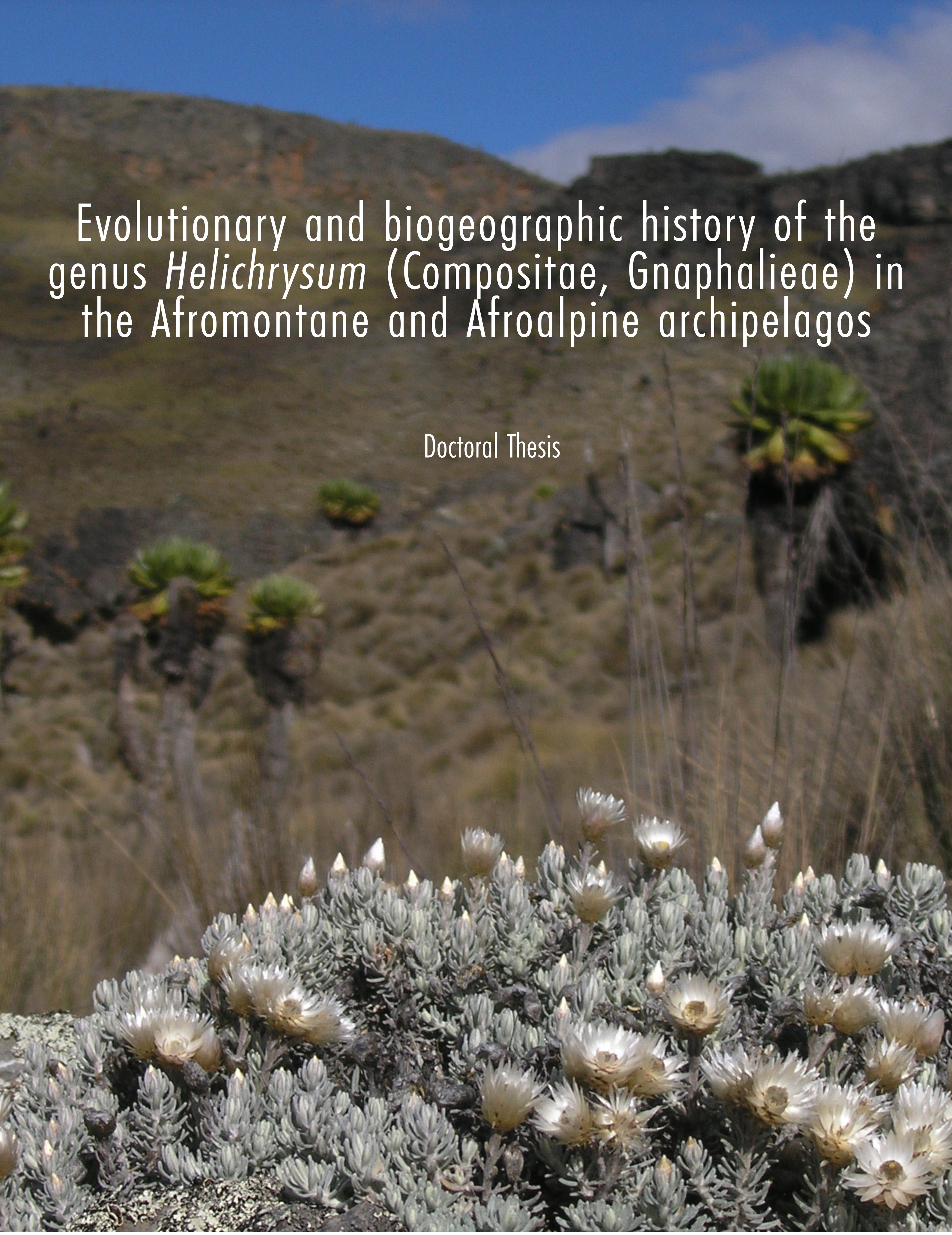
UAB
Universitat Autònoma de Barcelona

EVOLUTIONARY AND BIOGEOGRAPHIC HISTORY OF THE GENUS HELICHRYSUM (COMPOSITAE, GNAPHALIEAE) IN THE AFROMONTANE AND AFROALPINE ARCHIPELAGOS

CARME BLANCO GAVALDÀ



Evolutionary and biogeographic history of the genus *Helichrysum* (Compositae, Gnaphalieae) in the Afromontane and Afroalpine archipelagos

Doctoral Thesis

CARME BLANCO GAVALDÀ

2025

EVOLUTIONARY AND BIOGEOGRAPHIC HISTORY OF THE GENUS *HELICHRYSUM* (COMPOSITAE, GNAPHALIEAE) IN THE AFROMONTANE AND AFROALPINE ARCHIPELAGOS

UNIVERSITAT AUTÒNOMA DE BARCELONA · FACULTY OF BIOSCIENCIES
DEPARTMENT OF ANIMAL BIOLOGY, PLANT BIOLOGY AND ECOLOGY
UNIT OF BOTANY

INSTITUT BOTÀNIC DE BARCELONA
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS · AJUNTAMENT DE BARCELONA

Doctoral Thesis

Carme Blanco Gavaldà

Barcelona 2025
Doctorate Program in Plant Biology and Biotechnology

EVOLUTIONARY AND BIOGEOGRAPHIC HISTORY OF THE GENUS *HELICHRYSUM* (COMPOSITAE, GNAPHALIEAE) IN THE AFROMONTANE AND AFROALPINE ARCHIPELAGOS

Dissertation presented by Carme Blanco Gavaldà in fulfillment of the
requirements for the degree of Doctor
With the approval of:

Dr. Mercè Galbany Casals (supervisor and tutor)	Dr. Cristina Roquet Ruíz (supervisor)	Dr. Alfonso Susanna de la Serna (supervisor)

Universitat Autònoma de Barcelona
Doctorate Program in Plant Biology and Biotechnology

This thesis was funded by the Spanish Ministry of Science, Innovation and Universities under the project “Afromontane and Afroalpine archipelagos: unravelling the drivers of diversification and biogeographic connections of two species rich Compositae radiations (AFROMONT)”, grant number PD2019-105583GB-C22/AEI/10.13039/501100011033 and the Catalan government with “Ajuts a grups consolidats” 2021SGR00315 and Ph.D. FI-AGAUR grant 2022FI_B00150.

Cover images throughout the thesis by Mercè Galbany Casals.

1. Main cover: *Helichrysum newii*.
2. Acknowledgments cover: *Helichrysum chionosphaerum*.
3. Table of Contents cover: *Helichrysum xylocladum*.
4. General Introduction cover: *Helichrysum heterotrichum.*
5. Objectives cover: *Helichrysum ellipticifolium.*
6. Structure of the Thesis cover: *Pseudognaphalium jaliscense.*
7. Chapter 1 cover: *Helichrysum betsiliense.*
8. Chapter 2 cover: *Helichrysum hirtum.*
9. Chapter 3 cover: *Helichrysum marojejyense.*
10. General Discussion cover: *Helichrysum helothamnus.*
11. Conclusions cover: *Helichrysum geniorum.*

Acknowledgments

Agraïments / Acknowledgments

Fa gairebé nou anys vaig posar el peus per primera vegada a l'Autònoma, en aquell moment no podia imaginar el que m'esperava. Vull començar expressant el meu profund agraïment a totes les persones que, d'una manera o altra, han fet possible el desenvolupament d'aquesta tesi. La llista és molt llarga i no em voldria deixar a ningú, però us tinc ben presents.

Cristina, Juancho, Llorenç i Mercè, gràcies per transmetre aquesta passió i compromís amb la botànica. Ràpidament us vau convertir en els meus referents i em sento molt afortunada d'haver estat alumna vostra. Gràcies per confiar en mi des del primer dia!

Mercè, gràcies per ser la meva mentora i gairebé "mare" dins del món de la recerca. Juntament amb la Cristina, us dec l'oportunitat de fer el TFM amb vosaltres i continuar amb la tesi dins del projecte AFROMONT, apassionant i ambiciós a parts iguals. La vostra paciència ha estat infinita: heu suportant els meus correus eterns, la meva tossudesa i els meus constants dubtes. Gràcies per acollir-me, guiar-me, recordar-me els meus punts forts i animar-me a perdre la por a lo desconegut.

També us agraeixo haver-me portat a l'Institut Botànic, on he tingut la immensa sort de conèixer la Núria i l'Alfonso. Gràcies a vosaltres també per confiar en mi des del primer moment, formant plegats un gran equip. I, per descomptat, la Lucía, amb qui he parlat i rigut molt, però també he compartit intenses hores de treball. No només has estat la meva companya d'aventures durant aquests quatre anys, també t'has convertit en una gran amiga, consellera i font d'inspiració. Al Genís i a l'Òscar, que vau arribar més tard, però amb una energia i dedicació insuperables. Ha estat un plaer treballar al vostre costat. Quan penso en el futur de la botànica, us veig a vosaltres. En quatre anys passen moltes coses, bones i no tan bones, esperades i inesperades... Quan tot sembla anar malament, és fàcil voler llençar la tovallola, però també hem format un gran equip en aquest aspecte, cooperatiu, flexible i adaptable (com *Helichrysum*!). Tot i tenir personalitats molt diferents, ens hem complementat i ajudat en els moments més complicats. Sense cap dubte, puc dir que sou l'equip més humà amb qui hauré treballat mai.

Among the unexpected twists in this journey, Naturalis marked a true turning point. I am deeply grateful to Renske Onstein for welcoming me into her group, offering guidance and introducing me to Luís Valente. There, I had the privilege of meeting excellent researchers and collaborators from whom I gained valuable insights into research and evolutionary biology. Those months were undoubtedly challenging, yet they turned out to be an incredibly rewarding and enriching experience on many levels.

Semblava que això havia de durar per sempre, però ara començo a ser conscient que s'acaba i no puc imaginar com serà el meu futur sense vosaltres, fent una altra cosa. Al vostre costat he complert un dels meus somnis i alhora, he crescut com a persona.

Gràcies per tot!

Table of Contents

Table of Contents

General Introduction

General Introduction

1. Biodiversity: its origins and distribution

The term Biodiversity comprises the variety of forms on Earth across multiple scales, from genetic variations within populations to complex communities that constitute ecosystems (Wilson, 1988). This diversity includes both current and past forms of life, as well as the complex interactions between species and their environments, reflecting millions of years of evolution (Geeta et al., 2014). As the pioneering evolutionary biologist Theodosius Dobzhansky stated: "Nothing in biology makes sense except in the light of evolution." From this perspective, understanding the processes that have shaped and sustained biodiversity over millennia is essential for its conservation and the effective management of ecosystems. This urgency is heightened by the accelerated loss of biodiversity, driven by human-induced factors, such as habitat loss, invasive species, climate change, pollution and overexploitation. These threats have led to an unprecedented global biodiversity crisis (Sala et al., 2000; Pimm et al., 2014). The consequences of biodiversity loss are profound, directly impacting humanity, as biodiversity and ecosystems are an essential source of resources and services for human well-being (Cardinale et al., 2012).

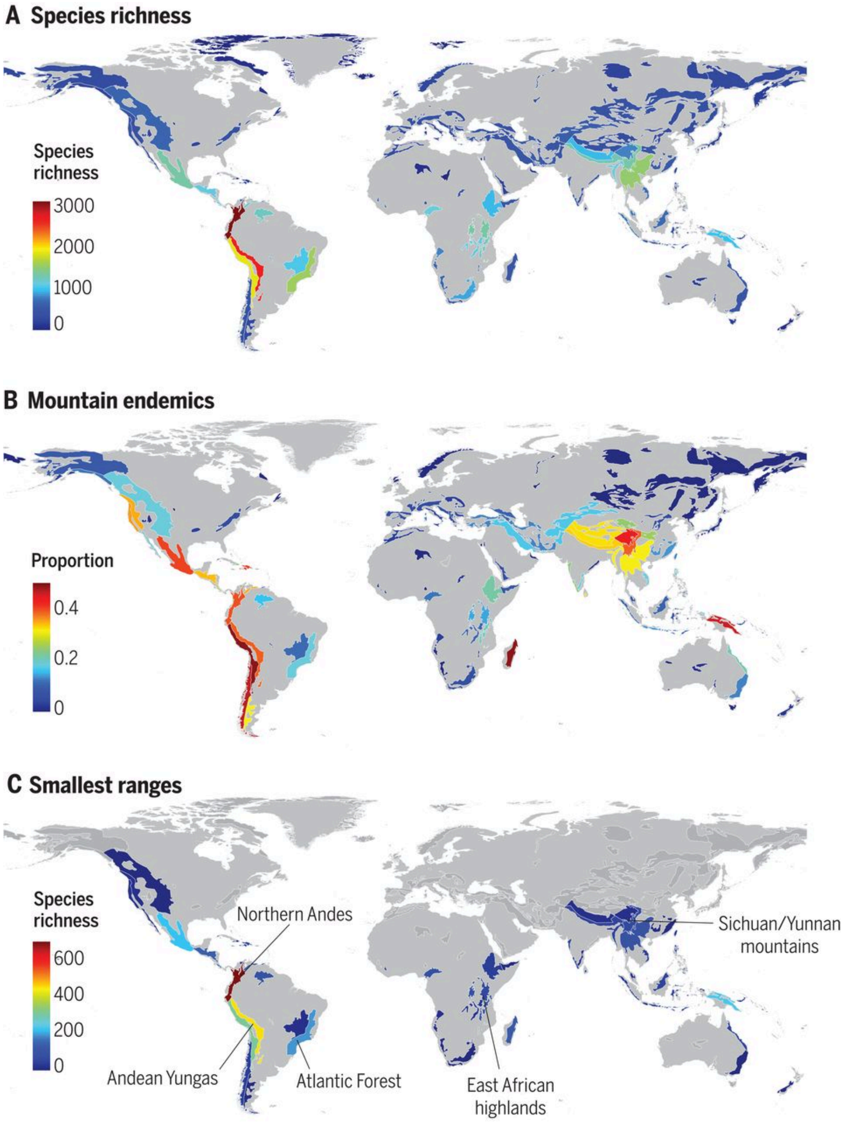
Biodiversity is not evenly distributed across the planet. Instead, it is concentrated in "biodiversity hotspots", which are highly threatened, irreplaceable and vulnerable regions harboring disproportionally high levels of endemism and species richness (Myers et al., 2000; Mittermeier et al., 2005). Understanding the geological, paleoclimatic and evolutionary origins of these hotspots is crucial, as this knowledge can be transferred into strategies to preserve ecosystems across different spatial and temporal scales (Willis et al., 2007). Among the most significant biodiversity hotspots are islands and island-like systems such as mountains, which serve as natural laboratories for studying evolutionary processes due to their relative isolation, ecological heterogeneity and concentration of diverse habitats within compact areas (Warren et al., 2015; Rahbek et al., 2019a; Muellner-Riehl et al., 2024).

Islands and mountains frequently host evolutionary radiations, driven by a combination of immigration processes (dispersal and establishment, related to their age, area and distance from source population), local diversification and ecological opportunities (Losos & Ricklefs, 2009). While dispersal plays a crucial role in the initial stages of colonization, once populations are established, in-situ diversification through adaptive and non-adaptive radiations may occur (Rundell & Price, 2009; Thomas et al., 2023).

If we define radiations as the rapid diversification of a single lineage into multiple species (Schluter, 2000), adaptive radiations involve the evolution of multiple descendant species that are ecologically differentiated, i.e. diversification is mainly driven by ecological speciation (Cerca et al., 2023). Classic examples include radiations on the Hawaiian archipelago or the Galápagos, where lineages rapidly diversified to exploit unoccupied ecological niches (Losos & Mahler, 2010). Mountain systems, with their sharp elevational gradients, also provide a spectrum of ecological conditions that can facilitate this process (Bentley et al., 2014; Hughes & Atchison, 2015). In contrast, non-adaptive radiations, often associated with dispersal and allopatric speciation, occur when species diversify primarily due to geographical isolation rather than ecological differentiation (Rundell & Price, 2009). In mountain systems, this can occur at multiple scales, within a single mountain due to complex topography or between mountains due to isolation. This is especially true for some lineages with advantageous preadaptations, in which case migrating might be easier and quicker than adapting to new conditions (Donoghue, 2008; Edwards & Donoghue, 2013). However, mounting evidence supports the idea that diversification in these systems is rarely purely adaptive or non-adaptive, but rather a mix of both, although the relative contribution of each depends on the studied system (Rundell & Price, 2009; Nürk et al., 2020; Thomas et al., 2023).

Diversification rates in islands and mountains are shaped by the intricate interplay of abiotic and biotic factors, influencing speciation and extinction dynamics (see Nürk et al., 2020 and references therein). Extrinsic factors such as past climatic oscillations (e.g. Pleistocene glaciation cycles; Nevado et al., 2018) and geological dynamics (e.g. mountain uplift, Antonelli et al., 2018; volcanism, Gillespie, 2007) promote population fragmentation and secondary contact, creating opportunities for diversification (Rahbek et al., 2019b). Meanwhile, important intrinsic factors such as key innovations (Miller et al., 2023), hybridization (Abbott et al., 2013) and polyploidization (Combrink et al., 2025) generate novel genetic and functional diversity, enhancing the adaptive potential of lineages in these systems.

Despite significant progress, a fundamental question in evolutionary biology remains unresolved: why do different biogeographic regions harbor contrasting levels of biodiversity and endemism? Understanding large- and small-scale drivers of biodiversity patterns remains a challenge, with a consensus yet to be reached (Ricklefs, 2004). These questions have been central to the study of biodiversity for centuries. For example, some researchers used the term "Humboldt's Enigma" to refer to the challenge of explaining the extraordinary levels of diversity found in mountain regions (Figure 1), especially within the tropics, a phenomenon that remains only partially understood (Rahbek et al., 2019b; Trew & Maclean, 2021). Resolving these long-standing questions is not just of theoretical importance but also crucial for biodiversity conservation, especially in the face of accelerated global environmental change.


A Species richness

B Mountain endemics

C Smallest ranges

Figure 1. Maps illustrating the global patterns of diversity in mountain regions based on key metrics (extracted from Rahbek et al., 2019b). (A) Species richness, considering the total number of overlapping amphibian, bird and mammal species ranges in each mountain region. (B) Proportion of species that are mountain endemics, highlighting regions with high levels of unique biodiversity. (C) Richness of small-ranged species, defined as the 25% of species with the smallest geographic ranges.

2. Sky islands: the Afromontane and Afroalpine archipelagos

Mountains worldwide harbor half of the planet's biodiversity hotspots despite covering less than one-fifth of the Earth's surface (Fjeldså et al., 2012). Among these hotspots is the Afromontane region, which is key for understanding the evolutionary and biogeographical history of the African continent (White, 1983). The Afromontane region consists of small, fragmented and isolated patches of cool, humid environments scattered across African mountains (Figure 2). These patches are surrounded by low-elevation habitats with entirely different climates and biotas, creating a sharp ecological contrast and earning the region the nickname "archipelago" (Gehrke & Linder, 2009, 2014). The regions span north to south of Africa, from the Ethiopian highlands to the Cape mountains, and east to west, from Madagascar's highest peaks to Mount Cameroon (Davis et al., 1994; Van Wyk & Smith, 2001; Burgoyne et al., 2005). Its remarkable biodiversity includes over 4000 plant species, more than 75% of which are endemic (White, 1981).

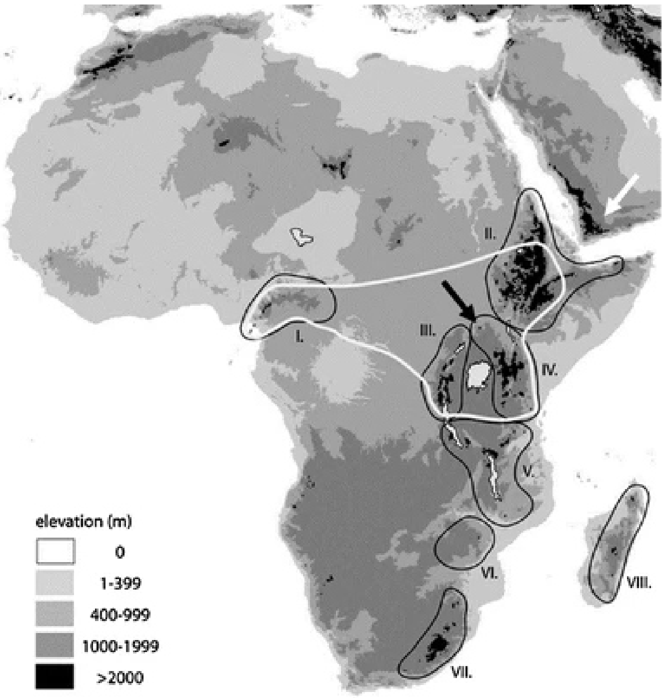


Figure 2. Afromontane and Afroalpine regions in Africa (extracted from Gehrke & Linder, 2014). The gray scale indicates elevation in meters. Regions: I. West Africa, II. Ethiopian Highlands, III. Kivu-Rwenzori, IV. Imatong-Usambara, V. Uluguru-Mlanje, VI. Chimanimani, VII. Southern Africa, VIII. Madagascar. The Tropical Afroalpine region is circumscribed with a white circle, including regions I, II, III and IV. The black arrow indicates Imatong Mountains and the white arrow indicates the high mountains of Yemen.

The evolutionary processes shaping this region's flora remain under discussion, but the geological and climatological contexts are widely acknowledged as pivotal. The mountains forming the Afromontane archipelago have diverse origins: the oldest date back to the Oligocene, while the youngest emerged in the Pleistocene (Gehrke & Linder 2014), most of them as a result of the intense volcanism and tectonic activity associated with the formation of the East African Rift (Chorowicz, 2005). These geological events coincided with mid-Miocene global cooling and aridification, which resulted in the emergence of Africa's vast deserts and savannas (Burgoyne et al., 2005). In parallel, mountains became cool, humid refuges, restricting the Afromontane flora to relatively small, isolated areas (Kebede et al., 2007; Chala et al., 2017). Along the Pleistocene's glacial and interglacial periods, this flora expanded and contracted its range repeatedly, often descending to lower altitudes during unfavourable times (Casas-Gallego et al., 2023). During climatically favorable periods, temporary montane belt corridors likely connected most of the Afromontane region (Allen et al., 2021), facilitating gene flow and species migration between mountains. Biogeographically, the flora of this region presents elements from southern Africa, northern hemisphere mountains and surrounding tropical lowlands (Hedberg, 1961, 1965, 1970). Floristic affinities between these regions have been explained through a bidirectional migratory route (Galley et al., 2007), with Eurasian elements migrating southward through the Yemeni mountains (Ehrich et al., 2007) and South African elements moving northward from the Cape, in which the Drakensberg would have played an important role as a stepping-stone area (Galley et al., 2007; Carlson et al., 2012; Pirie et al., 2016).

The vegetation of the Afromontane archipelago can be broadly divided into three elevational belts, each hosting distinct plant communities: the montane, the ericaceous and the alpine belts. The alpine belt, located above 3200 meters, represents another center of endemism, the Afroalpine archipelago (Davis et al., 1994). It features shrub communities dominated by giant *Dendrosenecio* (Hauman ex Hedberg) B.Nord. and *Lobelia* Mill., *Alchemilla* L., *Helichrysum* Mill. and open *Erica* L. formations (Gehrke & Linder, 2014). These plants are adapted to extreme climatic conditions, with diurnal rather than seasonal temperature cycles, high solar radiation during the day, freezing nights, and low precipitation (Hemp, 2006). Although less diverse, with only 520 species, 78% of these are endemic (Gehrke & Linder, 2014). Strictly Afromontane lineages have a complex history, many are ancient and trace their origins primarily to vicariance, either from surrounding tropical forests (Couvreur et al., 2008; Dimitrov et al., 2012) or the Rand Flora (Mairal et al., 2015), while some others are younger, coming from southern Africa or Eurasia (Galley et al., 2007; Gehrke & Linder, 2014). Contrastingly, Afroalpine lineages are more recent, dating back to the Pliocene-Pleistocene and generally arrived via long-distance dispersal from northern temperate regions (Hedberg, 1970; Gehrke & Linder, 2009; reviewed by Brochmann et al., 2022), often through repeated colonization events (e.g. Assefa et al., 2007; Kandziora et al., 2016). Recent research suggests that the Afroalpine biome is unstable and unsaturated, with the capacity to accommodate new lineages, resulting in a dynamic and relatively young flora (Kandziora et al., 2022).

The Afromontane archipelago also includes Madagascar's mountains (Figure 2), though they do not reach the minimum elevation to also be considered part of the Afroalpine archipelago (Burgoyne et al., 2005; Gehrke & Linder, 2014). This continental island is renowned for its extraordinary diversity and high endemism rates, distributed across a wide range of ecosystems. Madagascar's mountains are mostly placed on a north-to-south axis, which, combined with the northeast-southwest precipitation gradient, results in high habitat heterogeneity (Vences et al., 2009). Of the 11516 plant species described on the island, over 82% are endemic (Lowry et al., 2018; Antonelli et al., 2022). Despite its geographical isolation, dating back to its complete separation from other fragments of Gondwana c. 90 million years ago (De Wit, 2003), Madagascar exhibits notable floristic affinities with Africa (Yoder & Nowak, 2006). This observation aligns with the Theory of Island Biogeography, which predicts that most lineages on an island originate from the nearest mainland (MacArthur & Wilson, 1967). Recent studies indicate that the majority of Madagascar's endemic genera arose during the Miocene, with most colonizers coming from Africa (Buerki et al., 2013), often through multiple dispersal events (e.g. Razafimandimbinson et al., 2017). Similarities with continental Africa are especially pronounced in the mountains, where continental montane systems are mirrored, becoming particularly evident above 1300 meters on the island's central and eastern regions. There is evidence that these areas were colonized by some lineages from comparable environments in the African continent (e.g. Gehrke et al., 2016; Hackel et al., 2018).

Despite the wide expanse of the African "sky-islands", most research has focused on the tropical Afroalpine archipelago, beginning in the mid-20th century with the foundational contributions of the Swedish botanists Inga and Olov Hedberg (e.g. Hedberg & Hedberg, 1979). Their work inspired numerous studies on Afroalpine lineages, predominantly those of Eurasian origin, giving a misleading impression that the region is well-studied (Brochmann et al., 2022). However, significant knowledge gaps remain, especially regarding lineages with southern African origins and flora unique to the Afromontane region.

3. The genus *Helichrysum* and the HAP clade

The family Compositae (also known as Asteraceae) is the most diverse plant family in the Afromontane and Afroalpine regions (Gehrke & Linder, 2014) and ranks among the five dominant plant families in Madagascar (Callmander et al., 2011). Within this family, *Helichrysum*, a member of the Gnaphalieae tribe, stands out as one of the most species-rich genera, comprising between 500 and 600 species (Hilliard, 1983; Anderberg, 1991). Commonly known as "everlasting daisies", *Helichrysum* is renowned for its papery inflorescences, that retain their shape and color long after harvesting, as well as for its medicinal properties (e.g. Marini et al., 2023). This genus is particularly relevant to our study areas, as it is one of the most diverse in the Afromontane and Afroalpine regions and the most species-rich within its family in Madagascar. In these regions, it plays a key role in shaping the landscape, dominating open and slightly disturbed shrubby communities (Humbert, 1962; Hedberg, 1970; Brochmann et al., 2022).

It is hypothesized that the genus *Helichrysum* originated in southwestern Africa during the early Miocene (Galbany-Casals et al., 2014; Andrés-Sánchez et al., 2019). From there, it expanded northward and dispersed out of Africa, reaching tropical Africa, Madagascar, Eurasia and the Americas. Multiple independent colonization events of the tropical Afromontane archipelago have been proposed, with the Drakensberg mountains acting as a bridge and center of diversification. Previous studies suggest that the Afromontane region served as a source area for the adjacent Afroalpine region (Galbany-Casals et al., 2014), challenging the common hypothesis that most Afroalpine lineages originated in Eurasia (Brochmann et al., 2022). Similarly, research indicates that *Helichrysum* colonized Madagascar five times independently from Africa (Galbany-Casals et al., 2014). However, these findings should be interpreted cautiously, because they are based on phylogenies with limited sampling and low resolution. For instance, the most comprehensive phylogeny of *Helichrysum* to date (Galbany-Casals et al., 2014), includes 145 species (approximately 25% of the genus) and exhibits low overall supports, likely due to its reliance on just two nuclear markers (ITS and ETS). Moreover, the evolutionary history of *Helichrysum* is particularly complex, as it is partly shaped by hybridization and polyploidy. Finally, its relatively recent origin, its sub-cosmopolitan distribution, which includes mountain systems worldwide, and its high number of species, further complicate its study.

Southern Africa, where *Helichrysum* originated, accounts for most of the genus diversity, hosting over 240 species (Hilliard, 1983). While this number is based on the most recent floristic treatment of Southern African *Helichrysum* (Hilliard, 1983), newer research keeps increasing the number of species (e.g. Andrés-Sánchez et al., 2024). Beyond South Africa, the genus comprises more than 70 species across different African regions (Lisowski, 1989; Beentje, 2002), although these numbers may be imprecise due to outdated floristic references. Madagascar alone accounts for 111 documented species, almost all endemic (Humbert, 1962). The rest of the genus' diversity extends into the Northern Hemisphere, encompassing the Mediterranean basin, Macaronesia, the Arabian Peninsula and Asia, where approximately 45 species occur (Galbany-Casals et al., 2006; Ward et al., 2009; Kilian et al., 2010; Şenol et al., 2011; Herrando-Moraira et al., 2016, 2017).

Helichrysum stands out for its extraordinary ecological adaptability, thriving in a wide range of environments, including rocky sites, wetlands, various types of grasslands and shrubby communities, coastal dunes, semi-deserts, savannas and forests (Figure 3). While most species are subshrubs or shrubs, the genus displays diverse life forms (Figure 4), ranging from annual herbs to small trees and even lianas, fulfilling various ecological and community roles. Its morphological diversity is equally remarkable, manifesting in variations in the number, size and arrangement of capitula as well as in differences in color and size of involucral bracts (Anderberg, 1991). This diversity is particularly pronounced in mountain ecosystems, which host more than half of the African species of the genus, including approximately 60 strictly Afromontane species and 12 Afroalpine species (Lisowski, 1989; Beentje, 2002; Mesfin Tadesse, 2004). Southern African mountains harbor around 132 species (Hilliard, 1983), with 24 endemics to the highest parts of the Drakensberg range (Carbutt & Edwards, 2004). In Madagascar, nearly 44% of species inhabit montane habitats (Humbert, 1962).

Figure 3. Some of the African landscapes where *Helichrysum* inhabits: (a) *Helichrysum splendidum* Less. in southern African grasslands; (b) *Helichrysum globosum* Sch.Bip. in humid tropical Afromontane swamps; (c) Ericoid thicket in Mount Marojejy (Madagascar); (d) Tropical Afromontane *Hagenia* forest; (e) Malagasy highlands with Tapia forest in a markedly rocky substrate in Isalo, bordering the arid southwestern region of Madagascar; (f) *Dendrocenecio erici-rosenii* (R.E.Fr. & T.C.E.Fr.) E.B. Knox in the Afroalpine belt of Virunga mountains (Rwanda). (Images by Mercè Galbany-Casals).

Figure 4. Examples of morphological diversity in the HAP clade: (a) *Helichrysum dunense* Hilliard, a prostrate chamaephyte inhabiting coastal dunes in South Africa; (b) *Pseudognaphalium oxyphyllum* (DC.) Kirp., a hemicryptophyte inhabiting Central American open habitats; (c) *H. betsiliense* Klatt, a chamaephyte found in Malagasy grasslands; (d) *H. schimperi* (Sch.Bip. ex A.Rich) Moeser, a lianescent species spanning the forests of tropical Africa and the Arabian Peninsula; (e) *H. formosissimum* Sch.Bip., a treelet widely distributed in the upper belt of montane forests of tropical Africa; (f) *Anaphalis contorta* Hook.f., a chamaephyte found in Asian mountain grasslands. (Images by Mercè Galbany-Casals, except for (a) by Santiago Andrés-Sánchez).

The systematics of *Helichrysum* has long been a subject of debate, with many researchers considering the genus to be artificial due to its polyphyletic and paraphyletic status. Recent studies propose that it would be more appropriate to refer to the main *Helichrysum* clade as the HAP clade, which includes several genera such as *Achyrocline* (Less.) DC. —mostly diversified in America—, *Anaphalis* DC. —mostly diversified in Asia—, *Pseudognaphalium* Kirp. —mostly diversified in America—, among others (Smissen et al., 2011; Galbany-Casals et al., 2014; Nie et al., 2016). However, this proposal does not solve the issue of several southern African *Helichrysum* species falling outside of the HAP clade, being more closely related to other genera, such as *Anaxeton* Gaertn. and *Petalacte* D.Don. (Galbany-Casals et al., 2014 and references therein). This underscores the urgent need for a comprehensive taxonomic revision of the group, as suggested by Galbany-Casals et al. (2014). The process has already begun with the transference of Australian *Helichrysum* species into other genera like *Xerochrysum* Tzvekev and *Ozothamnus* R.Br. (Schmidt-Lebuhn et al., 2015) and the reassignment of South African species to *Syncarpha* DC. (Nordenstam, 2003; Bergh et al., 2015), but further taxonomic rearrangements are needed to achieve a well-defined and coherent generic delimitation in the group.

4. New sequencing methods and the advent of plant phylogenomics

The emergence of Sanger sequencing in the 1970s (Sanger et al., 1977) marked a milestone in the history of molecular genetics. This technique played a pivotal role in the development of major projects such as the Human Genome Project (Venter et al., 2001) and was key in characterizing and shedding light on the origins of biodiversity. It transformed our understanding of evolutionary relationships and diversification patterns. However, as research questions became increasingly complex and ambitious, the limitations of Sanger sequencing – such as low throughput, limited scalability and high costs (Mardis, 2013) – became more evident. These constraints often led to the formulation of incorrect phylogenetic hypotheses (Nicholls et al., 2015). The development of high-throughput sequencing (HTS) techniques that are faster, more cost-effective and capable of processing larger volumes of data allowed for genome-scale analyses instead of single-gene approaches, radically changing the approach of evolutionary studies (Escudero et al., 2020).

The advent of Next-Generation Sequencing (NGS) platforms has revolutionized genomics, providing tools that allow us to tackle complex biological questions. NGS enables sequencing across various scales, from reconstructing entire genomes and transcriptomes to recovering thousands of specific genomic regions, all within a short timeframe and at increasingly affordable costs. These options have had an enormous impact across multiple scientific disciplines, making NGS indispensable (Goodwin et al., 2016). In plant evolution and systematics, NGS offers the opportunity to obtain genomic data useful to address challenges inherent to the complex nature of plant genomes, such as polyploidy, hybridization, and incomplete lineage sorting (Rokas & Abbot, 2009; Rothfels, 2021) as well as providing versatile data for the study of both deep and shallow phylogenetic relationships, bridging between analyses at macro and microevolutionary scales (Combrink et al., 2025). Importantly, this technological revolution has been accompanied by significant advances in bioinformatics pipelines tailored to specific methods, as well as improvements in computational and storage resources (Koboldt et al., 2013; Barrett et al., 2016; Andermann et al., 2020).

A cost-effective NGS technique for non-model organisms is Hyb-Seq (Weitemier et al., 2014). This approach combines genome skimming —a low-resolution sequencing technique particularly useful for recovering organellar genomes— with target enrichment, which enables targeted sequencing of hundreds to thousands of phylogenetically informative markers (Mamanova et al., 2010; Grover et al., 2012). By integrating these two complementary techniques, Hyb-Seq generates high-quality data for addressing phylogenomic questions at various taxonomic levels while minimizing costs and reducing the complexity of bioinformatic analyses required for processing raw data (Cronn et al., 2012; Weitemier et al., 2014). Its affordability and low complexity compared to whole-genome sequencing make it particularly suited for studying hyper-diverse groups while accumulating comparable and reusable data, a huge advantage over other genomic enrichment techniques such as RADseq (Harvey et al., 2016). Moreover, since Hyb-Seq is effective with fragmented DNA, it can be successfully applied to historical herbarium specimens, which often contain highly degraded DNA (Lemmon & Lemmon, 2013; Bakker et al., 2016).

The development of specialized target-enrichment kits for certain groups of organisms, which enrich pre-defined DNA regions using baits (Herrando-Moraira et al., 2018), has further expanded the use of these techniques. In 2014, the first plant-specific kit, Compositae1061, was designed for the Compositae family (Mandel et al., 2014). This kit has successfully been used to obtain a well-resolved family tree backbone (Mandel et al., 2019) and reconstruct highly resolved phylogenies for numerous Compositae lineages, such as Cardueae (Herrando-Moraira et al., 2019), Vernonieae (Siniscalchi et al., 2019), *Dendrosenecio* (Gizaw et al., 2021) and Scorzoneriae (Hatami et al., 2022). A decade later, the complementary Compositae-ParaLoss-1272 kit (Moore-Pollard et al., 2024) has been released to improve on its predecessor and address unresolved phylogenetic challenges associated with paralogy. This widespread phenomenon in Compositae is often a reflection of the family's complex evolutionary history, shaped by multiple polyploidization and hybridization events as well as genomic duplications. Over time, other group-specific kits have been developed, such as the one for Euphorbiaceae (Villaverde et al., 2018), Orchidaceae (Eserman et al., 2021) and Rubiaceae (Ball et al., 2023) along with kits with a broader taxonomic scope such as Angiosperms353 (Johnson et al., 2019) and GoFlag (Breinholt et al., 2021). These kits can be successfully combined, as demonstrated in numerous studies (Larridon et al., 2020; Hendricks et al., 2021; Siniscalchi et al., 2021). Furthermore, custom kits based on known genomes or transcriptomes from related species can be designed to address specific evolutionary questions (e.g. Chau et al., 2018; Straub et al., 2020).

Phylogenetic inference primarily relies on orthology assumption, for which single-copy genes are expected to reflect speciation events. However, this assumption can be violated due to paralogy —the presence of multiple and diverging gene copies arising after gene or whole-genome duplication events (Fitch, 1970). The inclusion of paralogous sequences can introduce misleading signals, leading to misinterpretations of evolutionary relationships. This issue has been amplified with NGS methodologies. Although NGS has significantly increased the volume and resolution of genetic data available for analysis, it has also raised the likelihood of capturing paralogous copies. Although target-enrichment kits often attempt to minimize the inclusion of paralogous copies during probe design (e.g. Schmickl et al., 2016), a substantial proportion of genes with potential paralogs may be recovered depending on the study group (e.g. Siniscalchi et al., 2021). Furthermore, distinguishing between orthologous and paralogous copies remains challenging due to their high sequence similarity (Altenhoff et al., 2019). The most conservative strategy involves filtering out loci including paralogs, with pipelines designed for strict paralog removal (e.g. Zhou et al., 2022). However, this can lead to dramatically reduced datasets, especially in lineages in which whole-genome duplications, polyploidization or hybridization events are common (Morales-Briones et al., 2021). Some studies prove that adding more data, even genes including potential paralogs, can lead to convergence between different analytical approaches (Gardner et al., 2021). Thus, alternative approaches should be considered, such as ignoring paralog warnings and building alignments containing a mix of orthologous and paralogous sequences, followed by applying incomplete lineage sorting-aware species tree methods (Yan et al., 2022). Another option is retrieving all gene copies and constructing separate orthologous alignments for each divergent copy of a given locus. The latter is aligned with Smith & Hahn (2021), who propose that properly identified paralogs can enhance both the robustness and support of phylogenetic trees by capturing valuable evolutionary information. The pipeline ParalogWizard (Ufimov et al., 2022), facilitates this approach by identifying paralogs based on sequence divergence and integrating the most divergent copies as additional markers into separate alignments, increasing the total number of markers used and offering a more nuanced way to handle gene duplication events in phylogenomic datasets.

References

Abbott, R., Albach, D., Ansell, S., et al. (2013). Hybridization and speciation. *J. Evol. Biol.*, 26(2), 229–246.

Allen, K.E., Tapondjou, W.P., Freeman, B., et al. (2021). Modelling potential Pleistocene habitat corridors between Afromontane forest regions. *Biodivers. Conserv.*, 30, 2361–2375.

Altenhoff, A.M., Glover N.M., & Dessimoz, C. (2019). Inferring orthology and paralogy. In: *Evolutionary Genomics. Methods in Molecular Biology*, volume 1910, pp. 149-175. Anisimova M., (ed.). Humana: New York.

Anderberg, A.A. (1991). Taxonomy and phylogeny of the tribe Gnaphalieae (Asteraceae). *Opera Bot.*, 104, 1–195.

Andermann, T., Torres-Jiménez, M.F., Matos-Maraví, P., et al. (2020). A Guide to Carrying Out a Phylogenomic Target Sequence Capture Project. *Front. Genet.*, 10, 1407.

Andrés-Sánchez, S., Verboom, G.A., Galbany-Casals, M., & Bergh, N.G. (2019). Evolutionary history of the arid climate-adapted *Helichrysum* (Asteraceae: Gnaphalieae): Cape origin and association between annual life-history and low chromosome numbers. *J. Syst. Evol.*, 57(5), 468–487.

Andrés-Sánchez, S., Verboom, G.A., Galbany-Casals, M., & Bergh, N.G. (2024). Resurrection of two endemic Cape species of *Helichrysum* (Asteraceae, Gnaphalieae) from amongst the arid-adapted annual lineages. *S. Afr. J. Bot.*, 170, 417–424.

Antonelli, A., Kissling, W.D., Flantua, S.G.A., et al. (2018). Geological and climatic influences on mountain biodiversity. *Nat. Geosci.*, 11, 718–725.

Antonelli, A., Smith, R.J., Perrigo, A.L., et al. (2022). Madagascar's extraordinary biodiversity: evolution, distribution and use. *Science*, 378(6623), eabf0869.

Assefa, A., Ehrich, D., Taberlet, P., et al. (2007). Pleistocene colonization of afro-alpine 'sky islands' by the arctic-alpine *Arabis alpina*. *Heredity*, 99, 133–142.

Bakker, F.T., Lei, D., Yu, J., et al. (2016). Herbarium genomics: plastome sequence assembly from a range of herbarium specimens using an Iterative Organelle Genome Assembly pipeline. *Bot. J. Linn.*, 117(1), 33–43.

Ball, L.D., Bedoya, A.M., Taylor, C.M., & Lagomarsino L.P. (2023). A target enrichment probe set for resolving phylogenetic relationships in the coffee family, Rubiaceae. *Appl. Plant Sci.*, 11(6), e11554.

Barrett, C.F., Bacon, C.D., Antonelli, A., Cano, A., & Hofmann, T. (2016). An introduction to plant phylogenomics with a focus on palms. *Bot. J. Linn.*, 182(2), 234–255.

Beentje, H.J. (2002). *Helichrysum*. In: *Flora of Tropical East Africa*, part 2, pp. 403–452. Beentje, H.J., (ed.). Balkema: Rotterdam.

Bentley, J., Verboom, G.A., & Bergh, N.G. (2014). Erosive processes after tectonic uplift stimulate vicariant and adaptive speciation: evolution in an Afrotemperate-endemic paper daisy genus. *BMC Evol. Biol.*, 14, 27.

Bergh, N.G., Haiden, S.A., & Verboom, G.A. (2015). Molecular phylogeny of the 'Cape snow' genus *Syncarpha* (Asteraceae: Gnaphalieae) reveals a need for generic re-delimitation. *S. Afr. J. Bot.*, 100, 219–227.

Breinholt, J.W., Carey, S.B., Tiley, G.P., et al. (2021). A target enrichment probe set for resolving the flagellate land plant tree of life. *Appl. Plant Sci.*, 9(1), e11406.

Brochmann, C., Gizaw, A., Chala, D., et al. (2022). History and evolution of the afroalpine flora: in the footsteps of Olov Hedberg. *Alp. Bot.*, 132, 65–87.

Buerki, S., Devey, D.S., Callmander, M.W., Phillipson, P.B., & Forest, F. (2013). Spatio-temporal history of the endemic genera of Madagascar. *Bot. J. Linn.*, 171(2), 304–329.

Burgoyne, P.M., Van Wyk, A.E., Anderson, J.M., & Schrire, B.D. (2005). Phanerozoic evolution of plants on the African Plate. *J. Afr. Earth Sci.*, 43, 13–52.

Callmander, M.W., Phillipson, P.B., Schatz, G.E., et al. (2011). The endemic and non-endemic vascular flora of Madagascar updated. *Plant Ecol. Evol.*, 144(2), 121–125.

Carbutt, C., & Edwards, T.J. (2004). The flora of the Drakensberg Alpine Centre. *Edinburgh J. Bot.*, 60, 581–607.

Cardinale, B., Duffy, J., Gonzalez, A., et al. (2012). Biodiversity loss and its impact on humanity. *Nature*, 486, 59–67.

Carlson, S.E., Linder, H.P., & Donoghue, M.J. (2012). The historical biogeography of *Scabiosa* (Dipsacaceae): implications for Old World plant disjunctions. *J. Biogeogr.*, 39, 1086–1100.

Casas-Gallego, M., Hahn, K., Neumann, K., et al. (2023). Cooling-induced expansions of Afromontane forests in the Horn of Africa since the Last Glacial Maximum. *Sci. Rep.*, 13, 10323.

Cerca, J., Cotoras, D.D., Bieker, V.C., et al., (2023). Evolutionary genomics of oceanic island radiations. *Trends Ecol. Evol.*, 38(7), 631–642.

Chala, D., Zimmermann, N.E., Brochmann, C., et al. (2017). Migration corridors for alpine plants among the 'sky islands' of eastern Africa: do they, or did they exist? *Alp. Bot.*, 127, 133–144.

Chau, J.H., Rahfeldt, W.A., & Olmstead, R.G. (2018). Comparison of taxon-specific versus general locus sets for targeted sequence capture in plant phylogenomics. *Appl. Plant Sci.*, 6(3), e1032.

Chorowicz, J. (2005). The East African rift system. *J. Afr. Earth Sci.*, 43, 379–410.

Combrink, L.L., Golcher-Benavides, J., Lewanski, A.L., Rick, J.A., Rosenthal, W.C., & Wagner, C.E. (2025). Population Genomics of Adaptive Radiation. *Mol. Ecol.*, 34, e17574.

Couvreur, T.L., Chatrou, L.W., Sosef, M.S., & Richardson, J.E. (2008). Molecular phylogenetics reveal multiple tertiary vicariance origins of the African rain forest trees. *BMC Biol.*, 6, 54.

Cronn, R., Knaus, B.J., Liston, A., et al. (2012). Targeted enrichment strategies for next-generation plant biology. *Am. J. Bot.*, 99(2), 291–311.

Davis, S.D., Heywood, V.H., & Hamilton, A.C. (1994). Centres of plant diversity. In: *A Guide and Strategy for their Conservation. Europe, Africa, South West Asia and the Middle East*, volume 1. IUCN Publications Unit: Cambridge.

De Wit, M.J. (2003). Madagascar: heads it's a continent, tails it's an island. *Annu. Rev. Earth Planet. Sci.*, 31(1), 213–248.

Dimitrov, D., Nogués-Bravo, D., & Scharff, N. (2012). Why Do Tropical Mountains Support Exceptionally High Biodiversity? The Eastern Arc Mountains and the Drivers of *Saintpaulia* Diversity. *PLoS ONE*, 7(11), e48908.

Donoghue, M.J. (2008). Colloquium paper: a phylogenetic perspective on the distribution of plant diversity. *Proc. Natl. Acad. Sci. U.S.A.*, 105 (Suppl 1), 11549–11555.

Edwards, E.J., & Donoghue, M.J. (2013). Is it easy to move and easy to evolve? Evolutionary accessibility and adaptation. *J. Exp. Bot.*, 64 (13), 4047–4052.

Ehrich, D., Gaudeul, M., Assefa, A., et al. (2007). Genetic consequences of Pleistocene range shifts: contrast between the Arctic, the Alps and the East African mountains. *Mol. Ecol.*, 16, 2542–2559.

Escudero, M., Nieto Feliner, G., Pokorny, L., Spalink, D., & Viruel, J. (2020). Editorial: Phylogenomic Approaches to Deal With Particularly Challenging Plant Lineages. *Front. Plant Sci.*, 11, 591762.

Eserman, L.A., Thomas, S.K., Coffey, E.E.D., & Leebens-Mack, J.H. (2021). Target sequence capture in orchids: developing a kit to sequence hundreds of single-copy loci. *Appl. Plant Sci.*, 9, e11416.

Fitch, W.M. (1970). Distinguishing homologous from analogous proteins. *Syst. Biol.*, 19, 99–113.

Fjeldså, J., Bowie, R.C.K., & Rahbek, C. (2012). The Role of Mountain Ranges in the Diversification of Birds. *Annu. Rev. Ecol. Evol. Syst.*, 43, 249–265.

Galbany-Casals, M., Sáez, L., & Benedí, C. (2006). A taxonomic revision of *Helichrysum* Mill. sect. *Stoechadina* (DC.) Gren. & Godr. (Asteraceae, Gnaphalieae). *Canad. J. Bot.*, 84, 1203–1232.

Galbany-Casals, M., Unwin, M., Garcia-Jacas, N., et al. (2014). Phylogenetic relationships in *Helichrysum* (Compositae: Gnaphalieae) and related genera: Incongruence between nuclear and plastid phylogenies, biogeographic and morphological patterns, and implications for generic delimitation. *Taxon*, 63, 608–624.

Galley, C., Bytebier, B., Bellstedt, D.U., & Linder, H. P. (2007). The Cape element in the Afrotemperate flora: from Cape to Cairo? *Proc. Biol. Sci.*, 274(1609), 535–543.

Gardner, E.M., Johnson, M.G., Pereira, J.T., et al. (2021). Paralogs and off-target sequences improve phylogenetic resolution in a densely sampled study of the breadfruit genus (*Artocarpus*, Moraceae). *Syst. Biol.*, 70(3), 558–575.

Geeta, R., Lohmann, L.G., Magallón, S., et al. (2014). Biodiversity only makes sense in the light of evolution. *J. Biosci.*, 39, 333–337.

Gehrke, B., Kandziora, M., & Pirie, M.D. (2016). The evolution of dwarf shrubs in alpine environments: a case study of *Alchemilla* in Africa. *Ann. Bot.*, 117(1), 121–131.

Gehrke, B., & Linder, H.P. (2009). The scramble for Africa: pan-temperate elements on the African high mountains. *Proc. Biol. Sci.*, 276(1667), 2657–2665.

Gehrke, B., & Linder, H.P. (2014). Species richness, endemism and species composition in the tropical Afroalpine flora. *Alp. Bot.*, 124, 165–177.

Gillespie, R.G. (2007). Oceanic Islands: Models of Diversity. In: *Encyclopedia of Biodiversity*. Levin, S.A., (ed.). Elsevier: Amsterdam.

Gizaw, A., Gorospe, J.M., Kandziora, M., et al. (2021). Afro-alpine flagships revisited II: elucidating the evolutionary relationships and species boundaries in the giant senecios (*Dendrosenecio*, Asteraceae). *Alp. Bot.*, 132, 89–105.

Goodwin, S., McPherson, J.D., & McCombie, W.R. (2016). Coming of age: ten years of next-generation sequencing technologies. *Nat. Rev. Genet.*, 17(6), 333–351.

Grover, C.E., Salmon, A., & Wendel, J.F. (2012). Targeted sequence capture as a powerful tool for evolutionary analysis. *Am. J. Bot.*, 99, 312–319.

Hackel, J., Vorontsova, M.S., Nanjarisoa, O.P., et al. (2018). Grass diversification in Madagascar: *In situ* radiation of two large C3 shade clades and support for a Miocene to Pliocene origin of C4 grassy biomes. *J. Biogeogr.*, 45, 750–761.

Harvey, M.G., Smith, B.T., Glenn, T.C., Faircloth, B.C., & Brumfield, R.T. (2016). Sequence Capture versus Restriction Site Associated DNA Sequencing for Shallow Systematics. *Syst. Biol.*, 65(5), 910–924.

Hatami, E., Jones, K.E., & Kilian, N. (2022). New Insights Into the Relationships Within Subtribe Scorzonerinae (Cichorieae, Asteraceae) Using Hybrid Capture Phylogenomics (Hyb-Seq). *Front. Plant Sci.*, 13, 851716.

Hedberg, O. (1961). The phytogeographical position of the Afroalpine Flora. *Rec. Adv. Bot.*, 1, 914–919.

Hedberg, O. (1965). Afroalpine flora elements. *Webbia*, 19, 519–529.

Hedberg, O. (1970). Evolution of the Afroalpine Flora. *Biotropica*, 2, 16–23.

Hedberg, I., & Hedberg, O. (1979). Tropical-alpine life-forms of vascular plants. *Oikos*, 33(2), 297.

Hemp, A. (2006). Continuum or zonation? Altitudinal gradients in the forest vegetation of Mt. Kilimanjaro. *Plant Ecol.*, 184, 27–42.

Hendriks, K.P., Mandáková, T., Hay, N.M., et al. (2021). The best of both worlds: Combining lineage-specific and universal bait sets in target-enrichment hybridization reactions. *Appl. Plant Sci.*, 9(7), 10.1002/aps3.11438.

Herrando-Moraira, S., Blanco-Moreno, J.M., Sáez, L., & Galbany-Casals, M. (2016). Reevaluación del complejo *Helichrysum italicum* (Compositae: Gnaphalieae): una nueva especie en Mallorca (Islas Baleares). *Collect. Bot.*, 35, e009.

Herrando-Moraira, S., Calleja, J.A., Carnicero, P., et al. (2018). Exploring data processing strategies in NGS target enrichment to disentangle radiations in the tribe Cardueae (Compositae). *Mol. Phylogenetics Evol.*, 128, 69–87.

Herrando-Moraira, S., Calleja, J.A., Galbany-Casals, M., et al. (2019). Nuclear and plastid DNA phylogeny of tribe Cardueae (Compositae) with Hyb-Seq data: A new subtribal classification and a temporal diversification framework. *Mol. Phylogenetics Evol.*, 137, 313–332.

Herrando-Moraira, S., Carnicero, P., Blanco-Moreno, J.M., et al. (2017). Systematics and phylogeography of the Mediterranean *Helichrysum pendulum* complex (Compositae) inferred from nuclear and chloroplast DNA and morphometric analyses. *Taxon*, 66, 909–933.

Hilliard, O.M. (1983). *Helichrysum* Mill. In: *Flora of Southern Africa*, volume 33, part 7, fascicle 2, pp. 61–310. Leistner; O.A., (ed.). Department of Agriculture: Pretoria.

Hughes, C.E., & Atchison, G.W. (2015). The ubiquity of alpine plant radiations: from the Andes to the Hengduan Mountains. *New Phytol.*, 207, 275-282.

Humbert, H. (1962). *Flore de Madagascar et des Comores (Plantes vasculaires)*; 189e famille. Composées, part 2. Leroy, J.F., (ed.). Muséum National d'Histoire Naturelle: Paris.

Johnson, M.G., Pokorny, L., Dodsworth, S., et al. (2019). A Universal Probe Set for Targeted Sequencing of 353 Nuclear Genes from Any Flowering Plant Designed Using k-Medoids Clustering. *Syst. Biol.*, 68(4), 594–606.

Kandziora, M., Gehrke, B., Popp, M., Gizaw, A., Brochmann, C. & Pirie, M.D. (2022). The enigmatic tropical alpine flora on the African sky islands is young, disturbed, and unsaturated. *Proc. Natl. Acad. Sci. U.S.A.*, 119(22), e2112737119.

Kandziora, M., Kadereit, J.W., & Gehrke, B. (2016). Frequent colonization and little in situ speciation in *Senecio* in the tropical alpine-like islands of eastern Africa. *Am. J. Bot.*, 103, 1483–1498.

Kebede, M., Ehrich, D., Taberlet, P., Nemomissa, S., & Brochmann, C. (2007). Phylogeography and conservation genetics of a giant lobelia (*Lobelia giberroa*) in Ethiopian and Tropical East African mountains. *Mol. Ecol.*, 16, 1233–1243.

Kilian, N., Galbany-Casals, M., & Oberprieler, C. (2010). *Helichrysum nicolai* (Compositae, Gnaphalieae), a dwarf local endemic of the Cape Verde Islands, W Africa. *Folia Geobot.*, 45, 183–199.

Koboldt, D.C., Steinberg, K.M., Larson, D.E., Wilson, R.K., & Mardis, E.R. (2013). The next-generation sequencing revolution and its impact on genomics. *Cell*, 155(1), 27–38.

Larridon, I., Villaverde, T., Zuntini, A.R., et al. (2020). Tackling Rapid Radiations With Targeted Sequencing. *Front. Plant Sci.*, 10, 1655.

Lemmon, E.M., & Lemmon, A.R. (2013). High-throughput genomic data in Systematics and Phylogenetics. *Annu. Rev. Ecol. Evol. Syst.*, 44, 99–121.

Lisowski, S. (1989). *Helichrysum* Compositae (deuxième partie: Tribu Inuleae). In: *Flore d'Afrique Centrale* (Zaire, Rwanda, Burundi), pp. 68–192. National Botanic Garden of Belgium: Meise.

Losos, J.B., & Mahler, D.L. (2010). Adaptive radiation: The interaction of ecological opportunity, adaptation, and speciation. In: *Evolution Since Darwin: The First 150 Years*, pp: 381–420). Bell, M.A., Futuyma, D.J., Eanes, W.F. & Levinton, J.S. (eds). Sinauer Associates: Sunderland.

Losos, J.B., & Ricklefs, R.E. (2009). Adaptation and diversification on islands. *Nature*, 457(7231), 830–836.

