

Research Article

Ecological and geological processes impacting speciation modes drive the formation of wide-range disjunctions within tribe Putorieae (Rubiaceae)

Mario Rincón-Barrado^{1,2*} , Sanna Olsson³ , Tamara Villaverde^{1,4} , Belén Moncalvillo⁵, Lisa Pokorny^{1,6} , Alan Forrest⁷ , Ricarda Riina^{1†} , and Isabel Sanmartín^{1*} 

¹Real Jardín Botánico (RJB-CSIC), Madrid 28014, Spain

²Area of Biodiversity and Conservation, School of Experimental Sciences and Technology (ESCET), Universidad Rey Juan Carlos, Móstoles, Madrid 28933, Spain

³Forest Research Centre, INIA-CSIC, Madrid 28040, Spain

⁴Current address: Departamento de Biodiversidad, Ecología y Evolución, Facultad de Ciencias Biológicas, Universidad Complutense de Madrid, C/ José Antonio Novais, 12, Madrid 28040, Spain

⁵Department of Ecology, Faculty of Biology, Philipps-University Marburg, Marburg D-35043, Germany

⁶Current address: Centre for Plant Biotechnology and Genomics (CBGP UPM-INIA), Pozuelo de Alarcón, Madrid 28223, Spain

⁷Centre for Middle Eastern Plants, Royal Botanic Garden Edinburgh, 20a Inverleith Row, Edinburgh EH3 5LR, UK

[†]Ricarda Riina and Isabel Sanmartín are co-senior authors.

*Authors for correspondence. Mario Rincón-Barrado. E-mail: mariorincomb@gmail.com; Isabel Sanmartín. E-mail: isanmartin@rjb.csic.es

Received 1 February 2021; Accepted 31 March 2021; Article first published online 6 April 2021

Abstract Wide-range geographically discontinuous distributions have long intrigued scientists. We explore the role of ecology, geology, and dispersal in the formation of these large-scale disjunctions, using the angiosperm tribe Putorieae (Rubiaceae) as a case study. From DNA sequences of nuclear ITS and six plastid markers, we inferred a phylogeny with 65% of all known Putorieae species. Divergence times, ancestral ranges, and diversification rate shifts were then estimated using Bayesian inference. We further explored species climatic tolerances and performed ancestral niche reconstruction to discriminate among alternative speciation modes, including geographical and ecological vicariance, and ecogeographical, ecological, and dispersal-mediated speciation. As a result, we identified seven major clades in Putorieae, some of which exhibit striking geographical disjunctions, matching the Rand Flora pattern, with sister species in the Canary Islands and eastern and southern Africa. Initial diversification within the tribe occurred in the early Miocene, coincident with a period of climate warming; however, most clades diverged within the last 10 Myr. Aridification and high extinction rates, coupled with ecological vicariance, explain the oldest disjunctions. Adaptation to new environmental conditions, after allopatry, is observed in several clades. Dispersal, either long-distance or via corridors made available by mountain uplift, is behind the most recent disjunctions. Some of these events were followed by ecological speciation and rapid diversification, with species becoming adapted to xeric or increasingly colder continental climates. We show that an integrative approach may help discriminate among speciation modes invoked to explain disjunctions at macroevolutionary time scales, even when extinction has erased the signature of past events.

Key words: climate change, disjunct distribution, ecological vicariance, extinction rates, niche evolution, Rand Flora.

1 Introduction

The Swiss botanist Augustin de Candolle (1834) coined the term “endemic” to refer to geographically or climatically restricted plant groups. He also noted that some seemingly closely related plant groups showed remarkable disjunct distributions, and concluded that

endemic and disjunct species could not expand to other areas due to their strong connection with the environmental conditions of their region of origin (cf. Cox et al., 2016). It was not until the concepts of evolution, plate tectonics, and the molecular clock were introduced into biogeography that time became an essential dimension to understand the origin of disjunct distributions (Axelrod &

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Raven, 1978; Donoghue, 2003). A species might be endemic (i) because it has high environmental specificity and long ago evolved to only live in a few sites (paleoendemic), or it can be endemic (ii) because it evolved in recent times and did not have time yet to expand its range (neoendemic, Cox et al., 2016). Similarly, disjunct distributions can be (i) the result of recent events of dispersal across a pre-existing geographical barrier, for example, by vector-mediated long-distance dispersal (LDD; Popp et al., 2011; Lewis et al., 2014; Villaverde et al., 2017a; Mogle et al., 2018), or they can be (ii) the result of ancient events, for example, the fragmentation of an ancestral distribution by the formation of a new geographical barrier (Axelrod & Raven, 1978). In “geographical vicariance,” the barrier is geological and linked to plate tectonics, such as the opening of an ocean basin or the rise of a new mountain range (Sanmartín & Ronquist, 2004). However, in the “ecological vicariance” speciation mode, the barrier is environmental (Wiens, 2004): vicariance begins when environmental change creates conditions within a species’ geographical range that are outside of its ancestral ecological niche (e.g., a desert for a moist-adapted species); individuals are unable to persist in or disperse through these extreme ecological conditions, and the species’ range becomes fragmented (Couvreur et al., 2008; Mairal et al., 2015; Pokorný et al., 2015; Noben et al., 2017). When climatic change is global or at a large scale (continental), an

across-clade signature of ecological vicariance might be found (Crisp & Cook, 2007; Pokorný et al., 2015).

Discriminating among the alternative speciation modes mentioned above requires information about the temporal origin of the disjunction, as well as the climatic tolerances of the disjunct species with respect to their ancestors. Table 1 summarizes the expectations of different speciation modes in terms of spatial and temporal patterns, diversification trajectories, and niche dynamics, and depicts their phylogenetic signature. For example, in “geographical vicariance” and “ecological vicariance,” we expect to find allopatry between the two sister species and temporal congruence between their age of divergence and the driving tectonic/climatic events. In these cases, population divergence is typically slow, driven by random mutations and genetic drift (Czekanski-Moir & Rundell, 2019; Table 1). In “ecological vicariance,” we expect to also find a signature of high extinction rates, even though Wiens (2004) did not explicitly mention extinction in defining ecological vicariance. Yet, his model implies that reduced fitness and maladaptation of the marginal populations close to the barrier lead to decreasing effective population sizes and, eventually, to the extinction of the intervening populations, leaving only the surviving populations at the extremes of the original distribution. The end result is two disjunct distributed sister species separated by a long branch with no speciation events (Antonelli & Sanmartín, 2011; Mairal et al., 2015; Table 1).

Table 1 Speciation modes analyzed in this study and their corresponding expectations in terms of spatial and temporal patterns, diversification trajectories, and niche dynamics

Speciation modes	Spatial patterns	Temporal patterns	Diversification trajectories	Niche dynamics	Expected patterns
Geographical vicariance	Allopatric	Temporal congruence with barrier formation	Slow speciation by genetic drift	Typically, conserved niches	
Ecological vicariance	Allopatric	Temporal congruence with barrier formation	High extinction rates associated with speciation	Similar niches	
Ecological speciation	Sympatric	No temporal congruence with barrier formation	Rapid speciation by divergent selection	Divergent niches	
Ecogeographical speciation	Allopatric	Temporal congruence with barrier formation	Depends on the strength of adaptation	Divergent niches	
Dispersal-mediated speciation	Allopatric	No temporal congruence with barrier formation	Depends on dispersal mode (land bridge migration or LDD)	Similar niches, but gradual divergence may occur	

The last column represents the sum of expected patterns. Shape represents geographic distribution; a similar shape indicates sympatry. The filling (white or black) of shapes represents niche dynamics; same color indicates similar niches. Relative branch lengths represent diversification rates. The yellow bar line represents the dispersal barrier, which in some speciation models triggers species divergence. LDD, long-distance dispersal.

Phylogenetic niche conservatism (PNC), the conservation of the ancestral climatic tolerances over evolutionary time (Peterson, 2006), is often associated with “geographical vicariance” (Wiley & Lieberman, 2011; Sanmartín, 2012), but it is a prerequisite for “ecological vicariance,” that is, in this latter speciation mode, population divergence is the result of the species inability to adapt to the changing conditions within their range (Wiens, 2004). The opposite pattern is found in “ecological speciation,” in which population divergence is driven by the adaptation to different ecological settings within the species geographical range (Smith et al., 2001; Couvreur et al., 2011; Schnitzler et al., 2012). In the “ecological speciation” mode, intense ecological selection leads to rapid interruption of gene flow between the “budding” species, so a relatively short temporal gap is expected between stem and crown age estimates (Czekanski-Moir & Rundell, 2019; Riina et al., 2020), resulting in sister species with different ecological niches but sympatric geographical ranges (Table 1). An intermediate speciation mode would be “ecogeographical vicariance” (Gittenberger, 1991; Rundell & Price, 2009), which involves both allopatry and ecological adaptation. Speciation is driven initially by geographical isolation (allopatry), but one of the two descendant populations adapts *in situ* to a change in the local environmental conditions; in other words, niche divergence does not drive speciation as in ecological speciation, but postdates it (Czekanski-Moir & Rundell, 2019, Table 1). Depending on the strength of adaptation, speciation may be faster or slower, with some gene flow among the diverging populations (Couvreur et al., 2011), but it is usually slower than in “ecological speciation” (Table 1).

In the “dispersal-mediated” speciation mode, no temporal congruence is expected between the geological/climatic barrier and lineage divergence; as in vicariance, divergence between the sister species is driven by geographical isolation (i.e., allopatry, Table 1). Speciation might be rapid in the absence of gene flow, as in vector-mediated LDD dispersal, or it might be slow as in the case of dispersal over land corridors. PNC is often associated with LDD, determining which locations are suitable for colonization in the first place (Wiens, 2004; Crisp et al., 2011; Villaverde et al., 2017b). A low to moderate PNC scenario is expected for migration over land corridors, with niches gradually diverging as species adapt to the slightly different environments they disperse to (Pyron et al., 2015; Table 1). Finally, it is important to note that these speciation modes are not mutually exclusive, and processes acting at a given spatial and temporal scale can coexist with others acting at different scales; for example, ecological adaptation to two different habitats driving speciation is compatible with some level of geographic isolation among populations within each habitat.

The Rand Flora (Sanmartín et al., 2010; Pokorný et al., 2015) is a continental biogeographic pattern in which sister taxa (species or clades) are distributed on opposite edges of the African continent (e.g., Western Sahara–Morocco, Ethiopia–Somalia, and Namibia), the adjacent Macaronesian and Socotran Archipelagos, and southern Arabia. Rand Flora taxa exhibit subtropical, temperate, or sub-xeric climatic preferences, but are never found in desert or tropical areas. The northern (Sahara) and southern deserts of Kalahari and Namibia, and the tropical lowlands of Central Africa act as climatic barriers, restricting Rand

Flora lineages to the African continental margins forming a “ring” pattern (see Appendix SI for a more detailed description of this pattern). Recent studies have explained the origins of the Rand Flora as the result of climate-driven extinction and vicariance, linked to aridification waves that fragmented a once widespread distribution (Mairal et al., 2015; Pokorný et al., 2015).

A known example of the Rand Flora pattern is tribe Putorieae, in the family Rubiaceae (Pokorný et al., 2015), which, together with the Theligoneae and Rubieae tribes, forms the Spermacoceae alliance within the Rubioideae subfamily (Backlund et al., 2007). This tribe comprises 34 species, which are all included in a single genus, *Plocama* Ait. (Backlund et al., 2007). Treated by de Candolle in its *Prodromus Systematis Naturalis Regni Vegetabilis* (1830), *Plocama* originally included only one species, the Canarian “balo,” *Plocama pendula* Ait. However, in the most recent taxonomic treatment, Backlund et al. (2007) expanded the genus to include 33 additional species belonging to other African, Arabian, and Asian genera of Putorieae, which were synonymized with *Plocama*, for example, *Choulettia* Pomel, *Gaillonia* A.Rich. ex DC., or *Jaubertia* Guill. Thus, the distribution of tribe Putorieae (*Plocama*) presently extends from the Atlantic subtropical Canary Islands, through the sub-xeric Mediterranean Basin and the semiarid regions of eastern Africa–southern Arabia, all the way to the desert of Namibia in southwest Africa, and the continental mountainous regions of Iran and Afghanistan in western–central Asia (Fig. 1); see Supporting Information (Appendix SI, Table 1.1).

Yet, at the species level, Putorieae exhibits striking examples of continental disjunctions, some conforming to the Rand Flora pattern. In Backlund et al.’s (2007) phylogeny, based on three plastid markers (*rbcl*, *rps16*, *trnTF*) and 13 species (40% species diversity), the Canarian *Plocama pendula* is sister to a clade formed by a group of species endemic to Yemen and the Socotra Archipelago, the “yemenensis group” (*P. yemenensis* (Thulin) M.Backlund & Thulin and *P. tinctoria* (Balf.f.) M.Backlund & Thulin), as well as to the southern African *P. crocyllis* (Sond.) M.Backlund & Thulin (Fig. 1, *rfl*). Backlund et al. (2007) described another remarkable disjunction in the clade formed by *P. reboudiana* (Coss. & Durieu) M.Backlund & Thulin in northwest Africa and *P. aucheri* (Guill.) M.Backlund & Thulin in southern Arabia and the Irano-Turanian region, implying a disjunction within the genus of ~5000 km² across the Sahara Desert (Fig. 1, *naf*). *Plocama calabrica* (L.f.) M.Backlund & Thulin is the most widespread species, extending from northwestern Africa along the Mediterranean shores to Egypt and Irak (Fig. 1, *med*); its sister species *P. brevifolia* Coss. & Durieu ex Pomel is, however, narrowly distributed in northern Morocco–Algeria (Backlund et al., 2007).

Using Backlund et al.’s (2007) data set and Bayesian relaxed molecular clock analysis, Pokorný et al. (2015) dated the split between *P. pendula* and the “yemenensis group” at ~6 million years ago (Ma) and between the latter and *P. crocyllis* at ~4 Ma; the split between *P. reboudiana* and *P. aucheri* was dated also at ~6 Myr, placing these disjunctions within the Miocene–Pliocene warm intervals (Fedorov et al., 2013). In addition, Pokorný et al. (2015) inferred a significantly lower diversification rate in Putorieae (0.006 Myr⁻¹) than expected under the baseline diversification rate in angiosperms, in agreement with historically high extinction rates in the tribe.

