzed under different connectivity schemes (M0-M3) with the MCC from the constrained BEAST analysis (H1: *Selkirkia* is sister to *Myosotidium*).

**Fig. S1.** Maximum clade credibility (MCC) tree from the unconstrained BEAST analysis. The 95% Highest posterior density (HPD) bars of divergence time are provided for those nodes with  $\geq 0.90$  Bayesian Posterior Probability (BPP).

**Fig. S2.** Phylogenetic reconstruction of Omphalodeae. (a) Maximum likelihood inference through RaxML. Bootstrap support for each node is provided. (b) Bayesian inference through Mr.Bayes.

**Fig. S3.** Maximum clade credibility (MCC) tree from the constrained BEAST analysis. The 95% Highest posterior density (HPD) bars of divergence time are provided for those nodes with  $\geq 70$  Bootstrap support (BS).

**Fig. S4.** Time-calibrated phylogeny from the unconstrained BEAST analysis and dispersal-extinction-cladogenesis (DEC) reconstruction of Omphalodeae under M1 scheme.

**Fig. S5.** Time-calibrated phylogeny from the constrained BEAST analysis and dispersal-extinction-cladogenesis (DEC) reconstruction of Omphalodeae under M1 scheme.

**Fig. S6.** Time-calibrated phylogeny from the constrained BEAST analysis and dispersal vicariance analysis (DIVA) reconstruction of Omphalodeae under M1 scheme.

**Fig. S7.** Ancestral character state reconstruction of LDD syndrome in Omphalodeae using stochastic mapping (SIMMAP) with the constrained BEAST tree.