Lowry, P.P., Phillipson, II, P.B., Andriamahefarivo, L., Schatz, G.E., Rajaonary, F., & Andriambololonera, S. (2018). Flora. In: *The terrestrial protected areas of Madagascar: Their history, description, and biota,* pp. 243–255. Goodman, S.M., Raherilalao, M.J., & Wohlhauser, S., (eds.). Assn. Vahatra: Antananarivo.

MacArthur, R.H., & Wilson, E.O. (1967). *The Theory of Island Biogeography*. Princeton University Press: Princeton.

Mairal, M., Pokorny, L., Aldasoro, J.J., Alarcón, M., & Sanmartín, I. (2015). Ancient vicariance and climate-driven extinction continental-wide disjunctions in Africa: the case of the Rand Flora genus *Canarina* (Campanulaceae). *Mol. Ecol.*, 24(6), 1335–1354.

Mamanova, L., Coffey, A.J., Scott, C.E., et al. (2010). Target-enrichment strategies for next-generation sequencing. *Nat. Methods*, 7(2), 111–118.

Mandel, J.R., Dikow, R.B., Funk, V.A., et al. (2014). A target enrichment method for gathering phylogenetic information from hundreds of loci: An example from the Compositae. *Appl. Plant. Sci.*, 2(2), apps.1300085.

Mandel, J.R., Dikow, R.B., Siniscalchi, C.M., Thapa, R., Watson, L.E., & Funk. V.A. (2019). A fully resolved backbone phylogeny reveals numerous dispersals and explosive diversifications throughout the history of Asteraceae. *Proc. Natl. Acad. Sci. U.S.A.*, 116(28), 14083–14088.

Mardis, E.R. (2013). Next-generation sequencing platforms. *Annu. Rev. Anal. Chem.*, 6, 287–303.

Marini, L., Bini, L., Gori, M., et al. (2023). Genetic and morphological assessment of *Helichrysum* Mill. from Tuscan Archipelago (Italy). *Sci. Hortic.*, 321, 112360.

Mesfin Tadesse. (2004). *Helichrysum* Mill., Asteraceae (Compositae). In: *Flora of Ethiopia and Eritrea*, part 2, pp. 163–178. Hedberg ,I., Friis, I., Edwards, S., (eds.). Addis Abeba University: Addis Ababa.

Miller, A.H., Stroud, J.T., & Losos, J.B. (2023). The ecology and evolution of key innovations. *Trends Ecol. Evol.*, 38(2), 122–131.

Mittermeier, R.A., Robles-Gil, P., Hoffman, M., et al. (2005). *Hotspots Revisited: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions*. Conservation International: Washington D.C.

Moore-Pollard, E.R., Jones, D.S., & Mandel, J.R. (2024). Compositae-ParaLoss-1272: A complementary sunflower-specific probe set reduces paralogs in phylogenomic analyses of complex systems. Appl. Plant Sci., 12(1), e11568.

Morales-Briones, D.F., Gehrke, B., Huang, C.H., et al. (2021). Analysis of Paralogs in Target Enrichment Data Pinpoints Multiple Ancient Polyploidy Events in *Alchemilla* s.l. (Rosaceae). *Syst. Biol.*, 71(1), 190–207.

Muellner-Riehl, A.N., Anthelme, F., & Ibanez. T. (2024). Past, present, and future of mountain and island systems. *J. Syst. Evol.*, 62(2), 195-200.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, 403(6772), 853–858.

Nevado, B., Contreras-Ortiz, N., Hughes, C., & Filatov, D.A. (2018). Pleistocene glacial cycles drive isolation, gene flow and speciation in the high-elevation Andes. *New Phytol.*, 219(2), 779–793.

Nicholls, J.A., Pennington, R.T., Koenen, E.J.M., et al. (2015). Using targeted enrichment of nuclear genes to increase phylogenetic resolution in the neotropical rain forest genus *Inga* (Leguminosae: Mimosoideae). *Front. Plant Sci.*, 6, 710.

Nie, Z.-L., Funk, V.A., Meng, Y., Deng, T., Sun, H., & Wen, J. (2016). Recent assembly of the global herbaceous flora: evidence from the paper daisies (Asteraceae: Gnaphalieae). *New Phytol.*, 209, 1795–1806.

Nordenstam, B. (2003). Further contributions to the genus *Syncarpha* (Compositae-Gnaphalieae). *Compos. Newsl.*, 39, 52–57.

Nürk, N.M., Linder, H.P., Onstein, R.E., et al. (2020). Diversification in evolutionary arenas-Assessment and synthesis. *Ecol. Evol.*, 10(12), 6163–6182.

Pimm, S.L., Jenkins, C.N. Abell, R., et al. (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, 344, 1246752.

Pirie, M.D., Oliver, E.G.H., Mugrabi de Kuppler, A., et al. (2016). The biodiversity hotspot as evolutionary hot-bed: spectacular radiation of *Erica* in the Cape Floristic Region. *BMC Evol. Biol.*, 16, 190.

Rahbek, C., Borregaard, M.K., Antonelli, A., et al. (2019a). Building mountain biodiversity: Geological and evolutionary processes. *Science*, 365(6458), 1114–1119.

Rahbek, C., Borregaard, M.K., Colwell, R.K., et al. (2019b). Humboldt's enigma: What causes global patterns of mountain biodiversity? *Science*, 365(6458), 1108–1113.

Razafimandimbison, S.G., Kainulainen, K., Wikström, N., & Bremer, B. (2017). Historical biogeography and phylogeny of the pantropical Psychotrieae alliance (Rubiaceae), with particular emphasis on the Western Indian Ocean Region. *Am. J. Bot.*, 104, 1407–1423.

Ricklefs, R.E. (2004). A comprehensive framework for global patterns in biodiversity. *Ecol. Lett.*, 7, 1–15.

Rokas, A., & Abbot, P. (2009). Harnessing genomics for evolutionary insights. *Trends Ecol. Evol.*, 24(4), 192–200.

Rothfels, C.J. (2021). Polyploid Phylogenetics. *New Phytol.*, 230, 66–72.

Rundell, R.J., & Price, T.D. (2009). Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends Ecol. Evol.*, 24(7), 394–399.

Sala, O.E., Chapin, F.S., Armesto, J.J., et al. (2000). Global biodiversity scenarios for the year 2100. *Science*, 287(5459), 1770–1774.

Sanger, F., Nicklen, S., & Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.*, 74(12), 5463–5467.

Schluter, D. (2000). *The Ecology of Adaptive Radiations*. Oxford University Press: New York.

Schmickl, R., Liston, A., Zeisek, V., et al. (2016). Phylogenetic marker development for target enrichment from transcriptome and genome skim data: the pipeline and its application in southern African *Oxalis* (Oxalidaceae). *Mol. Ecol. Resour.*, 16(5), 1124–1135.

Schmidt-Lebuhn, A.N., Bruhl, J.J., Telford, I.R.H., & Wilson, P.G. (2015). Phylogenetic relationships of *Coronidium*, *Xerochrysum* and several neglected Australian species of "*Helichrysum*" (Asteraceae: Gnaphalieae). *Taxon*, 64, 96–109.

Şenol, S.G., Secmen, Ö., Öztürk, B., & Galbany-Casals, M. (2011). *Helichrysum unicapitatum* (Asteraceae), a new species from Turkey. *Ann. Bot. Fenn.*, 48, 145–154.

Siniscalchi, C.M., Hidalgo, O., Palazzesi, L., et al. (2021). Lineage-specific vs. universal: A comparison of the Compositae1061 and Angiosperms353 enrichment panels in the sunflower family. *Appl. Plant Sci.*, 9(7), e11422.

Siniscalchi, C.M., Loeuille, B., Funk, V.A., Mandel, J.R., & Pirani, J.R. (2019). Phylogenomics Yields New Insight Into Relationships Within Vernonieae (Asteraceae). *Front. Plant Sci.*, 10, 1224.

Smissen, R.D., Galbany-Casals, M., & Breitwieser, I. (2011). Ancient allopolyploidy in the everlasting daisies (Asteraceae: Gnaphalieae) - complex relationships among extant clades. *Taxon*, 60, 649–662.

Smith, M.L., & Hahn, M.W. (2021). New Approaches for Inferring Phylogenies in the Presence of Paralogs. *Trends Genet.*, 37(2), 174–187.

Straub, S.C.K., Boutte,J. Fishbein, M., & Livshultz, T. (2020). Enabling evolutionary studies at multiple scales in Apocynaceae through Hyb-Seq. *Appl. Plant Sci.*, 8(11), e11400.

Thomas, A., Meudt, H.M., Larcombe, M.J., et al., (2023). Multiple origins of mountain biodiversity in New Zealand's largest plant radiation. *J. Biogeogr.*, 50, 947–960.

Trew, B.T., & Maclean, I.M.D. (2021). Vulnerability of global biodiversity hotspots to climate change. *Global Ecol. Biogeogr.*, 30, 768–783.

Ufimov, R., Gorospe, J.M., Fér, T., et al. (2022). Utilizing paralogues for phylogenetic reconstruction has the potential to increase species tree support and reduce gene tree discordance in target enrichment data. *Mol. Ecol. Resour.*, 22, 3018–3034.

Van Wyk, A.E., & Smith, G.F. (2001). *Regions of Floristic Endemism in Southern Africa: A Review with Emphasis on Succulents*. Umdaus Press: Hatfield.

Vences, M., Wollenberg, K.C., Vieites, D.R., & Lees, D.C. (2009). Madagascar as a model region of species diversification. *Trends Ecol. Evol.*, 24(8), 456–465.

Venter, J.C., Adams, M.D., Myers, E.W., et al. (2001). The sequence of the human genome. *Science*, 291(5507), 1304–1351.

Villaverde, T., Pokorny, L., Olsson, S., et al. (2018). Bridging the micro- and macroevolutionary levels in phylogenomics: Hyb-Seq solves relationships from populations to species and above. *New Phytol.*, 220, 636–650.

Ward, J.M., Bayer, R.J., Breitwieser, I., et al. (2009). Gnaphalieae - Systematic and phylogenetic review. In: *Systematics, evolution, and biogeography of Compositae*, pp. 537–585. Funk, V.A., Susanna, A., Stuessy, T.F., & Bayer, R.J., (eds.). International Association for Plant Taxonomy: Vienna.

Warren, B.H., Simberloff, D., Ricklefs, R.E., et al., (2015). Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. *Ecol. Lett.*, 18(2), 200–217.

Weitemier, K., Straub, S.C., Cronn, R.C., et al. (2014). Hyb-Seq: Combining target enrichment and genome skimming for plant phylogenomics. *Appl. Plant Sci.*, 2(9), apps.1400042.

Willis, K. J., Gillson, L., & Knapp, S. (2007). Biodiversity hotspots through time: an introduction. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 362(1478), 169–174.

Wilson, E.O. (1988). *Biodiversity*. National Academy Press: Washington D.C.

White, F. (1981). The history of the Afromontane archipelago and the scientific need for its conservation. *Afr. J. Ecol.*, 19, 33–54.

White, F. (1983). *Vegetation of Africa: A Descriptive Memoir to Accompany the UNESCO/AETFAT/UNSO Vegetation Map of Africa*. UNESCO: Paris.

Yan, Z., Smith, M.L., Du, P., Hahn, M.W., & Nakhleh, L. (2022). Species Tree Inference Methods Intended to Deal with Incomplete Lineage Sorting Are Robust to the Presence of Paralogs. *Syst. Biol.*, 71(2), 367–381.

Yoder, A.D., & Nowak, M.D. (2006). Has vicariance or dispersal been the predominant biogeographic force in Madagascar? Only time will tell. *Annu. Rev. Ecol. Evol. Syst.*, 37(1), 405–431.

Zhou, W., Soghigian, J., & Xiang, Q.J. (2022). A New Pipeline for Removing Paralogs in Target Enrichment Data. *Syst. Biol.*, 71(2), 410–425.

Objectives

Objectives

The primary goal of this doctoral thesis is to unravel the evolutionary and biogeographic history of the HAP clade, with a particular focus on the Afromontane and Afroalpine *Helichrysum*, whose unique characteristics make it an ideal model for studying diversification processes in one of the most iconic and diverse plant lineages originated in the Cape Floristic Region and expanded throughout the globe. *Helichrysum* successfully colonized and diversified across the full range of the African "sky-islands", offering valuable insights into the assembly of these floras, their connections with South Africa and the macroevolutionary processes underlying their exceptional diversity. It also dispersed to other continents and diversified in many mountain systems and temperate lowlands, providing an interesting example for the study of global plant diversification dynamics. The thesis intends to address the limitations identified in previous studies by increasing both taxonomic and geographic sampling and generating high-quality molecular data using the Hyb-Seq approach with the Compositae1061 target-enrichment probe set. This thesis combines multiple analytical approaches to navigate the complexities of the group's evolutionary history, resolving long-standing phylogenomic questions and laying the ground for future works.

Specific Objectives

1. **To collect an extensive sampling of *Helichrysum* and other members of the HAP clade**, covering their full geographic distribution range, morphological and ecological diversity, as well as taxonomic groups, with a particular focus on Africa and the Afromontane and Afroalpine regions.
2. **To construct robust phylogenies of the HAP clade** using high-throughput sequencing data generated following Hyb-Seq protocols.
3. Using the newly generated phylogenies, **conduct molecular dating, ancestral range reconstruction and diversification dynamics analyses** to clarify the origins and evolutionary processes underlying the distribution and origins of its diversity. Specific research questions include:
    - Determine whether **Afromontane species on the African continent** result from o**ne or multiple colonization events**. Identify their **region of origin and assess the temporal, geological and climatological contexts** in which their ancestors dispersed.
    - Investigate whether **Afroalpine species** originated **from adjacent Afromontane ancestors** that adapted and diversified in-situ **or from the long-distance dispersal** of ancestors from the summits of the **Drakensberg mountains** in southern Africa.
    - Assess whether **Afromontane species in Madagascar** arose from **single or multiple colonization events**. Determine their **region of origin and the temporal, geological, and climatological contexts** in which their ancestors dispersed.
    - Evaluate whether **diversification rates in the HAP clade remained constant over time or experienced significant changes**. If shifts are detected, determine whether they correspond temporally and spatially with **climatic events or the acquisition of key functional traits**.
    - Investigate whether **mountain radiations** in this group **align with the concept of nested (or complex) radiations**, in which interconnected parallel radiations influence each other, ultimately becoming a single radiation. In this context, explore:

        (i) The **role of geographic and ecological speciation** in shaping the group's diversification patterns.

        (ii) The **evolution** of specific **functional traits** and their **association with environmental factors or particular habitats**.

Structure of the Thesis

Structure of the Thesis

The thesis is divided into three chapters, each corresponding to a scientific article. These chapters are presented in the typical format of scientific papers.

Chapters I and II are closely related, as both mainly address the biogeographic history of various clades within *Helichrysum*. **Chapter I** has been published in *Plants* (Blanco-Gavaldà et al., 2023). It specifically investigates the temporal and biogeographical origins of montane and alpine species in Tropical Africa based on a highly resolved time-calibrated phylogeny inferred using 833 loci from 304 samples generated with the Compositae1061 probe set. The findings of this chapter were also presented at the European Conference of Tropical Ecology in Montpellier (France, 2022), the congress of the European Society for Evolutionary Biology in Prague (Czech Republic, 2022), in a scientific session at the Faculty of Biosciences of the Autonomous University of Barcelona (2023) and as a poster in the IV Symposium of Spanish Botany in León (2023). The results of Chapter I set the groundwork for the subsequent chapters.

Chapter II has been published in *Molecular Phylogenetics and Evolution* (Blanco-Gavaldà et al., 2025). It explores the evolutionary and biogeographic origins of Malagasy *Helichrysum* lineages identified in Chapter I. Sampling efforts in Madagascar were increased to cover 50% of the species. Using a probabilistic inference framework based on the comparison of different models of ancestral range evolution, we inferred the ancestral source areas of Malagasy lineages and estimated the number of long-distance dispersal events from Africa to Madagascar. These findings allowed us to propose hypotheses about the predominant speciation modes, focusing on the roles of bioclimatic niche conservation and both allopatric speciation and in-situ ecological adaptation in the diversification of *Helichrysum* on the island. The results were also presented at the Netherlands Society for Evolutionary Biology Annual Meeting in Ede (The Netherlands, 2024) and the XX International Botanical Congress in Madrid (Spain, 2024).

Chapter III is under revision (submitted to the journal *Annals of Botany*). It investigates the diversification patterns of the HAP clade in high-elevation environments worldwide. Following an integrative approach that combines multiple sources of data, we found evidence of nested radiations, in which mountain radiations not only influenced one another but also contributed to lowland diversity. These dynamics are reflected in an increase in net diversification rates during the Pleistocene. Our results suggest that Pleistocene climate oscillations shaped diversification patterns, driving allopatric speciation across different spatial scales through fluctuating connectivity. Adaptation to the varied microhabitats and the elevational gradient found in mountains also contributed to diversification through ecological speciation. The evolutionary history of several functional traits and ecological preferences, along with evidence of correlated evolution among them, highlights the clade's evolutionary flexibility, marked by high niche lability and dispersibility. These findings were also presented as a poster at the V Symposium of Spanish Botany in Granada (2024) and in the TICA Talks hosted online by the International Compositae Alliance (2025).

Chapter 1

Repeatedly Northwards and Upwards: southern African grasslands fuel the colonization of the African Sky Islands in *Helichrysum* (Compositae)

CHAPTER 1: Repeatedy Northwards and Upwards: southern African grasslands fuel the colonization of the African Sky Islands in *Helichrysum* (Compositae)

The content of this chapter has already been published in the following research article: Blanco-Gavaldà C, Galbany-Casals M, Susanna A, Andrés-Sánchez S, Bayer RJ, Brochmann C, Cron GV, Bergh NG, Garcia-Jacas N, Gizaw A, Kandziora M, Kolář F, López-Alvarado J, Leliaert F, Letsara R, Moreyra LD, Razafimandimbison SG, Schmickl R & Roquet C. (2023). **Repeatedly Northwards and Upwards: Southern African Grasslands Fuel the Colonization of the African Sky Islands in *Helichrysum* (Compositae).** *Plants*, 12(11), 2213.

Abstract

The Afromontane and Afroalpine areas constitute some of the main biodiversity hotspots of Africa. They are particularly rich in plant endemics, but the biogeographic origins and evolutionary processes leading to this outstanding diversity are poorly understood. We performed phylogenomic and biogeographic analyses of one of the most species-rich plant genera in these mountains, *Helichrysum* (Compositae-Gnaphalieae). Most previous studies have focused on Afroalpine elements of Eurasian origin, and the southern African origin of *Helichrysum* provides an interesting counterexample. We obtained a comprehensive nuclear dataset from 304 species (≈50% of the genus) using target-enrichment with the Compositae1061 probe set. Summary-coalescent and concatenation approaches combined with paralog recovery yielded congruent, well-resolved phylogenies. Ancestral range estimations revealed that *Helichrysum* originated in arid southern Africa, whereas the southern African grasslands were the source of most lineages that dispersed within and outside Africa. Colonization of the tropical Afromontane and Afroalpine areas occurred repeatedly throughout the Miocene–Pliocene. This timing coincides with mountain uplift and the onset of glacial cycles, which together may have facilitated both speciation and intermountain gene flow, contributing to the evolution of the Afroalpine flora.

Keywords: Afroalpine; Afromontane; biogeography; Asteraceae; long-distance dispersal; evolution; *Helichrysum*; phylogeny; target-enrichment

1. Introduction

Some mountain systems are highly isolated from each other by valleys and lowlands with drastically different climatic conditions, and such mountains are commonly referred to as "sky islands" (Heald, 1967). Clusters of sky islands are similar in several respects to oceanic archipelagos, providing natural replicates for studies of evolutionary and biogeographic processes that have resulted in current biodiversity patterns. A prominent example of sky island archipelagos is found in Africa, where the highest mountains host enigmatic floras with outstanding levels of endemism (White, 1978, 1983; Gehrke & Linder, 2014).

The Afromontane area is one of the biodiversity hotspots of Africa and has long been recognized as an independent floristic region (Hedberg, 1969; Rutherford & Westfall, 1986; Gehrke & Linder, 2014). It holds ≈4000 vascular plant species, of which ≈3000 are endemic (White, 1983). This area extends from the escarpment mountains of the Yemen to the high massifs of the South African Western Cape, with some outliers such as Mount Cameroon to the west, Jebel Marra to the north, and the highest mountains of Madagascar to the southeast (Davis et al., 1994; Burgoyne et al., 2005; but see Carbutt & Edwards, 2015 for a discussion). The formation of many of the highest areas was coupled with the beginning of the mid-to-late Miocene aridification and cooling, which contributed to the development of deserts and semi-desert conditions along the southwestern parts of southern Africa, as well as to the predominance of savannah and subtropical landscapes across large parts of the continent. At present, these mountains constitute temperate and humid refugia, often isolated from each other by vast expanses with markedly different environmental conditions, such as drier and warmer lowlands (Gehrke & Linder, 2009). Despite their current isolation, patches of montane forest in some of these elevated areas may have been connected during interglacial periods (Hedberg, 1969; Kebede et al., 2007). There are only a few studies focusing on the evolution of strictly Afromontane lineages (e.g. Mairal et al., 2015).

In contrast to the Afromontane flora, the Afroalpine flora has received much more attention. The Afroalpine area comprises the upper part of the highest mountains of Africa's mainland, embedded within the Afromontane area. It occurs above the tree line, usually at elevations around 3200 m above sea level (m.a.s.l.) and is characterized by alpine grasslands with gigantic *Dendrosenecio* (Hauman) B. Nord. and *Lobelia* Mill., *Alchemilla* L. shrublands, and open communities with *Erica* L. and *Helichrysum* Mill. (Gehrke & Linder, 2014). Afroalpine plants are adapted to an environment with extreme diurnal temperature variation (warm days and frosty nights), high insolation, and less precipitation than in Afromontane areas (Hemp, 2006). The Afroalpine flora harbors only 520 plant species but shows one of the highest proportions of endemism in the world, possibly due to the interplay of small areas and extreme isolation (Kandziora et al., 2022) Many species are, however, shared between massifs; single-mountain endemics are rare, representing only 15% of the flora (Gehrke & Linder, 2014).

It has long been recognized that the tropical Afromontane and Afroalpine areas have strong phytogeographic affinities to temperate areas in other parts of the world (Engler, 1904; Hedberg, 1969; Killick, 1978). A two-way migration-dispersal route connecting Eurasia and South Africa via the highest mountains of Yemen and the Drakensberg highlands of eastern South Africa, through the mountain systems of eastern Africa, has been proposed (Killick, 1978; White, 1983b; Galley et al., 2007; Bentley et al., 2014) and confirmed in many phylogenetic studies (e.g. *Alchemilla*, *Carex* L., *Ranunculus* L.: Brochmann et al., 2022). Concerning the assembly of the Afroalpine flora, it seems that immigration of cold-adapted lineages from remote temperate regions has been more frequent than immigration of lineages from the surrounding tropical regions, which necessitates niche shifts (Brochmann et al., 2022; Kandziora et al., 2022). Long-distance dispersal (LDD) appears to be the main driver shaping the current floristic and genetic diversity of the Afroalpine flora (Brochmann et al., 2022; Kandziora et al., 2022). Nevertheless, gradual expansion during glacial times may have been possible through corridors of grasslands or open forests for drought-tolerant Afroalpine species, or for Afroalpine species with low elevation boundaries (Chala et al., 2017), followed by contraction into refugia during warmer periods (Popp et al., 2008).

Colonization of tropical Afromontane and Afroalpine areas from the south following the aforementioned dispersal route has been reported for some plant groups (Galley et al., 2007; Kandziora et al., 2022). However, only a few studies have focused on this, yielding poorly resolved phylogenies (e.g. Galbany-Casals et al., 2014). Due to the lack of phylogenetic resolution, it is currently unclear whether such Afroalpine lineages originated in southern Africa and dispersed towards the north through LDD from the highest areas of the Drakensberg Mountains, which may have constituted a stepping-stone area (Popp et al., 2008), or whether they colonized the tropical Afroalpine area after expanding to the surrounding tropical lowlands and tropical Afromontane areas (Gehrke, 2018).

The Compositae is the most diverse plant family in the tropical Afromontane and Afroalpine areas, and *Helichrysum* (Gnaphalieae) is one of the most species-rich genera, with ≈60 tropical Afromontane and 12 strictly tropical Afroalpine species. The genus is considered an iconic component of these floras, dominating many plant communities, and often having beautiful and showy capitula (Figure 1). Globally, the genus comprises around 500–600 species in total, with its highest species diversity found in Africa (≈314 species (Hilliard, 1983; Anderberg, 1991; Galbany-Casals et al., 2009, 2014), where more than half of the species are montane. Southern African mountains harbor 132 species, of which 24 are endemic to highest part of the Drakensberg mountains, called the Drakensberg Mountain Centre (Carbutt, 2019), while others grow in the southern African grasslands and some extend into tropical Afromontane areas. The other species are found in various ecosystems, including savannahs, shrublands in the Cape Region, deserts, semi-deserts, and humid coastal forests. In Madagascar, there are 115 species of *Helichrysum*, almost all of them endemic, and a considerable proportion (44%) occupy montane habitats (Humbert 1962).

Figure 1. Species of *Helichrysum* illustrating the wide variety of morphologies present in the genus (Photos: Mercè Galbany-Casals; (b): Filip Kolář; (c): Cristina Roquet): (a) *H. newii* Oliv. & Hiern (tropical Afroalpine; Aberdare Mts., Mt. Elgon, Mt. Kenya, Mt. Kilimanjaro, Mt. Meru, Virunga Mts.); (b) *H. gofense* Cufod (tropical Afroalpine; Bale Mts.); (c) *H. stuhlmannii* O. Hoffm. (tropical Afroalpine; Rwenzori Mts.); (d) *H. formosissimum* Sch. Bip. (widely distributed across tropical Afromontane areas); (e) *H. argyranthum* O. Hoffm. (widely distributed across tropical Afromontane areas); (f) *H. splendidum* Less. (southern African grasslands, and widely distributed across tropical Afromontane and Afroalpine areas); (g) *H. confertifolium* Klatt (southern African grasslands); (h) *H. helothamnus* Moeser (tropical Afromontane); (i) *H. gymnocephalum* Humbert (Madagascar); (j) *H. mahafaly* Humbert (Madagascar); (k) *H. elegantissimum* DC. (southern African grasslands); (l) *H. chionosphaerum* DC. (southern African grasslands).

Previous studies suggested that the genus *Helichrysum* originated in western South Africa during the early to middle Miocene, with subsequent northward migrations to the rest of Africa including Madagascar, as well as to Eurasia (Galbany-Casals et al., 2009, 2014; Andrés-Sánchez et al., 2019). These studies hypothesized that several independent colonizations of the tropical Afromontane and Afroalpine areas occurred along a dispersal route through the Drakensberg range. The closest relatives of the tropical Afroalpine lineages of *Helichrysum* were inferred to be of tropical Afromontane, Malagasy, and southern African origin (Galbany-Casals et al., 2014; Andrés-Sánchez et al., 2019), not Eurasian as it is the most frequent scenario for tropical Afroalpine groups (see examples in Brochmann et al., 2022). However, these hypotheses were based on small taxonomic datasets and phylogenies that were poorly resolved, likely because they were based on sequences of few DNA markers, i.e., the largest *Helichrysum* phylogeny included 145 species, representing around 25% of the genus, and was based only in two nuclear DNA markers (ITS and ETS) (Galbany-Casals et al., 2014).

Our main goal is to investigate the evolutionary and biogeographic history of the genus *Helichrysum*, as an example of a plant group with southern African origins that colonized the tropical Afromontane and Afroalpine areas, in order to contribute to the understanding of the origins of the enigmatic Afromontane and Afroalpine floras. To overcome the taxonomic and geographic sampling limitations of previous studies of *Helichrysum*, we obtained samples of 304 species (representing 50% of the genus) with emphasis on the African species, and generated sequences of nuclear loci using the target-enrichment approach (Cronn et al., 2012) with the Compositae1061 probe set (Mandel et al., 2014). Moreover, we applied a methodological approach that accounts for paralogy to infer phylogenetic relationships (Ufimov et al., 2022), which we used to determine the geographic and temporal origin of the tropical Afromontane and Afroalpine *Helichrysum*. Our specific aims were to determine (1) whether tropical Afromontane and Afroalpine lineages originated from a single or multiple colonization events; (2) whether they originated from geographically close tropical lowland ancestors, which would imply a niche shift to tolerate cold in the case of Afromontane species and frost in the case of Afroalpine species, or from frost (cold)-tolerant temperate lineages that reached tropical Afroalpine mountains through LDD; and (3) in which geological and climatic context the main biogeographic processes took place.

2. Materials and Methods

2.1. Taxon Sampling

We sampled approximately 50% of the genus *Helichrysum*, including 304 species from all taxonomic groups and geographic areas, which doubles the number of species included in the largest nuclear phylogeny of the genus published up to now (Galbany-Casals et al., 2014). We also included taxa from smaller genera nested within the *Helichrysum–Anaphalis–Pseudognaphalium* (HAP) clade (Galbany-Casals et al., 2014; Smissen et al., 2020): *Achyrocline* Less. (five species), *Anaphalis* DC. (nine species), *Pseudognaphalium* Kirp. (fifteen species), *Syncephalum* DC. (one species), and the monospecific genus *Humeocline* Anderb. We chose *Syncarpha staehelina* (L.) B. Nord. as the outgroup based on previous phylogenetic results at the tribal level (Nie et al., 2016; Smissen et al., 2020).

When possible, samples were collected in the field and dried in silica-gel by team members or collaborators, and a voucher of each was deposited in the correspondent herbaria. For species that we could not collect in the field, we sampled herbarium material with the authorization of curators from various herbaria: BC, BCN, BNRH, BR, CANB, CONC, E, LP, MA, MADJ, MEXU, MO, NBG, O, P, PRE, RSA, S, SALA, SI, US, and W (acronyms from Thiers (2021), access: http://sweetgum.nybg.org/ih/; studied materials in Table S1). The identities of field-collected and herbarium-sampled specimens were determined and checked by reference to taxonomic literature (Humbert, 1962; Adams, 1963; Guillarmod, 1971; Hilliard, 1983; Lisowski, 1989; Wood, 1997; Friss & Vollesen, 1998; Beentje, 2002; Carbutt & Edwards, 2004; Tadesse, 2004; Koekemoer & Steyn, 2010; SANBI Red List, 2012; Bingham et al., 2021; Hyde et al., 2021a, 2021b, 2021c, 2021d), with type images from JSTOR Global Plants online herbarium databases and specimens at the herbaria listed above. We also added to our dataset the raw sequence reads of 21 specimens previously generated for the Czech Science Foundation GAČR project No. 20-10878S.

2.2. DNA Extraction, Library Preparation, Target Capture, and Sequencing

Between 10 and 30 mg of dried plant material per sample was homogenized with a Mixer Mill MM 301 (Retsch, Haan, Germany). DNA extractions were performed using the E.N.Z.A® SP Plant DNA Kit (Omega Bio-Tek Inc., Norcross, GA, USA) following the manufacturer's instructions. Extracted DNA was sent in parafilm-sealed 1.5 mL tubes to Daicel Arbor Biosciences (Ann Arbor, MI, USA), where library preparation, target enrichment, and sequencing was performed as follows: Total DNA was quantified using an intercalating fluorescent dye (Quant-iT PicoGreen dsDNA Assay Kit, Thermo Scientific, Waltham, MA, USA). DNA quality of a subset of samples was assessed via Agilent Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). A total of 80% of the available mass up to 4 µg was brought to 100 µL with nuclease-free water and taken through a sonication and size-selection protocol targeting an average insert length of 500 bp. Next, up to 200 ng DNA was used in a library preparation method optimized for target enrichment. Unique dual-index combinations were added to each sample via 6–8 cycles of PCR amplification. The indexed libraries were quantified with both a spectrofluorimetric assay and a quantitative PCR assay. Prior to enrichment, up to 125 ng of each library, was pooled (10-plex reactions) and dried down to 7 µL by vacuum centrifugation.

Target enrichment was performed following the myBaits v.5 protocol with an overnight hybridization using the myBaits Expert COS Compositae 1Kv1 kit (Daicel Arbor Biosciences, Ann Arbor, MI, USA). For each sample, half of the volume of beads in the elution buffer were PCR-amplified for 10 cycles. Sequencing pools were prepared by pooling the enriched libraries with the unenriched libraries at a 60:40 ratio. The final library pools were quantified again using a spectrofluorimetric assay and a quantitative PCR assay. The samples were sequenced on the Illumina NovaSeq 6000 platform on partial S4 PE150 lanes.

2.3. Molecular Data Processing and Phylogenetic Analyses

Given the paleopolyploid origin of the Compositae (Baker et al., 2008, 2016) and the evidence of allopolyploidization in the origin of part of the HAP clade (Smissen et al., 2011), accounting for paralogy was key for the analysis of nuclear loci. Accordingly, we used ParalogWizard (Ufimov et al., 2022), available at https://github.com/rufimov/ParalogWizard [accessed on 22 February 2022], a workflow that separates orthologs and paralogs of a given locus into two subgroups based on sequence similarity and creates orthologous alignments for each subgroup, which can then be used for phylogenetic tree inference (ParalogWizard is hereafter referred to as PW followed by the number of the respective script). For phylogenetic inference, we used the HybPhyloMaker workflow (Fér & Schmickl, 2018; indicated hereafter as HPM followed by the number of the respective script), available at https://githubp.com/tomasfer/HybPhyloMaker [accessed on 18 November 2021].

As a pre-processing step, we removed the adaptors and low-quality reads, trimmed the raw reads using Trimmomatic v.0.39 (Bolger et al., 2014) and removed duplicates using BBMap v.38.42 (Bushnell, 2014) applying HPM1 (reads recovered per species can be found in Table S2). The Compositae1061 probe set (Mandel et al., 2014) was used to generate the reference file for initial read mapping by selecting all sunflower sequence representatives of the targeted loci. Contigs were obtained performing a *de novo* assembly of the reads previously mapped to targets (PW1a and 1b). In order to increase mapping specificity, we then generated a customized reference file based on the best matching longest exonic sequences from our ingroup samples (PW2b). A second round of ParalogWizard was conducted, with paralogy detection turned on, and contigs were assembled to this *Helichrysum*-customized reference (PW1a and 1b). To identify paralogous sequences, pairwise exonic sequence divergence was calculated, resulting in two peaks in the resulting histogram (PW2a). The first peak represents putative allelic variation (exonic contigs with low divergence), while the second peak represents highly divergent exonic contigs, presumably originated from duplication events. Paralogous loci were then separated into two different alignments, based on a reference where the two copies of a paralogous locus were separated (PW2b). The retrieved exonic contigs were matched to this "paralog-prone" reference using BLAT v.34 (Kent, 2002), and the resulting BLAT hits were filtered (for details see Ufimov et al., 2022). Orthologous alignments were generated using MAFFT v.7.475 (Katoh & Standley, 2013), and exons were concatenated to loci (PW3). We excluded sequences missing more than 70% data for the locus, and we removed loci for which less than 75% of all samples were represented (HPM5).

Given the ongoing debate on whether concatenation or coalescence-based approaches yield more accurate phylogenetic trees (see Lozano-Fernández, 2022), we followed the standard practice of exploring both. First, we concatenated all loci in a single supermatrix and estimated the best nucleotide substitution model for each locus with ModelTest-NG (Darriba et al., 2020). Phylogenetic inference under maximum likelihood (ML) was conducted with ten independent tree searches using RAxML-NG v.1.0.3 (Kozlov et al., 2019; modified HPM8f), specifying the best-fit model for each partition and starting from five random and five parsimony-based trees. We assessed branch support through 100 bootstrap replicates and drew support values on the best-scoring ML tree using both Felsenstein's bootstrap support (BS; Felsenstein, 1985) and transfer bootstrap expectation (TBE; Lemoine et al., 2018). Clades with BS and TBE ≥ 70% were considered to be supported. For the summary-coalescent approach, individual gene trees were estimated using RAxML v.8.212 (Stamatakis, 2014) with the general-time-reversible (GTR) substitution model with a gamma distributed rate variation among sites "GTRGAMMA" (following the recommendations; Stamatakis, 2006) and 500 bootstrap replicates (HPM6a). Based on the best ML tree of these gene trees, a species tree was generated using ASTRAL III v.5.7.8 (Zhang et al., 2018; HPM7 and HPM8a). Support values were computed as local posterior probabilities (LPP) based on gene tree quartet frequencies (Sayyari & Mirarab, 2016). Branches with LPP ≥ 0.95 were considered to be supported. Resulting trees were visualized in FigTree v.1.4.3 (Rambaut, 2016).

2.4. Divergence Time Estimation

We used the penalized likelihood method (Sanderson, 2002) implemented in the software treePL v.1.0 (Smith & O'Meara, 2012) to time-calibrate the highest-likelihood phylogenetic tree. This approach is particularly suited for large genomic datasets (Smith & O'Meara, 2012). Three calibration points (CP) in total were used as node age constraints. The lack of relatively old fossils from the Gnaphalieae makes it impossible to include any primary calibration point based on fossil evidence. Two secondary CPs were extracted from Nie et al. (2016): the HAP clade crown node (CP1) 15.4 million years ago (mya) and the *Anaphalis* + Mediterranean-Asiatic *Helichrysum* crown node (CP2) 7.04 mya. Both CP1 and CP2 were used as fixed ages. Although for secondary calibration points it would be preferable to apply a normal or lognormal distribution (Ho & Phillips, 2009; Heath, 2020), this is not possible in treePL. CP3 was a primary calibration point based on geological evidence, corresponding to the emergence of the archipelago of Madeira 5.6 mya (Ramalho et al., 2015), which was applied as a maximum age at the stem node of the clade grouping the four Madeiran endemic *Helichrysum* species (Galbany-Casals et al., 2014). The dating process consisted of an initial run to detect the optimal parameter settings (opt = 4, optad = 2, and optcvad = 1). We ran a second analysis, replicating the random cross-validation procedure to identify the best smoothing value (allowing it to vary from 0.001 to 1000) using a chi-squared test. Finally, the time-calibrated tree was obtained by applying the optimal smoothing parameter (0.01).

2.5. Ancestral Range Estimation

We collected information on distribution area, altitudinal range, and habitat for each species from available floras, online resources, and databases (Humbert, 1962; Adams, 1963; Guillarmod, 1971; Hilliard, 1983; Lisowski, 1989; Wood, 1997; Friss & Vollesen, 1998; Beentje, 2002; Carbutt & Edwards, 2004; Tadesse, 2004; Koekemoer & Steyn, 2010; SANBI Red List, 2012; Bingham et al., 2021; Hyde et al., 2021a, 2021b, 2021c, 2021d), as well as from the labels of the studied herbarium vouchers (see Table S1).

We defined 15 geographic areas reflecting the main distribution patterns of the sampled species (Figure 2). For Africa, areas were mostly based on the natural vegetation regions shown in Burgoyne et al. (2005) with small modifications.

First, being aware of the phytogeographical and ecological mismatches of the Afromontane concept exposed in Carbutt & Edwards (2015), we delimited the global Afromontane area following the criteria of Gehrke & Linder (2014), adding some massifs (Chimanimani, Mount Cameroon, and part of the

Ethiopian highlands). However, for the biogeographic analysis, we subdivided it into five smaller geographical areas (from north to south): (P) the mountains of the southern Arabian peninsula; (T) the tropical Afromontane area, for species found between 1800 and 3200 m.a.s.l. in the intertropical zone; (A) the tropical Afroalpine area, for species mostly found above 3200 m.a.s.l. or, if they grow below this limit, whose upper altitudinal limit is above 4500 m.a.s.l. in the intertropical zone; (G) the southern African grasslands (below the tropic of Capricorn), including the Drakensberg montane belt, locally known as the "highveld" grasslands, for species mostly growing between 900 and 2600 m.a.s.l.; (D) the high Drakensberg area, for species whose lower altitudinal limit is above 1800 m.a.s.l. and their upper altitudinal limit is above 2600 m.a.s.l. Therefore, our high Drakensberg area is equivalent to the Drakensberg Mountain Centre (DMC) (Carbutt & Edwards, 2015; Carbutt, 2019), which comprises the upper montane belt (above 1800 m.a.s.l., previously considered the subalpine belt; Van Wyk & Smith, 2002; Rutherford et al., 2006) and the Drakensberg Alpine Center (DAC, above 2800 m.a.s.l.; Van Wyk & Smith, 2002; Rutherford et al., 2006). In this work, we considered Madagascar as a whole unique independent area (M) due to computational limitations of BioGeoBEARS, despite the fact that the highest mountainous areas of Madagascar are generally considered to be part of the Afromontane area (Gehrke & Linder, 2014). Additionally, non-mountainous areas within continental Africa were defined as follows, from north to south: (L) the tropical African lowlands, for species mostly growing below 1800 m.a.s.l. in the intertropical zone; (S) the southern African savannah; (N) the arid to semi-arid southern African area, which includes the Namib Desert, Succulent Karoo, and Nama Karoo; (C) the Indian Ocean coastal belt forest area; and (F) the Fynbos biome area. To precisely delimit the proposed southern African areas, we checked the maps included in (Van Wyk & Smith, 2002; Rutherford et al., 2006). Finally, we defined more broadly four additional areas outside continental Africa: (E) the Mediterranean area; (I) Asia; (O) the Macaronesian area; and (R) the Americas.

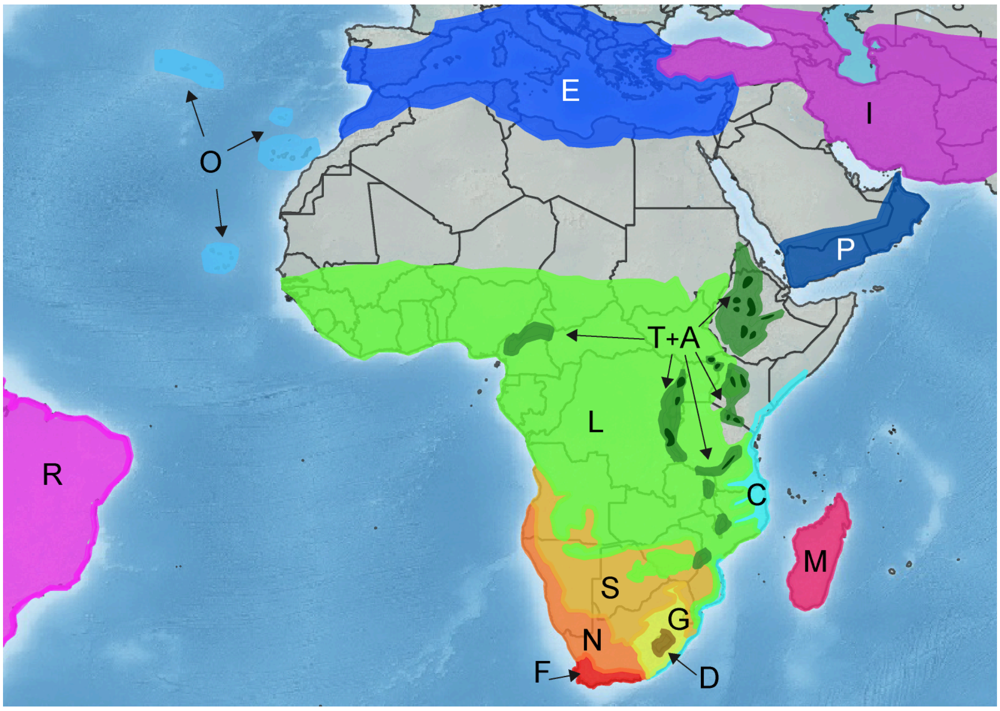


Figure 2. Geographical representation of the 15 geographic areas defined in the current study. These include six areas outside of continental Africa: E: the Mediterranean area; I: Asia; M: Madagascar; O: the Macaronesian area; P: the southern Arabian Peninsula; and R: the Americas. We defined nine areas within continental Africa (from north to south): L: the tropical African lowlands; T: the tropical Afromontane area; A: the tropical Afroalpine area; S: the southern African savannah; N: the arid to semi-arid southern African area; G: the southern African grasslands; D: the high Drakensberg area; C: the Indian Ocean coastal belt forest; and F: the Fynbos biome area. The dark green spots within the tropical Afromontane area are an overrepresentation of the tropical Afroalpine area. The colors and letters correspond to those in the biogeographical reconstruction (Figure 3).

Species occurrence in the geographical areas was assigned considering habitat information in combination with distribution and altitudinal range. Some species are only occasionally or marginally present in other areas outside of their main distribution range. In these cases, we only considered the core area(s) or altitudinal belt in which there are a significant number of occurrences of the species, or that can be interpreted as the central distribution area of the species.

We estimated geographic range evolution with the R package BioGeoBEARS v.1.1.2 (Matzke, 2013) available at https://github.com/nmatzke/BioGeoBEARS [accessed on 25 August 2022], using as input the time-calibrated tree after pruning the outgroup. The maximum number of areas for any node was set to three, which is the highest number of areas occupied by the most widespread extant taxon in our study. We tested the fit of three biogeographical models: the dispersal–extinction–cladogenesis model (DEC; Ree et al., 2005; Ree & Smith, 2008); a likelihood implementation of the dispersal–vicariance analysis model (DIVA, hereafter DIVAlike; Ronquist, 1997); and BAYAREAlike (Landis et al., 2013), a model that decouples range evolution from cladogenesis. In addition, a more complex version of each of the mentioned models was tested through the addition of the jump-dispersal parameter j (DEC+j, DIVAlike+j, BAYAREAlike+j), which depicts founder-event speciation (range switch that occurs at a lineage-splitting event, resulting in one daughter lineage in a new range and the other daughter lineage retaining the ancestral range; Matzke, 2014). We assessed model fit with the Akaike Information Criterion (AIC) and AIC weights (AICwt). We also conducted a Biogeographic Stochastic Mapping (BSM) analysis (Dupin et al., 2017) with 100 replicates to obtain the overall means of anagenetic and cladogenetic events conditional on the geographic distributions, the phylogeny, and the best-fitting model.

3. Results

3.1. Alignment Processing and Filtering

For each locus, we considered sequence divergence values between 7.6% and 18.9% to indicate paralogy (Figure S1). Using the *Helichrysum*-customized reference, 225 paralogous loci were detected on average (Table S3), with little variation among taxa (sd ± 3.4). We recovered 927 loci, of which 833 were retained and used to build orthologous alignments after filtering for missing data (Table S4). Alignment length for each locus averaged 318 bp (range 49–816 bp; Table S5). On average, each alignment had 116 variable sites and 73 parsimony informative sites.

3.2. Phylogenetic Analyses

Concatenation (RAxML-NG; Figure S2) and summary-coalescent (ASTRAL; Figure S3) approaches yielded species trees with similar topologies, especially for the deepest and shallowest nodes. Deep nodes were well supported in the concatenation-based analyses (BS and TBE metrics), but less supported in the summary-coalescent tree. Most of the shallow nodes were strongly supported by all metrics (in both the concatenation and summary-coalescent analyses). In sum, 79% of the nodes were supported with $BS \geq 70\%$, and 69% of these nodes received 100% BS support; 90.2% of the nodes were TBE supported, and 60.7% of these nodes received a TBE support of 1. In the phylogeny obtained using ASTRAL, 59% of the nodes were supported ($LPP \geq 0.95$), and 82.2% of these nodes obtained the maximum support ($LPP = 1$).

3.3. Divergence Time and Ancestral Range Estimation

The treePL analysis estimated that the HAP clade originated 15.4 mya and started to diversify 14.2 mya (Figures 3 and S2; the former figure is a partially collapsed phylogenetic tree to make it easier to read, Figure S2 is the complete tree). In the biogeographic analysis, DEC+j was the best fitting model according to AIC (AICc > 2 points lower than the second-best model; AICcwt = 1). Founder–event speciation

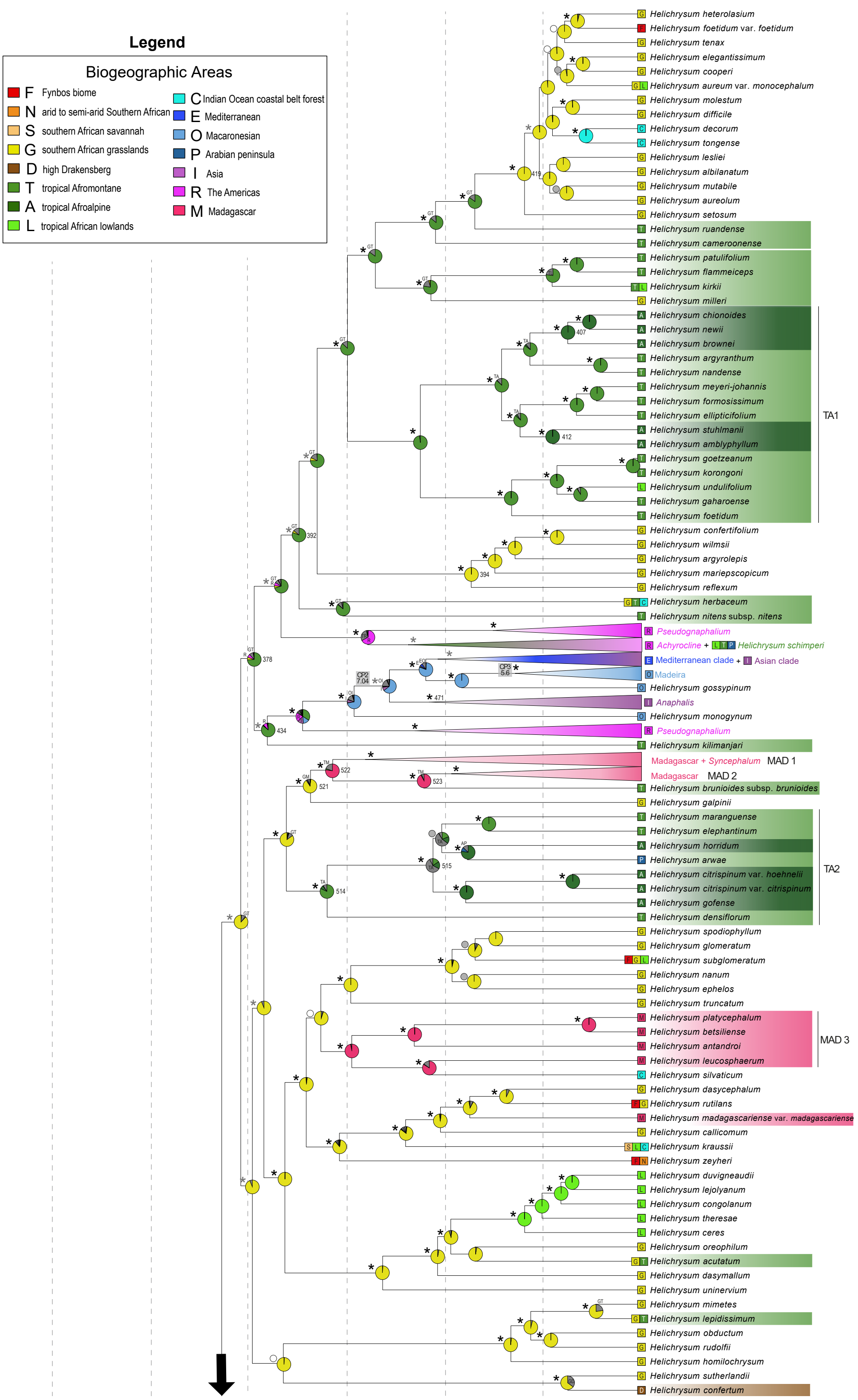
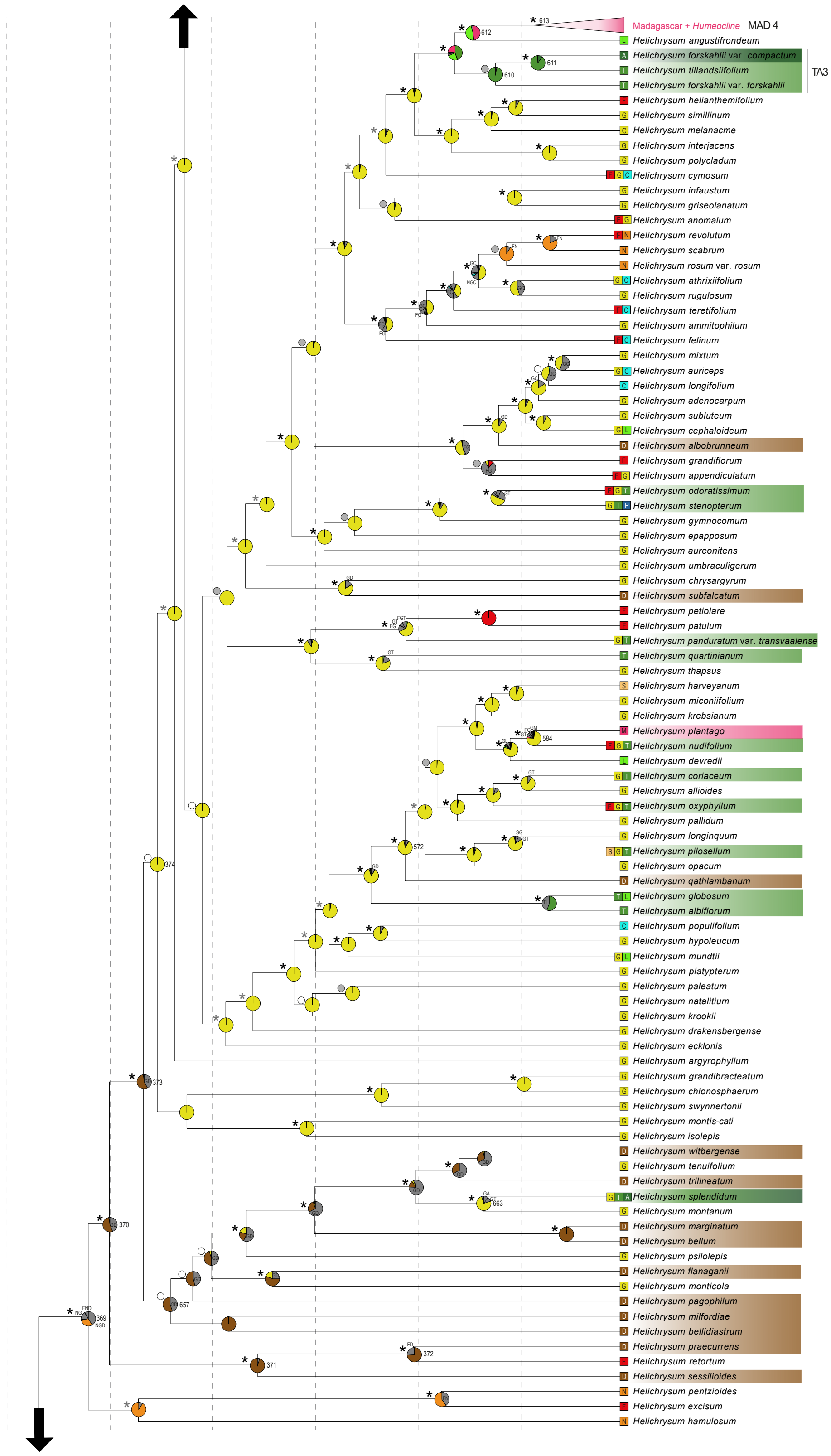
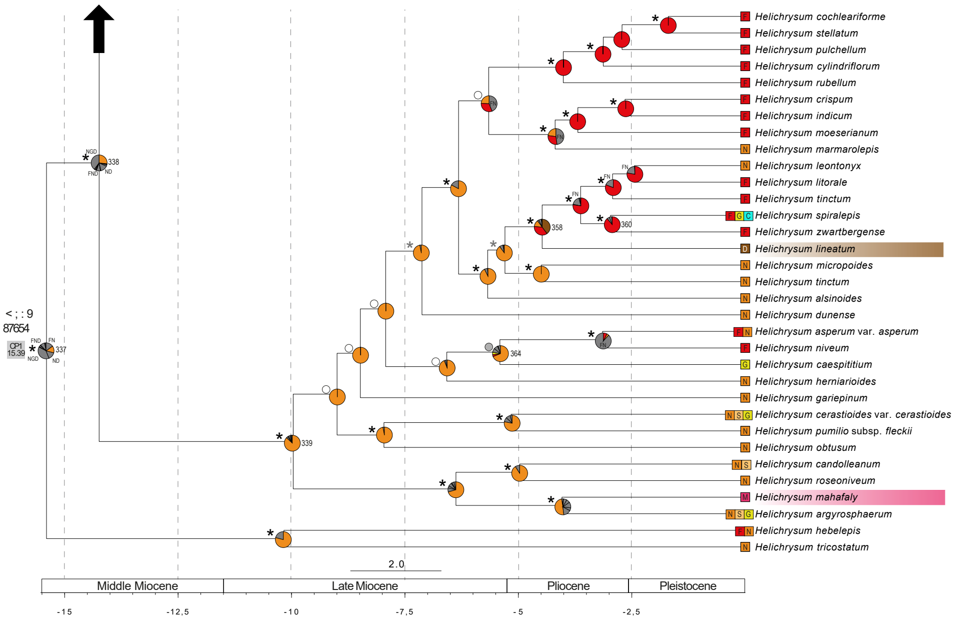


Figure 3. Ancestral range estimation of the HAP clade using the best-fitting model DEC+j based on a time-calibrated phylogeny generated under the concatenation approach using target-enrichment data (Compositae1061 probe set). Pie charts at nodes show the relative probability of the possible states (areas in primary colors; combinations of areas in grey). Relevant node numbers are indicated to the left of the node. Black asterisks indicate nodes that were strongly supported by both metrics (BS and TBE ≥ 95%/0.95), grey asterisks indicate nodes strongly supported by only one metric (BS or TBE ≥ 95%/0.95), grey circles indicate nodes moderately supported by both metrics (BS and TBE from 70%/0.70 to 94%/0.94), and empty circles indicate nodes moderately supported by only one metric (BS or TBE from 70%/0.70 to 94%/0.94). Highlighted species correspond to tropical Afromontane clades in green, which include tropical Afroalpine clades in dark green (labelled as TA1, TA2, TA3); high Drakensberg species in brown; Malagasy species in magenta (main clades labelled as MAD1, MAD2, MAD3, MAD4). Other genera and large Malagasy clades are collapsed. The complete tree obtained in this analysis is shown in Supplementary Figure S2.