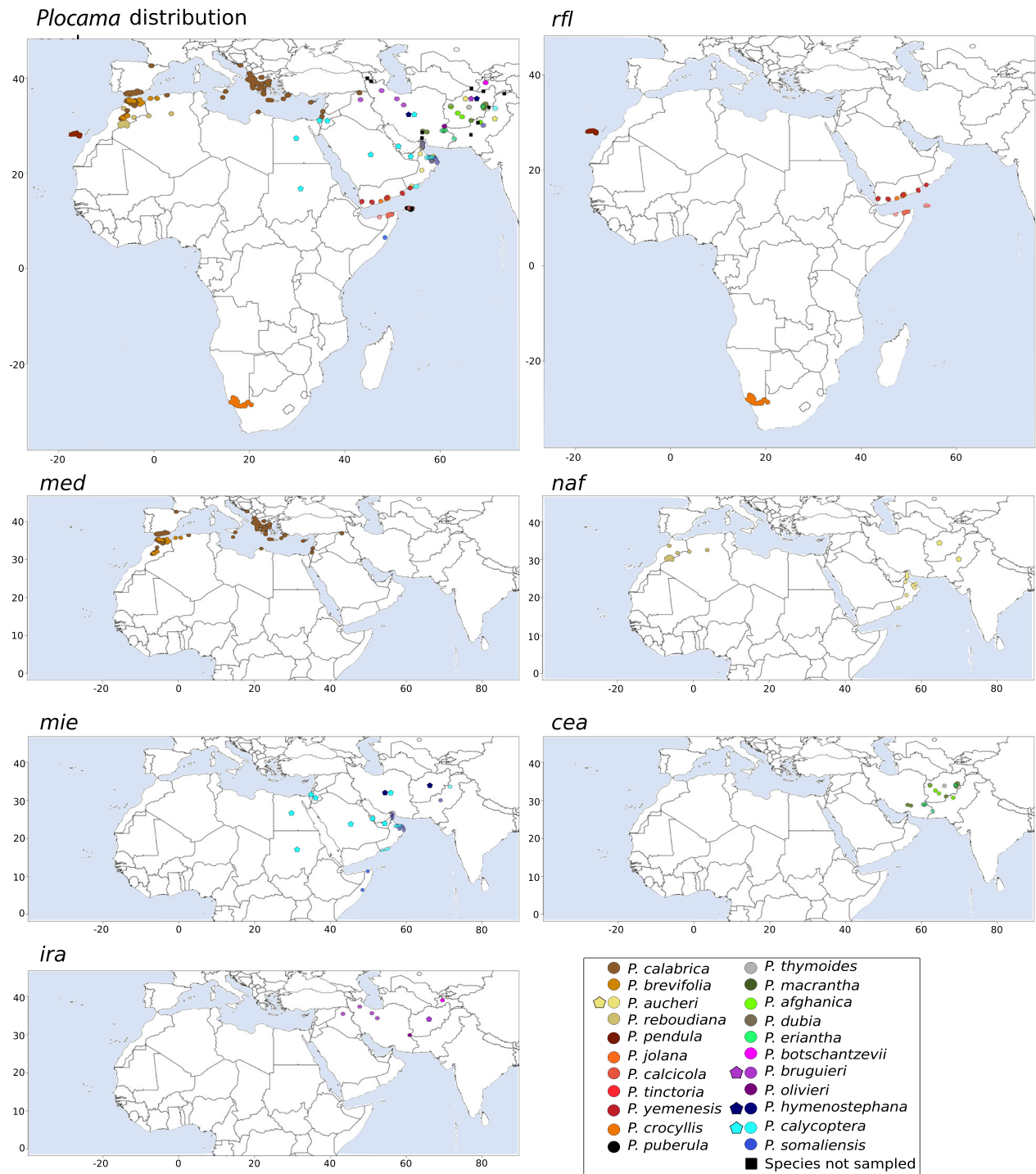


Fig. 1. Geographical distribution of species of tribe Putorieae included in this study, based on occurrence data from herbaria vouchers and GBIF (Appendix S1, Table I.1). Pentagonal symbols show the presence of species in countries mentioned in Backlund et al. (2007) for which precise locations were not available. Squares show geographical locations that are mentioned in georeferenced data sets or previous studies but are not represented in our sampling. The first left panel depicts the distribution of the tribe as a whole; the remaining panels represent the distribution of the six major clades recovered here. Species are colored according to the clade they belong to, but with slight variations to distinguish species within the same clade (*med*: brown, *naf*: yellow, *rfl*: red, *soc*: gray, *mie*: blue, *cie*: purple, *ira*: green). Note the striking geographical disjunctions present in clades *naf* and *rfl*.

Unlike other Rand Flora taxa, such as genera *Canarina* (Mairal et al., 2015) and *Camptoloma* (Culshaw et al., in press) or *Euphorbia* sect. *Balsamis* (Villaverde et al., 2018; Riina et al., 2020), the distribution of *Plocama* extends beyond the Rand Flora range, exhibiting two minor centers of diversity in the Middle East (Fig. 1, *cea*) and, especially, in the mountains and plateaus of Iran–Afghanistan, where a large number of montane-adapted species with nearly sympatric ranges occur (Fig. 1, *ira*).

Here, we used tribe Putorieae (*Plocama*)—with a distribution spanning three continents (Europe, Africa, Asia) and different climatic belts (subtropical, Mediterranean, subxeric, montane), and including examples of both wide-range disjunctions and narrow sympatry—as a case study to examine the role played by geographical isolation, ecology (niche dynamics), and climatic extinction in the formation of new species. To do so, we inferred a new phylogeny of Putorieae, representing 65% of all known species, from DNA sequences of nuclear ITS and six plastid regions. Lineage divergence times, ancestral geographical ranges, and diversification rate shifts were estimated using Bayesian inference (BI) methods. We further reconstructed species climatic tolerances and performed ancestral state reconstruction and niche similarity tests to detect cases of conservation versus divergence of climatic niches. Finally, we compared the inferred phylogenetic patterns, divergence times, niche dynamics, and diversification trajectories across clades with the expectations of different speciation modes (Table 1) to gauge their importance in shaping the current distribution of Putorieae.

We conclude that ecological vicariance and ecogeographical speciation, coupled with climate-driven high extinction rates, explain the origin of the oldest disjunct clades, but that ecological speciation and dispersal-mediated speciation are behind the most recent patterns. Exploring the role of ecology and geography in speciation (Wiens & Donoghue, 2004) is difficult when dealing with time scales of millions of years (Myr) because extinction is likely to have erased the signature of past events in the phylogeny and extant distributions. Yet, combining information from different data sources (spatial, temporal, ecological, geological, and climatic), as done here, can be a useful approach at these macroevolutionary levels, especially if extinction rates are hypothesized to have been historically high, as is the case with the Rand Flora.

2 Material and Methods

2.1 Taxon sampling and DNA sequencing

We generated new sequences for ITS and six plastid markers (*rps16*, *petBD*, *psbJ-petA*, *trnLF*, and *trnTL*) for a data set including 22 species of *Plocama* (Putorieae), some represented by multiple individuals (99 specimens, in total). We also included seven outgroup species, each represented by a single specimen and belonging to the closely related tribes Theligoneae (one *Theligonium* L.) and Rubieae (two species of *Galium* L. and two species of *Rubia* L.). Genus *Paederia* L. (tribe Paederieae), with two species, was used as the most distantly related outgroup, following Backlund et al. (2007). Taxonomic identifications were verified for dubious

specimens using relevant literature and compared with type specimens. GenBank accession numbers and voucher information are listed in Appendix SI (Table 1.1). The complete data set comprised 106 specimens. DNA extractions were done with DNeasy® Plant Kit (Qiagen, Venlo, the Netherlands) following the manufacturer's protocol. Primers and amplification conditions are listed in Appendix SI (Table 1.2). PCR products were sequenced by Macrogen Inc., South Korea (www.macrogen.com). Sequences were edited and aligned using PhyDE® v1.0 (Müller et al., 2005) based on the criteria laid out in Kelchner (2000). Two hotspot regions (438 bp long in *trnTL* and 75 bp in *rps16*) were excluded from downstream analyses, due to uncertain homology assessment. In addition, a 34-bp-long inversion in the *petBD* region was detected and included in the analysis as reverse complement (as discussed by Borsch & Quandt, 2009).

2.2 Phylogenetic analyses

Due to amplification failure or poor-quality sequence product, several specimens have missing markers (see Table 1.1). To avoid bias in phylogenetic reconstruction due to missing data, we built individual gene matrices that included only those specimens that were successfully sequenced for at least three markers. The final matrices comprised 56 sequences each (336 new sequences in total), including 49 specimens representing 22 *Plocama* species, and seven outgroup taxa. Each specimen was represented by both nuclear and plastid markers, with the exception of *P. crocyllis* IS356 and *P. yemenensis* IS226 for which we lack ITS sequences. The individual gene matrices were analyzed under BI with Markov Chain Montecarlo (MCMC) simulations, implemented in the software MrBayes v3.2.6 (Ronquist et al., 2012); we used the best-fit substitution model (i.e., GTR + G) selected using the AIC criterion in MrModelTest v2.3 (Nylander, 2004). Comparison of inferred phylogenies among individual markers revealed a few cases where clades were incongruent with high statistical support—defined as BI posterior probability (PP) values > 0.95 or ML bootstrap support (BS) values > 0.70 (Alfaro et al., 2003). None of these affected the backbone of the tree, which generally lacked resolution for the individual markers (Appendix SII, Figs. II.1–II.6).

Therefore, we constructed a concatenated plastid–nuclear data set (56 sequences) combining the individual matrices. This supermatrix was partitioned by region and analyzed under BI, with settings as above, allowing each individual gene its own GTR + G model, while the overall mutation rate was unlinked across markers. Two analyses with four chains of 10^7 generations each were run, sampling every 1000th generation. Convergence and mixing were assessed in Tracer v1.6 (Rambaut et al., 2014) and MrBayes using several criteria: the effective sample size (ESS) per parameter reached at least 200, potential scale reduction factor (PSRF) values converged to 1.0, and differences among split frequencies were < 0.01. A majority-rule consensus tree was estimated from the sampled trees after removing 25% as burnin. In addition, maximum-likelihood (ML) analyses were performed with RAxML (Stamatakis et al., 2008) on the concatenated supermatrix, using the CIPRES Science Gateway (Miller et al., 2010), under a GTR + CAT model with 1000 bootstrap replicates.

2.3 Divergence time estimation

Divergence times were estimated in BEAST v.1.8.3 (Drummond et al., 2012). The combined data set was partitioned into nuclear and chloroplast partitions to reflect differences in substitution rates (Wolfe et al., 1987), with individual GTR + G substitution models for each partition as selected by MrModelTest. The uncorrected lognormal clock model (UCLD; Drummond et al., 2006) and the birth–death model with incomplete sampling (BDIS; Stadler, 2009) were selected as clock and tree priors, respectively, and compared with the strict clock and the Yule tree priors using Bayes Factor (BF) comparisons based on 100 path sampling (PS) and stepping-stone sampling (SS) power posteriors (Baele et al., 2012). There are no known fossils of Putorieae, so we used a secondary time constraint calibration strategy. The root node—the divergence between tribe Paederiae and the clades formed by tribes Rubieae, Theligoneae, and Putorieae—was assigned a normal prior (Ho & Phillips, 2009) with mean = 41.2 Ma and standard deviation (SD) = 5.5, spanning a 95% highest posterior density (HPD) credibility interval of 32.15–50.25 Ma. This secondary age estimate was obtained from the study of Bremer & Eriksson (2009), which provides a time tree of family Rubiaceae. These authors used four fossils—confidently assigned to each of the three subfamilies and to crown node Rubiaceae—as minimum age priors (the oldest is of Paleogene age), to reconstruct divergence times for all major subclades (tribes and subfamilies) within the family. There are other older fossils (Late Cretaceous) attributed to Rubiaceae, but they have been discarded due to their uncertain taxonomic assignment (Bremer et al., 2004; Bremer & Eriksson, 2009; Wikström et al., 2015). To calibrate the stem age of Rubiaceae (78 Ma, 95% HPD: 95–62 Ma), Bremer & Eriksson (2009) relied on Bremer et al.'s (2004) fossil-calibrated phylogeny of the asterid clade, which provides the crown age of order Gentianales.

Younger molecular estimates for the stem age of Rubiaceae than the one used by Bremer & Eriksson (2009), have been reported in several angiosperm-wide molecular dating studies (Wikström et al., 2001; Smith et al., 2010; Magallón et al., 2015). However, Bremer & Eriksson (2009) used a much more extensive sampling of the family (534 species, 329 genera, and 44 tribes) than any of the aforementioned studies (typically one species per family). Furthermore, a recent angiosperm-wide dating study (Ramírez-Barahona et al., 2020)—with 238 fossil calibrations, more extensive sampling, and topology-constrained relaxed clock models—reported an older age for family Rubiaceae than that of Wikström et al. (2001), Smith et al. (2010), or Magallón et al. (2015). This age (Ramírez-Barahona et al., 2020) is much closer to the Late Cretaceous estimate by Bremer & Eriksson (2009): 82.63 Ma (95% HPD CI: 98.60–62.22 Ma). In addition, the secondary age estimates inferred by Bremer & Eriksson (2009) for the age of tribes Paederiae, Putorieae, Theligoneae, and Rubieae are in agreement with later molecular dating studies that included samples from these tribes (Huang et al., 2013; Nie et al., 2013; Wikström et al., 2015). Therefore, we are confident that our reliance on the secondary age estimate inferred by Bremer & Eriksson (2009) for the divergence of Paederiae from Theligoneae–Rubieae–Putorieae, to root our tree is adequate, as their estimates of the origin of Rubiaceae agree

well with the most up-to-date dating (Ramírez-Barahona et al., 2020). It should be noted that we did not use the crown age estimates for Theligoneae, Rubieae, and Putorieae from Bremer & Eriksson (2009) to avoid introducing a bias in the inferred phylogeny and lineage divergence times.

