Review

The Andes through time: evolution and distribution of Andean floras

Oscar Alejandro Pérez-Escobar, 1,16, Alexander Zizka 2,3,16, Mauricio A. Bermúdez 4, Andrea S. Meseguer, 5, Fabien L. Condamine 6, Carina Hoorn 7, Henry Hooghiemstra, 7, Yuanshu Pu, Diego Bogarín, 8,9 Lydian M. Boschman 8,10 R. Toby Pennington, 11,12 Alexandre Antonelli 1,13,14 and Guillaume Chomicki 15,16,,*

The Andes are the world’s most biodiverse mountain chain, encompassing a complex array of ecosystems from tropical rainforests to alpine habitats. We provide a synthesis of Andean vascular plant diversity by estimating a list of all species with publicly available records, which we integrate with a phylogenetic dataset of 14,501 Neotropical plant species in 194 clades. We find that (i) the Andean flora comprises at least 28,691 georeferenced species documented to date, (ii) Northern Andean mid-elevation cloud forests are the most species-rich Andean ecosystems, (iii) the Andes are a key source and sink of Neotropical plant diversity, and (iv) the Andes, Amazonia, and other Neotropical biomes have had a considerable amount of biotic interchange through time.

The multiple facets of Andean floras

The Andes are thought to contain ~10% of the world’s vascular plant diversity (30,000 species) in only 0.6% of its land surface [1]. With only 25% of the original vegetation remaining, the Andes are the world’s most species-rich plant biodiversity conservation hotspot [2]. The Andean mountains played a pivotal role in generating the biodiversity that colonized various regions of the Neotropics across timescales, notably contributing to the rich plant diversity of Amazonia and Central America [3–7]. Andean ecosystems also provide livelihoods and essential ecosystem services, sustaining millions of people [8]. Despite this, research on the evolution of Andean plants has been sporadic. Three major factors hinder our understanding of the origin and evolution of the Andean flora: (i) insufficient or incomplete knowledge of the Andean orogeny, with sometimes conflicting hypotheses [9,10], (ii) poor understanding of plant species richness and plant distribution patterns across the Andes, largely because of insufficient floristic surveys [11], and (iii) the scarcity of genetic data and time-calibrated phylogenies for most Andean lineages [12].

Our synthesis has five major aims: (i) to review the geological history of the Andes throughout its entire range to inform biological research, (ii) to estimate plant species diversity across the Andes, (iii) to synthesize our understanding of the species richness and ages of Andean ecosystems, and (iv) to use our new estimate of Andean plant diversity to dissect the migration routes of Andean plants, as well as (v) to identify priority groups for which few sequence data are available. Our work reveals key knowledge gaps which can inform future research and conservation work in the Andes.

Geological history of the Andes

The Andes extend over 7000 km in South America from ~10°N to 50°S. This mountain range was formed as a result of subduction (see Glossary) of the oceanic Nazca and Caribbean plates under the South American continental plate. The South American subduction zone is one of

Highlights

We present an evolutionary and floristic synthesis of Andean plant diversity and evolution across time and space.

Uplift of the Andes varied across time and space. Particularly, the fast uplift rates between 8 and 5 Ma in the Northern Andes may have favoured plant diversification.

Using online specimen databases, we suggest that the Andean flora comprises at least 28,691 species. We identify North Andean montane forests as the potential species richest area.

Using a biogeographic analysis on a dataset of 14,501 Neotropical species in 194 clades, we reveal that the Andes are both a key source and sink of Neotropical vascular plant biodiversity. We unveil strong biogeographical links between the Andes, Amazonia, and Central America.

We highlight a number of critical research gaps, notably major Andean plant groups are still understudied, and fewer studies exist for the Central and Southern Andes. Filling these gaps will allow a more holistic understanding of Andean floras and provide essential tools for their conservation.

1Royal Botanic Gardens, Kew TW9 3AB, Surrey, UK
2Biodiversity of Plants, Philipps University Marburg, 35043 Marburg, Germany
3German Center for Integrative Biodiversity Research Halle-Jena-Leipzig (iDiv), 04103 Leipzig, Germany
4Escuela de Ingeniería Geológica, Universidad Pedagógica y Tecnológica de Colombia, Tunja, Colombia
The rise of the Andes as a gradual process from the Eocene (40 Ma) onwards [10,28–32]. Since this time the Caribbean plate has been attached to South America, while moving toward the northwestern corner of the South American Plate at ~100 Ma. Collision of the trailing edge of the Caribbean Plate, at ~80–70 Ma, led to initiation of uplift in the Ecuadorian Andes, and since this time the Caribbean plate has been attached to South America, while moving toward its present-day position [21].