(***Figure 3.***, *cont.*)



***(Figure 3.**, cont.)*

processes (j = 0.0065) had a larger contribution than range expansion by dispersal (d = 0.0027) and range contraction (e = 1 × 10^{-12}) (Table S6, Figure S2). In the last years, there has been a controversy regarding the inclusion of the founder-event parameter (+j) in event-based models. Specifically, Ree & Sanmartín (2018) claimed that the comparison of DEC and DEC+j models is statistically invalid, based on two small hand-constructed datasets. Recently, Matzke (2022) presented a thorough simulation-based study that validates the log-likelihood comparison of DEC and DEC+j. In our case, the results based on the DEC model provided in some cases unrealistic range estimations (Figures S5 and S6, Table S9): for instance, the ancestor of the clade which includes numerous Malagasy species (from *H. stilpnocephalum* to *H. chamaeyucca,* MAD1 and MAD2 in Figure 3) and only one Afromontane species (*H. brunioides*), is inferred to have a distribution in the Afromontane area + Madagascar. Another example concerns the ancestor of *Pseudognaphalium+Achyrocline*, which is inferred to have a distribution in the Afromontane area + America in the DEC model (whereas the DEC+j model infers only America as the ancestral range). These two cases, among others, exemplify the relevance of the +j parameter in a study case as the one presented here, in which many disjunct distributions are much more probably originated by dispersion and founder event given the recent age of the studied group and the long-term isolation of some of the territories: the separation of America from Africa dates from around 100 mya (Seton et al., 2012) and the separation of Madagascar from Africa dates from 165 mya onwards (Yoder & Nowak, 2006). In a similar way, most of the current Afromontane areas have been isolated for a long time, and Afroalpine areas from different massifs have never been in contact. For all these reasons, we only discuss the results obtained under the DEC+j model, the best-fitting model for our data according to AICc.

Based on the inferred ancestral range of the basal nodes (probabilities in Table S7), the origin and initial diversification of the genus took place in the arid to semi-arid southern Africa area (N), possibly involving one or more other areas (the southern African grasslands, G; the Indian Ocean coastal belt forest, C; the high Drakensberg area, D; and/or the Fynbos biome, F). The basal-most node separated a clade of two arid southern African taxa (N) from the remaining taxa, which formed two main lineages (Figure 3). The smallest lineage (node 339, BS = 100%, TBE = 1) diversified within the ancestral area (N) but also spread to the Fynbos area (F) around 5.65 mya (BS = 57%, TBE = 0.785). The largest lineage (node 369, BS = 100%, TBE = 1) split into two clades approximately 13 mya, a small one that remained in western South Africa, and another (node 370, BS = 98%, TBE= 0.99) that diversified in the high Drakensberg area and the southern African grasslands. Diversification in the southern African grasslands included multiple dispersals towards the east and the north that took place in different time periods, comprising *Helichrysum* species from all geographic areas, including Madagascar and other continents. In this clade, the analyses also recovered back-colonizations from Madagascar to mainland Africa, as well as dispersals from the tropical Afromontane area to the surrounding arid tropical African lowlands.

Biogeographic stochastic mapping (BSM) analyses show that *Helichrysum* reached the tropical Afromontane area at least 13 times, with the mean number of events being 17.77 (Figure 4; BSM frequency distributions in Figure S4; BSM summaries in Table S8), mainly from the southern African grasslands (G). This number included colonization events that resulted in the origin of new tropical Afromontane endemic species and range expansions of widespread species. Most of the tropical Afromontane species were found in three lineages (Figure 3). The largest lineage (TA1) started to diverge 9.9 mya (node 378, unsupported) and gave rise to new lineages that colonized regions outside Africa (mainly Eurasia, the Americas, and Macaronesia). The TA1 lineage diversified in the tropical Afromontane area as well (node 392, BS = 88%, TBE = 0.9), reached the tropical Afroalpine area twice independently (nodes 407 and 412, both with BS = 100%, TBE = 1), and returned twice to the southern African grasslands (nodes 394 and 419, both with BS = 100%, TBE = 1). Another tropical Afromontane lineage (TA2) originated 8 mya (node = 514, BS = 100%, TBE = 1); it included several tropical Afroalpine species in two lineages and one species from the Arabian Peninsula. The third tropical Afromontane lineage (TA3) originated 3.13 mya (node = 610, BS = 80%, TBE = 0.9) and gave rise to the tropical Afroalpine *H. forskahlii* var. *compactum* (Vatke) Mesfin. Some tropical Afromontane species occurred scattered across the tree (e. g. *H. quartinianum* A. Rich, *H. albiflorum* Moeser) and mainly originated from southern African grasslands ancestors. Additional colonizations of the tropical Afromontane area took place recently, exemplified by species with a large distribution area extending from South Africa to the tropical Afromontane area (e.g. *H. stenopterum* DC., *H. odoratissimum* Sweet, *H. oxyphyllum* DC., *H. coriaceum* Harv., *H. pilosellum* Less., *H. lepidissimum* S.Moore, *H. acutatum* DC.).

Helichrysum reached the tropical Afroalpine area at least four times from tropical Afromontane ancestors. One colonization led to two tropical Afroalpine lineages found in clade TA1: one with *H. amblyphyllum* Mattf. and *H. stuhlmannii* O.Hoffm. (node 412, BS = 100%, TBE = 1), which split 2.3 mya, and one with *H. newii, H. chionoides* Philipson and *H. brownei* S.Moore (node 407, BS = 100%, TBE = 1), which split 1.9 mya. A third tropical Afroalpine lineage was nested within TA2 and comprised the Ethiopian species *H. gofense* Cufod. and *H. citrispinum* Delile. A fourth tropical Afroalpine lineage was only represented by the Ethiopian *H. horridum* Sch.Bip., also embedded in clade TA2 and sister to the Arabian species *H. arwae* J.R.I.Wood; this lineage diversified 5.3 mya (node 515, BS = 100%, TBE = 1). *Helichrysum forskahlii* var. *compactum* (TA3) represented the fourth colonization of the tropical Afroalpine area from a tropical Afromontane ancestor. Lastly, *H. splendidum* has a wide disjunct distribution area in the mountains from the southern African grasslands to the tropical Afroalpine area.

Notably, the tropical Afroalpine species were not found to be closely related to the alpine taxa in the high Drakensberg area (D). Some of the high Drakensberg species were resolved in a lineage (node 657, BS = 24%, TBE = 0.924) that started to diversify 11 mya and also included southern African grasslands species. Another lineage included two high Drakensberg species (*H. sessilioides* Hilliard and *H. praecurrens* Hilliard) plus the Fynbos species *H. retortum* Willd. (node 372, BS = 100%, TBE = 1). Five other high Drakensberg species were found scattered across the tree (*H. albobrunneum* S.Moore, *H. subfalcatum* Hilliard, *H. qathlambanum* Hilliard, *H. confertum* N.E.Br., and *H. lineatum* Bolus). Four of them originated from ancestors in the surrounding grasslands, whereas *H. lineatum* probably diverged from a Fynbos (F) ancestor 4.5 mya (node 358, BS = 100%, TBE = 1). In the high Drakensberg area, most speciation events took place between 11 and 1.4 mya.

Helichrysum colonized Madagascar at least six times. The oldest Malagasy (MAD1 + MAD 2) lineage dispersed there from Africa ca. 7.9 mya (node 522, BS = 100%, TBE = 1). The youngest Malagasy lineage

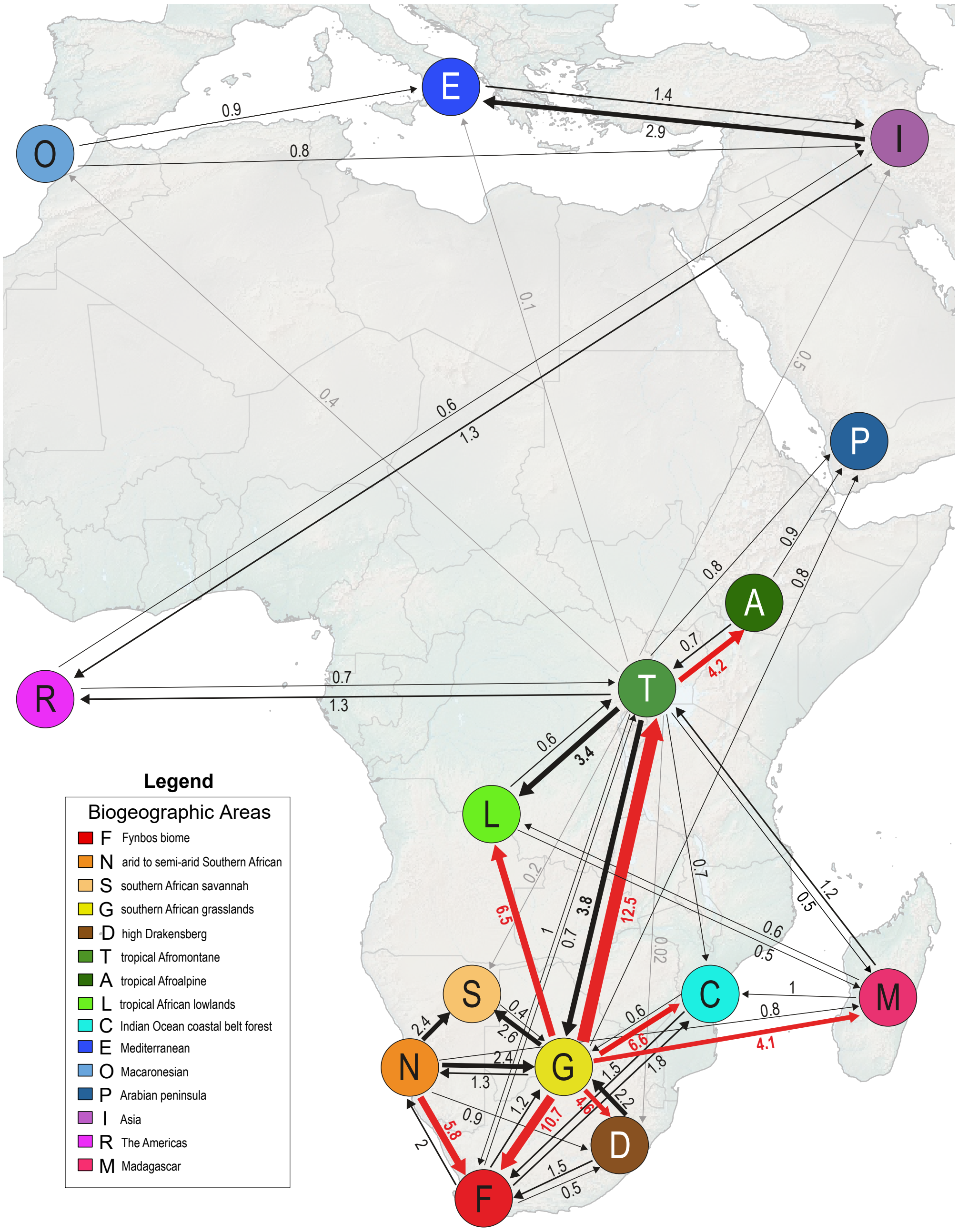


Figure 4. Summary of mean dispersal events estimated from 100 biogeographic stochastic mappings in *Helichrysum* (see all event counts in Table S8). Arrow tips indicate the directionality of the dispersals. Numbers on the arrows are the mean of dispersal event counts. Arrow thickness is proportional to the mean number of dispersals: gray arrows correspond to mean counts below 0.5 (for readability, only the ones involving the eastern African montane area are shown), black arrows represent mean dispersal events ranging from 0.5 to 4 counts, and red arrows show the most common dispersal pathways (mean above 4 counts).

(MAD4, node 613, BS = 100%, TBE = 1) dispersed from Africa around 3.7 mya and notably diversified in Madagascar 2.2 mya.

3.4. Number, Type, and Directionality Estimation of Biogeographical Events

We estimated from a summary of 100 BSM replicates that within-area speciation events predominated in *Helichrysum* (62.8% of speciation events), followed by dispersal events leading to diversification (28.4%), whereas vicariant events were rare (Table 1). Within-area speciation events appeared to have occurred mainly in the most species-rich areas, the southern African grasslands (G, 32%), and Madagascar (M, 23.5%). Founder events (12.5%) were almost as frequent in the history of *Helichrysum* as range expansions (15.9%).

Table 1. Summary of 100 biogeographic stochastic mapping counts for all *Helichrysum* ranges using the DEC+j model. Mean number of events estimated are shown along with standard deviations (sd). Range-switching dispersals (a) and range contractions (e) are not shown because those parameters were not required in the best fitting model (DEC and related). *Within-area speciation, subset type: the new species occupies a subset of the ancestral range. For more information on the modes, visit http://phylo.wikidot.com/biogeobears (accessed on 5 September 2022).

Mode	Type	Mean (sd)	% of Biogeographic Events
Within-area speciation	Sympatry	250.2 (3.95)	62.8
	Subset *	23.38 (5.2)	5.9
Dispersal	Founder events	49.9 (4.28)	12.5
	Range expansions	63.18 (2.64)	15.9
Vicariance	Vicariance	11.56 (2.85)	2.9
Total		398.2 (2.64)	100

The relative importance of an area as source or sink for dispersals was similar, independently of the type of dispersal (range expansion vs. founder events). The most frequent dispersals occurred from the southern African grasslands (G) towards the adjacent areas, with the Fynbos biome (F) and the tropical Afromontane area (T) being the most common sinks. Madagascar (M) was the only area where lineages appeared to have originated mainly by founder event speciation rather than range expansion, and these founder events seemed to have originated mainly from southern African grassland ancestors. Movements between the tropical Afromontane area (T) and the southern African grasslands (G) were highly asymmetric, with migrations from G to T being four times more common than from T to G. The high Drakensberg species (D) seemed to derive mainly via colonization from the surrounding grasslands and were never inferred to have spread further north. All tropical Afroalpine species descended from tropical Afromontane ancestors.

4. Discussion

4.1. Utility of Target-Enrichment Strategies in Reconstructing the Radiation of Helichrysum

Resolving relationships of young rapidly radiating groups is a well-known difficult task in the field of plant systematics. Researchers have to deal with high levels of tree discordance, consequences of evolutionary processes (i.e., ILS, hybridization, whole-genome duplication), and methodological artifacts. Hyb-Seq and related target-enrichment approaches have proven to be efficient in generating genomic data to deal with complex groups (Weitemier et al., 2014). Recent studies, aware of the gap between advances in sequencing techniques and data analysis approaches to deal with such incongruences, suggest some guidelines to yield more accurate phylogenomic inferences and improve our understanding of evolution (e.g. *Burmeistera* H.Karst. & Triana, Bagley et al., 2020; *Dendrosenecio*, Gizaw et al., 2021; Lobelioideae, Lagomarsino et al., 2022; and *Loricaria* Wedd. Kandziora et al., 2022b). The target-enrichment dataset generated here, coupled with a recent pipeline that detects and uses paralogs for phylogenetic reconstruction (ParalogWizard. Ufimov et al., 2022), allowed us to build the first highly resolved phylogeny of the HAP clade. This phylogeny is based on 833 loci and wide taxonomic sampling (304 species, representing ≈50% of the diversity in *Helichrysum*, and 31 species from smaller, closely related genera). Since the taxonomic sampling is not yet complete, we cannot discard that increased taxon sampling may affect phylogenetic inference and biogeographic estimation; however, to minimize this potential bias, our current sampling was designed to evenly represent all major taxonomical and biogeographical groups and was highly increased in relation to previous works. The most comprehensive nuclear phylogeny of *Helichrysum* previously published included only 145 species, which represented ≈25% of the diversity in *Helichrysum* and was based on nrDNA ITS and ETS markers (Galbany-Casals et al., 2014). Despite the different sampling, a similar general pattern has been recovered: (1) the early divergent lineages of *Helichrysum* are from SW Africa; (2) the Afromontane and Afroalpine species appear in several independent clades, i.e., they have multiple independent origins; (3) the Malagasy species appear in several clades, i.e., they have multiple independent origins; (4) the genera *Achyrocline*, *Anaphalis*, *Pseudognaphalium*, and *Humeocline* are nested in the main HAP clade; (5) the genus *Anaphalis* is sister to the Mediterranean–Asiatic–Macaronesian *Helichrysum*. However, the statistical branch supports are notably higher in the present approach, and some differences have emerged in the relationships obtained, among which we highlight the following: (1) although the early divergent lineages are from SW Africa in both phylogenies, they are grouped differently, i.e., they constituted several small clades placed in a grade at the base of the previous phylogenies; (2) the sister lineages of the Afromontane species were not resolved or statistically supported in the previous study, whereas in the current work they are; (3) the genus *Pseudognaphalium* constituted a unique clade in previous phylogenetic studies, whereas in the present one the species of this genus are distributed in two main clades, one of which is grouped with *Achyrocline*, and the other with *Anaphalis* and the Mediterranean–Asian–Macaronesian species of *Helichrysum*.

The topologies obtained with the concatenation approach (RAxML-NG tree; Figure S2) and the summary-coalescent approach (ASTRAL tree; Figure S3) are largely congruent, especially for the deepest and shallowest nodes. The main topological differences involve intermediate nodes with ages 12–8 mya, but these are not or only weakly supported in both topologies and have very short branches, suggesting rapid diversification events that are difficult to disentangle (Whiterfield & Lockhart, 2007). Therefore, the divergence times and biogeographic estimations obtained for these nodes should be taken with caution. These events seem to have occurred during the major uplift of the East African Rift system and the subsequent aridification of eastern Africa in the late Miocene (Sepulchre et al., 2006). Successive rapid speciation events are often characterized by strong ILS signatures at many loci (Maddison, 1997), which along with hybridization reduce phylogenetic signals. This effect magnifies with genome-scale datasets (Linkem et al., 2016). Moreover, the Malagasy clades MAD2 (node 523) and MAD4 (node 613) appear to have undergone rapid, but more recent (<2 mya) diversification, resulting in incongruences (Figures S2 and S3). Notably, the relationships between the tropical Afromontane and tropical Afroalpine species are virtually identical in both phylogenies.

4.2. The Early History of Helichrysum and Colonization of Madagascar

Our analyses place the ancestor of the HAP clade in the arid to semi-arid southern African areas during the mid-Miocene (ca. 15 mya; Figure 3). This dating coincides with the beginning of a global cooling trend (Westerhold et al., 2020) and the onset of more arid climates in the western parts of southern Africa. The dry summers likely contributed to a notable floristic change from the former subtropical humid forests and woody grasslands (Coetzee & Rogers, 1982; Dupont et al., 2011). Previous studies inferred the Greater Cape Floristic Region (constituted by the Fynbos biome and the Succulent Karoo biome) as the center of origin of the HAP clade (Andrés-Sánchez et al., 2019). Our new result arises because we added two newly sequenced species from the arid areas of western southern Africa that appear to be at the base of the group, and also because we delimited the biogeographic areas differently, uniting the arid winter-rainfall Succulent Karoo with the arid summer-rainfall regions of southern Africa (Nama Karoo, Namib desert) and not with the Fynbos as the former study did. However, there is no discrepancy between the two reconstructions if the northern part of the Succulent Karoo ("Namaqualand") is considered as the ancestral area for the HAP clade.

One of the two main *Helichrysum* clades first diversified in the arid areas of western southern Africa and subsequently colonized and diversified in the adjacent Fynbos biome. In contrast, the second main clade dispersed rapidly eastwards in South Africa in the middle to late Miocene, colonizing the southern African grasslands (G) and the high Drakensberg area (D). Several additional colonizations of these grasslands from arid-adapted ancestors occurred more recently. The colonizations of both the southern African grasslands and the high Drakensberg were followed by in situ diversifications, likely fostered by the opening of new habitats following the Miocene–Pliocene Drakensberg uplift (Partridge & Maud, 1987, 2000; Partridge, 1998; Carbutt & Edwards, 2004) and coinciding with many other major radiations during that time (see Hughes & Atchison (2015) for a global review on montane and alpine plant groups; also in other organisms such as rodents (Mikula et al., 2021) and butterflies (Aduse-Poku et al., 2022)). The biogeographic connection of southwestern Africa and the eastern African mountain systems is not surprising because even today, the arid southern African areas experience local frosts (Schulze, 2001), and this was likely more severe during past times, meaning that previous *Helichrysum* lineages could already be tolerant to cold conditions, which potentially would have facilitated the colonization of mountain habitats, characterized by cold climatic conditions.

Helichrysum's grasslands ancestors also dispersed multiple times to adjacent southern African areas, lowland tropical Africa, and Madagascar during the late Miocene to Early Pliocene. Dispersals from the southern African grasslands to other areas were often followed by in situ diversification and in some cases by back-colonization events, suggesting a highly dynamic biogeographic history for *Helichrysum* (Figure 4, Table S8).

Our study also provides new insights into the biogeographic history of the Malagasy species. A previous study only included 18 species from Madagascar and was based on few DNA markers, resulting in poorly resolved phylogenies (Galbany-Casals et al., 2014). Here, using many markers and including 65 Malagasy species (57% of the island's species), we identified at least six independent dispersal events from mainland Africa to Madagascar, with the southern African grasslands as the main source area (Figure 4). The first colonization of Madagascar is inferred at around 7.9 mya, and the most recent one around 3.7 mya. Since Madagascar was well separated from the African continent already in the early Cretaceous (Yoder & Nowak, 2006), all colonizations by *Helichrysum* probably occurred via wind- or water-mediated LDD across the Mozambique Channel (Masters et al., 2021). Most Malagasy *Helichrysum* colonized the cool temperate highlands. However, *H. mahafaly* colonized Madagascar around 4 mya from the arid southwestern African area. It is the only Malagasy species occurring in the arid southwestern lowlands, demonstrating niche conservatism. Notably, we identified two back-colonizations from Madagascar to mainland Africa around 5.5 mya, giving rise to *H. brunioides* Moeser in the tropical Afromontane area and *H. silvaticum* Hilliard in the forests of the southern African east coast.

4.3. Repeatedly Northwards

The tropical Afromontane area has been colonized independently at least 12 times from the southern African grasslands (Figure 4, Table S8), and at least once from Madagascar (*H. brunioides*, node 523). Thus, our results support the hypothesis that the southern African grasslands area harbored a fruitful floristic radiation that directly fed the tropical Afromontane flora (Linder & Verboom, 2015) and played an important role as a transitional environment (Gehrke, 2018; Carbutt, 2019). This biogeographical connection from the Cape region northwards, in which the southern African grasslands act as a stepping-stone area (Killick, 1978: White, 1983b; Galley et al., 2007; Bentley et al., 2014) has been reported for other plant groups, such as *Disa* Bergius (Orchidaceae) and the tribe Irideae (Iridaceae), with species numbers decreasing west- and northwards on the African continent as a general pattern (Galley et al., 2007) as found in *Helichrysum.*

The biogeographic assembly of the tropical Afromontane flora was probably deeply influenced by orogenic events and past climatic changes, as suggested for several plant groups (Gehrke & Linder, 2014). *Helichrysum* colonized the tropical Afromontane area repeatedly from the late Miocene to the Pleistocene. Some of the high mountains in eastern Africa were already present during the Oligocene (Sepulchre et al., 2006; Gehrke & Linder, 2014). However, major uplift events occurred in the Pleistocene, associated with the Great Rift Valley formation (Chorowicz, 2005; Sepulchre et al., 2006). Considering the habitat currently occupied by tropical Afromontane *Helichrysum* species (i.e., open forest glades), the climatic cooling trend during the late Miocene (Zachos et al., 2008) and the cold and arid conditions during glacial periods might have fostered a gradual expansion and later isolation of several lineages from the southern African grasslands towards the eastern African mountains.

During cold and dry periods, the extent and altitudinal range of forested areas decreased because the lower altitudinal limit was pushed up by the expanding arid open vegetation, while the expanding glaciers pushed the alpine and ericaceous belts downhill (Kebede et al., 2007). The increased geographical isolation between the tropical Afromontane areas and the southern African grasslands during unfavorable periods probably fostered the differentiation of tropical Afromontane endemic lineages through allopatric speciation. However, most tropical Afromontane *Helichrysum* species have wide distribution ranges that span several currently isolated mountain massifs, which is consistent with cyclic formation of forest bridges or at least reduced dispersal distances between patches of Afromontane formations during recurrent interglacials (Kebede et al., 2007). In a few cases, expansion from southern Africa appears to have been more recent, since several species have wide distribution ranges that include both the southern African grasslands and the tropical Afromontane area. For instance, *H. splendidum* (node 663, 3.42 mya) is nested within the main alpine-adapted high Drakensberg lineage and occurs in the southern African grasslands and tropical Afromontane and Afroalpine areas, and *H. nudifolium* (L.) Less. (node 584, 2.22 mya) occurs in the southern African grasslands as well as in the tropical Afromontane area and the Fynbos biome area.

In sum, the present-day distribution of *Helichrysum* in the tropical Afromontane area is probably a result of the increased topographical complexity during mountain uplift periods, followed by cycles of colonization, speciation, and extinction during the Pleistocene climatic oscillations, including extinction of ancestral transitional populations (Brochmann et al., 2022; Kandziora et al., 2022). One of the main characteristics of *Helichrysum* species is their broad niches, potentially hosting the variation needed for rapid adaptation. These features probably confer these species with adaptive versatility that has facilitated repeated environmental switches and geographic movements.

Interestingly, the tropical Afromontane area, apart from being the recipient of multiple immigrant lineages and hosting several in situ radiations, has probably been the main source of extra-African *Helichrysum* lineages. Our results indicate that both the northern hemisphere *Helichrysum* species (the Asian, Mediterranean and Macaronesian clade; node 434, BS = 36%, TBE = 0.985) and the Asian genus *Anaphalis* (node 471, BS = 100%, TBE = 1) descend from tropical Afromontane ancestors. However, as it is clearly visible from the number of connections among areas (Figure 4) that the Saharan and Arabian deserts have been crossed rarely, and always northwards in this plant group. Such result emphasizes the role of these deserts as biogeographic barriers and indicates that the mountains located in them (Hoggar Mountains, Tassili n'Ajjer Range, Tibesti Mountains, among others) did not play a significant role as steppingstones. Other extra-African radiations that possibly were sourced from tropical Afromontane area include one Malagasy group (MAD4, node 613) and the mostly American genera *Pseudognaphalium* and *Achyrocline*. The large genus *Senecio* (Asteraceae, tribe Senecioneae) shares a similar biogeographic historical pattern with *Helichrysum* (Kandziora et al., 2016, 2017), since it also originated in arid areas of western southern Africa and has been hypothesized to have colonized the Palearctic and the Americas multiple times from different ecosystems of the African continent.

4.4. Repeatedly Upwards

All tropical Afroalpine clades recovered in our *Helichrysum* tree are inferred to have originated from tropical Afromontane ancestors, not from ancestors preadapted to alpine conditions in Drakensberg or Eurasia. This scenario contrasts with that of most tropical Afroalpine lineages already adapted to cold climates with a common presence of frost and closely related to northern lineages (e.g. *Arabis alpina* Krock. ex. Steud., Assefa et al. (2007); *Cardamine* L., Carlsen et al. (2009); *Ranunculus* Gehrke & Linder (2009); *Carex* Gehrke & Linder (2009, 2014); *Alchemilla* Gehrke et al. (2016); *Lychnis* L. Gizaw et al. (2016);

Anthoxanthum L. Tusiime et al. (2017)). In contrast to all these examples, the Mediterranean–Asiatic *Helichrysum* species, which are shown to constitute a unique and highly supported clade (this work and Galbany-Casals et al., 2009) have not fed the Afromontane and Afroalpine floras but the opposite. Our results indicate that colonization of the alpine areas in the Drakensberg as well as in tropical Africa happened several times independently from *Helichrysum* ancestors occurring in the surrounding lower-elevation montane areas. Thus, repeated parallel adaptation from montane to alpine habitats has prevailed in *Helichrysum*. Indeed, most tropical Afroalpine *Helichrysum* species from different clades share particular morphological traits that suggest convergent adaptive evolution (Figure 1a–c): (i) their capitula are larger than in non-alpine species, possibly evolved to attract scarce pollinators; (ii) their capitula are white (instead of yellow, brown, pink, or reddish), the most attractive color for Diptera and Coleoptera, which were the dominating pollinators we observed in the Afroalpine area during the field expeditions; (iii) their dense pilosity, which protects against high U.V. radiation; and (iv) their nyctinastic capitula (i.e., capitula that close during the night), thus protecting the florets from night frost.

Both the high Drakensberg area and the tropical Afroalpine areas are mainly colonization sinks (as reported for *Erica* in the Drakensberg range; Pirie et al., 2019) because dispersal to other areas has not been detected or is exceptional. However, three of the tropical Afroalpine lineages of *Helichrysum* diversified in situ after colonization, probably by allopatric speciation. The age estimated for tropical Afroalpine species endemic to a single mountain range suggest that they emerged after periods of volcanic mountain-building activity. For instance, *H. brownei* diverged from its sister species approximately 1.9 mya on Mt. Kenya, which underwent an uplift 2.6 mya (Chorowicz, 2005; Sepulchre et al., 2006).

Establishment of *Helichrysum* in Afroalpine areas could only have happened from the time the alpine habitat became available after mountain uplift, estimated from the Miocene onwards (Axelrod & Raven, 1978; Harmsen et al., 1991). The alpine habitat was first available in the Drakensberg in the mid-Miocene after a major uplift period (Partridge & Maud, 1987, 2000; Partridge, 1998), and the formation of alpine patches accelerated during the Plio-Pleistocene, when the eastern African mountains reached their current altitudes (Gehrke & Linder, 2014). The establishment of *Helichrysum* in alpine habitats was probably favored by its preference for open and arid habitats rather than dense forests. The repeated shifts in *Helichrysum* from a geographically proximate but non-alpine biome rather than recruitment via LDD from remote but ecologically more similar alpine biomes suggest that adaptation to extreme daily temperature fluctuations, frosty nights, and high insolation may be more straightforward than often thought (Gehrke, 2018; Brochmann et al., 2022).

Long-distance dispersal of *Helichrysum* appears nevertheless to have been important within the tropical Afroalpine area. LDD was probably facilitated during glacial maxima, when the alpine belts were found at lower elevations (up to 1000 m lower than today; Chala et al., 2017) diminishing dispersal distances and increasing the total area of the alpine habitat (Brochmann et al., 2022). This allowed for increases in population size, resulting in increased production and exchange of propagules (Hedberg, 1969; Sepulchre et al., 2006; Gehrke & Linder, 2014). The disjunct occurrence of several tropical Afroalpine *Helichrysum* species in mountains in Kenya, Tanzania, Rwanda, and Uganda (e.g.,*H. argyranthum, H. formosissimum, H. newii*) probably results from LDD, given that the alpine belts of most of these mountains have never been connected (Kebede et al., 2007; Chala et al., 2017). Given this, it is likely that wind is the most important dispersal vector for *Helichrysum*—although their pappus are soon caducous and do not seem to have a significant role in dispersion, their achenes are tiny and light (Nathan et al., 2008; Bergh & Linder, 2009; Nie et al., 2016). In this regard, our results are in line with many biogeographic studies suggesting that LDD has played a major role in shaping colonization and genetic structuring of tropical Afroalpine plant communities (Brochmann et al., 2022). For instance, the widespread *Erica arborea* Brot. recently colonized the tropical East African mountains, likely from Ethiopia (Gizaw et al., 2013); gene flow has occurred among distant populations of tropical Afroalpine *Alchemilla* species (Gehrke et al.,

2016) and intermountain admixture and hybridization have been reported in the recently radiated *Dendrosenecio* genus (Tusiime et al., 2020; Gizaw et al., 2021).

5. Conclusions

The southern African grasslands are for the first time identified as having played a crucial role in the biogeographic history of *Helichrysum*. Our results indicate that these grasslands constitute the most important source of lineages colonizing other African areas, in particular the tropical Afromontane area. Similarly, the tropical Afromontane area spawned extra-African HAP clade radiations, including dispersal events to Eurasia, America, and Madagascar, and back-colonizations to the southern African grasslands. In general, the ability of *Helichrysum* to span broad niches may have favored repeated niche shifts and multiple colonizations of most areas. Range expansions of montane habitats during favorable periods of the Miocene-Pliocene, followed by phases of retraction and habitat isolation between mountain groups, likely fostered diversification and shaped the current Afromontane species diversity. Regarding the origin of the tropical Afroalpine species, our data do not provide evidence for dispersal from other alpine areas. Instead, we detected several independent colonization events of the tropical Afroalpine area from the geographically close tropical Afromontane area by ancestors that were likely preadapted to open habitats, suggesting repeated niche shifts to adapt to the extreme diurnal temperature oscillation and night frost of the Afroalpine area. Downwards elevational expansion of the tropical Afroalpine belt during glacial periods (Gehrke & Linder, 2009; Brochmann et al., 2022) likely favored dispersal among Afroalpine "islands" due to the increase in target habitat size (Gehrke & Linder, 2009, 2014). The combination of allopatric speciation and range expansion of some species is likely to have shaped the current diversity and distribution of tropical Afroalpine *Helichrysum*. In general, our results agree with studies suggesting that tropical Afroalpine plant radiations were both rapid and recent (Pliocene–Pleistocene) and triggered by the appearance of new ecological opportunities provided by mountain uplift and climatic oscillations (Hemp, 2006; Westerhold et al., 2020).

References

Adams, C.D. (1963). Compositae. In: *Flora of West Tropical Africa*, volume 2, pp. 225–297, 2nd ed. Hutchinson, J., Dalziel, J.M., revised by Keay, R.W.Y. & Hepper F.N.; with annotations from African Plants Database (Version 3.4.0). Crown Agents: London.

Aduse-Poku, K., van Bergen, E., Sáfián, S., et al. (2022). Miocene climate and habitat change drove diversification in *Bicyclus,* Africa's largest radiation of Satyrine butterflies. *Syst. Biol.*, 71, 570–558.

Anderberg, A.A. (1991). Taxonomy and phylogeny of the tribe Gnaphalieae (Asteraceae). *Opera Bot.*, 104, 1–195.

Andrés-Sánchez, S., Verboom, G.A., Galbany-Casals, M., & Bergh, N.G. (2019). Evolutionary history of the arid climate-adapted *Helichrysum* (Asteraceae: Gnaphalieae): Cape origin and association between annual life-history and low chromosome numbers. *J. Syst. Evol.*, 57, 468–487.

Assefa, A., Ehrich, D., Taberlet, P., Nemomissa, S., & Brochmann, C. (2007). Pleistocene colonization of afro-alpine "sky islands" by the arctic-alpine *Arabis alpina*. *Heredity*, 99, 133–142.

Axelrod, D.I., & Raven, P.H. (1978). Late Cretaceous and Tertiary vegetation history of Africa. In: *Biogeography and Ecology of Southern Africa*, pp. 77–139. Werger, M.J.A., (ed.). Springer: The Hague.

Bagley, J.C., Uribe-Convers, S., Carlsen, M.M., & Muchhala, N. (2020). Utility of targeted sequence capture for phylogenomics in rapid, recent angiosperm radiations: Neotropical *Burmeistera* bellflowers as a case study. *Mol. Phylogenet. Evol.*, 152, 106769.

Baker, M.S., Kane, N.C., Matvienko, M., et al. (2008). Multiple paleopolyploidizations during the evolution of the Compositae reveal parallel patterns of duplicate gene retention after millions of years. *Mol. Biol. Evol.*, 25, 2445–2455.

Baker, M.S., Li, Z., Kidder, T.I., et al. (2016). Most Compositae (Asteraceae) are descendant of a paleohexaploid and all share a paleotetraploid ancestor with the Calyceraceae. *Am. J. Bot.*, 103, 1203–1211.

Beentje, H.J. (2002). *Helichrysum*. In: *Flora of Tropical East Africa*, part 2, pp. 403–452. Beentje, H.J., (ed.). Balkema: Rotterdam.

Bentley, J., Verboom, G.A., & Bergh, N.G. (2014). Erosive processes after tectonic uplift stimulate vicariant and adaptive speciation: Evolution in an Afrotemperate-endemic paper daisy genus. *BMC Evol. Biol.*, 14, 27.

Bergh, N.G.M., & Linder, H.P. (2009). Cape diversification and repeated out-of-southern-Africa dispersal in paper daisies (Asteraceae- Gnaphalieae). *Mol. Phylogenet. Evol.*, 51, 5–18.

Bingham, M.G., Willemen, A., Wursten, B.T., Ballings, P., & Hyde, M.A. (2021). Flora of Zambia, based on Flora Zambesiaca. Available online: http://www.zambiaflora.com [Accessed: 21April 2021].

Bolger, A.M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30, 2114–2120.

Brochmann, C., Gizaw, A., Chala, D., et al. (2021). History and evolution of the afroalpine flora: In the footsteps of Olov Hedberg. *Alp. Bot.*, 132, 65–87.

Burgoyne, P.M., Van Wyk, A.E., Anderson, J.M., & Schrire, B.D. (2005). Phanerozoic evolution of plants on the African Plate. *J. Afr. Earth Sci.*, 43, 13–52.

Bushnell, B. (2014). *BBMap: A Fast, Accurate, Splice-AWARE Aligner*; Lawrence Berkeley National Lab (LBNL): Berkeley, CA, USA.

Carbutt, C. (2019). The Drakensberg Mountain Centre: A necessary revision of southern Africa's high-elevation centre of plant endemism. *S. Afr. J. Bot.*, 124, 508–529.

Carbutt, C., & Edwards, T.J. (2004). The flora of the Drakensberg Alpine Centre. *Edinburgh J. Bot.*, 60, 581–607.

Carbutt, C., & Edwards, T.J. (2015). Reconciling ecological and phytogeographical spatial boundaries to clarify the limits of the montane and alpine regions of sub-Sahelian Africa. *S. Afr. J. Bot.*, 98, 64–75.

Carlsen, T., Bleeker, W., Hurka, H., Elver, R., & Brochmann, C. (2009). Biogeography and phylogeny of *Cardamine* (Brassicaeae). *Ann. Missouri Bot. Gard.*, 96, 215–236.

Chala, D., Zimmermann, N.E., Brochmann, C., & Bakkesstuen, V. (2017). Migration corridors for alpine plants among the "sky islands" of eastern Africa: Do they, or did they exist? *Alp. Bot.*, 127, 133–144.

Chorowicz, J. (2005). The East African rift system. *J. Afr. Earth Sci.*, 43, 379–410.

Coetzee, J.A., & Rogers, J. (1982). Palynological and lithological evidence for the Miocene palaeoenvironment in the Saldanha region (South Africa). *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, 39, 71–85.

Cronn, R., Knaus, B.J., Liston, A., et al. (2012). Targeted enrichment strategies for next-generation plant biology. *Am. J. Bot.*, 99, 291–311.

Darriba, D., Posada, D., Kozlov, A.M., Stamatakis, A., Morel, B., & Flouri, T. (2020). ModelTest-NG: A new and scalable tool for the selection of DNA and protein evolutionary models. *Mol. Biol. Evol.*, 37, 291–294.

Davis, S.D., Heywood, V.H., & Hamilton, A.C. (1994). Centres of plant diversity. In: *A Guide and Strategy for their Conservation. Europe, Africa, South West Asia and the Middle East*, volume 1. IUCN Publications Unit: Cambridge.

Dupin, J., Matzke, N.J., Särkinen, T., et al. (2017). Bayesian estimation of the global biogeographical history of the Solanaceae. *J. Biogeogr.*, 44, 887–899.

Dupont, L.M., Linder, H.P., Rommerskirchen, F., & Schefuß, E. (2011). Climate driven rampant speciation of the Cape flora. *J. Biogeogr.*, 38, 1059–1068.

Engler, A. (1904). Plants of the northern temperate zone in their transition to the high mountains of tropical Africa. *Ann. Bot.*, 18, 523–540.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, 39, 783–791.

Fér, T., & Schmickl, R.E. (2018). HybPhyloMaker: Target Enrichment data analysis from raw reads to species trees. *Evol. Bioinform. Online*, 14, 1176934317742613.

Friis, I., & Vollesen, K. (1998). Flora of the Sudan-Uganda border area east of the Nile. In: *Catalogue of Vascular Plants,* volume 1, 1st part. Kongel Danske Vidensk Selsk: Copenhagen.

Galbany-Casals, M., Garcia-Jacas, N., Sáez, L., Benedí, C., & Susanna, A. (2009). Phylogeny, Biogeography and character evolution in Mediterranean, Asiatic and Macaronesian *Helichrysum* (Asteraceae, Gnaphalieae) inferred from nuclear phylogenetic analyses. *Int. J. Plant. Sci.*, 170, 365–380.

Galbany-Casals, M., Unwin, M., Garcia-Jacas, N., Smissen, R.D., Susanna, A., & Bayer, R.J. (2014). Phylogenetic relationships in *Helichrysum* (Compositae: Gnaphalieae) and related genera: Incongruence between nuclear and plastid phylogenies, biogeographic and morphological patterns, and implications for generic delimitation. *Taxon*, 63, 608–624.

Galley, C., Bytebier, B., Bellstedt, D.U., & Linder, H.P. (2007). The Cape element in the Afrotemperate flora: From Cape to Cairo? *Proc. R. Soc. B: Biol. Sci.*, 274, 535–543.

Gehrke, B. (2018). Staying cool: Preadaptation to temperate climates required for colonizing tropical alpine-like environments. *PhytoKeys*, 96, 111–126.

Gehrke, B., & Linder, H.P. (2009). The scramble for Africa Pan-temperate elements on the African high mountains. *Proc. R. Soc. B: Biol. Sci.*, 276, 2657–2665.

Gehrke, B., & Linder, H.P. (2014). Species richness, endemism and species composition in the tropical Afroalpine flora. *Alp. Bot.*, 124, 165–177.

Gehrke, B., Kandziora, M., & Pirie, M.D. (2016). The evolution of dwarf shrubs in alpine environments: A case study of *Alchemilla* in Africa. *Ann. Bot.*, 117, 121–131.

Gizaw, A., Brochmann, C., Nemomissa, S., et al. (2016). Colonization and diversification in the African 'sky islands': Insights from fossil-calibrated molecular dating of *Lychnis* (Caryophyllaceae). *New Phytol.*, 211, 719–734.

Gizaw, A., Gorospe, J.M., Kandziora, M., et al. (2021). Afro-alpine flagships revisited II: Elucidating the evolutionary relationships and species boundaries in the giant senecios (*Dendrosenecio*, Asteraceae). *Alp. Bot.*, 132, 89–105.

Gizaw, A., Kebede, M., Nemomissa, S., et al. (2013). Phylogeography of the heathers *Erica arborea* and *E. trimera* in the afro-alpine 'sky islands' inferred from AFLPs and plastid DNA sequences. *Flora: Morphol. Distrib. Funct. Ecol. Plants*, 208, 453–463.

Guillarmod, A.J. (1971). *Flora of Lesotho (Basutoland).* Verlag von Cramer: Lehre.

Harmsen, R., Spence, J.R., & Mahaney, W.C. (1991). Glacial interglacial cycles and development of the Afroalpine ecosystem on East-African mountains, II. Origins and development of the biotic component. *J. Afr. Earth Sci.*, 12, 512–523.

Heald, W.F. (1967). *Sky Island.* Van Nostrand: Princeton.

Heath, T.A. (2020). Dating Species Divergences with the Fossilized Birth-Death Process. Divergence time Estimation using BEAST v2. BEAST v2 Tutorial. [Accessed: 23 May 2023].

Hedberg, O. (1969). Evolution and speciation in a tropical high mountain flora. *Biol. J. Linn. Soc.*, 1, 135–148.

Rutherford, M.C., & Westfall, R.H. (1986). Biomes of southern Africa—An objective categorization. *Memo. Bot. Surv. South Afric*a, 54, 1–98.

Hemp, A. (2006). Continuum or zonation? Altitudinal gradients in the forest vegetation of Mt. Kilimanjaro. *Plant Ecol.*, 184, 27–42.

Hilliard, O.M. (1983). *Helichrysum* Mill. In: *Flora of Southern Africa*, volume 33, part 7, fascicle 2, pp. 61–310. Leistner; O.A., (ed.). Department of Agriculture: Pretoria.

Ho, S.Y.M., & Phillips, M.J. (2009). Accounting for Calibration Uncertainty in Phylogenetic Estimation of Evolutionary Divergence Times. *Syst. Biol.*, 58, 367–380.

Hughes, C.E., & Atchison, G.W. (2015). The Ubiquity of alpine plant radiations: From the Andes to the Hengduan Mountains. *New Phytol.*, 207, 275–282.

Humbert, H. (1962). *Flore de Madagascar et des Comores (Plantes vasculaires)*; 189e famille. Composées, part 2. Leroy, J.F., (ed.). Muséum National d'Histoire Naturelle: Paris.

Hyde, M.A., Wursten, B.T., Ballings, P., & Coates Palgrave, M. (2021a). Flora of Botswana. Available online: http://www.botswanaflora.com/index.php [Accessed: 22 April 2021].

Hyde, M.A., Wursten, B.T., Ballings, P., & Coates Palgrave, M. (2021b). Flora of Malawi. Available online: http://www.malawiflora.com [Accessed: 22 April 2021].

Hyde, M.A., Wursten, B.T., Ballings, P., & Coates Palgrave, M. (2021c). Flora of Mozambique. Available online: http://www.mozambiqueflora.com [Accessed: 23 April 2021].

Hyde, M.A., Wursten, B.T., Ballings, P., & Coates Palgrave, M. (2021d). Flora of Zimbabwe. Available online: http://www.zimbabweflora.co.zw [Accessed: 23 April 2021].

Kandziora, M., Kadereit, J.W., & Gehrke, B. (2017). Dual colonization of the Palaearctic from different regions in the Afrotropics by *Senecio*. *J. Biogeogr.*, 44, 147–157.

Kandziora, M., Gehrke, B., Popp, M., Gizaw, A., Brochmann, C., & Pirie, M.D. (2022a). The enigmatic tropical alpine flora on the African sky islands is young, disturbed, and unsaturated. *Proc. Natl. Acad. Sci. USA*, 119, e2112737119.

Kandziora, M., Kadereit, J.W., & Gehrke, B. (2016). Frequent colonization and little in situ speciation in *Senecio* in the tropical alpine-like islands of eastern Africa. *Am. J. Bot.*, 103, 1483–1498.

Kandziora, M., Sklenář, P., Kolář, F., & Schmickl, R. (2022b). How to tackle phylogenetic discordance in recent and rapidly radiating groups? Developing a workflow using *Loricaria* (Asteraceae) as an example. *Front. Plant Sci.*, 12, 765719.

Katoh, K., & Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.*, 30, 772–780.

Kebede, M., Ehrich, D., Taberlet, P., Nemomissa, S., & Brochmann, C. (2007). Phylogeography and conservation genetics of a giant lobelia (*Lobelia giberroa*) in Ethiopian and Tropical East Africa Mountains. *Mol. Ecol.*, 16, 1233–1243.

Kent, W.J. (2002). BLAT—the BLAST-like alignment tool. *Genome Res.*, 12, 656–664.

Killick, D.J.B. (1978). The Afro-Alpine region. In: Biogeography and Ecology of Southern Africa, Volume 31, pp. 515–542. Wegner, M.J.A. (ed.). Monographiae Biologicae: The Hague.

Koekemoer, M., & Steyn, H.M. (2010). Helichrysums of Southern Africa, version 1. South African National Biodiversity Institute. Available online: https://keys.lucidcentral.org/keys/v3/helichrysum/key/Helichrysum/Media/Html/home.htm [Accessed: 25 March 2021].

Kozlov, A.M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAxML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, 35, 4453–4455.

Lagomarsino, L.P., Frankel, L., Uribe-Convers, S., Antonelli, A., & Muchhala, N. (2022). Increased resolution in the face of conflict: Phylogenomics of the Neotropical bellflowers (Campanulaceae: Lobelioideae), a rapid plant radiation. *Ann. Bot.*, 129, 723–736.

Landis, M.J., Matzke, N.J., Moore, B.R., & Huelsenbeck, J.P. (2013). Bayesian analysis of biogeography when the number of areas is large. *Syst. Biol.*, 62, 789–804.

Lemoine, F., Domelevo Entfellner, J.B., Wilkinson, E., et al. (2018). Renewing Felsenstein's phylogenetic bootstrap in the era of big data. *Nature*, 556, 452–456.

Linder, H.P., & Verboom, G.A. (2015). The Evolution of Regional Species Richness: The History of the Southern African Flora. *Annu. Rev. Ecol. Evol. Syst.*, 46, 393–412.

Linkem, C.W., Minin, V.N., & Leaché, A.D. (2016). Detecting the anomaly zone in species trees and evidence for a misleading signal in higher-level skink phylogeny (Squamata: Scincidae). *Syst. Biol.*, 65, 465–477.

Lisowski, S. (1989). *Helichrysum* Compositae (deuxième partie: Tribu Inuleae). In: Flore d'Afrique Centrale (Zaire, Rwanda, Burundi), pp. 68–192. National Botanic Garden of Belgium: Meise.

Lozano-Fernández, J. (2022). A practical guide to design and assess a phylogenomic study. *Genome Biol. Evol.*, 14, evac129.

Maddison, W.P. (1997). Gene trees in species trees. *Syst. Biol.*, 46, 523–536.

Mairal, M., Pokorny, L., Aldasoro, J.J., Alarcón, M., & Sanmartín, I. (2015). Ancient vicariance and climate-driven extinction explain continental-wide disjunctions in Africa: The case of the Rand Flora genus *Canarina* (Campanulaceae). *Mol. Ecol.*, 24, 1335–1345.

Mandel, J.R., Dikow, R.B., Funk, V.A., et al. (2014). A target enrichment method for gathering phylogenetic information from hundreds of loci: An example from the Compositae. *Appl. Plant Sci.*, 2, 1300085.

Masters, J.C., Génin, F., Zhang, Y., et al. (2021). Biogeographic mechanisms involved in the colonization of Madagascar by African vertebrates: Rifting, rafting and runways. *J. Biogeogr.*, 48, 492–510.

Matzke, N.J. (2013). BioGeoBEARS: Biogeography with Bayesian (and likelihood) evolutionary analysis in R scripts. R package. Available at: https://github.com/nmatzke/BioGeoBEARS [Accessed: 2 September 2021]

Matzke, N.J. (2014). Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Syst. Biol.,* 63, 951–970.

Matzke, N.J. (2022). Statistical comparison of DEC and DEC+J is identical to comparison of two ClaSSE submodels and is therefore valid. *J. Biogeogr.*, 49, 1805–1824.

Mikula, O.; Nicolas, V.; Šumbera, R.; et al. (2021). Nuclear phylogenomics, but not mitogenomics, resolves the most successful Late Miocene radiation of African mammals (Rodentia: Muridae: Arvicanthini). *Mol. Phylogenet. Evol.*, 157, 107069.

Nathan, R., Schurr, F.M., Spiegel, O., Steinitz, O., Trakhtenbrot, A., & Tsoar, A. (2008). Mechanisms of long-distance seed dispersal. *Trends Ecol. Evol.*, 23, 638–647.

Nie, Z.-L., Funk, V.A., Meng, Y., Deng, T., Sun, H., & Wen, J. (2016). Recent assembly of the global herbaceous flora: Evidence from the paper daisies (Asteraceae: Gnaphalieae). *New Phytol.*, 209, 1789–1806.

Partridge, T.C. (1998). Of diamonds, dinosaurs and diastrophism: 150 years of landscape evolution in southern Africa. *S. Afr. Geogr. J.*, 101, 167–184.

Partridge, T.C., & Maud, R.R. (1987). Geomorphic evolution of southern Africa since the Mesozoic. *S. Afr. Geogr. J.*, 90, 179–208.

Partridge, T.C., & Maud, R.R. (2000). Macro-scale geomorphic evolution of southern Africa. *Oxf. Monogr. Geol. Geophys.*, 40, 3–18.

Pirie, M.D., Kandziora, M., Nürk, N.M., et al. (2019). Leaps and bounds: Geographical and ecological distance constrained the colonization of the Afromontemperate by *Erica*. *BMC Evol. Biol.* 19, 222.

Popp, M., Gizaw, A., Nemomissa, S., Suda, J., & Brochmann, C. (2008). Colonization and diversification in the African "sky islands" by Eurasian *Lychnis* L. (Caryophyllaceae). *J. Biogeogr.*, 35, 1016–1029.

Ramalho, R.S., Brum da Silveira, A., Fonseca, P.E., et al. (2015). The emergence of volcanic oceanic islands on a slow-moving plate: The example of Madeira Island, NE Atlantic. *Geochem. Geophys. Geosystems*, 16, 522–537.

Rambaut, A. (2016). FigTree ver. 1.4.3. Department of Zoology, University of Oxford: Oxford.

Ree, R.H., & Sanmartín, I. (2018). Conceptual and statistical problems with the DEC+J model of founder-event speciation and its comparison with DEC via model selection. *J. Biogeogr.*, 45, 741–749.

Ree, R.H., & Smith, S.A. (2008). Maximum likelihood inference of geographic range evolution by dispersal, local extinction and cladogenesis. *Syst. Biol.*, 57, 4–14.

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

Ronquist, F. (1997). Dispersal–vicariance analysis: A new approach to the quantification of historical biogeography. *Syst. Biol.*, 46, 195–203.

Rutherford, M.C., Mucina, L., & Powrie, L.W. (2006). Biomes and bioregions of Southern Africa. In: *The vegetation of South Africa, Lesotho and Swaziland*, pp. 30–51. Mucina, L., & Rutherford, M.C., (eds.). South African National Biodiversity Institute: Pretoria.

Sanderson, M.J. (2002). Estimating absolute rates of molecular evolution and divergence times: A Penalized Likelihood approach. *Mol. Biol. Evol.*, 19, 101–109.

Sayyari, E., & Mirarab, S. (2016). Fast coalescent-based computation of local branch support from quartet frequencies. *Mol. Biol. Evol.*, 33, 1654–1668.

Schulze, R.E. (2021). South African Atlas of Agrohydrology and Climatology. Available online: http://fred.csir.co.za/project/tmg/agrohydrology_atlas/atlas_toc.htm [Accessed: 29 January 2023].

Sepulchre, P., Ramstein, G., Fluteau, F., Schuster, M., Tercelin, J.J., & Brunet, M. (2006). Tectonic uplift and Eastern Africa aridification. *Science*, 313, 1419–1423.

Seton, M., Müller, R.D., Zahirovic, S., et al. (2012). Global continental and ocean basin reconstructions since 200 Ma. *Earth Sci. Rev.*, 113, 212–270.

Smissen, R.D., Bayer, R.J., Berg, N.G., et al. (2020). A revised subtribal classification of Gnaphalieae (Asteraceae). *Taxon*, 6, 778–806.

Smissen, R.D., Galbany-Casals, M., & Breitwieser, I. (2011). Ancient allopolyploidy in the everlasting daisies (Asteraceae: Gnaphalieae): Complex relationship among extant clades. *Taxon*, 60, 649–662.

Smith, S.A., & O'Meara, B.C. (2012). treePL: Divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics,* 28, 2689–2690.

Stamatakis, A. (2006). RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 22, 2688–2690.

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30, 1312–1313.

Tadesse, M. (2004). *Helichrysum* Mill., Asteraceae (Compositae). In: *Flora of Ethiopia and Eritrea*, part 2, pp. 163–178. Hedberg ,I., Friis, I., Edwards, S., (eds.). Addis Abeba University: Addis Ababa.

Thiers, B. (2021). Index Herbariorum. Available online: http://sweetgum.nybg.org/ih/ [Accessed: 20 April 2021].

Threatened Species Programme. Red List of South African Plants. South African National Biodiversity Institute (2012). Available online: http://redlist.sanbi.org/index.php [Accessed: 21 April 2021].

Tusiime, F.M., Gizaw, A., Gussarova, G., et al. (2020). Afro-alpine flagships revisited: Parallel adaptation, intermountain admixture and shallow genetic structuring in the giant senecios (*Dendrosenecio*). *PLoS ONE*, 15, e0228979.

Tusiime, F.M., Gizaw, A., Wondimu, T., et al. (2017). Sweet vernal grasses (*Anthoxanthum*) colonized African mountains along two fronts in the Late Pliocene, followed by secondary contract, polyploidization and local extinction in the Pleistocene. *Mol. Ecol.*, 26, 3513–3532.

Ufimov, R., Gorospe, J.M., Fér, T., et al. (2022). Utilizing paralogues for phylogenetic reconstruction has the potential to increase species tree support and reduce gene tree discordance in target enrichment data. *Mol. Ecol. Resour.*, 22, 3018–3034.

Van Wyk, A.E., & Smith, G.F. (2001). *Regions of Floristic Endemism in Southern Africa: A Review with Emphasis on Succulents*. Umdaus Press: Hatfield.

Weitemier, K., Straub, S.C., Cronn, R.C., et al. (2014). Hyb-Seq: Combining target enrichment and genome skimming for plant phylogenomics. *App. Plant Sci.*, 2, apps.1400042.