Schenk (2016) argued that the use of normal priors for secondary age estimates might lead to a false impression of precision or a distribution of ages that sometimes dramatically shifts away from those inferred in the original study from which the secondary calibrations were obtained. To test this, we followed a strategy similar to Lavor et al. (2018): we compared in Tracer the normal prior distribution and the estimated posterior distribution for the single calibrated node in our phylogeny, and corroborated that there was no significant difference in their credible interval widths (i.e., there was overlap in their 95% HPD interval, which was broader for the posterior distribution), and therefore no major bias. We also enforced the position of genus *Paederia* as the most external outgroup in the MCMC search, by constraining the monophyly of a clade formed by all samples of the remaining tribes, Theligoneae, Rubioideae, and Putorieae, in our data set using a Boolean prior. The MCMC chain was run with 5×10^7 generations, sampled every 2500th generation. Tracer was used to ensure adequate mixing and that all parameters had reached an ESS > 200. The software TreeAnnotator v1.8.3 (Drummond et al., 2006) was used to generate the maximum clade credibility (MCC) tree, summarizing clade support as posterior probability values.

2.4 Environmental niche reconstruction

Occurrence records for each species within Putorieae were compiled from data from GBIF (www.gbif.org, DOI:10.15468/dl.rynoer) and from herbarium label data. After removing human observations, incorrect records, and omitting duplicate points, we obtained 258 georeferenced localities, listed in Appendix SI, Table I.1. Some species were poorly represented in our final data set, such as *Plocama botschanzevii* (Lincz.) M.Backlund & Thulin (one specimen from Uzbekistan) and *P. olivieri* (A.Rich. ex DC.) M.Backlund & Thulin (one specimen from Iran). The opposite pattern was found in Mediterranean species *P. calabrica* and *P. reboudiana*, with 151 and 117 records, respectively. Records for which we only have a “country” location instead of a georeferenced locality (e.g., “Sudan” and “Egypt” for *P. calycoptera* (Decne.) M.Backlund & Thulin; cf. Backlund et al., 2007) were not included in the environmental niche reconstruction to avoid potential bias. Table I.1 provides more details on species geographic representation in our data set. To explore differences in climatic tolerances among extant species of Putorieae, for each locality included in our occurrence data set, we downloaded climatic data for 19 bioclimatic variables (Table 2) included in the WorldClim repository (<http://www.worldclim.org/>), using the five arcminutes spatial resolution (Hijmans et al., 2005). We then performed an environmental Principal Component Analysis (ePCA) using the R package “pca3d” (Weiner & Weiner, 2017) in the free software environment R v3.4.1 (R Development Core Team, 2017) to segregate species in the climatic space defined by these bioclimatic variables and the first three PCA axes.

Table 2 Clade posterior probabilities (PP), mean ages, and 95% highest posterior density (HPD) credibility intervals for the age of the stem and crown node of Putorieae, crown nodes of major clades, and other nodes marked in Fig. 2

Nodes/Clades	PP	Mean (Ma)	95% HPD
Stem node	0.83	30.260	19.046–42.751
crown node a	1	24.435	13.899–35.182
<i>med</i>	1	3.828	2.102–8.617
<i>b</i>	1	10.254	5.627–15.651
<i>naf</i>	1	4.950	2.102–8.617
<i>c</i>	0.512	9.991	5.716–15.252
<i>rfl</i>	1	7.315	3.932–11.585
<i>d</i>	0.966	9.248	5.191–14.334
<i>soc</i>	1	3.079	1.421–5.399
<i>e</i>	1	5.879	3.099–9.295
<i>f</i>	1	3.723	1.975–6.009
<i>ira</i>	1	1.808	0.855–3.026
<i>cea</i>	1	2.311	1.119–3.928
<i>mie</i>	1	2.009	0.916–3.553

To reconstruct the climatic tolerances of the ancestors of the extant species, we used continuous trait ancestral niche reconstruction (ANR) methods implemented in the R package *phytools* (Revell, 2012), using the *fastAnc* function, ML inference, and a Brownian motion (BM) model (Felsenstein, 1973). The BM model describes the random evolution of a trait over the branches of a phylogenetic tree and is governed by the rate of trait evolution and the elapsed time of evolution (Silvestro et al., 2015). We did not use the “adaptive selection” Ornstein–Uhlenbeck (OU) model (Lande, 1976) due to potential non-identifiability issues and inconsistency of the maximum-likelihood estimator (MLE) of some parameters, such as the selective optimum and root state; this may bias ancestral state estimation for deep-time nodes (Meseguer et al., 2018). As input phylogeny, we used the MCC chronogram pruned to exclude the outgroup taxa and leaving only one specimen per species of Putorieae (22 tips). In the case of species recovered as non-monophyletic in the phylogeny (see Results), we ran the ANR analysis multiple times, each one changing the individual used as representative for the species; this was done to evaluate potential biases in the inference of ancestral niches along the phylogeny resulting from different specimen selection. We performed ANR analyses for each bioclimatic variable and using the same occurrence and climatic data sets as in the ePCA. Silvestro et al. (2015) showed that failure to account for the intraspecific variability of natural populations leads to over-estimation of trait evolutionary rates in BM models. To avoid this, for each bioclimatic variable, we performed ANR analyses using (i) the mean climatic tolerances of the extant species and (ii) the SD relative to the mean, which was computed over all valid occurrences per species in our data set. In *P. botschanzevii* and *P. olivieri*, we only have one record, so we used instead the “average” genus SD, computed over all 22 species. The SD value serves here as a proxy for the species climatic niche width, which is likely to be smaller in narrow endemics (such as *P. brevifolia* Coss. & Durieu ex Pomel) than in wide-ranging species (such as *P. calabrica*). We preferred to use the SD value, instead of the absolute range, as this makes comparison among species with widely different climatic tolerances possible.

2.5 Niche overlap

To infer the ecological differences between species, we used the R package *ecospat* (Di Cola et al., 2017) in R 3.4.1 (R Development Core Team, 2017) to perform the following analyses. A PCA was built using the data of all nineteen bioclimatic variables. Then, the environmental space was divided into a grid of 100 × 100 cells (Broennimann et al., 2012) to correct for sampling bias and environmental availability. The *ecospat.grid.clim.dyn* function of the *ecospat* package (Di Cola et al., 2017) is a kernel smoother function that was applied to measure the frequency of species occurrences for each combination of environmental conditions in each grid cell of the environment.

The function *ecospat.niche.overlap* was used to calculate the differences in occurrence densities between species, and the Schöner D metric (Warren et al., 2008) was used to calculate the degree of overlap, from no overlap to complete overlap (i.e., from 0 to 1). To measure the equivalency and similarity among niches, we performed a test of niche equivalency and similarity (Broennimann et al., 2012), with 100 random permutations of occurrences between two species, using functions *ecospat.niche.equivalency.test* and *ecospat.niche.similarity.test*. In both cases, we tested for niche conservation and niche divergence (i.e., alternative “greater” and “lower”). Histograms of the observed and randomly simulated overlaps, with *p*-values for the equivalency or similarity tests, were plotted. These tests were only run for those species that were well represented in our occurrence data set (>20 records) to avoid bias.

2.6 Biogeographic analysis

To reconstruct the biogeographic history of Putorieae, we inferred ancestral geographic ranges and rates of geographic evolution under the Dispersal–Extinction–Cladogenesis (DEC) model (Ree & Smith, 2008). This model includes widespread states in the continuous-time Markov chain (CTMC) model that governs range evolution and is therefore appropriate when dealing with continental scenarios in which areas share an edge (Sanmartín, 2012); it also allows us to reconstruct

scenarios of range division, which is one aim in our study. We used the Bayesian implementation of the DEC model (Landis et al., 2018) in the free software platform RevBayes (Höhna et al., 2016). The advantage of using a Bayesian framework in the DEC analysis is the possibility to account for uncertainty in the values of anagenetic parameters (i.e., rates of range expansion and range contraction) and ancestral ranges at cladogenetic nodes by estimating marginal posterior probabilities (Landis et al., 2018; Lator et al., 2018). We defined nine discrete areas, based on distribution patterns in extant *Plocama* species (Fig. 1) but also trying to maximize geographical congruence with previous studies in the region (Moharrek et al., 2019): European Mediterranean Basin (U); Canaries (C); Northwest Africa (D); Southern Africa (S); Northeast Africa, including the coasts of Libya and Egypt, the Red Sea, and the Persian Gulf (N); Middle East (M); Eastern Africa, with adjacent southern Arabia and Socotra (E); Iran–Afghanistan plateau, including Iran, Afghanistan, and Western Pakistan (I); and Central Asia, comprising Uzbekistan, Tajikistan, and Kazakhstan (A). The DEC analysis was run on the MCC tree with one representative specimen per species. Each tip was assigned the full range of the species rather than the locality where the specimen was found. For example, specimen IS340 belonging to *P. hymenostephana* (Jaub. & Spach) Lincz. was coded as present in the entire distribution range of this species (MIE). This criterion seemed the most conservative to avoid biased inferences, as we did not have complete geographic representation of widespread species in the phylogeny. Exceptions to the “single-specimen” rule were those species that were recovered as non-monophyletic in the reconstructed phylogenies (Results). For example, wide-ranging *P. calycoptera* was represented in the DEC analysis by two specimens, IS338 and IS337, coded for the entire range EMIN. Similarly, single-area endemics *P. eriantha* (Jaub. & Spach) M. Backlund & Thulin and *P. macrantha* (Blatt. & Hallb.) M. Backlund & Thulin were represented by two specimens each, coded as area I. Unlike in the ANR analysis, “country” records (represented by pentagons in Fig. 1) were included as part of the geographic distribution of a species, for example, Egypt in *P. calycoptera*. Two independent analyses using default priors in RevBayes were run for 5000 generations, sampling every 10th generation, and final results were summarized in the MCC chronogram (Appendix S1 provides the script for the analyses).

2.7 Diversification rate analysis

To infer changes in speciation and extinction rates over time, we used the CoMET approach (May et al., 2016) implemented in the R package TESS (Höhna, 2013) and run in R v3.4.1. This method uses a hierarchical Bayesian approach to simultaneously infer the time and magnitude of rate shifts in speciation and extinction rates, and the number and timing of mass extinction events (MEEs). MEEs are modeled under the single-pulse model, as a point in time when a significant fraction of the standing diversity (e.g., 90%) is removed from the phylogeny. BF comparisons were used to detect well-supported MEEs and diversification rate shifts (May et al., 2016). Analyses were run on the DEC tree above. We first estimated rate shifts in net diversification (r = speciation – extinction) and background extinction (ϵ = extinction/speci-

ation) without accounting for MEEs. We then ran a second analysis using reversible-jump MCMC model selection to estimate MEEs while integrating out rate shifts in diversification and background extinction (May et al., 2016). This was done to avoid issues with parameter non-identifiability, as under the single pulse model, it becomes difficult to distinguish scenarios with constant diversification interrupted by a mass extinction event from another with a significant upturn in the rate of diversification preceded by slow speciation (Culshaw et al., 2019). We set the sampling fraction at present to 0.65, to account for our incomplete taxon sampling (65% of known Putorieae species), and used an empirical Bayesian approach to estimate the hyperprior shapes for speciation (normal distribution: mean = 0.22, SD = 0.20) and extinction rates (lognormal distribution: meanlog = -3.33, SDlog = 1.36). For other parameters, default priors were used: initial speciation rate = 2.0, initial extinction rate = 1.0, and number of expected rate events = 2.0, which assigns a prior probability of 0.5–0 rate shift events and MEEs (Höhna, 2013). The two non-MEE and MEE analyses were run for 100 million generations with a sampling frequency of 10, burnin = 10,000, and conditioned on survival of the process. Each analysis was run with two chains. Mixing and convergence between chains was assessed by estimating MCMC diagnostics in TESS, using the Rubin–Gelman statistic with ESS values (>500), and by comparing posterior density plots between chains.

3 Results

3.1 Phylogenetic inference and divergence time estimation

Summary statistics for the single markers and the concatenated data set can be found in Appendix SII (Table II.1). The individual gene topologies are shown in Appendix SII, Figs. II.1–II.6. Phylogenies inferred under ML and BI based on the concatenated data set are shown in Appendix SII (Figs. II.7–II.8, respectively). These inferred topologies were very similar, with high support for many backbone branches (BS > 0.70, PP > 0.95, Figs. II.7, II.8). *Galium* and *Rubia* (Rubieae) formed a strongly supported clade, sister to a *Theligonum* (Theligoneae) and *Plocama* (Putorieae) clade. The latter relationship is strongly supported in the ML phylogeny (BS = 100, Appendix SII, Fig. II.7), but weakly supported in our BI analyses (PP = 0.63, Fig. II.7). Seven geographically segregated clades were recovered within Putorieae (Appendix SII, Fig. II.7): Mediterranean (*med*), North African (*naf*), Rand Flora (*rfl*), Socotran (*soc*), Middle Eastern (*mie*), Irano-Turanian (*ira*), and Central Asian (*cea*). Clade *med*, which corresponds to former genus *Putoria* Pers., was inferred as sister to all other clades, followed by a polytomy among clades *naf*, *rfl*, and a clade consisting of all other remaining clades: *soc*, *mia*, *ira*, and *cea*. *Plocama pendula* was inferred as sister to southern African *P. crocyllis* and the East African–Southern Arabian “yemenensis group” (Appendix SII, Fig. II.7). Most species were recovered as monophyletic, except for some taxa in clades *mie*, *cea*, *ira*, and *rfl*, that is, *P. calycoptera*, *P. bruguieri* (A. Rich. ex DC.) M. Backlund, *P. eriantha*, *P. macrantha*, and *P. yemenensis* (Appendix SII, Fig. II.7).

