https://doi.org/10.35662/unine-thesis-2480

Historical biogeography of a species rich plant genus from the montane forests of the tropical Andes, *Macrocarpaea* (Gentianaceae)

Thèse présentée à la Faculté des Sciences

Institut de Biologie

Université de Neuchâtel, Suisse

Par

Julien Vieu

Composition du jury :

PD Dr Grant Jason, directeur de thèse

Prof Dr Willi Yvonne, Co-directeur et rapporteur

Prof Dr Kessler Félix, rapporteur

Dr. Struwe Lena, rapporteur

Dr. Hughes Colin, rapporteur

PD Dr Parisod Christian, rapporteur

Université de Neuchâtel, Switzerland

2015

UNIVERSITÉ DE NEUCHÂTEL

Faculté des sciences
Secrétariat-décanat de Faculté
Rue Emile-Argand 11
2000 Neuchâtel - Suisse
Tél: + 41 (0)32 718 2100
E-mail: secretariat.sciences@unine.ch

IMPRIMATUR POUR THESE DE DOCTORAT

La Faculté des sciences de l'Université de Neuchâtel autorise l'impression de la présente thèse soutenue par

Monsieur Julien VIEU

Titre:

"Historical biogeography of a species rich plant genus from the montane forests of the tropical Andes, *Macrocarpaea* (Gentianaceae)"

sur le rapport des membres du jury composé comme suit:

- PD Dr Jason Grant, directeur de thèse, Université de Neuchâtel
- Prof. Yvonne Willi, co-directrice de thèse, Université de Neuchâtel
- Prof. Felix Kessler, Université de Neuchâtel
- PD Dr Christian Parisod, Université de Neuchâtel
- Dr Colin Hughes, Université de Zürich
- Prof. Lena Struwe, Rutgers University, New Brunswick, USA

Neuchâtel, le 9 juin 2015

Le Doyen, Prof. B. Colbois

B. Colbois

Table of content

Acknowledgments

After nearly five years, my PhD thesis has come to its end. It has been a tremendous adventure punctuated sometimes by laborious and sometimes by pleasant episodes that all have contributed to my development as a young researcher but also as a person. The completion of this work would have never been made possible without the help of several persons and institutions, and now I would like to warmly thank them for their precious help.

I would like to thank Jason Grant, my PhD advisor for giving me the chance to conduct this research. You also transmitted me your love and passion for the wonderful Andean plant biodiversity and I am grateful for that. Thanks to Colin Hughes who kindly accepted to be member of my PhD jury but also for its generous availability for discussions and invaluable comments on several manuscripts of this PhD thesis. Colin, I learned a lot during our scientific exchanges and I only regret that we did not have more discussions. Thanks to the other members of the jury for the interesting and debated discussion we had during my PhD defense. More specifically, I thank Lena Struwe for discussions during my PhD and comments on the manuscript, and Christian Parisod for insightful discussions here in Neuchâtel. Your door was always open and I appreciated our discussions about various evolutionary topics. This helped me to think about my work on a broader scientific context.

Thanks to my friend David Dérousseaux for helping during the two months of field work in Peru. This fieldwork was a real adventure with its good moments (ie: La cañas) and misfortunes (ie: Los ronderos, the car). Your sense of logistics was extremely valuable and I also appreciated to share these experiences with a good friend. Thanks to Eric Franck Rodriguez who kindly accepted to guide us throughout Peru for one month and for housing David and me a few days at your house.

Thanks to the present and past members of the Laboratoire de botanique évolutive de l'Université de Neuchâtel for discussions, support during hard times and fun during good times. Specifically, thanks to all my office-mates Antoine Paccard, Emmanuel Bonjour, Marco Fracassetti and Guillaume Wos. Special thanks also to the other side of the corridor to Olivier Bachmann and the youngster Benjamin Dauphin.

Thanks to the Université de Neuchâtel for giving me logistics and financial support to conduct this PhD. Thanks also to the Basler Stiftung für biologische Forschung, the FNS overhead of the Université de Neuchâtel and the Fondation Dr. Joachim de Giacomi for financial support.

Last but not least, thank to my beloved Caroline Cloché who supported me all along these five years.

Abstract

Differences in species diversity among regions and among taxa have long fascinated evolutionary biologists, but the mechanisms that generated these patterns remain insufficiently understood. The middle elevation montane forests of the tropical Andes (MMF) stand out as the region of the world harbouring the highest level of both species diversity and endemism which make this region an appropriate place to address mechanisms that favour high species diversity. The main goal of this thesis was to test whether the high plant diversity of the MMF was the result of the slow accumulation of species or through episodes of rapid diversification. Additionally this thesis investigated the contribution of geographical processes such as dispersal and allopatric speciation vs. adaptive processes such as ecological speciation to the diversification of plants in the region. We used the plant genus *Macrocarpaea* (Gentianaceae) as a study system to tackle these questions. *Macrocarpaea* occurs in most montane regions of the Neotropics but its diversity (96 species) is within the MMF, suggesting that the genus radiated in the region. Also, the genus has a broad altitudinal distribution (500-3500m) but most species have a much narrower altitudinal range indicating that adaptive divergence along elevation gradients might have been common.

We first reconstructed a dated phylogenetic hypothesis that we subsequently used to infer the historical biogeography of the genus and to estimate diversification rates through time and among lineages. We found a pattern of diversification consistent with the signature of a radiation for *Macrocarpaea* in the Andean MMF. The radiation coincides with a period of rapid colonisation and range expansion across the entire extant distribution of the genus in the Andes starting some 7.2 million years ago. Furthermore, analyses support allopatric founder-event speciation as the dominant process contributing to the geographic and phylogenetic structure of the genus. We suggest that the likely rapid establishment of the MMF in the late Miocene when the Andes attained critical elevation to modify regional climates, provided large new areas of suitable habitat for *Macrocarpaea* to quickly colonise through repeated founder-events. This wave of colonisation triggered rapid diversification and as the range of the MMF became progressively occupied, the diversification rate slowed.

Then we investigated the contribution of adaptive divergence and niche conservatism in this radiation. We used morphological and climatic data in a phylogenetic comparative method framework to compare a set of evolutionary scenarios of various levels of complexity. We showed that the hypothesis of an adaptive radiation for *Macrocarpaea* in the MMF is very unlikely. The genus has remained confined to the upper montane forests (UMF>1800m) during more than the half of its history possibly due to evolutionary constraints. Later, and coincidental with the beginning of the Pleistocene climatic oscillations, a derived clade, the *micrantha* alliance, successfully colonized and radiated in the lower montane forests (LMF<1800m). This colonisation has been accompanied by the evolution of a new leaf phenotype unique to the species of the *micrantha* alliance that likely represents adaptation to life in this new adaptive zone. Therefore, our results suggested that niche conservatism and geographical processes have dominated most of the diversification history of *Macrocarpaea*, but that a rare adaptive divergence event allowed a transition into a new adaptive zone and enabled a second subsequent radiation in this zone through geographical processes.

Finally, we investigated the potential impact of the Pleistocene climatic oscillation (PCO) on the population structure in lower montane forest species. We used amplified fragment length polymorphism molecular makers (AFLP's) to investigate the phylogeography of 12 species in the *micrantha* alliance. We showed that some species identified on the basis of cryptically morphological differences, displayed complex and potentially reticulated relationships, requiring further examination. We also showed that populations of a species endemic to a single valley connected to the dry system of the Rio Marañón in northern Peru have low levels of gene diversity. We suggested that this is due to a recent demographic bottleneck resulting from the contraction of the montane wet forests into refugia as consequence of the expansion of the dry system into the valley during the glacial cycles of the PCO. This hypothesis needs further study, but its confirmation would imply that different valleys of the Andes might have responded in very different manners to the PCO.

Abbreviations

AFLP: Amplified fragment length polymorphism

AHZ: Amotape-Huancabamba zone

BM: Brownian motion

BMM: Brownian motion with multiple rates

DTT: Disparity through time

HEG: High elevation grassland

LDG: Latitudinal diversity gradient

LGM: Last glacial maximum

LMF: Lower montane forest

Ma: Million years ago

MCC tree: Maximum clade credibility tree

MDI: Morphological disparity index

MMF: Middle elevation montane forest

MRCA: Most recent common ancestor

Myr: Million years

OU: Ornstein-Uhlenbeck

OUM: Ornstein-Uhlenbeck with multiple optima

PCO: Pleistocene climatic oscillations

SDTF: Seasonally dry tropical forest

SGS: Spatial genetic structure

SLA: Specific leaf area

UFL: Upper forest line

UMF: Upper montane forest

Ya: Years ago

Introduction

The exceptional biodiversity of the tropics or the latitudinal diversity gradient

Why are there so many species in the tropics? The exceptional biodiversity of the tropics in comparison to temperate regions has long fascinated and questioned ecologists and the evolutionary biologists. The latitudinal diversity gradient (LDG) is certainly one of the most striking biological patterns by its pervasiveness among taxa and regions (Brown, 2014). The LDG has been at the centre of numerous programs of research for decades and have resulted in an impressive number of hypotheses (Willig *et al.*, 2003).

On a regional scale, the distribution of the species diversity as we can admire it today is essentially the by-product of the interaction of three processes: speciation, extinction and migration (Jablonski *et al.*, 2006; Wiens *et al.*, 2009). On the local scale, the coexistence of species is controlled by ecological processes (Ricklefs, 2004). Two main hypotheses have been proposed to explain the great diversity of the tropics. The "museum" hypothesis considers that the high diversity of the tropics is the outcome of the slow and progressive accumulation of species in a historically stable environment in which species were exposed to relatively lower extinction rates than at temperate latitudes (Mittelbach *et al.*, 2007). According to this scenario, the factors that limit the diversity would principally be the dispersal and adaptation abilities of species and the carrying capacity of the habitats (Wiens & Donoghue, 2004). On the other hand, the "cradle" hypothesis considers that the higher diversity of the tropics is the outcome of higher speciation rates that can arise from a variety of factors, such as: 1) the stronger genetic drift in populations of smaller size in the tropics (Dick & Heuertz, 2008); 2) the larger area of continuous continental mass exposed to tropical climate that offer more opportunities of isolation (Chown & Gaston, 2000); 3) the higher climatic productivity; 4) the higher amplitude of the climatic variations in the tropics (Jansson & Davis, 2008); 5) or the higher rate of UV irradiation that could stimulate the molecular evolution (Lee & Lowry, 1980)

More recently, it has also been emphasized that the higher diversity of the tropics could be simply explained by historical factors such as the time of residence or colonization of a region. The rational for the "tropical conservatism" hypothesis, is that much of the world was tropical for a long period before the present (Wiens & Donoghue, 2004). Many extant tropical groups that have high diversity originated in the tropics and thus had more time to accumulate diversity. Adding to that the rarity of successful adaptive transition from tropical

regions to temperate ones (Donoghue, 2008), there is no need to invoke differences in speciation and/or extinction rate between regions.

Other have also suggested that the dichotomy "museum" vs. "cradle" is misleading as the tropics are likely to be at the same time "cradle" and "museum" of biodiversity (Jablonski *et al.*, 2006). Contrarily to the "tropical conservatism" hypothesis, the "out of the tropic" hypothesis consider that there is more speciation, less extinction in the tropics but also that most of the current biodiversity including extra tropical biodiversity originated in the tropics.

Beside the generality of the LDG at the global scale it should be stressed that biodiversity is not evenly distributed among regions. For example, the Neotropics have about three times more tree species than the Paleotropics (Fine & Ree, 2006). There are also marked variations in species diversity inside regions. For example, montane ranges inside tropical regions harbour in general more species than the lowlands (Orme *et al.*, 2005). The ultimate processes that explain the variation in diversity at this more local scale remain speciation, extinction and migration. But the factors influencing these processes are necessarily more contingent (Gaston, 2000). Therefore, it is thought that historical factors such as the history of the dynamism of the landscape are of primary importance to explain the differences in species diversity at this scale (Linder, 2008).

The phylogenetic approach a window into deep to intermediate evolutionary time

A classical approach to study the factors responsible for the pattern of diversity consists in using time-calibrated phylogenies. Time-calibrated phylogenies are built from molecular DNA markers with methods that make use of time calibration information (usually fossils) and statistical models of molecular clock to estimate the time of divergence between related organisms (Ho & Philips, 2009).

Time-calibrated phylogenies not only provide hypothesis concerning the relatedness and the age of organism, but they can also be used in combination with statistical model to infer the rate of diversification of a group of organisms. Afterward it is thus possible to compare the rhythm of diversification (diversification= speciation-extinction) between organisms or between regions and to test some of the hypotheses mentioned above. Unfortunately, the diversification rate estimated from phylogenies alone gives only to date partial information concerning the contribution of speciation and extinction to its variation. While speciation rates are relatively

well-estimated from time calibrated phylogenies, extinction rates have been proven to be unreliable (Rabosky, 2010). Indeed, extinctions rate estimated from phylogenies are often null or at least very low which is in strong contradiction with fossil record that indicate that extinction are extremely common (Quental & Marshall, 2010).

Nevertheless, diversification rates estimated from time calibrated phylogenies remains extremely useful to identify the factors that might enhance or lessen species diversity. For example, this approach is commonly used to study radiations, when a group of species quickly diversify from a common ancestor (Grant & Grant 2006; Seehausen, 2006). This large accumulation of species in a short amount of time let a signal of massive increase in the diversification rate also known as "early burst", that can be detected with statistical methods that allow the rate varying through time and among branches of a phylogeny (Rabosky *et al.*, 2014). A second signature of radiation that is commonly detected in diversification rates is a "slowdown" that follows in general relatively quickly the "early burst". The "early burst" and the "slowdown" in diversification rates have been interpreted in the context of the adaptive radiation theory as evidence for the phase of ecological niche space exploration and the phase of progressive ecological niche space saturation predicted by the theory (Gavrilets & Losos, 2009). But more recently it has been emphasised that many other processes can let the same signal in phylogenies (Moen & Morlon, 2014). For example, non-adaptive radiations ensuing from processes of dispersal associated with allopatric speciation can also produce the early burst" and the "slowdown" patterns in the diversification rate (Comes *et al.*, 2008). In order to disentangle the adaptive or the non-adaptive (geographic) nature of radiations additional data are required.

An additional prediction of the adaptive radiation theory is the rapid accumulation of morphological and ecological differences ("early burst" of disparity) as the species adapt to different part of the available ecological space (Gavrilets & Losos, 2009). Such a disparity is not expected under a non adaptive radiation, or at least to a lesser extent (Gittenberger, 1991). Time calibrated phylogenies here again are very useful. They can be used in combination with morphological and/or ecological (ie: climate) information of extant species to estimates with statistical approach the evolution of disparity through the history of a particular taxa. In principle it is thus possible to test both the prediction of "early burst" of diversification and/or "early burst" of disparity and to rule on the presence or absence of a radiation then on its adaptive or non-adaptive nature.

The population genetics approach, a complementary approach

A common critique of the phylogenetic approach to estimate the contribution of adaptive (induced by natural selection) and non-adaptive evolution (geographically induced) is that it is not possible to verify if the morphological or ecological (hereafter, traits) differences between close related species evolved at or only after speciation (Losos & Glor, 2003). Thus it is not clear whether trait divergences inferred from phylogenies are involved in the speciation process. Others have argued that because several processes can lead to identical phylogenetic patterns, the phylogenetic approach can only help to circumscribe or exclude a list of potential scenario but rarely allow pinpointing the evolutionary processes responsible for species divergence (Losos, 2011).

In the prospect to identify the processes leading to reproductive isolation, population genetics is a promising complementary approach to molecular phylogenies (Via, 2009). Indeed, this approach allows studying the mechanism of reproductive isolation before their completion that might help to identify with greater accuracy the processes that can ultimately lead to species divergence. Population genetics give also access to the genetic diversity and genetic connectivity of the populations which allow addressing demographic and historical processes such as: 1) demographic bottleneck that can result from founder events or range contractions (Hewitt, 2000), or 2) gene flow that can impede in some case population divergence and in other case rescue populations with low diversity from inbreeding (Lenormand, 2002).

The middle elevation montane forests of the tropical Andes, the hottest hotspot in the world.

Among the tropical regions, the middle-elevation montane forests of the tropical Andes (MMF) are a fascinating place to address the question related to the origin and the maintenance of the biodiversity. The Andes formed mainly through a crustal thickening associated with Cenozoic (65-0 Ma) subduction and convergence between the Nazca Plate and the South American Plate (Poulsen *et al.*, 2010). However, the most intense peaks of Andean uplift took place during the late middle Miocene (~12 Ma) and early Pliocene (~4.5 Ma; Hoorn *et al.*, 2010). It is only between the Miocene and the early Pliocene that the Andes reach sufficient elevations (~50-70% to its modern elevation) to profoundly modifying the climate of the entire continent (Poulsen *et al.*, 2010).

From that time onward the Andes blocked the air masses loaded with moisture coming from the Amazon that triggered convective precipitations and a massive increase in precipitations along the eastern flank of the Andes. Likely, it is only at this time that continuous band of tropical montane forest developed on the eastern slopes of the Andes. Pollen record indicates the appearance of emblematic elements of the MMF, such as the gymnosperm genus *Podocarpus* (Podocarpaceae) at this time (Hooghiemstra *et al.*, 2006).

The MMF are humid to per-humid forests (>1500 mm per year) found between 1000 to 3500 m from northern Venezuela to northern Bolivia. These forests are marked by steep environmental gradients that vary with elevation. For example temperature, atmosphere pressure, soil pH and litter quality decrease while diurnal temperature range, precipitation, solar radiations and wind speed increase with increasing elevation (Beck & Richter 2008, Gerold 2008, Kessler et al. 2011). These environmental factors in turns induce in an important stratification of the vegetation in elevation belts with different physiognomy and floristic composition. For example, leaf size, canopy height and stem diameter decrease while leaf thickness, stem density, and canopy gaps increase with increasing elevation (Leuschner & Moser 2008, Asner et al., 2014). A variety of vegetation belts classification have been proposed (Webster *et al.*, 1993), here we will principally refer to the lower montane forests (LMF<1800m), the upper montane forests (UMF, ~1800-3200m) and the subparamo shrub forest (>3200m) as they represents the more "marked" discontinuity.

The MMF houses an outstanding diversity in general and diversity of plant in particular. It is one of the rare regions of the world that meet at same time the three criteria used for the recognition of hotspot of biodiversity at least in birds (Orme *et al.*, 2007): 1) total species richness; 2) threatened species richness; 3) endemic species richness. In comparison, it is relatively hard to find estimates of the total number of plant species in the MMF. A compilation of the estimations from Beck & Richter (2008) for Ecuador and Peru (10,500-16,000 species), from Gerold (2008) for Bolivia (10,000 species) and a "guess estimation" of comparable figures for Colombia and Venezuela together (10,000 species?) gives a rough total of 30,500 to 36,000 species. But to have a real idea of the number of species of the MMF, it would be necessary to verify which species occur in several countries. Probably 20,000 species is a very conservative number for the total diversity of the MMF. If we now consider that the great majority of this species are endemic of the MMF, this mean that a considerable number of species have originated in these forests in a relatively short amount of time.

Two main historical factors have been put forward to explain the exceptional diversity of the MMF (Gentry, 1982). The Andean uplift, which may have promoted speciation through: a) vicariant fragmentation and isolation of lowland taxa by mountain uplift; and b) allopatric speciation during range expansion of taxa adapted to newly available and greatly expanded montane habitats; and c) parapatric speciation along the steep and extended ecological gradients that characterize the Andes. Later, the Pleistocene climatic oscillations might have further enhanced speciation through allopatric speciation resulting from repeated glacial contractions and interglacial expansions repeatedly fragmenting and reuniting species distributions (Rull, 2011).

Studies using phylogenetic approach on a variety of organism groups from the MMF have shown that the Andean uplift profoundly affected the rhythm of diversification in this taxa (e.g. butterflies: *Lymanopoda*, Casner & Pyrcz 2010; birds: *Pionus*, Ribas *et al.*, 2007; frogs: Dendrobatidae, Santos *et al.*, 2009, plants: Bromeliaceae, Givnish *et al.*, 2014). However, the mechanisms behind this diversification rates increase remain poorly understood. Are these diversification rate increases induced by adaptive or geographical processes or both? Are this group still diversifying at high rate or are they subject to diversity dependence negative controls? Did the Pleistocene climatic oscillation impact the diversity in the MMF? If yes, did it stimulate or inhibit it? Clearly more studies using multidisciplinary approach are needed to help understanding the origin of the great diversity of the MMF.

The study model, the plant genus Macrocarpaea

In order to answer these questions, I use the genus *Macrocarpaea* (Gentianaceae) as a study system. *Macrocarpaea* is a group of 118 species of shrubs and small trees species restricted to the montane regions of the Neotropics. Eight species occur in the montane Atlantic forests of Brazil, six in the Pantepui of the Guayana Shield, five in Mesoamerica (Costa Rica and Panama), and three in the Greater Antilles of the Caribbean (one species on each island of Cuba, Hispaniola, and Jamaica). The majority of species (96 or 81%) occur in the middle elevation montane forests of the tropical Andes. This very high number of species in the MMF suggests that the genus radiated in the MMF which makes this plant genus an excellent system to study the factor and the processes that regulate the diversity in this forests. Moreover the genus spans the totality of the altitudinal range of the MMF (500-3500m) but individual species are restricted to much narrower range, indicating potential altitudinal adaptive divergence.

Relatively little is known about the biology of the species of *Macrocarpaea*. Chromosome number is remarkably stable among species (N=21) despite measures performed on species representative the whole distribution of the genus (Trunz *et al.*, 2012). All the species have proteandrous flowers (Grant, 2014) with yellow to green thick funnel-shaped corolla typical of the bat pollination syndromes. The nectar sugar composition of four species from Ecuador were analysed and revealed high hexose content characteristic of others bat pollinated plant species (Wolff, 2006). Flowers attracts diurnal visitors such as hummingbirds, ants, butterflies (Grant, 2004). Also, the seeds are small (0.2–2.2 mm) and numerous (at least 10,000 per fruit).

Outline of the chapters and context

These PhD thesis aims to better understand the factors that control the dynamic of diversification of *Macrocarpaea*, a species rich plant genus form the megadiverse montane forests of the tropical Andes (MMF). It is organized in three complementary chapters.

In the **chapter 1**, we tested the hypothesis of a rapid radiation of *Macrocarpaea* in the Andes and whether the radiation was initiated by the late Miocene Andean uplift or by the Pleistocene climatic oscillations. We reconstructed a biogeographical hypothesis for the genus and we estimated the potential variations of the diversification rate. We identified a pattern of diversification consistent with the signature of a radiation. The radiation started some 7Ma and coincided with the rapid colonization of the Andes by the genus. We proposed that the radiation was induced by repeated allopatric speciation associated with dispersal during the rapid range expansion. We also show that the timing and the rate of diversification of other plant group from the same region are strikingly similar indicating a potential common trajectory for the flora of the region.

In the **chapter 2**, we tested the potential adaptive nature of the radiation of *Macrocarpaea* in the Andes. We also investigated potential evolutionary constraint that might limit adaptation between vegetations belts. We used a set morphological and climatic trait on a phylogenetic context to test several scenarios of macroevolution. We first showed that the radiation was not adaptive. Then, we demonstrate that most *Macrocarpaea* species are historically constrained to occur in the upper montane forests, but that a clade, the *micrantha* alliance successfully invaded lower montane forests. This adaptive transition has been accompanied by an unusual leaf

phenotype in *Macrocarpaea*, that likely represents an adaption to life in the lower montane forests. It might have also allowed a second wave of range expansion that possibly induced an additional burst of diversification. The age of adaptive transition corresponded to the beginning of the Pleistocene climatic oscillations (PCO). We suggested that the fluctuation of the altitudinal position of vegetation belts induced by the PCO facilitated the adaptive transition

In the **chapter 3**, we addressed the potential influence of the Pleistocene climatic oscillation (PCO) on the population demography of several species from the *micrantha* alliance. We used amplified fragment length polymorphism makers (AFLP) to investigate the phylogeography of these species. We showed that a species endemic to a valley connected with dry systems had significantly lower gene diversity than the other species. We suggest that this species went through a demographic bottleneck induced by the aridification of the valley during the PCO. We also showed that several species had complex and potentially reticulated history.

REFERENCES

Antonelli, A., Nylander, J. A., Persson, C., & Sanmartín, I. (2009). Tracing the impact of the Andean uplift on Neotropical plant evolution. *Proceedings of the National Academy of Sciences*, **106**, 9749-9754.

Asner G.P., Anderson C.B., Martin R.E., Knapp D.E., Tupayachi R., Sinca F., & Malhi 631 Y. (2014) Landscape scale changes in forest structure and functional traits along an Andes-to-Amazon elevation gradient. *Biogeosciences*, **11**, 843-856.

Brown, J. H. (2014). Why are there so many species in the tropics?. Journal of Biogeography, **41**, 8-22.

Beck E. & Richter M. (2008). Ecological aspects of a biodiversity hotspot in the Andes of southern Ecuador. *In: Gradstein, R., Homeier, J., Gansert, D. (Eds.), The Tropical Mountain Forest. Patterns and Processes in a Biodiversity Hotspot. Biodiversity and Ecology series, 2.* Gottingen Centre for Biodiversity and Ecology, Gottingen, pp. 195–217.

Casner, K. L., & Pyrcz, T. W. (2010). Patterns and timing of diversification in a tropical montane butterfly genus, *Lymanopoda* (Nymphalidae, Satyrinae). *Ecography*, **33**, 251-259.

Chown, S. L., & Gaston, K. J. (2000). Areas, cradles and museums: the latitudinal gradient in species richness. *Trends in Ecology & Evolution*, **15**, 311-315.

Comes, H. P., Tribsch, A., & Bittkau, C. (2008). Plant speciation in continental island floras as exemplified by Nigella in the Aegean Archipelago. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **363**, 3083-3096.

Dick, C. W., & Heuertz, M. (2008). The complex biogeographic history of a widespread tropical tree species. *Evolution*, **62**, 2760-2774.

Donoghue, M. J. (2008). A phylogenetic perspective on the distribution of plant diversity. *Proceedings of the National Academy of Sciences*, **105**, 11549-11555.

Fine, P. V., & Ree, R. H. (2006). Evidence for a time-integrated species-area effect on the latitudinal gradient in tree diversity. *The American Naturalist*, **168**, 796-804.

Gaston, K. J. (2000). Global patterns in biodiversity. *Nature*, **405**, 220-227.

Gavrilets, S., & Losos, J. B. (2009). Adaptive radiation: contrasting theory with data. *Science*, **323**, 732-737.

Gentry, A. H. (1982). Neotropical floristic diversity: phytogeographical connections between Central and South America, Pleistocene climatic fluctuations, or an accident of the Andean orogeny?. *Annals of the Missouri Botanical Garden*, **69**,557-593.

Gerold G. (2008) Soil, climate and vegetation of tropical montane forests – a case study from the Yungas, Bolivia. *In: Gradstein, R., Homeier, J., Gansert, D. (Eds.), The Tropical Mountain Forest. Patterns and Processes in a Biodiversity Hotspot. Biodiversity and Ecology series, 2.* Gottingen Centre for Biodiversity and Ecology, Gottingen, pp. 137–162.

Gittenberger, E. (1991). What about non-adaptive radiation?. *Biological Journal of the Linnean Society*, **43**, 263-272.

Givnish T.J., Barfuss M.H.J., Van E.B., Riina R., Schulte K., Horres R., Gonsiska P.A., Jabaily R.S., Crayn D.M., Smith J.A.C., Winter K., Brown G.K., Evans T.M., Holst B.K., Luther H., Till W., Zizka G., Berry P.E., Sytsma K.J. (2014). Adaptive radiation, correlated and contingent evolution, and net species diversification in Bromeliaceae. *Molecular Phylogenetics and Evolution*, **71**, 55–78.

Grant, P. R., & Grant, B. R. (2006). Evolution of character displacement in Darwin's finches. *Science*, **313**, 224-226.

Grant, J. R. (2014). A monographic revision of the neotropical genus *Macrocarpaea* (Gentianaceae) in Ecuador. *In: The Gentianaceae-Volume 1: Characterization and Ecology* , Springer Berlin Heidelberg. pp 37-147.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, **405**,907-913.

Ho S.Y.W. & Phillips M.J. (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology*, **58**, 367-380.

Hooghiemstra, H., Wijninga, V. M., & Cleef, A. M. (2006). The paleobotanical record of Colombia: implications for biogeography and biodiversity 1. *Annals of the Missouri Botanical Garden*, **93**, 297-325.

Hoorn, C., Wesselingh, F.P., ter Steege, H., Bermudez, M.A., Mora, A., Sevink, J., Sanmartı´n, I., Sanchez-Meseguer, A., Anderson, C.L., Figueiredo, J.P., Jaramillo, C., Riff, D., Negri, F.R., Hooghiemstra, H., Lundberg, J., Stadler, T., Särkinen, T. & Antonelli, A. (2010) Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science*, **330**, 927–931.

Jablonski, D., Roy, K., & Valentine, J. W. (2006). Out of the tropics: evolutionary dynamics of the latitudinal diversity gradient. *Science*, **314**, 102-106.

Jansson, R., & Davies, T. J. (2008). Global variation in diversification rates of flowering plants: energy vs. climate change. *Ecology Letters*, **11**, 173-183.

Kessler M., Grytnes J.-A., Halloy S.R.P., Kluge J., Krömer T., León B., Macía M.J., & Young K.R. (2011) Gradients of Plant Diversity: Local Patterns and Processes. *In: Climate change and biodiversity in the tropical andes inter-american institute for global change research (IAI) and scientific committee on problems of the environment (SCOPE)*, Sao Jose dos Campos and Paris, pp. 204-219.

Lee, D. W., & Lowry, J. B. (1980). Solar ultraviolet on tropical mountains: can it affect plant speciation?. *American Naturalist*, **115**,880-883.

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, **17**, 183-189.

Leuschner, C., Zach, A., Moser, G., Homeier, J., Graefe, S., Hertel, D., Wittich, B., Soethe, N., Iost, S., Röpderstein, M., Horna, V., and Wolf, K. (2013) The carbon balance of tropical mountain forests along an altitudinal transect. *In: Bendix, J., Springer Verlag, Berlin, Heidelberg (Eds.), Ecological Studies Vol. 221*,New York, pp. 117–139.

Linder, H. P. (2008). Plant species radiations: where, when, why?. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **363**, 3097-3105.

Losos, J. B., & Glor, R. E. (2003). Phylogenetic comparative methods and the geography of speciation. *Trends in Ecology & Evolution*, **18**, 220-227.

Losos, J. B. (2011). Seeing the forest for the trees: the limitations of phylogenies in comparative biology. *The American Naturalist*, **177**, 709-727.

Mittelbach, G.G., Schemske, D.W., Cornell, H.V., Allen, A.P., Brown, J.M., Bush, M.B., Harrison, S.P., Hurlbert, A.H., Knowlton, N., Lessios, H.A., McCain, C.M., McCune, A.R., McDade, L.A., McPeek, M.A., Near, T.J., Price, T.D., Ricklefs, R.E., Roy, K., Sax, D.F., Schluter, D., Sobel, J.M. & Turelli, M. (2007). Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology letters*, **10**, 315-331.

Moen, D., & Morlon, H. (2014). Why does diversification slow down?. *Trends in ecology & evolution*, **29**, 190-197.

Moritz, C., Patton, J. L., Schneider, C. J., & Smith, T. B. (2000). Diversification of rainforest faunas: an integrated molecular approach. *Annual Review of Ecology and Systematics*, **31**, 533-563.

Orme, C.D.L., Davies, R.G., Burgess, M., Eigenbrod, F., Pickup, N., Olson, V.A., Webster, A.J., Ding, T.-S., Rasmussen, P.C., Ridgely, R.S., Stattersfield, A.J., Bennett, P.M., Blackburn, T.M., Gaston, K.J. & Owens,

I.P.F. (2005). Global hotspots of species richness are not congruent with endemism or threat. *Nature*, **436**, 1016–1019.

Poulsen, C. J., Ehlers, T. A., & Insel, N. (2010). Onset of convective rainfall during gradual late Miocene rise of the central Andes. *Science*, **328**, 490-493.

Quental, T. B., & Marshall, C. R. (2010). Diversity dynamics: molecular phylogenies need the fossil record. *Trends in Ecology & Evolution*, **25**, 434-441.

Rabosky, D. L. (2010). Extinction rates should not be estimated from molecular phylogenies. *Evolution*, **64**, 1816-1824.

Rabosky, D. L., Donnellan, S. C., Grundler, M., & Lovette, I. J. (2014). Analysis and visualization of complex macroevolutionary dynamics: an example from Australian scincid lizards. *Systematic biology*, **63**, 610-627.

Ribas, C. C., Moyle, R. G., Miyaki, C. Y., & Cracraft, J. (2007). The assembly of montane biotas: linking Andean tectonics and climatic oscillations to independent regimes of diversification in Pionus parrots. *Proceedings of the Royal Society B: Biological Sciences*, **274**, 2399-2408.

Santos, J. C., Coloma, L. A., Summers, K., Caldwell, J. P., Ree, R., & Cannatella, D. C. (2009). Amazonian amphibian diversity is primarily derived from late Miocene Andean lineages. *PLoS biology*, **7**, e1000056.

Ricklefs, R. E. (2004). A comprehensive framework for global patterns in biodiversity. *Ecology Letters*, **7**, 1-15.

Rull V. (2011) Neotropical biodiversity: timing and potential drivers. *Trends in Ecology & Evolution*, **26**, 508–844 513.

Seehausen, O. (2006). African cichlid fish: a model system in adaptive radiation research. *Proceedings of the Royal Society B: Biological Sciences*, **273**, 1987-1998.

Trunz, V., Zeltner, L., Küpfer, P., Grant, J. R. (2012). Cytotaxonomic notes on the tribe Helieae. *Harvard Papers in Botany*, **17**, 43-46.

Webster, G. L., Churchill, S. P., Balslev, H., Forero, E., & Luteyn, J. L. (1995). The panorama of Neotropical cloud forests. In *Biodiversity and conservation of Neotropical montane forests. Proceedings of a symposium, New York Botanical Garden, 21-26 June 1993*. New York Botanical Garden. pp. 53-77.

Wiens, J. J., Sukumaran, J., Pyron, R. A., & Brown, R. M. (2009). Evolutionary and biogeographic origins of high tropical diversity in Old World frogs (Ranidae). *Evolution*, **63**, 1217-1231.

Wiens, J. J., & Donoghue, M. J. (2004). Historical biogeography, ecology and species richness. *Trends in ecology & evolution*, **19**, 639-644.

Willig, M. R., Kaufman, D. M., & Stevens, R. D. (2003). Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. *Annual Review of Ecology, Evolution, and Systematics*, **34**,273-309.

Wolff, D. (2006). Nectar sugar composition and volumes of 47 species of Gentianales from a southern Ecuadorian montane forest. *Annals of Botany*, **97**, 767-777.

CHAPTER 1: Evolutionary diversification in the hyper-diverse montane forests of the tropical Andes: range expansion drives radiation of the plant genus *Macrocarpaea* (Gentianaceae)

AUTHORS NAMES: Vieu Julien C.[1], Hughes Colin E.[2] & Grant Jason R.[1]

AUTHORS AFFILIATIONS: [1]Institut de Biologie, Université de Neuchâtel, Rue Emile Argand 11, 2000 Neuchâtel, Switzerland; [2]Institute of Systematic Botany, University of Zurich, Zollikerstrasse 107, 8008 Zurich, Switzerland

This chapter is in review at Molecular Phylogenetics and Evolution

ABSTRACT

The humid mid-elevation montane forests (MMF) of the tropical Andes harbour high levels of plant species diversity contributing to the exceptional overall diversity of the tropical Andean biodiversity hotspot. However, little is known about the diversification dynamics of MMF plant lineages. Here, we use the plant genus *Macrocarpaea* (Gentianaceae) as an example to investigate patterns of plant diversification in the MMF. We sequenced 76 of 118 recognized *Macrocarpaea* species for six genetic makers to reconstruct a time-calibrated phylogeny for the genus using a secondary calibration approach and an uncorrelated relaxed clock model in BEAST. Dated trees were used to reconstruct the historical biogeography of *Macrocarpaea* with maximum likelihood methods implemented in the R package BioGeoBEARS and estimate diversification rates through time and among lineages with BAMM. We found a pattern of diversification consistent with the signature of a radiation for *Macrocarpaea* in the Andean MMF. The radiation coincides with a period of rapid colonisation and range expansion across the entire extant distribution of the genus in the Andes starting some 7.2 million years ago. Furthermore, analyses support allopatric founder-event speciation as the dominant process contributing to geographic phylogenetic structure of the genus. We suggest that the likely rapid establishment of the MMF in the late Miocene when the Andes attained critical elevation to modify regional climates provided large new areas of suitable habitat for *Macrocarpaea* to quickly colonise through repeated founder-events. This wave of colonisation triggered rapid diversification and as the range of the MMF became progressively occupied, the diversification rate slowed. Our study also indicates that plant radiations in the MMF might be older and slower than those in the very recent and fast-evolving Andean high elevation grasslands.

KEYWORDS

diversity-dependence; diversity, diversification; founder-events; historical biogeography; phylogeny

1. Introduction

Disparities in species diversity between different regions of the world have long fascinated evolutionary biologists as exemplified by the debate surrounding the underlying causes of the latitudinal gradient of species diversity (Jablonski *et al.*, 2006; Mittelbach *et al.*, 2007). Despite the generality of this broad global scale pattern, finer regional contrasts are equally striking (Gaston, 2000; Whittaker *et al.*, 2001). In this respect, biodiversity hotspots are of special interest because their high endemism implies that these regions favour longer species persistence (lower extinction rates) and/or enhanced species origination (higher speciation rates) than elsewhere (Goldberg *et al.*, 2005; Tzedakis, 2009). Disentangling the evolutionary trajectories associated with the assembly of hotspot biodiversity can provide insights into the processes that generate and maintain the highly uneven distribution of species diversity across the planet.

Among terrestrial hotspots, the tropical Andes is the richest in terms of both species diversity and endemism (Orme *et al.*, 2005). The extraordinary diversity of the tropical Andes has been broadly attributed to exceptional overall physiographic (topographic and habitat) heterogeneity, providing a key landscape for diversification. More specifically two major historical events prompting species diversification have been highlighted. First, the later phases of Andean uplift during the late Miocene and early Pliocene are thought to have promoted diversification (Hoorn *et al.*, 2010; Antonelli & Sanmartín 2011a; Luebert & Weigend, 2014) through: a) vicariant fragmentation and isolation of lowland taxa by mountain uplift; b) allopatric speciation during range expansion of taxa adapted to newly available and greatly expanded montane habitats; and c) parapatric speciation along the steep and extended ecological gradients that characterize the Andes. Second, Pleistocene climatic oscillations may have further enhanced diversification through allopatric speciation resulting from repeated glacial contractions and interglacial expansions repeatedly fragmenting and reuniting species distributions (Hooghiemstra *et al.*, 2006; Rull, 2011; Madriñán *et al.*, 2013; Flantua *et al.*, 2014).

It is also clear that the tropical Andean biodiversity hotspot cannot be viewed as a single coherent unit characterized by similar trajectories of evolutionary diversification. The tropical Andes hotspot encompasses a tremendous variety of environmental conditions (dry to perhumid, warm to freezing, low to high elevation) and sharply contrasting vegetation types and associated biotas. In order to facilitate macroevolutionary comparisons, three broad-scale biomes have been proposed based on overall environmental and floristic similarities (see map

in Särkinen *et al.*, 2012). The seasonally dry tropical forests (SDTF; 0-2000m) encompass the dry adapted plant communities from the rain-shadowed inter-Andean valleys from Colombia to Bolivia and the western Andean foothills in northern Peru. The humid mid-elevation montane forests (MMF; *c.* 1000-3500 m) comprise the mesic plant communities on the eastern slopes of the Andes from Venezuela to Bolivia and the Pacific slopes of Colombia and northern Ecuador, sandwiched between the lowland tropical rain forests and the tree line (including lower and upper montane forests, subpáramo shrub forests, and cloud forests). Finally, the high-elevation grasslands (HEG; *c.* 3000-4800m) group the cold-adapted plant communities above the tree line (including páramo, super-páramo, jalca, puna) along mountain ridges from Venezuela to Bolivia (Särkinen *et al.*, 2012). Several recent studies have addressed the diversification dynamics of plant lineages in the SDTF and the HEG (Pennington *et al.*, 2004, 2006, 2010; Hughes & Eastwood, 2006; Särkinen *et al.*, 2012; Hughes *et al.*, 2013; Madriñán *et al.*, 2013; Hughes & Atchison, 2015). They revealed that the timing, tempo and trajectories of diversification in these two biomes, although both island-like in distribution, differ markedly in the timing and tempos of diversification, suggesting that the two biomes have evolved essentially in isolation from one another at. In contrast, and despite the fact that the MMF harbours the greatest diversity in the tropical Andes (Gentry, 1992; Richter *et al.*, 2009; Jørgensen *et al.*, 2011), the evolutionary trajectories of plant species diversification in the MMF biome are still surprisingly poorly known. This means that the MMF evolutionary trajectory remains a key missing element in the formulation of an overall "Andean biotic separation hypothesis" for the tropical Andean biodiversity hotspot (Särkinen *et al.*, 2012).

The MMF is characterized by humid to perhumid forests rich in epiphytes (orchids, bromeliads, ferns, mosses) with a well-developed understorey that receives a significant fraction of its precipitation in the form of water trapped by the vegetation from the almost permanent cloud cover (Hölscher 2008). Seasonality in precipitation and temperature is low but tend to increase with increasing distance from the Equator. This biome is marked by steep environmental gradients associated with elevation (Beck & Richter 2008, Gerold 2008, Kessler *et al.* 2011) that influence the physiognomy and the floristic composition of the vegetation (Leuschner & Moser 2008, Asner *et al.*, 2014). In general, plant species richness decreases, while endemism increases with increasing elevation (Jørgensen *et al.*, 2011). The combination of steep slopes and high precipitation causes frequent landslides that contribute to the overall physiographic heterogeneity and dynamism of the MMF (Mutke *et al.*, 2014). It is clear that the MMF biome harbours many species-rich plant groups (Beck & Richter, 2008), such as *Epidendrum* (Orchidaceae), *Piper* (Piperaceae), *Miconia* (Melastomataceae), *Solanum* (Solanaceae), *Schefflera* (Araliaceae),

and *Pilea* (Urticaceae), suggestive of possible radiations. Initial phylogenies that included Andean MMF taxa (*Fuchsia*, Berry *et al.*, 2004; *Renealmia*, Särkinen *et al.*, 2007; *Ceroxylon*, Trénel *et al.*, 2007; *Cinchona*, *Ladenbergia*, Antonelli *et al.*, 2009; Episcieae, Gloxinieae, Perret *et al.*, 2013; *Polystichum*, McHenry & Barrington 2014) suggested that MMF diversification coincided with the final phase of the Andean uplift during the middle Miocene to early Pliocene. However, none of these studies were densely enough sampled to investigate geotemporal trajectories of diversification in detail, in large part because the MMF biome covers an extensive area of often very inaccessible dense vegetation on steep slopes such that it has been difficult to densely sample in the field. Several more recent phylogenetic studies have provided more compelling evidence for elevated rates of species diversification associated with Andean MMF clades, notably subgenus *Tafalla* in the genus *Hedyosmum* (Chloranthaceae) (Antonelli & Sanmartín, 2011b), within the core Tillandsioid clade of Bromeliaceae (Givnish *et al.* 2014), the mostly Andean Neotropical Vaccinieae (Ericaceae) (Schwery *et al.*, 2014), the *Oreinotinus* clade of *Viburnum* (Adoxaceae) (Spriggs *et al.*, 2015), and Andean supersection *Tacsonia* of *Passiflora* (Passifloraceae) (Abrahamczyk *et al.*, 2014), suggesting that radiations may be common in the Andean MMF biome.

In order to investigate the evolutionary dynamics of plants in the Andean MMF biome we use the plant genus *Macrocarpaea* (Gentianaceae: Helieae) as an example to test the following hypotheses and questions: (i) species diversification in the Andean MMF is predominantly recent and involved rapid radiations (Schwery *et al.*, 2014); (ii) MMF plant radiations are older and slower than those in the higher elevation HEG biome, as suggested by Särkinen *et al.* (2012) and Schwery *et al.* (2014); (iii) whether diversification was driven primarily by opportunities associated with the emergence of new habitats and high physiographic heterogeneity linked to the later phases of Andean uplift, or with the impacts of Pleistocene climate oscillations on the distribution of Andean vegetation zones. *Macrocarpaea* comprises 118 species of shrubs and small trees restricted to montane regions of the Neotropics, though many new species are still being discovered, especially in the Andes (Grant, 2004, 2005). Eight species occur in the montane Atlantic forests of Brazil, six in the Pantepui of the Guayana Shield, five in Mesoamerica (Costa Rica and Panama), and three in the Greater Antilles of the Caribbean (one species on each island of Cuba, Hispaniola, and Jamaica). The majority of species (96 or 81%) occur in the tropical Andes from Venezuela to Bolivia. The genus has a broad altitudinal range in the Andes (from 500-3500 m), centred squarely within the distribution of the MMF biome, but most individual species occupy much narrower altitudinal ranges. The majority of species are narrowly restricted endemics, and even the most widely

distributed species do not exceed 250 km latitudinally. It is common to find several species in sympatry. A recent phylogenetic study that addressed patterns of diversification of *Macrocarpaea* was limited by low phylogenetic resolution, and the absence of dating (Struwe *et al.*, 2009b). Here, we improve taxon sampling (19 additional species), phylogenetic resolution through the use of additional molecular markers, and reconstruct the first time-calibrated phylogeny for the genus. We use the dated trees to infer the historical biogeography of the genus, compare a set of vicariance and dispersal scenarios regarding the range evolution of *Macrocarpaea* in the Andes, and estimate diversification rates through time and among lineages.

2. Materials and methods

2.1. Taxon sampling

76 of the 118 currently recognized species of *Macrocarpae*a (64% of the total) were sampled in the field or from herbarium specimens assembled as part of a detailed taxonomic revision of the genus (Grant 2004, 2005; Struwe *et al.*, 2009b; Grant *et al.*, in prep, see Appendix S1 in Supporting Information). These species represent the sectional classification of *Macrocarpaea* as well as the full spectrum of morphological, ecological and geographical variation: 75% (6/8) for the Atlantic forests of Brazil, 17% (1/6) for the Pantepui of the Guayana Shield, 60% (3/5) for Mesoamerica, 100% (3/3) for the Greater Antilles of the Caribbean, and 65% (62/96) for the Andes. One representative of each sampled species was included alongside five species of *Tachia* and one species of *Chorisepalum* as outgroups based on previous phylogenetic studies that indicate that these three genera comprise the monophyletic subtribe *Macrocarpaea* within the tribe Helieae (Struwe *et al.*, 2009a; 2009b).

2.2. DNA extraction, amplification and sequencing

Total genomic DNA was extracted from silica gel dried leaf samples or herbarium specimens with a standard cetyltrimethylammonium bromide (CTAB)-chloroform extraction followed by isopropanol precipitation and ethanol washing. Two ribosomal DNA intergenic spacers, ITS (internal transcribed spacer: ITS1, 5.8S, ITS2) and 5S-NTS (non transcribed spacer) were sequenced. The ITS region was amplified using the primers ITS4/ITS5 from White *et al.* (1990), and 5S-NTS using the primers PI/PII from Cox *et al.*, (1992). Four additional non-coding plastid DNA regions were sequenced. The *rpL16* introns and the *trnH-psbA* intergenic spacer were

amplified using the primers and temperature profiles of S*haw et al.* (2005). The *trnL-trnF* introns and exons and the *rpl32-trnL* intergenic spacer were amplified using the primers and temperature profiles of Taberlet *et al.* (1991) and Shaw *et al.* (2007) respectively.

2.3. *Sequence assembly, alignment and evolutionary model choice*

The two complementary DNA strands were assembled and edited using Geneious Pro version 5.5.3 (Biomatters Ltd., New Zealand). Alignments were performed with MAFFT (http://mafft.cbrc.jp/alignment/ server/) using FFT-NS-i settings and verified and adjusted manually in MEGA 5.05 (Tamura *et al.*, 2011). Regions of uncertain homology due to variable indel lengths in 5S-NTS, *trnL-trnF*, *trnH-psbA*, and *rpl32-trnL* were removed. The DNA substitution models that best fit the data was selected for each marker based on the Akaike Information Criterion (AIC), using jModelTest 2.1.3 (Darriba *et al.*, 2012). As no incongruence among individual gene trees was detected using the test of congruence among distance matrices (CADM, Campbell *et al.*, 2011) the DNA alignments of the six genes were concatenated using SequenceMatrix v1.7.8 (Vaidya *et al.*, 2011).

2.4. *Data analyses*

2.4.1. *Phylogenetic analyses and divergence time estimation*

Phylogenies and divergence times were simultaneously estimated using Bayesian MCMC searches in BEAST ver.1.7.5 (Drummond & Rambaut, 2007) on the online CIPRES Science Gateway (Miller *et al.*, 2010). We assumed a lognormal relaxed clock model (Drummond *et al.*, 2006). Both pure-birth and birth-death models of speciation were evaluated and compared as tree priors with an MCMC-based adaptation of the AIC (AICM, Raftery *et al.*, 2007). The data were partitioned and the substitution and clock models were unlinked between the different loci. For each locus the substitution model was set to the closest model as selected with jModelTest (Table 1). All other parameters and operators were set to default.