Westerhold, T., Marwan, N., Drury, A.J., et al. (2020). An astronomically dated record of Earth's climate and its predictability over the last 66 million years. *Science*, 369, 1383–1387.

White, F. (1978). The Afromontane region. In: *Biogeography and Ecology of Southern Africa*, volume 31, pp. 463–513. Wegner, M.J.A. (ed.). Springer: Dordrecht.

White, F. (1983a). *Vegetation of Africa: A Descriptive Memoir to Accompany the UNESCO/AETFAT/UNSO Vegetation Map of Africa*. UNESCO: Paris.

White, F. (1983b). Long-distance dispersal and the origins of the Afromontane flora. *Sonderbd. Des Nat. Ver. Hambg.*, 7, 87–116.

Whiterfield, J.B., & Lockhart, P.J. (2007). Deciphering ancient rapid radiations. *Trends Ecol. Evol.*, 22, 258–265.

Wood, J.R.I. (1997). *A Handbook of the Yemen Flora*. Kew Publishing: Kew.

Yoder, A.D., & Nowak, M.D. (2006). Has vicariance or dispersal been the predominant biogeographic force in Madagascar? Only time will tell. *Annu. Rev. Ecol. Evol. Syst.*, 37, 404–431.

Zachos, J., Dickens, G., & Zeebe, R. (2008). An early Cenozoic perspective on greenhouse warming and carbon cycle dynamics. *Nature*, 451, 279–283.

Zhang, C., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). ASTRAL-III: Polytomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinform.*, 19, 153.

Supplementary Materials

Supplementary materials associated with this article can accessed online.

1. Supplementary Figures

Supplementary Figure S1. ParalogWizard pairwise distance histogram.

Pairwise distance histogram generated with ParalogWizard. The distribution of pairwise distances between exonic contigs for each locus is shown; distances are given as divergence percentage values (7.6—18.9%). The position of the first and second peak is indicated. The mean divergence and the standard deviation (plus and minus sigma) for each peak are marked by dashed lines; values are provided in the legend.

Supplementary Figure S2. Maximum likelihood ancestral range estimation of *Helichrysum* using the best-fit model DEC+j from BioGeoBEARS.

Ancestral range estimation of the HAP clade using the best-fitting model DEC+j based on a time-calibrated phylogeny generated under the coalescence approach using target-enrichment data (Compositae1061 probe set). Pie charts at nodes show the relative probability of the possible states (areas in primary colors; combinations of areas in grey). Relevant node numbers are indicated to the left of the node. Black asterisks indicate nodes that were strongly supported by both metrics (BS and TBE ≥ 95%/0.95), grey asterisks indicate nodes strongly supported by only one metric (BS or TBE ≥ 95%/0.95), grey circles indicate nodes moderately supported by both metrics(BS and TBE from 70%/0.70 to 94%/0.94), and empty circles indicate nodes moderately supported by only one metric (BS or TBE from 70%/0.70 to 94%/0.94).

Supplementary Figure S3. ASTRAL species tree of *Helichrysum.*

ASTRAL species tree of the HAP clade based on nuclear DNA loci obtained from target-enrichment (Compositae1061 probe set). Numbers on the nodes indicate local posterior probabilities (LPP). The outgroup is colored in blue.

Supplementary Figure S4. Frequency distribution of biogeographic event counts (DEC+j model).

Frequency distributions of event counts from 100 Biogeographic Stochastic Mappings (DEC+j model) on the *Helichrysum* phylogeny.

Supplementary Figure S5. Maximum likelihood ancestral range estimation of *Helichrysum* using the model DEC from BioGeoBARS.

Ancestral range estimatiion of the HAP clade using the model DEC based on a time-calibrated phylogeny generated under the coalescence approach using target-enrichment data (Compositae1061 probe set). Pie charts at nodes show the relative probability of the possible states (areas in primary colors, combinations of areas in grey).

Supplementary Figure S6. Frequency distribution of biogeographic event counts (DEC model).

Frequency distributions of event counts from 100 Biogeographic Stochastic Mappings (DEC model) on the *Helichrysum* phylogeny.

2. Supplementary Tables

Supplementary Table S1. List of studied materials with BioSample accession numbers.

List of studied materials with herbaria codes and voucher information. BioSample assession numbers are provided (access through BioProject PRJNA936872 through: https://www.ncbi.nlm.nih.gov/sra/PRJNA936872).

Supplementary Table S2. Reads recovered per species after filtering (HPM1).

Sequenced species and their corresponding number reads with HybPhyloMarker target extraction method after filtering (script HPM1).

Supplementary Table S3. Statistics from paralogy analyses using ParalogWizard.

Statistics obtained in the paralogy analyses using ParalogWizard. For each sample, the number and percentage of paralogs loci within the selected divergence threshold (7.6—18.9%) is provided.

Supplementary Table S4. Summary of selected loci for subsequent phylogeny reconstruction.

Summary of the selected loci (including paralogs) for subsequent phylogeny reconstruction, after filtering for missing data (script HPM5).

Supplementary Table S5. Summary statistics resulting from the processing of nuclear DNA data.

Summary statistics resulting from the processing of nuclear DNA data obtained using the Compositae1061 probe set.

Supplementary Table S6. Summary statistics of biogeographic model testing in BioGeoBEARS.

Summary statistics of biogeographic model testing in BioGeoBEARS in which likelihood scores (LnL) and corrected Akaike Information Criterion (AICc) are detailed along with the values of the free parameters (dispersal (d), extinction (e), founder-event (j)). The best-fitting model (DEC+j) is highlighted in bold.

Supplementary Table S7. Ancestral range probabilities of the model DEC+J for *Helichrysum.*

Ancestral range probabilities of the DEC+j model for *Helichrysum*, corresponding to the pie charts in Figure 3 and Supplementary Figure S3.

Supplementary Table S8. DEC+j model Biogeographic Stochastic Mapping mean and standard deviation tables.

DEC+j Biogeographic Stochastic Mapping Counts. Mean number of dispersal events estimated in the history of *Helichrysum* with Biogeographic Stochastic Mapping. Counts of dispersal events were averaged across 100 BSMs, standard deviations can be found below each table. Rows are the source of dispersals and columns the sink. Colors indicate the frequency of events, (red-green, highest-lowest).

Supplementary Table S9. DEC model Biogeographic Stochastic Mapping mean and standard deviation tables.

DEC Biogeographic Stochastic Mapping Counts. Mean number of dispersal events estimated in the history of *Helichrysum* with Biogeographic Stochastic Mapping. Counts of dispersal events were averaged across 100 BSMs, standard deviations can be found below each table. Rows are the source of dispersals and columns the sink. Colors indicate the frequency of events, (red-green, highest-lowest).

Chapter 2

Biome conservatism prevailed in repeated long-distance colonization of Madagascar's mountains by *Helichrysum* (Compositae, Gnaphalieae)

CHAPTER 2: Biome conservatism prevailed in repeated long-distance colonization of Madagascar’s mountains by *Helichrysum* (Compositae, Gnaphalieae)

The content of this chapter has already been published in the following research article: Blanco-Gavaldà C, Roquet C, Puig-Surroca G, Andrés-Sánchez S, Razafimandimbison SG, Letsara R, Bergh N, Cron GV, Moreyra LD, Calleja JA, Castillo O, Bayer RJ, Leliaert F, Susanna A, & Galbany-Casals M. (2025). **Biome conservatism prevailed in repeated long-distance colonization of Madagascar's mountains by *Helichrysum* (Compositae, Gnaphalieae)**. *Molecular Phylogenetics and Evolution*, 204, 108283.

Abstract

Colonization and diversification processes are responsible for the distinctiveness of island biotas, with Madagascar standing out as a biodiversity hotspot exceptionally rich in species and endemism. Regardless of its significance, the evolutionary history and diversification drivers of Madagascar’s flora remain understudied. Here we focus on *Helichrysum* (Compositae, Gnaphalieae) to investigate the evolutionary and biogeographic origins of the Malagasy flora. We inferred a highly resolved phylogeny based on target-enrichment data from 327 species (including 51% of Malagasy endemics) and conducted ancestral range estimation analyses. Our results revealed at least six trans-oceanic dispersal events from different African regions to Madagascar during the Pliocene. In this process, biome conservatism prevailed, as evidenced by similarities between Malagasy lineages and their African relatives. The southern African grasslands, known to be the center of diversification and the main source of African *Helichrysum* lineages, played a key role in the colonization of Madagascar as the ancestral source area of at least three clades. The Tropical Afromontane region was revealed as the source of at least two montane Malagasy lineages that substantially radiated *in-situ*. Finally, a dispersal event from southwestern Africa led to a lineage represented by a single species adapted to the island’s southwestern arid conditions. The main radiations of *Helichrysum* in Madagascar’s mountains occurred within the last 2 My, coinciding with a transition towards cooler and drier conditions and the expansion of open habitats, likely driven by a combination of geographic and ecological speciation. Overall, our findings highlight the affinities between the montane floras of continental Africa and Madagascar.

Keywords: Afromontane flora, Asteraceae, biogeography, dispersal, *Helichrysum*, Madagascar, target-enrichment

1. Introduction

Unraveling the origins and evolution of island biotas has been of interest to scientists since Darwin’s times (Darwin, 1859). Due to their isolation, comparatively small size and varying distances from mainlands, islands are considered natural laboratories for the study of ecological and evolutionary processes. Species richness of islands can generally be explained through features such as island area and isolation as well as the age and formation history (Whittaker et al., 2008; Graham et al., 2017). Distinguishing between oceanic islands and continental islands is important as it dictates the mechanisms underlying species colonization. Unlike the relatively younger oceanic islands, which emerge devoid of any resident species and gradually accumulate their diversity through dispersal followed by speciation, older continental islands inherit a baseline biota upon isolation from the mainland (Matthews & Triantis, 2021).

Madagascar, once part of Gondwana, is an ancient continental island. It first broke apart from Africa 155 to 165 million years ago (mya) (Yoder & Nowak, 2006 and references therein) and later from India 84 to 91 mya (Wells, 2003) and has remained separated since then. Madagascar hosts a hyperdiverse biota in a wide range of ecosystems, being home to an estimated total flora of ca. 14900 vascular plant species, of which 87% are endemics (Callmander et al., 2011; Lowry et al., 2018; Antonelli et al., 2022), with 310 endemic plant genera and five endemic families (Buerki et al., 2013). Such exceptional species richness and endemicity are the result of a complex geological, climatic and evolutionary history. Despite hosting vicariant groups predating Gondwana's breakup (e.g. *Takhtajania*, Winteraceae; Thomas et al., 2014), most of the present-day diversity likely established on Madagascar through long-distance dispersal (Yoder & Novak, 2006; Buerki et al., 2013), probably assisted by wind and oceanic currents. The existence of land bridges and stepping-stones between Africa and Madagascar has been hypothesized, but not supported by any direct evidence, so their potential role and relevance in facilitating dispersal remains under debate (Warren et al., 2010; Masters et al., 2021; Ali & Hedges, 2023; Aslanian et al., 2023). Geographic proximity likely accounts for the Malagasy biota having the strongest taxonomic links to African lineages (Yoder & Novak, 2006; Buerki et al., 2013).

Madagascar's complex topography plays a key role in shaping the main climatic and vegetation zones of the island and has probably had a significant impact on *in-situ* diversification of many groups. The main massifs of the island, reaching 1800-2900 m a.s.l., are placed in a north-south axis and are connected through the Central Highlands, a plateau ranging 800 to 1300 m. The orography combined with the prevailing easterly trade winds results in a rainfall gradient from the humid tropical east-northeast to the sub-arid southwest. The Central Highlands have a subhumid climate and are dominated by a grassland-woodland mosaic, heavily modified and degraded by human activity (Yoder et al., 2016; Antonelli et al., 2022). The highest parts of the mountains, above 1800-1900 m, are mainly occupied by sclerophyllous shrublands dominated by Ericaceae and Compositae, in addition to open grasslands and rupicolous plant communities (Burgoyne et al., 2005; Yoder et al., 2016; Antonelli et al., 2022; and pers. obs.). Outside the mountains, landscapes are dominated by deciduous forests, sclerophyllous and succulent plants westwards and evergreen rainforests eastwards. This ecological and biological heterogeneity provides the perfect ground for various post-colonization scenarios, such as repeated colonization without subsequent diversification (e.g. the Inulea-Pulcheinae group, Nylinder et al., 2016) as well as colonization followed by *in-situ* diversification (e.g. Psychotrieae alliance, Razafimandimbison et al., 2017; Coffeeae alliance, Kainulainen et al., 2017; C3 grasses, Hackel et al., 2018). Nevertheless, there is a knowledge gap on the relative importance of these scenarios in the evolutionary assembly of the Malagasy biota, given that the biogeographic history of many highly diverse Malagasy plant groups remains largely unexplored (Antonelli et al., 2022).

The relative contribution of adaptive evolution to the generation of biodiversity is still under debate. Niche conservatism –i.e., the tendency of species to retain their ancestral niches– has been suggested to be predominant in allopatric (also known as geographic) speciation (Wiens & Graham, 2005; Wiens et al., 2010) and has been assumed to prevail in angiosperm evolution (Wiens & Graham, 2005; Crisp. et al., 2009; Wiens et al., 2010) since migrating might be easier and quicker to achieve than adapting to new environmental conditions (Donoghue, 2008). However, multiple studies support the idea that niche shifts within plant genera are not as rare as previously thought (e.g. *Lonicera* L., Smith & Donoghue, 2010; *Coccinia* Wight & Arn., Holstein & Renner, 2011; *Ranunculus* L., Hörandl & Emadzade, 2011; *Hakea* Schrad & J.C.Wendl., Cardillo et al., 2017) and indeed older and/or more extensive biomes (e.g. sclerophyllous shrubland, tropical rainforest) have often been the source of lineages found in younger and/or less extensive biomes (e.g. grasslands, alpine, Mediterranean shrublands; Crisp et al., 2009; Donoghue & Edwards, 2014). To date, few studies have explored the role that niche evolution may have played in the diversification of Malagasy plant groups (but see the case of *Bulbophyllum* Thouars in Gamisch et al., 2016, 2021).

In terms of species richness, Compositae are one of the five dominant plant families in the Malagasy flora (520 spp. and 83% endemism; Callmander et al., 2011; Catalogue of the Vascular Plants of Madagascar, 2024). Within Malagasy Compositae, the most species-rich genus is *Helichrysum* Mill. In the most recent floristic treatment of Malagasy *Helichrysum*, Humbert (1962) recognized 111 species –all but one endemic– and numerous infraspecific taxa at the subspecific and varietal levels. He did not propose a formal infrageneric classification but organized the species into eleven informal taxonomic groups based on morphological affinities. *Helichrysum* occurs in almost all Malagasy biomes, from the extremely arid south-western coastal dunes (e.g. *H. mahafaly* Humbert) to the humid eastern rainforests (e.g. *H. geayi* Humbert), adopting a wide range of growth forms (lianas, herbs, subshrubs, shrubs and small trees). However, most of the Malagasy diversity of *Helichrysum* is found in the Central Highlands and on the highest mountains, distributed across their full elevational and latitudinal gradients. The diversity of *Helichrysum* in Madagascar reflects the extraordinary variation and adaptability of the genus, already evident in the case of continental African lineages (Blanco-Gavaldà et al., 2023).

Previous studies place the origin of *Helichrysum* in southern Africa (Galbany-Casals et al., 2014; Andrés-Sánchez et al., 2019; Blanco-Gavaldà et al., 2023) and suggest that several independent and asynchronous dispersals occurred from the African continent to Madagascar. However, these works treated Madagascar as a single biogeographic area and the sampling of both mainland and Malagasy species was limited. Here, we aim to infer the biogeographic history of *Helichrysum* in Madagascar as a study-case to elucidate the relative contribution of *in-situ* speciation vs. colonization to the generation and maintenance of Malagasy plant diversity. For this purpose, we first generated a highly resolved time-calibrated phylogeny based on target-enrichment sequences, substantially increasing the sampling of African and Malagasy species compared to previous studies. We used the resulting phylogeny to estimate ancestral ranges and infer the source, number and age of colonization events of the genus on Madagascar. Finally, we evaluate whether the dominant mode of diversification within *Helichrysum* has been adaptive speciation associated with biome shifts or non-adaptive speciation associated with allopatry and biome conservatism.

2. Materials and Methods

2.1 Taxon Sampling

We sampled 327 *Helichrysum* species (ca. 60% of the genus, see Supplementary Table S1) including 57 (ca. 51%) of the Malagasy species with representatives of all Humbert's (1962) taxonomic groups, the latest taxonomic treatment of Malagasy *Helichrysum*. For some of these Malagasy species, we included several infraspecific taxa (29 additional samples). While polyploids are known in *Helichrysum* and related genera, previous studies have shown that clades of polyploid origin are confined to specific geographical regions and lineages (the Mediterranean, Macaronesian and Asian members of *Helichrysum*, and the smaller genera *Achyrocline* (Less.) DC., *Anaphalis* DC. and *Pseudognaphalium* Kirp. in America and Asia; Galbany-Casals et al., 2009, 2014). None of these is closely related to Malagasy lineages (Blanco-Gavaldà et al., 2023). Taking this into account, we excluded well-delimited polyploid clades to optimize read-mapping processes and minimize potential sources of phylogenetic discordance (Tiley et al., 2024a). With this, we also eliminated non-African biogeographic regions, allowing for the recognition of more precise Malagasy areas while keeping the analyses computationally feasible.

We also included 34 specimens of other genera representing the main lineages of the tribe (Nie et al., 2016; Smissen et al., 2020) so that we could implement secondary age calibrations. The complete dataset comprises a total of 386 samples, 289 from previous studies (273 from Blanco-Gavaldà et al., 2023. BioProject PRJNA936872; six from Mandel et al., 2019, BioProject PRJNAS40287; and ten from Schmidt-Lebuhn & Bovill, 2021, BioProject PRJNA665592). Here we sequenced for the first time 97 samples using the same baits set, representing mainly *Helichrysum* species from continental Africa and Madagascar.

2.2 Next-Generation Sequencing

DNA was extracted from 10–30 mg of dried leaf material obtained from herbarium specimens and from our field expeditions (see Supplementary Table S1) using the E.N.Z.A® SP Plant DNA Kit (Omega Bio-Tek Inc., Norcross, GA, USA) following manufacturer instructions. We measured DNA concentration using Qubit™ Flex Fluorometer (Thermo Scientific, Waltham, MA, USA). Then, we used a Qsonica Q800R3 Sonicator (Qsonica LLC, Newton, CT, USA) at 20% amplitude for 45 sec to 8 min to shear 0.2-1 ug in 50 uL of DNA into 300-400 bp fragments. We performed gel electrophoresis (1.2% agarose) to check fragment length.

We prepared target-enrichment libraries from 25ul of the sonicated DNA using the NEBNext Ultra II DNA Library Prep Kit for Illumina ® (New England Biolabs, Ipswich, MA, USA) employing half of the volumes specified by the manufacturer and 15 cycles of PCR amplification. We barcoded the libraries using NEBNext Multiplex Oligos sets with unique single or dual index combinations. Afterward, we pooled the indexed libraries with more than 17 ng of DNA in groups of up to 10 samples and around 250 ng of DNA per library. We evaporated or filled with water the arranged pools to 7 ul of total volume to perform target-enrichment (protocol from Mandel et al. 2014) using the Microarray MyBaits COS kit (Daicel Arbor Biosciences, Ann Arbor, MI, USA), specifically developed for the Compositae family. The final sequencing pools were prepared by pooling enriched libraries with unenriched libraries at a 60:40 ratio and the samples were sequenced (PE 150 bp) on Illumina HiSeq2500 and HiSeqX platforms.

We deposited newly generated raw sequence reads in the NCBI Short Read Archive database (SRA; access: https://www.ncbi.nlm.nih.gov/sra) under the BioProject accession number PRJNA1121119.

2.3 Molecular Data Processing

We used HybPhyloMaker, a bioinformatic workflow developed to process Hyb-Seq data (Fér & Schmickl 2018, available at https://github.com/tomas-fer/HybPhyloMaker, indicated hereafter as HPM, followed by the number of the corresponding script) in combination with ParalogWizard (Ufimov et al., 2022, available at https://github.com/rufimov/ParalogWizard), which detects and separates paralogs of a given locus based on sequence similarity to generate orthologous alignments.

Specifically, we used Trimmomatic v.0.39 (Bolger et al., 2014) to remove adaptors and low-quality reads, and BBMap v.38.42 (Bushnell 2014) to remove duplicates, both implemented in HPM1 (read statistics in Supplementary Table S2). We generated a reference file for initial read mapping with BWA (Li & Durbin, 2009) and SPAdes (Bankevich et al., 2012) based on sunflower genome sequences from the Compositae1061 probe set (Mandel et al., 2014). Then, we generated a customized reference based on our ingroup samples to increase mapping specificity. Pairwise exonic sequence divergence was calculated to identify paralogs. The first resulting peak represents putative allelic variation, while the second peak represents highly divergent sequences corresponding to putative paralogs. We used the value of the latter as the threshold to retrieve putative paralogous sequences. We then aligned orthologous matrices using MAFFT v.7.475 (Katoh & Toh, 2008) to finally concatenate exons into putative loci. To reduce missing data, we excluded sequences missing more than 70% of the data and removed loci for which less than 75% of all samples were represented (HPM5).

2.4 Phylogenetic Analyses

We applied concatenation as well as coalescence summary approaches to infer phylogenetic relationships. We concatenated all nuclear loci into a single supermatrix and conducted maximum likelihood (ML) partitioned analyses using RAxML-NG v.1.1.0 (modified HPM8f; Kozlov et al., 2019). Specifically, we first estimated the best nucleotide substitution model for each locus with ModelTest-NG (Darriba et al., 2020) and then performed 20 independent ML tree searches. We assessed branch support with 1000 bootstrap replicates and annotated the best-scoring ML tree with Felsenstein's Bootstrap (BS, Felsenstein, 1985) and Transfer Bootstrap Expectation (TBE, Lemoine et al., 2018) proportions considering branches with BS ≥ 70% and TBE ≥ 0.7 to be statistically supported (Hillis & Bull, 1993, Lemoine et al., 2018).

We performed summary-coalescence inference with ASTRAL-III v.5.7.8 (Zhang et al., 2018). To do so, we first inferred individual gene trees for each retrieved locus generated with RAxML v.8.2.12 (Stamatakis, 2014) and performed 100 bootstrap replicates (HPM6a, HPM7 and HPM8a). For the ASTRAL tree, we calculated branch support values as local posterior probabilities (LPP), considering well-supported branches those with LPP ≥ 0.95 (Sayyari & Mirarab, 2016). In addition, we conducted a second summary-coalescence analysis with the same parameters but using only the 25 most informative loci, which we selected with SortaDate (Smith et al., 2018) based on three criteria: bipartition support, clock-likeness, and tree length (which is proportional to information content).

2.5 Divergence Time Estimation

We applied the RelTime method (Tamura et al., 2012, 2018) implemented in MEGA 11 (Tamura et al., 2021) to time-calibrate the most likely ML tree, and computed confidence intervals using the method of Tao et al. (2020). This methodology is especially suited for large empirical genomic datasets and allows the use of calibration densities (Costa et al., 2022). The tribe Gnaphalieae lacks old enough fossils to be useful as primary calibration points. Therefore, we used four secondary calibration points (CP, Supplementary Figure S1) from previously inferred divergence time estimates (Nie et al., 2016) as age constraints. We applied a normal density to each calibration point providing appropriate mean and standard deviation values to reflect the 95% confidence interval values reported by Nie et al. (2016): the tribe crown node (mean age 25 mya, std ± 2.55, CP1), the "crown radiation" node (mean age 20.7 mya, std ± 2.55, CP2), the "HAP clade" crown node (mean age 15.39 mya, std ± 1.95, CP3) and the "FLAG clade" crown node (mean age 12.78 mya, std ± 1.85, CP4).

2.6 Ancestral Geographic Range Reconstruction

To infer the most probable ancestral geographic ranges of *Helichrysum* lineages, we defined 17 areas. Regarding continental Africa and Arabia, we applied the same criteria as in Blanco-Gavaldà et al. (2023) resulting in the following areas: (A) the tropical Afroalpine area, (C) the Indian Ocean Coastal Forest Belt, (D) the high Drakensberg area, (F) the Fynbos Biome, (G) the southern African grasslands, (L) the tropical African lowlands, (N) the arid to semi-arid southern African area, (P) the Arabian Peninsula in Asia, (S) the southern African savannah, and (T) the tropical Afromontane area.

For our delimitation of Malagasy regions (Figure 1) we consulted many bibliographic sources: the definitions of the main vegetation types described in Jenkins (1987), Burgess et al. (2004), Goodman & Benstead (2003) –which in turn integrates previous classifications by Faramalala (1995), Du Puy & Moat (1996) and Lowry et al. (1997)– and Gautier et al. (2018); the delimitation of Madagascar's six principal biomes by Yoder & Nowak (2006), based on Goodman & Benstead (2003); the map of predominant vegetation types shown in Antonelli et al. (2022) simplified from Moat & Smith (2007); and the bioclimatic zones by Rakotoarivelo et al. (2019) modified in Razafimandimbison et al. (2022). However, our final proposal is not completely coincident with any of the cited sources, as it also integrates distribution patterns of the genus and our field experience. We used the following seven distribution areas: (B) southern and southwestern arid lowlands, an area that ranges from 0 to 800 m and mainly includes Madagascar's arid spiny thicket, with drought-resistant species adapted to extreme aridity and secondary grasslands; (W) dry western lowlands, an area that ranges approximately from 0 to 800 m and mainly includes dry and highly seasonal forests, succulent woodlands and secondary grasslands; (O) humid eastern lowlands, an area that ranges approximately from 0 to 800 m and mainly includes very humid lowland rainforests with some herbaceous clearings; (H) subhumid central highlands, an area that ranges from 800 to 1300 m and is mainly constituted by grassland formations and small patches of Tapia forests; (M) subhumid central montane belt, an area that ranges from 1300 to 1800 m and includes a mosaic of Tapia forest, woodlands, grasslands and rocky habitats; (R) humid eastern and northern montane belt, an area that ranges from 800 to 1800 m and mainly constituted by moist, broadleaf forests that, because they grow at certain altitude, are not as tall as lowland rainforests; and (E) highest areas, which are disjoint small patches comprising all areas above 1800 m and up to the summit of Maromokotro (2876 m) and include a mosaic of ericoid thickets, mountain grasslands and rocky habitats (these correspond, broadly speaking, to the massifs of Tsaratanana [and its surrounding mountain chains], Marojejy, Anjanaharibe, Ankaratra, Vavavato, Ibity [and its neighbor mountains], Itremo, Andringitra [and its satellite peaks] and Beampingaratra). We excluded the Lowland Sambirano Rainforest biome (grey area in Figure 1) represented in some maps at the north-western tip of the island, and coastal mangroves (not shown in Figure 1), because none of the included *Helichrysum* species inhabits these regions.

Based on all these criteria, our proposed biogeographical areas can be viewed as macroecological units, which despite not being geographically continuous in some cases, share bioclimatic conditions de-

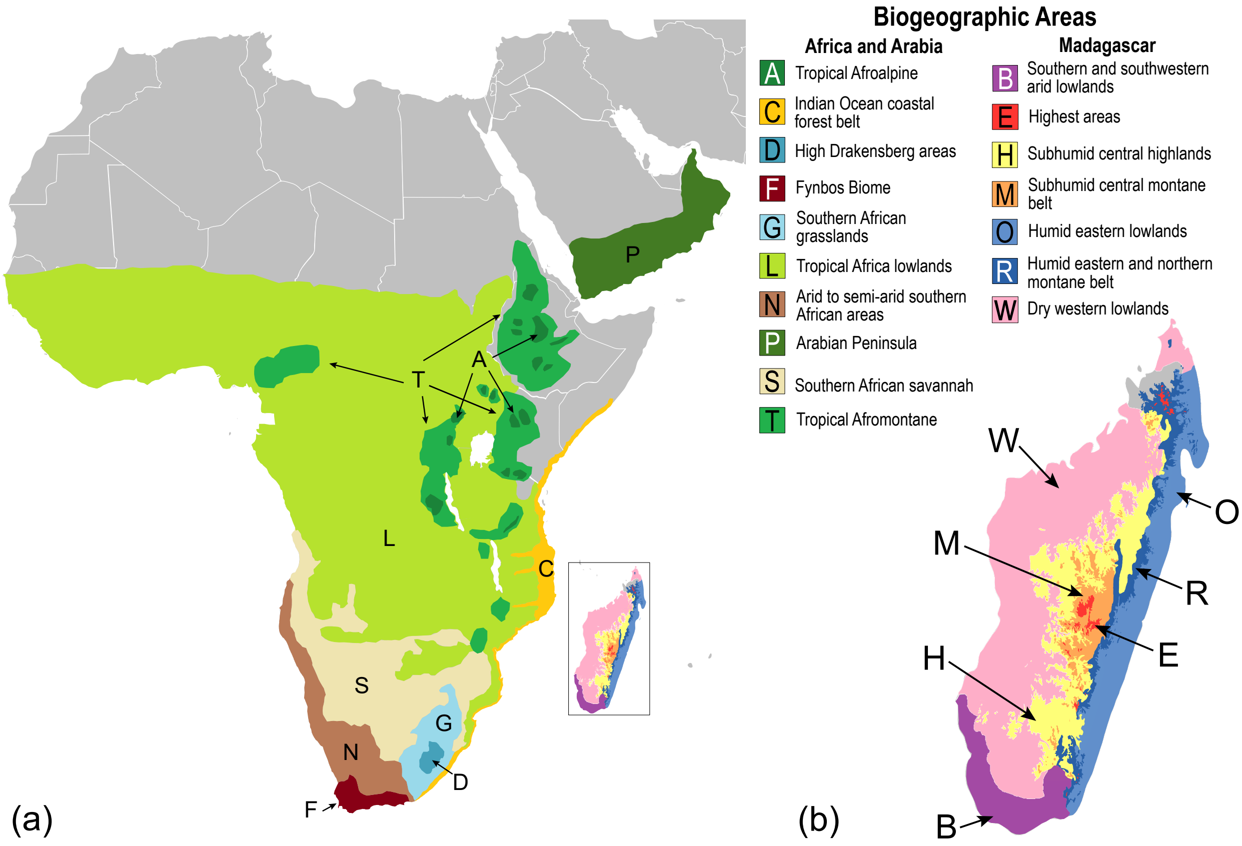


Figure 1. Maps illustrating the 17 geographic areas defined in this study. The colors and labels correspond to those in the biogeographical reconstruction analyses. (a) General map at scale, including continental Africa, Arabia and Madagascar. The dark green spots within the tropical Afromontane area are an overrepresentation of the tropical Afroalpine area. (b) Closeup of Madagascar's map to facilitate area distinction.

termined by factors such as temperature and precipitation that likely filtered out lineages according to their general bioclimatic preferences.

Species occurrence in the geographical areas was assigned considering information on distribution and elevational range (based on Humbert, 1962; Catalogue of the Vascular Plants of Madagascar, 2024; herbarium records and own field observations). Some species are only occasionally or marginally present in other areas outside of their main distribution range. In these cases, we only considered the core area/s or elevational belts in which there are a significant number of occurrences of the species, or that can be interpreted as the central distribution of the species.

We carried out biogeographic range evolution analyses in R with the package *BioGeoBEARS* (Matzke, 2013), using the ML time-calibrated tree as input. The maximum number of areas for any node was set to three, which is the highest number of areas occupied by the most widespread extant taxon in our study. We tested the fit of three biogeographical models: Dispersal-Extinction-Cladogenesis (DEC; Ree et al., 2005; Ree & Smith, 2008), a likelihood implementation of the Dispersal-Vicariance model (DIVAlike; Ronquist, 1997), and the BayArea model (BAYAREAlike; Landis et al., 2013). We also tested a more complex version of each model that accounts for founder-event speciation by adding the jump-dispersal parameter (+j). We compared and chose the best fitting model based on the Akaike Information Criterion (AIC) and AIC weight (AICw), considering that the comparison of DEC and DEC+j models is statistically valid based on Matzke's (2022) reply to Ree & Sanmartín (2018). We also performed Biogeographic Stochastic Mapping (BSM; Dupin et al., 2017) on the best fitting model (DEC+j) to estimate the frequency and types of biogeographical events, taking the mean and standard deviation of event counts from 100 BSM replicates.

3. Results

3.1 Molecular Data Processing

Out of the 1061 loci targeted, we initially recovered 929 loci. Based on pairwise sequence divergence histograms (Supplementary Figure S2), we estimated that values of 7.0 to 19.0 % of divergence indicated paralogy. Using the *Helichrysum*-customized reference tailored to our dataset, 228 (± 58) paralogous loci were detected on average (Supplementary Table S3). After filtering for missing data and splitting alignments containing paralogs, which resulted in 322 new alignments, we performed phylogenetic inference with a total of 971 loci. The aligned length of each locus averaged 278 bp (ranging from 37 to 735 bp, see Supplementary Table S4). Each locus had on average 77 (ranging from 4 to 264) parsimony informative sites and 113 (ranging from 5 to 383) variable sites. The average proportion of missing data was 2.8% (ranging from 0 to 65%). Concatenation of all loci resulted in a supermatrix with 295610 bp and 384 taxa.

3.2 Phylogenetic Analyses

The topologies of the phylogenetic trees inferred using the concatenation approach (hereafter, ML tree, collapsed version Figure 2 and full version Supplementary Figure S3) and the summary-coalescence approach (hereafter, ASTRAL tree, Supplementary Figure S4) based on 971 loci were congruent except for a few intermediate nodes, which are supported in both trees but show different relationships between species. In particular, two montane Malagasy clades are recovered as close relatives in the ML tree, but distant in the ASTRAL tree (ML tree node 367). In both trees, there are six lineages constituted by Malagasy species (MAD1-MAD6, as shown in Figure 2). Three of them are represented by a single species –*Helichrysum mahalafy* (MAD1), *H. plantago* DC. (MAD3) and *H. madagascariense* DC.(MAD5)–, each of them sister to an African species; a fourth one is constituted by four Malagasy species and a South African species; and the rest of Malagasy species constitute two speciose lineages, MAD2 and MAD6, the latter including some African species.

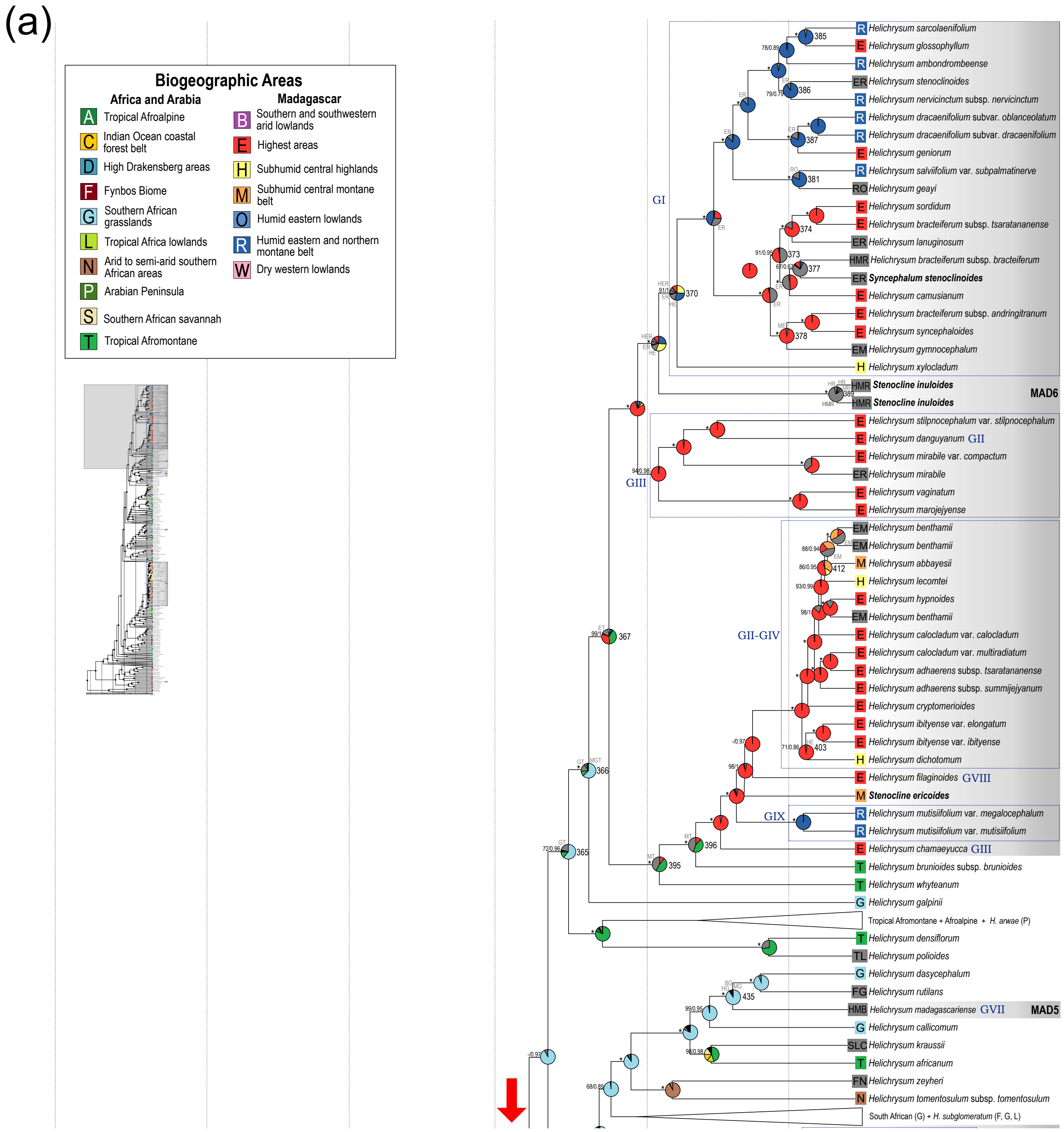
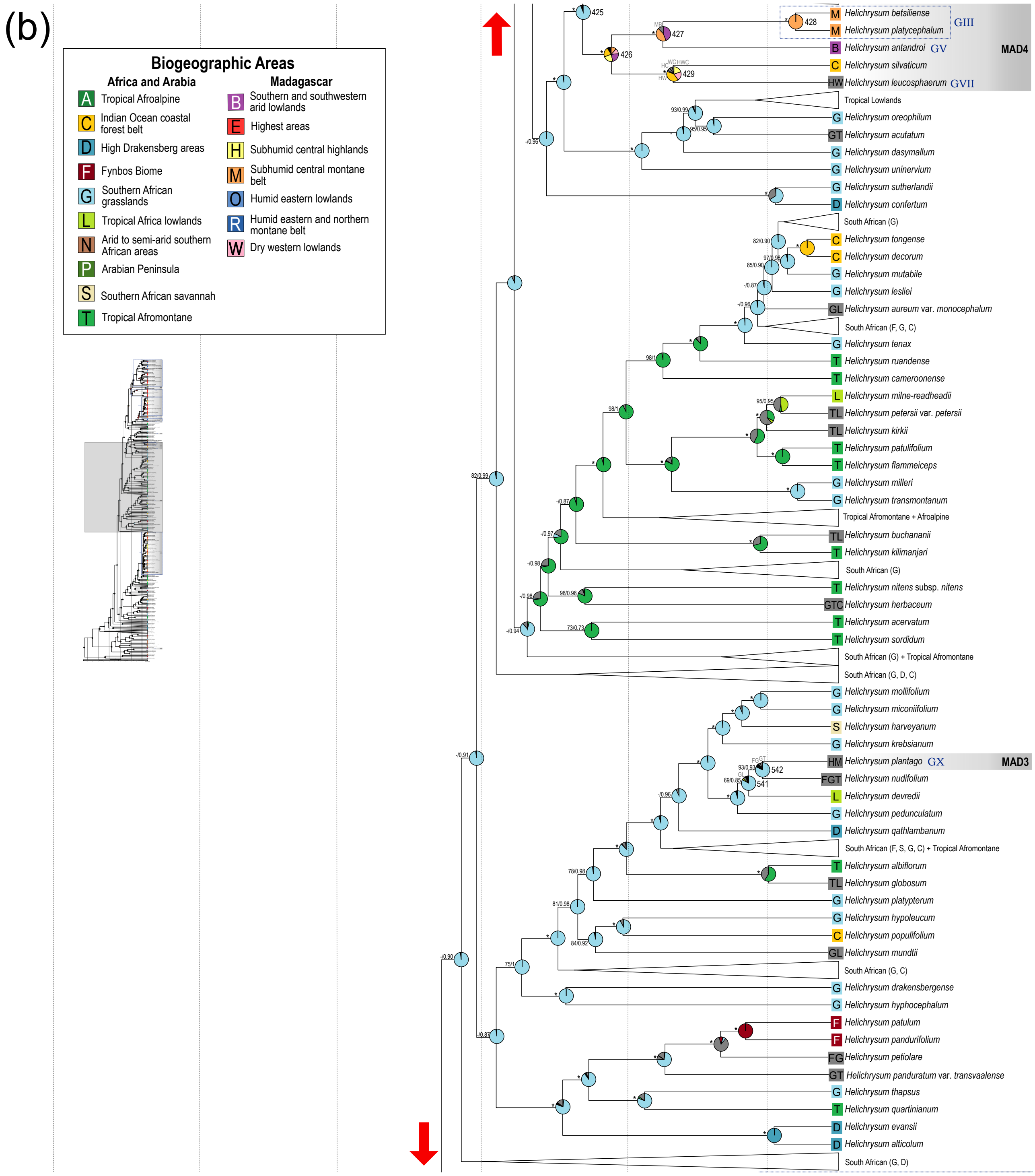
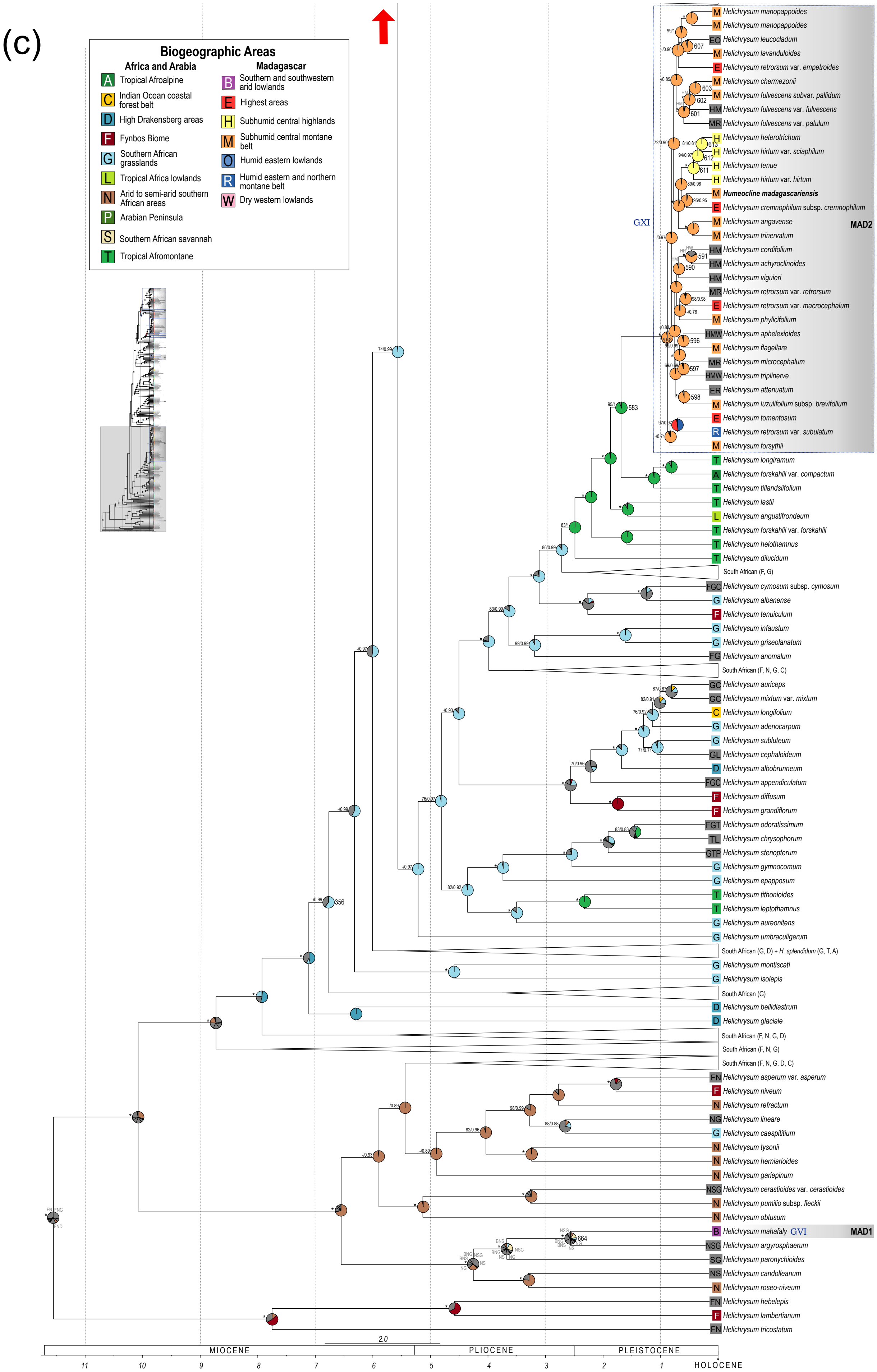


Figure 2. Ancestral range estimation of *Helichrysum* using the best-fitting model DEC+j. It is based on a time-calibrated phylogeny generated under the concatenation approach using target-enrichment data (Compositae1061 probe set). Pie charts at nodes show the relative probability of the possible states (areas in primary colors, combinations of areas in grey). Relevant node numbers are to the right of the node. Support values are to the left of the node. The first numerical value corresponds to the BS metric and the second to the TBE metric. Asterisks indicate nodes supported with a BS and TBE of 100 and 1, respectively. Names of species corresponding to genera other than *Helichrysum* are in bold. Malagasy clades highlighted in grey and labeled MAD1 to MAD6 are shown in pairs across panels: (a) includes clades MAD5 and MAD6; (b) includes clades MAD3 and MAD4; (c) includes clades MAD1 and MAD2. Humbert's (1962) taxonomic groups are indicated in blue. Unrelated African clades are collapsed and their geographical distribution is indicated to the right. The complete tree obtained from the ancestral range estimation analyses is shown in Supplementary Figure S6.



(**Figure 2.**, cont.)



(*Figure 2.*, cont.)

Lower overall clade support was obtained with the summary-coalescence approach: 61% of the nodes received significant support (LPP ≥ 0.95), whereas the percentage of significantly supported nodes in the ML tree was 85% and 96% according to BS and TBE metrics (≥70%), respectively. The summary-coalescence analysis based on the 25 most informative loci resulted in a tree (Supplementary Figure S5) that is congruent with the former ASTRAL tree, but even overall lower supports were obtained (only 25% of the nodes received significant support). Because of that, we will not further discuss the results of this tree.

3.3 Biogeographic Reconstruction

The best-fitting biogeographical model was DEC+j according to AIC values (Table 1). Regarding dispersal types, cladogenetic founder-event processes (j = 0.0085) had a slightly larger contribution than anagenetic range expansions (d = 0.0076).

Ancestral range reconstruction analyses (probabilities in Supplementary Table Table S5, full tree in Supplementary Figure S6) suggest that the genus most probably originated and initially diversified in the winter rainfall regions of western South Africa (Fynbos Biome, F and arid to semi-arid southern Africa, N) around 11.6 mya (11.4 – 15.0 95% CI), although the ancestral range may have also included some of the surrounding areas (the southern African grasslands, G; and/or the high Drakensberg area, D). The largest lineage in our tree notably diversified in the southern African grasslands and dispersals northwards and eastwards occurred from the late Miocene onwards (c. 8 mya, 6.5 – 11.7 95% CI, node 356). Repeated dispersals gave rise to the current global distribution of *Helichrysum*.

Table 1. Summary statistics of the biogeographic models tested in BioGeoBEARS. The best-fitting model used to infer the most likely area occupied by the ancestors of *Helichrysum* is highlighted in bold (DEC+j). Values of parameters for dispersal (d), extinction (e), founder effect (j), likelihood scores (lnL) and Akaike Information Criterion (AIC) are provided.

Model	lnL	No. free parameters	d	e	j	AIC	AIC weights
DEC	-1077	2	0.012	0.011	0	2157	2.6e-33
DEC+J	**-1001**	**3**	**0.0082**	**1e-12**	**0.0084**	**2007**	**1**
DIVALIKE	-1009	2	0.014	3.4e-09	0	2201	7.4e-43
DIVALIKE+J	-1028	3	0.0091	1 e-12	0.0087	2062	1.3e-12
BAYAREALIKE	-1088	2	0.010	0.23	0	2181	2.10e-38
BAYAREALIKE+J	-1012	3	0.0068	0.0077	0.011	2031	7.6e-6

Within Madagascar, Biogeographic Stochastic Mapping (BSM) analyses reveal at least six founder dispersal events from continental Africa to Madagascar (Figure 3; BSM summaries in Supplementary Table S6), all of them giving rise to endemic lineages. The main geographic source of Madagascar's colonization events was the southern African grasslands (G, with a mean of 3 colonization events) followed by the Tropical Afromontane area (T, with a mean of 2 colonization events).

According to our results, the oldest Malagasy clade (Figure 2, clade MAD6) is sister to *Helichrysum* galpinii N.E.Br., likely descending from an ancestor occurring in the southern African grasslands that dispersed at an inferred age of 3.7 mya (2.2 – 6.2 95% CI, node 366, BS = 100, TBE = 0.99). This clade comprises mainly Malagasy montane species found in the highest areas (area E), followed by species found in the subhumid northern and eastern montane belt (area R), but also includes two Tropical Afromontane species from mainland Africa. The probabilities of the ancestral range of the whole clade involve both Afromontane areas (tropical continental Africa and the highest areas of Madagascar). This clade is made up of two large and highly diversified Malagasy montane lineages (one inferred to have originated around 3.5 mya, 2.0 – 5.9 95% CI, node 367, BS = 99, TBE = 0.99 and the other at about 2.3 mya, 1.2 – 4.5 95% CI, node 396, BS = 100, TBE = 1). However, these two lineages are not closely related

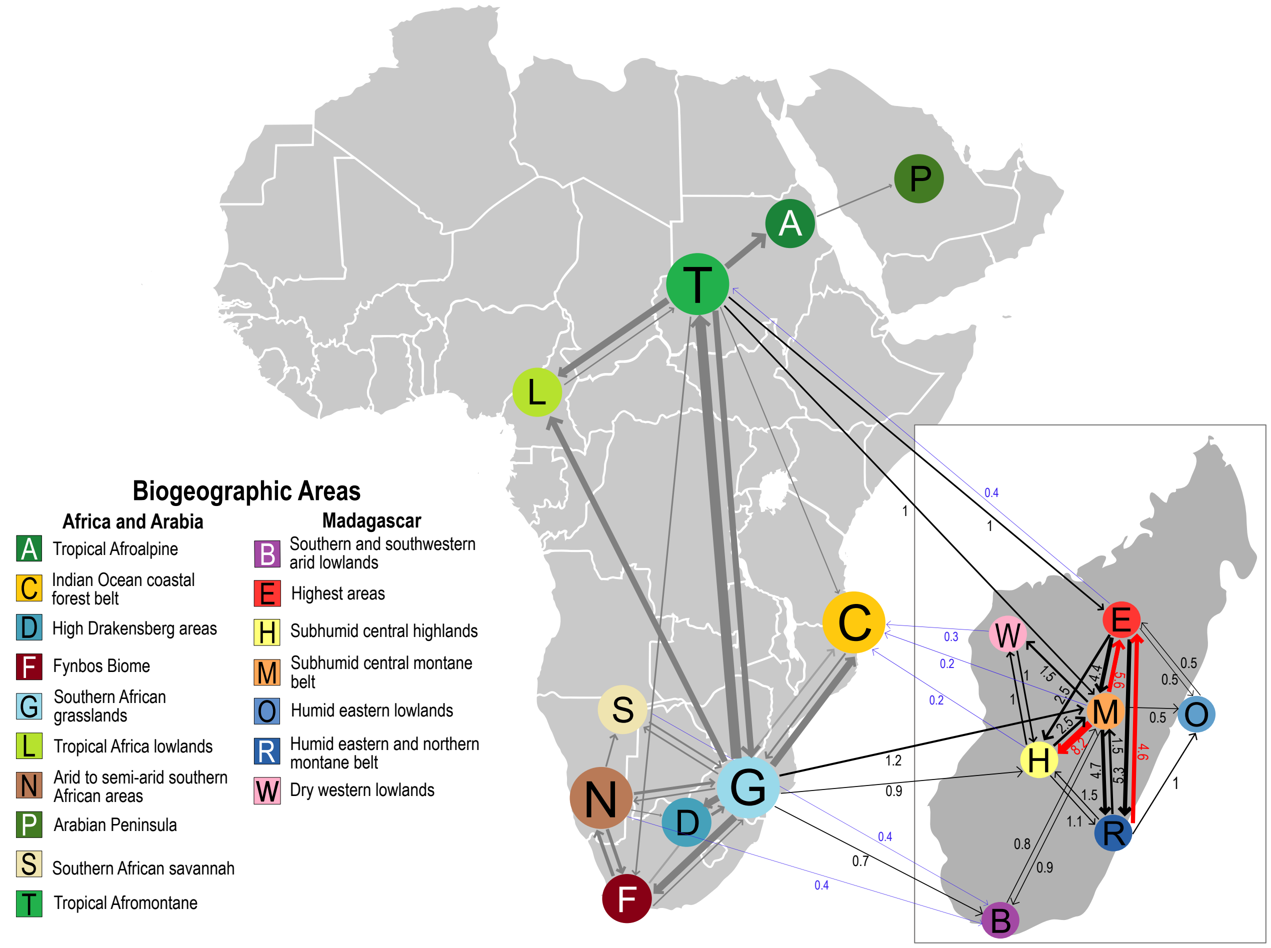


Figure 3. Summary of mean dispersal events estimated from 100 Biogeographic Stochastic Mappings (BSM) in *Helichrysum* (see all event counts in Supplementary Table S6). Arrow tips indicate the directionality of the dispersal. Numbers on the arrows are the mean of dispersal event counts (only given for dispersals related to Madagascar). Arrow thickness is proportional to the mean number of dispersals. For events involving Madagascar, either within Madagascar or between continental Africa and Madagascar: blue arrows represent dispersals with mean counts below 0.5; black arrows represent dispersals with mean counts between 0.5 and 4.9; red arrows represent dispersals with mean counts ≥ 5. Gray arrows represent dispersal events within continental Africa and the Arabian Peninsula. For readability, dispersals below 0.7 involving only these regions have not been represented.

to each other in the ASTRAL phylogeny. Therefore, these results should be taken with caution since concatenation and coalescence-based phylogenetic inferences provide conflicting relationships around these nodes. The two Malagasy montane clades notably diversified during the Pleistocene, with some species colonizing or extending their area into the subhumid central montane belt (area M, e.g. *H. abbayesii* Humbert), the grasslands of Madagascar's central highlands (area H, e.g. *H. lecomtei* R. Vig. & Humbert and *H. xylocladum* Baker) or the humid eastern lowlands (area O) in the case of *H. geayi* (node 381).

The second oldest colonization of Madagascar has an inferred age of 3.6 mya (2.2 – 5.9 95% CI, node 425, BS = 100, TBE = 1, MAD4), also from a southern African grassland ancestor. The ancestral range of the Malagasy descendant is highly uncertain due to the wide range of the extant species belonging to this clade, which include four areas in Madagascar (southern and southwestern arid lowlands, area B; subhumid central montane belt, area M; subhumid central highlands, area H and dry western lowlands, area W) and one area from continental Africa (Indian Ocean Coastal Forest Belt, area C). These results suggest a back-colonization to the southern African coast, inferred to have occurred 2.3 mya (1.2 – 4.5 95% CI, node 429, BS = 100, TBE = 1) giving rise to *H. silvaticum* Hilliard.

Our results suggest that a recent dispersal from the tropical Afromontane area took place at an inferred age of 1.7 mya (0.9 – 3.3 95% CI, node 583, BS = 100, TBE = 1, MAD2) giving rise to a recent radiation in the subhumid central montane belt of Madagascar (area M; ca. 0.9 mya, 0.4 – 2.0 95% CI, node 586, BS = 95, TBE = 1). Again, this was followed by several independent dispersals within Madagascar, mainly consisting of populations establishing in lower-elevation habitats. The distribution range of some species extended beyond the Central Highlands (e.g. *Helichrysum leucocladum* Humbert into the humid eastern lowlands and *H. aphelexioides* DC. into the western arid lowlands). Nevertheless, there were also a few instances of upward dispersal to higher elevations, such as the case of *H. cremnophilum* Humbert, whose distribution range reached the highest areas.

Three other independent dispersal events to Madagascar took place, each one giving rise to a single Malagasy species. *Helichrysum mahafaly* (MAD1) from the southern and south-western arid lowlands (area B) of Madagascar is nested within a clade of species from the geographically distant arid to semi-arid southwestern African area (area N). This dispersal event is inferred to have taken place 2.6 mya (1.4 – 4.7 95% CI, node 664, BS = 100, TBE = 1). The Malagasy widespread *H. madagascariense* (MAD5) diverged from its sister from the southern African grasslands around 1.8 mya (0.8 – 3.7 95% CI, node 435, BS = 100, TBE = 1). Lastly, the ancestor of *H. plantago* (MAD3) is also inferred to have dispersed from the southern African grasslands at about 1.1 mya (0.5 – 2.6 95% CI, node 542, BS = 93, TBE = 0.93).