The BEAST MCC chronogram (Fig. 2) showed a similar topology and levels of support as the ML and BI phylogenies (Appendix SII, Figs. II.7, II.8 and Fig. 2). The age of the stem

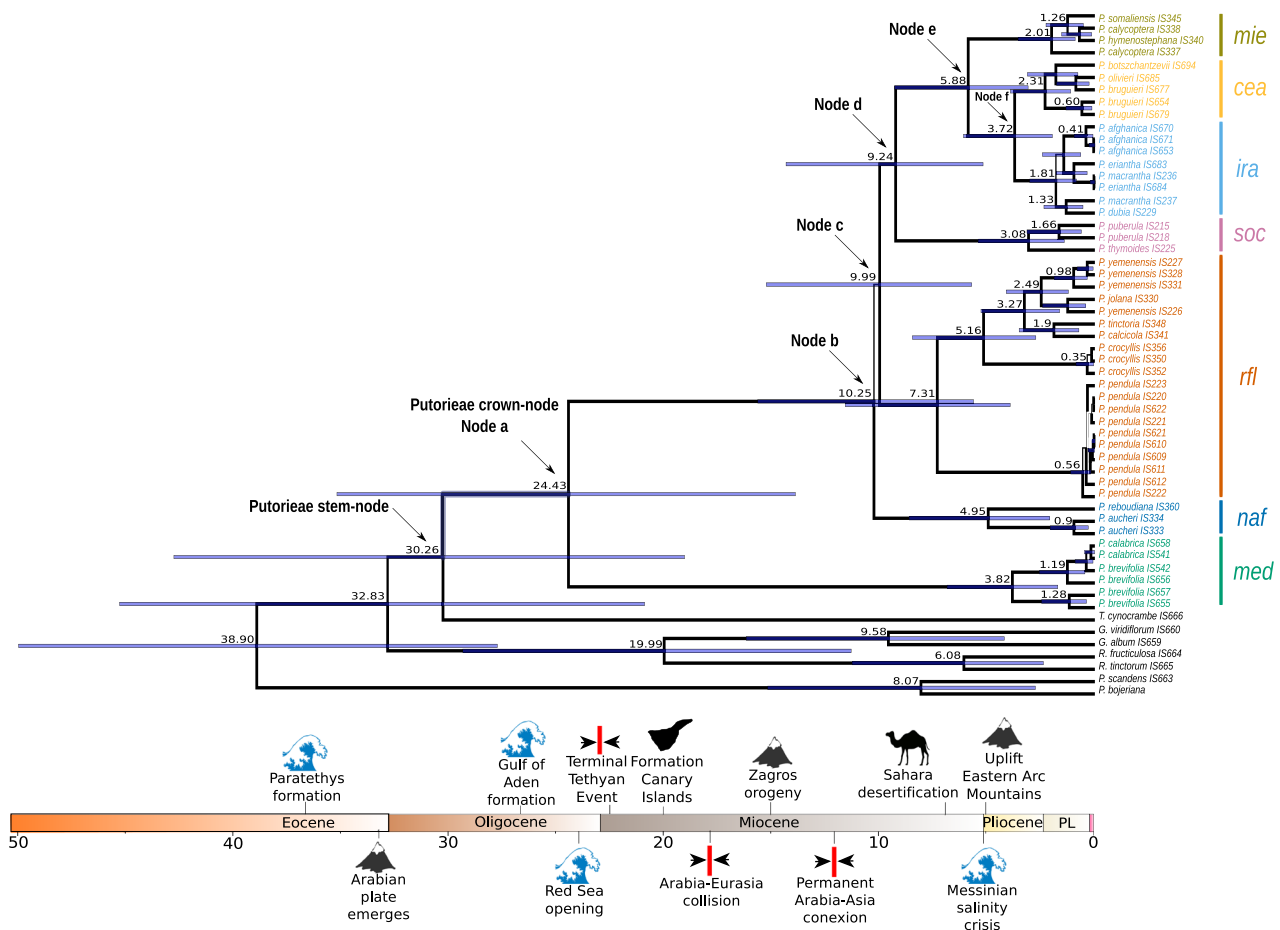


Fig. 2. Maximum clade credibility (MCC) tree from the BEAST analysis of Putorieae (*Plocama*) and allies, showing 95% HPD credibility intervals (blue bars). Branch thickness is proportional to clade posterior probability (PP), which is 1 for most clades. Clade ages are indicated at each node. Timeline shows some of the most important geological events in the biogeographic evolution of Putorieae (see text).

ancestor of Putorieae, that is, the divergence of *Plocama* from *Theligonum*, was estimated as early Oligocene (~30.2 Ma, Table 2), whereas the crown ancestor of all extant species of *Plocama*, the divergence of clade *med* from the rest of Putorieae (node a), was estimated at ~24.4 Myr, in the late Oligocene. The divergence of clade *naf* from the remaining clades (node b) was dated in the late Miocene (~10.2 Ma). The age of the split between *Plocama pendula* and the remaining species in the *rfl* clade was dated as Messinian (~7.31 Ma, Table 2), but the age of origin for this species—namely, the most recent common ancestor (MRCA) of all extant populations—was pushed forward to the late Pleistocene (Fig. 2). The crown age of the remaining clades was estimated as within the late Miocene–Pliocene interval.

3.2 Environmental niche analysis

The first three principal components (PCs) of the ePCA explained 80.96% of climatic variability (47.1%, 21.9%, and 12.1%, respectively; Appendix SIII, Table III.1). No clade was completely isolated from the others along the environmental space (Figs. 3A, 3B); however, some clade clusters could be discriminated. For example, compared with *ira* and *cea*, clades *mie*, *rfl*, and *soc* seem to tolerate higher mean temperatures (bio1, PC1, Fig. 3A),

lower variation in temperature seasonality and annual and diurnal temperature (bio2, PC2, Fig. 3A), and lower values of precipitation seasonality (bio15, PC3, Fig. 3B). Conversely, continental clades *ira* and *cea* showed tolerance to higher temperature seasonality (bio4, PC2, Figs. 3A, 3B). Clade *med* exhibits intermediate values for the variables above (Figs. 3A, 3B). Segregation of climatic niches in the PCA was observed within the North African clade *naf*: *Plocama reboudiana* (“reb” in Figs. 3A, 3B) exhibits similar tolerance to temperature and precipitation values as continental clades *cea* and *ira*, whereas *P. aucheri* (“auc”) appears closer in environmental space to desert-adapted clade *mie* (Figs. 3A, 3B). The insular clade *soc* exhibits the narrowest climatic niche among all clades. Within clade *rfl* (Fig. 3C; Appendix SIII, Table III.1), *Plocama pendula* exhibits lower tolerance to annual temperature oscillations (bio1, PC1) than the other species in *rfl*, especially in comparison with *P. crocyllis*, with larger tolerance to high-seasonality regimes (bio2, PC2, Fig. 3C).

Figure 4 shows the ancestral niche reconstruction for the three bioclimatic variables with the highest correlation values in the ePCA, using the species mean values (see Appendix SIII, Fig. III.2 for the corresponding analyses using the niche-width SD values). Appendix SIII (Fig. III.1) presents the ANR results for

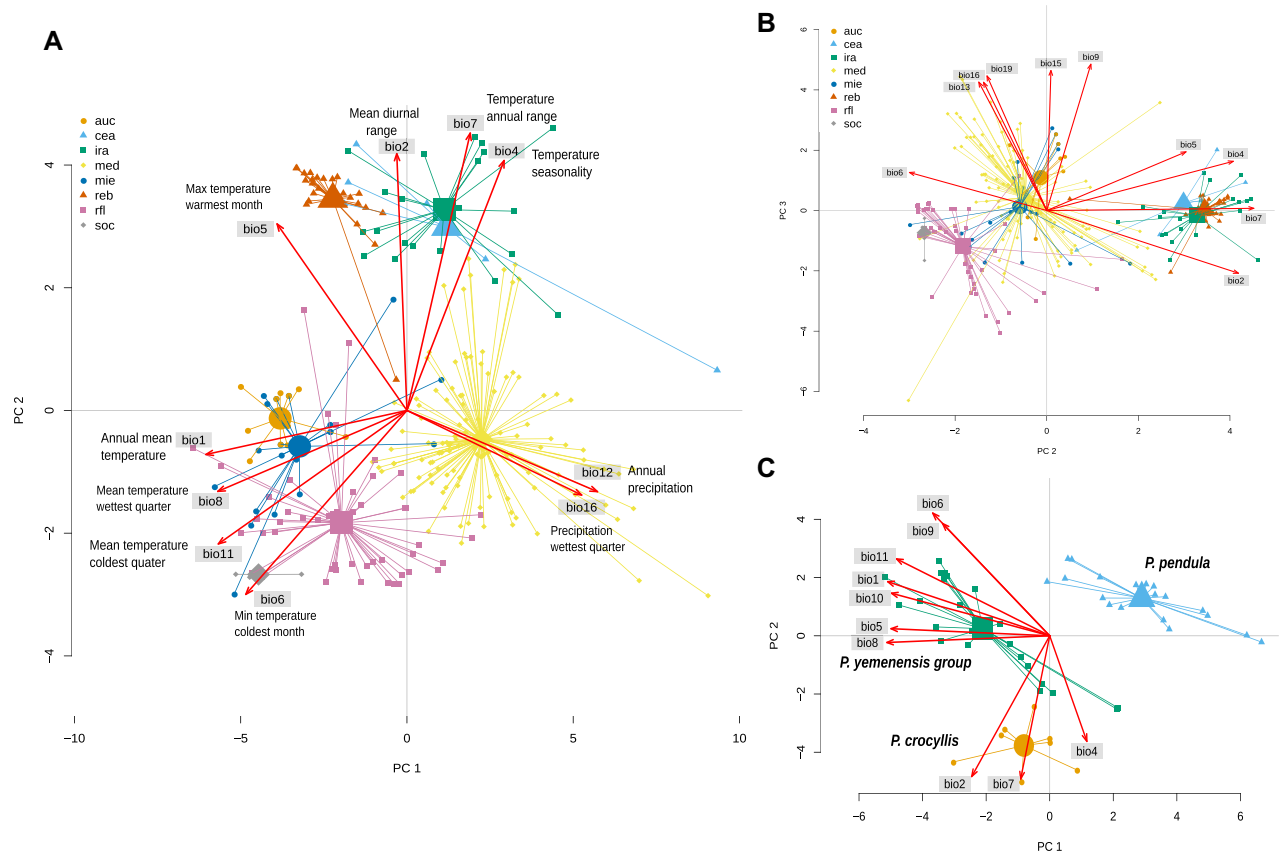


Fig. 3. Climatic niche space, derived from an environmental Principal Component Analysis (ePCA) based on the 19 bioclim variables and georeferenced localities for each species in Putorieae (*Plocama*). Species have been clustered according to the seven major clades recovered in Fig. 2; colored symbols represent these clades; clade abbreviations follow those in Fig. 2. Clade *naf* is not shown as a single clade; instead, sister species *P. aucheri* (labeled as “auc”) and *P. reboudiana* (labeled as “reb”) are shown separate to emphasize that they do not cluster together in the ePCA, that is, they do not share a common niche. **A, B,** Scatter plots of the first three principal components (PCs), explaining 80.9% of the observed variance. **C,** Scatter plot of the first two PCs, explaining 69.6% of the observed variance, for the species within the *rfl* clade.

the remaining bioclimatic variables. We found some evidence of divergence of climatic niches over time. For example, continental clades *ira* and *cea* evolved to adapt to a broader range in annual temperature (Fig. 4A), as well as lower temperature in the coldest quarter (Fig. 4B). Clade *rfl* also exhibits a gradual adaptation to lower annual temperature range (Fig. 4A) and lower precipitation from the ancestor of *P. pendula* to that of *P. crocyllis* (Fig. 4C). Clade *soc* exhibits a similar pattern to *rfl* (Figs. 4A, 4B). In contrast, clade *med* shows tolerance values that are somewhat intermediate to those of all other clades, and which resemble values in ancestral Putorieae (Figs. 4A, 4B). *Plocama calabrica* shows adaptation to slightly higher precipitation values in the coldest quarter than *P. brevifolia* (Fig. 4C). However, the niche overlap test indicated that *P. calabrica* and *P. brevifolia* present similar climatic niches (Fig. 4D); similarity test *p*-value ($p < 0.05$). Conversely, significant niche divergence between species pairs was found in the pair *P. aucheri* and *P. reboudiana* (Fig. 4E); equivalency test *p*-value < 0.05 . In spite of the large difference in range size between *P. calabrica* and *P. brevifolia*, niche divergence was not accompanied by a significant change in SD relative to the ancestral value (Appendix SIII, Fig. III.2).

3.3 Biogeographic and diversification analyses

The DEC biogeographic analysis (Fig. 5) shows large uncertainty levels for the nodes closest to the root, including the crown ancestor of Putorieae and node b (Appendix SIII, Table III.2). The MRCA of all extant Putorieae (node a) was reconstructed as widespread throughout the Mediterranean Basin (UMND, PP < 0.15). The MRCA (node b) of all clades excepting *med* was inferred as widespread across the southern Mediterranean Basin, the Arabian Peninsula, and the neighboring Middle East (EMND, PP = 0.19). Crown node *med* was inferred to have the same circum-Mediterranean distribution as the Putorieae crown node (node a; UMND, PP = 0.74); the MRCA of clade *naf* was also reconstructed as widespread across the southern Mediterranean, southern Arabia, and Middle East, albeit with low probability (EMND, PP = 0.25). Conversely, the next two nodes (c and d) were reconstructed with high probability accuracy as distributed in East Africa–Southern Arabia (E, PP = 0.86). Within clade *rfl*, a dispersal event to the Canary Islands from Eastern Africa/Southern Arabia, followed by allopatric speciation, explains the disjunct distribution of *Plocama pendula* (PP = 0.33). The distribution of the MRCA of the “yemenensis group” and