The limits of the Northern and Central Andes are mainly shaped by the complex configuration of the Nazca and Caribbean plates, by changes in the slope of subduction, and by interactions with precursor plates (i.e., Farallón and Phoenix) [14] (Figure 1A). By contrast, the Southern Andes are delineated by the interaction of the Antarctic, Scotia, and South American plates (e.g., [15–18]). The boundary between the Northern and Central Andes is marked by subduction of the Carnegie Ridge, a high on the Nazca Plate, which dives under the South American Plate in Ecuador. This geological phenomenon is geographically expressed by a depression across the Andes, known as the Huancabamba Depression [19,20]. In addition, the northern Andes have been shaped by interactions with the Caribbean Plate, of which the leading edge collided with the northwestern corner of the South American Plate at ~100 Ma. Collision of the trailing edge of the Caribbean Plate, at ~80–70 Ma, led to initiation of uplift in the Ecuadorian Andes, and since this time the Caribbean plate has been attached to South America, while moving toward its present-day position [21].

The Andean orogeny is the subject of intense study, and some issues remain contentious – including the timing, pace, and sequence of mountain building. Reconstructing mountain building is challenging because the geological record does not provide a direct measure of past elevation, paleo-altimetry methods contain large uncertainties and many caveats [22], and continental ranges are subject to erosion which results in a highly incomplete rock record [23]. Data from the sedimentary basins east of the Andes indicate that uplift in the Southern Andes started at ~100 Ma, in the Northern Andes at ~80 Ma, and in the Central Andes at ~70 Ma [24] (Figure 1C–K). These ages in line with exhumation ages, which are oldest for the Southern Andes (Campanian–Paleocene, 75–55 Ma) and younger for the Central and Northern Andes [5,23]. During those 100 million years of mountain building, however, uplift has not been constant or uniform across time and space, and there is debate in particular about the uplift history of the eastern domains of the Northern and Central Andes (the Eastern Cordillera of Colombia and the Altiplano). Some studies have presented evidence for remarkably rapid uplift during the Miocene [25,26], which, for the Northern Andes, has been associated with fast species diversification [6,27]. By contrast, other researchers regard the rise of the Andes as a gradual process from the Eocene (40 Ma) onwards [10,28–32].

A recent reconstruction of Andean mountain building, integrating paleo-altimetry data from 36 separate geomorphological domains across the Andes, shows that each of these domains has an independent history of surface uplift, and that uplift of the Andes has thus been a highly diachronous process [23]. The reconstruction shows that, since the Late Cretaceous, uplift generally migrated from the coastal and western cordilleras eastwards – toward the central and eastern cordilleras and sub-Andean zone (Figure 1C–K). Whereas uplift in the coastal and western cordilleras is generally old, slow, and constant, the central and eastern cordilleras, large parts of the Northern Andes, and the Altiplano all uplifted through young and rapid orogenesis with acceleration phases in the Oligocene and Miocene [9,25,33,34]. Most importantly, this reconstruction shows that drawing generalized conclusions about the history of uplift in the Andes as a whole is not warranted. We used this model [23] to present the main phases of Andean uplift (Figure 1C–K). In addition, we present a map of apatite fission track (AFT) ages that reveal the cooling ages of Andean rocks across its range (Figure 1B), which may
Glossary

Apatite fission track (AFT): a radiometric dating technique based on analyses of the damage trails, or tracks, left by fission fragments in particular uranium-bearing minerals and glasses such as apatite. The ages record the timing of cooling of the rocks on their journey from deep in the Earth toward the surface (i.e., exhumation). Assuming that exhumation is the result of uplift and erosion, AFT ages can be used to date mountain building. This assumption is not always warranted because cooling can be the result of many other tectonic processes, especially in active volcanic arcs. However, in non-volcanic regions, compilations of AFT ages may give a general overview of the timing of mountain building.

Diachronous: occurring in different geological periods.

Exhumation: the process by which rocks (that were formerly buried) approach the Earth’s surface.

Farallón and Phoenix: tectonic plates that existed in the Pacific Ocean during the early Paleozoic through to the late Cenozoic.