4. Discussion

Here, we explore the evolutionary dynamics of colonization and diversification on Madagascar using the species-rich genus *Helichrysum*. This is the first time that the origin and geographic diversification of a Malagasy plant group have been explored using a large phylogeny based on NGS data from over 300 taxa, including more than half of the species endemic to the island. We report at least six independent colonizations from continental Africa from the Pliocene onwards, most of these from the southern African grasslands and the tropical Afromontane areas. We also recover a single back-colonization to the southeastern African coastal region. Our results suggest a clear tendency to retain bioclimatic preferences and morphological characters in the colonization of Madagascar by *Helichrysum*. However, our findings indicate that once lineages established on the island, their diversification was partly associated with biome shifts, especially shifts from high-elevation to lower-elevation habitats. We also identify several simultaneous and recent *in-situ* radiations in the montane areas of the island, including the subhumid central montane belt and the highest areas, which enriched the diversity of their flora. Although some lineages are morphologically and functionally very uniform, others include a great variety of life forms and morphological characters.

4.1 Helichrysum colonized Madagascar multiple times from Africa since the Late Pliocene and likely returned once to the continent

Given the geological history of Madagascar, the two possible mechanisms by which Madagascar's biota was established are Gondwanan vicariance and long-distance dispersal. We infer a recent colonization of Madagascar by *Helichrysum*, with the earliest dispersal event estimated at around 3.7 mya (2.2 – 6.2 95% CI), followed by multiple independent colonizations until at least 1 mya (0.5 – 2.6 95% CI). Considering that Madagascar separated from Africa more than 150 mya, *Helichrysum* must therefore have colonized the island via long-distance dispersal across the Mozambique Channel. Short-lived land bridges across this ancient biogeographical barrier at different geological times have been proposed (Masters et al., 2021), but their existence has not been demonstrated, keeping their true biogeographical impact under debate (Ali & Hedges, 2023; Aslanian et al., 2023). In any case, the last potentially available land bridges would have probably disappeared before the first colonization of Madagascar by *Helichrysum*. Most Compositae species produce anemochorous fruits, and *Helichrysum* cypselae are extremely small (< 1 mm long), which enables wind dispersal (Nathan et al., 2009). Since the Miocene, dominant sea currents and winds in the southwestern Indian Ocean region have predominantly flowed westward, reducing the likelihood of dispersal from Africa to Madagascar, except through cyclones crossing the Mozambique Channel in the opposite direction. At the same time, this shift in ocean currents and wind directionality increased the probability of dispersal from Asia to Madagascar, as reflected in the strong floristic affinities between south-eastern Asia, India and Madagascar, especially in the humid northern and eastern regions (Schatz, 1996; Ali & Huber, 2010; Warren et al., 2010; Buerki et al., 2013). Another plausible mechanism that could have acted in parallel is avian dispersal, as there is evidence of direct seed dispersal by birds, either stuck in the plumage or contained in the digestive tract of the consumed prey (Padilla et al., 2012).

According to Buerki et al. (2013), during the Miocene there was an increase in colonizations of Madagascar by non-endemic genera, coinciding also with the emergence of most Malagasy endemic genera. Our dating indicates a much more recent colonization of Madagascar by *Helichrysum,* occurring during the Pliocene and Pleistocene. Buerki et al. (2013) identified a few genera with similarly recent inferred colonization dates, such as *Hibiscus* L. (Malvaceae), *Paracorynanthe* Capuron (Rubiaceae), *Colvillea* Bojer and *Lemuropisum* H.Perrier (Fabaceae). Another notable example within the tribe Gnaphalieae with a similar age of colonization is the genus *Stoebe* L. (Bergh & Linder, 2009). Nevertheless, unlike the diversification patterns observed in *Helichrysum*, all these genera are represented in Madagascar by few species.

The already mentioned prevailing westward currents and winds, which have dominated since the Miocene, should have increased the likelihood of dispersal from Madagascar to Africa, thus the potential of back-colonizations to the mainland. However, documented examples of such events are scarcely found in the literature. Some notable exceptions include Dombeyoideae (Malvaceae s.l.) with at least five migrations back to the continent (Skema et al., 2023), the Coffeeae and Psychotrieae alliances (Rubiaceae, Kainulainen et al., 2017; Razafimandimbison et al., 2017), *Croton* (Euphorbiacee, Haber et al., 2017) and grammitid ferns (Bauret et al., 2017). Here, we inferred at least one potential back-colonization from Madagascar to continental Africa in *Helichrysum*, involving *H. silvaticum*, which occurs in the Indian Ocean Coastal Forest Belt (area C, in Mozambique and South Africa). This biome, separated from Madagascar by the 415 km wide Mozambique Channel (Masters et al., 2021), is the geographically closest continental biome to the island. This species is sister to *H. leucosphaerum* Baker which inhabits Madagascar's subhumid central highlands (area H) and the dry western lowlands (area W). Both species belong to the MAD4 clade (Figure 2), which originated from a dispersal from the southern African grassland around 3.6 mya (2.2 – 5.9 95% CI), giving rise to five species, four of which are endemic to Madagascar. The most likely scenario in our reconstruction is a single dispersal to Madagascar, followed by diversification and a back-colonization of the ancestor of *H. silvaticum*. However, uncertainty at nodes 426 and 429 suggests an alternative possibility in which the ancestor of both species remained in the continent, in which case *H. leucosphaerum* would be the result of a second, independent colonization of Madagascar within the MAD4 clade. Given the clade's high morphological diversity, increased taxonomic sampling –particularly of continental African species, since Madagascar morphotypes and taxonomic groups are all represented– could reduce uncertainty regarding the biogeographic and evolutionary history of this lineage.

4.2 Biome conservatism prevails in the colonization of Madagascar by Helichrysum: the continental Afromontane region reveals as the main source of Malagasy lineages

Our results for *Helichrysum* are consistent with previous findings of biome conservatism following transoceanic dispersal (Crisp et al., 2009; Vences et al., 2009) as well as the maintenance of biome-related adaptations (Wiens & Graham, 2005; Wiens et al., 2010). Specifically, we infer that five out of the six hypothesized independent colonization events of Madagascar by *Helichrysum* happened in the island's highlands and mountains by ancestors from the continental Afromontane region *s.l.,* which apart from the tropical Afromontane area also includes the high southern African grasslands in the Drakensberg mountains. This suggests that *Helichrysum* dispersed from areas with similar broad-scale climatic conditions in continental Africa, reinforcing the idea that long-distance dispersal to Madagascar by *Helichrysum* was associated with biome conservatism. In greater detail, we identify two colonization events from the southern African montane grasslands to the Malagasy montane grasslands. In these cases, the colonizers preserved not only broad-scale bioclimatic preferences but also morphological traits. For example, *H. plantago* (MAD3) shares traits with its sister species *H. nudifolium* (L.) Less (Figure 4, more details in Table S7) and other closely related species. Likewise, *H. madagascariense* (MAD5) is strikingly similar to the species comprising its sister group, *H. dasycephalum* O.Hoffm. and *H. rutilans* D.Don, and to *H. callicomum* Harv., the sister to all three (Figure 4, more details in Table S7).

Malagasy grasslands, now extensive, have a complex history, and whereas some grassland formations are ancient, some others are much more recent and anthropogenic (Vorontsova et al., 2016; Hackel et al., 2018, Joseph & Seymour, 2020; Solofondranohatra et al., 2020; Joseph et al., 2021; Lehmann et al., 2022; Bond et al., 2023; Tiley et al., 2024b). Thus, some form of grassland environment was undoubtedly already available for *Helichrysum* with relevant adaptations to "dispersify" (Donoghue, 2008) from similar African environments. Dispersals from southern Africa to Madagascar have been reported for multiple grasses and sedge clades (e.g. Linder et al., 2014; Larridon et al., 2021). However, little is known about non-graminoid grassland species, although *Alchemilla* L. (Gehrke et al., 2016) and *Stoebe* (Bergh & Linder, 2009), currently found in the highest areas of Malagasy mountains, probably dispersed from the montane southern African grasslands.

Two ancestors from the tropical Afromontane area successfully established and later radiated in the mountains of Madagascar, one radiation resulting in most of the current species growing in the subhumid central montane belt (area M, Figure 2, clade MAD2) and the other resulting in most of the species occurring in the highest areas (area E, Figure 2, clade MAD6). The Tapia woodland mosaics in the subhumid central montane belt are structurally equivalent to the miombo woodland savannas of continental Africa (Alvarado et al., 2014) and the iconic montane ericoid shrublands found in Africa display similar physiognomy to those growing on the summits of the highest Malagasy mountains (Silander et al., 2024).

Examples of Afromontane migrations from tropical Africa to Madagascar involve several grass clades (Vorontsova et al., 2016; Hackel et al., 2018), sedge lineages (Larridon et al., 2021), *Kniphofia* Moench (Ramdhani et al., 2009), and the diverse *Erica* L. (Pirie et al., 2019). In the case of clade MAD2, morphological traits have been conserved in this jump between montane biomes in Africa and Madagascar by *Helichrysum* (Figure 4, detailed information in Table S7).

Lastly, we propose an alternative interpretation of the ancestral range reconstruction for the clade MAD6 (Figure 2). According to our ML phylogeny, *Helichrysum galpinii* (inhabiting the southern African grasslands, Figure 2, node 366) is sister to two Malagasy lineages, suggesting that the entire clade likely originated from an ancestor in the southern African grasslands. This clade also includes a couple of tropical Afromontane species, introducing uncertainty about the biogeographic range of the ancestor of these two Malagasy montane lineages. However, our ASTRAL phylogeny (Supplementary Figure S4) presents a different topology, where the two Malagasy montane clades are recovered as unrelated: one is sister to *H. galpinii* and the other is sister to *H. whyteanum* Britten and *H. brunioides* Moeser. This topology thus suggests two independent colonizations of Madagascar's montane area, one from the southern African grasslands and the other from the tropical Afromontane region. Indeed, the morphological distinctness of these two Malagasy montane lineages is consistent with the hypothesis of two independent colonizations.

Finally, we would like to note a striking case of conservatism in bioclimatic preferences in the colonization of Madagascar by the ancestor of *Helichrysum mahafaly* (Figure 2, clade MAD1). This species, which is endemic to the extremely hot and arid southwestern part of the island, is nested within a southern African clade from the western arid to semi-arid regions. In other words, *H. mahafaly* occupies a habitat characterized by environmental stressors strikingly similar to those experienced by its sister species in

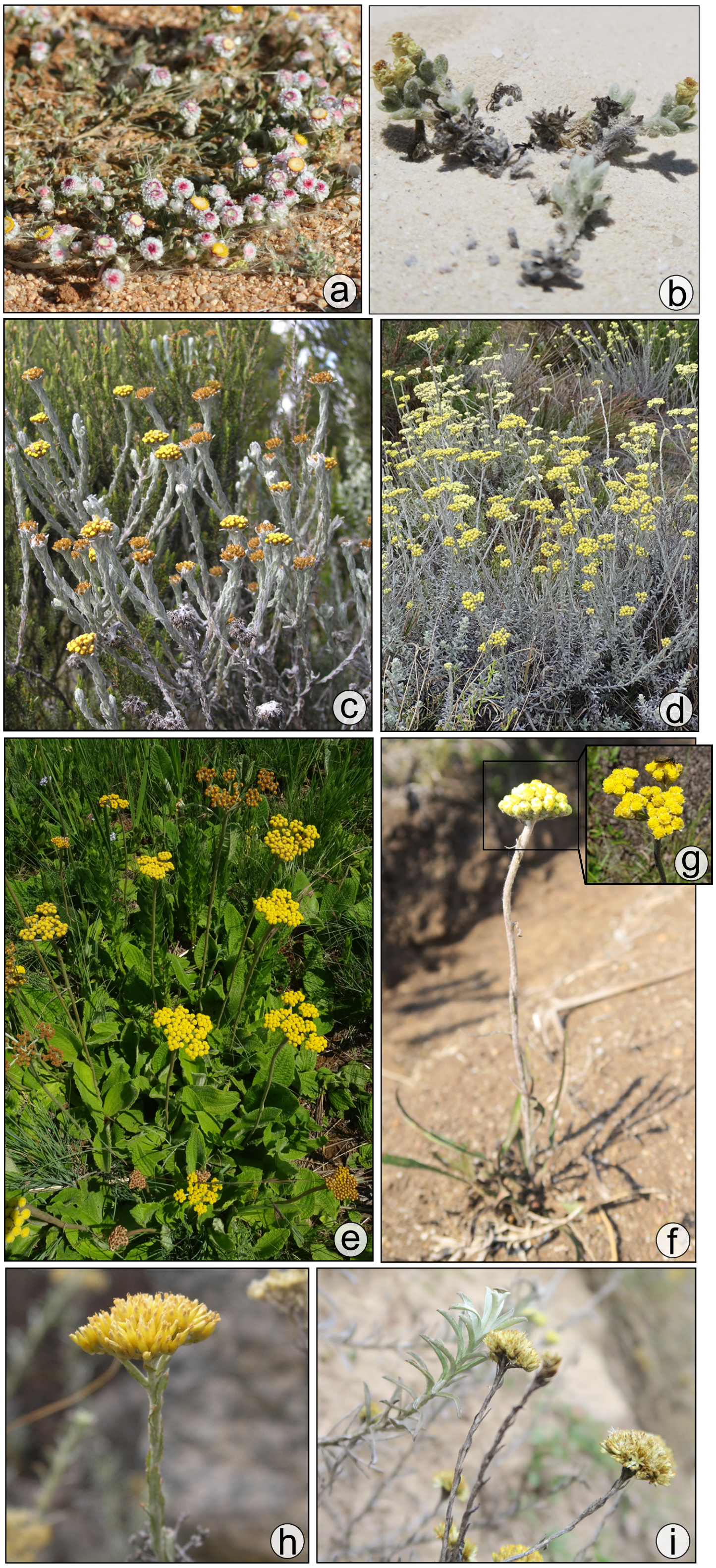


Figure 4. Species of *Helichrysum* illustrating the similarity between closely related continental Africa and Malagasy taxa. First row: (a) *Helichrysum argyrosphaerum* from South Africa, which is the inferred sister to (b) *Helichrysum mahafaly* (MAD1) from Madagascar. Both are prostrate herbs with solitary, subglobose capitula, inhabiting arid sandy habitats. Second row: (c) *Helichrysum forskahlii* var. *compactum* a tropical Afromontane species closely related to taxa from the Malagasy clade MAD2; (d) *Helichrysum fulvescens*, a Malagasy species from clade MAD2, mostly subshrubs with numerous small cylindrical capitula arranged in corymbs and fimbrilliferous receptacles. Third row: (e) *Helichrysum nudifolium* var. *pilosellum* from South Africa, inferred sister species of *Helichrysum plantago* (f) and (g) from Madagascar (MAD3). Both are perennial herbs with basal leaf rosette and dense corymbs of capitula with yellow involucral bracts. Fourth row: (h) *Helichrysum rutilans*, a southern African species closely related to (i) *Helichrysum madagascariense* (MAD5) from Madagascar, both subshrubs with dense corymbs of narrow cylindrical capitula. (Photos: Mercè Galbany-Casals, except (a) and (e): Marinda Koekemoer).

southern Africa (*H. argyrosphaerum*, Figure 4). Other examples of Malagasy plants adapted to arid environments and likely originating from African ancestors include baobabs (*Adansonia* L., Leong Pock Tsy et al., 2009), Portulacaceae members (Eggli, 1997; Hershkovitz & Zimmer, 2000) and *Neoapaloxylon* Rauschert (Leguminoseae, Choo et al., 2020).

4.3 Diversification within Madagascar

Over time, lineages that colonize new territories, such as islands, can lead to various evolutionary outcomes. These may include limited or extensive speciation, remaining relatively unchanged or adapting to new conditions. In *Helichrysum*, we observed instances of different diversification patterns. Some lineages have speciated considerably. The timing of colonization and speciation on Madagascar by the ancestors of these lineages coincided with the transition towards the so-called "Icehouse Climate State" (Westerhold et al., 2020). The cooling trend resulted in increased aridity, pronounced seasonality and lower atmospheric CO_2 concentrations, which have been linked to several plant radiations (especially in the two most species-rich grassland families Poaceae and Compositae, Palazzesi et al., 2022) in the open grassland biomes that had proliferated worldwide during the Miocene (Spriggs et al., 2014). In parallel, ocean circulation patterns changed around 3 – 4 mya, resulting in a decrease in rainfall and increased aridity in both East Africa and Madagascar (De Wit, 2003). Most speciation events leading to extant *Helichrysum* species are inferred to have occurred during the Pleistocene, as reported for other Malagasy groups such as orchid genera (*Aeranthes* Lindl.*, Angraecum* Bory and *Jumellea* Schltr., Andriananjamanantsoa et al., 2016) and scaly tree ferns (Cynthaeaceae, Janssen et al., 2008).

As is common for many groups that have diversified in archipelagos (Rundell & Price, 2009), *Helichrysum* likely experienced a combination of adaptive and non-adaptive radiations (summarized in Table S7). Some clades such as the clade MAD4 (Figure S7) and the youngest subclade in MAD6 (Figure 5)**,** comprise ecologically and morphologically well-differentiated species. This heterogeneity suggests ecological speciation (i.e. reproductive isolation arises through divergent natural selection in populations adapted to different ecological environments; Rundle & Nosil 2005) as the main evolutionary force. Similar evolutionary patterns have been reported for other groups, such as the radiation of the small Afromontane genus *Arrowsmithia* DC. (formerly *Macowania* Oliv., also belonging to the tribe Gnaphalieae) in the Drakensberg mountains (Bentley et al., 2014). Nonetheless, members of the more species-rich clades within *Helichrysum* tend to be morphologically uniform, making them easily identifiable as a cohesive group. One of such examples is a subclade in MAD6, which comprises all sampled species from Humbert's taxonomic group I (GI, Humbert, 1962). Despite their morphological similarities (Figure 5), species in this group show notable variation in geographic distribution and ecological preferences. This suggests that diversification in GI likely resulted from a combination of geographic (or allopatric) and ecological speciation events. However, despite the occurrence of several shifts in broad-scale bioclimatic preferences, the overall pattern within this clade points towards biome conservatism. Each of the two major subclades within GI has a distinct primary distribution area (see Figure 2): one diversified in the northern and eastern montane belt, while the other diversified on mountain summits. A similar pattern is observed in MAD2, which includes all sampled species of taxonomic group XI (GXI, Humbert, 1962). Although the group is morphologically well-defined (more details in Table S7), it exhibits considerable variation in growth form, capitula size, bract color, and ecological preferences.

Our findings indicate that biome shifts likely occurred in Madagascar*,* following speciation within the colonized area and/or expansion of the distribution range across Madagascar's biomes. Most shifts in bioclimatic preferences in Malagasy *Helichrysum* are associated with independent dispersals from higher to lower elevation environments. The colder glacial periods of the Pleistocene may have promoted this transition to mid-low elevations. For instance, the elevational range of many species in group XI (GXI, clade MAD2) appears to have expanded downward to the highland grasslands (e.g. *H. fulvescens* DC. and *H. viguieri* Humbert), while multiple species that are currently restricted to the highlands' grasslands have

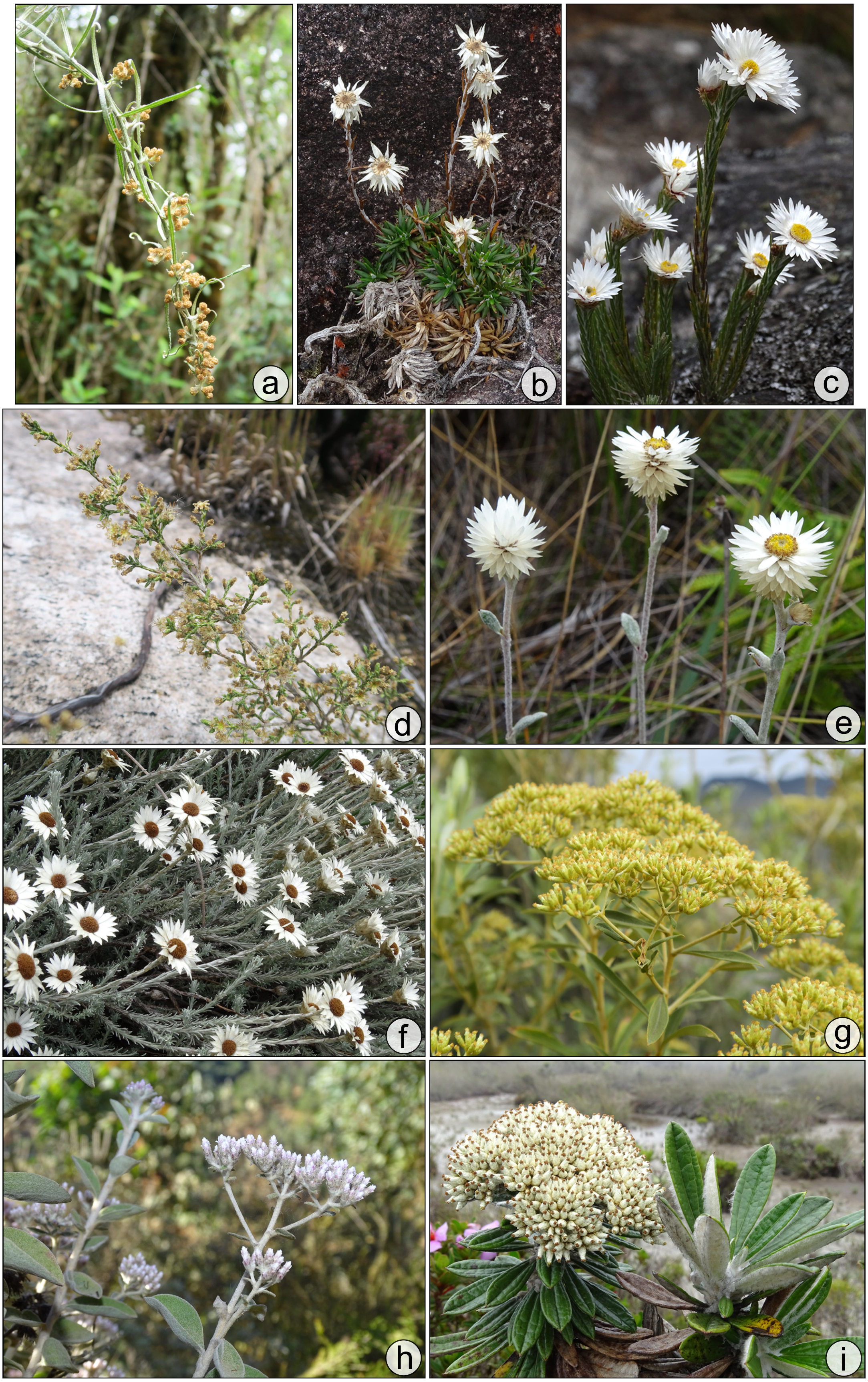


Figure 5. Morphological diversity of clade MAD6. (a) *Helichrysum mutisiifolium* (GIX), a lianioid species; (b) *Helichrysum chamaeyucca*, (e) *Helichrysum marojejyense*, and (f) *Helichrysum danguyanum*, the three from GIII, all with long white radiant involucral bracts; (c) *Helichrysum ibityense* var. *ibityense* (GII) and (d) *Helichrysum hypnoides* (GIV), both with ericoid leaves; (g) *Helichrysum gymnocephalum*, (h) *Helichrysum xylocladum* and (i) *Helichrysum geniorum*, the three from GI, shrubs or treelets with tiny capitula grouped in glomerules surrounded by leaves and gathered in big corymbose synflorescences. (Photos: Mercè Galbany-Casals).

montane ancestors. An example is the clade comprising *H. hirtum* Humbert*, H. tenue* Humbert and *H. heterotrichum* Humbert (GXI), in which biome conservatism prevailed after an initial shift to lower elevations. Exceptionally, a few montane species have colonized the dry western lowlands (e.g. *H. triplinerve* DC) or the humid eastern lowlands (e.g. *H. geayi*), two biomes with few *Helichrysum* species. Such repeated downward migrations are infrequently reported in the literature (e.g. *Dendrosenecio* (Hauman ex Hedberg) B.Nord., Knox & Palmer, 1995), the general evolutionary trend being migration from lower to higher elevations (see Gamisch et al., 2016 and references therein).

Lastly, two of the three Malagasy lineages represented by a single species, and descending from African ancestors, are morphologically distinct from all other Malagasy *Helichrysum* species. This was recognized in the most recent taxonomic treatment (Humbert, 1962), which placed these species into monospecific taxonomic groups: *H. mahafaly* in group VI (Figure 4, GVI, clade MAD1) and *H. plantago* in group X (Figure 4, GX, clade MAD3). This is not surprising as these species closely resemble their relatives in continental Africa, both morphologically and ecologically. Even if our sampling is not complete, we hypothesize that these two species are the only extant descendants of their respective colonizing ancestors. The lack of diversification in the lineage represented by *H. mahafaly* could be attributed to its highly specialized niche, which results in geographical isolation to the southwestern coast. Conversely, *H. plantago* is a widespread and opportunistic species that thrives in the heavily degraded highlands. Its recent origin and its ability to exploit this transformed landscape, which lacks significant geographical barriers, may have allowed it to spread widely, and all these factors may explain lack of further diversification in this clade. The sole member of the third lineage, *H. madagascariens*e (Figure 2, GVII, clade MAD5, see Figure 4) was classified by Humbert (1962) in group VII together with three other species. However, our tree shows that group VII is not monophyletic, as *H. leucosphaerum* (node 429) is placed in a separate Malagasy clade containing species from different taxonomic groups (clade MAD4 in Figure 2, see Figure S7). Without sampling the other two members of group VII, we cannot confirm whether *H. madagascariense* truly constitutes a monospecific lineage.

5. Conclusions

This study has focused on the Malagasy radiations of the highly diverse plant genus *Helichrysum* to unravel its biogeographic and evolutionary history within this long-isolated fragment of Gondwana. Our findings reveal a scenario of repeated recent transoceanic dispersal events, in which biome conservatism played a critical role in shaping speciation and diversification of Malagasy lineages. Some colonization events were followed by *in-situ* diversification, likely driven by a combination of allopatric and ecological speciation processes. The primary source of Malagasy lineages can be traced to the Afromontane region *s.l.*, including the highest mountains of tropical Africa and southern Africa. The ancestors were probably preadapted to high-elevation environments, which facilitated the colonization of the island's montane areas, harboring now most of the genus' diversity. Once established, the descendants of these African ancestors appear to have radiated *in-situ,* possibly through geographic speciation, giving rise to several high-elevation Malagasy endemics distributed across different massifs. In addition, our inferences suggest that some montane species shifted their ranges to lower-elevation environments, such as the highlands' grasslands, western dry deciduous forests or eastern evergreen forests. In some cases, these range shifts resulted in expanded distribution ranges, while in others, they promoted further speciation. In contrast, Malagasy clades coming from southern African grasslands ancestors consist of one or few widespread species. Notably, we identified a rare example of conservatism in bioclimatic preferences after long-distance dispersal from the arid southwestern Africa to Madagascar's similarly arid southwestern coastal dunes. Our findings also reveal a potential case of back-dispersal by *Helichrysum* to the southeastern coast of continental Africa, supporting the idea of sporadic biogeographical dispersals from Madagascar to mainland Africa.

References

Ali, J.R., & Hedges, S.B. (2023). The colonisation of Madagascar by land-bound vertebrates. *Biol. Rev.*, 98, 1583–1606.

Alvarado, S.T., Buisson, E., Rabarison, H., Rajariarisonm, C., Birkinsha, C., & Lowry II, P.P. (2014). Comparision of plant communities on the Ibity and Itremo massifs, Madagascar, with contrasting conservation histories and current status. *Plant Ecol. Divers.*, 7(4), 497–508.

Andrés-Sánchez, S., Verboom, G.A., Galbany-Casals, M., & Bergh, N.G. (2019). Evolutionary history of the arid climate-adapted *Helichrysum* (Asteraceae: Gnaphalieae): Cape origin and association between annual life-history and low chromosome numbers. *J. Syst. Evol.,* 57(5), 468–487.

Andriananjamanantsoa, H.N., Engberg, S., Louis, E.E., Jr., & Brouillet, L. (2016). Diversification of *Angraecum* (Orchidaceae, Vandeae) in Madagascar: revised phylogeny reveals species accumulation through time rather than rapid radiation. *PloS ONE*, 11(9), e0163194.

Antonelli, A., Smith, R.J., Perrigo, A.L., et al. (2022). Madagascar's extraordinary biodiversity: evolution, distribution and use. *Science*, 378(6623), eabf0869.

Aslanian, D., Pellen, R., Rabineau, M., et al. (2023). The postulation of intermittent land bridges as an explanation for reiterated colonization events of Madagascar by African vertebrates: An in-depth review and novel insights in honour of the late Judith Masters and Fabien Génin. *Earth-Sci. Rev.,* 246, 104585.

Bankevich, A., Nurk, S., Antipov, D., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.,* 19(5), 455–477.

Bauret, L., Gaudeul, M., Sundue, M.A., et al. (2017). Madagascar sheds new light on the molecular systematics and biogeography of grammitid ferns: New unexpected lineages and numerous long-distance dispersal events. *Mol. Phylogenet. Evol.*, 111, 1–17.

Bentley, J., Verboom, G.A., & Bergh, N.G. (2014). Erosive processes after tectonic uplift stimulate vicariant and adaptive speciation: evolution in an Afrotemperate-endemic paper daisy genus. *BMC Evol. Biol.*, 14, 27.

Bergh, N.G., & Linder, H.P. (2009). Cape diversification and repeated out-of-southern-Africa dispersal in paper daisies (Asteraceae-Gnaphalieae). *Mol. Phylogenet. Evol.,* 51(1), 5–18.

Blanco-Gavaldà, C., Galbany-Casals, M., Susanna, A., et al. (2023). Repeatedly northwards and upwards: southern African grasslands fuel the colonization of the African sky islands in *Helichrysum* (Compositae). *Plants*, 12(11), 2213.

Bolger, A.M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120.

Bond, W.J., Silander, J.A., & Ratsirarson, J. (2023). Madagascar's grassy biomes are ancient and there is much to learn about their ecology and evolution. *J. Biogeogr.*, 50, 614–621.

Buerki, S., Devey, D.S., Callmander, M.W., Phillipson, P.B., & Forest, F. (2013). Spatio-temporal history of the endemic genera of Madagascar. *Bot. J. Linn.*, 171(2), 304–329.

Burgess, N. (2004). *Terrestrial ecoregions of Africa and Madagascar: a conservation assessment.* Island Press: Washington, D.C.

Burgoyne, P.M., van Wyk, A.E., Anderson, J.M., & Schrire, B.D. (2005). Phanerozoic evolution of plants on the African plate. *J. Afr. Earth Sci.,* 43(1-3), 13–52.

Bushnell, B. (2014). BBMap: A fast, accurate, splice-aware aligner.

Callmander, M.W., Phillipson, P.B., Schatz, G.E., et al. (2011). The endemic and non-endemic vascular flora of Madagascar updated. *Plant Ecol. Evol.,* 144(2), 121–125.

Cardillo, M., Weston, P.H., Reynolds, Z.K.M., et al. (2017). The phylogeny and biogeography of *Hakea* (Proteaceae) reveals the role of biome shifts in a continental plant radiation. *Evolution*, 71, 1928–1943.

Choo, L.M., Forest, F., Wieringa, J.J., Bruneau, A., & de la Estrella, M. (2020). Phylogeny and biogeography of the Daniellia clade (Leguminosae: Detarioideae), a tropical tree lineage largely threatened in Africa and Madagascar. *Mol. Phylogenet. Evol.,* 146, 106752.

Crisp, M., Arroyo, M., Cook, L., et al. (2009). Phylogenetic biome conservatism on a global scale. *Nature*, 458, 754–756.

Costa, F.P., Schrago, C.G., & Mello, B. (2022). Assessing the relative performance of fast molecular dating methods for phylogenomic data. *BMC Genomics,* 23, 798.

Darriba, D., Posada, D., Kozlov, A. M., Stamatakis, A., Morel, B., & Flouri, T. (2020). ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. *Mol. Biol. Evol.,* 37(1), 291–294.

Darwin, C. (1859). *On the origin of species by means of natural selection or preservation of favoured races in the struggle for life.* John Murray: London.

De Wit, M.J. (2003). Madagascar: heads it's a continent, tails it's an island. *Annu. Rev. Earth Planet. Sci.,* 31(1), 213–248.

Donoghue, M.J. (2008). A phylogenetic perspective on the distribution of plant diversity. *Proc. Natl. Acad. Sci. U.S.A.,* 105, 11549–11555.

Donoghue, M.J., & Edwards, E.J. (2014). Biome shifts and niche evolution in plants. *Annu. Rev. Ecol. Evol. Syst.,* 45, 547–572.

Dupin, J., Matzke, N.J., Särkinen, T., et al. (2017). Bayesian estimation of the global biogeographical history of the Solanaceae. *J. Biogeogr.,* 44, 887–899.

Du Puy, D.J., & Moat, J. (1996). A refined classification of the primary vegetation of Madagascar based on the underlying geology: using GIS to map its distribution and to assess its conservation status. In: *Biogeography of Madagascar*. Lourenço, W.R., (ed.). ORSTROM, Paris.

Eggli, U. (1997). A synopsis of woody Portulacaceae in Madagascar. *Adansonia* sér. *3*, 19, 45–59.

Faramalala, M.H. & Conservation International (1995). *Formations Végétales et Domaine Forestier National de Madagascar.* Direction des Eaux et Forêts/CNRE/FTM: Antananarivo.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39(4), 783–791.

Fér, T., & Schmickl, R. E. (2018). HybPhyloMaker: target enrichment data analysis from raw reads to species trees. *Evol. Bioinform. Online,* 14, 1176934317742613.

Galbany-Casals, M., Garcia-Jacas, N., Sáez, L., Benedí, C., & Susanna, A. (2009). Phylogeny, biogeography and character evolution in Mediterranean, Asiatic and Macaronesian *Helichrysum* (Asteraceae, Gnaphalieae) inferred from nuclear phylogenetic analyses. *Int. J. Plant Sci.,* 170(3), 365–380.

Galbany-Casals, M., Unwin, M., Garcia-Jacas, N., Smissen, R.D., Susanna, A., & Bayer, R.J. (2014). Phylogenetic relationships in *Helichrysum* (Compositae: Gnaphalieae) and related genera: Incongruence between nuclear and plastid phylogenies, biogeographic and morphological patterns, and implications for generic delimitation. *Taxon*, 63, 608–624.

Gamisch, A., Fischer, G.A., & Comes, H.P. (2016). Frequent but asymmetric niche shifts in *Bulbophyllum* orchids support environmental and climatic instability in Madagascar over Quaternary time scales. *BMC Evol. Biol.,* 16, 14.

Gamisch, A., Winter, K., Fischer, G.A., & Comes, H.P. (2021). Evolution of crassulacean acid metabolism (CAM) as an escape from ecological niche conservatism in Malagasy *Bulbophyllum* (Orchidaceae). *New Phytol.,* 231, 1236–1248.

Gautier, L.G., Tahinarivony, J.A., Ranirison, P., & Hohlhauser, S. (2018). Vegetation. In: *The Terrestrial Protected Areas of Madagascar: Their History, Description, and Biota*, vol. 1., pp. 207–242. Goodman, S.M., Raharilalao, M.J., & Wohlhauser, S., (eds.). Assn. Vahatra: Antananarivo.

Graham, N., Gruner, D., Lim, J., & Gillespie, R. (2017). Island ecology and evolution: Challenges in the Anthropocene. *Environ. Conserv.*, 44(4), 323–335.

Gehrke, B., Kandziora, M., & Pirie, M.D. (2016). The evolution of dwarf shrubs in alpine environments: a case study of *Alchemilla* in Africa. *Ann. Bot.,* 117(1), 121–131.

Goodman, S.M., & Benstead, J.P. (eds.). (2003). *The Natural History of Madagascar.* University of Chicago Press: Chicago and London.

Hackel, J., Vorontsova, M.S., Nanjarisoa, O.P., et al. (2018). Grass diversification in Madagascar: In situ radiation of two large C_3 shade clades and support for a Miocene to Pliocene origin of C_4 grassy biomes. *J. Biogeogr.,* 45, 750–761.

Haber, E.A., Kainulainen, K., Van Ee, B.W., Oyserman, B.O., & Berry, P.E. (2017). Phylogenetic relationships of a major diversification of *Croton* (Euphorbiaceae) in the western Indian Ocean Region. *Bot. J. Linn.,* 183(4), 532–544.

Hershkovitz, M.A., & Zimmer, E.A. (2000). Ribosomal DNA evidence and disjunctions of western American Portulacaceae. *Mol. Phylogenet. Evol.,* 15(3), 419–439.

Hillis, D.M., & Bull, J.J. (1993). An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analyses. *Syst. Biol.,* 42, 182–192.

Holstein, N., & Renner, S.S. (2011). A dated phylogeny and collection records reveal repeated biome shifts in the African genus *Coccinia* (Cucurbitaceae). *BMC Evol. Biol.,* 11, 28.

Hörandl, E., & Emadzade, K. (2011). The evolution and biogeography of alpine species in *Ranunculus* (Ranunculaceae): A global comparison. *Taxon*, 60, 415–426.

Humbert, H. (1962). *Flore de Madagascar et des Comores (Plantes Vasculaires)*; 189e familia. Composées, part 2. Leroy, J.F., (ed.). Muséum National d'Histoire Naturelle: Paris.

IUCN/UNEP/WWF. (1987). *Madagascar, an environmental profile.* Jenkins M.D., (ed)., IUCN: Gland and Cambridge.

Janssen, T., Bystriakova, N., Rakotondrainibe, F., Coomes, D., Labat, J. N., & Schneider, H. (2008). Neoendemism in Malagasy scaly tree ferns results from recent, coincident diversification bursts. *Evolution*, 62(8), 1876–1889.

Joseph, G.S., Rakotoarivelo, A.R., & Seymour, C.L. (2021). How expansive were Malagasy Central Highland forests, ericoid, woodlands and grasslands? A multidisciplinary approach to a conservation conundrum. *Biol. Conserv.,* 261, 109282.

Joseph, G.S., & Seymour, C.L. (2020). Malagasy highlands: originally woodland and forest containing endemic grasses, not grazing-adapted grassland. *Proc. Biol. Sci.,* 287(1937), 20201956.

Kainulainen, K., Razafimandimbison, S.G., Wikström, N., & Bremer, B. (2017). Island hopping, long-distance dispersal and species radiation in the Western Indian Ocean: historical biogeography of the Coffeeae alliance (Rubiaceae). *J. Biogeogr.,* 44, 1966–1979.

Katoh, K., & Toh, H. (2008). Recent developments in the MAFFT multiple sequence alignment program. *Brief. Bioinform.,* 9(4), 286–298.

Knox, E.B., & Palmer, J.D. (1995). Chloroplast DNA variation and the recent radiation of the giant senecios (Asteraceae) on the tall mountains of eastern Africa. *Proc. Natl. Acad. Sci. U.S.A.,* 92(22), 10349–10353.

Kozlov, A.M., Darriba. D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAxML-NG: A fast, scalable, and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics,* 35(21), 4453–4455.

Larridon, I., Spalink, D., Jiménez-Mejías, P., et al. (2021). The evolutionary history of sedges (Cyperaceae) in Madagascar. *J. Biogeogr.,* 48, 917–932.

Landis, M.J., Matzke, N.J., Moore, B.R., & Huelsenbeck, J.P. (2013). Bayesian analysis of biogeography when the number of areas is large. *Syst. Biol.,* 62(6), 789–804.

Lehmann, C.E.R., Solofondranohatra, C.L., Morton, J.A., et al. (2022). The grassy ecosystems of Madagascar. In: *The New Natural History of Madagascar*, pp. 152–168. Goodman, S.M., Andrianarimisa, A., Armstrong, A.H., Cooke, A., De Wit, M., Ganzhorn, J.U., Gautier, L., Goodman, S.M., Jones, J.P.G., Jungers, W.L., Krause, D.W., Langrand, O., Lowry, P.P., Racey, P.A., Raselimanana, A.P., Safford, R.J., Sparks, J.S., Stiassny, M.L.J., Tortosa, P. & Vences, M., (eds.). Princeton University Press: Oxford-Princeton.

Lemoine, F., Domelevo Entfellner, J.B., Wilkinson, E., et al. (2018). Renewing Felsenstein's phylogenetic bootstrap in the era of big data. *Nature,* 556(7702), 452–456.

Leong Pock Tsy, J. M., Lumaret, R., Mayne, D., et al. (2009). Chloroplast DNA phylogeography suggests a West African centre of origin for the baobab, *Adansonia digitata* L. (Bombacoideae, Malvaceae). *Mol. Ecol.,* 18(8), 1707–1715.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754–1760.

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

Lowry, P.P., Phillipson, II, P.B., Andriamahefarivo, L., Schatz, G.E., Rajaonary, F., & Andriambololonera, S. (2018). Flora. In: *The terrestrial protected areas of Madagascar: Their history, description, and biota,* pp. 243–255. Goodman, S.M., Raherilalao, M.J., & Wohlhauser, S., (eds.). Assn. Vahatra: Antananarivo.

Lowry, P.P.I., Schatz, G.E., & Phillipson, P.B. (1997). The classification of natural and anthropogenic vegetation in Madagascar. In: *Natural change and human impact.* Goodman, S.M. & Patterson, B.D., (eds.). Smithsonian Institution Press: Washington, D.C.

Madagascar Catalogue. (2024). *Catalogue of the Vascular Plants of Madagascar.* Missouri Botanical Garden, St. Louis (USA) & Antananarivo (Madagascar). Available online: www.tropicos.org/Project/Madagascar [Accessed: 28 February 2024]

Mandel, J.R., Dikow, R.B., Funk, V.A., et al. (2014). A target enrichment method for gathering phylogenetic information from hundreds of loci: An example from the Compositae. *Appl. Plant. Sci.,* 2(2), apps.1300085.

Mandel, J.R., Dikow, R.B., Siniscalchi, C.M., Thapa, R., Watson, L.E., & Funk. V.A. (2019). A fully resolved backbone phylogeny reveals numerous dispersals and explosive diversifications throughout the history of Asteraceae. *Proc. Natl. Acad. Sci. U.S.A.,* 116(28), 14083–14088.

Masters, J.C., Génin. F., Zhang, Y., et al. (2021). Biogeographic mechanisms involved in the colonization of Madagascar by African vertebrates: Rifting, rafting and runways. *J. Biogeogr.,* 48, 492–510.

Matthews, T. J., & Triantis, K. (2021). Island biogeography. *Curr. Biol.,* 31(19), R1201–R1207.

Matzke, N. J. (2013). Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Front. Biogeogr.*, 5(4), 242–248.

Matzke, N. J. (2022). Statistical comparison of DEC and DEC+J is identical to comparison of two ClaSSE submodels and is therefore valid. *J. Biogeogr.,* 49, 1805–1824.

Moat, J., & Smith, P. (2007). *Atlas of the vegetation of Madagascar.* Royal Botanic Gardens: Kew Richmond.

Nathan, R., Bullock, J.M., Ronce, O., & Schurr, F.M. (2009). Seed Dispersal. In: *Encyclopedia of Life Sciences* (eLS), (Ed.).

Nie, Z.-L., Funk, V.A., Meng, Y., Deng, T., Sun, H., & Wen, J. (2016). Recent assembly of the global herbaceous flora: evidence from the paper daisies (Asteraceae: Gnaphalieae). *New Phytol.,* 209, 1795–1806.

Nylinder, S., Razafimandimbison, S.G., & Anderberg, A.A. (2016). From the Namib around the world: biogeography of the Inuleae–Plucheinae (Asteraceae). *J. Biogeogr.,* 43, 1705–1716.

Palazzesi, L., Hidalgo, O., Barreda, V.D., Forest, F., & Höhna, S. (2022). The rise of grasslands is linked to atmospheric CO_2 decline in the late Palaeogene. *Nat. Comm.,* 12, 293.

Pirie, M.D., Kandziora, M., Nürk, N.M., et al. (2019). Leaps and bounds: geographical and ecological distance constrained the colonisation of the Afrotemperate by *Erica. BMC Evol. Biol.,* 19, 222.

Padilla, D.P., González-Castro, A., & Nogales, M. (2012). Significance and extent of secondary seed dispersal by predatory birds on oceanic islands: the case of the Canary archipelago. *J. Ecol.*, 100, 416–427.

Rakotoarivelo, A.R., Goodman, S.M., Schoeman, M.C., & Willows-Munro, S. (2019). Phylogeography and population genetics of the endemic Malagasy bat, *Macronycteris commersoni* s.s. (Chroptera: Hipposideridae). *PeerJ,* 7, e5866.

Ramdhani, S., Barker, N.P., & Baijnath, H. (2009). Rampant nonmonophyly of species in *Kniphofia* Moench (Asphodelaceae) suggests a recent Afromontane radiation. *Taxon*, 58, 1141–1152.

Razafimandimbison, S.G., Kainulainen, K., Wikström, N., & Bremer, B. (2017). Historical biogeography and phylogeny of the pantropical Psychotrieae alliance (Rubiaceae), with particular emphasis on the Western Indian Ocean Region. *Am. J. Bot.,* 104, 1407–1423.

Razafimandimbison, S.G., Wikström, N., Khodabandeh, A., & Rydin, C. (2022). Phylogeny of the Madagascar-centred tribe Danaideae (Rubiaceae) as a precursor to taxonomic revision: insights into its generic and species limits, affinities and distribution. *Ann. Bot.,* 130(6), 849–867.

Ree, R.H., Moore, B. R., Webb, C.O. & Donoghue, M.J. (2005). A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution,* 45(11), 2299–2311.

Ree, R.H., & Sanmartín, I. (2018) Conceptual and statistical problems with the DEC+J model of founder-event speciation and its comparison with DEC via model selection. *J. Biogeogr.*, 45, 741–749.

Ree, R.H., & Smith, S.A. (2008). Maximum likelihood inference of geographic range evolution by dispersal, local extinction and cladogenesis. *Syst. Biol.,* 57(1), 4–14.

Ronquist, F. (1997). Dispersal-Vicariance Analysis: A New Approach to the Quantification of Historical Biogeography. *Syst. Biol.,* 46(1), 195–203.

Rundell, R.J., & Price, T.D. (2009). Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends Ecol. Evol.,* 24(7), 394–399.

Rundle, H.D. & Nosil, P. (2005), Ecological speciation. *Ecol. Lett.,* 8, 336–352.

Sayyari, E., & Mirarab, S. (2016). Fast Coalescent-Based computation of local branch support from quartet frequencies. *Mol. Biol. Evol.,* 33(7), 1654–1668.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), 1312–1313.

Schatz, G.E. (1996). Malagasy/Indo-australo-malesian phytogeographic connections. In: *Biogéographie de Madagascar*, pp. 73–84. Lourenço, W.R., (ed.). Editions ORSTOM: Paris.

Schmidt-Lebuhn, A.N., & Bovill, J. (2021), Phylogenomic data reveal four major clades of Australian Gnaphalieae (Asteraceae). *Taxon,* 70, 1020–1034.

Silander, J.A. Jr., Bond, W.J., & Ratsirarson, J. (2024). The grassy ecosystems of Madagascar in context: Ecology, evolution, and conservation. *Plants People Planet*, 6(1), 94–115.

Skema, C., Jourdain-Fievet, L., Dubuisson, J.Y., & Le Péchon, T. (2023). Out of Madagascar, repeatedly: The phylogenetics and biogeography of Dombeyoideae (Malvaceae s.l.). *Mol. Phylogenet. Evol.,* 182, 107687.

Smissen, R.D., Bayer, R.J., Bergh, N.G., et al. (2020). A revised subtribal classification of Gnaphalieae (Asteraceae). *Taxon*, 69, 778–806.

Smith, S.A., & Donoghue, M.J. (2010). Combining historical biogeography with niche modeling in the *Caprifolium* clade of *Lonicera* (Caprifoliaceae, Dipsacales). *Syst. Biol.,* 59(3), 322–341.

Smith, S.A., Brown, J.W., & Walker, J.F. (2018) So many genes, so little time: A practical approach to divergence-time estimation in the genomic era. *PLoS ONE,* 13(5), e0197433.

Solofondranohatra, C.L., Vorontsova, M.S., Hempson, G.P., et al. (2020). Fire and grazing determined grasslands of central Madagascar represent ancient assemblages. *Proc. Biol. Sci.,* 287(1927), 20200598.

Spriggs, E.L., Christin, P.A., & Edwards, E.J. (2014). C4 photosynthesis promoted species diversification during the Miocene grassland expansion. *PloS ONE,* 9(8), e105923.

Tamura, K., Battistuzzi, F.U., Billing-Ross, P., Murillo, O., Filipski, A., & Kumar, S. (2012). Estimating divergence times in large molecular phylogenies. *Proc. Natl. Acad. Sci. U.S.A.,* 109(47), 19333–19338.

Tamura, K., Stecher, G., & Kumar. S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.,* 38(7), 3022–3027.

Tamura, K., Tao, Q., & Kumar. S. (2018). Theoretical foundation of the RelTime methods for estimating divergence times from variable evolutionary rates. *Mol. Biol. Evol.,* 35(7), 1770–1782.

Tao, Q., Tamura, K., Mello, B., & Kumar, S. (2020). Reliable confidence intervals for RelTime estimates of evolutionary divergence times. *Mol. Biol. Evol.,* 37(1), 280–290.

Thomas, N., Bruhl, J.J., Ford, A., & Weston, P.H. (2014). Molecular dating of Winteraceae reveals a complex biogeographical history involving both ancient Gondwanan vicariance and long-distance dispersal. *J. Biogeogr.,* 41, 894–904.

Tiley, G.P., Crowl, A.A., Manos, P.S., et al. (2024a). Benefits and limits of phasing alleles for network inference of allopolyploid complexes. *Syst. Biol.,* 73(4), 666–682.

Tiley, G.P., Crowl, A.A., Almary, T.O.M., et al. (2024b). Genetic variation in *Loudetia* simplex supports the presence of ancient grasslands in Madagascar. *Plants People Planet,* 6(2), 315–329.

Ufimov, R., Gorospe, J.M., Fér, T., et al. (2022). Utilizing paralogues for phylogenetic reconstruction has the potential to increase species tree support and reduce gene tree discordance in target enrichment data. *Mol. Ecol. Resour.*, 22(8), 3018–3034.

Vences, M., Wollenberg, K.C., Vieites, D.R., & Lees, D.C. (2009). Madagascar as a model region of species diversification. *Trends Ecol. Evol.,* 24(8), 456–465.

Vorontsova, M.S., Besnard, G., Forest, F., et al. (2016). Madagascar's grasses and grasslands: anthropogenic or natural? *Proc. Biol. Sci.,* 283(1823), 20152262.

Warren, B.H., Strasberg, D., Bruggemann, J.H., Prys-Jonces, R.P., & Thébaud, C. (2010). Why does the biota of the Madagascar region have such a strong Asiatic flavour? *Cladistics,* 26, 526–538.

Wells, N.A. (2003). Some hypotheses on the Mesozoic and Cenozoic paleoenvironmental history of Madagascar. In: *The Natural History of Madagascar*. Goodman S.M., & Bestead, J.P., (eds.). The University of Chicago Press: Chicago.

Westerhold, W., Marwan, N., Drury, A.J., et al. (2020). An astronomically dated record of Earth's climate and its predictability over the last 66 million years. *Science*, 369, 1383–1387.

Whittaker, R.J., Triantis, K.A. & Ladle, R.J. (2008). A general dynamic theory of oceanic island biogeography. *J. Biogeogr.,* 35, 977–994.

Wiens, J.J., Ackerly, D.D., Allen, A.P., et al. (2010). Niche conservatism as an emerging principle in ecology and conservation biology. *Ecol. Lett.,* 13, 1310–1324.

Wiens, J.J., & Graham, C.H. (2005). Niche conservatism: integrating evolution, ecology and conservation biology. *Annu. Rev. Ecol. Evol. Syst.,* 36(1), 519–539.

Yoder, A.D., Campbell, C.R., Blanco, M.B., et al. (2016). Geogenetic patterns in mouse lemurs (genus *Microcebus*) reveal the ghosts of Madagascar's forests past. *Proc. Natl. Acad. Sci. U.S.A.,* 113(29), 8049–8056.

Yoder, A.D., & Nowak, M.D. (2006). Has vicariance or dispersal been the predominant biogeographic force in Madagascar? Only time will tell. *Annu. Rev. Ecol. Evol. Syst.,* 37(1), 405–431.

Zhang, C., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics,* 19(Suppl 6), 153.

Supplementary Materials

Supplementary materials associated with this article can accessed online.

1. Supplementary Figures

Supplementary Figure S1. Dated ML tree with 95% CI (calibration points indicated).

Dated ML tree of the genus *Helichrysum* generated with RelTime implemented in MEGA 11. Black numbers next to the nodes show mean ages (above) and grey numbers show 95% confidence intervals (below). Calibration points are indicated to the left of the nodes.

Supplementary Figure S2. Pairwise distances histogram.

Pairwise distance histogram generated with ParalogWizard. The distribution of pairwise distances between exonic contigs for each locus is represented: distances are given as divergence percentage values (6.96 – 18.98%). The position of the first and second peaks is indicated. The mean divergence and standard deviation (minus and plus sigma) for each peak are marked by dashed lines and values are provided in the legend.

Supplementary Figure S3. ML tree of *Helichrysum.*

ML tree (RAxML-NG) of the genus *Helichrysum* based on nuclear DNA loci generated with the target-enrichment probe set Compositae1061 (Mandel et al. 2014). Numbers next to the nodes correspond to the BS and TBE metric values (BS above, TBE below).

Supplementary Figure S4. ASTRAL tree of *Helichrysum.*

ASTRAL species tree of genus *Helichrysum* based on nuclear DNA loci generated with the target-enrichment probe set Compositae1061 (Mandel et al. 2014). Numbers next to the nodes indicate local posterior probabilities (LPP).

Supplementary Figure S5. Alternative ASTRAL tree of *Helichrysum* based on the top 25 most informative loci.

ASTRAL species tree of genus *Helichrysum* based on nuclear DNA loci generated with the target-enrichment probe set Compositae1061 (Mandel et al. 2014). The top 25 most informative loci were selected with SortaDate (Smith et al. 2018). Numbers next to the nodes indicate local posterior probabilities (LPP).

Supplementary Figure S6. Ancestral range reconstruction of *Helichrysum* with probabilities

Ancestral range reconstruction of the genus *Helichrysum* using the best-fitting model DEC+j on a time-calibrated phylogeny based on target-enrichment data (Compositae1061 probe set). Pie charts at nodes show the relative probability of the possible states (main areas in primary colors, combinations of areas in grey).

Supplementary Figure S7. Representative images of *Helichrysum* species from the clade MAD4

Helichrysum species from the clade MAD4: (a) and (b) *Helichrysum leucosphaerum,* (c) and (d) *Helichrysum betsiliense*, (e) *Helichrysum antandroi.* (Photos: Mercè Galbany-Casals). This Malagasy clade can be divided into two subgroups according to their morphology: *H. leucosphaerum*, a subshrub with many small and narrow white capitula resembles its southern African sister species *H. silvaticum* (not shown). *Helichrysum betsiliense* and *H. antandroi*, more closely related in the phylogeny, are perennial herbs with basal rosettes and corymbs of broad yellow capitula.

2. *Supplementary Tables*

Supplementary Table S1. Studied materials.

List of studied materials with herbaria codes and voucher information. BioSample assession numbers are provided (access through BioProject PRJNA1121119; https://www.ncbi.nlm.nih.gov/sra).

SupplementaryTable S2. Raw read statistics after filtering (HPM1).

Sequenced species and their corresponding number reads with HybPhyloMarker target extraction method after filtering (script HPM1).

Supplementary Table S3. Paralogy statistics from ParalogWizard.

Statistics obtained in the paralogy analyses using ParalogWizard. For each sample, the number of paralogs loci within the selected divergence threshold (7.6—18.9%) is provided.

Supplementary Table S4. Summary statistics of the selected loci.

Summary statistics of the selected loci (including paralogs) for subsequent phylogeny reconstruction, after filtering for missing data (script HPM5).

Supplementary Table S5. Ancestral range probabilities for each node.

Ancestral range probabilities of the DEC+j model for *Helichrysum*, corresponding to the pie charts in Figure 2 and Supplementary Figure S5.

Supplementary Table S6. Biogeographic Stochastic Mapping dispersal counts.

DEC+j Biogeographic Stochastic Mapping Counts. Mean number of dispersal events estimated in the history of *Helichrysum* with Biogeographic Stochastic Mapping. Counts of dispersal events were averaged across the 100 BSMs, standard deviations can be found below each table. Rows are the source of dispersals and columns the sink. Colours indicate the frequency of events, (red-green, highest-lowest).

Supplementary Table S7. Comparative summary between Malagasy *Helichrysum* clades and their closest African relatives.

Character retention and homogeneity assessment of Malagasy *Helichrysum* clades compared to African sister lineages. The table summarizes the morphological (M) and ecological (E) similarities between species constituting the six Malagasy clades identified in this study and their closest relatives, the degree of homogeneity within each clade and their fit to Humbert's taxonomic groups (1962).

Chapter 3

Climatic oscillations, dispersibility and adaptability behind worldwide mountain radiations of the HAP clade (Compositae)

CHAPTER 3: Climatic oscillations, dispersibility and adaptability behind worldwide mountain radiations of the HAP clade (Compositae)

Blanco-Gavaldà C, Onstein RE, Valente L, Janzen T, Andrés-Sánchez S, Bergh N, Calleja JA, Carnicero P, Castillo O, Cron GV, Leliaert F, Moreyra LD, Puig-Surroca G, Razafimandimbison SG, Susanna A, Roquet C & Galbany-Casals M.