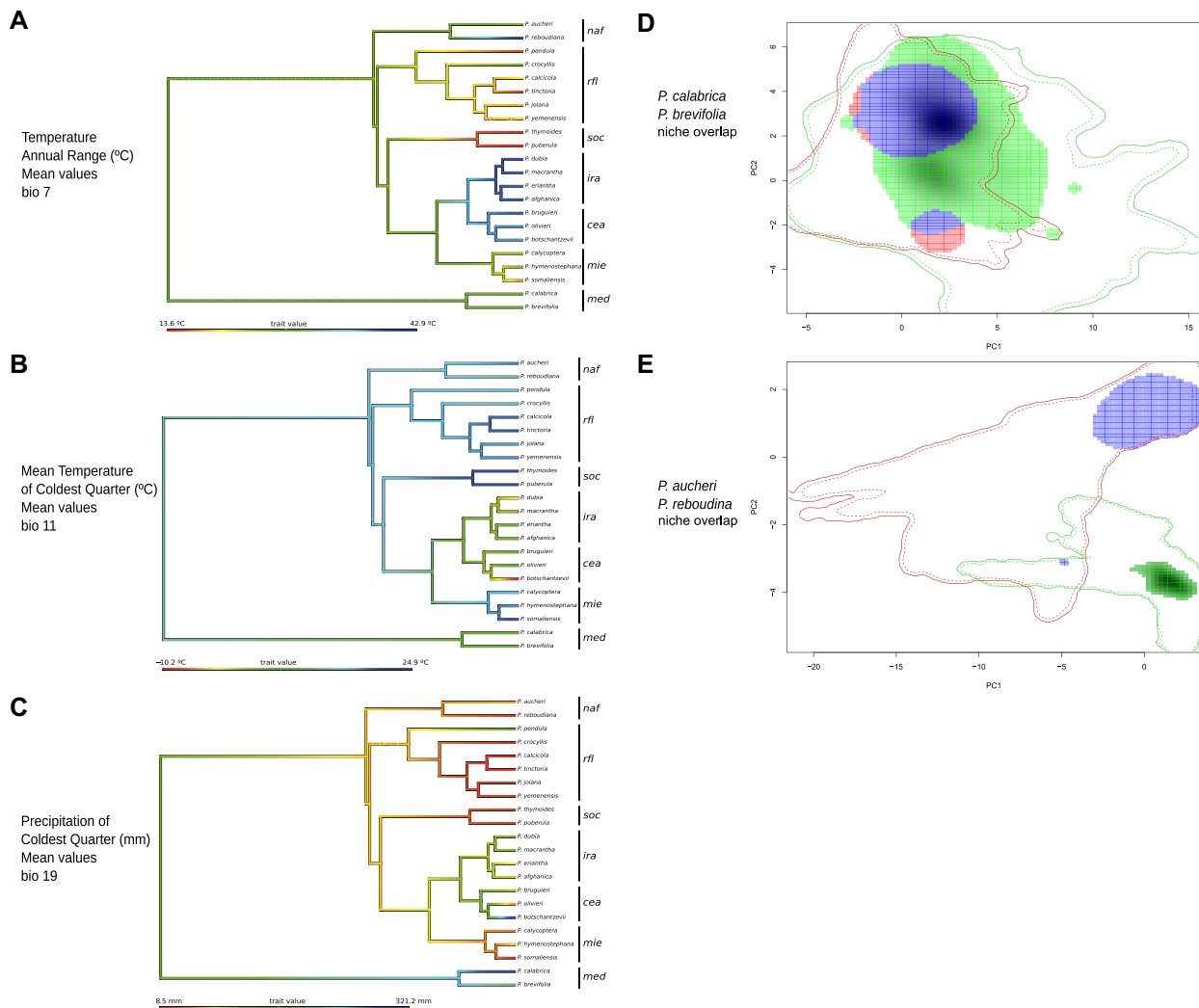


Fig. 4. Left: **A–C**, Ancestral niche reconstruction in Putorieae using mean values for bioclim variables: bio7 (temperature annual range), bio11 (mean temperature of coldest quarter), and bio19 (precipitation of coldest quarter). Temperature data are given in degrees Celsius ($^{\circ}\text{C}$) $\times 10$ and precipitation in mm^3 . Mean values of temperature and precipitation for each species were calculated using all georeferenced localities (Appendix S1, Table I.1). The phylogeny is the BEAST MCC tree in Fig. 2, pruned to leave one individual per species and removing outgroup taxa. Clade labels: cea (Central Asian), ira (Irano-Turanian), med (Mediterranean), mie (Middle East), naf (North African), rfi (Rand Flora), soc (Socotran). Right: **D, E**, Niche overlap analysis showing statistical comparison of climatic niches between two pairs of species. The analysis was performed using 19 bioclimatic variables from WorldClim and drawn using the first two principal components (PCs). **D**, *Plocama calabrica* (green area) and *P. brevifolia* (red area): this species pair has a niche more similar than expected (Similarity test p -value = 0.04). **E**, *Plocama aucheri* (green area) and *P. reboudiana* (red area): the niche of this species pair is less equivalent than expected (Equivalency test p -value = 0.00).

P. crocyllis (PP = 0.63) implied a preceding dispersal event from eastern to southern Africa (Fig. 5). Similarly, a dispersal event from southern Arabia to the Iranian plateau (I) before divergence explains the distribution of the ancestor of clades mie, cea, and ira (node e), with later dispersal to Central Asia (A), the Middle East (M), and the Eastern Mediterranean region (N) (Fig. 5).

The CoMET analysis without MEEs showed significant variation in speciation and extinction rates over time (Figs. 5A, 5B). A negative net diversification rate was inferred during the time interval starting at the crown node (24 Ma) until ~ 12 Ma (Fig. 5A), followed by a rapid upward shift in

diversification rates. Extinction rates were estimated as 1.5 times higher than speciation rates before 10 Ma; this was followed by an increase in speciation rates, which ended in a negative rate shift in the last ~ 1.5 Myr (Fig. 5B). The CoMET analysis with MEEs (Figs. 5C, 5D) recovered similar dynamics. The net diversification rate was inferred low, but positive, until ~ 10 Ma (Fig. 5D), suggesting that the negative diversification in the non-MEE analysis (Fig. 5A) could be the result of a mass extinction event. Indeed, BFs comparisons (Fig. 5D) revealed positive support (BF > 2) for a mass extinction event at ~ 10 Myr, with other minor (non-significant) MEEs before this date.

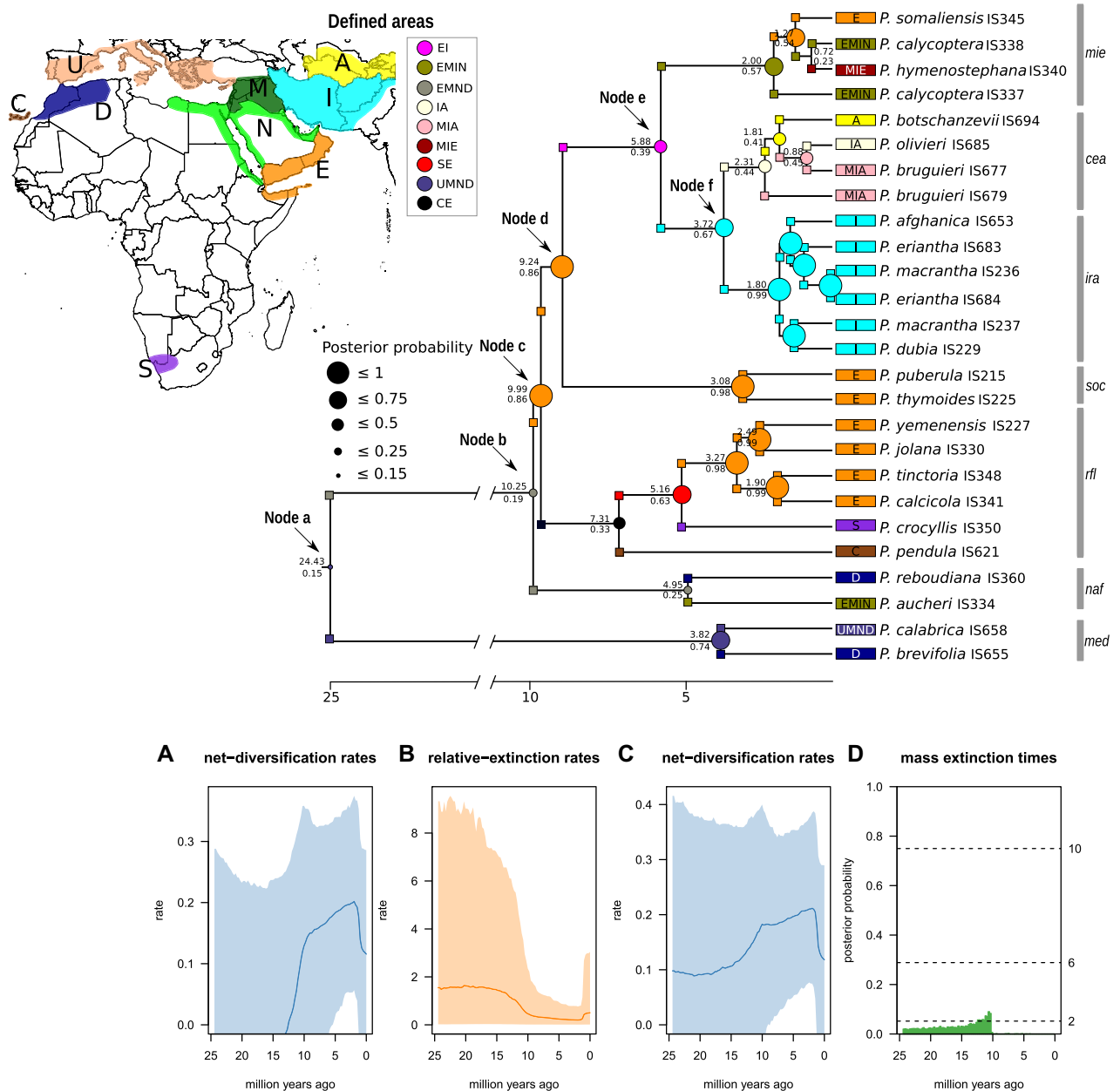


Fig. 5. Top: Geographical range evolution of tribe Putorieae (*Plocama*) inferred with a Bayesian implementation of the Dispersal–Extinction–Cladogenesis (DEC) model. Numbers above nodes represent mean age in Myr; those below nodes represent marginal posterior probabilities (PP) for ancestral ranges (circle size is proportional to PP). The inset map on the top left shows the current distribution of Putorieae, divided into nine areas, with corresponding colors in the phylogeny; color codes for widespread ranges are also indicated. Bottom: Changes in net diversification and relative extinction rates over the evolutionary history of Putorieae, inferred using the Bayesian CoMET approach. **A**, **B** show results allowing the rates of speciation and extinction to change at discrete times; **C**, **D** show analysis allowing for mass extinction events (MEEs): net diversification rates (**C**) and detection of the time of MEEs using Bayes factors (BFs) comparisons (**D**).

4 Discussion

4.1 A new molecular phylogeny for tribe Putorieae

Our phylogeny, based on an expanded data set of species and molecular markers (including nuclear ITS), supports major phylogenetic relationships recovered in previous works on genus *Plocama* (e.g., Backlund et al., 2007). The

relationship between Putorieae and the remaining tribes from the Spermacoceae alliance included in our study differs from previous works based on plastid sequences (Backlund et al., 2007; Bremer & Eriksson, 2009; Ehrendorfer & Barfuss, 2014; Deng et al., 2017), plastid and nuclear sequences (Rydin et al., 2009), or plastid (Wikström et al., 2020) and mitochondrion (Rydin et al., 2017) high-throughput se-

Table 3 Major clades and some species groups in Putorieae and the most probable speciation mode supported by spatial, temporal, diversification, and niche dynamics evidence; in some cases, evidence is not conclusive

Clade	Speciation mode	Sister species	Temporal Barrier	Diversification	Climatic Niches
<i>med</i>	Geographical vicariance/ Ecogeographical speciation	Sympatric?	Strait of Gibraltar opening	Slow drift	Slightly Divergent
<i>naf</i>	Ecogeographical speciation	Allopatric	Isthmus of Suez uplift and Messinian aridification	Moderate extinction rates	Divergent
<i>rfl</i>	Ecological vicariance	Allopatric	Onset of Sahara Desert	High extinction rates	Conserved
<i>soc</i>	Ecological speciation	Sympatric?	Unclear	Fast?	Divergent?
<i>cea</i>	Ecological speciation	Sympatric?	Iranian Mountains uplift	Fast	Divergent*
<i>ira</i>	Ecological speciation	Sympatric	Iranian Mountains uplift	Fast	Divergent*
<i>mie</i>	Medium range dispersal	Allopatric	Unclear	Fast?	Conserved
“yemenensis group”	Oversea LDD	Allopatric	MPWE marine regressions across the Gulf of Aden	Slow drift	Conserved
<i>P. crocyllis</i>	Land bridge migration	Allopatric	Eastern Arc Mountains uplift	High extinction rates	Slightly Divergent

Sympatric and allopatric are defined as occurring within the same large geographical region, as defined in Fig. 5. LDD, long-distance dispersal; MPWE, mid-Pliocene warming event. *With respect to the most recent common ancestor.

quencing (HTS) data. In these studies, *Theligonium*, *Rubia*, and *Galium* are grouped in a clade and separated from Putorieae, whereas in our work *Theligonium* is sister to Putorieae. To the best of our knowledge, the only study supporting this topology is a phylogenomic analysis of order Gentianales based on 353 low-copy nuclear ortholog markers, as well as partial plastomes (Antonelli et al., 2021). Different phylogenetic focus, taxonomic sampling, or the choice of genomic compartment are likely to explain these differences in phylogenetic relationships; future phylogenomic studies with denser sampling might resolve this issue.

Within tribe Putorieae, we recovered six major clades: *med*, *naf*, *rfl*, *mie*, *ira-cea*, and *soc*. These clades are mostly congruent with sections and clades recognized by Backlund et al. (2007), with the exception of clade *soc*, which is new for our analysis. Relationships among these clades also agree with those inferred by Backlund et al. (2007), with the exception of clade *naf*: this latter clade was reconstructed as sister to clade *rfl* in Backlund et al. (2007) but appears in our phylogeny as sister to all remaining clades in *Plocama*, excepting *med* (Fig. 2). In terms of individual species, we increased taxonomic sampling, relative to Backlund et al. (2007), by including five additional African and Southern Arabian species (*P. jolana* (Thulin) M.Backlund & Thulin, *P. calcicola* (Puff) M.Backlund & Thulin, *P. somaliensis* (Puff) M.Backlund & Thulin, *P. thymoides* (Balf.f.) M.Backlund & Thulin, and *P. puberula* (Balf.f.) M.Backlund & Thulin). On the

basis of morphology alone, Backlund et al. (2007) suggested that these species were part of the “yemenensis group.” However, in our molecular phylogeny, these taxa are grouped into three different clades, *rfl*, *soc*, and *mie* (Fig. 2). We also included four additional Iranian species relative to previous studies: *Plocama macrantha*, *P. dubia* (Aitch. & Hemsl.) M.Backlund & Thulin, and *P. afghanica* (Ehrend.) M.Backlund & Thulin, which are included in clade *ira* in our phylogeny, whereas *P. botzschantzevii* was placed in clade *cea*; the latter supports previous morpho-grouping of this species with *P. olivieri* (Backlund et al., 2007).

