



HAL
open science

Evolutionary origin and diversification of the European mountain flora

Jan Smyčka

► **To cite this version:**

Jan Smyčka. Evolutionary origin and diversification of the European mountain flora. Biodiversity and Ecology. Université Grenoble Alpes, 2019. English. NNT : 2019GREAV031 . tel-02963884

HAL Id: tel-02963884

<https://theses.hal.science/tel-02963884>

Submitted on 12 Oct 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTE UNIVERSITE GRENOBLE ALPES

Spécialité : **Biodiversité Ecologie Environnement**

Arrêté ministériel : 25 mai 2016

Présentée par

Jan SMYČKA

Thèse dirigée par **Sébastien LAVERGNE**, Chargé de
Recherche, Laboratoire d'Ecologie Alpine, et
codirigée par **Cristina ROQUET**, Chercheuse, **Departament de
Biologia Animal, Biologia Vegetal i Ecologia, Universitat
Autònoma de Barcelona**

préparée au sein du **Laboratoire d'Ecologie Alpine**
dans l'**École Doctorale de Chimie et Sciences du Vivant**

Origine évolutive et diversification de la flore des montagnes européennes

Thèse soutenue publiquement le **11/10/2019**
devant le jury composé de :

Dr Marianne ELIAS

Directrice de recherche à Institut de Systématique, Évolution,
Biodiversité, Paris, Rapporteuse

Prof David STORCH

Professeur à Center for Theoretical Studies, Charles University in
Prague, Rapporteur

Dr Wilfried THULLER

Directeur de recherche à Laboratoire d'Ecologie Alpine, Grenoble,
Président

Prof Catherine GRAHAM

Chercheuse à Swiss Federal Institute for Forest, Snow and
Landscape Research, Zurich, Examinatrice

Dr Philippe CHOLER

Directeur de recherche à Laboratoire d'Ecologie Alpine, Grenoble,
Examineur

Dr Sébastien LAVERGNE

Chargé de recherche à Laboratoire d'Ecologie Alpine, Grenoble,
Membre



**EVOLUTIONARY ORIGIN AND
DIVERSIFICATION OF THE EUROPEAN
MOUNTAIN FLORA**

**ORIGINE ÉVOLUTIVE ET
DIVERSIFICATION DE LA FLORE DES
MONTAGNES EUROPÉENES**

**EVOLUČNÍ PŮVOD A DIVERZIFIKACE
EVROPSKÉ HORSKÉ KVĚTENY**

Jan Smyčka

supervised by Sébastien Lavergne and Cristina Roquet



The illustration depicts a nunatak refugium, one of relevant hypotheses of high elevation specialist plant survival of glacial periods. The extent and shape of non-glaciated areas, as well as the presence and appearance of humans are an artistic rendition rather than evidence-based. Used with kind permission of author, Anna Vosolsobě.

CONTENTS

SUMMARY.....	1
Resumé.....	3
Shrnutí.....	5
ACKNOWLEDGEMENTS.....	7
GENERAL INTRODUCTION.....	8
Why are mountains species rich.....	10
How to study diversification.....	16
European mountains as a study system.....	21
Thesis aims and structure.....	28
References.....	31
CHAPTER 1.....	37
Abstract.....	38
Introduction.....	39
Methods.....	42
Results.....	48
Discussion.....	55
Acknowledgements.....	62
References.....	63
CHAPTER 2.....	67
Abstract.....	68
Introduction.....	69
Results.....	72
Discussion.....	81
Methods.....	86
Acknowledgements.....	91
References.....	92
CHAPTER 3.....	95
Abstract.....	96
Introduction.....	97
Methods.....	102
Results.....	108
Discussion.....	118
Future directions.....	125
Acknowledgements.....	129
References.....	130
GENERAL DISCUSSION.....	134
Past climatic dynamics and its influence on survival and diversification.....	136
Modes of diversification of the European mountain flora.....	140
Future directions.....	145
References.....	148
APPENDICES OF CHAPTER 1.....	152
APPENDICES OF CHAPTER 2.....	160
APPENDICES OF CHAPTER 3.....	180

SUMMARY

Mountain biotas across the world are rich in species and also in endemics, suggesting that specific processes of diversification and survival were involved in their formation. In my thesis I focus on dynamics and drivers of plant diversification in the European mountains - a mountain system that is not among the worlds richest diversity hotspots, but is strikingly rich in the regional context, especially given its recent glaciation history. A lot has been written about diversification and glacial survival of specific lineages in the European mountain system, but so far it has been difficult to summarize these findings into a general picture. Here I benefit from recent advances in molecular and computational biology, allowing me to tackle evolutionary questions at a broader extent than previously possible, without severely compromising the resolution of the study. My thesis consists of three chapters:

In the first chapter I explore relationships between patterns of endemism, evolutionary diversity, mountain topography and glaciation history in the Alps. I show here that there are two different types of endemic hotspots with dramatically different evolutionary histories – one in calcareous peripheral refugia conserving narrowly distributed and evolutionary isolated species, and another one in high elevations inhabited by phylogenetically clustered assemblages of species that are often endemic for the Alps, but relatively widespread within the Alpine arc.

In the second chapter I focus on six lineages that quickly diversified in the European mountain system to explore their diversification in a comparative way. The diversification dynamics across the six lineages were relatively little impacted by the onset of Quaternary, likely because climatic oscillations were buffered by elevational migrations to mid elevation areas. Due to little affinity to elevational belts, the majority of the six lineages diversified allopatrically or across bedrock types, rather than parapatrically across elevational zones. A notable exception to this general pattern is *Androsace* sect. *Aretia*, which hosts multiple lineages of strict high elevation specialists.

In the third chapter I focus on population structure and glacial survival of the high elevation species *Androsace argentea*. I show here that the population genetic structure of this species is driven by long distance dispersals followed by relatively little local gene flow, perhaps due

Summary

to inefficiency of pollination. The populations in the internal Alps have variable origin and survival histories, ranging from long term *in-situ* survival to possibly post-glacial recolonization from Pyrenees.

In conclusion, my results suggest that the diversification dynamics of the European mountain flora was surprisingly little impacted by the glaciation history. This is likely because a large portion of European mountain flora was surviving in peripheral refugia in mid- or low elevation areas, and has little affinity to the high elevation habitats severely impacted by glaciation. In contrast to this general pattern, there is a narrow set of true high elevation specialists that likely survived *in-situ* in high elevation areas. The major mode of speciation in the European mountain system was allopatry and our results from *A. argentea* suggest that at least in high elevation specialists it might have been stimulated by combination of frequent long distance dispersals coupled with little local gene flow.

Resumé

Les biotas de montagne à travers le monde sont riches en espèces, mais aussi en espèces endémiques, ce qui suggère que des processus spécifiques de diversification et de survie ont été impliqués dans leur formation. Dans ma thèse, je me concentre sur la dynamique et les moteurs de la diversification végétale dans les montagnes européennes - un système montagneux qui n'est pas l'un des points chauds de la diversité les plus riches du monde, mais qui est remarquablement riche dans le contexte régional, surtout étant donné son histoire glaciaire récente. On a beaucoup écrit sur la diversification et la survie glaciaire de lignées spécifiques dans le système montagneux européen, mais jusqu'à présent, il était difficile de résumer ces résultats dans un tableau général. Ici, je profite des progrès récents de la biologie moléculaire et computationnelle, ce qui me permet d'aborder les questions évolutives dans une plus large mesure qu'auparavant, sans compromettre gravement la résolution de l'étude. Ma thèse comprend trois chapitres :

Dans le premier chapitre, j'explore les relations entre les modèles d'endémisme, la diversité évolutive, la topographie des montagnes et l'histoire de la glaciation dans les Alpes. Je montre ici qu'il existe deux types différents de points chauds endémiques avec des histoires évolutives radicalement différentes - les uns dans des refuges périphériques calcaires conservant des espèces étroitement réparties et évolutivement isolées, et les autres en altitude, habités par des groupements phylogénétiques d'espèces qui sont souvent endémiques pour les Alpes, mais relativement répandues dans l'Arc Alpin.

Dans le deuxième chapitre, je me focalise sur six lignées qui se sont rapidement diversifiées dans le système montagneux européen pour explorer de manière comparative leur diversification. La dynamique de diversification des six lignées a été relativement peu touchée par l'apparition du Quaternaire, probablement parce que la dynamique climatique a été tamponnée par les migrations en altitude vers les zones de moyenne altitude. En raison de leur faible affinité aux étages altitudinales, la majorité des six lignées se sont diversifiées de façon allopatrique ou sur différents types de substratum rocheux, plutôt que de façon parapatricque sur un gradient altitudinal. Une exception à ce modèle général est *Androsace* sect. *Aretia*, qui héberge de multiples lignées de spécialistes de haute altitude.

Dans le troisième chapitre, je me concentre sur la structure génétique des populations et la survie glaciaire de l'espèce de haute altitude *Androsace argentea*. Je montre ici que la

Summary

structure génétique des populations de l'espèce est déterminée par des dispersions sur de longues distances suivies d'un flux génétique local relativement faible, peut-être en raison de l'inefficacité des pollinisateurs. Les populations des Alpes intérieures ont des antécédents de survie variables, allant de la survie *in-situ* à une recolonisation, probablement post-glaciaire, des Pyrénées.

En conclusion, mes résultats suggèrent que les dynamiques de diversification de la flore de montagne européenne ont été étonnamment peu influencées par l'histoire glaciaire. Cela est probablement dû au fait qu'une grande partie de la flore de montagne européenne survécut dans des refuges périphériques situés dans des zones de moyenne ou basse altitude, et a peu d'affinité pour les habitats de haute altitude gravement touchés par la glaciation. Contrairement à cette tendance générale, il existe un nombre restreint de véritables spécialistes de la haute altitude qui ont probablement survécu *in-situ* dans les zones de haute altitude. Le principal mode de spéciation dans le système montagneux européen fût l'allopatricité et nos résultats de *A. argentea* suggèrent qu'au moins chez les spécialistes de la haute altitude, elle pourrait avoir été stimulée par la combinaison de dispersions fréquentes sur de longues distances associées à un faible flux génétique local.

Shrnutí

Horské bioty napříč celým světem jsou druhově bohaté a také mají vysokou míru endemizmu, což naznačuje že při jejich vzniku hrály roli specifické diverzifikační procesy. Ve své dizertaci se zabývám dynamikou a hnacími silami diverzifikace rostlin v evropských horách, které nepatří mezi světové hotspoty biologické diverzity, ale jsou velmi bohaté v regionálním kontextu, zvláště když uvažíme jejich glaciální historii. O diverzifikaci a glaciální historii konkrétních evolučních linií v evropském horském systému toho již bylo napsáno mnoho, nicméně zatím se příliš nedařilo z těchto zjištění poskládat obecnější pohled. V této práci využívám nedávných pokroků v molekulární a výpočetní biologii, které mi umožňují zabývat se evolučními otázkami týkajícími se evropské horské flóry v širším rozsahu, a zároveň se zachováním potřebného detailu. Moje práce sestává ze tří kapitol:

V první kapitole zkoumám vztahy mezi patrnými endemizmu, evoluční diverzity, horské topografie, a historie zalednění v Alpách. Ukazují zde, že existují dva různé typy hotspotů endemizmu s dramaticky odlišnou evoluční historií. Jeden je na vápencových refugiích na periferii Alp a obývají jej evolučně izolované druhy s malými areály. Druhý je ve vysokých nadmořských výškách a obývají jej druhy evolučně mladé, které jsou často endemity Alp, ale v rámci Alp jsou široce rozšířené.

V druhé kapitole porovnávám šest evolučních linií horských rostlin, které rychle diverzifikovaly v evropském horském systému. Diverzifikační dynamika napříč těmito šestmi liniemi byla relativně málo ovlivněna nástupem čtvrtohor, pravděpodobně protože kvartérní klimatické oscilace byly utlumeny vertikální migrací do území ve středních nadmořských výškách. Vzhledem k malé specializaci na výškové stupně probíhal diverzifikační proces především alopaticky nebo napříč různými typy podloží, spíše než parapatricky napříč výškovými stupni. Význačnou výjimkou z výše popsaného je *Androsace* sect. *Aretia*, což je linie hostící několik podskupin, které jsou striktními specialisty na vysokohorské prostředí.

V třetí kapitole se zabývám populační strukturou a mechanismy přežívání dob ledových u vysokohorského specialisty pochybku Vandellova (*Androsace argentea*). Ukazují zde, že populační struktura tohoto druhu je silně formována dálkovými disperzemi semen kombinovanými s minimálním místním genovým tokem způsobeným pravděpodobně nízkou efektivitou opylování. Vnitroalpské populace mají velmi různé původy a historie přežívání

Summary

glaciálů, některé z nich přežívaly *in situ* v zaledněných Alpách a jiné jsou výsledkem dost možná postglaciální rekolonizace z Pyrenejí.

V souhrnu mé výsledky naznačují, že dynamika diverzifikace evropské horské květeny byla překvapivě málo ovlivněna glaciální historií. Pravděpodobným důvodem je, že většina evropské horské květeny přežívala zalednění v refugiích ve středních nadmořských výškách na úpatích hor, a není striktně vázaná na vysokohorské prostředí. Výjimkou z tohoto obecného pravidla je úzká skupina vysokohorských specialistů, kteří nejspíše přežívali glaciály *in situ* na vrcholcích zaledněných pohoří. Nejdůležitějším speciálním mechanismem v evropské horské flóře je alopatrie a výsledky z třetí kapitoly naznačují, že přinejmenším u vysokohorských specialistů mohl být tento mechanismus poháněn častými dálkovými disperzemi kombinovanými s minimálním lokálním genovým tokem.

ACKNOWLEDGEMENTS

First I would like to thank to my advisors, Seb and Cristina for the guidance and their, let's say, full contact approach to the project. And also for sharing all these funny fieldwork situations with forgotten keys, flat tires and lost tracks.

I'm also really grateful to many people from LECA, among them especially: Laura, for her insights and support; Joao and Océanne, for sharing the fate of doctoral students; and Thibaut, without his massive help, the third chapter would not happen.

I would also like to thank to Tomáš, Martin and Filip, the three guys who convinced me that leaving abroad with small kids to do a PhD is a good idea, and then supported me all the way.

Really crucial for me was the support from my family. Díky Maruško, Jiříku a Kájo.

GENERAL INTRODUCTION

Mountains across the world usually host higher species richness and higher levels of endemism than adjacent lowlands (Hughes and Atchinson, 2015; Steinbauer et al., 2016; see Box 1 for definition of mountain regions), and some mountain ranges are among the most important biodiversity hotspots of the world (Hughes, 2017; Myers et al., 2000). The origin and maintenance of biological diversity of mountain biotopes fascinate naturalists since the rise of modern biology, and some seminal works of founders of evolutionary biology and biogeography were inspired by mountains. The pioneering biogeographic and bioclimatologic works of von Humboldt were inspired by the observation of altitudinal zonation of vegetation in Northern Andes (von Humboldt, 1805), and also Darwin's insights on evolution of terrestrial plants were clearly connected to 19th century's discussions about the origin of the flora of the Alps (Ball, 1879; Friedman, 2009).

At present, I see two major motivations for studying the origins and maintenance of biodiversity (i.e. diversification in broad sense) in mountain regions. A practical one is that some mountain habitats, specifically the cold and wet ones, are particularly influenced by ongoing climatic changes (La Sorte and Jetz, 2010; Thuiller et al., 2014, 2005; Wang et al., 2014). Understanding the processes that led to formation of mountain biodiversity, especially in contexts past climatic fluctuations (Condamine et al., 2013) and evolution of climatic niches (Quintero and Wiens, 2013), may help us to estimate future threats and eventually optimize conservation planning and management for mountain biota.

The second motivation is that mountains are exciting study systems for testing fundamental ecological and evolutionary theories that aim to explain the origins of current-day biodiversity patterns. In particular, mountain regions are a notable high richness outlier from the latitudinal diversity gradient (Fjeldså et al., 2011). Moreover, the cold high elevation habitats in mountains exhibit faster diversification rates than their lower counterparts (Steinbauer et al., 2016), and have hosted some of the fastest documented radiations in the world (Hughes and Eastwood, 2006), which is in sharp contrast with global gradient of diversification rates growing towards warm regions (Pyron and Wiens, 2013; Rolland et al., 2014). As a notable exception, the diversity and diversification in mountain regions may thus help our understanding of global latitudinal gradients (Körner, 2000).

General introduction

Many studies of diversification in mountain biota across the world were conducted with these research motivations. In my thesis I focus on the European mountains system, a region where diversification processes have been surprisingly little explored given its accessibility, biogeographic research tradition and data availability. I believe that this effort contributes to ameliorate our understanding of the diversification process not only in European mountains, but in temperate mountain systems as a whole. In the following sections I review the potential mechanisms of origin and maintenance of diversity in mountain regions, the approaches to study them and the specifics of the European mountain system.

Box 1: Definition of mountain regions

Mountain regions are typically defined based on one of two criteria: ruggedness, i.e. high difference between highest and lowest point on spatial certain scale (e.g. 200 m on 2'30" *sensu* Körner et al., 2011), or presence of high elevation areas (e.g. above climatic timberline in Nagy and Grabherr, 2009). The majority of studies on global mountain biodiversity work with the ruggedness definition, while studies exploring the upper limits of life typically use the high elevation definition. The relationship between ruggedness and high elevation is theoretically triangular – areas with high elevational differences necessarily contain areas with high elevation, but high elevation areas are not necessarily rugged. Practically, the correlation between the two is much tighter, because flat high elevational plateaus are a relatively rare type of landscape across the world, the most notable exceptions being the Tibetan plateau and some parts of the Andes. In the European context, the coincidence between ruggedness (*sensu* Körner et al., 2011) and high elevation areas (above climatic timberline) is relatively strong, with major continuous rugged areas containing at least some habitats above the climatic timberline and vice versa (<http://www.mountainbiodiversity.org/>; Nagy and Grabherr, 2009). One exception to this patterns are rugged, but fairly warm areas in the Mediterranean region. These regions are rugged and often treeless, but the climatic limitation to the growth of trees is drought, eventually combined with ancient human management, rather than low temperatures. Another exception are lowland tundra regions in the European Arctic, being flat and above climatic timberline governed by low temperatures. In the following text, and throughout my thesis, I use both definitions, depending on the discussed phenomenon. To make a clear distinction among them, I refer to “mountains” when I write about ruggedness property, or otherwise to “high elevation areas” or “alpine vegetation”.

Why are mountains species rich

Multiple hypotheses have been proposed to explain the exceptional diversity of mountain regions. These explanations differ in the processes they rely on, but also in the diversity components and phenomena they explain. In particular, there are explanations for higher richness or endemism of mountains compared to surrounding flat regions, but also explanations for higher endemism and faster diversification specifically in high elevation areas (see Box 1 for definitions of mountains and high elevation areas). The majority of presented explanations are not mutually exclusive, and dominant ones are likely to differ in different mountain systems.

The most simplistic explanations of why are mountains species rich do not account for evolutionary processes and explain richness of the mountains systems solely by ecological mechanisms. One of such explanations is that a unit of map area in rugged regions contains more surface inhabitable by biota than a unit of map area in flat regions, leading to higher richness simply due to more physical space available (Fig. 1A). This explanation was originally developed for underwater systems (Kostylev et al., 2005), and up to my knowledge it has never been seriously considered in terrestrial ecosystems. However, this simple geometric property of rugged landscapes may be relevant for organisms that are limited by resources distributed according to surface rather than according to map area. An example of such organisms may be plants in rocky habitats, limited rather by availability of rooting opportunities in rock crevices than by light. Another purely ecological explanation is that mountains are rich because they host higher diversity of ecological conditions than flat areas, allowing the ecological assembly of more species specialized for different niches (*sensu* Chesson, 2000; Fig. 1B).

Both ecological mechanisms can explain higher richness of mountain regions, but they fail to tackle another notable pattern of mountain biodiversity, which is high endemism (Steinbauer et al., 2016). It suggests that in most situations, they must be coupled with other explanations accounting for evolutionary processes capable of producing endemism. For example, the higher richness of habitats may allow ecological assembly of more species, but on a temporal scale it may also trigger adaptive diversification to specific habitats (reviewed below), generating endemism. The main aim of this thesis is studying diversification, rather than ecological assembly, but the purely ecological explanations why are mountains species rich

General introduction

are an important null models to consider when exploring more complex explanations. Moreover, there are specific mountain regions with extremely low endemism, where the biota indeed resulted mostly from ecological assembly of immigrant species, for example the flora of arctic mountain regions emerged almost completely from post-glacial recolonization (Brochmann et al., 2003) and little in-situ diversification could take place after the glaciation.

Higher endemism in mountains may be connected to the fact that some mountain regions are more climatically stable than surrounding lowlands and may thus host paleoendemic species that got extinct anywhere else. Climatic stability in mountain regions in the past is related to two factors: first, rugged terrains provide species the opportunity to migrate across the altitudinal gradient to buffer temperature changes with little migration in space (Fig. 1C; Feng et al., 2016); second, during the Pleistocene glacial periods a very important limiting factor for vegetation was humidity, and in many regions, mountains constituted humid refugia that allowed a higher survival rate of species than in dry lowland landscapes (Fig. 1D; Birks and Willis, 2008; Schmitt, 2007). On the other hand, certain mountain ranges or their parts may have been climatically less stable than surrounding lowlands – an extreme example are the mountain ranges that were severely glaciated during the Pleistocene ice ages. Mountains may thus create a mosaic of regions that were more climatically stable during the Pleistocene than surrounding lowlands (i.e. refugia) and regions where almost all biota was eradicated in some phases of Pleistocene glacial cycles (Feng et al., 2016; Tribsch and Schönswetter, 2003). Where the stability explanation of high endemism is the case, a region should be a species museum (*sensu* Stenseth, 1984) with diversity in a large part consisting of rare ancient lineages.

The museum explanation may apply to some regions, but most mountain regions (including those that constitute global biodiversity hotspots) host many neoendemic species and lineages that diversified in recent times (Hughes and Atchinson, 2015; Pouchon et al., 2018; Xing and Ree, 2017), suggesting that such regions are species cradles rather than museums (*sensu* Stenseth, 1984). In fact, some radiations in tropical mountains, e.g. *Lupinus* in the Andes, are considered the fastest documented examples of diversification in the world (Hughes and Eastwood, 2006). There are several explanations of faster diversification in mountain regions, having specific ecological, geographical and temporal consequences.

General introduction

The first explanation is linked with the aforementioned explanation related to high diversity of environmental conditions, adding a dynamic evolutionary component. That is, diversification can be stimulated by the presence of many different habitats on a relatively small area, fostering adaptive differentiation of populations and ecological sympatric or parapatric speciation (Fig. 1E). If this mechanism is the predominant one, speciation events would take place across the major ecological gradients in the mountains, such as the elevational gradient (associated with gradients of temperature and precipitation), and for plants also across different bedrocks (calcareous vs siliceous bedrocks) or soil depths (deep soils vs bare rock). Indeed, it has already been documented that speciation events may be in some cases associated with elevational niche change (Hughes and Atchinson, 2015; Luebert and Weigend, 2014; Merckx et al., 2015; Pouchon et al., 2018) or with change of bedrock preference (Boucher et al., 2016; Moore and Kadereit, 2013; Pachschröll et al., 2015).

Another potential explanation is linked to insularity of mountains. Mountains can be considered as insular systems of specific habitats (“Sky Islands” *sensu* Heald, 1967), thus creating the opportunity for geographic isolation of populations and allopatric speciation (Fig. 1F). This mechanism may be further fostered by changing connectivity of habitats due to climatic changes, via a speciation pump effect *sensu* Haffer (1967). An important question concerning this explanation is at which spatial scale is the isolation relevant for allopatric speciation, given that mountains present insularity at different scales, from major mountain ranges of thousands to millions of square kilometers, to individual summits of massifs within the mountain ranges, or at even smaller scale, to patches of specific habitats such as rock crevices for plants. The expected pattern of biodiversity if the allopatric speciation predominates is that sister mountain lineages would present non-overlapping geographic distributions. For example, Boucher et al. (2016) explored the frequency of ecological vs allopatric speciation events in mountain Primulaceae in Europe, showing that allopatric speciation is indeed the dominant process. Importantly, different habitats present in the mountains may also exhibit different levels of insularity. For example high elevation habitats are typically more insular than lower elevation stands (although this might be the opposite in certain regions and geographic scales, such as in uplands divided by deep canyons in eastern Africa), and the more insular habitats should thus have higher rates of species diversification than the ones that are more connected.

General introduction

There are also other explanation of diversification in mountains, predicting that diversification should be faster specifically in high elevation areas. Interestingly, these explanations typically do not predict that high elevation areas would be more rich in species, which is consistent with observed patterns – the species richness is typically declining with elevation or is eventually hump shaped due to mid-domain effects (Lomolino, 2001; Quintero and Jetz, 2018; Rahbek, 1995), while proportion of endemism and evolutionarily young species grows with elevation (Aeschimann et al., 2011; Quintero and Jetz, 2018; Steinbauer et al., 2016). But in turn, high endemism in high elevation areas, despite lower total richness, may be an explanation for higher richness of whole mountain regions containing such high elevation areas.

One of such high-elevation-specific explanations assumes that cold high elevation habitats are relatively younger than warm low elevation habitats, and the fast diversifications of high-elevation biota is triggered by the opportunity to colonize vacant ecological space (Fig. 1H; Hughes and Eastwood, 2006; Stroud and Losos, 2016). Such radiations would present particular evolutionary signatures. Their beginning should be coincident with orogenetic processes or with the appearance of key innovations allowing the colonization of high altitude habitats. Due to this, the diversity proportions in high and low elevation habitats should be out of equilibrium that is to be reached after saturation of diversification in high elevation habitats, and stabilize with carrying capacities not necessarily larger in high than in lower elevation habitats. There is evidence that diversification dynamics match with orogenetic processes in some cases, as it has been shown for the flora of Hengduan mountains (Xing and Ree, 2017), bellflowers in the Andes (Lagomarsino et al., 2016) or for Apollo butterflies in Eastern Asia (Condamine et al., 2018). Other studies have shown that diversification slows down after initial radiation triggered by the evolution of morphological adaptations to cope with harsh environmental conditions found at high elevations (Roquet et al., 2013). In all these examples, it can be assumed that the vacant ecological space mechanism does not work by itself, but in concert with others mechanisms such as those related to habitat diversity, or insularity of mountain regions.

There are other hypotheses aiming to explain higher diversification at high elevation that are either little tested, or can only apply to certain groups of organisms. For example, diversification in high elevations may be stimulated by higher UV irradiance causing faster molecular evolution (Fig. 1G; Cortés et al., 2018; Willis et al., 2009). Similarly, it was

General introduction

proposed that fast diversification in some mountain lineages may be linked to genome duplications, that may confer advantages for life in colder temperature and trigger speciation at the same time (Fig. 1J; Jordon-Thaden and Koch, 2008; Schönswetter et al., 2007; Theodoridis et al., 2013). The fast diversifications in insect-pollinated high elevation plants may also be linked to lower availability and mobility of pollinators in high elevation habitats (García-Camacho and Totland, 2009; Körner, 1999), leading to smaller connectivity of populations, which can favor their differentiation and finally lead to speciation (Fig. 1I). Such specific mechanisms of diversification are fascinating subjects of study, but they do not provide a general answer to why high elevation areas are rich in endemics or why mountain biota diversifies faster, because they apply only to specific groups of organisms. The UV mechanism can only play an important role in organisms with germinal lineages cells long term exposed to outer environment, as are plants; polyploidisation is common in certain groups of vascular plants, but is very rare in others, and the later mechanism only applies to insect-pollinated plants. In contrast, high endemic richness and fast diversifications in high elevation areas are fairly general phenomena, taking place in biologically very disparate groups, as are vascular plants (Aeschimann et al., 2011; Hughes and Eastwood, 2006; Molina-Venegas et al., 2015; Pouchon et al., 2018), birds (Cai et al., 2018; Päckert et al., 2012; Quintero and Jetz, 2018), and insects (Chen et al., 2009; Merckx et al., 2015).

General introduction

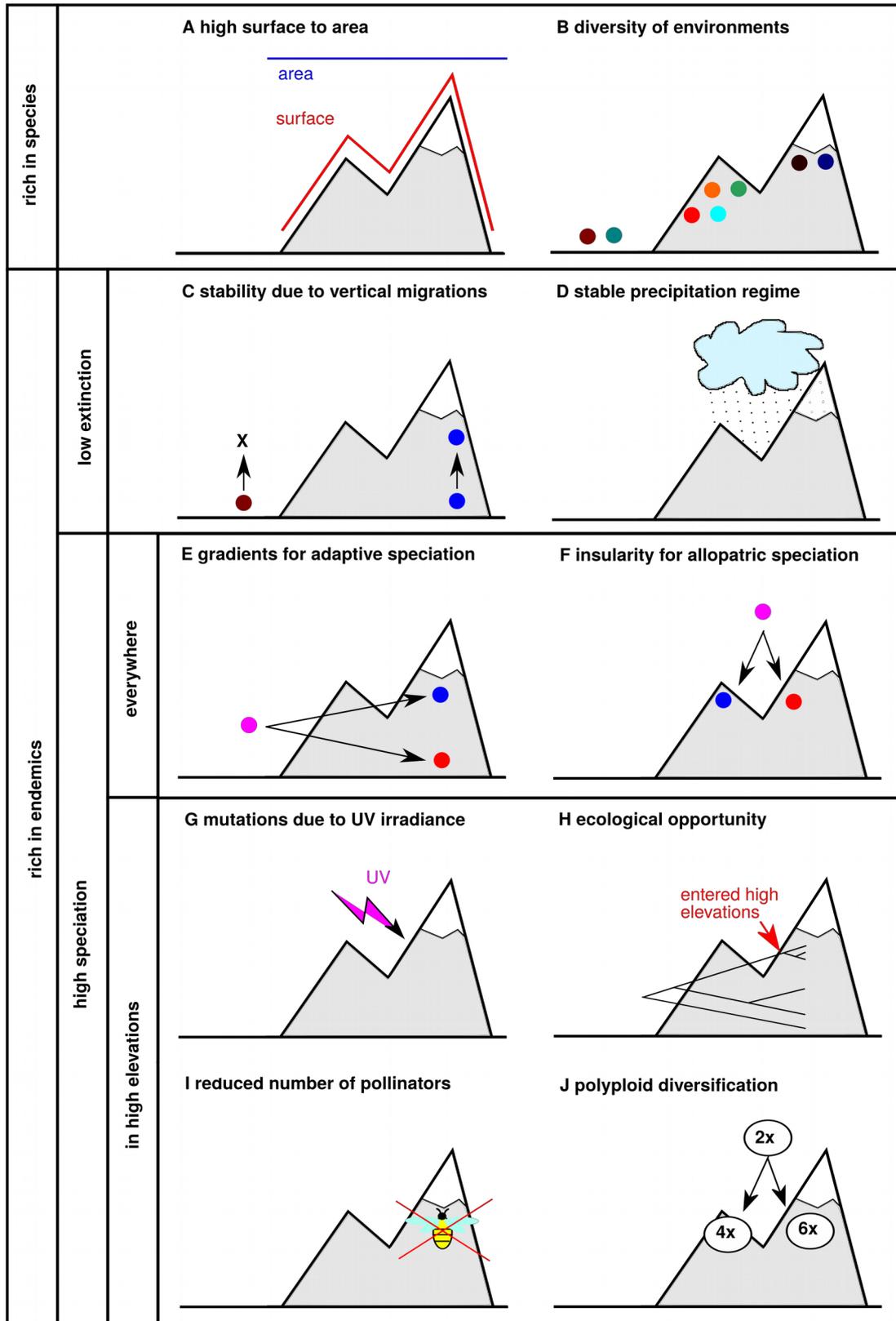


Figure 1: Explanations why mountains are rich in species (A-B) and in endemics (C-J). The explanations are further categorized based on whether they work with extinction (C-D) or speciation (E-J), and whether they are valid for mountain biotas as a whole (E-F) or only for high elevation biotas (G-J).

How to study diversification

The studies of diversification in mountains are often limited by availability of data, due to costs associated with gathering molecular phylogenetic datasets, and also due to limited accessibility of mountains complicating the collection of ecological or occurrence data in field. The choices of methodology, geographic and phylogenetic scale of the study are often driven by practical structure of available data, rather than theoretical assumptions about the study system. Such choices can however strongly influence the conclusions of evolutionary studies (Graham et al., 2018), and it is thus important to be aware of different methodological approaches, and of their strengths and weaknesses. In the literature, we can find two conceptual approaches regarding how to study diversification with data gathered from recent biota, each of them aiming on answering questions at different scales, i.e. with different amount of detail traded for potential of generalization.

The first approach, which could be called systematic or clade-oriented, consists in the study of diversification rates of a particular evolutionary lineage (Fig. 2A, e.g. genus *Androsace* in Roquet et al., 2013). A special case of systematic studies are phylogeographic studies, focusing on survival mechanisms within single species or diversification within single cladogenetic event, mapping for example divergence patterns between two newly emerging species at the level of individual populations (e.g. Kolář et al., 2016). The classical clade-oriented studies operating with multiple species aim on reconstructing their phylogenetic relationships using molecular data, optimally with a complete sampling of all species of the lineage. The phylogenetic trees are used to estimate diversification rates, their shifts and dynamics according to study-specific hypotheses. This can be done with process-based models considering diversification as a mechanistic branching process of the phylogenetic tree (reviewed in Morlon, 2014), as is Birth-Death model (BD; Nee et al., 1994), its modifications like State-dependent Speciation Extinction models (SSE; Maddison et al., 2007) or Bayesian Analysis of Macroevolutionary Mixtures (BAMM, Rabosky et al., 2014). The advantage of these modeling approaches is that they can often directly stand for tested hypotheses, and as other process-based models, they can also be used for hind- or forecasting. The process-based diversification models can also in theory estimate separately speciation and extinction rates, although this ability in real studies has been questioned (Rabosky, 2010). A complementary modeling approach is the use of non-parametric estimators of lineage

General introduction

diversification as the gamma statistic (Pybus and Harvey, 2000), various designs of comparisons between sister species or lineages (Boucher et al., 2016), visual inspection of shapes of phylogenies (Vargas, 2003) or lineage-through-time plots (LTT; Kadereit et al., 2004). A great example of the systematic approach to diversification of European alpine flora is the study of the genus *Phyteuma* (Schneeweiss et al., 2013). In this study, the authors use molecular data to build a species level phylogeny and analyze it via mechanistic diversification model to conclude that the diversification process in genus *Phyteuma* is strongly driven by geographic range splits at shallow scales and by ploidy dynamics at deeper scales. The power of systematic studies lies in the highly resolved phylogenetic relationships and the possibility for mechanistic modeling of diversification dynamics of the selected lineage. On the other hand, as the focal lineage is selected non-randomly, these studies lack potential for generalization of findings beyond the focal lineage.

The other approach, which could be called the community-oriented, is to study diversification across a broader selection of evolutionary lineages, either all lineages present in the biota in certain region, or their stratified subsample (Fig. 2B, e.g. Molina-Venegas et al., 2015). This typically means that evolutionary information for each of the lineages is much less precise than in systematic studies, because constructing well resolved phylogenies across many lineages would be too costly and time-consuming. The emphasis on the community also means that phylogenetic information is gathered only for species that occur in the focal area, leading to paraphyletic sampling of lineages, because most areas are systems open to immigration and emigration. These two properties of the data mean that mechanistic diversification models or their nonparametric alternatives cannot be used here, because they require high quality phylogenetic data and either complete species sampling of the lineage or at least random subselection of species (Fitzjohn et al., 2009). The analytical inference of community-oriented studies is thus limited to the exploration of spatial, ecological or temporal patterns of diversity indices that include the information about the evolutionary history, as are various phylogenetic diversity indices (reviewed in Tucker et al., 2016) such as phylogenetic diversity, mean terminal branch length (Feng et al., 2016) or more sophisticated measures like phylogenetic endemism (Mishler et al., 2014; Rosauer et al., 2009). A specific case are studies that test diversification hypotheses via patterns of diversity indices that do not explicitly include phylogenetic information, as are different measures of endemism (Pawłowski, 1970; Steinbauer et al., 2016; Tribsch and Schönswetter, 2003). To do this, they

General introduction

need to adopt the assumption that higher endemism indicates faster speciation or on the other hand slower extinction. An example of study using phylogenetic diversity indices to explore diversification patterns is Molina-Venegas et al. (2015). In this study the patterns of diversification across elevational, bedrock and geographic gradients in mountains of southern Spain are explored via phylogenetic alpha and beta diversity, to conclude that the strongest contributors to species diversification are geographic isolation and bedrock diversity, rather than elevational gradient. While the main power of the community-oriented approach lies in the study of a whole region and thus should allow generalizations, the main disadvantage lies in lower sensitivity due to phylogenetic data quality and also in the inability to efficiently separate different components generating diversity patterns, as are speciation, extinction or migration.

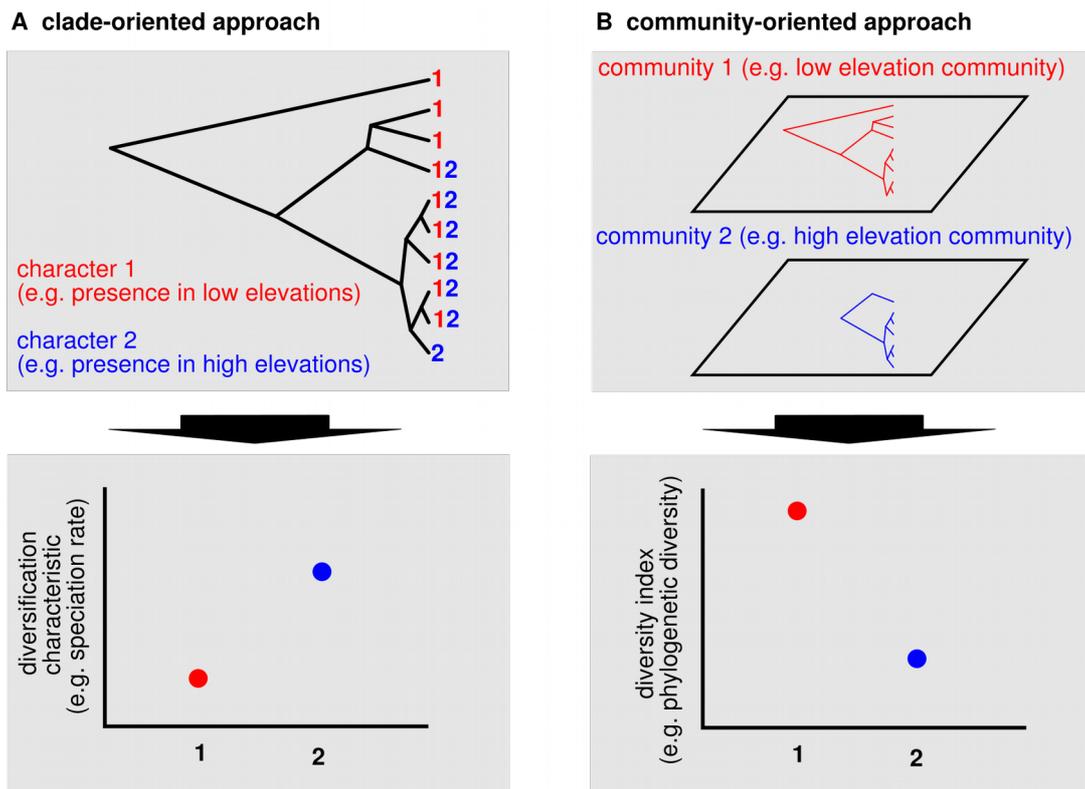


Figure 2: A schematic example comparing the diversification process in high and low elevation habitats using A clade-oriented approach or B community-oriented approach.

General introduction

Both clade- and community-oriented approaches have their strengths and weaknesses, making them more appropriate for answering different types of questions. The clade-oriented approach is well suited for testing existential statements and the clade-oriented studies constitute the case examples precisely mapping what is possible in biological evolution, rather than representative insights to region or ecosystem wide processes. In turn, the community-oriented approach is suitable for estimating dominant drivers of diversification in whole regions or ecosystems, but has little sensitivity for addressing details of this process or to identify its outliers. Recently, different strategies have emerged to fill in the gap between the two approaches, providing more generality to the clade-oriented approach or more precision to the community-oriented approach. These advances are fueled by the increasing availability (and decreasing cost) of molecular data on one hand, and recent methodological developments that allow to test complex evolutionary scenarios on the other hand.

The first set of approaches stems from clade-oriented studies, extending them to make their conclusions more general. One obvious way to attain this is simply making the clade-oriented study larger. For example, well resolved phylogenetic datasets containing large majority of birds or mammals across the world are available these days, and clade-oriented studies on such datasets (e.g. Machac and Graham, 2016) require no further generalization to draw conclusions about bird or mammal evolution. In the groups of organisms for which global phylogenies are not available, an obvious alternative is to conduct a meta-analysis comparing diversification dynamics from multiple systematic studies. For instance, Kadereit (2017) attempted to use meta-analysis of mountain plant diversification studies in the Alps; however, this type of approach is highly limited by the fact that systematic studies often use different sets of methods and molecular data for phylogenetic reconstructions, which makes their results difficult to compare in a formal way. This problem is particularly at stake when intending to compare the relative timing of diversification in different clades, as phylogeny dating will often be influenced by study-specific data, assumptions and methods. A way to go around this issue is to collect common ground molecular data for multiple lineages, and use them for constructing phylogenetic trees with one standardized method. Such set of trees can be analyzed one-by-one, with informal interpretations of common trends or differences (Boucher et al., 2016), or fitting a single diversification model to all lineages simultaneously (Anacker et al., 2011; Xing and Ree, 2017). This kind of approach is now being made

General introduction

possible by the massive rise of biodiversity genomics approaches, which intend to sequence large amount of species in parallel to biodiversity surveys.

The second set of approaches aims to construct a process-based community model that considers evolutionary processes, and fit this type of model to multi-lineage data, explicitly considering that evolutionary processes may involve migration and speciation outside of the focal community. An example of such a model is DAISIE (Valente et al., 2015), a methodological approach that was developed to study bird communities on Galapagos islands, consisting of multiple evolutionary lineages with independent migration history. A high potential in this direction lies also in the joint species distribution modeling framework that can currently accommodate competitive relationships among species in a community by parameterization of co-occurrence matrix, and there is room for accommodating species histories as well (Lawing and Matzke, 2014).

All these approaches, whether stemming from clade- or community- oriented approach, are adding precision to the estimates of overall diversification processes. But at the same time, they also open a new perspective, which is studying the variability of diversification. In particular, they provide the framework for testing whether the diversification dynamics are repeatable across different evolutionary lineages and to explore how the diversification processes in particular lineages translate into region-wide or ecosystem-wide diversification dynamics, which are the ones that can be directly linked the biodiversity patterns.

European mountains as a study system

The mountains in Europe (see Box 2 for definition) are rich in plant species compared to surrounding lowlands, and with the exception of arctic regions, also particularly rich in range-limited (endemic) species (Nagy and Grabherr, 2009). The high level of endemism suggests that processes involving better survival or faster speciation (reviewed above) took an important part in the formation of the European mountain flora. Although species-rich in European context, the European mountain system is poor in comparison to mountain ranges like the Northern Andes or Himalayas, and does not host the spectacularly fast diversifications known from these regions (Kadereit, 2017). There are many explanations why temperate mountains are less species rich than tropical ones, and in fact many general explanations for latitudinal diversity gradients can possibly be applied to mountains. A mountain-specific explanation is that temperate mountains are less species rich and their biota diversifies slower, because of seasonality causing less stratified altitudinal zonation (Janzen, 1967), affecting both ecological and allopatric speciation process. But there are also other specifics of European mountains, including the origin of mountain plant lineages or geological and glaciation history of the region. To better understand the diversification process in the European mountain system, it is thus necessary to take the historical perspective.

Most of the European mountain system emerged as a result of the Alpine orogeny that started in late Mesozoic (Moore and Fairbridge, 1997), peaked in early Cenozoic and, with lower intensity, still continues (Ager, 1975). There are however some exceptions: several middle mountain ranges resulted from earlier Caledonian and Hercynian orogenies, and others, for example Massif Central in France or southern Apennines, have a volcanic origin (Nagy and Grabherr, 2009). The orogenetic activity not only created rugged landscapes, but also originated high elevation areas where the early Cenozoic lowland flora could not live for climatic reasons (Fauquette et al., 2018). These vacant new habitats were colonized by plants from two main sources: via local adaptation of lineages already present in Europe (Comes and Kadereit, 2003), as is the case for e.g. *Globularia*, where the closest relatives to alpine species dwell in lowland areas of mediterranean Europe; and by migration of already adapted mountain lineages that originated in Asia and typically migrated through the mountain ranges of Middle East (Kadereit et al., 2008), which is the case of e.g. multiple lineages of *Gentiana*

General introduction

(Favre et al., 2016). Importantly, many of these lineages further diversified *in situ* in European mountains, but there is little consensus about the dominant speciation mechanisms and the timing of diversification. For example, it has been shown that mountain lineages of Primulaceae diversified mostly via allopatric speciations (Boucher et al., 2016), whereas ecological or polyploid speciation might have played an important role in other lineages (Moore and Kadereit, 2013; Pachschröll et al., 2015; Schneeweiss et al., 2013). It is generally assumed that the *in situ* diversification of European mountain flora started later than in Himalaya, because of multiple pre-adapted lineages migrating to Europe from this region and none in the opposite direction (Kadereit et al., 2008). The dating of the onset of plant diversification in European mountains is however fairly variable, with some lineages starting as early as 20 Ma BP (Favre et al., 2016), and other relatively recently.

A more recent factor that likely had an important influence in the diversity and the diversification dynamics of the European mountain flora are the Pleistocene climatic oscillations. As a mountain system lying in relatively high latitudes, the European mountain biota was significantly impacted by the changes of climate during the Quaternary. In fact, the Pleistocene climatic oscillations is probably the most discussed factor of current diversity patterns of European mountain flora, with discussions of this topic starting as early as 19th century (Brockmann-Jerosch and Brockmann-Jerosch, 1926; Darwin and Seward, 1903; de Candolle, 1875). The extent of glaciated areas varies between different cold periods (glacials), but for example in the maximum of the latest glacial period at cca 25 ka BP, the glaciers almost completely covered the Alps (Seguinot et al., 2018), the continental ice sheet covered the Scandinavian peninsula (Brochmann et al., 2003). In the other European mountains local glaciation appeared (Delmas, 2015; Kłapyta and Zasadni, 2018; Nývlt et al., 2011; see Fig. 3) and the permanent snow line shifted significantly downwards, often close to or below the current lower limit of high elevation habitats (Ronikier, 2011; Schönswetter et al., 2005). On the other hand, the foothills of European mountains constituted in some cases refugial areas with precipitation regimes more favorable for the interglacial flora, especially trees, than lowland dry steppes (Birks and Willis, 2008). It is important to point out that during the Quaternary, glacial periods prevailed, whereas interglacials periods (and the current Holocene period) were much shorter, and thus the latter periods can be seen as climatic deviations of the pre-Quaternary-like conditions from the generally cold and dry climate.

General introduction

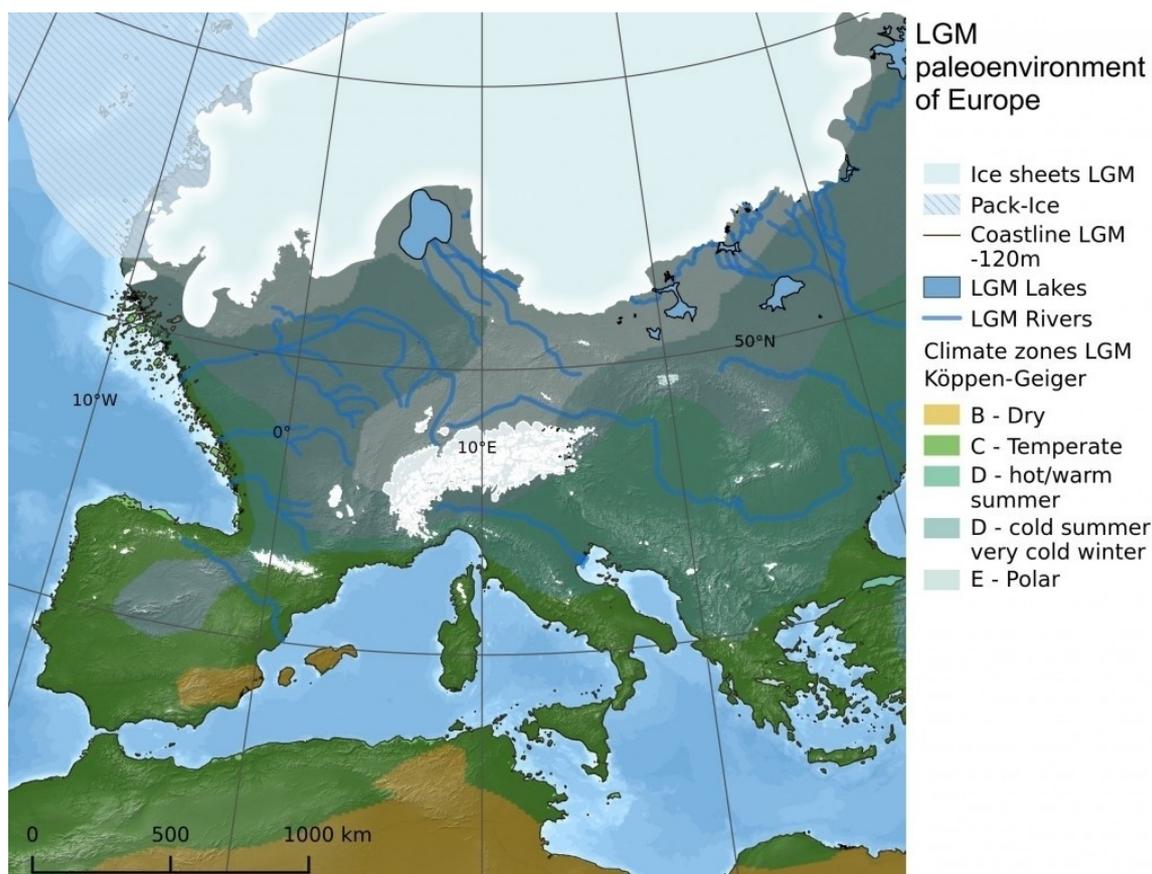


Figure 3: Glaciation extent and climatic zones in Europe during the last glacial maximum according to compilation from Becker et al. (2015). The glaciation in western Carpathians was modified according to Kłapyta and Zasadni (2018) and in Sudetes according to Nývlt et al. (2011).

How and where temperate mountain and high elevation plants survived ice ages of the Pleistocene is a long-standing debate. Three main mechanisms of survival have been proposed for mountain plants. The *in situ* glacial survival hypothesis, also known as the nunatak refugia hypothesis, suggests that mountain plants might have survived *in situ*, on rocky formations protruding above the ice sheet (so-called nunataks; Escobar García et al., 2012; Lohse et al., 2011; Schönswetter and Schneeweiss, 2019; Stehlik, 2003; Westergaard et al., 2019). Although the nunatak refugia hypothesis is formulated with regard to the Alps or Scandinavian mountains and glaciers, but *in situ* survival mechanism can be also hypothesized for less glaciated European mountains and survival above climatic snow line – it would suggest that mountain plants might have been surviving on microclimatic refugia above the identified extents of permanent snow line. The *in situ* survival implies that range and population sizes of surviving species would have shrunk considerably during glacial

General introduction

periods, with obvious consequences for population dynamics, local extinction rates and floristic composition. However, the *in situ* survival might have occurred in concert with other mechanisms, which could partially blur its floristic and population genetic signatures (Escobar García et al., 2012; Schönswetter and Schneeweiss, 2019). Importantly, it is often assumed that the *in situ* survival mechanism might have taken place only for species able to withstand harsh conditions in highest elevational belts, because such conditions presumably took place on the nunataks and high elevation microclimatic refugia during glaciation periods (Holderegger et al., 2011).

The second survival hypothesis, referred to as peripheral refugia hypothesis (sometimes called *tabula rasa* or *massif de refuge* hypothesis together with the third one), suggests that with the onset of glaciation periods, the mountain plants shifted the distribution towards the foothills of mountains, where the climatic conditions were suitable for their survival (Schönswetter et al., 2005; Tribsch and Schönswetter, 2003). Some of the mountain species might also inhabit broader regions in the glacial lowland, although the climatic conditions in glacial steppes were not paralleled to current alpine or arctic habitats, mostly due to low precipitation (Birks and Willis, 2008). The prevalence of glacial contraction vs expansion of ranges of cold adapted species is a subject of discussions (Stewart et al., 2010; Theodoridis et al., 2017), but a common assumption is that broad glacial ranges were prevalent in species with currently arctic-alpine distribution, because glacial lowland populations colonized not only mountains, but also adjacent arctic regions. On the other hand, for species endemic to central and southern European mountain systems, it is assumed that they survived at the foothills or in proximity of mountain ranges (Schönswetter et al., 2005; Tribsch and Schönswetter, 2003), with glacial range sizes comparable or smaller than interglacial ones. The subsequent recolonization of central part of the mountains was dominantly gradual, creating isolation-by-distance genetic signatures within species (Schönswetter et al., 2005), with break zones between areas recolonized from different refugia (Thiel-Egenter et al., 2011). The gradual recolonization process is detectable also in species composition, the formerly glaciated regions are less saturated by species than would correspond to their environmental conditions and this pattern gradually increases with distance from refugia (Dullinger et al., 2012). However in specific cases, the recolonization from peripheral refugia might have taken place via long distance dispersal events, reflecting the rugged topography and insularity of mountain environments (Schönswetter et al., 2005).