The content of this chapter is currently under review.

Abstract

Mountain ecosystems are recognised as biodiversity hotspots. However, the origins of their remarkable diversity remain unresolved. We explore this question by focusing on the HAP clade (*Helichrysum-Anaphalis-Pseudognaphalium*), a megadiverse lineage within the Compositae family that spans tropical and temperate mountain and lowland systems worldwide. The existence of multiple high-elevation clades provides an opportunity to address hypotheses regarding the impact of trait innovation, climatic oscillations, dispersal and niche lability in the diversification of mountain lineages. To investigate the biogeographic history and diversification dynamics of the HAP clade, we built a time-calibrated phylogeny of 560 taxa (62% of species) based on 989 nuclear loci. We examined the frequency of inter-mountain dispersal and lowland-to-mountain transitions and vice versa, tested whether diversification rates were dependent on time, climate or species diversity, and assessed the impact of bract colour on diversification rates using state-dependent speciation-extinction models. Additionally, we reconstructed the evolutionary history of two functional traits (bract colour, life form) and ecological preferences (elevational range, habitat) and explored potential correlations between them. The HAP clade extensively speciated during the Pleistocene, when net diversification rates nearly quadrupled, coinciding with parallel mountain radiations on multiple continents. The clade followed a pattern of nested radiations, with southern African mountains serving as initial diversity source and other mountain systems acting primarily as sinks. High-elevation ecosystems also contributed to lowland biodiversity. Diversification rates in high-elevations are independent of bract colour, yet significant trait-environment associations were supported. Functional traits and ecological preferences evolved repeatedly, with a tendency toward montane open habitat ecologies and chamaephytic life forms. Our findings suggest that mountains do not function as isolated evolutionary units for the HAP clade due to its high dispersibility and ability to thrive in both high- and low-elevations. However, the clade's evolutionary lability enabled repeated trait acquisition, niche shifts and microhabitat specialisation. This, coupled with Pleistocene climatic instability, likely played a significant role in driving allopatric and ecological speciation at different geographical scales.

Keywords: *Helichrysum*, HAP clade, *Anaphalis, Pseudognaphalium, Achyrocline*, diversification, high-elevation, Pleistocene, radiations, specialisation, adaptability, long-distance dispersal

1. Introduction

Mountain ecosystems worldwide are recognised as biodiversity hotspots, hosting disproportionately high levels of plant endemics relative to their small size (Antonelli, 2015; Hughes & Atchinson, 2015; Merckx et al., 2015) and contributing significantly to the global angiosperm richness. These ecosystems are often recognised as natural laboratories, offering great spatial and climatic heterogeneity and housing pronounced environmental gradients over short distances (Körner et al., 2011). Understanding the primary drivers of mountain biodiversity and diversity turnover is crucial for unravelling its origins and developing data-driven strategies for ensuring its future maintenance. Current research suggests that exceptional mountain diversity results from the combination of abiotic and biotic factors, whose additive interactions create unique conditions that promote speciation and species persistence over time (Nürk et al., 2020).

The main abiotic drivers of diversification in mountain ecosystems are associated with geological and climatic variation at small scales, including factors such as elevation, slope, soil (Antonelli et al., 2018; Perrigo et al., 2020), temperature and precipitation (Muellner-Riehl et al., 2019). Deep-time historical processes, such as orogeny and climatic oscillations that transformed previously mature landscapes, are also crucial to explain current biodiversity. Some studies indicate that net diversification rates declined during the Pleistocene due to climatic instability, extinction and range contractions, particularly in lineages with low dispersal ability or high specialisation (Sandel et al., 2017). However, substantial evidence supports the idea that Pleistocene climate oscillations had positive effects on diversification rates in both tropical (e.g. *Hypericum* L., Nürk et al., 2015) and temperate (e.g. bamboos, Xia-Ying Ye et al., 2018; *Jurinea*, Cass., Herrando-Moraira et al., 2023) mountain radiations, creating new opportunities and barriers that led to repeated cycles of migration and isolation, acting as a “species pump” (Flantua & Hooghiemstra, 2018). Many plant radiations (understood as episodes of rapid diversification, resulting in high species richness, Hughes & Eastwood, 2006) coincided with Pleistocene climate variations, while only a few predated the Mid-Miocene Climate Optimum (around 15 to 18 Ma, Böhme, 2003), none of them in temperate areas (Muellner-Riehl et al., 2019). Similarly, rapid mountain uplift has been associated with increased diversification rates (e.g. Andean Páramos, Madriñán et al., 2013). An additional mechanism to explain high mountain diversity is long-distance dispersal, which seems to have played a key role in providing preadapted lineages from geographically distant regions (Edwards & Donoghue, 2013) as observed in various mountain systems (e.g. Malaysia, Merckx et al., 2015; Afroalpine region, Kandziora et al., 2022). Tropical mountains are renowned for this, for which strong phylogenetic evidence suggests that both tropical, and more predominantly, distant temperate regions have served as important sources of their diversity (Galley & Linder, 2006; Gehrke & Linder, 2009; Sklenář et al., 2011; Nürk et al., 2018; Blanco-Gavaldà et al., 2023, 2025).

Among biotic drivers of diversification, key innovations, — i.e. newly evolved traits which enable organisms to exploit previously inaccessible niches—, are the most relevant (Miller et al., 2023). Although few studies have quantified the impact of key innovations on diversification rates in mountain plants (e.g. Merianieae, Dellinger et al., 2024; *Juniperus* L., Liu et al., 2024), the acquisition of specific traits seems crucial to face the extreme environmental conditions of high-elevation habitats, such as high UV radiation and freezing temperatures or low pollinator availability. The most studied innovations are related to life forms and include the acquisition of perenniality (*Lupinus* L. in the Andes, Drummond et al., 2012), cushion growth (*Androsace* L. in Eurasia, Roquet et al., 2013) and woodiness (in Afroalpine *Alchemilla* L., Gehrke et al., 2016). In tropical mountains, gigantism (e.g. *Lobelia* Mill. and *Dendrosenecio* (Hauman ex Hedberg) B.Nord in eastern tropical Africa mountains, *Espeletia* Mutis ex Bonpl. and *Puya* Molina in the Andes) provides adaptive advantages against abiotic stressors such as freezing and water scarcity. These species are often associated with damp microenvironments like stream borders and gullies (Hedberg, 1964). In contrast, temperate mountains species tend to develop compact cushion-like forms that protect persistent buds near the ground, offering greater resistance to cold and wind exposure (Hedberg, 1964; Nürk et al., 2019; Körner, 2021). Furthermore, certain floral traits, such as large white flowers, may have evolved not only to attract generalist pollinators present at high elevations (Diptera and Coleoptera) but also to reflect UV radiation (Hedberg, 1964; Lefebvre et al., 2018; Baumann et al., 2021; Körner, 2021) and even to enable nyctinastic movements to protect the reproductive organs during the night (Minorsky, 2019). Nevertheless, trait flexibility and lineage adaptability might exert a stronger influence on diversification rates than any single innovation (Onstein, 2020; Helmstetter et al., 2023).

Some plant groups have diversified more extensively than others in mountain systems, though the reasons for this disparity remain poorly understood. To date, diversification dynamics have been studied for a few species-rich mountain lineages, such as *Lupinus* (Drummond et al., 2012; Contreras-Ortiz et al., 2018), *Jurinea* (Herrando-Moraira et al., 2023) and Ericaceae (Schwery et al., 2015; Pirie et al., 2019). Moreover, recent studies suggest that diversification rates and species richness in plants are independent, with greater species richness in humid tropical areas and higher diversification rates in heterogeneous environments with high species turnover (Kandziora et al., 2022; Tietje et al., 2022; Tenorio et al., 2023). This underscores the need for further research on the biogeographic history and diversification drivers of other highly diversified plant clades, particularly those spanning both temperate and tropical mountains. It would also be important to include key biodiversity hotspots, such as Madagascar and the southern African region, whose mountains remain understudied from an evolutionary perspective (Rudbeck et al., 2022).

In this context, we focus on the HAP clade (comprising *Helichrysum* Mill., *Anaphalis* DC., *Pseudognaphalium* Kirp., and several smaller genera; Smissen et al., 2011, Galbany-Casals et al., 2014; Nie et al., 2016), which comprises a total of c. 800 species (Hilliard,,1983; Anderberg, 1991), as a relevant study-case. This clade belongs to the Compositae, the most diverse plant family in high-elevation environments such as the Andes (Pérez-Escobar et al., 2022), the Afroalpine area (Gehrke & Linder, 2014), the Asian mountains (Yu et al., 2020) and the southern African Drakensberg (Carbutt & Edwards, 2004; Carbutt 2019). The HAP clade is particularly suited for studying mountain plant diversification due to its wide geographical distribution across tropical and temperate mountain systems and its richness in widespread and also narrowly endemic species. *Helichrysum* is the largest genus in the HAP clade, particularly diverse in southern Africa (c. 250 species, over half of which are in mountain habitats; Hilliard, 1983), the tropical Afromontane and Afroalpine regions (c. 72 species, Lisowski, 1989; Beentje, 2002; Tadesse, 2004) and Madagascar (111 species, half of them montane; Humbert, 1962), although it is also present in the Mediterranean Basin and Macaronesia. *Anaphalis* (110 species, Anderberg, 1991) and *Pseudognaphalium* (90 species, Anderberg, 1991), predominate in the mountains of Asia and the Americas, respectively. This distribution pattern suggests multiple independent mountain radiations, but this has yet to be confirmed because a solid and comprehensive phylogenetic framework is lacking for this megadiverse pan-montane group. The HAP clade shows remarkable adaptability, with species growing in a wide range of climatic conditions, thriving in diverse biomes, ranging from montane grasslands and alpine culminal rocky areas to coastal dunes and semideserts. While shrubby forms dominate, the clade displays a wide range of growth forms, from annual herbs to small trees, and there is a high diversity in reproductive strategies reflected in the varied number, size and disposition of capitula, the number of florets and the colour and size of the involucral bracts (Anderberg, 1991). We observed that high-elevation species tend to have white involucral bracts, trait that could have enhanced diversification in those environments. Additionally, we also noted that specific life forms seem to be associated with certain habitats across the clade's geographical range.

This study aims to advance our understanding of plant diversification in high-elevation environments by applying an integrative approach. Combining data from multiple sources we investigate the biogeographic and diversification dynamics of the HAP clade and the roles of trait evolution and Pleistocene climatic oscillations in its evolutionary history. Based on previous works for tropical African and Malagasy mountain lineages (Blanco-Gavaldà et al., 2023, 2025), we hypothesise that: (H1) diversification in the HAP clade fits a pattern of nested radiations in mountain systems worldwide, in which southern African mountains act as main source and other mountain regions serve as a sink, with occasional crossseeding events (Linder, 2014; Linder & Verboom, 2015); (H2) the current biodiversity distribution patterns of the HAP clade result from the combination of long-distance dispersals of preadapted ancestors from distant regions with similar biomes (in line with the niche conservatism principle, Wiens et al., 2010) alongside with in-situ diversification due to habitat specialisation in regions with high environmental heterogeneity (Rundell & Price, 2009; Edwards & Donoghue, 2013) and regional allopatric speciation; (H3) in high-elevation HAP clade lineages, Pleistocene climatic oscillations had a higher impact in diversification rates than the acquisition of key functional traits that are nowadays common in high-elevation species and (H4) parallel adaptation to particular microhabitats shaped trait-environment associations in mountain systems, since certain functional traits are repeatedly shared among species inhabiting specific habitats, suggesting convergent adaptation to similar environments. To test these hypotheses, we built the most comprehensive time-calibrated phylogeny of the HAP clade to date, using a family-specific target-enrichment probe set and building on a wide taxonomic sampling that includes 62% of the species and a balanced representation of the morphological and ecological variation of the study-group throughout the globe and elevational belts. We inferred the clade's ancestral biogeographic ranges to assess its area of origin, its propensity for long-distance dispersal and the contribution of elevational and latitudinal transitions to mountain biodiversity. We also assessed the evolution and ancestral states of key functional traits frequent in high-elevation species (i.e. having whitish bracts and being a cushion-like or gigantic shrub, depending on the latitude) and ecological preferences. Finally, we evaluated macroevolutionary dynamics by testing whether diversification rates varied with time, climate, species diversity, or the acquisition of specific traits and whether functional traits (bract colour and life form) are correlated with certain ecological variables (elevation and habitat preferences).

2. Materials and Methods

2.1 Taxon sampling and laboratory workflow

To reconstruct the evolutionary history of the HAP clade, we used sequences from 526 samples (including 500 species and 24 infraspecific taxa) which represent c. 62% of the clade's diversity (c. 800 species in total) and cover all major clades, genera and morphological variations (Supplementary Table S1), as well as all distribution areas (Supplementary Table S2), and ecological and elevational preferences (Supplementary Table S3). The largest genus, *Helichrysum* is represented by 400 samples (393 species, 73% of species). All the smaller genera known to be nested within *Helichrysum* are also included (Smissen et al., 2011; Galbany-Casals et al., 2014; Nie et al., 2016): *Achyrocline* Less. (15 species, 47%), *Anaphalis* (39 species, 37%), *Galeomma* Rauschert (one species, 50%), *Humeocline* Anderb. (one species, 100%), *Pseudognaphalium* (62 species, 69%), *Stenocline* DC. (two species, 100%), *Stenophalium* Anderb. (two species, 40%) and *Syncephalum* DC. (one species, 20%).

Based on previous phylogenetic studies (Nie et al., 2016; Smissen et al., 2020), we added 34 additional species from the Gnaphalieae to allow calibration for divergence time estimation, resulting in a dataset of 560 samples (Supplementary Table S1). A total of 24 samples corresponding to 121 species were newly sequenced. The other 436 samples, corresponding to 418 species, were taken from previous studies (329 are from Blanco-Gavaldà et al., 2023, BioProject PRJNA936872; 91 are from Blanco-Gavaldà et al., 2025, BioProject PRJNA1121119; six are from Mandel et al., 2019, BioProject PRJNAS40287; and ten are from Schmidt-Lebuhn & Bovill, 2021, BioProject PRJNA665592). We targeted the set of genomic loci included in the Compositae COS 1061 loci kit (Mandel et al., 2014), which has been specifically developed for the family. We followed the target-enrichment library preparation and sequence capture protocols described in Blanco-Gavaldà et al. (2023). The newly generated DNA sequence reads have been deposited in the NCBI Sequence Read Archive database (SRA; access: https://www.ncbi.nlm.nih.gov/sra) under the BioProject accession number PRJNA1219322 (see unique sample accession numbers in Supplementary Table S1).

2.2 Molecular data processing and phylogenetic analyses

We followed the workflow described in Blanco-Gavaldà et al. (2023) based on the use of HybPhyloMaker (Fér & Schmickl, 2018, available at https://github.com/tomas-fer/HybPhyloMaker, indicated hereafter as HPM, followed by the number of the corresponding script) in combination with ParalogWizard (Ufimov et al., 2022, available at https://github.com/rufimov/ParalogWizard), to extract the targeted 1061 COS loci and detect and separate paralogues before generating orthologous alignments.

Trimmomatic v.0.39 (Bolger et al., 2014) was used to remove adaptors and low-quality reads and BBMap v.38.42 (Bushnell, 2014) to remove duplicates (HPM1). BWA (Li & Durbin, 2009) and SPAdes (Bankevich et al., 2012) were used to create an initial reference file for read mapping based on sunflower genome sequences from the Compositae1061 probe set (Mandel et al., 2014). Then we used a customised reference based on our ingroup samples to increase mapping specificity. ParalogWizard was used to identify potential paralogues based on a pairwise exonic sequence divergence calculation, which generated a histogram with two peaks: the first representing putative allelic variation and the second highly divergent sequences corresponding to putative paralogues (Supplementary Figure S1). We used the range value of the peaks as the threshold to retrieve putative paralogous sequences. Finally, we used MAFFT v.7.475 (Katoh & Toh, 2008) to align sequences in orthologous matrices and concatenate exons into putative loci. To reduce missing data, we excluded sequences missing more than 70% of the data and removed loci present in less than 75% of all samples (HPM5).

Phylogenetic analyses were based on both concatenation and coalescence-based methods. For the first approach, we concatenated all nuclear loci to generate a supermatrix and estimated the best nucleotide substitution model for each locus using ModelTest-NG (Darriba et al., 2019). This was followed by maximum likelihood (ML) partitioned analyses in which we used RAxML-NG v.1.1.0 (Kozlov et al., 2019) to run 20 independent ML tree searches (modified HPM8f), branch support being assessed via bootstrap (inference set to automatically determine a sufficient number of bootstrap replicates). We annotated the best-scoring ML tree with Felsenstein's Bootstrap (BS, Felsenstein, 1985) and Transfer Bootstrap Expectation (TBE, Lemoine et al., 2018) values, considering branches with BS ≥ 70% and TBE ≥ 0.7 to be significantly statistically robust (Hillis & Bull, 1993; Lemoine et al., 2018). For the coalescent-based approach, we performed summary-coalescence inference with ASTRAL-III v.5.7.8 (Zhang et al., 2018) based on individual gene trees for each locus with RAxML v.8.2.12 (Stamatakis, 2014; HPM6a, HPM7 and HPM8a). Branch support was assessed via local posterior probabilities (LPP), considering well-supported branches those with LPP ≥ 0.95 (Sayyari & Mirarab, 2016). In all cases, trees were rooted on four species representing four genera of the subtribe Relhaniinae, which is known to be the sister clade to the subtribe Gnaphaliinae (Bayer et al., 2000; Bergh & Linder 2009; Montes-Moreno et al., 2010; Smissen et al., 2020): *Athrixia phylicoides* DC., *Leysera gnaphalodes* Thunb., *Phagnalon sordidum* (L.) Rchb. and *Oedera pungens* (L'Hér.) N.G.Bergh.

2.3 Divergence time estimation

We used RelTime (Tamura et al., 2012, 2018) implemented in the software MEGA X (Kumar et al., 2018) to estimate divergence times by relaxing the assumption of a strict molecular clock on our best-scoring ML tree obtained with the concatenation approach. Confidence intervals were computed using the method of Tao et al. (2020), also implemented in RelTime. This dating method performs well for large empirical genomic datasets and allows the use of calibration densities (Costa et al., 2022). The Gnaphalieae tribe lacks sufficient ancient fossils to serve as primary calibration points. Thus, we used five secondary calibration points (CP; indicated on the tree in Supplementary Figure S2) from previously inferred divergence time estimates (Nie et al., 2016) and a geological evidence-based calibration point, corresponding to the emergence of the archipelago of Madeira (Ramalho et al., 2015), as age constraints. We applied a normal density distribution to the secondary calibration points, providing mean and standard deviation values to reflect the 95% confidence interval values reported in the original work: the tribe crown node (mean age 25 Ma, std ± 2.55, CP1), the "crown radiation" node (mean age 20.7 Ma, std ±

2.55, CP2), the "HAP clade" crown node (mean age 15.39 Ma, std ± 1.95, CP3), the "FLAG clade" crown node (mean age 12.78 Ma, std ± 1.85, CP4) and the "*Anaphalis* + Mediterranean-Asian *Helichrysum"* crown node (mean age 7.04 Ma, std ± 0.85, CP5). The emergence of the archipelago of Madeira 5.6 Ma (CP6) was set as a maximum age constraint at the stem node of the lineage grouping the four Madeiran endemic *Helichrysum* species (Galbany-Casals et al., 2014).

2.4 Ancestral biogeographic range inference

Taking into account the general distribution patterns of the HAP clade, we defined six large areas and subdivided them into two elevational belts (lowlands vs. mountains) resulting in 12 regions (species coded in Supplementary Table S2): (A) Southern Africa lowlands, (B) Southern Africa mountains; (C) Tropical Africa lowlands, (D) Tropical Africa mountains; (E) Madagascar lowlands, (F) Madagascar mountains; (G) Eurasia lowlands, (H) Eurasia mountains; (I) North and Central America lowlands, (J) North and Central America mountains; (K) South America lowlands and (L) South America mountains. Recognising that elevational belts heavily depend on latitude and environmental factors such as precipitation, orientation and continentality, among others (Körner et al., 2011; Körner, 2021), we applied a broad and simplified classification to identify general patterns on a global scale. Our classification mainly follows Kandziora et al. (2016), which distinguishes tropical and temperate regions to account for the latitudinal effect (i.e. elevational belts are shifted towards lower elevations at higher latitudes). Thus, we categorised species into two elevational stages: (a) lowlands, below 800 m above sea level in temperate latitudes and below 1800 m in tropical latitudes, and excluding those primarily inhabiting mountain massifs and (b) mountains (including the montane to alpine belts, considered along the study high-elevation species), above 800 m in temperate latitudes and above 1800 m in tropical latitudes; except in Madagascar, where the boundary was adjusted to 1300 m, based on species distribution patterns specific to the micro-continent (Humbert, 1962, as in Blanco-Gavaldà et al., 2025). Species were only classified as high-elevation if they primarily inhabited mountains or massifs, rather than highlands, despite elevation (Körner et al., 2011).

To assign an area to each species, we compiled distribution and elevational data for the HAP clade, recording lower and upper elevation bounds based on reference Floras (Supplementary References) and herbarium records from GBIF and institutional collections (mainly from BC, BCN, BNRH, BR, CANB, CONC, E, FM, LP, MA, MADJ, MBK, MEXU, MO, NBG, NY, O, P, PRE, RSA, S, SALA, SI, US and W). For species that occasionally extend beyond their core distribution range, we only considered the area or elevational stage where most occurrences are concentrated, disregarding marginal presences.

We carried out ancestral biogeographic range inference analyses on the ML time-calibrated tree to determine the area of origin of the HAP clade (H1) and the origin of colonisation events into high-elevation regions (H2). We used the R package BioGeoBEARS (Matzke, 2013), setting the maximum number of areas for any node to three, based on the highest number of areas occupied by the most widespread extant taxon in our study. We compared the fit of six biogeographical models: Dispersal-Extinction-Cladogenesis (DEC, Ree et al., 2005; Ree & Smith, 2008), a likelihood implementation of the Dispersal-Vicariance model (DIVAlike, Ronquis,t 1997), and the BayArea model (BAYAREAlike, Landis et al., 2013), and the corresponding version of each model that accounts for founder-event speciation through the addition of a jump-dispersal parameter (+j). We choose the best-fitting model based on the Akaike Information Criterion (AIC, Burnham & Anderson, 1998).

We performed Biogeographic Stochastic Mapping (BSM, Dupin et al., 2017) using 100 BSM replicates under the best-fitting biogeographic model to assess the role of mountain systems as sources or sinks of biodiversity (H1) across its whole distribution and the frequency of lineage exchanges (H2) between similar biomes (lowland to lowland, mountain to mountain) or biome shifts (lowland-to-mountain and vice versa). BSM analyses provide a probabilistic estimate of the timing, frequency, directionality and types of biogeographical events (anagenetic dispersals, cladogenetic range expansions, founder events and extinctions) between areas. Based on the BSM outputs, we plotted temporal diversity across the biogeographic regions using the R package *ltst*R v.0.1.0 (Skeels, 2019), extracting the number of species present in each region at a time point.

2.5 Ancestral state reconstruction of functional and ecological traits

To investigate how functional traits and ecological preferences evolved within the HAP clade (H2), we estimated ancestral states for two functional traits (involucral bract colour and life form) and two ecological preferences (elevational range and habitat preference). We scored 518 ingroup species (see classification in Supplementary Table S3) using reference Floras (Supplementary References), herbarium specimens, protologues and our field observations. Bract colour was categorised as either: (a) whitish sometimes with pink-red tints or (b) yellow-brown (Figure 1). Elevational range preference followed the criteria outlined in the ancestral biogeographic range inference section, but here additionally splitting mountain species in preference for montane or alpine belts. The alpine belt includes species that predominantly grow above the treeline, generally >2000 m in temperate latitudes and >3200 m in tropical latitudes (Körner et al., 2011, Körner 2021). Life forms were divided into seven states (Figure 1): (a) therophytes; (b) hemicryptophytes; (c) chamaephytes; (d) cushion-like chamaephytes; (e) nanophanerophytes, with persistent buds between 25 and 150 cm above ground; (f) macrophanerophytes, with persistent buds above 150 cm; and (g) lianescent forms. Species exhibiting plasticity between two states were assigned to the most common or representative category (e.g. some species may behave as therophytes or chamaephytes depending on environmental conditions, others can be classified either as chamaephytes or nanophanerophytes due to their variability in size). Habitat preference was treated as five discrete states (Figure 2): (a) permanently humid habitats, such as stream borders, wetlands and deep gullies; (b) open habitats like shrublands, grasslands, and disturbed habitats, including large forest clearings; (c) rocky habitats such as crevices or scree, considering here only strictly rupicolous species; (d) forests, considering here only species growing in the shaded understory or very small clearings, partly shaded; and (e) deep sands, mainly coastal dunes but also inner deserts.

To infer the evolutionary transitions between different trait states along the phylogeny, we used Markov models (Mk) implemented in the R package *diversitree* (FitzJohn, 2012). For each trait, we fitted and compared three Mk models using Akaike Information Criterion (AIC) values: (i) an all-rates-different model (ARD); (ii) a symmetrical model in which the transition rates between any two states are equal (SYM) and (iii) an equal rates model, constraining all transitions to a single rate (ER).

2.6 Time, climate and diversity-dependent macroevolutionary dynamics of the HAP clade

To assess the general drivers of diversification of the HAP clade (H3) we studied its macroevolutionary dynamics fitting and comparing diversification models based on explicit hypotheses, following a hypothesis-driven approach to avoid identifiability issues (Louca & Pennell, 2020; Morlon et al., 2022). We tested whether speciation and extinction rates varied linearly or exponentially with time or environmental changes (global temperature variation; Zachos et al., 2008) using the R package *RPANDA* (Morlon et al., 2016). We additionally used the R package *DDD* (Etienne et al., 2012) to test whether speciation was dependent on species diversity, fitting models in which the carrying capacity (i.e. the number of species that can exist at the same time) was added as an additional parameter (K). Lastly, to test whether there was a significant shift in diversification dynamics during the history of the HAP clade, we fitted diversity-dependent models in which all parameters were allowed to shift at a certain time, also implemented in *DDD*. All models received an analytical correction corresponding to the sampling fraction (62% of the clade's diversity) to account for incomplete taxon sampling. We identified the best-fit model as the one with the lowest AICc value (c, for corrected AIC, which modifies the standard AIC with a correction for small sample sizes to avoid model overfitting).

Figure 1. Specimens illustrating the seven most common life forms within the HAP clade defined in the study along with clear examples of the two most frequent involucral bract colours: (a) *Pseudognaphalium jaliscence* (Greenm.) Anderb. and (b) *Helichrysum micropoides* DC., both therophytes inhabiting open lowland habitats in North America and South Africa, respectively; (c) *H. mutisiifolium* Less., a lianecent form found in Malagasy montane forests; (d) *H. milbraedii* Moeser, a hemicryptophyte from tropical Africa open montane habitats; (e) *H. sessilioides* Hilliard, a cushion-like chamaephyte growing on Drakensberg rocks, South Africa; (f) *H. formosissium* Sch.Bip. a macrophanerophyte from tropical Africa montane forests; (g) *H. danguyanum* Humbert, a chamaephyte in Malagasy rocky montane habitats, exemplifying white involucral bracts; (h) *H. italicum* subsp. *microphyllum* (Willd.) Nym., a nanophanerophyte from Mediterranean open montane habitats and (i) *H. albilanatum* Hilliard, a chamaephyte from southern African open montane grasslands, illustrating yellow involucral bracts. (Photos: Mercè Galbany-Casals, except for (b) Santiago Andrés-Sánchez and (e) Deon du Plessis).

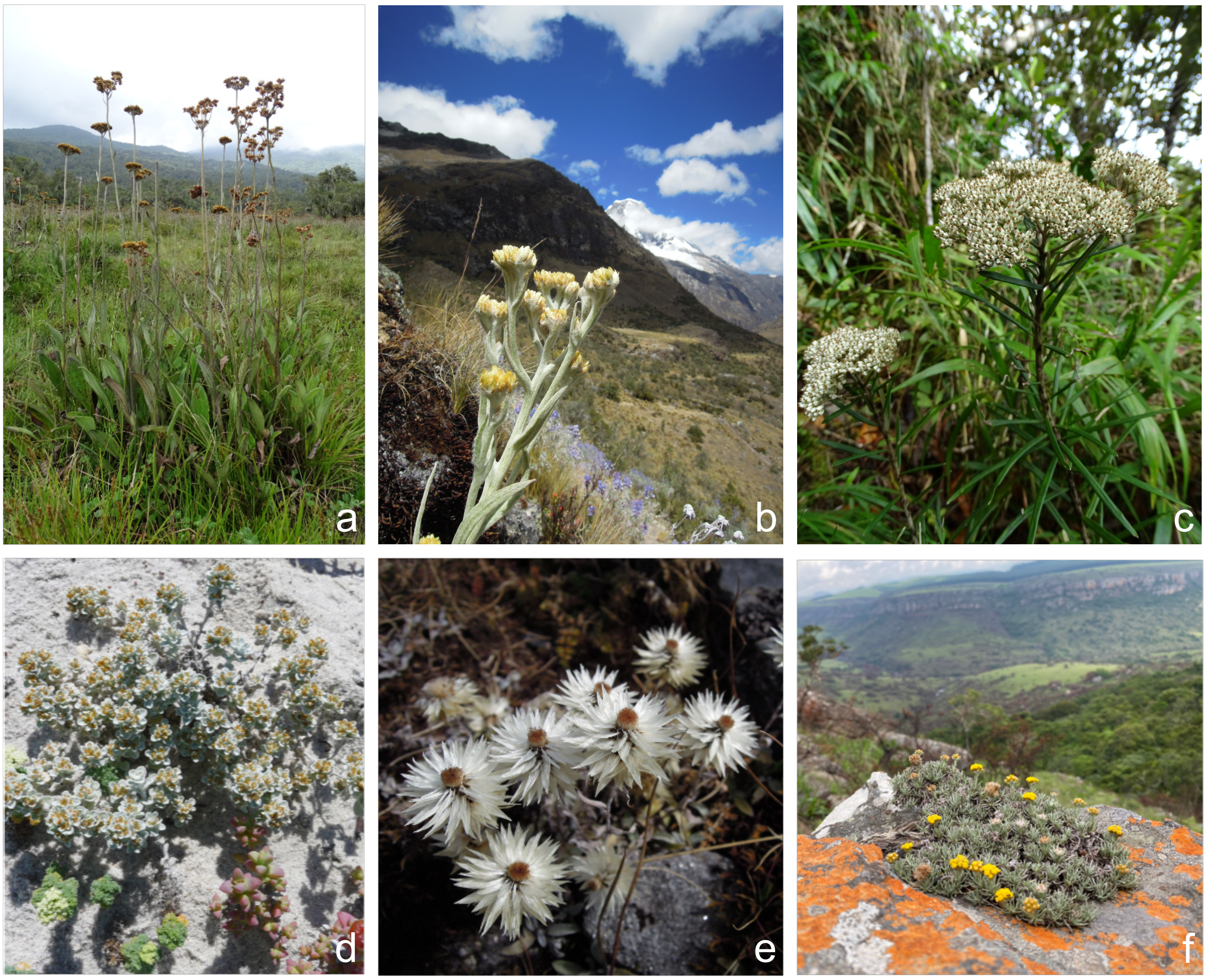


Figure 2. Representative species of the HAP clade are shown in their most common ecologies, illustrating the five categories of habitat preference defined in the study: (a) *Helichrysum globosum* Sch.Bip., a hemicryptophyte in humid montane habitats; (b) *Achryocline* spp. and (e) *Anaphalis nepalensis* (Spreng.) Hand.-Mazz., two chamaephytes growing in open high-elevation habitats, in South America and Asia respectively; (c) *H. dracaenifolium* Humbert, a macrophanerophyte from Malagasy montane forests and (d) *H. litorale* Bolus, a lowland chamaephyte in South Africa sandy environments (f) *H. galpinii* Schltr. & Moeser, a cushion-like chamaephyte characteristic of southern African rocky habitats. (Photos: Mercè Galbany-Casals, except for (f) Santiago Andrés-Sánchez and (f) Simmon Attwood).

2.7 Trait-dependent macroevolutionary dynamics of the HAP clade

We used the Several Examined and Concealed Trait-dependent Speciation and Extinction models implemented in the R package *SecSSE* v.3.1 (Herrera-Alsina et al., 2019), to determine whether the functional trait involucral bract colour and elevational preference impacted diversification rates (independently and combined) and to estimate net diversification rates associated with the character states (H3). We did not test for the impact of life form and habitat preferences to avoid overfitting, as a high number of character states would result in an excessive number of parameters to estimate, limiting model robustness. We fitted the following SecSSE models: (i) examined trait-dependent (ETD), (ii) concealed trait-dependent (CTD) and (iii) constant rate (CR), the latter equivalent to a null model where diversification parameters are not associated with any trait. ETD models test for heterogeneity in diversification parameters directly associated with specific observed ("examined") traits, while CTD models assess the impact of unknown ("concealed") traits that are not directly observed but may influence diversification rates. If a CTD model provides a better fit than an ETD model, it means that the observed trait may not be the primary factor influencing diversification. Instead, other factors could potentially explain diversification patterns better. For each combination of traits, we assessed model fit using AIC.

We tested nine SecSSE models on our independent binary trait (bract colour) that had different parameters (speciation, extinction and transitions between states) estimated. All the models had two concealed states and double transitions for concealed states were disallowed to simplify the models. We allowed transition rates of the concealed states to be different from the examined trait ones (argument diff.conceal=TRUE). Since the extinction rates for all models were estimated as zero, we built two simplified models, for which extinction rates (mu) were set to zero, and the estimated transition rates were set to be the same for the examined traits and concealed traits.

To test for a potential combined effect of bract colour and elevational preference on diversification, we additionally tested the fit of five SecSSE models. These models had four concealed states to match the model complexity of the four examined traits (Herrera-Alsina et al., 2019) that combined two bract colours with two elevational range preferences. In this case, we set the estimated transition rates to be the same for the examined traits and concealed traits (diff.conceal = FALSE), and again double transitions for concealed states were disallowed to simplify the models. Models with "full" in the name have all possible speciation (lambda, λ) and extinction (mu, μ) rates estimated, while those with "L1" and "M1" estimate just a single speciation or extinction rate respectively. The fitted models and estimated parameters are shown in Supplementary Table S4.

Our SecSSE models were fitted to a phylogeny comprising 62% of the total diversity in the HAP clade, with sampling fractions adjusted depending on the trait and state. Given that our sampling completeness was close to the lower threshold recommended for this type of analyses (Mynard et al., 2023), we conducted simulations to assess the potential impact of the sampling incompleteness on our results. We generated a total of 400 phylogenetic trees, simulating scenarios of both full (100%) and partial (62%) completeness. For the latter, 38% of the tips were randomly selected and pruned to generate trees with the same level of sampling as our empirical phylogeny. The trees were simulated based on parameter values estimated from the binary trait analyses in order to minimise the number of free parameters. After fitting the full ETD and CTD models to each of the complete and randomly sampled trees, diversification rates and likelihoods were calculated. The model with the best likelihood for each tree was selected, and the results from the complete and sampled trees were compared to assess the influence of sampling completeness on our conclusions.

2.8 Trait-environment parallel associations

We explored potential correlations between functional traits and ecological preferences to assess parallel evolution shaped by trait-environment associations in the HAP clade (H4). To do so, we used BayesTraits v.4.1.2 (Pagel & Meade, 2024; available from www.evolution.rdg.ac.uk). Using the discrete option and an ML approach, we compared the fit of independent and dependent models of trait-correlated evolution on our concatenated time-calibrated phylogeny. The independent model assumes that a pair of binary traits evolve without influencing each other's rate of evolution. Conversely, the dependent model assumes that the rate of change of one trait is influenced by the state of the other. In the first case, four rate parameters are estimated (alpha0, beta1, alpha1, and beta1, where alpha represents the rate of change from 0 to 1 and beta from 1 to 0, and the numbers after alpha and beta refer to the pair of analysed traits). In the second case, transition rates (q) between trait state combinations are estimated. To determine which model best explains our data, we performed likelihood-ratio tests and compared p-values from chi-squared tests for the following trait-environment pairs: (i) strictly alpine species with white bracts; (ii) cushion-like growth form with rocky habitats; (iii) cushion-like growth form with montane and alpine elevational belts (separately and combined); (iv) macrophanerophytes with forests habitats; (v) macrophanerophytes with humid habitats; (vi) lianescent plants with forests habitats; (vii) therophytes with lowlands; (viii) nanophanerophytes with lowlands; (ix) hemicryptophytes with montane elevational belt (but not alpine).

3. Results

3.1 Molecular data processing and phylogenetic analyses

Initial recovery of the 1061 COS targeted loci was near 76% (809 loci, before paralogue filtering). We considered loci with divergence values greater than 7.91% to be potential paralogues (Supplementary Figure S1). Using the *Helichrysum*-customised reference based on our dataset, we detected an average of 222 (± 59) paralogous loci per sample (Supplementary Table S5). After filtering for missing data and separating alignments containing paralogues, we obtained 224 new alignments from the paralogue set. Phylogenetic inference was then performed on a total of 989 loci. The average aligned length per locus was 283 bp, ranging from 37 to 732 bp. Each locus contained on average 88 (ranging from 4 to 377) parsimony informative sites and 129 (ranging from 5 to 453) variable sites. The mean proportion of missing data was 4%, varying from 0 to 66% (Supplementary Table S6). Concatenation of all loci resulted in a supermatrix consisting of 302173 bp across 560 taxa.

The phylogenetic trees reconstructed using the concatenation (thereafter ML tree, Supplementary Figure S3) and the summary-coalescent (thereafter ASTRAL tree, Supplementary Figure S4) approaches were largely congruent, with strong supports for most deeper nodes. However, topologies differed between the two methods at some intermediate and shallow nodes, particularly in the relationships involving the large polyploid genera (*Anaphalis, Achyrocline* and *Pseudognahalium*) and their relationship to *Helichrysum* species. Topology in these parts of the trees showed higher levels of uncertainty, often lacking statistical support. Lower overall clade support was obtained with the summary-coalescence approach: 51% of the nodes received high support (LPP ≥ 0.95), whereas the percentage of highly supported nodes in the ML tree was 72% and 88% according to BS and TBE metrics (≥70% and ≥0.7), respectively.

3.2 Divergence times and biogeographic history of the HAP clade

All divergence time estimations and 95% confidence interval values can be found in Supplementary Figure S2. The best-fitting model from the ancestral biogeographic range estimation analyses on the ML tree is DEC+j (see summary statistics in Supplementary Table S7; ancestral range reconstruction in Supplementary Figure S5; probabilities of ancestral areas for each node in Supplementary Table S8; and node numbers in Supplementary Figure S6). Biogeographical Stochastic Mapping (BSM) suggests that founder events speciation (represented by the parameter *j*) played a key role in shaping the current distribution of the HAP clade (Table 1).

Table 1. Summary of biogeographic stochastic mapping (BSM) events for the HAP clade based on the DEC+j model (see all events in Supplementary Table S9). The average number of biogeographic events estimated for 100 BSMs are shown along with standard deviations (sd). Range-switching dispersals and range contractions are omitted because they were not required in the best-fitting model (DEC+j). For further information on the modes, visit http://phylo.wikidot.com/biogeobears (last accessed January 2, 2025).

Mode	Type	Mean (sd)	% of biogeographic events
Within-area speciation	Sympatry	397.7 (4.2)	72.3
	Subset	21.2 (5.7)	5.9
Dispersal	Founder events	92.3 (3.7)	16.9
	Range expansions	31.1 (2)	6
Vicariance	Vicariance	5.8 (2)	1
Total		=500 (2)	100

Consistent with previous studies (Blanco-Gavaldà et al., 2025) our estimates place the origin of the HAP clade at 11.6 Ma (11.4-15.0 Ma 95% CI) with its most probable ancestral area being a combination of lowlands and mountains in southern Africa (areas AB, 70%). Ancestral area inference for nodes along the tree backbone indicates the importance of the southern African mountains (area B) as initial centre of diversification and a repeated source of lineages that dispersed to other regions, particularly during the Pliocene. We inferred a total of 52 dispersal events from area B (Supplementary Table S9 summarises all inferred dispersal event counts). Other relevant sources of dispersals were the southern African lowlands (A, 12 event counts), tropical African mountains (D, 12 event counts), the Malagasy mountains (F, 10 event counts) and the southern American mountains (L, 10 event counts). The top four dispersal sink areas are: the southern African lowlands (A, 28 event counts), tropical African mountains (D, 26 event counts; acting source and sink), the southern African mountains (B, 14 event counts; actings as source and sink), and the Malagasy lowlands (E, 12 event counts). We found a notable asymmetry between dispersals from lowland to mountain areas and vice versa. The general trend is dispersing downward (migrating from higher elevations towards lower elevations), for which the total event count is 62. This pattern is pronounced in southern Africa and Madagascar but also observed across the Americas and between southern and tropical Africa (Figure 3). In contrast, upward dispersals (from lowlands to mountain areas) were less common (24 events), yet still substantial, indicating permeability between elevational belts. In addition to elevational shifts, long-distance dispersals between geographically distant mountain regions were also frequent (35 events on average), especially common from southern Africa to tropical Africa, from tropical Africa to Madagascar, and between North and Central America and South America (bidirectional). Dispersal events between distant lowlands were extremely rare, with an inferred average of four events.

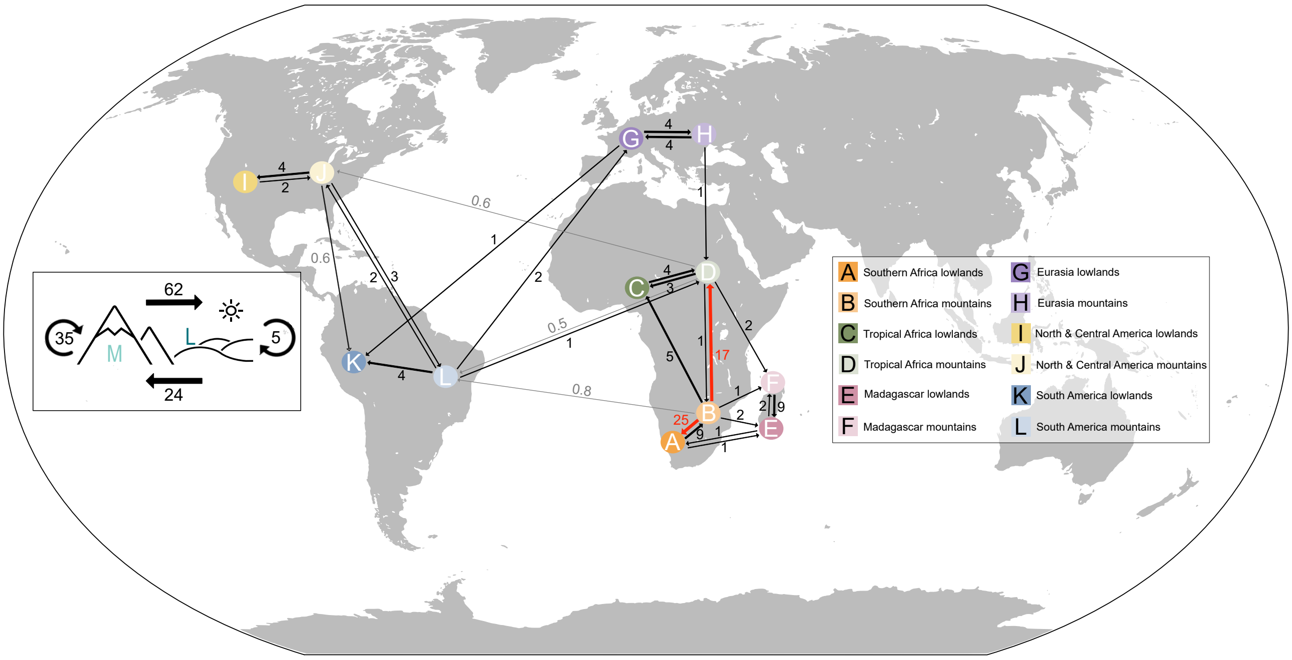


Figure 3. Average dispersal events between regions estimated from 100 biogeographic stochastic mappings (BSMs) using the DEC+j model for the HAP clade. Arrows indicate the direction of dispersal, with numbers representing the average count of events. Arrow thickness is proportional to dispersal frequency. For readability, arrows for average counts <1 are mainly shown when connecting African areas with other continents. The thickest arrows in red show the most frequent dispersal pathways (>12 events). The inset box to the left displays the number of events involving exchanges between similar elevational belts (lowland-to-lowland/mountain-to-mountain) and transitions between distinct elevational belts (lowland-to-mountain and vice versa). Summary tables for all event counts available in Supplementary Table S9.

During the Miocene, the southern African lowlands harboured a higher proportion of HAP lineages compared to the mountains (Figure 4). However, this trend shifted progressively, with high-elevation lineages dominating by the onset of the Pliocene (5.3 Ma). This transition coincided with the expansion of lineages into regions beyond southern Africa. By the end of the Pliocene (c. 2.6 Ma), the mountains of tropical Africa emerged as a significant centre of diversity. During the Pleistocene, southern Africa ceased to be the primary centre of diversification of the HAP clade, giving way to young high-elevation radiations mostly outside of the African continent at c. 1 Ma, particularly in Madagascar, Eurasia and the Americas.

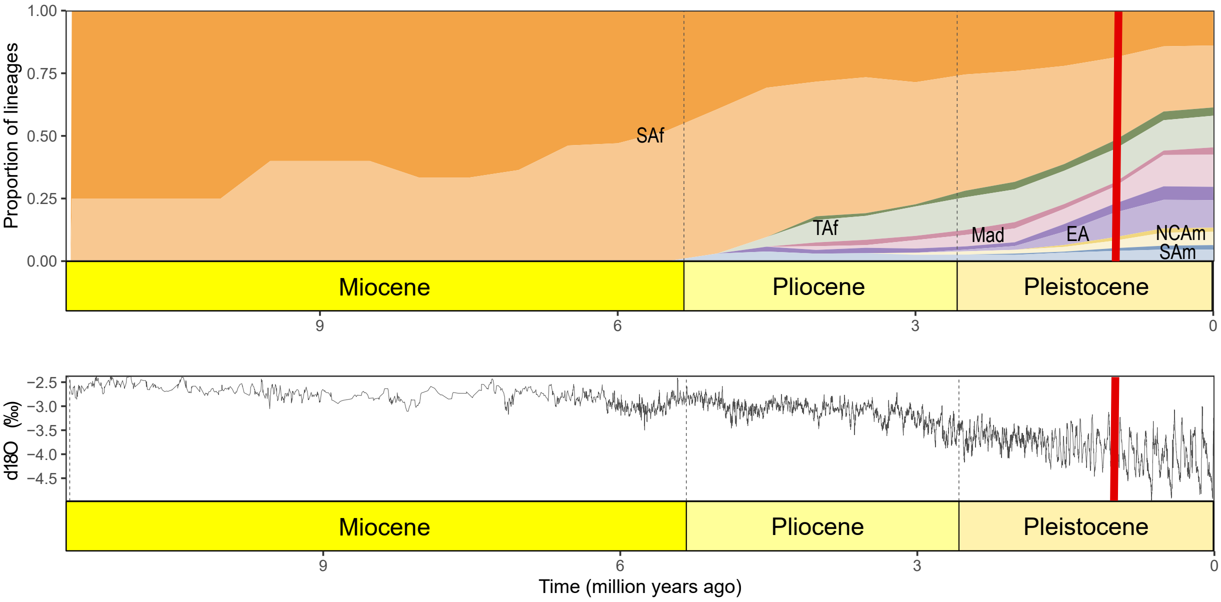


Figure 4. Relative proportion of lineages in different biogeographic regions over time for the HAP clade, based on average dispersal counts from the biogeographic stochastic mappings (BSMs) for the DEC+j model. The large shift in net diversification rates at approximately 1.2 Ma is indicated in red. Biogeographic regions: SAf, southern Africa (orange); TAf, tropical Africa (green); Mad, Madagascar (pink); EA, Eurasia (violet); NCAm, North and Centre America (yellow); SAm, South America (blue). For each geographical region, darker tones represent lower-elevation regions while lighter shades represent higher-elevation regions. The bottom panel shows the proxy curve for global temperature over the past 12 million years (from Zachos et al., 2001).

3.3 Ancestral state reconstruction of functional and ecological traits

Model selection according to AIC values indicated that the ARD model (allowing different transition rates between states) best fit the evolution of life form, elevational range and habitat preference, while involucral bract colour was better explained by equal rates (ER) or symmetrical (SYM) models (since both have the same transition rate for binary traits; see model summary tables in Supplementary Table S10). The ancestor of the HAP clade was reconstructed as a large shrub or small tree (macrophanerophyte) with yellow involucral bracts, occurring in sandy lowland habitats (Supplementary Figures S7 to S10). Over evolutionary time, alternative character states (white bracts, multiple growth forms and diverse ecological preferences) emerged independently multiple times, with some reversals to the ancestral states. Significant evolutionary transitions observed include: (i) bract colour, equal transition rates between yellow and white bracts (q_{01}=0.16, q_{10}=0.16); (ii) life form, the most frequent transition involved shifts to chamaephytes, particularly from therophytes (q_{13}=1.62), cushion-like forms (q_{43}=0.48) and nanophanerophyte (q_{53}=0.24). Other common transitions included shifts to nanophanerophyte (from lianescent forms, q_{75}=0.6 and macrophanerophytes, q_{65}=0.28) and to hemicryptophytes (from therophytes, q_{12} =0.43 and lianescent forms, q_{72}=0.26); (iii) elevational range preference, downward transitions from alpine to montane belts were the most common (q_{32}=0.2), with bidirectional exchanges between lowland and montane belts occurring equally (q_{12}= 0.11, q_{21}=0.11) and (iv) habitat preference, open habitats received the most transitions from all other habitats types, including forests (q_{42}=0.48), humid habitats, (q_{12}=0.45), rocky habitats (q_{32}=0.28) and sands (q_{52}=0.26).

3.4 Time, climate and diversity-dependent macroevolutionary dynamics of the HAP clade

In RPANDA, the preferred model was a time-dependent one, with speciation and extinction rates linearly correlated with time. In DDD, the favoured one was a diversity-dependent model, indicating a major shift in diversification dynamics during the Pleistocene (c. 1.16 Ma). At that time, the speciation rate quadrupled, shifting from $\lambda1$=0.8 to $\lambda2$=3.2 speciation events per lineage per million years, while the extinction rate remained negligible (μ=0). The estimated clade's carrying capacity also more than doubled, from K1 = 623 to K2 = 1370 species (see model inference summaries in Supplementary Table S11).

Although the RPANDA model had a slightly better fit based on AIC values, the signature of a pronounced shift in speciation rates in the Pleistocene seems evident in our dataset, explaining the position on the ranking of our second best-fitting model. Climate-dependent models had a worse fit overall, suggesting that climate alone was not the primary determinant of change in diversification dynamics.

3.5 Trait-dependent macroevolutionary dynamics of the HAP clade

Our SecSSE analyses aimed to determine whether bract colour (examined functional trait) or the interaction between bract colour and elevational preference influenced diversification within the HAP clade (Supplementary Tables S4). In both cases, models incorporating concealed traits (CTD) provided a better fit than examined-trait-dependent (ETD) ones, as indicated by their lower AIC values. This suggests that neither the colonisation of high-elevation habitats nor the acquisition of white bracts (examined traits) served as primary drivers of the observed diversification patterns. Furthermore, these results are robust to the level of taxon sampling in our phylogenetic tree: simulation analyses using both fully sampled and pruned phylogenies consistently supported the same models, indicating that incomplete sampling did not bias our findings.

3.6 Trait-environment parallel associations

Our analysis of 518 ingroup taxa, representing 62% of the HAP clade's diversity, revealed distinct patterns in character state distributions. Figure 5 illustrates the number of species in each character state (see full classification in Supplementary Table S3) while Figure 6 depicts their phylogenetic distribution across the HAP clade.

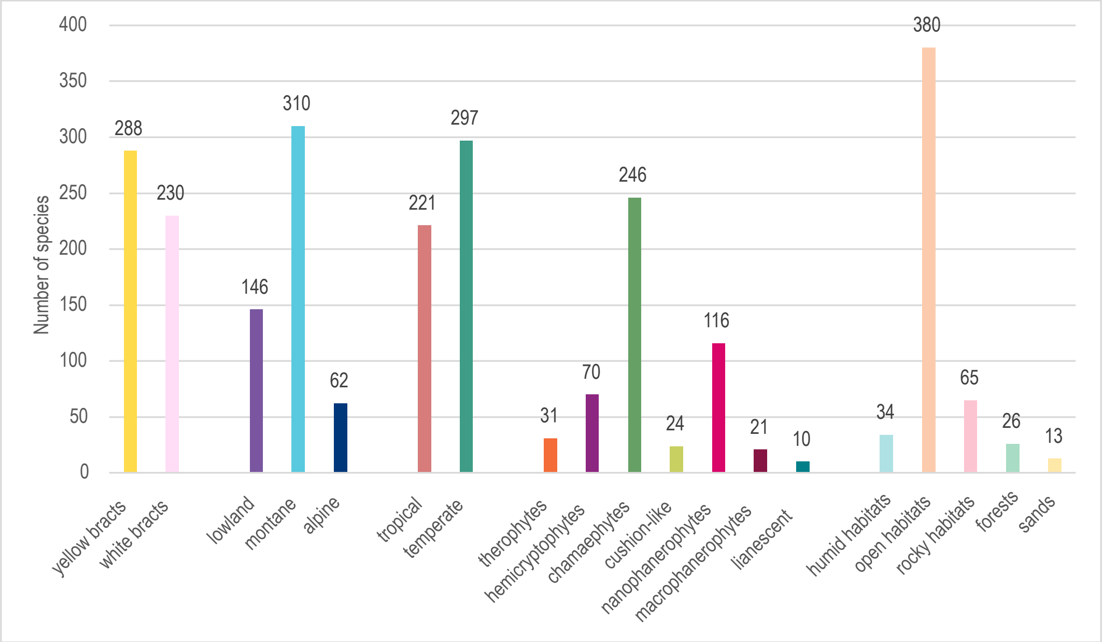


Figure 5. Bar chart presenting the number, out of a total of 518 HAP clade taxa of species within each of the following categories: involucral bract colour, elevational range preference, latitudinal range, life form and habitat preference. See functional trait and ecological preferences classification in Supplementary Table S3.

Our correlation tests performed using BayesTraits showed a significant association between most pairs of character states and habitat types analysed (p-value < 0.005, Table 2). These results suggest that the presence of one state is often linked with the existence of the other, probably reflecting parallel adaptation to particular habitats found in mountain systems (full model comparison in Supplementary Table S12).

Table 2. Results of correlation analyses for pairs of functional traits and ecological preferences (see classification in Supplementary Table S3) performed with BayesTraits, based on Likelihood Ratio test values. Positively correlated trait-environment are in black (*** for p < 0.001; ** for p < 0.01; * for p < 0.05); non-correlated trait combinations are in grey (full results in Supplementary Table S12).

Trait combination	p-value (<0.05)	correlated
Strictly alpine – white bracts	< 0.001	***
Rocky habitat – cushion-like growth	< 0.001	***
Strictly montane – cushion-like growth	< 0.001	***
Strictly alpine – cushion-like growth	< 0.001	***
High elevation (combination of montane and alpine) – cushion-like growth	0.08	No
Forest habitats – macrophanerophytes	< 0.001	***
Humid habitats – macrophanerophytes	< 0.001	***
Forest habitats – lianescent forms	< 0.001	***
Lowland – therophytes	< 0.001	***
Lowland – nanophanerophytes	< 0.001	***
Strictly montane – hemicryptophytes	< 0.001	***

4. Discussion

This study provides the most comprehensive and well-resolved phylogeny of the HAP clade to date, including representatives from across its global distribution. Our integrative approach combined data from multiple sources to test our four main hypotheses regarding the clade's diversification patterns. We confirmed a nested radiation pattern (H1), with the southern African mountains acting as the initial centre of diversification that fed other mountain systems. Our results support that long-distance dispersal repeatedly occurred among geographically distant mountain regions (H2). Multiple colonisations from high- to low-elevation habitats indicate significant elevational permeability. Pleistocene climatic oscillations likely played a crucial role in promoting regional allopatric speciation, as there was a notable rise in the net diversification rate during this period, which coincides with the emergence of young radiations within several mountain regions. In contrast, trait-dependent speciation and extinction analyses with SecSSE suggest that increased diversification rates in high-elevation environments are not associated with the acquisition of white bracts (H3). Instead, adaptability and niche lability seem to have contributed to local-scale diversification, as indicated by positive evolutionary correlations between functional traits and specific environmental preferences (H4), linked to microhabitat specialisation (H2). The wide geographical distribution and diversity of trait combinations paired with the varied ecological preferences of the HAP clade suggests a complex evolutionary history.