Admittedly, our sampling of *Plocama* (tribe Putorieae), though extended as compared with previous molecular studies, is far from complete: 12 species (35% of total richness) were not included in our analyses. Of these 12 missing species, only one (i.e., *P. rosea*, occurring in Afghanistan and Pakistan) has previously been sequenced (Backlund et al., 2007). This species is one of the most morphologically divergent within the genus, displaying traits associated to anemochory, but in Backlund et al. (2007), it was placed in a highly supported clade together with *P. eriantha* and *P. macrantha* (both in the terminal *ira* clade in our phylogeny) forming a clade with high support. Indeed, a preliminary analysis including all sequenced species, as well as the *trnTL* sequence of *P. rosea* from Backlund et al. (2007), grouped the latter species with *P. macrantha*, *P. eriantha*, and *P. dubia*. However, the resulting tree had many internal nodes with very low support, in addition to polytomies, so we decided to

base conclusions on those specimens that were successfully sequenced for at least three markers (see Methods section). To the best of our knowledge, there are no published molecular sequences of any of the other eleven missing species: *P. putorioides* (Radcl.-Sm.) M.Backlund & Thulin; *P. asperuliformis* (Lincz.) M.Backlund & Thulin; *P. bucharica* (B.Fedtsch. & Des.-Shost.) M.Backlund & Thulin; *P. crucianelloides* (Jaub. & Spach) M.Backlund & Thulin; *P. iljinii* (Lincz.) M.Backlund & Thulin; *P. inopinata* (Lincz.) M.Backlund & Thulin; *P. kandaharensis* (Ehrend. & Qarar ex Ehrend. & Schönb.-Tem.) M.Backlund & Thulin; *P. mestscherjakovii* (Lincz.) M.Backlund & Thulin; *P. trichophylla* (Popov) M.Backlund & Thulin; *P. szowitzii* (DC.) M.Backlund & Thulin, and *P. vassilzenkoi* (Lincz.) M.Backlund & Thulin. With the exception of *P. putorioides*, the remaining species occur in the Irano-Turanian region, with a distribution extending from the Caucasus Mountains to southern Uzbekistan and Tajikistan. On the basis of their morphology and geographic distribution, Ehrendorfer & Schönbeck-Temesy (2005) classified these 10 species as belonging to former genus *Gaillonia*, within taxonomic sections that also grouped Irano-Turanian species included in our study, such as *P. macrantha*, *P. eriantha*, *P. bostschantzevii*, *P. dubia*, *P. pruguieri*, or *P. olivieri*. These six species were grouped in our phylogeny within the two clades with Irano-Turanian distribution *ira* and *cea*; we predict that these 10 missing species will probably fall within these clades.

Plocama putorioides is endemic to Socotra Island. Ehrendorfer & Schönbeck-Temesy (2005) classified this species as part of genus *Gaillonia*, in a section together with *P. puberula* and *P. thymoides*, which in our phylogeny form the *soc* clade. Backlund et al. (2007) also supported this relationship and grouped these three taxa, based on morphology and geography, with the four Eastern African–Socotran species, *P. calcicola*, *P. yemenensis*, *P. tinctoria*, and *P. jolana*, which in our phylogeny form part of the *rfl* clade. Thus, it seems that *Plocama putorioides* might belong either to the *rfl* clade or to the *soc* clade.

4.2 Different speciation modes contributed to the widespread disjunct distribution of *Plocama*

4.2.1 Ecological vicariance and climate-driven extinction drove early evolution of Putorieae

The evolution of tribe Putorieae took place during a geological interval, the Neogene (23.03–2.58 Ma), characterized by global climatic and geological changes, which brought about profound transformations in vegetation composition (De Man et al., 2004; Plana et al., 2004; Thompson, 2005; Willis & MacDonald, 2011). The stem node was dated in the early Oligocene (30.26 Ma, Fig. 2; Table 3), with crown node divergence (node a, Fig. 2) in the late Oligocene (24.43 Ma). The first part of the Neogene was characterized by warm climates, with a peak in temperature at 24 Ma (Hansen et al., 2008), which culminated with the Middle Miocene Climate Optimum (17–14 Ma, Zachos et al., 2008), when temperatures rose nearly 10 degrees globally, coupled with drops in sea levels (Shevenell et al., 2008). During the late Oligocene, the Mesozoic Tethys Sea was still an open ocean at its eastern border, allowing a connection between the Indian and Atlantic oceans (Meulenkamp & Sissingh, 2003; Liu et al., 2018).

Starting in the early Miocene, ~22 Ma, the Arabian plate, which was isolated from the rest of Africa by the Red Sea

and the Gulf of Suez since 24 Ma (Bosworth et al., 2005), collided with the Eurasian Plate, closing the Tethys Seaway (ArRajehi et al., 2010). Final closure of the Tethys Sea in the Middle Miocene (16–14 Ma, Liu et al., 2018) triggered the rise of new mountain chains in eastern Africa (16–7 Ma) and southwestern Asia (Thomas et al., 1999; Ballato et al., 2010; Mouthereau et al., 2012), as well as the rise of the Iranian and eastern Anatolian plateaus (Biddle & Christie-Blick, 1985; Forte et al., 2010). The closure of the equatorial Tethys Seaway, together with the rise of eastern Africa, contributed to major climate cooling globally and in the African continent, from 14 Ma to ~5 Ma (the late Miocene cooling event, Zachos et al., 2008). This global cooling brought about a major change in vegetation composition in the southwest Palearctic: a Tethyan subtropical flora comprising temperate rainforests and woodland savannahs (Tiffney, 1985) was replaced by more xerophytic lineages and cold-adapted continental taxa throughout the newly uplifted plateaus and mountains of western–central Asia (Manafzadeh et al., 2014). In Africa, climate cooling led to widespread aridification and the expansion of savannahs in the north, east, and southwest (Senut et al., 2009).

Although uncertainty in our inferences is high (Fig. 5), the crown ancestor of tribe Putorieae (node a) is inferred as occurring along the shores of the Tethys Seaway (Fig. 5). Node b, the ancestor of all clades excepting *med*, is also reconstructed as widespread but in the southern Mediterranean Basin, and a similar distribution is inferred for the MRCA of the early-diverging *med* clade. Both node b and the *med* clade are preceded by long internode branches (Fig. 2)—defined here as a time interval between stem and crown nodes that is, at least, twice as long as the time interval between the crown node and the tips it subtends. Typically, long internode branches with no speciation events preceding a crown node are interpreted as the signature of high extinction rates (Antonelli & Sanmartín, 2011; Sanmartín & Meseguer, 2016). Support for the extinction hypothesis is also provided by the CoMET analysis, which detects a significant mass extinction event associated with high background extinction rates (> 90% magnitude) between ~24 Ma (crown node a) and ~10 Ma, the start of diversification of major clades within Putorieae (Fig. 5). This evidence suggests that the ancestors of extant Putorieae formed part of the subtropical vegetation that characterized the Early Tertiary Tethys Basin (Tiffney, 1985), and that the reduced/disjunct distribution of some of these major clades could be the result of extinction events associated with the closure of the Tethys Seaway and global aridification. The first half of the Miocene (23–13.8 Ma) has been suggested as a determinant period in African vegetation extinction based on the fossil record (Morley, 2000). Within the Spermacoceae alliance, Putorieae was not the only tribe affected by these major climatic and geological changes. *Theligonium*, which is believed to be also of Tethyan origin (Deng et al., 2017), probably underwent similarly high extinction rates during the Miocene, driven by the desertification of central Asia and the rise of the Qinghai–Tibetan Plateau. Nowadays, *Theligonium* has a disjunct distribution on either side of the Asian plateaus, where it could have found refugia after the Miocene climatic oscillations (Deng et al., 2017).

Another example of a clade subtended by a long branch is the Rand Flora (*rfl*) clade. The MRCA of all extant Macaronesian *P. pendula* populations is just 0.5 Ma (the species crown node), whereas the divergence from the eastern/southern African clade dates back to 7.32 Ma (the species stem node); that is, the *rfl* crown node is subtended by an internode branch that is nearly 14 times longer than the terminal branches within the *P. pendula* clade (Fig. 2). Such a large temporal gap between the western and eastern elements of the Rand Flora disjunction has been observed in other genera, such as *Canarina* (Mairal et al., 2015). As abovementioned, the collision of the Arabian and Eurasian plates, starting in the Middle Miocene, and the subsequent uplift of the Horn of Africa high plateaus and Eastern Arc Mountains, all led to a major aridification trend that climaxed with the formation of the Sahara Desert ~8–7 Ma (Sepulchre et al., 2006). A gradual replacement of moist subtropical vegetation by more xeric savannahs and grassland lineages took place in this period (Senut et al., 2009). Ancestral climatic preferences for the *rfl* clade (Fig. 4) correspond to those of a subtropical lineage, exhibiting adaptation to milder temperatures and less seasonal precipitation regimes, and to lower seasonality values, compared with those exhibited by species within the continental clades *mie* and *ira*. All this evidence supports the hypothesis that the divergence between the western and eastern (southern) elements of the *rfl* clade was driven by (aridification-driven) ecological vicariance and climatic extinction (Table 3), as inferred for other Rand Flora taxa (Mairal et al., 2015). Species within the *rfl* clade would thus conform to the definition of Rand Flora relicts: taxa isolated in climatic refugia on each side of the African continent (Pokorný et al., 2015; Mairal et al. 2017).

4.2.2 The interplay of ecology, tectonics, and geography

As mentioned before, the branch subtending clade *med* (former *Putoria*) is one of the longest in our phylogeny of Putorieae, nearly seven times longer (24.4 Ma) than the split between *Plocama calabrica* and *P. brevifolia* (3.8 Ma, Fig. 2). These two species exhibit similar climatic niches (Fig. 5), but *P. calabrica* shows a slightly wider niche than that of *P. brevifolia* (Fig. 5; Appendix SIII, Figs. III.1 and III.2). These broader climatic tolerances fit well the widespread distribution of *P. calabrica* along the shores of the Mediterranean Basin, from the Iberian Peninsula to Anatolia, including northern Africa. Indeed, the MRCA of clade *med* was inferred as occurring throughout the Mediterranean (UMND), with the split between the two species attributed to one species inheriting the entire ancestral range (*P. calabrica*) and the other (*P. brevifolia*, found in northwest Africa) inheriting only a subset of it (Fig. 5). Thus, these sister species are presently not allopatric. However, the split of *P. calabrica* from *P. brevifolia* (~3.8 Ma) is roughly congruent with the final opening of the Strait of Gibraltar in the Mid-Pliocene, which also marked the onset of the Mediterranean climate (Thompson, 2005). It might be that *P. calabrica* had originally a narrower distribution, spanning only the northwest Mediterranean Basin, and that its current pan-Mediterranean distribution stems from more recent dispersal events. If so, *P. calabrica* and *P. brevifolia* would be examples of the geographical vicariance speciation mode (Table 3), in

which allopatric divergence is driven by tectonic changes. The slightly climatically different (Mediterranean) tolerances of *P. calabrica*, compared with *P. brevifolia*, would also support some adaptation *in situ* (ecogeographical speciation, Table 1), but niche divergence was not supported by the Similarity test (Fig. 4). A population-level analysis of *P. calabrica* with more extensive sampling is needed to confirm this hypothesis.

The stem ancestor of clade *naf* (node b, Fig. 2) is dated as late Miocene (~10.2 Ma), with the split between *Plocama aucheri* and *P. reboudiana* taking place in the early Pliocene (~5 Ma, Fig. 2). Significant differences in climatic niches were detected between these two species (Figs. 3, 4e). This timing and niche preferences, together with their current geographical isolation—divided by the Sahara Desert (Fig. 1D)—support “ecogeographical speciation” as the primary mode underpinning their split (Table 3). An alternative explanation is that the extant species niches represent the extremes of a wider ancestral niche, fragmented by extinction. We cannot discard the latter scenario, as our climatic niche models are based on occurrence records and therefore only recover the realized niche (Peterson, 2006). Nevertheless, if this were the case, we would expect the two extant niches to be different from the ancestral niche. Instead, *P. aucheri* typically exhibits climatic tolerances resembling the inferred ancestral values (Figs. 4, III.1), whereas *P. reboudiana* consistently shows divergence in niche values from those of their ancestor. The time of the split between the two species, early Pliocene, coincides with the uplift of the Isthmus of Suez, which has been suggested as the underlying mechanism for the divergence of African and Arabian (western Asian) taxa (Sanmartín, 2003). The expansion of the Saharan and the Arab–Syrian deserts at the end of the Miocene (6 Ma), as a result of the Messinian aridification (Krijgsman et al., 1999), could also have acted as an additional vicariance barrier (Sanmartín, 2003).

4.2.3 Ecological speciation and mountain uplift drove the evolution of the youngest clades

In the “ecological speciation” mode, no geographical or climatic barrier fragments the ancestral distribution; instead, speciation is driven by niche evolution (Table 1). This niche evolution is often associated with events of rapid diversification by “ecological release” (Wiens et al., 2010). The uplifts of the Zagros, Kopet Dag, Alborz, and Pamir Mountains in the late Miocene were responsible for climate cooling and rain shadows on a large scale in western Asia. This tendency increased in the early Pliocene due to the Messinian Salinity Crisis, which introduced an even colder and drier climate (Meulenkamp & Sissingh, 2003). During this period, continental-adapted, xerophytic plant lineages became dominant in the Iranian and Afghanistan plateaus and adjacent regions (Manafzadeh et al., 2017). Clades *ira* and *cea* inhabit the Iranian plateau and the mountains of Afghanistan and western–central Asia (Fig. 5), and exhibit adaptations to lower mean annual temperatures and lower precipitation regimes, compared with other clades (Figs. 3, 4). The relatively short internode branches subtending these two clades (especially *ira*), together with the apparent non-monophyly of some species (Appendix SII, Fig. II.7), suggest incomplete lineage sorting (ILS) due to recent, rapid

speciation. All these characteristics fit well the “ecological speciation” mode (Table 3).

The stem ancestor of the soc clade, endemic to the Socotra Archipelago, was dated at ~9.2 Myr (Fig. 3), coinciding with the time when the sea level surrounding this archipelago was much lower and Socotra was closer to the Arabian plate, which would have facilitated colonization (Culek, 2013). Divergence between the extant species was dated back to the Mid-Pliocene (~3.1 Ma). The soc clade exhibits the narrowest climatic niche of all clades in the ePCA (Figs. 4A, 4B), and there is no evidence of species niche divergence in the ANR analyses, either (Figs. 5, III.1, III.2). However, these inferences might be biased by the lack of resolution in the bioclimatic data, that is, the WorldClim database is known to be less accurate for islands and mountain ranges, and also by our limited sampling. *Plocama thymoides* is confined to the western end of the Socotra Island, 360–700 m of altitude, whereas *P. puberula* is widespread in the island occurring along the coastal lowlands (Thulin, 1998). It is thus possible that there are also diverging climatic tolerances between these two species due to differences in humidity and temperature in their habitats. If niche divergence were to be observed using fine-scale climatic data, the soc clade would be yet another example of “ecological speciation.”