General introduction

The third hypothesis, referred to as southern refugia hypothesis (sometimes called *tabula rasa* or *massif de refuge* hypothesis together with the second one), suggests that the mountain ranges of northern and central Europe were recolonized by plants surviving glacial periods in mountain ranges in southern Europe, which were less impacted by glaciation. This mechanism was proposed for mid elevation mountain species (Zhang et al., 2001), and is basically equivalent to patterns observed in many European lowland species (Stewart et al., 2010). Due to insularity of European mountain system and large distance between southern and central European mountain ranges, an important assumption of this hypothesis is the high geographic mobility and capacity for long distance dispersal of species that have survived with this mechanism.

Under any of the three survival hypotheses it is likely that the Pleistocene climatic oscillations modified diversification dynamics of mountain plant clades. This could be due to increased extinction rates via range reductions and geographic shifts of suitable habitats, or on the other hand increased speciation via cyclic changes of habitat size and connectivity. One diversification scenario is that the majority of *in situ* speciation events in the European mountain flora took place before Pleistocene, whereas during the Pleistocene the speciation rates slowed down and/or the extinction rates increased due to the reduction of high elevation habitats, resulting in an impoverished flora compared to pre-Pleistocene times (Boucher et al., 2016; Kadereit et al., 2004; Ozenda, 1995). For example, *Primula* sect. *Auricula* and *Androsace* sect. *Aretia* were indeed found to slow down diversification during the Quaternary period (Boucher et al., 2016). However, it is difficult to disentangle whether such trends are specifically due to climate-triggered speciation slowdown, climate triggered extinction speedup, or intrinsic diversification slowdown due to saturation of ecological space.

Another scenario could be that Pleistocene climatic oscillations stimulated diversification, with a large portion of species emerging during the Pleistocene. The most likely mechanism stimulating the species origination would be the speciation pump (*sensu* Haffer, 1967), which suggests that temporal changes of habitat size and connectivity in island-like systems may promote allopatric speciations. While this mechanism was originally proposed for isolation in physical space, we can hypothesize that the isolation in mountain systems could also take place across ecological gradients – for instance, if both nunatak and peripheral survival occurred in one species, this could have favored the isolation of cold-adapted populations on nunataks from warm-adapted populations in peripheral refugia (DeChaine and Martin, 2006).

General introduction

The detection of diversifications during Pleistocene may be hindered by the fact that evolutionary lineages consisting of incipient species emerging from physical isolation events may be systematically poorly taxonomically described, or attributed to intraspecific taxonomic ranks (Vargas, 2003). This may be for example the case of *Soldanella*, which is a lineage that was often used as an example for rapid Pleistocene diversification (Boucher et al., 2016; Comes and Kadereit, 2003), but recent molecular and morphological evidence shows that the systematics and taxonomy of this genus needs to be substantially revised (Slovák, Smyčka et al., in prep.).

The mechanisms of survival and dynamics of diversification in Pleistocene may also differ for different ecological groups of mountain plants. For the European mountain flora, it is frequently discussed that diversification and survival processes might differ between plant species specialized to calcareous and siliceous bedrocks (Dullinger et al., 2012; Tribsch and Schönswetter, 2003). The bedrock preferences in European mountain plants are often quite distinct with many known specialists to each type of bedrocks (Moore and Kadereit, 2013; Pawłowski, 1970; Schneeweiss et al., 2013). At the same time, the calcareous and siliceous areas are non-randomly distributed across the European mountains, with calcareous areas mostly found at the periphery of mountain ranges and also in the southern part of the European mountain system. This together suggests that the survival and diversification dynamics in Pleistocene may have had fairly different regimes for calcareous and siliceous specialist species. Calcicolous lineages might have had relatively abundant refugia at the periphery of mountain ranges, in southern Europe, or in lowland steppes (Ewald, 2003) during the ice ages, whereas their habitats were more fragmented during interglacials; on the other hand, silicolous lineages might have had few rare refugia available during glacial periods, but on the other hand had large and continuous ranges with suitable habitat available in the central parts of the mountains during the interglacials. These different dynamics of calcicolous and silicolous lineages were supported by studies on patterns of endemism (Tribsch and Schönswetter, 2003) or species genetic structures (Alvarez et al., 2009; Thiel-Egenter et al., 2011) in the Alps, focusing on groups of species with clear bedrock affinities. However, the role of bedrock in differentiation in European mountain flora, including the stability of bedrock preferences and role of this gradient in stimulating diversification, is yet to be understood.

General introduction

Box 2 Delimitation of European mountain system

I delimit here the European mountain system as all the mountain ranges (using the ruggedness definition *sensu* Körner et al., 2011) on the European continent west from 30th meridian. With this definition, the Caucasus and Ural are not considered as a part of European mountain system, because of much tighter biogeographic connections with Central Asian and Middle Eastern mountain ranges than with the rest of European mountain system (Kadereit et al., 2008; Nagy and Grabherr, 2009), but also due to practical difficulties connected with conducting research in this region. Such defined European mountain system thus consists of several mountain ranges with the Alps as the central and highest mountain range (max. 4810 m a.s.l.); mountain ranges of southern Europe, as are Apennines, Pyrenees or Sierra Nevada; mountain ranges of northern Europe, as are Scandinavian mountains, with tight biogeographic connections to the flora of arctic lowlands; and mountain ranges of south eastern Europe, as are mountains in the Balkans peninsula or Carpathians, possessing a migration connection to the mountain of Middle East and further Asia.

Thesis aims and structure

In this thesis I explore the processes driving diversification in a broad sense, i.e. species origination and extinction, in the European mountain system. In particular I focus on patterns and dynamics of mountain plant survival and diversification in response to Pleistocene climatic oscillations, which is important for shedding more light on future of mountain flora in face of ongoing climatic changes. Further I focus on drivers of mountain plant speciation and their evolutionary assembly across environmental gradients. Understanding these drivers is useful for explaining why are the European mountains richer in species and in endemics than the surrounding lowlands, and also to compare whether these explanations are the same as in mountains less impacted by Pleistocene climatic oscillations.

As reviewed in previous sections, many studies have been published about survival or diversification of plants in the European mountain system, but all of them were fairly limited by their geographic and evolutionary resolution or scope (or grain and extent in ecological terminology). For example, there are studies that explored the spatial or ecological patterns of endemism in the Alps (Pawłowski, 1970; Tribsch and Schönswetter, 2003), or in other European mountains (Molina-Venegas et al., 2015; Mráz et al., 2016), but without more explicit evolutionary information, they could not resolve whether these patterns indicate specific processes of survival (e.g. glacial refugia) or speciation (e.g. diversification hotspots). The diversification was studied explicitly in particular plant lineages that diversified in the European mountain system, notably *Phyteuma* (Schneeweiss et al., 2013), *Androsace* (Boucher et al., 2016; Roquet et al., 2013) and *Primula* sect. *Auricula* (Boucher et al., 2016); however, the fact that these studies were limited to one or very few lineages and that different types of data and methods were used for phylogenetic reconstruction, do not allow for an analytical and quantitative comparative study (Kadereit, 2017). The case studies on population genetic diversity and structure of mountain species provide valuable information about diversification and survival mechanisms at even finer evolutionary scale (Escobar García et al., 2012; Lohse et al., 2011; Schönswetter et al., 2003). Recently there have been a few studies synthesizing genetic diversity patterns across different species and populations in the European mountains (Alvarez et al., 2009; Schönswetter et al., 2005; Taberlet et al., 2012), however they were limited by low data resolution and an analytical approach. In this thesis, I benefit from recent advances in computational and molecular

General introduction

biology, allowing me to address the questions concerning diversification in European mountain flora across multiple lineages and with better phylogenetic resolution than previously possible. My thesis consists of three chapters:

In chapter one, I investigate the drivers of plant endemism and phylogenetic diversity in the Alps using a phylogenetic and spatially explicit approach. More specifically, I use here a genus-level phylogeny of the flora of the Alps coupled with species distributional data on a grain of 20 by 20 km to explore the potential links between presence of ice-free areas during glacial periods, mountain topography, bedrock and different measures of endemism and phylogenetic diversity. The explicit phylogenetic information proves as extremely important, because it allows me to discriminate the areas of paleoendemism and neoendemism. This chapter 1 has been published as an article in the journal *Perspectives in Plant Ecology, Evolution and Systematics*.

In chapter two, I explore the mechanisms and dynamics of diversification in six plant lineages that were formally selected as representatives of large *in situ* diversifications within the European mountain system. To do so, I reconstructed highly resolved phylogenies based on chloroplast genome sequences, and I use these for modeling of past diversification dynamics, evolution of elevational and bedrock niche and geographic ranges. In order to investigate general diversification patterns, in addition to lineage-specific ones, I develop here a method to launch a single likelihood-based diversification model on several lineages simultaneously. This chapter corresponds to a manuscript that will be submitted to the journal *Proceedings of the National Academy of Sciences of the United States of America*.

In chapter three, I study the population genetic diversity and structure of *Androsace argentea* (former *A. vandelli*) based on reduced representation genomic data (ddRADseq) and a dated chloroplast genome phylogeny of different populations. This species is an interesting study case to test hypotheses about glacial refugia because it is a species occurring on compact cliffs from subalpine to nival zone in very high elevations in the Alps, and moreover it is a species widely distributed in the European mountain system (Alps, Pyrenees, Sierra Nevada), which means that all the three glacial survival hypothesis might take place in this species. I apply here a comparative approach to unravel the origin and past demography of three high elevation populations in the French Alps. This chapter corresponds to a manuscript in preparation, delayed due to methodological difficulties that appeared during the production of

General introduction

the ddRADseq dataset. The currently presented results identify the origins and overall demographic history of the focal populations, which are to be confirmed via process-based models as I indicate in the section “Future directions” of this chapter.

General introduction

References

- Aeschimann, D., Rasolofoa, N., Theurillat, J.P., 2011. Analyse de la flore des Alpes. 2: diversité et chorologie. *Candollea* 66, 225–253.
- Ager, D. V., 1975. The geological evolution of Europe. *Proc. Geol. Assoc.* 86, 127–154. [https://doi.org/10.1016/S0016-7878\(75\)80097-6](https://doi.org/10.1016/S0016-7878(75)80097-6)
- Alvarez, N., Thiel-Egenter, C., Tribsch, A., Holderegger, R., Manel, S., Schönschwetter, P., Taberlet, P., Brodbeck, S., Gaudeul, M., Gielly, L., Küpfer, P., Mansion, G., Negrini, R., Paun, O., Pellicchia, M., Rioux, D., Schüpfer, F., Van Loo, M., Winkler, M., Gugerli, F., 2009. History or ecology? Substrate type as a major driver of spatial genetic structure in Alpine plants. *Ecol. Lett.* 12, 632–640. <https://doi.org/10.1111/j.1461-0248.2009.01312.x>
- Anacker, B.L., Whittall, J.B., Goldberg, E.E., Harrison, S.P., 2011. Origins and consequences of serpentine endemism in the California flora. *Evolution* (N. Y.) 65, 365–376. <https://doi.org/10.1111/j.1558-5646.2010.01114.x>
- Ball, J., 1879. On the Origin of the Flora of the European Alps. *Proc. R. Geogr. Soc. Mon. Rec. Geogr.* 1, 564–589.
- Becker, D., Verheul, J., Zickel, M., Willmes, C., 2015. LGM paleoenvironment of Europe - Map. CRC806-Database. <https://doi.org/10.5880/SFB806.15>
- Birks, H.J.B., Willis, K.J., 2008. Alpines, trees, and refugia in Europe. *Plant Ecol. Divers.* 1, 147–160. <https://doi.org/10.1080/17550870802349146>
- Boucher, F.C., Zimmermann, N.E., Conti, E., 2016. Allopatric speciation with little niche divergence is common among alpine Primulaceae. *J. Biogeogr.* 43, 591–602. <https://doi.org/10.1111/jbi.12652>
- Brochmann, C., Gabrielsen, T.M., Nordal, I., Landvik, J.Y., Elven, R., 2003. Glacial survival or tabula rasa? The history of North Atlantic biota revisited. *Taxon* 52, 417–450. <https://doi.org/10.2307/3647444>
- Brockmann-Jerosch, M., Brockmann-Jerosch, M., 1926. Die Geschichte der Schweizerischen Alpenflora., in: Schröter, C. (Ed.), *Das Pflanzenleben Der Alpen*. Raustein, Zürich, pp. 1110–1215.
- Cai, T., Fjeldså, J., Wu, Y., Shao, S., Chen, Y., Quan, Q., Li, X., Song, G., Qu, Y., Qiao, G., Lei, F., 2018. What makes the Sino-Himalayan mountains the major diversity hotspots for pheasants? *J. Biogeogr.* 45, 640–651. <https://doi.org/10.1111/jbi.13156>
- Chen, I.-C., Shiu, H.-J., Benedick, S., Holloway, J.D., Chey, V.K., Barlow, H.S., Hill, J.K., Thomas, C.D., 2009. Elevation increases in moth assemblages over 42 years on a tropical mountain. *Proc. Natl. Acad. Sci.* 106, 1479–1483. <https://doi.org/10.1073/pnas.0809320106>
- Chesson, P., 2000. Mechanisms of Maintenance of Species Diversity. *Annu. Rev. Ecol. Syst.* 31, 343–366. <https://doi.org/10.1146/annurev.ecolsys.31.1.343>
- Comes, H.P., Kadereit, J.W., 2003. Spatial and Temporal Patterns in the Evolution of the Flora of the European Alpine Systems. *Taxon* 52, 451–562.
- Condamine, F.L., Rolland, J., Höhna, S., Sperling, F.A.H., Sanmartín, I., 2018. Testing the role of the Red Queen and Court Jester as drivers of the macroevolution of Apollo butterflies. *Syst. Biol.* 67, 940–964. <https://doi.org/10.1093/sysbio/syy009>
- Condamine, F.L., Rolland, J., Morlon, H., 2013. Macroevolutionary perspectives to environmental change. *Ecol. Lett.* 16, 72–85. <https://doi.org/10.1111/ele.12062>
- Cortés, A.J., Garzón, L.N., Valencia, J.B., Madriñán, S., 2018. On the Causes of Rapid Diversification in the Páramos: Isolation by Ecology and Genomic Divergence in Espeletia. *Front. Plant Sci.* 9, 1–17. <https://doi.org/10.3389/fpls.2018.01700>
- Darwin, F., Seward, A. (Eds.), 1903. More letters of Charles Darwin. John Murray, London.
- de Candolle, A., 1875. Sur les causes de l'inégale distribution des plantes rares dans la chaîne des Alpes. *Atti del Congr. Internazionale Bot. Tenuto Firenze.* 92–104.
- DeChaine, E.G., Martin, A.P., 2006. Using Coalescent Simulations To Test the Impact of Quaternary Climate Cycles on Divergence in an Alpine Plant-Insect Association. *Evolution* (N. Y.) 60, 1004. <https://doi.org/10.1554/05-672.1>
- Delmas, M., 2015. The last maximum ice extent and subsequent deglaciation of the Pyrenees: an overview of recent research. *Cuad. Investig. Geográfica* 41, 359. <https://doi.org/10.18172/cig.2708>

General introduction

- Dullinger, S., Willner, W., Plutzer, C., Englisch, T., Schrott-Ehrendorfer, L., Moser, D., Ertl, S., Essl, F., Niklfeld, H., 2012. Post-glacial migration lag restricts range filling of plants in the European Alps. *Glob. Ecol. Biogeogr.* 21, 829–840. <https://doi.org/10.1111/j.1466-8238.2011.00732.x>
- Escobar García, P., Winkler, M., Flatscher, R., Sonnleitner, M., Krejčíková, J., Suda, J., Hülber, K., Schneeweiss, G.M., Schönswetter, P., 2012. Extensive range persistence in peripheral and interior refugia characterizes Pleistocene range dynamics in a widespread Alpine plant species (*Senecio carniolicus*, Asteraceae). *Mol. Ecol.* 21, 1255–1270. <https://doi.org/10.1111/j.1365-294X.2012.05456.x>
- Ewald, J., 2003. The calcareous riddle: Why are there so many calciphilous species in the Central European flora? *Folia Geobot.* 38, 357–366. <https://doi.org/10.1007/BF02803244>
- Fauquette, S., Suc, J.-P., Médail, F., Muller, S.D., G., J.-M., Bertini, A., Martinetto, E., Popescu, S.-P., Zheng, Z., de Beaulieu, J.L., 2018. The Alps: a geological, climatic and human perspective on vegetation history and modern plant diversity, in: Hoorn, C., Perrigo, A., Antonelli, A. (Eds.), *Mountains, Climate and Biodiversity*. Wiley-Blackwell, p. 413.
- Favre, A., Michalak, I., Chen, C.H., Wang, J.C., Pringle, J.S., Matuszak, S., Sun, H., Yuan, Y.M., Struwe, L., Muellner-Riehl, A.N., 2016. Out-of-Tibet: the spatio-temporal evolution of *Gentiana* (Gentianaceae). *J. Biogeogr.* 43, 1967–1978. <https://doi.org/10.1111/jbi.12840>
- Feng, G., Mao, L., Sandel, B., Swenson, N.G., Svenning, J.C., 2016. High plant endemism in China is partially linked to reduced glacial-interglacial climate change. *J. Biogeogr.* 43, 145–154. <https://doi.org/10.1111/jbi.12613>
- Fitzjohn, R.G., Maddison, W.P., Otto, S.P., 2009. Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Syst. Biol.* 58, 595–611. <https://doi.org/10.1093/sysbio/syp067>
- Fjeldså, J., Bowie, R.C.K., Rahbek, C., Fjeldså, J., Bowie, R.C.K., Rahbek, C., 2011. The Role of Mountain Ranges in the Diversification of Birds. *Annu. Rev. Ecol. Evol. Syst.* 43, 249–265. <https://doi.org/10.1146/annurev-ecolsys-102710-145113>
- Friedman, W.E., 2009. The meaning of Darwin’s “abominable mystery.” *Am. J. Bot.* 96, 5–21. <https://doi.org/10.3732/ajb.0800150>
- García-Camacho, R., Totland, Ø., 2009. Pollen Limitation in the Alpine: A Meta-Analysis. *Arctic, Antarct. Alp. Res.* 41, 103–111. <https://doi.org/10.1657/1523-0430-41.1.103>
- Graham, C.H., Storch, D., Machac, A., 2018. Phylogenetic scale in ecology and evolution. *Glob. Ecol. Biogeogr.* 27, 175–187. <https://doi.org/10.1111/geb.12686>
- Haffer, J., 1967. Speciation in Colombian forest birds west of the Andes. *Am. Museum Novit.* 1–58.
- Heald, W., 1967. *Sky Islands*. Van Nostrand Company, Princeton.
- Holderegger, R., Thiel-Egenter, C., Parisod, C., 2011. Marie Brockmann-Jerosch and her influence on Alpine phylogeography. *Alp. Bot.* 121, 5–10. <https://doi.org/10.1007/s00035-010-0086-9>
- Hughes, C., Eastwood, R., 2006. Island radiation on a continental scale: Exceptional rates of plant diversification after uplift of the Andes. *Proc. Natl. Acad. Sci.* 103, 10334–10339. <https://doi.org/10.1073/pnas.0601928103>
- Hughes, C.E., 2017. Are there many different routes to becoming a global biodiversity hotspot? *Proc. Natl. Acad. Sci.* 114, 4275–4277. <https://doi.org/10.1073/pnas.1703798114>
- Hughes, C.E., Atchinson, G.W., 2015. The ubiquity of alpine plant radiations: from the Andes to the Hengduan Mountains. *New Phytol.* 207, 275–282. <https://doi.org/10.1111/nph.13201>
- Janzen, D.H., 1967. Why mountain passes are higher in the tropics. *Am. Nat.* 101, 233–249.
- Jordon-Thaden, I., Koch, M., 2008. Species richness and polyploid patterns in the genus *Draba* (Brassicaceae): A first global perspective. *Plant Ecol. Divers.* 1, 255–263. <https://doi.org/10.1080/17550870802349112>
- Kadereit, J.W., 2017. The role of in situ species diversification for the evolution of high vascular plant species diversity in the European Alps—A review and interpretation of phylogenetic studies of the endemic flora of the Alps. *Perspect. Plant Ecol. Evol. Syst.* 26, 28–38. <https://doi.org/10.1016/j.ppees.2017.03.002>
- Kadereit, J.W., Griebeler, E.M., Comes, H., 2004. Quaternary diversification in European alpine plants: pattern and process. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 359, 265–274. <https://doi.org/10.1098/rstb.2003.1389>
- Kadereit, J.W., Licht, W., Uhink, C.H., 2008. Asian relationships of the flora of the European Alps. *Plant Ecol. Divers.* 1, 171–179. <https://doi.org/10.1080/17550870802328751>

General introduction

- Kłapyta, P., Zasadni, J., 2018. Research history on the Tatra mountains glaciations. *Stud. Geomorphol. Carpatho-Balcanica* 43–85.
- Kolář, F., Fuxová, G., Závěská, E., Nagano, A.J., Hyklová, L., Lučanová, M., Kudoh, H., Marhold, K., 2016. Northern glacial refugia and altitudinal niche divergence shape genome-wide differentiation in the emerging plant model *Arabidopsis arenosa*. *Mol. Ecol.* 25, 3929–3949. <https://doi.org/10.1111/mec.13721>
- Körner, C., 2000. Why are there global gradients in species richness? Mountains might hold the answer. *Trends Ecol. Evol.* 15, 513–514. [https://doi.org/10.1016/S0169-5347\(00\)02004-8](https://doi.org/10.1016/S0169-5347(00)02004-8)
- Körner, C., 1999. *Alpine Plant Life*. Springer, Berlin.
- Körner, C., Paulsen, J., Spehn, E.M., 2011. A definition of mountains and their bioclimatic belts for global comparisons of biodiversity data. *Alp. Bot.* 121, 73–78. <https://doi.org/10.1007/s00035-011-0094-4>
- Kostylev, V.E., Erlandsson, J., Mak, Y.M., Williams, G.A., 2005. The relative importance of habitat complexity and surface area in assessing biodiversity: Fractal application on rocky shores. *Ecol. Complex.* 2, 272–286. <https://doi.org/10.1016/j.ecocom.2005.04.002>
- La Sorte, F.A., Jetz, W., 2010. Projected range contractions of montane biodiversity under global warming. *Proc. R. Soc. B Biol. Sci.* 277, 3401–3410. <https://doi.org/10.1098/rspb.2010.0612>
- Lagamarsino, L.P., Condamine, F.L., Antonelli, A., Mulch, A., Davis, C.C., 2016. The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). *New Phytol.* 210, 1430–1442. <https://doi.org/10.1111/nph.13920>
- Lawing, A.M., Matzke, N.J., 2014. Conservation paleobiology needs phylogenetic methods. *Ecography (Cop.)* 37, 1109–1122. <https://doi.org/10.1111/ecog.00783>
- Lohse, K., Nicholls, J.A., Stone, G.N., 2011. Inferring the colonization of a mountain range-refugia vs. nunatak survival in high alpine ground beetles. *Mol. Ecol.* 20, 394–408. <https://doi.org/10.1111/j.1365-294X.2010.04929.x>
- Lomolino, M. V., 2001. Elevation gradients of species-density: Historical and prospective views. *Glob. Ecol. Biogeogr.* 10, 3–13. <https://doi.org/10.1046/j.1466-822x.2001.00229.x>
- Luebert, F., Weigend, M., 2014. Phylogenetic insights into Andean plant diversification. *Front. Ecol. Evol.* 2, 1–17. <https://doi.org/10.3389/fevo.2014.00027>
- Machac, A., Graham, C.H., 2016. Regional Diversity and Diversification in Mammals. *Am. Nat.* 189, E1–E13. <https://doi.org/10.1086/689398>
- Maddison, P., Midford, E., Otto, P., 2007. Estimating a Binary Character's Effect on Speciation. *Syst. Biol.* 56, 701–710. <https://doi.org/10.1080/10635150701607033>
- Merckx, V.S.F.T., Hendriks, K.P., Beentjes, K.K., Mennes, C.B., Becking, L.E., Peijnenburg, K.T.C. a., Afendy, A., Arumugam, N., de Boer, H., Biun, A., Buang, M.M., Chen, P.-P., Chung, A.Y.C., Dow, R., Feijen, F. a. a., Feijen, H., Soest, C.F., Geml, J., Geurts, R., Gravendeel, B., Hovenkamp, P., Imbun, P., Ipor, I., Janssens, S.B., Jocqué, M., Kappes, H., Khoo, E., Koomen, P., Lens, F., Majapun, R.J., Morgado, L.N., Neupane, S., Nieser, N., Pereira, J.T., Rahman, H., Sabran, S., Sawang, A., Schwallier, R.M., Shim, P.-S., Smit, H., Sol, N., Spait, M., Stech, M., Stokvis, F., Sugau, J.B., Suleiman, M., Sumail, S., Thomas, D.C., van Tol, J., Tuh, F.Y.Y., Yahya, B.E., Nais, J., Repin, R., Lakim, M., Schilthuizen, M., 2015. Evolution of endemism on a young tropical mountain. *Nature* 524, 347–350. <https://doi.org/10.1038/nature14949>
- Mishler, B.D., Knerr, N., González-Orozco, C.E., Thornhill, A.D., Laffan, S.W., Miller, J.T., 2014. Phylogenetic measures of biodiversity and neo- and paleo-endemism in Australian *Acacia*. *Nat. Commun.* 5, 4473. <https://doi.org/10.1038/ncomms5473>
- Molina-Venegas, R., Aparicio, A., Slingsby, J.A., Lavergne, S., Arroyo, J., 2015. Investigating the evolutionary assembly of a Mediterranean biodiversity hotspot: Deep phylogenetic signal in the distribution of eudicots across elevational belts. *J. Biogeogr.* 42, 507–518. <https://doi.org/10.1111/jbi.12398>
- Moore, A.J., Kadereit, J.W., 2013. The evolution of substrate differentiation in *Minuartia* series *Laricifoliae* (Caryophyllaceae) in the European Alps: In situ origin or repeated colonization? *Am. J. Bot.* 100, 2412–2425. <https://doi.org/10.3732/ajb.1300225>
- Moores, E., Fairbridge, R. (Eds.), 1997. *Encyclopedia of European and Asian Regional Geology*. Springer, Dordrecht.
- Morlon, H., 2014. Phylogenetic approaches for studying diversification. *Ecol. Lett.* 17, 508–525. <https://doi.org/10.1111/ele.12251>

General introduction

- Mráz, P., Barabas, D., Lengyelová, L., Turis, P., Schmotzer, A., Janišová, M., Ronikier, M., 2016. Vascular plant endemism in the Western Carpathians: spatial patterns, environmental correlates and taxon traits. *Biol. J. Linn. Soc.* 119, 630–648. <https://doi.org/10.1111/bij.12792>
- Myers, N., Mittermeier, R., Mittermeier, C., da Fonseca, G., Kent, J., 2000. Biodiversity hotspots for conservation priorities. *Nature* 403: 853, 853–858.
- Nagy, L., Grabherr, G., 2009. *The Biology of Alpine Habitats*. Oxford University Press, Oxford.
- Nee, S., May, R.M., Harvey, P.H., 1994. The reconstructed evolutionary process. *Philos. Trans. R. Soc. B Biol. Sci.* 344, 305–311. <https://doi.org/10.1098/rstb.1994.0068>
- Nývlt, D., Engel, Z., Tyráček, J., 2011. Pleistocene glaciations of Czechia, in: Ehlers, J., Gibbard, P., Hughes, P. (Eds.), *Quaternary Glaciations - Extent and Chronology*. Elsevier, Amsterdam, pp. 37–46. <https://doi.org/10.1016/B978-0-444-53447-7.00004-0>
- Ozenda, P., 1995. L'endémisme au niveau de l'ensemble du Système alpin. *Acta Bot. Gall.* 142, 753–762. <https://doi.org/10.1080/12538078.1995.10515302>
- Pachschwöll, C., García, P.E., Winkler, M., Schneeweiss, G.M., Schönswetter, P., 2015. Polyploidisation and geographic differentiation drive diversification in a European high mountain plant group (*Doronicum clusii* aggregate, Asteraceae). *PLoS One* 10, 1–30. <https://doi.org/10.1371/journal.pone.0118197>
- Päckert, M., Martens, J., Sun, Y.H., Severinghaus, L.L., Nazarenko, A.A., Ting, J., Töpfer, T., Tietze, D.T., 2012. Horizontal and elevational phylogeographic patterns of Himalayan and Southeast Asian forest passerines (Aves: Passeriformes). *J. Biogeogr.* 39, 556–573. <https://doi.org/10.1111/j.1365-2699.2011.02606.x>
- Pawłowski, B., 1970. Remarques sur l'endémisme dans la flore des Alpes et des Carpates. *Vegetatio* 21, 181–243.
- Pouchon, C., Fernández, A., Nassar, J.M., Boyer, F., Aubert, S., Lavergne, S., Mavárez, J., 2018. Phylogenomic analysis of the explosive adaptive radiation of the Espeletia complex (Asteraceae) in the tropical Andes. *Syst. Biol.* 67, 1041–1060. <https://doi.org/10.1093/sysbio/syy022>
- Pybus, O.G., Harvey, P.H., 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. *Proc. R. Soc. B Biol. Sci.* 267, 2267–2272. <https://doi.org/10.1098/rspb.2000.1278>
- Pyron, R.A., Wiens, J.J., 2013. Large-scale phylogenetic analyses reveal the causes of high tropical amphibian diversity. *Proc. R. Soc. B Biol. Sci.* 280, 1–10. <https://doi.org/10.1098/rspb.2013.1622>
- Quintero, I., Jetz, W., 2018. Global elevational diversity and diversification of birds. *Nature* 555, 246–250. <https://doi.org/10.1038/nature25794>
- Quintero, I., Wiens, J.J., 2013. Rates of projected climate change dramatically exceed past rates of climatic niche evolution among vertebrate species. *Ecol. Lett.* 16, 1095–1103. <https://doi.org/10.1111/ele.12144>
- Rabosky, D.L., 2010. Extinction rates should not be estimated from molecular phylogenies. *Evolution (N. Y.)* 64, 1816–1824. <https://doi.org/10.1111/j.1558-5646.2009.00926.x>
- 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. *Syst. Biol.* 63, 610–627. <https://doi.org/10.1093/sysbio/syu025>
- Rahbek, C., 1995. The elevational gradient of species richness: a uniform pattern? *Ecography (Cop.)* 18, 200–205. <https://doi.org/10.1111/j.1600-0587.1995.tb00341.x>
- Rolland, J., Condamine, F.L., Jiguet, F., Morlon, H., 2014. Faster Speciation and Reduced Extinction in the Tropics Contribute to the Mammalian Latitudinal Diversity Gradient. *PLoS Biol.* 12. <https://doi.org/10.1371/journal.pbio.1001775>
- Ronikier, M., 2011. Biogeography of high-mountain plants in the Carpathians: An emerging phylogeographical perspective. *Taxon* 60, 373–389. <https://doi.org/10.1002/tax.602008>
- Roquet, C., Boucher, F.C., Thuiller, W., Lavergne, S., 2013. Replicated radiations of the alpine genus *Androsace* (Primulaceae) driven by range expansion and convergent key innovations. *J. Biogeogr.* 40, 1874–1886. <https://doi.org/10.1111/jbi.12135>
- Rosauer, D., Laffan, S.W., Crisp, M.D., Donnellan, S.C., Cook, L.G., 2009. Phylogenetic endemism: A new approach for identifying geographical concentrations of evolutionary history. *Mol. Ecol.* 18, 4061–4072. <https://doi.org/10.1111/j.1365-294X.2009.04311.x>
- Schmitt, T., 2007. Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Front. Zool.* 4, 1–13. <https://doi.org/10.1186/1742-9994-4-11>

General introduction

- Schneeweiss, G.M., Pachschröll, C., Tribsch, A., Schönschwetter, P., Barfuss, M.H.J., Esfeld, K., Weiss-Schneeweiss, H., Thiv, M., 2013. Molecular phylogenetic analyses identify Alpine differentiation and dysploid chromosome number changes as major forces for the evolution of the European endemic *Phyteuma* (Campanulaceae). *Mol. Phylogenet. Evol.* 69, 634–652. <https://doi.org/10.1016/j.ympev.2013.07.015>
- Schönschwetter, P., Lachmayer, M., Lettner, C., Prehler, D., Rechner, S., Reich, D.S., Sonnleitner, M., Wagner, I., Hülber, K., Schneeweiss, G.M., Trávníček, P., Suda, J., 2007. Sympatric diploid and hexaploid cytotypes of *Senecio carniolicus* (Asteraceae) in the Eastern Alps are separated along an altitudinal gradient. *J. Plant Res.* 120, 721–725. <https://doi.org/10.1007/s10265-007-0108-x>
- Schönschwetter, P., Schneeweiss, G.M., 2019. Is the incidence of survival in interior Pleistocene refugia (nunataks) underestimated? Phylogeography of the high mountain plant *Androsace alpina* (Primulaceae) in the European Alps revisited. *Ecol. Evol.* 9, 4078–4086. <https://doi.org/10.1002/ece3.5037>
- Schönschwetter, P., Stehlik, I., Holderegger, R., Tribsch, A., 2005. Molecular evidence for glacial refugia of mountain plants in the European Alps. *Mol. Ecol.* 14, 3547–3555. <https://doi.org/10.1111/j.1365-294X.2005.02683.x>
- Schönschwetter, P., Tribsch, A., Schneeweiss, G.M., Niklfeld, H., 2003. Disjunctions in relict alpine plants: Phylogeography of *Androsace brevis* and *A. wulfeniana* (Primulaceae). *Bot. J. Linn. Soc.* 141, 437–446. <https://doi.org/10.1046/j.0024-4074.2002.00134.x>
- Seguinot, J., Ivy-Ochs, S., Juvet, G., Huss, M., Funk, M., Preusser, F., 2018. Modelling last glacial cycle ice dynamics in the Alps. *Cryosphere* 12, 3265–3285. <https://doi.org/10.5194/tc-12-3265-2018>
- Stehlik, I., 2003. Resistance or emigration? Response of alpine plants to the ice ages. *Taxon* 52, 499–510. <https://doi.org/10.2307/3647448>
- Steinbauer, M.J., Field, R., Grytnes, J.A., Trigas, P., Ah-Peng, C., Attorre, F., Birks, H.J.B., Borges, P.A.V., Cardoso, P., Chou, C.H., De Sanctis, M., de Sequeira, M.M., Duarte, M.C., Elias, R.B., Fernández-Palacios, J.M., Gabriel, R., Gereau, R.E., Gillespie, R.G., Greimler, J., Harter, D.E.V., Huang, T.J., Irl, S.D.H., Jeanmonod, D., Jentsch, A., Jump, A.S., Kueffer, C., Nogué, S., Otto, R., Price, J., Romeiras, M.M., Strasberg, D., Stuessy, T., Svenning, J.C., Vetaas, O.R., Beierkuhnlein, C., 2016. Topography-driven isolation, speciation and a global increase of endemism with elevation. *Glob. Ecol. Biogeogr.* 25, 1097–1107. <https://doi.org/10.1111/geb.12469>
- Stenseth, N.C., 1984. The tropics: cradle or museum? *Oikos* 43, 417–420.
- Stewart, J.R., Lister, A.M., Barnes, I., Dalén, L., 2010. Refugia revisited: individualistic responses of species in space and time. *Proc. Biol. Sci.* 277, 661–71. <https://doi.org/10.1098/rspb.2009.1272>
- Stroud, J.T., Losos, J.B., 2016. Ecological Opportunity and Adaptive Radiation. *Annu. Rev. Ecol. Evol. Syst.* 47, 507–532. <https://doi.org/10.1146/annurev-ecolsys-121415-032254>
- Taberlet, P., Zimmermann, N.E., Englisch, T., Tribsch, A., Holderegger, R., Alvarez, N., Niklfeld, H., Coldea, G., Mirek, Z., Moilanen, A., Ahlmer, W., Marsan, P.A., Bona, E., Bovio, M., Choler, P., Cieřlak, E., Colli, L., Cristea, V., Dalmas, J.P., Frajman, B., Garraud, L., Gaudeul, M., Gielly, L., Gutermann, W., Jogan, N., Kagalo, A.A., Korbecka, G., Křpfer, P., Lequette, B., Letz, D.R., Manel, S., Mansion, G., Marhold, K., Martini, F., Negrini, R., Niřo, F., Paun, O., Pellicchia, M., Perico, G., Piekoř-Mirkowa, H., Prosser, F., Puřcař, M., Ronikier, M., Scheuerer, M., Schneeweiss, G.M., Schřnschwetter, P., Schrott-Ehrendorfer, L., Schřpfer, F., Selvaggi, A., Steinmann, K., Thiel-Egenter, C., van Loo, M., Winkler, M., Wohlgemuth, T., Wraber, T., Gugerli, F., 2012. Genetic diversity in widespread species is not congruent with species richness in alpine plant communities. *Ecol. Lett.* 15, 1439–1448. <https://doi.org/10.1111/ele.12004>
- Theodoridis, S., Randin, C., Broennimann, O., Patsiou, T., Conti, E., 2013. Divergent and narrower climatic niches characterize polyploid species of European primroses in *Primula* sect. *Aleuritia*. *J. Biogeogr.* 40, 1278–1289. <https://doi.org/10.1111/jbi.12085>
- Theodoridis, S., Randin, C., Szovenyi, P., Boucher, F.C., Patsiou, T.S., Conti, A.E., 2017. How do cold-adapted plants respond to climatic cycles? Interglacial expansion explains current distribution and genomic diversity in *Primula farinosa* L. *Syst. Biol.* 66, 715–736. <https://doi.org/10.1093/sysbio/syw114>
- Thiel-Egenter, C., Alvarez, N., Holderegger, R., Tribsch, A., Englisch, T., Wohlgemuth, T., Colli, L., Gaudeul, M., Gielly, L., Jogan, N., Linder, H.P., Negrini, R., Niklfeld, H., Pellicchia, M., Rioux, D., Schönschwetter, P., Taberlet, P., van Loo, M., Winkler, M., Gugerli, F., 2011. Break zones in the distributions of alleles and species in alpine plants. *J. Biogeogr.* 38, 772–782. <https://doi.org/10.1111/j.1365-2699.2010.02441.x>

General introduction

- Thuiller, W., Guéguen, M., Georges, D., Bonet, R., Chalmandrier, L., Garraud, L., Renaud, J., Roquet, C., Van Es, J., Zimmermann, N.E., Lavergne, S., 2014. Are different facets of plant diversity well protected against climate and land cover changes? A test study in the French Alps. *Ecography (Cop.)*. 37, 1254–1266. <https://doi.org/10.1111/ecog.00670>
- Thuiller, W., Lavorel, S., Araújo, M.B., Sykes, M.T., Prentice, I.C., 2005. Climate change threats to plant diversity in Europe. *Proc. Natl. Acad. Sci.* 102, 8245–50. <https://doi.org/10.1073/pnas.0409902102>
- Tribsch, A., Schönswetter, P., 2003. refugia for mountain plants: patterns of endemism and comparative phylogeography confirm palaeo-environmental evidence in the Eastern European Alps. *Taxon* 52, 477–497. <https://doi.org/10.2307/3647447>
- Tucker, C.M., Cadotte, M.W., Carvalho, S.B., Davies, T.J., Ferrier, S., Fritz, S.A., Grenyer, R., Helmus, M.R., Jin, L.S., Mooers, A.O., Pavoine, S., Purschke, O., Redding, D.W., Rosauer, D.F., Winter, M., Mazel, F., 2016. A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biol. Rev.* <https://doi.org/10.1111/brv.12252>
- Valente, L.M., Phillimore, A.B., Etienne, R.S., 2015. Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. *Ecol. Lett.* 18, 844–852. <https://doi.org/10.1111/ele.12461>
- Vargas, P., 2003. Molecular evidence for multiple diversification patterns of alpine plants in Mediterranean Europe. *Taxon* 52, 463–476.
- von Humboldt, A., 1805. *Essai sur la Géographie des Plantes*. Levrault, Schoell et compagnie., Paris.
- Wang, Q., Fan, X., Wang, M., 2014. Recent warming amplification over high elevation regions across the globe. *Clim. Dyn.* 43, 87–101. <https://doi.org/10.1007/s00382-013-1889-3>
- Westergaard, K.B., Zemp, N., Bruederle, L.P., Stenøien, H.K., Widmer, A., Fior, S., 2019. Population genomic evidence for plant glacial survival in Scandinavia. *Mol. Ecol.* 28, 818–832. <https://doi.org/10.1111/mec.14994>
- Willis, K.J., Bennett, K.D., Birks, H.J.B., 2009. Variability in thermal and UV-B energy fluxes through time and their influence on plant diversity and speciation. *J. Biogeogr.* 36, 1630–1644. <https://doi.org/10.1111/j.1365-2699.2009.02102.x>
- Xing, Y., Ree, R.H., 2017. Uplift-driven diversification in the Hengduan Mountains, a temperate biodiversity hotspot. *Proc. Natl. Acad. Sci.* 114, E3444–E3451. <https://doi.org/10.1073/pnas.1616063114>
- Zhang, L.B., Comes, H.P., Kadereit, J.W., 2001. Phylogeny and quaternary history of the European montane/alpine endemic *Soldanella* (Primulaceae) based on ITS and AFLP variation. *Am. J. Bot.* 88, 2331–2345. <https://doi.org/10.2307/3558393>

CHAPTER 1

Disentangling drivers of plant endemism and diversification in the European Alps - a phylogenetic and spatially explicit approach

Jan Smyčka^{1,*}, Cristina Roquet¹, Julien Renaud¹, Wilfried Thuiller¹, Niklaus E. Zimmermann², Sébastien Lavergne¹

published in *Perspectives of Plant Ecology, Evolution and Systematics* 28 (2017) 19–27 and 30 (2018) 31-40

¹ Université Grenoble Alpes, CNRS, Laboratoire d'Ecologie Alpine (LECA), 2233 rue de la Piscine, FR-38000 Grenoble, France

² Swiss Federal Research Institute WSL, Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland

*corresponding author, smyckaj@gmail.com

Abstract

Plant endemism in the European Alps is clustered into particular geographic areas. Two contrasted and non exclusive hypotheses have been suggested to explain these hotspots of endemism: (i) those areas were glacial refugia, where endemism reflects survival-recolonisation dynamics since the onset of Pleistocene glaciations, (ii) those are high elevation mountain areas, where endemism was fostered by local speciation events due to geographic isolation and harsh environmental niches, or by low dispersal ability of inhabiting species.

Here, we quantitatively compared these two hypotheses using data of species distribution in the European Alps (IntraBioDiv database), species phylogenetic relationships, and species ecological and biological characteristics. We developed a spatially and phylogenetically explicit modelling framework to analyze spatial patterns of endemism and the phylogenetic structure of species assemblages. Moreover, we analyzed interrelations between species trait syndromes and endemism.

We found that high endemism occurs in potential glacial refugia, but only those on calcareous bedrock, and also in areas with high elevation. Plant assemblages in calcareous refugia showed phylogenetic overdispersion – a signature of non-selective conservation forces, whereas those located in high mountain areas showed phylogenetic clustering – a signature of recent diversification and environmental filtering. Endemic species were either stress-tolerant, poorly dispersing species, or high elevation specialists with a wide distribution within the European Alps.

While both calcareous refugia and high-elevation hotspots harbour a large portion of plant endemism in the European Alps, the species they host have substantially different characteristics. Our results suggest that hotspots of endemism in calcareous refugia are more important for nature conservation planning, as they host many range restricted endemic species and rather isolated evolutionary lineages.

Introduction

Mountain ranges across the world are considered typical examples of endemic-rich regions (Hughes and Atchinson, 2015), but the evolutionary mechanisms and historical factors generating this high endemism are not fully understood. It was early recognized that mountains have a much richer endemic flora than the surrounding lowlands, and also that within mountain systems, there exist specific areas with exceptionally high endemism (Candolle, 1875; Pawłowski, 1970). Such hotspots of endemism in alpine regions were observed mostly in putative glacial refugia (Pawłowski, 1970; Tribsch and Schönswetter, 2003; Feng et al., 2016) or in high-elevation areas (Aeschmann et al., 2011; Nagy and Grabherr 2009; Tribsch and Schönswetter, 2003). This suggests that their occurrence is coupled with specific evolutionary dynamics: such regions may exhibit lower extinction rates due to climatic stability and reduced glacial extent, or higher speciation rates and poorer dispersal ability of high-elevation species.

Hotspots of plant endemism in the European Alps (Alps hereafter) have traditionally been explained by the presence of refugia on the periphery of glacial cover during the ice age periods. These refugia are assumed to have promoted long term population persistence of many species during glacial periods, which left imprints in the population structure of survivor species (Alvarez et al., 2009; Schönswetter et al., 2005; Stehlik, 2003). The distribution of glacial refugia has also shaped contemporary species distributions and endemism patterns, since some survivors could not recolonize all adjacent regions after the retreat of glaciers (Dullinger et al., 2012). It is often assumed that refugia with different bedrocks hosted different pool of species, as the majority of plant endemics of the Alps show either a clear affinity or a strong intolerance to calcareous bedrock, with endemic flora of calcareous bedrock being generally richer (Schönswetter et al., 2005; Tribsch and Schönswetter, 2003). It is likely that species survival during glacial periods has generated particular patterns of endemism, and also left a particular phylogenetic signature in local floras. We can thus hypothesize that species surviving glacial cycles in refugia were drawn from a pool of “pre-glacial” species, whereas species occurring outside refugia were filtered for ability of fast recolonisation, resulting in relative phylogenetic overdispersion of species assemblages in glacial refugia. Here, we aim at quantitatively testing across the whole Alps whether the areas predicted as inhabitable during glacial periods *sensu* Schönswetter et al.

Chapter 1

(2005) match the above described patterns of endemism and phylogenetic structure, and whether patterns of endemism in these potential refugia are influenced by other factors, as is refugium bedrock, topography, or geographic region.

In addition to the influence of Pleistocene historical processes, patterns of plant diversity and endemism in the Alps may also be linked to the characteristics of high-elevation ecosystems, such as geographic insularity and availability of free but hostile niches. Indeed, it was observed that endemic diversity in the Alps grows with elevation (Aeschimann et al., 2011; Tribsch and Schönswetter, 2003). This pattern could be explained by two processes that are not mutually exclusive. First, increased speciation rate in high elevation ecosystems (documented on a global scale and reviewed in Hughes and Atchison, 2015) could induce higher endemism in certain plant clades. This may be due to heterogeneity or niches in high mountain environments and specific life histories of mountain species favoring sympatric speciation (Dixon et al., 2009; Roquet et al., 2013), or due to topographic obstacles in high mountain environments stimulating allopatric speciation (Boucher et al., 2016; Comes and Kadereit, 2003). Repeated speciation events in high-mountain floras could then induce a phylogenetic signature of radiating lineages, producing a phylogenetic clustering in local species assemblages. Second, increased plant endemism in high elevations may have resulted from a reduced dispersal potential itself. Adaptation to high-alpine environments may imply stress tolerance, long lifespan, and preference for vegetative spread (Körner, 2003) to the detriment of dispersal capabilities (i.e. insularity syndrome). High endemism in mountains resulting from increased speciation rates or decreased dispersal capacities of high-elevation species can thus be expected to result in specific signatures of phylogenetic clustering in local species assemblages or in the presence of particular trait syndromes that have improved survival in high mountain environments at the expense of dispersal potential. This has never been tested to date.

We thus argue that patterns of endemism in the Alps could be explained both by local survival-recolonization dynamics following glaciations, and by dynamics of speciation and dispersal in high-elevation ecosystems. In this paper: (i) We quantitatively compare the relative importance of potential glacial refugia on different bedrocks on one side, and elevation on the other side, for patterns of plant endemism in the Alps. To do this, we use two measures of endemism, namely the proportion of endemic species and range size of endemic species, using a grid-based species occurrence data in the Alps (IntraBioDiv; Gugerli et al.,

Chapter 1

2008). (ii) We test whether the phylogenetic structure (richness-standardized phylogenetic diversity; Faith, 1992) of species assemblages differs between those different types of hotspots, according to hypothesized evolutionary processes. Importantly, we developed here a novel method based on Bayesian imperfect detection framework in order to overcome difficulties when calculating community phylogenetic indices from non-completely resolved phylogenies (Molina-Venegas and Roquet, 2014; Rangel et al., 2015). (iii) To shed more light on processes forming endemism in high elevations, we explore how endemism and range size of endemics are related to species elevational optimum (Landolt et al., 2010), ecological and functional characteristics related to high elevation adaptations and evolutionary distinctiveness of species (Isaac et al., 2007).

Methods

Study region

We focus on the European Alps, which corresponds to the great mountain range system stretching from south-eastern France to Slovenia. The tree line in the European Alps lies at cca. 2000 m a. s. l and upper limits of vascular plant occurrence lie in 3500–4500 m a.s.l., differing by region (Ozenda and Borel, 2003; Tribsch and Schönswetter, 2003). The European Alps belong to one of the coldest biomes on the planet at its highest elevations (Körner, 2011); nevertheless, this region appears to be relatively species rich, with a fairly high rate of plant endemism (about 13%; Aeschmann et al., 2004; Pawłowski, 1970). The Alps thus constitute a well known hotspot of biodiversity in Europe (Väre et al., 2003).

Species distribution and environmental data

Species distribution data originate from the mapping of the flora across the Alps produced by the IntraBioDiv consortium (IBD; Gugerli et al., 2008). This dataset contains census and expert based presences and absences of all plant species occurring above the tree line on a regular grid with cells of 20' longitude and 12' latitude (ca. 25 × 22 km). This grid was used for all subsequent spatial analyses. The restriction of the species pool to species occurring clearly above the treeline may be considered problematic for example for interpreting species richness across the dataset (but see Taberlet et al., 2012). Nevertheless, while investigating evolutionary processes, such restriction removes potential noise generated by lowland species that likely have reduced evolutionary histories related to mountains. We also excluded gymnosperms and ferns, because some of our working hypotheses may not be extended to them, trait definition for angiosperms are not easily applicable to those groups and sampling efforts for ferns was low compared to angiosperms.

To quantify which grid cells might serve as glacial refugia during glacial cycles, we overlaid the IBD grid with distribution of potential siliceous and calcareous refugia based on combination paleoclimatic model with geological data (adapted from Schönswetter et al., 2005) and estimated whether each grid cell contained calcareous, siliceous, none, or both types of potential refugia. The potential refugia, adapted from Schönswetter et al. (2005), are estimated for maximum of the last glacial period (110,000–12,000 years before present), but

Chapter 1

they can also be considered a proxy information for distribution of refugia in previous glacial periods. These potential refugia are mostly peripheral (along southwestern, southern, eastern and northern borders of the Alps), likely favouring isolation of plant populations for thousands of years (Schönswetter et al., 2005). The calcareous refugia are the ones lying on limestone or dolomitic bedrock, whereas siliceous refugia lie on variety of acidic bedrocks like granite or gneiss. Larger coherent areas of bedrock types not falling into these two categories are relatively rare within the Alps (Schönswetter et al., 2005; Tribsch and Schönswetter 2003). To separate the effect of refugia from the effect of bedrock itself, we also estimated whether each grid cell contained calcareous (limestone or dolomite) or siliceous bedrock (granite, diorite or gneiss) based on dominant parent material map (PARMADO) from European Soil database (resolution 1×1 km). To quantify the topography of grid cells, we calculated their mean elevation and their difference between highest and lowest elevation (elevation range, hereafter) based on Global digital elevation model by US Geological Survey (resolution $30'' \times 30''$, cca 1×1 km). All calculations were performed using the statistical environment R (R Core Team, 2016) and the R libraries raster (Hijmans, 2016), rgdal (Bivand et al., 2016) and spatialEco (Evans, 2016). All data used for our analyses are accessible in Dryad repository under accession number (to be completed after attribution of doi).