Flickering connectivity: a paleoecological model which posits that the contraction and expansion of distribution areas, as well as the connection and isolation of gene pools (vegetation) during glacial–interglacial cycles, are key drivers of diversification in the Andes.

High: an elevated topography that stands out from the rest of the area.

Paleo-altimetry: reconstruction of past elevations.

Night frost: temperature dipping below 0°C at night, occurring in the Andean Upper Montane Forests. In some areas of the range (varying with latitude and elevation) there are regular night frosts, whereas in other areas there are occasional frosts with low occurrences (sometimes every 10 years or so).

Subduction: geological process where the oceanic lithosphere of a tectonic plate plunges under the lithosphere of a second plate, either continental or oceanic.
generally be associated with uplift. Young AFT ages can be seen across the Northern Andes, mirroring recent uplift. Nevertheless, the whole range, and the Central and Southern Andes in particular, show interspersions of older and younger age (Figure 1B). This confirms that the timing and rate of Andean uplift have been highly uneven across its range. This new insight conflicts with what is often modeled in macroevolutionary studies attempting to link plant species diversification rate with Andean uplift [6,27,35]. Thus, future diversification models implementing Andean elevation as a time-dependent variable should avoid relying on a single uplift curve produced for an entire Cordillera, and should instead consider uplift heterogeneity as a function of species occurrences, whenever biological resolution allows [36].

The Andean orogeny has affected regional climate, hydrological conditions, nutrient cycling, landscape development, and thus potential plant evolution mechanisms at the continental scale. In the Northern and Central Andes, uplift increased rainfall east of the mountain range (and established a rain shadow with dry conditions in the west) and sediment flux into Amazonia [37–39]. This resulted in the current configuration of the Amazon drainage basin with precursors such as the Pebas and Acre depositional systems [5,40] and in the establishment of the “South American Dry Diagonal” consisting of the Caatinga, the Cerrado, and the Chaco biomes (e.g., [41,42]). It also led to the formation of an orographic rain shadow on the foothills of the Central and Southern Andes [43,44] from late Miocene (~11 Ma) onwards. As for the latter phenomenon, AFT data have revealed swift mountain uplift in the past 8–5 Ma in the Northern Andes (the Cocuy area of the Eastern Cordillera in particular [45–48]), but less so in the Central and Southern Andes. This scenario is supported by dated phylogenies from various plant groups showing young ages and rapid diversifications in the Northern Andes, but older ages in Central and Southern Andes [4,5,35,49,50]. Another possible explanation for this pattern is that erosion in the tropical Andes could have been substantially higher than in the Southern Andes (Figure 1A), where more extensive ice caps would have slowed erosion [38,51,52].

Thus, three take-home messages on Andean orogeny should be carefully considered in future studies of plant diversification and biogeography in the Andes: (i) Andean uplift was highly diachronous, starting in the Southern Andes at ~100 Ma, in the Northern Andes at ~80 Ma, and subsequently in the Central Andes at ~70 Ma. (ii) Uplift in the coastal and western cordilleras was generally old, slow, and constant, but the central and eastern cordilleras, large parts of the

---

**Figure 1. Geology and uplift history of the Andes.** (A) Geological map highlighting the three main Andean regions (Northern, Central, Southern) and the main Quaternary volcanic zones and tectonic processes. Adapted, with permission, from Ramos [17]. (B) Apatite fission tracks (AFTs) across the Andes, showing the cooling ages of the rocks (not their uplift) across the range. Red lines represent plate boundaries (convergent or subduction, line with red triangles; divergent and transform, red lines without symbols); white lines in the sea are seafloor fabric and magnetic lineations. Black lines in the continent correspond to country boundaries. (C–K) Reconstruction of Andean paleoelevation; adapted, with permission, from Boscman [23]. (C) At 80 Ma, the majority of the north-western edge of South America was still below sea level, the Central Andes were characterized by moderate topography in the continental margin arc, and the Southern Andes have already uplifted gradually since 100 Ma. (D) At the Cretaceous–Paleogene boundary (66 Ma), uplift has started in the Northern Andes (NAN) and is occurring in the Western Cordillera (WC) of the Central Andes. Large marine foreland basins cover the interior of the continent, and marine incursions flood in the Neuquén, San Gorge gulf, and Austral basin in the southeast of South America. (E) At 55 Ma, Western Patagonia has reached modern elevations. (F) By 34 Ma, uplift in the Northern Andes migrates northward, and uplift in the Central Andes (CAN) migrates eastward, resulting in the rise of local ranges in the Eastern Cordillera (EC). (G) At the start of the Miocene (23 Ma), marine incursions flood the San Gorge gulf and Austral basin again. (H) At 15 Ma, uplift in the Northern Andes continues to migrate north, and from now on also migrates east. Uplift in the Central Andes is focused in the Eastern Cordillera, and the foreland basins become briefly marine again. In the Southern Andes the orogenic front in Patagonia migrates east. (I) In the late Miocene uplift is focused in the Northern Andes and the eastern cordilleras and sub-Andean zone of the Central and Southern Andes, and the Altiplano experiences a particularly rapid phase of uplift. (J) The final stages of uplift are predominantly in three Andean regions: the Eastern Cordillera and the Mérida Andes in the Northern Andes, the sub-Andes and the Sierras Pampeanas in the Central Andes, and Eastern Patagonia in the Southern Andes. (K) Present-day elevation.
Northern Andes, and the Altiplano uplifted through young and rapid orogenesis with acceleration phases in the Oligocene and Miocene. (iii) This argues against using a single uplift curve in a diversification or biogeographic context.