Because no fossil record exists for *Macrocarpaea*, we used a secondary calibration strategy (Renner, 2005). To do this, we extended taxon sampling to cover the entire Gentianaceae for which fossil calibrations are available. We downloaded ITS and *TrnL* sequences from Genbank for species representing all major lineages of the Gentianaceae (except Voyrieae). Voucher data were extracted from Genbank to allow concatenation of the

markers as far as possible at the individual level and not just the species level. The data matrix includes 344 species representing 67 genera from the Gentianaceae, 72 species of *Macrocarpaea* (four species were omitted from these analyses), and seven outgroups from the Gentianales with an aligned length of 1447 nucleotides (Appendix S2, ITS 2% missing, *Trnl* 38.7% missing).

Our calibration strategy to estimate divergence times within Gentianaceae differs slightly from previous analyses (Yuan *et al.*, 2005; Merckx *et al.*, 2013). We deliberately set large priors on the calibration constraints to take into account uncertainty regarding the age of the fossils (Ho & Philips, 2009).Three calibration points were used to estimate the age of the Gentianaceae: (1) a normal prior distribution with a mean of 78 million years (Myr) and a large standard deviation of 15 Myr (95% confidence interval of the mean: 53.3–102.7 Myr) for the root of the tree, reflecting the age estimates for the crown node of Gentianales obtained by Janssens *et al.* (2009) and Bremer *et al.* (2004); (2) a lognormal prior probability with a minimum bound of 40 Myr (offset of 40 Myr) on the crown node of the Potalieae, based on fossil pollen similar to *Lisianthius* from the mid- to late Eocene of Panama (Graham, 1984); (3) a lognormal prior probability with a minimum bound of 5 Myr (offset of 5 Myr) on the *Gentiana* crown node, based on a fossil seed from the Pliocene of Thuringia in Germany, showing diagnostic synapomorphies of *Gentiana* (Mai & Walther, 1988). Both lognormal prior probabilities were assigned a mean of 5.0 and standard deviation of 0.5 as true ages are likely older than fossil ages (~10 Myr).

In a second step, we used posterior age estimates inferred from the time-calibrated Gentianaceae phylogeny as calibration priors for a more detailed analysis of *Macrocarpaea*. This detailed analysis was conducted using the six-gene dataset and included all 76 *Macrocarpaea* species sampled. Because divergence times estimated from the family level analysis were log-normally distributed, we imposed lognormal prior probabilities to the most recent common ancestor (MRCA) of subtribe *Macrocarpaea* (the root of the tree) and to the MRCA of *Macrocarpaea* itself. Offsets, means and standard deviations were set to cover the estimated 95% highest posterior density (HPD) intervals.

The Markov chains were run for 50 million generations sampling every 2000 generations and repeated four times under each speciation tree prior for the family and the genus-level analyses. The Akaike's information criterion through Markov chain Monte Carlo (AICM) was estimated using Tracer ver.1.6 (Baele *et al.*, 2012) for each replicate independently and did not favour a birth-death model over the Yule model. Stationarity,

convergence and effective sample sizes (ESS, >200) were verified using Tracer ver.1.6. removing a 10% burn-in, and BEAST log files were combined using LogCombiner v1.5.4. Mean evolutionary rates and divergence times were calculated using TreeAnnotator version 1.7.8, keeping median heights for nodes.

2.4.2. *Biogeographical analyses*

Geographical areas were defined following Struwe *et al.* (2009b), with the exception of the three cordilleras of the northern Andes that are here grouped together since no clear geographical structure arose from preliminary analyses of these regions, and because we specifically aimed to analyse migration patterns along the natural latitudinal axis of the Andes. A total of eight areas were delimited: A, Greater Antilles, Caribbean; B, Mesoamerica; C, Occidental, Central, Oriental, and Mérida Cordilleras of the northern Andes in Venezuela, Colombia, and northern Ecuador; D, the Amotape–Huancabamba zone in Ecuador and Peru; E, Cordillera Central in central Peru; F, Cordillera Central in southern Peru and Bolivian Yungas; G, Pantepui of the Guayana Shield; H, montane Atlantic forests of Brazil. The four Andean areas correspond to centres of endemism for *Macrocarpaea*, but we acknowledge that the gaps separating areas D, E and F most likely represent sampling, rather than real distribution gaps.

Parametric historical biogeographical analyses were conducted using the R package BioGeoBEARS (Matzke, 2013; 2014). This program implements likelihood optimization of the dispersal-extinction cladogenesis (DEC, Ree *et al.*, 2005) and dispersal-vicariance models (DIVA, Ronquist, 1997). It also allows users to define models by relaxing assumptions on cladogenetic (sympatry and vicariance) and anagenetic (dispersal, extirpation) range inheritance parameters (Matzke 2013). Another advantage of this method is that it implements a cladogenetic parameter that accounts for founder-event speciation, where a population establishes itself out of its ancestral range and subsequently differentiates into a new species (Tempelton, 2008). The founder-event parameter "*j*", is usually set as an additional free parameter and its *per event weight* is estimated together with the range expansion rate "*d*", the range contraction rate "*e*", and the overall likelihood of the model. The fit of different models to the data can be compared by computation of the corrected Akaike information criterion (AICc) score and the Akaike weights for each model. In this study, we compared four models: DEC, DIVAlike (for likelihood implementation of the DIVA model), DEC+j and DIVAlike+j.

In order to assess sensitivity to phylogenetic uncertainty, each model was run on the same set of 100 trees, sampled randomly from the posterior distribution of the BEAST analyses, after removing the outgroups. Maximum ancestral range size was set to 4, but we allowed wide distributions (>2 areas) only for exclusively Andean ranges (CDE, CDF, CEF, DEF, CDEF), thus limiting the total number of states to 42 instead of 163. We also set a dispersal matrix penalizing dispersal between non-contiguous areas to 0.5 while dispersal between contiguous areas was set to 1. The maximum likelihood estimates of the relative probability of the ancestral range at the nodes from the 100 trees, obtained under the best biogeographical model, are summarized on the maximum clade credibility (MCC) tree from BEAST, using a new R script and TreeAnnotator.

To further investigate the ancestral range of the Andean lineage, we tested additional models by constraining the likelihood of the ancestral range of the Andean most recent common ancestor (MRCA) using the arguments "fixnode" and "fixlike" from the function "define_BioGeoBEARS_run". We tested seven scenarios where the Andean MRCA was constrained to be: 1) endemic to the north (C=1), 2) endemic to the Amotape-Huancabamba zone (D=1); 3) endemic to central Peru (E=1); 4) endemic to southern Peru and the Bolivian Yungas (F=1); 5) widespread in the northern half of the range in the Andes (CD=1); 6) widespread in the southern half of the range of *Macrocarpaea* in the Andes (EF=1); and 7) widespread throughout the Andes (CDEF=1). We ran the seven scenarios under each of the four previous models on the set of 100 trees. The combination of the best scenario and the best model were evaluated by comparing the distribution of AICc scores for each area combination (N=28) over the set of trees. The DIVA+j model constantly returned better fit to the data under each scenario tested. For simplicity we present the results under this model.

2.4.3. *Diversification rate analyses*

Variation in species diversification rates through time and among lineages was investigated using the BAMM framework (Bayesian Analysis of Macroevolutionary Mixtures, Rabosky, 2014; Rabosky *et al.*, 2014a). BAMM implements a model of rate variation that assumes that phylogenetic trees are shaped by a discrete number of distinct and potentially dynamic evolutionary processes of speciation and extinction.

The model tests for a single time-varying diversification process (no events) across the entire phylogeny or two or more distinct time-varying processes (N events) governing evolutionary dynamics across the tree. BAMM

uses reversible jump Markov chain Monte Carlo to explore this spectrum of alternative models of lineage diversification. The relative probability of each diversification model sampled during the analysis can be calculated based on its posterior probability (Rabosky 2014). BAMM also estimates marginal distributions of speciation and extinction rates for every branch of a phylogenetic tree, thus allowing comparisons of clade-specific diversification trajectories.

We ran BAMM ver.2.1 on the MCC tree from the BEAST analysis. We corrected for sampling fractions by assuming that unsampled species belong to the lineage with which they share their geographical location. This assumption is subject to some uncertainty, and is imperfect given the non-monophyly of the northern Andes, but has the desirable effect of accounting for the asymmetry in taxon sampling, which is lower for the northern Andes than other areas. Six clades were used to assign unsampled diversity: 1) The genus *Chorisepalum*, with *f*=0.2 (1 species sampled for 5 described); 2) the genus *Tachia*, with *f*=0.36 (5/14); and within *Macrocarpaea* 3) the Brazilian clade, with *f*=0.75 (6/8); 4) the Caribbean clade with *f*=1 (3/3); 5) the clade including most of the species from the northern Andes, Mesoamerica and the Pantepui of the Guayana Shield, with f=0.36 (16/42); and 6) the clade including all species from the Amotape-Huancabamba zone, the central Andes and four species from the northern Andes, with f=0.77 (51/66). We also set f=0.60 (83/138) for the backbone part of the tree to account for the fact that *Chorisepalum* does not form a clade *per se* in our phylogeny (only one species sampled) and is therefore part of the backbone of the tree.

Priors were set using the "setBAMMpriors" function in BAMMtools (Rabosky *et al.*, 2014b). All other control parameters were set to their default values. We ran four independent BAMM analyses with 5 million generations of MCMC sampling each. Posterior and "event data" were sampled every 1000 generations. Convergence between runs was assessed by comparing the log likelihood traces from the MCMC outputs and the ESS for the log likelihood and the number of events, after 10% burn-in. Posterior probabilities of diversification models and evolutionary rates through time were extracted and summarized using BAMMtools.

When a slowdown was detected with BAMM, its significance was evaluated by measuring the γ-statistic and applying the Monte Carlo constant rates (MCCR) test in the R package LASER (Rabosky, 2006) using the number of known and unsampled species in the clade, fixing the number of replicates to 10,000.

3. Results

3.1. Phylogenetic relationships

Alignment lengths, numbers of variable and parsimony informative sites are reported for each DNA sequence locus together with the selected evolutionary model and the percentage of sequences available in the complete dataset (Table 1). Overall the nuclear marker 5S-NTS alone contributes half of the total number parsimony informative sites. The four chloroplast markers show useful levels of informative variation when compared with the popular nuclear marker ITS, with the *trnH-psbA* locus showing the highest ratio of parsimony informative sites to length (11%)

The MCC tree from the BEAST analysis of the Gentianaceae (Appendix S3) shows strong support (PP>95) for tribal relationships and the overall topology agrees with previous studies (Struwe *et al.*, 2002; Merckx *et al.*, 2013) except for Voyrieae which was not included here. The tribe Helieae is monophyletic (PP=0.94) with *Prepusa* as sister to the remaining genera. The MCC tree of the Helieae subtribe *Macrocarpaea* (Fig. 1) confirms previous results (Struwe *et al.*, 2009b) whereby *Macrocarpaea* species from Brazil (Section *Tabacifolieae*) form a monophyletic group (PP=1) sister to a large clade (PP=1) containing the remaining *Macrocarpaea* species. Within this large clade, a subclade composed of most of the species from the northern Andes (Colombia and northern Ecuador) with species from the Pantepui and Mesoamerica nested within it, is robustly supported (PP=0.96). The Caribbean species form a clade (PP=1) that is weakly supported (PP=0.36) as sister to a clade (PP=1) composed of almost all the species from the Amotape-Huancabamba zone and the central Andes plus a small subclade of northern Andean species nested within it. Relationships among species within the large central/north Andean subclade are generally weakly supported (PP<0.9).

3.2. Estimated ages of lineages

Age estimates indicate that subtribe *Macrocarpaea* diversified in the early Miocene and *Macrocarpaea* itself in the late Miocene and Pliocene (Fig. 1 & Table 2) The estimated age for the crown node of Gentianaceae is 76.8 Ma (HPD, 61.6-92.4 Ma), which is consistent with priors imposed on that node. The split between *Prepusa* and the remaining genera from the tribe Helieae is estimated to have occurred 36.2 Ma (HPD, 24.7-47.3 Ma), while

the MRCA of the Helieae excluding *Prepusa* is estimated at 24.0 Myr (HPD, 17.6-31.1 Ma). Within tribe Helieae, subtribe *Macrocarpaea* (*Chorisepalum*, *Macrocarpaea*, and *Tachia*) started to diversify 20.6 Ma (HPD, 14.0-27.5 Ma), corresponding also to the stem age of *Macrocarpaea*. The estimated crown age of *Macrocarpaea* is 11.1 Ma (HPD, 7.3-15.1 Ma). Both stem and crown estimates of the age of *Macrocarpaea* (mean and HPD) were used as priors in the secondary calibration analysis. Overall these divergence time estimations are slightly older than previous estimates (Merckx *et al.*, 2013), but we believe provide conservative age estimates for the genus *Macrocarpaea* and its close relatives.

3.3. *Historical biogeography*

The DIVALIKE+J (with founder-event) model outperformed all other models, returning consistently lower AICc scores (mean AICc= 159.9) and higher Akaike weights (mean ω=0.749) over the 100 sampled trees (Table 3). The second best model is the DEC+J model which according to Burnham and Anderson's (2002) criteria receives some support (Δ AICc mean=2.19). The mean evidence ratio (ωbest/ω*i*=2.992) indicates that the DIVALIKE+J model is almost three times more likely than the DEC+J model. In comparison, models that do not consider founder-event speciation (DEC and DIVALIKE) performed poorly (DEC: Δ AICc mean=63.607, DIVALIKE: Δ AICc mean=57.431), suggesting that founder-event speciation is an important process explaining the historical biogeography of *Macrocarpaea*. Under the DIVALIKE+J model, rates of range expansion and range contraction are estimated to be virtually null (mean *d*=1.92E-12, mean *e*=6E-11), indicating very little or no role for these two anagenetic processes in *Macrocarpaea* range evolution. Reconstructions of the ancestral areas of the MRCA of *Macrocarpaea* remain ambiguous (Fig. 2), with support for combinations of two non-contiguous areas, always including the montane Atlantic forests of Brazil (CH, *P*=0.39; DH, *P*=0.32). The ancestral area of the large clade including all Andean species is also reconstructed as ambiguous, with the Amotape-Huancabamba zone (D, *P*=0.35) and the northern Andes (C, *P*=0.29) obtaining the most support. Colonisation of the entire Andean range of the genus (areas C, D, E, F) apparently occurred rapidly over ca. 2 Myr (7.2-5.2 Ma) primarily from the northern (C, D) to the central Andes (E, F). colonisation of Mesoamerica (B) from the northern Andes (C) occurred about 4 Ma (HPD, 2.8-5.4 Ma), with two subsequent independent dispersal events back to the northern Andes. Despite their proximity and several exchanges between the Amotape-Huancabamba zone (D), the Cordillera Central in central Peru (E), the Cordillera Central in southern Peru and Bolivian Yungas (F), only one recent exchange (1.7 Ma, HPD, 1.15-2.41 Ma) has occurred between the

Amotape-Huancabamba zone in Ecuador and Peru (D) and the Occidental, Central, Oriental, and Mérida Cordilleras of the northern Andes in Venezuela, Colombia, and northern Ecuador (C) during the last 6 Myr.

Across the seven Andean ancestral area scenarios tested, the northern Andes (C) best fits the data (Table 4) for 61 of the sampled trees, followed by the Amotape-Huancabamba zone scenario (D) for 26 trees. Mean AICc differences (Δ AICc mean=-0.008) and Akaike weights for these two scenarios (C, mean ω=0.441; D, mean ω= 0.409) do not strongly favour one over the other. The mean evidence ratio indicates that the northern Andes scenario is only 1.8 times more probable than the Amotape–Huancabamba zone scenario. Among the other five scenarios, the only one gaining some support is the northern plus Amotape-Huancabamba combined area (CD, Δ AICc mean=3.578, mean ω=0.075). This confirms that even if it is not possible to unequivocally pinpoint a precise region as the ancestral range for the Andean MRCA, it is very likely that it was located in the north (either C, D, or CD).

3.4. *Diversification rates through time*

There is strong evidence for significant variation in species diversification rates through time and among lineages across subtribe *Macrocarpaea* (Fig. 3a). A two-rate regime model (one rate shift) was favoured by BAMM with a posterior probability of p=0.42 (Fig. 3 c). A single rate regime model (no rate shifts) obtained a posterior probability of p=0.02. A posterior odds ratio of 17.7 and a Bayes factor of 26 strongly favour a two-rate over a single rate regime model (Raftery, 1995). Models with 3 to 5 rate regimes have Bayes factor >26 compared with a single rate regime model, but a Bayes factor <1.7 when compared with a two-rate regime model, indicating limited support for further rate heterogeneity across the phylogeny. The location of the shift to higher rates of diversification is most likely located along the branch subtending the Andean clade (Bayes factor =169; Fig 3d).

The 95 % credible set of rate shift configurations, sampled using a conservative Bayes factor criterion of 30 to identify “true” shifts from shifts expected from the prior alone, contains four shift configurations (cumulative probability p=0.95). By far the most frequently sampled shift configuration (p=0.61) has a single shift at the Andean MRCA (Fig.3 b). The second-most frequently sampled configuration (p=0.27) is a single rate model. Two additional shift configurations were sampled with low frequency (p=0.051, p=0.027) and include the

original shift subtending the Andean clade as well as an additional rate increase, at slightly different positions, nested within the Andean clade close to the ancestor of a clade composed of predominantly lower elevation species (Fig. 3 b).

Diversification rate-through-time plots confirm that the marked increase in mean diversification rate around 7.2 Ma is attributable to the Andean clade (Fig. 4 a). Across the remainder of subtribe *Macrocarpaea*, the mean diversification rate increases only very slowly through time and remains low (max=0.13 sp Myr^{-1}, 90% CI: -0.03-0.29 sp Myr^{-1}) (Fig. 4 c), while the mean diversification rate of the Andean clade starts with a high rate of 0.81 species per Myr (90% CI: 0.31-1.39 sp Myr^{-1}) and decreases almost exponentially through time to the present-day rate of 0.34 species per Myr (90% CI: 0.20-0.46 sp Myr^{-1}) (Fig. 4 b).

The means of the posterior distribution of diversification rates for the genus *Tachia*, the Brazilian clade and the Andean clade are 0.11, 0.16 and 0.43 species per Myr respectively. Nevertheless, the estimation of the mean diversification rate for the Andean clade is somewhat misleading because diversification rates vary within this clade.

The γ-statistic and the MCCR test confirmed the significance of the rate slowdown detected within the Andean clade (known species= 111, unsampled species= 41) by BAMM (MCCR-test, γ-statistics = -3.02, critical value of γ = -2.64, p = 0.02).

4. Discussion

We present evidence for a rapid mid- to late-Miocene radiation in the Andean MMF biome using for the first time a densely sampled phylogeny for a species-rich MMF plant lineage. We show that the radiation of *Macrocarpaea* is associated with rapid colonisation and range expansion across the entire extant distribution of the genus in the Andes starting about 7.2 Ma (HPD: 9.7-5.2 Myr). We also show that Pleistocene climatic oscillation did not affect the diversification rate of *Macrocarpaea* in the Andes as a whole, but potentially further enhanced diversification in a derived subclade of predominantly lower elevation (<1800masl) Andean species.

4.1. Historical biogeography of the genus Macrocarpaea

Our results suggest a disjunct distribution including the montane Atlantic forests of Brazil, the northern Andes or the Amotape–Huancabamba zone for the area of origin of the genus. Connections between the montane Atlantic

forests of Brazil and the Andes are a common pattern for many Andean-centred plant taxa *sensu* Gentry (1982). Several well-documented examples (Berry *et al.*, 2004; Perret *et al.*, 2006; Givnish *et al.*, 2011) have indicated possible exchange through potential corridors of favourable habitat between Brazil and the Bolivian Andes in the past. Here we provide another example of these connections between the Atlantic forests of Brazil and the Andes, but the direction and the nature of this transition cannot be determined.

Based on the most likely ancestral state for the Andean clade and colonisation sequence, we hypothesize a predominantly southward wave of colonisation, range expansion and dispersification (*sensu* Moore & Donoghue, 2007) of *Macrocarpaea* along the tropical Andes. Strikingly, our results suggest that the entire extant distribution of the genus in the Andes was rapidly colonised within just ca. two Myr (5.2-7.2Ma). While many details of Andean uplift history remain controversial (Garzione *et al.*, 2008; Poulsen *et al.*, 2010; Luebert & Weigend, 2014), it is widely recognized that the northern South American climate was profoundly altered by the uplift of the Andes in the late Miocene (Hoorn *et al.*, 2010). It is thought that the onset of convective precipitation and the massive increase in rainfall along the eastern flanks of the Andes occurred once the Andes reached around 50-70% of their modern elevation, promoted by transport of Amazon moisture toward the Andes (Insel *et al.*, 2012). This regional climate change is likely to have triggered and enabled the establishment of widespread humid mid-elevation montane forests in the tropical Andes. Late Miocene palynological records from the northern Andes (Hooghiemstra *et al.*, 2006) and from the eastern cordillera in northern Bolivia (Graham *et al.*, 2001) together with the divergence time estimates for several plant groups currently inhabiting the MMF (*Ceroxylon*, Trénel *et al.*, 2007; *Cinchona* and *Ladenbergia*, Antonelli *et al.*, 2009; Episceae, Perret *et al.*, 2013; the *Tacsonia* clade of *Passiflora*, Abrahamczyk *et al.*, 2014; core Tillandsioids of Bromeliaceae, Givnish *et al.*, 2014; the *Oreinotinus* clade of *Viburnum*, Spriggs *et al.*, 2015;) support this idea (Table 5). The age we infer for the wave of colonisation and range expansion associated with dispersification of *Macrocarpaea* in the Andes corresponds to this period, suggesting that *Macrocarpaea* may have invaded the whole range of the Andean MMF soon after it became established. Species of *Macrocarpaea* produce numerous dust-like seeds (0.2-1.5 x 0.2-2 mm; Grant, 2005) that are readily transported by wind across considerable distances. This high dispersal ability might have contributed to the rapid southward colonisation of the Andean MMF by *Macrocarpaea*.

The multiple transitions inferred between the Amotape-Huancabamba zone and regions further south support the conclusions of Weigend *et al.* (2002) and Luebert & Weigend (2014) who found that the Amotape-Huancabamba

zone did not present a major dispersal barrier for mid-elevation taxa. Instead, we identify the Giron-Paute trans-Andean valley, which coincides with the northern boundary of the Amotape-Huancabamba zone (Keating, 2008), as a potentially important dispersal barrier for *Macrocarpaea*. If we exclude the transition that happened early in the history of the Andean clade (7.2 Ma), only one subsequent exchange is inferred, and rather recently, 1.4 Ma (HPD, 0.71-2.02 Ma), between the Amotape-Huancabamba zone and the northern Andes. North of the Giron-Paute valley lies the northern volcanic zone (NVZ) of the Andes, which has many volcanoes that have been active since the Quaternary (Beate *et al.*, 2001). Extensive areas surrounding the Giron-Paute valley and further north have been covered by a series of thick volcanic deposits dating to the Late Miocene and Early Pliocene (Pisayambo Formation, Coltori & Ollier, 2000). These massive deposits testify to intense volcanic activity since this time, which might have filtered *Macrocarpaea* migrations, by causing repeated catastrophic local extinction and/or by modifying regional soil composition (Foster *et al.*, 1998).

Another interesting feature of *Macrocarpaea* biogeography is the
colonisation of Mesoamerica from the northern Andes 4.3 Ma (HPD, 2.62-6.0 Ma), potentially predating closure of the Isthmus of Panama (c. 3 Ma), confirming that the Central American Seaway was not a major dispersal barrier for plants, or that earlier land connections existed between the two continents (Farris *et al.*, 2011; Hughes *et al.*, 2013; Luebert & Weigend. 2014).

4.2. *Processes involved in Macrocarpaea range evolution*

The biogeographical analyses strongly favour models integrating a parameter for founder-event speciation, which is known to be an important process in island-like habitats where dispersal to a new area is immediately accompanied by a strong reduction of gene flow because of the lack of habitat connections (Paulay & Meyer, 2002; Cowie & Holland, 2006). The new population, which is usually founded by a small number of individuals, endures a severe bottleneck followed by strong inbreeding that induces large changes in allele frequencies (Carson & Tempelton, 1984). If the population manages to survive and re-expand, the intricate action of drift and selection on the new genetic background may lead to reproductive isolation from the source population (Templeton, 2008). Recent studies have highlighted the island-like nature of the two biomes (SDTF and HEG) adjacent to the MMF in the tropical Andes (Hughes & Eastwood, 2006; Särkinen *et al.*, 2012). In contrast, the MMF has a generally continuous distribution along the eastern flank of the cordilleras. The important

contribution of founder-events to *Macrocarpaea* range evolution and the extremely low rates of anagenetic (along the branches) range expansion we infer, indicate that the different regions of the Andean MMF (as they were defined here) might have been less continuous in the past or that connections were narrow and constricted thereby promoting isolation (see below).

Finally, the better fit of the DIVAlike model over the DEC model might be explained by the fact that the subset sympatry parameter implemented in the DEC but not in the DIVAlike model, is not required to explain the data. As a consequence, the likelihood of the model inflates because the probabilities spread out over more inheritance scenarios (Matzke, personal communication).

4.3. *Diversification patterns in Macrocarpaea*

We found evidence for diversification rate heterogeneity through time and among lineages across the *Macrocarpaea* phylogeny. The root of the phylogenetic tree, the outgroups, and the Brazilian clade follow a trajectory of slow and only slowly increasing diversification rates (< 0.13 sp Myr^{-1}). A significant shift to higher rates of diversification is consistently detected along the branch leading to the ancestor of the Andean *Macrocarpaea* clade. This diversification rate regime starts with a relatively high diversification rate of 0.81species per Myr, a more than 6X increase over the background diversification rate. Subsequently, the diversification rate decreases almost exponentially to a current rate of 0.34 species per Myr (Fig. 4b). This pattern of an explosive burst of high diversification followed by a rate slowdown is typical of a diversity dependent model of diversification often considered to be the hallmark of adaptive radiations (Rabosky & Lovette, 2008; Givnish *et al.*, 2011). According to the ecological theory of adaptive radiation, the initial burst of diversification is driven by rapid adaptation to newly available ecological niches and the slowdown results from the progressive filling of niche space as species accumulate (Losos, 2010; Glor, 2010). Theory also predicts that diversification rate variation should be accompanied by adaptive ecological disparification (Gavrilets & Losos, 2009). However, non-adaptive geographical processes, such as rapid range expansion across multiple barriers can also cause a burst of diversification through the nearly simultaneous formation of many isolated populations (Moore & Donoghue 2007, Rundell & Price, 2009; Moen & Morlon, 2014). As the availability of suitable habitats is reduced, geographical opportunities for isolation will decrease and can lead to a decrease in the

diversification rate (Pigot *et al.*, 2010). Under this type of non-adaptive radiation, species diversification rate variation is expected to be decoupled from ecological disparification (Adams *et al.*, 2009).

It is notable that the diversification rate shift we infer for the Andean clade coincides with the start of the rapid colonisation and range expansion of *Macrocarpaea* throughout its current range in the Andean MMF some 7.2 Ma. The apparently rapid appearance and establishment of the MMF along the tropical Andes during the late Miocene discussed in the section 4.2 could have provided the opportunity for *Macrocarpaea* to extend its range. Furthermore, it is postulated that Andean palaeoelevations were at least 50% of the modern elevations with the complex physiography that characterizes this region already in place at that time (Insel *et al.*, 2012). Thus, *Macrocarpaea* range expansion along the Andes almost certainly occurred across an area of high physiographic heterogeneity that could have prompted isolation of newly formed populations and promoted rapid species diversification, as suggested for montane plant radiations more generally (Schwery *et al.*, 2014). This idea is consistent with the support we found for founder-event speciation as an important process involved in *Macrocarpaea* range evolution. The Andean MMF occupies a predominantly linear distribution along a latitudinal axis and can be regarded as a stepping-stone metapopulation system (Trénel *et al.*, 2008). Simulation studies have shown that range expansion through repeated founder-events in the presence of low inter-population gene flow in this type of metapopulation system tends to favour rapid genetic differentiation of populations along the expansion front (Le Corre & Kremer, 1998; Pigot *et al.*, 2010). Thus, the narrow linear distribution of the MMF is likely to have contributed to the accelerated rate of diversification rate during range expansion of *Macrocarpaea*. As *Macrocarpaea* progressively expanded throughout the Andean MMF, opportunities for isolation via the colonisation of unoccupied areas by the genus would have decreased, in turn slowing the rate of species diversification.

We have emphasized how the geographical processes that likely operated during range expansion in the Andes could have regulated the evolutionary dynamics of *Macrocarpaea*. However, the Andean MMF are also characterized by steep environmental gradients associated with altitude (Kessler *et al.*, 2011). Ecological gradients promote adaptive divergence (Doebeli & Dieckmann, 2003) and it seems likely that adaptation to different ecological conditions in conjunction with geographical processes have contributed to the rapid diversification of *Macrocarpaea* in the Andes, but rigorously testing this is beyond the scope of this study. The only clues to evaluate the relative contributions of ecological and geographical processes to the Andean clade

diversification come from species pair comparisons. Most putative sister species share similar habitats and are endemic to the same region, but are locally allopatric, favouring geography over ecology as the primary driver of species divergence. Nevertheless, it is also possible that ecological disparification could have occurred in the initial phase of adaptive radiation (Gavrilets & Losos, 2009), such that sister species comparisons which only capture the later phases of radiation might be misleading (Ackerley *et al.*, 2006). Additional studies to investigate niche evolution of *Macrocarpaea* in the Andes are needed in order to evaluate the adaptive component of this radiation.

Finally, the impacts of Pleistocene climatic oscillations on species distributions, by promoting repeated allopatric/parapatric speciation, have often been considered as a potential driver of diversification in the Andes (Rull, 2011; Turchetto-Zolet *et al.*, 2013). We did not detect any diversification pulse across the Andean clade as a whole during the last 2.6 Myr. Nevertheless, the two less frequently sampled rate shift configurations from the BAMM analyses suggest the possibility of a second nested shift to even higher rates of diversification in a derived subclade of mainly lower elevation sub-Andean forest species (Fig. 3b). The age estimates of these additional shifts are 3.1 Myr and 2.3 Myr respectively, i.e. close to the start of the Pleistocene. This suggests that Pleistocene climatic oscillations might have further enhanced the diversification of Andean species at lower elevations but not in the higher elevation vegetation belts. Palaeobotanical studies have revealed that among the MMF vegetation belts, it was the lower montane (sub-Andean) forests which experienced the greatest contraction at the last glacial maximum (LGM, Hooghiemstra & Van der Hammen, 2004; Flantua *et al.*, 2014). Cycles of strong contraction during glacial and re-expansion during interglacial periods throughout the Pleistocene, resulting in habitat fragmentation and re-colonisation, could have preferentially promoted allopatric speciation at lower elevations over the apparently less constricted higher elevation belt (Flantua *et al.*, 2014).

4.4. *The emerging picture of evolutionary plant diversification in the MMF biome.*

These findings for *Macrocarpaea* are remarkably congruent with recent results for other Andean MMF clades, and there is growing evidence that late Miocene / Pliocene radiation played a central role in generating the present hyperdiversity of the Andean MMF biome. In just the last year, a suite of other MMF plant radiations strikingly coincident in timing and geography, and with similar accelerated rates of species diversification, have been documented (Table 5).

The age we infer for the onset of radiation in *Macrocarpaea* in the Andes is intermediate between the ages of divergence reported for the older SDTF and younger HEG plant lineages (Särkinen *et al.*, 2012). The diversification rate is also intermediate between those of these adjacent biomes (Särkinen *et al.*, 2012; Madriñán *et al.*, 2013). The estimated diversification rate for *Macrocarpaea* at its peak (0.81sp $Myr^{-1,}$ 90% CI: 0.31-1.39 sp Myr^{-1}) is about three times lower than the diversification rate of the HEG genus *Lupinus* that was estimated using a time constant approach (Drummond *et al.*, 2012) that potentially minimizes the contrast. The high diversification rate inferred for the HEG has been explained by the fact that plant lineages in this very young biome might still be in the early explosive phases of radiation (Madriñán *et al.*, 2013; Hughes & Atchison, 2015). This is exemplified by the large Andean HEG radiation of *Lupinus*, where diversification rates are still accelerating towards the present suggesting a radiation in the early explosive phase (Hughes & Atchison, 2015). In contrast, *Macrocarpaea* presents a classical trajectory of adaptive radiation with a rapid initial burst of diversification followed by a rate slow down (Fig. 4). These results suggest that both the HEG and MMF biomes harbour significant plant radiations showing elevated rates of species diversification, but that the lower elevation MMF radiations are older, slower and more mature that those in the HEG. In sharp contrast, the Andean SDTF has been characterized as a "museum" of diversity where older lineages have slowly accumulated limited species diversity from the mid Miocene onwards (Pennington *et al.*, 2010; Särkinen *et al.*, 2012).

5. Conclusions

In this study we shown that the plant genus *Macrocarpaea* diversified in the Andean MMF following the classical trajectory of a radiation, starting with a high rate of diversification followed by an exponential decrease through time. The radiation of *Macrocarpaea* was likely triggered and enabled by the rapid establishment of MMF along the entire eastern flank of the tropical Andes once palaeoelevations were sufficient to cause high year-round precipitation. This establishment and expansion of the MMF biome would have represented an opportunity for pre-adapted *Macrocarpaea* palaeospecies to disperse, expand and diversify. Geographical processes arising from habitat constriction and fragmentation, and repeated rare dispersal events (founder-events), have likely played important roles in driving the radiation. Nevertheless, a potential concomitant role of ecological processes remains to be assessed. Our study also suggests that the MMF biome is a "mature cradle" of diversity that harbours plant lineages in the later phases of their radiations, thereby confirming that the three prominent biomes of the tropical Andes have contrasting evolutionary histories. This calls for careful consideration of geographical biodiversity units when planning conservation strategies or when trying to explain patterns of biodiversity distributions in this prominent global biodiversity hotspot.

AKNOWLEDGEMENTS

The work was supported by the Fonds des donations of the Université de Neuchâtel, the Basler Stiftung für biologische Forschung, and the Fondation Dr. Joachim de Giacomi. JV thanks David Desrousseaux and Eric F. Rodríguez R. for assistance with fieldwork in Peru. We also thank Lena Struwe for feedback and review of the manuscript, Nick Matzke for help with the software BioGeoBEARS.

REFERENCES

Abrahamczyk, S., Souto-Vilarós, D., & Renner, S. S. (2014) Escape from extreme specialization: passionflowers, bats and the sword-billed hummingbird. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 20140888.

Ackerley D.D., Schwilk D.W., & Webb C.O. (2006) Niche evolution and adaptive radiation: testing the order of trait divergence. *Ecology*, **87**, S50–S61.

Adams D.C., Berns C.M., Kozak K.H., & Wiens J.J. (2009) Are rates of species diversification correlated with rates of morphological evolution? *Proceedings of the Royal Society B: Biological Sciences*, **276**, 2729-2738.

Antonelli A., Nylander J.A., Persson C., & Sanmartín I. (2009) Tracing the impact of the Andean uplift on Neotropical plant evolution. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 9749–54.

Antonelli A. & Sanmartín I. (2011a) Why are there so many plant species in the Neotropics? *Taxon*, **60**, 403–414.

Antonelli A. & Sanmartín I. (2011b) 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.

Asner G.P., Anderson C.B., Martin R.E., Knapp D.E., Tupayachi R., Sinca F., & Malhi Y. (2014) Landscape-scale changes in forest structure and functional traits along an Andes-to-Amazon elevation gradient. *Biogeosciences*, **11**, 843-856.

Baele G., Lemey P., Bedford T., Rambaut A., Suchard M.A., & Alekseyenko A. V (2012) Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution*, **29**, 2157–2167.

Beck E. & Richter M. (2008) Ecological aspects of a biodiversity hotspot in the Andes of southern Ecuador. *In: Gradstein, R., Homeier, J., Gansert, D. (Eds.), The Tropical Mountain Forest. Patterns and Processes in a Biodiversity Hotspot. Biodiversity and Ecology series, 2. Gottingen Centre for Biodiversity and Ecology, Gottingen*, 195–217.

Beate B., Monzier M., Spikings R., Cotten J., Silva J., Bourdon E., & Eissen J.-P. (2001) Mio–Pliocene adakite generation related to flat subduction in southern Ecuador: the Quimsacocha volcanic centre. *Earth and Planetary Science Letters*, **192**, 561–570.

Berry P.E., Hahn W.J., Sytsma K.J., Hall J.C., & Mast A. (2004) Phylogenetic relationships and biogeography of *Fuchsia* (Onagraceae) based on noncoding nuclear and chloroplast DNA data. *American Journal of Botany*, **91**, 601–14.

Bremer K., Friis E., & Bremer B. (2004) Molecular phylogenetic dating of asterid flowering plants shows early Cretaceous diversification. *Systematic Biology*, **53**, 496–505.

Campbell, V., Legendre, P., & Lapointe, F.J. (2011) The performance of the Congruence Among Distance Matrices (CADM) test in phylogenetic analysis. *BMC evolutionary biology*, **11**: 64.

Carson H.L. & Templeton A.R. (1984) Genetic revolutions in relation to speciation phenomena: the founding of new populations. *Annual Review of Ecology and Systematics*, **15,** 97–131.

Coltorti M. & Ollier C. (2000) Geomorphic and tectonic evolution of the Ecuadorian Andes. *Geomorphology*, **32**, 1–19.

Le Corre V. & Kremer A. (1998) Cumulative effects of founding events during colonization on genetic diversity and differentiation in an island and stepping-stone model. *Journal of Evolutionary Biology*, **11**, 495–512.

Cox, A.V., Bennett, M.D., & Dyer, T.A. (1992) Use of the polymerase chain reaction to detect spacer size heterogeneity in plant 5S-rRNA gene clusters and to locate such clusters in wheat (*Triticum aestivum* L.) *Theoretical and Applied Genetics*, **83**, 684-690.

Cowie R.H. & Holland B.S. (2006) Dispersal is fundamental to biogeography and the evolution of biodiversity on oceanic islands. *Journal of Biogeography*, **33**, 193–198.

Crisp M.D., Trewick S.A., & Cook L.G. (2011) Hypothesis testing in biogeography. *Trends in Ecology & Evolution*, **26**, 66–72.

Darriba D., Taboada G.L., Doallo R., & Posada D. (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, **9**, 772–772.

Doebeli M. & Dieckmann U. (2003) Speciation along environmental gradients. *Nature*, **421**, 259–264.

Drummond A.J., Ho S.Y.W., Phillips M.J., & Rambaut A. (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biology*, **4**, e88.

Drummond A.J. & Rambaut A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Drummond C.S., Eastwood R.J., Miotto S.T.S., & Hughes C.E. (2012) Multiple continental radiations and correlates of diversification in *Lupinus* (Leguminosae): testing for key innovation with incomplete taxon sampling. *Systematic Biology*, **61**, 443–60.

Farris D.W., Jaramillo C., Bayona G., Restrepo-Moreno S.A., Montes C., Cardona A., Mora A., Speakman R.J., Glascock M.D., & Valencia V. (2011) Fracturing of the Panamanian Isthmus during initial collision with South America. *Geology*, **39**, 1007–1010.

Flantua S.G., Hooghiemstra H., Van Boxel J.H., Cabrera M., González-Carranza Z., & González-Arango C. (2014) Connectivity dynamics since the last glacial maximum in the northern Andes: a pollen-driven framework to assess potential migration. *In: Stevens W.D., Olga Martha Montiel O.M. & Raven P.H (eds), Paleobotany and biogeography : a festschrift for Alan Graham in his 80th year, Missouri Botanical Garden Press, St. Louis,* 98-123.

Foster D.R., Knight D.H., & Franklin J.F. (1998) Landscape Patterns and Legacies Resulting from Large, Infrequent Forest Disturbances. *Ecosystems*, **1**, 497–510.

Garzione C.N., Hoke G..D., Libarkin J.C., Withers S., MacFadden B., Eiler J., Ghosh P., & Mulch A. (2008) Rise of the Andes. *Science*, **320**, 1304–1307.

Gaston K.J. (2000) Global patterns in biodiversity. *Nature*, **405**, 220–227. Gavrilets S. & Losos J.B. (2009) Adaptive radiation: contrasting theory with data. *Science*, **323**, 732–737.

Gavrilets, S., & Losos, J. B. (2009) Adaptive radiation: contrasting theory with data. *Science*, **323**, 732-737.

Gentry A.H. (1982) Neotropical floristic diversity: phytogeographical connections between Central and South America, Pleistocene climatic fluctuations, or an accident of the Andean orogeny? *Annals of the Missouri Botanical Garden*, **69**, 557–593.

Gentry A.H. (1992) Tropical forest biodiversity: distributional patterns and their conservational significance. *Oikos*, **63**, 19–28.

Gerold G. (2008) Soil, climate and vegetation of tropical montane forests – a case study from the Yungas, Bolivia. *In: Gradstein, R., Homeier, J., Gansert, D. (Eds.), The Tropical Mountain Forest. Patterns and Processes in a Biodiversity Hotspot. Biodiversity and Ecology series, 2. Gottingen Centre for Biodiversity and Ecology, Gottingen*, 137–162.

Givnish T.J., Barfuss M.H.J., Van Ee B., Riina R., Schulte K., Horres R., Gonsiska P.A, Jabaily R.S., Crayn D.M., Smith J.A.C., Winter K., Brown G.K., Evans T.M., Holst B.K., Luther H., Till W., Zizka G., Berry P.E., & Sytsma K.J. (2011) Phylogeny, adaptive radiation, and historical biogeography in Bromeliaceae: Insights from an eight-locus plastid phylogeny. *American Journal of Botany*, **98**, 872–895.

Givnish T.J., Barfuss M.H.J., Van Ee B., Riina R., Schulte K., Horres R., Gonsiska P.A., Jabaily R.S., Crayn D.M., Smith J.A.C., Winter K., Brown G.K., Evans T.M., Holst B.K., Luther H., Till W., Zizka G., Berry P.E., Sytsma K.J. (2014) Adaptive radiation, correlated and contingent evolution, and net species diversification in Bromeliaceae. *Molecular Phylogenetics and Evolution*, **71**, 55–78.

Glor R.E. (2010) Phylogenetic insights on adaptive radiation. *Annual Review of Ecology, Evolution, and Systematics*, **41**, 251–270.

Goldberg E.E., Roy K., Lande R., Jablonski D., The S., Naturalist A., & June N. (2005) Diversity, endemism, and age distributions in macroevolutionary sources and sinks. *American Society of Naturalists*, **165**, 623–633.

Graham A. (1984) *Lisianthius* pollen from the Eocene of Panama. *Annals of the Missouri Botanical Garden*, **71**, 987–993.

Graham A., Gregory-Wodzicki K.M., & Wright K.L. (2001) Studies in Neotropical Paleobotany. XV. A Mio-Pliocene palynoflora from the Eastern Cordillera, Bolivia: implications for the uplift history of the Central Andes. *American Journal of Botany*, **88**, 1545–1557.

Grant J.R. (2004) De Macrocarpaeae Grisebach (ex Gentianaceis) speciebus novis V: Twenty-three new species largely from Peru, and typification of all species in the genus. *Harvard Papers in Botany*, **9**, 11–49.

Grant J.R. (2005) De Macrocarpaeae Grisebach (ex Gentianaceis) speciebus novis VI: seed morphology, palynology, an infrageneric classification, and another twenty-three new species largely from Colombia. *Harvard Papers in Botany*, **9**, 305–342.

Ho S.Y.W. & Phillips M.J. (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology*, **58**, 367–380.

Hölscher D. (2008) Hydrology of natural and anthropogenically altered tropical montane rainforests with special reference to rainfall interception. *In: Gradstein, R., Homeier, J., Gansert, D. (Eds.), The Tropical Mountain Forest. Patterns and Processes in a Biodiversity Hotspot. Biodiversity and Ecology series, 2. Gottingen Centre for Biodiversity and Ecology, Gottingen*, 129–136.

Hooghiemstra H. & Van der Hammen T. (2004) Quaternary Ice-Age dynamics in the Colombian Andes: developing an understanding of our legacy. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, **359**, 173–181.

Hooghiemstra H., Wijninga V.M., & Cleef A.M. (2006) The Paleobotanical record of Colombia : implications for biogeography and biodiversity 1. *Annals of the Missouri Botanical Garden* , **93**, 297–324.

Hoorn C., Wesselingh F.P., Ter Steege H., Bermudez M.A., Mora A., Sevink J., Sanmartín I., Sanchez-Meseguer A., Anderson C.L., Figueiredo J.P., Jaramillo C., Riff D., Negri F.R., Hooghiemstra H., Lundberg J., Stadler T., Särkinen T., & Antonelli A. (2010) Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science*, **330**, 927–931.

Hughes C.E. & Eastwood R.J. (2006) Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. *Proceedings of the National Academy of Sciences*, **103**, 10334–10339.

Hughes, C.E. & Atchison G.W. (2015). The ubiquity of alpine plant radiations: from the Andes to the Hengduan Mountains. *New Phytologist*, doi: 10.1111/nph.13230.

Hughes C.E., Pennington R.T., & Antonelli A. (2013) Neotropical plant evolution: assembling the big picture. *Botanical Journal of the Linnean Society*, **171**, 1–18.

Insel N., Poulsen C.J., Ehlers T. A., & Sturm C. (2012) Response of meteoric δ18O to surface uplift - implications for Cenozoic Andean Plateau growth. *Earth and Planetary Science Letters*, **317**, 262–272.

Jablonski D., Roy K., & Valentine J.W. (2006) Out of the tropics: evolutionary dynamics of the latitudinal diversity gradient. *Science*, **314**, 102–6.

Janssens S.B., Knox E.B., Huysmans S., Smets E.F., & Merckx V.S.F.T. (2009) Rapid radiation of *Impatiens* (Balsaminaceae) during Pliocene and Pleistocene: result of a global climate change. *Molecular Phylogenetics and Evolution*, **52**, 806–824.

Jørgensen P.M., Ulloa Ulloa C., León B., León-Yánez S., Beck S.G., Nee M., Zarucchi J.L., Celis M., Bernal R., & Gradstein R. (2011) Regional patterns of vascular plant diversity and endemism. *Climate Change and Biodiversity in the Tropical Andes Inter-American Institute for Global Change Research (IAI) and Scientific Committee on Problems of the Environment (SCOPE), Sao Jose dos Campos and Paris*, 192–203.

Keating P.L. (2008) The floristic composition and biogeographical significance of a megadiverse paramo site in the southern Ecuadorian Andes. *Journal of the Torrey Botanical Society*, **135**, 554–570.

Kessler M., Grytnes J.-A., Halloy S.R.P., Kluge J., Krömer T., León B., Macía M.J., & Young K.R. (2011) Gradients of Plant Diversity: Local Patterns and Processes. *Climate Change and Biodiversity in the Tropical Andes Inter-American Institute for Global Change Research (IAI) and Scientific Committee on Problems of the Environment (SCOPE), Sao Jose dos Campos and Paris*, 204-219.

Leuschner, C., Zach, A., Moser, G., Homeier, J., Graefe, S., Hertel, D., Wittich, B., Soethe, N., Iost, S., Röpderstein, M., Horna, V., and Wolf, K. (2013) The carbon balance of tropical mountain forests along an altitudinal transect. *In: Ecological Studies Vol. 221, edited by: Bendix, J., Springer Verlag, Berlin, Heidelberg, New York*, **10**, 117–139.

Losos J.B. (2010) Adaptive radiation, ecological opportunity, and evolutionary determinism. *The American Naturalist*, **175**, 623–639.

Luebert F. & Weigend M. (2014) Phylogenetic insights into Andean plant diversification. *Frontiers in Ecology and Evolution*, **2**, 1–17.

Madriñán S., Cortés A.J., & Richardson J.E. (2013) Páramo is the world's fastest evolving and coolest biodiversity hotspot. *Frontiers in Genetics*, **4**, 192.

Magallón, S., & Sanderson, M. J. (2001). Absolute diversification rates in angiosperm clades. *Evolution*, **55**, 1762-1780.

Mai D.H. & Walther H. (1988) Die pliozänen Floren von Thüringen, Deutsche Demokratische Republik. *Quartärpaläontologie*, **7**, 55–297.

Matzke N.J. (2013) Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography*, **5**, 242–248.

Matzke N.J. (2014) Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Systematic Biology*, syu056.

McHenry, M.A., & Barrington, D.S. (2014) Phylogeny and biogeography of exindusiate Andean *Polystichum* (Dryopteridaceae) *American Journal of Botany*, **101** (2), 365-375.

Merckx V.S.F.T., Kissling J., Hentrich H., Janssens S.B., Mennes C.B., Specht C.D., & Smets E.F. (2013) Phylogenetic relationships of the mycoheterotrophic genus *Voyria* and the implications for the biogeographic history of Gentianaceae. *American Journal of Botany*, **100**, 712–721.

Miller M.A., Pfeiffer W., & Schwartz T. (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *In: Proceedings of the Gateway Computing Environments Workshop (GCE) New Orleans*; **2010,** 1-8.

Mittelbach G.G., Schemske D.W., Cornell H. V, Allen A.P., Brown J.M., Bush M.B., Harrison S.P., Hurlbert A.H., Knowlton N., Lessios H.A., McCain C.M., McCune A.R., McDade L.A., McPeek M.A., Near T.J.,

Price T.D., Ricklefs R.E., Roy K., Sax D.F., Schluter D., Sobel J.M., & Turelli M. (2007) Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology Letters*, **10**, 315–31.

Moen D. & Morlon H. (2014) Why does diversification slow down? *Trends in Ecology & Evolution*, **29**, 190–7.

Moore, B.R. & Donoghue, M.J. (2007) Correlates of diversification in the plant clade Dipsacales: geographic movement and evolutionary innovations. *The American Naturalist*, **170**, S29–S55.

Mutke, J., Jacobs, R., Meyer, K., Henning, T., & Weigend, M. (2014) Diversity patterns of selected Andean plant groups correspond to topography and habitat dynamics, not orogeny. *Evolutionary and Population Genetics*, **5**, 180.