Phylogenetic data

A genus-level phylogeny was built for the Alpine flora using the workflow proposed by Roquet et al. (2013). We downloaded from Genbank three conserved chloroplastic regions (rbcL, matK and ndhF) plus eight regions for a subset of families or orders (atpB, ITS, psbA-trnH, rpl16, rps4, rps4-trnS, rps16, trnL-F), which were aligned separately by taxonomic clustering. All sequence clusters were aligned with three programs (MUSCLE, Edgar, 2004; MAFFT, Katoh et al., 2005; Kalign, Lassman and Sonnhammer, 2005a), then the best alignment for each region was selected using MUMSA (Lassmann and Sonnhammer, 2005b) and depurated with TrimAl (Capella-Gutiérrez et al., 2009) after visual checks. DNA matrices were concatenated to obtain a supermatrix. Maximum-likelihood phylogenetic inference analyses were conducted with RAxML (Stamatakis et al., 2008) applying a supertree constraint at the family-level based on Davies et al. (2004) and Moore et al. (2010). 100

Chapter 1

independent tree searches were performed. The 100 ML trees obtained were dated by penalized-likelihood using r8s (Sanderson, 2003) and 25 fossils for calibration extracted from Smith et al. (2010) and Bell et al. (2010).

To deal with unknown within-genera structures, we simulated 10 scenarios of within-genera random branchings for each of 100 genera-level trees using a Yule process as implemented in the R library apTreeshape (Bortolussi et al., 2012). This resulted in 1000 trees that represent a distribution of possible hypotheses about evolutionary relations in our dataset *sensu* Rangel et al. (2015).

Species ecological and biological features

For each species, we extracted from Flora Indicativa (Landolt et al., 2010) the following ecological indicator values and biological traits:

- inverted values of Landolt's T (expert based ordinal scale classification of species elevation, ranging from 1 for lowland to 5 for alpine species) as species level information about its elevational optimum domain,
- CSR strategy *sensu* Grime (1977) depicting stress tolerance (S), ruderal strategy (R) and competitive capacity (C) of each species (S and R were coded as independent ordinal variables with values between 0 a 3 referring to amount of "S" and "R" in three letter characteristic of a species, C is a linear combination of S and R and thus was not used separately; e.g. for CSS species, stress tolerance=2 and ruderal strategy=0),
- species dispersal capacity (a value of 1 was attributed to anemo- or zoochoric species, and 0 to those with a different dispersal strategy),
- vegetative reproduction (a value of 1 was attributed to species with the ability of any vegetative reproduction, and 0 to species with completely non-vegetative reproduction),
- cushion life form *sensu* Aubert (2014; a value 1 was attributed to species with vegetative reproduction forming tussocks or cushions, 0 for all other species)
- sexuality of the species (a value of 1 was given to species only capable of sexual reproduction, and 0 for species with facultative or obligate asexual seed generation mechanism, as is apomixis or cleistogamy)

Chapter 1

- Raunkiaer plant life-forms (a set of binary variables coding for species being a therophyte, geophyte, hemicryptophyte, chamaephyte or phanerophyte)

Species evolutionary distinctiveness (Isaac et al., 2007) was estimated as a so-called fair proportion measurement (as implemented in R package *picante*; Kembel et al., 2010) averaged across all 1000 phylogenetic trees. The evolutionary distinctiveness describes whether species are positioned in strongly or weakly branching parts of a phylogenetic tree (sometimes referred as “bushy” or “stemmy” subtrees), and the inverse value of evolutionary distinctiveness may be considered a species-level measure of diversification rate (Jetz et al., 2012).

Endemism distribution and its relationships with potential glacial refugia and topography

Species were classified as endemic or non-endemic to the Alps based on the Flora Alpina (variable referred to as endemic status hereafter; Aeschmann et al., 2011). To explore the relationship between spatial patterns of endemism, potential glacial refugia, bedrock and topography, we modeled the ratio of the number of endemics to the total number of plant species within each grid (proportion of endemics, hereafter) by a binomial process with the following predictors: mean elevation, elevation range, the presence of calcareous or siliceous bedrock, and presence of potential calcareous or siliceous refugia in grid cell. In order to locate region-specific patterns, spatial smoothing was added as an additional term to the model. To control for over-dispersion, a random effect from a Gaussian distribution was added. This was done by means of a generalized additive model (GAM), fitted with Bayesian inference. The MCMC sampling was performed using a JAGS sampler (Plummer, 2003; Plummer, 2016), using modified BUGS code generated by the *jagam* procedure in the R library *mgcv* (Wood, 2011). The model code is provided in appendix D. We ran the model on 5 chains for 70000 iterations, with a 20000 generation burn-in period and we thinned the resulting chains by 50. The convergence of the model was checked visually and by means of the Gelman-Rubin statistic (Gelman and Rubin, 1992).

The endemic rarity was defined as $\log(1/\text{species number of occurrences in IBD data})$ for all species endemic to the Alps. For each grid cell, we calculated the mean rarities of all occurring endemic species (variable further referred to as mean endemic rarity). To investigate

Chapter 1

links between mean endemic rarity and the presence of potential refugia and topography, we fitted a GAM similar to the one for proportion of endemics. To deal with the fact that mean endemic rarity is poorly estimated in grid cells with fewer endemics, we added a layer of hierarchy and modeled mean endemic rarity of each grid cell as a latent variable representing the mean of a Gaussian distribution from which rarities of different species in grid cell are drawn. The distribution of mean endemic rarity was modeled as a function of mean elevation, elevation range, presence of calcareous and siliceous bedrock, potential calcareous and siliceous refugia and spatial spline as predictors, using the Gaussian error structure. The model code is provided in appendix D. The JAGS sampler setup and convergence checking procedure was identical to the one used for modeling species endemism.

The GAM framework is useful for discovering region-specific patterns, does not rely on any pre-defined relationships, and provides easily interpretable visualizations. However, this framework may not necessarily be the most suitable way to control for spatial autocorrelation (Dormann et al., 2007). Because of that, we also tested other methods to account for spatial autocorrelation, namely spatial generalized least square models and conditional autoregressive models. These alternative methods provided qualitatively similar results to those obtained with the spatial GAM (see appendix A), and thus only GAM results are presented and discussed in the main text.

Phylogenetic diversity and its relation to potential glacial refugia and topography

We quantified the phylogenetic structure in each grid cell by calculating phylogenetic diversity (Faith, 1992) standardized for species richness effects (ses.PD), as implemented in R package PhyloMeasures (Tsirogiannis and Sandel, 2017). This standardization was used to remove implicit interdependence between species richness and phylogenetic diversity, and resulting values are further interpreted in this way.

The link between ses.PD and the presence of potential refugia and topography was explored with a bayesian GAM similar to the one for proportion of endemism or mean endemic rarity. In order to control for phylogenetic uncertainty, we performed ses.PD calculations for each of 1000 phylogenetic trees within each grid cell. The “real” value of ses.PD was modeled as a latent variable representing mean parameter of Gaussian distribution from which ses.PDs of different trees are drawn. The distribution of the “real” ses.PD per grid cell was modeled as a

function of mean elevation, elevation range, presence of calcareous or siliceous bedrock, potential calcareous or siliceous refugia and a spatial spline as predictors, using a Gaussian error structure. For the model code, see appendix D. The sampler setup and convergence checking was identical as for the previous Bayesian models.

Endemism-elevation relationships and its link to species traits and ecological properties

In order to explore relationships between species endemism and their elevational niche, we modeled species-level endemic status (endemism to the Alps coded as a binary variable) with a non-spatial generalized linear model (GLM) with a binomial error term, and used each species' elevational optimum (Landolt's T) as a predictor. To evaluate how the other species characteristics affected this endemism-elevation relationship, we fitted the same model with additional predictors: evolutionary distinctiveness, ecological strategies and species traits (see above "Species ecological and biological features"). All predictors were standardized prior to analyses. The important predictors were selected using forward-backward model selection based on Akaike's information criterion (AIC), as implemented in stepAIC procedure in R, starting from a full model containing all variables. In order to tease apart the relative effects of different species features, partial correlations for non-parametric Kendall's correlation coefficient were computed using the package ppcor in R (Kim, 2015). To explore relationships among endemics rarity, elevational optimum and other species characteristics, we fitted models similar to the ones used for endemic status. In this case we used Gaussian error structure in the GLM and parametric Pearson correlation coefficients for estimating partial correlations.

To ascertain that analyses were not biased by a phylogenetic structure in the dataset, we re-ran the final models obtained above by the stepAIC procedure within a phylogenetic regression framework, as implemented in the package phylolm (Ho and Ane, 2014). We obtained an Ornstein-Uhlenbeck $\alpha=0.193$ in the binomial model of endemic status and Pagel's $\lambda=0.037$ in the Gaussian model of endemics rarity. These values showed that the phylogenetic signal was weak in both cases, making non-phylogenetic analyses appropriate with our data set (for results of phylogenetic comparative models see Appendix C).

Results

Endemism distribution and its relationships with potential glacial refugia and topography

The proportion of endemics tended to be significantly (in the sense that Bayesian 95% credibility interval did not overlap 0) larger in grid cells that contained calcareous bedrock (effect size=0.245 ; (see Table 1 for 95% credibility intervals, and Appendix A for the results of alternative spatial models). This effect was further augmented by the presence of potential calcareous refugia (ES=0.195). The proportion of endemics was not affected by the presence of siliceous bedrock and was even significantly lower in cells containing potential siliceous refugia (ES=-0.070). The proportion of endemics increased significantly with mean elevation (ES=0.256) and elevation range (ES=0.161) in grid cells; for instance, an increment of 1000 m in mean elevation has an effect comparable to presence of potential calcareous refugia (Fig. 1 A). The spatial component of the model showed that the proportion of endemics declined towards the north-western edge of the Alps, and to a lower extent toward the eastern edge of the Alps (Fig. 1 B).

The mean endemic rarity was significantly positively associated with the presence calcareous bedrock (ES=0.107) and of potential calcareous refugia (ES=0.068; see Table 1 for 95% credibility intervals).The effect of siliceous bedrock was negative (ES=-0.077) and the effect of potential siliceous refugia was not distinguishable from 0. Mean endemic rarity was influenced neither by mean elevation nor by the elevation range (Fig. 1 C). Similar as for the proportion of endemics, we found a decline of mean endemic rarity towards the north-western corner of the Alps (Fig. 1 D).

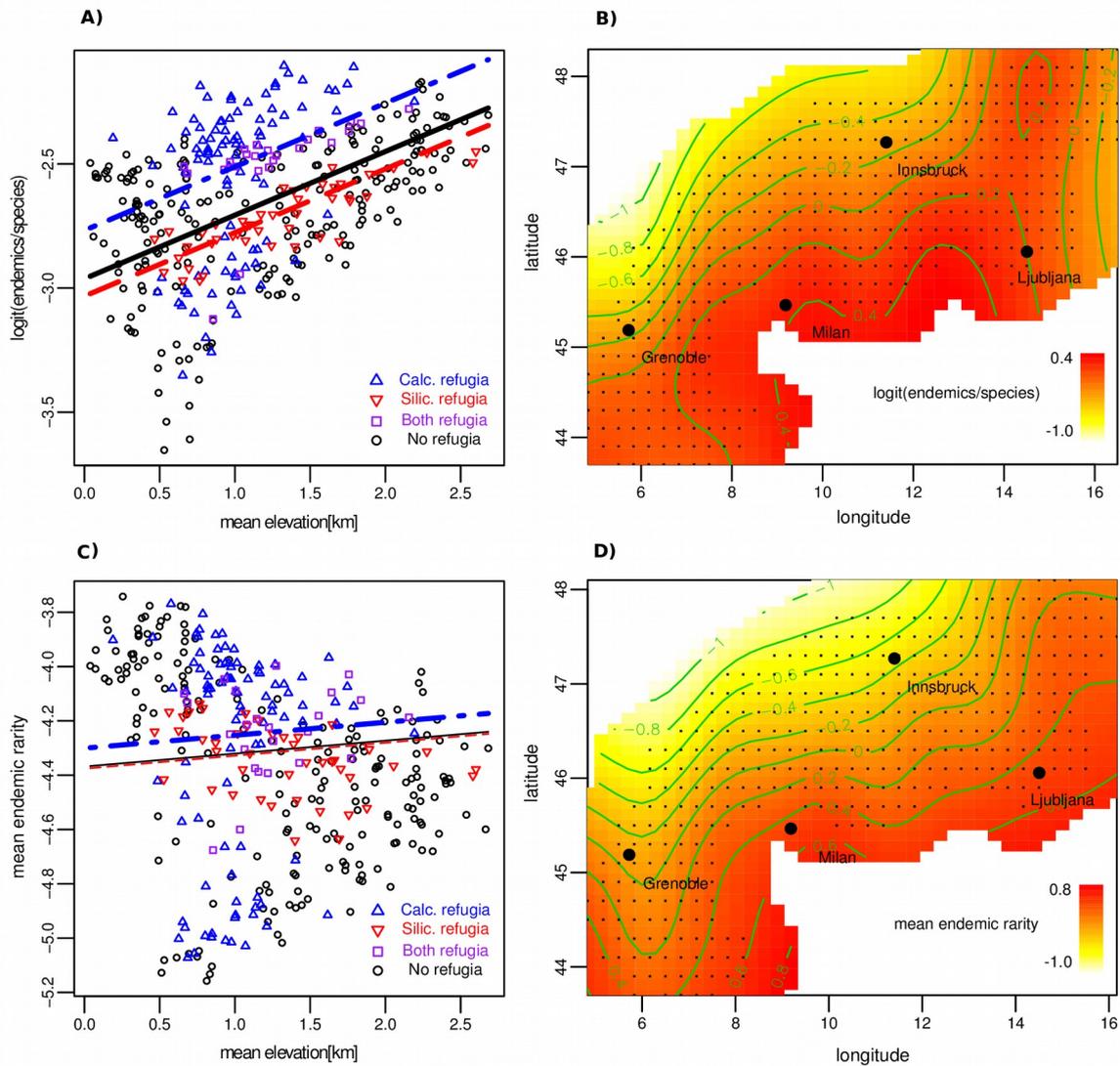


Figure 1: Spatial models depicting the geographic patterns of proportion of endemics (A, B) and the geographic patterns of mean rarity of endemics (C, D). Panels (A) and (C) show relationship of these two indices with mean elevation for grid cells without refugia (black, solid line, plotted in bold if effect of mean elevation is significant), with calcareous refugia (blue, double dashed line, plotted in bold if effect of calcareous refugia is significant) or with siliceous refugia (red, single dashed line, plotted in bold if effect of siliceous refugia is significant). Note that y-values of points are adjusted to account for the effect of not-displayed model variables and for the mean effect of smooth model component per group. Panels (B) and (D) represent the smooth component of each model, showing geographic areas with overall higher or lower proportion of endemics or mean endemic rarity when simultaneously accounting for parametric model components.

Phylogenetic diversity and its relation to potential glacial refugia and topography

Grid cells containing calcareous bedrock presented systematically larger ses.PD (ES=0.487, see Table 1 for 95% credibility intervals) and this effect was further augmented in potential calcareous refugia (ES=0.575). Effect of siliceous bedrock on ses.PD is not distinguishable from 0, but potential siliceous refugia exhibited marginally significant increase of ses.PD (ES=0.255). Large ses.PD in calcareous bedrock, calcareous refugia and possibly siliceous refugia indicates that these areas host species assemblages separated by longer evolutionary branches than the rest of dataset. Sites with a higher mean elevation hosted species assemblages with a relatively lower ses.PD (ES=-0.335); a decrease in 500 m of mean elevation is comparable to the presence of potential calcareous refugia (Fig. 2 a). This indicates that areas of higher mean elevation in the Alps host species assemblages that are more closely related than the rest of the dataset. Visualization of the spatial component of the model displayed particular areas with high (Savoy Prealps, Julian Alps) or low ses.PD (Ötztal and Rhaetian Alps). These spatial effects were, however, weak in comparison to the effects of the linear terms of the model (see Fig. 2 b).

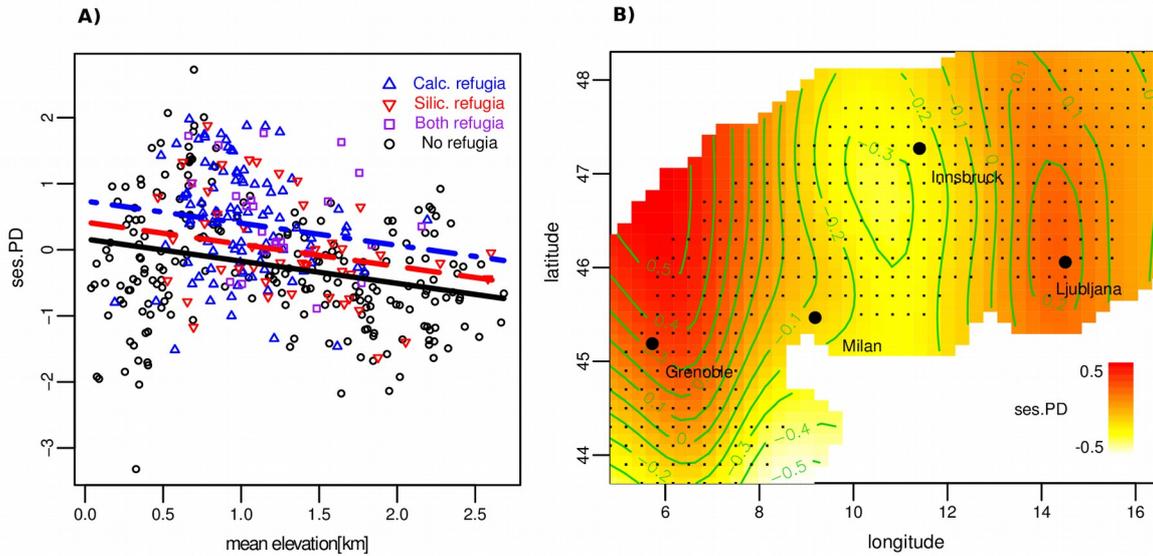


Figure 2: Spatial model depicting geographic patterns of species-richness standardized phylogenetic diversity (ses.PD). Panel (A) shows the relationship of ses.PD with mean elevation (black, solid line, plotted in bold if effect of mean elevation is significant) and the occurrence of calcareous (blue, double dashed line, plotted in bold if effect of calcareous refugia is significant), siliceous refugia (red, single dashed line, plotted in bold if effect of siliceous refugia is significant). Note that y-values of points are adjusted to account for the effect of not-displayed model variables and for the mean effect of smooth model component per group. Panel (B) represents the smooth component of the model, showing geographic areas with overall higher or lower ses.PD when simultaneously accounting for parametric model components.

	proportion of endemics			mean endemic rarity			ses.PD		
	lower	ES	upper	lower	ES	upper	lower	ES	upper
intercept	-3.102	-2.962	-2.822	-4.483	-4.368	-4.253	-0.171	0.162	0.501
calc. bedrock	0.173	0.245	0.318	0.051	0.107	0.164	0.247	0.487	0.727
calc. refugia	0.135	0.195	0.255	0.02	0.068	0.116	0.379	0.575	0.77
silic. bedrock	-0.057	0.025	0.108	-0.142	-0.077	-0.012	-0.471	-0.225	0.021
silic. refugia	-0.14	-0.07	-0.002	-0.064	-0.008	0.048	0	0.255	0.511
mean elevation	0.185	0.256	0.327	-0.008	0.048	0.103	-0.596	-0.335	-0.077
elevation range	0.096	0.161	0.226	-0.067	-0.015	0.036	-0.046	0.201	0.447

Table 1: Mean effect size (ES), lower and upper bounds of the 95% credibility intervals of the effects sizes of spatial models in which the proportion of endemics and the mean endemics rarity, as well as species-richness standardized phylogenetic diversity (ses.PD) are explained. Model estimates with credibility intervals not overlapping with 0 are given in bold.

Endemism-elevation relationship and its link to species traits and ecological properties

The elevational optimum of species (summarized with Landolt's T) was a significant positive predictor of species endemic status ($\tau=0.086$; Fig. 3A, see Table 2 for p-values). When mixed with other species trait predictors, the relationship with elevation became weaker but remained significant ($\tau=0.032$). Concerning other species-level predictors, the endemic status is significantly linked to poor dispersal capacity ($\tau=-0.137$), sexuality ($\tau=0.080$), and stress tolerance ($\tau=0.100$; Fig. 3A, Table 2). Evolutionary distinctiveness was negatively related to endemic status ($\tau=-0.054$), and its inclusion improved model AIC, but this effect did not reach statistical significance ($p=0.134$). The cushion life form was positively related to endemic status ($\tau=0.064$); the inclusion of this trait also improved the model AIC, but it did not reach statistical significance ($p=0.110$) either.

Endemics rarity was significantly negatively linked with elevational optimum ($R=-0.110$) and the strength of this relationship remained similar in a model that included species traits and ecological predictors ($R=-0.102$; Table 2, Fig. 3B). Endemics rarity was further associated with sexuality ($R=0.068$) and geophyte life form ($R=0.144$). Dispersal capacity ($R=-0.084$) had a weak but significant negative effect on rarity ($p=0.029$), but this would turn marginally non-significant if accounting for phylogenetic correction ($p=0.097$; see appendix C). Stress tolerance ($R=0.046$) and hemicryptohyte life form ($R=0.084$) were positively related to endemics rarity and improved model fit, yet these effects did not reach significance ($p=0.088$, $p=0.104$; Fig. 3B, Table 2).

Chapter 1

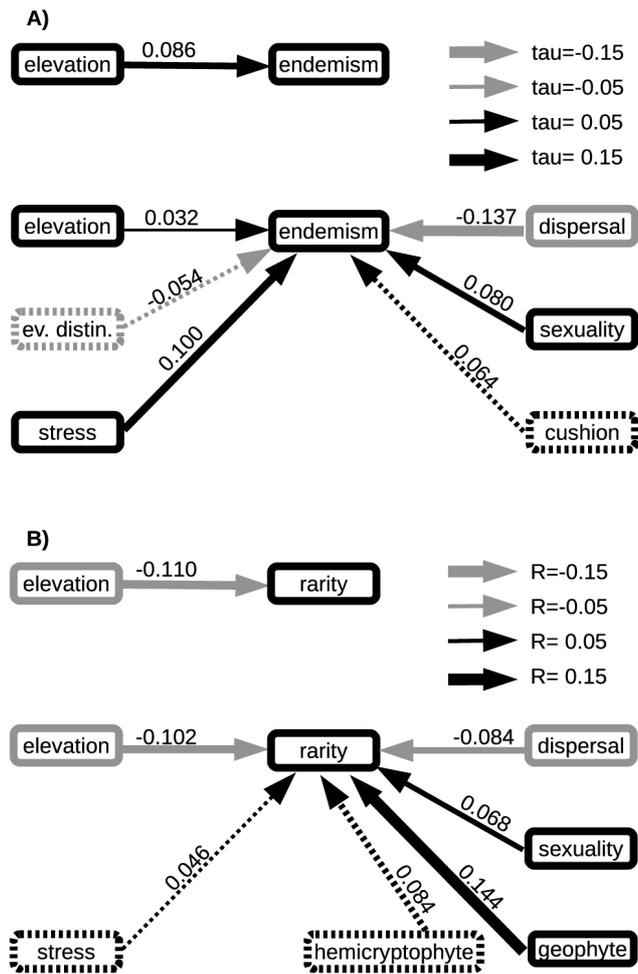


Figure 3: Relationship of elevational optimum (Landolt's T) and species endemic status (A) or the rarity of endemic species (B) in elevation only models and in models with elevation and the other species level predictors included. Tau is the partial Kendall correlation in binomial model of endemic status (A), R is the partial Pearson correlation in Gaussian model (B) of endemics rarity. All listed predictors contribute to model fit as measured by AIC, predictors significant at $p < 0.05$ appear with solid arrows, non-significant with dashed.

Chapter 1

	endemic status			endemics rarity		
	tau	ES	p	R	ES	p
elevational optimum alone	0.086	0.297	< 0.001	-0.11	-0.194	0.012
elevational optimum	0.032	0.179	0.043	-0.102	-0.206	0.009
evolutionary distinctiveness	-0.054	-0.132	0.134	N	N	N
ruderal strategy	N	N	N	N	N	N
stress strategy	0.1	0.231	0.009	0.046	0.144	0.088
dispersal	-0.137	-0.297	< 0.001	-0.084	-0.183	0.029
sexuality	0.08	0.258	0.023	0.068	0.193	0.01
vegetative reproduction	N	N	N	N	N	N
cushions	0.064	0.122	0.11	N	N	N
chamaephyte	N	N	N	N	N	N
geophyte	N	N	N	0.144	0.245	0.003
hemicryptophyte	N	N	N	0.084	0.136	0.104
phanerophyte	N	N	N	N	N	N
therophyte	N	N	N	N	N	N

Table 2: Effect size (ES), p-values and correlation coefficients (tau and R) of models explaining endemic status and endemics rarity species level characteristics. Model results are shown with species altitudinalelevational optima (Landolt's T) as a single predictor, and in combination with all other species characteristics. N stands for predictors that did not pass the stepwise AIC optimization procedure and thus did not contribute to the model. Bold are significant model terms ($p < 0.05$).

Discussion

In this study, we quantitatively compared the importance of endemism hotspots in potential glacial refugia and in areas with high elevation. We showed that the phylogenetic structure of plant assemblages occurring in these two types of hotspots is substantially different and reflects their contrasting evolutionary histories. Moreover, we explored so far undocumented interrelations between plant endemism, elevation, species ecological strategies, biological traits and evolutionary distinctiveness.

Hotspots of endemism in calcareous refugia and areas with high elevation

Areas of the Alps comprising potential calcareous glacial refugia hold a substantially high proportion of alpine endemics, and these endemics are typically narrowly distributed (with high rarity of endemics) within the Alps. This is partly caused by generally higher endemism in calcareous areas, but potential calcareous refugia host even larger proportion and rarer endemics than calcareous areas in general in our data. It suggests that potential calcareous refugia indeed host species that did not manage to recolonize larger areas after the retreat of glaciers (Dullinger et al., 2012). Increasing endemism with elevation has an effect per 1000 elevation meters comparable to the presence of calcareous refugia. However, high-elevation endemics are typically widespread within the Alps (with low rarity), in contrast to narrowly distributed endemics of potential calcareous refugia. While high endemism in glacial refugia and high elevations was previously reported from the Alps (Aeschmann et al., 2011; Tribsch and Schönswetter, 2003) and other mountain systems across the world (Feng et al., 2016; Mráz et al., 2015; Nagy and Grabherr, 2009), we quantitatively measured and tested those effects within a spatially explicit modeling framework. We show that endemism in potential calcareous glacial refugia and in high elevations are of comparable importance in the Alps, suggesting that similar patterns might be found in other mountain ranges affected by Pleistocene glaciation dynamics. Our quantitative approach does not only compare importance of refugia and elevation gradient for formation of endemic hotspots in the Alps, but also allows us to shed light on exceptions from this general trend.

Interestingly, we found that the overall patterns of high proportion of endemics and the high rarity of endemics in potential calcareous refugia are not paralleled in potential siliceous

Chapter 1

refugia. The mean rarity of endemic species occurring in potential siliceous refugia is comparable to the siliceous areas outside the refugia and the proportion of alpine endemics in potential siliceous refugia is even significantly lower than outside refugia. This suggests that high endemism previously reported in some potential siliceous refugia (Tribtsch and Schönswetter, 2003) could at least in some cases rather be attributed to the other predictors used in our model than to favourable conditions during glacial periods, i.e. those areas might be endemic-rich due to topographic roughness or because they lie in region that is overall richer in endemics (which is modeled by spatial autocorrelation in our model). A possible reason why potential siliceous refugia are not richer in endemics than non-refugial siliceous areas is that siliceous refugia are typically interconnected with non-refugial siliceous areas in the central Alps, in contrast to calcareous refugial and non-refugial areas, which are scattered at the margins of the Alps. Hence, most siliceous specialists species that survived glacial periods may have been able to recolonize broader areas after glacier retreat due to lower spatial isolation of habitats on siliceous bedrock (Alvarez et al., 2009; Dullinger et al., 2012). Following this hypothesis, postglacial migration might have erased patterns of endemism, even though the presence of siliceous refugia may be still visible in the spatial genetic structures of some particular species (Schönswetter et al., 2005; Stehlik, 2003).

Spatial models of proportion of endemics and mean endemic rarity suggest that the region of the north-western Alps hosts a lower proportion of endemics and also fewer rare endemics. A possible explanation is that the impact of glaciers was relatively stronger in this area due to combination of lower elevation and higher glacial extent towards the north-western part of the Alps (see maximum extent of glaciers in Schönswetter et al., 2005), making potential refugia in this area uninhabitable. This explanation is in agreement with studies tracking glacial refugia using intraspecific genetic patterns (Schönswetter et al., 2005), counts of endemic species (Tribtsch and Schönswetter, 2003), or species rarity (Taberlet et al., 2012); all of these studies found weak or no support for the existence of glacial refugia in the north-western Alps.

Calcareous refugia as species museums and high elevations as cradles

Areas with potential calcareous refugia contain systematically higher richness-standardized phylogenetic diversity (ses.PD), which suggests that they host a higher proportion of species

Chapter 1

with long and isolated branches than other sites in the Alps. The ses.PD in potential calcareous refugia is higher also in comparison with calcareous areas outside refugia. Such a phylogenetic signature may result from a random selection of refugial survivors from the pre-glacial species pool combined with selective re-colonizations of previously glaciated sites by well-dispersing species, resulting in assemblages at sites outside refugia being more clustered than in refugia. In addition, resource competition in refugia during glacial periods could have also contributed to the competitive exclusion of related species *sensu* Webb et al. (2002). The ses.PD in glacial refugia might have also been augmented by allopatric speciation among isolated glacial refugia, which would result in sister species seldom sharing one refugial site (Pigot and Etienne, 2015; Warren et al., 2014). With the current resolution of phylogenetic data and without evidence for past migration from and to the Alps it is not possible to estimate the relative importance of those three processes. Still, the observed evolutionary overdispersion in potential calcareous refugia strongly supports the idea that those areas constitute museums of phylogenetic diversity of the Alpine flora *sensu* Stenseth (1984), and served as islands of suitable condition during glacial periods. The long-term conservation of this local flora since glacial periods onwards has thus made them important zones for the protection of plant phylogenetic diversity in the Alps. For potential siliceous refugia, we found a pattern of phylogenetic overdispersion similar to potential calcareous refugia, but only marginally significant. It supports the interpretation that potential siliceous refugia also hosted species during glacial periods. Although the patterns of endemism were erased by postglacial migrations, the imprint of glacial survival may still be detectable in the phylogenetic structure of plant assemblages.

Areas in the Alps with high mean elevation exhibit significantly lower ses.PD than other sites. This means that with increasing elevation, species are more likely co-occur with their close relatives, and that this pattern is consistent throughout the Alps. Two classes of processes may explain this pattern. First, the high alpine flora could be strongly shaped by environmental filtering *sensu* Webb et al. (2002) by which related species tend to share adaptations to similar climatic environments and are, thus, jointly sorted across climatic gradients. Second, this first process may have almost certainly been amplified by recent and repeated speciations in clades that have largely diversified in alpine environments (Boucher et al., 2016; Mansion et al., 2012 ; Molina-Venegas et al., 2015; Roquet et al., 2013). Clearly, environmental filtering alone cannot explain the phylogenetic clustering of high alpine floras,

Chapter 1

since we found additional evidence that species richness increases with mean elevation (see Appendix B) and increased filtering should decrease species richness. Our results are thus in line with the extensive evidence that plant clades have rapidly diversified into high-mountain environments in general (reviewed in Hughes and Atchison, 2015). Given increasing evidence that speciation in alpine plants widely occur by allopatric speciation with little niche shift (Boucher et al., 2016; Comes and Kadereit, 2003), the observed signature of phylogenetic clustering suggests that migration during post-glacial recolonization caused closely related species to co-occur in similar mountain environments, a process known as secondary sympatry (Pigot and Tobias 2013).

Our analyses thus show that high elevation and refugial hotspots of endemism in the Alps are of similar importance, but were formed by different evolutionary processes. It suggests that in the Alps, but possibly also in other temperate mountains affected by Pleistocene glaciation dynamics, endemism should not be considered an indicator of one particular evolutionary process as is glacial survival, altered dispersal dynamics or faster speciation.

High elevation endemism is shaped by dispersal

Above we discussed that areas with high elevation and calcareous refugia are important hotspots of endemism and here we focus more closely on processes forming endemism in high elevations. Accordingly with spatial models, our species-level analysis provides evidence that species endemic status is positively related to elevational optimum and that endemic species with high elevational optimum have larger distribution ranges (lower rarity) than endemic species from lower elevations. However, once other predictors are included, the effect of species elevational optima on endemism becomes much weaker. The major pattern arising from this analysis is that poorly dispersing and/or stress tolerant endemics tend to be over-represented with increasing elevation. In addition, endemics with restricted distribution ranges (high rarity) seem to be rather poor-dispersing and stress tolerant species, although this last effect is weaker and marginally significant. The relationship between endemism, elevation optimum, dispersal and stress tolerance is in line with previous findings (Aeschimann et al, 2012; Hobohlm, 2008; Mráz et al., 2015), and suggests that some of the high-elevation species are endemic, because of their poor colonizing capabilities that are linked to their ecological and trait characteristics. These species are specialists of stressful

Chapter 1

habitats, such as of rocky outcrops or high-elevation habitats, possibly within glacial refugia, e.g. in Dolomites, Julian Alps or south-western Alps. Typical examples of narrowly distributed high-elevation endemics that are stress tolerant and poor dispersers are: *Rhizobotrya alpina*, *Braya alpina*, *Draba ladina*, *Sempervivum dolomiticum*, *Saxifraga florulenta*, *Myosotis gallica*, *Cardaminopsis pedemontana*, *Linaria tonzigii* or *Moehringia concarenae*.

Species life-forms and reproduction strategies also partly explained species endemic status and the rarity of endemic species. Cushion plants are more likely alpine endemics (e.g. *Androsace alpina*), and geophytes are more likely rare endemics (e.g. *Allium kermesinum*) than expected by chance. Our results also indicate that sexually reproducing species are more likely to be endemic and rare than asexual ones. This may appear surprising given that asexual lineages tend to create small but distinct populations, typically considered a species (e.g. *Alchemilla* or *Sorbus*). This result is thus likely an artefact derived from the fact that, in large compilations such as IntraBioDiv (Gugerli et al., 2008) or Flora Alpina (Aeschimann et al., 2004), asexual and taxonomically problematic species are typically merged to relatively broadly aggregated species definitions. For instance, the genus *Alchemilla* has only 5 species distinguished in the IntraBioDiv data. To shed more light on endemism-sexuality relations, it would be necessary to explicitly control for the species definition, e.g. by including data about species population structures, and perhaps by tackling the issue of species delimitation with modern sequencing techniques. However, given the large number of species of the Alps, this is likely a very demanding, if currently not impossible, task.

In contrast to endemics that are stress-adapted, poorly dispersing or having endemism-related life form, there is another large portion of species whose endemic status is best explained by elevational optimum itself. Such endemics are typically widespread within the Alps (have low rarity), and are likely well adapted to environments such as the subnival zone, that are relatively common within the Alps, but rare or non-existing in other European mountain systems. These species are thus expected to have a dispersal capacity sufficient to spread across the Alps. Yet, their dispersal capacity is not sufficient for colonizing and maintaining viable populations in isolated patches of suitable habitats outside the Alps. Typical examples of well-dispersing, widespread high-elevation endemics are for example: *Festuca interdecens*, *Cerastium pedunculatum*, *Crepis terglouensis*, *Adenostyles leucophylla*, *Erigeron neglectus* or *Artemisia genepei*.

Chapter 1

Surprisingly, the evolutionary distinctiveness of species is only weakly linked to endemic status; one would have expected that endemism should be common in species with little evolutionary distinctiveness resulting from repeated recent speciation events. Our analysis on the relationship between endemic status and evolutionary distinctiveness may have been partly weakened by the use of genus-level phylogenies, as recent speciation events would be most apparent at terminal branches, where we are missing accurate information. But if speciation events in general were a main driving force of high mountain endemism patterns, it is unlikely that missing information at the species level would completely remove the signal. The weak relationship between endemic status and evolutionary distinctiveness is interesting, because according to other results from our study, the flora of high elevation areas (including non-endemic species) shows a pattern of phylogenetic clustering – a possible indication of faster and recent diversification. Combining our findings, we hypothesize that the high-elevation flora may indeed have faster speciation rate, as was reported in studies from mountains all around the world (reviewed in Hughes and Atchison, 2015). However, even the youngest speciation events in our phylogeny are older than the last glaciation period (all terminal branch length estimates are longer than 110,000 years) and most of them is even older than the onset of glaciation-deglaciation cycles in the Quaternary (77% of terminal branch length estimates are longer than 2.58 million years). This suggests that although species may speciate faster in high elevations, the relationship between speciation and endemism is further erased by massive glacial-interglacial migrations and extinctions (see Kadereit et al., 2004). The temporal dynamics of the evolution of new species and drivers of endemism need to be further studied with a well-dated species-level phylogeny of the Alps and possibly beyond such as including the Carpathians and the Pyrenees or the Balkanic mountain systems.

Conclusions

Our analyses demonstrate that important hotspots of endemism in the European Alps are situated in potential calcareous glacial refugia, rather than siliceous ones, but also in areas with high elevation: an elevational change of 1000 m contributes to endemism similarly as the presence of potential calcareous refugium. Potential calcareous refugia harbour a large portions of endemics that are rare within the Alps, and plant assemblages with higher

Chapter 1

phylogenetic diversity than random expectations. It suggests that these areas indeed served as safe harbours for distinct evolutionary lineages during the periods of glaciation. The assemblages of high-elevation hotspots carry a signature of phylogenetic clustering, which indicates a combination of strong environmental filtering and faster speciation rate. More detailed analyses of high elevation endemism reveal that a large portion of endemics in the Alps are range restricted species with poor dispersal ability, and another large portion of endemics are high elevation specialists that are widespread within the Alps. Our results show that endemic hotspots both in calcareous refugia and high elevations are of high importance, but they are formed by contrasting evolutionary processes. In line with this, future studies of endemism and diversification in the Alps or other mountain ranges affected by past glaciation should consider that endemism in such areas is formed by interplay of migration and different diversification processes acting on different timescales, rather than one dominant force. From a point of view of nature conservation, calcareous refugia in the Alps deserve more attention than high elevation hotspots, because they host primarily range-restricted endemics and phylogenetically distinct species retaining a substantial evolutionary history overall.

Acknowledgements

We would like to acknowledge Tomáš Herben, Zdeněk Janovský, Tamara Münkemüller, Matthew Talluto and Martin Weiser for helpful comments at different stages of preparation of the manuscript, and David Aeschmann for providing digitalized Flora Alpina and useful advices concerning this data source. The research was funded by the ANR project Origin-Alps (ANR-16-CE93-0004), the European Research Council FP7/2007–2013 Grant Agreement no. 281422 (TEEMBIO), by a grant from Labex OSUG@2020 (Investissements d'avenir - ANR10LABX56) and a grant from the Pôle départemental de recherches sur la Biodiversité (Conseil Général de l'Isère). Jan Smyčka was supported by doctoral funding of French Ministry of Higher Education and Research.

Chapter 1

References

- Aeschimann, D., Lauber, K., Moser, D. M., Theurillat, J. P., 2004. *Flora Alpina*, Haupt Verlag, Bern, Stuttgart, Wien.
- Aeschimann, D., Rasolofo, N., Theurillat, J. P., 2011. Analyse de la flore des Alpes. 2: Diversité et chorologie. *Candollea*. 66, 193–219.
- Aeschimann, D., Rasolofo, N., Theurillat, J. P., 2012. Analyse de la flore des Alpes. 4: Ecologie. *Candollea*. 67, 193–219.
- Alvarez, N., Thiel-Egenter, C., Tribsch, A., Holderegger, R., Manel, S., Schönswetter, P., Taberlet, P., Brodbeck, S., Gaudeul, M., Gielly, L., Küpfer, P., Mansion, G., Negrini, R., Paun, O., Pellicchia, M., Rioux, D., Schüpfer, F., Van Loo, M., Winkler, M., Gugerli, F., 2009. History or ecology? Substrate type as a major driver of patial genetic structure in Alpine plants. *Ecol. Lett.* 12, 632–640. doi:10.1111/j.1461-0248.2009.01312.x
- Aubert, S., Boucher, F., Lavergne, S., Renaud, J., Choler, P., 2014. 1914–2014: A revised worldwide catalogue of cushion plants 100 years after Hauri and Schröter. *Alp. Bot.* 124, 59. doi:10.1007/s00035-014-0127-x
- Bell, C. D., Soltis, D. E., Soltis, P. S., 2010. The age and diversification of the angiosperms re-revisited. *Am. J. Bot.* 97, 1296–1303.
- Bivand, R., Keitt, T., Rowlingson, B., 2016. *rgdal: Bindings for the Geospatial Data Abstraction Library*. R package version 1.2-5.
- Bortolussi, N., Durand, E., Blum, M., Francois, O., 2012. *apTreeshape: Analyses of Phylogenetic Treeshape*. R package version 1.4-5.
- Boucher, F. C., Zimmermann, N. E., Conti, E., 2016. Allopatric speciation with little niche divergence is common among alpine Primulaceae. *J. Biogeogr.* 43, 591–602. doi:10.1111/jbi.12652
- Candolle, A. P. de., 1875. Sur les causes de l'inégale distribution des plantes rares dans la chaîne des Alpes. *Atti del Congresso Internazionale Botanico Tenuto in Firenze 1874*, 92–104.
- Capella-Gutiérrez, S., Silla-Martínez, J. M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*. 25, 1972–1973.
- Comes, H. P., Kadereit, J. W., 2003. Spatial and Temporal Patterns in the Evolution of the Flora of the European Alpine Systems. *Taxon*. 52, 451–562.
- Davies, T. J., Barraclough, T. G., Chase, M. W., Soltis, P. S., Soltis, D. E., Savolainen, V., 2004. Darwin's abominable mystery: Insights from a supertree of the angiosperms. *Proc. Natl. Acad. Sci. U.S.A.* 101, 1904–1909.
- Dixon, C. J., Schönswetter, P., Suda, J., Wiedermann, M. M., Schneeweiss, G. M., 2009. Reciprocal Pleistocene origin and postglacial range formation of an allopolyploid and its sympatric ancestors (*Androsace* affinis group, Primulaceae). *Mol. Phylogenet. Evol.* 50, 74–83. doi:10.1016/j.ympev.2008.10.009
- Dormann, C. F., McPherson, J. M., Araújo, M. B., Bivand, R., Bolliger, J., Carl, G., Davies, R. G., Hirzel, A., Jetz, W., Kissling, D. W., Kühn, I., Ohlemüller, R., Peres-Neto, P. R., Reineking, B., Schröder, B., Schurr, F. M., Wilson, R., 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography*. 30, 609–628. doi:10.1111/j.2007.0906-7590.05171.x
- Dullinger, S., Willner, W., Plutzer, C., Englisch, T., Schrott-Ehrendorfer, L., Moser, D., Ertl, S., Essl, F., Niklfeld, H., 2012. Post-glacial migration lag restricts range filling of plants in the European Alps. *Glob. Ecol. Biogeogr.* 21, 829–840. doi:10.1111/j.1466-8238.2011.00732.x
- Edgar, R. C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, 32, 1792–1797.
- Evans, J. S., 2016. *spatialEco*. R package version 0.0.1-4.
- Faith, D. P., 1992. Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* 61, 1–10.
- Feng, G., Mao, L., Sandel, B., Swenson, N. G., Svenning, J.-C., 2016. High plant endemism in China is partially linked to reduced glacial-interglacial climate change. *J. Biogeogr.* 43, 145–154. doi:10.1111/jbi.12613
- Gelman, A., Rubin, D.B., 1992. Inference from Iterative Simulation Using Multiple Sequences. *Stat. Sci.* 7, 457–511.

Chapter 1

- Grime, J. P., 1977. Evidence for the Existence of Three Primary Strategies in Plants and Its relevance for Ecological and Evolutionary Theory. *Am. Nat.* 111, 1169–1194.
- Gugerli, F., Englisch, T., Niklfeld, H., Tribsch, A., Mirek, Z., Ronikier, M., Zimmermann, N. E., Holderegger, R., Taberlet, P., 2008. Relationships among levels of biodiversity and the relevance of intraspecific diversity in conservation - a project synopsis. *Perspect. Plant Ecol. Evol. Syst.* 10, 259–281. doi:10.1016/j.ppees.2008.07.001
- Hijmans, R. J., 2016. raster: Geographic Data Analysis and Modeling. R package version 2.5-8.
- Ho, L. S. T., Ane, C. (2014) A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. *Syst. Biol.* 63, 397–408.
- Hobohm, C., 2008. Ökologie und Verbreitung endemischer Gefäßpflanzen in Europa. *Tuexenia.* 28, 10–22.
- Hughes, C. E., Atchinson, G. W., 2015. The ubiquity of alpine plant radiations: from the Andes to the Hengduan Mountains. *New Phytol.* 207, 275–282. doi:10.1111/nph.13201
- Isaac, N. J. B., Turvey, S. T., Collen, B., Waterman, C., Baillie, J. E. M., 2007. Mammals on the EDGE: Conservation priorities based on threat and phylogeny. *PLoS One.* doi:10.1371/journal.pone.0000296
- Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., Mooers, A. O., 2012. The global diversity of birds in space and time. *Nature.* 491, 444–448. doi:10.1038/nature11631
- Kadereit, J. W., Griebeler, E. M., Comes, H. P., 2004. Quaternary diversification in European alpine plants: pattern and process. *Phil. Trans. R. Soc. Lond. B.* 359, 265–274
- Katoh, K., Kuma, K., Toh, H., Miyata, T., 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* 33, 511–518.
- Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., Blomberg, S. P., Webb, C. O., 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics.* 26, 1463–1464.
- Kim, S., 2015. ppcor: Partial and Semi-Partial (Part) Correlation. R package version 1.1.
- Körner, C., 2003. *Alpine plant life - Functional plant ecology of high mountain ecosystems.* 2nd ed. Springer, Berlin. pp 274–290
- Körner, C., 2011. Coldest places on earth with angiosperm plant life. *Alp. Bot.* 121, 11–22.
- Landolt, F., Bäumler, B., Erhardt, A., Hegg, O., Klötzli, F., Lämmli, W., Nobis, M., Rudmann-Maurer, K., Schweingruber, F. H., Theurillat, J.-P., Urmi, E., Vust, M., Wolgemuth, T., 2010. *Flora indicativa - Ökologische Zeigerwerte und biologische Kennzeichen zur flora der Schweiz und der Alpen.* Haupt Verlag, Bern, Stuttgart, Wien.
- Lassmann, T., Sonnhammer, E. L., 2005a. Kalign - an accurate and fast multiple sequence alignment algorithm. *BMC Bioinformatics.* 6, 289–298.
- Lassmann T., Sonnhammer, E. L., 2005b. Automatic assessment of alignment quality. *Nucleic Acids Res.* 33, 7120–7128.
- Mansion, G., Parolly, G., Crawl, A., Mavrodiev, E., Cellinese, N., Oganessian, M., Fraunhofer, K., Kamari, G., Phitos, D., Haberle, R., Akaydin, G., İkinci, N., Raus, T., Borsch, T., 2012 How to Handle Speciose Clades? Mass Taxon-Sampling as a Strategy towards Illuminating the Natural History of Campanula (Campanuloideae), *PLoS ONE.* 7, e50076.
- Molina-Venegas, R., Aparicio, A., Slingsby, J. A., Lavergne, S., Arroyo, J., 2015. Investigating the evolutionary assembly of a Mediterranean biodiversity hotspot: Deep phylogenetic signal in the distribution of eudicots across elevational belts. *J. Biogeogr.* 42, 507–518. doi:10.1111/jbi.12398
- Molina-Venegas, R., Roquet, C., 2014. Directional biases in phylogenetic structure quantification: a Mediterranean case study. *Ecography.* 37, 572–580. doi:10.2478/s11535-008-0035-4.
- Moore, M. J., Soltis, P. S., Bell, C. D., Burleigh, J. G., Soltis, D. E., 2010. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4623–4628.
- Mráz, P., Barabas, D., Lengyelová, L., Turis, P., Schmotzer, A., Janišová, M., Ronikier, M., 2016. Vascular plant endemism in the Western Carpathians: spatial patterns, environmental correlates and taxon traits. *Biol. J. Linn. Soc.* 119, 1095–8312.
- Nagy, L., G. Grabherr, 2009. *The biology of alpine habitats.* Oxford University Press, Oxford. pp 37
- Ozenda, P., 1985. *La végétation de la chaîne alpine dans l'espace montagnard européen.* Masson, Paris.

Chapter 1

- Ozenda, P., Borel J.-L., 2003. The alpine vegetation of the Alps, in: Nagy, L., Grabherr, G., Körner, C., Thompson, D. B. A. (Eds.), *The biology of alpine habitats*. Oxford University Press, Oxford. pp 53–63.
- Pawlowski, B., 1970. Remarques sur l'endémisme dans la flore des Alpes et des Carpates. *Vegetatio* 21, 181–243.
- Pigot, A. L., Tobias, J. A., 2013. Species interactions constrain geographic range expansion over evolutionary time. *Eco. Lett.* 16, 330–338.
- Pigot, A. L., Etienne, R. S., 2015. A new dynamic null model for phylogenetic community structure. *Eco. Lett.* 18, 1461–1468.
- Plummer, M., 2003. JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*.
- Plummer, M., 2016. rjags: Bayesian Graphical Models using MCMC. R package version 4-6.
- R Core Team, 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rangel, T. F., Colwell, R. K., Graves, G. R., Fučíková, K., Rahbek, C., Diniz-Filho, J. A. F., 2015. Phylogenetic uncertainty revisited: Implications for ecological analyses. *Evolution*. 69, 1301–1312. doi:10.1111/evo.12644
- Roquet, C., Boucher, F. C., Thuiller, W., Lavergne, S., 2013. Replicated radiations of the alpine genus *Androsace* (Primulaceae) driven by range expansion and convergent key innovations. *J. Biogeogr.* 40, 1874–1886. doi:10.1111/jbi.12135
- Roquet, C., Thuiller, W., Lavergne, S., 2013. Building megaphylogenies for macroecology: taking up the Challenge. *Ecography*. 36, 013–026.
- Sanderson, M. J., 2003 r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics*. 19, 301–302.
- Schönswetter, P., Stehlik, I., Holderegger, R., Tribsch, A., 2005. Molecular evidence for glacial refugia of mountain plants in the European Alps. *Mol. Ecol.* 14, 3547–3555. doi:10.1111/j.1365-294X.2005.02683.x
- Stamatakis, A., Hoover, P., Rougemont, J., 2008. A Rapid Bootstrap Algorithm for the RAxML Web-Servers. *Syst. Biol.* 57, 758–771.
- Stenseth, N. C., 1984. The Tropics: Cradle or Museum? *Oikos*. 43, 417.
- Smith, S. A., Beaulieu, J. M., 2010 An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants. *Proc. Natl. Acad. Sci. U.S.A.* 107, 5897–5902.
- Stehlik, I., 2003. Resistance or emigration? Response of alpine plants to the ice ages. *Taxon*. 52, 499–510. doi:10.2307/3647448
- Taberlet, P., Zimmermann, N. E., Englisch, T., Tribsch, A., Holderegger, R., Alvarez, N., Niklfeld, H., Coldea, G., Mirek, Z., Moilanen, A., Ahlmer, W., Marsan, P. A., Bona, E., Bovio, M., Choler, P., Cieslak, E., Colli, L., Cristea, V., Dalmas, J. P., Frajman, B., Garraud, L., Gaudeul, M., Gielly, L., Gutermann, W., Jogan, N., Kagalo, A. A., Korbecka, G., Küpfer, P., Lequette, B., Letz, D.R., Manel, S., Mansion, G., Marhold, K., Martini, F., Negrini, R., Nino, F., Paun, O., Pellecchia, M., Perico, G., Piekos-Mirkowa, H., Prosser, F., Puscas, M., Ronikier, M., Scheuerer, M., Schneeweiss, G. M., Schönswetter, P., Schratl-Ehrendorfer, L., Schüpfer, F., Selvaggi, A., Steinmann, K., Thiel-Egenter, C., van Loo, M., Winkler, M., Wohlgemuth, T., Wraber, T., Gugerli, F., 2012. Genetic diversity in widespread species is not congruent with species richness in alpine plant communities. *Ecol. Lett.* 15, 1439–1448. doi:10.1111/ele.12004
- Tribsch, A., Schönswetter, P., 2003. refugia for mountain plants: patterns of endemism and comparative phylogeography confirm palaeo-environmental evidence in the Eastern European Alps. *Taxon*. 52, 477–497. doi:10.2307/3647447
- Tsirogiannis, K., Sandel, B., 2017. PhyloMeasures: Fast and Exact Algorithms for Computing Phylogenetic Biodiversity Measures. R package version 2.1.
- Warren, D. L., Cardillo, M., Rosauer, D. F., Bolnick, D. I., 2014. Mistaking geography for biology: Inferring processes from species distributions. *Trends Ecol. Evol.* 29, 572–580. doi:10.1016/j.tree.2014.08.003
- Väre, H., Lampinen, R., Humphries, C., Williams, P., 2003. Taxonomic diversity of vascular plants in the European alpine areas, in: Nagy, L., Grabherr, G., Körner, C., Thompson, D. B. A. (Eds.), *The biology of alpine habitats*. Oxford University Press, Oxford. pp 133–147

Chapter 1

- Webb, C. O., Ackerly, D. D., McPeck, M. a., Donoghue, M. J., 2002. Phylogenies and Community Ecology. *Annu. Rev. Ecol. Syst.* 33, 475–505. doi:10.1146/annurev.ecolsys.33.010802.150448
- Wood, S. N., 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 73, 3–36. doi:10.1111/j.1467-9868.2010.00749.x

CHAPTER 2

Tempo and drivers of plant diversification in the European mountain system

Jan Smyčka¹, Cristina Roquet^{1,2}, Eric Coissac¹, Marti Boleda¹, Frédéric Boyer¹, Rolland Douzet³, Christophe Perrier³, Maxime Rome³, Cédric Dentant⁴, Adrian Moehl⁵, Bogdan Hurdu^{6,7}, Luc Garraud⁸, Jeremie Van Es⁸, Jean Charles Villaret⁸, Adriana Alberti⁹, Kristýna Šemberová¹⁰, Niklaus Zimmermann¹¹, Wilfried Thuiller¹, Patrick Wincker⁹, the PhyloAlps consortium¹², Sébastien Lavergne¹

in preparation for *Proceedings of the National Academy of Sciences of the United States of America*

¹ Laboratoire d'Ecologie Alpine (LECA), Université Grenoble Alpes, CNRS, FR-38000 Grenoble, France

² Departament de Biologia Animal, Biologia Vegetal i Ecologia, Facultat de Biociències, Universitat Autònoma de Barcelona, ES-08193 Bellaterra, Spain

³ Jardin Alpine de Lautaret, Université Grenoble Alpes, FR-38000 Grenoble, France

⁴ Parc national des Ecrins, FR-05000 Gap, France

⁵ Botanische Garten Bern, CH-3013 Bern, Switzerland

⁶ Faculty of Biology and Geology, Babeş-Bolyai University, RO-400015 Cluj-Napoca, Romania

⁷ Institute of Biological Research, RO-400015 Cluj-Napoca, Romania

⁸ Conservatoire Botanique National Alpin, Domaine de Charance, FR-05000 Gap, France

⁹ Genoscope, Centre National de Séquençage, FR-91057 Evry Cedex, France

¹⁰ Department of Botany, Faculty of Science, Charles University in Prague, CZ-12801 Prague, Czech Republic

¹¹ Swiss Federal Research Institute WSL, CH-8903 Birmensdorf, Switzerland

¹² see appendix S5

Abstract

Temperate alpine floras harbor unusually high biodiversity and endemism given the harsh physical conditions prevailing at high elevations. This invites many fundamental questions that have long remained unresolved. How and when did plant clades speciate and diversify in such harsh environmental conditions that challenge physiological limits of vascular plant life? How did they survive Pleistocene climatic oscillations that severely altered high elevation environments? Here we reconstruct the tempo of species diversification and decipher the main speciation drivers in six angiosperm clades that have diversified in European mountains. To do so, we strictly applied the same approach of plastome phylogenomics based on shotgun sequencing and phylogenetic comparative analyses on all study clades. We found that diversification rates in most of the groups are not severely impacted by Pleistocene climatic oscillations, suggesting that high elevation habitat loss was buffered by migration towards lower elevations. The major driver of speciation is allopatry, important is also speciation associated with bedrock change, but we found no evidence for speciation associated with change of elevational niche. This suggests that the evolutionary assembly of European mountain flora was an interesting interplay between allopatric speciation across mountain islands of varying size and connectivity across the Pleistocene, and climate-induced elevational migrations of high elevation adapted species within these islands, likely suppressing the adaptive diversification across the elevational gradient known from tropical mountain systems.