**The Andean floras: their distribution, richness, and relationships**

To gain insights into the biotic assembly, diversity, and distribution of Andean floras, we investigated Andean plant species diversity using global distribution databases, and generated a working list of Andean vascular plants (see Materials and Methods in the supplemental information online) based on the list of Neotropical plants of Ulloa et al. [53], GBIF global distribution databases’ and taxonomic expertise. We identify 28,691 tentative Andean vascular plant species, defined as species currently occurring in the Andean cordillera at an elevational range between 100 and 6086 m. We suggest that this may be an underestimate; even if some species are lumped taxonomically in future, the Andes may house other species that have not yet been digitized and georeferenced, and others remain to be scientifically described.

The elevational delimitation of the Andes is contentious [54], and a multitude of studies rely on different elevational ranges starting at 100, 500, and 1000 m [55–57]. To assess the robustness of our elevational delimitation, we compiled additional lists of Andean species with elevation ranges starting at 500 and 1000 m (instead of 100 m) to 6086 m and found a difference of 3–20%, respectively, between the species richness reported when using a lower altitudinal bound of 100 m. This shows that the ‘lowland’ (100–500 m) and the ‘premontane’ (500–1000 m) intervals share many species, and that there is more floristic difference at elevations greater than 1000 m, consistent with previous biome reconstructions using pollen fossil data [58].

The Andean flora is a highly uneven assemblage of the plant tree of life. Only 10 plant families (Orchidaceae, Asteraceae, Leguminosae, Rubiaceae, Melastomataceae, Bromeliaceae, Piperaceae, Solanaceae, Araceae, and Poaceae) make up about half of all Andean plant species, while 226 plant families account for the remaining Andean plant diversity (Figure 2A and see Dataset S1 in the supplemental information online). The top 10 families in numbers of species are the same across the Andean elevation gradient up to >2000 m, but show turnover at >3000 m and >4000 m, where 30% and 50% of the families change, respectively, and four of the top 10 families are exclusive to the high elevation flora above 4000 m (Figure 2A). A suggested hyper-dominance of a reduced number of families on the diversity of Andean plants was first noted by Cuatrecasas [54], and later by Gentry [59], but a comparison of the 10 most species-rich families of the Neotropical plant list [53] and Neotropical dry forests [60] show a similar pattern where 10 dominant families account for half of the diversity, suggesting that this pattern is not specific to the Andean flora.

The classification, distribution, and diversity of such a rich array of Andean ecosystems have been investigated for decades. Numerous systems have been proposed mostly based on their altitudinal position, climatic characteristics, and floristic associations [54,61,62]. Nevertheless, which Andean ecosystems are the most species-rich and the similarities of the diversity they share remain open questions (see Outstanding questions). Gentry [59] suggested that Andean plant diversity is mostly concentrated in the Northern Andes, a geologically discrete section of the cordillera that hosts a wide diversity of vegetation types [62–64]. Our review is in line with these results, pointing to the hotspots of Andean vascular plants in the Northern Andes (Figure 2B). However, this pattern correlates with the number of collections and thus sampling effort, which likely bias the real floristic contribution of other regions (see Outstanding questions). Colombia has the highest number of Andean plants (10,932 species), whereas Ecuador (8,897 species) comes first when dividing species number by country area. Andean vascular plant diversity
shows a classic latitudinal gradient where species richness is highest at low latitudes (Northern Andes), and lowest at high latitudes (Southern Andes), with fluctuations at low latitudes (Figure 2C). Such oscillations in the Andes are likely the result of the superimposition of another conspicuous pattern of diversity, the altitudinal gradients, with a peak of diversity at mid-elevations (~1500 m) (Figure 2D), as first identified by Gentry [59].