4.1. High-elevation environments as sources and sinks of plant diversity

Our results suggest that the HAP clade originated in southwestern Africa in the Mid-Miocene (c. 11 Ma), coinciding with the onset of global cooling and increased seasonality, which replaced warmer subtropical conditions (Herbert et al., 2016; Westerhold et al., 2020). Particularly, southwestern Africa became more arid and shifted to a winter rainfall regime due to the development of the westerly wind system and the strengthening of the Benguela current (Roberts et al., 2013; Neumann & Bamford, 2015). As temperatures cooled, the proportion of southern African high-elevation HAP clade species increased (Figure 4) alongside the expansion of savannas and grasslands (Linder, 2003; Dupont et al., 2013), the latter determined by high-elevation cold conditions rather than drought in southern Africa (see Bredenkamp et al., 2002). In addition to pronounced global cooling, this period was marked by intense mountain uplift in eastern Africa, affecting the Great Escarpment, the Drakensberg and the mountains along the Rift (Partridge, 1998; Partridge & Maud, 2000; Sepulchre et al., 2006), driving significant biotic turnovers (Herbert et al., 2016). High-elevation HAP species started to dominate around the Miocene-Pliocene boundary with dispersals to tropical Africa leading to the occupation and diversification in even higher mountains (Figure 4), eventually colonising mountain systems on other continents in the Pleistocene (Figure 3).

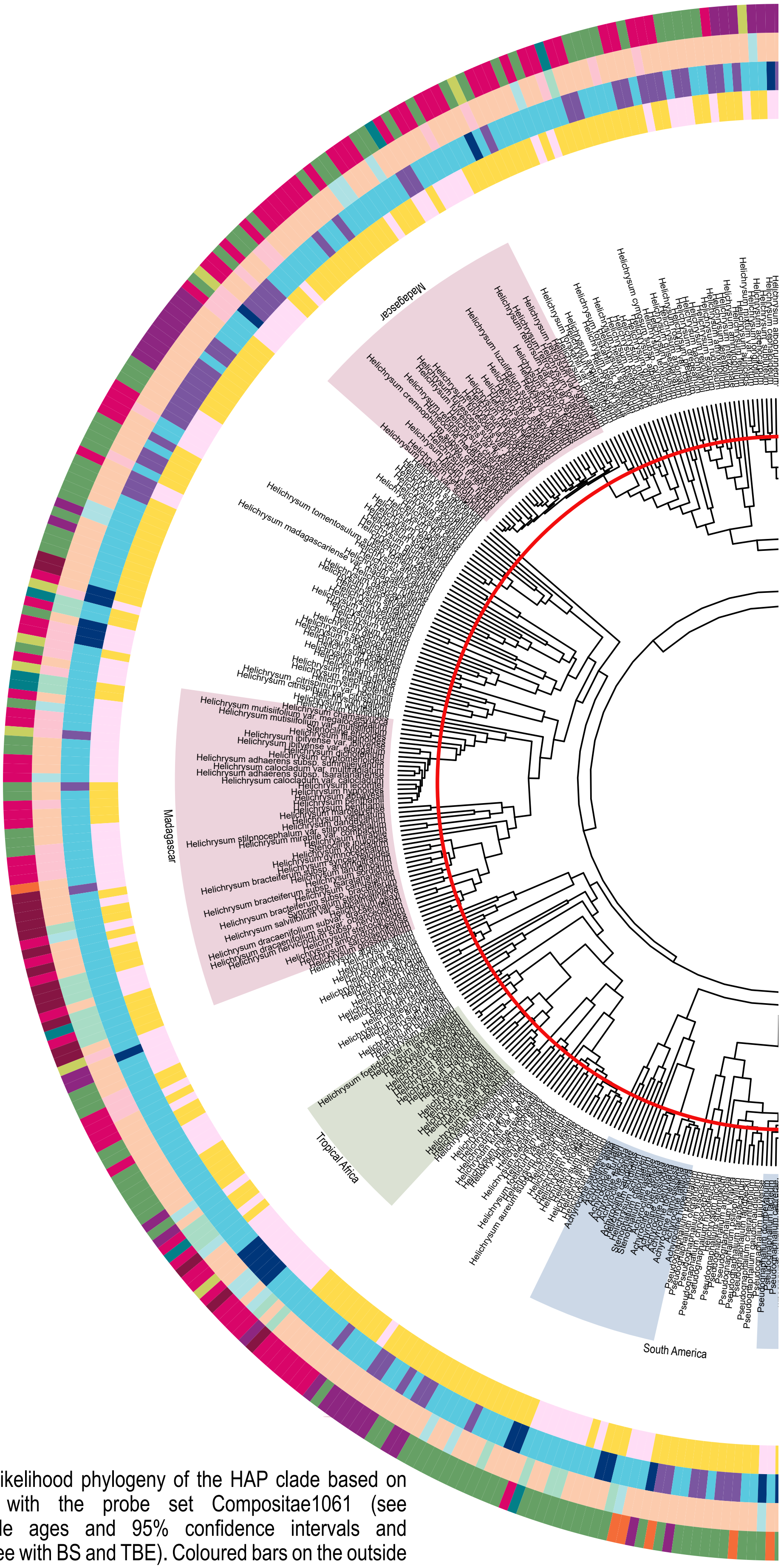


Figure 6. Time-calibrated maximum-likelihood phylogeny of the HAP clade based on target-enrichment data generated with the probe set Compositae1061 (see Supplementary Figure S2 for node ages and 95% confidence intervals and Supplementary Figure S3 for the full tree with BS and TBE). Coloured bars on the outside of the tree in four rows (inner to outer) show the current states of each species for: involucral bract colour (BC), elevational range preference (E), habitat preference (H) and life form (LF) (see full classification in Supplementary Table S3).

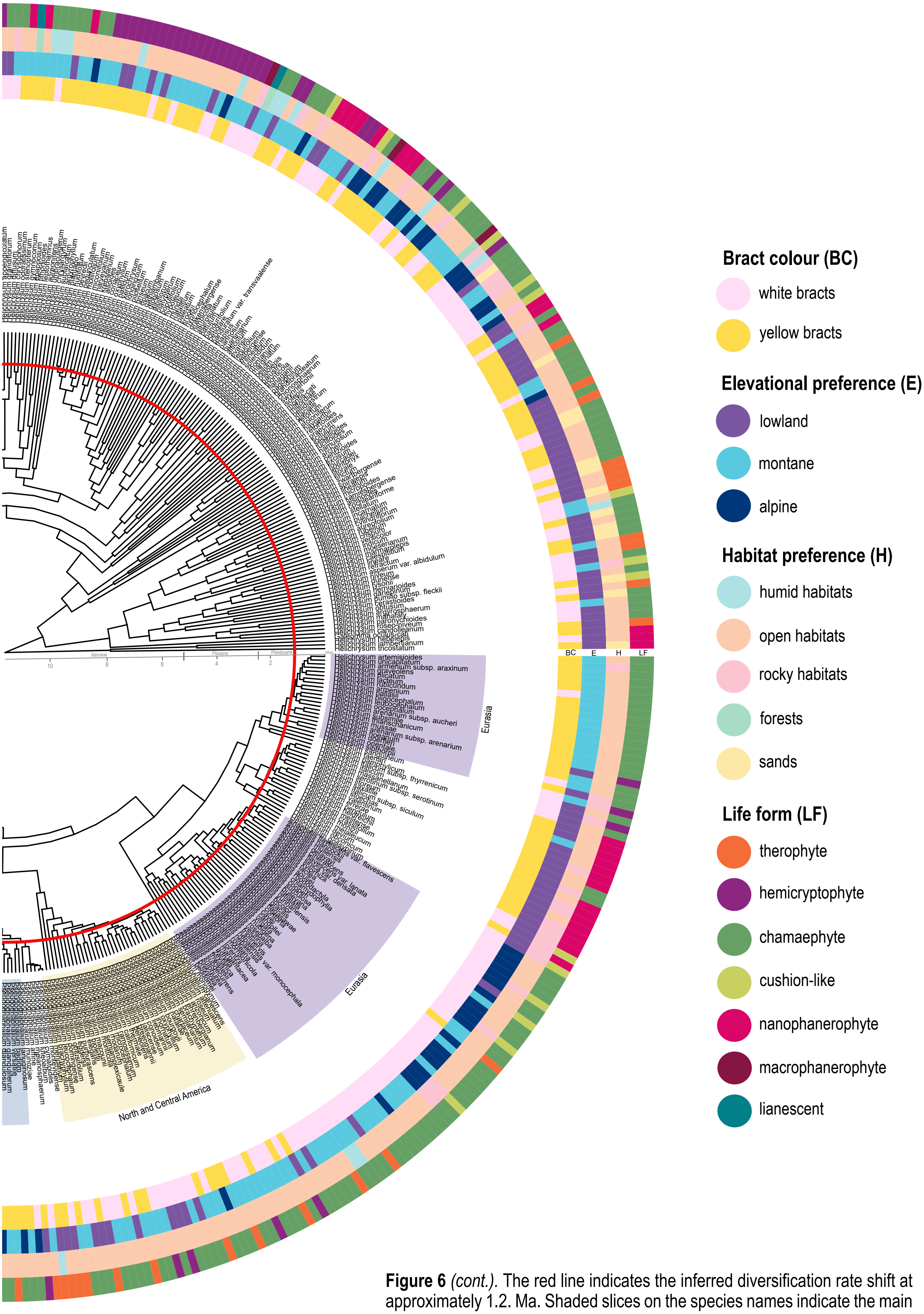


Figure 6 *(cont.).* The red line indicates the inferred diversification rate shift at approximately 1.2. Ma. Shaded slices on the species names indicate the main Pleistocene mountain radiations, colour-coded by biogeographic regions (geographic distribution range in Supplementary Table S2).

These findings support and extend the concept of nested evolutionary radiations described for southern African lineages (Linder, 2014; Linder & Verboom, 2015) demonstrating its relevance to mountain systems worldwide. Differences in clade ages, species richness and distribution patterns suggest that the southern African montane grasslands served as ancestral centre of diversification and source of preadapted high-elevation lineages (Donoghue, 2008). In contrast, mountains in the Americas, Eurasia and Madagascar acted primarily as sinks (Figure 3, Supplementary Table S9), harbouring younger and smaller radiations, with rapid accumulation of species occurring during the Pleistocene (Gytbes & McCain, 2007; Rabhek et al., 2019). Our biogeographic reconstruction reveals multiple instances of allopatric speciation driven by lineage dispersal between bioclimatically similar but distant mountain systems (Figure 3). This pattern, previously observed at smaller scales for Malagasy *Helichrysum* (Blanco-Gavaldà et al., 2025), highlights the role of niche conservatism in shaping global HAP clade distributions (Wiens & Graham, 2005; Wiens et al., 2010). Notably, high-elevation regions also served as a source of lineages for lower elevations, likely due to niche displacements during climatic fluctuations (Chala et al., 2017; Flantua et al., 2019), which eventually fostered niche shifts and subsequent specialisation (Donoghue & Edwards, 2014). This dynamic is reflected in a high permeability across elevational belts, with multiple recent colonisations of lowland regions throughout the clade's distribution, highlighting the ecological lability of HAP species and placing mountain systems worldwide as a source of lowland diversity. While upward shifts from low-to high-elevation habitats are more common in literature, for instance in *Espeletia* (Monasterio & Sarmiento, 1991), *Androsace* (Boucher et al., 2012), *Bulbophyllum* Thouars (Gamisch et al., 2016), *Saxifraga* L. (Carruthers et al., 2024), as well as in some butterflies (Chazot et al., 2016) and birds (Roy, 1997; Garcia-Moreno et al., 1998). Downward transitions also occur, such as those in *Dendrosenecio* (Knox & Palmer, 1995) and certain butterflies (Elias et al., 2009) and birds (Sedano & Burns, 2010; Van Els et al., 2021). These migrations across the elevational gradient were probably driven by preexisting adaptations, such as tolerance to high insolation and hydric stress (Donoghue, 2008; Edwards & Donoghue, 2013), considering that HAP clade species predominantly inhabit open habitats (Figures 5 and 6).

4.2. Pleistocene climate oscillations triggered parallel mountain radiations in the HAP clade

The evolutionary history of the HAP clade seems to be profoundly influenced by climate oscillations. The clade's expansion and radiation outside of Africa coincided with the onset of extreme glacial-interglacial cycles characteristic of the Mid-Pleistocene Transition c. 1.2 Ma (Clark et al., 2006). This timing aligns with our inference of a four-fold increase in the net diversification rate of the clade (Figure 4, Supplementary Table S11). As mentioned previously, these climatic fluctuations likely caused cycles of connectivity and isolation, driving major range expansions and contractions (Flantua et al., 2019). This dynamic facilitated the regional dispersal/migration of lineages preadapted to cool and dry conditions and the isolation or secondary contact of populations following elevational shifts of preferred climatic conditions. The formation of extensive continental ice sheets, mainly in the Northern Hemisphere and across the highest mountains worldwide, induced important environmental changes (Westerhold et al., 2020), forcing lineages to face extinction if they were unable to adapt in-situ or did not successfully colonise other regions through dispersal (Graae et al., 2018). The timing and environmental shifts associated with the Mid-Pleistocene Transition likely served as the primary driver of extensive diversification in Eurasian HAP lineages (Mediterranean, Asian and Macaronesian *Helichrysum* and Asian *Anaphalis*) as well as in the American lineages (*Achyrocline* and *Pseudognaphalium*). In Africa and Madagascar, glacial stages contributed to the expansion of cold ecosystems and increased connectivity between mountain ranges. This facilitated the descent and expansion of high-elevation lineages, both alpine and montane, with a stronger impact on the latter (Kebede et al., 2007; Chala et al., 2017). Similar processes have been documented in the Andes (Flantua et al., 2019) and the Guayanan tepuis (Rull, 2005). In African and Malagasy high-elevation lineages, these dynamics left a strong signature of allopatric speciation, exemplified by narrowly endemic species restricted to specific massifs, such as *H. brownei* S.Moore in Mt. Kenya, as well as widespread taxa spanning multiple mountains across tropical Africa, result of recent reconnections, such as *H. formosissimum* Sch.Bip. The species *H. arwae* J.R.I.Wood, a cushion-like chamaephyte in rocky habitats of Yemeni mountains, which is sister species to *H. horridium* Sch.Bip., a nanophanerophyte in open Ethiopian habitats, provide a clear example of allopatric speciation in mountain habitats (Figure 6). A similar pattern is observed in Madagascar, where several clades contain mountain-endemic species, such as *H. vaginatum* Humbert, found in several summits of the southeastern mountains, and its sister species, *H. marojejyense* Humbert, restricted to the summit of Mount Marojejy in the northeast (Humbert, 1962).

Recent studies suggest that speciation mechanisms are often interconnected, typically starting with an allopatric phase followed by ecological divergence, shaping macroevolutionary dynamics through both trait retention and trait changes (e.g. Aguilée et al., 2018; Gorel et al., 2022; Blanco-Gavaldà et al., 2025). Our findings align with the diversification patterns reported for *Veronica* sect. *Hebe* in New Zealand's mountains (Thomas et al., 2023). The lability of the HAP clade, which allowed for high permeability across heterogenous habitats, combined with the ability to persist and diversify in mountains favoured by climate oscillations, makes the clade fit the three stages of mountain diversity proposed in Thomas et al. (2023). Starting with long-distance dispersal from similar biomes, which provides preadapted colonisers (this would include both regional and local allopatric speciation), followed by in-situ speciation associated with niche shifts and ecological specialisation, rather than the acquisition of key innovations.

4.3 The history of a dynamic and highly adaptable group: heterogeneity within the HAP clade

Our trait-dependent speciation and extinction analyses with SecSSE revealed that models incorporating concealed traits provided a better fit than those where our examined states (bract colour and elevational range) affected diversification, indicating that unknown factors explain differences in net diversification rates across the HAP lineages (Supplementary Table S4). These results align with other studies showing that certain floral features might not directly drive rapid radiations (e.g. *Helianthemum*, Martin-Hernanz et al., 2023). While it is unclear whether other key innovations contributed to increased diversification rates in high-elevation species, the complex interplay between diverse trait combinations within the lineage and the heterogeneous environmental conditions across elevational and latitudinal gradients where the HAP clade occurs likely points to a more complex evolutionary pattern (see Figure 6).

Correlation analyses (Supplementary Table S12), although not indicating increased diversification rates, suggest associations between the evolution of functional traits (bract colour and life form) and ecological preferences (elevation and habitat preferences). Specifically, we found that the evolution of white bracts was associated with colonisation of higher elevations, where pollinator availability typically decreases (e.g. Blionis & Vokou, 2001). In such environments, larger and whiter flowers attract generalist pollinators, primarily Diptera (Lefebvre et al., 2018; Baumann et al., 2021; Körner, 2021). These larger flowers (capitula or aggregations of capitula, in the case of Compositae), not only enhance visibility but also retain warmth, providing pollinators with a warmer resting surface in addition to a food source (Dietrich & Körner 2014). Furthermore, shiny white bracts might have also been positively selected for their capacity to reflect harmful UV radiation (Hedberg, 1964; Körner, 2021).

Members of the HAP clade show a strong preference for open habitats (Supplementary Figure S9), which regardless of elevation, are devoid of tree cover and generally more exposed to climatic adversities. Plants in such environments often exhibit shorter statures (predominantly chamaephytes) an adaptation to water scarcity (Körner, 2021). Extreme forms include cushion-like chamaephytes, which dominate rocky high-elevation environments (Figure 6, Supplementary Tables S3 and S12). In contrast, larger arborescent forms tend to occupy more stable and humid environments, such as montane forests, streamside habitats or valleys and gullies (Körner, 2021, and in the HAP clade, see Figure 6, Supplementary Table S3 and S12). These life form–habitat associations detected in the HAP clade (Table 2, Figure 6), highlight how contrasting ecological niches within mountain systems promote diversification through ecological speciation, with steep elevational gradients creating a mosaic of microhabitats in close proximity, while remaining in the same broad-scale bioclimatic zones (Doebeli & Diechmann, 2003; Jones et al., 2011; Winkler et al., 2016; Antonelli et al., 2018; Garcia et al., 2020). Such local-scale adaptations and specialisation to microhabitats have likely contributed to diversification in mountainous regions, as demonstrated in European mountain systems (Smyčka et al., 2022). Microhabitat-driven speciation is illustrated by several pairs of sister taxa from our study, for example, the afroalpine sister species *H. newii* Oliv. & Hiern, a widespread nanophanerophyte common in high-elevation open habitats of eastern tropical Africa, coexisting with the partly sympatric *H. chionioides* Philipson, a macrophanerophyte restricted to humid habitats in Mount Aberdare and Mount Kenya, similar to the case of ecological speciation reported for two sympatric species of *Dendrosenecio* in Mount Kenya in Gorospe et al., (2025) and the case of parapatric speciation across the elevational gradient in Madagascar, where two morphotypes of *H. dracaenifolium* Humbert coexist in Mount Marojejy, one is a nanophanerophyte inhabiting open areas between 1550-1870 m and the other is a macrophanerophyte restricted to forests at 1150-1600 m.

Ancestral state reconstructions (Supplementary Figures S7 to S10) reflect a history of high adaptability and evolutionary dynamism in the HAP clade, with certain traits evolving independently multiple times across the phylogeny. The predominance of high-elevation perennial woody species (Figures 5 and 6) mirrors the well-documented evolutionary history of *Lupinus* (Fabaceae) in the New World (Hughes & Eastwood, 2006; Drummond, 2008). Woodiness has similarly been identified as a key innovation in other angiosperm mountain lineages, including *Alchemilla* in Africa (Gehrke et al., 2016), *Lachemilla* Rydb. in the Andes (Morales-Briones et al., 2018) and Delphinieae in the Himalayas (Jabbour & Renner, 2012). Nürk et al. (2019 and references therein) have shown the central role of disparity of growth forms in adaptive radiations on oceanic and tropical sky islands. The diversification of life forms observed in the HAP clade, particularly in tropical Africa and Madagascar (but not in Eurasia or the Americas, Figure 6), supports the idea that adopting diverse life forms is a key ecological strategy for thriving at high-elevations (Hedberg, 1964; Hedberg & Hedberg, 1979; Körner 2021). Interestingly, woody forms also dominate in lowland environments, since only a small proportion of HAP species are therophytes or hemicryptophytes (Figure 5). Ancestral state reconstructions suggest that the clade's ancestor was a macrophanerophyte (Supplementary Figure S10) adapted to the arid regions of southern Africa, consistent with links between the evolution of woodiness and paleoclimatic aridification (Hooft van Huysduynen et al., 2021; Zizka et al., 2022). Despite its origins in arid regions, this life form became dominant in tropical mountains and humid habitats.

Despite detecting a wide variation in traits linked to heterogeneous mountain habitats, trait evolution alone does not fully explain the HAP clade's rapid diversification during the Pleistocene. Other factors not considered in this study, such as ploidy (Combrick et al., 2025), may be important in explaining high-elevation radiations as observed in *Myosotis* Hill (Meudt et al., 2025). The HAP clade is also lineage shaped by a history of hybridisation and polyploidy, with known allopolyploid subclades predominantly in Eurasia and the Americas (Galbany-Casals et al., 2004; Galbany-Casals & Romo, 2008; Smissen et al., 2011; Galbany-Casals et al., 2014; Acosta-Maindo et al., 2018). The origin of these subclades may be closely tied to polyploid speciation, often associated with the colonisation of new territories (e.g. Ramsey, 2011; Carnicero et al., 2017) and genomic plasticity that supports resilience and adaptability (Barker et al., 2016; Van de Peer et al., 2017, 2021; Godfree et al., 2017). It is positively correlated with latitude (Rice et al., 2019) and has been shown to drive diversification in high-elevation and climatically unstable regions (Meundt et al., 2021), such as the Andes (Luebert & Weigend, 2014), the Pan-Himalayas (Wen et al., 2014) and European mountains (Pachschwöll et al., 2015; Slovák et al., 2023). Many polyploid species originated during the Pleistocene (Novikova et al., 2018; Han et al., 2022), likely driven by range contractions and reconnections that facilitated hybrid speciation (Folk et al., 2024). Polyploidy also contributes to morphological and ecological diversity (Baniaga et al., 2020), with habitat shifts influenced by microclimatic factors playing a crucial role in polyploid evolution (Marchant et al., 2016). Future research is needed to specifically investigate the impact of hybridisation and polyploidy on diversification patterns in the HAP clade, particularly regarding high-elevation radiations in Eurasia and the Americas.

5. Conclusions

This study of the evolutionary history of the HAP clade provides valuable insights into the assembly of high-elevation floras, highlighting the dual role of mountain systems as both sources and sinks of biodiversity. The diversification and distribution patterns of the HAP clade reflect the complex interplay of several factors, particularly related to climate and ecological adaptation, operating across various temporal and spatial scales. The clade's evolutionary success lies in its remarkable long-distance dispersal capacity and ability to colonise new regions as well as adaptability, being able to move across elevational gradients and thrive in different habitats, undergoing repeated radiations. While in-situ diversification is most evident in mountain systems, it also occurred at lower elevations. Our findings suggest that Pleistocene climatic fluctuations played a key role in recent parallel mountain radiations. These oscillations promoted range expansions and contractions, leading to cycles of isolation and reconnection between populations. Such dynamics likely fostered allopatric speciation at both regional and local scales, evident in the emergence of numerous regional endemics confined to specific massifs or regions, often constituting morphologically homogeneous clades. At a local scale, ecological speciation emerged as a significant driver of diversification. This is particularly evident in the parallel acquisition of life forms that enabled adaptation to specific microhabitats across different mountain massifs. The independent evolution of these traits within heterogeneous environments fostered niche differentiation, with African and Malagasy lineages showing especially diverse morphologies and ecological preferences. Although we did not consistently observe a clear association between functional traits and increased diversification rates at broader scales, our results indicate that niche lability was key in facilitating local-scale adaptations, potentially driving diversification within specific mountain systems. In contrast, the evolutionary trajectories of Eurasian and American radiations remain less clear but may involve polyploidisation and hybridisation, processes that enhanced genomic plasticity and resilience to climatic fluctuations. The HAP clade exemplifies how lineages can successfully persist and diversify in challenging high-elevation environments through flexible evolutionary strategies, combining dispersal, specialisation to microhabitats and genomic plasticity. These findings contribute to the understanding of diversification patterns in species-rich lineages inhabiting mountain systems worldwide.

References

Acosta-Maindo, A., Hinojosa-Espinosa, O., & Galbany-Casals, M. (2018). Polyploidy and new chromosome counts in *Pseudognaphalium* Kirp. (Compositae: Gnaphalieae). *Caryologia*, 71(4), 471–481.

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, 3013.

Anderberg, A.A. (1991). Taxonomy and phylogeny of the tribe Gnaphalieae (Asteraceae). *Opera Bot.*, 104, 1–195.

Antonelli, A. (2015). Biodiversity: Multiple origins of mountain life. *Nature*, 524(7565), 300–301.

Antonelli, A., Kissling, W.D., Flantua, S.G.A., et al. (2018). Geological and climatic influences on mountain biodiversity. *Nat. Geosci.*, 11, 718–725.

Baniaga, A.E., Marx, H.E., Arrigo, N., & Barker, M.S. (2020). Polyploid plants have faster rates of multivariate niche differentiation than their diploid relatives. *Ecol. Lett.*, 23, 68–78.

Bankevich, A., Nurk, S., Antipov, D., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.*, 19(5), 455–477.

Barker, M.S., Arrigo, N., Baniaga, A.E., Li, Z., & Levin, D.A. (2016). On the relative abundance of autopolyploids and allopolyploids. *New Phytol.,* 210, 391–398.

Baumann, K., Keune, J., Wolters, V., & Jauker, F. (2021). Distribution and pollination services of wild bees and hoverflies along an altitudinal gradient in mountain hay meadows. *Ecol. Evol.*, 11, 11345–11351.

Bayer, R.J., Puttock, C.F., & Kelchner, S.A. (2000). Phylogeny of South African Gnaphalieae (Asteraceae) based on two noncoding chloroplast sequences. *Am. J. Bot.*, 87(2), 259–272.

Beentje, H.J. (2002). *Helichrysum*. In: *Flora of Tropical East Africa*, part 2, pp. 403–452. Beentje, H.J., (ed.). Balkema: Rotterdam.

Bergh, N.G., & Linder, H.P. (2009). Cape diversification and repeated out-of-southern-Africa dispersal in paper daisies (Asteraceae-Gnaphalieae). *Mol. Phylogenet. Evol.*, 51(1), 5–18.

Blanco-Gavaldà, C., Galbany-Casals, M., Susanna, A., et al. (2023). Repeatedly Northwards and Upwards: Southern African Grasslands Fuel the Colonization of the African Sky Islands in *Helichrysum* (Compositae). *Plants*, 12(11), 2213.

Blanco-Gavaldà, C., Roquet, C., Puig-Surroca, G., et al. (2025). Biome conservatism prevailed in repeated long-distance colonization of Madagascar's mountains by *Helichrysum* (Compositae, Gnaphalieae). *Mol. Phylogenet. Evol.*, 204, 108283.

Blionis, G.J., & Vokou, D. (2001). Pollination ecology of *Campanula* species on Mt. Olympos, Greece. *Ecography*, 24, 287–297.

Böhme, M. (2003). The Miocene Climatic Optimum: evidence from ectothermic vertebrates of Central Europe. *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, 195(3-4), 389–401.

Bolger, A.M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120.

Boucher, F.C., Thuiller, W., Roquet, C., 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.

Bredenkamp, G., Spada, F., & Kazmierczak, E. (2002). On the origin of northern and southern hemisphere grasslands. *Plant Ecol.*, 163, 209–229.

Burnham, K.P., & Anderson, D.R. (1998). *Model Selection and Inference: A Practical Information-Theoretical Approach*. Springer-Verlag: New York.

Bushnell, B. (2014). BBMap: A Fast, Accurate, Splice-AWARE Aligner; Lawrence Berkeley National Lab (LBNL): Berkeley, CA, USA.

Carbutt, C. (2019). The Drakensberg Mountain Centre: A necessary revision of southern Africa's high-elevation centre of plant endemism. *S. Afr. J. Bot.*, 124, 508–529.

Carbutt, C., & Edwards, T.J. (2004). The flora of the Drakensberg Alpine Centre. *Edinb. J. Bot.*, 60, 581–607.

Carnicero, P., Sáez, L., Garcia-Jacas, N., & Galbany-Casals, M. (2017). Different speciation types meet in a Mediterranean genus: The biogeographic history of *Cymbalaria* (Plantaginaceae). *Taxon*, 66, 393–407.

Carruthers, T., Moerland, M.S., Ebersbach, J., et al. (2024). Repeated upslope biome shifts in *Saxifraga* during late-Cenozoic climate cooling. *Nat. Commun.*, 15(1), 1100.

Chala, D., Zimmermann, N.E., Brochmann, C., & Bakkestuen, V. (2017). Migration corridors for alpine plants among the 'sky islands' of eastern Africa: do they, or did they exist? *Alp. Bot.*, 127, 133–144.

Chazot, N., Willmott, K.R., Condamine, F., et al. (2016). Into the Andes: multiple independent colonizations drive montane diversity in the Neotropical clearwing butterflies Godyridina. *Mol. Ecol.*, 25(22), 5765–5784.

Clark, P.U., Archer, D., Pollard, D., et al. (2006). The middle Pleistocene transition: characteristics, mechanisms, and implications for long-term changes in atmospheric pCO2. *Quat. Sci. Rev.*, 25(23-24), 3150–3184.

Combrink, L.L., Golcher-Benavides, J., Lewanski, A.L., Rick, J.A., Rosenthal, W.C., & Wagner, C.E. (2025). Population Genomics of Adaptive Radiation. *Mol. Ecol.*, 34(2), e17574.

Contreras-Ortiz, N., Atchison, G.W., Hughes, C.E., & Madriñán, S. (2018). Convergent evolution of high elevation plant growth forms and geographically structured variation in Andean *Lupinus* (Fabaceae). *Bot. J. Linn. Soc.*, 187(1), 118–136.

Costa, F.P., Schrago, C.G., & Melle, B. (2022). Assessing the relative performance of fast molecular dating methods for phylogenomic data. *BMC Genomics*, 23, 798.

Crisp, M.D., Arroyo, M.T., Cook, L.G., et al. (2009). Phylogenetic biome conservatism on a global scale. *Nature*, 458(7239), 754–756.

Darriba, D., Posada, D., Kozlov, A.M., Stamatakis, A., Morel, B., & Flouri, T. (2020). ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. *Mol. Biol. Evol.*, 37(1), 291–294.

Dellinger, A.S., Lagomarsino, L., Michelangeli, F., Dullinger, S., & Smith S.D. (2024). The Sequential Direct and Indirect Effects of Mountain Uplift, Climatic Niche, and Floral Trait Evolution on Diversification Dynamics in an Andean Plant Clade. *Syst. Biol.*, 73(3), 594–612.

Dietrich, L., & Körner, C. (2014). Thermal imaging reveals massive heat accumulation in flowers across a broad spectrum of alpine taxa. *Alp. Bot.*, 124, 27–35.

Doebeli, M., & Dieckmann, U. (2003). Speciation along environmental gradients. *Nature*, 421(6920), 259–264.

Donoghue, M.J. (2008). Colloquium paper: a phylogenetic perspective on the distribution of Plant Divers. *Proc. Natl. Acad. Sci. U.S.A.*, 105 (Suppl 1), 11549–11555.

Donoghue, M.J., & Edwards, E.J. (2014). Biome Shifts and Niche Evolution in Plants. *Annu. Rev. Ecol. Evol. Syst.*, 45, 547–572.

Drummond, C.S. (2008). Diversification of *Lupinus* (Leguminosae) in the western New World: derived evolution of perennial life history and colonization of montane habitats. *Mol. Phylogenet. Evol.*, 48(2), 408–421.

Drummond, C.S., Eastwood, R.J., Miotto, S.T.S., & Hughes, C.E. (2012). Multiple Continental Radiations and Correlates of Diversification in *Lupinus* (Leguminosae): Testing for Key Innovation with Incomplete Taxon Sampling. *Syst. Biol.*, 61(3), 443–460.

Dupin, J., Matzke, N.J., Särkinen, T., et al. (2017). Bayesian estimation of the global biogeographical history of the Solanaceae. *J. Biogeogr.*, 44, 887–899.

Dupont, L.M., Rommerskirchen, F., Mollenhauer, G., & Schefuß, E. (2013). Miocene to Pliocene changes in South African hydrology and vegetation in relation to the expansion of C4 plants. Earth Planet. *Sci. Lett.*, 375, 408–417.

Edwards, E.J., & Donoghue, M,J. (2013). Is it easy to move and easy to evolve? Evolutionary accessibility and adaptation. *J. Exp. Bot.*, 64(13), 4047–4052.

Elias, M., Joron, M., Willmott, K., et al. (2009). Out of the Andes: patterns of diversification in clearwing butterflies. *Mol. Ecol.*, 18(8), 1716–1729.

Etienne, R.S., Haegeman, B., Stadler, T., et al. (2012). Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. *Proc. Biol. Sci.*, 279(1732), 1300–1309.

FitzJohn, R.G. (2012). Diversitree: comparative phylogenetic analyses of diversification in R. *Methods Ecol. Evol.*, 3, 1084–1092.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39(4), 783–791.

Fér, T., & Schmickl, R.E. (2018). HybPhyloMaker: Target Enrichment Data Analysis From Raw Reads to Species Trees. *Evol. Bioinform. Online*, 14, 1176934317742613.

Flantua, S.G.A., & Hooghiemstra, H. (2018). Historical connectivity and mountain biodiversity. In: *Mountains, climate and biodiversity*, 1st edition, pp. 171–185. Hoorn, C., Perrigo, A., Antonelli, A., (eds.). Wiley-Blackwell: Oxford.

Flantua, S.G.A., O'Dea, A., Onstein, R.E., Giraldo, C., & Hooghiemstra, H. (2019). The flickering connectivity system of the north Andean páramos. *J. Biogeogr.*, 46, 1808–1825.

Folk, R.A., Charboneau, J.L.M., Belitz, M., et al. (2024). Anatomy of a mega-radiation: Biogeography and niche evolution in Astragalus. *Am. J. Bot.*, 111(3), e16299.

Galbany-Casals, M., Garcia-Jacas, N., Susanna, A., Sáez, L., & Benedí, C. (2004). Phylogenetic relationships in the Mediterranean *Helichrysum* (Asteraceae, Gnaphalieae) based on nuclear rDNA ITS sequence data. *Aust. Syst. Bot.*, 17, 241–253.

Galbany-Casals, M., & Romo, A.M. (2008). Polyploidy and new chromosome counts in *Helichrysum* (Asteraceae, Gnaphalieae). *Bot. J. Linn. Soc.*, 158(3), 511–521.

Galbany-Casals, M., Unwin, M., Garcia-Jacas, N., Smissen, R.D., Susanna, A., & Bayer, R.J. (2014). Phylogenetic relationships in *Helichrysum* (Compositae: Gnaphalieae) and related genera: Incongruence between nuclear and plastid phylogenies, biogeographic and morphological patterns, and implications for generic delimitation. *Taxon*, 63, 608–624.

Galley, C., & Linder, H.P. (2006). Geographical affinities of the Cape flora, South Africa. *J. Biogeogr.*, 33, 236–250.

Gamisch, A., Fischer, G.A., & Comes, H.P. (2016). Frequent but asymmetric niche shifts in *Bulbophyllum* orchids support environmental and climatic instability in Madagascar over Quaternary time scales. *BMC Evol. Biol.*, 16, 14.

García, M.B., Domingo, D., Pizarro, M., Font, X., Gómez, D., & Ehrlén, J. (2020). Rocky habitats as microclimatic refuges for biodiversity. A close-up thermal approach. *Environ. Exp. Bot.*, 170, 103886.

Gehrke, B., Kandziora, M., & Pirie, M.D. (2016). The evolution of dwarf shrubs in alpine environments: a case study of *Alchemilla* in Africa. *Ann. Bot.*, 117(1), 121–131.

Gehrke, B., & Linder, H.P. (2009). The scramble for Africa: pan-temperate elements on the African high mountains. *Proc. Biol. Sci.*, 276(1667), 2657–2665.

Gehrke, B., & Linder, H.P. (2014). Species richness, endemism and species composition in the tropical Afroalpine flora. *Alp. Bot.*, 124, 165–177.

Godfree, R.C., Marshall, D.J., Young, A.G., Miller, C.H., & Mathews, S. (2017). Empirical evidence of fixed and homeostatic patterns of polyploid advantage in a keystone grass exposed to drought and heat stress. *R. Soc. Open Sci.*, 4(11), 170934.

Gorel, A.-P., Hardy, O.J., Dauby, G., et al. (2022). Climatic niche lability but growth form conservatism in the African woody flora. *Ecol. Lett.*, 25, 1164–1176.

Gorospe, J.M., Záveská, E., Chala, D., et al. (2025). Ecological speciation with gene flow followed initial large-scale geographic speciation in the enigmatic afroalpine giant senecios (*Dendrosenecio*). *New Phytol.* (Early View).

Graae, B.J., Vandvik, V., Armbruster, W.C., et al. (2018). Stay or go – how topographic complexity influences alpine plant population and community responses to climate change. *Perspect. Plant Ecol. Evol. Syst.*, 30, 41–50.

Grytnes, J.A., & Mccain, C.M. (2007). Elevational Trends in Biodiversity. *Encyclopedia of Biodiversity*, 2, 1–8.

Han, T.-S., Hu, Z.-Y., & Du, Z.-Q. (2022). Adaptive responses drive the success of polyploid yellowcresses (*Rorippa*, Brassicaceae) in the Hengduan Mountains, a temperate biodiversity hotspot. *Plant Divers.*, 44(5), 455–467.

Hedberg, O. (1964). Features of Afroalpine Plant Ecology. *Acta Phytogeogr. Suec.*, 49, 1–144.

Helmstetter, A.J., Zenil-Ferguson, R., Sauquet, H., et al. (2023). Trait-dependent diversification in angiosperms: Patterns, models and data. *Ecol. Lett.*, 26, 640–657.

Hedberg, I., & Hedberg, O. (1979). Tropical-alpine life forms of vascular plants. *Oikos*, 33, 297–307.

Herbert, T., Lawrence, K., Tzanova, A., Cleaveland Peterson, L., Caballero-Gill, R., & Kelly, C.S. (2016). Late Miocene global cooling and the rise of modern ecosystems. *Nat. Geosci.*, 9, 843–847.

Herrando-Moraira, S., Roquet, C., Calleja, J.A., et al. (2023). Impact of the climatic changes in the Pliocene-Pleistocene transition on Irano-Turanian species. The radiation of genus *Jurinea* (Compositae). *Mol. Phylogenet. Evol.*, 189, 107928.

Herrera-Alsina, L., van Els, P., & Etienne, R.S. (2019). Detecting the Dependence of Diversification on Multiple Traits from Phylogenetic Trees and Trait Data. *Syst. Biol.*, 68(2), 317–328.

Hillis, D.M., & Bull, J.J. (1993). An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analyses. *Syst. Biol.*, 42, 182–192.

Hilliard, O.M. (1983). *Helichrysum* Mill. In: *Flora of Southern Africa*, volume 33, part 7, fascicle 2, pp. 61–310. Leistner; O.A., (ed.). Department of Agriculture: Pretoria.

Hooft van Huysduynen, A., Janssens, S., Merckx, V., et al. (2021). Temporal and paleoclimatic context of the evolution of insular woodiness in the Canary Islands. *Ecol. Evol.*, 1, 12220–12231.

Hughes, C.E., & Atchison, G.W. (2015). The ubiquity of alpine plant radiations: from the Andes to the Hengduan Mountains. *New Phytol.*, 207(2), 275–282.

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. U.S.A.*, 103(27), 10334–10339.

Humbert, H. (1962). *Flore de Madagascar et des Comores (Plantes vasculaires)*; 189e famille. Composées, part 2. Leroy, J.F., (ed.). Muséum National d'Histoire Naturelle: Paris.

Jabbour, F., & Renner, S.S. (2012). A phylogeny of Delphinieae (Ranunculaceae) shows that *Aconitum* is nested within *Delphinium* and that Late Miocene transitions to long life cycles in the Himalayas and Southwest China coincide with bursts in diversification. *Mol. Phylogenet. Evol.*, 62, 928–942.

Jones, M.M., Szyska, B., & Kessler, M. (2011). Microhabitat partitioning promotes plant diversity in a tropical montane forest. *Glob. Ecol. Biogeogr.*, 20, 558–569.

Kandziora, M., Gehrke, B., Popp, M., Gizaw, A., Brochmann, C., & Pirie, M.D. (2022). The enigmatic tropical alpine flora on the African sky islands is young, disturbed, and unsaturated. *Proc. Natl. Acad. Sci. U.S.A.*, 119(22), e2112737119.

Kandziora, M., Kadereit, J.W., & Gehrke, B. (2016). Frequent colonization and little in situ speciation in *Senecio* in the tropical alpine-like islands of eastern Africa. *Am. J. Bot.*, 103(8), 1483–1498.

Kandziora, M., Vásquez, D.L.A., Brochmann, C., et al. (2024). Different resource partitioning explains plant species richness patterns in tropical alpine ecosystems. *bioRxiv,* 2024.06.24.599547.

Katoh, K., & Toh, H. (2008). Recent developments in the MAFFT multiple sequence alignment program. *Brief. Bioinform.*, 9(4), 286–298.

Kebede, M., Ehrich, D., Taberlet, P., Nemomissa, S., & Brochmann, C. (2007). Phylogeography and conservation genetics of a giant lobelia (*Lobelia giberroa*) in Ethiopian and Tropical East African mountains. *Mol. Ecol.*, 16, 1233–1243.

Knox, E.B., & Palmer, J.D. (1995). Chloroplast DNA variation and the recent radiation of the giant senecios (Asteraceae) on the tall mountains of eastern Africa. *Proc. Natl. Acad. Sci. U.S.A.*, 92(22), 10349–10353.

Körner, C. (2021). *Alpine plant life*: Functional plant ecology of high mountain ecosystems. Cham: Springer.

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.

Kozlov, A.M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAxML-NG: A fast, scalable, and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, 35(21), 4453–4455.

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.*, 35(6), 1547–1549.

Landis, M.J., Matzke, N.J., Moore, B.R., & Huelsenbeck, J.P. (2013). Bayesian analysis of biogeography when the number of areas is large. *Syst. Biol.*, 62(6), 789–804.

Lefebvre, V., Villemant, C., Fontaine, C., & Daugeron, C. (2018). Altitudinal, temporal and trophic partitioning of flower-visitors in Alpine communities. *Sci. Rep.*, 8, 4706.

Lemoine, F, Domelevo Entfellner, J.B., Wilkinson, E., et al. (2018). Renewing Felsenstein's phylogenetic bootstrap in the era of big data. *Nature*, 556(7702), 452–456.

Li, H., & Durbin, R. (2009). Fast and accurate short-read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754–1760.

Linder, H.P. (2003). The radiation of the Cape flora, southern Africa. *Biol. Rev.*, 78, 597–638.

Linder, H.P. (2014). The evolution of African Plant Diversity. *Front. Ecol. Evol.*, 2, 38.

Linder, H.P., & Verboom, G.A. (2015). The Evolution of Regional Species Richness: The History of the Southern African Flora. *Annu. Rev. Ecol. Evol. Syst.*, 46, 393–412.

Lisowski, S. (1989). *Helichrysum* Compositae (deuxième partie: Tribu Inuleae). In: *Flore d'Afrique Centrale (Zaire, Rwanda, Burundi)*, pp. 68–192. National Botanic Garden of Belgium: Meise.

Liu, K., Li, E., Cui, X., et al. (2024). Key innovations and niche variation promoted rapid diversification of the widespread *Juniperus* (Cupressaceae). *Commun. Biol.*, 7, 1002.

Louca, S., & Pennell, M.W. (2020). Extant time trees are consistent with a myriad of diversification histories. *Nature*, 580(7804), 502–505.

Luebert, F., & Weigend, M. (2014). Phylogenetic insights into Andean plant diversification. *Front. Ecol. Evol.*, 2, 27.

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

Mandel, J.R., Dikow, R.B., Funk, V.A., et al. (2014). A target enrichment method for gathering phylogenetic information from hundreds of loci: An example from the Compositae. *Appl. Plant Sci.*, 2(2), apps.1300085.

Mandel, J.R., Dikow, R.B., Siniscalchi, C.M., Thapa, R., Watson, L.E., & Funk, V.A. (2019). A fully resolved backbone phylogeny reveals numerous dispersals and explosive diversifications throughout the history of Asteraceae. *Proc. Natl. Acad. Sci. U.S.A.*, 116(28), 14083–14088.

Marchant, D.B., Soltis, D.E., & Soltis, P.S. (2016). Patterns of abiotic niche shifts in allopolyploids relative to their progenitors. *New Phytol.*, 212(3), 708–718.

Martín-Hernanz, S., Albaladejo, R.G., & Lavergne, S. (2023). Strong conservatism of floral morphology during the rapid diversification of the genus *Helianthemum*. *Am. J. Bot.*, 110(5), e16155.

Matzke, N.J. (2013). Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Front. Biogeogr.*, 5(4), 242–248.

Merckx, V.S., Hendriks, K.P., Beentjes, K.K., et al. (2015). Evolution of endemism on a young tropical mountain. *Nature*, 524(7565), 347–350.

Meudt, H.M., Albach, D.C., & Tanentzap, A.J. (2021). Polyploidy on Islands: Its Emergence and Importance for Diversification. *Front. Plant Sci.*, 12, 637214.

Meudt, H.M., Pearson, S.M., Ning, W., Prebble, J.M., & Tate, J.A. (2025). Forget-me-not phylogenomics: Improving the resolution and taxonomy of a rapid island and mountain radiation in Aotearoa New Zealand (*Myosotis*; Boraginaceae). *Mol. Phylogenet. Evol.*, 204, 108250.

Miller, A.H., Stroud, J.T., & Losos, J.B. (2023). The ecology and evolution of key innovations. *Trends Ecol. Evol.*, 38(2), 122–131.

Minorsky, P.V. (2019). The functions of foliar nyctinasty: a review and hypothesis. *Biol. Rev.*, 94, 216–229.

Monasterio, M., & Sarmiento, L. (1991). Adaptive radiation of *Espeletia* in the cold Andean tropics. *Trends Ecol. Evol.*, 6(12), 387–391.

Montes-Moreno, N., Sáez, L., Benedí, C., Susanna, A., & Garcia-Jacas, N. (2010). Generic delineation, phylogeny and subtribal affinities of *Phagnalon* and *Aliella* (Compositae, Gnaphalieae) based on nuclear and chloroplast sequences. *Taxon*, 59, 1654–1670.

Morales-Briones, D.F., Liston, A., & Tank, D.C. (2018). Phylogenomic analyses reveal a deep history of hybridization and polyploidy in the Neotropical genus *Lachemilla* (Rosaceae). *New Phytol.*, 218, 1668–1684.

Morlon, H., Lewitus, E., Condamine, F.L., Manceau, M., Clavel, J., & Drury, J.P. (2016). RPANDA: an R package for macroevolutionary analyses on phylogenetic trees. *Methods Ecol. Evol.*, 7(5), 589–597.

Morlon, H., Robin, S., & Hartig, F. (2022). Studying speciation and extinction dynamics from phylogenies: addressing identifiability issues. *Trends Ecol. Evol.*, 37(6), 497–506.

Muellner-Riehl, A.N., Schnitzler, J., Kissling, W.D., et al. (2019). Origins of global mountain plant biodiversity: Testing the 'mountain-geobiodiversity hypothesis'. *J. Biogeogr.*, 46, 2826–2838.

Mynard, P., Algar, A.C., Lancaster, L.T., et al. (2023). Impact of Phylogenetic Tree Completeness and Mis-specification of Sampling Fractions on Trait Dependent Diversification Models. *Syst. Biol.*, 72(1), 106–119.

Neumann, F.H., & Bamford, M.K. (2015). Shaping of modern southern African biomes: Neogene vegetation and climate changes. *Trans. R. Soc. S. Afr.*, 70(3), 195–212.

Novikova, P.Y., Hohmann, N., & Van de Peer, Y. (2018). Polyploid *Arabidopsis* species originated around recent glaciation maxima. *Curr. Opin. Plant Biol.*, 42, 8–15.

Nie, Z.-L., Funk, V.A., Meng, Y., Deng, T., Sun, H., & Wen, J. (2016). Recent assembly of the global herbaceous flora: evidence from the paper daisies (Asteraceae: Gnaphalieae). *New Phytol.*, 209, 1795–1806.

Nürk, N.M., Atchison, G.W., & Hughes, C.E. (2019). Island woodiness underpins accelerated disparification in plant radiations. *New Phytol.*, 224, 518–531.

Nürk, N.M., Linder, H.P., Onstein, R.E., et al. (2020). Diversification in evolutionary arenas—Assessment and synthesis. *Ecol. Evol.*, 10, 6163–6182.

Nürk, N.M., Michling, F., & Linder, H.P. (2018). Are the radiations of temperate lineages in tropical alpine ecosystems pre-adapted? *Glob. Ecol. Biogeogr.*, 27, 334–345.

Nürk, N.M., Uribe-Convers, S., Gehrke, B., Tank, D.C., & Blattner, F.R. (2015). Oligocene niche shift, Miocene diversification – cold tolerance and accelerated speciation rates in the St. John's Worts (*Hypericum*, Hypericaceae). *BMC Ecol. Evol.*, 15, 80.

Pachschwöll, C., Escobar García, P., 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(3), e0118197.

Pagel, M., & Meade, A. (2007). BayesTraits. Computer program and documentation, pp. 1216-1223.

Partridge, T.C. (1998). Of diamonds, dinosaurs and diastrophism: 150 years of landscape evolution in southern Africa. *S. Afr. Geogr. J.*, 101, 167–184.

Partridge, T.C., & Maud, R.R. (2000). Macro-scale geomorphic evolution of southern Africa. *Oxf. Monogr. Geol. Geophys.*, 40, 3–18.

Pérez-Escobar, O.A., Zizka, A., Bermúdez, M.A., et al. (2022). The Andes through time: evolution and distribution of Andean floras. *Trends Plant Sci.*, 27(4), 364–378.

Perrigo, A., Hoorn, C., & Antonelli, A. (2020). Why mountains matter for biodiversity. *J. Biogeogr.*, 47, 315–325.

Rahbek, C., Borregaard, M.K., Antonelli, A., et al. (2019). Building mountain biodiversity: Geological and evolutionary processes. *Science*, 365(6458), 1114–1119.

Ramalho, R.S., Brum da Silveira, A., Fonseca, P.E., et al. (2015). The emergence of volcanic oceanic islands on a slow-moving plate: The example of Madeira Island, NE Atlantic. *Geochem. Geophys. Geosyst.*, 16, 522–537.

Ramsey, J. (2011). Polyploidy and ecological adaptation in wild yarrow. *Proc. Natl. Acad. Sci. U.S.A.*, 108(17), 7096–7101.

Rice, A., Šmarda, P., Novosolov, M., et al. (2019). The global biogeography of polyploid plants. *Nat. Ecol. Evol.*, 3(2), 265–273.

Ree, R.H., Moore, B.R., Webb, C.O., & Donoghue, M.J. (2005). A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution*, 45(11), 2299–2311.

Ree, R.H., & Smith, S.A. (2008). Maximum likelihood inference of geographic range evolution by dispersal, local extinction and cladogenesis. *Syst. Biol.*, 57(1), 4–14.

Roberts, D., Cawthra, H., & Musekiwa, C. (2014). Dynamics of late Cenozoic aeolian deposition along the South African coast: a record of evolving climate and ecosystems. *Geol. Soc.*, 388(1), 353–387.

Ronquist, F. (1997). Dispersal-Vicariance Analysis: A New Approach to the Quantification of Historical Biogeography. *Syst. Biol.*, 46(1), 195–203.

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(10), 1874–1886.

Rudbeck, A.V., Sun, M., Tietje, M., et al. (2022). The Darwinian shortfall in plants: phylogenetic knowledge is driven by range size. *Ecography*, 2022(8), e06142.

Rull, V. (2005). Biotic diversification in the Guayana Highlands: a proposal. *J. Biogeogr.*, 32, 921–927.

Rundell, R.J., & Price, T.D. (2009). Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends Ecol. Evol.*, 24(7), 394–399.

Onstein, R.E. (2020). Darwin's second 'abominable mystery': trait flexibility as the innovation leading to angiosperm diversity. *New Phytol.*, 228, 1741–1747.

Sandel, B., Monnet, A.C., Govaerts, R., & Vorontsova, M. (2017). Late Quaternary climate stability and the origins and future of global grass endemism. *Ann. Bot.*, 119(2), 279–288.

Sayyari, E., & Mirarab, S. (2016). Fast Coalescent-Based computation of local branch support from quartet frequencies. *Mol. Biol. Evol.*, 33(7), 1654–1668.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), 1312–1313.

Schmidt-Lebuhn, A.N., & Bovill, J. (2021). Phylogenomic data reveal four major clades of Australian Gnaphalieae (Asteraceae). *Taxon*, 70, 1020–1034.

Schwery, O., Onstein, R.E., Bouchenak-Khelladi, Y., Xing, Y., Carter, R.J., & Linder, H.P. (2015). As old as the mountains: the radiations of the Ericaceae. *New Phytol.*, 207, 355–367.

Sedano, R.E., & Burns, K.J. (2010). Are the Northern Andes a species pump for Neotropical birds? Phylogenetics and biogeography of a clade of Neotropical tanagers (Aves: Thraupini). *J. Biogeogr.*, 37, 325–343.

Sepulchre, P., Ramstein, G., Fluteau, F., Schuster, M., Tercelin, J.J., & Brunet, M. (2006). Tectonic uplift and Eastern Africa aridification. *Science*, 313, 1419–1423.

Skeels, A. (2019). Lineages through space and time plots: visualising spatial and temporal changes in diversity. *Front. Biogeogr.*, 11, e42954.

Slovák, M., Melichárková, A., Štubňová, E.G., et al. (2023). Pervasive Introgression During Rapid Diversification of the European Mountain Genus *Soldanella* (L.) (Primulaceae). *Syst. Biol.*, 72(3), 491–504.

Smissen, R.D., Galbany-Casals, M., & Breitwieser, I. (2011). Ancient allopolyploidy in the everlasting daisies (Asteraceae: Gnaphalieae): Complex relationships among extant clades. *Taxon*, 60, 649–662.

Smissen, R.D., Bayer, R.J., Bergh, N.G., et al. (2020). A revised subtribal classification of Gnaphalieae (Asteraceae). *Taxon*, 69, 778–806.

Sklenář, P., Dušková, E., & Balslev, H. (2011). Tropical and Temperate: Evolutionary History of Páramo Flora. *Bot. Rev.*, 77, 71–108.

Smyčka, J., Roquet, C., Boleda, M., et al. (2022). Tempo and drivers of plant diversification in the European mountain system. *Nat. Commun.*, 13, 2750.

Tadesse, M. (2004). *Helichrysum* Mill., Asteraceae (Compositae). In: *Flora of Ethiopia and Eritrea,* part 2, pp. 163–178. Hedberg, I., Friis, I., Edwards, S., (eds.). Addis Abeba University: Addis Ababa.

Tamura, K., Battistuzzi, F.U., Billing-Ross, P., Murillo, O., Filipski, A., & Kumar, S. (2012). Estimating divergence times in large molecular phylogenies. *Proc. Natl. Acad. Sci. U.S.A.*, 109(47), 19333–19338.

Tamura, K., Tao, Q., & Kumar, S. (2018). Theoretical foundation of the RelTime methods for estimating divergence times from variable evolutionary rates. *Mol. Biol. Evol.*, 35(7), 1770–1782.

Tao, Q., Tamura, K., Mello, B., & Kumar, S. (2020). Reliable confidence intervals for RelTime estimates of evolutionary divergence times. *Mol. Biol. Evol.*, 37(1), 280–290.

Tenorio, E.A., Montoya, P., Norden, N., Rodríguez-Buriticá, S., Salgado-Negret, B., & Gonzalez, M.A. (2023). Mountains exhibit a stronger latitudinal diversity gradient than lowland regions. *J. Biogeogr.*, 50, 1026–1036.

Thomas, A., Meudt, H.M., Larcombe, M.J., et al. (2023). Multiple origins of mountain biodiversity in New Zealand's largest plant radiation. *J. Biogeogr.*, 50, 947–960.

Tietje, M., Antonelli, A., Baker, W.J., Govaerts, R., Smith, S.A., & Eiserhardt, W.L. (2022). Global variation in diversification rate and species richness are unlinked in plants. *Proc. Natl. Acad. Sci. U.S.A.*, 119(27), e2120662119.

Ufimov, R., Gorospe, J.M., Fér, T., et al. (2022). Utilizing paralogues for phylogenetic reconstruction has the potential to increase species tree support and reduce gene tree discordance in target enrichment data. *Mol. Ecol. Resour.*, 22, 3018–3034.

Van de Peer, Y., Ashman, T.L., Soltis, P.S., & Soltis, D.E. (2021). Polyploidy: an evolutionary and ecological force in stressful times. *Plant Cell*, 33(1), 11–26.

Van de Peer, Y., Mizrachi, E., & Marchal, K. (2017). The evolutionary significance of polyploidy. *Nat. Rev. Genet.*, 18, 411–424.

Van Els, P., Herrera-Alsina, L., Pigot, A.L., & Etienne, R.S. (2021). Evolutionary dynamics of the elevational diversity gradient in passerine birds. *Nat. Ecol. Evol.*, 5(9), 1259–1265.

Wen, J., Zhang, J.Q., Nie, Z.L., Zhong, Y., & Sun, H. (2014). Evolutionary diversifications of plants on the Qinghai-Tibetan Plateau. *Front. Genet.*, 5, 4.

Westerhold, T., Marwan, N., Drury, A.J., et al. (2020). An astronomically dated record of Earth's climate and its predictability over the last 66 million years. *Science*, 369, 1383–1387.