4.2.4 Dispersal-mediated speciation

The divergence between southern African *P. crocylis* and the “yemenensis group” (~5.2 Ma, Fig. 2) coincides with renewed tectonic activity in eastern Africa during the early Pliocene, which climaxed with the uplift of the Eastern Arc Mountains and the Horn of Africa highlands (Sepulchre et al., 2006). Rand Flora taxa have been hypothesized to use these montane corridors to migrate southward from east Africa (Sanmartín et al., 2010; Pokorný et al., 2015). Gradual niche divergence or adaptation to more continental, drier environments is observed in southern African species *P. crocylis* (Figs. 4C, 5, III.1, III.2). This relaxation of the PNC scenario is more congruent with a hypothesis of land bridge dispersal (Table 1) than with one of vector-assisted LDD (Pelsner et al., 2012). At the same time, we can observe a long branch subtending the crown population divergence within *P. crocylis* (Fig. 2). This suggests that, after southward migration (via land bridges) of the stem ancestor of *P. crocylis*, extinction wiped out the intervening populations, leaving only extant crown *P. crocylis* isolated in southern Africa. Steady aridification during late Pliocene–Pleistocene and climatic oscillations during Quaternary could have removed a former subtropical flora from southern Africa and replaced it by more arid-adapted lineages (Scott, 1999; Goldblatt & Manning, 2000; Scott & Vogel, 2000). It is thus likely that the ancestor of *P. crocylis* could have been part of that former flora, with the extant taxon now restricted to climatic refugia. Therefore, our results point to a “dispersal-mediated speciation” mode, in combination with climate-driven extinction (Table 3).

A second example of “dispersal-mediated speciation” is that of the “yemenensis group.” The four species in this group included in our study inhabit the coasts of the Gulf of Aden, including the Somalian Mountains, and coastal Yemen and Oman. Our analysis dates the split between *Plocama yemenensis* and *P. jolana* (in southern Arabia), from *P. calcicola* and *P. tinctoria* (in Somalia and Socotra), at 3.2 Ma, coincident with the Mid-Pliocene Warming Event marine regressions, which could have facilitated over-water migration across the Gulf of Aden.

Another candidate of the “dispersal-mediated speciation” is Middle East clade *mie*, sister to clades *ira* and *cea*. This clade has a widespread distribution spanning from eastern Africa through the Middle East to Iran–Afghanistan, typically occurring in coastal and lowland interior areas (Fig. 1). We did not find niche divergence within this clade (Fig. 4). This clade’s crown age is ~2 Myr, and it includes several non-monophyletic species (possible ILS, Appendix SII, Fig. II.7). We hypothesize that species in this clade were formed by tracking their ancestral niche, probably through short-range migration, thereby fitting the dispersal-mediated speciation mode (Tables 1, 3).

5 Conclusions

Lineages exhibiting geographically discontinuous distributions spanning different continents and climates are of special interest to scientists and conservation biologists, as they can provide information on the ability of species to either adapt to changing climates or migrate to other climatically similar regions, or, in the worst-case scenario, go extinct. Many of the wide-range disjunctions observed in Putorieae were driven by PNC and extinction associated with Neogene aridification, but dispersal and niche evolution via mountain uplift probably explained the youngest divergence patterns. We demonstrate that integrating information from lineage divergence times, ancestral ranges, ecological niche models, and diversification rates can shed light on the role of ecology, tectonics, and geography in speciation, and on the origin of wide-range geographic disjunctions like the Rand Flora.

Acknowledgements

Funding was provided by the Spanish Government, through grants CGL2015-67849-P (MINECO/FEDER) to IS. MR was supported by a MINECO FPI Fellowship (BES-2013-065389), jointly supervised by IS and RR. RR was supported by grant CGL2015-73621-JIN (AEI/FEDER). LP was supported through an EU-funded SYNTHESIS3 grant (GB-TAF-5153). The authors acknowledge the CIPRES portal for computational resources, Dietmar Quandt for assistance with *petBD* alignment, and the E, K, and MA herbaria for access to their collections. The authors declare that there is no conflict of interest.

Author contributions

I.S.: Conceptualization, formal analysis, investigation, writing (original draft). M.R.-B.: formal analysis, investigation, writing (original draft). R.R.: Investigation, writing (original draft). S.O.: Investigation, writing (review and editing). B.M.: Investigation. T.V.: Formal analysis, writing (review and editing). L.P.: Data resources, writing (review and editing). A.F.: Data resources.

References

Alfaro ME, Zoller S, Lutzoni F. 2003. A simulation study comparing the performance of Bayesian Markov chain Monte Carlo

- sampling and bootstrapping in assessing phylogenetic confidence. *Molecular Biology and Evolution* 20: 255–266.
- 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.
- Antonelli A, Clarkson AJ, Kainulainen K, Maurin O, Brewer GE, Davis AP, Eritawale N, Goyder DJ, Livshultz T, Persson C, Pokorny L, Straub SCK, Struwe L, Zuntini AR, Forest F, Baker WJ. 2021. Settling a family feud: A high-level phylogenomic framework for the Gentianales based on 353 nuclear genes and partial plastomes. *American Journal of Botany* 108: 1–23.
- ArRajehi A, McClusky S, Reilinger R, Daoud M, Alchalbi A, Ergintav S, Gomez F, Sholan J, Bou-Rabee F, Ogubazghi G, Haileab B, Fisseha S, Asfa L, Mahmoud S, Rayan A, Bendik R, Kogan L. 2010. Geodetic constraints on present-day motion of the Arabian Plate: Implications for Red Sea and Gulf of Aden rifting. *Tectonics* 29: TC3011.
- Axelrod DI, Raven PH. 1978. Late Cretaceous and Tertiary vegetation history of Africa. In: Werger MJA ed. *Biogeography and ecology of Southern Africa*. Dordrecht: Springer. 31: 77–130.
- Backlund M, Bremer B, Thulin M. 2007. Paraphyly of Paederieae, recognition of Putorieae and expansion of *Plocama* (Rubiaceae-Rubioideae). *Taxon* 56: 315–328.
- Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko AV. 2012. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution* 29: 2157–2167.
- Ballato P, Mulch A, Landgraf A, Strecker MR, Dalconi MC, Friedrich A, Tabatabaei SH. 2010. Middle to late Miocene Middle Eastern climate from stable oxygen and carbon isotope data, southern Alborz mountains, N Iran. *Earth and Planetary Science Letters* 300: 125–138.
- Biddle KT, Christie-Blick N. 1985. Strike-slip deformation, basin formation, and sedimentation: Based on a symposium. *Society of Economic Paleontologists and Mineralogists* 37: 375–384.
- Borsch T, Quandt D. 2009. Mutational dynamics and phylogenetic utility of noncoding chloroplast DNA. *Plant Systematics and Evolution* 282: 169–199.
- Bosworth W, Huchon P, McClay K. 2005. The Red Sea and Gulf of Aden Basins. *Journal of African Earth Sciences* 43: 334–378.
- Bremer K, Friis E, Bremer B. 2004. Molecular phylogenetic dating of asterid flowering plants shows early Cretaceous diversification. *Systematic Biology* 53: 496–505.
- Broennimann O, Fitzpatrick MC, Pearman PB, Petitpierre B, Pellissier L, Yoccoz NG, Thuiller W, Fortin M-J, Randin C, Zimmermann N, Graham CH, Guisan A. 2012. Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography* 21: 481–497.
- Bremer B, Eriksson T. 2009. Time tree of Rubiaceae: Phylogeny and dating the family, subfamilies, and tribes. *International Journal of Plant Sciences* 170: 766–793.
- de Candolle AP. 1830. *Prodromus Systematis Naturalis Regni Vegetabilis*. Paris: Treuttel and Würtz. 4–1–638.
- de Candolle AP. 1834. *Introduction à l'étude de la botanique, Tome 1 and 2*. Paris: Librairie Encyclopédique de Rovet.
- Couvreur TL, Chatrou LW, Sosef MS, Richardson JE. 2008. Molecular phylogenetics reveal multiple tertiary vicariance origins of the African rain forest trees. *BMC Biology* 6: 54.
- Couvreur TL, Forest F, Baker WJ. 2011. Origin and global diversification patterns of tropical rain forests: Inferences from a complete genus-level phylogeny of palms. *BMC Biology* 9: 44.
- Cox CB, Moore PD, Ladle RJ. 2016. *Biogeography: An ecological and evolutionary approach*. 9th ed. Hoboken, NJ: Wiley-Blackwell. 1–494.
- Crisp MD, Cook LG. 2007. A congruent molecular signature of vicariance across multiple plant lineages. *Molecular Phylogenetics and Evolution* 43: 1106–1117.
- Crisp MD, Trewick SA, Cook LG. 2011. Hypothesis testing in biogeography. *Trends in Ecology and Evolution* 26: 66–72.
- Culek M. 2013. Geological and morphological evolution of the Socotra Archipelago (Yemen) from the biogeographical view. *Journal of Landscape Ecology* 6: 84–108.
- Culshaw V, Stadler T, Sanmartín I. 2019. Exploring the power of Bayesian birth-death skyline models to detect mass extinction events from phylogenies with only extant taxa. *Evolution* 73: 1133–1150.
- Culshaw VMV, Villaverde T, Mairal M, Olsson SE, Sanmartín I. 2021. Rare and widespread: Integrating Bayesian MCMC approaches, Sanger sequencing and Hyb-Seq phylogenomics to reconstruct the origin of the enigmatic Rand Flora genus *Camptoloma*. *American Journal of Botany*, 00–00 (September issue).
- Czekanski-Moir JE, Rundell RJ. 2019. The ecology of nonecological speciation and nonadaptive radiations. *Trends in Ecology and Evolution* 34: 400–415.
- De Man E, Van Simaey S, De Meuter F, King C, Steurbaut E. 2004. Oligocene benthic foraminiferal zonation for the southern North Sea Basin. *Bulletin van het Koninklijk Belgisch Instituut voor Natuurwetenschappen, Aardwetenschappen* 74: 177–195.
- Deng T, Zhang JW, Meng Y, Volis S, Sun H, Nie ZL. 2017. Role of the Qinghai-Tibetan Plateau uplift in the Northern Hemisphere disjunction: evidence from two herbaceous genera of Rubiaceae. *Scientific Reports* 7: 1–12.
- Di Cola V, Broennimann O, Petitpierre B, Breiner FT, d'Amen M, Randin C, Engler R, Pottier J, Pio D, Dubuis A, Pellissier L, Mateo RG, Hordijk W, Salamin N, Guisan A. 2017. ecospat: An R package to support spatial analyses and modeling of species niches and distributions. *Ecography* 40: 774–787.
- Donoghue MJ. 2003. Toward an Integrative Historical Biogeography. *Integrative and Comparative Biology* 43: 261–270.
- Drummond AJ, Ho SY, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLOS Biology* 4: 88.
- 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.
- Ehrendorfer F, Barfuss MHJ. 2014. Paraphyly and Polyphyly in the Worldwide Tribe Rubieae (Rubiaceae): Challenges for Generic Delimitation. *Annals of the Missouri Botanical Garden* 100: 79–88.
- Ehrendorfer F, Schönbeck-Temesy E. 2005. Gaillonia, Jaubertia, Pterogaillonia. *Flora Iranica* 176: 22–35.
- Fedorov AV, Brierley CM, Lawrence KT, Liu Z, Dekens PS, Ravelo AC. 2013. Patterns and mechanisms of early Pliocene warmth. *Nature* 496: 43–49.
- Felsenstein J. 1973. Maximum likelihood and minimum-steps methods for estimating evolutionary trees from data on discrete characters. *Systematic Biology* 22: 240–249.
- Forte AM, Cowgill E, Bernardin T, Kreylos O, Hamann B. 2010. Late Cenozoic deformation of the Kura fold-thrust belt, southern Greater Caucasus. *Geological Society of America Bulletin* 122: 465–486.