Introduction

Understanding the historical assembly of biodiversity hotspots remains a fundamental question of evolutionary biology and biogeography. Mountain regions across the world are considered important biodiversity hotspots, owing their high species richness and endemism in large part to a faster pace of species diversification than the surrounding lowlands (1). Tropical regions are known to harbor by far the highest concentration of plant diversity in alpine habitats, but a notable characteristic of temperate mountain biota is their high biodiversity given the exceptionally harsh climatic and edaphic conditions prevailing at high elevations (2), with certain plant lineages occurring as high as 4,000 m in the European Alps (2–4). Many questions remain regarding how and when plant lineages have diversified in such harsh environmental conditions, sometimes seemingly in relatively short time-frames, and whether these diversification dynamics have been impacted by Pleistocene climate changes.

Most theories about why clades undergo rapid diversification involve some form of ecological opportunity. Indeed, the diversification of mountain biota is often described as a continuous process following the uplift of particular mountain range (5, 6), slowing down after saturation of the whole available physical and ecological space (e.g. ref. 7) and eventually being modulated by the Pleistocene climatic oscillations. Such modulation can be positive in terms of net diversification rate, with speciation stimulated by climate-induced range dynamics (speciation pump mechanism, refs. 8, 9), or negative with decreased speciation or increased extinction due to the impact of glaciations. This latter view is motivated by the observation of greater plant endemism in areas that have experienced relative climatic stability (10–12). Contrary to subtropical and tropical mountains such as Northern Andes where plant diversification has a lot occurred during Pleistocene (e.g. refs. 13, 14), it has long been considered that temperate mountains such as the European Alps were so severely glaciated that little plant diversification occurred during the Pleistocene (15, 16). However, the evidence for refugia at the peripheries of glaciated mountain ranges (11, 17–19) and also on ice-free microclimatic islands within the glaciated areas (so called nunataks, refs. 20–23) suggests that unglaciated high elevation habitats remained during glaciation periods and may have triggered speciation of high elevation specialist clades through range fragmentation. No consensus has yet emerged on whether and how did the Pleistocene

Chapter 2

climatic oscillations affected diversification of mountain plants. To date, this was never tested using broad samples of plant clades (see ref. 24 for Primulaceae), and never with common-ground genomic data and dating approaches, which would allow solid generalization.

The unique flora of mountain environments was assembled through a complex interaction between processes of lineage divergence, migration and local extinction. Like in other insular habitats or areas with rugged landscapes, it has long been considered that allopatric divergence was the main speciation driver in temperate mountains (for a recent example, see ref. 24). But this assumption seems to be at odds with the evidence that local adaptation across different elevation belts and bedrocks is pervasive in mountain plants (25, 26), and that parapatric speciation also occurred along these ecological gradients (27–30). In addition, Pleistocene climatic oscillations have differentially impacted rates of migration, speciation and extinction across the major ecological and geographical gradients of European mountains, depending on whether mountain ranges were heavily glaciated or remained ice-free during most Pleistocene – e.g. high elevation, siliceous ranges of central Alps vs. low and mid elevation, calcareous ones of peripheral Alps, respectively. Disentangling the effects of geography and environmental heterogeneity on the evolutionary assembly of mountain flora requires estimating past rates of migration and cladogenesis within and between mountain ranges and environments, based on sufficiently large amount of well-resolved phylogenetic data across replicated plant clades.

In this paper, we explore the diversification histories of European mountain plant clades, in order to reconstruct their past rates and tempo of species diversification with respect to Pleistocene climatic oscillations, and compare the relative importance of different speciation drivers such as geography, geology and elevation. Some of these questions have already been addressed on specific lineages, but using different genomic data and analytical approaches (e.g. refs. 24, 31, 32), thus unfortunately lacking scope to European mountain flora as a whole, and not permitting definitive conclusions regarding the aforementioned questions. Here we selected six mountain plant lineages considered as a representative cases of *in situ* diversification within the European Alpine system. We strictly applied the same approach of plastome phylogenomics based on shotgun sequencing and phylogenetic comparative analyses on all study clades.

Chapter 2

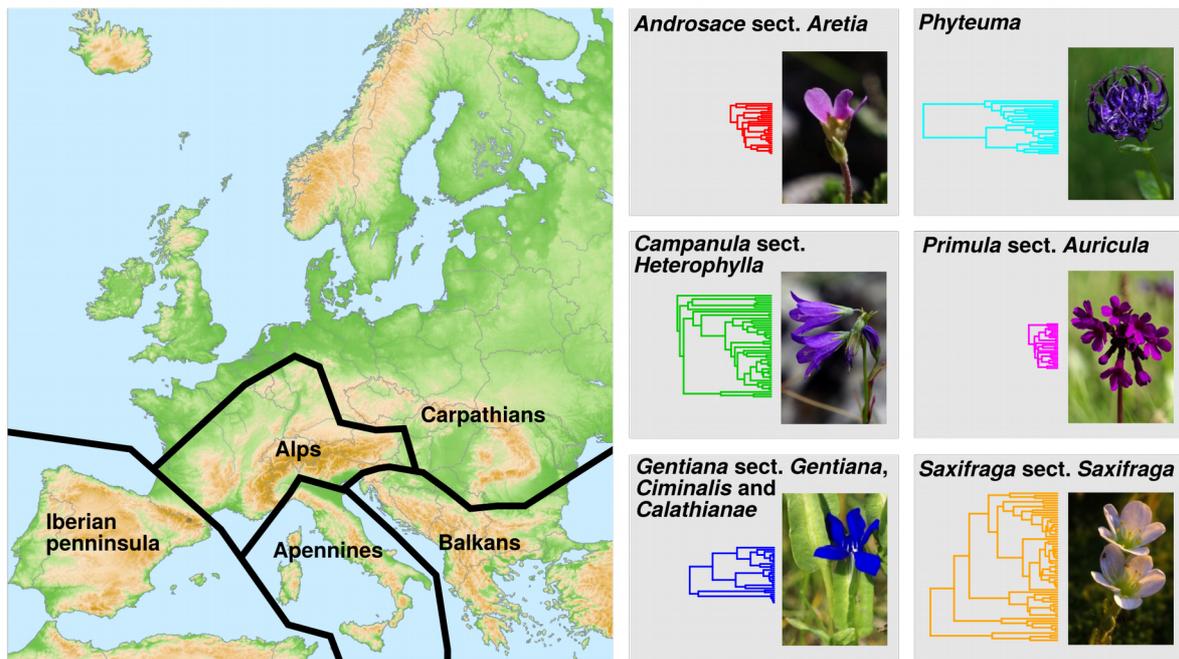


Figure 1: Map of European mountain system with division to 5 major geographic regions and the 6 mountain plant lineages used in this study. The height and width of maximum credibility phylogenies corresponds to time and number of species, color corresponds to color notation used throughout the paper.

Results

Phylogenomic reconstructions. We compiled biological, ecological and genomic data for six plant lineages considered as representative cases of in-situ diversification in the European mountain system (Fig. 1), along with an extensive set of outgroups. The study clades were *Androsace* sect. *Aretia* (*Androsace* hereafter), *Campanula* sect. *Heterophylla* (*Campanula* hereafter), *Gentiana* sections *Gentiana*, *Ciminalis* and *Calanthianae* (*Gentiana* hereafter), *Phyteuma*, *Primula* sect. *Auriculata* (*Primula* hereafter) and *Saxifraga* sect. *Saxifraga* (*Saxifraga* hereafter). For 212 ingroup species (84 % of total 251 ingroup species) we obtained whole chloroplast genomes by performing *de novo* assembly of low coverage shotgun sequencing data. Based on these plastome sequences we built alignment matrices of coding and non-coding sequences of the 4 plant families covering selection of our six lineages (Campanulaceae, Primulaceae, Gentianaceae, Saxifragaceae). This yielded high quality sequence matrices for each study group ranging between 35471-47102 bp for up to 72 coding regions and 2435-5112 bp for up to 17 non-coding regions, with less than 3% missing data overall. These alignments were coupled with dated fossil record and used for building dated phylogenies for each of the 4 families covering our 6 study clades. We obtained well resolved species-level phylogenies with 88% of nodes receiving >0.95 posterior probability.

Tempo of species diversification. The six reconstructed lineages started to diversify in Europe at variable times within the last 40 Ma, all of them however pre-date the onset of Pleistocene in 2.6 Ma before present (Fig. 2A). The oldest lineages of our collection were *Phyteuma* (95% highest probability density of crown age 17.9-32.0 Ma BP) and *Saxifraga* (12.5-38.1 Ma BP), followed by *Gentiana* (6.4-28.5 Ma BP) and *Campanula* (12.8-21.4 Ma BP), with *Androsace* (4.7-10.7 Ma BP) and *Primula* (3.2-7.9 Ma BP) being by far the youngest study clades. All the lineages accumulated considerable amount of diversity both before and during Pleistocene, so none of them is result of exclusively Tertiary or exclusively Quaternary diversification.

We explored different hypotheses about past dynamics of diversification by comparing five models depicting different temporal dynamics of diversification (Table 1). We first ran these models with parameters shared across the six lineages to estimate the general patterns. In order to explore the differences among the six lineages, we subsequently reran the set of five

Chapter 2

models for each lineage separately, and compared their fit to the general model with single set of parameters.

The best performing general model (with diversification dynamics shared across the six lineages) assumed constant speciation and extinction rates (Table 1), but the time- or temperature- dependent models could not be rejected with confidence. Their parameter estimates suggest that if the net diversification rates were variable, they would be decreasing towards the present or in colder geological periods as is the Pleistocene, rather than increasing (see Fig. 2B for estimates of model with temperature-dependent speciation).

The lineage-specific models were slightly outperforming the models with shared diversification dynamics (median $\Delta\text{AIC}=15.58$ on 10 degrees of freedom between the constant rate models), suggesting that the estimates of diversification parameters quantitatively differed across the six lineages (Fig. 2B). Importantly, there were also qualitative differences, i.e. diversification some of the lineages was better explained by other models than constant rates (Table 1). *Androsace* and *Campanula* showed better support for speciation slowdown in colder periods, though the model performance was indistinguishable from the constant rates model. The most notable outlayer was *Primula* showing significant slowdown of speciation either with time (median $\Delta\text{AIC}=2.04$ on 1 d. f.) or in colder geological periods (median $\Delta\text{AIC}=2.09$ on 1 d. f.).

Chapter 2

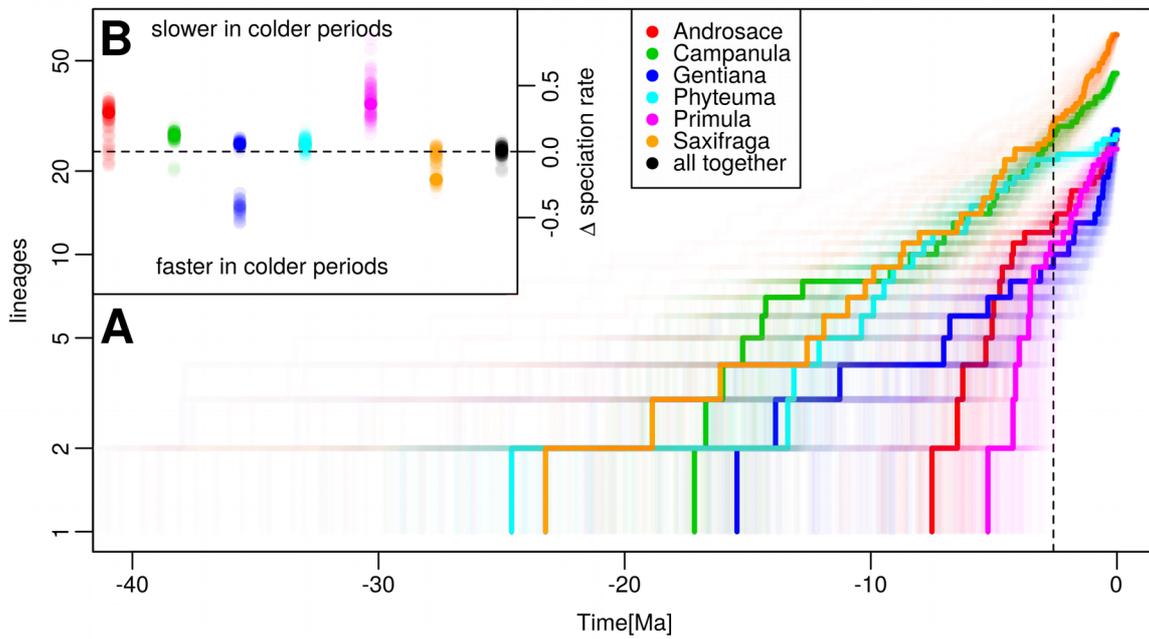


Figure 2: Temporal dynamics of diversification. The panel (A) shows the diversification of lineages through time in semi-logarithmic scale. Thick lines represent maximum credibility phylogenetic reconstructions, semi-transparent lines represent 100 trees from posterior distribution, dashed line marks the onset of Pleistocene at 2.6 Ma before present. The panel (B) shows estimates of the parameter controlling temperature dependence in the temperature-dependent speciation model, with higher values indicating lower speciation in colder geological periods. Thick dots are based on maximum credibility phylogenetic trees, semi-transparent dots on 100 trees from the posterior distribution.

Chapter 2

	<i>Androsace</i>	<i>Campanula</i>	<i>Gentiana</i>	<i>Phyteuma</i>	<i>Primula</i>	<i>Saxifraga</i>	all together
speciation constant, extinction constant	0.00	0.00	0.00	0.00	0.00	0.00	0.00
speciation dependent on past temperature, extinction constant	0.40	0.39	-1.42	-1.64	2.09	-1.70	-1.77
speciation constant, extinction dependent on past temperature	-2.00	-1.33	-0.25	-2.00	-2.01	-1.83	-1.65
speciation exponentially dependent on time, extinction constant	0.00	-0.43	-0.45	-1.86	2.04	-1.81	-1.62
speciation constant, extinction exponentially dependent time	-0.08	-1.18	-0.31	-2.00	1.62	-1.80	-1.69

Table 1: Median Δ AIC values of time- and temperature-dependent diversification models compared against constant speciation and constant extinction model. Values with median Δ AIC>2 are in bold and correspond to significant result in likelihood ratio test, all pairs of compared models differ by one degree of freedom.

Evolutionary assembly: the relative influence of bedrock, elevation and geography. We used cladogenetic state-dependent diversification models (ClaSSE, refs. 33, 34) to shed light on the evolutionary assembly of study lineages across bedrocks (calcareous vs. siliceous habitats), elevation belts (high elevation vs. mid-elevation habitats), and geographic regions (five major European mountain regions, see Fig. 1). To test the relative roles of ecology (bedrock, elevation) or geography on clades evolutionary assembly, we performed AIC comparisons to quantify rates of parapatric speciation (across bedrock or elevation) and allopatric speciation (between regions), as well as rates of in-situ speciation, extinction and migration among bedrocks, elevation belts and regions. The best model was then rerun in Bayesian framework in order to obtain probability envelopes around the parameter estimates. Similarly to diversification models, we fitted models with shared parameters across all lineages and subsequently with different parameters for each study lineage.

The parapatric speciation between siliceous and calcareous bedrocks appeared to be an important driver of plant diversification (Δ AIC=8.32 on 1 d.f.). The parameter estimates

Chapter 2

suggested that a bedrock generalist lineage splits to calcareous and siliceous descendant on average 0.407 times per Ma and such speciation events constitute 18.2% of all speciation events (Table 2). In contrast to this, we found no evidence for parapatric speciation between elevation belts ($\Delta\text{AIC}=-2$ on 1 d.f.) or for allopatric speciation across the five major mountain regions of Europe ($\Delta\text{AIC}=-2$ on 1 d.f.). The latter result however seems to be biased by the coarse scale of considered geographic regions (see sister species analysis below), and by the relative simplicity of the model which assumed only a single rate of allopatric speciation between regions, which is not realistic given the different nature of boundaries among them (e.g. sea between Balkans and Apennines vs hilly landscape between Balkans and Carpathians).

The rates of *in situ* speciation differed in different elevation, bedrock and regions, whereas we inferred a single in-situ extinction rates across all environments and regions. In particular, the speciation rate was higher in siliceous bedrocks than in calcareous ($\Delta\text{AIC}=3.39$ on 1 d.f., Bayesian credibility intervals in Fig. 3A), slightly higher in high elevation habitats than in mid-elevation habitats ($\Delta\text{AIC}=1.21$ on 1 d.f., Fig. 3B), and higher in the Alps and the Iberian mountains than any other European mountain region ($\Delta\text{AIC}=45.18$ on 4 d.f., Fig. S1). Interestingly enough, these differences in speciation rate were not the major force affecting proportions of species across elevational belts or bedrock types, i.e. higher speciation rate in certain habitats did not necessarily mean more species. These proportions, both present-day (Fig. 3C and 3D) and equilibrium proportions to which the models would converge with time $\rightarrow \infty$ (Fig. S2), were strongly influenced by directional migration.

The evolutionary assembly of study clades was strongly driven by directional migration between elevations and bedrocks. The rate of migration from siliceous to calcareous habitats was higher than in opposite direction ($\Delta\text{AIC}=5.30$ on 1 d.f., Fig. 3A), which explains comparable present-day proportions and also equilibrium proportions of species growing on silicate and limestone (Fig. 3C, Fig. S2) despite different speciation rates. We found even more important asymmetric migration rates between elevation belts ($\Delta\text{AIC}=17.79$ on 1 d.f., Fig. 3B), as we inferred strong directional migration from high elevation to mid-elevation habitats. This result captures events of secondary migration toward lower elevations from high altitude ancestors, leading to relatively high present-day and equilibrium proportions of species dwelling at either mid-elevations or both elevation belts (Fig. 3D, Fig. S2).

Chapter 2

Similarly as for temporal dynamics of diversification, the rates of evolutionary assembly differed quantitatively ($\Delta\text{AIC}=15.28$ on 30 d.f. for bedrock, $\Delta\text{AIC}=41.863$ on 20 d.f. for elevation) and also qualitatively among the six lineages. The most notable outliers concerning assembly on bedrock gradient were *Androsace* and *Phyteuma*, having higher proportion of species in siliceous habitats than calcareous (Fig. 3C). In *Androsace*, this was linked to faster diversification in siliceous habitats and symmetrical migration. *Phyteuma* was diversifying faster in calcareous bedrock, and the higher diversity in silicates was linked to strong directional migration (Fig. 3A). A notable exception concerning elevational assembly was *Androsace*, having higher proportion of species in high elevation habitats than in mid-elevations (Fig. 3D) due to directional migration from mid- to high elevations (Fig. 3B). Based on the ancestral state reconstructions (Fig. S3), *Androsace* was also the only lineage of our selection where large portion of evolutionary history took place exclusively in high elevation habitats. Another lineage having relatively high proportion of species in high elevation habitats was *Gentiana*, but the very broad credibility intervals of model parameters did not allow us to explore this pattern deeper.

Chapter 2

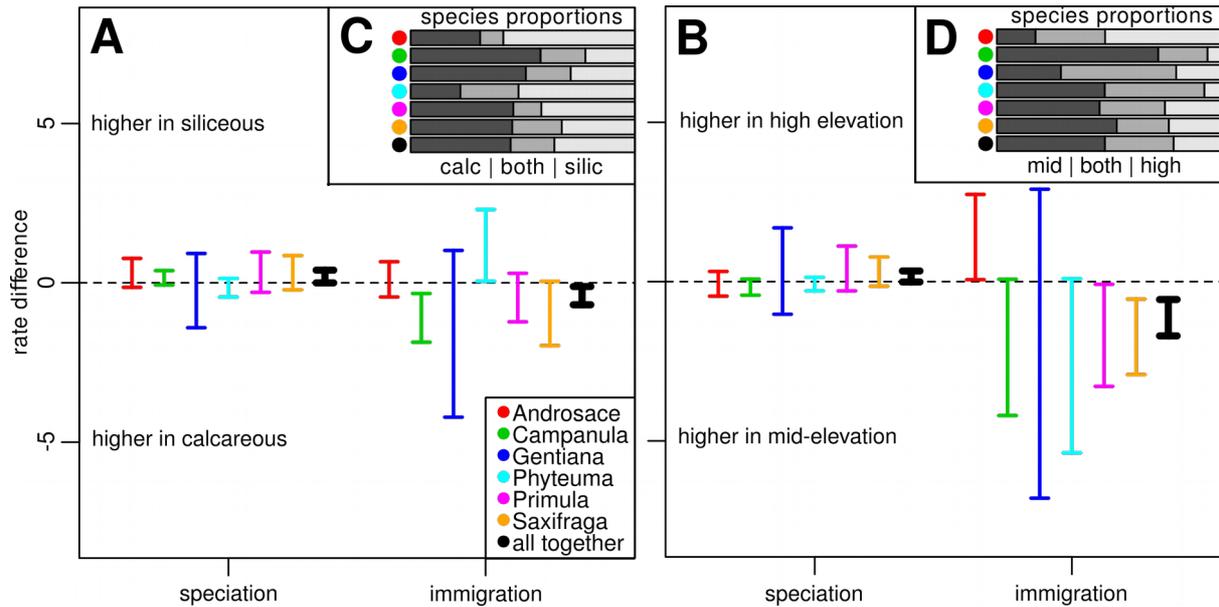


Figure 3: The evolutionary assembly across ecological gradients of bedrock (A, C) and elevation (B, D). The panel (A) shows differences of diversification and migration rates between siliceous and calcareous habitats. The panel (B) shows differences in diversification and migration rates between high and mid-elevation habitats. The bars indicate 95% credibility interval of Bayesian parameter estimates of ClaSSE model. The panel (C) represent proportions of species inhabiting calcareous, siliceous or both types of habitats, and the panel (D) represents proportion of species inhabiting high elevation, mid-elevation or both types of habitats.

	rate of parapatric or allopatric speciation	proportion of parapatric or allopatric to in-situ events (given present-day species proportions)	ΔAIC
bedrock parapatry	0.407	0.182	8.32
elevation parapatry	<0.001	<0.001	-2.00
coarse geographic allopatry	<0.001	<0.001	-2.00

Table 2: The best performing evolutionary assembly models containing parapatric or allopatric speciation, compared against their equivalent without parapatry or allopatry. Values with $\Delta AIC > 2$ are in bold and correspond to significant result in likelihood ratio test, all pairs of compared models differ by one degree of freedom.

Sister species overlap in geographical and ecological space. The SSE model results concerning the importance of allopatric speciation are strongly dependent on considered scale of geographic regions. To tackle this issue, we performed an analysis of elevational belt, bedrock type and geographic range overlaps between sister species, allowing us to use a much finer definition of geographic regions. The range overlap was estimated employing the previously used coarse geographic delimitation (five mountain regions, Fig. 1) or finer geographic delimitation (87 mountain regions, see Appendix S4 and Dataset S3 for definition). Our results showed (Fig. 4) that most of sister pairs diverge on fine geographic scale (39% pairs with no overlap). Similarly as in assembly models, sister species diverged less at the coarse geographic scale (five mountain regions), and across bedrock types (both 23% of pairs with no overlap). The divergence on elevation gradient was the most rare (9% of pairs with no overlap). Although these overlap estimates cannot be used as an absolute measure of different speciation rates, they remain useful for comparing the relative importance of different speciation mechanisms.

The order of speciation patterns remained among all six study lineages, however certain degrees of overlap slightly differed between lineages. Notably, the lineages with least geographic overlap on fine scale were *Primula* (62% of pairs with no overlap) and *Androsace* (43% of pairs with no overlap), suggesting that allopatric speciation was predominant for these two lineages. The lineage *Campanula* exhibited high variance in sister species overlap distribution on fine geographic scale, on one hand having high proportion of sister pairs with no overlap (43%), but on the other hand having the highest proportion of sister species with complete geographic overlap of the dataset (36%).

Chapter 2

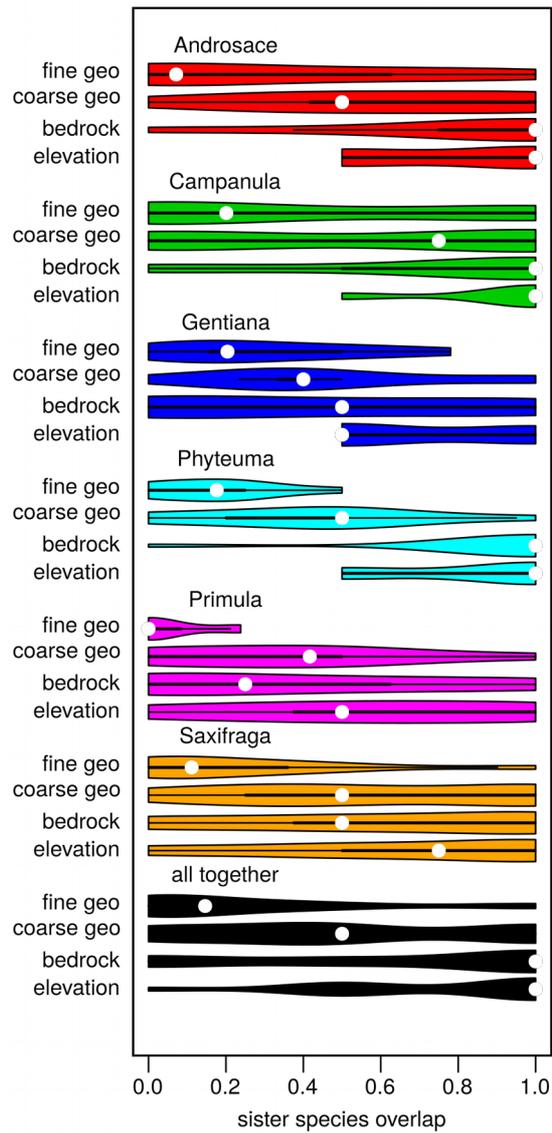


Figure 4: Sister species overlap in geographic and ecological space. Violin plots show distribution of Schoener overlap index at the level of 87 operational geographic units (fine geo), 5 European mountain regions (coarse geo), bedrock and elevational belts. The point within each plot represents median overlap value.

Discussion

Our study explores the main drivers of species diversification and draws interesting generalization into the evolutionary assembly of the flora of the European mountain system. Importantly, our analyses rely on highly resolved phylogenetic reconstructions of six mountain plant lineages, which constitute a representative and quality-balanced sample of plant diversification in European mountains. This study is a significant step given that very few studies to date focusing on the evolutionary history of European mountain clades have attempted to decipher speciation drivers quantitatively, and no comparative analytical study on that topic has yet been done with consistent data based on high-quality phylogenies.

Pleistocene glaciation cycles are thought to have had a crucial impact on the tempo of mountain plant diversification, due to glaciation and downward shift of permanent ice line dramatically reducing extent of mountain habitats, and causing massive extinctions or slowdowns of speciation (16, 24). Surprisingly, the effect of Pleistocene climatic oscillations is neither strong nor homogeneous across the six lineages. All lineages considered together show non-significant diversification slowdown with temperature cooling with minuscule effect size. Considered separately, three lineages show no signs of slowdown, while the other three yes. Moreover the high elevation habitats, and also the region of the Alps, which were most severely impacted by glaciation and shift of permanent snow line, do not show lower rates of diversification. Instead, high elevation lineages exhibit high rates of directional migration to lower elevations. This suggests that the favorable habitats for mountain plants were not severely reduced in glaciation periods but rather displaced. Such habitat displacement could cause much milder and more lineage-specific decline of diversification rates than reduction of habitats, and at the same time would explain the directional migration of lineages from high to mid elevations as a legacy of glacial range dynamics. During the glacial periods, the habitats hosting mountain plants could lie either in the lower peripheral, but still spatially extensive, parts of the mountain ranges, as indicated by high rates of endemism and genetic diversity in such areas (11, 17, 35), or even in the lowland plains, as suggested by evidence of alpine plants in lowland glacial palynological record (*Androsace* sp., *Gentiana* sp., *Saxifraga* sp., but also *Dryas octopetala*, *Polygonum viviparum* or *Saussurea* sp., ref. 36). In contrast to the major process of survival in lower elevations, our dataset also contains an exceptional lineage with stable affinity to high alpine habitats,

Chapter 2

presumably surviving in high elevation nunatak refugia (*sensu* ref. 20). The lineage *Androsace* hosts large proportion of species with exclusive affinity to high alpine habitats, and it can be hypothesized that the detected Pleistocene slowdown of diversification in this lineage might be linked to glacial range reduction.

An interesting question is which factors in general determine whether a lineage slows down or not. The answer can only be hypothesized given that we have only six independent points for comparison, but several possible explanations emerge from the patterns observed across the lineages. Under the ecological opportunity explanation (6, 13), we would expect that the diversification slowdown is strongest in the oldest and most diverse lineages that already saturated their ecological space. This is clearly not the case, with strongest slowdown observed in the youngest and relatively least diverse lineages (*Primula*, *Androsace*). The diversification slowdown could also be linked to the rate of allopatry: the three lineages that slowed down during the Pleistocene (*Primula*, *Androsace* and *Campanula*) also show the highest rates of allopatry (indicated by proportion of sister species with no geographical overlap). The prevalence of allopatric speciation can theoretically lead to intrinsic slowdown of speciation rate in case of low dispersal rate, when all the species have too small ranges to proceed further splitting (37). Another explanation is that Pleistocene climatic oscillations promoted migration between populations in areas that would otherwise stay isolated and thus inhibited allopatric diversification among them, a process antagonistic to the speciation pump mechanism. Simulations of diversification across dynamically fragmented landscapes indeed show that more connected landscapes or a faster pace of connection-disconnection events can lead to lower diversity under certain conditions (9, 38).

Allopatric speciation has long been regarded as a key mode of speciation in European mountains and elsewhere. The high importance of allopatric speciations indicated by our results is consistent with previous studies on European mountain plants (24, 39). An interesting finding of our study is the relatively small spatial scale of allopatric speciation, which is likely linked with low dispersal rate across the five major European regions, as indicated by the SSE model. Although the small scale allopatry is the most important mode of speciation in the European alpine flora, our results suggest that it is complemented by other mechanisms, in particular speciations associated with change of bedrock type. The previous findings that allopatric speciation is the almost exclusive mode in the European mountain flora (24) are thus likely biased by the fact that this study was performed only on Primulaceae

Chapter 2

- a family containing *Androsace* and *Primula*, the two lineages in our dataset for which allopatry is by far the most important mode of speciation.

Our analyses discovered that species divergence in European mountain plants is associated with change of bedrock in a substantial portion of events. This mechanism potentially generates about one fifth of speciation events in our dataset. Bedrock differentiation of species is well known from mediterranean floras (40–43), and in European mountain species it has been studied at the intraspecific level (26), and in a qualitative way (27), but up to our knowledge, we bring the first quantitative evidence of speciation related to bedrock change in mountain plants. It is not clear to what extent these speciation events are truly parapatric, with gene flow among the differentiating populations, and to what extent bedrock specialization takes place in the interplay with allopatry, given that calcareous and siliceous habitats rarely co-occur at small spatial scales. Experimental evidence of differentiation across bedrock types suggests that new species can indeed arise even in presence of gene flow due to divergent selection with severe fitness declines of calcareous and siliceous adapted ecotypes in non-matching habitat (44).

We detected strong migration between bedrocks, and intriguingly this migration was asymmetric, with silicate habitats being a source of limestone lineages. This is in contrast with the prevailing view in literature, that life on calcareous soils requires more specific adaptations than life on silicates (44–47). The stronger migration from silicates to limestones can be explained in two different ways and further evidence would be needed to decide among them. While silicate substrates may be relatively easier to adapt for in context of relatively deep soils, in rocky habitats this effect may be overcompensated by the almost complete lack of nutrients on bare siliceous rocks. Similar effect was shown in experiments with calcicole and silicicole wetland plants, where siliceous wetlands seem to be a more stressful environment (48). Another explanation of higher rate of migration from siliceous to calcareous bedrock could be due to the coincidence between siliceous bedrock and high elevation habitats in the central parts of the mountain systems, which is typical of the European mountain system (49), and thus constituting a confounding hidden factor in the SSE model (*sensu* ref. 50). Indeed, the adaptation from high elevation to mid-elevation habitats seems to be easier than in the opposite direction.

Chapter 2

The evolutionary assembly across the elevational belts seems to be driven by combination of higher diversification rate in high elevation habitats and strong directional migration from high to mid-elevations. The higher diversification rate in alpine habitats is often explained by ecological opportunity that newly emerged and unsaturated alpine habitats confer to colonizing species (6, 7, 13). Under this scenario, the proportion of species in high elevation habitats should be initially low and converge to equilibrium defined by diversification and migration rates across elevational belts. Our results suggest that the ecological opportunity currently does not drive diversification in high elevation habitats, because present-day species proportions are already very close to the predicted equilibrium by the evolutionary assembly model. An alternative explanation for faster diversification may be higher fragmentation of alpine habitats compared to lower elevations, fostering the allopatric speciation.

Interestingly, the faster diversification rate in high elevations is coupled with strong directional migration towards mid elevations, causing that the mid elevation habitats host more species. A part of explanation for this directional migration is that life in lower elevations requires less adaptation compared to the limit conditions in the alpine zone, which is in temperate mountains reinforced by seasonality (51). But the directional migration is also likely fostered by the Pleistocene dynamics causing downward shift of communities during glaciation periods, with alpine lineages leaving relict populations in lower elevations. Such climate-induced migration dynamics would have a contrasting double effect on diversification. On one hand, it would prevent species extinction due to habitat loss in glacial periods as discussed above. But on the other hand, intensive migration is the likely reason why we did not find any evidence for adaptive diversification across elevational belts, a mechanism involved in massive radiations known from Andes and other tropical mountains (1, 14, 29, 52). More specifically, high and mid elevation populations in the European mountain system may differentiate at the intraspecific level (25, 53), but strong migration from alpine to lower elevations is likely to prevent the emergence of reproductive barriers necessary for speciation, leading to lower diversity and lack of fast adaptive radiations in the European mountain system (16).

A notable exception from the assembly patterns across the elevational gradient is the lineage *Androsace*. This lineage has higher richness in high elevations due to directional migration from mid-elevation to high elevation habitats. In fact, several species of *Androsace* sect. *Aretia* occur at the highest limits of vascular plant life in the European mountain system

Chapter 2

(above 4000 m asl, ref. 3), and other species from the genus constitute the highest known occurrences of vascular plants in the world (6350 m a.s.l., ref. 54). Unlike the others in our dataset, the lineage *Androsace* could thus carry pre-adaptations facilitating the repeated entrance into high elevation habitats. A part of such pre-adaptation might be that *Androsace* is the only lineage in our dataset that repeatedly developed a dense cushion life form – an architecture that is linked with life in high elevations in this genus (55) and also in other groups (56). It can also be hypothesized that some of high elevation specialists from *Androsace* were surviving glaciation periods *in situ* in so called nunatak refugia – rocky islands with favourable microclimate protruding the glaciers (57), rather than in peripheral refugia prevalent for the majority of our dataset. The survival in the limited area of nunataks could also explain Pleistocene diversification slowdown observed in *Androsace*, but the importance and consequences of nunatak survival for diversification of high elevation subclades of *Androsace* are to be tested in future.

The diversification in European mountain system seems to be a complex process strongly interplaying with geographic and elevational migration. Importantly, the onset of Pleistocene climatic oscillations did not cause massive extinctions, but rather stimulated strong migration of biota across the elevational gradient. We can hypothesize that these massive elevational migration events on one hand buffered extinctions due to area loss, but on the other hand impeded the Pleistocene adaptive radiations across elevational gradient known for example from the Andes. The majority of speciation events thus took place allopatrically across islands of different mountain habitats, and about one fifth of diversification events can be explained by bedrock parapatry. The dynamics of the mostly allopatric diversification is likely not controlled by saturation of ecological opportunities, but other mechanisms of diversification slowdown may take place. The absence of adaptive diversification across elevational gradient and prevalence of allopatry thus likely contribute to lower richness and different diversification dynamics of the European mountain system compared to tropical mountain ranges.

Methods

Lineage selection, chloroplast genome and phylogeny reconstruction. To get a representative sample of high elevation diversifications in the European mountain system, we selected all the lineages that: are important contributors to European high elevation flora, contain at least 10 species inhabiting alpine and nival elevational belt in the European Alps (based on ref. 58); contain more than 20 species in total; were recently subjected to thorough taxonomic or phylogenetic analysis that revealed that they have center of diversity in Europe; do not follow derived life strategy as is mycotrophy, parasitism or carnivory. This selection resulted in six lineages. Within the PhyloAlps consortium, which is an extensive network of collaborating individuals and institutions (see Appendix S5), we sampled majority of ingroup species and well established subspecies (referred to as “species” throughout the paper) from these lineages. Majority of ingroup species were covered by one sample, but in several cases we included two or more samples to control for taxonomic consistency. For all groups we followed the most up-to-date taxonomic treatments and took into account phylogenetic studies about group circumscription (Appendix S1). As outgroups for our analyses, we sampled an extensive collection of species from families Campanulaceae, Primulaceae, Gentianaceae, Saxifragaceae + Grossulariaceae. The latter one was added to Saxifragaceae due to close relations (59) and lack of fossil calibration points within Saxifragaceae. For details on sample counts and identities, please refer to Appendix S1 and accession tables in Dataset S1.

We extracted DNA from collected samples using a modified protocol of the DNeasy Plant Mini Kit (QIAGEN) and prepared genomic shotgun libraries that were sequenced in Illumina HiSeq 2000. The resulting paired-end reads were used to reconstruct the chloroplast genomes, using a *de novo* organellar assembler based on De Bruijn graph Org.Asm 1.0.3 (60). We retrieved complete circular plastomes except for samples belonging to Campanulaceae (*Campanula* and *Phyteuma* lineages) where the assembly resulted in several contigs that covered the majority of chloroplast coding regions. We worked with these fragmented data in the same way as with the circular plastomes in the other lineages. That is, we annotated all chloroplast coding regions and identified selected non-coding regions available within the contigs. All coding and non-coding regions were aligned gene-by-gene at the family level and concatenated into an alignment matrix. See Dataset S2 for a list regions

Chapter 2

used in phylogenetic inference for each family , and Appendix S2 for more details on sequence processing.

The resulting coding and non-coding alignments for each of the four families were used for estimating dated phylogenies in BEAST 2 (61), using different site models for the non-coding alignment and each codon position in coding alignment, Yule tree prior and lognormal clock with uniform fossil dating priors on at least two nodes of each family. For details on phylogenetic and dating analysis, see Appendix S3. Maximum credibility trees and reduced posterior trees were pruned to only include ingroup species and one individual per species. None of our phylogenies contained polyphyletic species (individuals of one species scattered across phylogeny), in case of 5 paraphyletic species in our dataset (one species arising from within another), the tips were selected to conserve the most recent splitting node, for details of species tree construction see Appendix S1.

The phylogenies based on chloroplast genomes provide high resolution and accurate dating due to combination of large number of regions with variable mutation rates, and the universality of the bioinformatic pipeline allows to obtain common-ground phylogenetic information across virtually any seed plant groups. The limitation of this approach is that chloroplast-based phylogenies are only tracking evolution of maternal lineage, which might be problematic in systems with frequent hybridization, and also on shallow phylogenetic scales, due to incomplete lineage sorting. The absence of polyphyletic and low amount of identified paraphyletic species suggest that neither hybridization nor incomplete lineage sorting are frequent across our ingroups. Moreover, although these artifacts may influence specific topologies, they are unlikely to systematically bias diversification patterns across large datasets above extent that is common for phylogenies based on nuclear markers (e.g. protracted speciation bias *sensu* ref. 62).

Ecological and geographic characteristics of species. We attributed information about elevational and bedrock niche to each of the species from the focal six lineages. For each species we used local floristic literature (see Appendix S4) supplemented with expert knowledge of some groups to estimate their presence in habitats above timberline (referred to as high elevation habitats), below timberline (referred to as mid-elevation habitats),

Chapter 2

calcareous, dolomitic or ultrabasic habitats (referred to as calcareous habitats) and habitats with neutral or acidic soil reaction (referred to as siliceous habitats).

In order to compile information about species geographical occurrence, we defined smallest operational geographical units across Europe for which we could get the credible presence/absence information of every species in our dataset. We used this approach rather than formal processing of point occurrence data (as e.g. in Boucher 2015), because of imbalanced point data quality depending on the countries and broader European regions. For purpose of formal diversification models, operational geographic units were merged to the 5 major mountain regions in Europe and surrounding areas (Fig. 1). For detail on region or ecological niche definitions and data sources, please refer to Appendix S4 and Dataset S3.

Tempo of diversification. To explore hypotheses about timing and tempo of diversification, we fitted the phylogenies of six lineages with the five models of temporal dynamics of diversification presented in the results section. For the models with temperature-dependent speciation or extinction, we used ^{18}O isotope ratio timeline from Greenland ice cores (63). To accommodate for phylogenetic uncertainty, we calculated all the models across 100 trees randomly selected from the Bayesian posterior distribution of phylogenetic trees, and report either whole distribution (parameter estimates) or median of values (AIC comparisons).

In order to explore overall diversification dynamics across our representative sample of European mountain plant diversifications, we developed a method to fit the above described diversification models across multiple evolutionary lineages simultaneously. We assumed that each of our lineages is an independent realization of shared diversification dynamics, which allowed us to construct the overall likelihood functions of temporal diversification models as a product of likelihood functions of each of the six lineages, and harnessing the lineage-specific parameters into single values. Similar approach with overall likelihood was previously used for state-dependent diversification models (33, 64). We optimized the parameters of models using the overall likelihood function with a simplex routine, equivalent to default implementation of single-lineage models in R package RPANDA 1.5 (ref. 65, see Dataset S4 for used scripts).

To test for time- and temperature-dependence of diversification in each of the lineages and also overall, we calculated ΔAIC between constant speciation and constant extinction model

vs. the other respective models. To test whether the dynamics of diversification are indeed shared across the six lineages, or they quantitatively differ among them, we compared the overall model AIC with respective sum of AIC across the six lineage-specific models. Such comparison is meaningful, because sum of AIC of 6 lineage-specific models is equal to AIC calculated from model with overall likelihood function, but with each model parameter kept lineage-specific.

Evolutionary assembly on gradients of bedrock, elevation and geography. The patterns of evolutionary assembly across ecological and geographic gradients were explored using cladogenetic state-dependent speciation-extinction models (ClaSSE, ref. 34), used separately for each of the gradients. The elevation and bedrock gradients contain two habitats each, which allowed us to use the parametrization equivalent to GeoSSE model (33). Geographic gradient contains 5 regions, so we developed a generalization of GeoSSE model that can work with more than two regions (see Dataset S5 for scripts). All the analyses were run using the maximum credibility topologies for each of the 6 focal lineages. Similarly to time-dependent diversification models, we ran all the models for each lineage separately, and also an overall model for all lineages combined, using the likelihood multiplication as described above.

For each of the gradients, we performed a series of AIC comparisons of maximum likelihood model fits to test for importance of cladogenetic (referred to as parapatric or allopatric) speciation, differences in in-situ speciation and extinction rates among the habitats or regions, and directionality or symmetry of dispersal between the habitats or regions. The best model for each of the gradients was subsequently rerun in a Bayesian setup in order to obtain probability envelopes of parameter estimates, using slice MCMC sampler ran for 10000 iterations with 5000 burn-in. The proportions of parapatric and allopatric to in-situ speciation events were obtained by multiplication of present-day numbers of species in different model states (e.g. calcareous specialist, siliceous specialist, bedrock generalist) by respective speciation rates. The analyses were performed using R package diversitree 0.9-11 (66).

Sister species overlap in geographical and ecological space. We used analysis of overlaps between sister species in order to address importance of parapatric and allopatric

Chapter 2

speciations by complementary mean that can work with more geographic units (and thus finer geographic scales), than the ClaSSE models. To do this, we identified all the sister species pairs in maximum credibility phylogenies. For each of the pairs we calculated the Schoener niche overlap index (67) for operational geographic units (fine scale geography), five European mountain regions (coarse scale geography), elevational and bedrock niche. The niche overlaps were calculated using R package spaa 0.2.2 (68).

Acknowledgements

We would like to acknowledge Fabien Condamine, Catherine Graham, Gerald Schneeweiss, Nadir Alvarez, Filip Kolář, Marek Slovák, Tomáš Herben and Laura Pollock for helpful comments at different stages of the project and Marie Smyčková, Doubravka Požárová, Monika Kolářová, Jana Pilátová, Tomáš Figura, Jan Ponert, Petr Koutecký, Petr Macek, Louise Boulangeat, Anna Frattaroli and Llorenç Sáez for help with sampling. The National park of High Tatra, The DNA Bank of the Natural History Museum Oslo, St Andrews Botanic Garden and the Botanical Garden of the Charles University in Prague kindly provided us plant material from their collections.

The research was funded by ANR project Origin-Alps (ANR-16-CE93-0004), France Genomique project PhyloAlps, Labex OSUG@2020 project Nunataks and Cousins project of Pôle Biodiversité du Conseil Général de l'Isère. The sampling campaign in the Balkans region was supported by Bourse de Terrain of French Ecological Society to Jan Smyčka. Jan Smyčka was supported by ANR project Sphinx (ANR-16-CE02-0011) and by Doctoral school of Chemistry and Life Sciences of Université Grenoble Alps. Kristýna Šemberová was supported Mobility funds of Charles University in Prague.

Chapter 2

References

1. Hughes CE, Atchinson GW (2015) The ubiquity of alpine plant radiations: from the Andes to the Hengduan Mountains. *New Phytol* 207:275–282.
2. Körner C (1999) *Alpine Plant Life* (Springer, Berlin).
3. Dentant C, Lavergne S (2013) Plantes de haute montagne: état des lieux, évolution et analyse diachronique dans le massif des Écrins (France). *Bull Soc linn Provence* 64:83–98.
4. Marx HE, et al. (2017) Riders in the sky (islands): Using a mega-phylogenetic approach to understand plant species distribution and coexistence at the altitudinal limits of angiosperm plant life. *J Biogeogr* 44(11):2618–2630.
5. Xing Y, Ree RH (2017) Uplift-driven diversification in the Hengduan Mountains, a temperate biodiversity hotspot. *Proc Natl Acad Sci* 114(17):E3444–E3451.
6. Lagomarsino LP, Condamine FL, Antonelli A, Mulch A, Davis CC (2016) The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). *New Phytol* 210(4):1430–1442.
7. Roquet C, Boucher FC, Thuiller W, Lavergne S (2013) Replicated radiations of the alpine genus *Androsace* (Primulaceae) driven by range expansion and convergent key innovations. *J Biogeogr* 40(10):1874–1886.
8. Haffer J (1967) Speciation in Colombian forest birds west of the Andes. *Am Museum Novit* (2294):1–58.
9. Aguilée R, Claessen D, Lambert A (2013) Adaptive radiation driven by the interplay of eco-evolutionary and landscape dynamics. *Evolution (N Y)* 67(5):1291–1306.
10. Feng G, Mao L, Sandel B, Swenson NG, Svenning JC (2016) High plant endemism in China is partially linked to reduced glacial-interglacial climate change. *J Biogeogr* 43(1):145–154.
11. Smyčka J, et al. (2018) Reprint of: Disentangling drivers of plant endemism and diversification in the European Alps - a phylogenetic and spatially explicit approach. *Perspect Plant Ecol Evol Syst* 30(January):31–40.
12. Molina-Venegas R, Aparicio A, Lavergne S, Arroyo J (2017) Climatic and topographical correlates of plant palaeo- and neoendemism in a Mediterranean biodiversity hotspot. *Ann Bot* 119(2):229–238.
13. Hughes C, Eastwood R (2006) Island radiation on a continental scale: Exceptional rates of plant diversification after uplift of the Andes. *Proc Natl Acad Sci* 103(27):10334–10339.
14. Pouchon C, et al. (2018) Phylogenomic analysis of the explosive adaptive radiation of the Espeletia complex (Asteraceae) in the tropical Andes. *Syst Biol* 67(6):1041–1060.
15. Kadereit JW, Griebeler EM, Comes H (2004) Quaternary diversification in European alpine plants: pattern and process. *Philos Trans R Soc Lond B Biol Sci* 359(1442):265–274.
16. Kadereit JW (2017) The role of in situ species diversification for the evolution of high vascular plant species diversity in the European Alps—A review and interpretation of phylogenetic studies of the endemic flora of the Alps. *Perspect Plant Ecol Evol Syst* 26:28–38.
17. Schönswetter P, Stehlik I, Holderegger R, Tribsch A (2005) Molecular evidence for glacial refugia of mountain plants in the European Alps. *Mol Ecol* 14(11):3547–3555.
18. Escobar García P, et al. (2012) Extensive range persistence in peripheral and interior refugia characterizes Pleistocene range dynamics in a widespread Alpine plant species (*Senecio carniolicus*, Asteraceae). *Mol Ecol* 21(5):1255–1270.
19. Lohse K, Nicholls JA, Stone GN (2011) Inferring the colonization of a mountain range-refugia vs. nunatak survival in high alpine ground beetles. *Mol Ecol* 20(2):394–408.
20. Stehlik I (2003) Resistance or emigration? Response of alpine plants to the ice ages. *Taxon* 52(3):499–510.
21. Schneeweiss GM, Schönswetter P (2011) A re-appraisal of nunatak survival in arctic-alpine phylogeography. *Mol Ecol* 20(2):190–192.
22. Westergaard KB, et al. (2011) Glacial survival may matter after all: Nunatak signatures in the rare European populations of two west-arctic species. *Mol Ecol* 20(2):376–393.