To delve into the variation of species richness and connection between the distinct Andean floras, we used the ecoregions delineation adopted by World Wide Fund for Nature (WWF)iii (Figure 3A, B). This revealed that the Northern Andean montane forest is by far the richest environment, and that it shares many of its species with both Páramos and Central Andean Yungas (Figure 3A). The species richness of North Andean montane forests is especially striking given its small area (Figure 3B). The most species-poor environments scaled to their area include the Patagonian steppe and the Low Monte, both of which reach the Southern Andean foothills and have relatively low connectivity with other Andean floras, in addition to being conspicuously dry (Figure 3A,B).

Our new list of Andean plants allows quantification of which taxonomic groups are the least known. To identify potential DNA sequencing gaps in the Andean flora, we searched the US National Center for Biotechnology Information (NCBI) GenBank repositoryiv relying on widely used DNA markers in phylogenetic studies (see the supplemental information online). We found that only 27% of the 226 families, 79% of the 2537 genera, and overall 33% of the species have publicly available DNA sequences. Focusing on the eight families with the largest number of Andean plants, we found that species with available DNA sequences range from 17.4% (Orchidaceae) to 65.4% (Solanaceae). These sequencing gaps are priorities for future research on Andean plants (see Outstanding questions).

The assembly of Andean floras through space and time
To gain insights into the assemblage of Andean floras through time, we reviewed phylogenetic studies, expanding the framework of Luebert and Weigend [50] to works published up to April 2021, and evaluated the Andean plant fossil record (Box 1). We identify three emerging patterns regarding the origin of Andean floras. The emerging patterns are based on 37 studies – some of which include cross-taxonomic analyses – all cited below.

First, high-elevation Páramo taxa are relatively young and have diversified rapidly. Páramo is an alpine grassland with >3400 species, most of which are endemic [65]. The iconic Páramo-endemic genus *Espeletia* (Asteraceae) evolved – with key adaptations including pubescent leaves and persistent rosette leaves protecting the stem and water-storing pith – only at the onset of the Quaternary (2.58 Ma), according to phylogenetic data, followed by rapid diversification (with up to 3.1 speciation events per lineage and million years; [49,66,67]). However, the pollen fossil record suggests that it most possibly evolved in the Pliocene (5–4 Ma), and shows that the Páramo flora was in strong development at ~2.25 Ma [68] (Box 1). Other groups that, according to phylogenetic data, diversified rapidly and recently in the Andean alpine environment include the lupines (Leguminosae [35,69,70]), a clade of ~90 *Hypericum* species (Hypericaceae) [71], and high Andean *Astragalus* [72]. Cross-taxonomic analyses are consistent with recent ages of Páramo [5,49].
Figure 3. Distribution of species diversity across Andean ecosystems and their similarities. (A) Species richness and connectivity of Andean ecoregions as defined by the World Wide Fund for Nature. The size of the circle is directly proportional to the species richness reported for each ecoregion, whereas the thickness of the arrow represents species richness shared between two ecoregion. (B) Species richness of Andean ecoregions and their corresponding area sizes. Abbreviations: N.A., northern Andean; C.A., central Andean; S.A., southern Andean.

Box 1. Fossil Andean floras

The plant fossil record should provide evidence for both the history and turnover of floras among the major Andean regions. Using the Paleobiodb database, we generated a list of Andean plant fossils, and only retained Cenozoic records because they are the best-curated (see Materials and methods and Dataset S3 in the supplemental information online). Our compilation provides four important findings. However, it is important to note that the Andean fossil record is spatially biased, and systematically incomplete, owing to the geological activity and erosion in the region.

First, our database shows that most Cenozoic Andean fossils are from the Northern Andes, pointing to a lower number of studies in the Central and Southern Andes, or less digitization (Figure I).