Orme, C.D.L., Davies, R.G., Burgess, M., Eigenbrod, F., Pickup, N., Olson, V.A., Webster, A.J., Ding, T.-S., Rasmussen, P.C., Ridgely, R.S., Stattersfield, A.J., Bennett, P.M., Blackburn, T.M., Gaston, K.J. & Owens, I.P.F. (2005) Global hotspots of species richness are not congruent with endemism or threat. *Nature*, **436**, 1016–1019.

Paulay G. & Meyer C. (2002) Diversification in the tropical pacific: comparisons between marine and terrestrial systems and the importance of founder speciation. *Integrative and Comparative Biology*, **42**, 922–34.

Pennington R.T., Lavin M., Daríen P.E., Colin P.A., Susan P.K., & Charles B.A. (2004) Historical climate change and speciation: neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 515–537.

Pennington R.T., Lavin M., & Oliveira-Filho A. (2009) Woody plant diversity, evolution, and ecology in the tropics: perspectives from seasonally dry tropical forests. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 437–457.

Pennington R.T., Lavin M., Särkinen T., Lewis G.P., Klitgaard B.B., & Hughes C.E. (2010) Contrasting plant diversification histories within the Andean biodiversity hotspot. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 13783–13787.

Perret M., Chautems A., De Araujo A.O., & Salamin N. (2013) Temporal and spatial origin of Gesneriaceae in the New World inferred from plastid DNA sequences. *Botanical Journal of the Linnean Society*, **171**, 61–79.

Perret M., Chautems A., & Spichiger R. (2006) Dispersal-vicariance analyses in the tribe Sinningieae (Gesneriaceae): a clue to understanding biogeographical history of the Brazilian Atlantic forest. *Annals of the Missouri Botanical Garden*, **93**, 340–358.

Pigot A.L., Phillimore A.B., Owens I.P.F., & Orme C.D.L. (2010) The shape and temporal dynamics of phylogenetic trees arising from geographic speciation. *Systematic Biology*, **59**, 660–73.

Poulsen C.J., Ehlers T.A., & Insel N. (2010) Onset of convective rainfall during gradual late Miocene rise of the central Andes. *Science*, **328**, 490–493.

Rabosky, D.L. (2006) LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies. *Evol Bioinfo Online*, **2**,257–260.

Rabosky D.L. (2014) Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS ONE*, **9**, e89543.

Rabosky D.L., Donnellan S.C., Grundler M., & Lovette I.J. (2014a) Analysis and visualization of complex macroevolutionary dynamics: an example from Australian scincid lizards. *Systematic Biology*, **63**, 610–27.

Rabosky D.L., Grundler M., Anderson C., Title P., Shi J.J., Brown J.W., Huang H., & Larson J.G. (2014b) BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution*, 701–707.

Rabosky D.L. & Lovette I.J. (2008) Density-dependent diversification in North American wood warblers. *Proceedings of the Royal Society B: Biological Sciences*, **275**, 2363–2371.

Raftery, A. E. (1995) Bayesian model selection in social research. *Sociological methodology*, **25**, 111-164.

Raftery A.E., Newton M.A., Satagopan J.M., & Krivitsky P.N. (2007) Estimating the integrated likelihood via posterior simulation using the harmonic mean identity. *In: Bernardo J.M., Bayarri M.J., Berger J.O., Editors. Bayesian Statistics. Oxford: Oxford University Press*, 1–45.

Ree R.H., Moore B.R., Webb C.O., & Donoghue M.J. (2005) A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution*, **59**, 2299–2311.

Renner S.S. (2005) Relaxed molecular clocks for dating historical plant dispersal events. *Trends in Plant Science*, **10**, 550–558.

Richter M., Diertl K.-H., Emck P., Peters T., & Beck E. (2009) Reasons for an outstanding plant diversity in the tropical Andes of southern Ecuador. *Landscape Online*, **12**, 1–35.

Ronquist F. (1997) Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Biology*, **46**, 195–203.

Rull V. (2011) Neotropical biodiversity: timing and potential drivers. *Trends in Ecology & Evolution*, **26**, 508–513.

Rundell R.J. & Price T.D. (2009) Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends in Ecology & Evolution*, **24**, 394–399.

Särkinen T., Pennington T.R., Lavin M., Simon M.F., & Hughes C.E. (2012) Evolutionary islands in the Andes: persistence and isolation explain high endemism in Andean dry tropical forests. *Journal of Biogeography*, **39**, 884–900.

Särkinen T.E., Newman M.F., Maas P.J.M., Maas H., Poulsen A.D., Harris D.J., Richardson J.E., Clark A., Hollingsworth M., & Pennington R.T. (2007) Recent oceanic long-distance dispersal and divergence in the amphi-Atlantic rain forest genus *Renealmia* L.f. (Zingiberaceae) *Molecular Phylogenetics and Evolution*, **44**, 968–80.

Schwery, O., Onstein, R.E., Bouchenak-Khelladi, Y., Xing, Y., Carter, R.J., & Linder, H.P. (2014). As old as the mountains: radiations of the Ericaceae. *New Phytologist*, doi: 10.1111/nph.13234.

Shaw J., Lickey E.B., Beck J.T., Farmer S.B., Liu W., Miller J., Siripun K.C., Winder C.T., Schilling E.E., & Small R.L. (2005) The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany*, **92**, 142–166.

Shaw J., Lickey E.B., Schilling E.E., & Small R.L. (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany*, **94**, 275–288.

Spriggs, E.L., Clement, W.L., Sweeney, P.W., Madriñán, S., Edwards, E.J., & Donoghue, M.J. (2015) Temperate radiations and dying embers of a tropical past: the diversification of *Viburnum*. *New Phytologist*. doi: 10.1111/nph.13305.

Struwe, L., V. A. Albert, M. F. Calió, C. Frasier, K. B. Lepis, K. G. Mathews, & J.R. Grant. (2009a) Evolutionary patterns in neotropical tribe Helieae (Gentianaceae): Evidence from morphology, chloroplast and nuclear DNA sequences. *Taxon*, **58**: 479-499.

Struwe L., Haag S., Heiberg E., & Grant J.R. (2009b) Andean speciation and vicariance in neotropical *Macrocarpaea* (Gentianaceae–Helieae). *Annals of the Missouri Botanical Garden*, **96**, 450–469.

Struwe, L., Kadereit, J., Klackenberg, J., Nilsson, S., Thiv, M., Hagen, K.B. von & Albert, V.A. (2002) Systematics, character evolution, and biogeography of Gentianaceae, including a new tribal and subtribal classification. *In: Struwe, L. & Albert, V.A. (eds.), Gentianaceae—Systematics and Natural History, Cambridge University Press, Cambridge*, 21–309.

Taberlet P., Gielly L., Pautou G., & Bouvet J. (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, **17**, 1105–1109.

Tamura K., Peterson D., Peterson N., Stecher G., Nei M., & Kumar S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, **28**, 2731–2739.

Templeton A.R. (2008) The reality and importance of founder speciation in evolution. *Bioessays*, **30**, 470–479.

Trénel P., Gustafsson M.H.G., Baker W.J., Asmussen-Lange C.B., Dransfield J., & Borchsenius F. (2007) Mid-Tertiary dispersal, not Gondwanan vicariance explains distribution patterns in the wax palm subfamily (Ceroxyloideae: Arecaceae) *Molecular Phylogenetics and Evolution*, **45**, 272–288.

Trénel P., Hansen M.M., Normand S., & Borchsenius F. (2008) Landscape genetics, historical isolation and cross-Andean gene flow in the wax palm, *Ceroxylon echinulatum* (Arecaceae) *Molecular Ecology*, **17**, 3528–3540.

Turchetto-Zolet A.C., Pinheiro F., Salgueiro F., & Palma-Silva C. (2013) Phylogeographical patterns shed light on evolutionary process in South America. *Molecular Ecology*, **22**, 1193–1213.

Tzedakis P.C. (2009) Museums and cradles of Mediterranean biodiversity. *Journal of Biogeography*, **36**, 1033–1034.

Vaidya G., Lohman D.J., & Meier R. (2011) SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*, **27**, 171–180.

Weigend M. (2002) Observations on the biogeography of the Amotape-Huancabamba zone in northern Peru. *The Botanical Review*, **68**, 38–54.

White T.J., Bruns T., Lee S., & Taylor J.W. (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR protocols: a guide to methods and applications*, **18**, 315–322.

Whittaker R., Willis K., & Field R. (2001) Scale and species richness: towards a general, hierarchical theory of species diversity. *Journal of Biogeography*, **28**, 453–470.

Yuan Y.-M., Wohlhauser S., Möller M., Klackenberg J., Callmander M., & Küpfer P. (2005) Phylogeny and biogeography of *Exacum* (Gentianaceae): a disjunctive distribution in the Indian ocean basin resulted from long distance dispersal and extensive radiation. *Systematic Biology*, **54**, 21–34.

SUPPORTING INFORMATION

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

Appendix S1: List of taxa sampled and analysed in the *Macrocarpaea* subclade phylogeny and their corresponding Genbank accession numbers.

Appendix S2: List of taxa used in the Gentianaceae phylogeny and their corresponding Genbank accessions numbers.

Appendix S3: Time calibrated phylogeny of the Gentianaceae.

FIGURE CAPTIONS

Figure 1. Chronogram of subtribe *Macrocarpaea* with 95% highest posterior density bars, based on BEAST analyses of the six-gene dataset (ITS, 5S-NTS, *trnL, trnH-psbA, rpl16 & rpl32*) and a secondary calibration strategy. See main text for details concerning calibration. Posterior probabilities (PP) are displayed for nodes with PP>0.9. Nested in the northern Andean clade, three species from Mesoamerica are indicated by a square and one species from the Pantepui by a triangle. Four species from the northern Andes nested in the central Andean clade are indicated by stars.

Figure 2. Ancestral area reconstruction obtained using the R package BioGeoBEARS. The DIVALIKE+J model was applied to a sample of 100 dated trees from the BEAST posterior distribution (outgroups pruned) summarized and displayed on the *Macrocarpaea* MCC tree. Pie charts represent the average relative probabilities of the ancestral state at each node. Areas: A, Greater Antilles, Caribbean, green; B, Mesoamerica, pink; C, Cordilleras of the northern Andes in Venezuela, Colombia, and northern Ecuador, cyan; D, Amotape–Huancabamba zone in Ecuador and Peru, blue; E, Cordillera Central in central Peru, red; F, Cordillera Central in southern Peru and Bolivia, yellow; G, Guayana Shield, dark brown; H, montane Atlantic forests of Brazil, light brown. Ranges that encompass two or more areas have colours that are a mix of the colours of the areas that compose them (see the legend next to the map).

Figure 3. Species diversification rates through time and among lineages of *Macrocarpaea*. Results from Bayesian Analysis of Macroevolutionary Mixtures (BAMM) applied to the MCC time-calibrated tree of the subtribe *Macrocarpaea*. (a) Mean phylorate plot showing model-averaged diversification rates at any point along every branch. Nested in the northern Andean clade, three species from Mesoamerica are indicated by a square

and one species from the Pantepui by a triangle. Four species from the northern Andean nested in the central Andes clade are indicated by stars. (b) The 95% credible set of shift configurations obtained using a Bayes factor criterion of 30; "f" indicates the posterior probability for each shift configuration. Locations of shifts to higher rates of diversification are indicated by a circle with the size proportional to the marginal probability of the shift. (c) Posterior probabilities of the number of regimes in the model sampled. (d) Bayes factor support for a rate shift on each branch across the tree, with branch lengths scaled according to the Bayes factor support for a rate shift along the branch. The branch with the strongest support for a rate shift is the branch subtending the Andean clade (BF=169).

Figure 4. Variation in diversification rates through time. a. across subtribe *Macrocarpaea* as a whole; b. for the Andean *Macrocarpaea* clade; c. excluding the Andean clade. Black lines represent the mean and the grey polygons the 0.1 to 0.9 quantiles of the posterior distributions of the diversification rates through time

Table 1. Summary data on the sequence matrices used in the divergence time and phylogenetic analyses of Gentianaceae and subtribe *Macrocarpaea*. "Length" is the DNA region length retained in the concatenated matrix. *V* = variables characters, *PI* = parsimony-informative characters, % = percentage of taxa sampled for each DNA sequence locus. "Model" is the best evolutionary model selected with jModelTest.

DNA region	Aligned length	*V*	*PI*	%	Model
ITS	794	577	475	99	TN93+G
Trnl	653	333	190	61	HKY+G
Total Gentianaceae	1447	910	665		unlinked
ITS	681	158	85	85	GTR+G
5S-NTS	381	309	256	94	TN93+G+I
rpl16	769	72	35	80	TN93+G+I
Trnl	822	88	38	83	GTR+G
trnH-psbA	345	63	37	82	GTR+G
rpl32	896	115	51	55	GTR+G
Total subtribe *Macrocarpaea*	3894	805	502		unlinked

Table 2. Stem and crown ages (Ma) for Gentianaceae clades, obtained from BEAST analyses. Lower and higher 95% posterior densities are shown in parentheses. [a] results from the family-level and fossil-constrained analyses based on a two-gene dataset; [b] results from the secondary dating analyses of subtribe *Macrocarpaea* based on the six-gene dataset. We note that the Andean *Macrocarpaea* subclade includes nested within it species from Mesoamerica, the Caribbean and the Guayana Shield.

Clade	Stem age	Crown age
Gentianaceae [a]	81.99 (66.61, 99.26)	76.84 (61.56, 92.39)
Helieae [a]	46.63 (37.59, 57.11)	36.17 (24.67, 47.31)
Subtribe *Macrocarpaea* [a]	23.97 (17.61, 31.07)	20.55 (13.96, 27.48)
Macrocarpaea [a]	20.55 (13.96, 27.48)	11.04 (7.29, 15.13)
Tachia [b]	17.16 (12.23, 22.03)	7.98 (4.77, 12.04)
Brazilian *Macrocarpaea* subclade [b]	11.03 (7.82, 14.54)	5.11 (2.46, 8.20)
Andean *Macrocarpaea* subclade [b]	11.03 (7.82, 14.54)	7.16 (5.21, 9.77)

Table 3. Summary statistics for the biogeographical model selection procedure. Each model was applied to a sample of 100 posterior trees from BEAST (outgroups were removed), using a ML optimization approach (BioGeoBEARS, R). *d* anagenetic range-expansion; *e* range-contraction free parameters; *j* founder-event cladogenetic free parameter. AICc is the corrected Akaike information criterion for each model. Δ AICc is the AICc difference between the best model and each model. ω Akaike weight for each model. The evidence ratio is the ω of a model divided by the ω of the best model. % refers to the percentage of trees for which each model obtained the lower AICc. All values (except %) are means, estimated over the sample of 100 trees.

Model	*d*	*e*	*J*	AICc	Δ AICc	*Ω*	Evidence ratio	%
DIVALIKE	0.022	4E-03	NA	217.3	57.431	3E-11	3E+14	0
DEC	0.018	0.013	NA	223.5	63.607	1E-13	5E+14	0
DIVALIKE+J	2E-12	6E-11	0.035	159.9	0	0.749	1	100
DEC+J	0	2E-10	0.036	162.1	2.19	0.251	2.992	0

Table 4. Comparison of scenarios for the ancestral area of the Andean *Macrocarpaea* clade. Scenarios involved constraining the state of the MRCA as one of seven possible areas (on 100 trees, with BioGeoBEARS, R). C, D, E, F and their combinations refer to the areas depicted in Figure 3. For an explanation of AICc, Δ AICc, ω, Evidence ratio and % refer to Table 3. Parentheses indicate the percentage of trees for which two scenarios obtained identical AICc.

Scenario	AICc	Δ AICc	Ω	Evidence ratio	%
C	162.7	0	0.441	1	61 (9)
D	162.7	-0.008	0.409	1.792	26 (10)
E	170.9	7.610	0.012	133.796	0
F	169.6	6.832	0.047	209.383	3 (1)
CD	166.3	3.578	0.075	8.943	0
EF	176.5	13.788	10E-4	5554.218	0
CDEF	171.3	8.604	0.016	779.071	0

Table 5. Comparison of size, age and diversification rates for Andean MMF plant radiations for which there are well-sampled time-calibrated phylogenies. Diversification rates reported are estimates reported in the original study. Diversification rates were estimated by using Magallón & Sanderson (2001) equation 4 (r = log(n)−log(2)]/T, where n =standing diversity, and T =inferred crown clade age) in the R package GEIGER. Confidence intervals when available are displayed in parenthesis.

Clade	Number of species	Crown age (My)	Diversification reported ($sp.My^{-1}$)	Diversification rate estimated ($sp.My^{1}$)	Reference
Andean *Macrocarpaea*	96	7.2 (5.2-9.8)	0.43	0.54 (0.74-0.40)	This paper
Andean *Oreinotinus* clade of *Viburnum*	16	6? (NA-NA)	NA	0.35 (NA-NA)	Spriggs *et al.* (2015)
Core Tillandsioid clade	1236	9.6 (NA-NA)	0.67	0.67 (NA-NA)	Givnish *et al.* (2014)
Supersection *Tacsonia* of *Passiflora*	62-64	8.4 (6.2-11.2)	NA	0.41 (0.56-0.31)	Abrahamczyk *et al.* (2014)
Tafalla	40	15 (10.8-29.1)	0.13	0.20 (0.28-0.10)	Antonelli & Sanmartín (2011)

Figure 1.

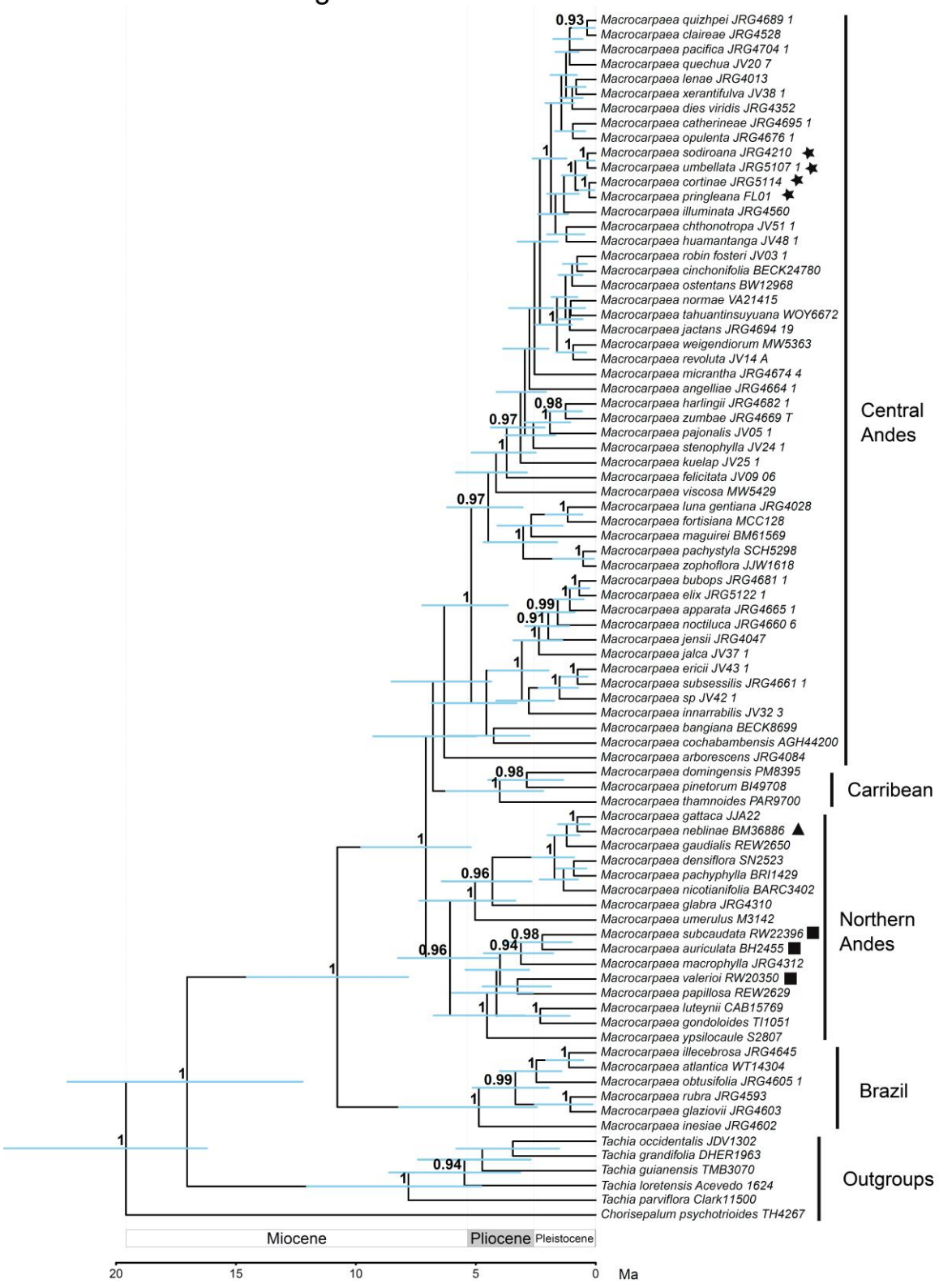

Macrocarpaea quizhpei JRG4689 1
Macrocarpaea claireae JRG4528
Macrocarpaea pacifica JRG4704 1
Macrocarpaea quechua JV20 7
Macrocarpaea lenae JRG4013
Macrocarpaea xerantifulva JV38 1
Macrocarpaea dies viridis JRG4352
Macrocarpaea catherineae JRG4695 1
Macrocarpaea opulenta JRG4676 1
Macrocarpaea sodiroana JRG4210
Macrocarpaea umbellata JRG5107 1
Macrocarpaea cortinae JRG5114
Macrocarpaea pringleana FL01
Macrocarpaea illuminata JRG4560
Macrocarpaea chthonotropa JV51 1
Macrocarpaea huamantanga JV48 1
Macrocarpaea robin fosteri JV03 1
Macrocarpaea cinchonifolia BECK24780
Macrocarpaea ostentans BW12968
Macrocarpaea normae VA21415
Macrocarpaea tahuantinsuyuana WOY6672
Macrocarpaea jactans JRG4694 19
Macrocarpaea weigendiorum MW5363
Macrocarpaea revoluta JV14 A
Macrocarpaea micrantha JRG4674 4
Macrocarpaea angelliae JRG4664 1
Macrocarpaea harlingii JRG4682 1
Macrocarpaea zumbae JRG4669 T
Macrocarpaea pajonalis JV05 1
Macrocarpaea stenophylla JV24 1
Macrocarpaea kuelap JV25 1
Macrocarpaea felicitata JV09 06
Macrocarpaea viscosa MW5429
Macrocarpaea luna gentiana JRG4028
Macrocarpaea fortisiana MCC128
Macrocarpaea maguirei BM61569
Macrocarpaea pachystyla SCH5298
Macrocarpaea zophoflora JJW1618
Macrocarpaea bubops JRG4681 1
Macrocarpaea elix JRG5122 1
Macrocarpaea apparata JRG4665 1
Macrocarpaea noctiluca JRG4660 6
Macrocarpaea jensii JRG4047
Macrocarpaea jalca JV37 1
Macrocarpaea ericii JV43 1
Macrocarpaea subsessilis JRG4661 1
Macrocarpaea sp JV42 1
Macrocarpaea innarrabilis JV32 3
Macrocarpaea bangiana BECK8699
Macrocarpaea cochabambensis AGH44200
Macrocarpaea arborescens JRG4084
Macrocarpaea domingensis PM8395
Macrocarpaea pinetorum BI49708
Macrocarpaea thamnoides PAR9700
Macrocarpaea gattaca JJA22
Macrocarpaea neblinae BM36886
Macrocarpaea gaudialis REW2650
Macrocarpaea densiflora SN2523
Macrocarpaea pachyphylla BRI1429
Macrocarpaea nicotianifolia BARC3402
Macrocarpaea glabra JRG4310
Macrocarpaea umerulus M3142
Macrocarpaea subcaudata RW22396
Macrocarpaea auriculata BH2455
Macrocarpaea macrophylla JRG4312
Macrocarpaea valerioi RW20350
Macrocarpaea papillosa REW2629
Macrocarpaea luteynii CAB15769
Macrocarpaea gondoloides TI1051
Macrocarpaea ypsilocaule S2807
Macrocarpaea illecebrosa JRG4645
Macrocarpaea atlantica WT14304
Macrocarpaea obtusifolia JRG4605 1
Macrocarpaea rubra JRG4593
Macrocarpaea glaziovii JRG4603
Macrocarpaea inesiae JRG4602
Tachia occidentalis JDV1302
Tachia grandifolia DHER1963
Tachia guianensis TMB3070
Tachia loretensis Acevedo 1624
Tachia parviflora Clark11500
Chorisepalum psychotrioides TH4267
Central Andes
Carribean
Northern Andes
Brazil
Outgroups
Miocene
Pliocene
Pleistocene
20
15
10
5
0
Ma

Figure 2.

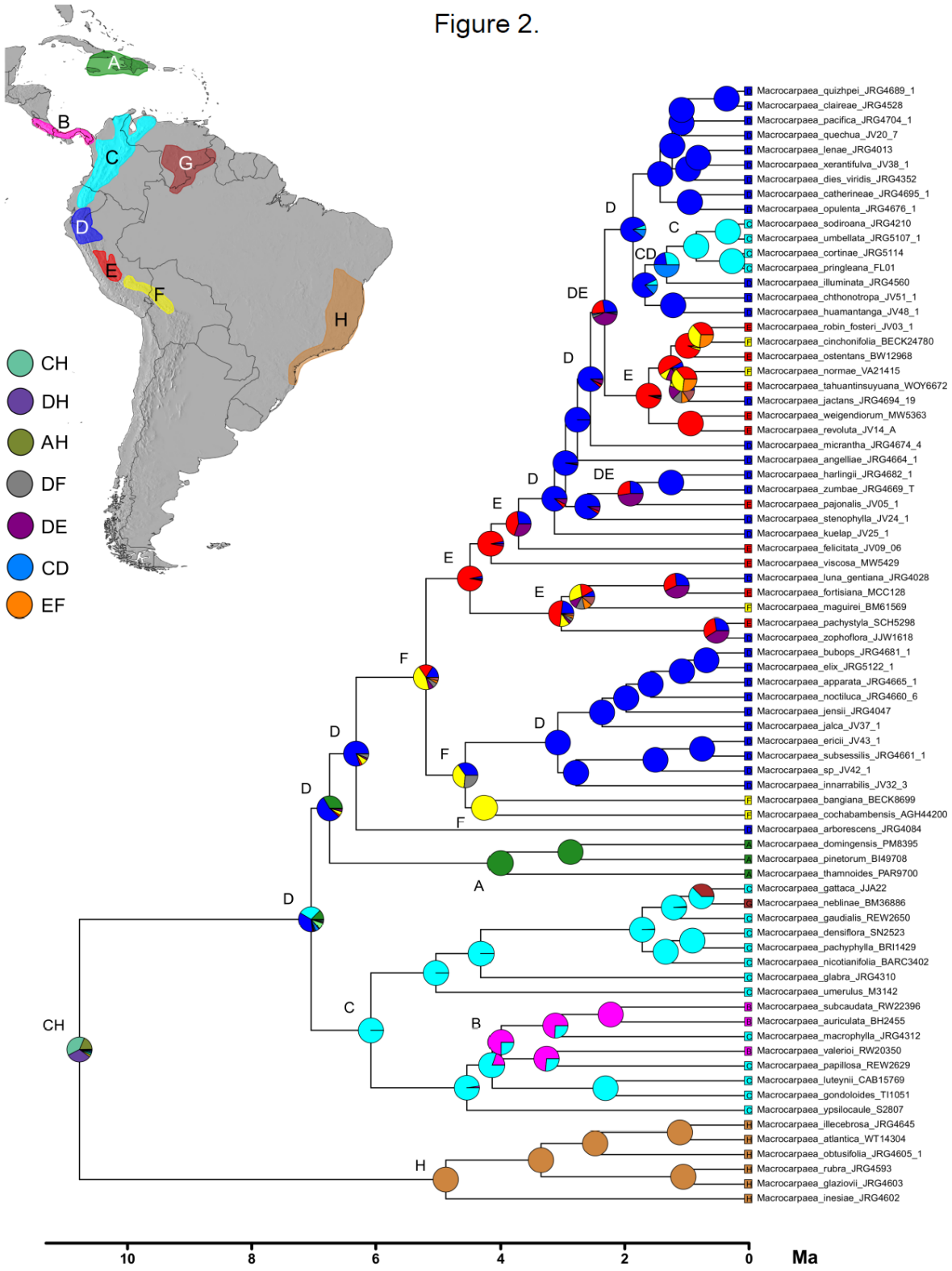

A
B
C
D
E
F
G
H
CH
DH
AH
DF
DE
CD
EF
Macrocarpaea_quizhpei_JRG4689_1
Macrocarpaea_claireae_JRG4528
Macrocarpaea_pacifica_JRG4704_1
Macrocarpaea_quechua_JV20_7
Macrocarpaea_lenae_JRG4013
Macrocarpaea_xerantifulva_JV38_1
Macrocarpaea_dies_viridis_JRG4352
Macrocarpaea_catherineae_JRG4695_1
Macrocarpaea_opulenta_JRG4676_1
Macrocarpaea_sodiroana_JRG4210
Macrocarpaea_umbellata_JRG5107_1
Macrocarpaea_cortinae_JRG5114
Macrocarpaea_pringleana_FL01
Macrocarpaea_illuminata_JRG4560
Macrocarpaea_chthonotropa_JV51_1
Macrocarpaea_huamantanga_JV48_1
Macrocarpaea_robin_fosteri_JV03_1
Macrocarpaea_cinchonifolia_BECK24780
Macrocarpaea_ostentans_BW12968
Macrocarpaea_normae_VA21415
Macrocarpaea_tahuantinsuyuana_WOY6672
Macrocarpaea_jactans_JRG4694_19
Macrocarpaea_weigendiorum_MW5363
Macrocarpaea_revoluta_JV14_A
Macrocarpaea_micrantha_JRG4674_4
Macrocarpaea_angelliae_JRG4664_1
Macrocarpaea_harlingii_JRG4682_1
Macrocarpaea_zumbae_JRG4669_T
Macrocarpaea_pajonalis_JV05_1
Macrocarpaea_stenophylla_JV24_1
Macrocarpaea_kuelap_JV25_1
Macrocarpaea_felicitata_JV09_06
Macrocarpaea_viscosa_MW5429
Macrocarpaea_luna_gentiana_JRG4028
Macrocarpaea_fortisiana_MCC128
Macrocarpaea_maguirei_BM61569
Macrocarpaea_pachystyla_SCH5298
Macrocarpaea_zophoflora_JJW1618
Macrocarpaea_bubops_JRG4681_1
Macrocarpaea_elix_JRG5122_1
Macrocarpaea_apparata_JRG4665_1
Macrocarpaea_noctiluca_JRG4660_6
Macrocarpaea_jensii_JRG4047
Macrocarpaea_jalca_JV37_1
Macrocarpaea_ericii_JV43_1
Macrocarpaea_subsessilis_JRG4661_1
Macrocarpaea_sp_JV42_1
Macrocarpaea_innarrabilis_JV32_3
Macrocarpaea_bangiana_BECK8699
Macrocarpaea_cochabambensis_AGH44200
Macrocarpaea_arborescens_JRG4084
Macrocarpaea_domingensis_PM8395
Macrocarpaea_pinetorum_BI49708
Macrocarpaea_thamnoides_PAR9700
Macrocarpaea_gattaca_JJA22
Macrocarpaea_neblinae_BM36886
Macrocarpaea_gaudialis_REW2650
Macrocarpaea_densiflora_SN2523
Macrocarpaea_pachyphylla_BRI1429
Macrocarpaea_nicotianifolia_BARC3402
Macrocarpaea_glabra_JRG4310
Macrocarpaea_umerulus_M3142
Macrocarpaea_subcaudata_RW22396
Macrocarpaea_auriculata_BH2455
Macrocarpaea_macrophylla_JRG4312
Macrocarpaea_valerioi_RW20350
Macrocarpaea_papillosa_REW2629
Macrocarpaea_luteynii_CAB15769
Macrocarpaea_gondoloides_TI1051
Macrocarpaea_ypsilocaule_S2807
Macrocarpaea_illecebrosa_JRG4645
Macrocarpaea_atlantica_WT14304
Macrocarpaea_obtusifolia_JRG4605_1
Macrocarpaea_rubra_JRG4593
Macrocarpaea_glaziovii_JRG4603
Macrocarpaea_inesiae_JRG4602
10
8
6
4
2
0
Ma

Figure 3.

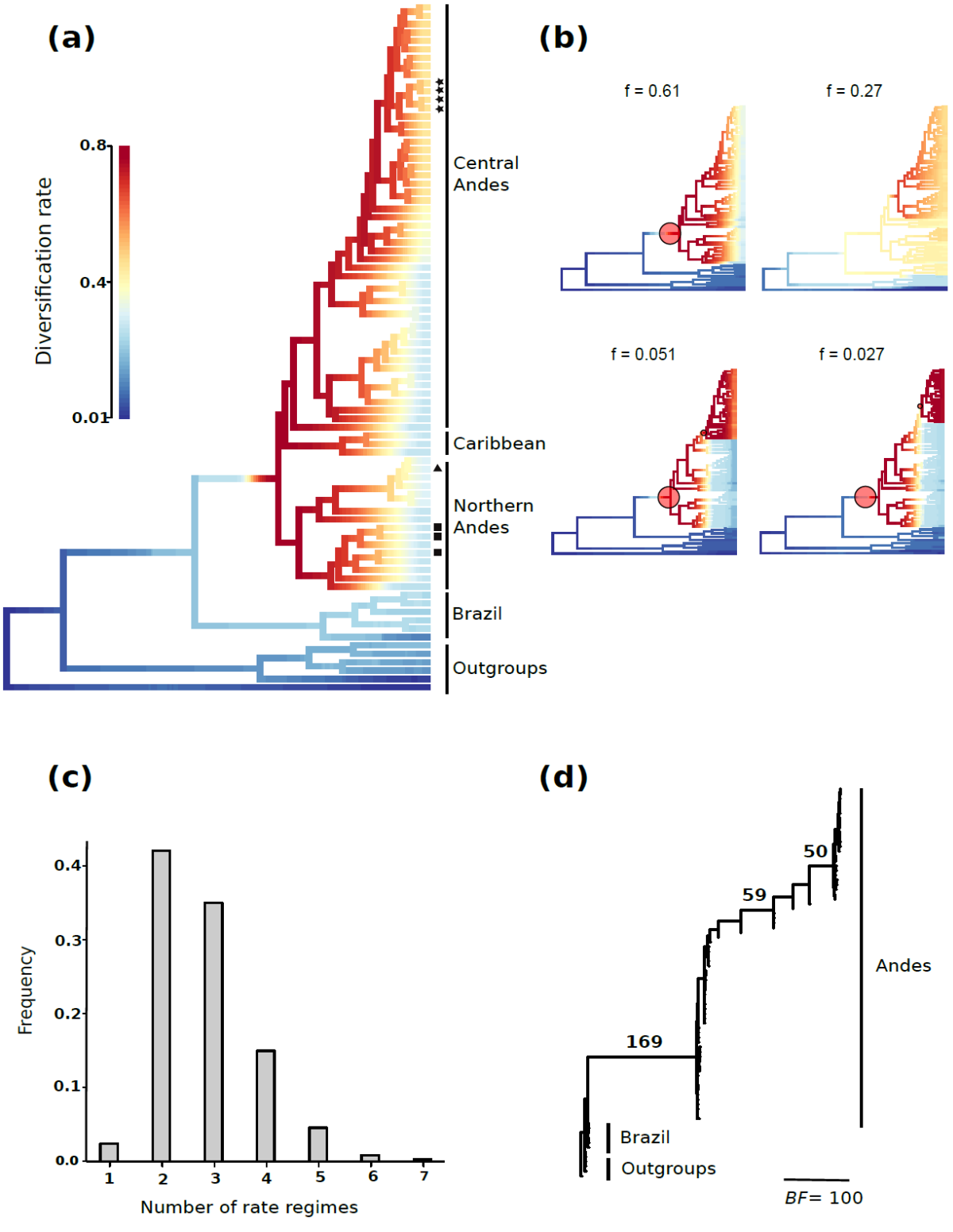
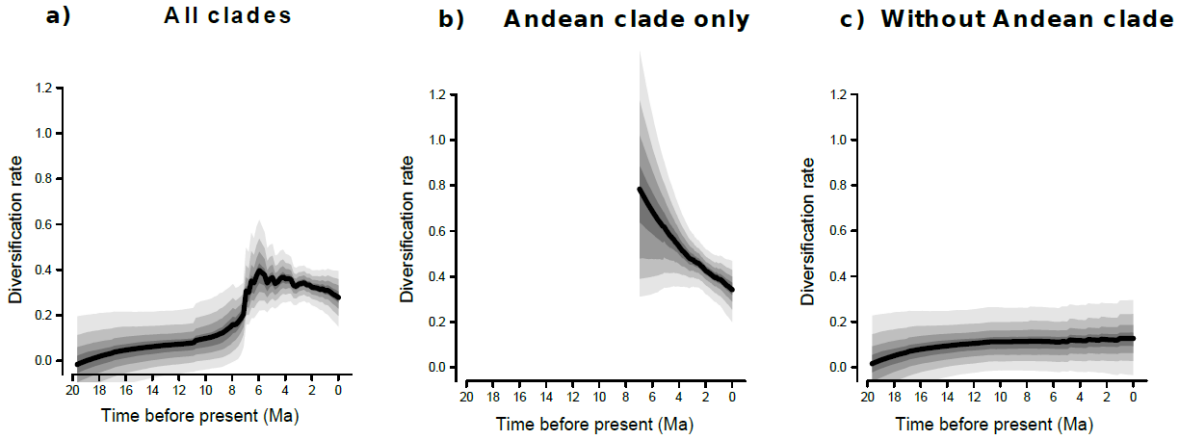

(a)
Diversification rate
0.8
0.4
0.01
Central Andes
Caribbean
Northern Andes
Brazil
Outgroups
(b)
f = 0.61
f = 0.27
f = 0.051
f = 0.027
(c)
Frequency
0.4
0.3
0.2
0.1
0.0
1
2
3
4
5
6
7
Number of rate regimes
(d)
50
59
169
Andes
Brazil
Outgroups
BF= 100

Figure 4.


a) All clades
b) Andean clade only
c) Without Andean clade
Diversification rate
Time before present (Ma)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
20 18 16 14 12 10 8 6 4 2 0

Evolutionary diversification in the hyper-diverse montane forests of the tropical Andes: range expansion drives radiation of the plant genus *Macrocarpaea* (Gentianaceae)

Vieu Julien C.[1], Hughes Colin E.[2] & Grant Jason R.[1]

SUPPORTING INFORMATION

Appendix S1: List of taxa used to reconstruct the time calibrated phylogeny of the subclade *Macrocarpaea* of the tribe Helieae (Gentianaceae) and their GenBank accession numbers. The molecular dataset is comprised of two ribosomal DNA intergenic spacers, ITS (internal transcribed spacer: ITS1, 5.8S, ITS2) and 5S-NTS (non transcribed spacer) and four non-coding plastid DNA regions, the *rpL16* introns, the *trnH-psbA* intergenic spacer, the *trnL-trnF* introns and exons and the *rpl32-trnL* intergenic spacer. Note for editors and reviewers: XXXX indicates sequences new to this study in GenBank submission process.

Organism	Isolate	Collected-by	Country	ITS	5S-NTS	*rpl16*	*trnL*	*trnH-psbA*	*rpl32*
Chorisepalum psychotrioides	TH4267	T.W. Hankel 4267 (NY)	Guyana	EU709793					
Macrocarpaea angelliae	JRG4664_1	J.R. Grant & J. Vieu 4664 (NY)	Ecuador	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea apparata	JRG4665_1	J.R. Grant & J. Vieu 4665 (NY)	Ecuador	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea arborescens	JRG4084	J.R. Grant 4084 (NY)	Ecuador	EU528076	EU541686	EU541801	EU528051	EU541781	XXXX
Macrocarpaea atlantica	WT14304	W. Thomas 14304 (NY)	Brazil	EU528125	EU541775	EU541810	EU528057	EU541790	
Macrocarpaea auriculata	BH2455	B. Haber 2455 (MO)	Costa Rica		EU541687	XXXX	XXXX	XXXX	
Macrocarpaea bangiana	BECK8699	S.G. Beck 8699 (NEU)	Bolivia	EU528079	EU541689	XXXX	XXXX	XXXX	
Macrocarpaea bubops	JRG4681_1	J.R. Grant & J. Vieu 4681 (NY)	Ecuador	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea catherineae	JRG4695_1	J.R. Grant & J. Vieu 4695 (NY)	Ecuador	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea chthonotropa	JV51_1	J. Vieu & D. Desrousseaux 47 (NEU)	Peru	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea cinchonifolia	BECK24780	S.G. Beck 24780 (NY)	Bolivia	EU528084	EU541696	EU541806		EU541786	XXXX
Macrocarpaea claireae	JRG4528	J.R. Grant 4528 (NY)	Ecuador	EU528068	EU541672				
Macrocarpaea cochabambensis	AGH44200	A. Gentry 44200 (NY)	Bolivia	EU528085	EU541697	EU541796	EU528047	EU541776	XXXX
Macrocarpaea cortinae	JRG5114	J.R. Grant & J. Cortina 5114 (NY)	Brazil	XXXX	XXXX	XXXX	XXXX	XXXX	
Macrocarpaea densiflora	SN2523	K. von Sneidern 2523 (S)	Colombia	EU528087	EU541700				

Macrocarpaea dies-viridis	JRG4352	J.R. Grant 4352 (NY)	Ecuador	EU528089	EU541702	EU541809	EU528056	EU541789	XXXX
Macrocarpaea domingensis	PM8395	P. Maas 8395 (NY)	Dominican Republic	EU528092	EU541705	XXXX	XXXX	XXXX	XXXX
Macrocarpaea elix	JRG5122_1	J.R. Grant & J. Cortina (NY)	Ecuador	XXXX	XXXX	XXXX	XXXX	XXXX	
Macrocarpaea ericii	JV43_1	J. Vieu & D. Desrousseaux 39 (NEU)	Peru	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea felicitata	JV09_6	J. Vieu & D. Desrousseaux 12x (NEU)	Peru	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea fortisiana	MCC128	D. McCarroll 128 (NY)	Peru	EU528095	EU541709	XXXX	XXXX	XXXX	
Macrocarpaea gattaca	JJA22	J. Jaramillo 22 (NY)	Ecuador	EU528096	EU541711	XXXX	XXXX	XXXX	XXXX
Macrocarpaea gaudialis	REW2650	R.E. Weaver 2650 (GH)	Colombia	EU528097	EU541713	XXXX	XXXX	XXXX	
Macrocarpaea glabra	JRG4310	J.R. Grant 4310 (NY)	Colombia	EU528098	EU541714	EU541812	EU528060	EU541792	XXXX
Macrocarpaea glaziovii	JRG4603	J.R. Grant, L. Zeltner, B. Resende Silva & V. Trunz 4603 (NY)	Brazil			XXXX	XXXX	XXXX	
Macrocarpaea gondoloides	TI1051	G. Tipaz 1051 (MO)	Ecuador	XXXX	EU541716				
Macrocarpaea harlingii	JRG4682_1	J.R. Grant & J. Vieu 4682 (NY)	Ecuador	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea huamantanga	JV48_1	J. Vieu & D. Desrousseaux 44 (NEU)	Peru	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea illecebrosa	JRG4645	J.R. Grant & V. Trunz 4645 (NY)	Brazil	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea illuminata	JRG4560	J.R. Grant, B. Angell, W. Grant & V. Trunz 4560 (LOJA)	Ecuador	XXXX			XXXX		
Macrocarpaea inesiae	JRG4602	J.R. Grant, L. Zeltner, F. Calió & V. Trunz 4602 (NY)	Brazil			XXXX	XXXX	XXXX	
Macrocarpaea innarrabilis	JV32_3	J. Vieu & D. Desrousseaux 30 (NEU)	Peru	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea jactans	JRG4694_19	J.R. Grant & J. Vieu 4694 (NY)	Ecuador	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea jalca	JV37_1	J. Vieu & D. Desrousseaux 34 (NEU)	Peru	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea jensii	JRG4047	J.R. Grant 4047 (NY)	Ecuador	EU528107	EU541724	XXXX	XXXX	XXXX	XXXX
Macrocarpaea kuelap	JV25_1	J. Vieu & D. Desrousseaux 21 (NEU)	Peru	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea lenae	JRG4013	J.R. Grant 4013 (NY)	Ecuador	EU528110	EU541727	XXXX	XXXX	XXXX	XXXX
Macrocarpaea luna-gentiana	JRG4028	J.R. Grant 4028 (NY)	Ecuador	EU528111	EU541728	EU541800	EU528050	EU541780	XXXX
Macrocarpaea luteynii	CAB15769	I. Cabrera & H. van der Werff (COL)	Colombia		EU541730	XXXX		XXXX	

Macrocarpaea macrophylla	JRG4312	J.R. Grant 4312 (NY)	Colombia	EU528113	EU541735	EU541813	EU528061	EU541795	XXXX
Macrocarpaea maguirei	BM61569	B. Maguire 61569 (NY)	Peru	EU528114	EU541736	XXXX	XXXX	XXXX	
Macrocarpaea micrantha	JRG4674_4	J.R. Grant & J. Vieu 4674 (NY)	Ecuador	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea neblinae	BM36886	B. Maguire 36886 (NY)	Venezuela		EU541739	XXXX			
Macrocarpaea nicotianifolia	BARC3402	A.S. Barclay 3402 (US)	Colombia	EU528118	EU541740	XXXX	XXXX	XXXX	
Macrocarpaea noctiluca	JRG4660_6	J.R. Grant & J. Vieu 4660 (NY)	Ecuador	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea normae	VA21415	C. Vargas 21415 (CUZ)	Peru		EU541744	XXXX	XXXX	XXXX	
Macrocarpaea obtusifolia	JRG4605_1	J.R. Grant, L. Zeltner, B. Resende Silva & V. Trunz 4605 (NY)	Brazil	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea opulenta	JRG4676_1	J.R. Grant & J. Vieu 4676 (NY)	Ecuador	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea ostentans	BW12968	B. Wallnöfer 12968 (U)	Peru	EU528128	EU541747	XXXX	XXXX	XXXX	
Macrocarpaea pachyphylla	BRI1429	M.L. Bristol 1429 (GH)	Colombia	EU528129	EU541748				
Macrocarpaea pachystyla	SCH5298	J. Schunke-Vigo 5298 (NY)	Peru	EU528130	EU541751			XXXX	
Macrocarpaea pacifica	JRG4704_1	J.R. Grant & J. Vieu 4704 (NY)	Ecuador	XXXX	XXXX	XXXX	XXXX	XXXX	
Macrocarpaea pajonalis	JV05_1	J. Vieu & D. Desrousseaux 05b (NEU)	Peru	XXXX	XXXX	XXXX	XXXX	XXXX	
Macrocarpaea papillosa	REW2629	R.E. Weaver 2629 (GH)	Venezuela	EU528131	EU541750	XXXX	XXXX	XXXX	
Macrocarpaea pinetorum	BI49708	J. Bisse 49708 (HAJB)	Cuba	EU528133	EU541753	EU541797	XXXX	EU541777	
Macrocarpaea pringleana	FL01	F. Luisier 1 (NY)	Ecuador	EU528134	EU541755	XXXX	XXXX	XXXX	XXXX
Macrocarpaea quechua	JV20_7	J. Vieu & D. Desrousseaux 18 (NEU)	Peru	XXXX			XXXX	XXXX	XXXX
Macrocarpaea quizhpei	JRG4689_1	J.R. Grant & J. Vieu 4689 (NY)	Ecuador	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea revoluta	JV14_A	J. Vieu & D. Desrousseaux 14 (NEU)	Peru	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea robin-fosteri	JV03_1	J. Vieu & D. Desrousseaux 03 (NEU)	Peru	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea rubra	JRG4593	J.R. Grant, L. Zeltner & V. Trunz 4593 (NY)	Brazil	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea sodiroana	JRG4210	J.R. Grant 4210 (NY)	Ecuador	EU528140	EU541758		XXXX	XXXX	XXXX
Macrocarpaea sp.	JV42_1	J. Vieu & D. Desrousseaux 38 (NEU)	Peru	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX

Macrocarpaea stenophylla	JV24_1	J. Vieu & D. Desrousseaux 22 (NEU)	Peru	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea subcaudata	RW22396	R.L. Wilbur & Almeda 16828 (DUKE)	Costa Rica	EU528143	EU541763	EU541805	XXXX	EU541785	XXXX
Macrocarpaea subsessilis	JRG4661_1	J.R. Grant & J. Vieu 4661 (NY)	Ecuador	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea tahuantinsuyuana	WOY6672	F. Woytkowski 6672A (MO)	Peru		EU541766	XXXX	XXXX		
Macrocarpaea thamnoides	PAR9700	P. Acevedo-Rodríguez 9700 (NY)	Jamaica		EU541767	XXXX	XXXX		
Macrocarpaea umbellata	JRG5107_1	J.R. Grant & J. Cortina 5107 (NEU)	Ecuador	XXXX	XXXX	XXXX	XXXX	XXXX	
Macrocarpaea umerulus	M3142	H. Mendoza 3142 (FMB)	Colombia		XXXX		XXXX		
Macrocarpaea valerioi	RW20350	R.L. Wilbur 20350 (NY)	Costa Rica	EU528147	EU541768	XXXX	XXXX	XXXX	
Macrocarpaea viscosa	MW5429	M. Weigend 5429 (NY)	Peru	EU528149	EU541770	EU541804	EU528054	EU541784	XXXX
Macrocarpaea weigendiorum	MW5363	M. Weigend 5363 (NY)	Peru	EU528150	EU541771	EU541803	EU528053	EU541783	XXXX
Macrocarpaea xerantifluva	JV38_1	J. Vieu & D. Desrousseaux 34bis (NEU)	Peru	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Macrocarpaea ypsilocaule	S2807	D. Stancik 2807 (COL)	Colombia		XXXX			XXXX	
Macrocarpaea zophoflora	JJW1618	J.J. Wurdack 1618 (NY)	Peru		EU541772	XXXX	XXXX	XXXX	
Macrocarpaea zumbae	JRG4669_T	J.R. Grant & J. Vieu 4669 (NY)	Ecuador	XXXX	XXXX	XXXX	XXXX	XXXX	XXXX
Tachia grandiflora	DHER1963	J.H.E. Rova 1963	French Guiana	DQ401418	DQ401429				
Tachia guianensis	TMB3070	T. McDowell 3670 (US)	Guyana	DQ401420	DQ401430				
Tachia loretensis	PA1624	P. Acevedo 1624 (NY)	Peru	DQ401421	XXXX				
Tachia occidentalis	JDV1302	J. Duivenvoorden 1302 (NY)	Colombia	DQ401423	DQ401427				
Tachia parviflora	Clark11500	J.L Clark 11500 (NY)	Peru	XXXX	XXXX	XXXX	XXXX	XXXX	