Chapter 2

23. Bettin O, Cornejo C, Edwards PJ, Holderegger R (2007) Phylogeography of the high alpine plant *Senecio halleri* (Asteraceae) in the European Alps: In situ glacial survival with postglacial stepwise dispersal into peripheral areas. *Mol Ecol* 16(12):2517–2524.
24. Boucher FC, Zimmermann NE, Conti E (2016) Allopatric speciation with little niche divergence is common among alpine Primulaceae. *J Biogeogr* 43(3):591–602.
25. Rolland J, Lavergne S, Manel S (2015) Combining niche modelling and landscape genetics to study local adaptation: A novel approach illustrated using alpine plants. *Perspect Plant Ecol Evol Syst* 17(6):491–499.
26. Alvarez N, et al. (2009) History or ecology? Substrate type as a major driver of spatial genetic structure in Alpine plants. *Ecol Lett* 12(7):632–640.
27. Moore AJ, Kadereit JW (2013) The evolution of substrate differentiation in *Minuartia* series *Laricifoliae* (Caryophyllaceae) in the European Alps: In situ origin or repeated colonization? *Am J Bot* 100(12):2412–2425.
28. Gao Y-D, Gao X-F, Harris A (2019) Species Boundaries and Parapatric Speciation in the Complex of Alpine Shrubs, *Rosa sericea* (Rosaceae), Based on Population Genetics and Ecological Tolerances. *Front Plant Sci* 10(March):1–16.
29. Merckx VSFT, et al. (2015) Evolution of endemism on a young tropical mountain. *Nature* 524:347–350.
30. Theodoridis S, Randin C, Broennimann O, Patsiou T, Conti E (2013) Divergent and narrower climatic niches characterize polyploid species of European primroses in *Primula* sect. *Aleuritia*. *J Biogeogr* 40(7):1278–1289.
31. Schneeweiss GM, et al. (2013) Molecular phylogenetic analyses identify Alpine differentiation and dysploid chromosome number changes as major forces for the evolution of the European endemic *Phyteuma* (Campanulaceae). *Mol Phylogenet Evol* 69(3):634–652.
32. Tkach N, et al. (2015) Molecular phylogenetics, morphology and a revised classification of the complex genus *Saxifraga* (Saxifragaceae). *Taxon* 64(6):1159–1187.
33. Goldberg EE, Lancaster LT, Ree RH (2011) Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Syst Biol* 60(4):451–465.
34. Goldberg EE, Igić B (2012) Tempo and mode in plant breeding system evolution. *Evolution (N Y)* 66(12):3701–3709.
35. Mráz P, et al. (2015) Vascular plant endemism in the Western Carpathians: relation to geography, environmental factors and taxon traits. *Biol J Linn Soc* 119(3):630–648.
36. Magyari EK, et al. (2014) Late Pleniglacial vegetation in eastern-central Europe: are there modern analogues in Siberia? *Quat Sci Rev* 95(95):60–79.
37. Moen D, Morlon H (2014) Why does diversification slow down? *Trends Ecol Evol* 29(4):190–197.
38. Aguilée R, Gascuel F, Lambert A, Ferriere R (2018) Clade diversification dynamics and the biotic and abiotic controls of speciation and extinction rates. *Nat Commun* 9(1):1–13.
39. Vargas P (2003) Molecular evidence for multiple diversification patterns of alpine plants in Mediterranean Europe. *Taxon* 52(3):463–476.
40. Kruckeberg AR (1986) An Essay: The Stimulus of Unusual Geologies for Plant Speciation. *Syst Bot* 11(3):455–463.
41. Cowling RM, Holmes PM (1992) Endemism and speciation in a lowland flora from the Cape Floristic Region. *Biol J Linn Soc* 47(4):367–383.
42. Lexer C, et al. (2014) Genomics of the divergence continuum in an African plant biodiversity hotspot, I: Drivers of population divergence in *Restio capensis* (Restionaceae). *Mol Ecol* 23(17):4373–4386.
43. Anacker BL, Strauss SY (2014) The geography and ecology of plant speciation: range overlap and niche divergence in sister species. *Proc R Soc B Biol Sci* 281(1778):20132980.
44. Guggisberg A, et al. (2018) The genomic basis of adaptation to calcareous and siliceous soils in *Arabidopsis lyrata*. *Mol Ecol* 27(24):5088–5103.
45. Davies MS, Snaydon RW (1973) Physiological Differences Among Populations of *Anthoxanthum odoratum* L. Collected from the Park Grass Experiment, Rothamsted. I. Response to Calcium. *J Appl Ecol* 10(1):33–45.

Chapter 2

46. Snaydon RW (1970) Rapid population differentiation in mosaic environment. I. The response of *Anthoxanthum odoratum* populations to soils. *Evolution (N Y)* 24(2):257–269.
47. Zohlen A, Tyler G (2004) Soluble inorganic tissue phosphorus and calcicole-calcifuge behaviour of plants. *Ann Bot* 94(3):427–432.
48. Clymo RS (1962) An Experimental Approach to Part of the Calcicole Problem. *J Ecol* 50(3):707–731.
49. Nagy L, Grabherr G eds. (2009) *The Biology of Alpine Habitats* (Oxford University Press, Oxford).
50. Rabosky DL, Goldberg EE (2015) Model inadequacy and mistaken inferences of trait-dependent speciation. *Syst Biol* 64(2):340–355.
51. Janzen DH (1967) Why mountain passes are higher in the tropics. *Am Nat* 101(919):233–249.
52. Luebert F, Weigend M (2014) Phylogenetic insights into Andean plant diversification. *Front Ecol Evol* 2(June):1–17.
53. Kolář F, et al. (2016) Northern glacial refugia and altitudinal niche divergence shape genome-wide differentiation in the emerging plant model *Arabidopsis arenosa*. *Mol Ecol* 25(16):3929–3949.
54. Dentant C (2018) The highest vascular plants on Earth. *Alp Bot* 128(2):97–106.
55. Boucher FC, et al. (2012) Reconstructing the origins of high-alpine niches and cushion life form in the genus *Androsace* s.l. (Primulaceae). *Evolution (N Y)* 66(4):1255–1268.
56. Boucher FC, Lavergne S, Basile M, Choler P, Aubert S (2016) Evolution and biogeography of the cushion life form in angiosperms. *Perspect Plant Ecol Evol Syst* 20:22–31.
57. Schönswetter P, Schneeweiss GM (2019) Is the incidence of survival in interior Pleistocene refugia (nunataks) underestimated? Phylogeography of the high mountain plant *Androsace alpina* (Primulaceae) in the European Alps revisited. *Ecol Evol* 9(7):4078–4086.
58. Aeschimann D, Rasolofo N, Theurillat JP (2011) Analyse de la flore des Alpes. 2: diversité et chorologie. *Candollea* 66(2):225–253.
59. Ebersbach J, et al. (2017) In and out of the Qinghai-Tibet Plateau: divergence time estimation and historical biogeography of the large arctic-alpine genus *Saxifraga* L. *J Biogeogr* 44(4):900–910.
60. Coissac E (2016) The ORGanelle ASSEMBLER 1.0.3. Available at: <https://git.metabarcoding.org/org-asm/org-asm/wikis/home> [Accessed September 10, 2018].
61. Bouckaert R, et al. (2014) BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Comput Biol* 10(4):1–6.
62. Morlon H (2014) Phylogenetic approaches for studying diversification. *Ecol Lett* 17(4):508–525.
63. Zachos JC, Dickens GR, Zeebe RE (2008) An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature* 451(7176):279–283.
64. Anacker BL, Whittall JB, Goldberg EE, Harrison SP (2011) Origins and consequences of serpentine endemism in the California flora. *Evolution (N Y)* 65(2):365–376.
65. Morlon H, et al. (2016) RPANDA: An R package for macroevolutionary analyses on phylogenetic trees. *Methods Ecol Evol* 7(5):589–597.
66. Fitzjohn RG, Maddison WP, Otto SP (2009) Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Syst Biol* 58(6):595–611.
67. Schoener TW (1970) Nonsynchronous spatial overlap of lizards in patchy habitats. *Ecology* 51(3):408–418.
68. Zhang J (2016) spaa: SPecies Association Analysis 0.2.2. Available at: <https://cran.r-project.org/package=spaa> [Accessed September 10, 2018].

CHAPTER 3

Origin and glacial survival of the cliff dwelling plant *Androsace argentea* C.F.Gaertn. on the rooftops of Western Alps

Jan Smyčka¹, Thibaut Capblancq^{1,2}, Cédric Dentant³, Sébastien Ibanez⁴, Peter Schönswetter⁵, Gerald Schneeweiss⁶, Martí Boleda¹, Wilfried Thuiller¹, Cristina Roquet^{1,7} & Sébastien Lavergne¹

¹ Laboratoire d'Ecologie Alpine (LECA), Université Grenoble Alpes, CNRS, FR-38000 Grenoble, France

² Department of Plant Biology, University of Vermont, VT-05405 Burlington, USA

³ Parc national des Ecrins, FR-05000 Gap, France

⁴ Laboratoire d'Ecologie Alpine (LECA), Université Savoie Mont Blanc, CNRS, FR-73376 Le Bourget-du-Lac, France

⁵ Institute of Botany, University of Innsbruck, A-6020 Innsbruck, Austria

⁶ Department of Botany and Biodiversity Research, University of Vienna, A-1030 Vienna, Austria

⁷ Departament de Biologia Animal, Biologia Vegetal i Ecologia, Facultat de Biociències, Universitat Autònoma de Barcelona, ES-08193 Bellaterra, Spain

Abstract

One of most enigmatic questions concerning the flora of the European Alps is how did the high elevation species survived Pleistocene glaciations. In particular, whether species and populations in the internal heavily glaciated part of the Alps survived *in situ* on nunataks (rocky formations protruding the glaciers), in the less glaciated peripheral regions of the Alps, or in southern ranges of the European mountain system. In this paper we explore the evolutionary history of the siliceous rock specialist *Androsace argentea* (former *A. vandelli*), with a special emphasis on colonization and survival in three regions of internal western Alps. We use here a combination of ddRADseq dataset and chloroplast genome phylogeny to address population processes with high precision and explicit time information, and confront them to past glaciation models. Our results suggest strikingly common long distance dispersal across the whole distribution range of the species, which was seldom followed by admixture among neighboring populations. In line with this, populations in two of the focal internal Alpine regions are resulting from long distance dispersal events that took place very likely after Riss glacial maximum (second most recent, 150 ka BP) and possibly even after last glacial maximum (25 ka BP), one from Pyrenees and the other likely from eastern Alps. The populations in the third internal Alpine region are a part of a larger population grouping that likely survived *in situ* on nunataks over multiple glacial periods. Our results thus suggest that the origins of internal Alpine populations may be fairly diverse, especially in species specialized on extremely insular habitats with common long distance dispersal.

Introduction

The origin and maintenance of high elevation flora has fascinated naturalists and biogeographers since the rise of modern biology (Bliss, 1971; Brockmann-Jerosch and Brockmann-Jerosch, 1926; Darwin and Seward, 1903; de Candolle, 1875; Körner, 1999; von Humboldt, 1805). Interesting discoveries of vascular plants at extremely high elevations were made early on, during the famous expedition of von Humboldt and Bonpland in the Andes at around 4600 m a.s.l., and this upper limit then rose to 6400 m a.s.l. during the 30s and the 50s with the early expeditions to the Everest (reviewed in Dentant, 2018). In the European Alps (the Alps hereafter), angiosperms are known to frequently thrive at elevations higher than 4,000 m a.s.l. (Dentant and Lavergne, 2013). At such altitudes, plants have to handle some of the harshest climatic conditions on Earth, including extremely low temperatures, increased solar radiation, repeated droughts, and tremendous daily temperature variation (Körner, 2011, 1999). High elevation plants not only had to handle above-mentioned stresses, but also had to adapt to past climate changes, that dramatically impacted species ranges during the Quaternary, especially in temperate mountains (Brockmann-Jerosch and Brockmann-Jerosch, 1926; de Candolle, 1875; Schönswetter et al., 2005; Stehlik, 2003). Understanding the evolutionary processes that shaped high elevation plant distributions is important for predicting the fate of these often endemic and patrimonial species, especially in the face of current climate changes (La Sorte and Jetz, 2010; Steinbauer et al., 2018; Thuiller et al., 2014).

One of the most enigmatic questions about the flora of the Alps is how species and populations inhabiting high elevation habitats in the internal part of the mountain range survived the Pleistocene glacial periods. The Alps were periodically covered by glaciers during the Pleistocene climatic oscillations, with non-glaciated areas on the peripheries of the mountain range (Schönswetter et al., 2005; Seguinot et al., 2018) or on high summits of internal Alps emerging from ice-sheets (so called “nunataks”; Carcaillet et al., 2018; Carcaillet and Blarquez, 2017; Delunel, 2010; Stehlik, 2003). In other European mountains, the extent of glaciation was less important (Delmas, 2015; Kłapyta and Zasadni, 2018; Nývlt et al., 2011), but the current ranges of many high elevation plants were still lying in the zone of permanent snow (Ronikier, 2011; Schönswetter et al., 2005). Three main hypotheses have been proposed to explain how high elevation specialist plants could have survived the glacial

Chapter 3

periods in the Alps. The first one (referred to as nunatak refugia hypothesis) suggests that mountain plants survived *in situ*, keeping viable populations on mountain summits protruding the ice-sheet, acting like islands of favorable microhabitat within the permanent snow zone (Escobar García et al., 2012; Stehlik, 2003). The second hypothesis (referred to as peripheral refugia hypothesis, sometimes also referred to as *tabula rasa* hypothesis together with the third one) suggests that mountain plants survived in the refugia on the periphery of the Alpine ice-sheet or the permanent snow zone (Schönswetter et al., 2005; Smyčka et al., 2018; Tribsch and Schönswetter, 2003). Finally, under the third hypothesis (referred to as southern refugia hypothesis), mountain plants would have survived in southern European mountain ranges, which have likely been less impacted by glaciation. This last mechanism was originally proposed for lowland species (Stewart et al., 2010; Tzedakis et al., 2013), but might have taken place in some mountain species as well (Zhang et al., 2001).

Each of these hypotheses is associated with particular recolonization mechanisms, which have likely left tractable imprints on the current genetic structure of populations. The southern refugia hypothesis assumes that the recolonization took place via long distance seed dispersal, over several hundreds of kilometers, leading to severe population bottlenecks (Kropf et al., 2006; Sanz et al., 2014). The colonization from peripheral refugia was likely more gradual, with larger subset of refugial populations migrating towards the central Alpine areas, generating isolation-by-distance genetic signatures (Dullinger et al., 2012; Escobar García et al., 2012; Schönswetter et al., 2005; Thiel-Egenter et al., 2011), even if the long distance dispersals might have played role as well (Schönswetter et al., 2005). The nunatak refugia hypothesis assumes that gradual recolonization occurred from nunataks toward surrounding areas (Escobar García et al., 2012; Marx et al., 2017; Stehlik, 2003), although long distance dispersal remains possible between more isolated summits.

The described survival and recolonization scenarios are the extreme cases that likely contributed in a combined way to build the current distribution and population genetic structures of high elevation species. For example, it was suggested that some nunatak populations could have remained in contact with population from peripheral refugia, washing out the genetic patterns of *in situ* survival (Escobar García et al., 2012; Lohse et al., 2011; Schönswetter and Schneeweiss, 2019; Stehlik, 2003). Because the Pleistocene has consisted of many glacial periods with different severity (Hansen et al., 2006; Zachos et al., 2001) it is also possible that the survival and recolonization mechanisms of particular plant species

Chapter 3

differed depending on the glacial period (Schneeweiss and Schönswetter, 2011). As a theoretical example, one can imagine a high elevation plant species surviving Würm glaciation (most recent; maximum in cca 25 ka BP) in both lowland and nunatak refugia, but more severe Riss glaciation (second most recent; maximum in cca 150 ka BP) only in peripheral or southern refugia.

It is important to note that previous evidence of Quaternary range dynamics of high alpine plant specialists mainly relied on indirect evidence provided by the analysis of spatial patterns of low information genetic markers (Escobar García et al., 2012; Schönswetter and Schneeweiss, 2019; Stehlik, 2003), and alternative glacial and post-glacial recolonization scenarios have not been explicitly tested with model-based and time-explicit approaches. Such knowledge is of paramount importance for understanding the fate of high alpine flora in the face of climatic changes. In particular, a firm demonstration of nunataks refugia in internal mountain ranges of the Alps would suggest that despite the overwhelming glacial cover, some species were able to survive on isolated mountain tops for long periods of time. This can perhaps radically modify our understanding of the possibility of long term population persistence in high elevation specialist plant species, which bears interesting insights into their conservation biology.

Androsace is a genus whose lineages repeatedly colonized the arctic and alpine regions of the Northern Hemisphere, developing a particular cushion life form enabling to handle low temperatures (Boucher et al., 2016; Roquet et al., 2013). In the European mountain system, the *Androsace* sect. *Aretia* is the only evolutionary lineage exhibiting elevational niche conservatism with several closely related species exclusively specialized to high elevation habitats (Smyčka in prep., chapter 2). Different species from the section exhibit different mechanisms of survival during glacial periods: for example *A. alpina*, currently confined to the internal mountain ranges of the Alps, likely survived glacial periods in both nunatak and peripheral refugia (Schönswetter and Schneeweiss, 2019); *A. wulfeniana* survived at the periphery of the Alps (Schönswetter et al., 2003); and widely distributed *A. lactea* survived in multiple regions across Europe, including southern European mountains (Schneeweiss and Schönswetter, 2010).

In this study, we investigate the origin of the internal Alpine populations of *Androsace argentea* (Fig. 1A; former *vandelli*, see Dentant et al., 2018) and explore their mechanisms of

Chapter 3

survival during the glacial periods of Pleistocene. *A. argentea* is a specialist of compact siliceous cliffs occupying broad elevational range including elevational limits of vascular plant occurrence in Europe, widely distributed in the Alps, Pyrenees and Iberian Sierra Nevada (Sierra Nevada hereafter). The species is often found in southern slopes of distinct peaks in internal Alps, suggesting the possibility of nunataks and microclimatic refugia during the glacial periods, even if all the three aforementioned survival scenarios might be valid. In order to identify the dynamics of dispersal and survival of *A. argentea* during the Pleistocene, we use here an extensive sampling of populations across the whole distribution of the species, coupled with denser sampling in the three internal region in western Alps where the species notably reaches its highest elevation records (namely the of Belledonne, Ecrins and Mont Blanc regions, see Fig. 2A). We use here a combination genomic information coming from a ddRADseq experiment and fossil-dated whole plastome (cpDNA) phylogeny, allowing to test different scenarios of origin and demographic histories of the focal internal Alpine populations more directly and time-explicitly.

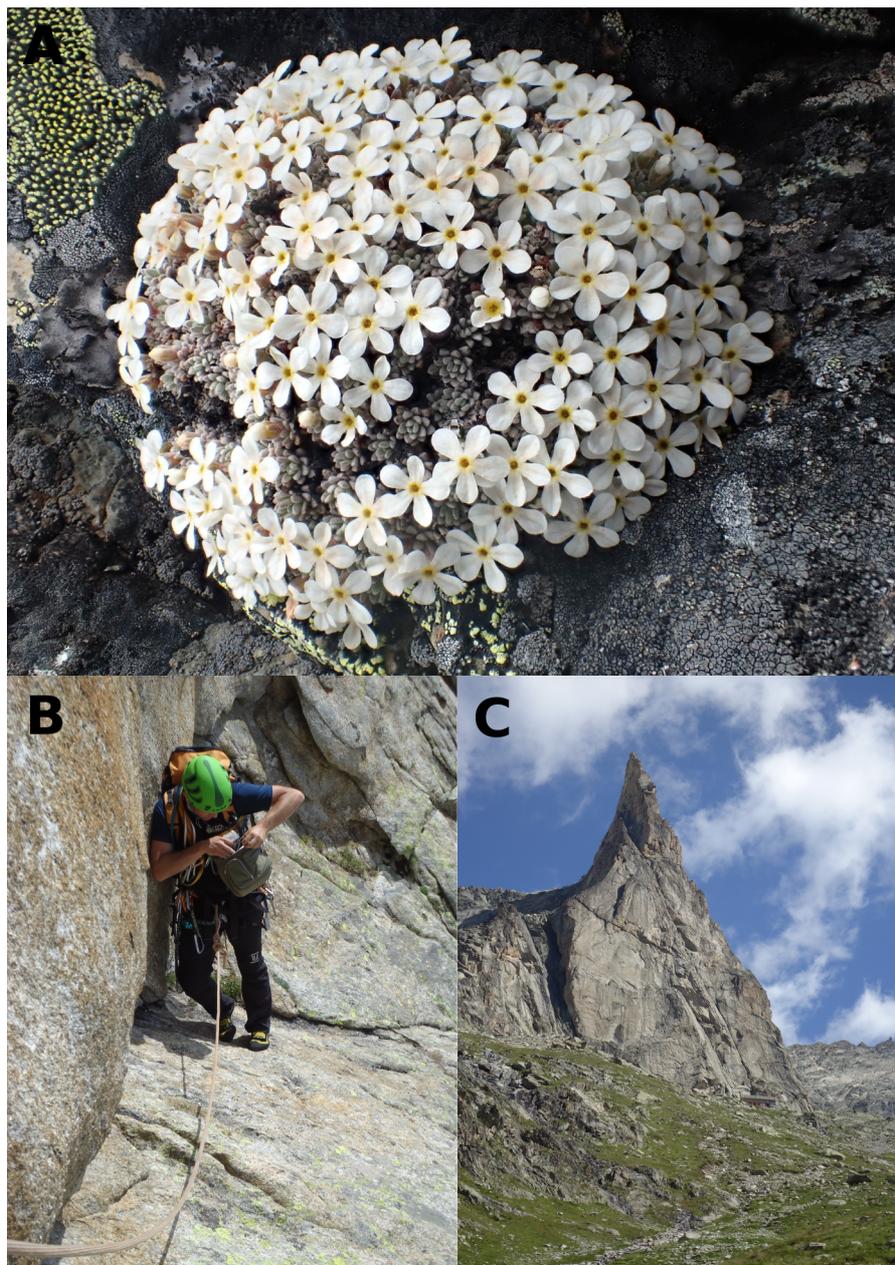


Figure 1: *Androsace argentea* and its habitat. Panel B shows sampling in of Dibona II population with use of mountaineering techniques. Panel C shows south face of Aiguille Dibona in Ecrins, as an example of putative glacial nunatak (Dibona I population at the foot, Dibona II population in the face, Dibona III population below the summit).

Methods

Field sampling, preparation of ddRADseq libraries and SNP calling

In order to obtain genetic data for *A. argentea*, we sampled 60 populations across the whole distribution range of the species in the Alps, Pyrenees and Sierra Nevada (see Fig. 2A for locations of sampled populations, Appendix S4 for population accessions table). In addition to this, we performed a denser sampling in three internal high elevation regions in western Alps (Belledune, Ecrins, Mt Blanc; see Fig. 2), with a particular focus on populations lying on putative glacial nunataks and climatic microrefugia, i.e. high and distinct peaks and steep south-oriented rock faces, often accessible only with the use of mountaineering techniques (Fig. 1B and C). Each of the population samples consisted of 1-10 individuals (median 5) collected at least 1 m apart to assure that different genetic individuals were sampled. In total we sampled 292 individuals (see Appendix S5 for sample list). We also included the following samples as outgroups for phylogenetic network inference: 7 samples of *A. alpina*, one sample of *A. helvetica* and one sample *A. lactea*, which were extracted and sequenced in the same way as the *A. argentea* samples. All plant material was dried and stored in silicagel.

The collected samples were extracted using a modified protocol of the DNeasy Plant Mini Kit (Quiagen, Germany). A double-digested RAD (restriction site associated DNA) experiment was conducted on 287 individuals (seven different libraries) using a modified version of the protocol described in Peterson et al. (2012). A total of 150 ng of DNA template from each individual were double-digested with 10 units each of PstI and MspI (New England Biolabs Inc.) at 37°C during one hour in a final volume of 34 µl using the CutSmart buffer provided with the enzymes. Digestion was further continued together with ligation of P1 and P2 adapters (see Peterson et al., 2012) by adding 10 units of T4 DNA ligase (New England Biolabs Inc.), adapters P1 and P2 in 10-fold excess (compared to the estimated number of restriction fragments) and 1 µl of 10 mM ribo-ATP (New England Biolabs Inc.) in each sample. This simultaneous digestion-ligation reaction was performed on a thermocycler using a succession of 60 cycles of 2 min at 37°C for digestion and 4 min at 16°C for ligation. An equal volume mixture of all the digested-ligated fragments were purified with Agencourt AMPure XP beads (Beckman Coulter, France). Fragments were size-selected between 200 and 600 bp on a Pippin Prep. The ddRAD libraries obtained were amplified using the following PCR mix: a volume of 20µl with 2 µl of DNA template, 10 mM of dNTPs, 2 µM of

Chapter 3

each PCR probe (Peterson et al., 2012) and 2 U/ μ l of Taq Phusion-HF (New England Biolabs Inc.); and the following PCR cycles: an initial denaturation at 98°C for 30 sec; 15 cycles of 98°C for 10 sec, 66°C for 30 sec and 72°C for 1 min; followed by a final extension period at 72°C for 10 min. The amplified ddRAD libraries were purified with QIAgen MinElute PCR Purification Kit (Qiagen, Germany). The seven libraries were then sequenced on a complete lane of Illumina Hi-Seq 2500 2x125 (Fasteris SA, Switzerland). DNA reads (~250 million reads) resulting from the sequencing were used to genotype single-nucleotide polymorphism (SNP).

We processed the bioinformatic treatment of the ddRAD sequences with the STACKS program (Catchen et al., 2013) using a Phred score of 10 for reads filtering (process_radtags function), a minimum coverage of 3 to create a stack (-m 3 in ustacks function) and a maximum of 6 different nucleotides to merge two different stacks (-M 6). Highly-repetitive stacks and over merged tags were dropped using both the “Removal algorithm” (-r) and the “Deleveraging algorithm” (-d). Furthermore, a maximum of 10 mismatches was allowed for considering two individual tags as the same locus and to merge them in the catalogue (-n 10 in the cstacks function). In order to evaluate the experimental success for each library and each individual, we used ProcessMyRAD scripts (Cumer et al., 2018) to produce graphical outputs at different steps of the treatment. This pipeline provides the proportion and number of reads retained after quality filtering, sequencing nucleotide quality along the reads, and the proportion of the different nucleotides. It also allows repeating the critical step of *de novo* individual locus assembly (ustacks) using a range of M values. By looking at the number of assembled RAD tags across M values it provides a way to find a pertinent value for this threshold (Fig. S3, M=6 has been chosen).

Population genetic structures

The SNP dataset obtained by the *de novo* calling procedure was filtered to contain only loci with less than 40% of missing data and samples with less than 40% of missing data. To get rid of paralogous regions and regions under selection, we also excluded all loci that were significantly ($p < 0.05$ in exact test) out of Hardy-Weinberg equilibrium. To reduce the risk of sequencing errors, we excluded all loci with minor allele frequency lower than 1%. After this filtering, the data matrix contained 4855 SNP for 240 individuals. The filtering and data

Chapter 3

manipulation were performed in R 3.6.0 (R Core Team, 2018), using packages *adegenet* 2.1 (Jombart and Ahmed, 2011), *vcfR* 1.8 (Knaus and Grünwald, 2017), *dartR* 1.1.11 (Gruber et al., 2019) and *HardyWeinberg* 1.6.2 (Graffelman, 2015).

The resulting dataset was used for running analyses aimed to describe population structure. We identified the number of genetically distinct population clusters and potential admixture among them using the sparse nonnegative matrix factorization approach implemented in the function *snmf* from the R package *LEA* (Frichot and Francois, 2014). We will refer to this analysis hereafter as the *snmf* analysis. We run *snmf* with 10 independent repetitions for a number of clusters (*K*) between 1 and 20. The optimal number of clusters was selected via cross-entropy criterion as implemented in *LEA*. To make sure that the *snmf* analysis was not biased by denser sampling in Belledonne, Ecrins and Mont Blanc regions, we reran the analysis with a sub-selection of only two populations from each of these regions. As the results were consistent between the two datasets (see Fig. S1), we used the original one. To explore the population genetic structures of our dataset using both a continuous and a hierarchical approach, we performed a PCA and constructed a neighbor joining tree based on the Nei genetic distance matrix among the populations.

To explore which of the identified population clusters in the Alps are lying in formerly glaciated or ice-free areas, we mapped all the populations in the Alps on two different glaciation extent estimates. The first estimate was an informal mapping of peripheral glacial refugia on siliceous bedrocks during last glacial maximum, as presented in Schönswetter et al. (2005). The second estimate were ice-free areas based on maximum ice extent in digital glaciation model from Seguinot et al. (2018). For all the populations in the Alps, we calculated a distance to the margins of closest ice-free areas based on these glaciation estimates. In addition to this, we mapped the populations in Ecrins and Belledonne clusters also on glaciation estimate from Delunel (2010) based on glacier trim lines and cosmogenic nuclids dating, that is available for Ecrins and Beledonne regions. This later analysis is presented in Future directions section rather than Results. The geoinformatic data processing was performed with R packages *raster* 2.8-4 (Hijmans, 2019) and *rgdal* 1.2-12 (Bivand et al., 2019).

Phylogenetic network

In order to explore the relationships of different populations of *A. argentea*, and to identify the closest relatives to the three focal mountain regions, we constructed a phylogenetic network of populations based on SNP data using the maximum likelihood TreeMix algorithm (Pickrell and Pritchard, 2012). We adopted a network approach, rather than a standard phylogenetic tree, because the evolution within a species cannot be assumed to be tree-like due to possible admixture among populations. The results of snmf analysis and also the analysis of chloroplast genomes (see Results) indeed suggest that some of our populations are a result of admixture, which would lead to artificial positioning of these population close to the root in standard tree methods, and biasing the tree topology as a whole (Pickrell and Pritchard, 2012). The optimal number of admixture edges was selected using the delta M statistic as implemented in the R package OptM 0.1.3 (Fitak, n.d.) based on 10 bootstrap replicates for numbers of edges between 0 and 8. The final topology was estimated using the estimated number of edges in a maximum likelihood model fitted without bootstrapping, and support values of nodes were generated by 100 bootstrap replicates of this final model. The network manipulation and plotting was performed using R packages ape 5.1 (Paradis et al., 2004) and phytools 0.6-44 (Revell, 2012).

The SNP dataset used for phylogenetic network inference was assembled and filtered in the same way as the one for inference of population structures, but in order to root the network topology with outgroups, we included seven samples of *A. alpina*, one sample of *A. helvetica* and one sample of *A. lactea* in the data matrix. In order to keep the outgroups in the dataset, we also modified the thresholds for missing loci and individuals from 40% to 50%. The data matrix after filtering contained 4617 SNP for 260 individuals including outgroups.

Demographic history

To investigate the demographic histories of *A. argentea* populations, we performed a Stairway plot analysis (Liu and Fu, 2015) on each of the clusters identified with the snmf analysis (using individuals with $q > 0.5$ affinity to one of the clusters). Stairway plot is an exploratory method to infer past population size dynamics from site frequency spectrum (SFS). However, the absolute timing of inferred demographic events is linearly dependent on the assumed mutation rate in absolute time units, and the absolute effective population sizes

Chapter 3

are linearly dependent on the assumed generation time (Liu and Fu, 2015) – two variables that cannot be easily approximated for *A. argentea*. To address this, we ran the Stairway plot analysis twice with values of mutation rate $5 \cdot 10^{-8}$ substitutions/site/year or $5 \cdot 10^{-10}$ substitutions/site/year representing relevant extent of mutation rates previously used for RADseq dataset in *Primula farinosa* (Theodoridis et al., 2017), and generation time 10 years or 50 years representing a conservative estimate based on experience with cultivation of cushion species of *Androsace* (based on collections of Lautaret Alpine Botanical Garden). The results obtained using the two pairs of parameters serve as rough estimates of minimum and maximum timing of obtained demographic events and minimum and maximum effective population sizes. Given the large variance in these, we however focus on interpretation of relative demographic dynamics across different snmf-defined clusters, rather than on the absolute timing and population sizes.

The SNP dataset used for Stairway plot reconstruction was filtered in the similar way as the one for inference of population structures. Instead of using filter for 1% MAF, we ran the Stairway plot procedure ignoring singletons. The filters for maximum 40% of missing loci and samples were applied within each snmf-defined clusters rather than globally and the site frequency spectra were calculated via subsampling to deal with the missing data (Bagley et al., 2017).

Chloroplast genome phylogeny

For one sample from the majority of populations (see Appendix S5 for sample list), we performed low-coverage sequencing of the DNA without the ddRAD treatment, in order to be able to reconstruct the chloroplast genome phylogeny of such samples, serving as independent information about the timing of divergence events of maternal lineages. In order to obtain the complete sequences of the chloroplast genomes, we performed *de novo* assembly of the obtained pair reads using an organellar assembler based on De Bruijn graph (Coissac, 2016). This procedure resulted in 41 circular reconstructions of chloroplast genomes from different populations of *A. argentea*, and we supplemented the data with previously reconstructed chloroplast genomes of other species of *Androsace* (Smycka in prep., chapter 2) and additional 12 chloroplast genomes of *A. lactea*, which is widely distributed and a close relative of *A. argentea*.

Chapter 3

We then extracted the sequences of all coding and selected non-coding plastome regions, and used them for building a dated phylogeny of maternal lineages. This was done in BEAST 2 (Bouckaert et al., 2014), using separate site models for the non-coding alignment and each codon position in coding alignment, coalescent tree prior and lognormal clock. In order to obtain dated trees, we applied an uniformly distributed secondary calibration of the root node (split between *A. lactiflora* and *A. septentrionalis* vs tips belonging to *Androsace* sect. *Aretia*) with an upper bound of 27.5 Ma BP and a lower bound of 12.5 Ma BP, based on the chloroplast genome phylogeny of Primulaceae dated with fossil calibration (Smycka in prep., chapter 2). This assumed secondary calibration interval is also consistent with previous dated phylogenies of the genus *Androsace*, putting the stem age of sect. *Aretia* to 12.5 Ma BP (Boucher et al., 2016) or 22 Ma BP (Roquet et al., 2013). The BEAST was run for 100 M iterations, with 30% burn-in, and convergence was checked visually and using ESS scores. The resulting maximum credibility tree was retrieved with TreeAnnotator procedure.

Results

ddRADseq experiment and SNP calling

The ddRAD libraries produced a mean of 44124 RAD-fragments with a mean coverage of 15.51 reads/fragment for the 292 individuals (240 after filtering). The polymorphism in these fragments allowed the calling of 7456 SNPs (4855 after filtering). The influence of the clustering parameter M on fragment reconstruction was visually inspected and guided the choice of $M=6$ (Fig. S3). We can observe in the figure S3 that M values lower than 6 return very small proportion of polymorphic RAD-fragments when values above 6 do not increase consequently the proportion of polymorphism but could increase the number of paralogues in the dataset.

Population genetic structures

The cross-entropy criterion of the snmf classification was decreasing with the increasing number of clusters until it reached a plateau at $K=13$ (Fig. S2). We thus define the genetic clusters based on $K=13$ throughout the manuscript. This relatively high number of clusters was selected as a conservative choice, in order to assure that clusters of populations used for demographic inference were not too genetically heterogeneous. Decreasing K below 13 would mean working with a large portion of admixed clusters, and also clusters of populations that are geographically disparate and overlapping with other clusters, hampering the assumption that the clusters are structural units with frequent gene flow within, but rare among.

The populations of Belledonne, Ecrins and Mont Blanc regions were identified as consistent clusters (Fig. 2A), and this result remained after thinning out these densely sampled areas (Fig. S1). The other identified clusters were also geographically consistent, with the exception of the cluster Mercantour 2, which contained two populations of southwestern Alps (Pierre-André and Mont Viso) and one population of eastern Alps (Stilfser Joch). We attributed a name to each cluster referring to its geographic position and use these names throughout the manuscript (see legend of Fig. 2). We also identified three populations with an important amount of admixture, with individuals exhibiting $q < 0.5$ affiliation to any of the clusters: the population of Vall de Cardós in eastern Pyrenees, the population of Valle Cervo

Chapter 3

south from central Alps, and a population of Dent Blanche in central Alps consisting of a single individual with an intermediate position between the Mont Blanc cluster and the neighboring Valais cluster.

To further explore the genetic structure of our dataset in a more continuous way, we performed a principal component analysis (Fig. 2B). The clusters identified by snmf were well separated along the first two axes of the PCA. The first axis (10.8% of explained variance) discriminated the individuals from the Ecrins cluster, together with the Lagorai cluster in eastern Alps, the eastern Pyrenees cluster and the western Sierra Nevada cluster, against the rest of the dataset. The second axis (7.8% of explained variance) discriminated individuals from the Ticino cluster against the rest of the dataset.

The hierarchical population structure was explored by a neighbor joining approach based on the genetic distances among populations (Fig. 2C). The grouping of populations was, here again, consistent with the clusters identified with snmf, the only exception were the clusters of Mont Blanc and Valais, given that the Mont Blanc cluster was here included within the variability of Valais cluster. The analysis suggests that the populations the most closely related to the Belledone cluster are the ones in western Pyrenees cluster, and the most closely related to the Ecrins cluster are the ones in Lagorai cluster in eastern Alps and the admixed population of Valle Cervo south from central Alps.

The majority of clusters in the Alps contained at least one population that was lying in a glacial refugium as identified in Schönswetter et al. (2005), with exception of the clusters Valais, Mont Blanc, Ecrins and Mercantour 2 (Fig. 3). In contrast to this, the majority of populations in the Alps were lying outside refugia as identified by Seguinot et al. (2018), with the exception of one population in the Belledonne cluster. This sharp difference between Schönswetter et al. (2005) and Seguinot et al. (2018) may suggest that the later source overestimated glaciation extent.

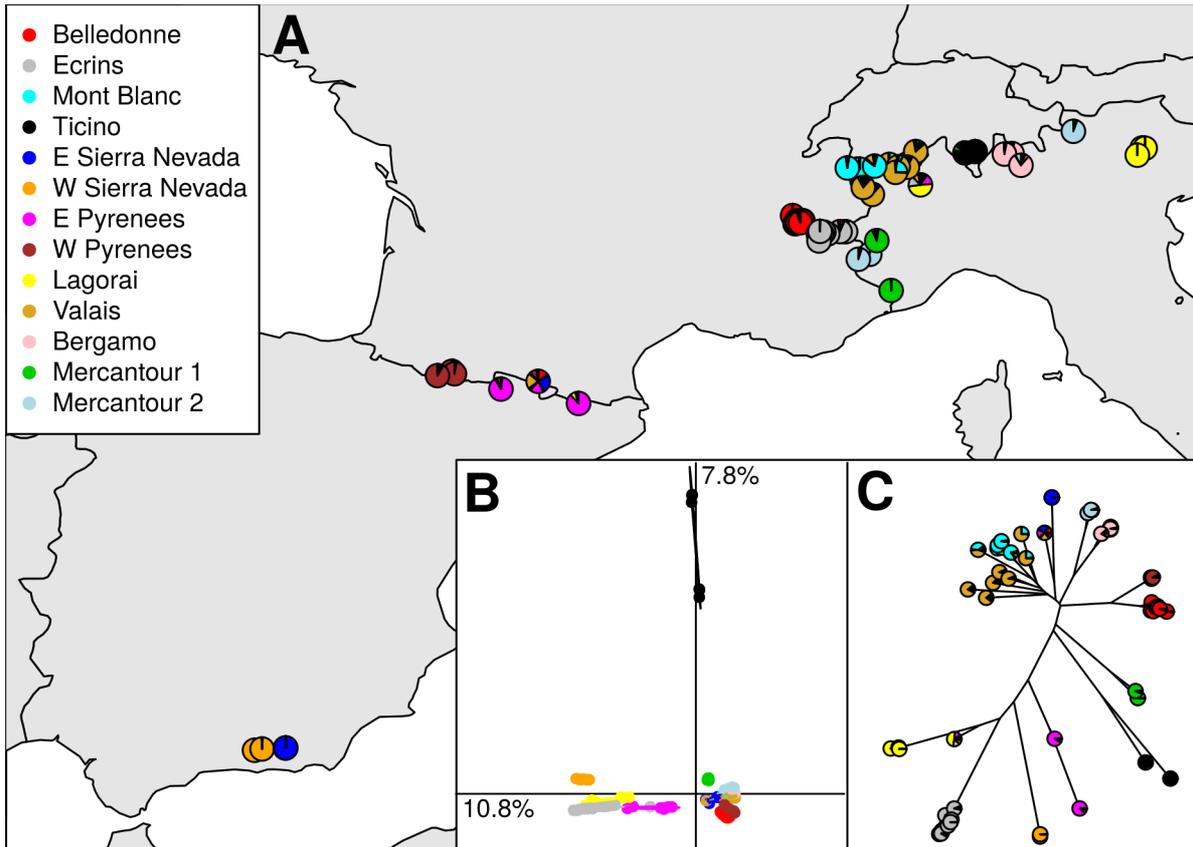


Figure 2: Range-wide population structure of *A. argentea*. Panel A shows geographic distribution of sampled populations; pies indicate positions of populations and reflect proportional assignment of populations to the respective snmf clusters. Panel B shows a first and second axis of principal component analysis of *A. argentea* individuals; colors are attributed based on maximum assignment to snmf clusters. Panel C shows a neighbor joining tree of *A. argentea* populations. The legend refers to the names of snmf clusters that are used throughout the manuscript.

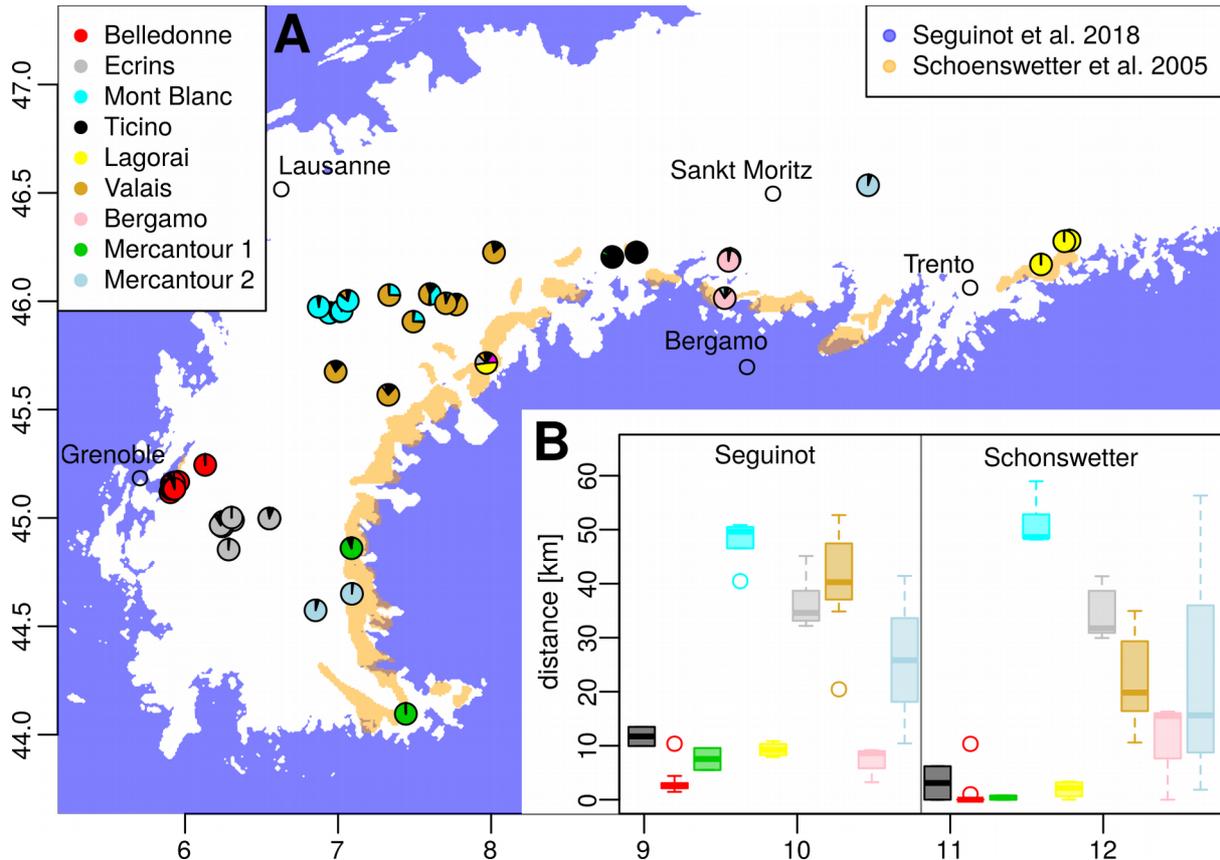


Figure 3: Populations in the Alps mapped on two different estimates of ice-free areas in last glacial maximum. Panel B shows distances of populations grouped in different clusters to the closest ice-free area. Pies indicate positions of populations and reflect proportional assignment of populations to the respective snmf clusters.

Phylogenetic network

The identified optimal number of admixture edges on the phylogenetic network resulting from TreeMix (Fig. 4) was six, two of them connecting the basal nodes of *A. argentea* with the outgroups, possibly reflecting a scenario of introgression from various lineages of *Androsace* sect. *Aretia* (see below). Two of the internal admixture edges connected geographically proximate clades in Pyrenees, another one connected geographically proximate clades south from central Alps, and the fourth internal admixture edge connected the ancestor of the clade formed by the Ecrins cluster in western Alps with the clade formed by the western Sierra Nevada cluster.

The tree topology suggested that the most ancestral populations in our dataset were the ones in Cima della Trosa and Cima dell'Uomo forming the Ticino cluster of central Alps

Chapter 3

(bootstrap support 100, Fig. 4). The clades of populations were generally consistent with the clusters identified by snmf, with the exception of the Valais cluster, which was paraphyletic (BS 100) given that it contained the monophyletic clade of Mont Blanc cluster (BS 100) and the western Sierra Nevada cluster (BS 100). The two independent clades formed by population clusters from the Pyrenees and two independent clades formed by the Sierra Nevada clusters (all BS 100) were positioned among the clades from various locations in the Alps.

The populations from Belledonne cluster formed a monophyletic clade, and their closest relative was a clade formed by the western Pyrenees cluster (BS 100). The populations from Ecrins cluster formed a monophyletic clade that was sister to the clade formed by the Lagorai cluster populations from eastern Alps, however this relationship was only moderately supported with a bootstrap of 77% . The closest relative of the monophyletic clade of Mont Blanc cluster was a population in Val de Bagnes from Valais cluster (BS 99).

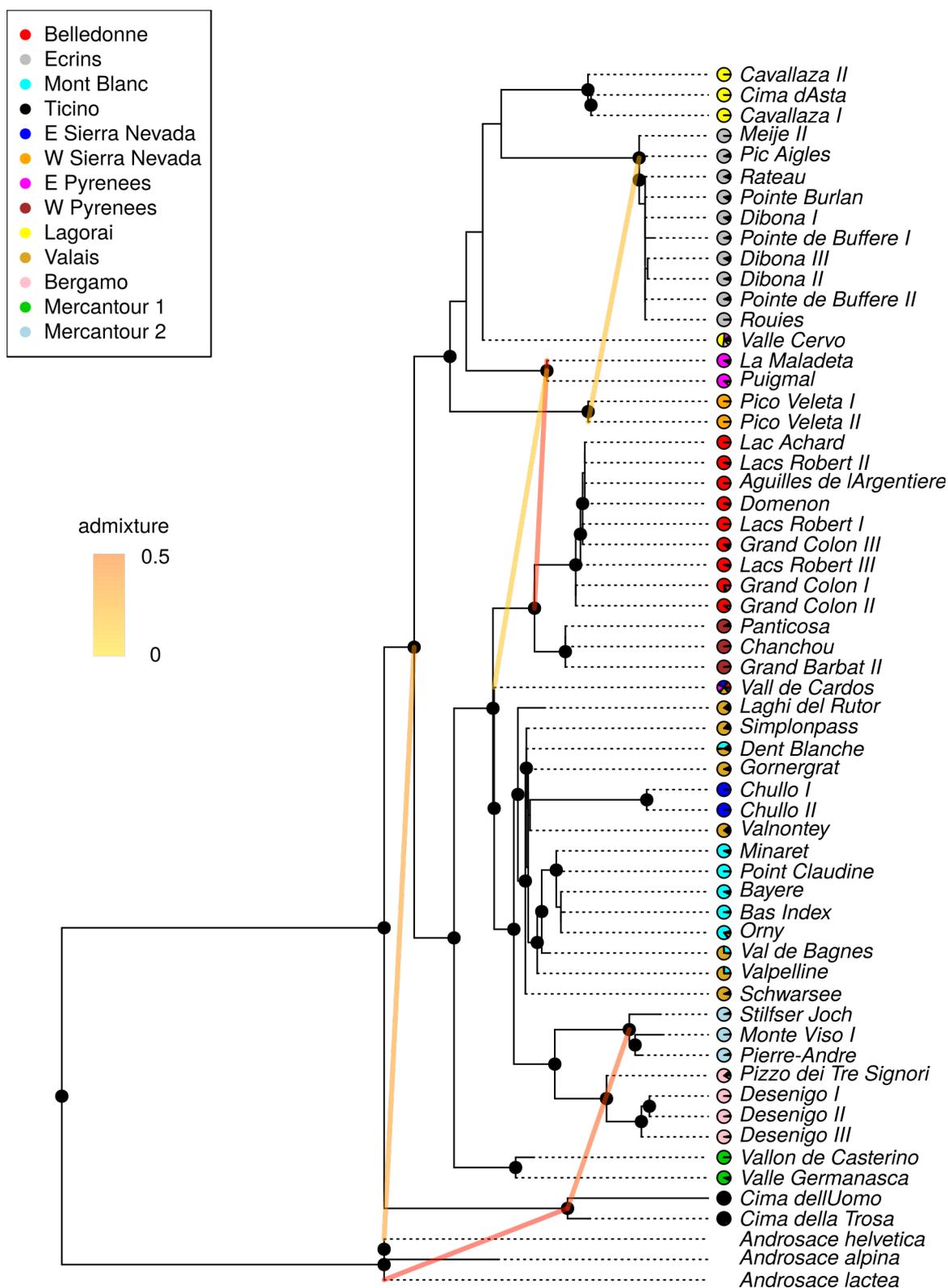


Figure 4: A TreeMix phylogenetic network of sampled populations of *A. argentea* and their outgroups. The dots indicate nodes with bootstrap support > 90, the intensity of admixture edges is indicated on heat scale, the pies reflect proportional assignment to the respective snmf clusters.

Demographic history

According to the Stairway plot analysis, the Belledone cluster (Fig. 5A) had experienced a bottleneck that reduced its effective population size to 50% of original size in mean estimate. This event dated to 10 ka BP under the assumption of fast mutation rate, or to 1 Ma BP under the assumption of low mutation rate. In the closely related western Pyrenees cluster, no such bottleneck was detectable, but its effective population size had been declining to since 35 ky BP or 35 Ma BP.

The Ecrins cluster (Fig. 5B) had experienced a severe decline of effective population size, resulting in the lowest estimated current effective population size in our dataset (i.e. 2% of Belledonne and 0.5% of Mont Blanc cluster effective population sizes). This decline started 6 ka BP or 600 ka BP. In the closely related Lagorai cluster, no such decline was detectable, but this can be partly attributed to low statistical power given that this cluster contained only seven individuals.

The effective population size of the Mont Blanc cluster (Fig. 5C) had grown to 350% of their original size between 35 and 10 ka BP or between 3.5 and 10 Ma BP. The Mont Blanc cluster was the one with the largest effective population size of the dataset, being 5.1 times larger than the Belledonne cluster and 207.52 times larger than Ecrins cluster. In the closely related Valais cluster, the effective population size had been declining since 35 ka BP or 35 Ma BP.