Second, the Andean plant fossil record supports the presence of both humid tropical forest and dry forest in the Andes for ~60 million years. Humid tropical forest taxa show records from the Paleocene onwards, confirming the idea that the origin of the tropical forest biome is old, and has existed at least since the early Cenozoic [97–101]. By comparing extensive fossil sequences, Carvalho et al. [101] identified more open canopy forests in the late Cretaceous of Colombia, and further pointed to substantial turnover at the Cretaceous–Tertiary boundary.

Third, the pollen fossil record of high-elevation Páramos, dating to 5 Ma [102], is consistent with its recent origin and fast diversification [69]. Additionally, this could mean that Espeletia, and potentially other high-altitude taxa, had an early diversification, followed by lineage extinctions and replacement once considerably colder conditions came into force with the onset of the Northern Hemisphere glaciations – as proposed by Silva et al. [103]. Puna-like ecosystems (high-elevation dry grasslands) were present as early as the Pliocene in the Central Andes [104].

Fourth, the Andean fossil record reveals plant assemblages that apparently lack extant analogs. For instance, in the Central Andes, a Miocene forest contained an intricate mix of both common montane taxa such as Podocarpus or Hedyscomum, and high-elevation taxa including Polyalthis and Valeriana, as well as plants typical of lowland ecosystems such as large legume trees and palms [104].
Páramos have been dynamic environments shifting over an elevational range of ~1500 m through Pleistocene glacials and interglacials, moving between ~2000–3500 m and ~3400–4900 m) [73]. This implied that Páramos were recurrently connected and disconnected over time. This flickering connectivity mediated by glacial–interglacial cycles has likely been a key driver for the diversification of Páramos plants, probably by facilitating allopatric speciation and secondary contact (sympatry) [39,64,74,75]. During glacial maximum conditions C4 plants were more abundant in the Páramo (linked to the low atmospheric partial pressure of carbon dioxide, pCO2), whereas today Páramo is dominated by C3 plants [76].

Second, seasonally dry Andean forests at lower elevations appear to show the opposite pattern—older groups that diversified slowly. Such florae appear to have been assembled gradually over the past ~20 Ma [77–79]. Seasonally dry forests comprise a diverse array of vegetation types, including tall forest on moister sites to cactus scrub on the driest parts [80]. Smaller in stature than a rainforest, seasonally dry forests are characterized by strongly seasonal ecological processes where many species flower synchronously at the transition between the dry and the wet seasons while still leafless [81]. These forests occupy inter-Andean valleys, and plant taxa show a high level of isolation. For instance, *Cyathostegia mathewsii* (Leguminosae), a shrub endemic to dry
Andean forests, shows that populations separated by only 600 km have been isolated for at least 5 Ma [82]. Andean seasonally dry forests are highly isolated, and climatically similar forests are found scattered in the Neotropics. Historical connectivity amongst these disjunct patches may be further limited because plants moving into these environments also require specific adaptations to drought to survive [82]. The high level of isolation of these forests is illustrated by Northern and Central inter-Andean valleys that have almost no floristic overlap in species [59].

However, there are some exceptions to this pattern. In Colombia, there is lower species endemism in inter-Andean dry forests than in Ecuador and Peru, and this may reflect the fact that they are not isolated from the Caribbean coastal dry forests by high cordilleras [83]. Similarly, the Huancabamba depression appears to provide a corridor for the migration of dry forest plants between inter-Andean valleys on the Pacific coast because of its relative low altitude [84]. These enclaves shrunk to the very bottom of the valleys during interglacial conditions, reinforcing their isolation [85]. By contrast, glacial conditions were optimal for lowland Andean seasonally dry forests because of the low pCO2 and dry atmosphere. However, such expansion did not drive the type of ‘flickering connectivity’ dynamics that occurred in the Páramos because it did not affect the connectivity of dry forests from distinct valleys [77,82].

Third, Andean cloud forests are the most speciose environments in the Andes (Figure 3A,B). Andean cloud forests occur from ~1200 m to the upper forest line at 3200–3500 m of elevation [86]. They are characterized by high humidity levels that are similar to lowland rainforests, but with lower temperatures. They comprise mostly evergreen small trees, shrubs, and epiphytes—termed Andean-centered taxa by Gentry [59]. Because these intermediate elevation areas connect the low elevation ‘Amazonian’ floras to the alpine floras in the Central and Northern Andes, one might expect them to show a patchwork of evolutionary histories with both fast and slow diversification rates. We find that the majority of cloud forest lineages have diversified from the early Miocene onwards [50], when cloud forests are thought to have first appeared [87]. Andean cloud forests have the highest level of vascular epiphyte diversity in the Neotropics [88]. As expected, various cloud forest lineages—including several key epiphytic lineages—show fast diversification rates [6,27,89–91]. However, certain plant groups such as Begonia show a pattern of diverse colonization of cloud forests without subsequent rapid diversification [92], whereas others, such as Cyatheaceae tree ferns, show slow diversification [93].