Wiens, J.J., Ackerly, D.D., Allen, A.P., et al. (2010). Niche conservatism as an emerging principle in ecology and conservation biology. *Ecol. Lett.*, 13, 1310–1324.

Wiens, J.J., & Graham, C.H. (2005). Niche Conservatism: Integrating Evolution, Ecology, and Conservation Biology. *Annu. Rev. Ecol. Evol. Syst.*, 36, 519–539.

Winkler, M., Lamprecht, A., Steinbauer, K., et al. (2016). The rich sides of mountain summits – a pan-European view on aspect preferences of alpine plants. *J. Biogeogr.*, 43, 2261–2273.

Xing, Y., & Ree, R.H. (2017). Uplift-driven diversification in the Hengduan Mountains, a temperate biodiversity hotspot. *Proc. Natl. Acad. Sci. U.S.A.*, 114(17), E3444–E3451.

Ye, X.-Y., Ma, P.-F., Yang, G.-Q., et al. (2019). Rapid diversification of alpine bamboos associated with the uplift of the Hengduan Mountains. *J. Biogeogr.*, 46, 2678–2689.

Yu, H., Miao, S., Xie, G., Guo, X., Chen, Z., & Favre, A. (2020). Contrasting Floristic Diversity of the Hengduan Mountains, the Himalayas and the Qinghai-Tibet Plateau Sensu Stricto in China. *Front. Ecol. Evol.*, 8, 136.

Zachos, J., Dickens, G., & Zeebe, R. (2008). An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature*, 451, 279–283.

Zachos, J., Pagani, M., Sloan, L., et al. (2001). Trends, Rhythms, and Aberrations in Global Climate 65 Ma to Present. *Science*, 292, 686–693.

Zhang, C., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, 19 (Suppl 6), 153.

Zizka, A., Onstein, R.E., Rozzi, R., et al. (2022). The evolution of insular woodiness. *Proc. Natl. Acad. Sci. U.S.A.*, 119(37), e2208629119.

Supplementary Materials

Supplementary materials associated with this article can accessed online.

1. List of Supplementary Figures

Supplementary Figure S1. Pairwise distance histogram for paralogy inference.

Pairwise distance histogram from ParalogWizard showing the distribution of pairwise distances between exonic contigs for each locus (in divergence percentages, 7.9 – 20.3%). Dashed lines indicate mean divergence values for each peak, and standard deviation (minus and plus sigma) values are provided in the legend.

Supplementary Figure S2. Dated ML tree with 95% confidence intervals (calibration points indicated).

Dated Maximum-Likelihood phylogeny of the HAP clade generated with RelTime implemented in MEGA 11. Blue numbers next to the nodes show mean ages (above) and grey numbers show 95% confidence intervals (below). Calibration points are indicated to the left of the nodes.

Supplementary Figure S3. ML tree of the HAP clade.

Maximum-Likelihood phylogeny (RAxML-NG) of the HAP clade based on nuclear DNA loci generated with the target-enrichment probe set Compositae1061. Node support values are indicated: BS (in blue, above) and TBE (in red, below).

Supplementary Figure S4. ASTRAL tree of the HAP clade.

ASTRAL species tree of the HAP clade based on nuclear DNA loci generated with the target-enrichment probe set Compositae1061. Numbers next to the nodes indicate local posterior probabilities (LPP).

Supplementary Figure S5. Ancestral range inference of the HAP clade.

Ancestral biogeographic range inference of the HAP clade using the best-fitting model DEC+j. It is based on a time-calibrated phylogeny generated under the concatenation approach using target-enrichment data (Compositae1061 probe set). Pie charts at nodes show the relative probability of the possible states. Primary colours indicate single areas; grey denotes area combinations. The shift in net diversification rates is marked in red (approx. 1.2 Ma). Coloured squares indicate the main Pleistocene mountain radiations.

Supplementary Figure S6. Node numbers on the time-calibrated phylogeny of the HAP clade.

Node numbers on a time-calibrated phylogeny of the HAP clade generated under the concatenation approach using target-enrichment data (Compositae1061 probe set). These node numbers correspond to the ones in the ancestral range probabilities table (Supplementary Table S6).

Supplementary Figure S7. Bract colour evolution in the HAP clade.

Ancestral state reconstruction analysis for bract colour evolution in the HAP clade based on a time-calibrated ML tree generated with target-enrichment data (Compositae1061 probe set). Pie charts on tree nodes indicate the probability of the most likely ancestral morphology. Circles next to the tips are coloured according to the current morphotype of each species.

Supplementary Figure S8. Elevational range preference evolution in the HAP clade.

Ancestral state reconstruction analysis for elevational range preference evolution in the HAP clade based on a time-calibrated ML tree generated with target-enrichment data (Compositae1061 probe set). Pie charts on tree nodes indicate the probability of the most likely ancestral elevational range. Circles next to the tips are coloured according to the current preference of each species.

Supplementary Figure S9. Habitat preference evolution in the HAP clade.

Ancestral state reconstruction analysis for habitat preference evolution in the HAP clade based on a time-calibrated ML tree generated with target-enrichment data (Compositae1061 probe set). Pie charts on tree nodes indicate the probability of the most likely ancestral habitat. Circles next to the tips are coloured according to the current preference of each species.

Supplementary Figure S10. Life form evolution in the HAP clade.

Ancestral state reconstruction analysis for life form evolution in the HAP clade based on a time-calibrated ML tree generated with target-enrichment data (Compositae1061 probe set). Pie charts on tree nodes indicate the probability of the most likely ancestral morphology. Circles next to the tips are coloured according to the current morphotype of each species.

2. List of Supplementary Tables

Supplementary Table S1. Studied materials.

List of studied materials, including herbaria codes and voucher information. BioSample accession numbers are provided for each specimen (accessible at https://www.ncbi.nlm.nih.gov/sra). Samples taken from other studies are associated with the following BioProjects: Blanco-Gavaldà et al., 2023, BioProject PRJNA936872; from Blanco-Gavaldà et al., 2025, BioProject PRJNA1121119; Mandel et al., 2019, BioProject PRJNAS40287; and Schmidt-Lebuhn and Bovill 2021, BioProject PRJNA665592.

Supplementary Table S2. Biogeographic classification of HAP clade taxa.

Geographical distribution of HAP clade species for biogeographic ancestral state reconstruction analyses with BioGeoBEARS. The first row lists the 12 biogeographic areas defined in the study with their corresponding letter codes in brackets.

Supplementary Table S3. Functional trait and ecological preference classification of HAP clade taxa.

Functional trait (bract colour and life form) and ecological preference (elevation and habitat) classification table for HAP clade species.

Supplementary Table S4. SecSSE models and parameter estimates.

Summary of the defined SecSSE models and parameter estimates. The best-fitting models, identified by the highest AIC scores are highlighted in bold. Abbreviations: L (lambda, speciation rate); M (mu, extinction rate); Q (transition rate); LL (likelihood); k (number of parameters); AIC (Akaike Information Criterion). Model types: ETC (examined trait-dependent); CTD (concealed trait-dependent); CR, (constant rate). Variations: full (all possible speciation and extinction rates are estimated); L1 (a single speciation rate is estimated); M1 (a single extinction rate is estimated); simple (extinction rate fixed at 0).

Supplementary Table S5. Paralogy statistics from ParalogWizard.

Statistics from paralogy analyses using ParalogWizard. For each sample, the number of paralogues loci within the selected divergence threshold (7.9—20.4%) is provided.

Supplementary Table S6. Summary statistics of the selected loci.

Summary statistics of the selected loci (including paralogues) after filtering for missing data for subsequent phylogeny reconstruction (script HPM5).

Supplementary Table S7. Summary statistics of biogeographic model testing in BioGeoBEARS.

Summary statistics of biogeographic model testing in BioGeoBEARS. Likelihood scores (LnL), corrected Akaike Information Criterion (AICc) and free parameter values are detailed. Parameters include: dispersal (d), extinction (e) and founder-event (j). The best-fitting model (DEC+j) is highlighted in bold.

Supplementary Table S8. Ancestral range probabilities for each node.

Ancestral range probabilities for the HAP clade on the DEC+j model, corresponding to the pie charts in Supplementary Figure S5. The first column contains node numbers from Supplementary Figure S6.

Supplementary Table S9. Biogeographic Stochastic Mapping dispersal counts.

Average number of dispersal events estimated in the history of the HAP clade using 100 biogeographic stochastic mappings under the DEC+j model. Counts are averaged across and standard deviations are provided below each table. Rows represent source areas and columns the sink areas. Cell colours indicate event frequency (red: highest to green: lowest). Additional tables summarise exchanges between similar elevational belts (lowland-to-lowland, mountain-to-mountain) and transitions between bioclimatically distinct elevational belts (lowland-to-mountain and vice versa), reflecting broad-scale niche conservatism and niche shifts.

Supplementary Table S10. Ancestral state reconstruction models and parameter estimates.

Summary of the fitted models for ancestral state reconstruction for two functional traits (bract colour and life form) and two ecological preferences (elevation and habitat). Model abbreviations correspond to: ER, model with all rates equal; ARD, model with different transition rates for all possible transitions; SYM, symmetrical model in which rates between any two states do not differ. The best-fitting model according to AIC is indicated in bold.

Supplementary Table S11. Time, climate and diversity dependent diversification model fit.

Summary of the fit of various diversification models applied to the HAP clade's phylogeny. The best-fitting model based on AICc is highlighted in bold and green. Abbreviations: NP (number of free parameters); logL, (log-likelihood); AICc, (corrected AIC scores); lambda (speciation rate); alpha (rate of variation in speciation over time or with temperature, depending on the model); mu (extinction rate); K (carrying capacity in diversity-dependent models); beta (rate of variation of extinction over time or with temperature, depending on the model).

Supplementary Table S12. Summary of correlated evolution testing for trait-environment pairs.

Maximum likelihood correlation tests on pairs of binary traits (functional trait vs. ecological preference) performed with BayesTraits. The independent model (four rate parameters, alpha1, beta1, alpha2, beta2) assumes traits evolve independently on the phylogenetic tree, while the dependent model (eight rate parameters) assumes correlated evolution, where the rate of change in one trait depends on the state of the other (transition rates, q). In the dependent model, double transitions are set to zero. Likelihood ratio (LR) tests were used to compare model likelihoods (Lh), and significance was assessed from chi-squared tests (p-values). Probabilities for the root being in specific state combinations are also indicated.

3. Supplementary References

List of reference floras and online resources consulted for extracting geographic range distribution (Supplementary Table S2), functional trait data and ecological preference data (Supplementary Table S3) used to classify the studied HAP clade species.

General Discussion

General Discussion

The research carried out in this doctoral thesis has resulted in advances in the understanding of plant systematics and evolution across several domains: methodology, phylogenetic relationships, biogeography, diversification and evolutionary processes. These contributions are particularly significant for the study of Afromontane and Afroalpine floras, as well as the HAP clade. In this section, the most relevant findings and their implications are discussed.

1. Methodology

Our research demonstrates that the Compositae1061 probe kit (Mandel et al., 2014) successfully retrieves a sufficient number of phylogenetically informative markers to address complex evolutionary questions. These markers notably enhance phylogenetic resolution and support for recently radiated groups as is the case of *Helichrysum* Mill. (over 500 species) within the broader HAP clade (over 800 species). As outlined in the Introduction, NGS techniques outperform Sanger sequencing in data quantity and quality. However, NGS can also exacerbate other issues, such as paralogy and incomplete lineage sorting, which can lead to biased phylogenetic inference (Liu et al., 2015; Kong & Kubatko, 2021).

Given the complexity of the evolutionary history of the Compositae in general (Barker et al., 2016) and the HAP clade in particular (Galbany-Casals et al., 2014), we anticipated the retrieval of a high proportion of paralogous copies among the targeted loci using the Compositae1061 probe set. Although the Compositae1061 set was designed to recover low-copy orthologous loci, potential paralogs often remain, thus, it is important to use appropriate methodologies to address this issue. Evidence suggests that removing paralogs and restricting data to single-copy genes may result in an extremely reduced dataset, especially in lineages that have repeatedly experienced whole genome duplications and are prone to hybridization (Morales-Briones et al., 2021). This massive loss of sequenced information can reduce phylogenetic inference resolution. Instead, retaining paralogous copies can maximize the amount of informative data (Smith & Hahn, 2021; Ufimov et al., 2022; Yan et al., 2022). Some recent studies adopted this approach (e.g. *Alchemilla* L., Morales-Briones et al., 2021; Sapindales, Joyce et al., 2023). Nevertheless, there is still a lack of consensus on the best practice (Kandziora et al., 2022b) and few bioinformatic tools exist to effectively detect allelic variation in Hyb-Seq data (reviewed in Ning et al., 2025).

Some widely used pipelines do not incorporate paralogous copies in phylogenetic analyses. For example, PHYLUCE (Faircloth, 2015) discards all loci with potential paralogs and HybPiper (Johnson et al., 2016) retains only the longest copy for each locus or, if multiple copies of the same locus have equal read lengths, the copy with higher coverage or the most similar sequence to the target is retained. Instead, we addressed paralogy using a recently developed pipeline, ParalogWizard (Ufimov et al., 2022) alongside HybPhyloMaker (Fér & Schmickl, 2018). ParalogWizard identifies potential paralogs based on sequence divergence peaks, distinguishing allelic variation from higher sequence divergence linked to putative paralogy. Putative paralogous copies are then treated as additional markers to build orthologous alignments for phylogenetic inference. This strategy increases the size and the phylogenetic signal of the final dataset, which is particularly relevant for our study group, as excluding paralogs would significantly reduce the dataset. However, paralog selection can influence concordance across methodologies and may not fully resolve relationships affected by reticulate evolution. It is worth noting that hybrid copies derived from different parental species behave differently from those occasioned by gene or whole-genome duplications, introducing different biases that might be more challenging to manage (Moore-Pollard & Mandel, 2023). Moreover, increasing the number of markers also amplifies the limitations of the concatenation method, because it can be understood as a simplification of the coalescent model that treats all genes as a single variable (Kubatko & Degnan, 2007; Liu et al., 2015). While concatenating all sequences in a single supermatrix may yield highly resolved phylogenies, this approach overlooks underlying biological processes. Consequently, higher support values may be an artefact from analytical biases rather than indicators of phylogenetic reliability (Herrando-Moraira et al., 2018). Furthermore, relying on bootstrap support has been criticized as outdated for HTS datasets (Salichos & Rokas et al., 2013). Originally designed to assess inference robustness when data was limited, it might no longer be the most appropriate measure given the scale of modern genomic datasets. To address this, we also assessed Transfer Bootstrap Expectation (TBE) support values, which provides a more reliable assessment of phylogenetic inference (Lemoine et al., 2018). Despite these potential issues, by separating orthologous and pseudo-orthologous copies, we generated well-resolved and supported phylogenies mostly consistent with morphological and distributional data.

As exposed previously, methodological assumptions affect phylogenetic inference accuracy across taxonomic levels, depending on the evolutionary nature of the group (Lozano-Fernández, 2022). Since it is still controversial which method provides more accurate trees (Edwards et al., 2016; Springer & Gatesy, 2016), we chose to employ concatenation, coalescence and even hybrid approaches to ensure robustness and to highlight the unique insights provided by each methodology. Some recent clades show low branch support across different phylogenetic inference methods, particularly the position of those of known allopolyploid origins. Relationships among many closely related species resulting from recent radiations were weakly supported, independently of the inference methods, likely due to high gene tree discordance associated with incomplete lineage sorting (ILS, the persistence of ancestral polymorphisms after speciation) and reticulation events. This outcome was expected given the use of a large number of markers (Degnan & Rosenberg, 2009). Similarly, the placement of clades with hypothetical allopolyploid origin varied between concatenation and coalescent-based phylogenies, with higher supports observed in the concatenated trees. Considering that allopolyploids inherit genomes from different parental species with distinct evolutionary histories, distinguishing homoeologous gene copies poses an extra significant challenge, leading to conflicting phylogenetic results (Smith & Hahn, 2021; Ning et al., 2025).

While concatenation and coalescence-based trees provided largely congruent and consistent results, coalescence-based trees, although generally yielded lower branch supports, often resulted in biologically more plausible phylogenetic relationships. In evolutionary scenarios involving polyploidization and hybridization, branch support may be lower, but the inferred relationships provide valuable information, especially when compared to concatenated phylogenies (e.g. Herrando-Moraira et al., 2018). For example, the dataset analyzed in Chapter 3 suggests that the genus *Pseudonaphalium* Kirp. is monophyletic based on the concatenation method. However, coalescence-based analyses split it into two independent clades, one of which appears more closely related to *Achyrocline* (Less). DC. Nevertheless, both methodologies are highly congruent with biogeographical patterns, separating a North and Central American clade and a South American clade, and their combined results hint at a possible allopolyploid origin as previously hypothesized (Smissen et al., 2011; Galbany-Casals et al., 2014) as well as the potential parental species involved. Our results also indicate a potential allopolyploid origin for a Malagasy clade (clade MAD6 in Chapter 2). Similarly, this large montane clade appears monophyletic under the concatenation approach but splits into two distinct lineages in the coalescence-based inference.

Given these complexities, exploring both concatenation and coalescence-based approaches provides complementary insights that can help detect and understand the origin of polyploid groups. However, since the primary goal of this thesis was to analyze the general biogeographic and evolutionary patterns of the whole genus, we based our analyses on the concatenated tree, which provided overall higher supported branch lengths. We performed our divergence time estimations and downstream analyses on it, as branch length estimates in concatenation tend to be more accurate in large datasets than coalescent ones (Schwartz & Mueller, 2010). Nevertheless, more refined dating methodologies for large coalescent-based phylogenies are currently being developed (see Tabatabaee et al., 2025). Further research, including reticulation analyses and finer-scale phylogenetic studies, is required to fully unravel the evolutionary history of allopolyploid clades and address taxonomic challenges within the HAP clade. Advances in molecular systematics, including new bioinformatic tools (e.g. Nauheimer et al., 2021; Zhang & Mirarab, 2022; Jackson et al., 2023) and improved target-enrichment kits that account for paralogy (e.g. Moore-Pollard et al., 2024) will likely facilitate and improve future research on complex evolutionary groups.

Despite generating Hyb-Seq data of both nuclear and chloroplast molecular markers, we chose not to include the latter in our studies due to several limitations, including low resolution, incongruent topologies and insufficient DNA concentration. The latter issue was exacerbated by the use of older herbarium samples, for which DNA quantity was often barely sufficient to perform target-capture. However, the primary reason for excluding the analyses of plastid marker was their inheritance pattern (usually maternally transmitted), which produces organellar phylogenies at the infrageneric level —especially between relatively close species — that often do not reflect species boundaries due to the effects of ILS and hybridization (e.g. Scheunert & Heubl, 2017). In the context of this thesis and the HAP clade, previous studies have shown that plastid data was incongruent with nuclear markers and yielded biogeographically inconsistent patterns that were difficult to interpret (Galbany-Casals et al., 2014). Nonetheless, the plastid data generated remains highly valuable, particularly as a first approach for future works investigating the origin of the allopolyploid clades, as cytonuclear phylogenetic discordance may be indicative of reticulate evolution (Rieseberg & Soltis, 1991). Additionally, the Hyb-Seq approach proved to be highly effective for herbarium specimens with sufficient DNA, enabling the sequencing of samples up to 102 years old (although most of them were from the 1950-1980s) and giving the chance to include species from remote and inaccessible regions that otherwise could not have been sampled. DNA quality and quantity of herbarium specimens are highly influenced by factors such as plant chemistry and preservation conditions (drying method, chemical treatment, environmental conditions, etc.; Drábková, 2014), which were out of our control.

2. Phylogenetic relationships and biogeographic history

The use of Next Generation Sequencing methods (NGS) coupled with theoretical and methodological improvements in phylogenetic inference (Bouckaert et al., 2019; Kozlov et al., 2019; Zhang et al., 2020), molecular dating (Smith & O'Meara, 2012; Tamura et al., 2018) and biogeographic inference (Matzke et al., 2013; Dupin et al., 2017) have facilitated the study of distribution patterns in large and complex plant groups. These innovations have led to an exponential increase in high-quality biogeographic studies over recent years (e.g. *Jurinea* Cass., Herrando-Moraira et al., 2023; *Cypripedium* L., Lagou et al., 2024; *Astragalus* L., Folk et al., 2024; *Panax* L., Zhang et al., 2025).

In Chapters I and II, we explored the origins and the temporal, geological and climatological contexts in which the Afromontane and Afroalpine *Helichrysum* lineages arose. Based on the aforementioned methods, we constructed highly resolved phylogenetic trees that incorporated extensive taxonomic and biogeographic sampling of *Helichrysum* and related genera within the HAP clade. Overcoming the evolutionary and taxonomic complexities of this hyperdiverse and widespread group required substantial sampling efforts, which became possible thanks to international collaborations and herbarium collections. Additionally, the obtention of sequences for hundreds of nuclear loci enabled us to address the main limitations of previous studies based on Sanger sequencing (Galbany-Casals et al., 2014; Andrés-Sánchez et al., 2019). These efforts allowed us to test previously formulated hypotheses on the origin and evolution of these floras and propose new ones (Table 1 and 2).

Table 1. Summary of the specific hypotheses evaluated and newly formulated regarding the evolutionary and biogeographic history of *Helichrysum* in Tropical Africa.

Hypotheses concerning Tropical Africa	
Tropical Afromontane and Afroalpine *Helichrysum* species have southern African origins (Galbany-Casals et al., 2014; Andrés-Sánchez et al., 2019) and not Eurasian origins as frequently reported for Afroalpine groups (Brochmann et al., 2022).	Confirmed
Tropical Afromontane and Afroalpine lineages originated from multiple colonization events instead of a single dispersal (Galbany-Casals et al., 2014; Kandziora et al., 2022a).	Confirmed
Tropical Afromontane and Afroalpine species descend from preadapted temperate ancestors that reached these biomes by LDD. The Drakensberg mountains played a key role as a stepping-stone area in their dispersal (Galley et al., 2007; Kandziora et al., 2022a).	Partially confirmed, only for Tropical Afromontane species
Tropical Afroalpine species descend from multiple independent colonizations by ancestors distributed in the adjacent regional Afromontane belts. Thus, their origin involved adaptation to harsher climatic conditions.	New hypothesis, confirmed
Climatic oscillations since the late Miocene fostered repeated range expansion and contraction of montane and alpine habitats, promoting dispersal and diversification and shaping current biodiversity patterns in those regions (Chala et al., 2017; Muellner-Riehl et al., 2019).	Confirmed

Table 2. Summary of the specific hypotheses evaluated and newly formulated regarding the evolutionary and biogeographic history of *Helichrysum* in Madagascar.

Hypotheses concerning Madagascar	
Helichrysum colonized Madagascar from Africa several times independently (Galbany-Casals et al., 2014), as expected due to its physical proximity (Yoder & Nowak, 2006).	Confirmed
Helichrysum colonized Madagascar from Africa several times independently (Galbany-Casals et al., 2014), as expected due to its physical proximity (Yoder & Nowak, 2006).	Partially confirmed
Both repeated long-distance dispersal and in-situ diversification played a significant role in shaping *Helichrysum* biodiversity patterns on Madagascar.	New hypothesis
Biome conservatism played a key role in the establishment of new lineages on Madagascar, resulting in allopatric speciation (Wiens, 2004; Wiens & Graham, 2005; Crisp & Cook, 2012).	Confirmed
Colonization and establishment were followed by local-scale biome shifts (Donoghue & Edwards, 2014), mainly involving descends from high to low-elevation belts and adaptation to particular microhabitats, presumably through ecological speciation.	New hypothesis

In Chapter I, we confirmed that the Tropical Afromontane and Afroalpine regions have been colonized several times independently. Our analyses revealed that the southern African grasslands, included in the broad definition of the Afromontane region, are the primary source area of Tropical Afromontane species. However, the adjacent Tropical Afromontane region emerged as the main source of several independent Afroalpine clades, contrasting with studies conducted on other Afroalpine lineages that place their origin in distant temperate alpine-like regions (reviewed in Brochmann et al., 2022), or in this case, the summit of the Drakensberg in southern Africa. In Chapter II, we inferred that Madagascar has been colonized at least six times independently from different African biomes. Biome conservatism played a key role in the successful establishment of lineages on the island, with the southern African grasslands identified as the main source of Malagasy highland species and the Tropical Afromontane region as the primary source of high-elevation Malagasy lineages. Our results confirm a strong Afromontane affinity between mainland Africa and Madagascar's mountain floras, with similar patterns as those observed in *Erica* L. (Pirie et al., 2019). *Helichrysum* and *Erica* together constitute the iconic Afromontane shrublands, providing further support for the inclusion of Madagascar in the broad definition of the Afromontane region (Gehrke & Linder, 2014). Remarkably, we also report the retention of arid bioclimatic preferences for one lineage restricted to the most arid part of the island, with an ancestor originating in the arid regions of southwestern Africa.

In general, Afromontane and Afroalpine *Helichrysum* species had been poorly sampled and investigated in previous studies. Aside from the few species included in Galbany-Casals et al. (2014), large knowledge gaps persisted for one of the most relevant genera in these biodiversity hotspots, which constitutes the base of many communities and landscapes along the elevational gradient (Hedberg, 1970; Brochmann et al., 2022). Our results support and provide a more detailed understanding of what previous authors already hypothesized a decade ago based on the poorly resolved phylogenetic trees obtained with Sanger markers. It is worth noting that the results of this thesis provide novel insights into an overlooked biodiversity hotspot. While prior studies on Tropical Africa's high-elevation flora primarily focused on Afroalpine lineages with Eurasian origins (e.g. *Anthoxanthum* L., Pimentel et al., 2013; *Lychnis* L., Gizaw et al., 2016; *Alchemilla*, Gehrke et al., 2016), our research shed light on the biogeographic history of the Afromontane flora as well, using a lineage with southern African origins as a study-case. Similarly, this is the first time the evolutionary history of a large Malagasy genus that repeatedly colonized the island and highly diversified in the mountains is explored in depth (but see Cyperaceae, Larridon et al., 2021 and Danaideae, Razafimandimbinson et al., 2022) since most phylogenetic studies on the island have focused on lowland lineages, either from the xeric western side (e.g. *Adansonia* L., Wan et al., 2024) or the eastern humid rainforest (e.g. *Canarium* L., Federman et al., 2015). Our findings offer new perspectives on the origins and general diversification patterns of Afromontane and Afroalpine lineages, contributing to a broader understanding of fragmented and isolated biodiversity hotspots and large widespread lineages.

Lastly, while this thesis did not explicitly focus on taxonomical issues, one of our goals in Chapter II was to test whether the most recent taxonomical revision of Malagasy *Helichrysum* (Humbert, 1962), which classified Malagasy species into 11 groups based on morphological traits, matched molecular-based phylogenetic relationships. In most cases, it did, showing that morphological studies are a relevant complementary source of information to molecular data. However, notable discrepancies were observed, such as Humbert's group III, which was fragmented into several clades, while groups II and IV mostly merged into a single clade (see Chapter 2, Figure 2, clade MAD6). Additionally, three endemic genera, *Humeocline* Anderb., *Syncephalum* DC. and *Stenocline* DC., were found to be nested within Malagasy *Helichrysum*, reinforcing the proposal presented in the Introduction that a revised generic delimitation is required. In Chapter III, we present the most comprehensive and highly resolved phylogeny of the HAP clade to date. Our findings confirmed *Helichrysum*'s paraphyly, as it includes at least eight other genera: *Achyrocline*, *Anaphalis* DC., *Galeomma* Rauschert, *Humeocline, Pseudognaphalium*, *Stenocline* DC., *Stenophalium* Anderb. and *Syncephalum*. *Helichrysum* is also polyphyletic as suggested in previous studies (Smissen et al., 2011; Galbany-Casals et al., 2014; Nie et al., 2016) given that, apart from *Helichrysum dasyanthum* Sweet, shown to be nested within *Anaxeton* Gaertn. in previous works (Galbany-Casals et al., 2014 and references therein) and here, additional southern African *Helichrysum* species are shown for the first time to be more closely related to other southern African Gnaphalieae genera. The inclusion and release of over 500 newly sequenced taxa using the NGS Hyb-Seq approach, covering more than 60% of the HAP clade and including the first-ever sequenced genus*, Stenocline*, provides a robust phylogenomic framework for further research. The Hyb-Seq data generated in this thesis can be reanalyzed using emerging approaches, offering new insights into the highly controversial taxonomic classification of the HAP clade and the Compositae family.

3. Diversification patterns and evolutionary history of the HAP clade

The development of advanced diversification modeling methods (e.g. Etienne et al., 2012; Rabosky 2014; Morlon et al., 2016; Beaulieu & O'Meara, 2016; Herrera-Alsina et al., 2019; Barido-Sottani et al., 2020) has provided a powerful framework for exploring the processes and drivers of plant diversification. These models allow us to understand how evolutionary and ecological dynamics shaped current plant distribution patterns.

In Chapter III, we investigated the diversification dynamics of the HAP clade (c. 800 species) in high-elevation ecosystems worldwide. Our study focused on biogeographic aspects, the temporal and environmental context in which the lineage diversified, and key functional traits that might have influenced changes in diversification rates and predominant speciation modes. The strongest points of this work are: (i) comprehensive taxonomic and biogeographic sampling; (ii) highly resolved time-calibrated phylogenetic trees with strongly supported nodes for major clades; (iii) phenotypic data collection and classification of species based on life forms and bract color and (iv) distribution range and habitat data collection and classification of species taking into account elevation, latitude and habitat preferences.

As highlighted previously, Chapter III presents the most extensive phylogeny of the HAP clade to date, covering taxa across its entire distribution range. The clade, in addition to the predominantly African genus *Helichrysum*, includes other specious genera such as *Anaphalis* (primarily found in Asia), *Pseudognaphalium* (mainly found in the Americas) and smaller ones (*Achyrocline* and *Stenophalium* in the Americas; *Humecline*, *Stenocline* and *Syncephalum* in Madagascar; *Galeomma* in South Africa). These genera predominantly thrive in high-elevation environments, yet the remarkable radiations of some of them in such habitats and the evolutionary processes behind those have not been thoroughly explored. To address this knowledge gap, we combined multiple sources of data, compiling distributional, morphological and ecological information for 62% of the clade's diversity, represented by 516 samples (including infraspecific taxa). This comprehensive dataset allowed us to investigate the drivers shaping the clade's current biodiversity patterns.

We confirmed that the HAP clade likely originated in the arid regions of southern Africa during the Mid-Late Miocene (Galbany-Casals et al., 2014; Andrés-Sánchez et al., 2019). Initially inhabiting lowland habitats, the clade expanded and diversified into higher-elevation environments as the climate became increasingly arid and cooler (Sepulchre et al., 2006; Westerhold et al., 2020). By the Pliocene, it had expanded its range from the southern African montane grasslands (its initial center of diversification) into Tropical Africa, where it underwent significant speciation in the mountains. The onset of the Pleistocene, marked by the alternance of glacial and interglacial periods (Lisiecki & Raymo, 2007), triggered a nearly fourfold increase in net diversification rates. This period coincided with parallel dispersal events towards Madagascar, Eurasia and the Americas, followed by extensive in-situ radiations in each of these regions. Notably, the evolutionary history of these dispersal events followed slightly different trajectories. Malagasy *Helichrysum* lineages originated through multiple independent colonizations and parallel radiations of varied magnitudes, whereas in the Americas (where the clade is represented by *Achyrocline*, *Pseudognaphalium* and *Stenophalium*) and Eurasia (where the clade includes *Anaphalis* and *Helichrysum*), one, or at most two colonization events gave rise to substantial in-situ radiations. However, the exact number of colonization events remains uncertain due to low support and conflicting topologies, likely associated with the allopolyploid origin of these lineages as discussed in previous sections. This pattern would align with the hypothesis that polyploidy facilitates the successful colonization of new territories (e.g. Linder & Barker, 2014) and provides adaptive advantages in front of harsh and unstable environmental conditions (Van der Peer et al., 2020).

Our results also reveal a pattern of nested radiations (Linder & Verboom, 2015), in which mountain diversity feeds other mountain systems both within and across continents, resulting in a single major radiation. Additionally, we report downward migrations that may have contributed to lower elevation diversity mostly in temperate latitudes. These patterns suggest that both long-dispersal capacity and permeability across elevational gradients were key in fostering allopatric and microallopatric speciation.

While no specific key innovations were directly associated with increased diversification rates in high-elevations, several functional traits and habitat preferences exhibited strong correlations and were repeatedly acquired across the phylogeny, indicating potential microhabitat specialization (Velásquez-Puentes et al., 2023). This suggests that the clade's evolutionary success and capacity for radiating are likely linked to its exceptional adaptability and niche lability, which have facilitated convergent evolution in specific environments. Therefore, both geographic and ecological speciation appear to have contributed to increased net diversification rates at different spatial and temporal scales. Ancestral niche conservatism, associated with the dispersal of preadapted ancestors, likely facilitated the initial stages of colonization and establishment in new regions (Wiens & Graham, 2005; Donoghue, 2008; Wiens et al., 2010; Edwards & Donoghue, 2013). Subsequent habitat expansions and contractions, driven by Pleistocene climatic oscillations, pushed plants up and down the elevational gradient, promoting isolation and secondary contacts on a local scale. This dynamic, consistent with the flickering connectivity hypothesis (Flantua et al., 2019), led to allopatric speciation at regional and local scales, to be followed by in-situ diversification. The speciation patterns observed in Malagasy clades, as detailed in Chapter II (i.e. initial biome conservatism followed by repeated biome shifts along the elevational gradient as well as adaptation to particular microhabitats) reflect this multi-stage diversification dynamics on a regional scale, aligning with other studies performed on mountain island floras (e.g. Thomas et al., 2023).

This research deepens our understanding of the evolutionary processes underlying high-elevation plant radiations. By integrating phylogenetic, biogeographic, morphological and ecological data at multiple scales, we provide new insights into the complex interplay of allopatric and ecological speciation processes shaping biodiversity (and likely hybrid and polyploid speciation if the allopolyploid origin of mainly American and Eurasian clades is confirmed). Our findings not only enhance our comprehension of the HAP clade's evolutionary history but also highlight the importance of past climatic oscillations, dispersal dynamics and ecological flexibility in global plant radiations. This integrative approach sets the ground for future studies to unravel similar complex evolutionary histories in other large plant groups.

References

Andrés-Sánchez, S., Verboom, G.A., Galbany-Casals, M., & Bergh, N.G. (2019). Evolutionary history of the arid climate-adapted *Helichrysum* (Asteraceae: Gnaphalieae): Cape origin and association between annual life-history and low chromosome numbers. *J. Syst. Evol.*, 57(5), 468–487.

Barido-Sottani, J., Vaughan, T.G., & Stadler, T. (2020). A Multitype Birth-Death Model for Bayesian Inference of Lineage-Specific Birth and Death Rates. *Syst. Biol.*, 69(5), 973–986.

Barker, M.S., Li, Z., Kidder, T.I., et al. (2016). Most Compositae (Asteraceae) are descendants of a paleohexaploid and all share a paleotetraploid ancestor with the Calyceraceae. *Am. J. Bot.*, 103(7), 1203–1211.

Beaulieu, J.M., & O'Meara, B.C. (2016). Detecting Hidden Diversification Shifts in Models of Trait-Dependent Speciation and Extinction. *Syst. Biol.*, 65(4), 583–601.

Blanco-Gavaldà, C., Galbany-Casals, M., Susanna, A., et al. (2023). Repeatedly Northwards and Upwards: Southern African Grasslands Fuel the Colonization of the African Sky Islands in *Helichrysum* (Compositae). *Plants*, 12(11), 2213.

Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., et al. (2019). BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.*, 15(4), e1006650.

Brochmann, C., Gizaw, A., Chala, D., et al. (2022). History and evolution of the afroalpine flora: in the footsteps of Olov Hedberg. *Alp. Bot.*, 132, 65–87.

Chala, D., Zimmermann, N.E., Brochmann, C., et al. (2017). Migration corridors for alpine plants among the 'sky islands' of eastern Africa: do they, or did they exist? *Alp. Bot.*, 127, 133–144.

Crisp, M.D., & Cook, L.G. (2012). Phylogenetic niche conservatism: what are the underlying evolutionary and ecological causes? *New Phytol.*, 196, 681–694.

Degnan, J.H., & Rosenberg, N.A. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol.*, 24(6), 332–340.

Donoghue, M.J. (2008). Colloquium paper: a phylogenetic perspective on the distribution of plant diversity. *Proc. Natl. Acad. Sci. U.S.A.*, 105(Suppl 1), 11549–11555.

Donoghue, M.J., & Edwards, E.J. (2014). Biome Shifts and Niche Evolution in Plants. *Annu. Rev. Ecol. Evol. Syst.*, 45, 547–572.

Drábková, L.Z. (2014). DNA extraction from herbarium specimens. *Methods Mol. Biol.*, 1115, 69–84.

Dupin, J., Matzke, N.J., Särkinen, T., et al. (2017). Bayesian estimation of the global biogeographical history of the Solanaceae. *J. Biogeogr.*, 44, 887–899.

Edwards, E.J., & Donoghue, M.J. (2013). Is it easy to move and easy to evolve? Evolutionary accessibility and adaptation. *J. Exp. Bot.*, 64(13), 4047–4052.

Edwards, S.V., Xi, Z., Janke, A., et al. (2016). Implementing and testing the multispecies coalescent model: A valuable paradigm for phylogenomics. *Mol. Phylogenetics Evol.*, 94(Pt A), 447–462.

Etienne, R.S., Haegeman, B., Stadler, T., et al. (2012). Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. *Proc. Biol. Sci.*, 279(1732), 1300–1309.

Faircloth, B.C. (2015). PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics*, 32(5), 786–788.

Federman, S., Dornburg, A., Downie, A., Richard, A.F., Daly, D.C., & Donoghue, M.J. (2015). The biogeographic origin of a radiation of trees in Madagascar: implications for the assembly of a tropical forest biome. *BMC Evol. Biol.*, 15, 216.

Fér, T., & Schmickl, R.E. (2018). HybPhyloMaker: Target Enrichment Data Analysis From Raw Reads to Species Trees. *Evol. Bioinform. Online*, 14, 1176934317742613.

Flantua, S.G.A., O'Dea, A., Onstein, R.E., Giraldo, C., & Hooghiemstra, H. (2019). The flickering connectivity system of the north Andean páramos. *J. Biogeogr.*, 46, 1808–1825.

Folk, R.A., Charboneau, J.L.M., Belitz, M., et al. (2024). Anatomy of a mega-radiation: biogeography and niche evolution in *Astragalus*. *Am. J. Bot.*, 111(3), e16299.

Galbany-Casals, M., Unwin, M., Garcia-Jacas, N., et al. (2014). Phylogenetic relationships in *Helichrysum* (Compositae: Gnaphalieae) and related genera: Incongruence between nuclear and plastid phylogenies, biogeographic and morphological patterns, and implications for generic delimitation. *Taxon*, 63, 608–624.

Galley, C., Bytebier, B., Bellstedt, D.U., & Linder, H.P. (2007). The Cape element in the Afrotemperate flora: from Cape to Cairo?.*Proc. Biol. Sci.*, 274(1609), 535–543.

Gehrke, B., Kandziora, M., & Pirie, M.D. (2016). The evolution of dwarf shrubs in alpine environments: a case study of Alchemilla in Africa. *Ann. Bot.*, 117(1), 121–131.

Gehrke, B., & Linder, H.P. (2014). Species richness, endemism and species composition in the tropical Afroalpine flora. *Alp. Bot.*, 124, 165–177.

Gizaw, A., Brochmann, C., Nemomissa, S., et al. (2016). Colonization and diversification in the African 'sky islands': insights from fossil-calibrated molecular dating of *Lychnis* (Caryophyllaceae). *New Phytol.*, 211(2), 719–734.

Hedberg, O. (1970). Evolution of the Afroalpine Flora. *Biotropica,* 2, 16–23.

Herrando-Moraira, S., & Cardueae Radiations Group. (2018). Exploring data processing strategies in NGS target enrichment to disentangle radiations in the tribe Cardueae (Compositae). *Mol. Phylogenetics Evol.*, 128, 69–87.

Herrando-Moraira, S., Roquet, C., Calleja, J.A., et al. (2023). Impact of the climatic changes in the Pliocene-Pleistocene transition on Irano-Turanian species. The radiation of genus *Jurinea* (Compositae). *Mol. Phylogenetics Evol.*, 189, 107928.

Herrera-Alsina, L., van Els, P., & Etienne, R.S. (2019). Detecting the Dependence of Diversification on Multiple Traits from Phylogenetic Trees and Trait Data. *Syst. Biol.*, 68(2), 317–328.

Humbert, H. (1962). *Flore de Madagascar et des Comores (Plantes vasculaires)*; 189e famille. Composées, part 2. Leroy, J.F., (ed.). Muséum National d'Histoire Naturelle: Paris.

Jackson, C., McLay, T., & Schmidt-Lebuhn, A.N. (2023). hybpiper-nf and paragone-nf: containerization and additional options for target capture assembly and paralog resolution. *Appl. Plant Sci.*, 11(4), e11532.

Johnson, M.G., Gardner, E.M., Liu, Y., et al. (2016). HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Appl. Plant Sci.*, 4(7), apps.1600016.

Joyce, E.M., Appelhans, M.S., Buerki, S., et al. (2023). Phylogenomic analyses of Sapindales support new family relationships, rapid Mid-Cretaceous Hothouse diversification, and heterogeneous histories of gene duplication. *Front. Plant Sci.*, 14, 1063174.

Kandziora, M., Gehrke, B., Popp, M., Gizaw, A., Brochmann, C., & Pirie, M.D. (2022a). The enigmatic tropical alpine flora on the African sky islands is young, disturbed, and unsaturated. *Proc. Natl. Acad. Sci. U.S.A.*, 119(22), e2112737119.

Kandziora, M., Sklenář, P., Kolář, F., & Schmickl, R. (2022b). How to Tackle Phylogenetic Discordance in Recent and Rapidly Radiating Groups? Developing a Workflow Using *Loricaria* (Asteraceae) as an Example. *Front. Plant Sci., 12*, 765719.

Kong, S., & Kubatko, L.S. (2021). Comparative Performance of Popular Methods for Hybrid Detection using Genomic Data. *Syst. Biol.*, 70(5), 891–907.

Kozlov, A.M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, 35(21), 4453–4455.

Kubatko L.S., & Degnan, J.H. (2007). Inconsistency of Phylogenetic Estimates from Concatenated Data under Coalescence. *Syst. Biol.*, 56(1), 17–24.

Lagou, L.J., Kadereit, G., & Morales-Briones, D.F. (2024). Phylogenomic analysis of target enrichment and transcriptome data uncovers rapid radiation and extensive hybridization in the slipper orchid genus *Cypripedium*. *Ann. Bot.*, 134(7), 1229–1250.

Larridon, I., Spalink, D., Jiménez-Mejías, P., et al. (2021). The evolutionary history of sedges (Cyperaceae) in Madagascar. *J. Biogeogr.*, 48, 917–932.

Lemoine, F., Domelevo Entfellner, J.B., Wilkinson, E., et al. (2018). Renewing Felsenstein's phylogenetic bootstrap in the era of big data. *Nature*, 556(7702), 452–456.

Linder, H.P., & Barker, N.P. (2014). Does polyploidy facilitate long-distance dispersal? *Ann. Bot.*, 113(7), 1175–1183.

Linder, H.P., & Verboom, G.A. (2015). The Evolution of Regional Species Richness: The History of the Southern African Flora. *Annu. Rev. Ecol. Evol. Syst.*, 46, 393–412.

Lisiecki, L.E., & Raymo, M.E. (2007). Plio–Pleistocene climate evolution: trends and transitions in glacial cycle dynamics. *Quat. Sci. Rev.*, 26 (1-2) 56–69.

Liu, L., Xi, Z., Wu, S., Davis, C.C., & Edwards, S.V. (2015). Estimating phylogenetic trees from genome-scale data. *Ann. N. Y. Acad. Sci.*, 1360, 36–53.

Lozano-Fernandez, J. (2022). A Practical Guide to Design and Assess a Phylogenomic Study. *Genome Biol. Evol.*, 14(9), evac129.

Mandel, J.R., Dikow, R.B., Funk, V.A., et al. (2014). A target enrichment method for gathering phylogenetic information from hundreds of loci: An example from the Compositae. *Appl. Plant. Sci.*, 2(2), apps.1300085.

Matzke, N.J. (2013). Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Front. Biogeogr.*, 5(4), 242–248.

Moore-Pollard, E.R., Jones, D.S. & Mandel, J.R. (2024). Compositae-ParaLoss-1272: A complementary sunflower-specific probe set reduces paralogs in phylogenomic analyses of complex systems. *Appl. Plant Sci.*, 12(1), e11568.

Moore-Pollard, E.R., & Mandel, J.R. (2023). From paralogy to hybridization: Investigating causes of underlying phylogenomic discordance using the complex genus *Packera* (Senecioneae; Asteraceae). *bioRxiv*, 2023.08.14.553290.

Morales-Briones, D.F., Gehrke, B., Huang, C.H., et al. (2021). Analysis of Paralogs in Target Enrichment Data Pinpoints Multiple Ancient Polyploidy Events in *Alchemilla* s.l. (Rosaceae). *Syst. Biol.*, 71(1), 190–207.

Morlon, H., Lewitus, E., Condamine, F.L., Manceau, M., Clavel, J., & Drury, J. (2016). RPANDA: an R package for macroevolutionary analyses on phylogenetic trees. *Methods Ecol. Evol.*, 7, 589–597.

Muellner-Riehl, A.N., Schnitzler, J., Kissling, W.D., et al. (2019). Origins of global mountain plant biodiversity: Testing the 'mountain-geobiodiversity hypothesis'. *J. Biogeogr.*, 46, 2826–2838.

Nauheimer, L., Weigner, N., Joyce, E., Crayn, D., Clarke, C., & Nargar, K. (2021). HybPhaser: A workflow for the detection and phasing of hybrids in target capture data sets. *Appl. Plant Sci.*, 9(7), e11441.

Nie, Z.-L., Funk, V.A., Meng, Y., Deng, T., Sun, H., & Wen, J. (2016). Recent assembly of the global herbaceous flora: evidence from the paper daisies (Asteraceae: Gnaphalieae). *New Phytol.*, 209, 1795–1806.

Ning, W., Meudt, H.M., & Tate, J.A. (2025). A roadmap of phylogenomic methods for studying polyploid plant genera. *Appl. Plant Sci.*, 12(4), e11580.

Pimentel, M., Sahuquillo, E., Torrecilla, Z., Popp, M., Catalán, P., & Brochmann, C. (2013). Hybridization and long-distance colonization at different time scales: towards resolution of long-term controversies in the sweet vernal grasses (*Anthoxanthum*). *Ann. Bot.*, 112(6), 1015–1030.

Pirie, M.D., Kandziora, M., Nürk, N.M., et al. (2019). Leaps and bounds: geographical and ecological distance constrained the colonisation of the Afrotemperate by *Erica*. *BMC Evol. Biol.*, 19, 222.

Rabosky, D.L., Grundler, M., Anderson, C., et al. (2014). BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods Ecol. Evol.*, 5, 701–707.

Razafimandimbison, S.G., Wikström, N., Khodabandeh, A., & Rydin, C. (2022). Phylogeny of the Madagascar-centred tribe Danaideae (Rubiaceae) as a precursor to taxonomic revision: insights into its generic and species limits, affinities and distribution. *Ann. Bot.*, 130(6), 849–867.

Rieseberg, L.H., & Soltis, D.E. (1991). Phylogenetic consequences of cytoplasmic gene flow in plants. *Am. J. Bot.*, 5, 565–84.

Salichos, L., & Rokas, A. (2013). Inferring ancient divergences requires genes with strong phylogenetic signals. *Nature*, 497, 327–331.

Scheunert, A., & Heubl, G. (2017). Against all odds: reconstructing the evolutionary history of *Scrophularia* (Scrophulariaceae) despite high levels of incongruence and reticulate evolution. *Org. Divers. Evol.*, 17, 323–349.

Schwartz, R.S., & Mueller, R.L. (2010). Branch length estimation and divergence dating: estimates of error in Bayesian and maximum likelihood frameworks. *BMC Evol. Biol.*, 10, 5.

Sepulchre, P., Ramstein, G., Fluteau, F., Schuster, M., Tiercelin, J.J., & Brunet, M. (2006). Tectonic uplift and Eastern Africa aridification. *Science*, 313(5792), 1419–1423.

Smissen, R.D., Galbany-Casals, M., & Breitwieser, I. (2011). Ancient allopolyploidy in the everlasting daisies (Asteraceae: Gnaphalieae) - complex relationships among extant clades. *Taxon*, 60, 649–662.

Smith, M.L., & Hahn, M.W. (2021). New Approaches for Inferring Phylogenies in the Presence of Paralogs. *Trends Genet.*, 37(2), 174–187.

Smith, S.A., & O'Meara, B.C. (2012). treePL: divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics*, 28(20), 2689–2690.

Springer, M.S., & Gatesy, J. (2016). The gene tree delusion. *Mol. Phylogenetics Evol.*, 94(Pt A), 1–33.

Tabatabaee, Y., Claramunt, S., & Mirarab, S. (2025). Coalescent-based branch length estimation improves dating of species trees. *bioRxiv*, 2025.02.25.640207.

Tamura, K., Tao, Q., & Kumar, S. (2018). Theoretical Foundation of the RelTime Method for Estimating Divergence Times from Variable Evolutionary Rates. *Mol. Biol. Evol.*, 35(7), 1770–1782.

Thomas, A., Meudt, H.M., Larcombe, M.J., et al. (2023). Multiple origins of mountain biodiversity in New Zealand's largest plant radiation. *J. Biogeogr.*, 50, 947–960.

Ufimov, R., Gorospe, J.M., Fér, T., et al. (2022). Utilizing paralogues for phylogenetic reconstruction has the potential to increase species tree support and reduce gene tree discordance in target enrichment data. *Mol. Ecol. Resour.*, 22(8), 3018–3034.

Van de Peer, Y., Ashman, T.L., Soltis, P.S., & Soltis, D.E. (2021). Polyploidy: an evolutionary and ecological force in stressful times. *Plant Cell*, 33(1), 11–26.

Velásquez-Puentes, F.J., Torke, B.M., Barratt, C.D., et al. (2023). Pre-adaptation and adaptation shape trait-environment matching in the Neotropics. *Glob. Ecol. Biogeogr.*, 32, 1760–1772.

Wan, J.N., Wang, S.W., Leitch, A.R., et al. (2024). The rise of baobab trees in Madagascar. *Nature*, 629, 1091–1099.

Westerhold, T., Marwan, N., Drury, A.J., et al. (2020). An astronomically dated record of Earth's climate and its predictability over the last 66 million years. *Science*, 369(6509), 1383–1387.

Wiens J.J. (2004). Speciation and ecology revisited: phylogenetic niche conservatism and the origin of species. *Evolution*, 58(1), 193–197.

Wiens, J.J., Ackerly, D.D., Allen, A.P., et al. (2010). Niche conservatism as an emerging principle in ecology and conservation biology. *Ecol. Lett.*, 13, 1310–1324.

Wiens, J.J., & Graham, C.H. (2005). Niche Conservatism: Integrating Evolution, Ecology, and Conservation Biology. *Annu. Rev. Ecol. Evol. Syst.*, 36, 519–539.

Yan, Z., Smith, M.L., Du, P., Hahn, M.W., & Nakhleh, L. (2022). Species Tree Inference Methods Intended to Deal with Incomplete Lineage Sorting Are Robust to the Presence of Paralogs. *Syst. Biol.*, 71(2), 367–381.

Yoder, A.D., & Nowak, M.D. (2006). Has vicariance or dispersal been the predominant biogeographic force in Madagascar? Only time will tell. *Annu. Rev. Ecol. Evol. Syst.*, 37(1), 405–431.

Zhang, C., & Mirarab, S. (2022). ASTRAL-Pro 2: ultrafast species tree reconstruction from multi-copy gene family trees. *Bioinformatics*, 38(21), 4949–4950.

Zhang, M.-H., Nie, Z.-L., Fairbanks, R.A., et al. (2025). Phylogenomic insights into species relationships, reticulate evolution, and biogeographic diversification of the ginseng genus *Panax* (Araliaceae), with an emphasis on the diversification in the Himalayan-Hengduan Mountains. *J. Syst. Evol.*, 63(1), 99–114.

Zhang, C., Scornavacca, C., Molloy, E.K., & Mirarab, S. (2020). ASTRAL-Pro: Quartet-Based Species-Tree Inference despite Paralogy. *Mol. Biol. Evol.*, 37(11), 3292–3307.

Conclusions

Conclusions

1. General phylogenetic insights: the analysis of over 500 taxa using Hyb-Seq in combination with biogeographic, morphological and ecological data reveals that *Helichrysum* is not monophyletic but paraphyletic and polyphyletic. Both concatenated and coalescent-based phylogenies yield similar results, except for the clades with a potential allopolyploid origin. These findings support the adoption of the term "HAP clade" to refer to *Helichrysum* and at least eight other nested genera. A recircumscription of the genera integrating the HAP clade should be considered and further studies employing appropriate methodologies are needed to disentangle the evolutionary relationships of allopolyploid lineages. Additionally, several southern African species fall outside the HAP clade, being more closely related to other Gnaphalieae genera, further reinforcing the need for a taxonomic reassessment.

2. Origins and early expansion: *Helichrysum* originated during the Miocene in the arid lowland regions of southern Africa. It rapidly expanded into the Drakensberg montane grasslands, becoming a predominantly high-elevation clade. This expansion coincided with a period of increasing aridification and cooling that fostered the rise of open habitats and grasslands, which are the genus' preferred habitats.

3. The main source region: the southern African montane grasslands were the primary center of diversification and the main source of lineages of *Helichrysum*. From this region, lineages spread across low- and high-elevation habitats in southern Africa, Tropical Africa and Madagascar.

4. Speciation in the Tropical Afromontane region: the Tropical Afromontane region was colonized multiple times independently by ancestors from the southern African grasslands. In some cases, colonization was followed by in-situ diversification, producing lineages with endemic species confined to specific mountain ranges or areas, while in others, dispersal resulted in range expansions, with widespread species inhabiting both tropical and southern African montane areas.

5. Colonization and adaptation to the Tropical Afroalpine region: colonization of the Tropical Afroalpine region occurred independently on several occasions, primarily from the adjacent Tropical Afromontane areas rather than from alpine-adapted ancestors from the southern African Drakensberg or Eurasia. These events reflect repeated upward migrations, followed by in-situ adaptation to extreme alpine conditions.

6. Colonization of Madagascar: *Helichrysum* colonized Madagascar at least six times independently from the Pliocene onwards. Ancestral biome conservatism played a key role in early colonization stages: Tropical Afromontane ancestors established in higher elevations, southern African grassland ancestors thrived in the Central Highlands, and arid-adapted ancestors from southern Africa settled along the island's arid southwestern coast. In general, the descendant Malagasy species retained ancestral morphological traits, with many extant species highly resembling their closest continental relatives.

7. Speciation in Madagascar: following biome-driven colonization and establishment on the island, in-situ ecological speciation became relevant together with allopatric speciation at a regional scale. Both processes played a key role in driving two large parallel Pleistocene mountain radiations, characterized by multiple downward migrations associated with shifts in bioclimatic preferences and morphology, the origin of local mountain endemics and the origin of several species adapted to particular microhabitats such as wetlands and rock crevices.

8. Morphological and molecular coherence: the phylogenetic relationships inferred from molecular data are highly consistent with the latest taxonomic classification of Malagasy *Helichrysum* (Humbert 1962) for most taxa and groups, although there are some exceptions. This coherence underscores the value of morphological and ecological data as complementary tools for phylogenetic inference, encouraging an integrative approach to Systematics.

9. Mountain systems as sinks and sources of diversity: frequent long-distance dispersal among distant mountainous regions reveals a pattern of nested radiations, in which multiple radiations are interconnected. The clade initially diversified in the mountains of southern Africa before expanding upwards and spreading beyond Africa. This expansion led to multiple parallel radiations across mountain systems worldwide, including Madagascar, Eurasia and the Americas. Additionally, downward migrations along the elevational gradient, implying the colonization of lowland, mainly temperate, habitats by high-elevation ancestors, suggest high elevational permeability.

10. The impact of Pleistocene climatic fluctuations: our dataset suggests a nearly fourfold increase in net diversification rates around 1.2 Ma, coinciding with the emergence of the youngest mountain radiations. Therefore, Pleistocene climatic oscillations likely played a key role in shaping the clade's species distribution and diversification dynamics. Periods of habitat expansion probably facilitated dispersal and secondary contacts among populations, while habitat contractions likely promoted allopatric speciation by isolating populations at regional and local scales in mountainous regions.

11. The role of ecological speciation in diversification: following colonization and successful establishment in distant regions, the clade's high adaptability and niche lability played a key role in microhabitat specialization. This is supported by the sympatry of species with different habitat preferences, the strong correlations between pairs of functional traits and habitat preferences, as well as the repeated independent acquisition of such traits, which enabled the establishment in specific environments. However, the acquisition of traits potentially considered key innovations for high-elevation conditions does not appear to directly drive higher diversification rates in mountainous environments.

12. Flexible evolutionary strategies: the diversification history of the HAP clade reflects a complex interplay of geographic, ecological and allopolyploid speciation processes operating at different temporal and spatial scales. Its evolutionary success likely stems from its high dispersibility and adaptability, which together help explain its contribution to mountain biodiversity. These enabled the colonization of remote geographic regions and led to a high morphological and ecological diversity within the clade. All these factors provide insights into the clade's current global distribution and remarkable diversity across both low- and high-elevation environments.