Evolutionary diversification in the hyper-diverse montane forests of the tropical Andes: range expansion drives radiation of the plant genus *Macrocarpaea* (Gentianaceae)

Vieu Julien C.[1], Hughes Colin E.[2] & Grant Jason R.[1]

SUPPORTING INFORMATION

Appendix S2: ***List of taxa used to reconstruct the time-calibrated phylogeny of the Gentianaceae and their GenBank accession numbers. ITS full means that the sequence contained the internal transcribed spacer 1 and 2 (ITS1 & ITS) and the 5.8S ribosomal DNA region. Note for editors and reviewers: XXXX indicates sequences new to this study in GenBank submission process.***

Organism	voucher	*trnL*	ITS full	ITS1	ITS2
Adenolisianthus arboreus	Sequences obtained from Genbank		EU709784		
Anthocleista amplexicaulis	Sequences obtained from Genbank	AF102375	DQ449914		
Anthocleista grandiflora	Sequences obtained from Genbank	AJ490190	AJ489864		
Anthocleista scandens	Sequences obtained from Genbank	AF102376	DQ449917		
Anthocleista vogelii	Sequences obtained from Genbank	AF102377	DQ449915		
Apocynum cannabinum	Sequences obtained from Genbank	AF102380	DQ005966		
Aripuana cullmaniorum	Sequences obtained from Genbank		EU709785		
Bartonia paniculata	Sequences obtained from Genbank		EU812472		
Bartonia verna	Sequences obtained from Genbank		EU812473		
Bartonia virginica	Sequences obtained from Genbank	AF102383		AJ306078	AJ306106
Bisgoeppertia scandens	Sequences obtained from Genbank		FJ232556		
Blackstonia acuminata	Sequences obtained from Genbank	AY251743		AY251685	AY251715
Blackstonia grandiflora	Sequences obtained from Genbank	AY251742		AY251684	AY251714
Blackstonia perfoliata	Sequences obtained from Genbank	AF402254		AY251686	AY047878
Calolisianthus pulcherrimus	Sequences obtained from Genbank	AF102388	EU709786		
Calolisianthus speciosus	Sequences obtained from Genbank		EU709787		
Canscora alata	Sequences obtained from Genbank	AJ490191	AJ489865		
Canscora andrographioides	Sequences obtained from Genbank	AJ490192	AJ489866		
Canscora diffusa	Sequences obtained from Genbank	AJ490193	AJ489867		
Centaurium bianoris	Sequences obtained from Genbank	DQ166305		DQ166323	DQ166368
Centaurium centaurioides	Sequences obtained from Genbank	AY879913		AY879810	AY879861
Centaurium chloodes	Sequences obtained from Genbank	AY879914		AY879811	AY879862
Centaurium gypsicola	Sequences obtained from Genbank	AY251731		AY251671	AY251701
Centaurium malzacianum	Sequences obtained from Genbank	AY879932		AY879831	AY879883
Centaurium serpentinicola	Sequences obtained from Genbank	AY879942		AY879841	AY879892
Centaurium turcicum	Sequences obtained from Genbank	AY879954		AY879853	AY879903
Chelonanthus alatus	Sequences obtained from Genbank	AJ490194	AJ489868		
Chelonanthus albus	Sequences obtained from Genbank	AF102397	EU709789		
Chelonanthus angustifolius	Sequences obtained from Genbank	AJ490195	AJ489869		
Chelonanthus grandiflorus	Sequences obtained from Genbank		EU709788		
Chelonanthus purpurascens	Sequences obtained from Genbank		FJ232584		
Chelonanthus sp	J. Vieu & D. Desrousseaux JV09 (NEU)	XXXX	XXXX		
Chelonanthus sp	J. Vieu & D. Desrousseaux JV28 (NEU)	XXXX	XXXX		
Chelonanthus viridiflorus	Sequences obtained from Genbank	AF102399	EU709792		

Chionogentias muelleriana	Sequences obtained from Genbank		AY136507		
Chionogentias polysperes	Sequences obtained from Genbank		AY136505		
Chironia baccifera	Sequences obtained from Genbank	AJ490197	AJ489871		
Chironia laxa	Sequences obtained from Genbank	AJ490198	AJ489872		
Chironia linoides	Sequences obtained from Genbank	AJ490199	AJ489873		
Chironia purpurascens	Sequences obtained from Genbank	FJ014139	FJ666016		
Chorisepalum psychotrioides	Sequences obtained from Genbank		EU709793		
Cicendia filiformis	Sequences obtained from Genbank	AF102403		AJ011463	AJ011473
Cicendia quadrangularis	Sequences obtained from Genbank			AY251682	AY251712
Coffea Arabica	Sequences obtained from Genbank	AF102405	EU650386		
Comastoma pedunculatum	Sequences obtained from Genbank		AF346009		
Comastoma pulmonarium	Sequences obtained from Genbank		AF346008		
Comastoma tenellum	Sequences obtained from Genbank	AJ580518	AJ580552		
Comastoma traillianum	Sequences obtained from Genbank		DQ317493		
Coutoubea spicata	Sequences obtained from Genbank		EU709780		
Crawfurdia campanulacea	Sequences obtained from Genbank		GU251014		
Crawfurdia delavayi	Sequences obtained from Genbank	AY563391	AY562176		
Crawfurdia sessiliflora	Sequences obtained from Genbank		GU251020		
Crawfurdia speciosa	Sequences obtained from Genbank	AY858682	AY858675		
Curtia tenuifolia	Sequences obtained from Genbank			AJ242613	AJ242614
Enicostema axillare	Sequences obtained from Genbank		FJ232582		
Enicostema verticillatum	Sequences obtained from Genbank	AF102414	FJ232580		
Erithalis fruticosa	Sequences obtained from Genbank	AY763824	AY763892		
Eustoma exaltatum	Sequences obtained from Genbank	AJ490200	AJ489875		
Exaculum pusillum	Sequences obtained from Genbank	AJ490201	AJ489876		
Exacum affine	Sequences obtained from Genbank	AJ490211	AJ489886		
Exacum appendiculatum	Sequences obtained from Genbank	FJ014140	FJ666036		
Exacum bulbilliferum	Sequences obtained from Genbank	AJ490206	AJ489881		
Exacum exiguum	Sequences obtained from Genbank	AJ490209	AJ489884		
Exacum marojejyense	Sequences obtained from Genbank	AJ490218	AJ489893		
Exacum pedunculatum	Sequences obtained from Genbank	AJ490224	AJ489899		
Exacum sutaepense	Sequences obtained from Genbank	AJ490231	AJ489906		
Exacum trinervium	Sequences obtained from Genbank	AJ490234	AJ489909		
Exacum walkeri	Sequences obtained from Genbank	AJ490237	AJ489912		
Exacum wightianum	Sequences obtained from Genbank	AJ490238	AJ489913		
Exochaenium baumianum	Sequences obtained from Genbank	FJ014148	FJ666023		
Exochaenium grande	Sequences obtained from Genbank	FJ014151	FJ666026		
Exochaenium platypterum	Sequences obtained from Genbank	FJ014157	FJ666030		
Exochaenium teucszii	Sequences obtained from Genbank	FJ014162	FJ666021		
Fagraea elliptica	Sequences obtained from Genbank	AF102420	FJ232579		
Fagraea gracilipes	Sequences obtained from Genbank		FJ232576		
Fagraea racemosa	Sequences obtained from Genbank		FJ232577		
Fagraea salticola	Sequences obtained from Genbank		FJ232571		
Frasera albicaulis	Sequences obtained from Genbank			AJ294587	AJ294647
Frasera parryi	Sequences obtained from Genbank			AJ306083	AJ306111
Frasera speciosa	Sequences obtained from Genbank			Z48146	Z48124
Gelsemium sempervirens	Sequences obtained from Genbank	AF102428	DQ358881		

Gentiana asclepiadea	Sequences obtained from Genbank	AJ580515	AJ580549		
Gentiana atropurpurea	Sequences obtained from Genbank	AY858686	AY858678		
Gentiana crassicaulis	Sequences obtained from Genbank	AY858684	AY858676		
Gentiana dahurica	Sequences obtained from Genbank	DQ398700	DQ398632		
Gentiana decora	Sequences obtained from Genbank	EU834122	EU812468		
Gentiana decumbens	Sequences obtained from Genbank	DQ398714	DQ398655		
Gentiana macrophylla	Sequences obtained from Genbank	DQ398724	DQ398653		
Gentiana oreodoxa	Sequences obtained from Genbank	DQ398728	DQ398657		
Gentiana scabra	Sequences obtained from Genbank	GQ864088	GQ864015		
Gentiana walujewii	Sequences obtained from Genbank	DQ398717	DQ398646		
Gentianella amarella	Sequences obtained from Genbank	AJ580540	AJ580573		
Gentianella anisodonta	Sequences obtained from Genbank	AJ580524	AJ580558		
Gentianella astonii	Sequences obtained from Genbank		AY136494		
Gentianella azurea	Sequences obtained from Genbank		AF346014		
Gentianella bohemica	Sequences obtained from Genbank	AJ580537	AJ580570		
Gentianella campestris	Sequences obtained from Genbank	AJ580523	AJ580557		
Gentianella cerina	Sequences obtained from Genbank		AY136502		
Gentianella corymbifera	Sequences obtained from Genbank		AY136492		
Gentianella diemensis	Sequences obtained from Genbank		AY136504		
Gentianella engadinensis	Sequences obtained from Genbank	AJ580526	AJ580559		
Gentianella grisebachii	Sequences obtained from Genbank		AY136489		
Gentianella lineate	Sequences obtained from Genbank		AY136503		
Gentianella magellanica	Sequences obtained from Genbank		AY160219		
Gentianella narcissoides	Sequences obtained from Genbank		AY136486		
Gentianella peruviana	Sequences obtained from Genbank	AF102432		AJ294618	AJ294678
Gentianella quinquefolia	Sequences obtained from Genbank		EU812469		
Gentianella saxosa	Sequences obtained from Genbank		AY136499		
Gentianella spenceri	Sequences obtained from Genbank		AY136496		
Gentianella styriaca	Sequences obtained from Genbank	AJ580544	AJ580577		
Gentianella umbellata	Sequences obtained from Genbank	AJ580519	AJ580553		
Gentianodes tianschanica	Sequences obtained from Genbank	DQ398721	DQ398650		
Gentianopsis barbata	Sequences obtained from Genbank		AF346007		
Gentianopsis ciliata	Sequences obtained from Genbank	AJ580517	AJ580551		
Gentianothamnus madagascariensis	Sequences obtained from Genbank	AJ490240	AJ489914		
Gyrandra brachycalyx	Sequences obtained from Genbank	AF402241		AY047771	AY047855
Gyrandra tenuifolia	Sequences obtained from Genbank	AF402242		AY047772	AY047857
Halenia aquilegiella	Sequences obtained from Genbank			AJ411740	AJ411709
Halenia brevicornis	Sequences obtained from Genbank			AJ318540	AJ410319
Halenia corniculata	Sequences obtained from Genbank			AJ306087	AJ306114
Halenia decumbens	Sequences obtained from Genbank			AJ411749	AJ411718
Halenia deflexa	Sequences obtained from Genbank			AJ411750	AJ411719
Halenia elliptica	Sequences obtained from Genbank		AF346012		
Halenia pulchella	Sequences obtained from Genbank			AJ411760	AJ411729
Halenia schiedeana	Sequences obtained from Genbank			AJ411764	AJ411733
Halenia weddelliana	Sequences obtained from Genbank			AJ318542	AJ410321
Helia oblongifolia	Sequences obtained from Genbank		EU709794		
Irlbachia poeppigii	Sequences obtained from Genbank	AF102441	EU709796		

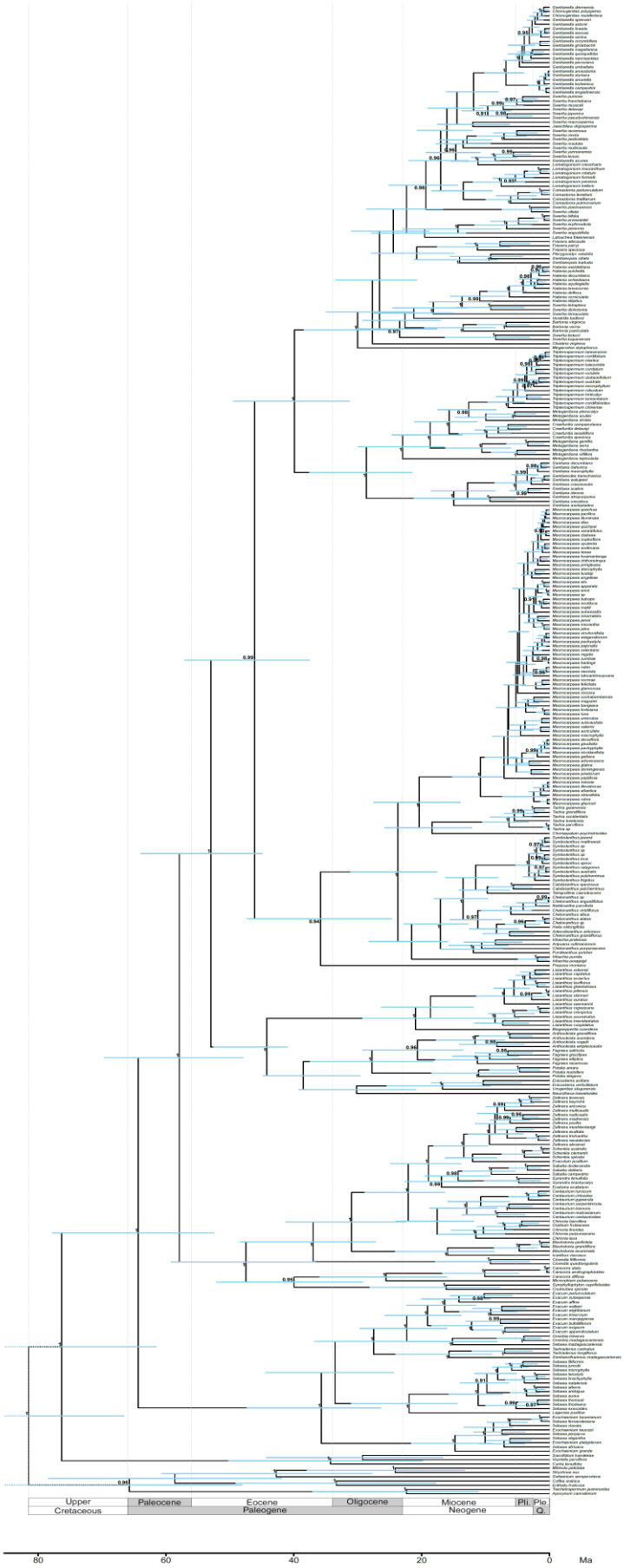
Irlbachia pratensis	Sequences obtained from Genbank		EU709797		
Irlbachia pumila	Sequences obtained from Genbank		EU709798		
Ixanthus viscosus	Sequences obtained from Genbank	AY251741		AY251683	AY251713
Jaeschkea oligosperma	Sequences obtained from Genbank	AF102444		AJ294633	AJ294693
Lagenias pusillus	Sequences obtained from Genbank	FJ014181	FJ666015		
Latouchea fokienensis	Sequences obtained from Genbank			AJ318544	AJ410323
Lisianthius acuminatus	Sequences obtained from Genbank		FJ232566		
Lisianthius adamsii	Sequences obtained from Genbank		FJ232547		
Lisianthius auratus	Sequences obtained from Genbank		FJ232554		
Lisianthius brevidentatus	Sequences obtained from Genbank		FJ232568		
Lisianthius capitatus	Sequences obtained from Genbank		FJ232549		
Lisianthius cuspidatus	Sequences obtained from Genbank		FJ232567		
Lisianthius exsertus	Sequences obtained from Genbank		FJ232550		
Lisianthius glandulosus	Sequences obtained from Genbank		FJ232551		
Lisianthius jefensis	Sequences obtained from Genbank	AF102448	EU709782		
Lisianthius laxiflorus	Sequences obtained from Genbank	AF102449	FJ232552		
Lisianthius nigrescens	Sequences obtained from Genbank		FJ232563		
Lisianthius oreopolus	Sequences obtained from Genbank		FJ232565		
Lisianthius seemannii	Sequences obtained from Genbank		FJ232555		
Lisianthius skinneri	Sequences obtained from Genbank		FJ232553		
Lomatogonium bellum	Sequences obtained from Genbank			AJ318545	AJ410324
Lomatogonium forrestii	Sequences obtained from Genbank			AJ306091	AJ306118
Lomatogonium macranthum	Sequences obtained from Genbank		AF346011		
Lomatogonium oreocharis	Sequences obtained from Genbank	AF102452		AJ294635	AJ294695
Lomatogonium perenne	Sequences obtained from Genbank			AJ318547	AJ410326
Lomatogonium rotatum	Sequences obtained from Genbank		AF346010		
Macrocarpaea angelliae	J.R. Grant & J. Vieu JRG4664 (NY)	XXXX	XXXX		
Macrocarpaea apparata	J.R. Grant 4002 (NY)	EU528058	DQ401413		
Macrocarpaea arborescens	J.R. Grant & J. Vieu JRG4659 (NY)	XXXX	XXXX		
Macrocarpaea atlantica	W. Thomas 14304 (NY)	EU528057	EU528125		
Macrocarpaea auriculata	B. Haber 2455 (MO)	XXXX			
Macrocarpaea bangiana	S.G. Beck 8699 (NEU)	XXXX	EU528079		
Macrocarpaea bubops	J.R. Grant & J. Vieu JRG4681 (NY)	XXXX	XXXX		
Macrocarpaea chthonotropa	J. Vieu & D. Desrousseaux JV47 (NEU)	XXXX	XXXX		
Macrocarpaea cinchonifolia	S.G. Beck 24780 (NY)		EU528084		
Macrocarpaea claireae	J.R. Grant & J. Vieu JRG4666bis (NY)	XXXX	XXXX		
Macrocarpaea cochabambensis	A. Gentry 44200 (NY)	EU528047	EU528085		
Macrocarpaea densiflora	K. von Sneidern 2523 (S)		EU528087		
Macrocarpaea dies-viridis	J.R. Grant & J. Vieu JRG4679 (NY)	XXXX	XXXX		
Macrocarpaea domingensis	P. Maas 8395 (NY)	XXXX	EU528092		
Macrocarpaea elix	G. Harling & Andersson 23442 (MO)		EU528094		
Macrocarpaea ericii	J. Vieu & D. Desrousseaux JV39 (NEU)	XXXX	XXXX		
Macrocarpaea felicitata	J. Vieu & D. Desrousseaux JV11x (NEU)	XXXX	XXXX		
Macrocarpaea fortisiana	D. McCarroll 128 (NY)	XXXX	EU528095		
Macrocarpaea gattaca	J. Jaramillo 22 (NY)	XXXX	EU528096		
Macrocarpaea gaudialis	R.E. Weaver 2650 (GH)	XXXX	EU528097		
Macrocarpaea glabra	J.R. Grant 4310 (NY)	EU528060	EU528098		

Macrocarpaea glamorosa	J.R. Grant & J. Vieu JRG4668 (NY)	XXXX	XXXX
Macrocarpaea glaziovii	Grant, J.R., L. Zeltner, B. Resende Silva & V. Trunz 4603 (NY)	XXXX	
Macrocarpaea harlingii	J.R. Grant & J. Vieu JRG4682 (NY)	XXXX	XXXX
Macrocarpaea huamantanga	J. Vieu & D. Desrousseaux JV44 (NEU)	XXXX	XXXX
Macrocarpaea illecebrosa	J.R. Grant & V. Trunz 4645 (NY)	XXXX	XXXX
Macrocarpaea illuminata	J.R. Grant & J. Vieu JRG4693 (NY)	XXXX	XXXX
Macrocarpaea inesiae	J.R. Grant, L. Zeltner, F. Calió & V. Trunz 4602 (NY)	XXXX	
Macrocarpaea innarrabilis	J. Vieu & D. Desrousseaux JV29 (NEU)	XXXX	XXXX
Macrocarpaea jactans	J.R. Grant & J. Vieu JRG4694 (NY)	XXXX	XXXX
Macrocarpaea jalca	J. Vieu & D. Desrousseaux JV34 (NEU)	XXXX	XXXX
Macrocarpaea jensii	J.R. Grant & J. Vieu JRG4683 (NY)	XXXX	XXXX
Macrocarpaea kuelap	J. Vieu & D. Desrousseaux JV21 (NEU)	XXXX	XXXX
Macrocarpaea lenae	J.R. Grant 4013 (NY)	XXXX	EU528110
Macrocarpaea luna-gentiana	J.R. Grant 4028 (NY)	EU528050	EU528111
Macrocarpaea macrophylla	J.R. Grant 4312 (NY)	EU528061	EU528113
Macrocarpaea maguirei	B. Maguire 61569 (NY)	XXXX	EU528114
Macrocarpaea mattii	F. Matt 14 (ER)		EU528123
Macrocarpaea micrantha	J.R. Grant & J. Vieu JRG4674 (NY)	XXXX	XXXX
Macrocarpaea nicotianifolia	A.S. Barclay 3402 (US)	XXXX	EU528118
Macrocarpaea noctiluca	J. Vieu & D. Desrousseaux JV45 (NEU)	XXXX	XXXX
Macrocarpaea normae	C. Vargas 21415 (CUZ)	XXXX	
Macrocarpaea obtusifolia	J.R. Grant, L. Zeltner, B. Resende Silva & V. Trunz 4605 (NY)	XXXX	XXXX
Macrocarpaea opulenta	J.R. Grant & J. Vieu JRG4676 (NY)	XXXX	XXXX
Macrocarpaea ostentans	B. Wallnöfer 12968 (U)	XXXX	EU528128
Macrocarpaea pachyphylla	M.L. Bristol 1429 (GH)		EU528129
Macrocarpaea pachystyla	J. Schunke-Vigo 5298 (NY)		EU528130
Macrocarpaea pacifica	J.R. Grant & J. Vieu JRG4704 (NY)	XXXX	XXXX
Macrocarpaea pajonalis	J. Vieu & D. Desrousseaux JV05b (NEU)	XXXX	XXXX
Macrocarpaea papillosa	R.E. Weaver 2629 (GH)	XXXX	EU528131
Macrocarpaea pinetorum	J. Bisse 49708 (HAJB)	XXXX	EU528133
Macrocarpaea pringleana	F. Luisier 1 (NY)	XXXX	EU528134
Macrocarpaea quechua	J. Vieu & D. Desrousseaux JV17 (NEU)	XXXX	XXXX
Macrocarpaea quizhpei	J.R. Grant & J. Vieu JRG4689 (NY)	XXXX	XXXX
Macrocarpaea revoluta	J. Vieu & D. Desrousseaux JV13 (NEU)	XXXX	XXXX
Macrocarpaea robin-fosteri	J. Vieu & D. Desrousseaux JV05 (NEU)	XXXX	XXXX
Macrocarpaea rubra	J.R. Grant, L. Zeltner & V. Trunz 4593 (NY)	XXXX	XXXX
Macrocarpaea sodiroana	P. Maas 6528 (U)		EU528141
Macrocarpaea sp.	J. Vieu & D. Desrousseaux JV38 (NEU)	XXXX	XXXX
Macrocarpaea stenophylla	J. Vieu & D. Desrousseaux JV22 (NEU)	XXXX	XXXX
Macrocarpaea subcaudata	R.L. Wilbur & Almeda 16828 (DUKE)	XXXX	EU528143
Macrocarpaea subsessilis	J.R. Grant & J. Vieu JRG4661 (NY)	XXXX	XXXX
Macrocarpaea tahuantinsuyuana	F. Woytkowski 6672A (MO)	XXXX	
Macrocarpaea umerulus	H. Mendoza 3142 (FMB)	XXXX	
Macrocarpaea valerioi	R.L. Wilbur 20350 (NY)	XXXX	EU528147
Macrocarpaea viscosa	M. Weigend 5429 (NY)	EU528054	EU528149
Macrocarpaea weigendiorum	J. Vieu & D. Desrousseaux JV16 (NEU)	XXXX	XXXX

Macrocarpaea xerantifulva	J.R. Grant & J. Vieu JRG4667 (NY)	XXXX	XXXX		
Macrocarpaea zophoflora	J.J. Wurdack 1618 (NY)	XXXX			
Macrocarpaea zumba	J.R. Grant & J. Vieu JRG4669 (NY)	XXXX	XXXX		
Megacodon stylophorus	Sequences obtained from Genbank	AY858687	AY858679		
Metagentiana gentilis	Sequences obtained from Genbank	AY563386	AY562177		
Metagentiana leptoclada	Sequences obtained from Genbank	AY858681	AY858674		
Metagentiana pterocalyx	Sequences obtained from Genbank	AY563384	AY562171		
Metagentiana rhodantha	Sequences obtained from Genbank	GQ864096	GQ864023		
Metagentiana serra	Sequences obtained from Genbank	AY563387	AY562175		
Metagentiana souliei	Sequences obtained from Genbank	AY563388	AY562170		
Metagentiana striata	Sequences obtained from Genbank	AY563389	AY562173		
Metagentiana villifera	Sequences obtained from Genbank	AY858680	AY858672		
Microrphium pubescens	Sequences obtained from Genbank	AJ490241	AJ489916		
Mitreola petiolata	Sequences obtained from Genbank	AF102460	AF054635		
Neblinantha parvifolia	Sequences obtained from Genbank	AF102461			
Neurotheca loeselioides	Sequences obtained from Genbank	AF102463	FJ232570		
Obolaria virginica	Sequences obtained from Genbank			AJ306094	AJ306121
Ornichia madagascariensis	Sequences obtained from Genbank	AJ490242	AJ489917		
Ornichia trinervis	Sequences obtained from Genbank	AJ490243	AJ489918		
Orphium frutescens	Sequences obtained from Genbank	FJ014163	FJ666017		
Potalia amara	Sequences obtained from Genbank		DQ449919		
Potalia elegans	Sequences obtained from Genbank		DQ449920		
Potalia resinifera	Sequences obtained from Genbank	AF102471	DQ449922		
Prepusa montana	Sequences obtained from Genbank		EU709805		
Pterygocalyx volubilis	Sequences obtained from Genbank		GU251036		
Purdieanthus pulcher	Sequences obtained from Genbank		EU709800		
Sabatia campestris	Sequences obtained from Genbank	AY255692		AY256382	AY256387
Sabatia dodecandra	Sequences obtained from Genbank	AY255693		AY256383	AY256388
Sabatia stellaris	Sequences obtained from Genbank	AY255694		AY256384	AY256389
Saccifolium bandeirae	Sequences obtained from Genbank			AJ242611	AJ242612
Schenkia australis	Sequences obtained from Genbank	AY315684		AY251679	AY251709
Schenkia clementii	Sequences obtained from Genbank	AY251739		AY251680	AY251710
Schenkia spicata	Sequences obtained from Genbank	AF402253		AY047791	AY047877
Sebaea africana	Sequences obtained from Genbank	FJ014147	FJ666025		
Sebaea albens	Sequences obtained from Genbank	FJ014164	FJ666003		
Sebaea ambigua	Sequences obtained from Genbank	FJ014165	FJ666002		
Sebaea aurea	Sequences obtained from Genbank	FJ014166	FJ666004		
Sebaea brachyphylla	Sequences obtained from Genbank	AJ490245	AJ489920		
Sebaea clavata	Sequences obtained from Genbank	FJ014149	FJ666018		
Sebaea exacoides	Sequences obtained from Genbank	FJ014170	FJ665999		
Sebaea fernandesiana	Sequences obtained from Genbank	FJ014150	FJ666027		
Sebaea filiformis	Sequences obtained from Genbank	FJ014171	FJ665991		
Sebaea junodii	Sequences obtained from Genbank	FJ014172	FJ665990		
Sebaea leiostyla	Sequences obtained from Genbank	FJ014173	FJ665994		
Sebaea madagascariensis	Sequences obtained from Genbank	AJ490247	AJ489921		
Sebaea microphylla	Sequences obtained from Genbank	FJ014177	FJ665987		
Sebaea natalensis	Sequences obtained from Genbank	FJ014180	FJ665993		

Sebaea oligantha	Sequences obtained from Genbank	FJ014155	FJ666024		
Sebaea perparva	Sequences obtained from Genbank	FJ014156	FJ666031		
Sebaea thodeana	Sequences obtained from Genbank	FJ014189	FJ666012		
Sebaea thomasii	Sequences obtained from Genbank	FJ014190	FJ666014		
Strychnos nux	Sequences obtained from Genbank	AF102485	JF938015		
Swertia angustifolia	Sequences obtained from Genbank		FJ010793		
Swertia bifolia	Sequences obtained from Genbank		DQ317490		
Swertia bimaculata	Sequences obtained from Genbank		JF978820		
Swertia ciliata	Sequences obtained from Genbank		FJ010798		
Swertia cincta	Sequences obtained from Genbank		JF978823		
Swertia delavayi	Sequences obtained from Genbank		GQ848486		
Swertia dichotoma	Sequences obtained from Genbank		DQ317488		
Swertia erythrosticta	Sequences obtained from Genbank		AF251122		
Swertia franchetiana	Sequences obtained from Genbank		AF255916		
Swertia japonica	Sequences obtained from Genbank		GQ848488		
Swertia leducii	Sequences obtained from Genbank		DQ317486		
Swertia luquanensis	Sequences obtained from Genbank		DQ317487		
Swertia macrosperma	Sequences obtained from Genbank		FJ010804		
Swertia multicaulis	Sequences obtained from Genbank		FJ010799		
Swertia mussotii	Sequences obtained from Genbank		AF255915		
Swertia pedicellata	Sequences obtained from Genbank		FJ010800		
Swertia perennis	Sequences obtained from Genbank		AJ580550		
Swertia pianmaensis	Sequences obtained from Genbank		DQ317485		
Swertia przewalskii	Sequences obtained from Genbank		AF255913		
Swertia pseudochinensis	Sequences obtained from Genbank		GQ848487		
Swertia punicea	Sequences obtained from Genbank		GQ848485		
Swertia racemosa	Sequences obtained from Genbank		FJ010807		
Swertia rosulata	Sequences obtained from Genbank	AJ490248	AJ489922		
Swertia tenuis	Sequences obtained from Genbank		DQ317489		
Swertia tetraptera	Sequences obtained from Genbank		JF978834		
Swertia yunnanensis	Sequences obtained from Genbank		JF978836		
Symbolanthus australis	Sequences obtained from Genbank	AF102489	EU709801		
Symbolanthus calygonus	Sequences obtained from Genbank		FJ232585		
Symbolanthus frigidus	Sequences obtained from Genbank	AF102498	EU709802		
Symbolanthus inca	J.R. Grant 4618 (NY)		XXXX		
Symbolanthus jasonii	J.R. Grant 4350 (NY)	EU528055	EU528151		
Symbolanthus JV23_1	J. Vieu & D. Desrousseaux JV20 (NEU)	XXXX	XXXX		
Symbolanthus JV40_1	J. Vieu & D. Desrousseaux JV40 (NEU)	XXXX	XXXX		
Symbolanthus JV50_1	J. Vieu & D. Desrousseaux JV46 (NEU)	XXXX	XXXX		
Symbolanthus matthewsii	J.R. Grant 4344 (NY)		EU528152		
Symbolanthus pulcherrimus	Sequences obtained from Genbank	AF102488	EU709803		
Symbolanthus spnov	J.R. Grant 4634 (NY)		XXXX		
Symphyllophyton caprifolioides	Sequences obtained from Genbank			AJ011462	AJ011472
Tachia grandiflora	Sequences obtained from Genbank		DQ401415		
Tachia guianensis	Sequences obtained from Genbank		DQ401419		
Tachia loretensis	Sequences obtained from Genbank		DQ401421		
Tachia occidentalis	Sequences obtained from Genbank		DQ401423		

Tachia parviflora	J.L Clark 11500	XXXX	XXXX		
Tachia sp	no voucher available	XXXX	XXXX		
Tachiadenus carinatus	Sequences obtained from Genbank	AJ490249	AJ489923		
Tachiadenus longiflorus	Sequences obtained from Genbank	AJ490250	AJ489924		
Tetrapollinia caerulescens	Sequences obtained from Genbank	AF102494	EU709804		
Trachelospermum jasminoides	Sequences obtained from Genbank	AB710087	AB610208		
Tripterospermum alutaceifolium	Sequences obtained from Genbank		GU251037		
Tripterospermum australe	Sequences obtained from Genbank		GU251038		
Tripterospermum chinense	Sequences obtained from Genbank		AY858668		
Tripterospermum cordatum	Sequences obtained from Genbank		AY562172		
Tripterospermum cordifolioides	Sequences obtained from Genbank		GU251041		
Tripterospermum cordifolium	Sequences obtained from Genbank		GU251042		
Tripterospermum hirticalyx	Sequences obtained from Genbank		GU251043		
Tripterospermum lanceolatum	Sequences obtained from Genbank		GU251044		
Tripterospermum luteoviride	Sequences obtained from Genbank		GU251045		
Tripterospermum microphyllum	Sequences obtained from Genbank		GU251047		
Tripterospermum nienkui	Sequences obtained from Genbank		GU251048		
Tripterospermum robustum	Sequences obtained from Genbank		GU251052		
Tripterospermum taiwanense	Sequences obtained from Genbank		GU251053		
Tripterospermum volubile	Sequences obtained from Genbank		AY858667		
Urogentias ulugurensis	Sequences obtained from Genbank	AF102495	FJ232583		
Veratrilla baillonii	Sequences obtained from Genbank		AF251123		
Voyriella parviflora	Sequences obtained from Genbank			AJ242615	AJ242616
Zeltnera abramsii	Sequences obtained from Genbank	AF402201		AY047712	AY047797
Zeltnera arizonica	Sequences obtained from Genbank	AY318874		AY047725	AY047809
Zeltnera beyrichii	Sequences obtained from Genbank	AF402216		AY047733	AY047817
Zeltnera exaltata	Sequences obtained from Genbank	AF402210		AY047723	AY047807
Zeltnera madrensis	Sequences obtained from Genbank	AF402222		AY047742	AY047827
Zeltnera muehlenbergii	Sequences obtained from Genbank	AY315687		AY315668	AY047869
Zeltnera multicaulis	Sequences obtained from Genbank	AF402213		AY047727	AY047813
Zeltnera nevadensis	Sequences obtained from Genbank	AF402206		AY047719	AY047803
Zeltnera nudicaulis	Sequences obtained from Genbank	AF402229		AY047745	AY047830
Zeltnera pusilla	Sequences obtained from Genbank	AF402230		AY047755	AY047841
Zeltnera texensis	Sequences obtained from Genbank	AF402217		AY047735	AY047819
Zeltnera trichantha	Sequences obtained from Genbank	AF402199		AY047709	AY047795



Appendix S3: Chronogram of the Gentianceae with 95% highest posterior density bars, based on BEAST analyses of a two gene dataset (ITS & *Tmt*). See main text for details concerning calibration. Posterior probabilities (PP) are displayed for nodes with PP>0.9. See Apendix S2 for a list of taxa included.

CHAPTER 2: The evolution of trait disparity during the radiation of *Macrocarpaea* (Gentianaceae) in the tropical Andes

AUTHORS NAMES: Vieu Julien C. & Grant Jason R.

AUTHORS AFFILIATIONS: Institut de Biologie, Université de Neuchâtel, Rue Emile Argand 11, 2000 Neuchâtel, Switzerland;

ABSTRACT

The evolutionary processes responsible for the extraordinary plant diversity in the middle elevation montane forests of the tropical Andes (MMF) remain poorly understood. Here, we use the plant genus *Macrocarpaea* (Gentianaceae) as an example to investigate the contribution of adaptive divergence and niche conservatism to the radiation of plant lineages in the MMF. We use morphological and climatic data in comparative framework, to compare a set of evolutionary scenarios of various levels of complexity. We show that the hypothesis of an *adaptive radiation* for *Macrocarpaea* in the MMF is very unlikely. The genus has remained confined to the upper montane forests (UMF>1800m) during more than the half of its history possibly due to evolutionary constraints. Later, coinciding with the beginning of the Pleistocene, a derived clade, the *micrantha* alliance, successfully colonized and radiated in the lower montane forests (LMF<1800m). This colonisation was accompanied by the evolution of a new leaf phenotype unique to the species of the *micrantha* alliance that likely represents an adaptation to life in this new adaptive zone. Therefore, our results suggests that *niche conservatism* and *geographical processes* have dominated most of the diversification history of *Macrocarpaea*, but that a rare adaptive divergence event allowed a transition into a new adaptive zone and enabled a progressive radiation in this zone through geographical processes.

KEYWORDS

Adaptive zone, evolutionary constraint, phylogeny, phylogenetic comparative methods

INTRODUCTION

Understanding the evolutionary mechanisms responsible for the species richness of some lineages compared with others is a fundamental and active domain of research in evolutionary biology. The ecological theory of adaptive radiation predicts that ecological opportunity is a primary factor regulating the tempo of lineage diversification (Mahler *et al.* 2010). Ecological opportunity can be perceived as the availability of under-exploited resources that can arise from the extinction of competitors, dispersal into a new habitats, or the acquisition of a key innovation that allow organisms to explore new ecological niches (Losos, 2010, Yoder *et al.*, 2010). During the phase of niche exploration, divergent selection is thought to promote bursts of species divergence (diversification) together with rapid accumulation of phenotypic diversity (disparification, Losos & Mahler 2010). As niche space becomes progressively saturated, the rates of diversification and of disparification are expected to slow down (Gavrilets & Losos 2009). Classical examples of adaptive radiations often involve organisms located on island or island-like habitats (i.e.: lake, mountain tops) such as Galapagos Darwin's finches on the Galapagos island (Grant & Grant 2006), the Caribbean *Anolis* lizards (Losos et al. 1998), and the African rift lake cichlid fishes (Seehausen 2006), but more recently, cases of adaptive radiations on continents have also been documented (Burbrink & Pyron 2010; Slater *et al.*, 2010; Schweizer *et al.*, 2014).

Adaptive radiations have been proposed to be responsible for much of the biodiversity on Earth (Schluter, 2000; Benkman, 2003). However, the ubiquity of the link between diversification and disparification rates predicted by the adaptive radiation theory has been recently challenged (Adams *et al.*, 2009; Harmon *et al.*, 2010). Several studies have revealed that bursts of diversification are not necessarily associated with bursts of morphological disparity (Rowe *et al.*, 2011; Burbrink *et al.*, 2012; Hipslay *et al.*, 2014; Rabosky *et al.*, 2014). This suggests that adaptive divergence is not the only evolutionary process that can lead to episodes of rapid diversification in a clade. Indeed, geographical processes arising from the interactions between habitat spatial fragmentation, lineage ecological niche stasis (niche conservatism), and low dispersal ability also have the potential to produce bursts of diversification by inducing repeated allopatric speciation (Wiens 2004; Rundell & Price 2009). Under this type of non-adaptive radiation, diversification and disparification rates are expected to be decoupled from one another (Gittenberger, 1991) such that rapid species diversification is not accompanied by rapid trait disparification. As noted by Wiens (2004), natural selection is not absent from this process. Rather, stabilizing selection constrains the evolution of the ecological niche over time, and thus contributes to prevent secondary contact between

allopatric populations separated by unfavourable habitats that can ultimately lead to speciation. To date, non-adaptive radiations have perhaps received less interest than adaptive ones but compelling examples have been documented in North American *Plethodon* salamanders (Kozak et al., 2006) or Aegean *Nigella* (Comes *et al.*, 2008; but see Rundell & Price 2009). Despite recent advances, how the relative contribution of adaptive (ecological based divergent selection) and non-adaptive processes (stabilizing selection and drift) to shaping biodiversity poorly understood.

The middle elevation mountain forests of the tropical Andes (MMF, 1000-3500m, including: lower and upper montane forests and sub-paramo shrub forests) offer an excellent study system to address these questions. First, they harbour tremendous amount biodiversity including many diverse lineages rich in endemics (Orme *et al.*, 2005) suggesting that radiations have been common in the MMF (Vieu *et al.*, in review). Secondly, the complex physiography of the Andes with deep valleys and high peaks represents a large area with steep and extended ecological gradients associated with elevation as well as physical barriers to gene flow. Ecological gradients are thought to promote adaptive divergence (Doebeli & Dieckmann, 2003) whereas physical barriers to gene flow can promote geographic isolation and may lead to allopatric speciation (Coyne 1992). Third, the MMF apparently evolved relatively recently, likely during the late Miocene as the Andes reached sufficient elevation to trigger high year-round precipitation on the eastern slopes(van der Hammen & Hooghiemstra, 2000; Hoorn *et al.*, 2010; Insel *et al.*, 2012). These vast and recently emerged habitats might have represented a considerable new ecological opportunities for the regional biodiversity pool to colonize and to adapt to. Finally, the recent history of the MMF has been marked by the Pleistocene climatic oscillations (PCO, 2.6Ma-15.000Ya) that caused dramatic fluctuations of the altitudinal range and cycles of contraction and re-expansion of the main vegetation belts (Hooghiemstra *et al.*, 2006; Flantua *et al.*, 2014). As a consequence, species inhabiting the MMF are thought to have alternated between phases of population fragmentation during the glacial periods and re-expansion during interglacials (Rull, 2011). Thus the PCO potentially induced ideal conditions to promote extensive allopatric speciation at the scale of the whole MMF territory.

Several studies focusing on a variety of animal groups have used the MMF study system to study the evolutionary processes responsible of species divergence and to broader evolutionary diversification. While examples of adaptive divergence have been documented (butterflies: Elias *et al.*, 2009), most studies point to a preponderance of niche conservatism and non-adaptive geographical processes (allopatric speciation) as the

main driver of species divergence in the MMF (birds: Gutiérrez-Pinto *et al.*, 2012; Benham *et al.*, 2014; butterflies: Casner & Pyrcz, 2010, frogs: Hutter *et al.*, 2014, mice: Patton & Smith 1992). By comparison, the processes and the mechanisms underlying involved in plant diversification in the MMF have been neglected. Recently, several studies have investigated plant diversification dynamic in the MMF and have suggested that colonization and/or range expansion across emerging MMF could have represented a key opportunity for these plant group to diversify through rapid allopatric speciation, thus privileging nonadaptive over adaptive processes (Antonelli & Sanmartín, 2011; Givnish *et al.*, 2014, Spriggs *et al.*, 2015; Vieu *et al.*, in review). While there is compelling evidence for accelerated species diversification rates in these clades, none of these studies estimated rates of ecological and/or morphological disparification through time to investigate a potential concomitant role for adaptive divergence during diversification of these clades.

In this study, we explore the potential contributions of adaptive divergence and niche conservatism to plant diversification in the MMF using the genus *Macrocarpaea* (Gentianaceae). *Macrocarpaea* has 118 species exclusively found in mountainous regions of the Neotropics, 96 of which are endemic to the Andean MMF. A recent study showed that the clade including all *Macrocarpaea* species from the MMF diversified following the classical diversification trajectory expected under a radiation (early burst followed by a slowdown) and suggested that rapid allopatric speciation during rapid range expansion of the genus across the MMF facilitated the radiation (Vieu *et al.*, in prep). *Macrocarpaea* displays substantial morphological variation, with species ranging from small shrubs hardly taller than 1m, with small coriaceous leaves to 6m tall trees with very large leaves. Although, the genus spans a broad elevational range (800-3500m), most species are confined within 500m elevation bands. If these trait (morphology and ecology) differences evolved synchronously with the burst of species diversification that would indicate that the radiation of *Macrocarpaea* in the MMF has an important adaptive component.

Alternatively, it is possible that traits differences accumulated later. For example, niche evolution could have been constrained or slow in the early history of the MMF clade but during the PCO with repeatedly changing climates species might have adapted to different elevation bands instead of migrating, hereby enhancing trait evolution toward the tip of the tree.

It is notable that the large majority of *Macrocarpaea* MMF species occur in the upper montane forest (UMF>1800m, following Leuschner & Moser 2008), but a derived clade, the *micrantha* alliance, has most of its species (19 out of 25) in the lower montane forest (LMF<1800m). This could indicate that the ancestor of *micrantha* alliance invaded a new adaptive zone *sensu* Simpson (1953) relatively depauperate from others *Macrocarpaea* lineages (the LMF). In consequence trait evolutions in the *micrantha* alliance might have been decoupled from the rest of the MMF clade. Alternatively, trajectories of trait evolution could have been elevation belt-specific and all the lineages occurring in the LMF including the few species outside of the *micrantha* alliance might have follow the same evolutionary trajectories.

Finally, it is also possible that the pace and trajectories of trait evolution have been influenced by region-specific conditions. The Amotape-Huancabamba zone (AHZ), a region located in southern Ecuador and northern Peru (between 3° and 8°S), is the centre of diversity for many Andean plant lineages in the MMF, including *Macrocarpaea* (Weigend, 2002; Mutke *et al.*, 2014). It has been proposed that the exceptional habitat heterogeneity of the AHZ enhanced adaptive divergence more than in any other region of the MMF (Weigend, 2002; Mutke *et al.*, 2014). According to this hypothesis, *Macrocarpaea* lineages from the AHZ should display faster trait evolution than lineages from elsewhere.

Here, we use a previously generated time-calibrated phylogeny for the genus *Macrocarpaea* and compile morphological and climatic datasets to test the following hypotheses in a phylogenetic comparative method (PCM) framework: 1) assess the contribution of adaptive divergence during the radiation of *Macrocarpaea* in the MMF (if adaptive divergence fuelled the radiation, we expect to detect a burst of disparification in the early history of the MMF clade); 2) test whether traits evolution for *Macrocarpaea* in the MMF followed a constrained or an unconstrained mode of evolution; and 3) evaluate whether the constraint or the rate of evolution was constant in time and uniform across all species, or varied according to: (i) climate dynamism history, (ii) regional or (iii) elevation band-specific condition, or (iv) clade-specific history.

MATERIALS AND METHODS

Time-calibrated phylogeny.

We used a previously generated time-calibrated phylogeny of *Macrocarpaea* that included 76 of the 118 currently recognized species in the genus (Vieu *et al.*, in review). The maximum clade credibility tree (MMC tree) and a random sample of 100 trees from the posterior distribution of the BEAST analyses were pruned to remove all non-Andean species. After pruning, the phylogeny included 63 Andean species representing 64% of the MMF diversity of the genus (63/98).

Morphological data

Morphological data for 55 Andean *Macrocarpaea* species were obtained from measurements of herbarium specimens that were performed as part of species descriptive taxonomic research. Measurements were done by one investigator (Jason Grant) based on all the available herbarium specimens (2154 sheets). The number of specimens examined varied greatly between species (min=2, max=215, median=28) reflecting a combination of the variation in sampling effort and the abundance of each species. Six traits were selected using following criteria: 1) reflecting the morphological variation displayed by the genus; 2) have data measured for a high proportion of species included in the phylogeny; 3) and have relatively straightforward plant functional interpretations. Plant height, leaf length and leaf width describe the vegetative components of the plants and their variation and likely reflect evolution driven by adaptation to the environment, while calyx width, corolla width at mouth and corolla length describe reproductive biological components and their variation potentially reflecting evolution driven by pollinators. The measurements used are estimates of ranges of size (minimum and maximum) for the different traits, from which we estimated the mean by assuming that it is equal to the centre of the range. We acknowledge a potential source of inaccuracy around this assumption. The overall proportion of missing data in the dataset is 6%. Data are missing for 9 species (55 included) and the species with missing data are *M angelliae*, *M luteynii*, *M catherineae*, *M chthonotropa*, *M gondoloides*, *M pacifica*, *M quechua*, *M arborescens* and *M micrantha* that all have two traits missing. Traits with missing data are corolla width at mouth (9 species), corolla length (5 species) and calyx width (4 species). In order to avoid removal of species with missing data, we imputed missing data using a method described by Penone *et al.* (2014) that takes into account phylogenetic relationship to improve the imputation process. Data were log transformed prior the imputation in order to reduce scale variations of the different variables. We used the PVRdecomp function from

the R software package PVR (Santos *et al.*, 2012) to decompose the phylogenetic distance matrices into a set of orthogonal eigenvectors. We kept the first ten eigenvectors (representing 64% of the variation in the phylogenetic distances among species) and appended them to the trait matrix that already contained 12 columns (minimum and maximum values for 6 traits). Imputation was done on this hybrid matrix in the R package missForest (Stekhoven & Buehlmann, 2012) using the options maxiter=20 and ntree=1000. MissForest stopped after 6 iterations and returned a relatively low error estimate (normalized root squared error, NRMSE= 0.042). After the imputation process, species means were computed from the log-scale minimum and maximum values for each trait. Phylogenetic correlations between plant size and the other morphological traits were tested by applying a correlation test on the phylogenetic independent contrasts (PIC). Among the five traits, four displayed a significant positive correlation with plant size ($r = 0.31\text{-}0.45$, $p = 0.0004\text{-}0.019$) while the correlation with calyx width was close to be significant ($r = 0.23$, $p = 0.094$). The five shape trait measurements were size-corrected using phylogenetic regression on plant size that account for the phylogenetic non-independence of species, within the R package "phytools" (Revell 2012). To do this, the phylogeny was pruned to include only species with morphological data. Residuals from the phylogenetic regression were extracted to conduct a phylogenetic principal components analysis (pPCA), which also accounted for the phylogenetic non-independence, on the covariance matrix (Revell, 2009). We retained the two first axes that collectively explained 88% of the variation and have relatively straightforward biological meaning. Leaf traits have very high negative loadings on the first pPCA axis which most likely describe variations in leaf size, while floral traits have relatively high negative loadings on the second pPCA axis which likely describes variations in floral shape. The Table1 reports the Traits loadings and the proportion of variance explained by the two first pPCA axes for the analysis based on the MCC tree are shown in Table 1. We note that the preparatory steps described above were repeated for each tree (100) in the posterior sample. All subsequent analyses relative to morphological evolution were conducted on: 1) log plant size (in log-mm), 2) the morphology pPCA1 (leaf size) and 3) the morphology pPCA2 (flower size).