We suggest that these species richness and diversification patterns result from rapid in situ diversification and frequent immigration events. Andean cloud forests can be separated into Lower Montane Forest (LMF; ca 1200–2300 m, depending on latitude) and Upper Montane Forest (UMF; ca 2300 up to the Upper Forest Line at 3200–3500 m, depending on latitude). The latter is environmentally mainly separated by the phenomenon of night frost. Although LMF species cannot normally resist night frost, UMF species can. The LMF, in particular, includes a condensation zone (cooling of ascending air masses) which decreases the chance of night frost and has given rise to a rich epiphyte flora. In contrast to the Páramo, the UMF and LMF have not suffered fragmentation and had almost continuous connectivity during glacial cycles [39,74]. Moreover, this high connectivity of Andean cloud forests—a crossroads of lowlands and higher elevation areas—is likely crucial for both the immigration and accumulation of lineages [94].

Origins and main migration routes of Andean Plants
By what routes did Andean floras assemble, and how did the Andes contribute to other Neotropical floras? A recent analysis of 4450 animal and plant species revealed that Amazonia has contributed over 2800 lineages to other Neotropical regions, making it the primary source of Neotropical biodiversity [4]. Nevertheless, the assembly of Amazonia is tightly linked to the uplift
of the Andes [5], and phylogenetic studies have revealed numerous interchanges in biodiversity between Amazonia and the Andes [4,6,95], as suggested by earlier floristic work [59]. These analyses also revealed that species colonized Andean grasslands from Amazonia, Central America, Cerrado and Chaco, and the Patagonian Steppes [4].

To quantify the number of transitions into and out of the Andes, we relied on a large-scale vascular plant dated phylogeny [96] and used our working Andean plant list described above to estimate the biogeographic dispersal events from other Neotropical regions to the Andes, and vice versa (see Materials and methods in the supplemental information online). We also retrieved a dataset of 329 536 georeferenced records for 89 736 species, and we were able to match 14 501 species between the occurrence records and the phylogeny. Using a recently developed approach [4], we identified 194 clades in which at least 85% of the species occur within the Neotropics, resulting in clade sizes ranging between 9 and 100 species. We then extracted all clades that showed at least one bioregion shift (171 clades) and used ancestral range estimation

Figure 4. Vascular plant interchange across Neotropical regions. The analyses are based on 194 clades representing 14 501 Neotropical vascular plant species (see Materials and methods in the supplemental information online). (A) Map of the Neotropical regions considered in this study. Red arrows show shifts involving at least one Andean region; turquoise arrows show shifts not involving the Andes. (B) Vascular plant interchanges in the Andes (including only interchanges involving Andean taxa). Arrows indicate the direction and number of dispersal events; line thickness is proportional to the number of events. Orange arrows show shifts toward an Andean region, and blue arrows show shifts away from an Andean region. Selected Andean ecoregions are shown: (C) Northern Andean cloud forests, (D) Central Andean Yungas, (E) Tumbes Andean dry forest, (F) Páramo, (G) Patagonian steppe. Photos: O. Pérez-Escobar.
implemented in the dispersal–extinction–cladogenesis model to estimate the number of shifts among regions (see Materials and methods and Dataset S2 in the supplemental information online).

We identified 1795 shifts across Neotropical bioregions, including 483 into or out of the Andes. We found that the Northern Andes showed more interchanges than any other Andean region (215 in vs 235 out), followed by the Central Andes (127 in vs 117 out), and the Southern Andes (12 in vs five out) (Figure 4A,B). Overall, the Andes are the third most important biogeographical source and sink of vascular plant interchange after Amazonia and the Atlantic Forest. However, the Andes become second, immediately after the Atlantic Forest but before Amazonia, when normalized by area. This shows that the Andes are both a key source and sink of Neotropical plant diversity, with a similar number of lineages colonizing and dispersing from the region (Figure 4B). Our results unveil a more prominent role of the Andes in the biotic interchange than reported by Antonelli et al. [4]. This is because we use a broader definition of the Andes, we include regions not considered in their analysis, we split the Andes into three regions, and, importantly, we include many more lineages (171 vs 104 clades). However, this does not bias the analyses to give a more prominent role to the Andes because our methodology is not enriched for Andean species (see Materials and Methods in the supplemental information online).