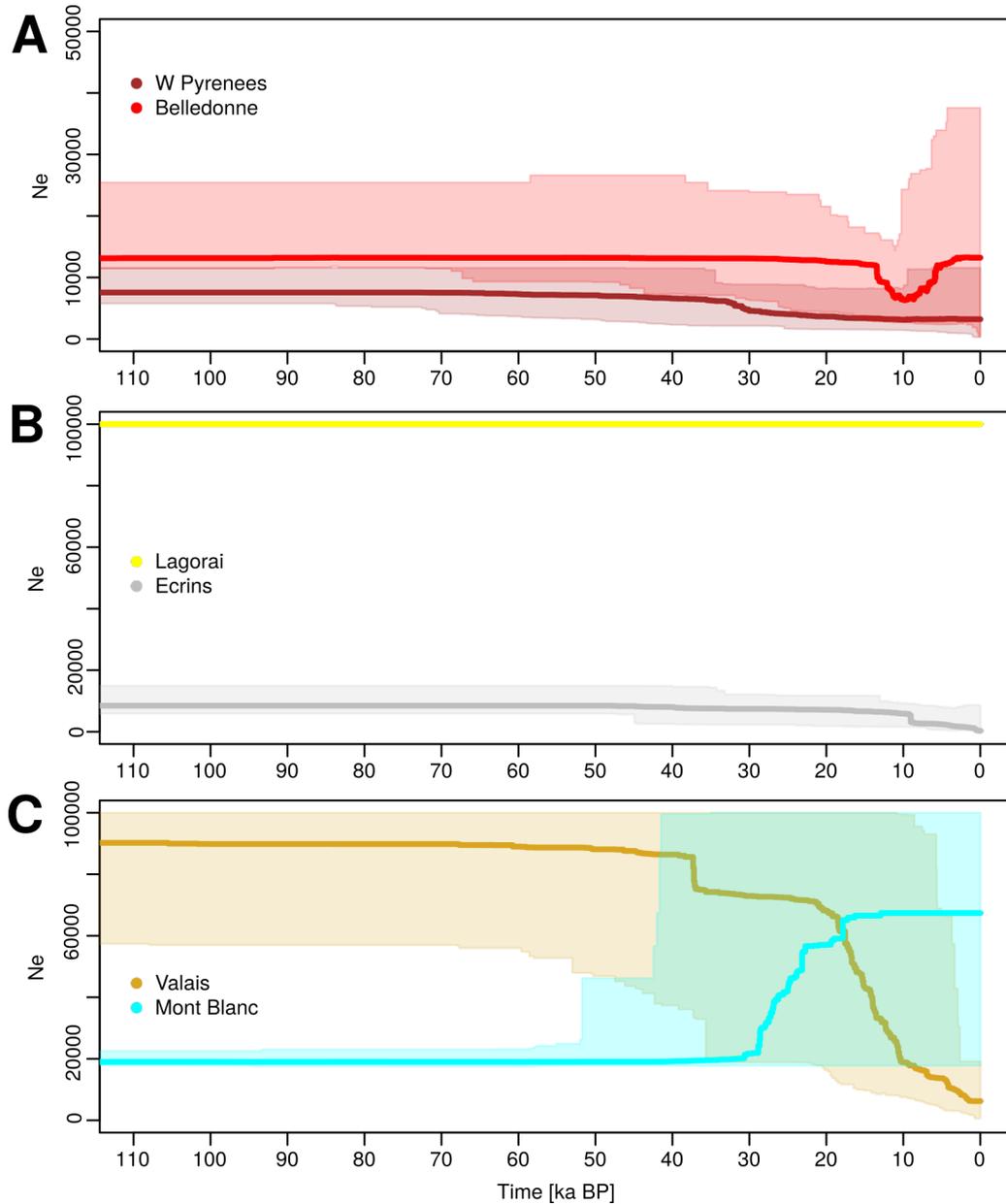


Figure 5: The past demographic profiles of focal internal Alpine clusters as reconstructed by the Stairway plot. The lines represent median estimates and bands 95% highest probability density envelopes. Note that the absolute values on both axes depend on assumed mutation rate ($5 \cdot 10^{-8}$ substitutions/site/year) and generation time (10 years), both timing of population events and effective population sizes should thus be interpreted only relatively, in comparison to the other clusters.

Chloroplast genome phylogeny

In order to get information about the timing of population divergence events, and also in order to shed more light on the origins of *A. argentea* within the whole *Androsace* sect. *Aretia*, we constructed a fossil dated phylogeny of the section *Aretia* and selected individuals

Chapter 3

of *A. argentea* based on complete sequences of chloroplast genomes (Fig. 6). The sampled individuals of *A. argentea* formed two separate clades in the chloroplast phylogeny, stemming from different parts of the section *Aretia*. The closest relative of the first clade was *A. lactea*, and the closest relatives of the other clade were *A. hirtella* and *A. wilkommi*. The stem age of the clade related to *A. lactea* was 1.785 Ma BP (0.664-3.226 95% highest probability density interval), the stem age of the clade related to *A. hirtella* and *A. wilkommi* was 2.911 Ma BP (1.109-5.055 HPD). Most of the clusters identified by snmf were confined exclusively to one of these clades, with the exception of the Valais cluster. The individuals from Belledonne and Ecrins clusters formed monophyletic clades, whereas the Mont Blanc cluster individuals were intermixed with the Valais clade. The sister clade of individuals from the Ecrins cluster was an individual from Valle Cervo, which is an admixed population related to Lagorai cluster based on nuclear dataset. The stem age of individuals from the Ecrins cluster was 240 ka BP (61-480 HPD) and their crown age was 49 ka BP (3-114 HPD). The sister clade to individuals from the Belledonne cluster was an individual from La Maladeta from eastern Pyrenees cluster. The stem age of the individuals from Belledonne cluster was 69 ka BP (11-149 HPD) and their crown age was 25 ka BP (1-65 HPD). The most recent common ancestor (mrca) of the individuals from Mont Blanc cluster dated back to 44 ka BP (4-100 HPD), the mrca of the individuals from Valais cluster belonging to *A. lactea* clade dated to 215 ka BP (69-397 HPD). Please note that the tips of chloroplast genome phylogeny corresponded to single individuals, and not to whole populations as represented in the nuclear phylogenetic network. Because of that, the chloroplast phylogeny may reflect incomplete lineage sorting on shallow scales, which should be taken in account when exploring its consistency with nuclear phylogeny.

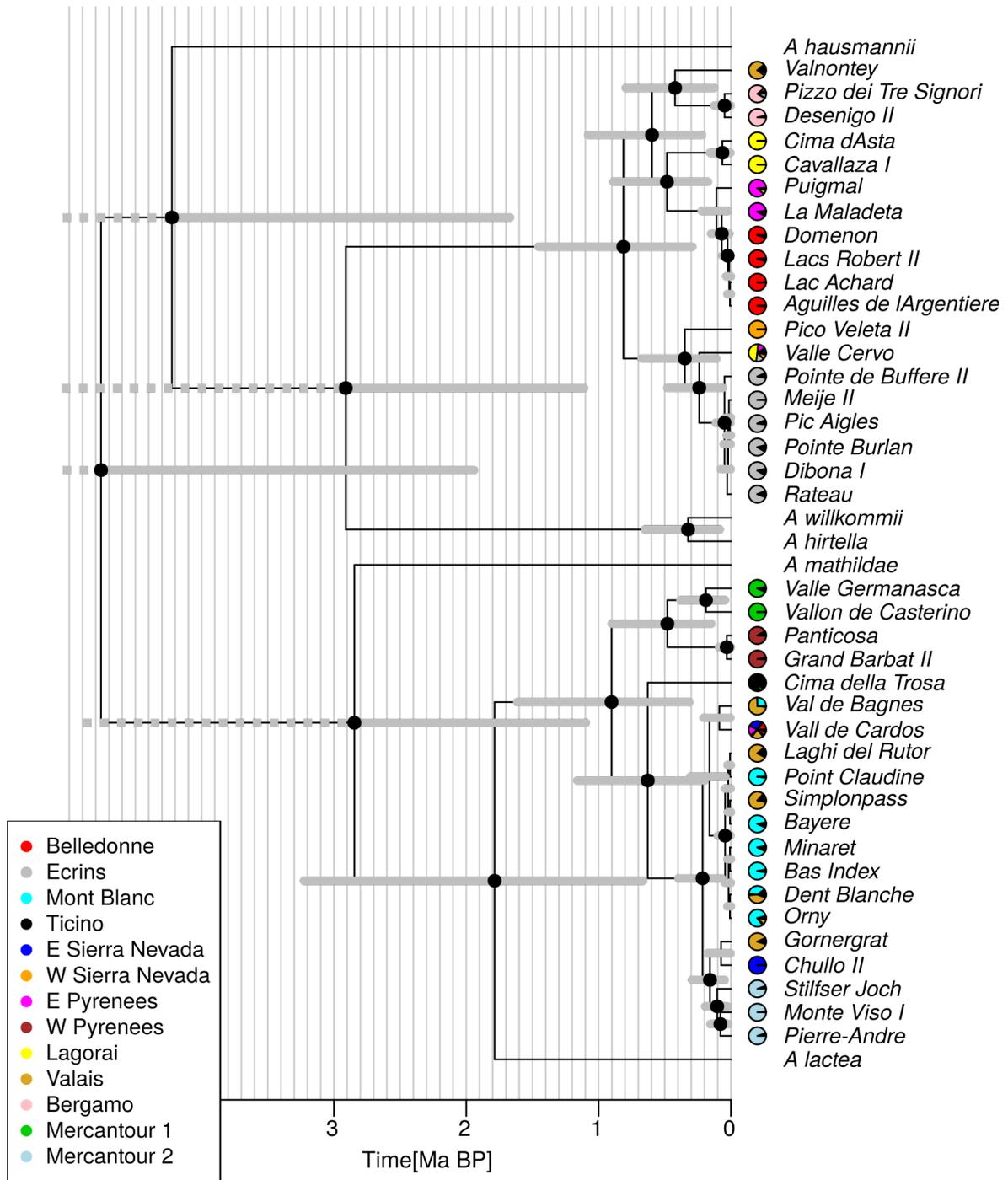


Figure 6: Dated chloroplast genome phylogeny of *Androsace* sect. *Aretia* (species distant to populations of *A. argentea* have been pruned) reconstructed by Bayesian inference with BEAST. Dots indicate nodes with posterior probability > 0.9, grey bars represent 95 % highest probability density of node dating, grid span is 100 ky and pies reflect proportional assignment to the respective snmf clusters.

Discussion

In this study, we explored the genetic structure of *Androsace argentea* with the main aim to investigate the dispersal and glacial survival dynamics of this high elevation species widely distributed in the European mountains. We paid here special attention to populations that are occurring at high elevations in internal Alps in order to reconstruct their specific history. Previous studies on Quaternary range and population dynamics in high elevation plant species mostly relied on low information genetic markers, as is AFLP, allowing only indirect evidence via spatial patterns of genetic structure, and not allowing to disentangle more complex scenarios (Escobar García et al., 2012; Schönswetter and Schneeweiss, 2019; Stehlik, 2003). The combination of ddRADseq and chloroplast genome data used in this paper allowed us to address dispersal and survival dynamics of *A. argentea* in more explicit way and with better information about timing of population events. Due to the stochastic nature of life in arctic-alpine environments, phylogeographic studies sometimes reveal complex and unexpected histories after refining sampling and the inference methods (Winkler et al., 2012), and our study is such a case. We show here that the phylogeographic history of *A. argentea* consists of a strikingly large amount of long distance seed dispersal events across the European mountain system, and also that different internal Alpine populations likely experienced rather divergent demographic trajectories and survival scenarios through the Quaternary.

The evolutionary history of *Androsace argentea*

Our analyses identified two broader regions where *A. argentea* very likely survived on the long term, i.e. over multiple glaciation cycles – one is the peripheral areas of the Alps and the other the Pyrenees. The nuclear phylogenetic network suggests that the species most likely originated in central Alps and the most basal populations are the ones grouped in Ticino cluster. The chloroplast phylogeny reveals that at least one introgression event with other species took place during the evolutionary history of *A. argentea*, resulting in two separate *A. argentea* lineages in the chloroplast phylogeny, that are sister to different species of the section *Aretia*. The chloroplast phylogeny also suggests an alternative scenario about a Pyrenean origin of the species, given that *A. wilkommii* and *A. hirtella*, which are both narrow endemics of the Pyrenees, appear as sister species of one of the two main clades of *A.*

Chapter 3

argentea. Moreover, these two species resemble morphologically to *A. argentea* more than the third possible sister *A. lactea* (Castroviejo, 2012). Exploring the geographic and phylogenetic origin of *A. argentea* with higher confidence is out of the scope of this study and would require denser sampling of outgroups for nuclear phylogenetic network inference. In any case, the dating of chloroplast phylogenies suggests that the lineages leading to *A. argentea* separated from the other species between 0.664 and 5.055 Ma BP, and the species origin thus clearly predates at least the last three glacial periods (Würm, maximum in cca 25 ka BP; Riss, maximum in cca 150 ka BP; Mindel, maximum in cca 420 ka BP). It seems also clear that *A. argentea* radically changed the bedrock preference during or after its origination, because all the three candidate sister species are strict calcareous specialists, whereas *A. argentea* only grows on silicates.

In the evolutionary history of *A. argentea*, multiple events of long distance seed dispersal occurred between the Alps, Pyrenees and Sierra Nevada, and likely also among distant parts of the Alps. According to the most likely scenario that the species originated in the Alps, our data suggest that the colonization of the Pyrenees occurred with at least two independent dispersal events, one resulting in the population cluster of western Pyrenees, and another one in the population cluster of eastern Pyrenees. The alternative scenario (Pyrenean origin of *A. argentea*) would imply that the Alps were colonized with four or more independent events. In any case, two independent colonization events lead to the colonization of Sierra Nevada, one of them very likely from central Alps over a distance of more than 1300 km. Long distance dispersal events seems to also occur among various parts of the Alps, some of them likely have taken place in recent times. For example, the population in Stilsfer Joch forms the Mercantour 1 cluster together with populations in Mont Viso and Pierre-André more than 300 km apart, which cannot be explained by ancient range fragmentation, because of multiple non-admixed population clusters lying between these location.

Interestingly, these long distance dispersal events are seldom followed by admixture with neighboring populations. Most notably, the two clusters in Sierra Nevada show no evidence for local admixture although they are isolated only by 30 km. The same is true for various neighboring clusters in the Alps, for example there is no evidence of admixture between Ecrins and Belledonne clusters lying cca 30 km apart, or between Mercantour 1 and Mercantour 2 clusters that are even overlapping in their geographic extent. The admixture was however detected between some of the most basal clusters, specifically between the

Chapter 3

Ticino and the Mercantour 1 cluster and also between the two clusters in the Pyrenees, suggesting that the occurrence of local admixture is extremely stochastic, but possibly depends on time of coexistence.

The importance of long distance seed dispersal and the distance at which acts the gene flow due to pollen exchange are often discussed in mountain plants (Escobar García et al., 2012; Kropf et al., 2006; Scheepens et al., 2012; Thiel-Egenter et al., 2011). For example Thiel-Egenter et al. (2011) assume that the pollen transport acts on longer distances than seed dispersal, and use this for explaining incoherence of genetic and taxonomic patterns of beta diversity across the Alps. However, the assumption that pollen dispersal acts at longer distances than seed dispersal remains poorly tested, and in entomogamous high elevation specialist, one can imagine that it could be the other way round. The seed dispersal could occur on relatively long distances due to intense wind transport among mountain summits similar to anemo-orographic systems (Jeník, 1997), whereas pollen transport would be limited by the discontinuity of species distribution and the low mobility of pollinators (Scheepens et al., 2012). In such system, the local geneflow would be reinforced neither by seed dispersal on short distances, because such events would require dispersal into proximity of existing population, which may rarely happen in species with sparse and discontinuous distribution. Our results indeed strongly suggest that the combination of frequent long distance seed dispersals and limited reach of local gene flow characterize well the phylogeographic structure of *A. argentea*. It would be interesting to test whether similar patterns are found in other mountain specialists with broad European distribution and complex genetic structure, for example *Androsace lactea* (Schneeweiss and Schonswetter, 2010) or *Gentiana nivalis* (Alvarez et al., 2012). The little local gene flow can be explained by inefficiency of insect pollination in high elevation habitats, and more insight could thus be brought by comparison of population genetic structures of entomo- and anemogamous species (as discussed in Ley et al., 2018). If the combination of long distance seed dispersal and little local gene flow is common also in other species and lineages, it might explain why European mountain plants diversified mostly by allopatry (Smycka in prep., chapter 2; Boucher et al., 2016). For a species spreading mostly by long distance seed dispersal with little probability of contact with neighboring populations, each long distance dispersal event constitutes a possible allopatric speciation, given that the new population survives long enough to develop morphological differences and become recognizable morphological

species (De Queiroz, 2007) or to develop intrinsic reproductive isolation and become biological species (Mayr, 1942).

The origin and glacial survival of internal Alpine populations

The studied population clusters in three focal internal Alpine areas (Mont Blanc, Belledonne, Ecrins) likely have different histories of origin and glacial survival. It has been shown that different mountain plant species may have survived glacial periods with contrasting mechanisms, for example *Eritrichium nanum* was likely surviving in nunatak refugia in the Alps, while *Erinus alpinus* and *Rumex nivalis* was surviving in peripheral refugia (Stehlik, 2003). It has also been shown that current high elevation populations may result from admixture of lineages with different survival history (Escobar García et al., 2012; Schönswetter and Schneeweiss, 2019). However to our knowledge, we bring here the first evidence that different high elevation populations within one species were following distinct strategies of glacial survival.

The populations of Mont Blanc cluster forms a monophyletic clade nested within the geographically neighboring Valais cluster. The demographic profiles suggest that the Valais populations formed a large cluster in the past that started to decline at certain time, and at the same time the Mont Blanc cluster massively expanded. The absolute timing of this turnaround cannot be estimated from the timing of demographic profiles, but the dated chloroplast phylogeny suggests that the expansion of Mont Blanc cluster started between 4 and 100 ka BP (95% highest probability density interval of crown age), clearly postdating the Riss glacial maximum (second most recent glaciation, cca 150 ka BP), and possibly also postdating the last glacial maximum (cca 25 ka BP; Seguinot et al., 2018). The demographic stability of the Mont Blanc cluster following the growth suggests that the expansion might indeed have taken place after the last glacial maximum – the Mont Blanc cluster consists exclusively of high elevation populations far from glacier margins and such populations could hardly survive the glacial cycle without detectable demographic change towards the present where the Stairway plot method is the most sensitive (Liu and Fu, 2015). Without more direct evidence it is however not possible to conclude whether the colonization of Mont Blanc area took place after the last glacial maximum or before and whether these populations survived last glaciation *in situ*. Interestingly, the source cluster of Valais, that is very likely

Chapter 3

older (crown age 69 to 397 BP) than the last glacial maximum and possibly also older than the Riss glacial maximum, and had large effective population sizes in the past, is also lying out of the peripheral refugia according to both of our glaciation models for the Alps. It suggests that if not the Mont Blanc cluster, the Valais cluster survived one or more glacial periods in the internal Alpine nunatak refugia. This interpretation is to be confirmed by comparing population genetic patterns of Valais and Mont Blanc populations with reconstructions of past glaciation on finer scales (see Future directions), but the previous studies suggest that the Valais region might indeed contain nunatak refugia, given that high genetic diversities of populations have been detected for multiple species (Schönswetter et al., 2005; Stehlik, 2003), and moreover this areas presents high endemic diversity (Smyčka et al., 2018).

The cluster of populations in Belledonne is most closely related to the western Pyrenees cluster. This could mean that either the western Pyrenees cluster gave rise to the Belledonne cluster via long distance dispersal, or the other way around. The first option is more likely, because the long distance seed dispersal means temporary reduction of effective population size to single individual, and should thus create a population-genetic signature of severe bottleneck. Such bottleneck is indeed visible in the past demographic profile of Belledonne cluster, suggesting that this cluster originated via long distance seed dispersal from Pyrenees. The absolute timing of this event cannot be estimated from the demographic profiles, but from the dated chloroplast phylogeny it can be assumed that the Belledonne cluster splitted from the western Pyrennean one between 11 and 149 ka BP and started to expand in Belledonne between 1 and 65 ka BP. The colonization of Belledonne from Pyrenees thus took place either after Riss glacial maximum (second most recent, cca 150 ka BP) or even after last glacial maximum (cca 25 ka BP). According to both glaciation models, the Belledonne region indeed remained partly non-glaciated during last glaciation, possibly allowing survival of *A. argentea* populations, but the Pyrenees definitely served as southern refugium of populations in Belledone region at least in the more severe Riss glaciation. The evidence for survival in southern refugia is often assumed for thermophylous lowland species (Stewart et al., 2010; Tzedakis et al., 2013) or for mid elevation mountain plants (Zhang et al., 2001), but was never seriously considered for high elevation plants. Our results suggest that this mechanism should be taken in account, especially for species with high dispersal capacity as is *A. argentea*.

Chapter 3

The cluster of populations in Ecrins is likely related to the Lagorai cluster. The Lagorai cluster is localized in eastern Alps, occupying a possible peripheral refugium as identified by Schönswetter et al. (2005). Although there is one population in central Alps related to the Lagorai cluster (Valle Cervo), the number of old and non-admixed populations located between Lagorai or Valle Cervo and Ecrins suggests that the colonization of Ecrins took place via long distance dispersal, rather than via gradual colonization with subsequent range fragmentation. Unlike the common idea that the peripheral refugia were lying close to the areas to be colonized and allowed gradual colonization process (Escobar García et al., 2012; Schönswetter et al., 2005; Thiel-Egenter et al., 2011), this is likely not the case of *A. argentea* in Ecrins region. The demographic profile of Ecrins cluster does not show a bottleneck signature that could be associated with long distance dispersal, but rather a continuous decline. The most likely explanation for this is that the signature of recovery from the bottleneck event is not captured by the demographic profile, perhaps due to very small current species abundances in Ecrins region (Lavergne and Dentant, pers. obs.). The timing of colonization of Ecrins seems to be similar or slightly younger than Belledonne in the demographic profiles. The dates from chloroplast phylogeny suggest that the Ecrins cluster separated between 61 and 480 ka BP, but the sampled populations started to split between 3 and 114 ka BP. We thus cannot exclude that *A. argentea* persisted multiple glaciations in Ecrins region, but the current species distribution in the massif was shaped definitely after Riss glacial maximum (150 ka BP) and maybe even after the last glacial maximum (cca 25 ka BP).

The high elevation plants in the internal Alps had three possibilities to survive the Pleistocene climatic oscillations: in southern mountain ranges, in peripheral refugia or *in situ* on nunatak refugia. It is often assumed that the first two possibilities can take place in species with good dispersal capacities, while the third one would imply a severe reductions of ranges and population sizes (and thus a drastic genetic diversity reduction) on the microclimatic refugia (Lohse et al., 2011; Schneeweiss and Schönswetter, 2011; Tribsch and Schönswetter, 2003; Westergaard et al., 2011). The large portion of studies that support nunatak survival in the Alps indeed suggest that during the interglacial periods, there was an intensive admixture between nunatak and peripheral populations (Escobar García et al., 2012; Lohse et al., 2011; Schönswetter and Schneeweiss, 2019; Stehlik, 2003), possibly constituting a rescue effect for nunatak populations both in terms of genetic diversities and population sizes. In our study, we

Chapter 3

did not find such interglacial admixture, and moreover our demographic analyses suggest that the populations surviving on long term on nunatak refugia (Valais and Mont Blanc clusters) managed to persist in population sizes comparable to peripheral refugia. On the other hand, our results suggest that it is rather the populations that recolonized internal Alpine areas from the periphery or southern refugia, which recently experienced small population sizes and reduced genetic diversity, because these colonizations took place via long distance seed dispersal. A key for understanding the differences in glacial survival patterns and mechanisms among different high elevation species might be their habitat requirements at small scales. Species for which the nunatak survival in the Alps was explored previously, like *Senecio carniolicus* (Escobar García et al., 2012), *Androsace alpina* (Schönswetter and Schneeweiss, 2019) or *Eritrichium nanum* (Stehlik, 2003) often grow on scree fields or looser rock formations, whereas *A. argentea* is a narrow specialist of compact vertical cliffs. Although all these species are high elevation plants specialized to siliceous bedrocks, the habitat requirements of *A. argentea*, and thus also the requirements on refugia type and size, may thus differ from the others.

Future directions

Although the presented manuscript provides insights into population genetic structure, migration history and survival of different populations of *A. argentea*, there are two additional analyses that could be performed with the gathered data and that would bring firmer evidence for the discussed processes. These analyses, which are very time-demanding, have not been performed yet due to the unexpected delay in the production of ddRADseq data (the final dataset arrived in February 2019), and we plan to conduct them in the near future. While these two analyses are not central for our interpretation of the data and the main presented conclusions, they would allow to explore the past population processes of *A. argentea* in greater detail, as we discuss below.

Genetic structure of internal Alpine populations with respect to glaciation and past climate

In order to further explore the mechanism of *in situ* survival of the Mont Blanc and Valais clusters (and also Belledonne and Ecrins clusters if they survived Würm glaciation *in situ*), it would be interesting to compare their internal genetic structures with last maximum glaciation models coupled with hindcasts of species climatic niche. This can be done by comparing patterns of genetic alpha or beta diversity with past glaciation and climatically suitable spots, or by process based coalescence models (dadi, Gutenkunst et al., 2009); ABC, Csilléry et al., 2010). These approaches would shed more light on position, size and morphology of nunataks that were suitable for survival, and would also allow us to explore the pathway of post-glacial recolonization. The reason why we did not take this direction so far is that there are multiple models of past glaciation of the Alps, that are often contradictory and operate on different scales. Before adopting this direction, it is thus necessary to review different sources of glaciation models, and in collaboration with contacted local glaciologists choose the most relevant one for further analysis.

At the moment there are two relevant glaciation estimates for the whole Alps, presented above in this study. One of them is the informal estimate based on classical glaciological literature used in Schönswetter et al. (2005) , another one is an estimate of process based model of past glaciation from Seguinot et al. (2018). Unfortunately, both these sources

Chapter 3

predict dramatically different extent of glaciers in the last glacial maximum, and the one from Schönswetter et al. (2005) seems to better match the population clusters that we identified as refugial in our genetic analyses (Fig. 3). None of these estimates is however suitable for exploring nunatak refugia. Schönswetter et al. (2005) explicitly focuses only on large peripheral refugia. Seguinot et al. (2018) should theoretically be able to predict extent of nunataks, but it identifies almost no such areas, either due to the fact that this model is systematically overestimating volumes of ice, or because it operates on scale 1 km² and the nunataks were typically smaller than that.

In addition to these glaciation estimates for the whole Alps, there are also ones that were made only for certain regions. One of such estimates is performed by Delunel (2010), which is based on process based modeling, and has been calibrated with local occurrences of glacier trim lines and cosmogenic nuclids dating. This model predicts a much lower last glacial maximum ice extent than Seguinot et al. (2018) and predicts large nunatak areas in the internal Alps (Fig. 7). Given that this model is intended to work on lower scale and works with much more detailed information than the estimates for the whole Alps, it seems to be more suitable for exploring the nunatak survival. We plan to compare the internal genetic structures in Belledonne and Ecrins clusters with Delunel (2010) glaciation estimate in near future, to test whether they survived Würm glaciation *in situ* or not. Even more interesting would be performing the same also in Valais and Mont Blanc clusters, where the nunatak survival had a long term character according to the population genetic analyses. The reason why these analyses are not presented in the current state of the manuscript is that a deeper review of local paleoglaciological literature is needed for the identification of a suitable glaciation model, possibly in collaboration with experts on glaciology.

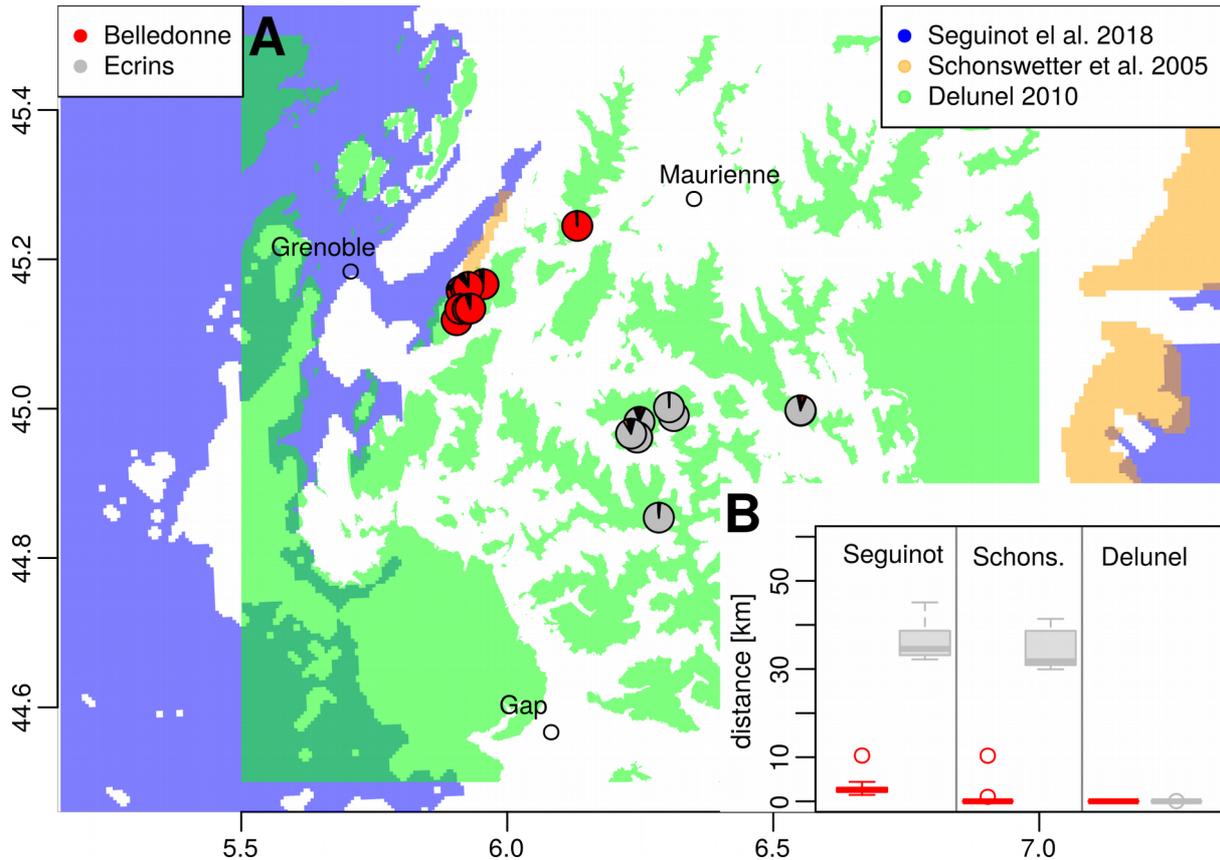


Figure 7: A comparison of glaciation estimates in the Ecrins and Belledonne regions based on Schönswetter et al. (2005), Seguinot et al. (2018) and Delunel (2010). The pies indicate positions of populations and reflect proportional assignment of populations to the respective snmf clusters.

Integrating mutation rates of nuclear SNP dataset and fossil dating into one analysis

The current uncertainty in dating of survival and colonization events is partly due to the quality of primary input information – Primulaceae fossils with broad date extent and relatively vague information about mutation rates in ddRADseq datasets. However, a partial improvement may be attained by integrating these two sources within one formal evolutionary model. This can be done by identifying with confidence multiple nodes that are equivalent in plastome and ddRADseq tree (e.g. mrcas of clusters forming monophyletic clades both on plastome and ddRADseq tree), and using the height posterior from plastome tree as a secondary date for the ddRADseq topology to be built by algorithm using information about generation times and mutation rates. An optimal way to do this would be the SNAPP procedure in BEAST. Results of this analysis are not presented in the current state

Chapter 3

of the manuscript, due to extremely long running time of the analysis that is expected to take up to 6 months.

Acknowledgements

We would like to acknowledge Nadir Alvarez and Filip Kolář for valuable advices, Océane Bartholomé, Camille Martinez, Jan Bayer, Louise Boulangeat and Laure Barrabé for help with fieldworks, Filippo Prosser for providing species occurrence data from eastern Alps.

The research was funded by ANR project Origin-Alps (ANR-16-CE93-0004), France Genomique project PhyloAlps, Labex OSUG@2020 project Nunataks and Cousins project of Pôle Biodiversité du Conseil Général de l'Isère. Jan Smyčka was supported by ANR project Sphinx (ANR-16-CE02-0011) and by Doctoral school of Chemistry and Life Sciences of Université Grenoble Alps.

References

- Alvarez, N., Manel, S., Schmitt, T., 2012. Contrasting diffusion of Quaternary gene pools across Europe: The case of the arctic-alpine *Gentiana nivalis* L. (Gentianaceae). *Flora Morphol. Distrib. Funct. Ecol. Plants* 207, 408–413. doi:10.1016/j.flora.2012.03.006
- Bagley, R.K., Sousa, V.C., Niemiller, M.L., Linnen, C.R., 2017. History, geography and host use shape genomewide patterns of genetic variation in the redheaded pine sawfly (*Neodiprion lecontei*). *Mol. Ecol.* 26, 1022–1044. doi:10.1111/mec.13972
- Bivand, R., Keitt, T., Rowlingson, B., 2019. rgdal: Bindings for the “Geospatial” Data Abstraction Library [WWW Document]. URL <https://cran.r-project.org/package=rgdal> (accessed 5.12.19).
- Bliss, L.C., 1971. Arctic and Alpine Plant Life Cycles. *Annu. Rev. Ecol. Syst.* 2, 405–438. doi:10.1146/annurev.es.02.110171.002201
- Boucher, F.C., Zimmermann, N.E., Conti, E., 2016. Allopatric speciation with little niche divergence is common among alpine Primulaceae. *J. Biogeogr.* 43, 591–602. doi:10.1111/jbi.12652
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Comput. Biol.* 10, 1–6. doi:10.1371/journal.pcbi.1003537
- Brockmann-Jerosch, M., Brockmann-Jerosch, MC, 1926. Die Geschichte der Schweizerischen Alpenflora., in: Schröter, C. (Ed.), *Das Pflanzenleben Der Alpen*. Raustein, Zürich, pp. 1110–1215.
- Carcaillet, C., Blarquez, O., 2017. Fire ecology of a tree glacial refugium on a nunatak with a view on Alpine glaciers. *New Phytol.* 216, 1281–1290. doi:10.1111/nph.14721
- Carcaillet, C., Latil, J.L., Abou, S., Ali, A., Ghaleb, B., Magnin, F., Roiron, P., Aubert, S., 2018. Keep your feet warm? A cryptic refugium of trees linked to a geothermal spring in an ocean of glaciers. *Glob. Chang. Biol.* 24, 2476–2487. doi:10.1111/gcb.14067
- Castroviejo, S. (Ed.), 2012. *Flora Iberica*. Real Jardín Botánico CSIC.
- Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., Cresko, W.A., 2013. Stacks: an analysis tool set for population genomics. *Mol. Ecol.* 22, 3124–40. doi:10.1111/mec.12354
- Coissac, E., 2016. The ORGanelle ASSEMBLER 1.0.3 [WWW Document]. URL <https://git.metabarcoding.org/org-asm/org-asm/wikis/home> (accessed 9.10.18).
- Csilléry, K., Blum, M.G.B., Gaggiotti, O.E., François, O., 2010. Approximate Bayesian Computation (ABC) in practice. *Trends Ecol. Evol.* 25, 410–418. doi:10.1016/j.tree.2010.04.001
- Cumer, T., Pouchon, C., Boyer, F., Yannic, G., Rioux, D., Bonin, A., Capblancq, T., 2018. Double-digest RAD-sequencing: do wet and dry protocol parameters impact biological results? *bioRxiv* 33. doi:10.1101/445122
- Darwin, F., Seward, A. (Eds.), 1903. *More letters of Charles Darwin*. John Murray, London.
- de Candolle, A., 1875. Sur les causes de l'inégale distribution des plantes rares dans la chaîne des Alpes. *Atti del Congr. Internazionale Bot. Tenuto Firenze*. 92–104.
- De Queiroz, K., 2007. Species concepts and species delimitation. *Syst. Biol.* 56, 879–86. doi:10.1080/10635150701701083
- Delmas, M., 2015. The last maximum ice extent and subsequent deglaciation of the Pyrenees: an overview of recent research. *Cuad. Investig. Geográfica* 41, 359. doi:10.18172/cig.2708
- Delunel, R., 2010. Evolution géomorphologique du massif des Ecrins-Pelvoux depuis le Dernier Maximum Glaciaire – Apports des nucléides cosmogéniques produits in-situ. Université Joseph Fourier Grenoble.
- Dentant, C., 2018. The highest vascular plants on Earth. *Alp. Bot.* 128, 97–106. doi:10.1007/s00035-018-0208-3
- Dentant, C., Lavergne, S., 2013. Plantes de haute montagne: état des lieux, évolution et analyse diachronique dans le massif des Écrins (France). *Bull. Soc. linn. Provence* 64, 83–98.
- Dentant, C., Lavergne, S., Malécot, V., 2018. Taxonomic revision of West-Alpine cushion plant species belonging to *Androsace* subsect. *Aretia*. *Bot. Lett.* 165, 337–351. doi:10.1080/23818107.2018.1450784
- Dullinger, S., Willner, W., Plutzer, C., Englisch, T., Schratt-Ehrendorfer, L., Moser, D., Ertl, S., Essl, F., Niklfeld, H., 2012. Post-glacial migration lag restricts range filling of plants in the European Alps. *Glob. Ecol. Biogeogr.* 21, 829–840. doi:10.1111/j.1466-8238.2011.00732.x

Chapter 3

- Escobar García, P., Winkler, M., Flatscher, R., Sonnleitner, M., Krejčíková, J., Suda, J., Hülber, K., Schneeweiss, G.M., Schönswetter, P., 2012. Extensive range persistence in peripheral and interior refugia characterizes Pleistocene range dynamics in a widespread Alpine plant species (*Senecio carniolicus*, Asteraceae). *Mol. Ecol.* 21, 1255–1270. doi:10.1111/j.1365-294X.2012.05456.x
- Fitak, R., n.d. optM: an R package to optimize the number of migration edges using threshold models. *J. Hered.*
- Frichot, E., Francois, O., 2014. LEA: an R package for Landscape and Ecological Associations studies [WWW Document]. URL <http://membres-timc.imag.fr/Olivier.Francois/LEA/index.htm> (accessed 4.10.19).
- Graffelman, J., 2015. Exploring Diallelic Genetic Markers: The HardyWeinberg Package. *J. Stat. Softw.* 64, 1–16. doi:10.18637/jss.v064.i03
- Gruber, B., Georges, A., Unmack, P., Clark, L., Berry, O., 2019. dartR 1.1.11 [WWW Document]. URL <https://cran.r-project.org/package=dartR> (accessed 5.10.19).
- Gutenkunst, R.N., Hernandez, R.D., Williamson, S.H., Bustamante, C.D., 2009. Inferring the Joint Demographic History of Multiple Populations from Multidimensional SNP Frequency Data. *PLoS Genet.* 5, e1000695. doi:10.1371/journal.pgen.1000695
- Hansen, J., Sato, M., Ruedy, R., Lo, K., Lea, D.W., Medina-Elizade, M., 2006. Global temperature change. *Proc. Natl. Acad. Sci.* 103, 14288–14293. doi:10.1073/pnas.0606291103
- Hijmans, R., 2019. raster: Geographic Data Analysis and Modeling [WWW Document]. URL <https://cran.r-project.org/package=raster> (accessed 5.12.19).
- Jeník, J., 1997. Anemo-orographic systems in the Hercynian Mts. and their effects on biodiversity. *Acta Univ. Wratislaviensis* 4, 9–21.
- Jombart, T., Ahmed, I., 2011. adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* 27, 3070–3071. doi:10.1093/bioinformatics/btr521
- Kłapyta, P., Zasadni, J., 2018. Research history on the Tatra mountains glaciations. *Stud. Geomorphol. Carpatho-Balcanica* 43–85.
- Knaus, B.J., Grünwald, N.J., 2017. vcfr: a package to manipulate and visualize variant call format data in R. *Mol. Ecol. Resour.* 17, 44–53. doi:10.1111/1755-0998.12549
- Körner, C., 2011. Coldest places on earth with angiosperm plant life. *Alp. Bot.* 121, 11–22. doi:10.1007/s00035-011-0089-1
- Körner, C., 1999. *Alpine Plant Life*. Springer, Berlin.
- Kropf, M., Comes, H.P., Kadereit, J.W., 2006. Long-distance dispersal vs vicariance: The origin and genetic diversity of alpine plants in the Spanish Sierra Nevada. *New Phytol.* 172, 169–184. doi:10.1111/j.1469-8137.2006.01795.x
- La Sorte, F.A., Jetz, W., 2010. Projected range contractions of montane biodiversity under global warming. *Proc. R. Soc. B Biol. Sci.* 277, 3401–3410. doi:10.1098/rspb.2010.0612
- Ley, A.C., Nissen, J., Wölk, A., Röser, M., 2018. Glacial refugia and speciation in a group of wind-pollinated and -dispersed, endemic Alpine species of *Helictotrichon* (Poaceae). *PLoS One* 13, 1–14. doi:10.1371/journal.pone.0205354
- Liu, X., Fu, Y.X., 2015. Exploring population size changes using SNP frequency spectra. *Nat. Genet.* 47, 555–559. doi:10.1038/ng.3254
- Lohse, K., Nicholls, J.A., Stone, G.N., 2011. Inferring the colonization of a mountain range-refugia vs. nunatak survival in high alpine ground beetles. *Mol. Ecol.* 20, 394–408. doi:10.1111/j.1365-294X.2010.04929.x
- Marx, H.E., Dentant, C., Renaud, J., Delunel, R., Tank, D.C., Lavergne, S., 2017. Riders in the sky (islands): Using a mega-phylogenetic approach to understand plant species distribution and coexistence at the altitudinal limits of angiosperm plant life. *J. Biogeogr.* 44, 2618–2630. doi:10.1111/jbi.13073
- Mayr, E., 1942. *Systematics and the origin of species*. Columbia University Press, New York.
- Nývlt, D., Engel, Z., Tyráček, J., 2011. Pleistocene glaciations of Czechia, in: Ehlers, J., Gibbard, P., Hughes, P. (Eds.), *Quaternary Glaciations - Extent and Chronology*. Elsevier, Amsterdam, pp. 37–46. doi:10.1016/B978-0-444-53447-7.00004-0
- Paradis, E., Claude, J., Strimmer, K., 2004. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* 20, 289–290. doi:10.1093/bioinformatics/btg412

Chapter 3

- Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E., 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One* 7, e37135. doi:10.1371/journal.pone.0037135
- Pickrell, J.K., Pritchard, J.K., 2012. Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. *PLoS Genet.* 8, e1002967. doi:10.1371/journal.pgen.1002967
- R Core Team, X., 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing [WWW Document]. URL <https://www.r-project.org/> (accessed 5.10.19).
- Revell, L.J., 2012. phytools: An R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* 3, 217–223. doi:10.1111/j.2041-210X.2011.00169.x
- Ronikier, M., 2011. Biogeography of high-mountain plants in the Carpathians: An emerging phylogeographical perspective. *Taxon* 60, 373–389. doi:10.1002/tax.602008
- Roquet, C., Boucher, F.C., Thuiller, W., Lavergne, S., 2013. Replicated radiations of the alpine genus *Androsace* (Primulaceae) driven by range expansion and convergent key innovations. *J. Biogeogr.* 40, 1874–1886. doi:10.1111/jbi.12135
- Sanz, M., Schönswetter, P., Vallès, J., Schneeweiss, G.M., Vilatersana, R., 2014. Southern isolation and northern long-distance dispersal shaped the phylogeography of the widespread, but highly disjunct, European high mountain plant *Artemisia eriantha* (Asteraceae). *Bot. J. Linn. Soc.* 174, 214–226. doi:10.1111/boj.12132
- Scheepens, J.F., Frei, E.S., Armbruster, G.F.J., Stöcklin, J., 2012. Pollen dispersal and gene flow within and into a population of the alpine monocarpic plant *Campanula thyrsoidea*. *Ann. Bot.* 110, 1479–1488. doi:10.1093/aob/mcs131
- Schneeweiss, G.M., Schönswetter, P., 2010. The wide but disjunct range of the European mountain plant *Androsace lactea* L. (Primulaceae) reflects Late Pleistocene range fragmentation and post-glacial distributional stasis. *J. Biogeogr.* 37, 2016–2025. doi:10.1111/j.1365-2699.2010.02350.x
- Schneeweiss, G.M., Schönswetter, P., 2011. A re-appraisal of nunatak survival in arctic-alpine phylogeography. *Mol. Ecol.* 20, 190–192. doi:10.1111/j.1365-294X.2010.04927.x
- Schönswetter, P., Schneeweiss, G.M., 2019. Is the incidence of survival in interior Pleistocene refugia (nunataks) underestimated? Phylogeography of the high mountain plant *Androsace alpina* (Primulaceae) in the European Alps revisited. *Ecol. Evol.* 9, 4078–4086. doi:10.1002/ece3.5037
- Schönswetter, P., Stehlik, I., Holderegger, R., Tribsch, A., 2005. Molecular evidence for glacial refugia of mountain plants in the European Alps. *Mol. Ecol.* 14, 3547–3555. doi:10.1111/j.1365-294X.2005.02683.x
- Schönswetter, P., Tribsch, A., Schneeweiss, G.M., Niklfeld, H., 2003. Disjunctions in relict alpine plants: Phylogeography of *Androsace brevis* and *A. wulfeniana* (Primulaceae). *Bot. J. Linn. Soc.* 141, 437–446. doi:10.1046/j.0024-4074.2002.00134.x
- Seguinot, J., Ivy-Ochs, S., Jouvét, G., Huss, M., Funk, M., Preusser, F., 2018. Modelling last glacial cycle ice dynamics in the Alps. *Cryosphere* 12, 3265–3285. doi:10.5194/tc-12-3265-2018
- Smyčka, J., Roquet, C., Renaud, J., Thuiller, W., Zimmermann, N.E., Lavergne, S., 2018. Reprint of: Disentangling drivers of plant endemism and diversification in the European Alps - a phylogenetic and spatially explicit approach. *Perspect. Plant Ecol. Evol. Syst.* 30, 31–40. doi:10.1016/j.ppees.2017.08.003
- Stehlik, I., 2003. Resistance or emigration? Response of alpine plants to the ice ages. *Taxon* 52, 499–510. doi:10.2307/3647448
- Steinbauer, M.J., Grytnes, J.-A., Jurasinski, G., Kulonen, A., Lenoir, J., Pauli, H., Rixen, C., Winkler, M., Bardy-Durchhalter, M., Barni, E., Björkman, A.D., Breiner, F.T., Burg, S., Czortek, P., Dawes, M.A., Delimat, A., Dullinger, S., Erschbamer, B., Felde, V.A., Fernández-Arberas, O., Fossheim, K.F., Gómez-García, D., Georges, D., Grindrud, E.T., Haider, S., Haugum, S. V., Henriksen, H., Herreros, M.J., Jaroszewicz, B., Jaroszynska, F., Kanka, R., Kapfer, J., Klanderud, K., Kühn, I., Lamprecht, A., Matteodo, M., di Cella, U.M., Normand, S., Odland, A., Olsen, S.L., Palacio, S., Petey, M., Piscová, V., Sedlakova, B., Steinbauer, K., Stöckli, V., Svenning, J.-C., Teppa, G., Theurillat, J.-P., Vittoz, P., Woodin, S.J., Zimmermann, N.E., Wipf, S., 2018. Accelerated increase in plant species richness on mountain summits is linked to warming. *Nature* 556, 231–234. doi:10.1038/s41586-018-0005-6
- Stewart, J.R., Lister, A.M., Barnes, I., Dalén, L., 2010. Refugia revisited: individualistic responses of species in space and time. *Proc. Biol. Sci.* 277, 661–71. doi:10.1098/rspb.2009.1272

Chapter 3

- Theodoridis, S., Randin, C., Szovenyi, P., Boucher, F.C., Patsiou, T.S., Conti, A.E., 2017. How do cold-adapted plants respond to climatic cycles? Interglacial expansion explains current distribution and genomic diversity in *Primula farinosa* L. *Syst. Biol.* 66, 715–736. doi:10.1093/sysbio/syw114
- Thiel-Egenter, C., Alvarez, N., Holderegger, R., Tribsch, A., Englisch, T., Wohlgemuth, T., Colli, L., Gaudeul, M., Gielly, L., Jogan, N., Linder, H.P., Negrini, R., Niklfeld, H., Pellicchia, M., Rioux, D., Schönswetter, P., Taberlet, P., van Loo, M., Winkler, M., Gugerli, F., 2011. Break zones in the distributions of alleles and species in alpine plants. *J. Biogeogr.* 38, 772–782. doi:10.1111/j.1365-2699.2010.02441.x
- Thuiller, W., Guéguen, M., Georges, D., Bonet, R., Chalmandrier, L., Garraud, L., Renaud, J., Roquet, C., Van Es, J., Zimmermann, N.E., Lavergne, S., 2014. Are different facets of plant diversity well protected against climate and land cover changes? A test study in the French Alps. *Ecography (Cop.)*. 37, 1254–1266. doi:10.1111/ecog.00670
- Tribsch, A., Schönswetter, P., 2003. refugia for mountain plants: patterns of endemism and comparative phylogeography confirm palaeo-environmental evidence in the Eastern European Alps. *Taxon* 52, 477–497. doi:10.2307/3647447
- Tzedakis, P.C., Emerson, B.C., Hewitt, G.M., 2013. Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends Ecol. Evol.* 28, 696–704. doi:10.1016/j.tree.2013.09.001
- von Humboldt, A., 1805. *Essai sur la Géographie des Plantes*. Levrault, Schoell et compagnie., Paris.
- Westergaard, K.B., Alsos, I.G., Popp, M., Engelskjø, T., Flatberg, K.I., Brochmann, C., 2011. Glacial survival may matter after all: Nunatak signatures in the rare European populations of two west-arctic species. *Mol. Ecol.* 20, 376–393. doi:10.1111/j.1365-294X.2010.04928.x
- Winkler, M., Tribsch, A., Schneeweiss, G.M., Brodbeck, S., Gugerli, F., Holderegger, R., Abbott, R.J., Schönswetter, P., 2012. Tales of the unexpected: Phylogeography of the arctic-alpine model plant *Saxifraga oppositifolia* (Saxifragaceae) revisited. *Mol. Ecol.* 21, 4618–4630. doi:10.1111/j.1365-294X.2012.05705.x
- Zachos, J.C., Pagani, M., Sloan, L., Thomas, E., Billups, K., 2001. Trends, Rhythms, and Aberrations in Global Climate 65 Ma to Present. *Science (80-)*. 292, 686–693. doi:10.1126/science.1059412
- Zhang, L.B., Comes, H.P., Kadereit, J.W., 2001. Phylogeny and quaternary history of the European montane/alpine endemic *Soldanella* (Primulaceae) based on ITS and AFLP variation. *Am. J. Bot.* 88, 2331–2345. doi:10.2307/3558393

GENERAL DISCUSSION

The general aim of my thesis has been to explore the main drivers and temporal dynamics of diversification (in a broad sense, i.e. speciation and extinction) for the flora of European mountains. Diversification drivers and dynamics can be studied at various scales: one may aim to investigate diversification processes explaining the evolutionary assembly of whole biotas or floras (Merckx et al., 2015; Molina-Venegas et al., 2015; Xing and Ree, 2017), explore drivers of speciation in specific clades (Pouchon et al., 2018; Roquet et al., 2013; Schneeweiss et al., 2013), or study the mechanisms of survival in populations of one particular species as a study-case (Escobar García et al., 2012; Schönswetter and Schneeweiss, 2019; Stehlik, 2003). Our results suggest that the stochastic nature of diversification in the European mountains does not allow to extrapolate findings across these scales without difficulties. As shown in chapter two, drivers and dynamics of diversification differ across plant lineages, suggesting that the diversification processes unraveled for a specific European mountain plant lineage cannot be directly extrapolated to in the European mountain flora as a whole. Similarly, as shown in chapter three, the mechanisms of glacial survival in one high elevation population is far from providing exhaustive information about the mechanisms involved in the survival of the whole species. When talking about the drivers and dynamics of diversification, it is thus important to explicitly specify to which phylogenetic and geographic scale we do refer to and also purpose they will serve.

When discussing diversification in the European mountain system, it is important to bear in mind that the majority of plants in the European mountains are not the high elevation specialists (i.e. species occurring only in the alpine zone or above), but rather mid or low elevation species (Aeschimann et al., 2011; Mráz et al., 2016). Referring to all the plant species confined in the mountain areas of Europe, such statement is trivial and fairly consistent with elevational richness gradients observed for various parts of biota across the mountains of the world (Lomolino, 2001; Quintero and Jetz, 2018; Rahbek, 1995). In the European mountain system, this statement is however true even if we use much narrower definitions of what is a mountain species. We find more species in mid and low elevations if we consider only species endemic for these mountain systems (Aeschimann et al., 2011; Mráz et al., 2016). Strikingly, mid elevation species and elevation generalist are predominant

General discussion

even if we delimit the mountain species specifically to fast diversifying lineages that gave rise to the European high elevation flora, as we did in chapter two. It is thus important to clearly distinguish between characteristics of diversification in European mountain flora as a whole, linked mostly to species and lineages with capacity to grow in mid or low elevations, and characteristics of diversification in high elevation specialists.