Ecological data

Occurrence data were compiled from geo-referenced herbarium specimens. After excluding duplicated localities the dataset contained a total of 464 occurrences for 62 species. Overall, the number of occurrences varied greatly between species (min=1, max=36, median=5). The occurrences were used to extract elevation and climatic data from the WorldClim dataset, which comprises the 19 BioClim variables that describes average and seasonal

variation in temperature and precipitation at a 30-s resolution (Hijmans *et al.*, 2005). The extraction process was performed using the R package RASTER (Hijmans *et al.*, 2014). Species means were computed for Latitude, Longitude, Elevation and the 19 BioClim variables. This dataset of 21 variables and the phylogeny pruned to include species with climatic data only (62 species) were used to conduct a pPCA on the correlation matrix, because sizes and units of the variables included were different. The two first axes of the pPCA explained 68% of the variation and were retained. Table2 reports the traits loadings and the proportion of variance explained by the two first pPCA axes for the analysis based on the MCC tree. The variables with strong loadings on the first axis are Altitude and BioClim variables representing average, minimum and maximum of temperatures and precipitation. The variables with strong loadings on the second axis are Latitude, Longitude and the BioClim variables representing seasonal variation in temperatures. Thus the pPCA axis 1 represents the climatic variation related to the elevation range of the species and most likely describes the altitudinal niche of species that potentially constraint their migration toward adjacent elevations zones. On the other hand, the pPCA axis 2 represent climatic variations related to the latitudinal position (as the Andes have a banana shape, absolute latitude and absolute longitude are strongly negatively correlated, r = -81) and likely describes the regional environmental niche of species which potentially constrained their migration toward adjacent regions. We acknowledge that climatic variables represent extrinsic factors that are not directly inherited by descent (Grandcolas *et al.*, 2011), nevertheless they indirectly reflect intrinsic physiological tolerances of species which are of evolutionary relevance. Again, all the preparatory steps were repeated for each tree (100) in the posterior sample. Subsequent analyses related to the climatic niche evolution of *Macrocarpaea* species were conducted on: 1) the climatic pPCA1 (Altitudinal niche), and 2) the climatic pPCA2 (Latitudinal niche).

Functional data

Among plant functional traits, specific leaf area (SLA, defined as fresh leaf area/ dry mass), which is associated with carbon capture, is commonly used as a proxy for contrasting plant physiological strategies (Wright *et al.*, 2004). Plant species with high SLA (thin leaves) tend to have higher photosynthetic rates and shorter leaf life spans and in general characterize fast-growing species in water and nutrient rich environments (Reich *et al.*, 1997). In contrast, plant species with low SLA (thick/dense leaves) tend to have higher nutrient retentions and better protection from desiccation and thus characterize slower growing species in nutrient and/or water poor environments. In general, SLA tends to decrease with increasing elevation (Scheepens & Stöcklin, 2010) particularly in the tropics, most likely as a consequence of lower temperatures and lower nutrient availability

rather than water deficiency (Leuschner & Moser, 2008). SLA was measured for 31 species from specimens collected in the field by a single investigator (Julien Vieu). A total of 774 individuals were measured for an average of 20 individuals per species (range: 3-87). Mean SLA was computed for each species from the log transformed individual data (log $mm^2.mg^{-1}$) and was used in subsequent analysis.

Disparification, evolutionary models and phylogenetic signal.

In order to investigate the evolution of trait and ecological disparification through time we computed the Morphological Disparity Index (MDI, Harmon *et al.*, 2003) for the six traits described above on the MCC tree with the R package GEIGER (Harmon *et al.*, 2008). The MDI statistic compares the observed trait disparity through time (DTT) to the median of a distribution of DTT simulated under a Brownian motion model of evolution (BM) on the phylogeny. Negative values of MDI indicate that disparity is principally distributed among subclades and is considered to be the hall mark of adaptive radiation (Harmon *et al.*, 2003). On the other hand positive values of MDI indicate that disparity is distributed within subclades. The evolutionary interpretation of this second pattern is less clear but would tend to indicate that each subclade explored a higher proportion of the niche/trait space than expected from a BM model. The statistical significance of deviations from the BM expectation was assessed using 2000 simulations for each trait. Plots of the observed and simulated disparities against relative time from the root of the tree for the Climate pPCA1 and pPCA2 are displayed in Figure 1.

In order to compare different scenarios for the evolution of traits and climatic niches of *Macrocarpaea* in the Andes we considered a set of 11 models with different levels of complexity: 1) a simple Brownian model (BM) model which describes trait evolution driven by neutral processes such as genetic drift (Butler & King, 2004). We expect this model to perform better if trait evolution was un-constrained and did not show any specific pattern related to time, elevation belts, taxonomy or geography; 2) an Ornstein–Uhlenbeck model with a single optimum (OU) which describes trait evolution driven by stabilizing selection toward a single adaptive zone (Hansen 1997, Butler & King, 2004). A preference for this model would indicate that evolutionary constraints homogeneous across the whole phylogeny have governed traits evolution; 3) an Early burst model (EB) which corresponds to a niche-filling process during an adaptive radiation (Harmon *et al.*, 2010). If the radiation of *Macrocarpaea* in the Andes was an adaptive one, we expect this model to perform better; 4) a two-rate BM model (SHIFT) with rate of evolution allowed to change after a fixed point in time set to be 2.6 Ma (O'Meara *et*

al., 2006). The preference of this model would indicate that Pleistocene climatic oscillations have induced a slow down or an increase in the rate of evolution; 5) an Ecological Release model (ER) that assumes a constrained evolution under an OU model prior to a fixed point in time, followed by un unconstrained evolution under a BM model (Slater, 2013). The selection of this model would suggest that Pleistocene climatic oscillations have released species from their ancestral adaptive optima and allowed them to explore wider morphospace and ecospace; 6) a two-rate BM model (BMM Regime) where species from the lower montane forest (LMF, <1800m) and the upper montane forest (UMF, >1800m) can have different rates of trait evolution. A better fit of this model would indicate that species from the two main vegetation belts of MMF have distinct rates of trait evolution (Leuschner & Moser, 2008); 7) an OU model with two optima (OUM Regime), where species from the LMF and UMF can have different adaptive optima. The selection of this model would indicate that species from the two vegetation belts have evolved under divergent selection (convergence); 8) a two-rate BM model (BMM Clade) were species from the *micrantha* alliance can have a different rate of evolution from the rest of the phylogeny. The choice of this model would indicate a change in the tempo of trait evolution for the species of the *micrantha* alliance; 9) an OU model with two optima (OUM Clade) where species from the *micrantha* alliance can have a different adaptive optimum from the rest of the phylogeny. A better fit of this model would indicate that the species of the *micrantha* alliance have shifted towards a new adaptive optimum; 10) a three rate BM model (BMM Geo) where species from the northern Andes, the Amotape-Huancabamba zone and the central Andes can have different rates of evolution. The choice of this model would suggest that the rate of trait evolution is region specific; 11) finally, an OU model with three optima (OUM Geo) where species from the three regions of the MMF can have divergent adaptive optima. The selection of this last model would indicate that the direction of stabilizing selection is region-specific. We note that the BM or OU are particular case of the complex model described above where rates or optima for the different part of the tree are equal. The BM model is also a particular case of the OU model when the selection parameter (α) tends toward zero.

Inference of the habitat (LMF vs. UMF) and the geographical area (northern Andes, Amotape-Huancabamba zone, central Andes) across the phylogenies were performed using a marginal ancestral state reconstruction method in the R package phytools with a symmetric transition matrix (Revell, 2012). Time slices, regimes, clade and geography were mapped onto the trees using the "paintBranches" function from phytools. Figure 2 illustrates mapping of traits onto the MMC tree including all the species for which we have morphological data. The different models were fitted for each trait/climate niche in the R package MvMORPH (Clavel *et al.*, 2014) that takes as input trees with mapped characters. All the analyses and preparation steps were performed on the

MMC tree and repeated on the posterior sample of 100 trees to account for phylogenetic uncertainty. Model performances were compared using corrected Akaike information criterion (AICc) scores.

In order to evaluate the statistical power of our phylogeny to discriminate among the different models, we applied the Monte Carlo-based approach developed by Boettiger *et al.* (2012). The method first estimates parameters for both models to be compared with the real data. Then two datasets are generated by simulating a large number of trait samples under each model and using the parameters estimated using the real data. For each sample in each dataset, the parameters are re-estimated using the two models, which allow computing a likelihood ratio δ (see Equation 3 in Boettiger *et al.*, 2012). Finally the two distributions of δ generated are compared between them and with the likelihood ratio observed on the real data (δ_0). The probability of rejecting the "simpler" model (or the poorer fit based on AICc) is obtained by computing the proportion of the distribution of δ_{s1} (computed on the dataset simulated under model 1, the poorer fit model) that is greater than δ_0. The "power" of the test is evaluated by computing the proportion δ_{s2} (computed on the dataset simulated under the model 2, the fittest model) that is greater than the 0.95 quantiles of the distribution of δ_{s1}. For each trait, we applied this approach to verify the ability of our data to discriminate between the two best fitting models according to the AICc. For each model considered, 2000 samples of traits were simulated on the MCC tree with the function "OUwie.sim" in the R package OUwie (Beaulieu & O'Meara, 2012).

Finally, we extracted the selective parameter (α) estimated under the OU model to compute the phylogenetic half-life ($t_{1/2} = \ln 2/\alpha$) for each trait. The half-life is expressed in the same time units as the time-calibrated trees used to compute it (here in My) and represents the amount of time necessary for a clade to move half way out from the adaptive optimum of its ancestor (Hansen 1997). The half-life can be used as a proxy to estimate the phylogenetic signal in a trait (Hansen 2014, Münkemüller *et al.*, 2014). A half-life greater than the age of the MRCA of the clade under consideration indicates that the evolution of the trait is almost indistinguishable from evolution under a BM model and therefore the phylogenetic signal is strong (Uyeda & Harmon, 2014). In contrast, when the half-life is much smaller than the age of the youngest node in the clade, the evolution of the trait is not phylogenetically correlated and the phylogenetic signal is null. If the best model chosen for a trait during the model comparison process was an OUM model (multiple optima), we also report the estimated half-life under this model. This allows comparing how much phylogenetic signal remains after we account for divergent adaptive regimes. In addition, we report the stationary variance ($V_y = \sigma/2*\alpha$) estimated under the OU

model (and OUM when preferred) that quantifies the relative strength of drift around the adaptive optimum (σ) vs. directional selection towards the optimum (α).

RESULTS

Disparity through time plots indicate that in the initial history of the Andean *Macrocarpaea* clade, disparification is not different from the expectation under a BM model of evolution for all the traits investigated (Fig.1). Latter (~3.2Ma), the disparity becomes much greater than the BM expectation for the altitudinal niche (Fig.1). Disparity also becomes moderately greater towards more recent times for leaf size, flower size and SLA (Fig.1). Disparity bellow the median of the BM expectation is observed only in the initial history for the latitudinal niche but it still remains within the 95% interval for the BM expectation. This is also reflected by the MDI statistic which is moderately negative only for the latitudinal niche, but it is not statistically different from the BM expectation (MDI== -0.011, P= 0.74). Negative MDI is the expected pattern under an adaptive radiation (Harmon *et al.*, 2003). Clearly here none of the variables investigated match this prediction. Positive and statistically different MDIs from the BM expectation were recovered for the altitudinal niche (MDI=0.399, P= 0.98). This indicates that species within the different subclades occupy a larger proportion of the altitudinal niche than expected from a BM evolution. The MDI for the other variables are moderately positive to positive (MDI=0.007-0.185) but none are significantly different from the BM expectation (P=0.74-0.91).

Maximum likelihood ancestral state estimation for the regime consistently reconstructs the MRCA of Andean *Macrocarpaea* as occurring in the UMF (Figure 2) and the MRCA of the *micrantha* alliance as occurring in the LMF. Concerning the geography, the Andean MRCA is ambiguously reconstructed as occurring either in the northern Andes or the AHZ in agreement with previous work (Struwe *et al.*, 2009; Vieu *et al.*, in review).

Model comparison for the evolution of leaf size preferred the OUM Clade model that accounts for different adaptive optima for the *micrantha* alliance clade and the rest of the phylogeny (Table 3). The second best model is the OU1, which has a single optimum for the whole phylogeny, but the AICc difference between these two models is large (MCC tree: ΔAICc=7.66; median of the posterior sample: ΔAICc=7.68), indicating a strong support for the OUM Clade model over the alternative models (Burnham & Anderson, 2002). The optimum for the *micrantha* alliance clade (θ_2, full parameter estimates are available in the appendix S4) is toward larger leaves than for the rest of the phylogeny (θ_1, see Fig 3a). The simulation procedure to estimate the power of our

data to discriminate between the two best models based on the MCC tree indicates that we can reject the OU1 in favour of the OUM Clade model with very high probability (P=0.005, see Appendix S3 for plots of the distributions of simulations) and very strong power (98.9%).

Considering flower size, the best fit is to the BMM Clade model that accounts for different rates of evolution for the *micrantha* alliance and the rest of the phylogeny. The second best model is the OU1, but again the AICc difference is large (MCC tree: ΔAICc=7.03; median of the posterior sample: ΔAICc=5.31). The estimated rate of flower size evolution for the *micrantha* alliance (σ_2) is more than four times higher than the rate for the rest of the phylogeny (σ_1, Fig 3b). The power test indicates that we can reject the OU1 in favour of the BMM Clade with a high probability (P=0.003) and strong power (94.2%).

For plant size, the BMM Geo model that allows a region-specific rate of evolution (three rates) obtains the lowest AICc and is followed by the ER that accounts for a release of stabilizing selection during the last 2.6 My (MCC tree: ΔAICc=1.84; median of the posterior sample: ΔAICc=2.59). Simulations rejected the ER in favour of the BMM Geo (P=0.024) with substantial power (88.6%). Under the BMM Geo model the rate of plant size evolution estimated in the Amotape-Huancabamba region is almost the double (σ_2) that of the rate in the central Andes (σ_3) and more than five times larger than in the northern Andes (σ_1, Fig 3c).

For the evolution of SLA, the OUM Regime model that accounts for different adaptive optima for the LMF and the UMF, and the OUM Clade model obtain almost equivalent support (MCC tree: ΔAICc=1.74; median of the posterior sample: ΔAICc=0.091). In comparison all the other models are outperformed (MCC tree: ΔAICc>8; median of the posterior sample: ΔAICc>7). Simulations based on the MCC tree shown that we can't reject the OUM Regime in favour of the OUM Clade at the 5% level (P=0.063) with a power of 88.4%. The estimates of the optimum for the UMF (θ_1) under the OUM Regime and the optimum for the species outside of the *micrantha* alliance (θ_1) under the OUM Clade are very close and correspond to a SLA of 9.90 $mm^2.mg^{-1}$ and 10.00 $mm^2.mg^{-1}$ respectively (once transformed back to their original scales). The optimum estimated for the LMF (θ_2) is a bit larger than the optimum for the micrantha alliance (θ_2) with a SLA of 21.20 $mm^2.mg^{-1}$ and 18.65 $mm^2.mg^{-1}$ respectively, but imply in both cases selection towards leaves that are twice as thin.

The best model for the altitudinal niche is, as expected, the OUM Regime model that account for different optima between the LMF and the UMF. All the other models have much poorer fit (MCC tree: ΔAICc>50; median of the posterior sample: ΔAICc>50). The second best model is the OUM Clade model that also has a large AICc difference with the next best fitting model (MCC tree: ΔAICc>9; median of the posterior sample: ΔAICc>8). Not surprisingly, the OUM Clade model is rejected in favour of the OUM Regime model with high probability (P=0.000) and maximal power (100%).

For the latitudinal niche, the OUM Geo model that accounts for region-specific adaptive optima (three θ) is returned, as expected, as the best model (MCC tree: ΔAICc>33; median of the posterior sample: ΔAICc>42). The second best model is the BMM Geo model, but in comparison the remaining models have much poorer fit (MCC tree: ΔAICc>42; median of the posterior sample: ΔAICc>36). The BMM Geo model is rejected in favour of the OUM Geo model with high probability (P=0.00) and strong power (P=100%). However, the observed likelihood ratio using the real data (δ_0) is much smaller than the distribution of ratio obtained on the dataset simulated with the OUM Geo model. This indicates that the OUM Geo model might not describe properly the data. Under the BMM Geo model the rate of evolution in the central Andes (σ_3) is 12 times higher than in the northern Andes (σ_1) and more than 30 times higher than in the Amotape-Huancabamba region (σ_2).

Overall, the EB model obtains extremely low support for all the variables investigated and its rate change parameter (r) is always returned as null. Models that account for potential influences of the Pleistocene climatic oscillations (SHIFT, ER) obtain some support but only for the evolution of plant size.

The phylogenetic half-life estimated under the OU1 model for the different traits (see Table 4) is in general much smaller than the age of the MRCA of *Macrocarpaea* in the Andes (7.16 Myr). This suggests relatively low phylogenetic signal. The only exception is the regional niche that has a half-life comparable with the age of the MRCA of the clade ($t_{1/2}$= 5.44My). This is consistent with the model comparison that preferred the BM1 over the OU1 for this trait. Interestingly leaf size, flower size and SLA have similar half-lives ($t_{1/2}$=1.635-1.978Myr), while the half-life of plant size is substantially greater ($t_{1/2}$=2.925 Myr). Under the OUM Clade model (best fit), the half-life for leaf size ($t_{1/2}$=0.266 Myr) collapses to a value smaller than the youngest node in the phylogeny (age=0.360 My). This indicates that once we account for divergent adaptive optima for the *micrantha* alliance and the rest of the phylogeny, the phylogenetic signal becomes null and leaf size evolution around one optima or

the other is not phylogenetically correlated. The same type of observation but with a greater magnitude is seen for the evolution of altitudinal niche under the OUM Regime ($t_{1/2}$=0.027 Myr) and to the regional niche under the OUM Geo ($t_{1/2}$=0.125 Myr) with a reduction of the phylogenetic signal of 33 and 43 times respectively.

DISCUSSION

Is the radiation of *Macrocarpaea* in the Andes, an adaptive radiation?

Based on the traits investigated in this study, we found that it is very unlikely that the radiation of *Macrocarpaea* in the Andes was associated with niche partitioning (adaptation). Neither the disparification analyses with the MDI statistic nor the model comparison supports the pattern of an "early burst" of disparification expected under an adaptive radiation (Harmon *et al.*, 2003; Gavrilets & Losos, 2009). In fact, the EB model performed poorly in the model comparisons for all the traits investigated and its parameter that describes the time dependency of trait evolution is always returned null which collapse this model to a simple Brownian motion model (BM) penalized for an additional parameter. Despite the purported prevalence of adaptive radiations in nature (Schluter, 2000), the early burst pattern of disparification predicted by theory has been inferred only for a relatively small number of study clades (Harmon *et al.*, 2003; Slater *et al.*, 2010; Schweizer *et al.*, 2014; but see Harmon *et al.*, 2010). Recently, Slater and Pennell (2014) suggested that the scarcity of evidence for early burst models found to date was probably often the consequence of lack of sufficient statistical power of PCM instead of evidence for the rarity of adaptive radiations in evolution. In our approach, model comparison is complemented with disparification through time methods that allow graphical inspection of a trend in the disparification of traits. For most traits investigated, there is no signal of an early burst of disparification thus confirming the results from the model comparison procedure.

The only potential exception is for the latitudinal niche. The latitudinal niche reflects the north to south climatic variations (principally temperature seasonality) experienced by species occurring in different regions of the Andes. In the initial history of the clade, the observed disparity curve for the latitudinal niche, while not reaching the significance level, is below the median of simulations under a BM process. Vieu *et al.* (in review) proposed that the radiation of *Macrocarpaea* in the Andes was driven by dispersification (diversification associated with dispersal) or range expansion during rapid colonization (<2My) of the MMF biome some 7Ma. The trajectory of the disparity curve when is below the median of the BM simulations suggests that at the time of the burst of species diversification of *Macrocarpaea* in the Andes, latitudinal niche disparity was predominantly partitioned

between lineages (Harmon *et al.*, 2003). It has been recently suggested that climatic niche evolution reflects more the history of species migration rather than the mode of evolution of their climatic tolerances (Boucher *et al.*, 2014). Accordingly, if we assume that changes in the latitudinal niche were predominantly associated with dispersal along the Andes, this result is consistent with diversification associated with colonization of previously unoccupied regions. Later the disparity curve tracks back to the median of the BM simulation, most likely as a consequence of more recent exchanges between the different regions of the Andes.

Pleistocene climatic oscillation, an engine for disparity?

We did not find support for a dramatic impact of Pleistocene climatic oscillations (PCO) on trait or niche evolution across the Andean *Macrocarpaea* clade as a whole. It is known that PCO caused dramatic fluctuations in altitudinal distributions of the vegetation in the MMF (Hooghiemstra & Van der Hammen, 2004, Flantua *et al.*, 2014). It has been proposed that this recent dynamic history could have promoted allopatric speciation in the Andes through repeated cycles of species range contraction and re-expansion (Rull, 2011). PCO might also have enhanced ecological divergence by driving populations unable to track the pace of the displacement of their habitat, to adapt to the new local ecological conditions in order to survive (Donoghue, 2008). In our study, the SHIFT and ER model that specifically account for an influence of the PCO on the tempo and the mode of evolution respectively (Slater, 2013) performed relatively poorly in comparison with others models explaining *Macrocarpaea* evolution for almost all the traits we investigated. The only exception is plant size for which the SHIFT and the ER performed substantially better than most of the other models considered but which still did obtain the lower AICc (best fit) in any of the trees from the posterior sample.

It should be noted that the parameter values estimated under the SHIFT and ER models are often biologically unrealistic (Table S4). The pre-PCO evolutionary rate (σ_1) of the SHIFT model is returned null for all traits excepted latitudinal niche. Also the selection parameter of the ER model which constrained the evolution around the ancestral optimum for the pre-PCO part of the tree, is estimated to be very large ($\alpha>146$) for leaf size, flower size and altitudinal niche. According to these estimates, no evolution would have occurred in these traits before the PCO, either because of the absence of evolutionary potential ($\sigma_1 = 0$) or due to very strong stabilizing selection pressures (very large α). The fossil record documents numerous examples of lineages that experienced trait stasis (retention of the ancestral phenotype) over million years (Eldredge *et al.*, 2005). However, we think that a complete stasis over almost 4.5My as suggested here, is unlikely and rather reflects a statistical limitation

of models that partition phylogenetic trees into time slices using PCM. It has been shown through simulations that in the absence of fossil data, PCM have a low ability to converge toward the parameter values that generated the simulated data with these types of models (Slater, 2013). The latitudinal niche is the only trait that has a non null pre-PCO rate of evolution ($\sigma_1 = 2.42$) under the SHIFT model and a realistic estimate for selection parameter (α=0.081) under the ER model. Among the traits we investigated, the latitudinal niche displays the stronger phylogenetic signal (Table 4). Thus, in the absence of fossil data, it seems possible that substantial phylogenetic signal could be a required to obtain reliable parameter estimates with time slice models.

Evidence for elevation belt, clade history and region-specific evolution?

We found that leaf size evolution is better explained by a model of stabilizing selection that accounts for divergent adaptive optimum between the *micrantha* alliance and the rest of the phylogeny (OUM Clade). Concerning the SLA, the OUM Clade model obtains also the best fit, but the model that accounts for divergent optima between the two elevations belts (OUM Regime) obtains equivalent support. This is likely because the lower sampling fraction for SLA make these two models very similar and thus statistically hard to discriminate (Fig S2). Taken together the preference of the OUM Clade model for leaf trait evolution indicates that the *micrantha* alliance lineage escaped from the ancestral adaptive optimum that constrained leaf evolution in the others lineages of *Macrocarpaea* in the Andes and evolved toward a new optimum. These results are consistent with the hypotheses that the colonization of the lower montane forest (LMF<1800m) by the *micrantha* alliance represented a shift into a new adaptive zone. The LMF experiences higher vertical precipitations (rainfall) and higher temperatures than the UMF but also occurs on soils with on average higher nutrient contents (Webster 1993, Beck & Richter 2008, Gerold 2008, Leuschner & Moser, 2008). These abiotic factors likely enhance competition for light and thus tend to favour fast-growing species which are often characterized by larger, thinner and shorter lived leaves than species found in the UMF (Moser *et al.*, 2007,). Thus, the evolution of larger leaves and higher SLA in the *micrantha* alliance likely represent functional adaptations to the LMF conditions that potentially improve the competitive ability of species in this clade in the LMF plant communities.

Interestingly, the few species in the other parts of the phylogeny that occur in the LMF for which we have morphological (4 species) and SLA (2 species) data do not have especially large leaves nor high SLAs (Figs. 3A and 3D). This likely explains why the OUM Regime model that accounts for divergent adaptive optima for the LMF and the UMF obtains only the third best score to explain leaf size evolution, far behind the OUM Clade

model but also the OU1 model that have a single adaptive optimum for the whole phylogeny. One might have expected that natural selection would have lead to phenotypic convergence of all the species that occur in the LMF due to the similar environmental conditions they experience. Phenotypic convergence has been reported in a wide range of organisms (Pupo *et al.*, 2000; Losos 2011; Frédérich *et al.*, 2013) and is thought to reflect the deterministic nature of evolution on a macroevolutionary adaptive landscape (Mahler *et al.*, 2013). Here, the phenotypic deviation from the adaptive optimum of the *micrantha* alliance for the LMF species that do not belong to this clade suggests that: 1) either these species were wrongly assigned as occurring in LMF, for example if the LMF/ UMF transition, which is more a diffuse than a strict discontinuity, is located bellow 1800m where these species occur; 2) or that they do not share the adaptive strategy of the species of the *micrantha* alliance, possibly because their leaf evolution remain constrained around the ancestral adaptive optimum of *Macrocarpaea* in the Andes. Evolutionary constraints can arise from a variety of processes such as via genetic correlations between several traits (i.e.: pleiotropic effects, linkage disequilibrium) but also due to lack of genetic variation or the limited efficiency of selection compared to drift in small populations (Arnold, 1992; Agrawal & Stinchcombe, 2009; Futuyma, 2010). Such constraints on leaf evolution might have prevented these *Macrocarpaea* species to respond in a predictable manner to the novel selective pressure of the LMF abiotic environment, which in turn could have put them at competitive disadvantage in the LMF plant communities.

The limited niche overlap between the species from the LMF and UMF observed along the altitudinal niche axes (Fig 3e) indicates that the hypothesis of wrong assignments of species to the LMF is plausible. However, it is striking that there is only a single species divergence event within the LMF outside the *micrantha* alliance (sampled and non-extinct) in our phylogeny. While we acknowledge that our low sampling fraction for the species from the northern Andes might exacerbate this pattern, the lack of species diversification we infer in the LMF for lineages outside the *micrantha* alliance indicates that they have not been particularly successful within the LMF. By comparison, the *micrantha* alliance forms a relatively large clade of 25 species sampled here, 19 of which occurs in the LMF from northern Ecuador to central Peru. The study of Vieu *et al.* (in prep) detected a potential shift in the diversification rate along the branches leading to the basal part of the *micrantha* alliance. It is tempting to see a causal link between the switch toward the new adaptive zone and the burst of species diversification.

The theory of the phenotypic adaptive landscape predicts that the ecological release following the entry into a new adaptive zone should facilitate diversification through the partitioning of the broad adaptive zone into several more specialized sub zones (Simpson, 1953 p.209). In the example of the *micrantha* alliance it is not clear whether the potential burst of species diversification has been promoted by finer niche specialization within the LMF. Model comparisons for the evolution of flower size strongly favour the BMM Clade model and the rate of evolution estimated for the *micrantha* alliance is more than four times higher than the rates estimated for the others lineages. It is perhaps notable that the *micrantha* alliance includes both the species with the largest and the smallest flowers reported in our phylogeny. These results are consistent with the hypothesis of a release on flower size evolution associated with entry of the clade into the LMF. However, it is hard to evaluate how much the faster flower size evolution is responsible for the burst of species diversification detected at the base of this clade. All species of *Macrocarpaea* have greenish, whitish to yellowish bell-shaped flowers rich in nectar mostly released at night (Wolff, 2006) which are predominantly pollinated by nectar feeding bats (but are also visited by hummingbirds during the day and geometrid moths during the night). Only 10 species of bats specialized to nectar feeding (Phyllostomidae) are known for the whole South America, (Fleming *et al.*, 2009). These bats are generalists and tend to feed on any available nectar sources (Muchhala, 2008), which suggests that pollinator specialization is an unlikely mechanism driving species divergence in *Macrocarpaea*. This also suggests that different *Macrocarpaea* species compete for the same pollinators and the evolution of different sized flowers might have limited interspecific pollen transfer through the spatial segregation of the pollen deposit on bats heads (Muchhala, 2008). In the absence of *Macrocarpaea* competitors in the LMF, species of the *micrantha* alliance might have explored de novo the whole *Macrocarpaea* flower size spectrum. Thus, pollinator competition is a credible explanation for how flower size variation could potentially provide a mechanism for prezygotic isolation in *Macrocarpaea* that might translate into species divergence. However, we think that the faster flower size evolution is unlikely to be the main driver of the pulse of diversification detected at the base of the *micrantha* clade. Instead, the rapid range expansion of the clade that followed its entry in the LMF (Fig. 2d, Vieu *et al.*, in prep) seems to be a more plausible explanation. In other words, the colonization of the LMF, lacking other *Macrocarpaea* lineages, likely represented an opportunity for a second wave of range expansion associated with species diversification at lower elevations for the *micrantha* alliance.

Adaptive zone or niche shifts are usually thought to be accompanied or facilitated by the acquisition of a key innovation (Simpson, 1953). A key innovation is often referred as a new trait that allows organism to interact with new environments in a novel way (for detailed definitions of key innovation see Hunter, 1998). There are several examples in the literature of the acquisition of novel traits that allowed plant lineages to invade a new adaptive zone, such as the cushion life form of the *Androsace* (Primulaceae; Boucher *et al.*, 2012) or the pachycaul stem rosette life form of *Espeletineaa* (Asteraceae; Monasterio & Sarmiento, 1991) that allowed these clades to colonize and diversify in harsh alpine habitats. In the case of the *micrantha* alliance, no new trait is associated with adaptation to LMF, but instead a change in leaf size, a leaf fundamental trait, seems to have been made possible by the liberation from an ancestral evolutionary constraint. It has recently been suggested that constraint breaking mutations are the most prominent agents of biological innovation (Nei, 2013). Such a constraint breaking mutation(s) could have arisen in the ancestor of the *micrantha* alliance and been maintained with standing genetic variation. Then, the entry into the LMF through the dispersal of propagules or the fluctuation of the vegetation belts during the PCO might have turned this mutation into an advantageous one that allowed evolution toward a new phenotype adapted to the LMF way of life. The estimated age of the MRCA of the *micrantha* alliance (2.6 Myr) coincided remarkably closely with the beginning of the PCO. Finally, the fact that a successful entry in the LMF only happened once in *Macrocarpaea* instead of repeatedly along the Andes is consistent with the hypothesis of an advantageous mutation or a set of advantageous mutations unique to the *micrantha* alliance.

The last trait that we have not yet considered is plant size which broadly describes the different growth habits found in *Macrocarpaea* (small shrubs to tall trees). Model comparisons favoured a model that accounts for region-specific rates of evolution (BMM Region) with the rate estimated for the AHZ inferred to be two and four times faster than for the central Andes and the northern Andes respectively. This result supports the hypothesis that the AHZ stimulated morphological (adaptive?) evolution more than any other region of the Andes (Weigend, 2002; Mutke *et al.*, 2014). However, this assertion should be taken with some caution as our low sampling fraction for the northern Andes might again alter the pattern. Nevertheless, the fact that the *micrantha* alliance likely originated in the AHZ (Fig. 2d; Vieu *et al*., in prep.) tells us that there may be indeed something special about the AHZ.

Note on phylogenetic half-life (phylogenetic signal).

Our estimation of the phylogenetic signal for the various traits we investigated is in general relatively low. The phylogenetic half-life estimated under the OU model, which represents the amount of time necessary for a clade to move half way out from the adaptive optimum of its ancestor (Hansen 1997) is about a third of the height of the phylogenetic tree for leaf size, SLA and flower size and about a sixth for the altitudinal niche. This tells us that phylogenetic relationships are a weak predictor of these traits values for a particular species and that closely related species are often more dissimilar than expected under a BM model in regard to their leaf and flower morphologies but also their altitudinal position (Boucher *et al.*, 2014). The half-life for plant size is substantially greater than for the other morphological traits suggesting that this particular trait might be more phylogenetically conserved. Concerning the latitudinal niche, the half-life is very close to the BM expectation (almost equal to the height of the phylogenetic tree), which indicates that closely related species tend to occur in the same geographical area (Vieu *et al.*, in review). However, once we accounted for divergent adaptive optima for each main region the half-life collapsed to a very low value suggesting that within each region, the geographical position of a species is random with regard to the phylogeny (Uyeda & Harmon, 2014). This means that while closely related species tend to co-occur in the same major region, within each region they tend to be less closely adjacent geographically than expected under a BM model. The same applies to leaf size, SLA and the altitudinal niche. Once we accounted for the divergent adaptive optima, the phylogenetic signal collapsed to very low values indicating that within the boundaries of an adaptive optimum, or an adaptive zone, the position of a particular species is relatively random in regard to the phylogeny. Thus the phylogenetic signal we measured before accounting for the divergent adaptive optima was very likely a measure of the phylogenetic signal for the belonging to particular optima or to the other. This implies that within the boundaries of an adaptive zone, the position of a species is not phylogenetically constrained but the fact to belong to one adaptive zone or another is. We guard ourselves from interpreting the phylogenetic signal in terms of niche conservatism as the link between the former and the latter has been shown to be complex (Revell *et al.*, 2008; Münkemüller *et al.*, 2014). We believe that results from the model comparisons discussed above provide a good idea of the importance of niche conservatism and niche evolution during the diversification of *Macrocarpaea* in the Andes but also of which traits are constrained or not.

Potential caveats of the study.

The quality of the data we used for the morphological traits is questionable and it is clear that several individual measures for each species would have been preferable by allowing direct estimation of species means and standards deviation. However, we think that the large differences in model fit we obtain with these data suggest that the patterns we infer are probably relatively robust. Also, while we tested a relatively wide spectrum of models and specific hypothesis, we acknowledge that this set is not exhaustive and an evolutionary picture with a finer resolution for *Macrocarpaea* in the Andes is totally possible.

CONCLUSIONS

In this study we have shown that it is very unlikely that adaptive divergence fostered the main radiation of *Macrocarpaea* in the MMF. Instead, geographical processes arising from the rapid range expansion of the genus along the Andes, ie a pattern of dispersification is favoured. We also showed that during more than the half of the history of the genus in the Andes, species were restricted almost exclusively to the UMF possibly because of evolutionary constraints. More recently and coinciding with the beginning of the PCO, the *micrantha* lineage successfully established and diversified in the LMF. This adaptive zone shift has been accompanied by evolution of larger and thinner leaves that likely represent adaptations to warmer and wetter conditions in the LMF. We suggest that the development of this new phenotype potentially originated via constraint breaking mutation(s) acquired by the MRCA of the *micrantha* alliance. Colonization of the LMF was quickly followed by a new burst of range expansion associated with renewed dispersification along the Andes corresponding to the secondary burst of diversification detected at the basis of the clade in a previous study (Vieu *et al.*, in prep). Finally, while leaf traits and altitudinal niche variations seem to be bounded within each adaptive zone, variation within these bounds is relatively random in regard with the phylogeny. Overall, the results we report here are in close agreement with the theory of quantum evolution of Simpson (1953) that predicts that evolution is relatively stable and bounded during relatively long periods of time but that occasionally adaptive transitions occur.

AKNOWLEDGMENTS

The work was supported by the Fonds des donations of the Université de Neuchâtel, the Basler Stiftung für biologische Forschung, and the Fondation Dr. Joachim de Giacomi.

REFERENCES

Adams, D. C., Berns, C. M., Kozak, K. H. & Wiens, J. J. (2009) Are rates of species diversification correlated with rates of morphological evolution? *Proceedings of the Royal Society B: Biological Sciences*, **276**, 2729-2738.

Agrawal, A. F., & Stinchcombe, J. R. (2009) How much do genetic covariances alter the rate of adaptation?. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 1183-1191.

Antonelli A. & Sanmartín I. (2011b) 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.

Arnold, S. J. (1992) Constraints on phenotypic evolution. *American Naturalist*, **140**, S85-S107.

Beaulieu J. M., O'Meara B. C. (2012) OUwie: analysis of evolutionary rates in an OU framework. R package version 1.21.

Benham, P. M., Cuervo, A. M., McGuire, J. A., & Witt, C. C. (2014) Biogeography of the Andean metaltail hummingbirds: contrasting evolutionary histories of tree line and habitat-generalist clades. *Journal of Biogeography*, doi: 10.1111/jbi.12452

Benkman, C. W. (2003) Divergent selection drives the adaptive radiation of crossbills. *Evolution*, **57**, 1176-1181.

Boettiger, C., Coop, G., & Ralph, P. (2012) Is your phylogeny informative? Measuring the power of comparative methods. Evolution, 66, 2240-2251

Boucher, F. C., Thuiller, W., Roquet, C., Douzet, R., Aubert, S., Alvarez, N., & Lavergne, S. (2012) Reconstructing the origins of high-alpine niches and cushion life form in the genus *Androsace* s.l. (Primulaceae). *Evolution*, **66**, 1255-1268.

Boucher, F. C., Thuiller, W., Davies, T. J., & Lavergne, S. (2014) Neutral biogeography and the evolution of climatic niches. *The American Naturalist*, **183**, 573-584.

Burbrink, F. T., & Pyron, R. A. (2010). How does ecological opportunity influence rates of speciation, extinction, and morphological diversification in New World ratsnakes (tribe Lampropeltini)?. *Evolution*, **64**, 934-943.

Burbrink, F. T., Chen, X., Myers, E. A., Brandley, M. C., & Pyron, R. A. (2012) Evidence for determinism in species diversification and contingency in phenotypic evolution during adaptive radiation. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 4817-4826.

Burnham, K.P. & Anderson, D.R. (2002) *Model selection and multimodel inference: a practical information-theoretic approach*, 2nd Ed. Springer-Verlag, New York.

Butler, M. A., & King, A. A. (2004). Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *The American Naturalist*, **164**, 683-695.

Casner, K. L., & Pyrcz, T. W. (2010) Patterns and timing of diversification in a tropical montane butterfly genus, *Lymanopoda* (Nymphalidae, Satyrinae). *Ecography*, **33**, 251-259.

Clavel, J., King, A. & Paradis, E. (2014) MvMORPH: multivariate comparative tools for fitting evolutionary models to morphometric data. v. 1.0.2. Available at: http://cran.rproject.org/web/packages/raster/index.html

Comes, H. P., Tribsch, A., & Bittkau, C. (2008) Plant speciation in continental island floras as exemplified by Nigella in the Aegean Archipelago. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **363**, 3083-3096.

Coyne, J. A. (1992) Genetics and speciation. *Nature*, **355**, 511-515.

Doebeli, M., & Dieckmann, U. (2003) Speciation along environmental gradients. *Nature*, **421**, 259-264.

Donoghue, M. J. (2008). A phylogenetic perspective on the distribution of plant diversity. *Proceedings of the National Academy of Sciences*, **105**, 11549-11555.

Eldredge, N., Thompson, J.N., Brakefield, P.M., Gavrilets, S., Jablonski, D., Jackson, J.B.C., Lenski, R.E., Lieberman, B.S., McPeek, M.A. & Miller, W. III (2005) The dynamics of evolutionary stasis. *Paleobiology*, **31**, 133–145.

Flantua S.G., Hooghiemstra H., Van Boxel J.H., Cabrera M., González-Carranza Z., & González-Arango C. (2014) Connectivity dynamics since the last glacial maximum in the northern Andes: a pollen-driven framework to assess potential migration. *In: Stevens W.D., Montiel O.M. & Raven P.H (eds), Paleobotany and biogeography: a festschrift for Alan Graham in his 80th year*, Missouri Botanical Garden Press, St. Louis, pp. 98-123.

Fleming, T. H., Geiselman, C., & Kress, W. J. (2009) The evolution of bat pollination: a phylogenetic perspective. *Annals of Botany*, **104**, 1017-1043.

Frédérich, B., Sorenson, L., Santini, F., Slater, G. J., & Alfaro, M. E. (2013) Iterative ecological radiation and convergence during the evolutionary history of damselfishes (Pomacentridae). *The American Naturalist*, **181**, 94-113.

Futuyma, D. J. (2010) Evolutionary constraint and ecological consequences. *Evolution*, **64**, 1865-1884.

Gavrilets, S. & Losos, J.B. (2009) Adaptive radiation: contrasting theory with data. *Science*, **323**, 732–737.

Gerold G. (2008) Soil, climate and vegetation of tropical montane forests – a case study from the Yungas, Bolivia. *In: Gradstein, R., Homeier, J., Gansert, D. (Eds.), The Tropical Mountain Forest. Patterns and Processes in a Biodiversity Hotspot. Biodiversity and Ecology series, 2.*, Gottingen Centre for Biodiversity and Ecology, Gottingen, pp. 137–162

Gittenberger, E. (1991) What about non-adaptive radiation?. *Biological Journal of the Linnean Society*, **43**, 263-272.

Givnish T.J., Barfuss M.H.J., Van Ee B., Riina R., Schulte K., Horres R., Gonsiska P.A., Jabaily R.S., Crayn D.M., Smith J.A.C., Winter K., Brown G.K., Evans T.M., Holst B.K., Luther H., Till W., Zizka G., Berry P.E., Sytsma K.J. (2014) Adaptive radiation, correlated and contingent evolution, and net species diversification in Bromeliaceae. *Molecular Phylogenetics and Evolution*, **71**, 55–78.

Grandcolas, P., Nattier, R., Legendre, F., & Pellens, R. (2011) Mapping extrinsic traits such as extinction risks or modelled bioclimatic niches on phylogenies: does it make sense at all?. *Cladistics*, **27**, 181-185.

Grant J.R. (2004) De Macrocarpaeae Grisebach (ex Gentianaceis) speciebus novis V: Twenty-three new species largely from Peru, and typification of all species in the genus. *Harvard Papers in Botany*, **9**, 11–49.

Grant J.R. (2005) De Macrocarpaeae Grisebach (ex Gentianaceis) speciebus novis VI: seed morphology, palynology, an infrageneric classification, and another twenty-three new species largely from Colombia. *Harvard Papers in Botany*, **9**, 305–342.

Gutiérrez-Pinto, N., Cuervo, A. M., Miranda, J., Pérez-Emán, J. L., Brumfield, R. T., & Cadena, C. D. (2012) Non-monophyly and deep genetic differentiation across low-elevation barriers in a Neotropical montane bird (*Basileuterus tristriatus*; Aves: Parulidae). *Molecular Phylogenetics and Evolution*, **64**, 156-165.

Hansen, T.F. (1997). Stabilizing selection and the comparative analysis of adaptation. *Evolution*, **51**,1341-1351.

Hansen, T. F. (2014) use and misuse of comparative methods in the study of adaptation. *In: modern phylogenetic comparative methods and their application in evolutionary biology*, Springer-Verlag, Berlin, pp. 351-379.

Harmon L.J., Schulte J.A., Larson A., Losos J.B. (2003) Tempo and mode of evolutionary radiation in iguanian lizards. *Science*, **301**, 961–964.

Harmon, L.J., Weir, J.T., Brock, C.D., Glor, R.E., & Challenger, W. (2008) GEIGER: investigating evolutionary radiations. *Bioinformatics*, **24**, 129-131.

Harmon, L.J., Losos, J.B., Davies, T.J., Gillespie, R.G., Gittle- man, J.L., Jennings, W.B., Kozak, K.H., McPeek, M.A., Moreno-Roark, F., Near, T.J., Purvis, A., Ricklefs, R.E., Schluter, D., Schulte, J.A., Seehausen, O., Sidlauskas, B.L., Torres-Carvajal, O., Weir, J.T. & Mooers, A.Ø. (2010) Early bursts of body size and

shape evolution are rare in comparative data. *Evolution*, **64**, 2385–2396.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hijmans, R.J., van Etten, J., Mattiuzzi, M., & Hijmans, M.R.J. (2014). Package 'raster'. Available at: http://cran.rproject.org/web/packages/raster/index.html

Hipsley, C. A., Miles, D. B., & Müller, J. (2014) Morphological disparity opposes latitudinal diversity gradient in lacertid lizards. *Biology letters*, **10**, 20140101.

Hooghiemstra, H. & Van der Hammen, T. (2004) Quaternary ice-age dynamics in the Colombian Andes: developing an understanding of our legacy. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, **359**, 173–180.

Hoorn, C., Wesselingh, F.P., Ter Steege, H., Bermudez, M.A., Mora, A., Sevink, J., Sanmartín, I., Sanchez-Meseguer, A., Anderson, C.L., Figueiredo, J.P., Jaramillo, C., Riff, D., Negri, F.R., Hooghiemstra, H., Lundberg, J., Stadler, T., Särkinen, T., & Antonelli, A. (2010) Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science*, **330**, 927–931.

Hutter, C. R., Guayasamin, J. M., & Wiens, J. J. (2013) Explaining Andean megadiversity: the evolutionary and ecological causes of glassfrog elevational richness patterns. *Ecology letters*, **16**, 1135-1144.

Insel, N., Poulsen, C.J., Ehlers, T. A., & Sturm, C. (2012) Response of meteoric δ18O to surface uplift - implications for Cenozoic Andean Plateau growth. *Earth and Planetary Science Letters*, **317**, 262–272.

Kreft, H., & Jetz, W. (2007). Global patterns and determinants of vascular plant diversity. *Proceedings of the National Academy of Sciences*, **104**, 5925-5930.

Leuschner, C., & Moser, G. (2008) Carbon allocation and productivity in tropical mountain forests. *In: Gradstein SR, Homeier J, Gansert D (eds.) The tropical mountain forest - patterns and processes in a biodiversity hotspot. Göttingen Centre for Biodiversity and Ecology.* Biodiversity and Ecology Series 2, pp. 109-128.

Leuschner, C., Zach, A., Moser, G., Homeier, J., Graefe, S., Hertel, D., Wittich, B., Soethe, N., Iost, S., Röpderstein, M., Horna, V., and Wolf, K. (2013) The carbon balance of tropical mountain forests along an altitudinal transect, *In: Ecological Studies Vol. 221, edited by: Bendix, J., Springer Verlag*, Berlin, Heidelberg, New York, pp. 117–139

Losos, J. B., Jackman, T. R., Larson, A., de Queiroz, K., & Rodríguez-Schettino, L. (1998) Contingency and determinism in replicated adaptive radiations of island lizards. *Science*, **279**, 2115-2118.

Losos, J. B. (2010) Adaptive radiation, ecological opportunity, and evolutionary determinism. *American*

Naturalist, **175**, 623-639.

Mahler, D. L., Revell, L. J., Glor, R. E., & Losos, J. B. (2010) Ecological opportunity and the rate of morphological evolution in the diversification of Greater Antillean anoles. *Evolution*, **64**, 2731-2745.

Mahler, D. L., Ingram, T., Revell, L. J., & Losos, J. B. (2013) Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science*, **341**, 292-295.

Monasterio, M., & Sarmiento, L. (1991) Adaptive radiation of *Espeletia* in the cold Andean tropics. *Trends in Ecology & Evolution*, **6**,387-391.

Moser, G., Hertel, D., & Leuschner, C. (2007) Altitudinal change in LAI and stand leaf biomass in tropical montane forests: a transect study in Ecuador and a pan-tropical meta-analysis. *Ecosystems*, **10**, 924-935.

Mutke, J., Jacobs, R., Meyers, K., Henning, T., & Weigend, M. (2014). Diversity patterns of selected Andean plant groups correspond to topography and habitat dynamics, not orogeny. *Frontiers in genetics*, **5**:351.

Nei, M. (2013) *Mutation-driven evolution*. Oxford University Press, New York.

O'Meara, B.C., Ané, C., Sanderson, M.J., & Wainwright, P.C. (2006) Testing for different rates of continuous trait evolution using likelihood. *Evolution*, **60**, 922-933.

Orme, C.D.L., Davies, R.G., Burgess, M., Eigenbrod, F., Pickup,N., Olson, V.A., Webster, A.J., Ding, T.-S., Rasmussen, P.C., Ridgely, R.S., Stattersfield, A.J., Bennett, P.M., Blackburn, T.M., Gaston, K.J. & Owens, I.P.F. (2005) Global hotspots of species richness are not congruent with endemism or threat. *Nature*, **436**, 1016–1019.