We unveil important biogeographical links between the Northern and Central Andes (but not the Southern Andes) and Amazonia, as well as between the Northern and to a lesser extent the Central Andes and Central America (Figure 4B). Specifically, the colonization of the Andes from Amazonia and dispersals from the Andes to Amazonia were high and similar in both directions (Figure 4B), which highlights their strong connectivity. The pattern is similar for Central America, where 65 Andean colonization events and 63 dispersal events were out of the Andes. Given the fluctuation of boundaries between ecosystems following glacial cycles, more dispersal events may have occurred during interglacial periods as ecosystem barriers shifted upslope. Our cross-taxonomic analysis of 194 plant clades also reveals slightly more dispersal outside the Andes than within (203 in vs 229 out). This suggests that in situ diversification has been the dominant evolutionary process in the Andes.

Concluding remarks
Our review identifies key patterns and processes underlying today’s outstanding levels of plant richness in the world’s longest mountain chain – the Andes. This compilation also aids in the identification of areas of future research, the type and source of data needed to address remaining questions (see Outstanding questions), and the methodological advances that are required. It is now particularly crucial to integrate past and present geological and climatic factors with biotic interactions so as to increase our understanding of the evolution of Andean biodiversity through time. By providing a new framework for the temporal assembly of Andean taxa and their migration routes throughout the Neotropics, we hope that our review may inspire further research on Andean plant diversity.

Acknowledgments
We thank Victor Alberto Ramos for permission to reuse the map in Figure 1A, Mario Coiro for help with plant fossil records curation, and Colin Hughes, two anonymous reviewers, and the editor Susanne Brink for comments that helped to improve the manuscript. Eve J. Lucas and Gwilym P. Lewis are thanked for providing useful insights into the taxonomy of the Andean plant list. O.A.P-E is funded by a Sainsbury Orchid Fellowship at the Royal Botanic Gardens Kew and the Swiss Orchid Foundation. A.Z. acknowledges funding by DvI via the German Research Foundation (FZT-118, DGF), specifically through sDvI, the Synthesis Centre of DvI. F.L.C. is supported by an Investissements d’Avenir grant managed by the Agence Nationale de la Recherche [ANR; Center for the Study of Biodiversity in Amazonia (CEBA) grant ANR-10-LABX-25-01] and by the ANR GAARAnt project (ANR-17-CE31-0009). M.A.B. thanks the Universidad Pedagógica y Tecnológica de Colombia

Outstanding Questions
How does sampling effort bias the general patterns of Andean plant species richness? Our working list of Andean plants reveals that mid-elevation North Andean forests are the species-richest Andean ecosystems, but it also shows that the Northern Andes have substantially more botanical collections than Central and Southern Andean regions. Working out how this sampling bias affects Andean plant diversity patterns, and how to correct for those, is an important future goal.

Are the biogeographic and diversification patterns of Andean floras biased by the uneven sequencing of Andean plants? We identified large differences of sampling between Andean plant families and genera, including in the most diverse ones. These sequencing gaps are priorities for future research on Andean plants.

How does Andean plant diversity compare with plant diversity in other tropical mountains? Recent data resources such as the IUCN Global Ecosystem Typology 2.0 and the Global Mountain Biodiversity Assessment open up possibilities to perform comparison of diversity in the Andes to other tropical mountains in a comparable way. To do so, new species lists such as the one we generated for the Andes here needs to be produced.

Is most of the Andean diversity already extinct? Recent data pushed the start of the Andean orogenesis in the Jurassic, with documented uplift in the Cretaceous, and fossil plants document presence of Andean fossil tropical forest plants in the early Cenozoic. By contrast, most Andean plant lineages are fairly young, which raise the question of whether the majority of Andean plants that once evolved have already gone extinct due to natural events, such as increased glaciation in the Quaternary. New paleobotanical data have suggested that some extinct Andean ecosystems lack current analogues, but more data are needed, particularly from the Central and Southern Andes.
No interests are declared.

Supplemental information

Supplemental information associated with this article can be found, in the online version, at https://doi.org/10.1016/j.trendsplant.2021.09.010.

References

378 Trends in Plant Science, April 2022, Vol. 27, No. 4