Past climatic dynamics and its influence on survival and diversification

The majority of mountain species in Europe likely survived the Pleistocene climatic oscillations in mid elevational habitats at the periphery of mountain ranges or even in lowlands, rather than *in situ* in the high elevation areas. This is supported by our findings from chapter one, where we show that in the Alps, the range restricted species are more concentrated in peripheral refugia and in lower elevation areas. Similar patterns of endemism have been found in other European mountain ranges (Molina-Venegas et al., 2015; Mráz et al., 2016). The survival in relatively extensive mid or low elevation refugia is the likely reason why we did not observe a strong and consistent Pleistocene diversification slowdown in the six lineages selected in chapter two. In particular, we observed a diversification slowdown in *Primula* sect. *Auriculata*, in line with previous findings of Boucher et al. (2016) and Kadereit et al. (2004), and sub-significant also in *Androsace* sect. *Aretia* and *Campanula* sect. *Heterophylla*. However, if this slowdown was due to slower speciation or faster extinction caused by glacial range contractions and habitat loss, it should either be observed also in the other three lineages, or should be strongest in the lineages with most high elevation species.

Of course, the absence of diversification slowdown in the lineages studied in chapter two is biased by our study group selection (we selected them on the basis that they include many extant species and high proportion of alpine species) and thus it does not mean that there was not an extinction increase in Pleistocene in European mountain flora as a whole. It however suggests that there was a way to adapt to Pleistocene climatic changes, and the lineages that managed to do so are among the richest at present. Results from chapter one indeed suggest that, at least in the Alps, peripheral refugia host evolutionary assemblages with long phylogenetic branches, likely of pre-Pleistocene origin. We can easily imagine that the warm and humid climate adapted relatives of Tertiary relict species, as are *Berardia subacaulis* or *Ramonda pyrenaica*, got extinct with the onset of Pleistocene or even earlier as suggested by the paleontological record from the Alps (Fauquette et al., 2018), creating the observed species-museum phylogenetic pattern. With the current analytical means and available data there is little possibility to infer extinction rates of currently species-poor lineages. A

General discussion

promising way forward would be to collect precise phylogenetic information also for species-poor lineages of European flora and analyze it via community-oriented process-based models.

An interesting question is which are the characteristics that determine post glacial expansion of species from peripheral refugia to formerly uninhabitable areas, and in particular to high elevation habitats. Such species definitely should have good dispersal capacities and be able to live in high elevation habitats (Dullinger et al., 2012). Results from chapter one indicate that another important factor for postglacial colonization is bedrock niche, which is supported also by the work of Dullinger et al. (2012). The calcareous refugia often host range restricted species, but this pattern is not paralleled in siliceous refugia. A possible explanation for this is the particular geological arrangement of European mountains, and especially the Alps, where the calcareous massifs are more often at the periphery, whereas the central part is siliceous. During the interglacials, species in calcareous refugia are thus confined to the edges of mountain ranges, whereas silicicolous species can expand through the internal parts of mountain ranges. On the other hand, higher species richness and higher endemism in calcareous habitats are not only known for mountains, but also from lowland regions, where this pattern is traditionally explained by other mechanisms (Ewald, 2003). In order to get a more general picture about the evolutionary assembly in calcareous vs siliceous habitats, it would be interesting to study these patterns across multiple regions of the world in a comparative way, controlling for regional and historical influences.

In contrast to the majority diversification and survival patterns of mid elevation or elevation generalist lineages, there is an exception constituted by few high elevation specialist lineages that do not occur in mid elevation habitats. Analyses from chapter two suggest that the European mountain plant lineages that significantly diversified and only occur in high elevation habitats are two lineages in *Androsace* sect. *Aretia*, one consisting of *A. brevis*, *A. wulfeniana* and three subspecies of *A. vitaliana* and the other of *A. alpina*, *A. obtusifolia*, *A. pubescens*, *A. helvetica* and *A. ciliata*. In the other lineages considered in chapter two, high elevation specialist species occur as well, but the maximum size of purely high elevational subclades is two species. These high elevation lineages could survive glaciations in two ways, either in mid elevation refugia with relict populations gone extinct after the end of glacial periods, or *in situ* of high elevation areas, likely inhabiting scattered small microclimatic refugia, e.g. the ones on nunataks in the Alps. Population genetic studies suggest that mixture of these two mechanisms might have taken place (Escobar García et al., 2012; Schneeweiss

General discussion

and Schönswetter, 2011), for example the high elevation specialist *Androsace alpina* was likely surviving both in peripheral and high elevational refugia on nunataks, and the admixture of these two lineages took place during the Holocene (Schönswetter and Schneeweiss, 2019).

An interesting insight into the processes occurring in high elevation areas is provided by the results of chapter three concerning *Androsace argentea*. Technically, this species is not a strict high elevation specialist as defined in chapter two, because it sometimes occurs on cliffs in the subalpine zone as well, but the processes taking place in its high elevational populations are still very informative about the survival of high elevation specialists. The results from the study on *A. argentea* demonstrate that different high elevation populations within this species may have dramatically different origins, some of them likely surviving the glaciation at the periphery or in relatively distant mountain ranges, while others in high elevational refugia, likely on nunataks protruding the Alpine ice sheet. The demographic analyses also show that the nunatak survival does not necessarily mean stronger bottleneck and lower genetic diversity than survival in peripheral refugia. The survival mechanisms in high elevation specialists thus can be fairly complex and for their better understanding, it would be useful to integrate precise and time explicit genetic information with detailed geomorphological evidence of ice free areas and past climatic conditions, and possibly work comparatively across multiple species.

The findings about the ability of European mountain plants to cope with past climatic fluctuations have important implications for predicting the fate of the European mountain flora under future climatic changes. It is generally assumed that species inhabiting high elevation habitats are more threatened by climate warming than species in lower elevations, because vertical movement of elevational belt caused by warming will reduce the extent of their habitats (La Sorte and Jetz, 2010), especially in pyramid-shaped mountain systems as are most of the mountains in Europe (Elsen and Tingley, 2015). This view however does not account for the fact that biota in different elevations may have different breadths of climatic niche (Thuiller et al., 2014), different dispersal capacities and different survival history. According to the latter point, we can distinguish three main groups of species in the European mountain system :

General discussion

Least threatened by upward shift of elevational belts are the species that survived Pleistocene climatic oscillations in mid elevation refugia or in lowlands, and managed to expand either to high elevation areas or horizontally during the interglacials. These are species that have a relatively large dispersal capacity and can survive in the warmer climate of lower elevational belts. Some of these species likely survived in peripheral refugia at the foothills of mountains (e.g. *Campanula cochlearifolia* or *Gentiana acaulis* or both subspecies of *Phyteuma globulariifolium* from chapter two), and thus had smaller ranges in the past, which indicates that they should be less vulnerable to future range contractions; whereas other species were possibly widely distributed in glacial lowlands during the Pleistocene (e.g. *Polygonum viviparum* or *Dryas octopetala*, Magyari et al., 2014). On the other hand, species with wide glacial distributions often have a wide arctic-alpine distribution (Stewart et al., 2010), and their survival may thus also depend on the processes taking place in the arctic.

Relatively more threatened is the small group of high elevation specialists that survived glaciations *in situ*. These species might be threatened by the complete disappearance of high elevation habitats, but given that these are species that survived past glaciation in small microclimatic refugia, they are unlikely to be seriously harmed by milder range size reduction. An interesting example of how high elevation populations can cope with different climatic contexts is *A. argentea*. The very typical high elevation habitat of this species in the Alps and Pyrenees is south facing rock faces; but in lower elevations it can also grow in other orientations than south (Smyčka, pers. obs.); and for instance, in Sierra Nevada, which has a much warmer climate, the species is confined exclusively to north orientated localities.

The species of the European mountain flora that are definitely most threatened by warming are the range-restricted species that are still confined to mid or low elevation refugia. These species did not manage to recolonize broader areas after the onset of Holocene, either because of poor dispersal capacities, or due to the fact that they require specific conditions present only in the refugium. For example, these can be species from calcareous refugia surrounded by silicates (e.g. *Phyteuma cordatum* or *Gentiana ligustica* from chapter two), or species bound to more specific substrates, as are serpentinites (e.g. *Campanula bertolae* from chapter two). In any case, these species with already limited distribution range are unlikely to move upwards with ongoing warming, which makes them more vulnerable to future climatic changes than the first two groups.

Modes of diversification of the European mountain flora

The dominant mode of speciation in the European mountain system is allopatry, as suggested by the results of chapter two and also previous works (Boucher et al., 2016; Vargas, 2003). Importantly, almost one fifth of speciation events in the six lineages studied in chapter two is also associated with bedrock differentiation, a mechanism that was shown to drive intraspecific genetic structures in mountain plants (Alvarez et al., 2009) or divergence of particular sister species (Moore and Kadereit, 2013; Pachschwöll et al., 2015), but never qualitatively tested. In contrast to this, we found no evidence for parapatric speciation across elevational belts, a mechanism that seems to be important in tropical mountain biodiversity hotspots, often associated with morphological evolution (Luebert and Weigend, 2014; Merckx et al., 2015; Pouchon et al., 2018).

The absence of elevational differentiation in temperate mountains was traditionally attributed to seasonality, given that temperate species at any elevational belt have to handle much larger extent of climatic conditions than mountain species in the tropics (Janzen, 1967). Our results from chapter two suggest that another important factor impeding elevational differentiation might be oscillations on a much longer time scale: the glacial-interglacial oscillations during Pleistocene. The less developed and generally shorter elevational gradient than in tropics may have a double effect on diversification – it prevents lineages from diversifying parapatrically across the elevational gradient, and at the same time the less developed elevational zonation also means less insularity, making the allopatric speciation more difficult (Luebert and Weigend, 2014; Steinbauer et al., 2016).

Moreover, Pleistocene climatic oscillations in tropical mountains are known to have changed connectivity of different habitats, and stimulated diversification by the speciation-pump mechanism, which likely triggered some of rapid plant radiations in the Andes (Madriñán et al., 2013; Pouchon et al., 2018). According to simulation studies, the speciation pump mechanism is dependent on changes of connectivity of middle intensity (Aguilée et al., 2013). In contrast, for example in the Alps and Carpathians the last glacial maximum permanent snowline likely descended below 1900 m a.s.l. (Ronikier, 2011; Schönswetter et al., 2005), which is close to or below current lower limits of high elevation habitats. The Pleistocene climatic oscillations in European mountains have thus likely to caused dramatic habitat displacements, rather than mild changes of connectivity that could stimulate

General discussion

speciation. Our results suggest that Quaternary climatic dynamics did not cause severe diversification slowdown or increased extinction in the lineages studied in chapter two, but at the same time, Quaternary climatic oscillations likely did not stimulate faster diversification as documented from tropical mountain ranges. Unraveling the reasons why floras of temperate mountains are less rich and have slower rates of diversification than subtropical or tropical is likely complex, especially because diversification dynamics differ as well among non-temperate mountains (Hughes, 2017), and would need to be addressed comparatively, such as a few studies have done it (Quintero and Jetz, 2018; Steinbauer et al., 2016). Our results however suggest that diversification processes in the European mountains, and potentially also in other temperate mountain systems, may be driven by fairly specific mechanisms, and the findings from non-temperate mountains cannot be easily translated here.

European mountains are richer in species and endemics than surrounding lowlands, and the fact that the allopatric speciation predominates here suggests that such richness and endemism patterns emerged thanks to the fact that they constitute insular systems. The classical view was that allopatric speciation in the European mountain systems has occurred between the islands formed by major European mountain ranges, as are the Alps, the Pyrenees or the Carpathians (Kadereit, 2017). Our results however suggest that the insular units critical for allopatric speciation lie rather within these major mountain ranges. With the currently available spatial information of European mountain plants we can only hypothesize what forms such islands. For species confined to peripheral refugia, these can be different refugia separated from each other by glaciers in the Alps (Kadereit et al., 2004; Zhang et al., 2001), or by valleys with dry glacial lowland conditions in less glaciated mountain systems. For mid and low elevation species it can also be valleys separated from each other by mountain ridges with hostile high elevation conditions (Thiel-Egenter et al., 2011). Such speciation islands may be also defined by bedrock conditions and the diversification among them may be an interplay between allopatry and adaptive divergence across bedrock types, as suggested by our results from chapter two concerning importance of bedrock-linked speciation.

An even more interesting question is how are allopatric boundaries defined for species inhabiting high elevation areas. The sister pairs of high elevation specialists and elevation generalists identified in chapter two do not co-occur on small scales, similarly as mid elevation sister pairs. It is possible that different high elevation areas are isolated from each

General discussion

other by the valleys (Thiel-Egenter et al., 2011), but on the other hand, wind transport of propagules among high elevation areas isolated by lowlands is possibly more efficient than among valleys isolated by mountain chains (Jeník, 1997). The results of chapter one suggest that, at least in the Alps, the species inhabiting high elevations have generally larger ranges than species of lower elevations, and the idea that high elevation biota was filtered for high geographic mobility is also in line with elevational migration patterns inferred in chapter two.

The findings from chapter three may serve as an explanation of how allopatry can occur even in lineages with high geographic mobility. The population structure of *Androsace argentea* suggests a high prevalence of long distance seed dispersal events across European mountain system. For example, Sierra Nevada, a mountain range on the edge of distribution of the species, was colonized twice independently: once likely from Pyrenees and once from central Alps, which are separated by more than 1300 km. Similarly, we found evidence for multiple long-distance dispersal events among the Alps and the Pyrenees, and also among different regions of the Alps. At the same time, geographically close populations of different origins are often genetically isolated, for example populations in Ecrins and Belledune massifs that are 30 km apart show no evidence of geneflow, possibly due to the stochastic nature of seed dispersal and limited flow of pollen. The boundaries for allopatric speciation may thus be linked to pollen mobility, which may be more restrictive than mobility of seeds in the case of entomogamous plants living in high elevation habitats with low pollinator availability. In such lineages, each successful long distance dispersal event constitutes a potential speciation event, given that resulting populations survive long enough to differentiate morphologically to be recognized as morphological species (De Queiroz, 2007), and eventually develop incompatibility mechanisms to be recognized as biological species (Mayr, 1942). It would be interesting to find out whether the combination of long-distance dispersal and little admixture among neighbouring populations are paralleled in other widespread mountain species in Europe. So far the results from other species rather do not support such idea, long distance dispersal seems to be rare among multiple high elevation species (Schönswetter et al., 2005), and for example in *A. alpina*, local populations of different origins frequently admix (Schönswetter and Schneeweiss, 2019). However it is not clear how much these results are influenced by use of relatively non-informative genetic markers (AFLP, RADP and plastome haplotyping based on few regions in both cited studies). More studies on high elevation

General discussion

plants with use of high-resolution molecular data would thus be needed to shed more light on this problematic.

The rate of speciation is slightly faster in high elevation habitats than in mid and low elevations, as suggested by results of chapter two. This pattern is consistent with findings from studies on different mountains across the world (Hughes, 2017; Madriñán et al., 2013; Quintero and Jetz, 2018), but what is specific in the European mountain system is that faster diversification in high elevation habitats is coupled with strong directional migration, causing that the majority of lineages emerging in high elevations become elevation generalists. An interesting question is what are the main factors that explain the elevated rate of diversification in high elevation habitats. It is possible that such habitats have increased allopatric speciation due to the absence of efficient pollen dispersal as discussed above. We can hypothesize that this mechanism could be coupled with faster molecular evolution due to UV irradiation, causing faster divergence of isolated populations. This idea would need to be tested in future, properly addressing the problems connected to exploring link between molecular and cladogenetic evolution (Lanfear et al., 2010). Fast molecular evolution can also take place through genome duplications. Indeed, polyploid diversification has often been invoked in mountain plants (Jordon-Thaden and Koch, 2008; Schneeweiss et al., 2013; Schönswetter et al., 2007; Theodoridis et al., 2013), and significant ploidy variability is known for three of the six lineages studied in chapter two - *Campanula*, *Saxifraga* and *Phyteuma*.

Another explanation for faster speciation in high-elevation habitats would be that these are younger and less ecologically saturated than mid-elevation and lowland habitats, providing ecological opportunities for diversification. In such a case, the proportion of species in high elevation habitats should be initially low and converge to equilibrium state given by the diversification and migration rates of the ecological assembly process (a similar pattern is discussed for SSE models of morphological evolution in O'Meara et al., 2016). In the results of chapter two we find no support for diversification driven by ecological opportunity, because the currently observed diversity proportions across elevational belts are very close to the equilibrium of the evolutionary assembly model. Moreover, we detected little evidence for diversification slowdown in the six lineages as a whole, and neither older nor dominantly mid elevation clades of these lineages are slowing down faster than the others, as would be predicted by the ecological opportunity scenario. It would be interesting to explore what are

General discussion

the diversification limits of the predominant allopatric speciation processes we observed in the European mountain plants. In other words, whether and when the so far almost exponential diversification observed across the majority of the studied lineages in chapter two is likely to slow down due to factors like minimum viable population sizes (Storch et al., 2018), range sizes (Moen and Morlon, 2014) or hybridizations of allopatrically diverging species.

Future directions

In this thesis I explored the dynamics and drivers of diversification of the European mountain flora. Apart from above discussed answers for some of the research questions, this effort has also brought new motivations for future research, and also produced data and approaches that are to be further explored.

One of the findings of this thesis, and in particular of chapter two, is that the temporal dynamics of diversification and evolution of elevational and bedrock niche are not shared across different mountain lineages. However, it is possible that there are more basal rules shared across different lineages that we did not explore. One of such might be the relationship between species range sizes and rates of diversification. There is a theoretical study that suggests that the rate of diversification should be faster in species with large geographic ranges, especially when the major process of diversification is allopatric speciation (Pigot et al., 2010). The phylogenetic dataset from chapter two, coupled with the geographic ClaSSE framework we developed, could be an optimal playground for testing whether this is really the case. Going even a step forward, it would be of interest to investigate whether diversification dynamics are controlled by current range sizes or glacial-period range sizes, which species had for most of the Pleistocene. Similarly, it would be interesting to explore which types of landscape borders enhance allopatric speciation. A limiting factor for such investigations is the availability of quality-balanced species distribution data across the whole European mountain system, and also information about past climate and glaciation, which is needed to optimally infer current and past species geographic distribution through niche modeling. At the moment, satisfactory occurrence data of mountain plants exist for the Alps, Iberian peninsula, Carpathians and Scandinavia, whereas such data is more incomplete for species occurring in Italy, and is virtually missing for the Balkans region. It is an interesting paradox that nowadays high-quality phylogenetic data is available, or easy to obtain, for many European mountain plant species, whereas for the very same species as basal characteristic as geographic occurrence is still missing.

Allopatric speciation requires geographical isolation of populations, followed by morphological differentiation or emergence of reproductive barriers, so that diverging populations can be recognized as separate species (Boucher et al., 2016; Coyne and Orr, 2004). This differentiation process might theoretically be accelerated in high elevation

General discussion

habitats due to higher UV irradiance causing faster mutation rates (Cortés et al., 2018; Willis et al., 2009). On the other hand, the rate of molecular evolution is also dependent on temperature (Davies et al., 2004; Gillman et al., 2010; Wright et al., 2006), suggesting that it may be slower in high elevation habitats. No consensus has emerged yet whether lineages inhabiting high elevation areas have faster rates of molecular evolution, and little is known also about the link between cladogenetic and molecular evolution in general (Goldie et al., 2011; Lanfear et al., 2010). The chloroplast and nuclear genomic data we collected for chapter two, and eventually genomic data from a broader sample of mountain plant lineages would be an extremely suitable dataset for exploring these relationships.

An important limitation of the approaches used in my thesis is the ability to infer rates of extinction, despite the dynamics of extinction is part of the central concepts I use. The reasons for such limitation is partially fundamental, given that estimates of extinction rates based only on molecular phylogenies combined with recent data will always be dependent on *a priori* assumptions about the diversification process (Morlon, 2014; Rabosky, 2010). However, another source of limitations is related to data availability and the sampling schemes we used. In particular, the sample of six lineages in chapter two allowed us to explore the extinction patterns in fast diversifying lineages in the European mountain system, but the processes taking place in species-poor lineages may also be important for understanding overall extinction dynamics. Such lineages were covered in chapter one, but in turn, the resolution of phylogenetic data only to the level of genera did not allow us to address extinction dynamics in greater detail. The optimal method here would thus to couple highly resolved flora-wide phylogenetic data with community-oriented process based diversification models that explicitly work with extinction, as DAISIE does (Valente et al., 2015). Such flora-wide phylogenetic data will soon be available thanks to the ongoing PhyloAlps initiative, which is gathering common ground genomic data of the same type as used in chapter two for the whole floras of the Alps, Carpathians, Scandinavian mountains and significant portions of floras of other European mountain regions. Once ready, these data will allow us to obtain highly resolved flora-wide phylogenies and infer diversification processes without compromising the scope by lineage selection.

Going even further in generalization, it is extremely tempting to compare the evolutionary patterns we observe in European mountains with other mountain systems of the world. This comparison is however limited by the quality of the information against which we could

General discussion

compare. Especially evolutionary assembly across elevational gradients and between calcareous and siliceous bedrocks would deserve to be tested in a global context, disentangling local and general drivers. Recently, there has been multiple studies of mountain diversification across the world ((Quintero and Jetz, 2018; Steinbauer et al., 2016), limited however by phylogenetic, occurrence and ecological informations available at these scales. For example in Chinese Himalaya or on Indonesian Mountains, similar floristic and phylogenetic initiatives as is PhyloAlps are emerging (Cheng et al., 2018; Merckx et al., 2015; Xing and Ree, 2017), promising massive amelioration in this direction. Such efforts may be complicated in other regions, due to limited accessibility, lower level of systematic knowledge or for example extremely fast diversification making quantitative common-ground methods unsuitable (Pouchon et al., 2018), but I believe that this is an important direction towards understanding diversification processes (not only) in mountain regions.

An important challenge to future diversification studies is dealing with systematic concepts and definitions of species. In this thesis I worked with the flora of European mountains, which is one of the best explored floras of the world, at least in terms of classical systematics. Despite this, we had to pay extreme attention to systematic concepts and species definitions we were using, especially in chapter two. For example we originally wanted to include the genus *Soldanella* as the 7th lineage, but after comparing plastome and nuclear phylogenies with previous studies (Boucher et al., 2016; Zhang et al., 2001; Zhang and Kadereit, 2004) we found out that species delimitations in this group has to be completely reworked (Slovák, Smyčka et al., in prep.). Species definition is crucial for modeling dynamics of diversification– defining too broad species leads to slowdown due to protracted speciation artifact (Moen and Morlon, 2014), whereas defining very narrow species means working with unrealistic tree topologies, due to incomplete lineage sorting. One solution for such issues is to work with firm species definitions that are comparable across the groups, which may not always be achievable in flora-wide studies. Another approach is to develop analytical methods that can control for species definition. A pioneer study in this direction has proposed the phylogenetic endemism index (Rosauer et al., 2009), which is designed to be invariant to merging the neighboring tree tips and their range information. Process-based models of diversification adopting similar approaches are yet to come in the future.

General discussion

References

- Aeschimann, D., Rasolofoa, N., Theurillat, J.P., 2011. Analyse de la flore des Alpes. 2: diversité et chorologie. *Candollea* 66, 225–253.
- Aguilée, R., Claessen, D., Lambert, A., 2013. Adaptive radiation driven by the interplay of eco-evolutionary and landscape dynamics. *Evolution* (N. Y.) 67, 1291–1306. <https://doi.org/10.1111/evo.12008>
- Alvarez, N., Thiel-Egenter, C., Tribsch, A., Holderegger, R., Manel, S., Schönswetter, P., Taberlet, P., Brodbeck, S., Gaudeul, M., Gielly, L., Küpfer, P., Mansion, G., Negrini, R., Paun, O., Pellicchia, M., Rioux, D., Schüpfer, F., Van Loo, M., Winkler, M., Gugerli, F., 2009. History or ecology? Substrate type as a major driver of spatial genetic structure in Alpine plants. *Ecol. Lett.* 12, 632–640. <https://doi.org/10.1111/j.1461-0248.2009.01312.x>
- Boucher, F.C., Zimmermann, N.E., Conti, E., 2016. Allopatric speciation with little niche divergence is common among alpine Primulaceae. *J. Biogeogr.* 43, 591–602. <https://doi.org/10.1111/jbi.12652>
- Cheng, S., Melkonian, M., Smith, S.A., Brockington, S., Archibald, J.M., Delaux, P.M., Li, F.W., Melkonian, B., Mavrodiev, E. V, Sun, W., Fu, Y., Yang, H., Soltis, D.E., Graham, S.W., Soltis, P.S., Liu, X., Xu, X., Wong, G.K.S., 2018. 10KP: A phylodiverse genome sequencing plan. *Gigascience* 7, 1–9. <https://doi.org/10.1093/gigascience/giy013>
- Cortés, A.J., Garzón, L.N., Valencia, J.B., Madriñán, S., 2018. On the Causes of Rapid Diversification in the Páramos: Isolation by Ecology and Genomic Divergence in Espeletia. *Front. Plant Sci.* 9, 1–17. <https://doi.org/10.3389/fpls.2018.01700>
- Coyne, J., Orr, H., 2004. *Speciation*. Sinauer Associates Inc, Sunderland.
- Davies, T.J., Savolainen, V., Chase, M.W., Moat, J., Barraclough, T.G., 2004. Environmental energy and evolutionary rates in flowering plants. *Proc. R. Soc. B Biol. Sci.* 271, 2195–2200. <https://doi.org/10.1098/rspb.2004.2849>
- De Queiroz, K., 2007. Species concepts and species delimitation. *Syst. Biol.* 56, 879–86. <https://doi.org/10.1080/10635150701701083>
- Dullinger, S., Willner, W., Plutzer, C., Englisch, T., Schratz-Ehrendorfer, L., Moser, D., Ertl, S., Essl, F., Niklfeld, H., 2012. Post-glacial migration lag restricts range filling of plants in the European Alps. *Glob. Ecol. Biogeogr.* 21, 829–840. <https://doi.org/10.1111/j.1466-8238.2011.00732.x>
- Elsen, P.R., Tingley, M.W., 2015. Global mountain topography and the fate of montane species under climate change. *Nat. Clim. Chang.* 5, 772–776. <https://doi.org/10.1038/nclimate2656>
- Escobar García, P., Winkler, M., Flatscher, R., Sonnleitner, M., Krejčíková, J., Suda, J., Hülber, K., Schneeweiss, G.M., Schönswetter, P., 2012. Extensive range persistence in peripheral and interior refugia characterizes Pleistocene range dynamics in a widespread Alpine plant species (*Senecio carniolicus*, Asteraceae). *Mol. Ecol.* 21, 1255–1270. <https://doi.org/10.1111/j.1365-294X.2012.05456.x>
- Ewald, J., 2003. The calcareous riddle: Why are there so many calciphilous species in the Central European flora? *Folia Geobot.* 38, 357–366. <https://doi.org/10.1007/BF02803244>
- Fauquette, S., Suc, J.-P., Médail, F., Muller, S.D., G., J.-M., Bertini, A., Martinetto, E., Popescu, S.-P., Zheng, Z., de Beaulieu, J.L., 2018. The Alps: a geological, climatic and human perspective on vegetation history and modern plant diversity, in: Hoorn, C., Perrigo, A., Antonelli, A. (Eds.), *Mountains, Climate and Biodiversity*. Wiley-Blackwell, p. 413.
- Gillman, L.N., Keeling, D.J., Gardner, R.C., Wright, S.D., 2010. Faster evolution of highly conserved DNA in tropical plants. *J. Evol. Biol.* 23, 1327–1330. <https://doi.org/10.1111/j.1420-9101.2010.01992.x>
- Goldie, X., Lanfear, R., Bromham, L., 2011. Diversification and the rate of molecular evolution : no evidence of a link in mammals. *BMC Evol. Biol.* 11. <https://doi.org/10.1186/1471-2148-11-286>
- Hughes, C.E., 2017. Are there many different routes to becoming a global biodiversity hotspot? *Proc. Natl. Acad. Sci.* 114, 4275–4277. <https://doi.org/10.1073/pnas.1703798114>
- Janzen, D.H., 1967. Why mountain passes are higher in the tropics. *Am. Nat.* 101, 233–249.
- Jeník, J., 1997. Anemo-orographic systems in the Hercynian Mts. and their effects on biodiversity. *Acta Univ. Wratislawensis* 4, 9–21.
- Jordon-Thaden, I., Koch, M., 2008. Species richness and polyploid patterns in the genus *Draba* (Brassicaceae): A first global perspective. *Plant Ecol. Divers.* 1, 255–263. <https://doi.org/10.1080/17550870802349112>

General discussion

- Kadereit, J.W., 2017. The role of in situ species diversification for the evolution of high vascular plant species diversity in the European Alps—A review and interpretation of phylogenetic studies of the endemic flora of the Alps. *Perspect. Plant Ecol. Evol. Syst.* 26, 28–38. <https://doi.org/10.1016/j.ppees.2017.03.002>
- Kadereit, J.W., Griebeler, E.M., Comes, H., 2004. Quaternary diversification in European alpine plants: pattern and process. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 359, 265–274. <https://doi.org/10.1098/rstb.2003.1389>
- La Sorte, F.A., Jetz, W., 2010. Projected range contractions of montane biodiversity under global warming. *Proc. R. Soc. B Biol. Sci.* 277, 3401–3410. <https://doi.org/10.1098/rspb.2010.0612>
- Lanfear, R., Welch, J.J., Bromham, L., 2010. Watching the clock: studying variation in rates of molecular evolution between species. *Trends Ecol. Evol.* 25, 495–503. <https://doi.org/10.1016/j.tree.2010.06.007>
- Lomolino, M. V., 2001. Elevation gradients of species-density: Historical and prospective views. *Glob. Ecol. Biogeogr.* 10, 3–13. <https://doi.org/10.1046/j.1466-822x.2001.00229.x>
- Luebert, F., Weigend, M., 2014. Phylogenetic insights into Andean plant diversification. *Front. Ecol. Evol.* 2, 1–17. <https://doi.org/10.3389/fevo.2014.00027>
- Madriñán, S., Cortés, A.J., Richardson, J.E., 2013. Páramo is the world's fastest evolving and coolest biodiversity hotspot. *Front. Genet.* 4, 1–7. <https://doi.org/10.3389/fgene.2013.00192>
- Magyari, E.K., Kuneš, P., Jakab, G., Sümegi, P., Pelánková, B., Schäbitz, FrankBraun, M., Chytrý, M., 2014. Late Pleniglacial vegetation in eastern-central Europe: are there modern analogues in Siberia? *Quat. Sci. Rev.* 95, 60–79. <https://doi.org/10.1016/j.quascirev.2014.04.020>
- Mayr, E., 1942. *Systematics and the origin of species*. Columbia University Press, New York.
- Merckx, V.S.F.T., Hendriks, K.P., Beentjes, K.K., Mennes, C.B., Becking, L.E., Peijnenburg, K.T.C. a., Afendy, A., Arumugam, N., de Boer, H., Biun, A., Buang, M.M., Chen, P.-P., Chung, A.Y.C., Dow, R., Feijen, F. a. a., Feijen, H., Soest, C.F., Geml, J., Geurts, R., Gravendeel, B., Hovenkamp, P., Imbun, P., Ipor, I., Janssens, S.B., Jocqué, M., Kappes, H., Khoo, E., Koomen, P., Lens, F., Majapun, R.J., Morgado, L.N., Neupane, S., Nieser, N., Pereira, J.T., Rahman, H., Sabran, S., Sawang, A., Schwallier, R.M., Shim, P.-S., Smit, H., Sol, N., Spait, M., Stech, M., Stokvis, F., Sugau, J.B., Suleiman, M., Sumail, S., Thomas, D.C., van Tol, J., Tuh, F.Y.Y., Yahya, B.E., Nais, J., Repin, R., Lakim, M., Schilthuizen, M., 2015. Evolution of endemism on a young tropical mountain. *Nature* 524, 347–350. <https://doi.org/10.1038/nature14949>
- Moen, D., Morlon, H., 2014. Why does diversification slow down? *Trends Ecol. Evol.* 29, 190–197. <https://doi.org/10.1016/j.tree.2014.01.010>
- Molina-Venegas, R., Aparicio, A., Slingsby, J.A., Lavergne, S., Arroyo, J., 2015. Investigating the evolutionary assembly of a Mediterranean biodiversity hotspot: Deep phylogenetic signal in the distribution of eudicots across elevational belts. *J. Biogeogr.* 42, 507–518. <https://doi.org/10.1111/jbi.12398>
- Moore, A.J., Kadereit, J.W., 2013. The evolution of substrate differentiation in *Minuartia* series *Laricifoliae* (Caryophyllaceae) in the European Alps: In situ origin or repeated colonization? *Am. J. Bot.* 100, 2412–2425. <https://doi.org/10.3732/ajb.1300225>
- Morlon, H., 2014. Phylogenetic approaches for studying diversification. *Ecol. Lett.* 17, 508–525. <https://doi.org/10.1111/ele.12251>
- Mráz, P., Barabas, D., Lengyelová, L., Turis, P., Schmotzer, A., Janišová, M., Ronikier, M., 2016. Vascular plant endemism in the Western Carpathians: spatial patterns, environmental correlates and taxon traits. *Biol. J. Linn. Soc.* 119, 630–648. <https://doi.org/10.1111/bij.12792>
- O'Meara, B.C., Smith, S.D., Armbruster, W.S., Harder, L.D., Hardy, C.R., Hileman, L.C., Hufford, L., Litt, A., Magallón, S., Smith, S.A., Stevens, P.F., Fenster, C.B., Diggle, P.K., 2016. Non-equilibrium dynamics and floral trait interactions shape extant angiosperm diversity. *Proc. R. Soc. B Biol. Sci.* 283, 20152304. <https://doi.org/10.1098/rspb.2015.2304>
- Pachschwöll, C., García, P.E., Winkler, M., Schneeweiss, G.M., Schönswetter, P., 2015. Polyploidisation and geographic differentiation drive diversification in a european high mountain plant group (*Doronicum clusii* aggregate, Asteraceae). *PLoS One* 10, 1–30. <https://doi.org/10.1371/journal.pone.0118197>
- 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. *Syst. Biol.* 59, 660–673. <https://doi.org/10.1093/sysbio/syq058>

General discussion

- Pouchon, C., Fernández, A., Nassar, J.M., Boyer, F., Aubert, S., Lavergne, S., Mavárez, J., 2018. Phylogenomic analysis of the explosive adaptive radiation of the Espeletia complex (Asteraceae) in the tropical Andes. *Syst. Biol.* 67, 1041–1060. <https://doi.org/10.1093/sysbio/syy022>
- Quintero, I., Jetz, W., 2018. Global elevational diversity and diversification of birds. *Nature* 555, 246–250. <https://doi.org/10.1038/nature25794>
- Rabosky, D.L., 2010. Extinction rates should not be estimated from molecular phylogenies. *Evolution (N. Y.)* 64, 1816–1824. <https://doi.org/10.1111/j.1558-5646.2009.00926.x>
- Rahbek, C., 1995. The elevational gradient of species richness: a uniform pattern? *Ecography (Cop.)* 18, 200–205. <https://doi.org/10.1111/j.1600-0587.1995.tb00341.x>
- Ronikier, M., 2011. Biogeography of high-mountain plants in the Carpathians: An emerging phylogeographical perspective. *Taxon* 60, 373–389. <https://doi.org/10.1002/tax.602008>
- Roquet, C., Boucher, F.C., Thuiller, W., Lavergne, S., 2013. Replicated radiations of the alpine genus *Androsace* (Primulaceae) driven by range expansion and convergent key innovations. *J. Biogeogr.* 40, 1874–1886. <https://doi.org/10.1111/jbi.12135>
- Rosauer, D., Laffan, S.W., Crisp, M.D., Donnellan, S.C., Cook, L.G., 2009. Phylogenetic endemism: A new approach for identifying geographical concentrations of evolutionary history. *Mol. Ecol.* 18, 4061–4072. <https://doi.org/10.1111/j.1365-294X.2009.04311.x>
- Schneeweiss, G.M., Pachschröll, C., Tribsch, A., Schönswetter, P., Barfuss, M.H.J., Esfeld, K., Weiss-Schneeweiss, H., Thiv, M., 2013. Molecular phylogenetic analyses identify Alpine differentiation and dysploid chromosome number changes as major forces for the evolution of the European endemic *Phyteuma* (Campanulaceae). *Mol. Phylogenet. Evol.* 69, 634–652. <https://doi.org/10.1016/j.ympev.2013.07.015>
- Schneeweiss, G.M., Schönswetter, P., 2011. A re-appraisal of nunatak survival in arctic-alpine phylogeography. *Mol. Ecol.* 20, 190–192. <https://doi.org/10.1111/j.1365-294X.2010.04927.x>
- Schönswetter, P., Lachmayer, M., Lettner, C., Prehler, D., Rechnitzer, S., Reich, D.S., Sonnleitner, M., Wagner, I., Hülber, K., Schneeweiss, G.M., Trávníček, P., Suda, J., 2007. Sympatric diploid and hexaploid cytotypes of *Senecio carniolicus* (Asteraceae) in the Eastern Alps are separated along an altitudinal gradient. *J. Plant Res.* 120, 721–725. <https://doi.org/10.1007/s10265-007-0108-x>
- Schönswetter, P., Schneeweiss, G.M., 2019. Is the incidence of survival in interior Pleistocene refugia (nunataks) underestimated? Phylogeography of the high mountain plant *Androsace alpina* (Primulaceae) in the European Alps revisited. *Ecol. Evol.* 9, 4078–4086. <https://doi.org/10.1002/ece3.5037>
- Schönswetter, P., Stehlik, I., Holderegger, R., Tribsch, A., 2005. Molecular evidence for glacial refugia of mountain plants in the European Alps. *Mol. Ecol.* 14, 3547–3555. <https://doi.org/10.1111/j.1365-294X.2005.02683.x>
- Stehlik, I., 2003. Resistance or emigration? Response of alpine plants to the ice ages. *Taxon* 52, 499–510. <https://doi.org/10.2307/3647448>
- Steinbauer, M.J., Field, R., Grytnes, J.A., Trigas, P., Ah-Peng, C., Attorre, F., Birks, H.J.B., Borges, P.A.V., Cardoso, P., Chou, C.H., De Sanctis, M., de Sequeira, M.M., Duarte, M.C., Elias, R.B., Fernández-Palacios, J.M., Gabriel, R., Gereau, R.E., Gillespie, R.G., Greimler, J., Harter, D.E.V., Huang, T.J., Irl, S.D.H., Jeanmonod, D., Jentsch, A., Jump, A.S., Kueffer, C., Nogué, S., Otto, R., Price, J., Romeiras, M.M., Strasberg, D., Stuessy, T., Svenning, J.C., Vetaas, O.R., Beierkuhnlein, C., 2016. Topography-driven isolation, speciation and a global increase of endemism with elevation. *Glob. Ecol. Biogeogr.* 25, 1097–1107. <https://doi.org/10.1111/geb.12469>
- Stewart, J.R., Lister, A.M., Barnes, I., Dalén, L., 2010. Refugia revisited: individualistic responses of species in space and time. *Proc. Biol. Sci.* 277, 661–71. <https://doi.org/10.1098/rspb.2009.1272>
- Storch, D., Bohdalková, E., Okie, J., 2018. The more-individuals hypothesis revisited: the role of community abundance in species richness regulation and the productivity–diversity relationship. *Ecol. Lett.* 21, 920–937. <https://doi.org/10.1111/ele.12941>
- Theodoridis, S., Randin, C., Broennimann, O., Patsiou, T., Conti, E., 2013. Divergent and narrower climatic niches characterize polyploid species of European primroses in *Primula* sect. *Aleuritia*. *J. Biogeogr.* 40, 1278–1289. <https://doi.org/10.1111/jbi.12085>
- Thiel-Egenter, C., Alvarez, N., Holderegger, R., Tribsch, A., Englisch, T., Wohlgemuth, T., Colli, L., Gaudeul, M., Gielly, L., Jogan, N., Linder, H.P., Negrini, R., Niklfeld, H., Pellicchia, M., Rioux, D., Schönswetter,

General discussion

- P., Taberlet, P., van Loo, M., Winkler, M., Gugerli, F., 2011. Break zones in the distributions of alleles and species in alpine plants. *J. Biogeogr.* 38, 772–782. <https://doi.org/10.1111/j.1365-2699.2010.02441.x>
- Thuiller, W., Guéguen, M., Georges, D., Bonet, R., Chalmandrier, L., Garraud, L., Renaud, J., Roquet, C., Van Es, J., Zimmermann, N.E., Lavergne, S., 2014. Are different facets of plant diversity well protected against climate and land cover changes? A test study in the French Alps. *Ecography (Cop.)*. 37, 1254–1266. <https://doi.org/10.1111/ecog.00670>
- Valente, L.M., Phillimore, A.B., Etienne, R.S., 2015. Equilibrium and non-equilibrium dynamics simultaneously operate in the Galapagos islands. *Ecol. Lett.* 18, 844–852. <https://doi.org/10.1111/ele.12461>
- Vargas, P., 2003. Molecular evidence for multiple diversification patterns of alpine plants in Mediterranean Europe. *Taxon* 52, 463–476.
- Willis, K.J., Bennett, K.D., Birks, H.J.B., 2009. Variability in thermal and UV-B energy fluxes through time and their influence on plant diversity and speciation. *J. Biogeogr.* 36, 1630–1644. <https://doi.org/10.1111/j.1365-2699.2009.02102.x>
- Wright, S., Keeling, J., Gillman, L., 2006. The road from Santa Rosalia: a faster tempo of evolution in tropical climates. *Proc. Natl. Acad. Sci. U. S. A.* 103, 7718–7722. <https://doi.org/10.1073/pnas.0510383103>
- Xing, Y., Ree, R.H., 2017. Uplift-driven diversification in the Hengduan Mountains, a temperate biodiversity hotspot. *Proc. Natl. Acad. Sci.* 114, E3444–E3451. <https://doi.org/https://doi.org/10.1073/pnas.1616063114>
- Zhang, L.B., Comes, H.P., Kadereit, J.W., 2001. Phylogeny and quaternary history of the European montane/alpine endemic *Soldanella* (Primulaceae) based on ITS and AFLP variation. *Am. J. Bot.* 88, 2331–2345. <https://doi.org/10.2307/3558393>
- Zhang, L.B., Kadereit, J.W., 2004. Nomenclature of *Soldanella* L. (Primulaceae). *Taxon* 53, 741–752. <https://doi.org/10.2307/4135448>

APPENDICES OF CHAPTER 1

Appendix A – Comparison of models accounting for spatial autocorrelation

We tested whether other methods for accounting for spatial autocorrelation could yield similar results as the generalized additive model (GAM) we used. To do so, we fitted a generalized least square model (GLS) with an linear correlation structure and conditional autoregressive model (CAR) accounting for closest neighbour cells. For comparison we present also non-spatial linear model. For the sake of simplicity, all models were fitted in a frequentist framework using mean values per grid cell as the response variables. The proportion of endemics was logit transformed prior to analyses. The tables below report the results for three study variables, namely the proportion of endemic species, the mean endemic rarity, and species richness standardized phylogenetic diversity. The estimated effect size (ES) is given, along with its p-value. Results in bold indicate $p < 0.05$.

The main difference between bayesian GAM presented in results and GLS or CAR presented here is that CAR and GLS depicts a growth of the proportion of endemics with elevation range rather than with mean elevation. This difference would not change the overall interpretation of the results, and can likely be attributed to a high correlation between mean elevation and elevation range predictors, and to the absence of hierarchic structures in CAR and GLS models, compared to bayesian GAM.

Appendices of chapter 1

Proportion of endemics

	GLS		CAR		non-spatial	
	ES	p	ES	P	ES	p
intercept	-5.392	<0.001	-5.158	<0.001	-5.374	<0.001
calcareous bedrock	0.407	0.055	0.329	0.126	0.443	0.038
calcareous refugia	0.511	0.004	0.374	0.039	0.596	<0.001
siliceous bedrock	-0.285	0.173	-0.297	0.185	-0.244	0.230
siliceous refugia	0.197	0.391	0.038	0.872	0.289	0.182
mean elevation	0.293	0.206	0.354	0.143	0.333	0.127
elevation range	1.251	<0.001	1.144	<0.001	1.169	<0.001

Mean endemic rarity

	GLS		CAR		non-spatial	
	ES	p	ES	p	ES	p
intercept	-4.229	<0.001	-4.360	<0.001	-3.991	<0.001
calcareous bedrock	0.112	0.071	0.146	0.019	0.135	0.107
calcareous refugia	0.129	0.014	0.059	0.249	0.189	0.004
siliceous bedrock	-0.074	0.267	-0.032	0.637	-0.209	0.009
siliceous refugia	0.114	0.100	0.071	0.311	0.522	<0.001
mean elevation	0.020	0.790	0.110	0.143	-0.116	0.155
elevation range	0.114	0.064	0.100	0.104	0.050	0.510

Richness standardized phylogenetic diversity (ses.PD)

	GLS		CAR		non-spatial	
	ES	p	ES	p	ES	p
intercept	0.179	0.256	0.248	0.172	0.140	0.346
calcareous bedrock	0.454	<0.001	0.365	0.004	0.484	<0.001
calcareous refugia	0.529	<0.001	0.415	<0.001	0.605	<0.001
siliceous bedrock	-0.199	0.103	-0.179	0.175	-0.170	0.149
siliceous refugia	0.292	0.030	0.190	0.176	0.260	0.039
mean elevation	-0.285	0.034	-0.293	0.039	-0.381	0.003
elevation range	0.162	0.174	0.191	0.124	0.220	0.057

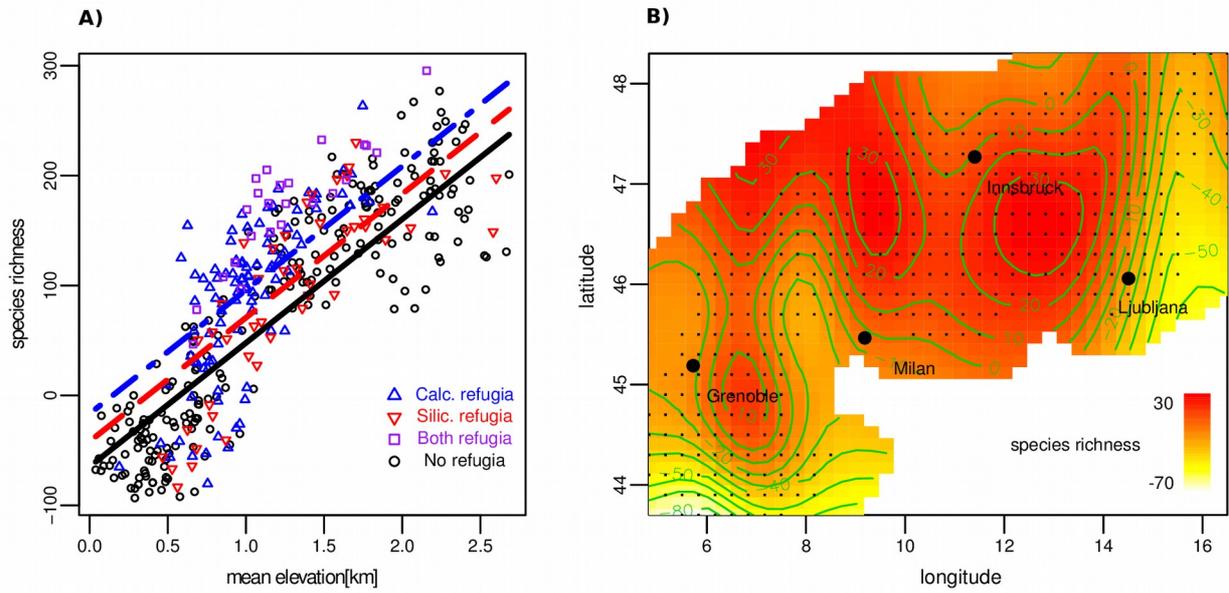
Appendix B – The spatial model of species richness

Spatial patterns of species richness should not be interpreted independently, as this variable is biased by accounting only for species that can occur above treeline. However, for interpretation of the phylogenetic diversity or the ratio of endemics to species, it is useful to know the trends of species richness in the dataset. In order to explore this, we fitted a bayesian generalized additive model (GAM) similar to spatial models used for the other variables. The species richness was modelled as having a Gaussian distribution of errors.

Panel (A) shows the relationship of species richness with mean elevation (black, solid line, plotted in bold if effect of mean elevation is significant) and the occurrence of calcareous refugia (blue, double dashed line, plotted in bold if effect of calcareous refugia is significant), siliceous refugia (red, single dashed line, plotted in bold if effect of siliceous refugia is significant). Note that y-values of points are adjusted to account for the effect of not-displayed model variables and for the mean effect of smooth model component per group (and hence the species richness may become negative). Panel (B) represents the smooth component of the model, showing geographic areas with overall higher or lower species richness when simultaneously accounting for parametric model components. Estimates and credibility intervals of model parameters are given in table.

	lower	ES	upper
intercept	-84.162	-64.784	-45.087
calc. bedrock	36.563	49.039	61.538
calc. refugia	37.747	48.121	58.536
silic. bedrock	5.317	19.356	33.262
silic. refugia	9.231	22.949	36.668
mean elevation	97.762	112.558	127.268
elevation range	38.298	51.402	64.433

Appendices of chapter 1



Appendix C – Species level analysis with phylogenetic corection

Endemic status

Species endemic status was modeled against species-level predictors using a a binomial model with a phylogenetic correlation structure. The Ornstein-Uhlenbeck alpha=0.193 suggests very weak phylogenetic signal. The estimated effect size (ES) is given, along with its p-value, results in bold indicate $p < 0.05$.

	ES	p
O-U alpha	0.193	
elevational optimum	0.238	0.009
evolutionary distinctiveness	-0.113	0.228
ruderal strategy	N	N
stress strategy	0.269	0.003
dispersal	-0.296	0.001
sexuality	0.236	0.044
vegetative reproduction	N	N
cushions	0.018	0.833
chamaephyte	N	N
geophyte	N	N
hemicryptophyte	N	N
phanerophyte	N	N
therophyte	N	N

Endemic rarity

The rarity of endemic species was modeled against species-level predictors using a log-normal model with a phylogenetic correlation structure. The Pagel lambda=0.037 suggests very weak phylogenetic signal. The estimated effect size (ES) is given, along with its p-value, results in bold indicate $p < 0.05$.

	ES	p
Pagel lambda	0.037	
elevational optimum	-0.199	0.012
evolutionary distinctiveness	N	N
ruderal strategy	N	N
stress strategy	0.141	0.106
dispersal	-0.146	0.097
sexuality	0.200	0.007
vegetative reproduction	N	N
cushions	N	N
chamaephyte	N	N
geophyte	0.259	0.002
hemicryptophyte	0.157	0.071
phanerophyte	N	N
therophyte	N	N

Appendix D - BUGS code of the bayesian models

Proportion of endemics

```

model {
  #linear regression
  eta <- X %*% b ## linear regression

  #random effect
  for (i in 1:n) { rand[i]~dnorm(0,tau) }
  tau ~ dgamma(.05,.005)

  #expected response
  for (i in 1:n) { mu[i] <- ilogit(eta[i]+rand[i]) }

  #response
  for (i in 1:n) { end[i] ~ dbin(mu[i],div[i]) }

  #parametric effect priors
  for (i in 1:7) { b[i] ~ dnorm(0,0.039) }

  #priors for smooth terms
  K1 <- S1[1:29,1:29] * lambda[1] + S1[1:29,30:58] *
lambda[2]
  b[8:36] ~ dnorm(zero[8:36],K1)

  #smoothing parameter priors
  for (i in 1:2) {
    lambda[i] ~ dgamma(.05,.005)
    rho[i] <- log(lambda[i])
  }
}

```

Mean endemic rarity

```

model {
  #imperfect detection layer
  rartau~dgamma(.05,.005)
  for (p in 1:n){
    #uncertainty
    for (t in 1:nend[p]){
      lrar[t,p]~dnorm(rarmean[p],rartau)
    }
  }
}

```

Appendices of chapter 1

```
#linear regression
mu <- X %*% b

#response
for (i in 1:n) { rarmean[i] ~ dnorm(mu[i],tau) }

#precision parameter prior
tau ~ dgamma(.05,.005)
scale <- 1/tau

#parametric effect priors
for (i in 1:7) { b[i] ~ dnorm(0,0.00052) }

#priors for smooth terms
K1 <- S1[1:29,1:29] * lambda[1] + S1[1:29,30:58] *
lambda[2]
b[8:36] ~ dnorm(zero[8:36],K1)

#smoothing parameter priors
for (i in 1:2) {
  lambda[i] ~ dgamma(.05,.005)
  rho[i] <- log(lambda[i])
}
}
```

Richness standardized phylogenetic diversity (ses.PD)

```
model {
  #imperfect detection layer
  mntdtau~dgamma(.05,.005)
  for (p in 1:n){
    #phylogenetic uncertainty
    for (t in 1:1000){
      mntd[p,t]~dnorm(mntdmean[p],mntdtau)
    }
  }

  #linear regression
  mu <- X %*% b

  #response
  for (i in 1:n) {
    mntdmean[i] ~ dnorm(mu[i],tau)
  }
}
```

Appendices of chapter 1

```
}

#precision parameter prior
tau ~ dgamma(.05,.005)
scale <- 1/tau

#parametric effect priors
for (i in 1:7) { b[i] ~ dnorm(0,0.019) }

#priors for smooth terms
K1 <- S1[1:29,1:29] * lambda[1] + S1[1:29,30:58] *
lambda[2]
b[8:36] ~ dnorm(zero[8:36],K1)

#smoothing parameter priors
for (i in 1:2) {
  lambda[i] ~ dgamma(.05,.005)
  rho[i] <- log(lambda[i])
}
}
```

APPENDICES OF CHAPTER 2

Supplementary materials for “Tempo and drivers of plant diversification in the European mountain system”

Contents:

Figure S1 – Rates of diversification across 5 major mountain regions in European mountain system.