Patton, J. L., & Smith, M. F. (1992) MtDNA phylogeny of Andean mice: a test of diversification across ecological gradients. *Evolution*, **46**, 174-183.

Penone, C., Davidson, A.D., Shoemaker, K.T., Di Marco, M., Rondinini, C., Brooks T.M., Young B.E., Graham C.H. & Costa, G.C. (2014) Imputation of missing data in life-history trait datasets: which approach performs the best?. *Methods in Ecology and Evolution*, **5**, 961-970.

Pupo, G. M., Lan, R., & Reeves, P. R. (2000) Multiple independent origins of *Shigella* clones of *Escherichia coli* and convergent evolution of many of their characteristics. *Proceedings of the National Academy of Sciences*, **97**, 10567-10572

Rabosky, D. L., Donnellan, S. C., Grundler, M., & Lovette, I. J. (2014) Analysis and Visualization of Complex Macroevolutionary Dynamics: An Example from Australian Scincid Lizards. *Systematic biology*, **63**, 610-627.

Reich, P. B., Walters, M. B., & Ellsworth, D. S. (1997) From tropics to tundra: global convergence in plant

functioning. *Proceedings of the National Academy of Sciences*, **94**, 13730-13734.

Revell, L. J., Harmon, L. J., & Collar, D. C. (2008) Phylogenetic signal, evolutionary process, and rate. *Systematic Biology*, ***57***, 591-601.

Revell, L. J. (2009) Size-correction and principal components for interspecific comparative studies. *Evolution*, **63**, 3258-3268.

Revell, L. J. (2012) phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, **3**, 217-223.

Rowe, K. C., Aplin, K. P., Baverstock, P. R., & Moritz, C. (2011) Recent and rapid speciation with limited morphological disparity in the genus *Rattus*. *Systematic Biology*, **60**, 188-203.

Rundell, R. J., & Price, T. D. (2009) Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends in Ecology & Evolution*, **24**, 394-399.

Rull V. (2011) Neotropical biodiversity: timing and potential drivers. *Trends in Ecology & Evolution*, **26**, 508–513.

Santos, T., Diniz-Filho, J.A.F., Rangel, T. & Bini, L.M. (2012) Package 'PVR'. Available at: http://cran.r-project.org/web/packages/PVR/index.html

Scheepens, J.F., Frei, E.S., & Stöcklin, J. (2010) Genotypic and environmental variation in specific leaf area in a widespread Alpine plant after transplantation to different altitudes. *Oecologia*, **164**, 141-150.

Schluter, D. (2000) *The ecology of adaptive radiation* .Oxford UK: Oxford University Press.

Schweizer, M., Hertwig, S. T., & Seehausen, O. (2014) Diversity versus disparity and the role of ecological opportunity in a continental bird radiation. *Journal of Biogeography*, **41**, 1301-1312.

Seehausen, O. (2006) African cichlid fish: a model system in adaptive radiation research. *Proceedings of the Royal Society B: Biological Sciences* **273**, 1987-1998.

Simpson, G. G. (1953) *The major features of evolution*. Columbia Univ. Press, New York.

Slater, G. J., Price, S. A., Santini, F., & Alfaro, M. E. (2010) Diversity versus disparity and the radiation of modern cetaceans. *Proceedings of the Royal Society B: Biological Sciences*, **277**, 3097-3104.

Slater, G. J. (2013). Phylogenetic evidence for a shift in the mode of mammalian body size evolution at the Cretaceous-Palaeogene boundary. *Methods in Ecology and Evolution*, **4**, 734-744.

Slater, G. J., & Pennell, M. W. (2014) Robust regression and posterior predictive simulation increase power to detect early bursts of trait evolution. *Systematic Biology*, **63**, 293-308.

Stekhoven, D.J., & Buehlmann, P. (2012) MissForest - non-parametric missing value imputation for mixed-type data. *Bioinformatics*, **28**, 112-118.

Uyeda, J. C., Hansen, T. F., Arnold, S. J., & Pienaar, J. (2011) The million-year wait for macroevolutionary bursts. *Proceedings of the National Academy of Sciences*, **108**, 15908-15913.

Uyeda, J., & Harmon, L.J. (2014) A novel bayesian method for inferring and interpreting the dynamics of adaptive landscapes from phylogenetic comparative data. *Systematic Biology*, **63**, 902-918.

van der Hammen, T., & Hooghiemstra, H. (2000) Neogene and Quaternary history of vegetation, climate, and plant diversity in Amazonia. *Quaternary Science Reviews*, **19**, 725-742.

Vieu, J., Hughes, C. H., & Grant, J. R. Evolutionary diversification in the hyper-diverse montane forests of the tropical Andes: range expansion drives radiation of the plant genus *Macrocarpaea* (Gentianaceae). *In review at Molecular Phylogenetics and Evolution*.

Webster, G. L. (1995) The panorama of Neotropical cloud forests. *In: Churchill, S. P., Balslev, H., Forero, E., & Luteyn, J. L. (Eds.). Biodiversity and conservation of Neotropical montane forests. Neotropical Montane Forest Biodiversity and Conservation Symposium 1,* New York Botanical Garden, New York, New York, pp. 53–77

Weigend, M. (2002) Observations on the biogeography of the Amotape-Huancabamba zone in northern Peru. *The Botanical Review*, **68**, 38-54.

Wiens, J. J. (2004) Speciation and ecology revisited: phylogenetic niche conservatism and the origin of species. *Evolution*, **58**, 193-197.

Wolff, D. (2006) Nectar sugar composition and volumes of 47 species of Gentianales from a southern Ecuadorian montane forest. *Annals of Botany*, **97**, 767-777.

Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.L., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J. & Villar, R. (2004) The worldwide leaf economics spectrum. *Nature*, **428**, 821–827.

Yoder, J.B., Clancey, E., Des Roches, S., Eastman, J.M., Gentry, L., Godsoe, W., Hagey, T.J., Jochimsen, D., Oswald, B.P., Robertson, J., Sarver, B.a.J., Schenk, J.J., Spear, S.F. & Harmon, L.J. (2010) Ecological opportunity and the origin of adaptive radiations. *Journal of Evolutionary Biology*, **23**, 1581–1596.

FIGURE CAPTIONS

Figure 1. Mean subclade disparity through time (DTT) for the 6 traits investigated in this study (solid line), based on the maximum clade credibility (MCC) tree of the Andean species of *Macrocarpaea*. The dashed lines indicate the median subclade DTT based on 2000 simulations of character evolution under a Brownian motion of evolution. The grey shaded areas indicate the 95% DTT range for the simulated data. The x axis is the relative time with x=0 being the root of the tree and x=1 the tips.

Figure 2. Mapping of time slices (a), regimes (b), taxonomy (c) and geography (d) on the MCC tree of the Andean species of *Macrocarpaea* that have morphological data (55 species) that were used to build the different model of trait evolution. a) The mapping depict part of the tree older than the Pleistocene climatic oscillations (PCO, >2.6Ma) in black from those contemporaneous to the PCO in red. The SHIFT is a two rate Brownian model (BM) of trait evolution were black portions of the tree share the same rate of evolution σ_1 and red portions have a different rate σ_2. The ER (for ecological release) is an Ornstein-Uhlenbeck model (OU) applied to the black part of the tree, followed by a BM applied to the red branches. b) Black branches correspond to species occurring in the upper montane forest (UMF, >1800m) and red branches to species occurring in the lower montane forest (LMF, <1800m). States for internal branches were estimated using maximum likelihood methods (ML, see Methods for details). The BMM Regime is a two-rate regime specific BM model, were black branches have a rate σ_1 and red branches have another rate σ_2. The OUM Regime is a two optimum regime specific OU model, were black branches have an optimum θ_1 and red branches have another optimum θ_2. c) Red branches correspond to the clade of the *micrantha* alliance and black branches correspond to parts of the tree outside the *micrantha* alliance. The BMM Clade is a two rate clade specific BM model (black= σ_1, red= σ_2) and the OUM Clade is a two optima clade specific OU model (black= θ_1, red= θ_2). d) Black, red and green branches correspond to species occurring in the northern Andes, the Amotape-Huancabamba zone and central Andes respectively. States for internal branches were inferred using ML methods. The BMM Geo is a three rates region specific BM model (black= σ_1, red= σ_2, green= σ_3) and the OUM Geo is a three optima region specific OU model (black= θ_1, red= θ_2, green= θ_3).

Figure 3. Phenograms reconstructed using the R package phytools (Revell, 2012) for the 6 traits investigated in the study. The x axes represent the time in My and the y axes the trait values. The method uses a maximum likelihood approach to infer the ancestral value at internal nodes. For each trait, two phenograms are displayed with branch colours representing alternative mappings that were used to fit the two best fitting models (see Tab3. for models scores). Vertical lines on the right of the phenograms represent the trait range value for each mapped state. a) For leaf size the best fitting model is the OUM Clade (see caption of the Fig2. and methods for details concerning the models) followed by the OU1, but here we decided to represent the OUM Regime instead (3rd best model), for discussion purpose. IN and OUT mean inside and outside the *micrantha* alliance respectively. UMF and LMF refers to the upper (>1800m) and lower (<1800m) montane forest respectively. b) For flower size the two best models are the BMM Clade and OU1. c) For plant size, the BMM Geo and the ER are the two best models. North, AHZ, and Centre refer to the northern Andes, the Amotape-Huancabamba zone and the central Andes respectively while >2.6Ma and <2.6Ma refer to part of the trees older and contemporaneous with the Pleistocene climatic oscillations respectively. d) For specific leaf area the OUM Regime and OUM Clade obtain almost equivalent fit. e) For the altitudinal niche the OUM Regime obtained the best fit and is followed by the OUM Clade. f) For the latitudinal niche the best fitting model is the OUM Geo followed by the BMM Geo.

.

Table 1. Loadings, eigenvalues and variance explained by the two firsts principal components from an analysis of 5 morphological traits for 55 Andean *Macrocarpaea* species. Results displayed are from the phylogenetic PCA based on the MCC tree.

	pPC1	pPC2
blade length	-0.96	0.08
blade width	-0.97	0.09
calyx width	-0.40	-0.78
corolla length	-0.28	-0.65
corolla width	-0.11	-0.89
Eigenvalue	0.19	0.05
Proportion of Variance	0.70	0.18
Cumulative proportion	0.70	0.88

Table 2. Loadings, eigenvalues and variance explained by the two firsts principal components from an analysis of the 19 BioClim variables together with 3 geo-position variables for 62 Andean *Macrocarpaea* species. Results displayed are from the phylogenetic PCA based on the MCC tree.

	pPC1	pPC2
LAT	0.06	0.88
LONG	0.06	-0.82
ALT	-0.92	0.11
BIO1	0.89	-0.19
BIO2	-0.48	-0.54
BIO3	-0.03	0.69
BIO4	0.02	-0.82
BIO5	0.86	-0.33
BIO6	0.92	0.09
BIO7	-0.41	-0.77
BIO8	0.88	-0.23
BIO9	0.92	-0.08
BIO10	0.90	-0.24
BIO11	0.89	-0.10
BIO12	0.82	0.07
BIO13	0.71	-0.14
BIO14	0.71	0.26
BIO15	-0.36	-0.46
BIO16	0.71	-0.15
BIO17	0.72	0.26
BIO18	0.70	-0.14
BIO19	0.68	0.30
Eigenvalue	10.67	4.27
Proportion of Variance	0.48	0.19
Cumulative proportion	0.48	0.68

Variables are as follows: LAT= latitude, LONG= longitude, ALT= altitude, BIO1 = annual mean temperature; BIO2 = mean diurnal temperature range [mean of monthly (maximum temperature) minimum temperature)]; BIO3 = isothermality (BIO2⁄BIO7 x 100); BIO4 = temperature seasonality (standard deviation of monthly temperature); BIO5 = maximum temperature of the warmest month; BIO6= minimum temperature of the coldest month; BIO7 = temperature range (BIO6 - BIO5); BIO8 = mean temperature of the wettest quarter; BIO9 = mean temperature of the driest quarter; BIO10 = mean temperature of the warmest quarter; BIO11 = mean temperature of the coldest quarter; BIO12 = annual precipitation; BIO13 = precipitation of the wet- test month; BIO14 = precipitation of the driest month; BIO15 = precipitation seasonality (standard deviation of monthly precipitation); BIO16 = precipitation of the wettest quarter; BIO17 = precipitation of the driest quarter; BIO18 = precipitation of the warmest quarter; BIO19 = precipitation of the coldest quarter.

Table 3. Results from the model comparison analyses for the evolution of the 6 traits investigated in the study. Values displayed are ΔAICc estimated from the MCC tree together with in parenthesis the 10%, 50%, 90% quantiles of the distribution of ΔAICc estimated from the posterior distribution of trees (100 trees). We ask the reader to refer to the main text for a detailed explanation of the 11 models compared. For each trait, the model(s) that obtain the best fit is (are) in bold.

Model	Leaf size	Flower size	Plant Size	SLA	Altitudinal niche	Latitudinal niche
BM1	14.25 (8.13;15.36;23.30)	13.97 (5.96;12.61;17.78)	4.40 (3.23;5.04;7.20)	9.99 (5.72;8.34;20.27)	77.93 (63.48;77.38;100.04)	79.40 (60.24;80.58;91.53)
EB	16.49 (10.37;17.60;25.54)	16.21 (8.20;14.85;20.02)	6.64 (5.47;7.28;9.44)	12.45 (8.18;10.80;22.73)	80.14 (65.69;79.59;102.26)	81.61 (62.45;82.79;93.74)
OU1	7.66 (0.00;7.68;11.42)	7.03 (0.50;5.31;9.07)	6.73 (4.31;6.07;7.85)	9.17 (6.37;8.06;11.78)	59.96 (55.68;59.31;63.19)	82.57 (58.71;82.73;93.85)
SHIFT	9.60 (3.27;11.13;16.00)	9.06 (2.09;7.81;11.88)	2.16 (0.97;2.94;5.16)	9.36 (5.17;7.46;17.95)	70.22 (57.76;69.77;90.41)	81.17 (60.62;82.01;92.99)
ER	9.378 (2.95;10.76;15.97)	9.06 (1.36;7.82;11.96)	1.84 (0.73;2.59;4.33)	8.75 (4.90;7.01;17.72)	70.22 (57.76;69.77;90.41)	81.47 (61.31;82.60;93.43)
BMM Regime	14.01 (7.16;14.42;20.22)	14.37 (4.88;13.31;17.19)	5.58 (4.51;6.35;8.81)	10.66 (7.02;9.51;15.76)	58.95 (51.94;61.23;74.03)	76.40 (59.06;78.32;88.14)
OUM Regime	9.04 (1.22;8.60;12.62)	9.29 (2.74;7.45;11.20)	9.06 (6.61;8.39;10.07)	**1.74 (0.00;0.00;4.28)**	**0.00 (0.00;0.00;0.00)**	81.46 (57.20;80.35;91.47)
BMM Clade	15.77 (8.04;15.69;22.33)	**0.00 (0.00;0.00;1.05)**	3.74 (3.16;6.39;8.67)	11.46 (7.71;10.15;17.33)	80.13 (63.68;79.59;101.79)	81.60 (59.39;81.21;92.58)
OUM Clade	**0.00 (0.00;0.00;1.72)**	9.36 (0.69;7.48;11.19)	8.59 (6.35;8.12;9.80)	**0.00 (0.00;0.091;9.24)**	50.92 (43.44;50.43;58.57)	84.80 (60.10;84.43;95.96)
BMM Geo	9.99 (2.31;10.06;16.67)	14.45 (6.80;12.38;19.10)	**0.00 (0.00;0.00;0.00)**	12.35 (8.10;10.63;21.75)	76.03 (65.59;75.68;94.45)	33.94 (30.36;42.10;57.67)
OUM Geo	11.28 (3.36;11.24;14.90)	10.13 (3.33;8.11;11.80)	7.04 (4.32;6.37;8.47)	11.80 (8.93;10.70;14.43)	60.40 (56.80;59.81;63.16)	**0.00 (0.00;0.00;0.00)**

Table 4. Phylogenetic half-life and stationary variance for the 6 traits considered in the study estimated under the evolutionary model OU1 and OUM if for a particular trait an OUM model obtained the best fit. For leaf size, the OUM is the OUM Clade. For SLA and altitudinal niche, the OUM is the OUM Regime. For the latitudinal niche the OUM is the OUM Geo. The values displayed are estimated from the MCC tree together with in parenthesis the 10%, 50%, 90% quantiles of the distribution of ΔAICc estimated from the posterior distribution of trees (100 trees). The phylogenetic half-life ($t_{1/2} = \ln 2/\alpha$) is the time in millions of years (Ma) taken for an OU process to erase half the phylogenetic covariance between sister taxa. A half-life equal to the age of the MRCA of a clade (here 7.2Ma) indicates a phylogenetic signal equivalent to the expectation under a BM model of evolution (strong signal) while a half-life smaller than the age of the youngest node in a clade (here 0.36Ma) indicates no phylogenetic correlations (no signal). The stationary variance ($V_y = \sigma/2*\alpha$) quantify the relative strength of drift and directional selection around an adaptive optimum.

Trait	half-life OU1	stationary variance OU1	half-life OUM	stationary variance OUM
Leaf size	1.938 (1.294;1.635;2.415)	0.066 (0.043;0.094;0.151)	0.312 (0.054;0.266;1.47)	1.69 (0.106;2.389;56.093)
Flower size	1.797 (1.340;1.729;2.359)	0.019 (0.011;0.021;0.034)	NA	NA
Plant Size	3.658 (2.266;2.925;4.385)	0.009 (0.006;0.013;0.021)	NA	NA
SLA	1.951 (1.117;1.978;2.557)	0.028 (0.017;0.028;0.089)	0.022 (0.014;0.888;1.828)	123.654 (0.026;0.094;316.945)
Altitudinal niche	1.011 (0.086;0.920;1.497)	9.483 (5.454;10.156;769.061)	0.027 (0.008;0.020;0.194)	4676.83 (88.81;8948.43;39845.79)
Latitudinal niche	6.758 (2.764;5.435;8.854)	0.279 (0.195;0.368;0.858)	0.139 (0.033;0.125;0.563)	64.658 (4.228;81.014;1139.458)

Figure 1.

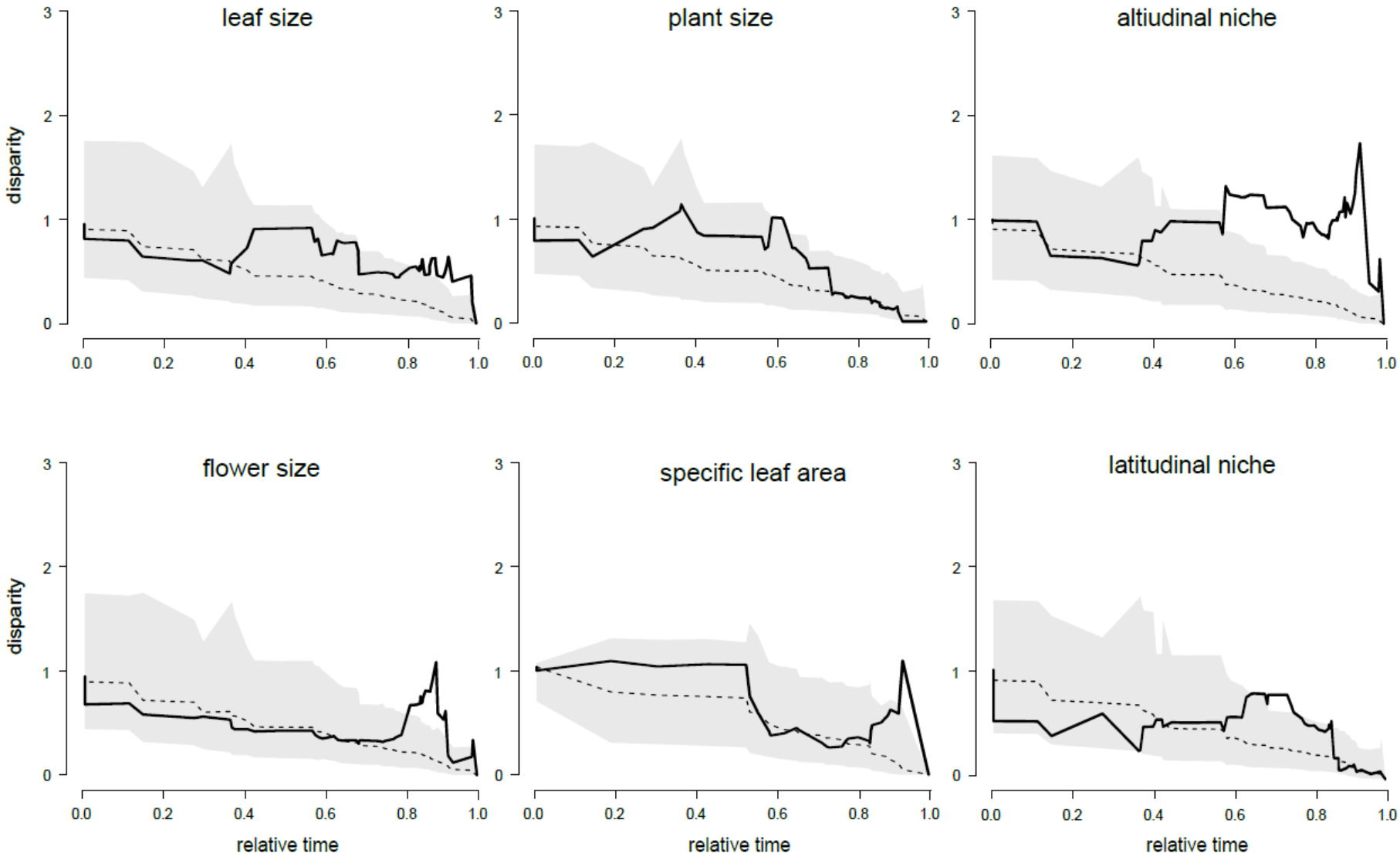

leaf size
plant size
altiudinal niche
flower size
specific leaf area
latitudinal niche
disparity
disparity
relative time
relative time
relative time
0
1
2
3
0.0
0.2
0.4
0.6
0.8
1.0

Figure 2.

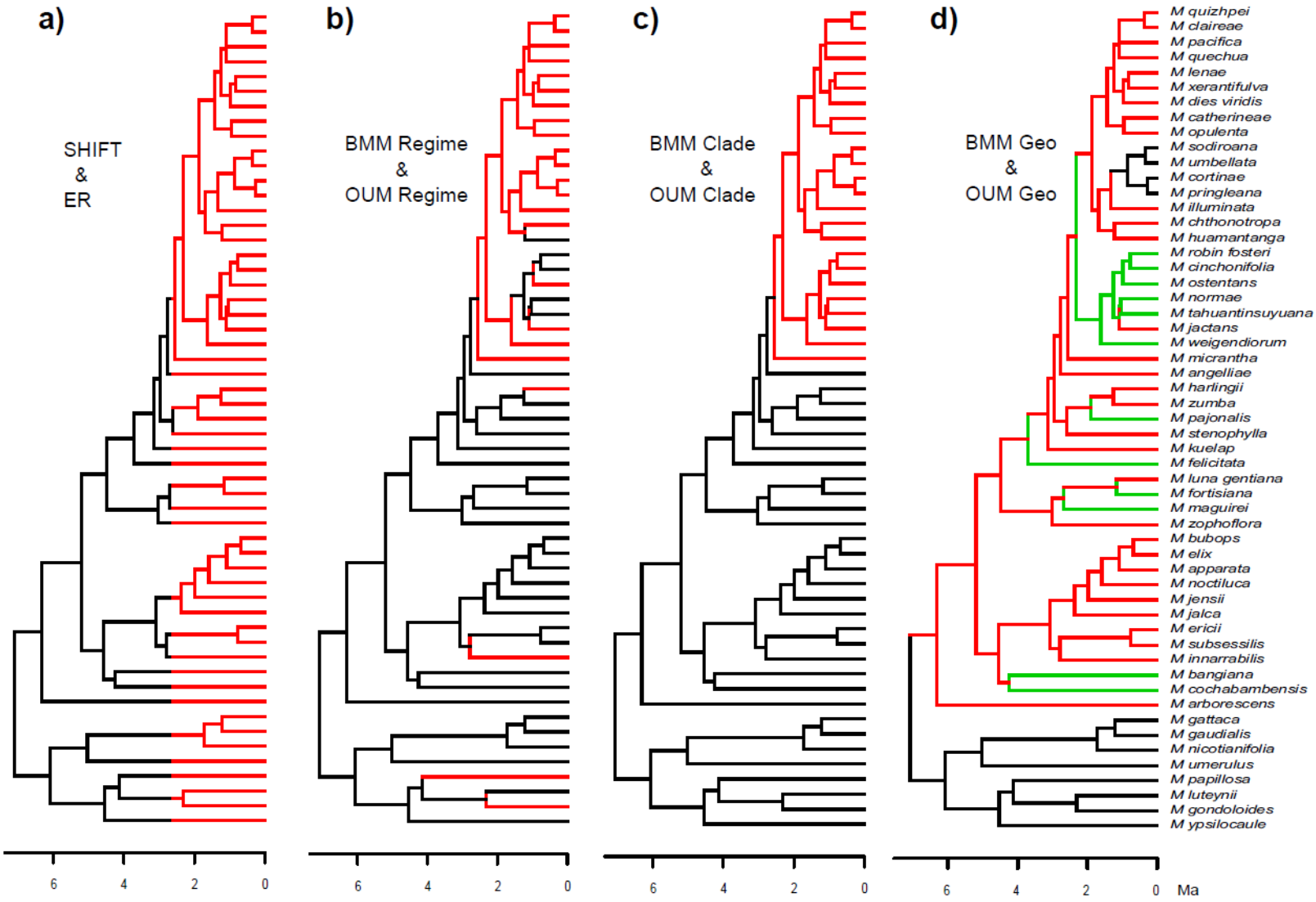
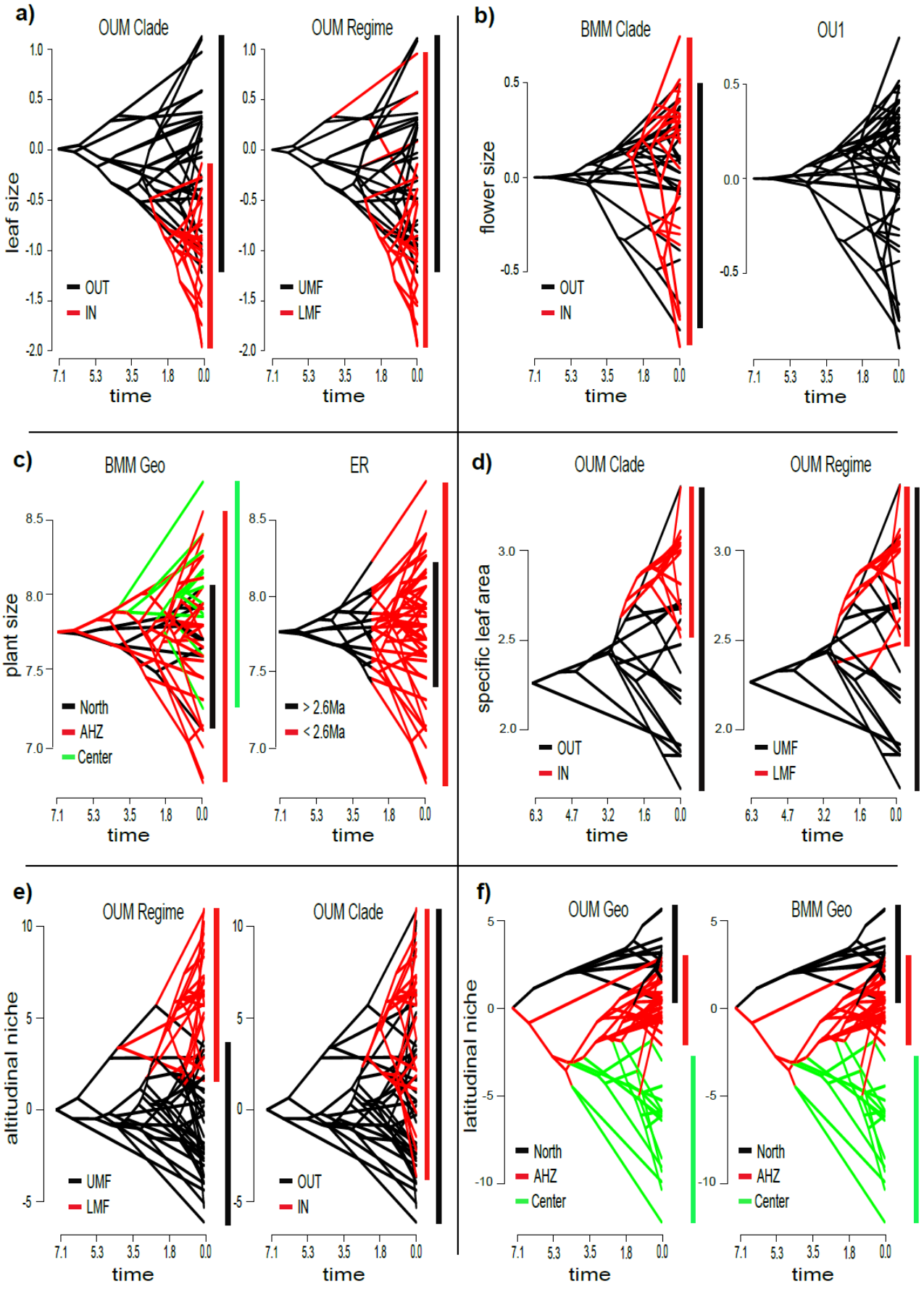

a)
SHIFT
&
ER
b)
BMM Regime
&
OUM Regime
c)
BMM Clade
&
OUM Clade
d)
BMM Geo
&
OUM Geo
M quizhpei
M claireae
M pacifica
M quechua
M lenae
M xerantifulva
M dies viridis
M catherineae
M opulenta
M sodiroana
M umbellata
M cortinae
M pringleana
M illuminata
M chthonotropa
M huamantanga
M robin fosteri
M cinchonifolia
M ostentans
M normae
M tahuantinsuyuana
M jactans
M weigendiorum
M micrantha
M angelliae
M harlingii
M zumba
M pajonalis
M stenophylla
M kuelap
M felicitata
M luna gentiana
M fortisiana
M maguirei
M zophoflora
M bubops
M elix
M apparata
M noctiluca
M jensii
M jalca
M ericii
M subsessilis
M innarrabilis
M bangiana
M cochabambensis
M arborescens
M gattaca
M gaudialis
M nicotianifolia
M umerulus
M papillosa
M luteynii
M gondoloides
M ypsilocaule
6
4
2
0
Ma

Figure 3.


a)
OUM Clade
OUM Regime
leaf size
OUT
IN
UMF
LMF
b)
BMM Clade
OU1
flower size
OUT
IN
c)
BMM Geo
ER
plant size
North
AHZ
Center
> 2.6Ma
< 2.6Ma
d)
OUM Clade
OUM Regime
specific leaf area
OUT
IN
UMF
LMF
e)
OUM Regime
OUM Clade
altitudinal niche
UMF
LMF
OUT
IN
f)
OUM Geo
BMM Geo
latitudinal niche
North
AHZ
Center
time

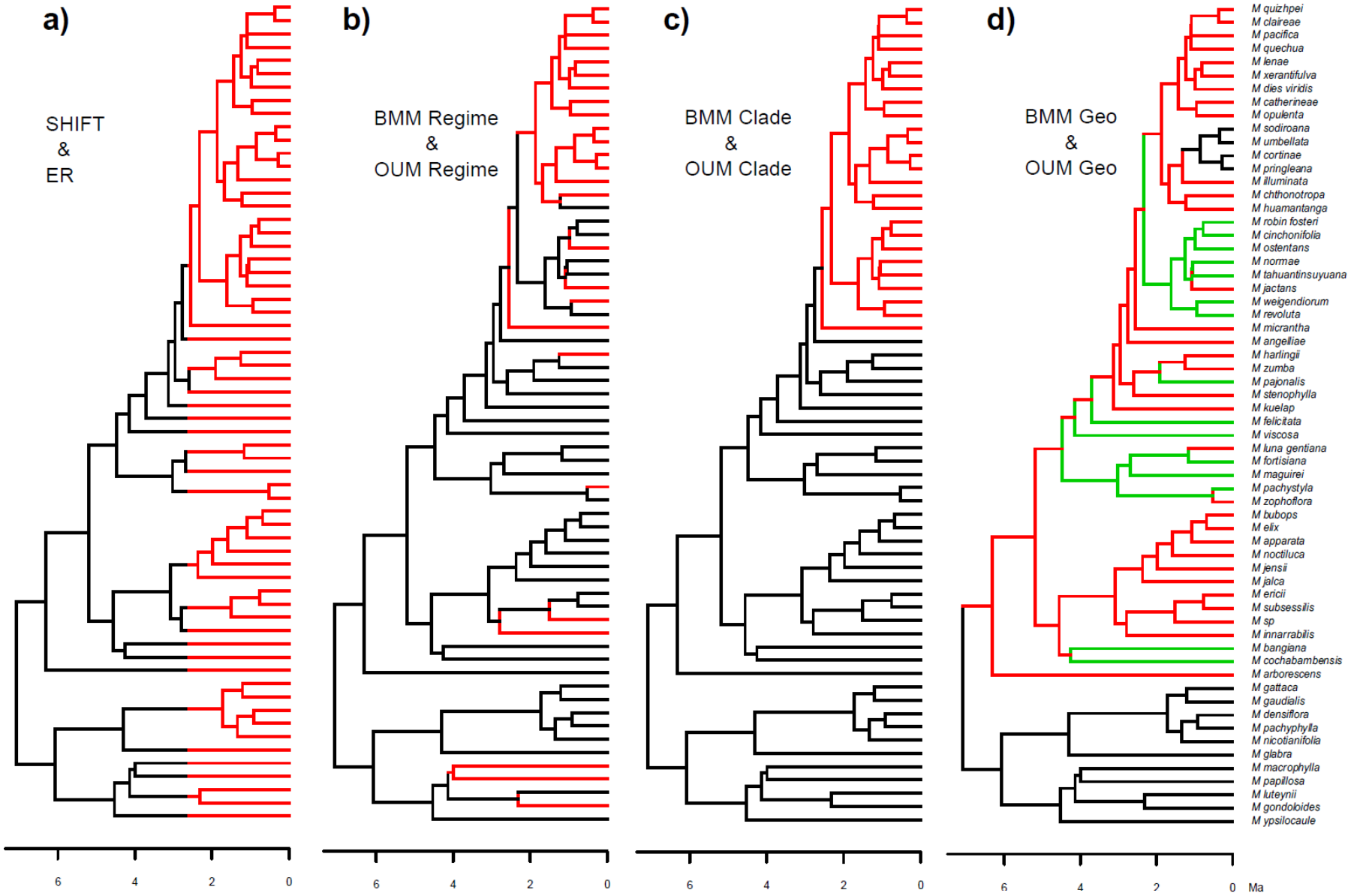


Figure S1. Mapping of time slices (a), regimes (b), taxonomy (c) and geography (d) on the MCC tree of the Andean species of *Macrocarpaea* that have climatic data (62 species). See Fig2. caption for further details.

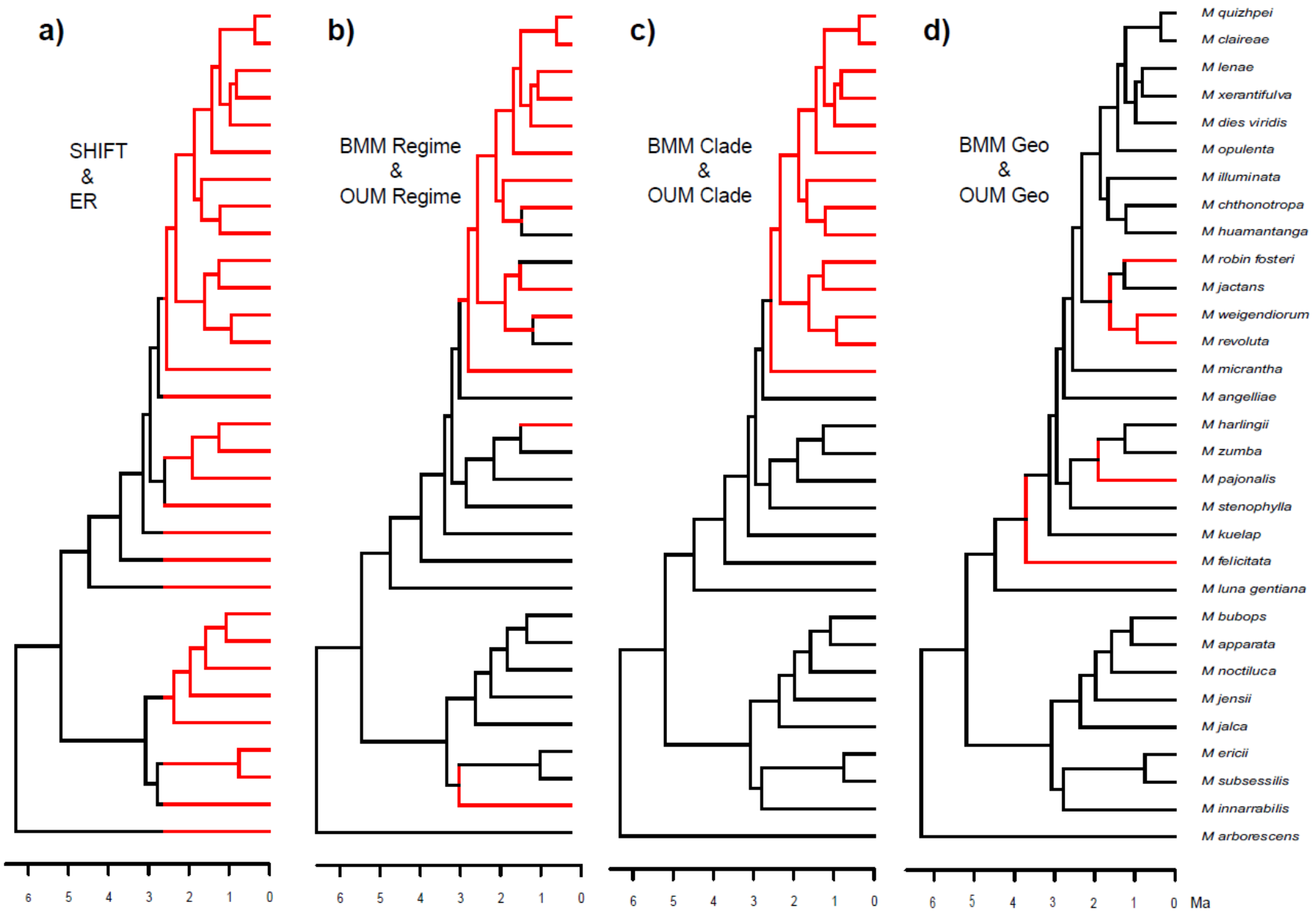


Figure S2. Mapping of time slices (a), regimes (b), taxonomy (c) and geography (d) on the MCC tree of the Andean species of *Macrocarpaea* that have specific leaf area data (31 species). See Fig2. caption for further details.

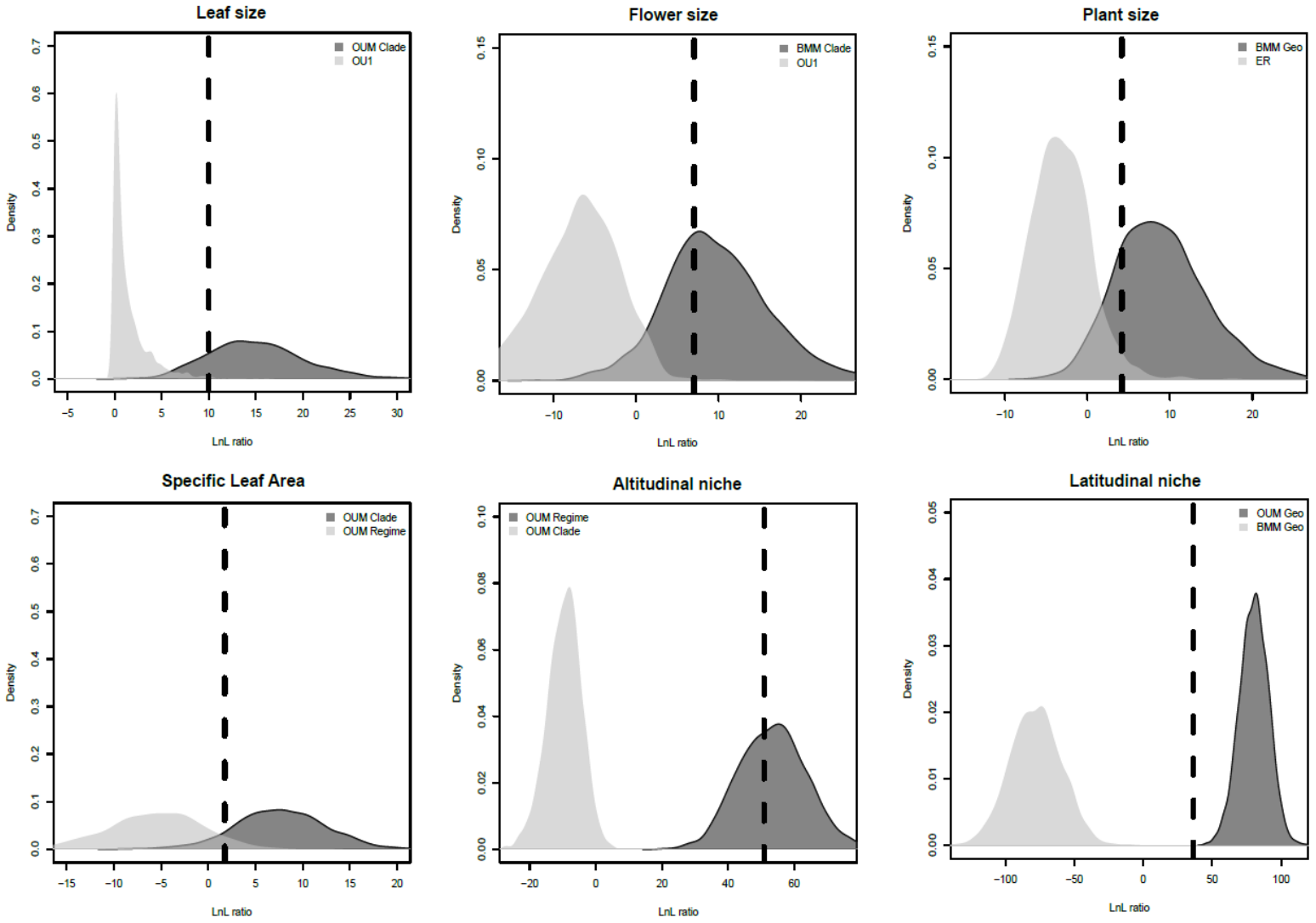


Figure S3. Distributions of the likelihood ratio statistic (δ) for the two best fitting models for each of 6 traits investigated in the study. In each case, the lighter distribution shows the distribution of δ values obtained by bootstrapping under the less fit of the two models, whereas the darker distribution shows the distribution under the best fitting model. A total of 2000 replicates are used for each distribution. The dashed vertical line indicates the observed value of δ when the models are fit to the real dataset. The models compared for each traits are indicated in the legend at the top right or the top left corner of individual figures

Table S4. Results from the evolutionary model comparison for the six traits investigated in the study and corresponding parameter estimates. The values displayed are estimated on the MCC tree together with in parenthesis, the 10%, 50%, and 90% quantiles of the distribution estimated from the posterior distribution trees (100 trees). X_0 is the estimated root value under BM1, SHIFT and BMM models. The θ_1 is the unique optimum of the OU1 and the ER model or the first optimum of the OUM models. The σ_1 is the unique evolutionary rate of the BM1, ER and OUM models or the first evolutionary rate of the SHIFT and BMM models. The r is the variation through time parameter of the EB model. The α is the selection strength parameter of the OU1, ER and OUM models. The σ_2 and σ_3 are the second and third drifting rate parameters of the SHIFT and BMM models. The θ_2 and θ_3 are the second and third adaptive optima of the OUM models. NA indicates that the corresponding parameter is not included in the model.