- Gittenberger E. 1991. What about non-adaptive radiation? *Biological Journal of the Linnean Society* 43: 263–272.
- Goldblatt P, Manning J. 2000. *Cape plants: A conspectus of the Cape flora of South Africa*. Pretoria, South Africa: National Botanical Institute.
- Hansen J, Sato M, Kharecha P, Beerling D, Berner R, Masson-Delmotte V, Pagani M, Raymo M, Royer DL, Zachos JC. 2008. Target atmospheric CO₂: Where should humanity aim? *The Open Atmospheric Science Journal* 2: 217–231.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965–1978.
- Ho SYW, Phillips MJ. 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology* 58: 367–380.
- Höhna S. 2013. Fast simulation of reconstructed phylogenies under global time-dependent birth–death processes. *Bioinformatics* 29: 1367–1374.
- Höhna S, Landis MJ, Heath TA, Boussau B, Lartillot N, Moore BR, Huelsenbeck JP, Ronquist F. 2016. RevBayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. *Systematic Biology* 65: 726–736.
- Huang WP, Sun H, Deng T, Razafimandimbison SG, Nie ZL, Wen J. 2013. Molecular phylogenetics and biogeography of the eastern Asian–eastern North American disjunct *Mitchella* and its close relative *Damnacanthus* (Rubiaceae, Mitchelleae). *Botanical Journal of the Linnean Society* 171: 395–412.
- Kelchner SA. 2000. The evolution of non-coding chloroplast DNA and its application in plant systematics. *Annals of the Missouri Botanical Garden* 87: 482.
- Krijgsman W, Hilgen FJ, Raffi I, Sierro FJ, Wilson DS. 1999. Chronology, causes and progression of the Messinian salinity crisis. *Nature* 400: 652–655.
- Lande R. 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30: 314–334.
- Landis MJ, Freyman WA, Baldwin BG. 2018. Retracing the Hawaiian silversword radiation despite phylogenetic, biogeographic, and paleogeographic uncertainty. *Evolution* 72: 2343–2359.
- Lavor P, Calvente A, Versieux LM, Sanmartín I. 2018. Bayesian spatio-temporal reconstruction reveals rapid diversification and Pleistocene range expansion in the widespread columnar cactus *Pilosocereus*. *Journal of Biogeography* 46: 238–250.
- Lewis LR, Behling E, Gousse H, Qian E, Elphick CS, Lamarre J-F, Bêty J, Liebezeit J, Rozzi R, Goffinet B. 2014. First evidence of bryophyte diaspores in the plumage of transequatorial migrant birds. *PeerJ (Corta Madera, CA and London)* 2: 424.
- Liu H, Li S, Ugolini A, Momtazi F, Hou Z. 2018. Tethyan closure drove tropical marine biodiversity: Vicariant diversification of intertidal crustaceans. *Journal of Biogeography* 45: 941–951.
- Magallón S, Gómez-Acevedo S, Sánchez-Reyes LL, Hernández-Hernández T. 2015. A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. *New Phytologist* 207: 437–453.
- Mairal M, Pokorný L, Aldasoro JJ, Alarcón M, Sanmartín I. 2015. Ancient vicariance and climate-driven extinction explain continental-wide disjunctions in Africa: The case of the Rand Flora genus *Canarina* (Campanulaceae). *Molecular Ecology* 24: 1335–1354.
- Mairal M, Sanmartín I, Pellissier L. 2017. Lineage-specific climatic niche drives the tempo of vicariance in the Rand Flora. *Journal of Biogeography* 44: 911–923.
- Manafzadeh S, Salvo G, Conti E. 2014. A tale of migrations from east to west: The Irano-Turanian floristic region as a source of Mediterranean xerophytes. *Journal of Biogeography* 41: 366–379.
- Manafzadeh S, Staedler YM, Conti E. 2017. Visions of the past and dreams of the future in the Orient: The Irano-Turanian region from classical botany to evolutionary studies: Visions of past and dreams of future in Orient. *Biological Reviews* 92: 1365–1388.
- May MR, Höhna S, Moore BR. 2016. A Bayesian approach for detecting the impact of mass-extinction events on molecular phylogenies when rates of lineage diversification may vary. *Methods in Ecology and Evolution* 7: 947–959.
- Meseguer AS, Lobo JM, Cornuault J, Beerling D, Ruhfel BR, Davis CC, Josselin E, Sanmartín I. 2018. Reconstructing deep-time palaeoclimate legacies in the clusioid Malpighiales unveils their role in the evolution and extinction of the boreotropical flora. *Global Ecology and Biogeography* 27: 616–628.
- Meulenkaamp JE, Sissingh W. 2003. Tertiary palaeogeography and tectonostratigraphic evolution of the Northern and Southern Peri-Tethys platforms and the intermediate domains of the African–Eurasian convergent plate boundary zone. *Palaeogeography, Palaeoclimatology, Palaeoecology* 196: 209–228.
- Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 *Gateway Computing Environments Workshop (GCE)* 1–8.
- Mogle MJ, Kimball SA, Miller WR, McKown RD. 2018. Evidence of avian-mediated long-distance dispersal in American tardigrades. *PeerJ (Corta Madera, CA and London)* 6: 5035.
- Moharrek F, Sanmartín I, Kazempour-Osaloo S, Nieto-Feliner G. 2019. Morphological innovations and vast extensions of mountain habitats triggered rapid diversification within the species-rich Irano-Turanian genus *Acantholimon* (Plumbaginaceae). *Frontiers in Genetics* 9: 698.
- Morley RJ. 2000. *Origin and evolution of tropical rain forests*. Pretoria, South Africa: John Wiley and Sons.
- Mouthereau F, Lacombe O, Vergés J. 2012. Building the Zagros collisional orogen: Timing, strain distribution and the dynamics of Arabia/Eurasia plate convergence. *Tectonophysics* 532–535: 27–60.
- Müller J, Müller K, Neinhuis C, Quandt D. 2005. PhyDE-Phylogenetic data editor. Available from <http://www.Phyde.De> [accessed February 12, 2021].
- Nie ZL, Deng T, Meng Y, Sun H, Wen J. 2013. Post-Boreotropical dispersals explain the pantropical disjunction in *Paederia* (Rubiaceae). *Annals of botany* 111: 873–886.
- Noben S, Kessler M, Quandt D, Weigand A, Wicke S, Krug M, Lehnert M. 2017. Biogeography of the Gondwanan tree fern family Dicksoniaceae—A tale of vicariance, dispersal and extinction. *Journal of Biogeography* 44: 2648–2659.
- Nylander JAA. 2004. *MrModeltest v2*. Uppsala: Evolutionary Biology Center, Uppsala University.
- Pelster PB, Abbott RJ, Comes HP, Milton JJ, Möller M, Looseley ME, Cron GV, Barcelona JF, Kennedy AH, Watson LE, Barone R, Hernández F, Kadereit JW. 2012. The genetic ghost of an invasion past: Colonization and extinction revealed by historical hybridization in *Senecio*. *Molecular Ecology* 21: 369–387.
- Peterson AT. 2006. Uses and requirements of ecological niche models and related distributional models. *Biodiversity Informatics* 3: 369–387.
- Plana V, Gascoigne A, Forrest LL, Harris D, Pennington RT. 2004. Pleistocene and pre-Pleistocene *Begonia* speciation in Africa. *Molecular Phylogenetics and Evolution* 31: 449–461.

- Pokorny L, Riina R, Mairal M, Meseguer AS, Culshaw V, Cendoya J, Serrano M, Carbajal R, Ortiz S, Heuertz M, Sanmartín I. 2015. Living on the edge: Timing of Rand Flora disjunctions congruent with ongoing aridification in Africa. *Frontiers in Genetics* 6: 154.
- Popp M, Mirre V, Brochmann C. 2011. A single Mid-Pleistocene long-distance dispersal by a bird can explain the extreme bipolar disjunction in crowberries (*Empetrum*). *Proceedings of the National Academy of Sciences of the United States of America* 108: 6520–6525.
- Pyron RA, Costa GC, Patten MA, Burbrink FT. 2015. Phylogenetic niche conservatism and the evolutionary basis of ecological speciation: Niche conservatism and speciation. *Biological Reviews* 90: 1248–1262.
- Rambaut A, Drummond A, Suchard M. 2014. Tracer v1. 6. Available from <http://beast.bio.ed.ac.uk> [accessed February 12, 2021].
- Ramírez-Barahona S, Sauquet H, Magallón S. 2020. The delayed and geographically heterogeneous diversification of flowering plant families. *Nature Ecology & Evolution* 4: 1232–1238.
- Ree RH, Smith SA. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology* 57: 4–14.
- Revell LJ. 2012. phytools: An R package for phylogenetic comparative biology (and other things): phytools: R package. *Methods in Ecology and Evolution* 3: 217–223.
- Riina R, Villaverde T, Rincón-Barrado M, Molero J, Sanmartín I. 2020. More than one sweet tabaiba: Disentangling the systematics of the succulent dendroid shrub *Euphorbia balsamifera*. *Journal of Systematics and Evolution*. <https://doi.org/10.1111/jse.12656>
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61: 539–542.
- Rundell RJ, Price TD. 2009. Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends in Ecology and Evolution* 24: 394–399.
- Rydin C, Razafimandimbison SG, Khodabandeh A, Bremer B. 2009. Evolutionary relationships in the Spermacoceae alliance (Rubiaceae) using information from six molecular loci: insights into systematic affinities of *Neohymenopogon* and *Mouretia*. *Taxon* 58: 793–810.
- Rydin C, Wikström N, Bremer B. 2017. Conflicting results from mitochondrial genomic data challenge current views of Rubiaceae phylogeny. *American Journal of Botany* 104: 1522–1532.
- R Development Core Team. 2017. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. <http://www.R-project.org/>
- Sanmartín I. 2003. Dispersal vs. vicariance in the Mediterranean: Historical biogeography of the Palearctic Pachydeminae (Coleoptera, Scarabaeoidea). *Journal of Biogeography* 30: 1883–1897.
- Sanmartín I. 2012. Historical biogeography: Evolution in time and space. *Evolution: Education and Outreach* 5: 555–568.
- Sanmartín I, Anderson CL, Alarcón M, Ronquist F, Aldasoro JJ. 2010. Bayesian island biogeography in a continental setting: The Rand Flora case. *Biology Letters* 6: 703–707.
- Sanmartín I, Meseguer AS. 2016. Extinction in phylogenetics and biogeography: From timetrees to patterns of biotic assemblage. *Frontiers in Genetics* 7: 35.
- Sanmartín I, Ronquist F. 2004. Southern Hemisphere biogeography inferred by event-based models: Plant versus Animal patterns. *Systematic Biology* 53: 216–243.
- Schenk JJ. 2016. Consequences of secondary calibrations on divergence time estimates. *PLOS One* 11: 0148228.
- Schnitzler J, Graham CH, Dormann CF, Schifffers K, Linder H. 2012. Climatic niche evolution and species diversification in the Cape flora, South Africa. *Journal of Biogeography* 39: 2201–2211.
- Scott L. 1999. Vegetation history and climate in the Savanna biome South Africa since 190,000 ka: a comparison of pollen data from the Tswaing Crater (the Pretoria Saltpan) and Wonderkrater. *Quaternary International* 57: 215–223.
- Scott L, Vogel JC. 2000. Evidence for environmental conditions during the last 20 000 years in Southern Africa from 13c in fossil hyrax dung. *Global and Planetary Change* 26: 207–215.
- Senut B, Pickford M, Ségalen L. 2009. Neogene desertification of Africa. *Comptes Rendus Geoscience* 341: 591–602.
- Sepulchre P, Ramstein G, Fluteau F, Schuster M, Tiercelin JJ, Brunet M. 2006. Tectonic uplift and Eastern Africa aridification. *Science* 313: 1419–1423.
- Shevenell AE, Kennett JP, Lea DW. 2008. Middle Miocene ice sheet dynamics, deep-sea temperatures, and carbon cycling: A Southern Ocean perspective. *Geochemistry, Geophysics, Geosystems* 9: Q02006.
- Silvestro D, Cascales-Miñana B, Bacon CD, Antonelli A. 2015. Revisiting the origin and diversification of vascular plants through a comprehensive Bayesian analysis of the fossil record. *New Phytologist* 207: 425–436.
- Smith SA, Beaulieu JM, Donoghue MJ. 2010. An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants. *Proceedings of the National Academy of Sciences of the United States of America* 107: 5897–5902.
- Smith TB, Schneider CJ, Holder K. 2001. Refugial isolation versus ecological gradients. In: Hendry AP, Kinnison MT eds. *Micro-evolution rate, pattern, process*. Dordrecht: Springer. 383–398.
- Stadler T. 2009. On incomplete sampling under birth–death models and connections to the sampling-based coalescent. *Journal of Theoretical Biology* 261: 58–66.
- Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML web servers. *Systematic Biology* 57: 758–771.
- Thulin M. 1998. *Gaillonia* (Rubiaceae-Paederieae) in Africa and Arabia. *Nordic Journal of Botany* 18: 31–38.
- Thomas JC, Grasso J-R, Bossu R, Martinod J, Nurtaev B. 1999. Recent deformation in the Turan and South Kazakh platforms, western central Asia, and its relation to Arabia-Asia and India-Asia collisions. *Tectonics* 18: 201–214.
- Thompson JD. 2005. *Plant evolution in the Mediterranean*. Pretoria, South Africa: Oxford University.
- 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.
- Villaverde T, Escudero M, Martín-Bravo S, Jiménez-Mejías P, Sanmartín I, Vargas P, Luceño M. 2017a. Bipolar distributions in vascular plants: A review. *American Journal of Botany* 104: 1680–1694.
- Villaverde T, González-Moreno P, Rodríguez-Sánchez F, Escudero M. 2017b. Niche shifts after long-distance dispersal events in bipolar sedges (*Carex*, Cyperaceae). *American Journal of Botany* 104: 1765–1774.
- Villaverde T, Pokorny L, Olsson S, Rincón-Barrado M, Johnson MG, Gardner EM, Wickett NJ, Molero J, Riina R, Sanmartín I. 2018. Bridging the micro- and macroevolutionary levels in

- phylogenomics: Hyb-Seq solves relationships from populations to species and above. *New Phytologist* 220: 636–650.
- Warren DL, Glor RE, Turelli M. 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution* 62: 2868–2883.
- Weiner J, Weiner MJ. 2017. R Package ‘pca3d’.
- Wiens JJ. 2004. Speciation and ecology revisited: Phylogenetic niche conservatism and origin of species. *Evolution* 58: 193–197.
- Wiens JJ, Ackerly DD, Allen AP, Anacker BL, Buckley LB, Cornell HV, Damschen EI, Davies T, Grytnes JA, Harrison SP, Hawkins BA, Holt RD, McCain CM, Stephens PR. 2010. Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters* 13: 1310–1324.
- Wiens JJ, Donoghue MJ. 2004. Historical biogeography, ecology and species richness. *Trends in Ecology and Evolution* 19: 639–644.
- Wikström N, Bremer B, Rydin C. 2020. Conflicting phylogenetic signals in genomic data of the coffee family (Rubiaceae). *Journal of Systematics and Evolution* 58: 440–460.
- Wikström N, Kainulainen K, Razafimandimbison SG, Smedmark JE, Bremer B. 2015. A revised time tree of the asterids: Establishing a temporal framework for evolutionary studies of the coffee family (Rubiaceae). *PLoS One* 10: e0126690.
- Wikström N, Savolainen V, Chase MW. 2001. Evolution of the angiosperms: Calibrating the family tree. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 268: 2211–2220.
- Wiley EO, Lieberman BS. 2011. *Phylogenetics: Theory and practice of phylogenetic systematics*. Hoboken, NJ: John Wiley & Sons.
- Willis KJ, MacDonald GM. 2011. Long-term ecological records and their relevance to climate change predictions for a warmer world. *Annual Review of Ecology, Evolution, and Systematics* 42: 267–287.
- Wolfe KH, Li WH, Sharp PM. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences of the United States of America* 84: 9054–9058.
- Zachos JC, Dickens GR, Zeebe RE. 2008. An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature* 451: 279–283.

Supplementary Material

The following supplementary material is available online for this article at <http://onlinelibrary.wiley.com/doi/10.1111/jse.12747/supinfo>:

Appendix SI. Expanded Introduction and Materials and Methods, including Tables I.1 (list of specimens studied) and I.2 (primer and PCR amplification conditions).

Appendix SII. Extended results from the phylogenetic analyses.

Appendix SIII. Extended results from the climatic niche reconstruction and biogeographic analyses.