Figure S2 – Current and equilibrium species proportions across the bedrock and elevational gradient.

Figure S3 – Ancestral reconstructions of bedrock and elevational niches.

Appendix S1 – Species sampling and taxonomic treatment

Appendix S2 – Genomic regions selection and sequence processing.

Appendix S3 – Phylogenetic analyses and fossil calibrations.

Appendix S4 – Compilation of ecological and geographic information.

Appendix S5 – The PhyloAlps consortium

Dataset S1 – Accession tables.

Dataset S2 – Genomic regions.

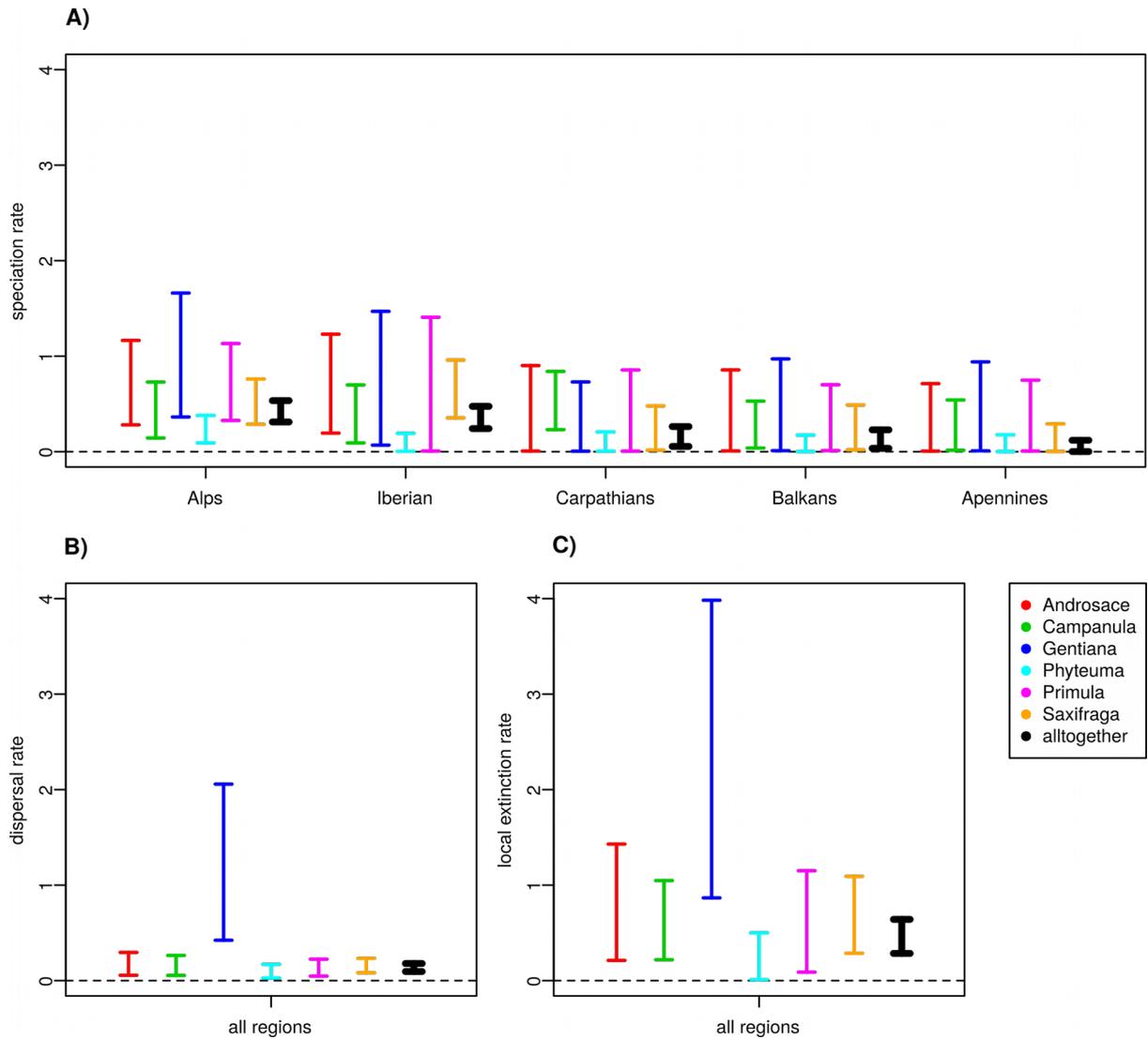
Dataset S3 – Ecological and geographic information.

Dataset S4 – Multi-lineage time-dependent diversification model script

Dataset S5 – Multi-regions GeoSSE script

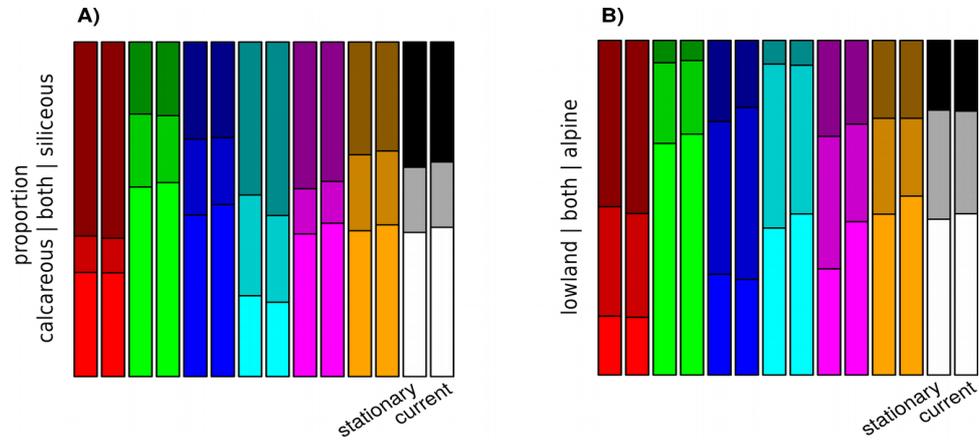
Appendices of chapter 2

Figure S1: Rates of (A) speciation across 5 regions of European mountain system, (B) dispersal and (C) local extinction common for all the 5 regions. The bars indicate 95% credibility interval of Bayesian parameter estimate of ClaSSE model with the best AIC (no allopatric speciation, region-specific speciation rates, single extinction rate, symmetric dispersal).



Appendices of chapter 2

Figure S2: Current and equilibrium proportions of (A) calcareous specialists, siliceous specialists or bedrock generalist species, respectively (B) high elevation specialists, mid-elevation specialists, and elevation generalist species. The color scheme follows one used in Fig. S1, but includes shading to differentiate niche states. Equilibrium proportions represent a limit distribution of states of the model for $\text{time} \rightarrow \infty$, and are calculated by eigendecomposition of ClaSSE model matrices constructed from mean posterior parameter estimates, via function *stationary.freq.classe* in R package *diversitree*.



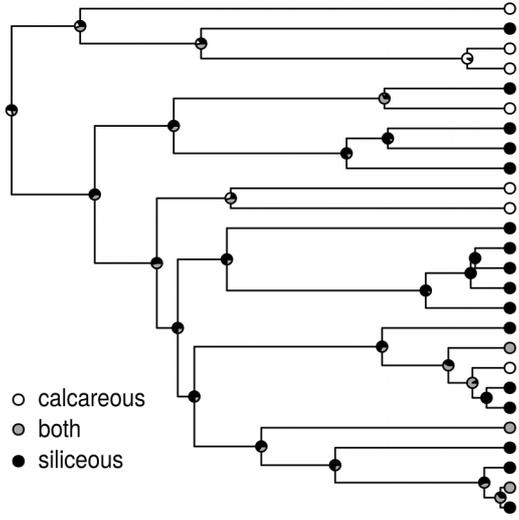
Appendices of chapter 2

Figure S3: The ancestral state reconstruction of (A) bedrock and (B) elevation niches for the six lineages. The reconstruction was performed using median posterior parameter estimate from respective ClaSSE models using the marginal reconstruction algorithm provided in R package HiSSE.

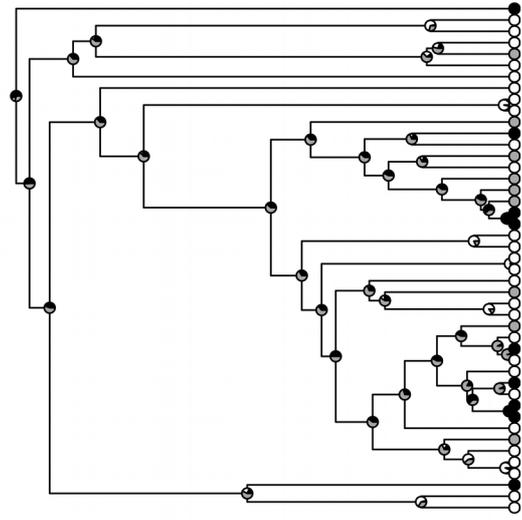
Appendices of chapter 2

A)

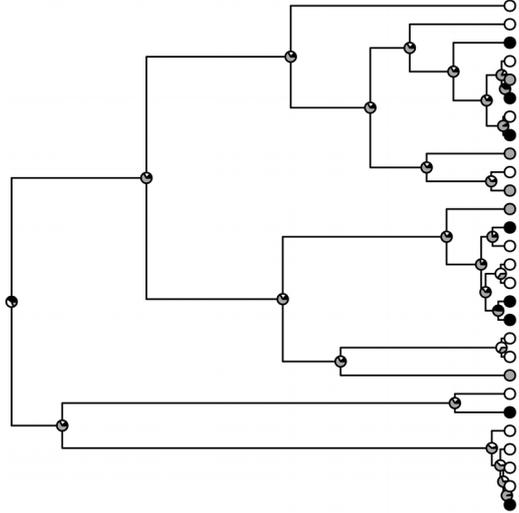
Androsace



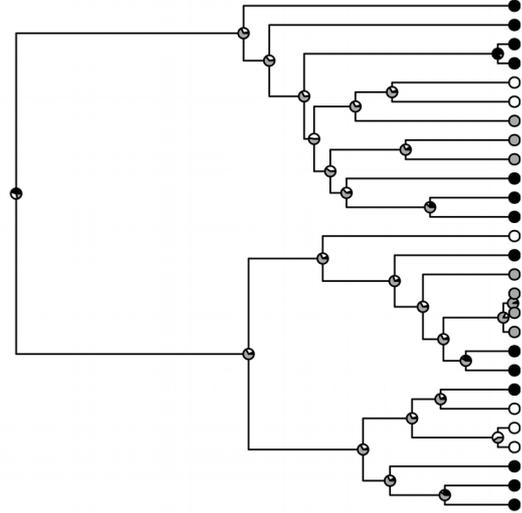
Campanula



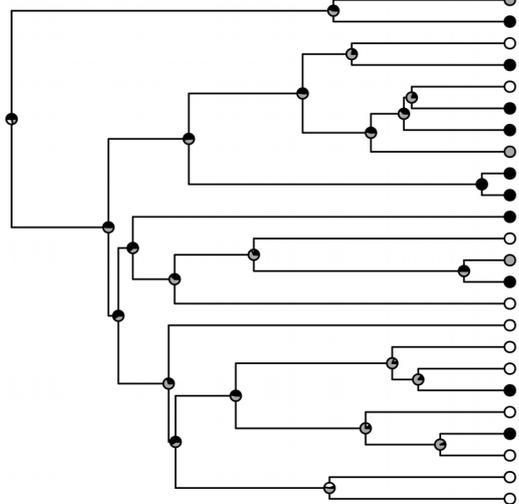
Gentiana



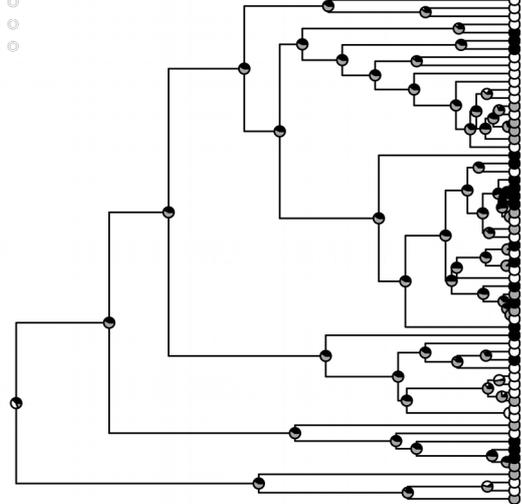
Phyteuma



Primula



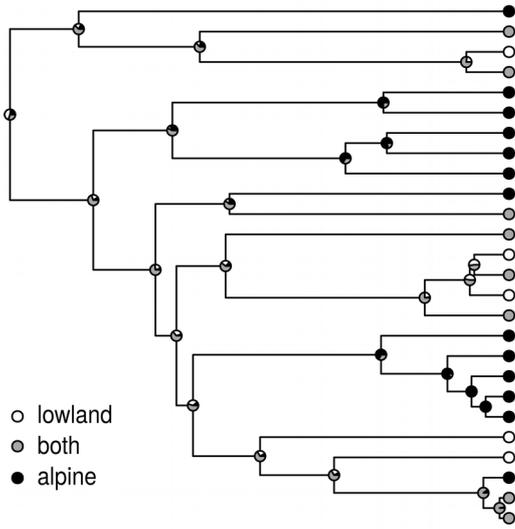
Saxifraga



Appendices of chapter 2

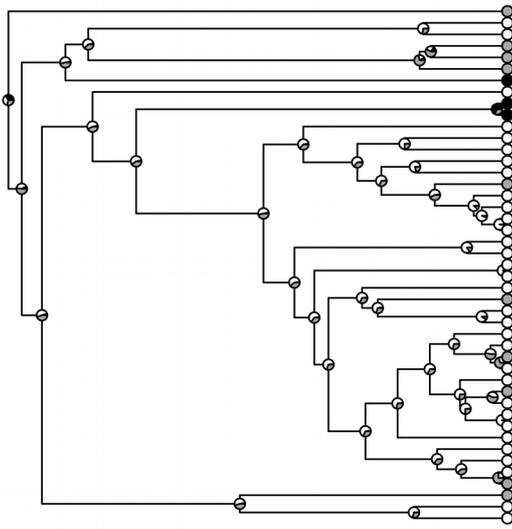
B)

Androsace

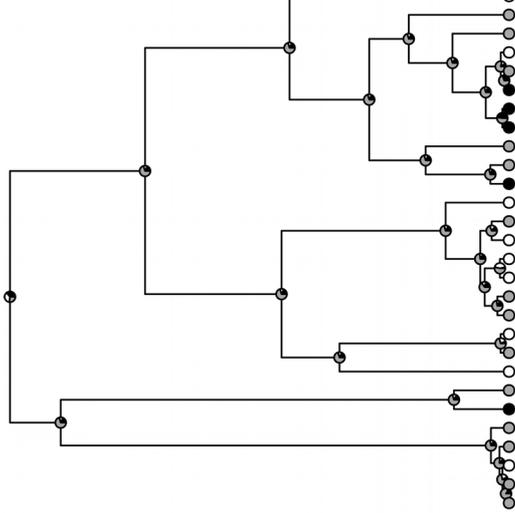


- lowland
- ◐ both
- alpine

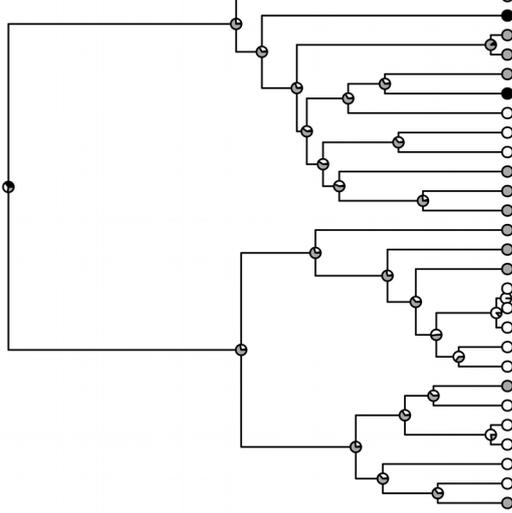
Campanula



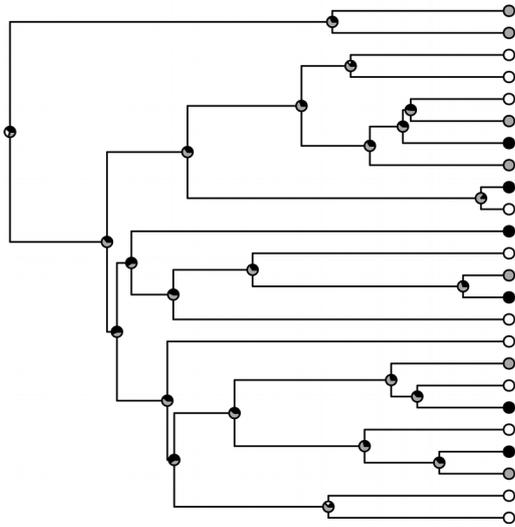
Gentiana



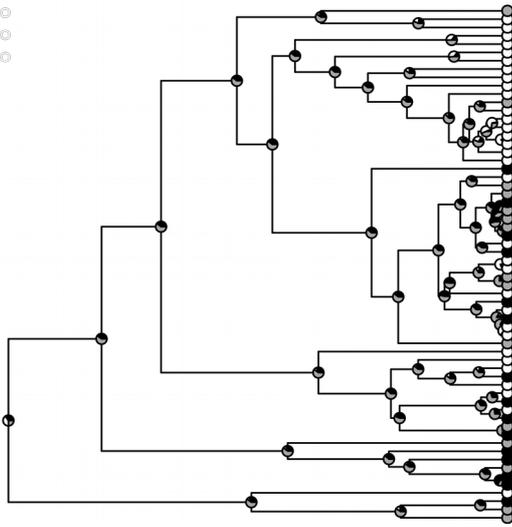
Phyteuma



Primula



Saxifraga



Appendix S1 Species sampling and taxonomic treatment

Our ingroup sampling included 38 samples of 26 species for *Androsace* sect. *Aretia* (total 29 species; counts excluding subgenus *Douglasia*, see below), 80 samples of 45 species for *Campanula* sect. *Heterophylla* (total 50 species), 31 samples of 28 species for *Gentiana* sections *Gentiana*, *Ciminalis* and *Calanthianae* (total 35 species; all 3 sections considered together, see below), 33 samples of 27 species for *Phyteuma* (total 27 species), 28 samples of 24 species for *Primula* sect. *Auriculata* (total 24 species), 69 samples of 62 species for *Saxifraga* sect. *Saxifraga* (total 86 species). Concerning outgroups, we included 49 outgroup samples of Campanulaceae, 14 outgroup samples of Gentianaceae, 95 outgroup samples of Primulaceae, and 49 and 6 outgroup samples of Saxifragaceae and Grossulariaceae, respectively. In total our dataset contains 492 samples. The large majority of dataset was collected in the field, but 3 samples come from recent (<20 years) herbarium specimens, and 21 samples come from individuals cultivated in botanical gardens from seeds collected in the field. Please see the accession tables in Dataset S1 for details.

We detail below the taxonomic literature and molecular phylogenetic studies that we took in account to generate the list of taxa that constitute each of the six study lineages. For all the groups, we took in account information from regional floristic literature: Flora Alpina (1), Flora Iberica (2), Flóra Slovenska (3), Wildpflanzen Siebenbürgens (4) and Flora Srbije (5). In addition to this, we took in account specific taxonomic literature for each of the lineages, specifically:

- For *Androsace* sect. *Aretia* we used Schneeweiss et al. 2004 (6), Schonswetter et al. 2009 (7), Boucher et al. 2012 (8), Schonswetter et al. 2015 (9) and Boucher et al. 2015 (10).
- For *Campanula* sect. *Heterophylla* we used Mansion et al. 2012 (11), Kovačić 2004 (12) and Fenaroli et al. 2013 (13). As this lineage is the most complex of ones included in this study from the taxonomic point of view, we established a specific collaboration with Kristýna Šemberová (co-author of this study) because she is currently conducting a detailed taxonomic revision of the species of this group. For controverted species, we included several individuals to check for species monophyly. In addition, a specific study integrating molecular and morphological data of

Appendices of chapter 2

Pyrenean species has been conducted for the taxonomic revision of controverted species endemic to this area (Roquet et al. in prep.).

- For *Gentiana* we used Favre 2015 (14), and specifically for section *Gentiana* we used Rossi 2011 (15) and for section *Calathianae* we used Hämmerli 2007 (16).
- For *Phyteuma* we used Schneeweiss et al. 2013 (17).
- For *Primula* sect. *Auriculata* we used Zhang and Kadereit 2004 (18) and Boucher et al. 2015(10).
- For *Saxifraga* sect. *Saxifraga* we used Vargas 2000 (19), Webb and Gornall 1989 (20) and Tkach et al. 2015 (21).

Where available and specifically for controverted taxa, we used more than one sample per species for constructing phylogenies. To turn the sample topologies into species trees, we applied several specific treatments:

- We selected the tips of paraphyletic species to conserve most recent splitting node: 5 species in our phylogenies (*Campanula jaubertiana*, *Gentiana lutea* ssp. *lutea*, *Androsace halleri* ssp. *halleri*, *Androsace cylindrica* ssp. *hirtella*, *Saxifraga glabella*) proved to be paraphyletic, i.e. with another species arising from within subtree of their samples. This phenomenon may be caused by incomplete lineage sorting, but may also reflects evolutionary reality of budding speciation within these lineages. In order to keep most realistic splitting date between such hierarchically organized species, we selected the samples of paraphyletic species in order to conserve most recent splitting node.
- We pruned the sample of *Campanula witasekiana* from Eastern Alps: *Campanula witasekiana* was a seemingly polyphyletic species in our dataset, with one sample from Dinaric mountains in Bosnia and another from Austrian Alps, lying in different parts of *Campanula* sect. *Heterophylla* phylogeny. Given the inconsistency between species descriptions by botanical communities in the Balkans and in Austria, and the fact that the type locality of species is in Bosnia, we assumed that the description *C. witasekiana* from eastern Alp is erroneous and decided to keep only the sample of *C. witasekiana* from Bosnia in our dataset.

Appendices of chapter 2

- We pruned multiple samples determined in field that could not be verified from herbarium sheets. Multiple samples collected in field as known complex species *Campanula rotundifolia*, *C. scheuchzerii*, *Gentiana verna* or *Primula auricula* and sequenced, could not be without doubts attributed to this, but also any other species. These *sensu lato* samples (marked by “sl” following scientific name in accession tables of family-level phylogenies) were thus not considered for construction of species trees.

In addition to this, we took two specific taxonomic treatments:

- We considered *Gentiana* sect. *Gentiana*, *Ciminalis* and *Calathianae* a single lineage. Monophyly of these lineages was previously suggested (14) . We found that these three European sections of *Gentiana* form monophyletic clade with posterior probability 1, so we treat the as a single lineage.
- We pruned the subgenus *Douglasia* from the *Androsace* sect. *Aretia* phylogeny: The subgenus *Douglasia* is a part of *Androsace* sect. *Aretia*, is monophyletic and endemic to north America. It thus represents a single speciation event in the European mountain system and the rest of diversification took place in North America, which is out of the geographic scope of our study.

Appendix S2 Gene regions selection and sequence processing

We aimed to work with all chloroplast coding regions that were present in our chloroplast genome reconstructions. To do so, we detected all open reading frames in our reconstructions and automatically compared them with already annotated genes available in GenBank, as implemented in Org.Annot procedure in Org.Asm (Coissac in prep., <https://git.metabarcoding.org/org-asm/org-asm/wikis/home>). The non-coding regions we used in our study were selected on the basis of their universal phylogenetic informativeness (22) or previous use in one of the focal lineages (11, 14–17, 21, 23, 24). The non-coding regions were identified and extracted based on their positioning relative to the neighboring coding regions.

Coding and non-coding regions of ingroup and outgroup samples were further filtered in order to minimize missing data in our family-level alignment matrices. The filtering resulted in 72 coding regions, of which 40 are shared across all study families, and 17 non-coding regions of which 5 are shared across all the families, see Dataset S2 for details. Samples with missing data for some of the regions were kept in specific cases: when they corresponded to the only representative of an ingroup species or to an outgroup species defining the dating node and at the same had so large portion of missing data that excluding missing regions would severely limit the resolution of the whole family phylogeny. To make sure that inclusion/exclusion had no impact on resulting topology or dating (with exception of dating-node outgroups), we reran phylogenetic reconstruction analyses without these samples and compared the resulting trees with all-samples phylogenies. The samples with missing data are indicated as such in the accession tables in Dataset S1.

For each of the four families, we aligned the coding regions gene-by-gene using MACSE (25) procedure acknowledging the triplet structure of codon alignment. All the positions were quality filtered by Gblocks (26) in codon setup, and concatenated together using FasConCat (27). The noncoding regions were aligned by Mafft (28), quality filtered by Gblocks in DNA setup and concatenated together.

Appendix S3 Phylogenetic analyses and fossil calibrations

Coding and non-coding alignments for each of the four families were used to estimate dated phylogenies in BEAST 2 (29). Non-coding alignments and each codon position in coding alignments were modeled with separate averaging-site models as implemented in bModeltest (30). We used Yule tree prior and log-normal clock model for each of the families. As fossil calibrations, we used uniform priors restricting age of at least two nodes in each family:

For Campanulaceae we restricted, in accordance with Mansion et al. 2011 (11):

- crown age of *C. carpatica*+*C. pulla*+*C. pyramidalis* to 16.5-80 Ma BP. The minimum bound is based on fossil seed of *C. palaeopyramidalis* (31). The maximum bound is based on inferred split between Rouseaceae and Campanulaceae (32).
- root of Campanulaceae to 16.5-80 Ma BP.

For Gentianaceae we restricted, in accordance with Favre et al. 2016 (14) for minimum bounds:

- crown age of *Gentiana*+*Gentianopsis*+*Lomatogonium*+*Swertia* to 5-100 Ma BP. The minimum bound is based on fossil seeds of *Gentiana* sect. *Cruciata* (33). The maximum bound is based on conservative estimate from Rybczinski et al. 2014 (34).
- root of Gentianaceae to 33.6-100 Ma BP. The minimum bound is based on fossil pollen of *Lisianthus* (35).

For Primulaceae we restricted, in accordance with Xing and Ree 2017 (36):

- crown age of *Cortusa*+*Hottonia*+*Primula*+*Soldanella* to 15.97-72 Ma BP. The minimum bound is based on fossil seeds of *Primula rosiae* (37). The maximum bound is based on primuloid fossil flower that cannot be attributed directly to Primulaceae (38).
- crown age of *Androsace*+*Cortusa*+*Douglasia*+*Hottonia*+*Primula*+*Soldanella* to 5.3-72 Ma BP. The minimum bound is based on fossil seeds of *Androsace* (39).
- crown age of *Anagallis*+*Androsace*+*Cortusa*+*Douglasia*+*Hottonia*+*Lysimachia*+*Primula*+*Soldanella*+*Trientalis* to 28-72 Ma BP. The minimum bound is based on fossil seeds of *Lysimachia angulata* (40).

Appendices of chapter 2

- root of Primulaceae to 48.6-72 Ma BP. The minimum bound is based on fossil *Ardisia* (41).

For Saxifragaceae+Grossulariaceae, we restricted, in accordance with Ebersbach et al. 2017 (42) for minimum bounds:

- crown age of *Ribes* to 14.5-125 Ma BP. The minimum age is based on fossil leaves of *Ribes webbia* attributable to *Ribes* group *Calobotrya* (43). The maximum age is based on earliest angiosperm fossil evidence (44).
- root of Saxifragaceae+Grossulariaceae to 42-125 Ma BP. Minimum age is based on fossil leaves of *Ribes axelrodii* (43).

The BEAST analysis for each of the families was run for 100M iterations, with 30% burn-in, and the convergence was evaluated both visually and checking whether ESS>100 for all the parameters. The full tree posterior was used for building the maximum credibility trees, and also reduced posterior of 100 randomly selected trees, which were used for subsequent analyses.

Appendix S4 Compilation of ecological and geographic information

Information about bedrock, elevational niche and geographical occurrences was obtained from regional floristic literature: Flora Alpina (1) for the Alps, Flora Iberica (2) for species of Iberian peninsula, Flóra Slovenska (3) and Wildpflanzen Siebenbürgens (4) for Carpathians, and Flora Srbije (5) for the Balkans. Additional information about *Androsace* (7, 9, 45), *Phyteuma* (17) and *Saxifraga* (20) was obtained from literature about ecology and biogeography of these lineages. The information about *Campanula* was in part compiled based on knowledge of herbarium specimens and local taxonomic literature. In several cases, we contacted local taxonomists and made categorization based on provided information. Where possible, the information from floristic literature was compared and supplemented with previously published (10) or publicly accessible (GBIF, <https://www.gbif.org/>) sources of point occurrence data.

The calcareous bedrock niche was defined by regular presence of species on calcareous, dolomitic or ultrabasic bedrocks, the siliceous bedrock niche was defined by regular presence of species on any bedrocks that do not fall in the calcareous category. E.g. in case of species covered by Flora Alpina (1), species inhabits calcareous niche if regular presence is indicated either from limestones (ca) or serpentinites (ser), and siliceous niche if regular presence is indicated from silicates (si), intermediate substrates (ca/si) or volcanic rocks (bas).

The mid-elevation niche was defined by regular presence of species in habitat below timberline, i.e. up to subalpine elevational zone sensu Flora Alpina (1). High elevation niche was defined by regular presence above timberline, i.e. in alpine and nival zone sensu Flora Alpina (1). In case of floristic information that did not refer to elevational belts terminology or timberline, specifically Flora Iberica (2), we marked species presence in mid-elevation niche if the species was inhabiting habitats lower than 200 m below regional timberline and in high elevation niche if the species was inhabiting habitats 200 m above regional timberline. No such treated species was restricted to the range ± 200 m around timberline.

As the small scale geographic regions, we used operative geographic units in Flora Alpina (1) for the Alps, and mountain regions based on Körner et al. 2017 (46) for other European mountains, with subsequent modifications that better reflect structuring of biogeographic information in local floristic literature.

Appendices of chapter 2

Specifically, we merged:

- Pennines and Cambrian Mountains to England
- MacGillycuddy's Reeks and Wicklow Mountains to Ireland
- Vosges, Black Forest and Jura Mountains to Rhine Valley
- Grampian Mountains, Northwestern Highlands and Southern Uplands to Scotland

We joined:

- Basque Mountains and Tras-os-montes to Cantabrian Mountains
- Korab, Sar Planina and Jablanica to Dinaric Alps
- Kontovounia, Tayetos Oros, Crete and Parnon Oros to Peloponnisos
- Grike, Mali i Gjere, Mount Nemercke, Mount Olympus and Mount Othris to Pindos
- Osogovo, Maleshevo, Belasica, Voras Mountains, Rila, Pirin and Pangaion Hills to Rhodope mountains

We renamed:

- Balkan Mountains to Stara Planina

We divided:

- Carpathian Mountains to Western Carpathians and South Eastern Carpathians along the border between Slovakia and Ukraine

We newly defined:

- lower mountain ranges in Czechia not included into Western Carpathians as Sudetes

In addition to this, whole Africa, Middle East (including Caucasus) and Arctic (including Siberian mountains) were represented by one region each, accommodating species with ranges extending out of Europe. Species occurring in lowlands out of mountain regions were always attributed to geographically closest mountain region.

Appendices of chapter 2

The 5 major geographic regions were defined as sets of small scale geographic regions, specifically:

- Alps: Alps, Massif central and Rhine valley
- Apeninnes: Apennines, Corsica, Sardinia and Sicily
- Balkans: Dinaric Alps, Rhodopes, Stara planina, Peloponnisos, Pindos and Middle East (the latter region was merged within the Balkans because no focal endemic species and strong floristic connections)
- Carpathians: Carpathians, Sudetes, all northern European regions and Arctic (the latter region was merged within the Carpathians because very few focal endemic species and strong floristic connections)
- Iberian mountains: all mountains on Iberian Peninsula, Mallorca, Madeira and northern Africa (the latter two regions were merged within Iberian mountains one because very few focal endemic species and strong floristic connections)

For table of ecological niches and regions for each species, see Dataset S3.

Appendix S5 – The PhyloAlps consortium

- Laboratoire d'Ecologie Alpine (LECA), Université Grenoble Alpes, CNRS, FR-38000 Grenoble, France
- Swiss Federal Research Institute WSL, CH-8903 Birmensdorf, Switzerland
- Jardin Alpine de Lautaret, Université Grenoble Alpes, FR-38000 Grenoble, France
- Conservatoire Botanique National Alpin, Domaine de Charance, FR-05000 Gap, France
- Conservatoire Botanique National Méditerranéen, FR-83400 Hyères, France
- Parc National des Ecrins, FR-05000 Gap, France
- Parc National de la Vanoise, FR-73000 Chambéry, France
- Parc National du Mercantour, FR-06006 Nice Cedex 1, France
- Genoscope, Centre National de Séquençage, FR-91057 Evry Cedex, France
- Faculty of Biology and Geology, Babeş-Bolyai University, RO-400015 Cluj-Napoca, Romania

Dataset S1 Accession tables

NOTE: This dataset is in preliminary version.

<https://github.com/smyckaj/Thesis-Datasets/blob/master/6%20clades/Dataset%20S1%20Accessions.csv>

Dataset S2 Genomic regions

<https://github.com/smyckaj/Thesis-Datasets/blob/master/6%20clades/Dataset%20S2%20Genomic%20regions.csv>

Dataset S3 Ecological and geographic information

<https://github.com/smyckaj/Thesis-Datasets/blob/master/6%20clades/Dataset%20S3%20Niches%20and%20geography.csv>

Dataset S4 Multi-lineage time-dependent diversification model script

The multi-lineage diversification models were implemented as a modification of `fit_bd` and `fit_env` functions in RPANDA 1.3. A likelihood function optimized within these functions was constructed as a product of functions for multiple input phylogenies.

<https://github.com/smyckaj/Thesis-Datasets/blob/master/6%20clades/Dataset%20S4%20Multilineage%20TDD.R>

Dataset S5 Multi-regions GeoSSE script

A generalization of GeoSSE for multiple regions was implemented as a helper functions identifying different types of biogeographic processes from provided table of state numbers and biogeographic regions using set relations. The output of this function is a list of restriction rules that can be inputted into standard ClaSSE procedure. Optionally, the function allows to restrict the parameters of certain type (e.g. dispersal parameters) into one value or several values based on receptor region. Note that in addition to normal GeoSSE, our function can

Appendices of chapter 2

also define founder-event speciations, but this option was turned off in analyses performed in this study.

<https://github.com/smyckaj/Thesis-Datasets/blob/master/6%20clades/Dataset%20S5%20Multiregion%20GeoSSE.R>

Appendices of chapter 2

References

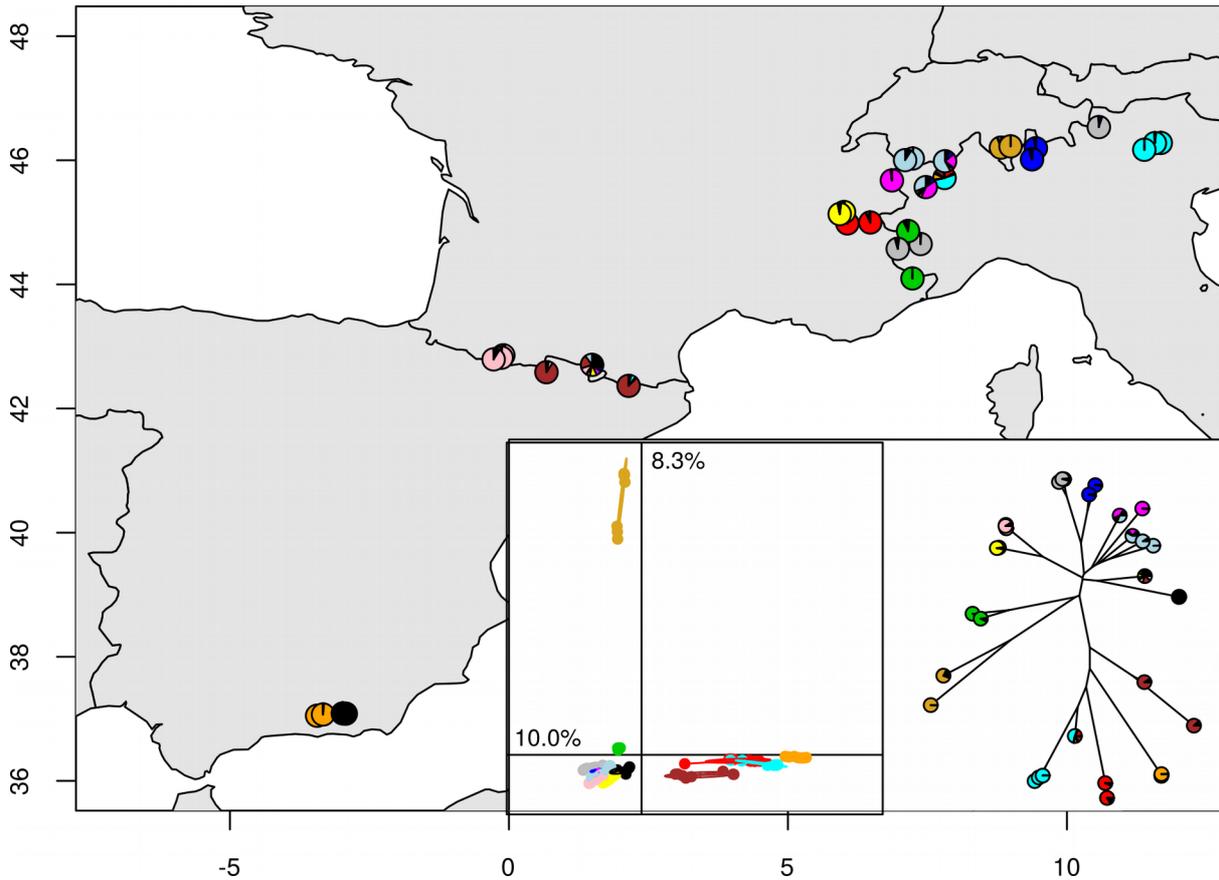
1. Aeschimann D, Lauber K, Moser DM, Theurillat JP (2004) *Flora Alpina* (Editions Belin).
2. Castroviejo S ed. (2012) *Flora Iberica* (Real Jardín Botánico CSIC).
3. Goliášová K, Michalková E eds. (2012) *Flóra Slovenska* (Vydavateľstvo Slovenskej akadémie vied).
4. Speta E, Rákosy L (2010) *Wildpflanzen Siebenbürgen* (Naturhistorisches Museum Wien).
5. Sarić MR ed. (1992) *Flora Srbije* (Srpska akademija nauka i umetnosti).
6. Schneeweiss GM, Schönswetter P, Kelso S, Niklfeld H (2004) Complex biogeographic patterns in Androsace (Primulaceae) and related genera: Evidence from phylogenetic analyses of nuclear internal transcribed spacer and plastid trnL-F sequences. *Syst Biol* 53(6):856–876.
7. Schönswetter P, Schneeweiss GM (2009) *Androsace komovensis* sp. nov., a long mistaken local endemic from the southern Balkan Peninsula with biogeographic links to the Eastern Alps. *Taxon* 58(2):544–549.
8. Boucher FC, et al. (2012) Reconstructing the origins of high-alpine niches and cushion life form in the genus *Androsace* s.l. (Primulaceae). *Evolution* 66(4):1255–1268.
9. Schönswetter P, Magauer M, Schneeweiss GM (2015) *Androsace halleri* subsp. *nuria* Schönsw. & Schneew. (Primulaceae), a new taxon from the eastern Pyrenees (Spain, France). *Phytotaxa* 201(3):227–232.
10. Boucher FC, Zimmermann NE, Conti E (2016) Allopatric speciation with little niche divergence is common among alpine Primulaceae. *J Biogeogr* 43(3):591–602.
11. Mansion G, et al. (2012) How to Handle Speciose Clades? Mass Taxon-Sampling as a Strategy towards Illuminating the Natural History of *Campanula* (Campanuloideae). *PLoS One* 7(11).
12. Kovacic S (2004) The genus *Campanula* L. (Campanulaceae) in Croatia, circum-Adriatic and west Balkan region. *Acta Bot Croat* 63(2):171–202.
13. Fenaroli F, Pistarino A, Peruzzi L, Cellinese N (2013) *Campanula martinii* (Campanulaceae), a new species from northern Italy. *Phytotaxa* 111(1):27–38.
14. Favre A, et al. (2016) Out-of-Tibet: the spatio-temporal evolution of *Gentiana* (Gentianaceae). *J Biogeogr* 43(10):1967–1978.
15. Rossi M (2011) *Taxonomy, phylogeny and reproductive ecology of Gentiana lutea* L. Dissertation (University in Bologna).
16. Hämmerli M (2007) *Molecular Aspects in Systematics of Gentiana Sect. Calathianae Froel.* Dissertation (Université de Neuchâtel).
17. Schneeweiss GM, et al. (2013) Molecular phylogenetic analyses identify Alpine differentiation and dysploid chromosome number changes as major forces for the evolution of the European endemic *Phyteuma* (Campanulaceae). *Mol Phylogenet Evol* 69(3):634–652.
18. Zhang LB, Kadereit JW (2004) Classification of *Primula* sect. *Auricula* (Primulaceae) based on two molecular data sets (ITS, AFLPs), morphology and geographical distribution. *Bot J Linn Soc* 146(1):1–26.
19. Vargas P (2000) Plant Systematics and Evolution A phylogenetic study of *Saxifraga* sect. *Saxifraga* (Saxifragaceae) based on nrDNA ITS sequences. *Plant Syst Evol* 223:59–70.
20. Webb DA, Gornall RJ (1989) *Saxifrages of Europe* (Timber Press).
21. Tkach N, et al. (2015) Molecular phylogenetics, morphology and a revised classification of the complex genus *Saxifraga* (Saxifragaceae). *Taxon* 64(6):1159–1187.
22. Shaw J, et al. (2014) Chloroplast DNA sequence utility for the lowest phylogenetic and phylogeographic inferences in angiosperms: The tortoise and the hare IV. *Am J Bot* 101(11):1987–2004.
23. Hungerer KB, Kadereit JW (1998) The phylogeny and biogeography of *Gentiana* L. sect. *Ciminalis* (Adans.) Dumort.: A historical interpretation of distribution ranges in the European high mountains. *Perspect Plant Ecol Evol Syst* 1(1):121–135.
24. Roquet C, Boucher FC, Thuiller W, Lavergne S (2013) Replicated radiations of the alpine genus *Androsace* (Primulaceae) driven by range expansion and convergent key innovations. *J Biogeogr* 40(10):1874–1886.

Appendices of chapter 2

25. Ranwez V, Harispe S, Delsuc F, Douzery EJP (2011) MACSE: Multiple alignment of coding SEquences accounting for frameshifts and stop codons. *PLoS One* 6(9).
26. Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17(4):540–52.
27. Kück P, Meusemann K (2010) FASconCAT: Convenient handling of data matrices. *Mol Phylogenet Evol* 56(3):1115–1118.
28. Katoh K, Kuma KI, Toh H, Miyata T (2005) MAFFT version 5: Improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res* 33(2):511–518.
29. Bouckaert R, et al. (2014) BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Comput Biol* 10(4):1–6.
30. Bouckaert R, Drummond AJ (2017) bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evol Biol* 17(1):1–11.
31. Łańcucka-Środoniowa M (1979) Macroscopic plant remains from the freshwater Miocene of the Nowy Sącz Basin (West Carpathians, Poland). *Acta Palaeobot* 1(20):3–117.
32. Bell CD, Soltis DE, Soltis PS (2010) The age and diversification of the angiosperms re-revisited. *Am J Bot* 97(8):1296–1303.
33. Pirie MD, Litsios G, Bellstedt DU, Salamin N, Kissling J (2015) Back to gondwanaland: Can ancient vicariance explain (some) Indian ocean disjunct plant distributions? *Biol Lett* 11(6).
34. Rybzcynki JJ, Davey MR, Mikula A eds. (2014) *The Gentianaceae - Volume 1: Characterization and Ecology* (Springer).
35. Graham A (1984) Lisianthus pollen from the Eocene of Panama. *Ann Missouri Bot Gard* 71(4):987–993.
36. Xing Y, Ree RH (2017) Uplift-driven diversification in the Hengduan Mountains, a temperate biodiversity hotspot. *Proc. Natl. Acad. Sci. U.S.A.* 114 (17), 3444–3451.
37. Czaja A (2003) *Paläokarpologische Untersuchungen von Taphozöosen des Unter- und Mittelmiozäns aus dem Braunkohlentagebau Berzdorf/Oberlausitz (Sachsen)*. (Schweizerbart).
38. Friis EM, Pedersen KR, Crane PR (2010) Cretaceous diversification of angiosperms in the western part of the Iberian Peninsula. *Rev Palaeobot Palynol* 162(3):341–361.
39. Dorofeev PI, Takhtajan AL (1963) Primulaceae. *Oznovij Paleontologii* (Izdavatelstvo Akademii nauk SSSR), pp 517–518.
40. Dorofeev PI (1963) *Treticnye flory zapadnoj Sibiri* (Izdavatelstvo Akademii nauk SSSR).
41. Collinson ME (1984) *Fossil plants of the London Clay* (Wiley-Blackwell).
42. Ebersbach J (2017) In and out of the Qinghai-Tibet Plateau: divergence time estimation and historical biogeography of the large arctic-alpine genus *Saxifraga* L. *J Biogeogr* 44(4):900–910.
43. Hermsen E (2005) *The fossil record of Iteaceae and Grossulariaceae in the cretaceous and tertiary of the United States and Canada*. Dissertation (Cornell University).
44. Hughes NF (1994) *The Enigma of Angiosperm Origin* (Cambridge Paleobiology Series).
45. Schneeweiss GM, Schonswetter P (2010) The wide but disjunct range of the European mountain plant *Androsace lactea* L. (Primulaceae) reflects Late Pleistocene range fragmentation and post-glacial distributional stasis. *J Biogeogr* 37(10):2016–2025.
46. Körner C, et al. (2017) A global inventory of mountains for bio-geographical applications. *Alp Bot* 127(1):1–15.

APPENDICES OF CHAPTER 3

Figure S1: Range-wide population structure of *A. argentea* after subsampling the number of Belledonne, Ecrins and Mont Blanc populations to two. **A** Geographic distribution of sampled populations. The pies indicate positions of populations and reflect proportional assignment of populations to the respective snmf clusters. **B** The first and second axis of principal component analysis of *A. argentea* individuals. Colors are attributed based on maximum assignment to sNMF clusters. **C** Neighbor joining tree of *A. argentea* populations.



Appendices of chapter 3

Figure S2: The relationship between cross-entropy criterion and number of population clusters (K) in the snmf analysis.

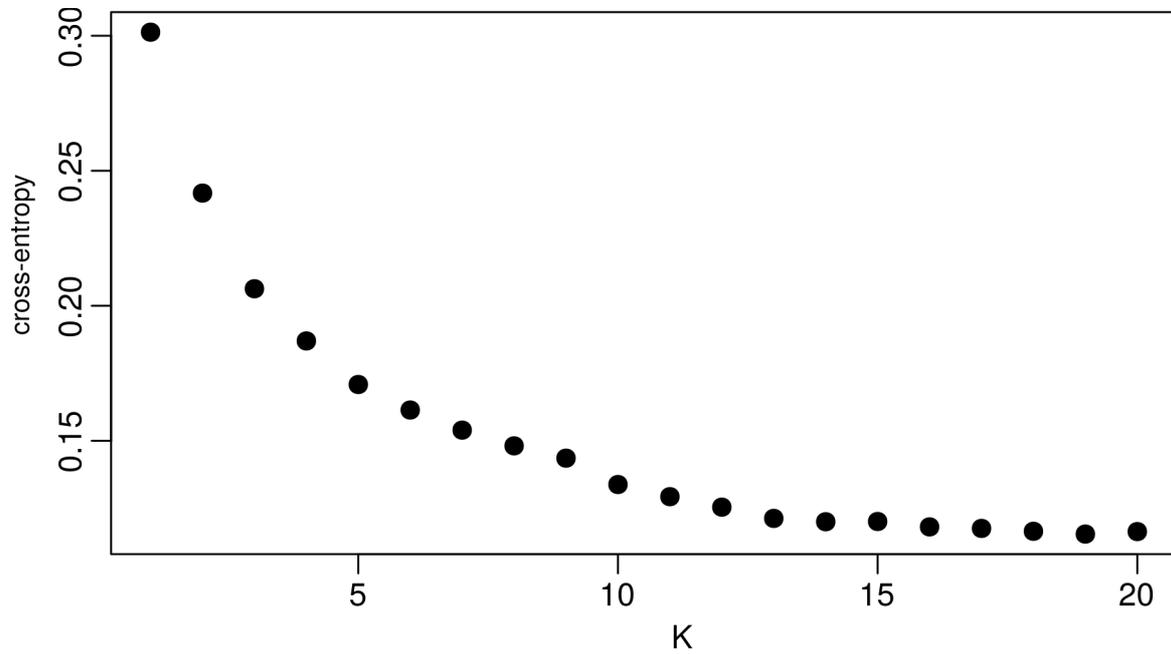
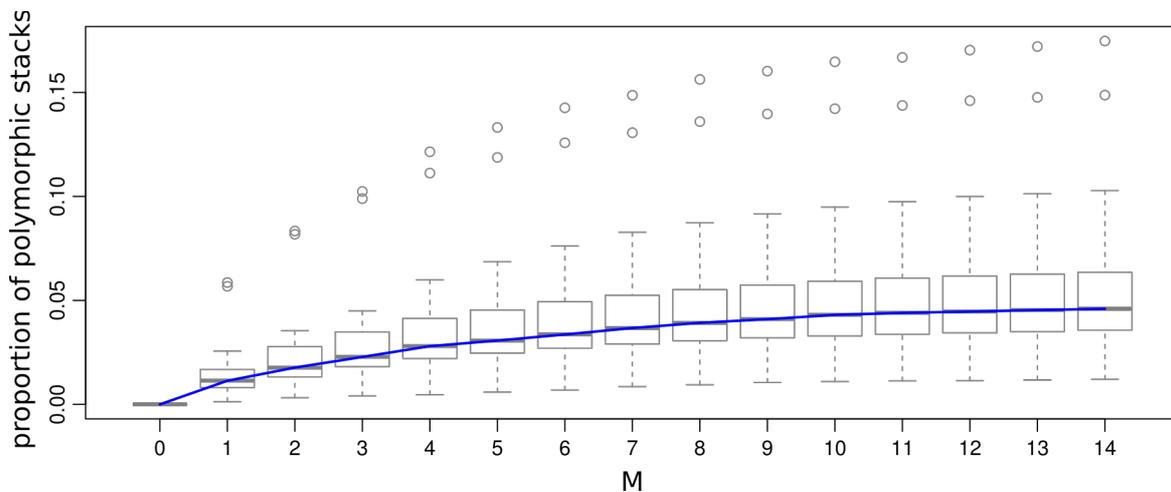


Figure S3: The relationship between proportion of polymorphic stacks and clustering parameter M in the Stack *de novo* assembly procedure.



Appendices of chapter 3

Dataset S4 Population accession table

<https://github.com/smyckaj/Thesis-Datasets/blob/master/Argentea/Dataset%20S4%20Populations%20list.csv>

Dataset S5 Sample list

<https://github.com/smyckaj/Thesis-Datasets/blob/master/Argentea/Dataset%20S5%20Sample%20list.csv>