Trait	Model	LnL	AICc	delta AICc	X0	θ1	σ1	r	α	σ2	σ3	θ2	θ3	half-life	stationary variance
Leaf shape	BM	-57.932 (-63.965; -60.118 ;-57.455)	120.095 (119.14; 124.467; 132.161)	14.251 (8.130; 15.360; 23.303)	0 (0;0;0)	NA	0.19 (0.156; 0.22; 0.266)	NA	NA	NA	NA	NA	NA	NA	NA
	EB	-57.932 (-63.965; -60.118; -57.455)	122.335 (121.38; 126.707; 134.401)	16.491 (10.371; 17.6; 25.543)	0 (0;0;0)	NA	0.19 (0.156; 0.22; 0.266)	0 (0;0;0)	NA	NA	NA	NA	NA	NA	NA
	OU	-53.519 (-56.626; -54.823; -53.248)	113.508 (112.968; 116.117; 119.723)	7.664 (0; 7.686; 11.428)	NA	-0.15 (-0.226; 0.164; 0.196)	0.367 (0.305; 0.446; 0.568)	NA	0.358 (0.287;0.424;0.536)	NA	NA	NA	NA	1.938 (1.294; 1.635; 2.415)	0.066 (0.043; 0.094; 0.151)
	SHIFT	-54.484 (-59.978; -56.582; -54.329)	115.439 (115.129; 119.635; 126.426)	9.595 (3.273; 11.139; 15.998)	-0.187 (-0.24; 0.161; 0.205)	NA	0 (0; 0; 0.009)	NA	NA	0.223 (0.204; 0.249; 0.295)	NA	NA	NA	NA	NA
	ER	-54.375 (-59.979; -56.58; -53.962)	115.221 (114.394; 119.631; 126.428)	9.377 (2.946; 10.755; 15.965)	NA	-0.178 (-0.237; 0.162; 0.21)	0.217 (0.188; 0.245; 0.295)	NA	4.145 (1.102; 146.239; 5364.064)	NA	NA	NA	NA	NA	NA
	BMM Regime	-56.692 (61.505; -58.593; -55.93)	119.854 (118.331; 123.657; 129.481)	14.01 (7.157;14.424;20.216)	0.006 (-0.008; 0.003; 0.013)	NA	0.142 (0.112; 0.151; 0.193)	NA	NA	0.263 (0.211; 0.318; 0.403)	NA	NA	NA	NA	NA
	OUM Regime	-53.041 (-56.042; -54.175; -52.924)	114.883 (114.648; 117.15;120.883)	9.039 (1.221; 8.597; 12.621)	NA	-0.101 (-0.181; -0.117; 0.15)	0.38 (0.322; 0.479; 0.645)	NA	0.386 (0.322; 0.486; 0.636)	NA	NA	-0.507 (-0.642; -0.499; 0.565)	NA	1.798 (1.09; 1.425; 2.155)	0.073 (0.052; 0.116; 0.207)
	BMM Clade	-57.572 (61.717;-58.993;-56.597)	121.615 (119.665; 124.456; 129.905)	15.771 (8.035; 15.685; 22.33)	0.004 (-0.007 ;0.003; 0.009)	NA	0.162 (0.12; 0.166; 0.215)	NA	NA	0.228 (0.184; 0.284; 0.407)	NA	NA	NA	NA	NA

	OUM Clade	-48.522 (-54.923;-48.946;-48.295)	105.844 (105.39;106.693;118.646)	0 (0;0;1.715)	NA	-0.103 (-0.185;-0.113;0.148)	1.523 (0.453;1.817;8.705)	NA	2.22 (0.471;2.609;12.745)	NA	NA	-1.007 (-1.073;-0.999;1.048)	NA	0.312 (0.054;0.266;1.47)	1.69 (0.106;2.389;56.093)
	BMM Region	-53.517 (-58.125;-55.149;-52.64)	115.833 (114.079;119.098;125.049)	9.989 (2.311;10.057;16.666)	-0.005 (-0.079;0.01;0.104)	NA	0.25 (0.18;0.28;0.441)	NA	NA	0.248 (0.199;0.277;0.352)	0.016 (0.008;0.017;0.026)	NA	NA	NA	NA
	OUM Region	-52.948 (-55.68;-54.07;-52.691)	117.12 (116.606;119.365;122.585)	11.276 (3.359;11.238;14.896)	NA	0.056 (-0.097;0.013;0.085)	0.389 (0.319;0.466;0.601)	NA	0.402 (0.325;0.475;0.613)	NA	NA	-0.234 (-0.304;-0.229;0.28)	-0.416 (-0.508;-0.427;0.459)	1.725 (1.13;1.458;2.134)	0.078 (0.051;0.111;0.184)
Flower shape	BM	-20.546 (-24.302;-21.513;-17.388)	45.322 (39.007;47.257;52.835)	13.972 (5.957;12.607;17.777)	0 (0;0;0)	NA	0.049 (0.039;0.051;0.061)	NA	NA	NA	NA	NA	NA	NA	NA
	EB	-20.546 (-24.302;-21.513;-17.388)	47.562 (41.247;49.497;55.075)	16.212 (8.197;14.847;20.017)	0 (0;0;0)	NA	0.049 (0.039;0.051;0.061)	0 (0;0;0)	NA	NA	NA	NA	NA	NA	NA
	OU	-15.956 (-17.79;-16.176;-14.373)	38.382 (35.216;38.822;42.05)	7.032 (0.496;5.312;9.071)	NA	0.029 (-0.007;0.022;0.04)	0.099 (0.077;0.103;0.133)	NA	0.386 (0.294;0.401;0.517)	NA	NA	NA	NA	1.797 (1.34;1.729;2.359)	0.019 (0.011;0.021;0.034)
	SHIFT	-16.969 (-20.136;-17.558;-14.735)	40.408 (35.941;41.586;46.742)	9.058 (2.086;7.814;11.881)	0.047 (-0.005;0.036;0.053)	NA	0 (0;0;0)	NA	NA	0.057 (0.05;0.059;0.068)	NA	NA	NA	NA	NA
	ER	-16.969 (-20.208;-17.558;-13.991)	40.408 (34.452;41.586;46.886)	9.058 (1.361;7.815;11.956)	NA	0.047 (-0.008;0.034;0.053)	0.057 (0.048;0.058;0.068)	NA	206114.844 (1.461;1903.144;55219.122)	NA	NA	NA	NA	NA	NA
	BMM Regime	-19.623 (22.927;-20.156;-16.582)	45.717 (39.635;46.783;52.325)	14.367 (4.88;13.308;17.194)	0 (-0.003;0;0.002)	NA	0.038 (0.031;0.039;0.048)	NA	NA	0.066 (0.049;0.068;0.087)	NA	NA	NA	NA	NA
	OUM Regime	-15.918 (-17.731;-16.143;-14.319)	40.636 (37.439;41.086;44.262)	9.286 (2.735;7.447;11.203)	NA	0.021 (-0.008;0.011;0.028)	0.1 (0.079;0.105;0.14)	NA	0.396 (0.299;0.409;0.535)	NA	NA	0.079 (-0.01;0.077;0.141)	NA	1.75 (1.297;1.696;2.318)	0.02 (0.012;0.022;0.036)
	BMM Clade	-12.44 (-17.114;-13.667;-11.053)	31.35 (28.577;33.805;40.699)	0 (0;0;1.046)	0.001 (-0.013;0;0.002)	NA	0.019 (0.015;0.02;0.031)	NA	NA	0.089 (0.07;0.09;0.114)	NA	NA	NA	NA	NA
	OUM Clade	-15.956 (-17.498;-16.011;-	40.711 (35.938;40.822;43.	9.361 (0.691;7.478;11.1	NA	0.029 (-0.028;0.015;0.032)	0.099 (0.077;0.105;0.13	NA	0.386 (0.294;0.408;0.531)	NA	NA	0.032 (-0.06;0.109;0.477)	NA	1.796 (1.305;1.698;2.35	0.019 (0.012;0.022;0.036)

		13.569)	795)	86)			5)							4)	
	BMM Region	-18.5 (-22.768;-18.968;-14.771)	45.8 (38.342;46.736;54.337)	14.45 (6.799;12.376;19.098)	-0.005 (-0.012;-0.004;0.003)	NA	0.025 (0.017;0.027;0.048)	NA	NA	0.046 (0.031;0.049;0.061)	0.084 (0.054;0.085;0.104)	NA	NA	NA	NA
	OUM Region	-15.129 (-16.527;-15.279;-13.656)	41.483 (38.536;41.783;44.279)	10.133 (3.33;8.106;11.798)	NA	0.038 (0.009;0.042;0.072)	0.113 (0.085;0.126;0.192)	NA	0.486 (0.377;0.534;0.813)	NA	NA	0.097 (0.034;0.081;0.119)	-0.162 (-0.225;-0.179;-0.117)	1.427 (0.853;1.299;1.838)	0.027 (0.016;0.033;0.073)
Plant size	BM	-25.477 (-28.245;-26.688;-25.456)	55.184 (55.142;57.606;60.721)	4.399 (3.226;5.039;7.2)	7.76 (7.745;7.772;7.797)	NA	0.058 (0.049;0.063;0.076)	NA	NA	NA	NA	NA	NA	NA	NA
	EB	-25.477 (-28.245;-26.688;-25.456)	57.424 (57.382;59.846;62.961)	6.639 (5.466;7.279;9.439)	7.76 (7.745;7.772;7.797)	NA	0.058 (0.049;0.063;0.076)	0 (0;0;0)	NA	NA	NA	NA	NA	NA	NA
	OU	-25.524 (-27.235;-26.21;-25.273)	57.519 (57.016;58.891;60.94)	6.734 (4.308;6.07;7.848)	NA	7.768 (7.757;7.776;7.793)	0.093 (0.079;0.11;0.138)	NA	0.19 (0.158;0.237;0.306)	NA	NA	NA	NA	3.658 (2.266;2.925;4.385)	0.009 (0.006;0.013;0.021)
	SHIFT	-23.235 (-25.802;-24.518;-23.186)	52.94 (52.843;55.506;58.075)	2.155 (0.97;2.942;5.162)	7.756 (7.753;7.766;7.789)	NA	0 (0;0;0)	NA	NA	0.072 (0.069;0.077;0.087)	NA	NA	NA	NA	NA
	ER	-23.078 (-25.702;-24.249;-22.962)	52.627 (52.395;54.969;57.875)	1.842 (0.726;2.588;4.334)	NA	7.775 (7.758;7.774;7.795)	0.066 (0.058;0.074;0.085)	NA	1.067 (0.533;2.218;4645.89)	NA	NA	NA	NA	NA	NA
	BMM Regime	-24.947 (-27.927;-26.238;-24.952)	56.365 (56.375;58.946;62.324)	5.58 (4.509;6.346;8.81)	7.762 (7.744;7.772;7.8)	NA	0.067 (0.055;0.07;0.084)	NA	NA	0.044 (0.039;0.052;0.065)	NA	NA	NA	NA	NA
	OUM Regime	-25.523 (-27.21;-26.198;-25.268)	59.846 (59.337;61.195;63.219)	9.061 (6.606;8.386;10.07)	NA	7.766 (7.758;7.782;7.799)	0.093 (0.079;0.11;0.138)	NA	0.188 (0.155;0.238;0.309)	NA	NA	7.788 (7.679;7.745;7.824)	NA	3.683 (2.244;2.915;4.458)	0.009 (0.007;0.013;0.021)
	BMM Clade	-24.028 (-27.992;-25.919;-24.316)	54.526 (55.102;58.308;62.455)	3.741(3.161;6.378;8.666)	7.762 (7.747;7.773;7.798)	NA	0.074 (0.052;0.069;0.089)	NA	NA	0.038 (0.036;0.052;0.071)	NA	NA	NA	NA	NA
	OUM Clade	-25.285 (-27.107;-26.057;-25.087)	59.371 (58.974;60.914;63.015)	8.586 (6.351;8.123;9.801)	NA	7.756 (7.746;7.773;7.797)	0.092 (0.078;0.107;0.137)	NA	0.189 (0.154;0.236;0.316)	NA	NA	8.2 (7.526;8.041;8.271)	NA	3.677 (2.197;2.943;4.49)	0.009 (0.006;0.012;0.022)
	BMM	-20.992 (-	50.785	0 (0;0;0)	7.746	NA	0.014	NA	NA	0.078	0.047	NA	NA	NA	NA

	Region	23.552;-21.806;-20.677)	(50.154;52.412;55.904)		(7.735;7.759;7.787)		(0.012;0.015;0.06)			(0.02;0.084;0.106)	(0.035;0.044;0.056)				
	OUM Region	-23.299 (-24.884;-23.933;-23.023)	57.822 (57.271;59.09;60.993)	7.037 (4.32;6.371;8.466)	NA	7.695 (7.691;7.706;7.72)	0.102 (0.091;0.122;0.156)	NA	0.267 (0.24;0.327;0.431)	NA	NA	7.688 (7.665;7.686;7.706)	8.36 (8.226;8.287;8.349)	2.593 (1.607;2.121;2.891)	0.014 (0.011;0.02;0.033)
SLA	BM	-19.762 (-24.903;-19.303;-17.794)	43.953 (40.017;43.034;54.234)	9.993 (5.723;8.335;20.274)	2.26 (2.234;2.262;2.292)	NA	0.079 (0.068;0.079;0.121)	NA	NA	NA	NA	NA	NA	NA	NA
	EB	-19.762 (-24.903;-19.303;-17.794)	46.413 (42.477;45.494;56.695)	12.453 (8.184;10.795;22.734)	2.26 (2.234;2.262;2.292)	NA	0.079 (0.068;0.079;0.121)	0 (0;0;0)	NA	NA	NA	NA	NA	NA	NA
	OU	-18.12 (-20.032;-17.98;-16.996)	43.129 (40.88;42.849;46.953)	9.169 (6.372;8.059;11.782)	NA	2.415 (2.374;2.427;2.487)	0.16 (0.126;0.157;0.29)	NA	0.355 (0.271;0.35;0.62)	NA	NA	NA	NA	1.951 (1.117;1.978;2.557)	0.028 (0.017;0.028;0.089)
	SHIFT	-18.215 (-22.511;-17.797;-16.129)	43.318 (39.147;42.484;51.91)	9.358 (5.172;7.46;17.95)	2.434 (2.384;2.439;2.479)	NA	0 (0;0;0)	NA	NA	0.092 (0.081;0.09;0.13)	NA	NA	NA	NA	NA
	ER	-17.909 (-22.393;-17.512;-16.032)	42.707 (38.952;41.912;51.675)	8.747 (4.9;7.005;17.715)	NA	2.409 (2.383;2.421;2.453)	0.087 (0.076;0.087;0.13)	NA	2.44 (1.249;4.63;2363.464)	NA	NA	NA	NA	NA	NA
	BMM Regime	-18.867 (-22.069;-18.737;-17.268)	44.624 (41.424;44.362;51.027)	10.664 (7.02;9.508;15.763)	2.219 (2.197;2.221;2.265)	NA	0.052 (0.046;0.055;0.065)	NA	NA	0.121 (0.088;0.114;0.211)	NA	NA	NA	NA	NA
	OUM Regime	-13.08 (-14.739;-13.105;-11.903)	35.698 (33.345;35.749;39.016)	1.738 (0;0;4.277)	NA	2.278 (2.25;2.293;2.355)	0.16 (0.117;0.166;11.068)	NA	0.545 (0.396;0.559;36.522)	NA	NA	3.075 (2.903;3.054;3.196)	NA	1.272 (0.019;1.239;1.752)	0.044 (0.023;0.047;202.315)
	BMM Clade	-19.266 (-22.537;-18.999;-17.589)	45.421 (42.066;44.887;51.963)	11.461 (7.705;10.151;17.334)	2.242 (2.212;2.247;2.285)	NA	0.059 (0.052;0.062;0.075)	NA	NA	0.108 (0.074;0.1;0.207)	NA	NA	NA	NA	NA
	OUM Clade	-12.211 (-17.589;-12.211;-11.907)	33.96 (33.352;33.96;44.716)	0 (0;0.091;9.242)	NA	2.303 (2.285;2.303;2.378)	7.979 (0.142;0.24;12.766)	NA	30.994 (0.379;0.781;49.459)	NA	NA	2.926 (2.639;2.926;3.183)	NA	0.022 (0.014;0.888;1.828)	123.654 (0.026;0.094;316.945)
	BMM Region	-19.711 (-24.412;-19.208;-17.758)	46.31 (42.405;45.305;55.712)	12.35 (8.097;10.625;21.751)	2.264 (2.239;2.271;2.306)	NA	0.082 (0.07;0.082;0.129)	NA	NA	0.063 (0.047;0.059;0.079)	NA	NA	NA	NA	NA

	OUM Region	-18.109 (-20.022;-17.972;-16.995)	45.757 (43.529;45.482;49.583)	11.797 (8.932;10.703;14.431)	NA	2.42 (2.377;2.422;2.478)	0.16 (0.126;0.158;0.29)	NA	0.354 (0.269;0.349;0.619)	NA	NA	2.353 (2.296;2.432;2.554)	NA	1.96 (1.12;1.985;2.576)	0.028 (0.017;0.028;0.088)
Altitudinal niche	BM	-189.02 (-194.917;-188.098;-182.501)	382.244 (369.204;380.399;394.037)	77.934 (63.479;77.376;100.044)	0 (0;0;0)	NA	10.495 (9.925;10.481;11.234)	NA	NA	NA	NA	NA	NA	NA	NA
	EB	-189.02 (-194.917;-188.098;-182.501)	384.454 (371.415;382.609;396.248)	80.144 (65.689;79.586;102.255)	0 (0;0;0)	NA	10.495 (9.925;10.481;11.234)	0 (0;0;0)	NA	NA	NA	NA	NA	NA	NA
	OU	-178.926 (-184.455;-177.272;-166.789)	364.265 (339.992;360.958;375.323)	59.955 (55.68;59.308;63.185)	NA	1.023 (-1.267;0.984;1.44)	27.662 (22.944;28.229;199.778)	NA	0.686 (0.463;0.753;8.106)	NA	NA	NA	NA	1.011 (0.086;0.92;1.497)	9.483 (5.454;10.156;769.061)
	SHIFT	-184.058 (-188.87;-182.591;-178.79)	374.53 (363.993;371.595;384.155)	70.22 (57.756;69.768;90.413)	0.693 (-0.834;0.465;1.04)	NA	0 (0;0;0)	NA	NA	11.782 (10.968;11.717;12.643)	NA	NA	NA	NA	NA
	ER	-184.059 (-188.872;-182.592;-178.791)	374.532 (363.997;371.597;384.157)	70.222 (57.759;69.772;90.414)	NA	0.692 (-0.834;0.464;1.04)	11.78 (10.968;11.707;12.624)	NA	468.62 (481.586;2821.801;31046.794)	NA	NA	NA	NA	NA	NA
	BMM Regime	-178.425 (-184.83;-176.924;-170.481)	363.265 (347.376;360.261;376.073)	58.955 (51.937;61.225;74.031)	-1.041 (-1.314;-1.038;1.153)	NA	2.855 (1.917;2.758;3.795)	NA	NA	30.94 (24.832;29.841;37.913)	NA	NA	NA	NA	NA
	OUM Regime	-147.804 (-153.633;-145.983;-133.81)	304.31 (276.321;300.668;315.968)	0 (0;0;0)	NA	-1.184 (-1.482;-0.997;1.003)	358.989 (51.191;481.74;905.369)	NA	26.056 (3.567;34.922;88.788)	NA	NA	6.322 (-6.18;5.763;7.153)	NA	0.027 (0.008;0.02;0.194)	4676.828 (88.814;8948.43;39845.788)
	BMM Clade	-189.011 (-194.878;-187.633;-181.968)	384.436 (370.35;381.679;396.17)	80.126 (63.683;79.586;101.789)	-0.004 (-0.057;-0.003;0.042)	NA	10.284 (7.854;10.355;13.451)	NA	NA	10.821 (7.55;10.376;14.995)	NA	NA	NA	NA	NA
	OUM Clade	-173.264 (-181.219;-171.185;-158.177)	355.229 (325.056;351.072;371.14)	50.919 (43.436;50.432;58.568)	NA	0.136 (-0.348;0.134;0.436)	174.485 (29.523;227.041;943.444)	NA	5.57 (0.737;9.066;52.032)	NA	NA	4.649 (-4.332;3.866;6.221)	NA	0.124 (0.013;0.076;0.942)	485.925 (11.326;1154.361;22289.729)
	BMM Region	-185.818 (-	380.338 (361.811;	76.028 (65.589;7	0.407 (-0.395;0.2	NA	6.796 (4.378;6.	NA	NA	6.586 (4.478;6.	26.82 (15.92;26	NA	NA	NA	NA

		191.478;-184.316;-176.555)	377.334; 391.659)	5.684;94.447)	93;0.601)		522;9.247)			172;9.099)	.441;42.061)				
	OUM Region	-176.82 (-182.281;-174.922;-164.439)	364.711 (339.95;360.916;375.633)	60.401 (56.801;59.805;63.156)	NA	2.258 (-2.155;2.084;2.605)	27.214 (23.147;27.87;198.499)	NA	0.728 (0.517;0.793;8.174)	NA	NA	1.73 (-1.99;1.561;2.217)	-1.455 (-2.259;-0.943;0.923)	0.952 (0.085;0.874;1.341)	9.908 (6.145;10.965;848.366)
Regional niche	BM	-160.604 (-166.832;-159.262;-153.487)	325.412 (311.178; 322.728; 337.868)	79.395 (60.24;80.583;91.53)	0 (0;0;0)	NA	4.197 (3.886;4.183;4.42)	NA	NA	NA	NA	NA	NA	NA	NA
	EB	-160.604 (-166.832;-159.262;-153.487)	327.623 (313.388; 324.938; 340.078)	81.606 (62.45;82.793;93.74)	0 (0;0;0)	NA	4.197 (3.886;4.189;4.423)	0 (0;0;0)	NA	NA	NA	NA	NA	NA	NA
	OU	-161.084 (-165.657;-158.684;-152.889)	328.583 (312.192; 323.782; 337.729)	82.566 (58.705;82.731;93.847)	NA	-0.166 (-0.287;-0.125;0.019)	5.442 (4.976;5.63;6.866)	NA	0.103 (0.078;0.128;0.251)	NA	NA	NA	NA	6.758 (2.764;5.435;8.854)	0.279 (0.195;0.368;0.858)
	SHIFT	-160.388 (-165.869;-158.298;-153.133)	327.19 (312.681; 323.01;338.151)	81.173 (60.622;82.01;92.986)	-0.012 (-0.069;0.006;0.11)	NA	2.719 (1.242;2.426;3.145)	NA	NA	4.466 (4.055;4.498;4.971)	NA	NA	NA	NA	NA
	ER	-160.536 (-166.265;-158.669;-153.145)	327.486 (312.704; 323.752; 338.944)	81.469 (61.305;82.597;93.427)	NA	-0.111 (-0.276;-0.087;0)	4.303 (3.991;4.318;4.69)	NA	0.064 (0.005;0.081;0.228)	NA	NA	NA	NA	NA	NA
	BMM Regime	-158.001 (-164.929;-157.228;-151.481)	322.417 (309.376; 320.871; 336.273)	76.4 (59.06;78.32;88.14)	0.047 (-0.032;0.062;0.218)	NA	5.783 (3.381;5.337;6.32)	NA	NA	1.802 (1.553;2.328;5.327)	NA	NA	NA	NA	NA
	OUM Regime	-159.386 (-164.122;-157.02;-150.863)	327.473 (310.428; 322.741; 336.946)	81.456 (57.201;80.347;91.469)	NA	-0.845 (-1.217;-0.788;-0.402)	5.372 (4.927;5.594;6.991)	NA	0.119 (0.095;0.147;0.317)	NA	NA	8.169 (1.575;6.14;10.263)	NA	5.837 (2.183;4.731;7.329)	0.319 (0.232;0.422;1.144)
	BMM Clade	-160.603 (-	327.621 (312.851;	81.604 (59.389;8	0 (-	NA	4.224 (3.395;4.	NA	NA	4.155 (3.118;4.	NA	NA	NA	NA	NA

		166.763;-158.377;-153.219)	323.168; 339.94)	1.213;92.578)	.021)		12;4.981)			182;5.056)					
	OUM Clade	-161.056 (-165.647;-158.499;-152.721)	330.814 (314.144; 325.7;339.996)	84.797 (60.1;84.427;95.963)	NA	-0.142 (-0.443;-0.164;0.014)	5.438 (4.975;5.616;6.875)	NA	0.103 (0.078;0.127;0.253)	NA	NA	-2.441 (-3.354;0.419;9.822)	NA	6.751 (2.741;5.477;8.902)	0.279 (0.19;0.351;0.866)
	BMM Region	-135.627 (-154.669;-139.755;-126.986)	279.955 (262.674; 288.211; 318.039)	33.938 (30.363;42.101;57.667)	1.332 (0.822;1.455;2.013)	NA	1.615 (0.753;1.856;3.408)	NA	NA	0.47 (0.372;0.729;2.548)	26.214 (13.447;23.991;28.852)	NA	NA	NA	NA
	OUM Region	-117.473 (-127.418;-117.416;-107.78)	246.017 (226.631; 245.903; 265.907)	0 (0;0;0)	NA	3.011 (1.104;2.986;3.667)	25.882 (6.738;28.956;109.798)	NA	4.996 (1.231;5.563;20.82)	NA	NA	0.29 (-0.219;0.598;3.02)	-7.237 (-7.833;-6.787;-5.365)	0.139 (0.033;0.125;0.563)	64.658 (4.228;81.014;1139.458)

CHAPTER 3: Exploration of the population genetic structure and genetic diversity of a *Macrocarpaea* species complex from the tropical Andes

AUTHORS NAMES: Vieu Julien C. & Grant Jason R.

AUTHORS AFFILIATIONS: Institut de Biologie, Université de Neuchâtel, Rue Emile Argand 11, 2000 Neuchâtel, Switzerland;

ABSTRACT

The Pleistocene climatic oscillations (PCO) are thought to have profoundly affected species demography around the world. While the effect of the PCO on population dynamics at temperate latitudes is relatively well understood, considerable gaps remain concerning its impact on the biodiversity of Neotropical mountains.

Here, we use amplified fragment length polymorphism molecular makers (AFLP's) to investigate the phylogeography of 12 species of shrubs to small trees endemic to the tropical Andes that belong to the same species complex in the genus *Macrocarpaea* (Gentianaceae). We show that some species identified on the basis of cryptically morphological differences, display complex and potentially reticulated relationships, requiring further examination. We also show that populations of a species endemic to a single valley connected to the dry system of the Rio Marañón in northern Peru have low levels of gene diversity. We suggest that this is due to a recent demographic bottleneck resulting from the contraction of the montane wet forests into refugia as consequence of the expansion of the dry system into the valley during the glacial cycles of the PCO. This hypothesis needs further study, but its confirmation would imply that different valleys of the Andes might have responded in very different manners to the PCO.

KEYWORDS

AFLP, cryptic species, phylogeography, Pleistocene climatic oscillations, refugia

INTRODUCTION

Mountain ranges are often regarded as "species pumps", in the sense that they are speciation centres that significantly contribute to enhance species diversity around the globe (Smith *et al.*, 2007; Fjeldså *et al.*, 2012). Mountains are thought to favour speciation through their high overall physiographic heterogeneity that facilitates both allopatric divergence between populations occurring on different montane systems and adaptive divergence along elevational gradients within montane systems (Moritz *et al.*, 2000, Wollenberg et *al.*, 2008). Numerous studies document radiations (rapid species divergence) of taxa taking place almost immediately after the colonization of a particular montane region (Hughes & Eastwood, 2006; Linder *et al.*, 2014, Schwery *et al.*, 2014; Hughes & Atchison, 2015). While the evidence for the contribution of adaptive divergence remains nuanced (Elias *et al.*, 2009), it is clear that allopatric speciation plays a determinant role in driving many montane radiations (Fjeldså *et al.*, 2012, Hutter *et al.*, 2014, Vieu *et al.*, in review (see chapter 1)).

Climate change has also been proposed as potential important driver of speciation in montane regions, particularly in non vagile organisms like plants (Hewitt, 1996, 2004). It is relatively well acknowledged that the Pleistocene climatic oscillations (PCO, 2.6Ma-15.000Ya) have profoundly affected the demographic evolution of montane elements by repeatedly moving down species during the glacial periods and up during interglacial periods (Hewitt, 1996). At temperate latitudes, the PCO is generally considered as a time of range expansion and secondary contact in the lowlands for montane cold-adapted taxa (Hewitt, 1996; Crespi *et al.*, 2003; Schmitt, 2007) but examples of divergence or at least isolation in peripheral refugia valleys have also been documented (Knowles, 2000; Schönswetter *et al.*, 2005). In contrast, the role of the PCO on Neotropical montane biodiversity has received much less attention.

Among the Neotropical mountains, the tropical Andes are of particular interest because they have exceptional levels of diversity and endemism (Orme *et al.*, 2005). It has been proposed that the PCO significantly contributed to the overall diversity found in the region (Gentry, 1982; Rull, 2008) but currently its potential consequences remain poorly understood. The dramatic altitudinal lowering of the upper forest line (UFL) during glacial periods is well documented by the palynological record sampled across the tropical Andes but yet shows considerable variations between regions, but also often inside regions (but see Brunschön & Behling, 2010, Flantua *et al.*, 2014).

What is less known and is at the centre of an intense debate is whether the PCO affected the level of precipitation in the tropical Andes and in the Neotropics in general (Ramírez-Barahona & Eguiarte, 2013). On one hand, the tenant of the "refugia" model proposes that glacial periods were characterised by an increase in aridity (van der Hammen & Hooghiemstra, 2000; Mourguiart & Ledru, 2003). On the other hand, the opponent of this model considers that the PCO did not change the overall precipitation and that glacial periods remained wet (Colinvaux et al., 2000; Baker et *al.*, 2003). The "dry refugia" model predicts that during glacial periods, montane forests populations were compressed into refugia as a consequence of both the lowering of the UFL resulting from cooling and the expansion of dry conditions on montane foothills (Ramírez-Barahona & Eguiarte, 2013). Under this scenario it is expected that the action of genetic drift in small populations separated into distinct refugia would result in an overall loss of genetic diversity but also in a potential genetic differentiation between refugia. The periods of rapid re-expansion of the populations during the interglacial period could have also induced a gradient of loss of genetic diversity between the source populations in the refugia and the expansion front. On the contrary, the moist forests model do not predict the compression by the foothills during the glacial periods but instead the migration of montane elements down in the lowlands, resulting in extensive secondary contact between populations, followed by subsequent upslope migration and re-fragmentation during the inter-glacial periods. Under this scenario, the PCO should not induce a significant loss of genetic diversity and the patterns of geographical genetic structure are expected to be more complex.

Palynological evidence has shown that the Amazon facing slopes of the eastern Andean cordilleras remained fairly wet during the glacial episodes of the PCO (Colinvaux et al., 2000; Bush & Oliveira, 2006). However, it is less known whether the montane forests that occur in the inter-Andean valleys at the proximity of where seasonally dry forests occurs nowadays suffered more extensively from compressions induced by the potential intensification of drought during the glacial periods of the PCO. Also, according to the prediction of the moist forests model, the potential secondary contacts initiated through the down slope migration of the populations in the wet Amazon facing slopes of the eastern cordilleras might have represented a challenge for the maintenance of incipient species integrity if the populations differentiating in allopatry did not have enough time to complete reproductive isolation. Unfortunately, the scarcity of phylogeography study on tropical montane plant currently limits our understanding about the potential complex impact the PCO had on these forests (Turchetto-Zolet *et al.*, 2013).

Here, we use amplified fragment length polymorphism genetic markers (AFLP's) to examine the phylogeography and the landscape genetics of 12 species belonging to the same species complex in the plant genus *Macrocarpaea* (Gentianaceae). All these species are endemic to the lower montane forests (LMF<1800m) of Ecuador and northern Peru. Previous studies strongly supported the group as monophyletic, but the relationships between species were globally poorly resolved (Vieu *et al.*, in review). The age of the most recent common ancestor (MRCA) of the group has been dated at 2Myr making all subsequent divergence events in the group contemporaneous with the PCO. Ten species: *M. cortinae*, *M. dies-viridis*, *M. illuminata*, *M. lenae*, *M. pacifica*, *M. pringleana*, *M. quechua*, *M. quizhpei*, *M. sodiorana*, and *M. umbellata* are small trees (3-6m) cryptically differentiated in their morphology while two species: *M. claireae* and *M. xerantifulva* are herbaceous/woody herbs (1-3 m; Grant, 2014a; 2014b). Three species occur on Pacific-facing slopes of the Andes (*M. pacifica*, *M. sodiroana*, and *M. umbellata*), while the remaining seven occur on Amazon facing slopes of the eastern cordilleras in the Amotape-Huancabamba zone in southern Ecuador and northern Peru (see map Fig. 1). *Macrocarpaea claireae* and *M. xerantifulva* are endemic to a single more internal valley extending from southern Ecuador and northern Peru between the Cordillera Real at west and the Cordillera del Condor at east, the Rio Chinchipe valley. This valley is connected in the south to the main Rio Marañon valley, which is associated with dry vegetation (the seasonally dry tropical forests: SDTF). The objectives of the study are: 1) to test whether the genetic diversity of the populations from *M. claireae* and *M. xerantifulva* is lower than in the other species as one would expect if the internal Andean valleys were dryer during the PCO; 2) to evaluate the validity of the species boundaries that were defined based on taxonomic and phylogenetic knowledge; and 3) to explore and compare the spatial genetic structure in this group of species.

MATERIALS AND METHODS

Sampling

Samples were collected between 2011-2013 in Ecuador and Peru. Since *Macrocarpaea* species generally occur in relatively small, sparse populations that are often hard to locate, it was not possible to follow a predefined sampling design. Most species were sampled from few and sometimes small populations (Figure 1, Table 1.). The exception is *M. xerantifulva* that we sampled more intensively throughout its distribution range. As no preliminary information was available about the population structure for any species in the genus we arbitrarily decided that plants from patches separated by a least 500m belonged to different populations.

DNA extraction and genotyping

Total genomic DNA was extracted from silica gel dried leaf samples with a standard cetyltrimethylammonium bromide (CTAB)-chloroform extraction followed by isopropanol precipitation and ethanol washing. Water eluted DNA samples went through a second step of purification using the DNeasy plant mini kit (Qiagen Ltd, Crawley, UK) and taking the Mini protocol at the step (11). DNA quality was checked on agarose gels and DNA concentrations were measured with Nanodrop 2000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA). DNA concentrations were homogenised between 10 and 20 $\eta g.\mu l^{-1}$.

Amplified fragment length polymorphism markers (AFLP) were generated for 345 individuals and 103 duplicates (29.7% of the dataset) following the protocol of Vos *et al.* (1995) with the modifications of Parisod & Christin (2008). All cycles of digestion ligation were performed on the same thermocycler machine. After a preliminary test on twenty four selective primers combinations, three selective primers combinations were retained based on their levels of polymorphism and reproducibility: M-CGA/E-AGC, M-CCA/E-ACA, M-CCG/E-ATC, labelled with the fluorescent dye FAM, VIC and NED respectively. PCR products amplified with these primers were pooled with GeneScan 500 LIZ Size Standard. Pooled product were sent to Macrogen Inc. and were separated on a ABI 3730XLs DNA Analyser (Applied Biosystems, Foster City, CA,USA). Fragment sizes were estimated with GeneMapper v4.0 (Applied Biosystems, Foster City, CA,USA) using the AFLP default peak detection parameters. Marker bins were set manually for each primer combinations and only fragments within the range of 50–500 bp with a height peak of a minimum of 50 relative fluorescent unit (rfu) were considered for generating a marker (a bin).

Peak scoring on the bin sets were performed automatically by the software and resulting tables with peak heights and sizes were exported and processed through the step (2) and (3) of the R script scanAFLP using the default threshold options (Hermann *et al.*, 2010). This script assesses the presence/absence within a particular locus relying on locus-specific fluorescent signal intensity thresholds estimated from qualities of the fragment signal intensity distribution. For each primer combination, we exported the binary data matrix referred as MatrixB in Hermann *et al.* (2010) that we assembled in a single binary matrix containing a total of 522 loci. At this step, loci present in less than 10 individuals and more than all the individuals minus 10 were discarded. The matrix with 306 remaining loci was used to estimate locus specific error rate on duplicated samples applying the mismatch approach of Bonin *et al.* (2004). Every locus with an error rate larger than 10% was removed from the binary matrix resulting in a dataset of 203 loci for 345 individuals with a mean genotyping error rate of 4.1%.

Data analyses

Population structure

The hierarchical structure in the AFLP dataset was investigated using the Bayesian genotype clustering method adapted to dominant markers implemented in the program STRUCTURE version 2.3.4 (Pritchard *et al.*, 2000; Falush *et al.*, 2007). The program allows to estimate the most likely number of genetic clusters in a particular dataset by grouping individuals such that Hardy-Weinberg equilibrium is maximized within clusters. The optimal number of genetic clusters *K* was determined using the modal distribution of the ΔK statistic of Evanno *et al.* (2005) with STRUCTURE HARVESTER (Earl, 2012). In datasets with complex structures, the Evanno approach tends to detect the upper most level of hierarchical structure, but not necessarily the optimal number of populations. As our dataset is composed of potentially fully differentiated populations (species), when dichotomic clusters were identified (*K=2*), the dataset was subset according to the results of the individual assignments for *K*=2 and analyses were repeated inside each cluster independently. For the analysis on the complete dataset, *K* values ranging from 1 to 16 were considered while analyses on the subset were performed for *K* values ranging from either 1 to 12 or 1 to 8. For each type of analysis and each value of *K*, three replicate chains of 175 000 MCMC iterations were ran and the first 25 000 iterations were discarded as a burn-in. The programme STRUCTURE was set with the recessive allele option and the admixture model with correlated allele frequencies was applied, any other options were let to default otherwise. The average membership coefficients for the 3 simulation runs of a given *K* value were generated with CLUMPAK (Kopelman *et al.*, 2015).

Descriptive statistics

Allele frequencies were estimated for each population from AFLP fragment frequencies with the Bayesian method of Zhivotovsky (1999) implemented in the software AFLPsurv (Vekemans *et al.*, 2002) using a non-uniform prior distribution and assuming that populations are at the Hardy-Weinberg equilibrium. The resulting allele frequencies were used to estimate the expected heterozygosity or Nei's gene diversity (H_j; Lynch & Milligan**,** 1994) for each population and to compute the genetic differentiation across all populations (F_{st}) using 1000 permutations. Thousand matrices of Nei's unbiased genetic distances were computed by bootstrapping over AFLP loci in AFLPsurv and exported to PHYLIP software package (Felsenstein, 1993) to build a neighbour-joining (NJ) dendrogram with bootstrap support on branches. In order to compare the portioning of the genetic variation between our a priori species hypothesis and the clusters identified at the issue of the two steps clustering analyses with STRUCTURE, two hierarchical analyses of molecular variance (AMOVA) were performed in Arlequin v3.5.1.3 (Excoffier *et al.*, 2010). Populations were first grouped by species, then by genetic clusters (see Table 1).

Spatial genetic structure

In order to estimate the spatial genetic structure (SGS) in the two main genetic clusters identified by STRUCTURE, the pairwise kinship coefficient was calculated with the software SPAGeDi v1.5.a (Hardy & Vekemans, 2002). The estimator for the kinship coefficient for dominant genetic markers developed by Hardy (2003) requires the user to specify an estimate of the inbreeding coefficient. Here we assumed a null inbreeding coefficient. Then, Kinship coefficients were associated with paired spatial distance in correlograms with 9 distance classes for the cluster composed by *M. xerantifulva* and *M. claireae* (0.1, 0.5, 2, 5, 10, 20, 50, 100, 200 km) and with 11 distances classes for the cluster composed by the other species (0.1, 0.5, 2, 5, 10, 20, 50, 100, 200, 400, 800 km). This relatively large distance classes compared with what is found in the literature was selected in order to accommodate our sparse sampling design. While it does not allow us to determine the fine SGS, it will already give us an idea of how genetic structure is influenced by distance. The significance of the average kinship coefficient in every distance class was tested using 1000 permutations of spatial locations and individuals while standard errors were estimated by Jacknifing across loci. The permutation procedure also allowed testing if the regressions slope of the kinship coefficient over the distance was significantly negative, as expected under a scenario of isolation by distance.

Both the kinship coefficient and the regression slopes are affected by the sampling design and thus are not appropriate for comparison between species. In order to compare the strength of SGS between genetic clusters, the *Sp* statistic (Vekemans and Hardy, 2004) was computed as $Sp = -b_F/(1-F_{(1)})$ were b_F is the slope of the regression of kinship against ln distance and $F_{(1)}$ is the average kinship of individuals in the first distance class here (0 -100 m). Because this statistic is calculated on the logarithm of the distance it is less affected by the sampling design and thus has the desirable advantage to be comparable across studies.

RESULTS

Population structure

The ΔK statistic supported an optimal value of K=2 for the STRUCTURE analyses including the genotypes of all specimens sampled in the study (Fig. S1A). On the other hand, estimates of mean log likelihood (mean L(K)) increased steadily after K=2 before to reach a plateau at K=8. This discrepancy between the two statistics indicates that the ΔK statistic identified the highest level of hierarchical structure in our dataset but that additional layers of genetic structure are likely present inside the clusters. In Figure 2A, Cluster B is composed of populations of *M. claireae* and *M. xerantifulva*, while cluster A is composed of the populations of all the remaining species. Levels of admixture between these two clusters are generally very low.

In the subsequent analyses performed inside the cluster B, both ΔK and mean L(K) favoured K=4 (Fig. S1B). The barplot shows the probability of assignment of individuals to a particular genetic cluster, with the individuals sorted according to increasing distance from the northern most individual (north: left, south: right, Fig. 1B, Fig. 2B) shows that two genetic groups are dominant. The sky blue (A1) in the north is highly dominant in the populations of *M. claireae 1* and *M. xerantifulva 20*, while the red (A3) is dominant in the populations found between the centre and the south. In between, the green cluster (A2) is often found in populations mixed with A1 and A3. Its prevalence tends to decrease toward the south and it is absent from the two northern-most populations. The only pure A2 populations are *M. xerantifulva 19* and *M xerantifulva 2* but both are formed by only two individuals. Last, the violet cluster (A4) is relatively rare but is present all along the geographic distribution. The only pure A4 population is *M xerantifulva 9*. Overall the level of admixture in individual is weak, except in Peruvian populations close to the border with Ecuador.

Inside the cluster B the ΔK statistic favoured *K*=3 while the mean L(K) only reach a plateau for *K*=6 (Fig. S1C). This indicates that several layers of genetic structure might remains inside the clusters identified.

One species have its populations that fall into distinct genetic clusters. *M lenae 1* falls into the cluster green (B1, Fig. 2C) while M. lenae 2 falls into the yellow cluster (B2). All the populations of *M. dies-viridis* fall in the cluster blue (B3). The cluster B1 regroups populations from eight different species and covers a very large geographic distribution (Fig 1A.). The level of admixture in individuals is very low except for few individuals from the three populations of *M. quizhpei*.

Descriptive statistics of diversity

Gene diversity (H_j, see Tab1.) is almost two times larger in the populations of the cluster B (mean H_j= 0.150; sd= 0.025) than in those of the cluster A (mean H_j= 0.089; sd= 0.026) and the difference in mean between the two clusters is highly significant ($T_{(34.5)}$= 7.076, P<0.001).

The results of the hierarchical AMOVAs performed either by grouping populations by species or by their prevailing assignment with Bayesian clustering (Tab. 2) revealed that in both case more than the half of total genetic variation (54.88% and 57.07 % respectively) can be explained by the variation between individuals inside the populations. In both design, the proportion of the variation explained by differences among groups (species or cluster) is two times higher than the variation explained by the difference among populations within groups. This indicates that the groups boundaries explain a substantial proportion of the genetic variation. However, the proportion of the variation explained by the species (33.06%) is higher than the proportion of the variation explained by STRUCTURE genetic clusters (27.25%), suggesting that the species boundaries better capture the overall genetic variation. In both design, all types of F statistics tested with permutations were highly significant (Tab.2). The fact that F_{SC} statistics are significant indicates that populations inside the groups are still appreciably differentiated.

The NJ dendrogram computed from Nei's genetic distance (Fig. 3) display the same basal dichotomy as inferred with STRUCTURE. The cluster A has relatively high bootstrap support (849/1000) and the population *M. xerantifulva 9* which is the only population predominantly assigned to the sub cluster A4 is well supported as being the basal most population inside the cluster. Then inside the cluster A, populations do not group conforming to their STRUCTURE assignment. Also, internodes are short and bootstrap support weak, indicating

potentially little divergence among populations. Inside the cluster B, the basal most position of *M. pacifica* is weakly supported (486/1000). Bootstrap supports are in general higher and the internodes much longer than in the cluster A indicating very likely a larger divergence between populations inside the cluster B. The only sub cluster identified by STRUCTURE strongly supported as monophyletic is the cluster B3.

In agreement with STRUCTURE, *M lenae* is not returned as monophyletic. In the other hand, *M. dies-viridis* (893/1000), *M. illuminata* (845/1000) and *M. quizhpei* are strongly supported as monophyletic (920/1000).

Spatial genetic structure

The correlograms showing the spatial autocorrelation of the kinship (F_{ij}) measured for the cluster A (Fig. 4A) and the cluster B (Fig. 4B) independently are similar in the way they reveal significant higher kinship than expected by chance across relatively large distance. For the cluster A, the kinship becomes null only between 6.5 and 33 km. For the cluster B, the kinship becomes null between approximately 6 and 36 km. The slope of the regression of kinship against ln of the distance is significantly more negative than expected under a scenario of absence of spatial genetic structure both for the cluster A (b_F = -0.0147; P<0.01) and the cluster B (b_F = -0.0148; P<0.01). The two clusters have relatively similar Sp statistic with *Sp* = 0.0158 and *Sp* =0.0172 for the cluster A and the cluster B respectively. Finally, it should be noted that the intra-individual kinship ($F_{(0)}$) which correspond to the F_{IT} when individual from different populations are compared like in this study (Hardy & Vekemans, 2002), is much higher for the cluster B ($F_{(0)}$=0.22) than for the cluster A ($F_{(0)}$= 0.13). This indicates that the sum of inbreeding inside the populations and the overall differentiation between populations is higher in the group B. As the gene diversity is higher in the populations of the cluster B (see above and Tab. 1), the difference is likely a consequence of a higher differentiation between populations in the cluster B rather than a higher inbreeding.

DISCUSSION

Deep genetic structure

The dendrogram constructed from matrices of genetic distances between populations (Fig. 2) and the Bayesian clustering analyses performed at the individual level (STRUCTURE) identified the same basal dichotomy in the genetic structure of the group of species we studied. The cluster B is composed by species that have their populations occurring on the Amazon facing slopes of the eastern cordilleras from northern Ecuador to northern Peru plus the unique population of *M. pacifica* that occurs on the pacific facing slopes of the western cordillera in southern Ecuador and the unique populations of *M. cortinae* and *M. umbellata* that occurs on the pacific facing slopes of the western cordillera in northern Ecuador. The cluster B is composed by the only population of *M. claireae* and the 19 populations we sampled for *M. xerantifulva* that all occur in the Rio Chinchipe valley. Interestingly, one of the rare studies that addressed the phylogeography of a montane plant group from the same region identified the Rio Chinchipe valley as a source population in *Ceroxylon echinulatum* (Arecaceae) for the colonization of the eastern and western cordilleras of the Andes followed by subsequent migration toward the north during the Quaternary (Trénel *et al.*, 2008). Here, instead we show that the cluster A likely remained trapped in this valley and the low level of admixture detected between the cluster A and B indicate that they have evolved in relative isolation (spatial and/or genetic) from one another. Of course the confirmation of this isolation await closer inspection of the valleys surrounding the Rio Chinchipe valley, but it should be noted that despite several visits on the eastern side of the Rio Chinchipe valley (populations were sampled on the western side of the river) we were not able to find populations from the cluster A while others *Macrocarpaea* species from different sections in the genus were collected at higher elevation. Visits further south, on the other side of the Rio Marañon were also unfruitful.

The concept of species inside the species complex

Several elements from our results rise important questions regarding the definition of the species boundaries inside this *Macrocarpaea* species complex. It is notable that neither the genetic distance based dendrogram nor the Bayesian clustering analyses support the species boundary defined based on taxonomic knowledge for several species. This is for example the case of *M. claireae* in the cluster A that is relatively deeply nested in the middle of the populations of *M. xerantifulva* on the dendrogram. Also the Bayesian clustering analyses reveal that *M. claireae* 1 and the northern most population of *M. xerantifulva* that were sampled few kilometres apart from one another (<3km) are members of the same genetic cluster (cluster A1). The fact that individuals

belonging to the cluster A1 are also found mixed with other genetic clusters in others populations close to the border between Ecuador and Peru indicate that *M. claireae* and *M. xerantifulva* are essentially the same species that display a geographic genetic structure. The relatively low branch length at the internodes for the clade A on the dendrogram also suggests a relatively low genetic differentiation between populations inside the cluster which support the hypothesis of the presence of a single species inside the cluster A. From this point onward we will refer to the cluster A as *M. xerantifulva*.

Inside the cluster B, *M. lenae* seems also problematic as its populations are neither monophyletic in the distance based dendrogram nor inferred as member of the same cluster by STRUCTURE. These results are likely the outcome of wrong identifications and illustrate the difficulty inerrant to the taxonomic assignment of populations from relatively young cryptic species relying solely on morphological character that can be subject to plasticity to a certain extant and DNA molecular makers displaying low level of genetic variation (Beheregaray & Caccone, 2007; Bickford *et al.*, 2007). A population genetic approach like we employ in this study has a great potential in order to first detect the genetic boundaries of cryptic species then to assist the taxonomic assignment of individuals in one species or another.

Surprisingly, the ΔK statistic (Evanno *et al.*, 2005) only favoured 3 genetic groups inside the cluster B where we recognized 10 taxonomic species (Grant, 2014a; 2014b). We think that this small number of group identified is very likely the outcome of the detection of a deep genetic structure inside the cluster B rather than the presence of a maximum of three differentiated genetic groups (Coulon *et al.*, 2008). This is illustrated by the low level of admixture detected inside each sub-cluster (Fig. 2C). In comparison, the mean log likelihood for the number of cluster (mean L(K); Sup. 1C) only started to reach a plateau around K=6 and thus indicate the potential presence of additional levels of genetic structure. The comparison of the partitioning of the genetic variance (AMOVA, Tab.2) either assuming the species boundaries or the genetic clusters as the highest level of genetic structure revealed that the species boundaries explained a greater proportion of the genetic variance which also the support the idea that the genetic clusters identified by STRUCTURE did not capture the totality of the genetic structure present in our dataset. In order to better understand the finer genetic structure inside the cluster B, it would be required to repeat the Bayesian clustering analyse on each sub-cluster independently.

However, the three sub-clusters are not fully supported by the genetic distance based dendrogram, which indicate complex and potentially reticulated relationship between the different sub-clusters. In the current state, we cannot take position concerning the taxonomic status of the different species inside the cluster B, but it is very likely that they have a complex history.

Difference in gene diversity and the dry refugia.

We found a significant deficit of about 40% in gene diversity in the populations of *M. xerantifulva* (mean H_j =0.089) in comparison with the populations from the cluster B (mean H_j =0.150). Compared with estimates in literature survey for others perennial and predominantly out crossing plant (Nyboom, 2004), gene diversity is relatively low for the cluster B and very low for *M. xerantifulva*. The very low levels of gene diversity found in *M. xerantifulva* indicate that this species very likely went through a demographic bottleneck which fits the prediction of the dry refugia model (Ramírez-Barahona & Eguiarte, 2013). This in turn suggests that the Rio Chinchipe valley might have been drier during the cold cycles of the PCO than nowadays and that the populations of *M. xerantifulva* re-colonized the valley only recently. An alternative hypothesis to explain the low gene diversity in *M. xerantifulva* could be that it has recently been founded by a small amount of individual (founder event; Dlugosch & Parker, 2008). This hypothesis is quickly falsified by the fact that the basal divergence between *M. xerantifulva* and the cluster B likely date back their common ancestor some 2Myr.

According the prediction of the dry refugia hypothesis it should be possible to locate the position of the refuge by identifying the population or the region that have the highest level of gene diversity. In *M. xerantifulva* there is no clear candidate population that distinguish from the other. The only exception is *M. xerantifulva* 12, but it is a very small population (N=3, *H*j=0.132) which limits its significance. Another approach to identify refugia could consist in looking at the diversity of genetic clusters inside populations (genotypes diversity). Using this approach, *M. xerantifulva* 18 which is located in northern Peru at proximity of the border with Ecuador and *M. xerantifulva* 18 which is located about 20 kilometres further south are the best candidates. Considering that if an intensification of drought occurred in the Rio Chinchipe valley, it should have been more severe in the south, the position of refugia in that part of the valley is not illogical. Nevertheless this pattern of genotypes richness cannot be taken as the ultimate evidence for the position of refugia. Secondary contact zone can also generate hotspot of genotypes diversity (genetic cline) and both scenarios have been shown to be relatively hard to set apart (Durand *et al.*, 2009).

Spatial genetic structure

Analyses of spatial genetic structure (SGS) revealed strikingly similar pattern of SGS between *M. xerantifulva* and the genetic cluster B. Both groups show significantly high kinship over relatively long distance (6-30km) and thus have a relatively week SGS. The regression slopes of the relationship between pairwise kinships and pairwise logarithms of distances (b_F) were significantly more negative than expected by chance which indicate a significant effect of isolation by distance. However this last assertion should be taken with caution as the maximal distances we considered were very large in comparison with other studies (124 km for M. xerantifulva and 847km for the cluster B!). Therefore it is not surprising to detect an effect of isolation by distance at such a large spatial scale. In addition, despite their similarity the SGS estimates for the cluster B is likely to be meaningless as this group is potentially composed by several species while SGS is usually estimated at the intraspecific level (Vekemans & Hardy, 2004). In consequence, we will limit the interpretation of SGS to *M. xerantifulva*. The significantly high kinship we found at relatively large distance in *M. xerantifulva* suggests that the populations of this species are relatively well connected genetically. *Macrocarpaea* species are usually predominantly pollinated by nectar feeding bats (Wolf, 2006; Grant, 2014a) which have foraging radius of 30 to 50 km and thus can disperse pollen across considerable distance (Fleming *et al.*, 2009). It is therefore possible that the high population connectivity we detect is enhanced by bat pollination. Finally the *Sp* statistic for *M. xerantifulva* (*Sp* = 0.0158) is extremely similar to the rare estimate that exist in the literature for bat pollinated plant species (Hardy *et al.*, 2006; Collevatti *et al.*, 2010).

CONCLUSION

We shown that the species delineation in this young *Macrocarpaea* species complex composed by several morphological cryptic species requires a new consideration. In that perspective the use of population genetic tool is expected to be particularly valuable. This study is a first attempt in that direction in the genus *Macrocarpaea* but suffers from several limitations. The highly asymmetrical sampling in terms of number of populations and number of individual included for the different species do not allow proper comparison of their genetic structure. Also, a greater effort in terms of analyses is obviously required to disentangle the complex genetic structure of the cluster B that regroups the essential of the supposed taxonomic diversity in this species complex. Nevertheless, we were able to demonstrate that *M. claireae* and *M. xerantifulva* very likely form a single species endemic of the Rio Chinchipe valley that evolved in relative isolation from the rest of the species complex.

Relying on a more dense sampling, we show that this species have a relatively low gene diversity that could have resulted from a recent demographic bottleneck. We suggest that intensification of drought induced by the PCO in the Rio Chinchipe valley might have compressed this species in refugia. While remaining highly speculative this statement suggests that the different Andean valley might have responded in dramatically different way to the PCO. The eastern most valleys remaining wet while the more internal ones, which are currently in contact with dry systems such as the Rio Marañon, suffering from drought.

AKNOWLEDGEMENTS

The work was supported by the Fonds des donations of the Université de Neuchâtel, the Basler Stiftung für biologische Forschung, and the Fondation Dr. Joachim de Giacomi.

REFERENCES

Baker, P. A., Bush, M., Fritz, S., Rigsby, C. A., Seltzer, G., & Silman, M. (2003). Last Glacial Maximum in an Andean cloud forest environment (Eastern Cordillera, Bolivia). *Geology*, **31**, 195-198.

Beheregaray, L. B., & Caccone, A. (2007). Cryptic biodiversity in a changing world. *Journal of Biology*, **6**, 6-9.

Bickford, D., Lohman, D.J., Sodhi, N.S., Ng, P.K.L., Meier, R., Winker, K., Ingram, K.K. & Das, I. (2007). Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, **22**, 148-155.

Bonin, A., Bellemain, E., Bronken Eidesen, P., Pompanon, F., Brochmann, C., & Taberlet, P. (2004). How to track and assess genotyping errors in population genetics studies. *Molecular ecology*, **13**, 3261-3273.

Bush, M. B., & Oliveira, P. E. D. (2006). The rise and fall of the refugial hypothesis of Amazonian speciation: a paleoecological perspective. *Biota Neotropica*, **6**, bn00106012006.

Brunschön, C., & Behling, H. (2010). Reconstruction and visualization of upper forest line and vegetation changes in the Andean depression region of southeastern Ecuador since the last glacial maximum—A multi-site synthesis. *Review of Palaeobotany and Palynology*, **163**, 139-152.

Colinvaux, P. A., De Oliveira, P. E., & Bush, M. B. (2000). Amazonian and neotropical plant communities on glacial time-scales: the failure of the aridity and refuge hypotheses. *Quaternary Science Reviews*, **19**, 141-169.

Collevatti, R. G., Lima, J. S., Soares, T. N., & Telles, M. P. C. (2010) Spatial genetic structure and life history traits in Cerrado tree species: inferences for conservation. *Natureza & Conservação*, **8**, 54-59.

Coulon, A., Fitzpatrick, J. W., Bowman, R., Stith, B. M., Makarewich, C. A., Stenzler, L. M., & Lovette, I. J. (2008). Congruent population structure inferred from dispersal behaviour and intensive genetic surveys of the threatened Florida scrub-jay (Aphelocoma cœrulescens). *Molecular Ecology*, **17**, 1685-1701.

Crespi, E. J., Rissler, L. J., & Browne, R. A. (2003). Testing Pleistocene refugia theory: phylogeographical analysis of Desmognathus wrighti, a high-elevation salamander in the southern Appalachians. *Molecular Ecology*, **12**, 969-984.

Dlugosch, K. M., & Parker, I. M. (2008). Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, **17**, 431-449.

Earl, D. A. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation genetics resources*, **4**, 359-361.

Elias, M., Joron, M., Willmott, K., Silva-Brandão, K.L., Kaiser, V., Arias, C.F., Piñerez, L.M.G., Uribe, S., Brower, A.V.Z., Freitas, A.V.L. & Jiggins, C.D. (2009). Out of the Andes: patterns of diversification in clearwing butterflies. *Molecular Ecology*, **18**, 1716–1729.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular ecology*, **14**, 2611-2620.

Falush, D., Stephens, M., & Pritchard, J. K. (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular ecology notes*, **7**, 574-578.

Felsenstein, J. 1993. PHYLIP (phylogeny inference package), version 3.6a2. Distributed by the author, Department of Genetics, Univ. Washington, Seattle.

Fjeldså, J., Bowie, R. C., & Rahbek, C. (2012). The role of mountain ranges in the diversification of birds. *Annual Review of Ecology, Evolution, and Systematics*, **43**, 249-265.

Fleming, T. H., Geiselman, C., & Kress, W. J. (2009) The evolution of bat pollination: a phylogenetic perspective. *Annals of Botany*, **104**, 1017-1043.

Flantua S.G., Hooghiemstra H., Van Boxel J.H., Cabrera M., González-Carranza Z., & González-Arango C. (2014) Connectivity dynamics since the last glacial maximum in the northern Andes: a pollen-driven framework to assess potential migration. *In: Stevens W.D., Montiel O.M. & Raven P.H (eds), Paleobotany and biogeography : a festschrift for Alan Graham in his 80th year*, Missouri Botanical Garden Press, St. Louis, pp. 98-123.

Gentry, A. H. (1982). Neotropical floristic diversity: phytogeographical connections between Central and South America, Pleistocene climatic fluctuations, or an accident of the Andean orogeny?. *Annals of the Missouri Botanical Garden*, **69**,557-593.

Grant, J. R. (2014). A Monographic Revision of the Neotropical Genus *Macrocarpaea* (Gentianaceae) in Ecuador. *In: The Gentianaceae-Volume 1: Characterization and Ecology*, Springer Berlin Heidelberg, pp. 37-147.

Grant, J. R. (2014). De *Macrocarpaeae* Grisebach (Ex Gentianacels) speciebus novis XI: five new species from the Andes of Ecuador and Colombia. *Harvard Papers in Botany*, **19**, 227-239.

Hardy, O.J., Vekemans, X. (2002). SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618-620.

Hardy, O. J. (2003). Estimation of pairwise relatedness between individuals and characterization of isolation-by-distance processes using dominant genetic markers. *Molecular Ecology*, **12**, 1577-1588.

Hardy, O.J., Maggia, L., Bandou, E., Breyne, P., Caron, H., Chevallier, M.H., Doligez, A., Dutech, C., Kremer, A., Latouche-Halle, C., Troispoux, V., Veron, V. & Degen, B (2006). Fine-scale genetic structure and gene dispersal inferences in 10 Neotropical tree species. *Molecular ecology*, **15**, 559-571.

Herrmann, D., Poncet, B. N., Manel, S., Rioux, D., Gielly, L., Taberlet, P., & Gugerli, F. (2010). Selection criteria for scoring amplified fragment length polymorphisms (AFLPs) positively affect the reliability of population genetic parameter estimates. *Genome*, **53**, 302-310.

Hewitt, G. M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological journal of the Linnean Society*, **58**, 247-276.

Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **359**, 183-195.

Hughes, C., & Eastwood, R. (2006). Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. *Proceedings of the National Academy of Sciences*, **103**,10334-10339.

Hughes, C. E., & Atchison, G. W. (2015). The ubiquity of alpine plant radiations: from the Andes to the Hengduan Mountains. *New Phytologist*. doi: 10.1111/nph.13230.

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular ecology resources*, doi: 10.1111/1755-0998.12387.

Knowles, L. L. (2000). Tests of Pleistocene speciation in montane grasshoppers (genus *Melanoplus*) from the sky islands of western North America. *Evolution*, **54**, 1337-1348.

Linder, H. P., Rabosky, D. L., Antonelli, A., Wüest, R. O., & Ohlemüller, R. (2014). Disentangling the influence of climatic and geological changes on species radiations. *Journal of Biogeography*, ***41***, 1313-1325.

Lynch, M., & Milligan, B. G. (1994). Analysis of population genetic structure with RAPD markers. *Molecular ecology*, **3**, 91-99.

Moritz, C., Patton, J. L., Schneider, C. J., & Smith, T. B. (2000). Diversification of rainforest faunas: an integrated molecular approach. *Annual Review of Ecology and Systematics*, **31**, 533-563.

Mourguiart, P., & Ledru, M. P. (2003). Last glacial maximum in an Andean cloud forest environment (Eastern Cordillera, Bolivia). *Geology*, **31**, 195-198.

Nybom, H. (2004). Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular ecology*, **13**, 1143-1155.

Orme, C.D.L., Davies, R.G., Burgess, M., Eigenbrod, F., Pickup,N., Olson, V.A., Webster, A.J., Ding, T.-S., Rasmussen, P.C., Ridgely, R.S., Stattersfield, A.J., Bennett, P.M., Blackburn, T.M., Gaston, K.J. & Owens, I.P.F. (2005) Global hotspots of species richness are not congruent with endemism or threat. *Nature*, **436**, 1016–1019.

Parisod, C., & Christin, P. A. (2008). Genome-wide association to fine-scale ecological heterogeneity within a continuous population of Biscutella laevigata (Brassicaceae). *New Phytologist*, **178**, 436-447.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959.

Ramírez-Barahona, S., & Eguiarte, L. E. (2013). The role of glacial cycles in promoting genetic diversity in the Neotropics: the case of cloud forests during the Last Glacial Maximum. *Ecology and evolution*, **3**, 725-738.

Rull, V. (2008). Speciation timing and neotropical biodiversity: the Tertiary–Quaternary debate in the light of molecular phylogenetic evidence. *Molecular Ecology*, **17**, 2722-2729.

Schmitt, T. (2007). Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in zoology*, **4**, 1-13.

Schönswetter, P., Stehlik, I., Holderegger, R., & Tribsch, A. (2005). Molecular evidence for glacial refugia of mountain plants in the European Alps. *Molecular Ecology*, **14**, 3547-3555.

Schwery, O., Onstein, R. E., Bouchenak-Khelladi, Y., Xing, Y., Carter, R. J., & Linder, H. P. (2014). As old as the mountains: the radiations of the Ericaceae. *New Phytologist*. doi: 10.1111/nph.13234.

Smith, S. A., De Oca, A. N. M., Reeder, T. W., & Wiens, J. J. (2007). A phylogenetic perspective on elevational species richness patterns in Middle American treefrogs: why so few species in lowland tropical rainforests?. *Evolution*, **61**, 1188-1207.

Trénel, P., Hansen, M. M., Normand, S., & Borchsenius, F. (2008). Landscape genetics, historical isolation and cross-Andean gene flow in the wax palm, Ceroxylon echinulatum (Arecaceae). *Molecular Ecology*, **17**, 3528-3540.

Turchetto-Zolet, A. C., Pinheiro, F., Salgueiro, F., & Palma-Silva, C. (2013). Phylogeographical patterns shed light on evolutionary process in South America. *Molecular Ecology*, **22**, 1193-1213.

van der Hammen, T., & Hooghiemstra, H. (2000). Neogene and Quaternary history of vegetation, climate, and plant diversity in Amazonia. *Quaternary Science Reviews*, **19**, 725-742.

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

Vekemans, X., & Hardy, O. J. (2004). New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology*, **13**, 921-935.

Vieu, J., Hughes, C. H., & Grant, J. R. Evolutionary diversification in the hyper-diverse montane forests of the tropical Andes: range expansion drives radiation of the plant genus Macrocarpaea (Gentianaceae). *In review at Molecular Phylogenetics and Evolution.*

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van de Lee, T., Hornes, M., Fritjers, A., Pot, J., Peleman, J., Kuiper, M. & Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic acids research*, **23**, 4407-4414.

Wolff, D. (2006). Nectar sugar composition and volumes of 47 species of Gentianales from a southern Ecuadorian montane forest. *Annals of Botany*, **97**, 767-777.

Wollenberg, K. C., Vieites, D. R., Van Der Meijden, A., Glaw, F., Cannatella, D. C., & Vences, M. (2008). Patterns of endemism and species richness in Malagasy cophyline frogs support a key role of mountainous areas for speciation. *Evolution*, **62**, 1890-1907.

Table 1. Properties of the sampled populations including: the voucher; country (E: Ecuador, P: Peru); the region code: C: Carchi, EO: El Oro, J: Jaén, MS: Morona-Santiago, P: Pichincha, S: Sucumbíos, SA: San Ignacio, SM: San Martin, ZC: Zamora-Chinchipe; Cluster, the main genetic cluster identified by STRUCTURE are indicated by A and B, sub- clusters inside these clusters are indicated by numbers from 1 to 4 for cluster A and from 1 to 3 for cluster B; N, the sample size; PLP, the proportion of polymorphic markers; *Hj* and S.E.(*Hj*), the gene diversity and its standard error respectively.

Population	Voucher	Country	Region	Cluster	Latitude	Longitude	N	PLP	*Hj*	S.E.(*Hj*)
M. claireae 1	JRG4666	E	ZC	A1	-4.5422	-79.1304	19	21.7	0.090	0.009
M. cortinae 1	JRG5114	E	S	B1	0.3671	-77.4923	12	38.9	0.136	0.011
M. dies-viridis 1	JRG4693	E	ZC	B3	-3.9009	-78.5117	9	43.8	0.119	0.011
M. dies-viridis 2	JRG4679	E	ZC	B3	-3.9373	-78.6245	10	48.3	0.143	0.012
M. dies-viridis 3	JRG4679	E	ZC	B3	-3.9263	-78.6235	20	49.8	0.136	0.011
M. dies-viridis 4	JRG4693	E	ZC	B3	-3.9120	-78.5040	6	44.3	0.127	0.011
M. illuminata 1	JRG4687	E	ZC	B1	-3.9673	-78.6863	10	54.7	0.157	0.012
M. illuminata 2	JRG4679	E	ZC	B1	-3.9378	-78.7200	7	41.9	0.139	0.012
M. lenae 1	JRG4700	E	ZC	B1	-3.5363	-78.5288	14	44.8	0.152	0.011
M. lenae 2	JRG4675	E	ZC	B2	-4.0891	-78.9653	18	56.2	0.168	0.011
M. pacifica 1	JRG4704	E	EO	B2	-3.6509	-79.7469	11	58.6	0.162	0.012
M. pringleana 1	JRG5121	E	MS	B1	-2.2073	-78.1985	5	37.4	0.128	0.012
M. quizhpei 1	JRG4688	E	ZC	B2	-4.1809	-78.6441	13	44.8	0.162	0.013
M. quizhpei 2	JRG4689	E	ZC	B2	-4.2503	-78.6613	18	52.2	0.180	0.013
M. quizhpei 3	JRG4688	E	ZC	B2	-4.2597	-78.6481	5	45.8	0.191	0.014
M. quechua 1	JV20	P	SM	B1	-6.4560	-76.2914	7	39.9	0.122	0.011
M. sodiorana 1	JRG5106	E	P	B1	0.3033	-78.8689	8	43.3	0.125	0.011
M. umbellata 1	JRG5107	E	C	B1	0.9156	-78.2013	2	36.9	0.207	0.015
M. xerantifulva 1	JV27	P	J	A3	-5.9771	-79.0516	4	3.9	0.032	0.005
M. xerantifulva 2	JV26	P	J	A2	-5.6830	-78.8715	2	34	0.116	0.011
M. xerantifulva 3	JV26	P	J	A3	-5.6777	-78.8779	5	8.9	0.061	0.008
M. xerantifulva 5	JV26	P	J	A3	-5.6840	-78.9044	8	36.9	0.079	0.009
M. xerantifulva 6	JV26	P	J	A3	-5.6824	-78.9032	17	35	0.112	0.012
M. xerantifulva 7	JV26	P	J	A3	-5.6754	-78.9116	6	36.9	0.099	0.010
M. xerantifulva 8	JV46bis	P	J	A3	-5.6857	-78.9328	5	32.5	0.088	0.010
M. xerantifulva 9	JV46	P	J	A4	-5.7051	-78.9443	3	32	0.099	0.010
M. xerantifulva 10	JV46	P	J	A3	-5.7023	-78.9335	3	3.4	0.036	0.005
M. xerantifulva 11	JV44	P	J	A3	-5.5770	-78.8976	10	42.4	0.107	0.011
M. xerantifulva 12	JV44	P	J	A3	-5.5816	-78.9428	3	34.5	0.132	0.012
M. xerantifulva 13	JV44	P	J	A3	-5.5788	-78.9494	14	11.3	0.060	0.008
M. xerantifulva 14	JV44	P	J	A3	-5.5748	-78.9710	5	34.5	0.086	0.010
M. xerantifulva 15	JV44	P	J	A3	-5.5655	-78.9730	5	37.4	0.089	0.009
M. xerantifulva 16	JV28	P	J	A2	-5.3705	-78.9104	17	41.4	0.105	0.010
M. xerantifulva 17	JV28	P	J	A3	-5.3747	-78.9378	12	35.5	0.082	0.009
M. xerantifulva 18	JV29	P	SI	A2	-5.1384	-79.0337	13	41.9	0.121	0.010
M. xerantifulva 19	JRG4666bis	E	ZC	A2	-4.8409	-79.2464	2	32.5	0.073	0.009
M. xerantifulva 20	JRG4667	E	ZC	A1	-4.5590	-79.1371	17	39.9	0.116	0.011

Table 2. Summary of the hierarchical analyses of molecular variance (AMOVA) performed in Arlequin by grouping: 1) the populations by species (top half of the table) and 2) by genetic clusters identified by the two steps STRUCTURE analyses. P-values were obtained with 1023 permutations

Source of variation	Df	Variance components	Percentage of variation	*F* statistics	*P-value*
Among species (1)	10	6.25	33.06	F_{CT} 0.33	<0.001
Among populations/species	26	2.28	12.06	F_{SC} 0.18	<0.001
Within populations	308	10.38	54.88	F_{ST} 0.45	<0.001
Total	344	18.91			
Among clusters (2)	6	4.96	27.25	F_{CT} 0.27	<0.001
Among populations/clusters	30	2.85	15.68	F_{SC} 0.22	<0.001
Within populations	308	10.38	57.07	F_{ST} 0.43	<0.001
Total	344	18.19			

Note: F_{CT}= variation among groups divided by total variation, F_{SC}= variation among sub-groups within groups divided by the sum of variation among sub-groups within groups and variation within sub-groups, F_{ST}= the sum of variation among groups and variation among sub-groups within groups divided by total variation.

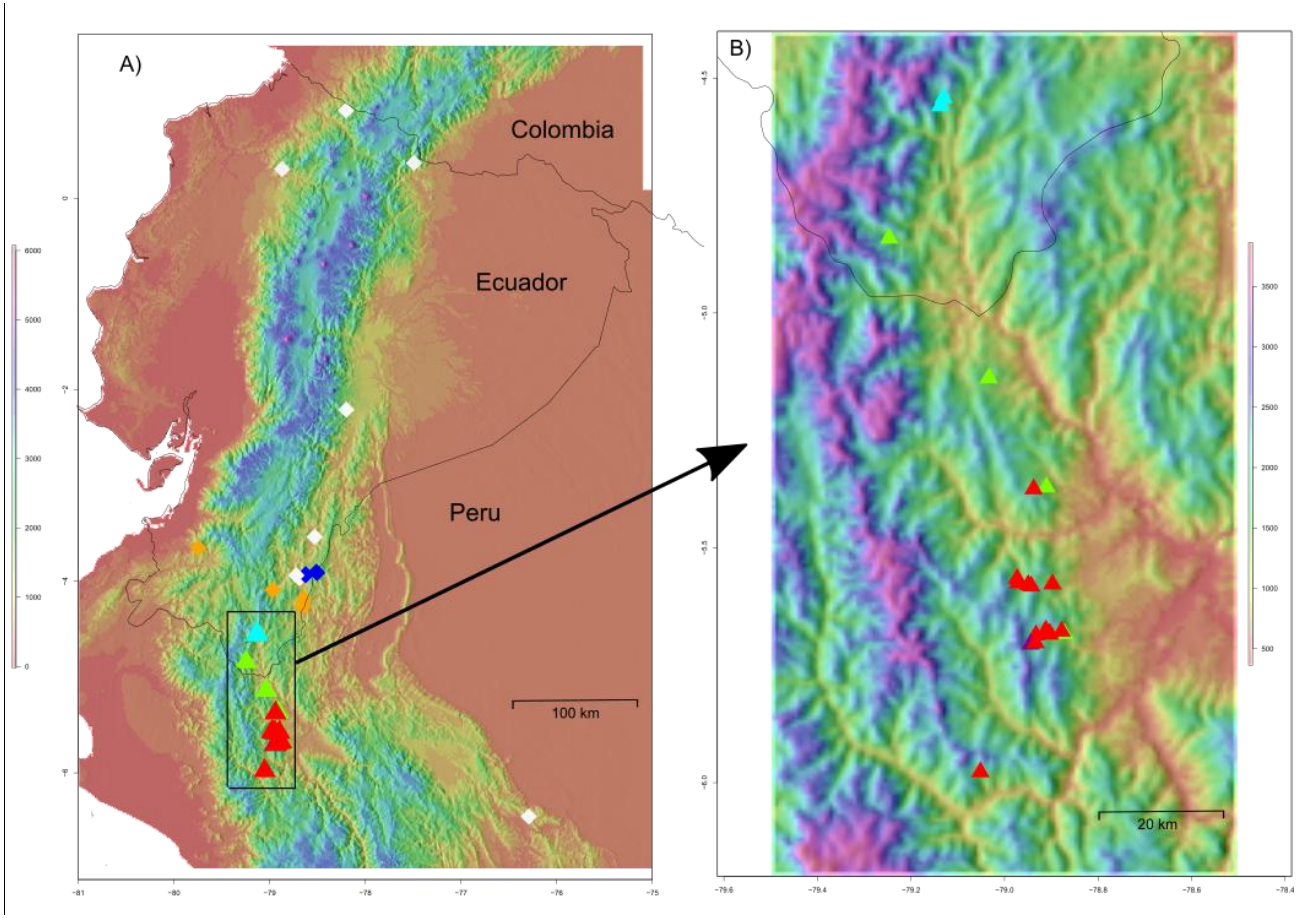


Figure 1. Map of the distribution of the populations sampled in the study. A) Distribution of all the populations sampled. Diamonds indicate the populations of the genetic cluster B inferred with the Bayesian clustering program STRUCTURE. The white diamond corresponds to the populations of the sub-cluster B1 (see Tab.1), the orange to the sub-cluster B2, the blue to the sub-cluster B3. Triangles indicate the populations of the cluster A. The sky blue triangle corresponds to the sub-cluster A1, the light green to the sub-cluster A2, the red to the sub-cluster A3, the violet to the sub-cluster A4. B) Zoom on the Rio Chinchipe valley and the distribution of the cluster A (*M. xerantifulva*).

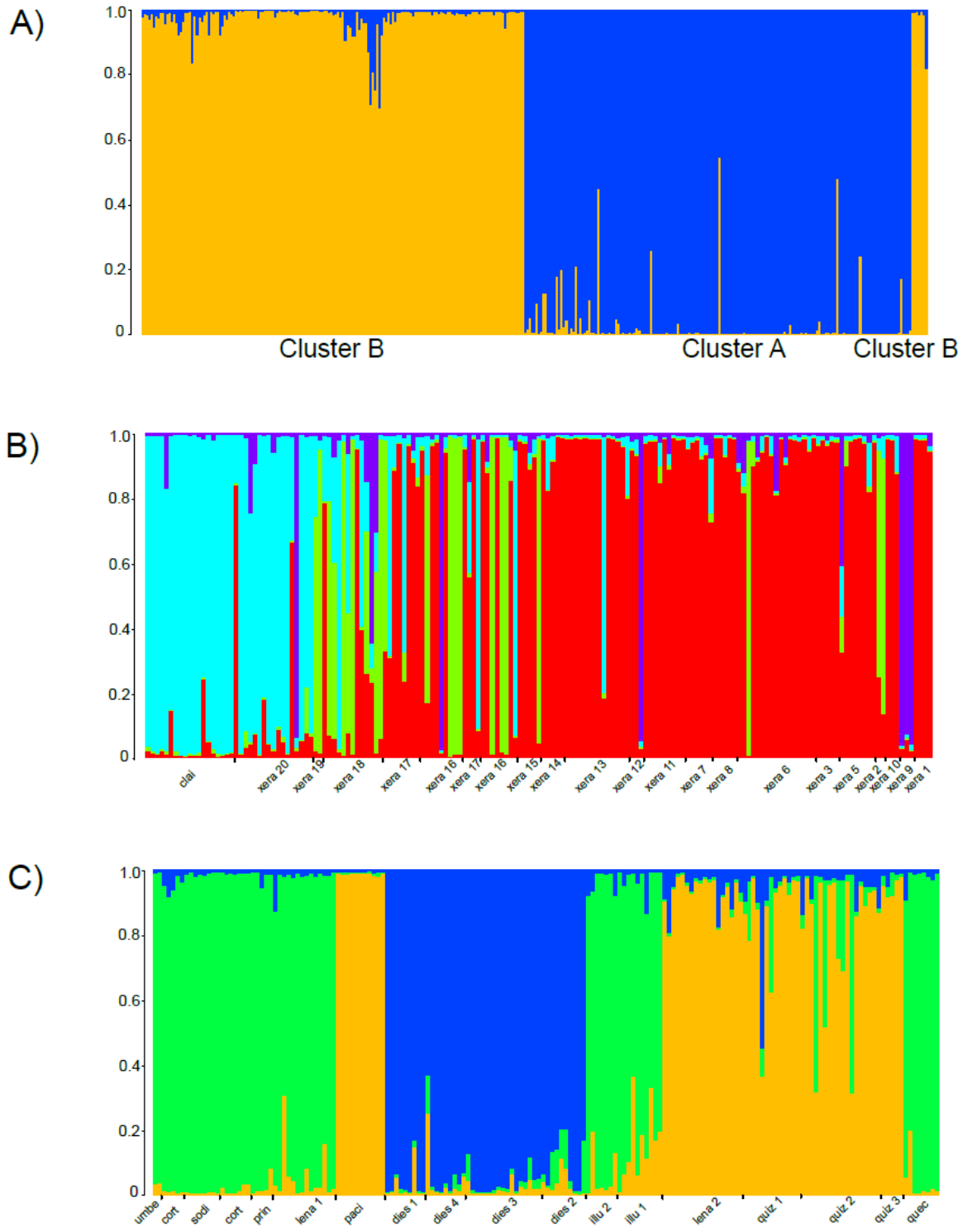


Figure 2. Barplots showing the average cluster assignment probability in nested STRUCTURE analyses of AFLP data for individuals from different species of *Macrocarpaea*. Each bar represents one individual and the colour indicates the source of the genetic cluster of its genotype. Bars are sorted according to decreasing latitude. A) Analyses performed including all the individuals sampled in the study. B) Analyses preformed on the individuals assigned to the cluster A in the analyses A, representing *M. claireae* and *M. xerantifulva* .C) Analyses preformed on the individual assigned to the cluster B in the analyses A, representing all the remaining species. The abbreviations of the names of the populations (see Tab. 1) are indicated below the bars.

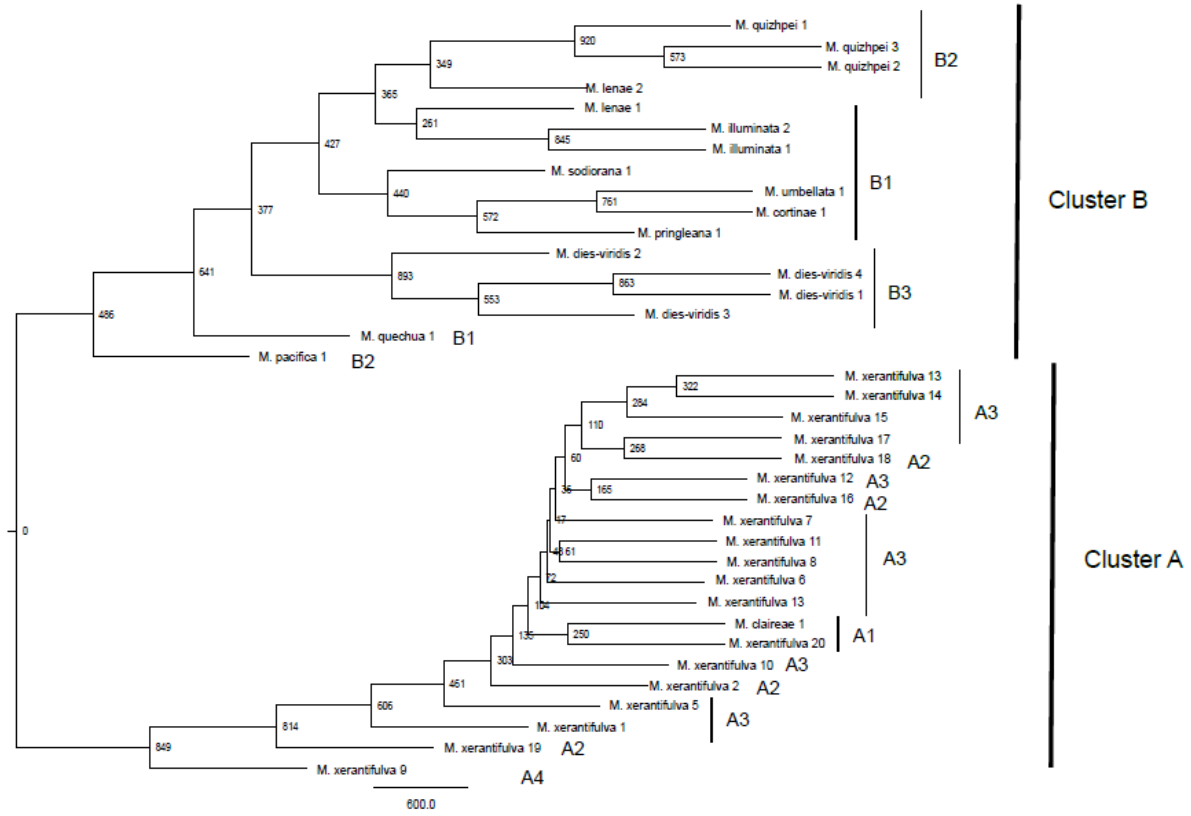


Figure 3. Neighbour-joining (NJ) dendrogram computed from Nei's genetic distance from AFLP loci. Values at node correspond to bootstrap support (N/1000) obtained by recalculating genetic distance from bootstrapped loci. The code A, A1 to A4, B and B1 to B3 refers to genetic clusters identified by Bayesian clustering method performed with STRUCTURE.

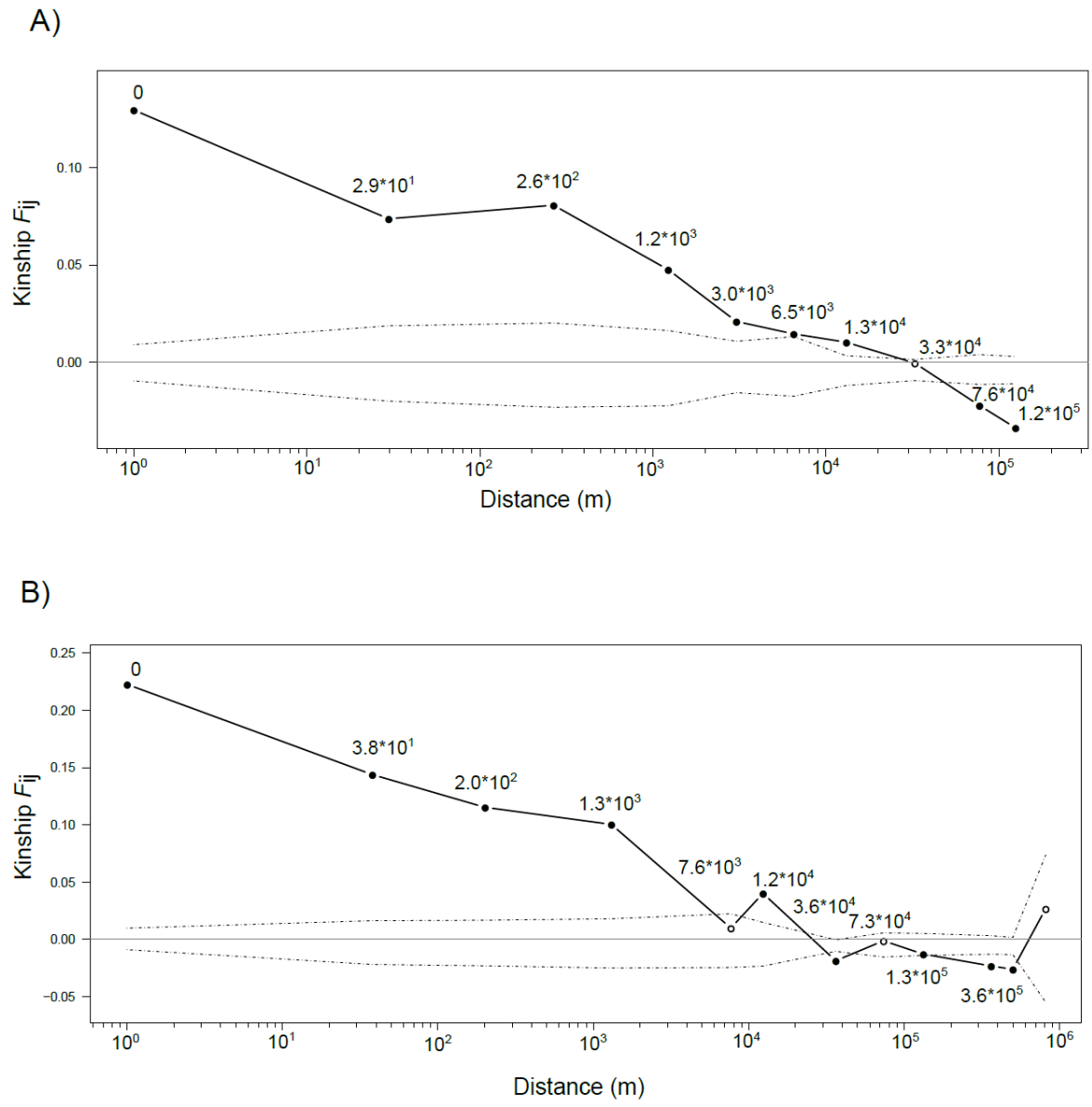


Figure 4. Correlograms of kinship coefficients (*Fij*) per distance classes represented on a logarithm scale; A) for the cluster A identified by STRUCTURE; B) for the cluster B. The plain line represents the value measured on the data. The dotted line represents the confidence interval of the permutations and indicates the null expectation of no spatial genetic structure. The circles show the mean distance estimated inside a particular distance class (see the methods). Closed circles indicate value significantly different from the null expectation while open circles correspond to the null expectation.

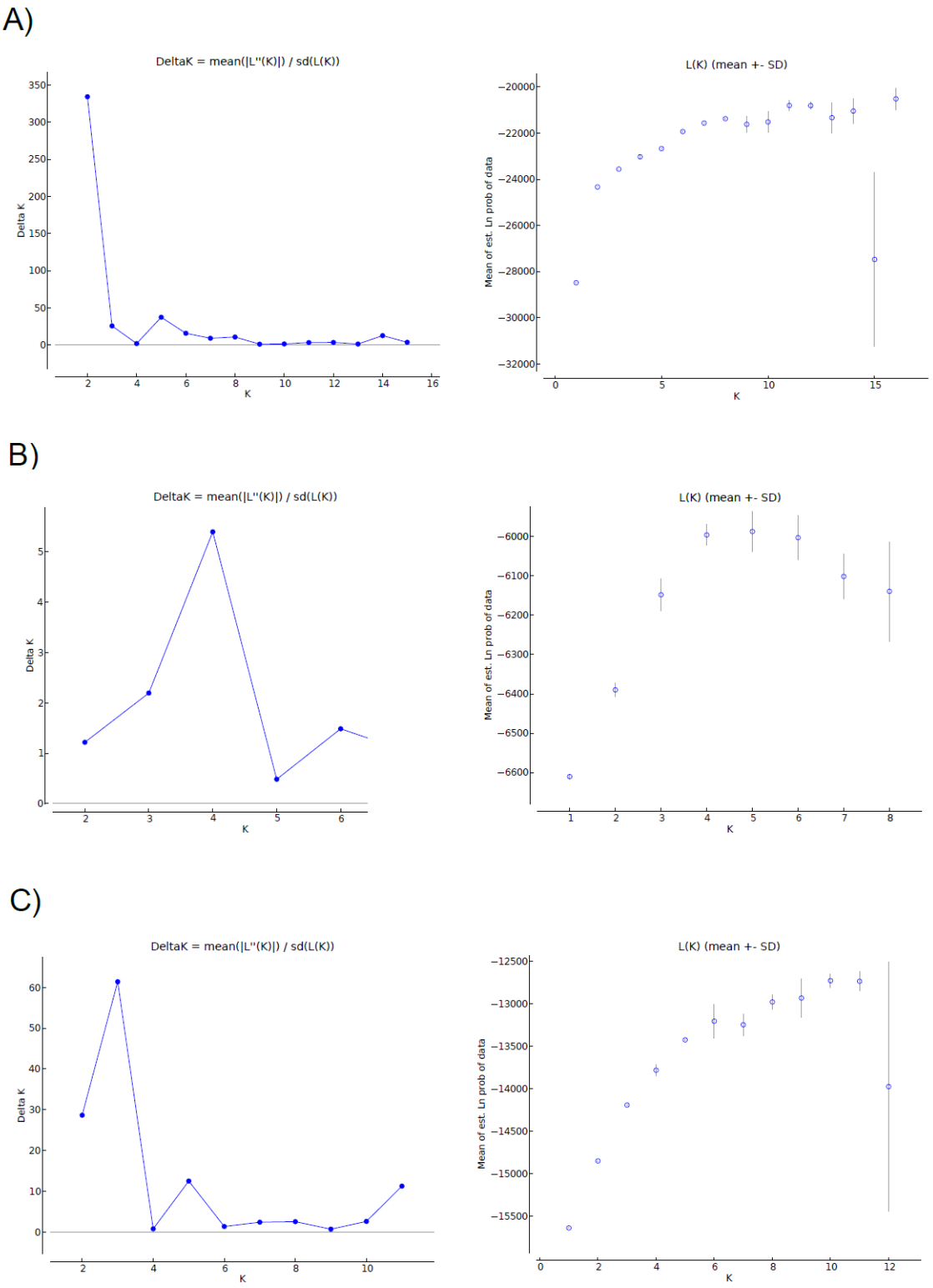


Figure S1. On the left side, plots of the ΔK statistic of Evanno *et al.* (2005) for detecting the number of K groups that best fit the data. On the right side, plot of mean likelihood L(K) and variance per K value from STRUCTURE. A) Analyses performed including all the individuals. B) Analyses performed inside the genetic cluster A. C) Analyses performed inside the genetic cluster B.

Conclusions

The middle elevation montane forests of the tropical Andes, a "mature" cradle of species diversity

The debate surrounding the origin of the exceptional diversity of the tropics has been historically centred on whether the tropics are regarded as a "museum" (long and slow accumulation of species), or as a "cradle" (recent and rapid accumulation of species) of diversity (Jablonski *et al.*, 2006). In the first chapter of this PhD thesis we show that the plant genus *Macrocarpaea* diversified recently (7Ma), in the middle elevation montane forests of the tropical Andes (MMF) with a diversification rate (0.43 $sp.my^{-1}$) two and a half times greater than a clade of the same genus in Brazil (0.16 $sp.my^{-1}$), and about four times larger than the closest relative of *Macrocarpaea*, the largely Amazonian plant genus *Tachia* (0.11 $sp.my^{-1}$). In consequence, we propose that the MMF is a "cradle" of diversity for *Macrocarpaea* and possibly for plant diversity in general. However, we also demonstrate that *Macrocarpaea* did not diversify at the same pace along its history in the MMF. The diversification rate rapidly increased to a very high rate at the time of the common ancestor (0.81 $sp.my^{-1}$), then, it decreased exponentially to reach a much lower rate in the modern time (0.31 $sp.my^{-1}$; this is still higher than elsewhere). This exponential decrease is a hallmark of a negative control induced by mechanisms of diversity dependence and indicates that the species diversity of *Macrocarpaea* in the MMF might be on the way to reach diversity equilibrium (Simberloff, 1974). We therefore suggest that the MMF is a "mature cradle" of diversity where species are in a later phase of their radiation. The existence of equilibrium also suggests that ecological interactions between species and/or space availability might in the long run limit species diversity (Ricklefs, 2004). In the second chapter we show that a clade of *Macrocarpaea*, the *micrantha* alliance, more recently colonized and possibly radiated in the lower montane forests (LMF), while the genus is ancestrally constrained to occur in the upper montane forests (UMF). These results indicate that the colonization of the LMF by the *micrantha* alliance has allowed this clade to escape from the diversity dependence diversification dynamic that affect the other sections of the genus in the Andes. This implies that the diversity dependence diversification dynamic we observed might be the outcome of: 1) either intra-generic competition interactions (this would mean that plant diversity in general does not necessarily affect the diversification dynamics of the genus, but the diversity of *Macrocarpaea* species in a particular ecological zone do); or 2) the diversity saturation of the UMF but not of the LMF. The distinction between these two hypotheses would require comparing the diversification dynamic of several plant group restricted to the LMF or the UMF.

On the prevalence of geographical processes and the stochastic nature of adaptive transitions

The contribution of adaptive (induced by divergent selection) and geographical processes (induced by isolation by distance and drift) in the origin of species diversity is a highly debated topic (Schluter, 2009). In Chapter 1 we showed that the radiation of *Macrocarpaea* in the Andes coincided with a period of rapid colonization of the entire extant distribution of the genus in the MMF. We also showed that dispersal associated with allopatric speciation (founder event) shaped the geographical phylogenetic structure of the genus. The second burst of diversification detected at the basis of the *micrantha* alliance is associated with a second wave of range expansion in the LMF. Taken together, these results indicate that geographical processes are determinant for the radiations of *Macrocarpaea* in the Andes. In Chapter 2 we explicitly tested the potential concomitant role of adaptive divergence to explain the main radiation of *Macrocarpaea* in the Andes and we refuted the hypothesis of an adaptive radiation. On the contrary we found that morphological and ecological evolution has been predominantly constrained, a pattern consistent with the hypothesis of "niche conservatism" (Wiens & Graham, 2005). Nevertheless, we also showed that a major adaptive transition occurred, the *micrantha* alliance shifted toward the LMF, an adaptive zone relatively depauperate of other *Macrocarpaea* species. The adaptive transition opened new space available for colonization and diversification through geographical processes. While the LMF and the UMF are adjacent habitats, the adaptive transition likely happened only once. This indicates that the nature of such adaptive transition might be stochastic and determined at the same time by intrinsic and extrinsic factors as also suggested by other studies (Onstein *et al.*, 2014).

The importance of historical factors

Historical factors such as mountain uplifts or climate dynamism have been put forward to explain the variation in species diversity between and within regions (Linder, 2008). In Chapter 1 we show that the main radiation of *Macrocarpaea* has occurred through repeated allopatric speciation coincided with the late phase of the Andean uplift in the late Miocene. It is at this time that the Andes attained sufficient elevation to modify the precipitation regime of the whole continent (Hoorn *et al.*, 2010). Precipitation increased on the eastern flank of the Andes, which likely triggered the rapid establishment of the MMF. We suggest that the potentially pre-adapted ancestor

of *Macrocarpaea* in the Andes took the opportunity to quickly colonize and diversify in this new habitat. The timing of diversification of several other plant groups that occur in the MMF strikingly coincided with the timing

we infer for *Macrocarpaea* which suggest a common diversification trajectory for the flora of the MMF and a primordial influence of the Andean uplift for the MMF diversity.

In Chapter 1 we also detected a potential secondary shift in the diversification rate at the basis of the *micrantha* alliance which occurs predominantly in the LMF, that coincides with the start of the Pleistocene climatic oscillations (PCO). Relying on pollen records for plants of the MMF (Flantua et al., 2014) we suggest that the greater contraction of the LMF compared to others vegetation belts during the PCO potentially stimulated diversification in the LMF. In Chapter 2, we instead suggest that the PCO facilitated the entry into the LMF of the *micrantha* alliance but not necessarily directly promoted species divergence. However, the PCO is a young event and it is hard to estimate its potential impact from molecular phylogeny if molecular markers used do not display a high mutation rate. In Chapter 3, we then use amplified fragment length polymorphism (AFLP) to study the potential impact of the PCO on the population demography and structure of 12 species inside a clade of the *micrantha* alliance. We showed that a species located in valleys connected to a dry system have significantly lower gene diversity than species from the eastern slopes of the Andes that are known for remaining wet during the glacial episodes of the PCO. We suggest that the low gene diversity of this species has resulted from a demographic bottleneck induced by the contraction of the montane forest as the dry system expands to this valley during the glacial cycle of the PCO. We also showed that the species from the slopes that likely remained wet often have complex relationships that could indicate a reticulated history facilitated by secondary contacts. Lastly, the impact of the PCO on *Macrocarpaea* diversity is not really clear, but it is possible that different valleys of the Andes have responded in very different way to the PCO.

The limitation of the phylogenetic approach and the complementarity of the population genetic approach

The phylogenetic approaches used in this PhD thesis have been useful to investigate the timing of the events of dispersal and divergence but also to investigate the dynamics of diversification and of morphological/ecological evolution. From these patterns observed on the phylogeny (which it is important to acknowledge, are subject to a substantial uncertainty inherent of these methods and the quality of the data we used) and correlations with what

is known from others discipline about historical events, we were able to propose an evolutionary scenario that can explain the diversification history of the plant genus *Macrocarpaea* in the MMF and that might be extended to others plant groups in the Andes. However, it is also important to acknowledge that the links between the

processes we propose and the patterns we infer remain largely verbal arguments and thus subject to potential alternative explanations. We thus agree with Losos (2011) when he says:" Phylogenies are powerful tools for understanding the past, but like any tool, they have their limitations… phylogenies are much more informative about patterns than they are about processes…they need to be integrated with direct studies of evolutionary process". The population genetic approach has much to offer for studying evolutionary processes. Not to say that it does not suffer from its own limitations. In this PhD thesis its use is extremely preliminary but it certainly deserves to be studied further.

Perspectives

Next-generation sequencing has brought evolutionary biology into a new area that allows combining phylogenetic and population genetic approach in an integrative way. It is now possible to study the mechanism of species divergence along a speciation continuum using the same type of data that facilitates the comparison of the processes acting at different hierarchical levels of diversity, from population, to clades, to species. In this thesis we showed that geographical processes were of primary importance for the diversification of *Macrocarpaea* in the Andes and we minimized the role of a potential adaptive divergence along the steep elevation gradients that are found in the Andes. Nevertheless our experience of field botanists tell us in our heart that adaptive evolution along the elevation gradient could have occurred, but the phylogenetic approach that captures the general pattern did not detect it (or we were not able to interpret it). *Macrocarpaea* species very often replace each other along elevation gradients. Future research could use the NGS approach for example:

1) to study in detail the potential effect of isolation by the environment (IBE, here the elevation) versus the isolation by distance (IBD) on the genomic differentiation among populations within species. The IBD should in principle affects the entire genome in the same way (action of genetic drift) while it is expected to find islands of strong genetic differentiation in some regions of the genome under the IBD (action of divergent selection).

2) Then, exactly the same approach could be applied to compare the genomic differentiation between closely-related species. Correlations between the processes of genomic divergence at this two hierarchical scale would allow rolling on the importance of adaptive divergence.

3) the NGS data could also be used to address a variety of other question such as the phenomenon of hybridization (at least two species of *Macrocarpaea* are natural hybrids), or introgression between species.

References

Flantua S.G., Hooghiemstra H., Van Boxel J.H., Cabrera M., González-Carranza Z., & González-Arango C. (2014) Connectivity dynamics since the last glacial maximum in the northern Andes: a pollen-driven framework to assess potential migration. *In: Stevens W.D., Montiel O.M. & Raven P.H (eds), Paleobotany and biogeography : a festschrift for Alan Graham in his 80th year, Missouri Botanical Garden Press, St. Louis*, pp. 98-123.

Hoorn, C., Wesselingh, F.P., ter Steege, H., Bermudez, M.A., Mora, A., Sevink, J., Sanmartı´n, I., Sanchez-Meseguer, A., Anderson, C.L., Figueiredo, J.P., Jaramillo, C., Riff, D., Negri, F.R., Hooghiemstra, H., Lundberg, J., Stadler, T., Särkinen, T. & Antonelli, A. (2010) Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science*, **330**, 927–931.

Jablonski, D., Roy, K., & Valentine, J. W. (2006). Out of the tropics: evolutionary dynamics of the latitudinal diversity gradient. Science, **314**, 102-106.

Linder, H. P. (2008). Plant species radiations: where, when, why?. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **363**, 3097-3105.

Losos, J. B. (2011). Seeing the forest for the trees: the limitations of phylogenies in comparative biology. *The American Naturalist*, **177**, 709-727.

Onstein, R. E., Carter, R. J., Xing, Y., & Linder, H. P. (2014). Diversification rate shifts in the Cape Floristic Region: The right traits in the right place at the right time. *Perspectives in Plant Ecology, Evolution and Systematics*, **16**, 331-340.

Ricklefs, R. E. (2004). A comprehensive framework for global patterns in biodiversity. *Ecology Letters*, **7**, 1-15.

Rundle, H. D., & Nosil, P. (2005). Ecological speciation. *Ecology letters*, **8**, 336-352.

Schluter, D. (2009). Evidence for ecological speciation and its alternative. *Science*, **323**, 737-741.

Simberloff, D. S. (1974). Equilibrium theory of island biogeography and ecology. *Annual Review of Ecology and Systematics*, **5**,161-182.

Uyeda, J. C., Hansen, T. F., Arnold, S. J., & Pienaar, J. (2011). The million-year wait for macroevolutionary bursts. *Proceedings of the National Academy of Sciences*, **108**, 15908-15913.

Wiens, J. J., & Graham, C. H. (2005). Niche conservatism: integrating evolution, ecology, and conservation biology. *Annual review of ecology, evolution, and systematics*, 519-539.